The utilisation of Fmoc solid phase chemistry as a novel approach to the generation of duocarmycin analogues

Michael James Stephenson

A thesis submitted for the degree of Doctor of Philosophy

August 2015

School of Pharmacy

University of East Anglia

Declaration

This thesis is submitted to the University of East Anglia for the Degree of Doctor of Philosophy and has not been previously submitted at this or any university for assessment or for any other degree. Except where stated, and reference or acknowledgment is given, this work is original and has been carried out by the author alone.

Michael James Stephenson

Acknowledgements

I would first of all like to thank my supervisors Prof. Mark Searcey, Dr Maria O'Connell and Dr Steve Collingwood for all of the support and guidance they have given to me during the course of my PhD studies. I feel privileged to have been afforded the opportunity to work on such a fascinating family of natural products, and one which captured my interest as an undergraduate (for which I must thank Mark again). I would also like to offer my deepest thanks to Dr Lesley Howell for her additional support and training.

My thanks also go to Prof. Keith Fox for conducting the DNA thermal cleavage assay reported in chapter three, to Dr Jenny Kingston for performing the supercritical fluid chromatography reported in chapter two, and to Dr G. Richard Stephenson for helping me perfom the DFT calculations reported in chapter three.

For their support and friendship, in and out of the lab, I thank Michael Austin, Sarah Goffin, and Dr Richard Steel. My thanks also go to everyone I met during the course of my time at UEA. I would also like to thank everyone I met during my placement at Novartis for making me feel so welcome, particularly Dr Claire Adcock, Dr Richard Robinson, and Dr Julia Hatto for their guidance and training, and the Barber family for welcoming me in to their home.

I thank the School of Pharmacy and Novartis for funding, and the EPSRC mass spectrometry service in Swansea for accurate mass analysis.

Finally I must thank my family for their support, and in particular I thank my partner Laura for allowing me to pursue my passion for science, and for her unwavering support at home.

Abstract

Duocarmycin SA is an archetypal member of an ultrapotent family of antitumor antibiotics. The family is characterised by a common spirocyclopropylcyclohexadienone pharmacophore. This unusual structural motif is responsible for exceptionally efficient alkylation of adenine bases following activation through conformational changes induced by non-covalent recognition of DNA's minor groove.

This thesis describes the conception and multi-gram synthesis of a duocarmycin SA alkylation subunit suitably substituted to serve as a building block for Fmoc based solid phase synthesis and initial investigations into its application.

Chapter two describes the pilot and subsequent large scale racemic synthesis of the desired duocarmycin building block, and its preparative chiral resolution by supercritical fluid chromatography. The synthesis includes a novel route to an early indole intermediate, and represents one of the shortest available strategies to access a previously reported di-Boc-protected duocarmycin structure. The large scale synthesis afforded over 8 g of the racemic Fmoc-protected building block, representing a 3 % overall yield over 13 steps.

In chapter three, application of the building block to the conjugation of amino acids through solid phase synthesis is explored. This work highlights the importance of careful resin selection and the need to optimise cleavage conditions. A small library of duocarmycin analogues was generated, and subsequent assays revealed the effects of *C*-terminal amino acid substituents on biological activity. This work has demonstrated the potential utility of this building block for the future development of novel peptide linked antibody drug conjugates of duocarmycin SA.

Finally, chapter four explores incorporation of the building block into polypyrrole structures, highlighting the potential for the direct solid phase synthesis of sequence selective bifunctional hairpin polyamides which contain the alkylation subunit of duocarmycin SA.

Published work within this thesis

Michael J Stephenson, Lesley A. Howell, Maria A. O'Connell, Keith R Fox, Claire Adcock, Jenny Kingston, Helen Sheldrake, Klaus Pors, Stephen P Collingwood, and Mark Searcey. The solid phase synthesis of duocarmycin analogues and the effect of C-terminal substitution on biological activity. *Journal of Organic Chemistry*, Just Accepted Manuscript. DOI: 10.1021/acs.joc.5b01373

Table of contents

List of figures

List of Schemes

List of tables

Abbreviations

A	adenine
ACN	acetonitrile
ADC	antibody drug conjugate
AIBN	azobisisobutyronitrile
Ala	alanine
BCR	breakpoint cluster region
BnBr	benzyl bromide
Boc	*tert*-butyloxycarbonyl
Boc_2O	di-*tert*-butyl dicarbonate
Bu	butyl
C	carbon or cytosine (depending on context)
CBI	cyclopropabenzo[e]indolone
CD	cluster of differentiation
CML	chronic myelogenous leukemia
COSY	correlation spectroscopy
CPI	cyclopropylpyrrolo[e]indolone
CYP	cytochrome P450 enzyme
DCM	dichloromethane
DEPT	distortionless enhancement by polarisation transfer
DFT	density functional theory
DIC	diisopropylcarbodiimide
DIPEA	*N*,*N*-diisopropylethylamine
DMAP	4-dimethylaminopyridine
DNA	deoxyribonucleic acid
DME	dimethoxyethane
DMF	*N*,*N*-dimethylformamide
DMSO	dimethylsulfoxide
DSC	differential scanning calorimetry
EDCI	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDT	1,2-ethanedithiol
ee	enantiomeric excess
Et	ethyl
Fmoc	fluorenylmethyloxycarbonyl
Fmoc-Cl	fluorenylmethyloxycarbonyl chloride

Fmoc-OSu	*N*-(fluorenylmethoxycarbonyloxy)succinimide
g	gram
G	guanine
Glu	glutamic acid
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HBTU	*O*-(Benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate
HCl	hydrochloric acid
HER2	human epidermal growth factor 2
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HOMO	highest occupied molecular orbital
Hp	*N*-methyl-3-hydroxypyrrole
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum correlation
IC_{50}	half maximal inhibitory concentration
ICR	intrinsic reaction coordinate
Im	*N*-methylimidazole
IPA	isopropanol
i-Pr	isopropyl
K	Kelvin or potassium (depending on context)
kg	kilogram
LC-MS	liquid chromatography–mass spectrometry
Lys	lysine
M	Molar
MeOH	Methanol
min	minute
MTS	3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium
NAD	nicotinamide adenine dinucleotide
NaH	sodium hydride
NIS	*N*-iodosuccinimide
nM	nanomolar
nm	nanometre
NMR	nuclear magnetic resonance spectroscopy

^{o}C	degrees celsius
PEG	polyethylene glycol
PES	phenazine ethosulfate
Ph	phenyl
pH	log hydrogen cation concentration
Phe	phenylalanine
pK_a	log acid dissociation constant
Py	*N*-methylpyrrole
PyBop	(Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
RNA	ribonucleic acid
SCF	supercritical fluid
Ser	serine
S_N	nucleophilic substitution
TBAF	tetrabutylammonium fluoride solution
***t*-butyl**	*tert*-butyl
TES	triethylsilane
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIPS	triisopropylsilane
TLC	thin-layer chromatography
TTMSS	tris(trimethylsilyl)silane
UV	ultraviolet
µg	microgram
µM	micromolar

1

Chapter One

Introduction

Cancer is a condition arising from the aberrant proliferation of an individual's own cells. The term encompasses hundreds of individual diseases characterised by their cell type of origin.[1] It is now generally accepted that the transformation of a normal somatic cell to its malignant counterpart is a progressive process. This progression has been likened to Darwinian evolution in that successive genetic mutations, each conferring a specific growth advantage, eventually leads to the presentation of a malignant phenotype.[2, 3]

The process is often segmented into three mechanistic phases.[4] The first is tumour initiation. In this phase, exposure to chemical or physical carcinogens produce DNA damage. Mutations arising from this damage, resulting in the unrestrained activation of oncogenes, or the ablation of tumour-suppressor genes, set the affected cells on their path to malignancy. During the second phase, tumour promotion, these initiated cells undergo a clonal expansion. The expansion is a result of disruption to regulatory circuits which govern normal cell proliferation and tissue homoeostasis. This increase in proliferation, coupled with genomic instability, can lead to the acquisition of yet more mutations conferring further growth advantages. The final phase, tumour progression, is characterised by the invasion of surrounding tissue, growth in tumour sizes, and eventual metastasis. Again further tumour promoting mutations can drive this stage.

The vast array of different malignancies could be a problem for the field of cancer research. It has led to a position of seemingly insurmountable complexity in the differing pathophysiological changes and genomic presentations governing the development of each individual disease. In 2000, Hanahan and Weinberg attempted to simplify the situation. In their seminal review, '*The Hallmarks of Cancer*', they outlined six characteristic features of the malignant cell,[5] which they believed likely to be common to the majority, if not all cancers; essentially suggesting that the differences between individual malignancies stem from the different mechanisms utilised to affect these abilities and the order in which they are acquired. Although an essentially accurate model, this view constrains consideration of the tumorigenic process to intrinsic properties of the transforming cells alone. An increasing body of evidence suggests that this may be an over simplification. It is suspected that normal cells play an important role in the promotion of malignancy, particularly the inflammatory component of many tumour microenvironments.[6] Such interactions may serve to supplement the progression of malignancy prior to the aberrant cell line acquiring particular abilities. This is reflected in Hanahan and Weinberg follow up review published in 2011.[7]

For example, healthy cells require external signals to move from a state of quiescent to that of proliferation. They are also subject to signals which inhibit growth. Malignant cells

must therefore acquire autonomy over proliferative control. This requires self-sufficiency in growth promotion and insensitivity to inhibitory signals. In many cases, the former is a result of oncogenic gene products providing autocrine growth signalling. However there is also evidence for inflammatory cells providing the necessary stimuli. The evasion of growth inhibition is often the result of mutations in tumour-suppressor genes which disrupted inhibitory signalling cascades.[5]

Rate of proliferation is not the only factor governing tumour development. Many cancerous tissues show a decrease in cellular attrition. Furthermore DNA damage inherent in tumour initiation should in many cases result in the activation of regulatory machinery, and subsequent repair or apoptosis. Thus the malignant cell must acquire mechanisms allowing the evasion of apoptopic commands.[5] Again this may be the result of mutations in tumour-suppressor genes whose products are pivotal in implementing programmed cell death, although anti-apoptopic NF-κB gene products activated through inflammatory stimulus have also been implicated.[6]

Normal somatic cells possess an intrinsic tumour suppressing property; they are unable to divide indefinitely.[5] This is a consequence of the end-replication problem which causes the gradual erosion of telomeres. Telomeres, which comprise non-coding regions at the end of chromosomes, have a protective function. Once telomere erosion reaches a critical point this protective effect is loss, and the cell enters senescence; a non-proliferatory state. Malignant cells must acquire telomere extension mechanisms conferring them limitless replicative capacity. In 85 % of cancers this is achieved by expression of the dormant enzyme telomerase.[8] This reverse transcriptase uses an endogenous RNA template to extend the tandemly repeated telomere sequence, using the single stranded 3' overhang as its primer.[9]

All cells, including cancer cells, require adequate perfusion from the vasculature to meet their demand for oxygen and nutrients.[5] The expansion of tumours appears to be initially curtailed by an inability to stimulate the required level of angiogenesis. However this is eventually overcome, in what has been described as an angiogenic switch;[10] wherein the aberrant malignant cells acquire the ability to alter the balance of pro and inhibitory factors in favour of stimulating the require vessel growth.

The ability of primary tumours to spawn pioneer cells which can relocate and establish secondary tumours, is perhaps the most lethal acquired ability of the malignant cell; being responsible for 90 % of cancer deaths.[11] The alteration of several protein classes, especially those involved in the tethering of cells to extra cellular matrixes, have been implicated in the mechanisms of metastasis.[5]

Despite the biological complexity of cancer, traditional therapy has remained relatively simplistic. Surgery to excise the diseased tissue is by far the most effective treatment for most solid tumours. However, radiotherapy and chemotherapy play important roles both prior to excision, to decrease tumour mass, and as adjuvants following surgery, to prevent reoccurrence from potential residual tumour cells.[12]

Cytotoxic drugs have long dominated pharmaceutical intervention in cancer therapy.[13] Many such treatments target DNA directly, through both non-covalent and covalent interactions, disrupting the cell's replicatory machinery, and leading to apoptotic responses. For example doxorubicin (figure 1.1) is widely prescribed in the treatment of many cancers, including various carcinomas, sarcomas, and haematology malignancies.[14] It is an intercalating agent, binding non-covalently with DNA through insertion of its planar chromophore between two adjacent base pairs, and projecting its glyosidic structure in to the minor groove.[15]

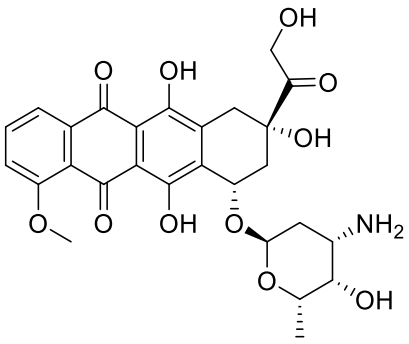
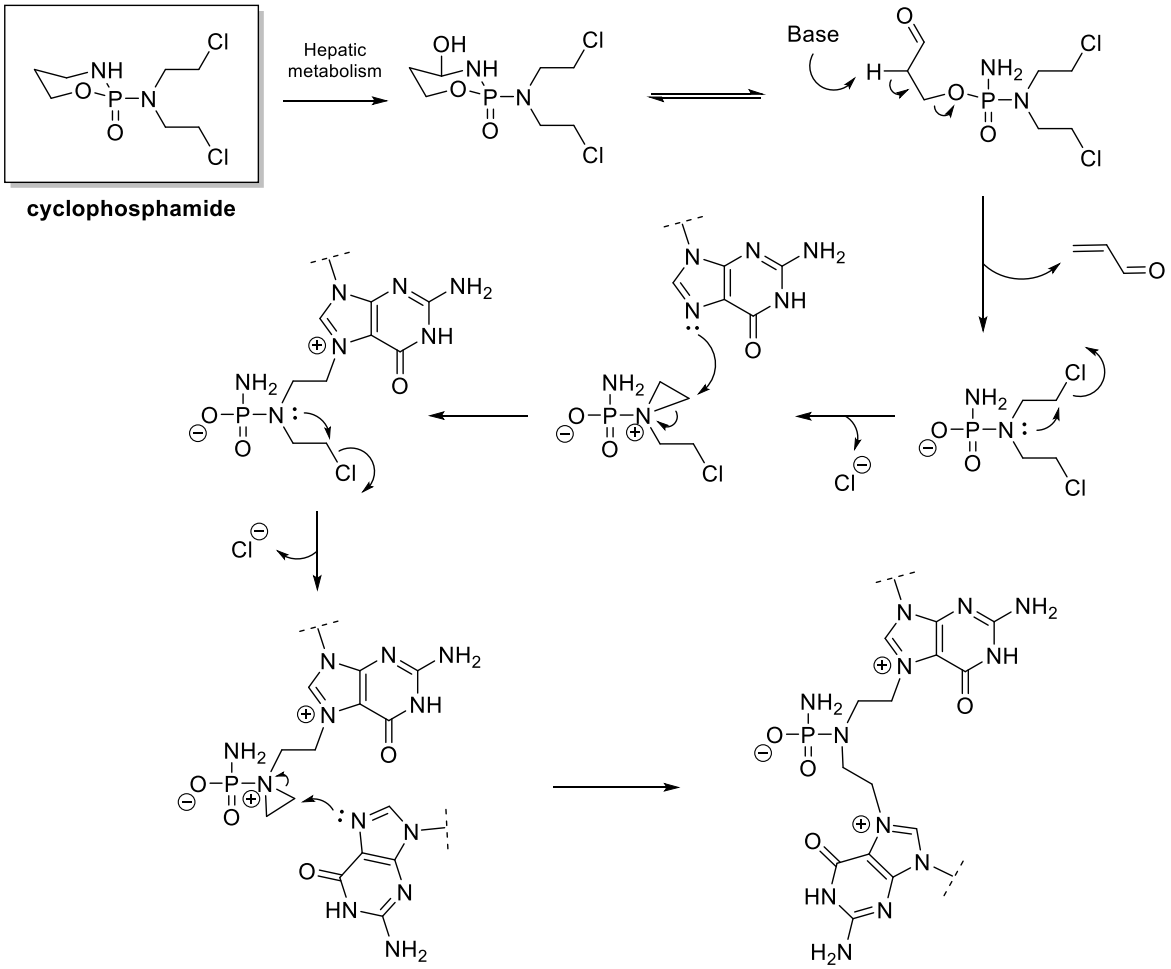


Figure 1.1 Structure of doxorubicin.

Cyclophosphamide is an alkylating agent, and an example of a treatment which targets DNA through covalent interaction.[16] It is arguably one of the most successful antineoplastic drugs in the history of cancer chemotherapy. It is effective against a broad spectrum of malignancies, and is still commonly included in many standard treatment regimes,[14] despite being first introduced over 50 years ago.[16]

Cyclophosphamide is an oxazaphosphorine prodrug of the original aliphatic nitrogen mustards. Hepatic cytochrome P-450 metabolism leads to oxidation of the oxazaphosphorine ring vincinal to the ring bound nitrogen. Tautomerization, and subsequent β-elimination reveal the active mustard, which is free to form the highly unstable aziridinium cation and ultimately alkylate DNA through nucleophilic attack by guanine bases. The presence of two chloroethyl groups means inter and intrastrand crosslinking is often observed (scheme 1.1).[16, 17]



Scheme 1.1 Structure and activation of cyclophosphamide.

As our understanding of the biology of cancer has evolved the scope of potential treatment strategies has widened. It is therefore unsurprising that the draw of DNA as a target has diminished over recent years, with many researchers turning their attention to the disruption of aberrant signalling pathways specific to particular malignant cell lines. Such approaches aim to produce more selective and therefore tolerable treatments. This strategy has already yielded significant successes in the clinic.

Imatinib (figure 1.2), was one of the first such treatments to gain regulatory approval, and has revolutionised the treatment of susceptible forms of chronic myeloid leukaemia (CML). It is an orally administered synthetic tyrosine kinase inhibitor, which targets the BCR-ABL oncogenic gene product;[18] a fusion gene resulting from a chromosomal translocation event which is common to the pathogenesis of the majority of CML cases. The BCR-ABL protein is a permanently activated tyrosine kinase, whose activity results in the uncontrolled proliferation of affected cells.[19]

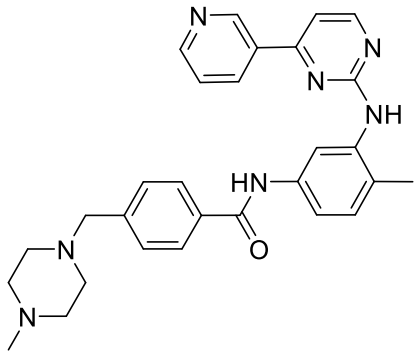
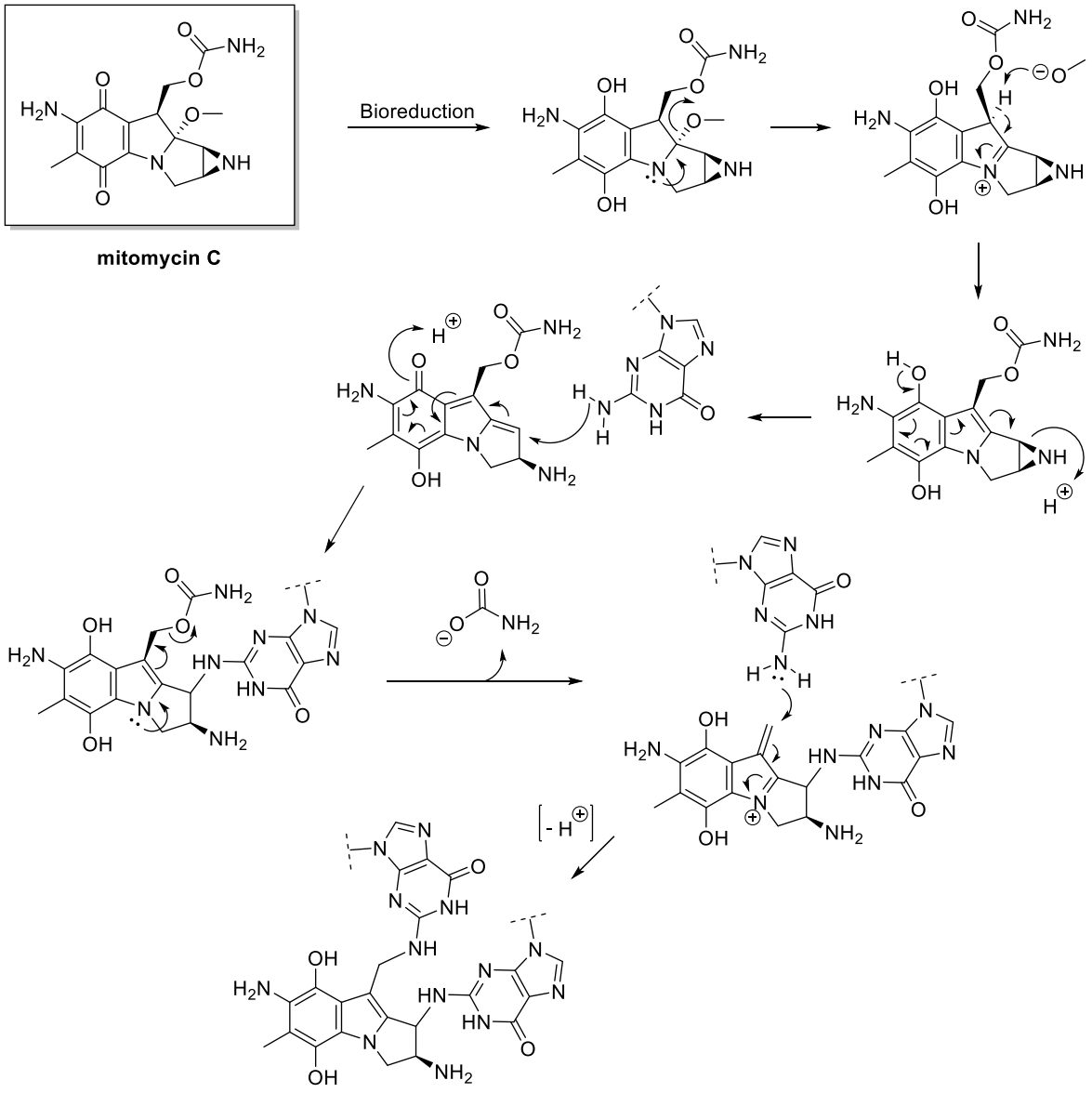


Figure 1.2 Structure of imatinib.

The success of imatinib has been said to mark a paradigm shift in the treatment of malignant disease,[20] and since its introduction in 2001 many similar molecular targeted therapies have gained regulatory approval.[21] For example, trastuzumab, a monoclonal antibody which targets the *HER2* growth receptor,[22] has had a major impact on the treatment *HER2* positive breast cancer.[23, 24]

It is likely that this field will contiune to grow and yield many more ground breaking treatments in the future. Nevertheless, susceptibility to resistance, the implict narrow spectrum of such treatments, and controversy surrounding the cost-effectivness of some less successful additions to the clinic, mean a reliance on conventional cytotoxins may remain for some time.[25-27] It is therefore still of value to explore the development of novel cytotoxic compounds, especially those with the necessary potency to be applied successfully to tumour directed delivery and/or prodrug strategies. Such approaches also hold the potential to yield more selective and tolerable treatments.

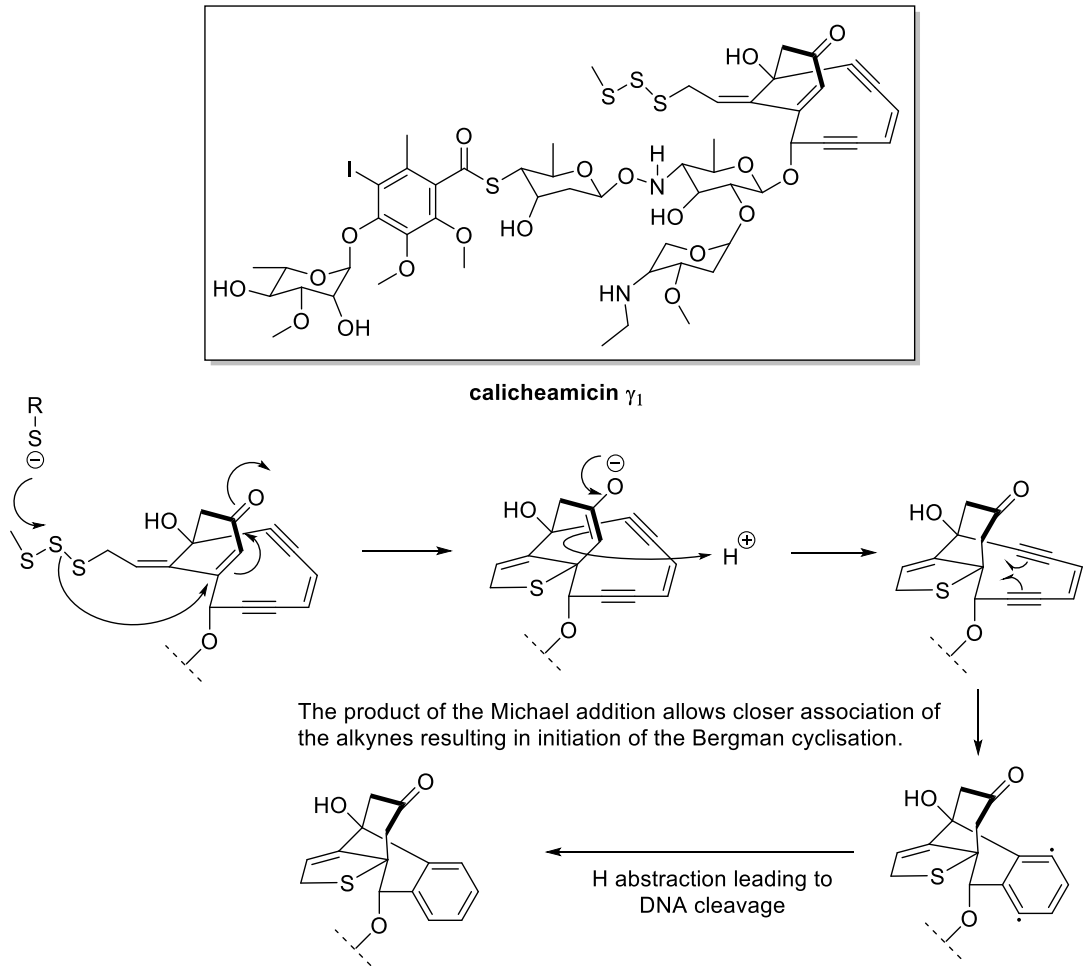
Perhaps contrary to public perception, particularly of those influenced by the market of alternative therapy, modern medicine has long recognised the therapeutic potential of natural products. However, unlike the pedlars of snake oils and folklore, this reasoned world of scientific exploration, has always viewed nature's rich pool of biologically active secondary metabolites as leads; understanding the need to optimise these compounds for uses to which they did not evolve. The selection pressures imposed upon bacteria have resulted in a vast array of compounds with obvious applications in the treatment of infectious and malignant disease. It was reported in 2010, that since the 1940s 51.5 % of all anticancer agents introduced were either natural products or directly derived analogues, many originally isolated from microbes.[28] Indeed, the aforementioned doxorubicin is a member of the anthracycline family of natural products; secondary metabolites isolated from various strains of *Streptomyces.* [29, 30]



Scheme 1.2 Structure and DNA crosslinking activity of mitomycin C.

Another pertinent example would be mitomycin C. Introduced in the 1960s,[31] mitomycin C is still commonly included in a number of treatment regimens.[14] It is a member of the aziridine family of antitumor antibiotics, isolated from *Streptomyces caespitosus*.[32] The cytotoxicity of mitomycin C is derived from its ability to crosslink DNA through a bifunctional covalent interaction. This is thought to occur through a complex series of intramolecular transformations, initiated from the bioreduction of mitomycin C to its hydroquinone counterpart (scheme 1.2).[33] Subsequent loss of the methoxy group, and tautomerization of the resulting imine leads to opening of the aziridine ring by the hydroquinone through the indole structure, and generation of the first alkylating species. This is attacked by guanine bases by a conjugate addition. The resulting reformation of

the hydroquinone indole structure, allows formation of the second alkylating species, *via* loss of the carbamate.



Scheme 1.3 Structure and mechanism of cytotoxicity of calicheamicin γ1.

It is a pleasing coincidence for the medicinal chemist, that two of the most potent families of antitumor antibiotics discovered to date also derive their biological activity from interesting mechanistic chemistry. Perhaps the most potent is the family comprising the enediyne antibiotics, such as calicheamicin originally isolated from *Micromonospora echinospora* (scheme 1.3).[34] Their cytotoxicity is developed from the diradical intermediate of a biologically triggered Bergman cyclisation.[35]

The second family is that consisting of CC-1065, yatakemycin, and the duocarmycins, whose structures are depicted in figure 1.3.[36] It has now been over 30 years since the first member of this family, CC-1065,[37, 38] was reported, and despite substantial interest from distinguished medicinal chemistry programmes, research is yet to yield a successful clinical candidate. However this unrealised therapeutic potential only adds to the allure of these already chemically pleasing entities.

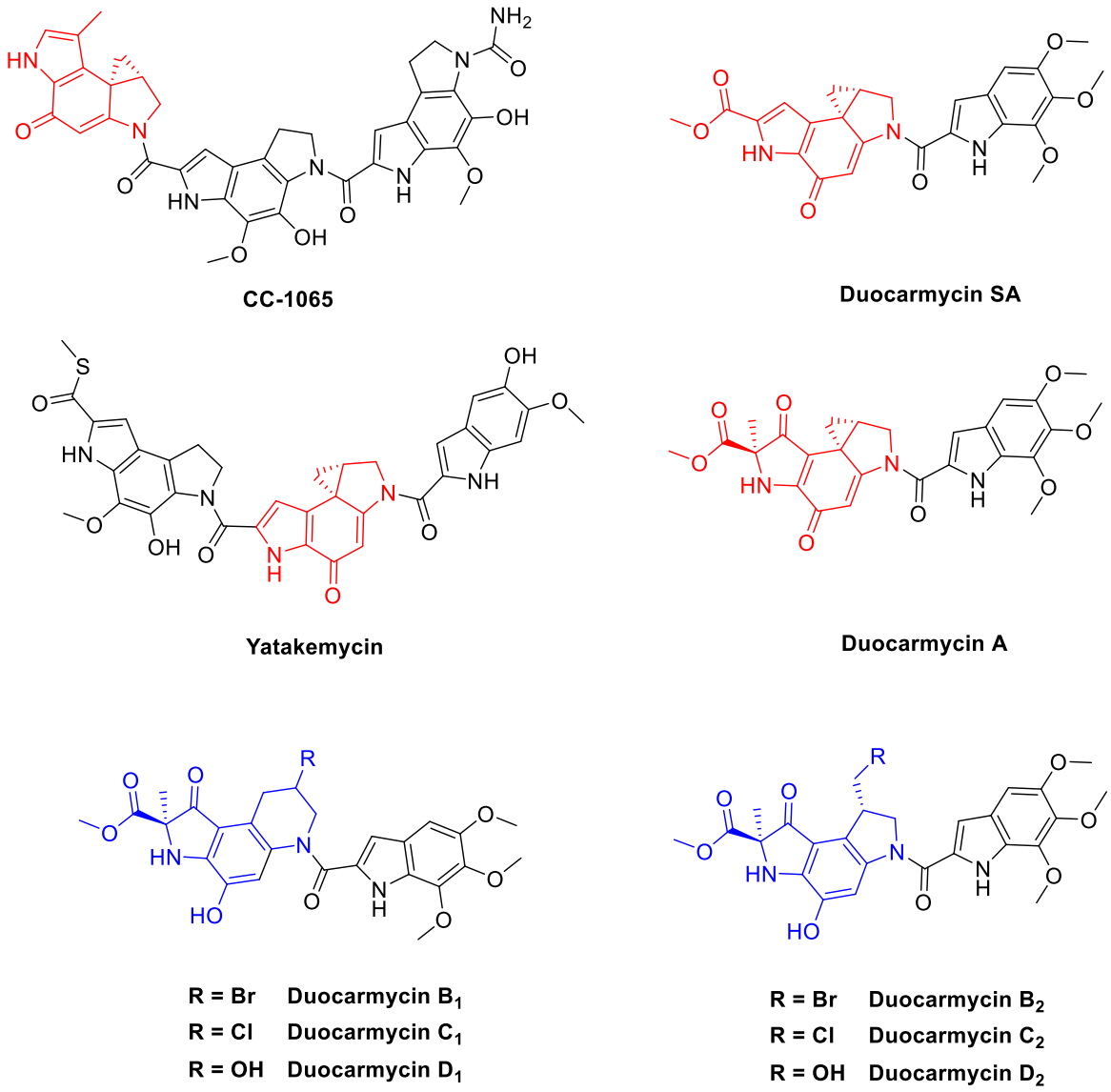
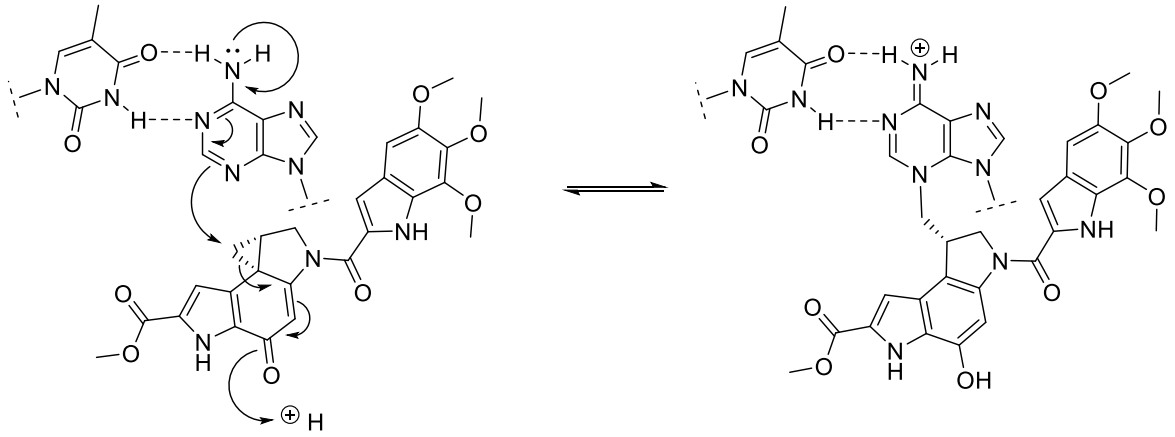


Figure 1.3 Structures of CC-1065, yatakemycin, and the duocarmycins. Red = alkylation subunit. Blue = alkylation subunit precursor – requiring spirocyclisation prior to the alkylation event.

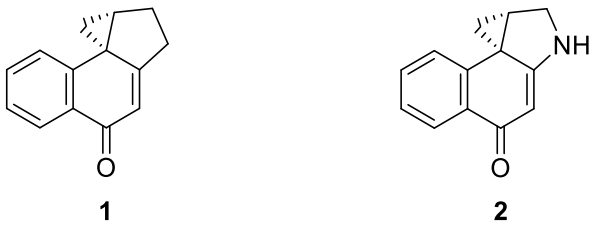
CC-1065, yatakemycin, and the duocarmycins all combine non-covalent recognition of AT rich regions of DNA's minor grove with exceptionally efficient alkylation of adenine,[39] the downstream effects of which result in their ultrapotent cytotoxic activity. The common pharmacophore responsible for this alkylation consists of a cyclopropane vicinal to the α, β unsaturated carbonyl of a cyclohexadienone. This spirocyclic system is flanked by the fusion of a further two 5 membered nitrogen heterocycles. Alkylation proceeds through nucleophilic attack of the cyclopropane by the N3 of adenine (Scheme 1.4). This occurs at the least substituted carbon, resulting in ring opening of the cyclopropane and subsequent development of an aromatic phenol from the conjugated and activating cyclohexadienone. It is the control of this seemingly facile reaction, imposed by the subtle interplay of molecular features contained within the compact structure of these compounds, that

defines their characteristic and arguably most interesting feature: *in situ* activation.[40] As will be described this is a quite unique and elegant process; one which does not require induction of chemical change, but simply a change in conformation induced from binding to their target.



Scheme 1.4 Alkylation of adenine by duocarmycin SA.

In solution these compounds demonstrate remarkable stability towards nucleophilic attack. It is only when bound within the minor groove of DNA that the true reactivity of the cyclopropane is revealed.[41] This stability results from the electronic influence of a vinylogous amide constructed from the carbonyl of the cyclohexadienone and the nitrogen of the neighbouring heterocycle. This feature serves to reduce the electrophilicity of the cyclohexadienone by conjugation of the nitrogen's lone pair with the carbonyl and stabilise the cyclopropane.



1 **2**

Figure 1.4 Structures of CBI analogues **1** and **2**.

The magnitude of this effect has been implicitly demonstrated by comparison of the reactivity of a carbocyclic analogue of the CBI alkylation subunit with its nitrogen containing counterpart (figure 1.4).[42] It was shown that **1** which lacks the vinylogous amide, was 3200 times more reactive than **2** towards solvolysis at pH 3; and at pH 7 where **2** is stable, **1** was revealed to be 104 times more reactive. It is disruption of the conjugation of this vinylogous amide which triggers these natural products *in situ* activation.

Disruption of this stabilising influence is made possible because the nitrogen of the vinylogous amide also forms part of a linking amide, through which the alkylation subunits of the natural products are connected to additional, non-alkylating, bi or polycyclic systems, known to contribute non-covalently to minor groove recognition. As such, the lone pair of this nitrogen is effectively cross conjugated with the carbonyls of both amides. When the natural products bind to DNA they undergo a conformational change. This change involves rotation around the linking amide so as to allow the adoption of a bent conformation which is more complementary to the helical rise of the minor groove. Structural studies of bound duocarmycin SA suggest that the bulk of this twist is carried by the $\chi 1$ dihedral angle (22.4 ^{o}), as opposed to that of the amide $\chi 2$ (11.0 ^{o}) (see figure 1.5).[43, 44] The result is a change in the balance of cross conjugation in favour of the linking amide, as the nitrogen's lone pair is no longer planner with the cyclohexadienone's π system.

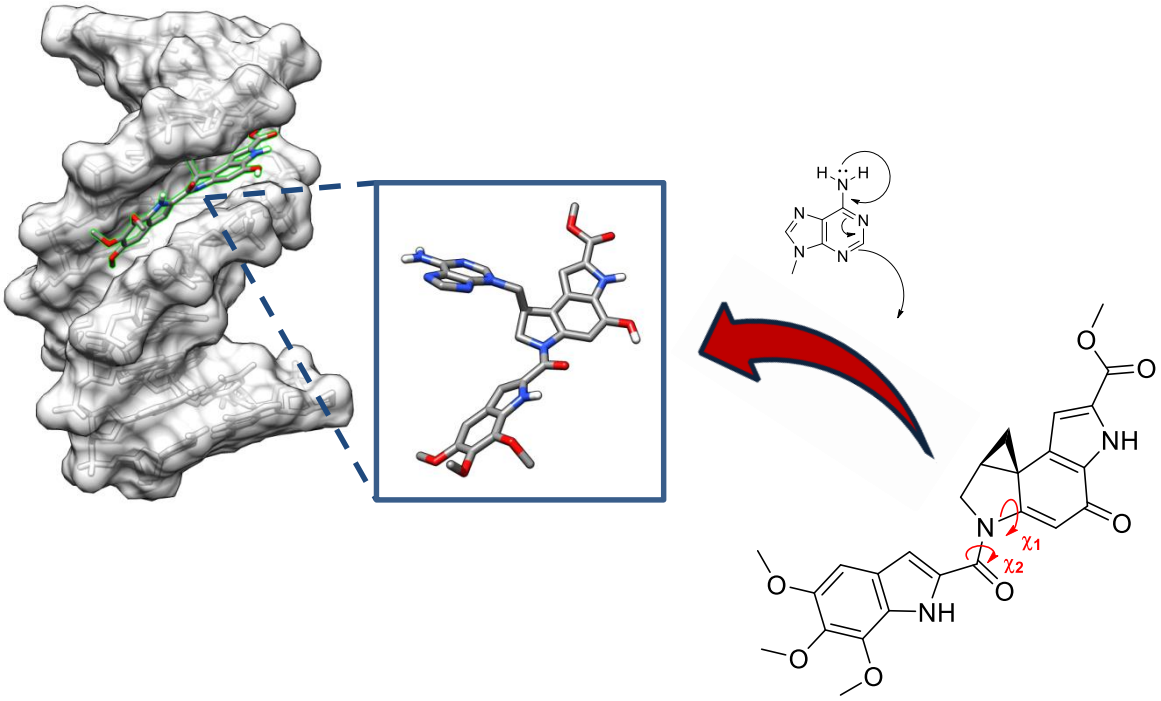


Figure 1.5 Visual representation of the *in situ* activation of duocarmycin SA (see main text for description). NMR Structure images generated using 'Chimera 1.10.1', coordinates from protein data bank (ID:1DSA).[42]

Comparison of the activity of CBI analogues **3** and **4** (figure 1.6) clearly shows the necessity of the linking amide in this process. While DNA alkylation is easily detectable with **3**, which possesses the linking amide, **4**, in which this amide is replaced by a methylene bridge is unable to alkylate DNA even under forcing conditions.[45]

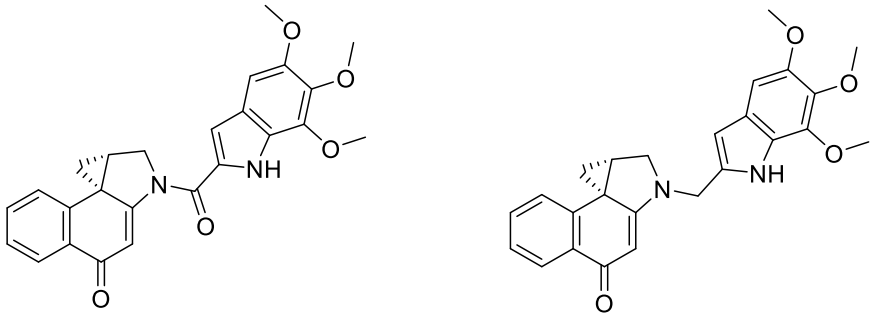


Figure 1.6 Structures of CBI analogues **3**, and **4**.

Comparison of the x-ray structures of a series of CBI analogues (figure 1.7), where the ring size of the nitrogen heterocycle was increased from 5 through to 7 atoms, further demonstrates this effect, and captures its incremental nature. As ring size increases the $\chi 1$ angle widens. This is accompanied by a lengthening in the bond between the nitrogen and cyclohexadienone, reflecting reduction in the conjugation of the vinylogous amide. These structural features are mirrored by an expected progressive increase in reactivity as the degree of vinylogous amide conjugation decreases. [46, 47]

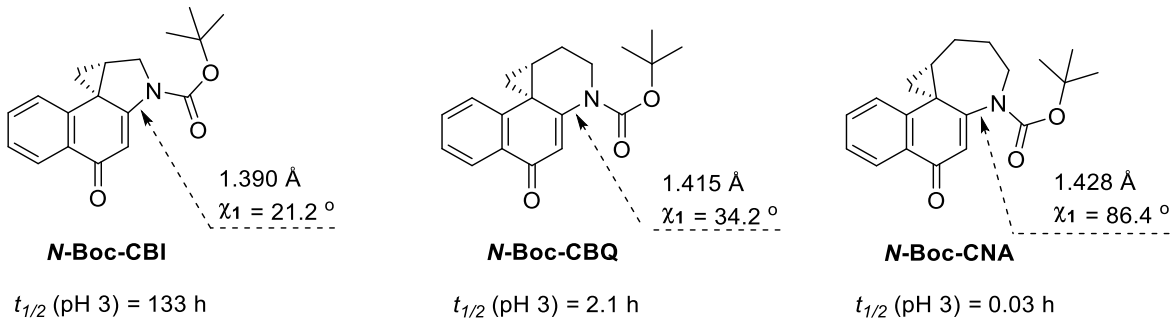


Figure 1.7 Structure of CBI analogues with increasing heterocycle size. Showing discussed crystallography and stability data.

As disruption of conjugation of the vinylogous amide is not an absolute process, and the degree of rotation induced at the $\chi 1$ angle can have an incremental effect on the reactivity of the spirocyclopropylcyclohexadienone, the rigid length of compounds containing this pharmacophore can have a significant impact on their relative DNA alkylation efficiencies. Essentially, longer structures can amplify the twist imposed by the helical rise of the minor groove, and thus increase the reactivity of the cyclopropane.

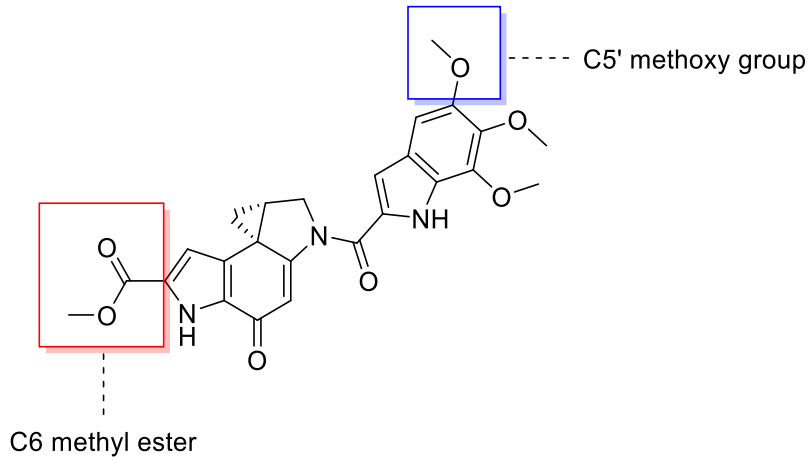


Figure 1.8 Structure of duocarmycin SA highlighting the C6 methyl ester, and C5' methoxy group.

This relationship is manifested in the observed importance of both the C6 methyl ester, and C5' methoxy substituent, on the DNA alkylation efficiency of duocarmycin SA (figure 1.8). Despite being shown to confer on the alkylation subunit greater stability towards solvolysis, the absence of the C6 methyl ester reduces the relative rate of DNA alkylation by 12 fold. Similarly, removal of the C5' methoxy substituent produces a 20 fold decrease in DNA alkylation rate. When both substituents are removed, a 250 fold decrease in DNA alkylation rate is observed.[48]

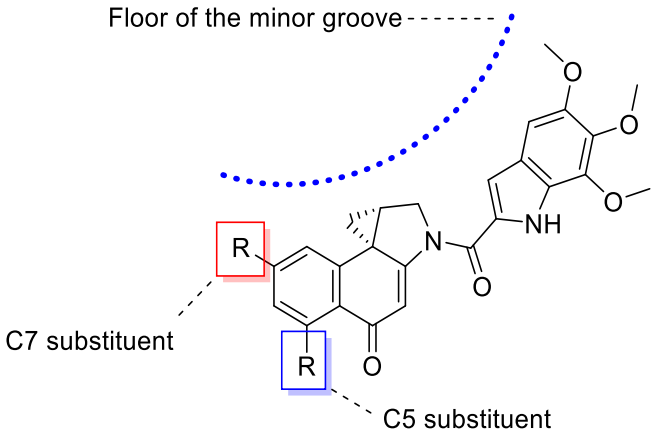


Figure 1.9 Structure of discussed CBI analogues, highlighting the C7, and C5 substituents, and their respective positions relative to the full of the minor groove.

These findings have been mirrored through investigation of the effects of substitution at the C7, and C5 position of the CBI alkylation subunit (figure 1.9).[49-51] It was observed that C7 substituents, which extend along the floor of the minor groove, conveyed a rate accelerating effect on DNA alkylation activity, and that this was independent of the substituent's electronic character, and hence corresponding effect on solvolytic stability of the respective alkylation subunits. Conversely, addition of a methyl ester at the C5

position, which extends out of the minor groove, significantly enhanced the solvolytic stability of the alkylation subunit, while also reducing the rate of DNA alkylation, in line with its electronic influence. In the C5 position, the substituent does not increase the rigid length of the structure relative to the minor groove, and hence its stabilising effect is not countered by an increase in binding induce activation.

The relationship between the intrinsic stability of the alkylation subunit and its corresponding biological activity is now well established. Many analogues of the spirocyclopropylcyclohexadienone system have been generated ranging vastly in their resistance towards solvolysis. A collative meta-analysis has revealed a strong parabolic relationship between solvolytic stability and anti-proliferation activity (figure 1.10).[52] This is intuitively unsurprising; intrinsic stability towards nucleophilic attack confers a greater chance that the alkylation subunit will reach its target intact, but this advantage is only maintained if upon binding in the minor groove, the stability can be sufficiently disrupted by *in situ* activation to produce an efficient alkylation event. Pleasingly, the two most potent natural products in this family, yatakemycin and duocarmycin SA, are positioned at the parabola of this trend. This beautifully highlights the power of natural selection to tailor the properties of secondary metabolites through the direction of biosynthetic pathways in microbes. Indeed, the S of duocarmycin SA literally stands for stable, and this ultrapotent antitumor antibiotic has been referred to as nature's prodrug. [53]

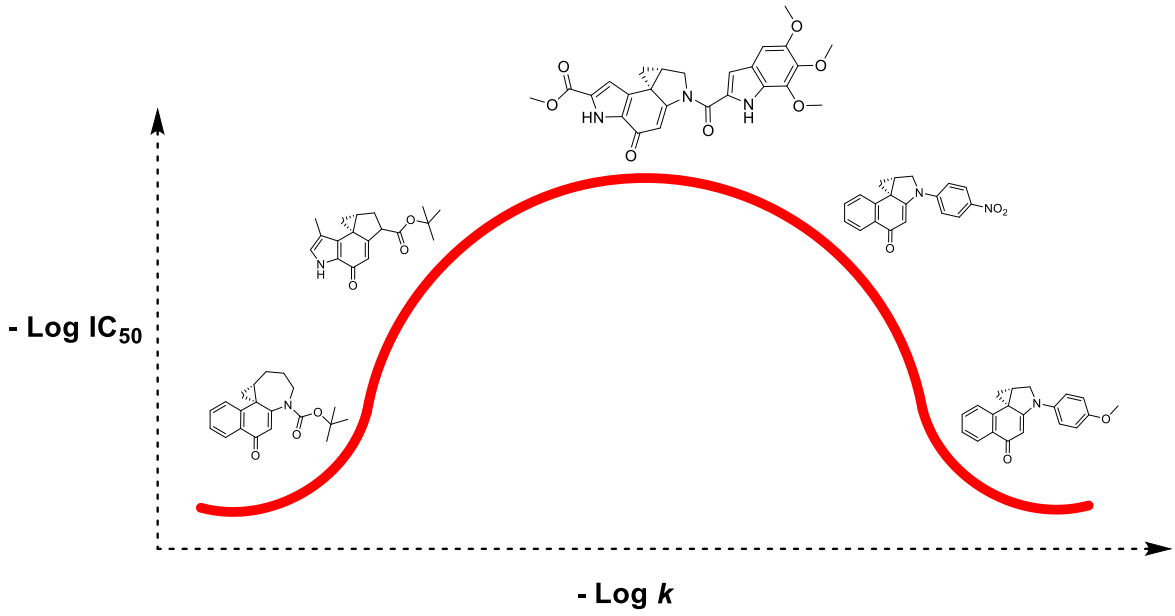


Figure 1.10 Schematic representation of the observed parabolic trend between intrinsic stability and biological potency of the spirocyclopropylcyclohexadienone family.

However, duocarmycin SA is only a prodrug in the sense that it selectively alkylates DNA over other biological nucleophiles. Like all classical alkylating agents, the natural products of this family are incapable of discriminating malignant cells from those of healthy tissue. They therefore have the potential to cause dose limiting toxicities apparent with all non-selective cyctotoxins. The ultra-potent nature of this family has to date prohibited their clinical use in the treatment of malignant disease. For example, the first member of this family to be discovered, CC-1065, quickly moved into pre-clinical evaluation, following identification of potent activity against a large range of malignant cell lines *in vitro*.[54] However, it could not progress in to clinical trials, due to exhibiting severe hepatotoxicity, which led to delayed death syndrome in mice at sub-therapeutic doses.[55]

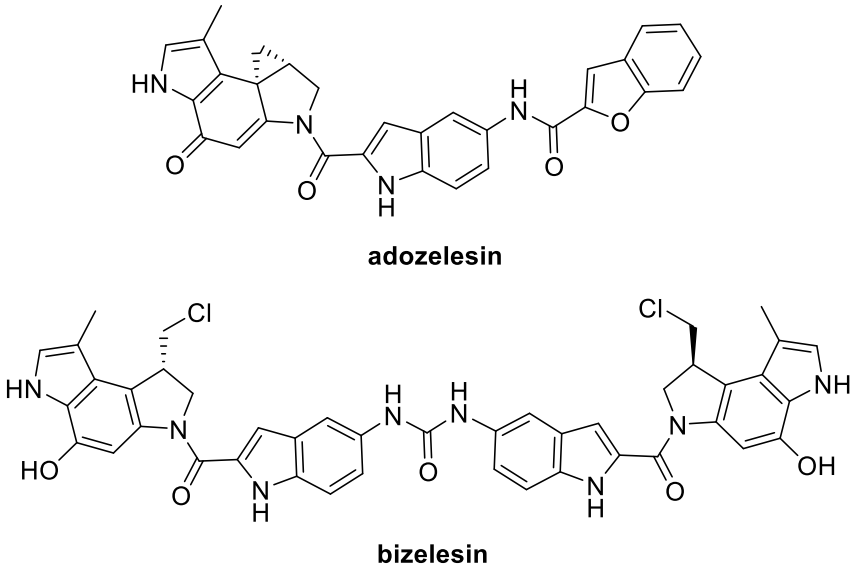
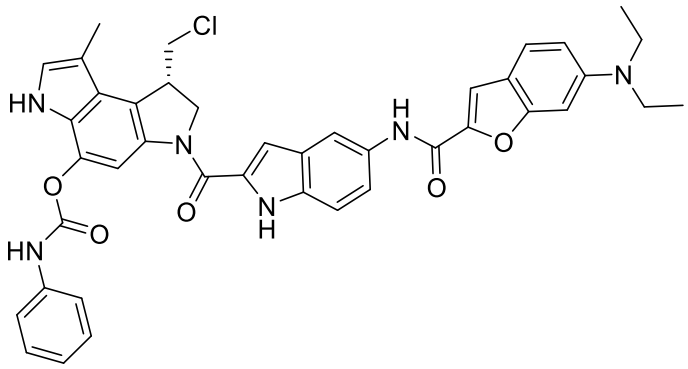


Figure 1.11 Structures of adozelesin and bizelesin.

This delayed death syndrome is not however a universal feature of the spirocyclopropylcyclohexadienone pharmacophore, but appears to be specific to the structure of CC-1065. This may be a consequence of the apparent irreversibility of the alkylation event, or specific sequence selectivity of CC-1065. Indeed, analogues of CC-1065 have been synthesised that do not exhibit the delayed death syndrome. Two such clinical candidates, adozelesin, and it's bifunctional analogue bizelesin (figure 1.11), have progressed into early clinical trials, but have ultimately failed to reach the clinic due to unacceptable toxicity profiles and limited efficacy at tolerated doses.[56-62]

It would seem clear that some form of targeting mechanism is necessary in order to harness the impressive antitumour activity of this unique family of natural products. One advantage of pharmacophores which possess ultrapotent cytotoxicity mediated through mechanistic chemistry, is that they provide the opportunity for the rational design of prodrugs.

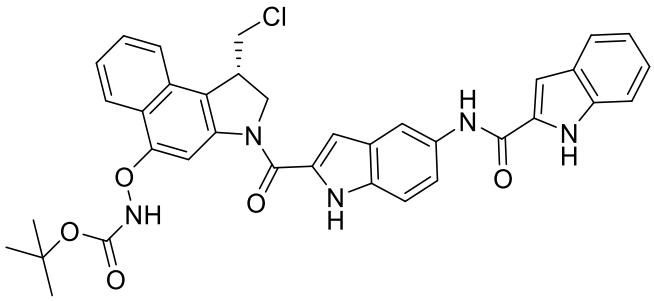


carzelesin

Figure 1.12 Structure of carzelesin.

A common strategy employed with the spirocyclopropylcyclohexadienone pharmacophore has been to inhibit the *in vivo* spirocyclisation of their seco precursors through biologically labile protection of the phenolic oxygen. Carzelesin (figure 1.12), one of the few compounds of this family to enter phase II clinical trials exploited this tactic.[63-66] Protection was achieved through a phenylurethane moiety which required serum dependent hydrolysis for cleavage.[67] More recent refinements of this approach have included attempts at introducing phenol protection which is cleavable under tumour selective conditions. The hypoxic nature of the tumour microenvironment has been key to this strategy.

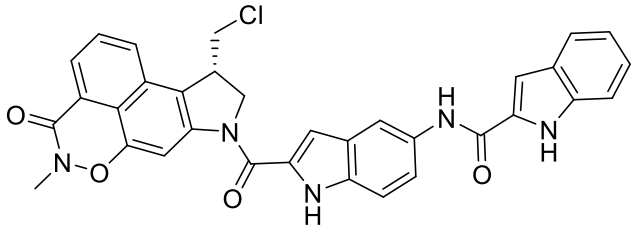
One example has been the development of *N*-acyl *O*-amino derivatives of the seco form of the CBI subunit. These were designed with the anticipation that the weak N-O bond could be cleaved by reducing nucleophiles of which a higher concentration can be found in hypoxic tissue.[68, 69] It was shown that the stability of this bond could be predictably tuned by the electronic and steric nature of *N*-substitution. The most effective derivative was a *tert*-Butyl carbamate, **5** (figure 1.13). It was shown to be incapable of alkylating DNA in cell free systems, but demonstrated potency approaching that of the free drug in functional cellular assays, with improved efficacy revealed from *in vivo* antitumor models. This suggests preferential release of the free drug at tumour sites. Given the average life span of the control mice and the length of the study, the number of long term survivors detected suggests, that despite the absence of dose optimisation, cures were detected.



5

Figure 1.13 Structure of *N*-acyl *O*-amino prodrug **5**.

This concept was later refined with the report of cyclic variants of **5**. The oxazinone prodrug **6** (figure 1.14),[70] was designed to be cleaved by the same reductive mechanisms, but to release no extraneous masking group as a by-product; a strategy that could simplify downstream regulatory approval should this species progress as a clinical candidate.



6

Figure 1.14 Structure of oxazinone prodrug **6**.

6, was shown to exhibit hypoxia selective cytotoxicity in cellular assays, returning a 10 fold increase in potency compared to its activity under aerobic conditions in a PC3/MM2 cell line. Furthermore it was shown to demonstrate good stability in human plasma with a half-life of one week. In addition, **6** was considerably more efficacious than a seco-CBI equivalent in an *in vivo* murine model despite returning higher IC_{50} values *in vitro*. At a dose of 2500 μg/kg, **6** produced six long-term survivors out of ten animals over a period of one year. In comparison treatment with the seco-CBI equivalent control, resulted in premature death due to toxicity in all animals at doses of between 100 and 500 μg/kg, and only one long-term survivor at a lower 60 μg/kg dose. Again this suggests a benefit from either hypoxia targeted or prolonged intercellular release of the free drug, most likely a combination of both. It was also noted that **6** was better tolerated at the injection site, than the seco-CBI equivalent.

Other ongoing attempts at hypoxia activated prodrugs of the seco-CBI alkylation subunit include replacement of the phenolic hydroxyl with a nitro group.[71-73] These compounds were envisioned to undergo reductive metabolism to their aniline counterparts, which in turn can undergo spirocyclisation to the active alkylating species through imine formation analogous to that of the phenol compounds. This approach has proved to be far from straight-forward. The initial simple nitro derivatives showed promising differential cyctotoxicity *in vitro*, but low potency *in vivo*. This was reasoned to result from suboptimal one electron reduction potential of the nitro group leading to insufficient bioreduction. Introduction of electron-withdrawing ring substituents placed *para* or *meta* to the carbon bearing the nitro group on the neighbouring ring were found to raise the reduction potential of the nitro group, however this did not have a consistent effect on hypoxia selective cytotoxicity. A clear correlation between increased reduction potential and hypoxic selectively was only apparent for a subset of electron-withdrawing groups. These were the sulfonamide and carboxamide substituents. The authors suggest this may result from specific interactions of these groups with the reductases responsible for reduction under hypoxic conditions, effectively suggesting that high reduction potential in compounds not containing these reductase specific interactions, could lead to increased oxic activation and reduced selectively. Despite inconsistencies in the structure activity trends of these analogues, a promising hypoxia selective lead was identified. This contained a *meta* SO_2NH_2 group. Further work has shown increased hypoxic selectivity by introduction of sulfonate spirocyclisation leaving groups in place of the usual halogens. Some of the best results were seen with analogues combining sterically demanding sulfonate leaving groups, neutral DNA binding subunits, and *meta* sulfonamide or carboxamide substituents (figure 1.15).

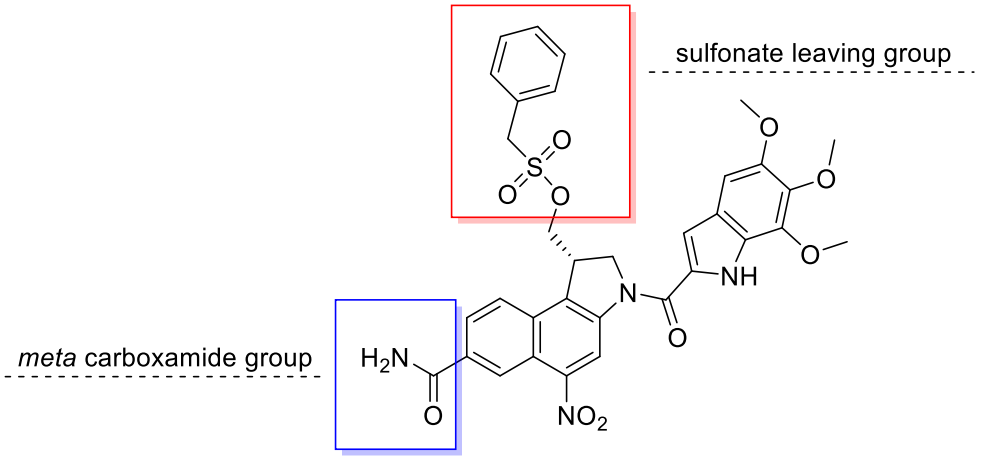
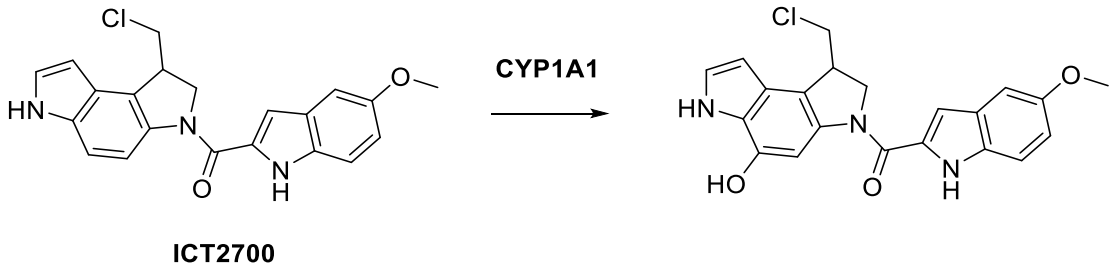


Figure 1.15 Annotated example of a nitro CBI prodrug structure.

Evidence of the over expression of particular isoforms of cyctochrome P450 oxidases in tumour cells, has led to a novel take on the inhibition of spirocyclisation prodrug strategy. Instead of masking this process by protection, the phenolic hydroxyl group was removed completely.[74, 75] The rationale being that the dehydroxylated CPI subunit (that of duocarmycin SA minus the C6 ester) would provide a substrate for cyctochrome P450 metabolism. It was envisioned that site specific oxidation by this enzyme system would reintroduce the hydroxyl group to reveal the active pharmacophore; a process hypothesised to mimic the evolution of this family of antibiotics.

The compound reported, named ICT2700 (figure 1.16), consists of the dehydroxylated seco-CPI alkylation subunit linked with the 5-methoxyindole DNA binding unit. It was shown to be 1000 fold less cytotoxic than its hydroxylated counterpart *in vitro*. The cellular distribution of the two compounds was shown to be comparable, thus confirming successful deactivation of the pharmacophore by removal of the hydroxyl group. Comparison of the cytotoxic effect of ICT2700 on CYP1A1 expressing and non-expressing CHO cells, showed successful differential activation of this compound in CYP1A1 positive cells, with an IC50 of <0.5 nM compared to >25000 nM in CYP1A1 negative cells.



ICT2700

Figure 1.16 Structure and activation of ICT2700 by CYP1A1 oxidation.

An alternative, but not mutually exclusive approach for improving the therapeutic window of the duocarmycins, is to add structural features which aid in the physical targeting of tumour cells. One such strategy involves the refinement of sequence selectivity, so as to preferentially bind in the minor groove at DNA sites which represent gene mutations specific to the malignant genome; a concept which will be discussed in chapter four. Another tactic is to separate the task of targeting and alkylation, and to consider the alkylation subunit of the duocarmycins as a ‘payload’ for a cell selective delivery vehicle.

The impressive recognition properties of antibodies have seen these highly engineerable proteins positioned as excellent candidates to serve as the delivery mechanism for cytotoxins. The principle is simple; multiple units of a cytotoxic compound are tethered to

an antibody which is engineered to recognise a tumour specific, or overexpressed antigen. On binding to the antigen, which is presented on the surface of the malignant cells, the antibody, together with its 'payload', is internalised *via* a process of receptor-mediated endocytosis. Upon internalisation, the cytotoxic compound is cleaved *via* a labile linker, or is released following lysosomal degradation of the antibody. [76]

Early generations of antibody drug conjugates (ADCs), proved disappointing in the clinic.[76] One reason for this was the common side effect of immunogenic responses to the mouse derived antibodies that were used.[77] This problem has largely been resolved through the development of fully humanised monoclonal antibodies, which retain only the complementarity determining region of the original murine protein (figure 1.17).[78]

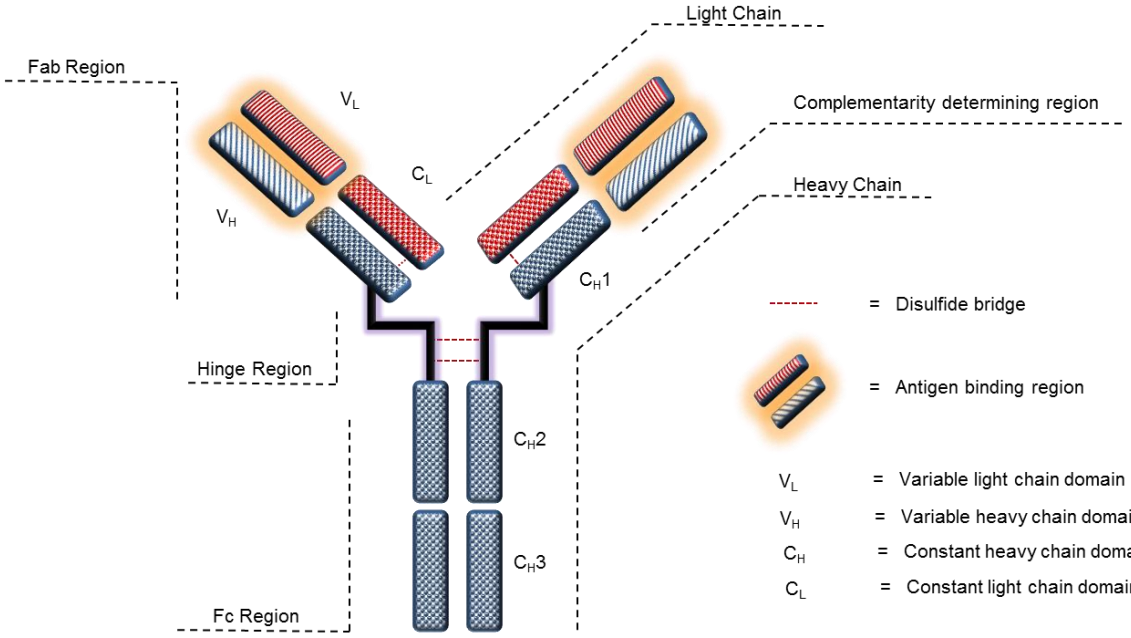


Figure 1.17 Schematic representation of the basic structure of a monoclonal antibody.[77]

A second problem was poor efficacy *in vivo*. ADCs can be very effective at selectively targeting cytotoxins to malignant cells, however the actual physical number of cytotoxic molecules which reach the interior of the targeted cell is thought to be small. The first generations of ADCs employed commonly prescribed antitumor agents, such as doxorubicin, as their payloads. These compounds simply did not possess the necessary potency to mount an effective antitumor response when delivered in the amount possible by antibodies.[76] Therefore subsequent generations of ADCs have focused on the delivery of ultrapotent cytotoxins, such as the enediyne antibiotics, maytansinoid and auristatin derived antimitotic agents, and of course the duocarmycins.[76]

One of the first ADCs reported which utilised the spirocyclopropylcyclohexadienone pharmacophore was anti-B4-DC1 (figure 1.18).[79] The 'payload' of anti-B4-DC1 was based on the CC-1065 analogue adozelesin, but saw the replacement of the CPI alkylation subunit with the more stable and easily synthesised seco-CBI structure. To allow conjugation to the antibody *via* disulfide exchange, the benzofuran group of adozelesin was also replaced by an indole unit terminally substituted with a reactive disulfide group.

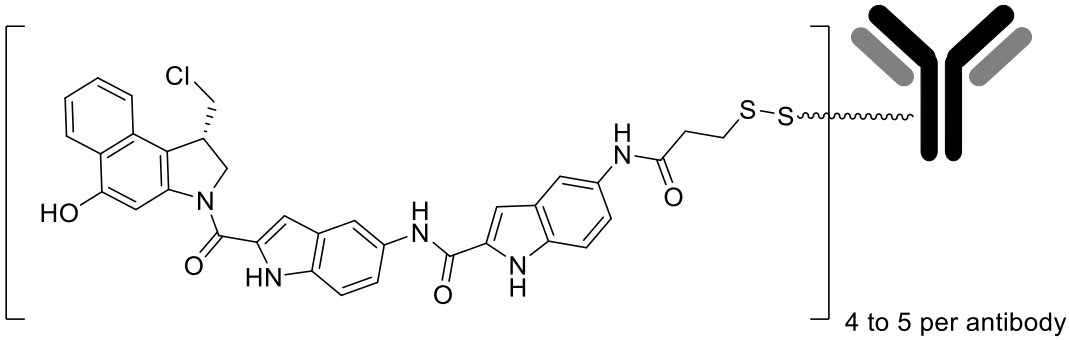


Figure 1.18 Structure of the anti-B4-DCI 'payload' and linker.

Anti-B4-DC1 demonstrated low picomolar activity against CD19-positive cell lines *in vitro*, and was effectivity nontoxic to antigen negative cells in the same assay. The selectivity of cytotoxicity was also demonstrated by the abolishment of activity when competing with an excess of unconjugated anti-B4 antibodies. Furthermore, the unconjugated payload was equally cytotoxic towards antigen positive and negative cell lines. Anti-B4-DC1 proved to be highly efficacious in murine human tumour models, significantly outperforming doxorubicin, cyclophosphamide, vincristine, and etoposide when dosed at their respective maximum tolerated levels.

Unfortunately, poor solubility of the anti-B4-DC1 'payload' in aqueous conjugation reaction solvents, and instability in physiological buffers, ultimately precluded its clinical development. However, recent efforts by the same laboratory, has seen the development of more soluble and stable phenolic phosphate prodrugs of the anti-B4-DC1 'payload'.[80] Interestingly, antibody conjugation of these prodrugs, led to a 250 fold increase in *in vitro* potency when compared to the unconjugated equivalents. This is suggestive of improved cellular uptake of the conjugates, and that passive diffusion of the free prodrugs is inhibited by the charged phosphate group. Again selective low picomolar activity against CD19-positive cell lines was demonstrated. Furthermore *in vitro* potency was only increased slightly by the co-incubation of phosphatases, demonstrating that the prodrugs are effectively dephosphorylated by endogenous enzymes. Conjugate huB4-SPP−DC4 (figure 1.19), showed promising antitumor activity in *in vivo* models, where the unconjugated prodrug was ineffective.[80]

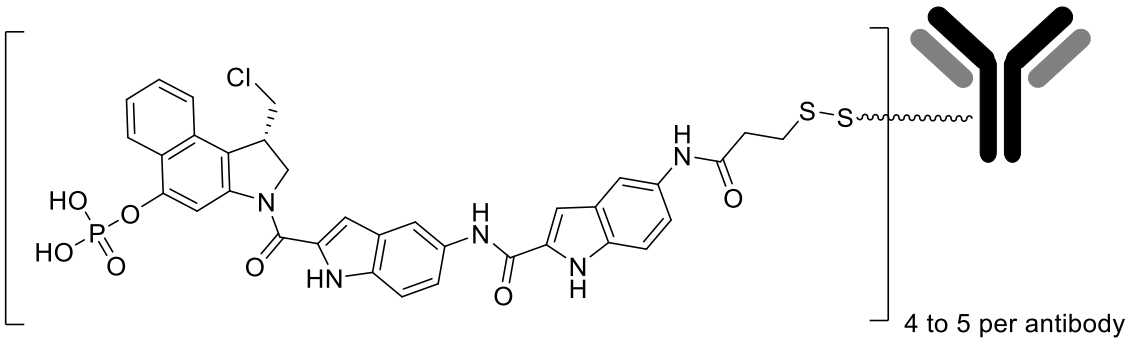


Figure 1.19 Structure of the huB4-SPP−DC4 'payload' and linker.

MDX-1203 (figure 1.20), in development by Bristol-Myers Squibb, represents another example of combining water solubilising phenolic protection of the CBI alkylation subunit with antibody conjugation.[81] Here, CBI is trapped in its seco form by a piperazino carbamate, and is tethered to a CD70 targeting antibody *via* a dipeptide linker which extends from substitution of a terminal aniline group. The *N*-terminus of the dipeptide possesses a maleimide group through which conjugation to thiol modified sidechains of antibody lysine residues form the attachment site.[82] This prevents the need to disrupt disulfide bonds between antibody cysteine residues, minimising disruption to protein structure. Following antigen specific internalisation, the peptide is cleaved, releasing the prodrug, which is ultimately activated by endogenous esterases. MDX-1203 has completed phase I clinical trials for renal cell carcinoma or non-Hodgkin's lymphoma.[83]

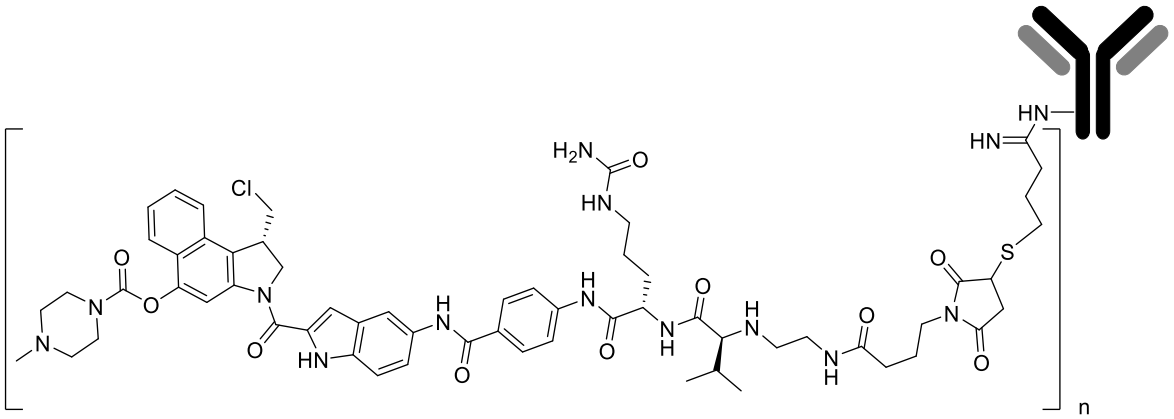


Figure 1.20 Structure of the MDX-1203 'payload' and linker.

SYD985 (figure 1.21) is a novel HER2-targeting ADC which has recently entered pre-clinical evaluation, and represents an alternative approach to antibody conjugation of the spirocyclopropylcyclohexadienone pharmacophore.[84, 85] In this strategy, the phenol group of the seco-alkylation subunit forms the attachment site of the linking group. In this way conjugation and seco-trapping phenol protection are combined functions of the same

structure. Linkage to the antibody is effected by conjugation of the linker's terminal maleimide group with cysteine thiols derived from intrachain disulphide bridges. The remaining linker structure consists of a dipeptide, and two well characterised self-eliminating spacers.[86, 87] Upon internalisation, the dipeptide is cleaved revealing a *p*-aminobenzyl alcohol from which the elimination cascade is initiated (scheme 1.5), ultimately releasing the free seco form of the alkylation subunit to undergo activating spontaneous spirocyclisation.

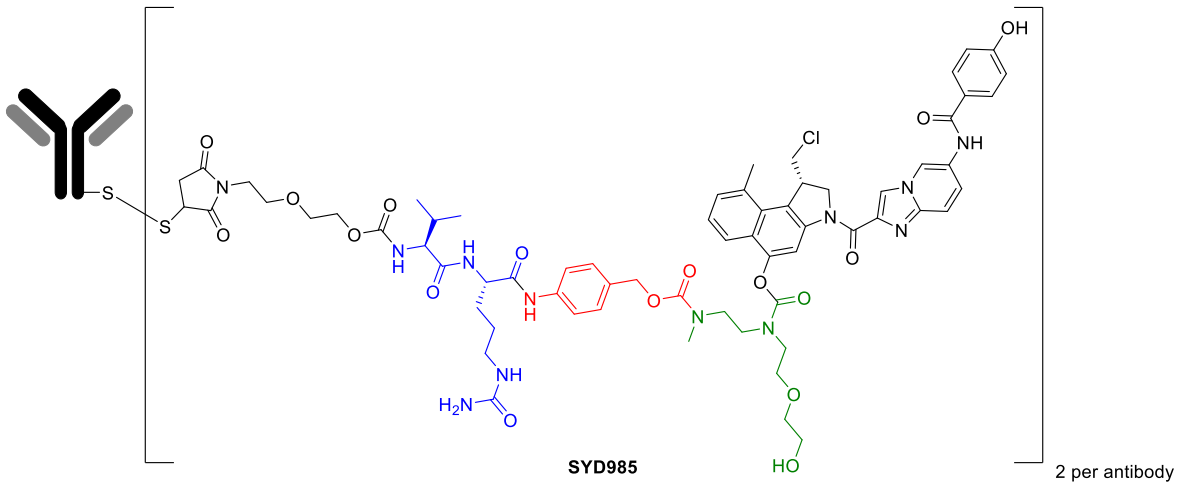
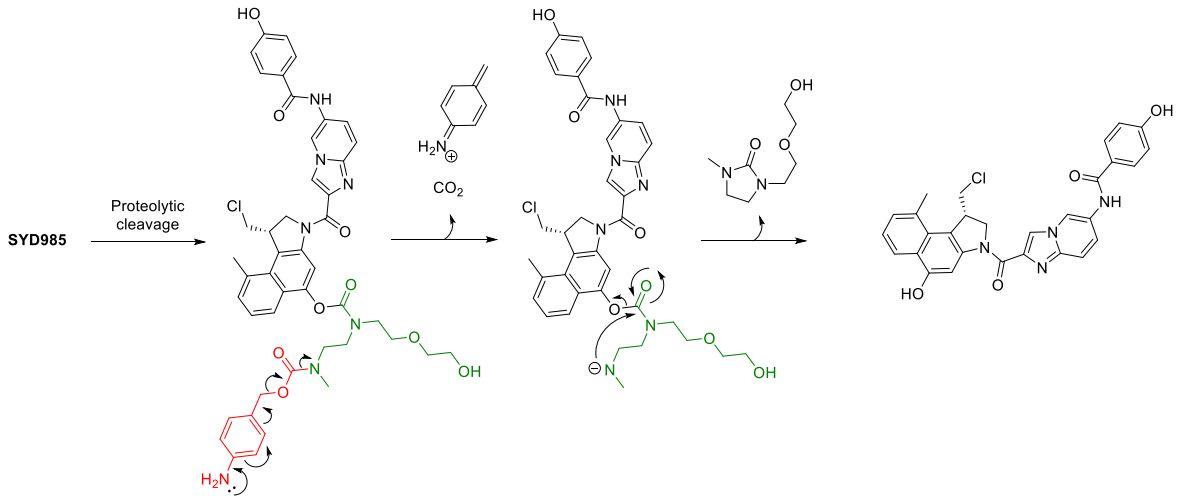


Figure 1.21 Colour coded structure of SYD985. Blue = dipeptide linker, Red = p-aminobenzyl alcohol spacer, Green = second spacer forming carbamate masking the phenol of the seco-alklyation subunit.



Scheme 1.5 Mechanism of self-elimination initiated by release of p-aminobenzyl alcohol.

To date it would appear that ADCs incorporating the spirocyclopropylcyclohexadienone pharmacophore, have been restricted to utilising derivatives of the CBI alkylation subunit, with linkage extending from either the phenol, or *N*-terminal substituents. However, the

amine side chains of antibody lysine residues provide the opportunity for conjugation through amide bonds. The terminal ester and amide functionality of the duocarmycin SA alkylation subunit would seem to afford excellent potential for incorporation into polyamide structures, including potentially short water solubilising peptides which could be applied as *C*-terminal linkers for the development of novel ADCs. Fmoc based solid phase synthesis is a convenient, and well established methodology for the construction of short peptides on a laboratory scale (see chapter 3). This thesis describes the conception and multigram synthesis of a duocarmycin SA alkylation subunit suitably substituted to serve as a 'building block' for Fmoc solid phase synthesis, and initial investigations in to its application.

2

Chapter Two

Synthesis of the duocarmycin alkylation subunit suitably substituted to serve as a 'building block' for Fmoc based solid phase synthesis.

2.1 Aims.

The work described in this chapter was carried out in order to address the following aims:

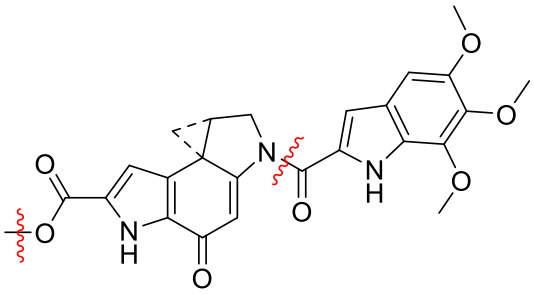
- To design an analogue of the duocarmycin SA alkylation subunit which is suitably substituted to allow incorporation of this functionality in to polyamide structures *via* Fmoc solid phase synthesis.

- To explore the synthesis of this target, and optimise a reliable route which could be practically conducted on the scale necessary to access a sufficient quantity of the final product to allow investigation of its use as a 'building block' for solid phase synthesis.

- To conduct the synthesis on a sufficient scale to access a multigram quantity of the 'building block'.

2.2 Design of the 'building block'.

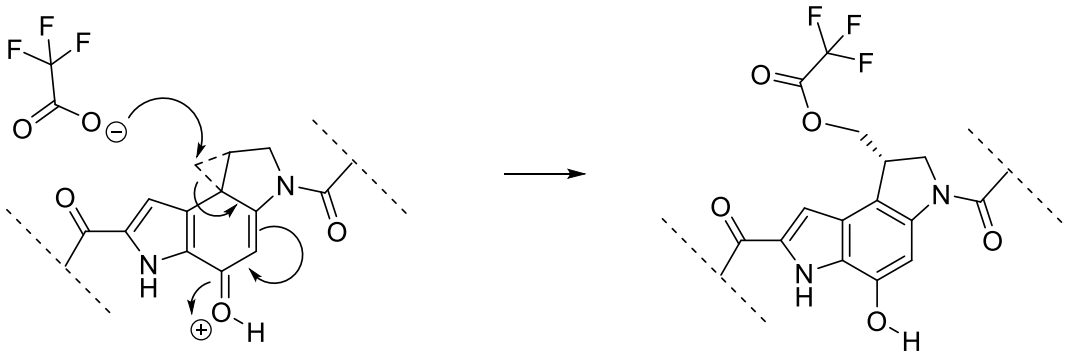
As discussed in the introduction, the alkylation subunit of duocarmycin SA is essentially an amino acid. The terminal ester and amide functionality conceal this, but clearly highlight its potential for incorporation into polyamide structures (Figure 2.1).



Duocarmycin SA

Figure 2.1 Disconnection of duocarmycin SA to reveal amino acid functionality.

Therefore it might be concluded that simply substituting the trimethoxyindole unit for Fmoc-protection of the indoline nitrogen, and hydrolysis of the ester to give the free carboxylic acid, would afford a suitable structure for application to solid phase synthesis. However, this would ignore several foreseeable issues.



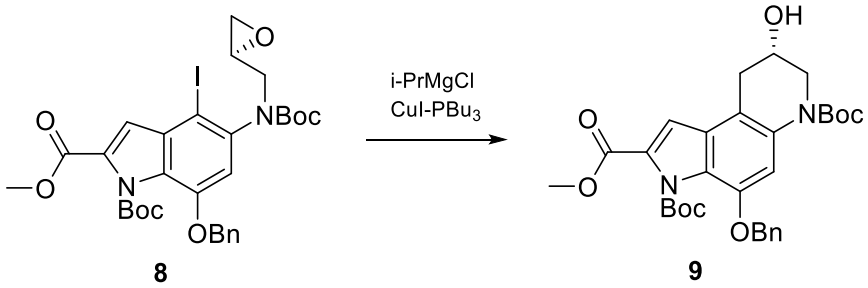
7

Scheme 2.1 Potential undesired by-product of resin cleavage by TFA.

Firstly it has been shown that amide bond formation at the indoline nitrogen is inhibited when the cyclopropane is in place.[88] Thus, such a building block would suffer from poor coupling efficiencies. The intact alkylation subunit introduces another problem. Commonly used resins in Fmoc based solid phase synthesis are cleaved under acidic conditions, typically with TFA.[89] Under such conditions it is likely that ring opening of the cyclopropane would be observed, and lead to the formation of undesired products such as **7** (scheme 2.1).

In addition to these synthetic concerns, use of the intact alkylation subunit represents a considerable safety hazard. It was envisioned that the 'building block' would eventually be synthesised on a multigram scale. Considering the formidable cytotoxicity of the duocarmycins, handling and storing large quantities of active compound safely represents a significant concern.

With these issues in mind, it was decided that a more suitable 'building block' would be one trapped in a seco form of the cyclopropane by protection of the phenol (a standard practice in duocarmycin synthesis). Different seco forms of the duocarmycin alkylation subunit are possible, and would be dictated by the chosen synthetic strategy (see scheme 2.2, and scheme 2.3).



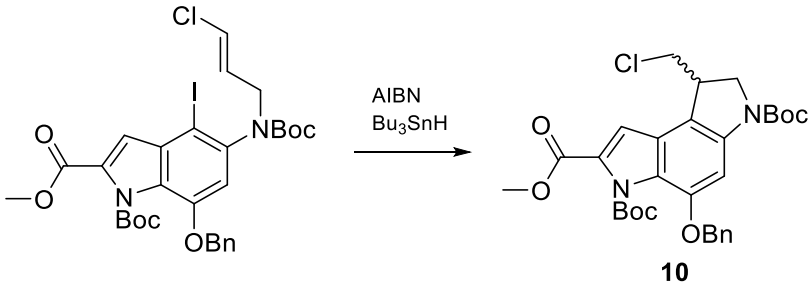
8 **9**

Scheme 2.2 Relevant step in Boger *et al*'s asymmetric synthesis of (+)-duocarmycin SA.

An elegant asymmetric route to the duocarmycin alkylation subunit has been reported by Boger *et al.*[90] The key step in this synthesis (Scheme 2.2) is the intramolecular epoxide opening of **8**. Metal-halogen exchange with i-PrMgCl affords the Grignard. Subsequent transmetalation with CuI-PBu_3 provides the cuprate, which in turn attacks the electrophilic epoxide at the least substituted position. The resulting seco duocarmycin alkylation subunit **9** is isolated with an impressive *ee* of 99 %. Although this strategy would remove the need for chiral resolution, the structure of **9** may not be ideal to form the core of the desired solid phase 'building block'. This is because the free secondary alcohol would need protecting to avoid acting as a competing nucleophile during couplings. Furthermore, spirocyclisation by displacement of the alcohol requires Mitsunobu activation. The chemistry could also prove challenging when performing the synthesis on a large scale.

An alternative to the epoxide opening is the radical 5-exo-trig cyclisation of a tethered vinyl chloride (Scheme 2.3), first reported by Patel *et al.*[91] in the synthesis of related heterocycles, and employed by Boger *et al.*[48, 49, 90, 92, 93] towards the racemic synthesis of various duocarmycins. The halide seco form of the duocarmycin alkylation subunit **10**, was deemed more suitable than **9** to provide the core structure of the solid phase building block. Not only does the chemistry appear more scalable, and the structure not contain a

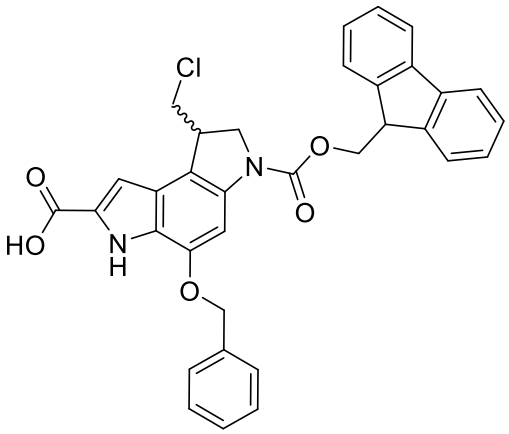
competing nucleophile, but the halide represents a better leaving group, and therefore spirocyclisation can occur spontaneously *in vivo* after removal of phenol protection. Indeed this has been shown not to be rate limiting, and halide seco duocarmcyins have comparable activity to their spirocyclised counterparts.[67, 94] Therefore this structure holds the advantage that compounds of biological interest would be obtained in fewer synthetic steps following cleavage from the solid phase.



10

Scheme 2.3 Racemic radical cyclisation.

To complete the design of the solid phase building block **11** (Figure 2.2) benzyl protection of the phenol was chosen. This protection strategy is commonly used during synthesis of duocarmycins. It is a robust group, unlikely to be lost unintentionally, and one which is orthogonal to the conditions used during solid phase synthesis.[95] This is an important consideration, as were the phenol protection to be lost under the basic conditions of Fmoc-deprotection, spirocyclisation could occur, and in turn this would reduce coupling efficiencies, and introduce the cleavage problems already discussed.

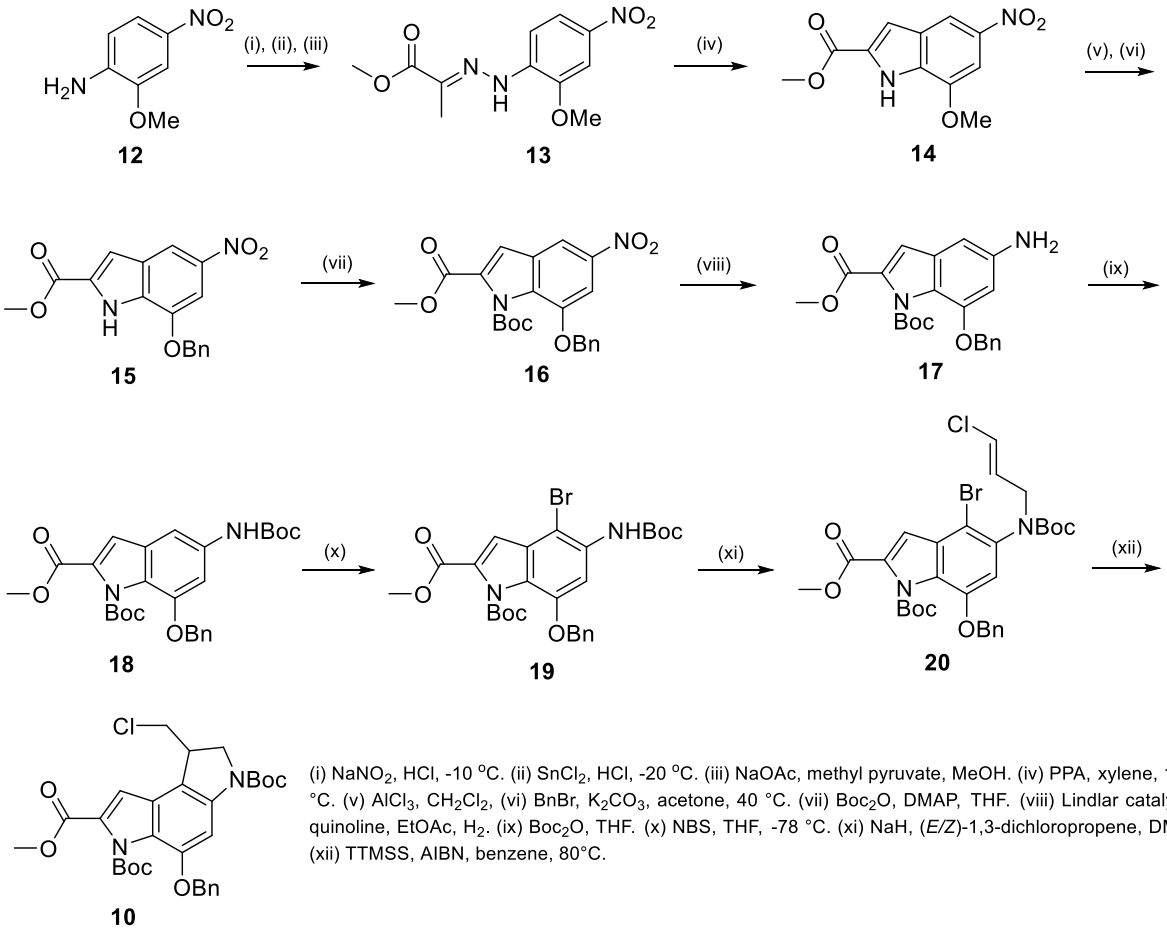


11

Figure 2.2 Structure of the desired solid phase 'building block'.

2.3 The pilot synthesis of 11.

It was envisioned that **11** would be easily accessible from the previously reported Boc-protected seco duocarmycin alkylation subunit **10** (scheme 2.4).[90, 96] This would be achieved by hydrolysis of the methyl ester, exhaustive Boc-removal, and regioselective protection of the indoline nitrogen with Fmoc-Cl.

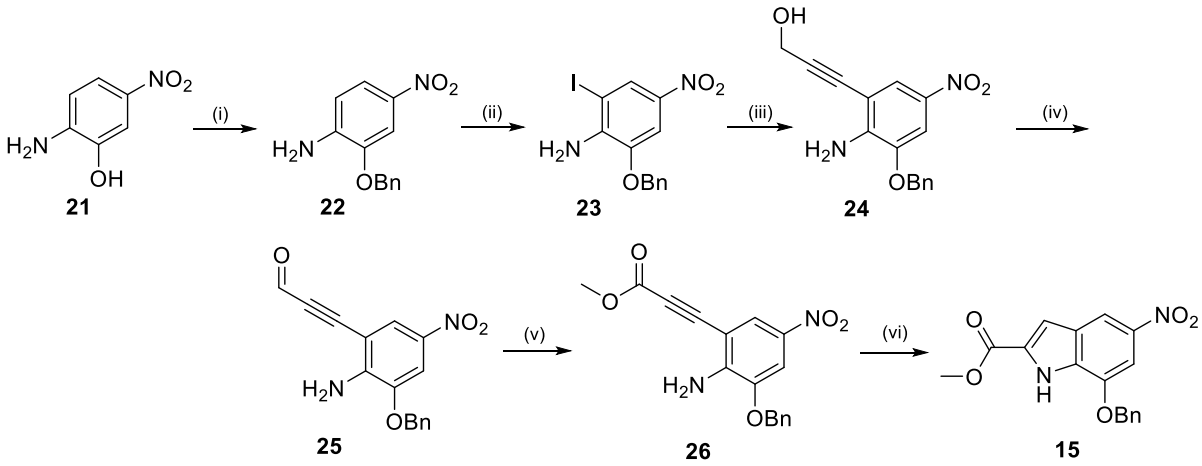


(i) $NaNO_2$, HCl, -10 °C. (ii) $SnCl_2$, HCl, -20 °C. (iii) NaOAc, methyl pyruvate, MeOH. (iv) PPA, xylene, 120 °C. (v) $AlCl_3$, CH_2Cl_2, (vi) BnBr, K_2CO_3, acetone, 40 °C. (vii) Boc_2O, DMAP, THF. (viii) Lindlar catalyst, quinoline, EtOAc, H_2. (ix) Boc_2O, THF. (x) NBS, THF, -78 °C. (xi) NaH, (*E/Z*)-1,3-dichloropropene, DMF. (xii) TTMSS, AIBN, benzene, 80°C.

Scheme 2.4 Tietze *et al.*'s synthesis of **10**.

One of the most concise routes to **10** is shown in scheme 2.4, and combines a Fischer indole synthesis to give key nitro indole intermediate **15**, with the radical 5-exo-trig cyclisation to complete the indoline ring.[96] However, the route contains an undesirable switching of protecting groups from methyl ether **14** to benzyl ether **15**. The authors explain this as unintentional. Their initial plan was to retain methyl ether protection throughout the synthesis, but that this proved incompatible with later steps. Previous efforts in our lab to repeat the Fischer indole reaction with benzyl ether protection in place from the start were unsuccessful. Furthermore attempts to repeat the published procedure

have produced unsatisfactory yields. Therefore an alternative strategy to access the nitro indole **15** was sought.

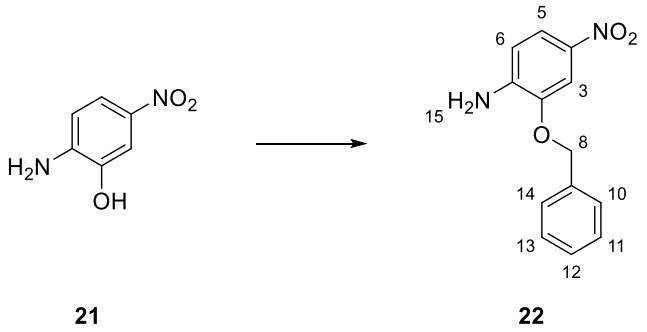


(i) BnBr, K_2CO_3, DMF. (ii) ICl, THF, 66 ^oC. (iii) $Pd(PPh_3)_2Cl_2$, CuI, propargyl alcohol, Et_3N, N_2. (iv) Oxalyl chloride, DMSO, Et_3N, THF, -78 ^oC. (v) NIS, K_2CO_3, MeOH. (vi) TBAF, THF, 66 ^oC.

Scheme 2.5 Unpublished route to indole **15**.

Personal communication with Dr Helen Sheldrake from the University of Bradford suggested that the currently unpublished route shown in scheme 2.5 was worth exploring. Here indole ring formation is achieved by a TBAF promoted annulation of the suitably substituted aniline **26**. The aniline is accessed by the Sonogashira coupling of propargyl alcohol to 2-iodoaniline **23** with two subsequent oxidation steps to give the desired ester. This route formed the basis of our initial strategy when beginning the pilot synthesis.

2.3.1 Introduction of the benzyl ether.



Scheme 2.6 Introduction of the benzyl ether.

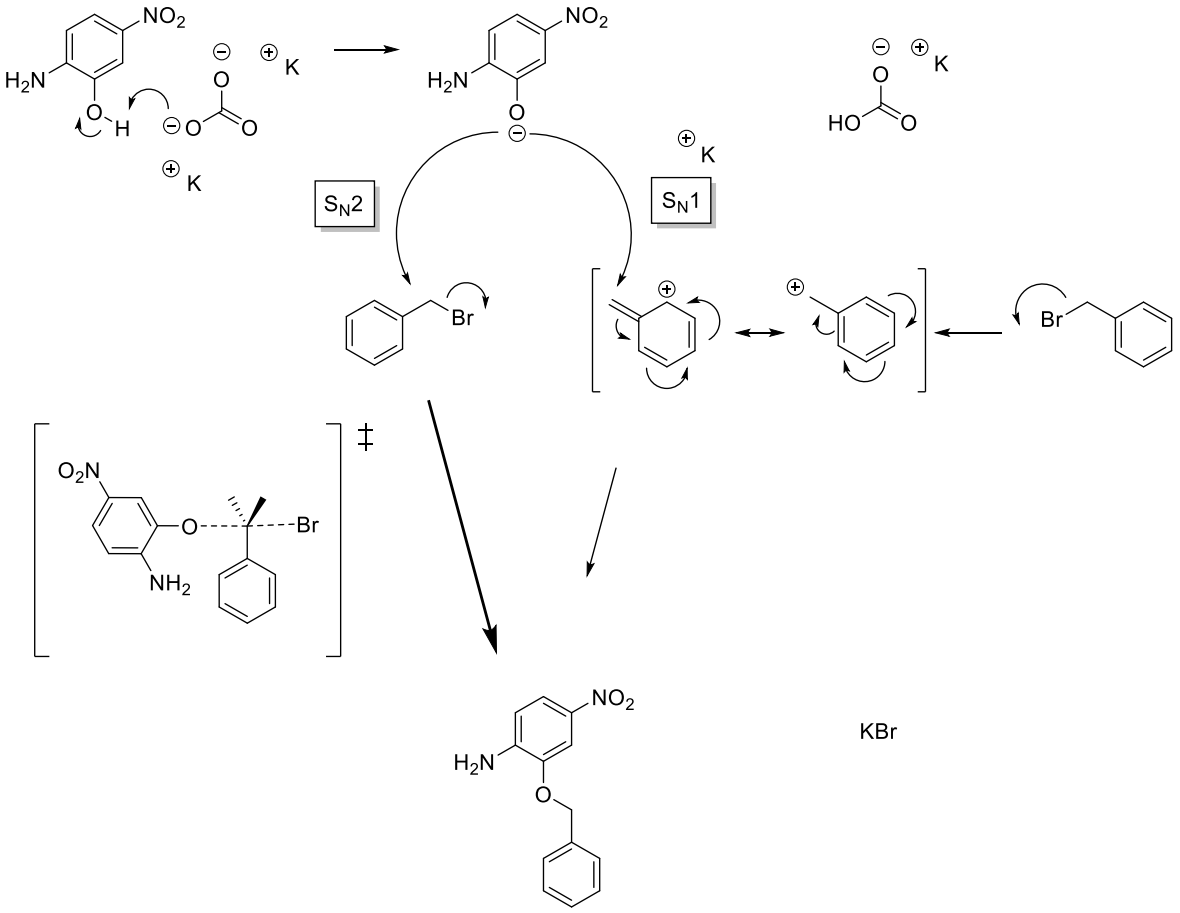
Cheap and readily available 2-amino-5-nitrophenol, was converted in to the benzyl ether **22** by treatment with K_2CO_3 and BnBr in DMF at room temperature (scheme 2.6). The product was isolated in sufficient purity to be used without further purification, by precipitation in a mixture of crushed ice and water, with near quantitative yield (99 %).

The product was confirmed by ^{1}H NMR (see scheme 2.6 for atom numbering). A distinctive aromatic splitting pattern characteristic of the substitution arrangement of the aniline ring was observed. The hydrogen at the C-6 position produced a doublet at 6.66 ppm, with a coupling constant of 8.7 Hz, characteristic of *ortho* coupling to its neighbouring hydrogen. The C-5 hydrogen itself gave rise to a doublet of doublets at 7.83 ppm with coupling constants of 8.7 Hz and 2.4 Hz, the smaller of these values representing additional *meta* coupling to the C-3 hydrogen. The reciprocal doublet of the C-3 hydrogen was observed at 7.77 ppm.

There is a dramatic difference in the chemical shift of the resonance signal for the C-6 hydrogen compared to that of the C-3 and C-5, with that of the C-6 being observed significantly upfield relative to the other two. This can be explained through the mesomeric effects of the ring substituents. The nitro group has an electron withdrawing effect which is most pronounced at *para* or *ortho* positions. Thus the C-5 and C-3 hydrogens experience greater deshielding relative to the C-6 hydrogen. This difference in shielding is amplified by the donating effect of the amine, which increases the electron density at the C-6 position.

The presence of the benzyl ether is confirmed by the appearance of a further 5 aromatic hydrogens signals overlapping to form a complex multiplet observed at 7.37-7.46 ppm, and a singlet at 5.15 ppm corresponding to the two aliphatic hydrogens at the C-8 position. The downfield position of the C-8 hydrogens, represents the electron withdrawing effect of the neighbouring oxygen atom and benzene ring. A broad singlet at 4.60 ppm, integrating for the two exchangeable hydrogens of the amine, confirms the regioselectivity of the reaction.

This reaction is a nucleophilic substitution (scheme 2.7). Deprotonation of the phenol by addition of the basic K_2CO_3 generates the more nucleophilic phenoxide. This species then most likely attacks the electrophilic centre of the polarised carbon bromine bond of BnBr directly by a concerted S_N2 pathway. Here the lone pair within the HOMO of the phenoxide approaches the antibonding orbital at the rear of the carbon bromine bond. As more electron density is donated to the antibonding orbital, the bromine carbon bond increasingly weakens. This results in a hypervalent trigonal bipyramidal carbon transition state, in which the carbon is partial bonded to both the oxygen and the bromine atom. BnBr is particularly reactive *via* this mechanism as the transition state is stabilised by orbital overlap of the neighbouring pi system.



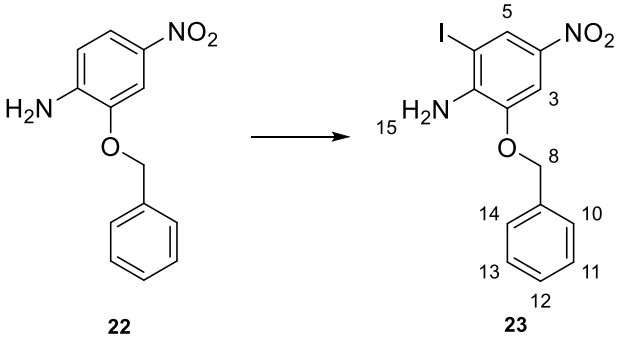
Scheme 2.7 Benzyl ether formation by nucleophilic substitution predominantly *via* the S_N2 mechanism.

It is possible for this reaction to also proceed *via* an S_N1 mechanism. Bromine is a good leaving group, and the resulting carbocation intermediate would be stabilised by resonance with the pi system of the benzene ring. It is likely that both pathways occur to some extent, but that the S_N2 route dominates. The fact that the reaction is performed in an aprotic polar solvent (DMF) also favours the S_N2 pathway. This is due to preferential solvation of the phenoxide's counter ion, increasing the electron density of oxygen, and hence reactivity. Polar protic solvents such as water would hydrogen bond with the oxygen and reduce its electron density. At the same time they would also be able to stabilise the carbocation of the S_N1 route, favouring this mechanism. As the reaction is achiral, which pathway is taken is of no consequence, because the same product is reached.

The amine group of 2-amino-5-nitrophenol has the potential to act as a competing nucleophile in this reaction. However its *para* position relative to the withdrawing nitro group, means it is heavily deactivated. A small more hydrophobic side product was observed by TLC, and this may represent the occurrence of some benzylation of the amine, most likely occurring after formation of the ether, by reaction of the product with

the small excess of BnBr. It is possible that the potential for a S_N1 reaction to occur, allows the benzylation of this deactivated nucleophile to proceed. This side product was present in very small quantities and did not precipitate with the product during the work up.

2.3.2 Iodination of 22



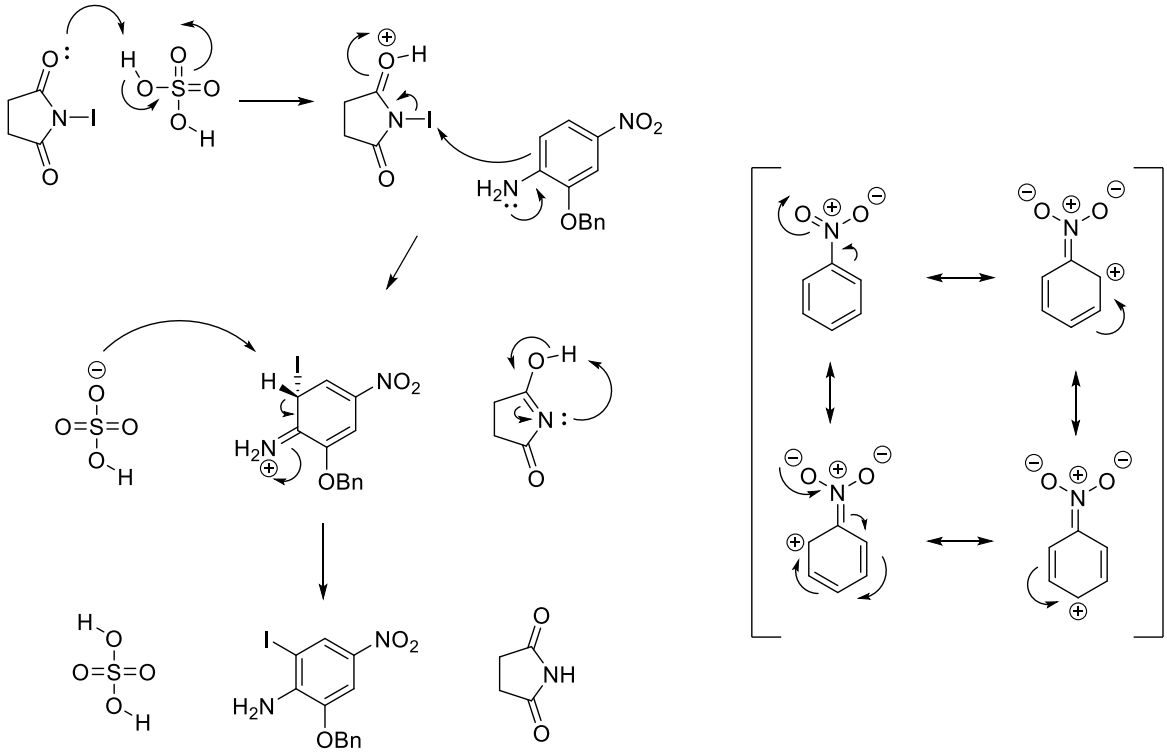
22 **23**

Scheme 2.8 Iodination of **22**.

Regioselective iodination of **22** provides the handle for subsequent Pd cross couplings. The unpublished route achieves this using ICl in THF at reflux. ICl is a troublesome solid to handle. It begins to melt at room temperature, and is sensitive to decomposition under atmospheric conditions, producing toxic and corrosive fumes.[97] With a view to the later large scale synthesis, it was decided to substitute ICl, for an alternative and more easily handled source of electrophilic iodine. Consequently iodination was achieved in this synthesis, by treatment of **22** with NIS, and catalytic H_2SO_4 in THF at room temperature overnight. Washing with sodium thiosulphate, and subsequent flash chromatography afforded the product in a yield of 88 %.

The regioselectivity of iodination was confirmed by ^1H NMR (see scheme 2.8 for numbering). Disappearance of the *ortho* coupled doublet at 6.66 ppm corresponding to the C-6 Hydrogen observed for **22** indicates that the desired position has been iodinated. Furthermore the C-5 hydrogen which was a doublet of doublets, is now observed as doublet with a coupling constant consistent with long range *meta* coupling to its C-3 ring partner. The chemical shifts of the C-3, and C-5 hydrogen signals have also changed in line with alteration in electronic character of the ring imposed by addition of the iodine substituent. The electronegative nature of the iodine atom has decreased the electron density of the ring. This is an inductive effect, and is most pronounced on the closest C-5 hydrogen. It has moved downfield from 7.83 to 8.29 ppm, due to decreased shielding. The more distant C-3 hydrogen has not moved to a more downfield position. This can be explained by mesomeric effects. Although the electronegativity of the iodine atom is

overall deactivating, its electronic configuration means it is able to donate electrons through resonance, and therefore increases the relative electron density at ring positions *ortho* and *para* to itself. Therefore the C-3 hydrogen benefits equally from this mesomeric shielding effect as the C-5 hydrogen, but is least effected by the inductive deshielding effect.

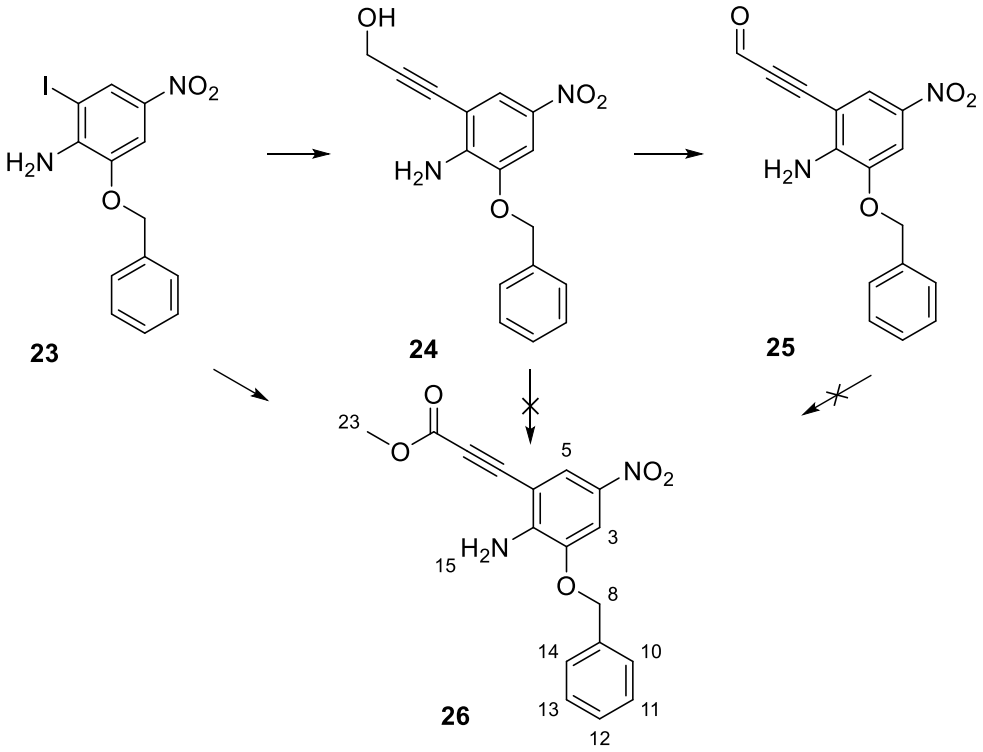


Scheme 2.9 Iodination of **22** by electrophillic aromatic substitution with NIS including the directing effect of the amine. In square brackets resonance structure of nitro benzene showing deactivation of the *ortho* and *para* positions.

The reaction is an electrophilic aromatic substitution, with NIS serving as a source of electrophilic iodine. The likely mechanism is depicted in scheme 2.9. NIS is activated by the presence of the acid catalyst. Protonation of the carbonyl leads to heterolytic cleavage of the nitrogen iodine bond to resolve the oxocarbenium ion. This can be as a result of direct nucleophilic attack of the aniline ring as depicted in scheme 2.9, or through spontaneous release of the iodine cation which is subsequently attacked by the aniline ring. It is also possible that the released iodine cation is first captured by the bisulfate with the resulting species acting as the actual source of electrophilic iodine.[98] Scheme 2.9 also shows the directing effect of the amine. Donation of the nitrogen's lone pair increases electron density at the *ortho* ring position, and hence the nucleophillicity of this position. Attack of the electrophilic iodine results in the resonance stabilised iminium ion. Abstraction of the hydrogen by the bisulfate restores aromaticity, resolving the iminium ion

to give the desired product **23**. This step also regenerates the acid catalyst. In addition to the directing effect of the amine, regioselectivity is also promoted by the electron-withdrawing effect of the nitro group which deactivates the remaining unsubstituted positions of the aniline ring. Resonance structures of nitrobenzene which show the deactivation of the *ortho* and *para* positions are shown in brackets in scheme 2.9.

2.3.3 Introduction of the alkyne.

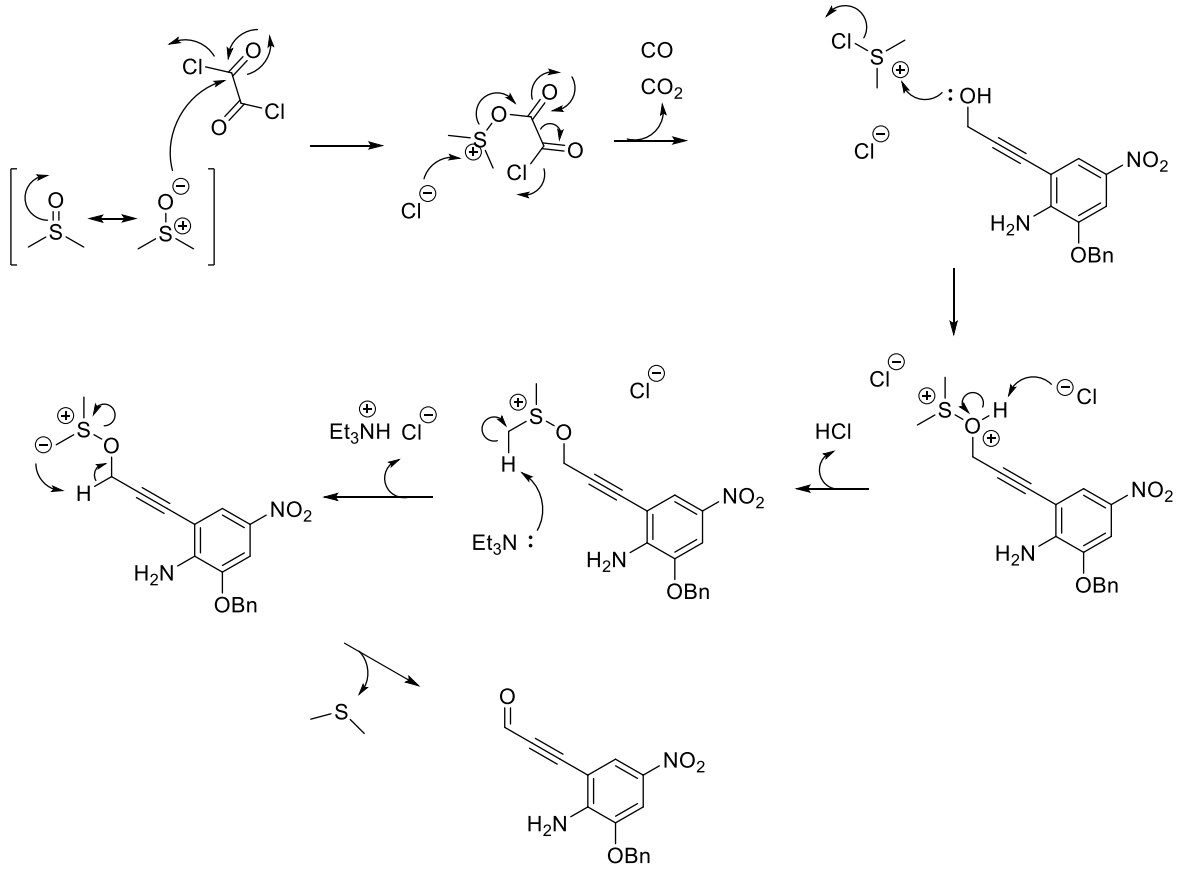


Scheme 2.10 Introduction of the alkyne.

With the iodine in place it provided a handle to investigate Pd cross couplings for introduction of the alkyne. Initially it was decided to continue to follow the unpublished route (scheme 2.5). In this route propargyl alcohol is introduced by Sonogashira coupling,[99] and subsequently oxidised in multiple steps to give the desired ester.

The Sonogashira coupling proceeded without issue. **23** was treated with $Pd(PPh_3)_2Cl_2$ (5 mol %), CuI (10 mol %), and propargyl alcohol (1.1 equiv.) stirring overnight, at room temperature, as a suspension in Et_3N, under N_2. Flash chromatography afforded the desired alkyne **24** with a yield of 93 %. However this route had to be abandoned as oxidation of the alcohol to the target ester could not be repeated.

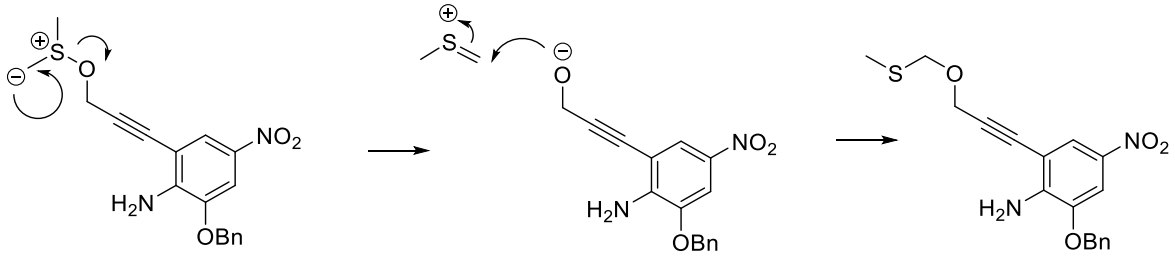
The first of these oxidations did work but at a very low yield of only 31 %. This step was a Swern oxidation[100] and proceeds *via* the mechanism shown in scheme 2.11.[101] DMSO reacts with the electrophilic oxalyl chloride forming a short lived intermediate which decomposes to the reactive dimethylchlorosulfonium salt, with the release of carbon dioxide and carbon monoxide. Attack of the alcohol yields the alkoxysulfonium salt. Subsequent treatment with Et_3N leads to deprotonation at one of the methyl groups forming the sulfur yilde. This species then decomposes *via* an intramolecular β-elimination to give the desired aldehyde and dimethylsulfide as a side product.



Scheme 2.11 Mechanism of the Swern oxidation.

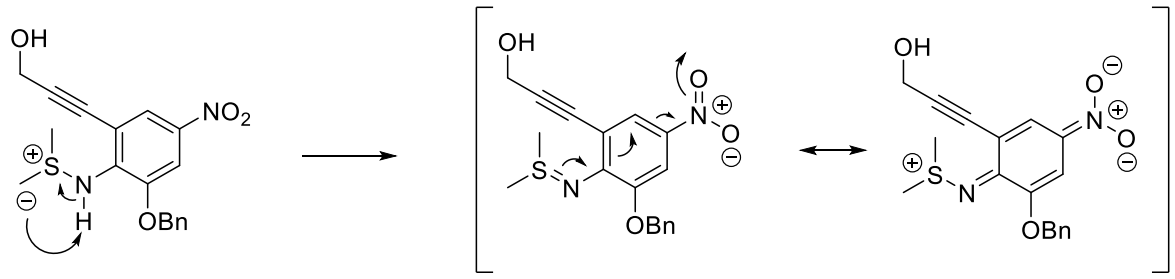
At first, TLC analysis suggested that the low yield was a result of the incomplete reaction of the starting material. The Swern oxidation is a sensitive reaction and can be influenced by the timing of addition of each reagent, and also by control of temperature. The reaction is typically conducted at -78 ^{o}C. Cryogenic operating conditions are required for several reasons. Firstly the reaction between DMSO and oxalyl chloride is extremely exothermic, and dangerous at room temperature. The dimethylchlorosulfonium salt is also unstable at elevated temperatures. Furthermore an alternative attack of the sulfur ylide can lead to a

Pummerer type rearrangement giving rise to mixed thioacetal side products if the temperature is not properly controlled (scheme 2.12).[102]



Scheme 2.12 Pummerer rearrangement type side reaction.

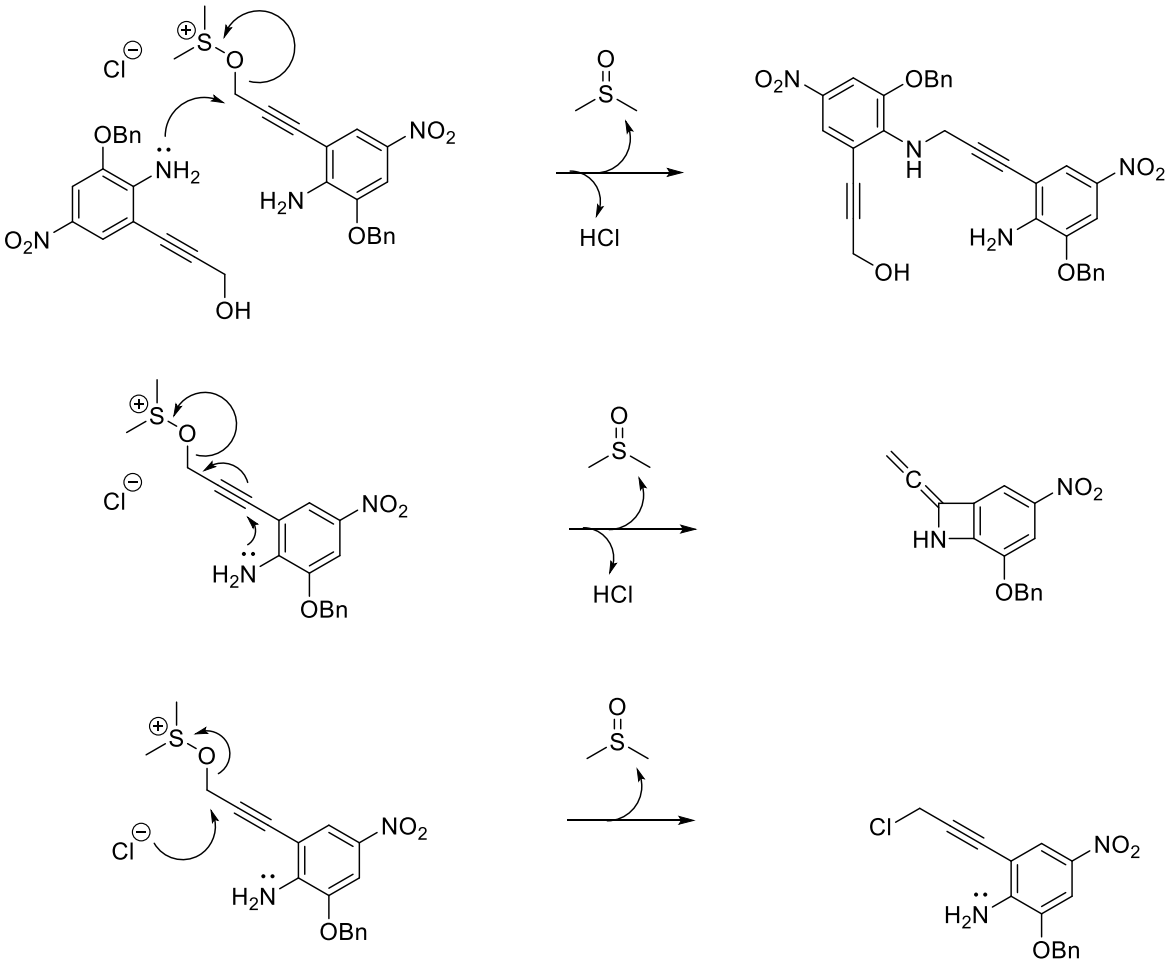
It is possible that addition of DMSO to the oxalyl chloride in THF was not slow enough and thus temperatures rose leading to some decomposition of the dimethylchlorosulfonium salt before it could react with the alcohol. Alternatively the delay of 20 mins prior to addition of the alcohol may have been too long. The activation of DMSO with oxalyl chloride is extremely facile even at -78 ^{o}C, thus addition of the alcohol after only 5 mins or at the cessation of gas evolution, could have been more appropriate.



Scheme 2.13 Possible but unlikely side product of Swern oxidation.

However, closer inspection of the TLC, revealed that the suspected unreacted starting material may have been a side product, as although it co-eluted with the starting material it produced a different colour spot under visualisation by staining with vanillin. This could be as a result of the thioacetal formation as already discussed (scheme 2.12) if the temperature was not sufficiently controlled.

Swern oxidation conditions have the potential to oxidase amines into imines.[103] However, the amine in this compound is an aniline. Not only is it deactivated by the *para* nitro group, but there is no hydrogen available for the necessary β-elimination. It is perhaps possible for the resulting sulfur ylide to abstract the hydrogen from the nitrogen itself as depicted in scheme 2.13, but this seems unlikely; although the resulting product could be stabilized by resonance with the nitro group. Indeed it is also conceivable for the Pummerer type rearrangement to occur here.

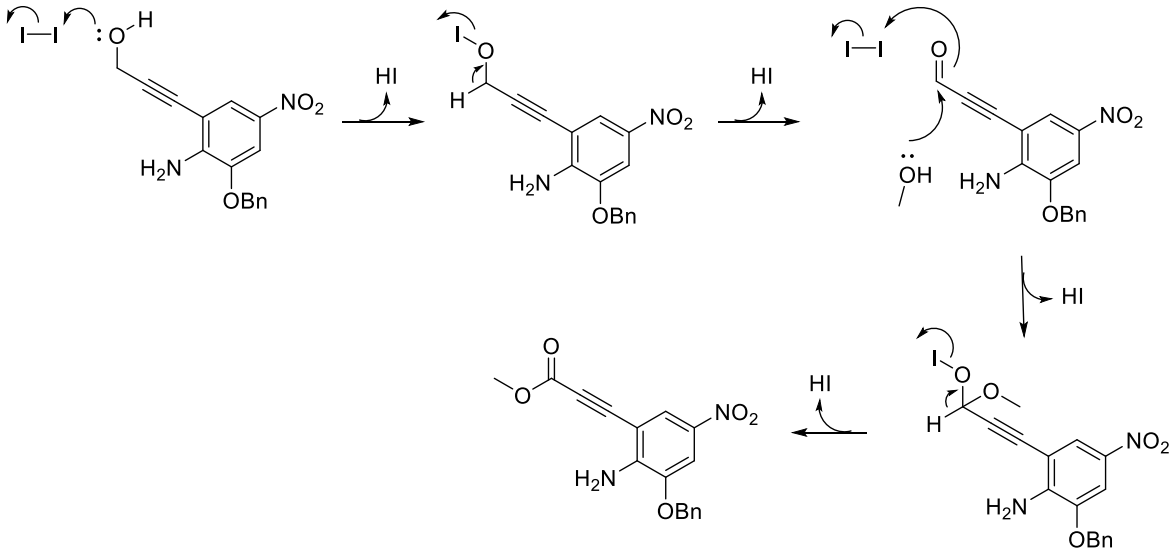


Scheme 2.14 Potential but unlikely displacement reactions.

Another timing consideration is the delay between addition of the alcohol and addition of the base. The alkoxysulfonium salt which exists prior to addition of the base is a potential leaving group, which could promote displacement of the alcohol (Scheme 2.14). Here the aniline could act as nucleophile. However this would have to be intermolecular, or involve participation of the alkyne and both seem unlikely. Participation of the alkyne would lead to a four membered heterocyclic ring and a cumulated diene so seems very unfavoured. Another possible displacement could come from the chloride counter ion of the alkoxysulfonium salt itself, leading to chlorination.

As this route was abandoned, the reaction was not repeated to isolate the side product and determine at which stage the reaction was failing. In retrospect the complicated operating conditions, and the notoriously malodourous dimethyl sulfide side product, would have proved troublesome on scale up anyway.

Nevertheless, the small amount of product isolated was reacted on through the planned route, but was unsuccessful. This reaction was a direct oxidation of the aldehyde to the desired ester using NIS in MeOH with K_2CO_3. Starting material was consumed but led to multiple products and the desired ester could not be isolated. To circumvent the Swern oxidation a number of unsuccessful attempts were also made to oxidise the alcohol directly to the ester by treatment with molecular iodine at reflux in MeOH with K_2CO_3.[104] The likely mechanism of such reactions is shown in scheme 2.15.



Scheme 2.15 Likely mechanism for direct oxidation of the alcohol to the ester by molecular iodine.

As with the NIS oxidation of the aldehyde, this reaction was very unselective and led to multiple products preventing isolation of the desired ester. The mechanism involves the formation of unstable hypoiodite species. Alternative homolytic cleavage of these intermediates would result in alkyoxy radicals which could be responsible for the undesired side reactions. In retrospect conducting the reaction in the dark and under anaerobic conditions would help to reduce this. The mechanism highlights another issue. The aldehyde intermediate is stabilized by conjugation with the alkyne, and this must be broken by subsequent attack of the MeOH to form the hemiacetal hypoiodite species. This probably represents the rate limiting step in the reaction. Conjugation is restored by formation of the ester. There is also the potential for competitive oxidation of the solvent. This concern meant the reaction was conducted at high concentration. This proved

incompatible with refluxing, and the long reaction times required, leading to formation of a slurry overnight. An attempt was made to substitute MeOH for 2,2,2-trifluroethanol.[104] It was expected that 2,2,2-trifluroethanol would be less susceptible to oxidation, and therefore allow the reaction to be conducted at lower a concentration. However this reaction also did not provide the desired product.

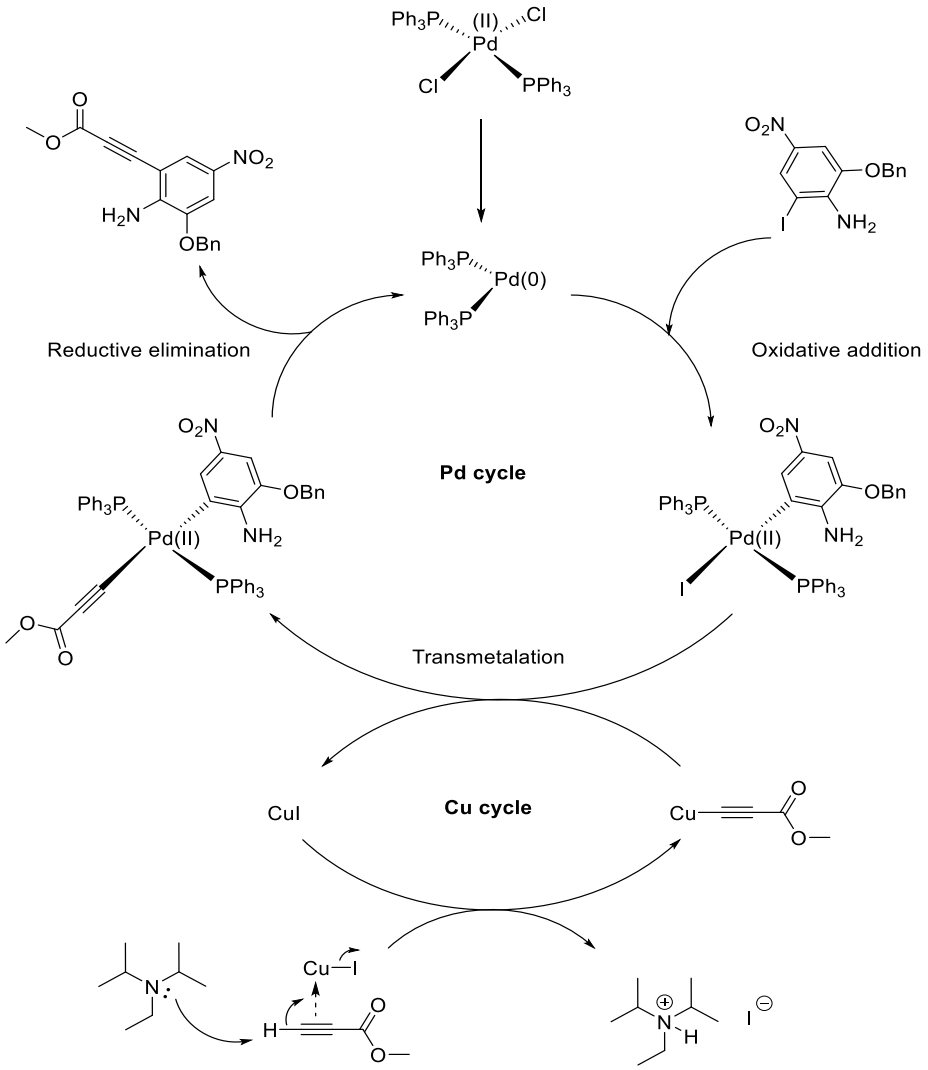
With the failure to find a reliable method to oxidase the propargyl alcohol, attention turned to the possibility of introducing the ester directly by coupling of methyl propiolate. This was initially attempted *via* the same Sonogashira conditions used to couple the alcohol. However it soon became apparent that using Et_3N as the solvent would not be successful. Dilution of the methyl propiolate in Et_3N prior to addition to the reaction mixture, led to the instantaneous formation of a black precipitate. This was most likely the result of a Michael type addition of the nucleophilic Et_3N at the terminal alkyne of the methyl propiolate.

An alternative protocol was trialled using an inorganic base (K_2CO_3) and refluxing in THF.[105] However no reaction was observed. This was reasoned to be a potential consequence of poor solubility of the base. Addition of a few drops of water as a co-solvent had no effect. As a result the reaction was repeated in DMF. The change was dramatic leading to full consumption of the starting material. Unfortunately the reaction produced multiple products by TLC with none appearing to dominate.

It was decided to repeat the reaction again in DMF but substitute DIPEA for K_2CO_3. It was anticipated that DIPEA, being a more hindered and therefore less nucleophilic organic base, would not undergo the troublesome Michael addition suspected with Et_3N, particularly when only used at 4 *equiv*. as opposed to the large excess available if used as the solvent. As a preliminary test methyl propiolate was added to a large excess of DIPEA. Indeed the instantaneous reaction seen with Et_3N was not apparent, and slow discoloration was only observed after about an hour. The substitution of DIPEA for K_2CO_3 turned out to be the key to the success of this coupling, and the desired ester was afforded with a yield of 79 % after flash chromatography. The reaction required 4 *equiv.* of methyl propiolate; attempts to reduce this led to incomplete couplings.

Appearance in the ^1H NMR spectra of a characteristic singlet at 3.86 ppm resulting from the methyl group of the ester confirmed the successful coupling of methyl propiolate. The doublet for the aromatic *meta* coupled C-5 hydrogen also shifted upfield from 8.29 ppm to 8.06 ppm (see scheme 2.10 for numbering). This is consistent with it no longer being so deshielded by the inductive withdrawing effect of the iodine atom. The ^{13}C NMR also helped to confirm the correct product with the appearance of four additional carbon

signals. A DEPT-edited HSQC experiment also showed that only one of these new carbon signals was correlated with protons, and that this was the CH_3 of the methyl ester.



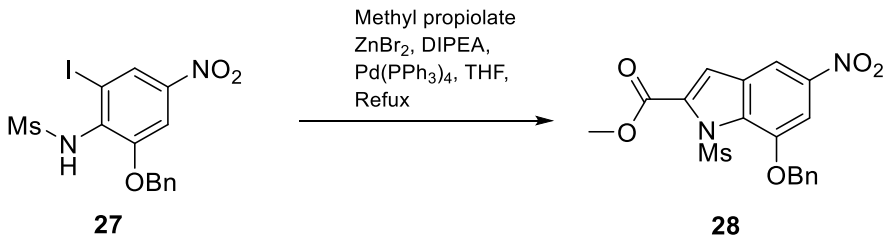
Scheme 2.16 Mechanism of Sonogashira coupling.

The exact mechanism of the Sonogashira coupling is not fully understood (scheme 2.16).[106] It is believed to involve two independent catalytic cycles: that of the Pd and that of the Cu co-catalyst. Although the protocol employed in this synthesis uses $Pd(PPh_3)_2Cl_2$, a Pd(II) species, it is actually a Pd(0) complex formed *in situ* that is believed to be the active catalyst. It is suggested that this can be formed from the addition and subsequent reductive elimination of two of the acetylene molecules, however there is also speculation surrounding involvement of the amine in this reduction. This might explain why DIPEA proved more effective than K_2CO_3. The first step of the Pd cycle is the oxidative addition of the aryl halide and this is believed to be rate limiting. Transmetalation of the methyl propiolate from the Cu cycle follows. The resulting Pd complex then undergoes *cis*/*trans*-

isomerization, followed by reductive elimination to give the final alkyne and regenerate the Pd(0) complex. The formation of the copper acetylene, required for this transmetalation, is promoted by deprotonation of methyl propiolate, which is catalysed by π-coordination of the Cu. Transmetalation regenerates the CuI.

Unfortunately the yield of the Sonogashira coupling dropped considerable on modest attempts to scale up the reaction, and although this ultimately provided enough material to complete the pilot synthesis, it was decided to explore alternative Pd cross couplings with a view to later large scale synthesis.

Electron deficient alkynes like methyl propiolate are known to be difficult substrates for Sonogashira couplings, and standard conditions can often result in low yielding reactions.[107] A good alternative in such cases is the Negishi coupling.[108] Both reactions essentially employ the same Pd chemistry, but whereas the Sonogashira methodology utilizes a copper co-catalyst, transmetalation in the Negishi coupling is effected directly from an organozinc compound; typically in a heteroleptic form, such as R-Zn-X, where X is Cl, or Br. In the standard Negishi coupling this species is formed prior to the coupling reaction itself. However, in recent years Negishi *et al.* have reported successful Pd-catalysed alkynylations *via* the *in situ* formation of the necessary alkynylzinc species, by treatment with $ZnBr_2$, and Et_3N.[109]



27 **28**

Scheme 2.17 Hiroya *et al*.'s Negishi coupling and *in situ* indole ring closure.

This chemistry has since been utilised by Hiroya *et al*. towards the total synthesis of duocarmycin SA, (scheme 2.17).[110] They found that when the mono-mesylated iodoaniline **27** was subjected to Negishi coupling conditions, the indole was formed *in situ*. Previous group members have tried to repeat this transformation but have found the reported yields of upwards of 56 % difficult to replicate. In addition they report that the indole was difficult to separate from uncyclised product. As following this route would also introduce two additional synthetic steps (dimesylation and mono-demesylation), it was decided to explore subjecting the already synthesised unprotected iodoaniline **23** directly to similar Nesghi coupling conditions.

Here **23** was treated with $Pd(PPh_3)_2Cl_2$ (5 mol %), $ZnBr_2$ (4 equiv), methyl propiolate (4 *equiv*.), and DIPEA (4 *equiv.*), in THF at reflux under N_2, overnight. The use of THF over DMF led to a more convenient work up, as the increased volatility of this solvent allowed the reaction mixture to be easily concentrated by rotary evaporation, and subjected directly to flash chromatography. The desired coupled product **26** was isolated in a good yield of 79 %. No evidence of cyclisation was observed. TLC analysis of the crude showed only one UV active product and a baseline spot.

It is unclear why under our Negishi coupling conditions no indole formation is observed. There are several differences between our reaction and that of Hiroya *et al.* Perhaps most notable is the absence of the mesylate group. It is therefore possible that this protecting group is necessary for ring closure to be promoted, but it is not obvious as to why this should be the case. If the mechanism of ring closure requires some form of nucleophilic attack of the aniline nitrogen, then the presence of a withdrawing mesylate group might be seen to hinder rather than aid this process. However, Hiroya *et al.* claim this reaction is catalyzed by a Pd species. They were not able to identify the active catalyst, but by probing the ring closure in separate reactions to the coupling, they did ascertain that both the $Pd(PPh_3)_4$, and methyl propiolate were necessary, and that although not essential $ZnBr_2$ did improve the efficiency of the reaction.[111]

They therefore speculate that a catalytic species formed from $Pd(PPh_3)_4$, and methyl propiolate, the formation of which is accelerated by the presence of $ZnBr_2$ is responsible for this annulation. If this is the case then the presence of the mesylate group may provide a handle for oxidative addition. However, this would seem inconsistent with the final product which retains this protecting group.

Our conditions also differ from Hiroya *et al.* in the ratio of reagents used. For their sequential coupling-cyclisation reactions Hiroya *et al.* typically employ 2 *equiv.* of methyl propiolate, 3 *equiv.* of $ZnBr_2$, and 6 equiv. of DIPEA. As a legacy from our previous Sonogashira conditions, which worked best with 4 *equiv.* of methyl propiolate, 4 *equiv.* of methyl propiolate, $ZnBr_2$, and DIPEA were used in our Negishi coupling. This was in an attempt to provide 4 *equiv.* of the alkynylzinc species for transmetalation. It is difficult to reason why this would have a deleterious effect the formation of the cyclisation catalyst. For example it seems unlikely that an excess of $ZnBr_2$ over methyl propiolate, or an excess of DIPEA over both is required.

The final difference is the starting form of the Pd species. Hiroya *et al.* used $Pd(PPh_3)_4$ while our conditions employ $Pd(PPh_3)_2Cl_2$. Although Hiroya *et al.* state that the PPh_3 ligand is required for cyclisation, $Pd(PPh_3)_2Cl_2$ is likely to form the same catalytic species,

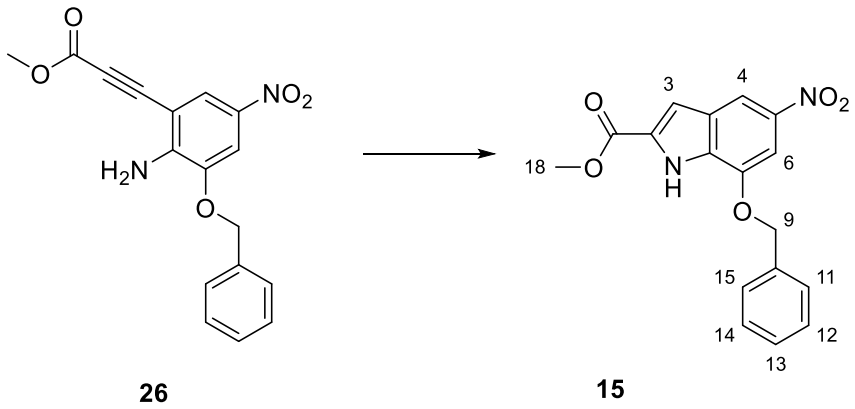
unless the additional two *equiv.* of this ligand are the important factor. However this speculation can be ruled out, as our Negishi conditions were later trialled with $Pd(PPh_3)_4$, and still no indole formation was observed.

Whatever the reason, it would seem to be an advantage to our purpose that no cyclisation was seen. This is because Hiroya *et al*.'s conditions seem to lead to a mixture of the indole and the uncyclised product, and a separate annulation step is preferable to a difficult chromatography separation on scale. Nevertheless, a number of unsuccessful attempts were made to promote cyclisation during the Negishi coupling.

Firstly, it was reasoned that increased temperature might drive ring closure. Therefore the reaction was repeated in toluene at reflux. However this only served to produce faster coupling at significantly reduced yield.

As treatment of the coupled product **26**, with TBAF would later be used to successfully form the indole, the use of TBAF over DIPEA as the base in the Negishi coupling was trialed. However, although the starting material was consumed this did not lead to formation of either the coupled product, or the indole. The product was not isolated in sufficient quantity or purity to make a confident assignment, but the presence of alkene signals in the ^1H NMR lead to speculation that TBAF may have in fact promoted an aza Michael type addition of the aniline to methyl propiolate.

2.3.4 Annulation to give the indole.



Scheme 2.18 Annulation to give the indole.

With the success of the direct Pd cross coupling of methyl propiolate, a reliable and significantly shortened route to **26** had been secured. Therefore attention turned to the formation of the indole as a separate step. The successful cyclisation of 2-ethynylaniline derivatives using TBAF has been reported by Sakamoto *et al.*[112] This chemistry is employed in the unpublished duocarmycin synthesis, and was repeated here to promote

the desired annulation. **26** was treated with a stoichiometric excess of TBAF at reflux in THF. The reaction was quick and selective, producing only one new UV active product and a baseline impurity. Sadly, this step appears to be intrinsically low yielding, with the best yield achieved being only 58 % after flash chromatography.

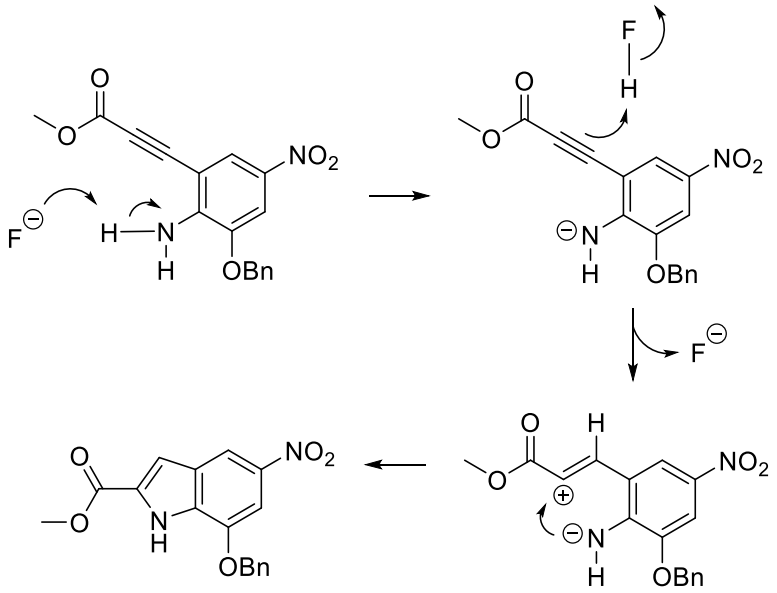
Successful annulation was confirmed by the product retaining the same mass, but producing a different ^{1}H NMR spectra. The changes were characteristic of the desired indole structure. Perhaps most notable is the disappearance of the aniline hydrogen signal at 5.32 ppm, which is now replaced by a broad singlet at 9.32 ppm integrating for the single hydrogen of the indole nitrogen. Its downfield position reflects the fact that the nitrogen atom's lone pair is now formally contributing to the aromaticity of the indole.

The aromatic hydrogens of the benzyl group still appear as a complex multiplet, but 3 rather than 2 additional aromatic signals are now apparent, this resulting from the new aromatic hydrogen signal at the C-3 position (see scheme 2.18 for numbering). The C-3, C-4, and C-6, aromatic hydrogen signals all appear as narrow doublets with coupling constants of around 2 Hz. This reflects long range coupling through the π system. The C-4 hydrogen which is coupled to both the C-3, and C-6 hydrogen does not appear as a doublet of doublets, this probably simply reflects the similarity in the coupling constant of both long range couplings.

The new C-3 hydrogen signal adopts the most upfield position at 7.35 ppm. This is consistent with this being the most nucleophilic position of the indole ring, receiving a formal negative charge in resonance with the indole nitrogen's lone pair.

Finally the C-18 methyl group hydrogen signal of the ester has shifted slightly downfield from 3.86 ppm, to 3.96 ppm. This most likely reflects a reduction in electron density of the ester carbonyl which is now conjugated with the larger aromatic indole system, as opposed to the previous alkyne.

How TBAF promotes the ring closure of **26** to the desired indole **15**, is not clear. Commercially available solutions of TBAF are known to be basic,[113] and the ring closure of 2-ethynylanilines *via* treatment with strong bases such as potassium *tert*-butoxide has been reported.[114] Such reactions presumably progress through the mechanism outlined in scheme 2.19. Here the basic species deprotonates the aniline, and subsequent protonation of the alkyne by the conjugate acid activates the alkyne for nucleophile attack. However, it seems unlikely that the fluoride ion, as present in a TBAF solution, is sufficiently basic to deprotonate the aniline. Although it is true that the acidity of the aniline is increased by the withdrawing effect of the *para* nitro group.

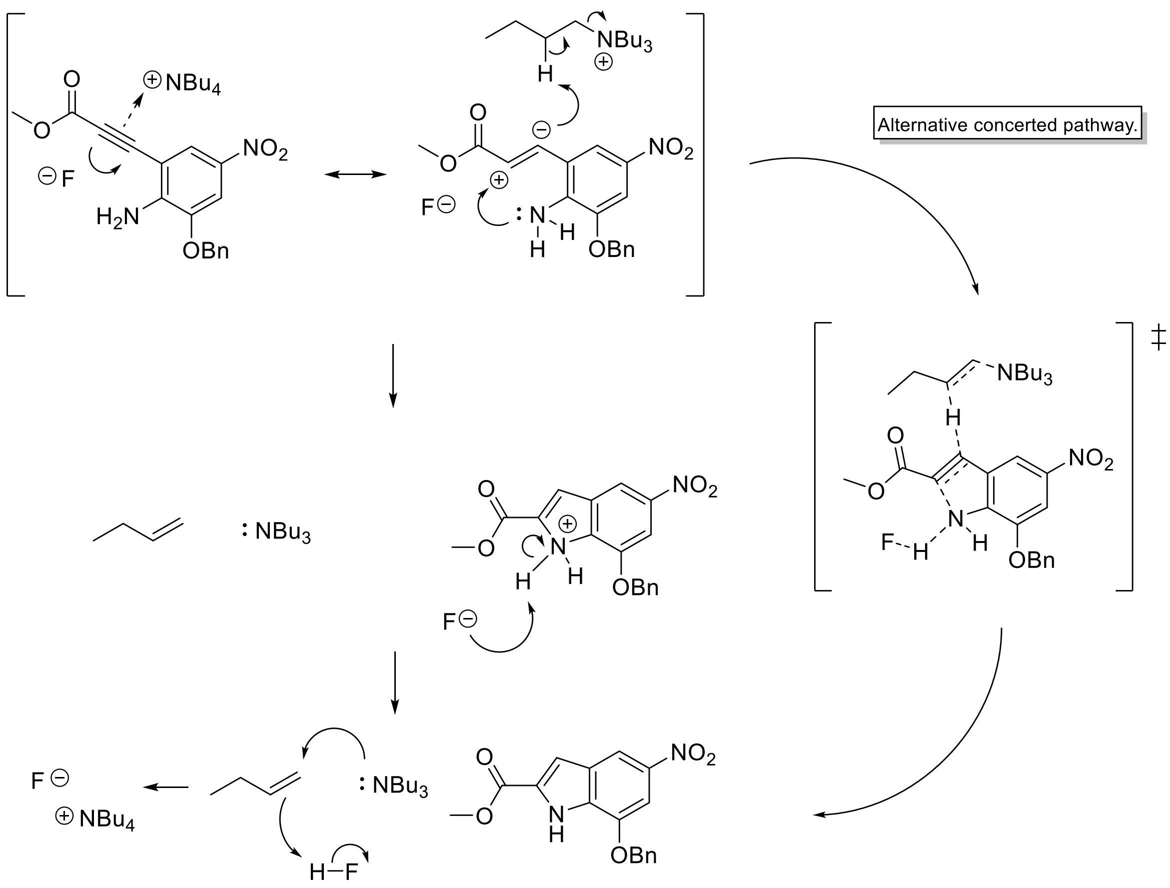


Scheme 2.19 Unlikely basic mechanism of ring closure.

Sakamoto *et al.* have explored the mechanism of similar TBAF promoted annulations of 2-alkynylbenzyl alcohol and 2-alkynylbenzylamine derivatives.[115] Their work clearly shows that both the fluoride anion and the tetrabutylammonium cation are essential to promote cyclisation. They also noticed the appearance of ^{1}H NMR signals corresponding to the formation of Bu_3N during the reaction. With this evidence in mind the mechanism depicted in scheme 2.20 seems more plausible.

In this proposal, Sakamoto *et al.* suggest that the alkyne is initially activated by association with the tetrabutylammonium cation and fluoride anion. This leads to polarization of the alkyne. The resulting isomerization brings the electrophilic end of this species in closer proximity to the aniline leading to nucleophilic attack of the nitrogen's lone pair. The nucleophilic end of the isomerized alkyne abstracts a hydrogen from one of the *tert*-butyl chains of the tetrabutylammonium cation *via* a Hoffmann type elimination.

This results in a protonated intermediate of the desired indole, and the formation of Bu_3N observed by Sakamoto *et al.* Subsequent deprotonation by the fluoride anion, gives the indole, and reaction of the Bu_3N, butylene, and HF regenerates TBAF. However, the resonance structure of the polarized alkyne proposed by Sakamoto *et al.* is very unusual. It seems perhaps more likely that this process is concerted and progresses through a transition state such as that also depicted in scheme 2.20.



Scheme 2.20 Possible reaction mechanism for TBAF promoted ring closure.

Such a mechanism would suggest that the reaction might be possible with a catalytic quantity of TBAF, and indeed Sakamoto *et al.* have shown that the annulation of 2-alkynylbenzylamine derivatives is possible with as little as 10 mol % TBAF.

As noted previously the TBAF promoted cyclisation as used in this synthesis appears to be intrinsically low yielding. It is not clear why, as the reaction is fast, and leads to only a base line impurity by TLC. Considering the water content of THF, it is possible that TBAF is sufficiently basic (*via* formation of hydroxide) under the reaction conditions to lead to partial hydrolysis of the methyl ester, which is then lost as the baseline impurity. In

retrospect it would be interesting to see if the yield could be improved by using a catalytic quantity of TBAF.

The annulation of 2-alkynylaniline derivatives is a common route to the formation of indoles, and many methods exist. Considering that the low yield of the TBAF cyclisation represented an early bottleneck in the synthesis, a number of alternative ring closures were trialed. However none were successful.

As Cu(I) salts are sometimes employed in such annulations,[116] the Sonogashira coupling was briefly revisited, and trialled using a stoichiometric amount of CuI, to see if this would induce *in situ* ring closure after coupling. Unfortunately this led to failure of the coupling reaction. This is perhaps due to homo-coupling of the methyl propiolate.[117]

Lewis acids such as $ZnBr_2$ have also seen success in similar ring closures,[118] and although no cyclisation was observed during the Negishi coupling, it was postulated that this may be because the $ZnBr_2$ is consumed in the formation of the alkynylzinc species. Therefore, separate treatment of the coupled product, with a stoichiometric excess of $ZnBr_2$ at reflux in toluene was trialled. This resulted in no reaction.

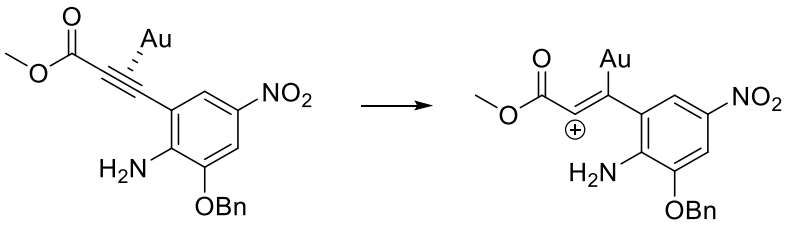
Finally it was decided to briefly explore transition metal catalysis. As no cyclisation was observed during the Negishi coupling, it was decided not to try separate treatment of the coupled product with $Pd(PPh_3)_2Cl_2$. Instead an attempt was made to cyclize **26** using Au catalysis. Au(III) species are known to be excellent activators of alkynes for nucleophilic attack, particularly for intramolecular hydroaminations such as this.[119]

Marinelli *et al.* report high yielding annulation of a wide range of 2-alkynylanilines, by treatment with the Au(III) species, $NaAuCl_4.2H_2O$ in EtOH at room temperature.[120] Their works showed promise, as one of their substrates contained a nitro group *para* to their aniline. The deactivating effect on the aniline of this group has been suspected to possibly contribute to the resistance of **26** to ring closure in this synthesis. They also use unprotected anilines. Furthermore they state that their method has successfully cyclized 2-alkynylanilines which have proved unreactive towards Pd catalysis.

Therefore **26** was treated with $NaAuCl_4.2H_2O$ in EtOH at room temperature. No reaction was observed. However, it was noted that **26** appeared to be only partially soluble in ethanol. The reaction was therefore heated to reflux, and although the starting material was consumed the reaction produced multiple products none of which matched the R_f of the desired indole by TLC. The reaction was repeated in THF, but no reaction was observed at room temperature or reflux.

The lack of reactivity at room temperature in ethanol is possibly due to poor solubility, but this is not true for the lack of reactivity in THF. It is still possible that the deactivating effect of the *para* nitro group is to blame, but as discussed one of Marinelli *et al.*'s substrates did contain this ring configuration. Although, it is also true that this proved to be their lowest yielding substrate.

It seems likely that the resistance of **26** to ring closure is in some way a consequence of the carbonyl of the ester being vicinal to alkyne. Au(III) activates the alkyne of 2-alkynylaniline derivatives for intramolecular hydroamination, by coordination as depicted in scheme 2.21. For **26**, this would lead to the formation of a positive charge on the carbon of the alkyne which is neighbouring the already partially positive carbon of the carbonyl, and this is likely unfavorable. It is true that this polarization of the alkyne is also required for cyclisation by TBAF. It may be that it is the contribution of the fluoride anion that makes that reaction possible, and why other annulation methods have failed.

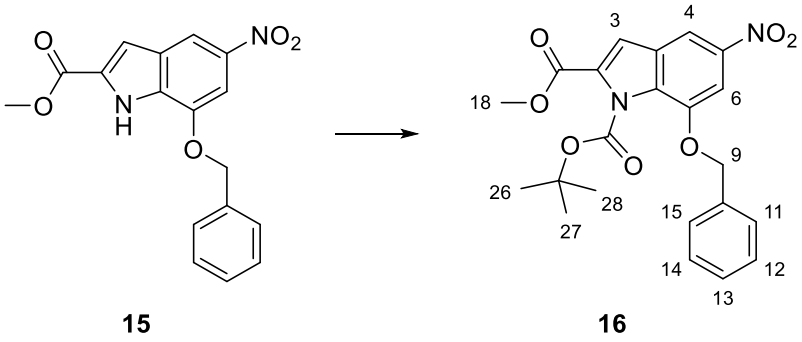


Scheme 2.21 Activation of alkyne by Au(III), disfavoured because of charge repulsion of the partially positively charged carbon of the neighbouring carbonyl.

The presence of the ester may introduce other problems for the Au catalyzed reaction. It is possible to imagine that the Au could coordinate with both the alkyne and the carbonyl oxygen leading to 5 membered metallic ring systems, or indeed that both oxygen atoms of the ester could chelate the Au. However, such processes would surely lead to a change by TLC analysis which was not observed.

As discussed, consumption of the starting material to form undesired products was observed at reflux in ethanol. As this was not seen at reflux in THF, this must be a consequence of the reactivity of ethanol, and may, or may not involve a contribution from the $NaAuCl_4.2H_2O$.

2.3.5 Boc protection of the indole.



Scheme 2.22 Boc-protection of the indole.

Successful formation of indole **15**, represents the convergence of this synthesis with the previously published route (scheme 2.4).[96] Here, Boc_2O and a catalytic quantity of DMAP is employed to introduce Boc-protection of the indole nitrogen. In this synthesis it was decided, following from previous experience to instead use 1 equiv. of DMAP allowing greatly reduced reaction times. The reaction was carried out in DCM and the reaction mixture subjected directly to flash chromatography, affording **16** in a yield of 75 %.

The most characteristic change in the ^1H NMR (see scheme 2.22 for numbering), was the disappearance of the indole nitrogen hydrogen signal, and the emergence of an upfield singlet at 1.47 ppm integrating for the 9 equivalent hydrogens of the Boc group (C-26, C-27, and C-28 hydrogen signals).

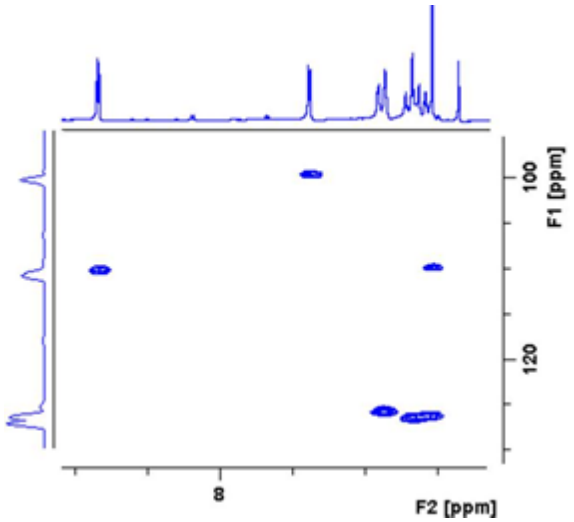
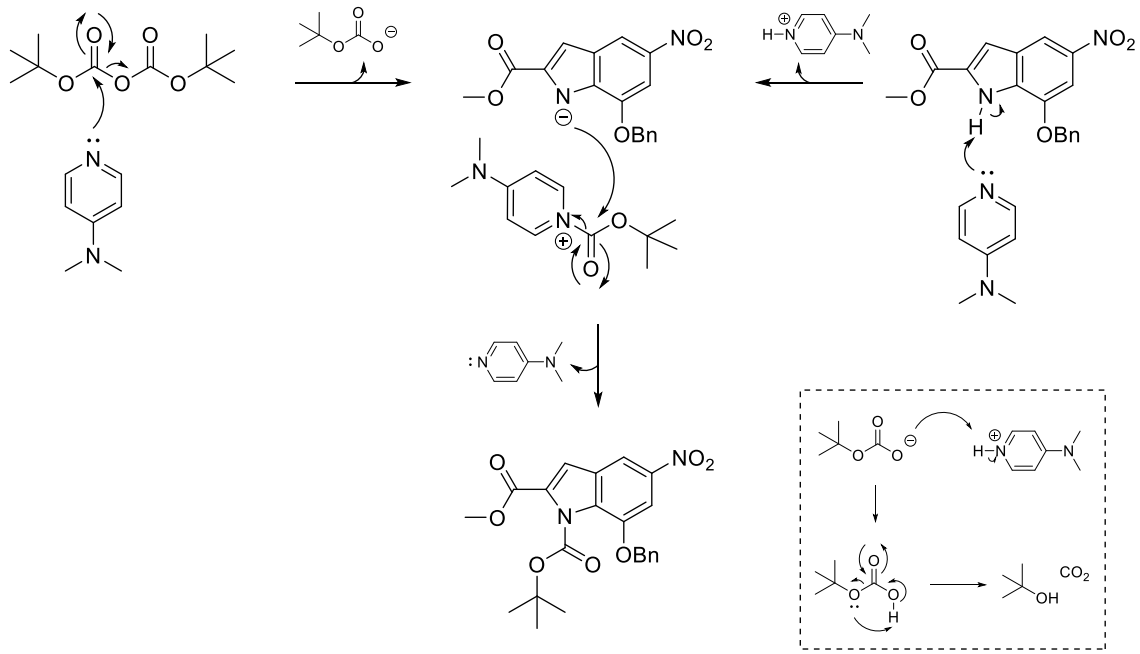


Figure 2.3 Aromatic region of the DEPT-edited HSQC of **16** at 298 K.

Interestingly the appearance of the aromatic signals of the benzyl group also changed. Where they appeared as single complex multiplet, they are now split in to 2 distinct multiplets, the most upfield of which partially overlaps with the C-3 indole aromatic

hydrogen signal at 7.33 ppm. Two dimensional DEPT-edited HSQC experiments confirmed this, as the C-3 proton signal was clearly correlated with a separate CH carbon distinct from those of the benzyl group, and one more similar in environment to the two carbon signals correlated with the C-4 and C-6 hydrogen signals (see figure 2.3). It is not obvious what causes this change. It may be a rotameric effect caused by steric clash between the benzyl ether and the newly introduced Boc group.

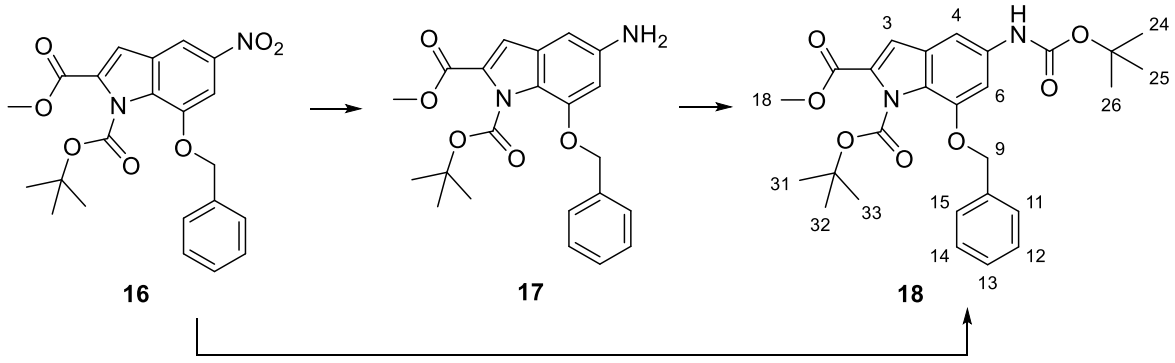


Scheme 2.23 Mechanism of Boc-protection of the indole nitrogen using DMAP as an acyl transfer agent.

The reaction is an acylation of the indole nitrogen by Boc_2O to give the carbamate. Although primary amines can react directly with Boc_2O, the acylation of indole nitrogens requires the use of a catalyst such as DMAP. The necessity for a catalyst is due to the reduced nucleophilicity of the indole nitrogen, a consequence of its lone pair formally contributing to the aromaticity of the indole ring system. The mechanism is outlined in scheme 2.23. Here DMAP is acting as both a base and an acyl transfer agent.

The nucleophilic pyridine nitrogen of DMAP attacks the electrophilic centre of one of the Boc_2O carbonyls. This leads to formation of the *tert*-butyl carbamate pyridinium cation, and a *tert*-butyl carbonate anion. DMAP also likely serves to deprotonate the indole. The pyridinium species is attacked by the deprotonated indole. Here DMAP serves as a leaving group, regenerating the catalyst to give the desired Boc-protected product.

2.3.6 Selective reduction of the nitro group and subsequent Boc-protection of the amine.



Scheme 2.24 Selective reduction of the nitro group and subsequent Boc-protection of the amine.

The next step in the published route[96] is the selective reduction of nitro indole **16** to the amine **17**, followed by Boc-protection to give **18**. Tietze *et al.* achieved the reduction *via* a heterogeneous Pd-catalysed hydrogenation. In order to prevent the competing benzyl ether cleavage they use a poisoned Pd species in the form of the Lindlar catalyst.

It was decided not to use this procedure, as the need for H_2 gas might prove troublesome when later scaling up the synthesis. Instead, a dissolving metal reduction using Zn and aqueous ammonium chloride, employed by Boger *et al.* for the same transformation,[90] was explored. The Boger group used 5 *equiv.* of Zn and 10 *equiv.* of ammonium chloride, stirring in a 5 to 1 mixture of acetone and water for 30 mins. This procedure was repeated but the results were disappointing. Although the starting material was consumed very quickly, many side products were observed by TLC, and the amine proved difficult to isolate.

It was decided to repeat the reaction, and subject the crude product directly to Boc-protection, with the expectation that the Boc-protected amine would be easier to isolate. Indeed, this proved to be the case, and the protected amine was easily isolated by flash chromatography. Sadly the yield was disappointing at 37 %. Boger *et al.* had reported 98 % for the reduction, and 95 % for the subsequent Boc-protection.

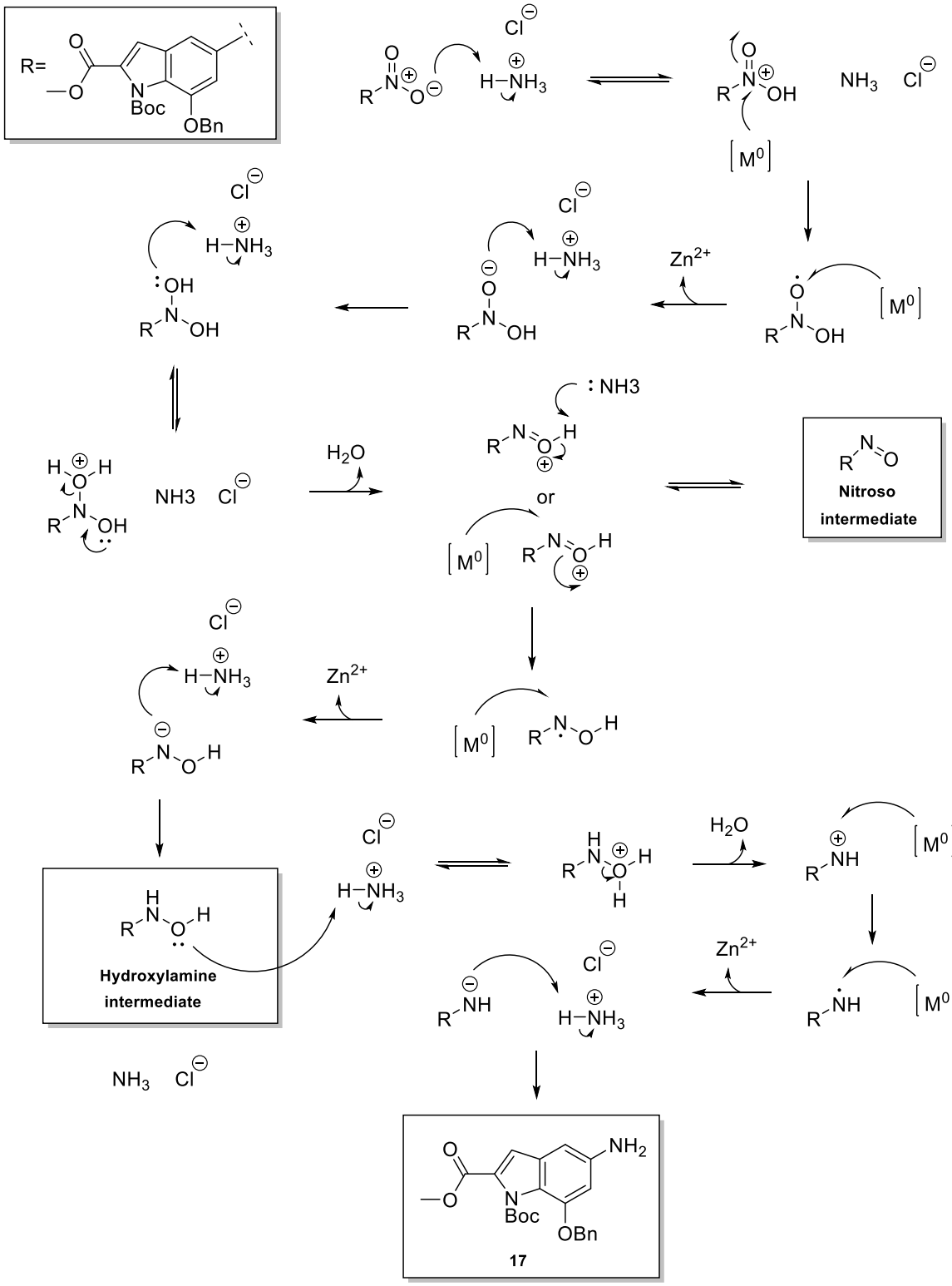
It was noted that during the chromatography, the silica column was left stained with a large green band spreading from the base line. The side product was not isolated, but the vivid colour led to speculation that it might be an azo type compound, formed from intermediates during the reduction. This suggested that, although the starting material was quickly consumed, formation of the amine might be slower than expected. It was noted on closer inspection of the Boger procedure, that their method used Zn nanopowder. The

particle size of Zn nanopowder is 3 times smaller than the standard Zn powder employed here. It therefore seemed possible that the larger surface area of the Zn nanopowder may have led to the more efficient reduction reported by the Boger group.

As a result the reaction was repeated, with the quantity of Zn powder being increased from 5 to 15 *equiv.* The product was again isolated as the Boc-protected amine by directly treating the crude reduction product with Boc_2O, and catalytic DMAP in THF. Pleasingly, this led to a significant improvement in yield to 70 %. Through further refinement, it was found that the process could be simplified, by trapping the forming amine *via in situ* Boc-protection, without a reduction in yield. This provided a convenient one pot procedure. For this process, THF was substituted for acetone. The change was made because THF was the solvent previously used for the Boc-protection reaction. However, it provided further unanticipated advantages. In acetone substantial aggregation of the zinc powder was observed, while in THF this was less pronounced, and thus the stirring properties of the suspension were improved. This was advantageous for scale up of the reaction, because significant aggregation of the zinc at larger scales, could lead to difficulties in stirring the reaction.

Successful nitro group reduction, and Boc-protection of the resulting amine can be confirmed from the characteristic changes observed in the ^{1}H NMR spectra of the product (see scheme 2.24 for numbering). Two upfield singlets both integrating for 9 hydrogens, observed at 1.51 ppm and 1.44 ppm, confirm the presence of the two Boc-protected nitrogens. The single hydrogen of the Boc-protected amine produces a broad singlet at 6.83 ppm. Loss of the strongly withdrawing effect of the nitro group, has led to a large upfield shift for the C-4 and C-6 aromatic hydrogen signals. Where they had appeared as a pair of narrow *meta* coupled doublets at 8.26 ppm, and 7.67 ppm in the Boc-protected nitro indole **16**, they are now observed as a broad singlet at 6.38 ppm, and a sharp singlet at 7.10 ppm respectively. The C-3 hydrogen signal remains relatively unchanged at 7.28 ppm and is still overlapping partially with the aromatic signals of the benzyl group. What causes the broadening of the C-4 hydrogen is not obvious.

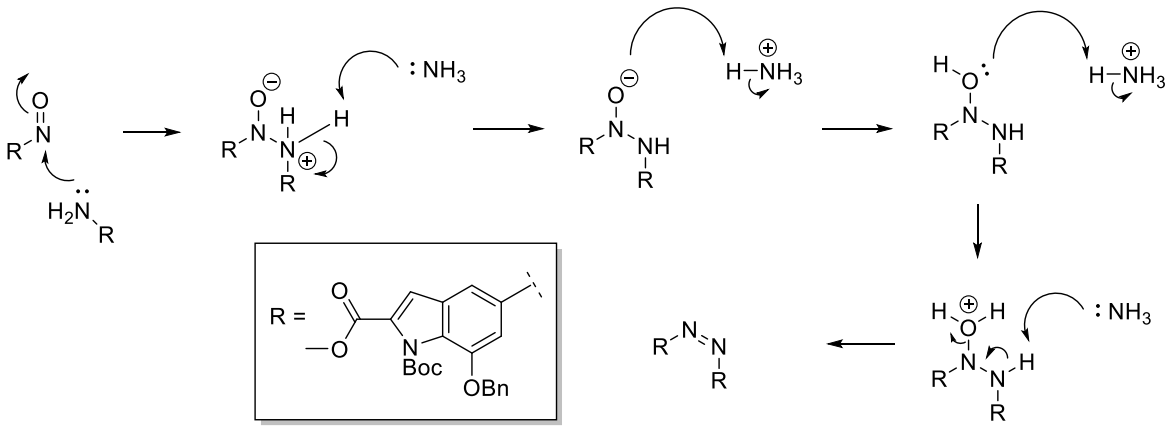
As already mentioned, the reduction of the nitro group using Zn and aqueous ammonium chloride is termed a dissolving metal reduction. Here the bulk metal serves as a source of electrons, and the mildly acidic ammonium chloride provides a source of protons. The exact molecular mechanism of such reductions is not fully understood, however scheme 2.25 shows a plausible mechanism which is consistent with the known intermediates of such reductions.[121]



Scheme 2.25 Plausible mechanism for nitro reduction.

The reduction is thought to progress through a series of single electron transfers and protonations. For every two single electron transfers that occur, a Zn^{2+} cation can be lost from the bulk metal, ultimately forming the soluble $ZnCl_2$ salt. This gives rise to the name 'dissolving metal reduction'.

The reduction is known to progress *via* a number of intermediates. First, the nitro group receives 2 electrons and 3 protons. The resulting loss of water gives the protonated nitroso. This species can be deprotonated to give the nitroso intermediate. Alternatively, the protonated nitroso can receive a further two electrons and another proton to give the hydroxylamine intermediate. The receipt of an additional 2 electrons and 2 protons, with the resulting loss of water, gives the desired amine.



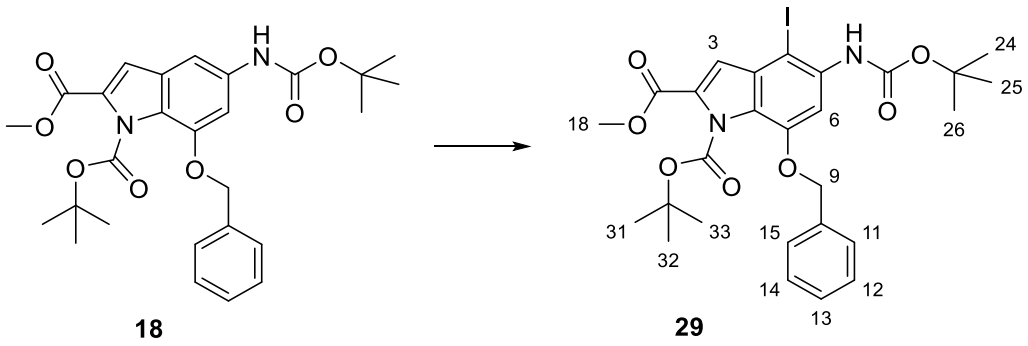
Scheme 2.26 Possible condensation reaction to give azo side product.

Such reductions commonly employ a strong acid such as HCl.[122] However, the risk of competing Boc-deprotection means this is not suitable for the reduction of **16**. Fortunately, ammonium chloride proves to be an acceptable source of protons. This is interesting because the treatment of nitrobenzene with Zn and ammonium chloride is known to give only the hydroxylamine.[123] For nitrobenzene, ammonium chloride is not sufficiently acidic to protonate the oxygen of the hydroxylamine over the nitrogen. The successful generation of the amine in this case, may be a consequence of the decreased availability of the nitrogen's lone pair, as it is delocalised into the larger aromatic system of the indole.

Given the complexity of the reduction mechanism, the suspected azo impurity observed when the reduction was carried out with only 5 *equiv.* of Zn powder, could have arisen in a number of ways. For example, it is possible to envision the condensation of the nitroso intermediate with the aniline (scheme 2.26).

Boc-protection of the resulting amine, is achieved *via* the same acyl transfer mechanism already described for the Boc-protection of the indole nitrogen in section 2.3.5. Trapping the amine by *in situ* Boc-protection in the one pot procedure, likely further reduces the risk of azo side products.

2.3.7 Iodination of the indole.



Scheme 2.27 Iodination of the indole.

Successful nitro group reduction with *in situ* Boc-protection had provided a convenient one pot procedure for the generation of **18**. The next step in the published route[96] is the bromination of **18** to give **19** (scheme 2.4). Introduction of the halogen provides the handle for the later radical 5-exo-trig cyclisation. In this synthesis, it was decided to again deviate from this route, and instead to iodinate **18**. The decision was made primarily due to reagent availability; however the weaker carbon iodine bond in theory should provide a more reactive site for initiation of the radical cyclisation.

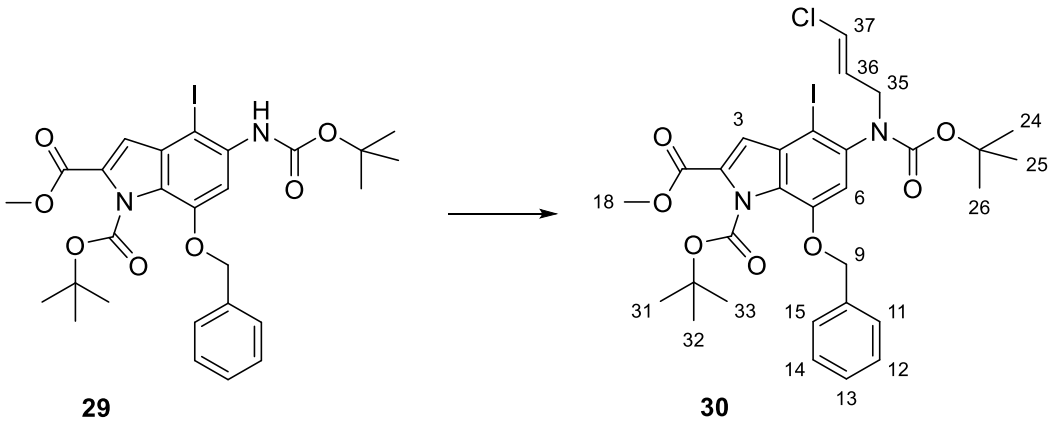
In this pilot synthesis iodination was achieved by following Boger *et al*.'s method,[90] *via* treatment of **18** with NIS and catalytic acetic acid in toluene overnight. The reaction mixture was subjected directly to flash chromatography, and the desired iodinated indole **29** isolated in a yield of 84 %.

Success of the iodination is evident in the ^{1}H NMR spectra, by the disappearance of the C-4 aromatic indole hydrogen signal (see scheme 2.26 for numbering). The withdrawing effect of the electronegative iodine atom has also caused a pronounced downfield shift in the C-3 hydrogen signal, which is now observed as broad singlet at 7.79 ppm. Again what is causing the broadening of this signal is not obvious. The C-6 hydrogen signal remains practically unchanged, and is still observed as a sharp singlet at 7.09 ppm. It is perhaps less affected by the withdrawing effect of the iodine atom, due to increased donation from the ether oxygen and Boc-protected amine.

The iodination occurs in an analogous fashion to that already described in section 2.3.2, with the NIS serving as a source of electrophilic iodine. The regioselectivity of iodination is likely a mixture of electronic and steric effects. The C-3 position (see scheme 2.26 for numbering), which would be the most nucleophile position of an unsubstituted indole, is deactivated by Boc-protection of the indole nitrogen and the vicinal ester. Furthermore, the C-4 and C-6 position are activated by mesomeric electron denotation by the ether oxygen, and also by the amine, although the amine's directing effect is reduced by the withdrawing affect of Boc-protection. It is not obvious that the C-4 position is more nucleophilic than the C-6 position, but this does appear to be the case given its more upfield resonance in the ^{1}H NMR of **18**. The decreased nucleophilicity of the C-6 position might be a consequence of an inductive withdrawing effect due to its closer proximity to the ether group. However preference for iodination at the C-4 position is most likely augmented by steric blocking of the C-6 position by the benzyl ether, and Boc-protected amine.

Confirmation of regioselectivity from the ^{1}H NMR alone is difficult. The fact that the two remaining indole hydrogen signals appear as singlets suggests a lack of *meta* coupling, and implicates the C-4 position as the site of iodination. However, the broadening effect of the C-3 hydrogen makes this difficult to rely on. This is especially true when considering that doublets arising from *meta* coupling were not observed in the starting material, due to broadening of the C-4 hydrogen signal in that case, and overlap of the C-3 hydrogen signal with that of the benzyl hydrogen signals. However, coupling is also not evident in COSY NMR analysis of **29**, which might still be expected despite broadening if iodination had occurred at the C-3 or C-6 position. The real confirmation of regioselectivity however, is the success of later reactions.

2.3.8 Introduction of the tethered vinyl chloride.



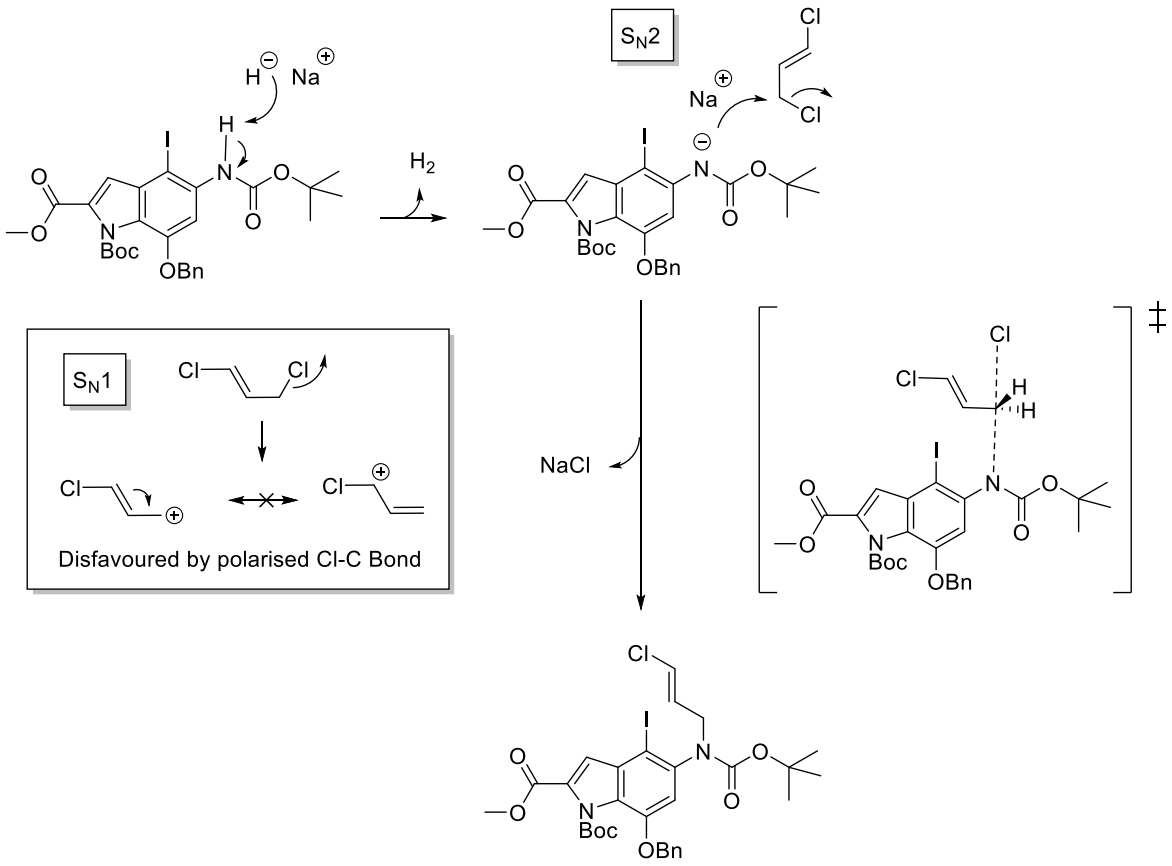
Scheme 2.28 Introduction of the tethered vinyl chloride.

Successful iodination had provided a handle for the formation of an aryl radical, and thus affords the starting point for the subsequent radical cyclisation to give the indoline ring. Introduction of the vinyl chloride completes the substrate for this cascade reaction.

The vinyl chloride is introduced by the *N*-alkylation of **29** with 1,3-dichloropropene. In the pilot synthesis this was achieved using the standard conditions of NaH in DMF.[90, 96] Like the formation of the benzyl ether discussed in section 2.2.1, this reaction is a nucleophilic substitution. The carbon of the polarised sp^3 hybridised carbon chlorine bond provides the electrophilic centre, and the nitrogen of the carbamate acts as the nucleophile. Deprotonation of the carbamate nitrogen by the NaH improves its nucleophilicity by providing a formal negative charge.

As with the reaction of the phenol with benzyl bromide this reaction most likely proceeds *via* a concerted S_N2 mechanism, but again the sp^2 hybridised nature of the vicinal carbon means the carbocation formed *via* an S_N1 mechanism could be stabilised by resonance with the double bond. However the second chloride makes this less likely, as this resonance structure would place a positive charge at the already partially positive carbon (scheme 2.29).

The deprotonation was conducted at 0 ^oC due the reactivity of the hydride. Success is evident by the evolution of H_2 gas, prompting the addition of the 1,3-dichloropropene, at which point the reaction was allowed to warm to room temperature. An excess of 3 *equiv.* was used to ensure the reaction progressed to completion, as it proved difficult to monitor by TLC. These conditions provided an acceptable yield of 60 % after purification by flash chromatography.



Scheme 2.29 *N*-alkylation to introduce the vinyl chloride most likely proceeding via S_N2 mechanism.

For economical reasons the 1,3-dichloropropene is used in the form of a mixture of *E* and *Z* isomers. This is of no consequence for the next step of the synthesis as the double bond is removed by the formation of the indoline ring. However, it does make assignment of the ^{1}H NMR of **30** difficult, as the product is also isolated as mixture of *E* and *Z* isomers. Therefore the product was initially confirmed by accurate mass analysis, which showed a mass consistent with the desired structure and the correct isotope pattern for a molecule containing 1 Cl atom. Use of DEPT-edited HSQC NMR experiment allowed for some interpretation of the ^{1}H NMR (see scheme 2.28 for numbering).

The five aromatic hydrogen signals of the benzyl group produce a complex multiplet at 7.28-7.44 ppm. A sharp singlet at 7.18 ppm, and an apparent multiplet at 6.65-6.47 ppm correspond to the two indole aromatic hydrogens at C-3 and C-6. DEPT-edited HSQC clearly shows the multiplet is correlated with one indole carbon. At 6.00-5.80 ppm another complex multiplet is seen. This integrates for two hydrogens. The HSQC shows that this is correlated with 4 similar carbons, and the DEPT editing shows these to be CH carbons. These must relate to the alkene hydrogen signals at C-36 and C-37, with the 4 carbon correlations being a result of the two isomers (see figure 2.4 a).

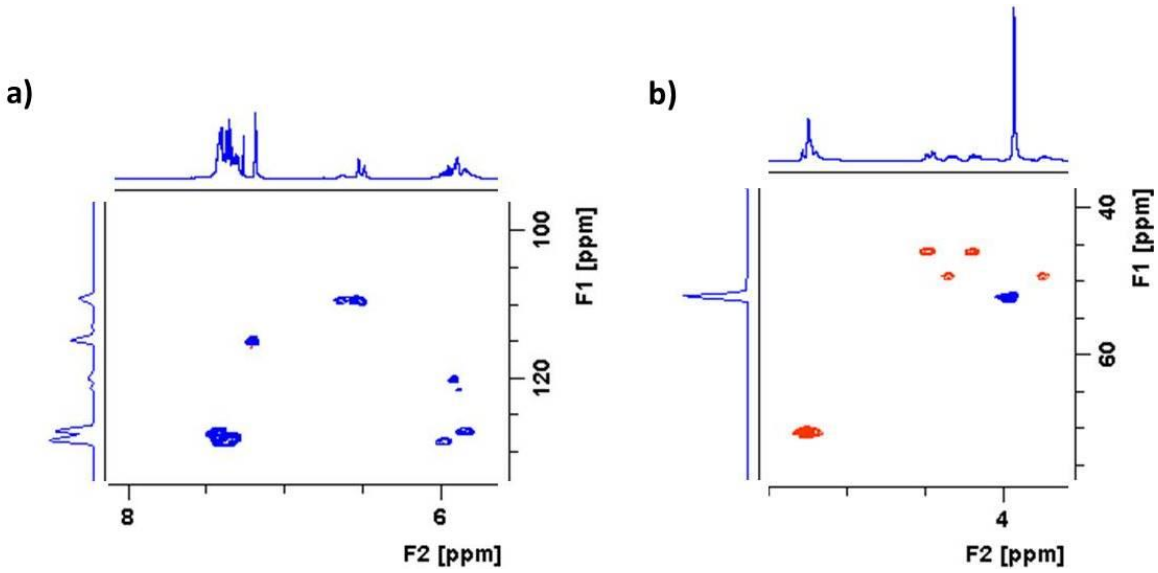
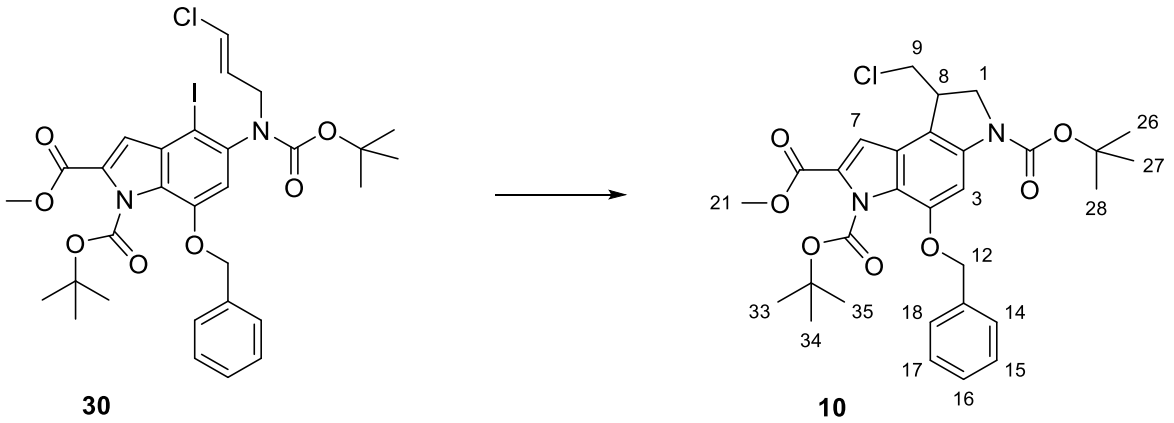


Figure 2.4 (a) Expansion of the 5.5 ppm to 8 ppm (F2) region of the DEPT-edited HSQC of **30** at 298 K. (b) Expansion of the 3.5 ppm to 5.5 ppm (F2) region of the DEPT-edited HSQC of **30** at 298 K. DEPT phasing: Blue = CH or CH_3 carbon. Red = CH_2.

At 5.17-5.28 ppm another apparent multiplet is observed. In the DEPT-edited HSQC this is correlated with one CH_2 carbon, and relates to the hydrogen signals at C-9 of the benzyl group. The 2 hydrogens of the C-35 carbon are appearing as 4 distinct multiplets each integrating to 0.5, corresponding to the geminal hydrogens at this position in each isomer. At 3.93 ppm, the C-18 methyl hydrogens still appear as a narrow singlet (see figure 2.4 b). Finally one Boc group is appearing as a singlet integrating for 9 hydrogens at 1.53 ppm. The other is distorted and split into two peaks together integrating for 9, at 1.29 ppm & 1.27 ppm.

2.3.9 Formation of the indoline ring.

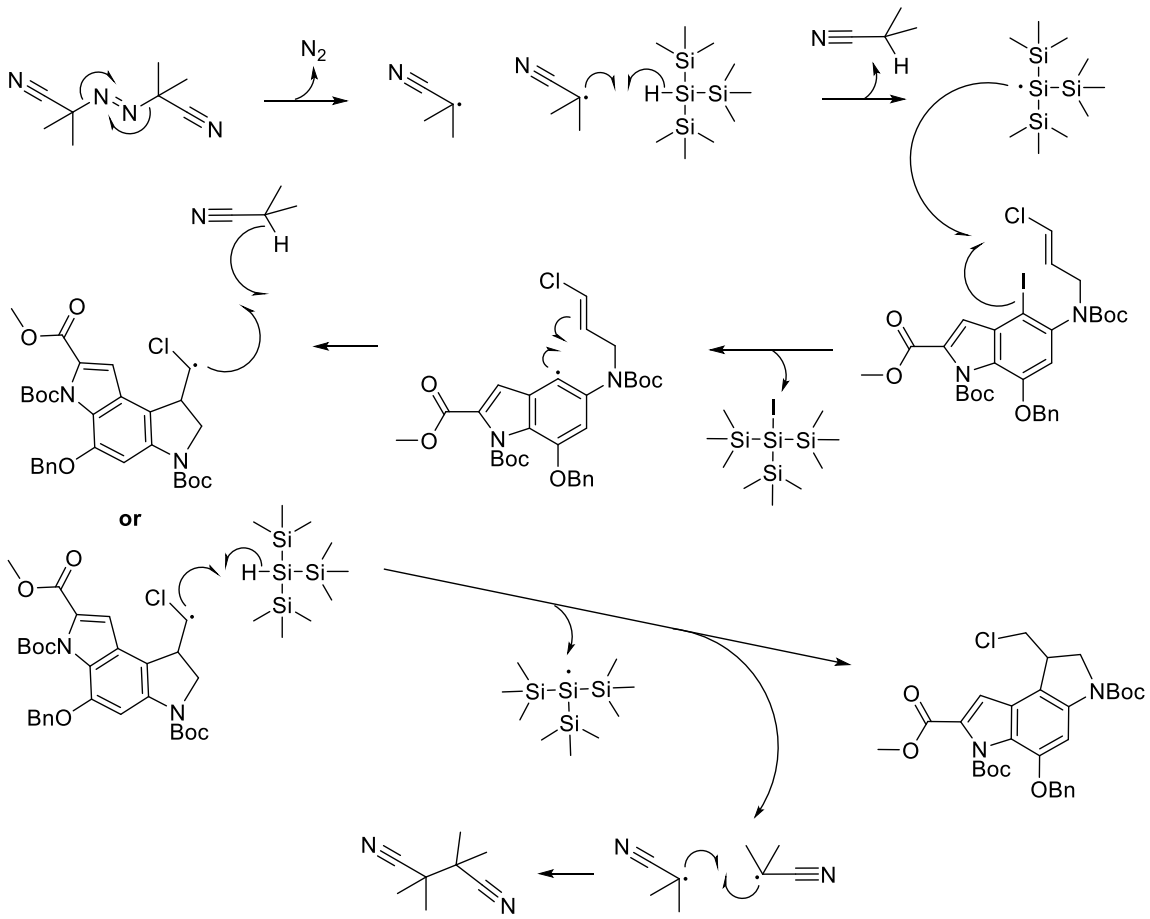


Scheme 2.30 Formation of the indoline ring.

Successful access to **30** provided the substrate for the 5-exo-trig radical cyclization which would form the indoline ring, and afford the benzyl protected seco form of the di-Boc-protected duocarmycin alkylation subunit **10**.

The original chemistry pioneered by the Patel and Boger groups makes use of Bu_3SnH to propagate the radical reaction.[90, 91] Here it was decided to follow the Tietze adaption which employs TTMSS in place of the tin species, as it is reportedly easier to remove during purification.[96] The Tietze group conducts the reaction in benzene. Here it was decided to substitute benzene for the less toxic toluene, as this was expected to be easier to handle during the scale up of the synthesis.

This step proceeded without issue, using 25 mol % of AIBN, and 1.1 equiv. of TTMSS in toluene at 90 ^{o}C. The presence of reactive O_2 is known to be potential deleterious to the success of radical reactions, therefore this step was performed under an atmosphere of N_2 and the solution degassed prior to heating. The reaction was carried out at a low concentration of 0.03 M. This is to favour the intramolecular cyclisation, and lessen the likelihood of any intermolecular side reactions. The product was easily isolated by subjection of the reaction mixture directly to flash chromatography, in excellent yield (90 %).



Scheme 2.31 Likely mechanism of the 5-exo-trig radical cyclisation.

As discussed the reaction is 5-exo-trig radical cyclisation, and is most likely to proceed *via* the mechanism depicted in scheme 2.31. Here AIBN serves as a radical initiator. Upon heating, the AIBN decomposes *via* the concerted homolytic cleavage of the two carbon nitrogen bonds of the azo group. This is driven by the thermodynamically favourable release of diatomic nitrogen gas. The process also forms two isobutyronitrile radicals. These radicals initially existed as a cage pair. This is to say they are solvated as one species *via* a 'cage' of surrounding solvent molecules. Some of the caged radical pairs couple together to form tetramethylsuccinonitrile. Others escape the cage by rapid diffusion and become available to initiate the reaction.

Although it is of course possible for the isobutyronitrile radicals to react directly with the substrate, the main pathway involves the propagation of the reaction *via* the stoichiometric excess of TTMSS. Here, the isobutyronitrile radical abstracts a hydrogen from the relatively weak hydrogen silicone bond of the TTMSS, producing the silane radical. Subsequently the silane radical attacks the weak iodine carbon bond of **30**, abstracting the halogen and providing the aryl radical. Formation of the aryl radical provides the starting point of the intramolecular cyclisation, and the subsequent attack of this species

at the alkene forms the indoline ring. This process is likely augmented by the donating effect of the *para* ether group, which increases the electron density at the radical centre. The resulting alkyl radical is presumably reduced by either a molecule of unreacted TTMSS, or the isobutyronitrile formed by the generation of the silane radical. Both routes afford the desired product, and propagate the reaction by producing more radicals. Various termination steps are possible, *via* the coupling of two radicals.

The reaction is selective for the 5-exo-trig cyclisation. No evidence of the completing 6-endo-trig reaction was observed. This is consistent with Baldwin's rules.[124] Although, the more substituted radical product of the 6-endo-trig cyclisation is often more stable, and hence thermodynamically favoured; the dominance of the 5-exo-trig pathway is typically rationalised as being more kinetically favourable, due to better overlap of the molecular orbitals in the transition state.[125]

In fact, in the case of this substrate, the 5-exo-trig pathway might be both kinetically, and thermodynamically favoured. This is because of the chlorine atom geminal to the radical centre of the cyclisation product. Although this group is electronegative, and thus might be seen to disfavour the electron deficient radical, it is important to remember that the alkyl radical is best described as being sp^2 hybridised. It is therefore conceivable, that the singularly occupied p orbital of the radical could overlap with a lone pair containing p orbital of the chloro group, and thus allow donation of electron density to the radical. This may have a net stabilising effect which is greater than the inductive stabilisation of the radical formed by the 6-endo-trig pathway.

As stated the substrate for this reaction is a mixture of *E* and *Z* isomers. This has no effect on the product as the carbon bonding to the chlorine becomes sp^3 hybridised and achiral in the product. Chirality is however introduced at C-8 carbon (see scheme 2.30 for numbering). The reaction is racemic, as the aryl radical can attack either face of the alkene.

Introduction of the chiral centre produces diastereotopic effects in the ^{1}H NMR. The geminal hydrogens of the C-1 and C-9 carbons, are non-equivalent. This leads to complex second order effects in the splitting patterns of the five hydrogens comprising the substituted indoline ring. These appear as triplet integrating for one hydrogen at 4.13 ppm, neighbouring a complex multiplet at 4.06-3.89 ppm integrating for the remaining 4 hydrogens.

A characterised correlation pattern is observed when **10** is analysed by DEPT-edited HSQC experiment. The triplet at 4.13 ppm is correlated to a CH_2 carbon at 52.3 ppm. This

carbon is also correlated with part of the neighbouring multiplet, approximately the 3.91 ppm region. These signals most likely correspond to the geminal diastereotopic pair of hydrogens on the C-9 carbon. This assumption is made based on the downfield position of the first hydrogen signal, as this pair is closest to the electronegative chlorine. Similarly the 4.07 ppm, and 3.92 ppm regions of the multiplet, are both correlated with the same CH_2 carbon at 47.6 ppm. This represents the second geminal diasterotopic pair of the C-1 carbon. Finally the 3.99 ppm region of the multiplet is correlated to a CH carbon at 40.7 ppm, corresponding to the hydrogen of the C-8 carbon of the chiral centre (see figure 2.5).

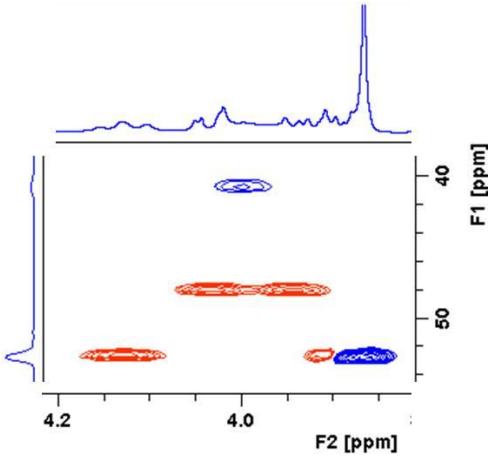
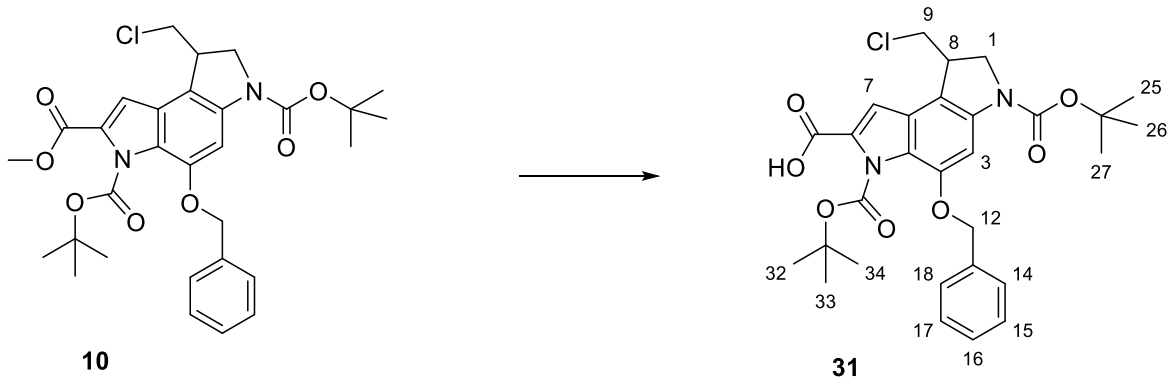


Figure 2.5 Expansion of the indoline region of the DEPT-edited HSQC of **10** at 298 K. DEPT phasing: Blue = CH or CH_3 carbon. Red = CH_2.

Partially overlapping with the upfield edge of the indoline multiple is a strong singlet integrating for the 3 hydrogens of the methyl ester, at 3.87 ppm. This correlated with a CH_3 carbon at 52.2 ppm.

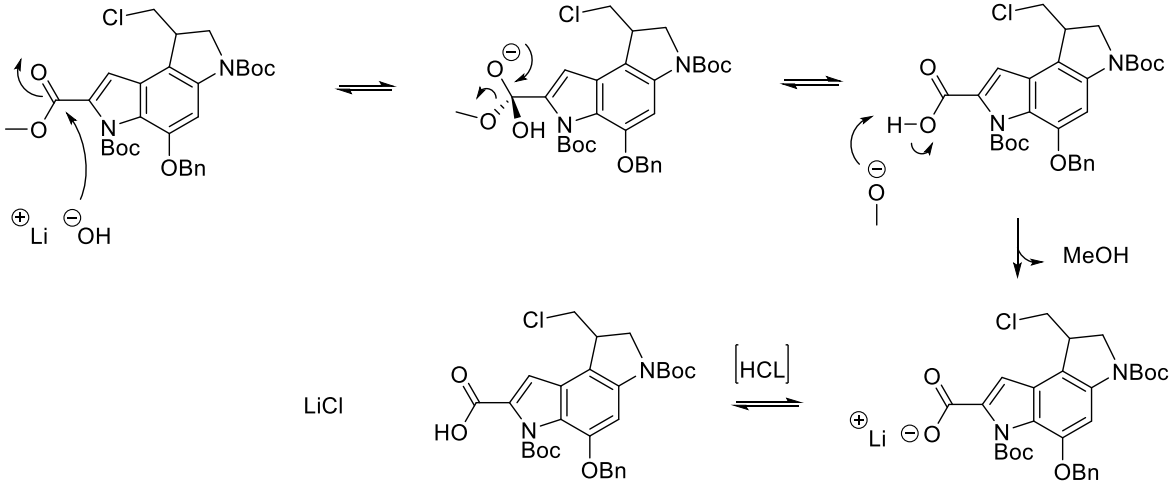
The two Boc groups are now again seen as two singlets at 1.48 ppm, and 1.39 ppm. A singlet at 5.27 ppm corresponds to the two alkyl hydrogens of the benzyl group. The aromatic benzyl hydrogens signals are overlapping with one of the indole hydrogens to form a complex multiplet 7.47-7.29 ppm, integrating for 6 hydrogens. This is again confirmed by the HSQC experiment. The second indole hydrogen is observed as broad singlet at 7.69 ppm. Again this signal was confirmed to correlate with a CH carbon at 97.4 ppm.

2.3.10 Ester hydrolysis.



Scheme 2.32 Ester hydrolysis.

The successful 5-exo-trig radical cyclisation had provided the previously reported racemic benzyl protected seco form of the di-Boc-protected duocarmycin alkylation subunit **10**.[96] It was now envisioned that hydrolysis of the methyl ester, exhaustive Boc-removal, and regioselective protection of the indoline nitrogen with Fmoc-Cl, would yield the desired solid phase building block.



Scheme 2.33 Mechanism of methyl ester hydrolysis with LiOH.

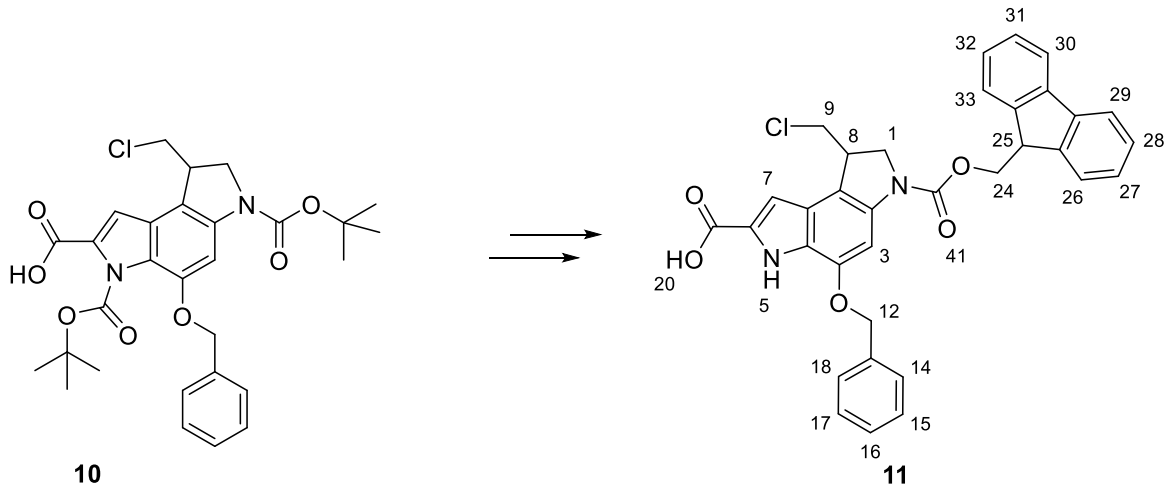
Although there was potential for the free carboxylic acid to interfere with Fmoc-protection of the indoline nitrogen, it was decided that ester hydrolysis would be performed first. This was because the Fmoc group would almost certainly be liable to cleavage under the basic conditions planned to hydrolyse the methyl ester.

Hydrolysis was affected at room temperature in a 3:2:1 mixture of THF, MeOH and a saturated aqueous solution of LiOH. This step proceeded without issue in a near quantitative yield of 96 %.

The product was confirmed by the disappearance of the methyl hydrogen signal in the ^{1}H NMR, and the corresponding carbon signal in ^{13}C NMR spectra of the HSQC. A broad singlet integrating for the exchangeable hydrogen of the carboxylic acid, was also observed at 13.43 ppm.

The mechanism of base promoted hydrolysis of the methyl ester is shown in scheme 2.33. Here, the hydroxide ion attacks the electrophilic carbon of the carbonyl. This forms a new α bond breaking the π bond, and forming an anionic tetrahedral intermediate. The anionic oxygen acts as the driving force for collapse of this intermediate, resulting in reformation of the carbonyl, and the loss of the alkoxide ion as a leaving group. This species deprotonates the newly formed carboxylic acid to give the Li salt. The negative charge is shared by both oxygen atoms of the carboxylate. Acidic work up gives the free carboxylic acid.

2.3.11 Boc cleavage and introduction of Fmoc-protection.

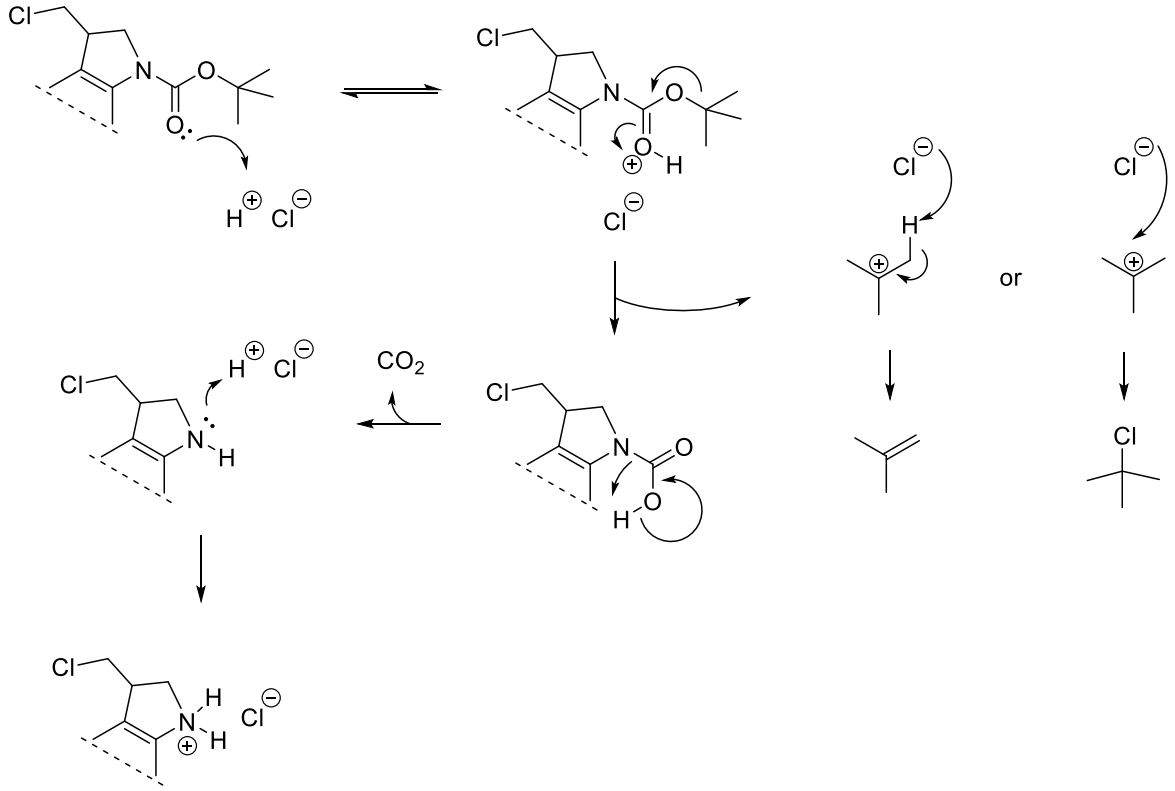


Scheme 2.34 Boc cleavage and introduction of Fmoc-protection.

Boc-deprotection of both nitrogens was affected simultaneously by treatment with 4 M HCl in ethyl acetate overnight. This was followed by regioselective Fmoc-protection of the crude product at the indoline nitrogen, using 1 *equiv.* of Fmoc-Cl and 3 *equiv.* of sodium bicarbonate at 0 ^{o}C. The product was isolated in a yield of 78 % after column chromatography.

Boc-deprotection of the indoline nitrogen under acidic conditions can be rationalised *via* the mechanism depicted in scheme 2.35. Here protonation of the carbamate results in decomposition of the Boc group to give the carbamic acid, and *tert*-butyl cation. Subsequent decarboxylation of the unstable carbamic acid affords the secondary amine driven by the release of carbon dioxide. Under the acidic conditions the amine is

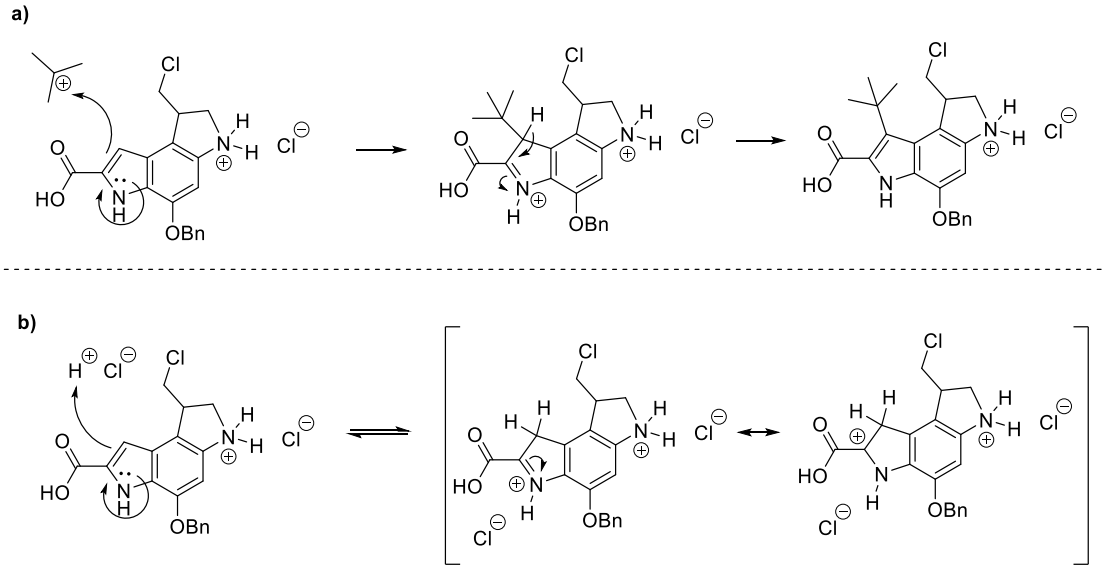
protonated to give the chloride salt. Boc-deprotection of the indole nitrogen proceeds in an analogous fashion.



Scheme 2.35 Mechanism of acidic Boc-deprotection of the indoline nitrogen, and two possible fates of the *tert*-butyl cation.

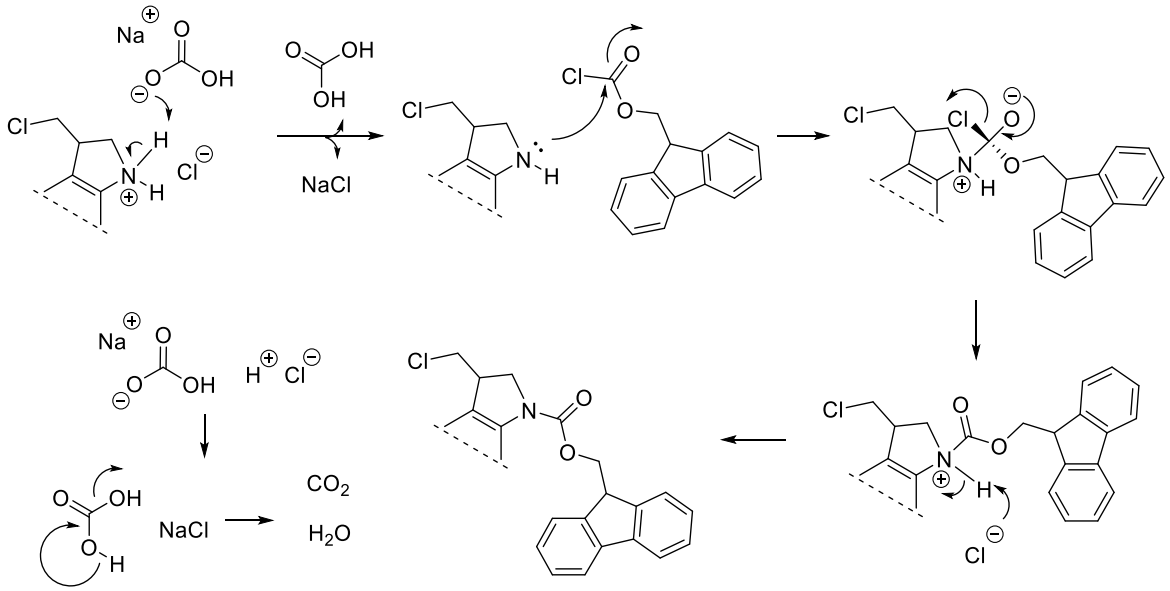
The fate of the *tert*-butyl cation in anhydrous conditions is debatable. Deprotonation by the chloride ion, and the subsequent release of isobutylene gas is one possibility, as is the formation of *tert*-butyl chloride.

There is also a risk of alkylation of the nucleophilic indole scaffold by the *t*-butyl cation at the electrophilic 3 position (standard indole numbering, C-7 in duocarmycin structure). This may have occurred to a small extent, and it is possible that the yield maybe improved by the addition of scavengers. However, it is also possible that the indole scaffold is somewhat protected from alkylation under the strongly acidic conditions, due to protonation of the indole. The aromatic nature of indole means the nitrogen is an extremely weak base. However, protonation at the 3 position to give the indolyl ion is possible with strong acids such as HCl (see scheme 2.36).[126]



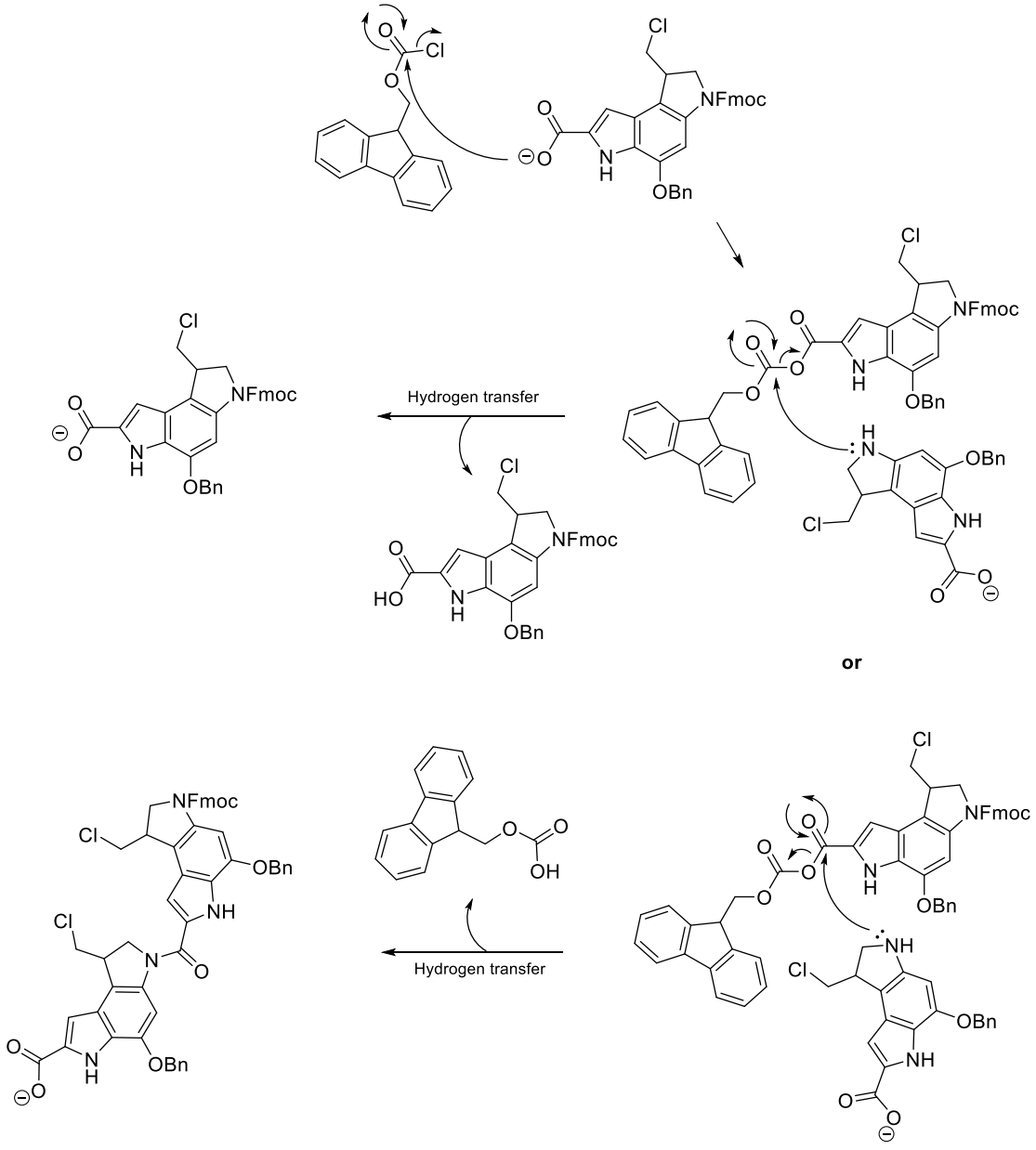
Scheme 2.36 (a) Possible alkylation of the indole by the *tert*-butyl cation. (b) Possible protection from alkylation by protonation of the indole under the strongly acidic conditions.

Fmoc-Cl is an acyl chloride, and Fmoc protection of the indoline nitrogen proceeds *via* a nucleophilic addition-elimination mechanism (see scheme 2.37). The sodium bicarbonate is added first to neutralise the chloride salts resulting from Boc-deprotection. Subsequent addition of the Fmoc-Cl, leads to nucleophile attack of its carbonyl by the indoline nitrogen. The resulting anionic tetrahedral intermediate collapses reforming the carbonyl, with the chloride ion acting as the leaving group, and subsequently deprotonating the carbamate, releasing HCl. This is neutralised by the sodium bicarbonate.



Scheme 2.37 Mechanism of Fmoc-protection.

Fmoc-Cl is extremely reactive and the reaction is facile even at 0 ^{o}C. No excess of the reagent is used in order to limit side reactions. Regioselectivity is achieved due to the increased nucleophilicity of the indoline nitrogen, over both the indole and carboxylate. Reaction at the carboxylate has the potential to lead to an interesting side reaction, as this would form a mixed anhydride. Subsequent attack of the indoline nitrogen could then take place at either carbonyl. This could lead to a second route to the desired product, or the coupling of two indole units (see scheme 2.38). These side reactions did not prove to be a problem. If they had, then use of the less reactive Fmoc-OSu in place of Fmoc-Cl would likely resolve the issue.



Scheme 2.38 Potential side reaction via formation of a mixed anhydride.

A mass consistent with the structure of the desired product was observed by accurate mass spectrometry, including the expected chlorine isotope peaks. However, success of the reaction was not immediately evident from the ^{1}H NMR spectrum, which contained unexpected second order effects making assignment difficult. However, analysis by DEPT-edited HSQC experiment confirmed the correct structure. The second order effects were suspected to have resulted from the presence of rotamers. This was confirmed by the observation of rotameric coalescence when ^{1}H NMR analysis was performed at 333 K. The rotamers most likely result from steric clash between the Fmoc and benzyl groups.

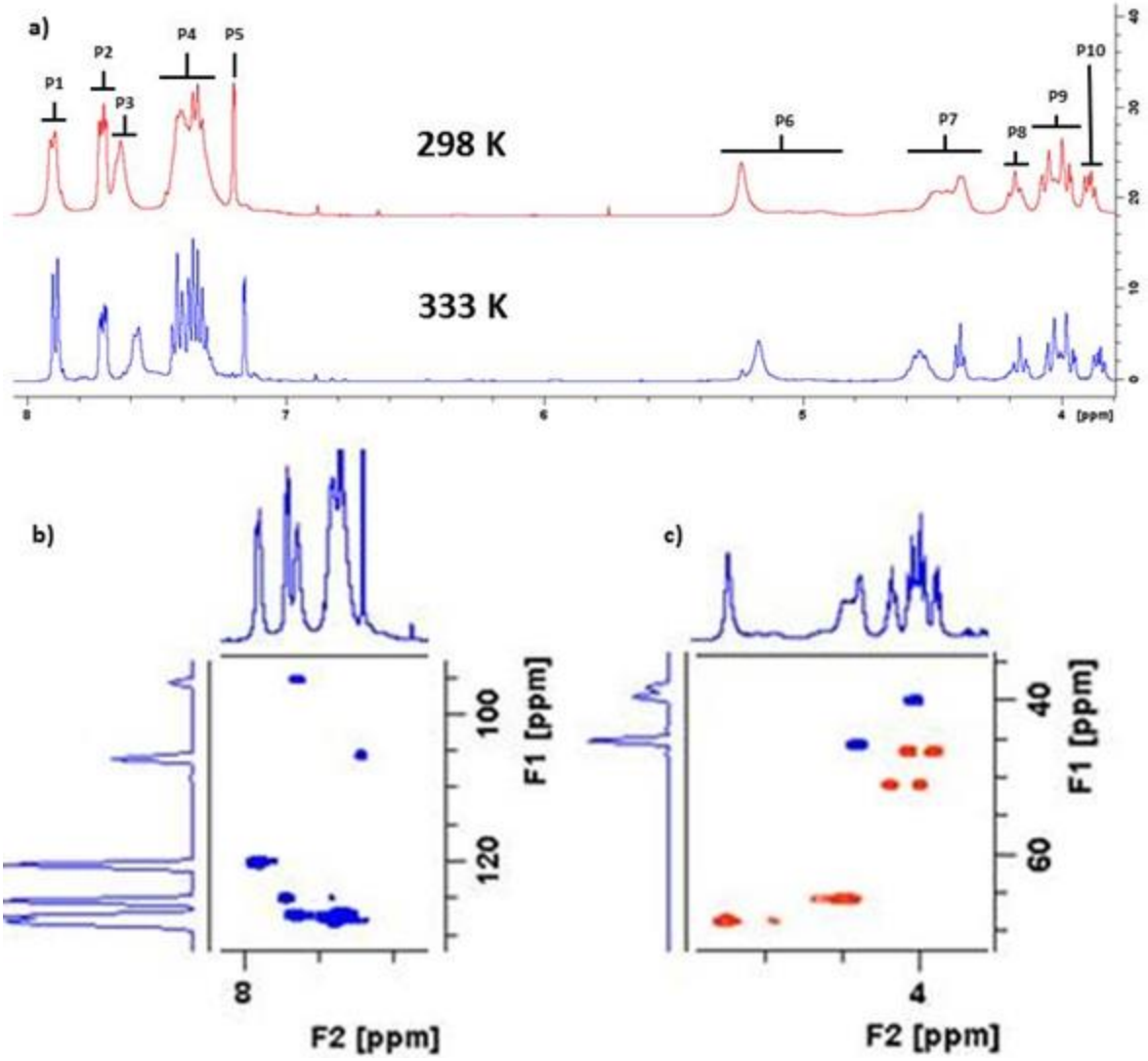


Figure 2.6 (a) Overlay of the ^{1}H NMR of **11** at 298 K and 333 K. (b) Aromatic region of the DEPT-edited HSQC of **11** at 298 K. (c) Aliphatic region of the DEPT-edited HSQC of **11** at 298 K. DEPT phasing: Blue = CH or CH_3 carbon. Red = CH_2.

Figure 2.6a shows an overlay of the relevant region of the ^{1}H NMR at 298 K and 333 K. Figure 2.6b and c, show expansions of the DEPT-edited HSQC at 298 K for the aromatic

and aliphatic regions respectively. Blue cross-peaks denote correlation with either CH or CH_3 carbons. Red cross-peaks denote correlation with CH_2 carbons.

The peak labelled P7 integrates to 3, and corresponds to the overlapping signals of the C-24 and C-25 aliphatic hydrogens of the Fmoc group (see scheme 2.34 for atom numbering). This can clearly be seen by correlation with a CH and CH_2 carbon in the DEPT-edited HSQC. These signals separate 333 K, to give the expected triplet for the C-25 Hydrogen, and an apparent broad quintet for the C-24 hydrogen where a doublet might be expected. What is causing the apparent broad quintet is not obvious.

Peak 6 is a singlet and integrates for 2. It correlates with a CH_2 carbon, and corresponds to the C-12 aliphatic hydrogens of the benzyl group. At 298 K, Peak 6 is significantly and unsymmetrically broadened, and this is also evident in the DEPT-edited HSQC. Significant but not complete narrowing of the singlet is observed at 333 K. It maybe that this is related to the apparent quintet observed at 333 K for the C-38 hydrogens of the Fmoc group. Perhaps some rotameric effect is still present.

P8, P9, and P10 remain consistent at 298 K and 333 K. These peaks represent the substituted indoline hydrogens. The characteristic diastereotopic pairs for the C-1 and C-9 hydrogens are again evident by the DEPT-edited-HSQC. P8 intergrates for 1 hydrogen. Its relative downfield position, suggests it is one of the C-9 hydrogens, as this pair is closest to the electronegative chlorine atom. P8 is correlated to a CH_2 carbon which is also correlated to part of the P9 signal, representing the second hydrogen of this diastereotopic pair.

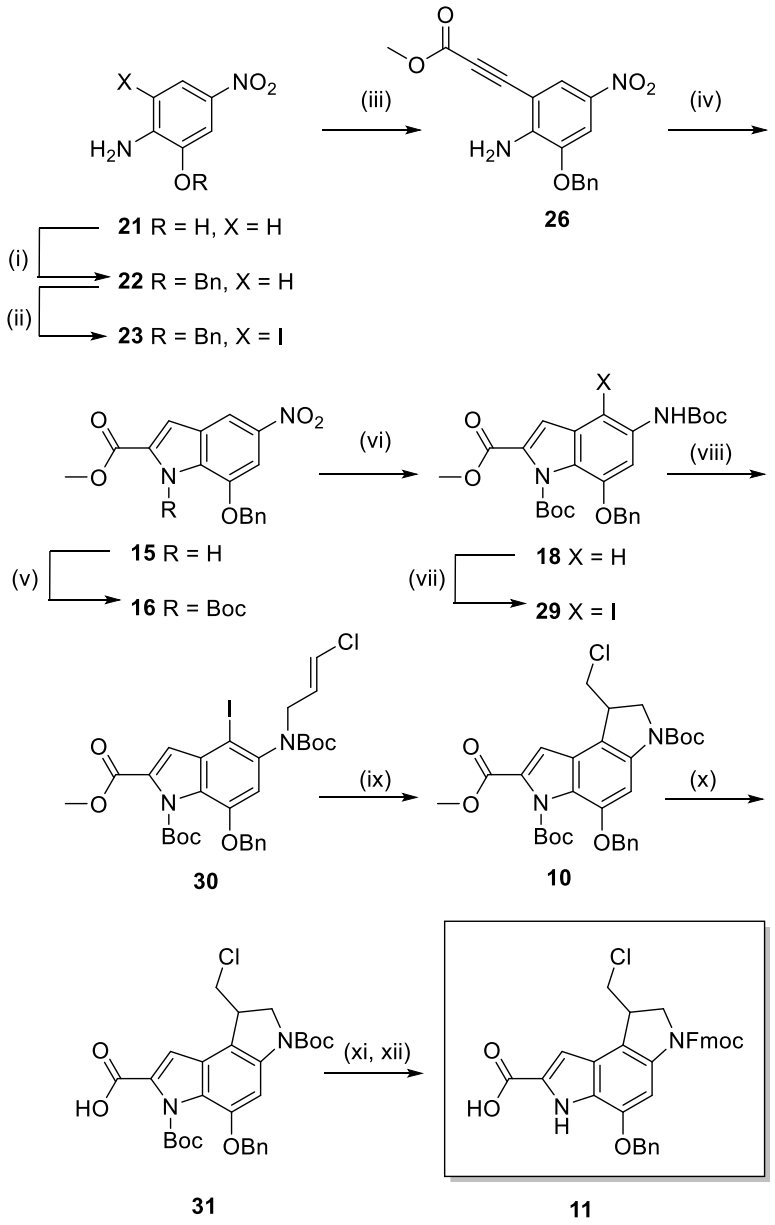
In total P9 integrates for 3 hydrogens. One of the additional signals corresponds to one of the hydrogens of the second diastereotopic pair on the C-1 carbon. Again this is evident from the DEPT-edited-HSQC, by correlation to a CH_2 carbon, which is also correlated to P10. P10 integrates for 1 hydrogen and is of course the second member of the C-1 diastereotopic pair. The remaining hydrogen signal of the P9 multiplet, is clearly correlated with a CH carbon, and represents the single hydrogen of the C-8 carbon.

The aromatic region again shows significant peak sharping at 333 K, further suggesting the rotameric effect is due to steric clash between the Fmoc and benzyl groups. P5 is a narrow singlet integrating to 1, and is clearly correlated with an indole type carbon in the DEPT-edited HSQC. It is most likely the C-3 Hydrogen. The C-7 indole hydrogen, is observed as part of the P3 signal. Again this is due to correlation with an indole type carbon. The remaining aromatic signals correspond to those of the Fmoc and benzyl protecting groups.

Finally the hydrogen of the indole nitrogen and the carboxylic acid, are observed as a narrow singlet at 11.90 ppm, and a broad singlet at 12.97 ppm respectively. Observation of these peaks confirms the regioselectivity of the reaction.

As discussed, the unexpected second order effects at 298 K, are most likely rotameric in nature, and caused by steric interaction between the Fmoc and benzyl groups. This is apparent due to the signals from these groups being most effected. It is also possible that it is an aggregation effect caused by pi-pi stacking interactions between molecules. However, as the second order effects persist even when the sample is diluted, it is more likely the intramolecular rotameric interaction discussed.

2.4 The scale up synthesis.



i: BnBr, K_2CO_3, DMF, 98.5% (mean 2 batches). **ii:** NIS, H_2SO_4, DMF, 91.5 % (mean 2 batches). **iii:** $ZnBr_2$, $Pd(PPh_3)_2Cl_2$, DIPEA, Methyl propiolate, DMF, 66 ^oC, N_2, 77 %. **iv:** TBAF, THF, 66 ^oC. **v:** Boc_2O, DMAP, DCM, 39 % (2 steps). **vi:** Zn, NH_4Cl, Boc_2O, DMAP, THF:Water. **vii:** NIS, H_2SO_4, DMF, 59 % (3 steps). **viii:** Potassium *tert*-butoxide, 1,3-dichloropropene, DMF, 62 %. **ix:** AIBN, TTMSS, Toluene, 90 ^oC, N_2, 70 % **x:** LiOH, THF:MeOH:Water, 100 %. **xi:** 4M HCl in Dioxane. **xii:** Fmoc-Cl, $NaHCO_3$, THF:Water, 80 % (2 Steps).

Scheme 2.39 The scale up synthesis.

Success of the pilot synthesis had provided a reliable synthetic route to the desired solid phase building block **11**. Although this work had afforded enough of the compound to begin preliminary solid phase experiments, it was clear that in order to access a sufficient quantity of **11,** to allow significant solid phase work to be conducted, the synthesis would need to be repeated, either several times, or on a much larger scale. It was decided that

one large scale synthesis would be a more time efficient option. The decision was made to undertake this process during a 3 month placement at the laboratories of the author's industrial sponsor. This location would provide a superior infrastructure for large scale synthetic work, including a purpose built scale up laboratory, and a team of experienced process chemists who would be available to advise.

Scheme 2.39 depicts the scale up synthesis. It can be seen that some of the conditions are different from those described for the pilot work. As will be discussed, these changes were made for either process, or safety reasons.

Introduction of the benzyl ether was conducted as per the conditions used in the pilot synthesis. Isolation by precipitation over crushed ice and water meant that 78.05 g of **22** was easily accessed in two batches with a mean yield of 98.5 %.

Iodination of **22** saw the first deviation from the pilot synthesis. This reaction was originally conducted in THF, and isolated after extractive work up by column chromatography. It had been observed that a thick and foaming emulsion initially formed on addition of the NIS when the reaction was conducted in this solvent. What caused this emulsion was not known, and it was deemed to be both, a practical, and safety concern for large scale work.

The reaction was therefore trialled on a small scale in DMF. In this solvent the foaming emulsion was not observed. The reaction also proceeded significantly faster than in THF, and at increased yield. Furthermore, isolation of the product was possible by precipitation, thus avoiding large scale chromatography at this early stage. DMF was therefore chosen as the solvent for the large scale reaction.

During the test reaction the internal temperature was monitored, and a small exotherm was observed. As a precaution the large scale reaction was performed in a room temperature water bath, and the internal temperature carefully monitored. The water bath could be used to cool the reaction by the addition of ice if a large rise in temperature was seen, although this proved to be unnecessary.

The improved conditions allowed the synthesis of 105.64 g of **23** in two batches with a mean yield of 91.5 %.

For introduction of the alkyne, the more reliable Negishi coupling would be used in place of the Sonogashira reaction. The Negishi coupling was seen to have the added advantage of proceeding in THF. The increased volatility of this solvent over DMF, was advantageous, because it allowed the reaction to be easily concentrated by rotary evaporation, and the product isolated directly by short column chromatography. This was

important because extractive work up had proved impossible, as the interface was obscured by the opaque black appearance of both phases.

Unfortunately, when the Negishi coupling was trialled on a modest 10 g scale in THF the reaction stalled. As the Sonagashira coupling had not worked at all in THF, and was only successful in DMF, it was questioned whether DMF could also serve as a better solvent for the Negishi coupling. It was also questioned whether $Pd(PPh_3)_4$ might be a better catalyst than $Pd(PPh_3)_2Cl_2$, as this was the preference of Hiroya *et al.*

As a result, an array of four 100 mg scale reactions was set up to compare the different reaction conditions. Monitoring by LC-MS showed the combination of DMF and $Pd(PPh_3)_2Cl_2$ to be superior to the other reaction conditions, proceeding faster and with fewer side products, although all reactions went to completion overnight. Figure 2.7, shows the HPLC trace for each set of conditions after 3 hrs. Figure 2.7c is the combination of DMF and $Pd(PPh_3)_2Cl_2$ and has clearly proceeded nearer to completion after 3 hours than all other conditions.

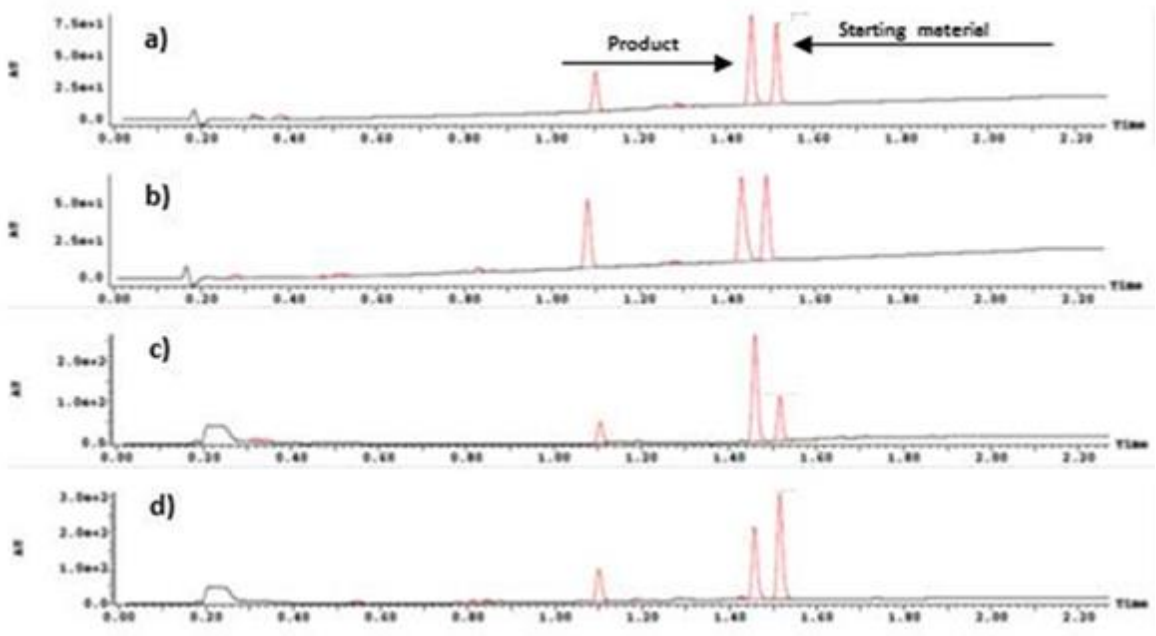


Figure 2.7 HPLC of Negishi coupling test reactions after 3 hours. (a) THF and $Pd(PPh_3)_2Cl_2$. (b) THF and $Pd(PPh_3)_4$. (c) DMF and $Pd(PPh_3)_2Cl_2$. (d) DMF and $Pd(PPh_3)_4$. HPLC conditions: Acquity CSH C18 50x2.1 mm column. Solvent A: Water + 0.1 % HCOOH. Solvent B: Acetonitrile +0.1 % HCOOH. Gradient, 0.0-0.2 min 5 % B, 0.2-1.8 min 5-98 % B, 1.8-2 min 98 % B. Flow rate 1 mL per min. Column temperature 50 ^{o}C.

The combination of DMF and $Pd(PPh_3)_2Cl_2$, was clearly the most reliable choice. However there were still concerns surrounding the work up. Anecdotal advice that DMF can often

be substituted for the more volatile DME, led to this solvent also being trialled. However, the reaction in DME proved to be even slower than in THF.

These results led to the replacement of THF with DMF for the Negishi coupling. The reaction was progressively repeated on increasing scales. First a 1 g scale reaction was trialled. This went to completion overnight without issue. However, a subsequent attempt to repeat the reaction on a 3.8 g scale led to the reaction failing to go to completion overnight. The only difference between the 1 g and 3.8 g scale had been the concentration. The 1 g scale reaction had been conducted at a concentration of 0.09 M. In an attempt to reduce the quantity of DMF, the 3.8 g scale reaction had been conducted at a concentration of 0.205 M. Repeating the 3.8 g scale reaction at the lower concentration of 0.09 M was successful. This concentration was then trialled on a 10 g scale, and was again successful. It therefore seemed clear that the reaction was concentration dependent.

Before the new conditions could be used to convert the main batch of **23**, the work up would require optimisation. Advice from the process chemists in the scale up laboratory, was that it would not be practical to concentrate the necessary volume of DMF by rotary evaporation. As discussed, direct extractive work up was not possible due to an inability to visually discrimination between the two phases. A number of alternative work ups were therefore trialled to recover the product from the DMF.

First an attempt was made to subject a sample of the reaction mixture directly to short column chromatography. Sadly, the product simply eluted with the DMF, and even after this process, subsequent extractive work up was still not possible.

Next a sample of the reaction mixture was treated with a 1:1 mixture of $MgSO_4$ and activated charcoal. The resulting slurry was filtered through a column of diatomaceous earth, before attempting the extractive work up. Again it still proved impossible to observe the interface.

Finally a sample of the reaction mixture was poured over crushed ice and water. The chocolate coloured precipitate was collected by filtration. LC-MS analysis showed the precipitate to be crude product, and that the black filtrate contained no product. The product could be isolated from the crude precipitate *via* short column chromatography.

The Negishi coupling was now ready to convert the remaining large batch of **23**. However, the scale of the reaction triggered an institutional safety protocol during the risk assessment. The presence of the nitro and halide substituents of the aniline raised

concerns about the thermal stability of the starting material. No exotherm had been detected during monitoring of the internal temperature of the small scale experiments. However, the planned increase in scale was large, and therefore assessment of a safe working temperature range was required to gain approval to conduct the reaction.

The assessment was made using heat flux differential scanning calorimetry (DSC). DSC is an analytical technique which measures the energy required to maintain a constant temperature between a sample and inert reference at a given temperature. In a heat flux instrument, the sample and reference cell are contained within the same furnace. Temperature differences between the sample and reference are measured as a function of heat flow between the two cells.[127] Figure 2.8 shows the DSC curve of **23**, as observed over a dynamic temperature range. Here the sample and reference were heated at a rate of 4 ^{o}C per min, from 10 ^{o}C to 400 ^{o}C, under an atmosphere of nitrogen.

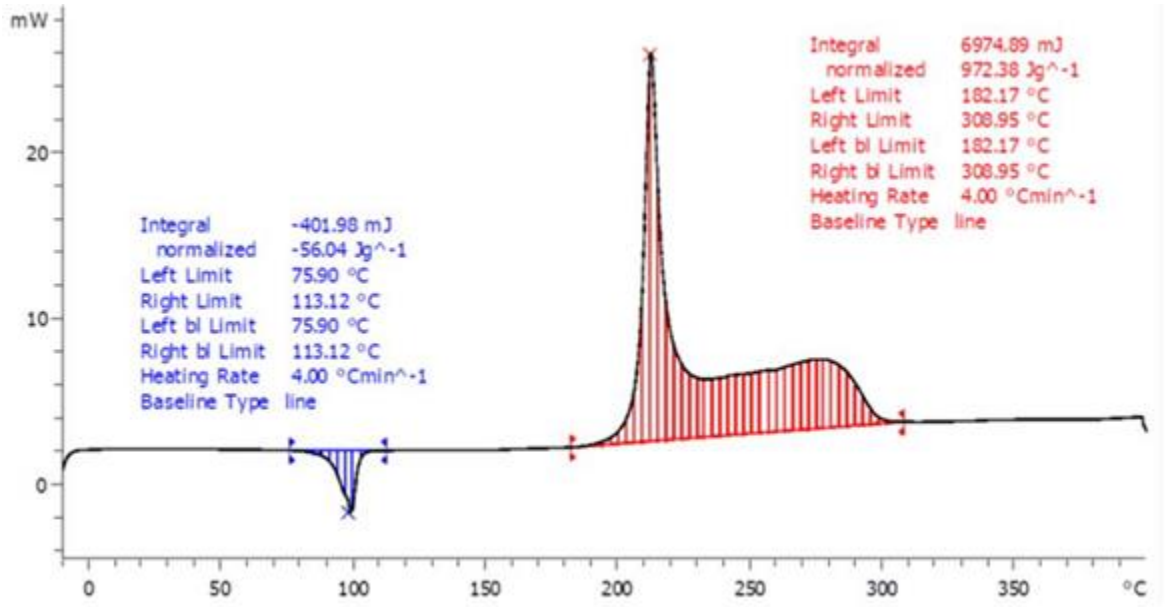


Figure 2.8 Dynamic DSC curve of **23**. Instrument: Mettler Toledo DSC 1. Pan type: 40µl gold high pressure pan. Reference: empty pan. Sample size: 8 mg. Temperature: 10 ^{o}C to 400 ^{o}C at 4 ^{o}C per min. Atmosphere: Nitrogen.

Melting of the solid is observed as a small endothermic event at 100 ^{o}C. The compound then appears to be stable up to 182 ^{o}C, where the onset of a large exothermic event is observed. This decomposition was well above the planned temperature of the reaction (66 ^{o}C). However, as an exothermic event was observed, a second precautionary experiment was performed. This involved measuring the DSC curve of **23** at a constant temperature of 130 ^{o}C for 12 hours, followed by a second dynamic run. No exothermic events were observed at 130 ^{o}C over the course of the experiment, and the onset of decomposition in

the following dynamic run was consistent with that of the initial experiment. These results led to approval of the reaction conditions. In retrospect, it seems likely that the thermal stability of the product of this reaction should also have been tested, as the alkyne is likely to increase the explosive potential of **26** over **23**.

In order to reduce the size of the necessary round bottom flask to 3 L, the large scale reaction was conducted in two batches and the crude precipitates combined. Short column chromatography afforded 62 g of **26** at a yield of 77 %. This was comparable to the best yield obtained for the reaction in DMF on a small scale which was 84 %.

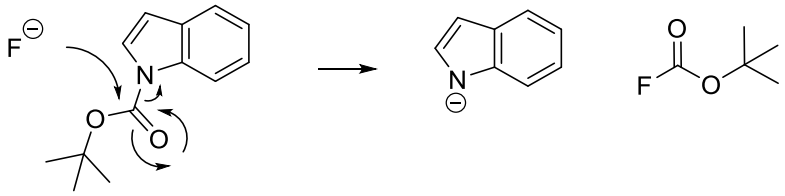
Due to the size of the glassware, in order to allow the heat source to be quickly removed and the reaction cooled if exothermic run away was observed, the heating block was placed on a jack. This would allow it to be lowered away from the clamped round bottom, and replaced with an ice bath if necessary.

The reason for the reaction being concentration dependent is not known. The dark colour of the reaction suggests the formation of Pd black.[128] It is possible that this process is accelerated at higher concentrations, reducing the efficiency of the reaction. Further optimisation may be possible to allow the reaction to proceed at higher concentrations; for example, lowering the catalyst loading. The reason why DMF is superior, is also not known. This may simply be the improved solubility of all the reagents. Although it is also possible that coordination of the polar aprotic solvent, is able to stabilise the Pd(0) species thought to be responsible for the coupling reaction. [129]

In the pilot synthesis, the next two products, the indole **15**, and the Boc-protected indole **16**, were both isolated by column chromatography. It was clear from the experience of purifying the product of the Negishi coupling, that chromatography on a large scale, even short column chromatography, is operationally demanding, and not a trivial step. It requires large volumes of solvent which can be difficult to handle, and the process requires significant planning.

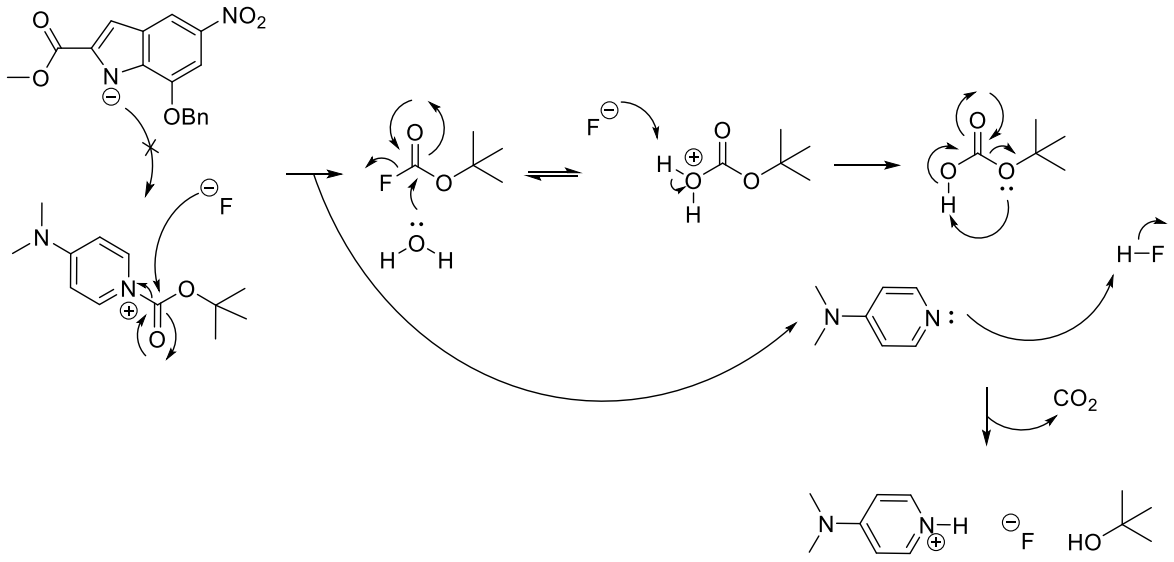
It was therefore decided that an attempt should be made to see if Boc-protection of **15** could be effected without prior chromatographic isolation. Initially a one pot procedure was trialled. Here, **26** was treated with 2 equiv. of TBAF in THF at reflux. On completion of the reaction, as monitored by LC-MS, the mixture was cooled to room temperature, and 1 equiv. of DMAP, and 3 equiv. of Boc_2O added. However, no reaction of the indole was observed.

Inspection of the literature revealed that TBAF has been reported as an effective reagent for the *N*-Boc deprotection, particularly of aromatic heterocyclics.[130] This most likely explains the failure of the one pot procedure. The authors postulate the mechanism depicted in scheme 2.40 to be the most likely pathway of this reaction. Here the fluoride anion acts as a nucleophile at the carbonyl of the carbamate, and the resulting tetrahedral intermediate collapses with the indolic amide acting as the leaving group. The alkyl fluorocarbonate generated is presumable hydrolysed, and the resulting HF protonates the indolic amide.



Scheme 2.40 Mechanism of *N*-Boc-deprotection by TBAF.

It is therefore possible that the presence of TBAF interferes with the Boc-protection, either by promoting the instantaneous deprotection of any product that does form, or perhaps more likely by competing with the indole for reaction with the *tert*-butyl carbamate pyridinium cation (scheme 2.41).



Scheme 2.41 Possible inhibition of Boc-protection reaction by the fluoride ion of TBAF.

With this in mind, it was decided to see if Boc-protection would proceed, if the TBAF was first removed by washing with water. Once the indole had formed, the THF was removed by rotary evaporation. The residue was dissolved in ethyl acetate and washed 3 times with water. After concentration of the ethyl acetate, the residue was dissolved in DCM,

and treated with 1 equiv. of DMAP, and 3 equiv. of Boc_2O. This proved successful and Boc- protection proceeded without issue.

This procedure was repeated on the remaining large batch of **26**, and afforded 31 g of the Boc-protected indole **16** after column chromatography This represents a yield of 40 % over the two steps. The low yield was expected, and is a result of the intrinsically low yielding TBAF cyclisation step. It is not a consequence of the 'telescoping procedure' used to avoid chromatographic isolation of indole **15**. The yield over the 2 steps is comparable to the mathematic product of the 2 best individual yields from small scale reactions (0.58 x 0.75 x 100 = 43 %).

The next steps were reduction of the nitro group and subsequent Boc-protection of the amine. As discussed, a convenient one pot procedure had already been developed during the pilot synthesis. It was decided that the reliability of the following iodination reaction might allow avoidance of the prior chromatographic isolation of the Boc-protected amine **18**. Therefore this was trialled on a small scale. A one pot procedure was not attempted as the acidic catalyst required to activate NIS, seemed incompatible with the use of DMAP. Instead crude **18** was subjected directly to iodination after aqueous work up. The conditions for iodination were also altered from the pilot synthesis. Here toluene and catalytic acetic acid, were replaced with DMF and catalytic H_2SO_4, as these conditions had provided faster reaction times for iodination of the aniline. It was also hoped that the product may be isolated by precipitation over crushed ice and water. Iodination proceeded quickly under the new conditions. Precipitation of the product was observed when the reaction mixture was poured over crushed ice and water, however it was very fine, thus in order to maximise yield it was decided to recover it by extracting with ether.

The new conditions were repeated on the remaining large scale batch of **16**. The reduction/Boc-protection step was conducted in two batches to avoid the need for mechanical stirring of the suspended zinc, and the crudes combined for iodination, affording 26 g of **29** after column chromatography. This represents a yield of 59 % over the 3 steps (nitro group reduction, Boc-protection, and iodination). In retrospect, it is likely the chromatography of **29**, could also have been avoided by collecting the precipitate by filtration. This seemed possible when conducted on the large scale as the precipitate was more substantial, however it was decided not to risk deviating from the trialled procedure with the large batch. Also at this stage in the synthesis, the quantities of material were now at a scale, where chromatography had become more practical. Although, still larger than the typical scale of work being conducted in the research laboratories, it was now

possible to use the widely accessible standard equipment, and booking of the specialist large scale flash chromatography instrument in the scale up laboratory was not required.

Alkylation of **29** by 1,3-dichloropropane, produced another hurdle. Although anticipated to be straightforward, the conditions used in the pilot synthesis, could not be approved for scaled up. The scale of the planned reaction, although considerably smaller than those at the beginning of the synthesis, still fell within the institution's classification of large scale for the purpose of risk assessment. The problem was the combination of DMF and NaH. This combination had recently (the previous week) been banned by the institution for work classified as large scale, and was strongly discouraged on a small scale.

This combination has been documented to sometimes give rise to a self-accelerating and uncontrollable exothermic reaction between the hydride and solvent.[131] The reaction results in the decomposition of the DMF, and produces CO and H_2 gas. There is a significant risk of explosion and fire if the runaway reaction occurs.

The combination is common in the literature,[132] and often proceeds without issue on a small scale. However, there have been several major incidents officially reported with this combination on a plant scale. The risk is greatest if the reaction is heated above 50 ^{o}C, however calorimetry data produced during an investigation after one industrial accident, suggested the onset of this process can be as low as room temperature.[133, 134]

The probability of experiencing such an incident when using DMF and NaH, is difficult to quantify, and probably very low particularly when conducted at low temperature. However, there appears to be a consensus amongst process chemists that the risk associated with this combination is real, dangerous, and unpredictable. There is argument the risk should be better publicised, as the specific danger of this combination is not listed in the material data safety sheets of either substance, and there are many published procedures which use this combination, which can be repeated by other researchers without a full appreciation of the risk.

To circumvent this problem an alternative method was sought. First, an attempt was made to change the solvent. The choice was limited due to the poor solubility of the base. NaH is often used as a suspension in THF.[135] Therefore, the reaction was trialled on a small scale in 2-MeTHF. Unlike THF and DMF, 2-MeTHF is immiscible in water, and this was deemed to be advantageous for the work up. However no alkylation was observed by LC-MS, although some hydrolysis of the ester was seen. Failure of the reaction is most likely due to poor solubility of the deprotonated carbamate intermediate. As such, an attempt was made to find an alternative base which could be used in DMF. It was found that 2

equiv. of potassium *tert*-butoxide could be used in place of 1 *equiv*. of NaH, and the reaction proceeded with comparable yields, at room temperature. As such these conditions were used to alkylate the remaining large batch **29**, providing 18 g of **30,** at a 62 % yield, after column chromatography.

Now attention turned to the radical cyclisation. Here the scale of the reaction again raised concerns. AIBN is known to be potentially explosive. The use of DSC to assess the danger is of course of no use, as the thermal decomposition of this reagent is the very reason it is being used. An institutional operational protocol existed for the use of AIBN, which limited the quantity of the reagent permitted within a single reaction. In order to comply with this protocol the reaction was conducted in two batches. As a precaution the reactions were performed behind a blast shield. The reaction is run under an atmosphere of nitrogen. In order to achieve this and allow for the escape of pressure, a constant steam of nitrogen was applied, and gas allowed to leave the system *via* a gas bubbler fitted to the top of the reflux condenser. This also produced a weak point in the apparatus which could disconnect in the event of a sudden increase in pressure, thus reducing the risk of explosion of the round bottom flask.

This procedure afforded 10.28 g of **10** at a combined yield of 70 %, after column chromatography. This is considerably lower than the 90 % yield achieved in the pilot synthesis on a 2 g scale. It is not clear what caused the drop in yield. The concentration was not increased on scale up, in order to retain the selectivity for the intramolecular ring closure. It is conceivable that on a larger scale, the greater volume of solvent takes longer to reach the desired temperature, and that this has a negative effect on yield. The 5-exo-trig cyclisation is known to be more efficient at higher temperatures.[93] It may be that the delay in heating from the temperature at which AIBN begins to decompose and yield the first primary radicals, to the optimum temperature for 5-exo-trig cyclisation, leads to more formation of side products, such as perhaps the reduction of the aryl radical before cyclisation can occur. However such a side product was not isolated. It is possible that the yield could have been improved by heating the reaction mixture to full temperature before addition of the AIBN. Although, this would seem operationally risky, as the sudden heat shock could promote explosion of this reagent. It may be that a higher concentration would be more optimal at larger scales.

The remaining steps proceeded without issue. Some changes were made from the pilot synthesis, but these were simply to improve the operational convince of the process.

Hydrolysis of the ester was achieved using the same reaction conditions as in the pilot synthesis. However, the work up was simplified. Instead of extracting the free acid after

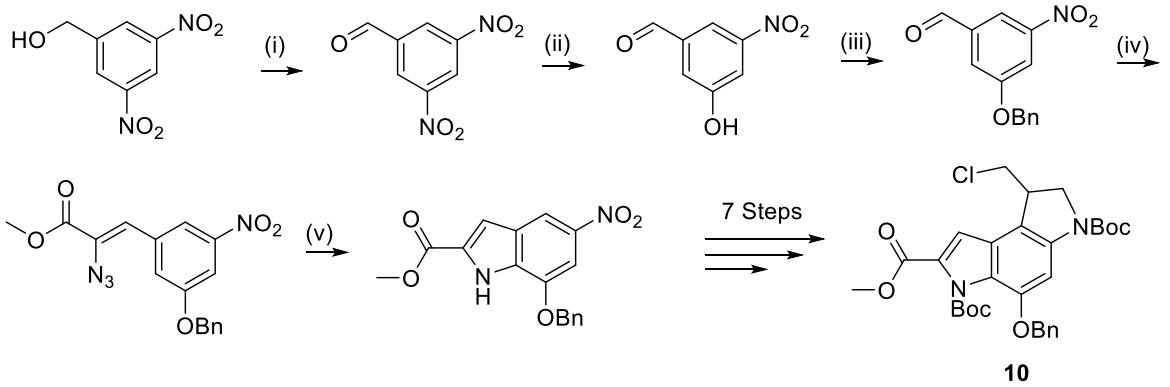
acidification of the Li salt, the resulting precipitate, which was more substantial on scale, was collected by filtration. The organic solvents were also removed prior to acidification, to avoid the risk of esterification by the MeOH, and to aid precipitation. This allowed for the avoidance of chromatography, and afforded 10 g of **31** in quantitative yield.

Boc-deprotection was affected using 4 M HCl in dioxane. This was chosen over 4 M HCl in ethyl acetate, used in the pilot synthesis, due to its commercial availability. This removed the inconvenience of having to make the required volume of acidic solvent. Subsequent Fmoc-protection was achieved as per the pilot synthesis, except for a minor adjustment to the work up, where ethyl acetate was replaced with 2-MeTHF for extraction. This was due to better solubility of the product in this solvent, making the extraction process more efficient. After column chromatography, 8.3 g of racemic **11** was isolated at a yield of 80 % over the 2 steps.

The synthesis had been successfully scaled up to provide 8.3 g of the racemate in an overall yield of 3 % over the 13 steps. A better overall yield could be given for the route itself, based on the best results achieved for each reaction over the course of both the pilot and scale up synthesis. However, this figure feels artificial and less valid. The yield given is a true result for the scale up synthesis itself, and should therefore be reproducible at the same scale.

The synthesis also represents one of the most concise routes to the common duocarmycin intermediate **10**, at 9 steps, (or 10 if you consider the one pot reduction Boc-protection to be two steps). Tietze *et al.*'s Fischer indole route, employs 11 steps to the same intermediate (or 12 if you consider their one pot diazonium formation and reduction to give the hydrazine to be two steps). Using the published data,[96] their route afforded 0.855 g of **10** at an overall yield of 14 %. In comparison our route afforded 10.28 g of **10** at an overall yield of 7 %. The published yield is superior despite involving more steps. However, this route has not been proven to be practical on a large scale. Our route has afforded 12 times the quantity of **10** at only a 2 fold decrease in yield.

The same comparison cannot be made with Hiroya *et al.*'s similar Negishi coupling strategy,[110] as they do not go on to synthesise the same intermediate. However, despite requiring a separate annulation step, our route is a step shorter to the first nitro indole intermediate. Furthermore their indole is protected by a mesylate group, so an additional 2 steps to change the protecting group would be require to access **10**. In addition the incomplete *in situ* cyclisation to the indole, would require a demanding large scale separation from the uncyclised product, which would seem to limit the scalability of this route.



(i) Swern oxidation, 95 %. **(ii)** benzaldehyde oxime, K_2CO_3, DMF, 90 °C. **(iii)** BnBr, DMF, 86 %. **(iv)** methyl azidoacetate, NaOMe, MeOH, 4 °C, 78%. **(v)** xylenes, 140 °C, 7 h, 68%.

Scheme 2.42 Boger *et al.*'s Hemetsberger-Rees indole synthesis route.

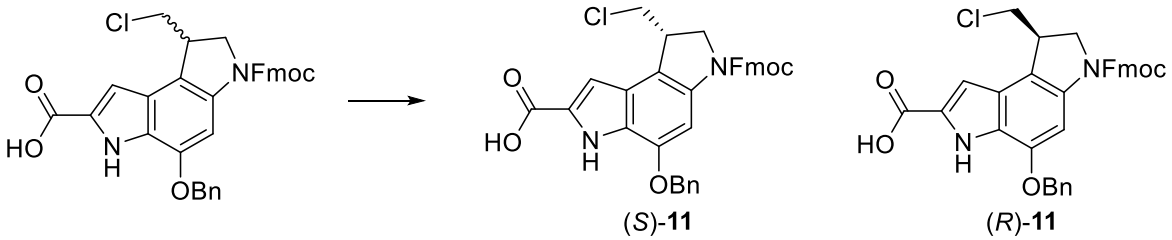
Another concise route to **10** which converges on the Tietze synthesis, is that of Boger *et al.* which employs a Hemetsberger-Rees indole synthesis to reach the first indole intermediate.[90] This route is also longer than ours at 11 steps (or 12 if you count their one pot nitro group displacement and benzyl ether protection as two steps). Again despite more steps they appear to have accessed **10** with a superior overall yield of 30 %. However, again this route has not been proven to be practical on a large scale. Boger *et al.* report the synthesis of 0.025 g of **10**. Therefore, based on this figure our synthesis has afforded 411 times the quantity at only a 4 fold drop in yield. Furthermore, a number of steps appear difficult to scale. Firstly, the temperature sensitive, and operationally demanding early Swern oxidation, might be troublesome on a large scale. Also, the safety of preforming the Hemetsberger-Rees step on a large scale might be questionable, as it involves heating a compound which contains both an azide and an aromatic nitro group to high temperatures (140 ^{o}C). This might present an unacceptable risk of explosion if this compound proved to be thermally unstable by DSC. Similarly, it is likely a large quantity of the explosive NaN_3 would need to be handled, to provide the methyl azidoacetate used to introduce this group.

Of course it is not possible to know for sure, that either Tietze's Fischer, or Boger's Hemetsberger-Rees indole route, could not be scaled up, and maintain their superior yields. However, our Negishi coupling strategy does at least provide a shorter route to the first common indole intermediate, and has been proven to be scalable to provide a multi gram quantity of **10**. Furthermore, subtle changes in the later steps of the Tietze/Boger route, have been made here which improved their scalability. Most significantly, the employment of a one pot procedure for nitro group reduction and Boc-protection.

There are clearly significant areas where further improvement could be made in this synthesis. Firstly, the intrinsically low yielding TBAF cyclisation represents a significant early bottleneck. It seems this is a hindered annulation, most likely due to both, the deactivating effect on the aniline of the *para* nitro group, and the necessity to polarise the alkyne in such a way that places the electron deficient end vicinal to the carbonyl of the ester. TBAF appears to possess specific properties which allow this ring closure to take place, where for example, gold(III) catalysis has failed. It may be possible to improve the yield of the TBAF method itself, for example by using a catalytic quantity of the reagent. A wider such for alternative catalysts may also find a higher yielding reaction. It may be possible to reduce the nitro group prior to the Negishi coupling, and if this was to improve the favourability of ring closure, it might lead to *in situ* cyclisation. If this was the case the indole and amine might be protectable in one step, reducing the length of the synthesis dramatically. Of course di-Boc-protecting the amine may be a problem, and the aryl halide may not be stable to reducing conditions.

Column chromatography has been avoided at several steps during the scale up synthesis. However, to scale the synthesis further, more of the chromatographic steps would need to be avoided. It seems likely that recrystallization could be used to replace the remaining chromatography, but this has not been proven. Furthermore, it may be that the limits of the 5-exo-trig radical cyclisation have been reached, as the yield dropped significantly during the scale up synthesis. In addition, the use of the explosive AIBN is not ideal.

2.5 Chiral resolution of 11.



Scheme 2.43 Chiral resolution of **11**.

The 8.3 g of **11**, accessed from the large scale synthesis, was combined with a second smaller batch (1.9 g), which was synthesised alongside the large batch to test scale up processes. 9.8 g of the combined racemate was taken forward for chiral resolution. This was achieved using preparative supercritical fluid chromatography.

A supercritical fluid, is a fluid state of matter which is distinct from both the liquid or gas phases. Such a state exists when a substance is heated, and compressed, to a point where the densities of the liquid and gas phases are equal. Above these critical temperatures and pressures, the liquid and gas phases coalesce to form a single fluid with no observable interface. Such fluids possess physical properties that lie somewhere between those that would be expected for a particular substance in either its liquid or gas phase. This typically results in fluids with liquid like densities, but viscosities that are more comparable to a gas.[136]

Such properties mean that supercritical fluids can often serve as excellent solvents for chromatographic separations. Their liquid like densities allow for the dissolution of solids, while their gas like viscosities lead to improved mass transfer kinetics, and permit the use of fast flow rates with high acuity columns. As such, highly efficient separations can be achieved. This makes supercritical fluid chromatography particularly attractive for preparative scale work, where the short analysis times, and low solvent consumption, are particularly advantageous.[137]

In order to be useful as solvent for chromatography, a supercritical fluid must be accessible at operationally feasible temperatures and pressures. The critical temperature and pressure of CO_2, is 304.12 K, and 73.74 bar respectively.[138] A supercritical fluid of CO_2 is therefore easily attainable, and it is a very common solvent for supercritical fluid chromatography. In addition it is relativity cheap, non-toxic, and non-flammable.

Supercritical CO_2 is very non-polar, and it is seldom used as the sole component of the mobile phase. Organic modifiers, such as polar co-solvents, are usually employed to increase elution and solvation. Acidic additives such as TFA are also common. They can be used to suppress ionisation of analytes and improve resolution. The critical temperature and pressure of such binary or tertiary mixtures is likely to be higher than the typical operating conditions of most separations. It has therefore been argued that the majority of supercritical fluid chromatography takes place under subcritical conditions. However, by convention, mobile phases which comprise a supercritical fluid as their major component are generally referred to as supercritical fluid chromatography.

The application of supercritical fluid chromatography for the chiral resolution of racemic mixtures is well established.[138] Enantioselective retention is most commonly affected using a chiral stationary phase, although chiral mobile phase additives are also used. The low viscosity of the mobile phase, means that pack columns are typically favoured of open tubular designs. Many stationary phases have been developed. Amongst the most common are the polysaccharide stationary phases, comprising of various derivatives of cellulose or amylose. The Chiralpak AD-H column is a typical example. Most columns in use today have not been specifically designed for supercritical fluid chromatography, and are also used in HPLC. However, in addition to the advantageous properties of supercritical fluids already discussed, the small size of CO_2 has been suggested to confer a specific advantage in chiral separations. This is because it is easily displaced from the hydrophobic pockets of chiral selectors, and leads to increased retention of the analyte.[138]

Here, the chiral resolution of **11** was affected using a Chiralpak AD-H column (250 x 30 mm, 5 micron), and an isocratic flow of 50 % CO_2, and 50 % IPA containing 0.1 % TFA, at 45 mL per min. The back pressure was regulated at 10 MPa, and column temperature controlled at 40 ^{o}C. The racemate of **11** (9.8 g) was dissolved in THF:MeOH 1:1 (100 mL), and 1.25 mL (125 mg) injected every 9 mins. Fractions were combined and dried to afford 2.82 g of peak 1, and 3.1 g of peak 2, both as a cream solid. Figure 2.9 shows the separation from one injection.

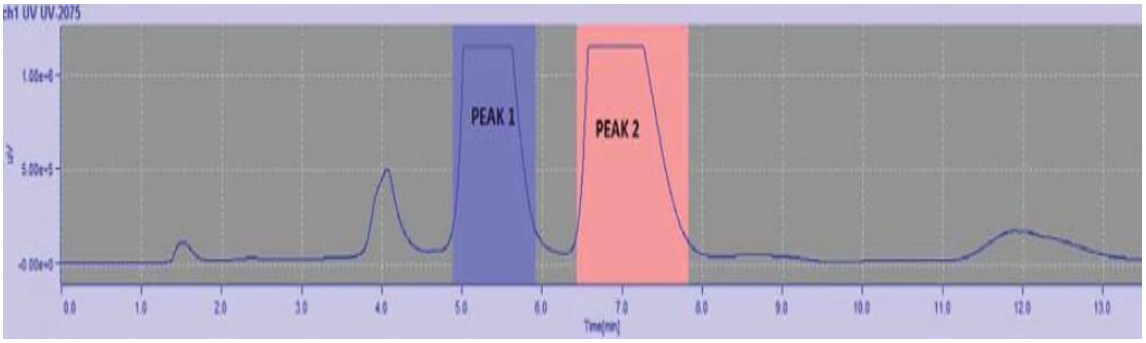
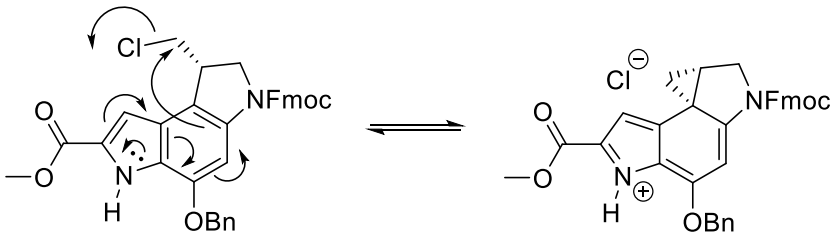


Figure 2.9 Chromatogram of preparative super critical fluid separation of racemic **11** monitored by UV at 220 nm. Method details are described in the main text.

Peak 1 was assigned as the natural (*S*) enantiomer based on the sign of specific rotation matching that of the well characterised seco Boc-protected derivative of the duocarmycin alkylation subunit. Peak 1 $[\alpha]_D$ –20 (*c* 0.05, DMF). Peak 2 $[\alpha]_D$ +20 (*c* 0.05, DMF). This is of course not proof of the absolute stereochemistry of **11**, but it is a reasonable assumption based on all previously reported seco-duocarmycin analogues exhibiting small (-) rotation for the (*S*) enantiomer.

Analytical supercritical fluid chromatography of peak 1, showed a 7 % impurity with a similar retention time to peak 2. The mass of this peak suggested the loss of Cl as opposed to racemization. NMR analysis showed no evidence of this impurity, with the ^{1}H NMR of both enantiomers being identical to that of the racemate, with the exception of a small amount of residue IPA. We propose that the observed impurity may be an artefact of supercritical fluid chromatography, and represent a transient alternative spirocyclisation through the indole nitrogen; perhaps promoted by the pressure (scheme 2.44). Such a spirocyclisation has been reported in limited cases, for seco forms of duocarmycin SA when spirocyclisation through the phenol is blocked by benzyl ether protection.[139]



Scheme 2.44 Possible transient alternative spirocyclistion during supercritical fluid chromatography.

2.6 Conclusion

In summary, the pilot and scale up synthesis has realised a shorter route than previously published to the common racemic duocarmycin intermediate **10**, and one which has been successfully conducted on a multigram scale. Conversion of **10** to the novel 'building block' **11** was straightforward, and chiral resolution *via* supercritical fluid chromatography provided both enantiomers to explore their application to the solid phase synthesis of novel duocarmycin analogues.

3

Chapter Three

*The first application of **11** towards the solid phase synthesis of duocarmycin analogues: coupling to resin bound amino acids.*

3.1 Aims.

The work described in this chapter was carried out in order to address the following aims:

- To explore the utility of the newly synthesised **11** as a 'building block' for Fmoc solid phase synthesis.

- To explore the effect of *C*-terminal amino acid substitution on the cytotoxic activity of the seco-duocarmycin alkylation subunit.

This work was carried out with a view towards the potential future application of using Fmoc solid phase synthesis to incorporate this functionality into water solubilising peptide based linkers for antibody drug conjugation (see chapter one).

3.2 Sold phase peptide synthesis.

The introduction of solid phase synthesis in 1963[140] significantly improved the synthetic accessibility of peptides. Since its inception the technique has evolved considerably, and has found far reaching applications beyond simply the construction of polyamide targets.[141] However, it is still within the field of peptide and oligonucleotide synthesis that this method dominates. When first introduced, the method was received with unveiled hostility by many within the peptide synthesis community.[142] It certainly represented a paradigm shift from the classical solution approach, and early work was not altogether convincing. However, many initial challenges were soon overcome, and the method was validated with results that only the most harden critic could deny as impressive.[143] It is now arguably the most common method for the synthesis of peptides on a laboratory scale. The importance of this technique saw its inventor Bruce Merrifield later recognised with the 1984 Nobel prize in chemistry.[144]

The basic concept is elegant in its simplicity.[145] Synthetic transformations are performed on a substrate which is covalently anchored to an insoluble support. The solid supported product is isolated by filtration, allowing unbound side products and reagents to be simply washed away. This operational simplicity allows for the use of large excesses of reagents to drive reactions to completion. As multiple sequential reactions can be performed without the need for intervening purification, this technique can significantly reduce the time required to perform multistep syntheses. Furthermore, the use of reaction vessels designed to allow *in situ* filtration, under pressure, or by the application of a vacuum, mean multistep syntheses can effectivity be performed as one pot processes. This can be particularly advantageous when working on a small scale, where in solution phase chemistry the additive attrition of material during the work up and purification of individual reactions can limit the effective starting scale of a long synthesis. The final product is isolated by cleavage from the solid support.

The need for orthogonal cleavage conditions, and a cleavage product that introduces a desired structural motif can limit the scope of solid phase synthesis. The repetitive nature of peptide synthesis and the relative simplicity of the synthetic strategy means it is an ideal area for application of solid phase methodology. It is therefore unsurprising that peptides were the first target of solid phase synthesis, and represent its most common application. The operational simplicity of solid phase peptide synthesis means automation

of the process was quickly introduced, and indeed peptide targets are now routinely synthesised by commercially available automated peptide synthesisers.[146]

In solid phase peptide synthesis the target peptide is constructed in a stepwise fashion *via* the sequential coupling of amino acid residues by amide bond formation. The first amino acid is anchored to the solid support through its carboxylic acid. Each subsequent amino acid is coupled with its amine protected. This prevents unintended homocoupling in solution. Subsequent deprotection of the amine after coupling, allows the next amino acid to be coupled, and so on, until the desired peptide is complete. Orthogonal side chain protection is also required to prevent the formation of solid supported impurities. Side chain protection is typical chosen that will be cleaved under the same conditions used to release the completed peptide from the solid support. Linkage to the solid support is usually, but not exclusively, designed to provide a free carboxylic acid after cleavage.

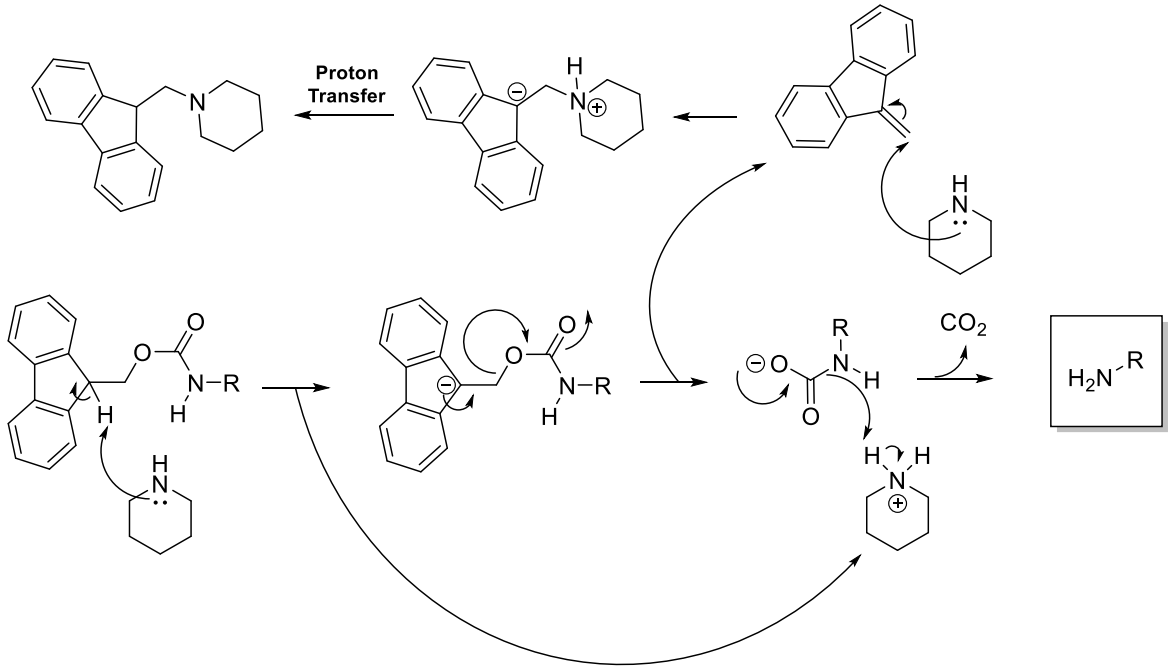
Solid supports typically take the form of cross-linked polymeric resins. Polystyrene based resins are perhaps the most commonly used.[145] The term solid phase synthesis is arguably a misnomer, as reactions do not take place at the surface of a truly heterogeneous system. Solvation, in fact, leads to swelling of the resin, and the formation of a gel like matrix in which the reactions take place. Studies have shown that this environment provides similar access to reagents, as if the individual polymers were free in solution.[147]

The classical Merrifield approach utilises a Boc and benzyl-protecting strategy. The *N*-terminal amino group of the growing peptide is protected by a Boc group. This is cleaved prior to coupling by treatment with TFA. The side chains of trifunctional amino acids are protected by benzyl derivatives, and the peptide anchored to the resin through a benzyl ester. Treatment with anhydrous hydrofluoric acid yields the cleaved unprotected peptide.

In skilled hands this method has proved an extremely powerful technique. However, it does suffer from several drawbacks. Firstly, the protection strategy is not truly orthogonal, and relies on differing acid sensitivity between the Boc and benzyl-protecting groups. As such some undesired loss of benzyl protection is often observed, which can result in the accumulation of resin bound impurities. The strongly acidic conditions required for cleavage can also limit the success of syntheses incorporating sensitive residues. However, perhaps the biggest deterrent to the use of this method, is the hazardous nature of hydrofluoric acid. This reagent is notoriously toxic, and requires demanding safety protocols. It is also incompatible with common laboratory equipment.[148]

Although still in use, the Merrifield approach has been largely superseded by the development of Fmoc solid phase peptide synthesis,[149] which allows the use of milder deprotection and cleavage conditions.[150] This technique utilises the base sensitive Fmoc group for *N*-terminal protection. As such, side chain protection and linkage to the resin can be effected by protecting groups and linkers that are cleavable under mildly acidic conditions.

Fmoc deprotection is typically achieved by treatment with piperidine (scheme 3.1). The Fmoc group is base sensitive due to the acidic nature of the proton at the bridgehead of the fluorenyl group. Acidity is a result of the aromatic stabilisation of the carbanion formed by deprotonation. In the Fmoc group, deprotonation results in the intramolecular attack of the anion at the neighbouring partially-positive carbon. Here the carbamic acid anion serves as a leaving group, and subsequently decomposes releasing CO_2 and the desired amine. A large excess of piperidine is typically used to scavenge the reactive dibenzofulvene by-product.

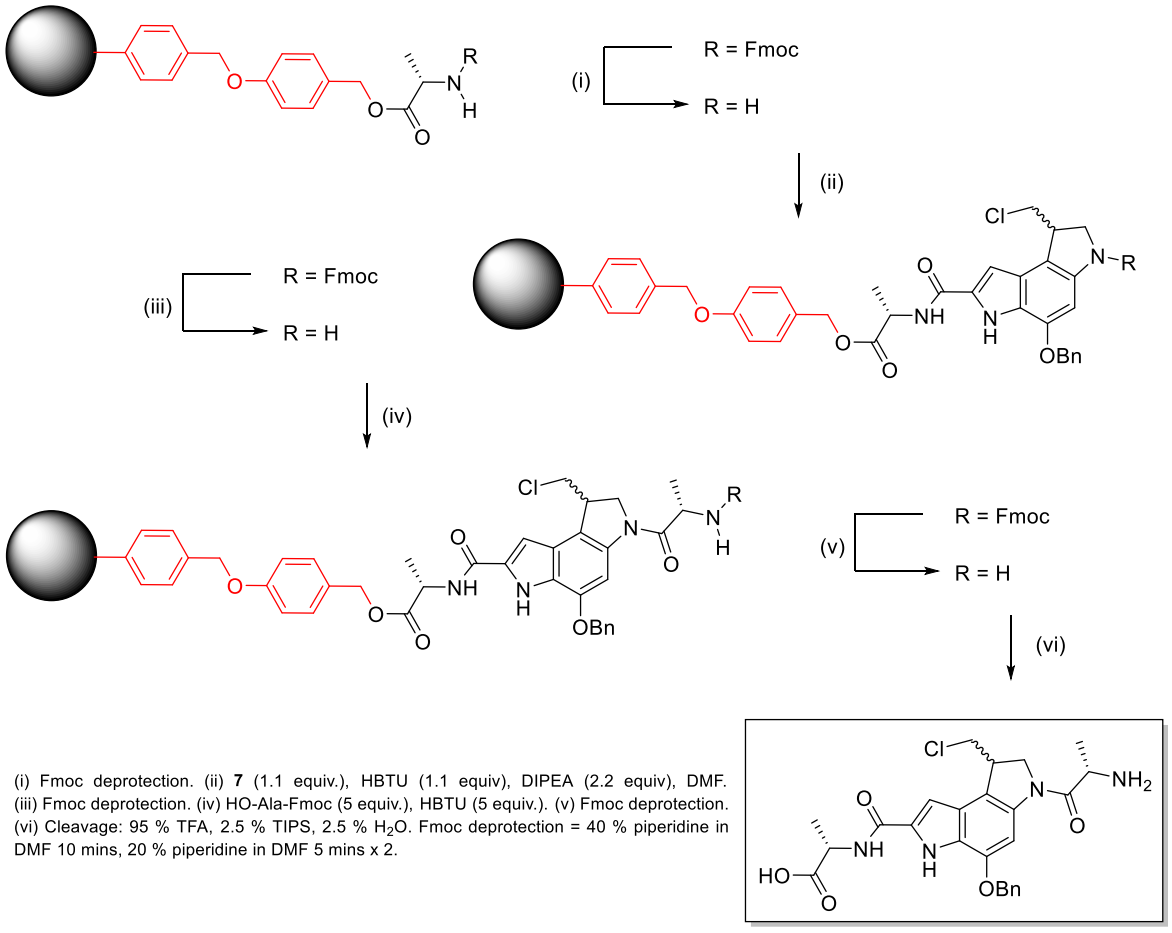


Scheme 3.1 Mechanism of Fmoc-deprotection by treatment with piperidine.

Many resins differing in both their linking strategy and polymeric construction have been developed for use in Fmoc solid phase synthesis. As will be seen, this would prove extremely useful in optimising the solid phase synthesis of duocarmycin analogues using the new Fmoc-protected alkylation subunit **11**.

3.3 The first solid phase experiment and resulting optimisation of cleavage conditions.

In order to assess the suitability of **11** for incorporation into polyamide structures through the application of solid phase synthesis, a simple target compound was envisioned. This consisted of the alkylation subunit of duocarmycin sandwiched between two alanine residues (**32**, scheme 3.2). Standard conditions used routinely in our lab for peptide synthesis were planned to prepare it (scheme 3.2). This preliminary work was undertaking prior to completion of the scale up process with racemic **11** accessed from the pilot synthesis.



(i) Fmoc deprotection. (ii) **7** (1.1 equiv.), HBTU (1.1 equiv), DIPEA (2.2 equiv), DMF. (iii) Fmoc deprotection. (iv) HO-Ala-Fmoc (5 equiv.), HBTU (5 equiv.). (v) Fmoc deprotection. (vi) Cleavage: 95 % TFA, 2.5 % TIPS, 2.5 % H_2O. Fmoc deprotection = 40 % piperidine in DMF 10 mins, 20 % piperidine in DMF 5 mins x 2.

32

Scheme 3.2 First solid phase experiment: synthesis of **32**. Red structure = structure of the Wang linker.

The synthesis began from a commercially available preloaded alanine Wang resin. As such, the manufacturer's specification provided a substitution level with respect to the

quantity of alanine loaded. This allowed for fine control of the stoichiometric ratios of reagents used for the following couplings.

The term Wang, refers to the structure of the linker, which anchors the alanine residue to the polystyrene based resin.[151] This is a very common linker for Fmoc solid phase peptide synthesis. Its structure is depicted in red in scheme 3.2. The amino acid is attached *via* a benzyl ester. However, unlike that used in the Merrifield approach, which requires treatment with hydrofluoric acid to affect cleavage, the structure of the Wang linker, allows cleavage by treatment with TFA. The increased acid sensitivity of the benzyl ester of the Wang linker is the result of the *para* benzyl ether. This provides mesomeric electron donation which stabilises the carbocation resulting from cleavage.

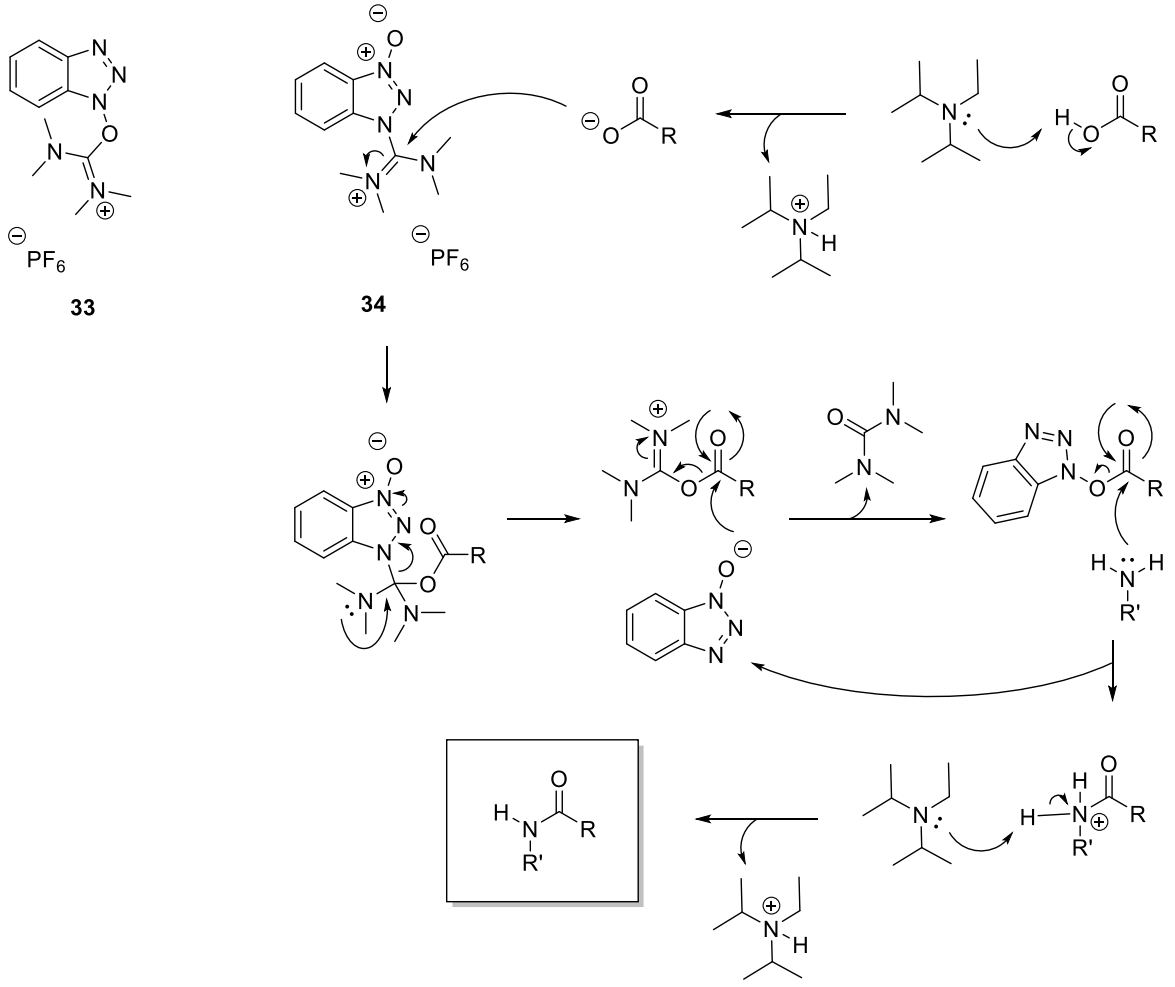
When conducting any solid phase synthesis. It is imperative that the resin is appropriately swelled to allow efficient access to the individual polymer chains. This was achieved by solvating the resin in DCM for 30 mins, followed by DMF for a further 30 mins.

As discussed one of the advantages of solid phase synthesis is the removal of by-products and excess reagents by washing of the resin. This occurs between every step. In the proceeding discussions, this process will not be highlighted, and unless otherwise stated the resin was washed between treatments with an appropriate volume of DMF at least 6 times.

The commercially available resin was supplied with Fmoc-protection of the loaded alanine. This was removed by treatment with 40 % piperidine in DMF for 10 min, followed by 20 % piperidine in DMF for 5 min twice. Separate treatments help to ensure full deprotection, but has not been proven to be necessary. The mechanism of Fmoc-deprotection by piperidine has already been discussed in section 3.1.

Next would follow the first coupling of **11**, and to the best of our knowledge the first solid phase amide coupling of a duocarmycin structure. Amide bond formation is typically promoted by formation of an activated ester of the carboxyl group. Many coupling reagents have been developed. Uronium salts are arguably amongst the most popular choice for solid phase couplings, of which HBTU is a typical example. It is one of the older uronium salts, and more efficient alternatives exist. However, it is relatively cheap when compared to the newer generations of uronium salts, and thus is still commonly used as cost effective option. Indeed, HBTU was the standard coupling reagent used in our lab for solid phase peptide synthesis, and as such was considered first for the application of **11** to the solid phase methodology.

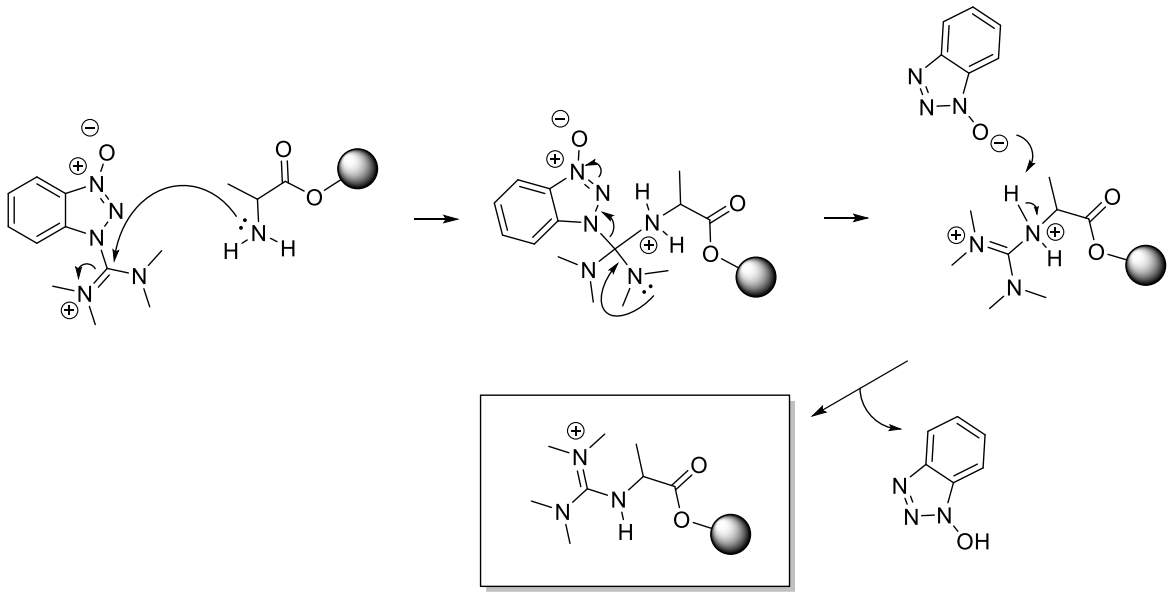
The term uronium salt is misleading. This refers to the structure proposed when the reagent was first introduced (**33**, scheme 3.3).[152] However, more recent evidence suggests it is more commonly found as the guanidinium isomer (**34**, scheme 3.3).[153] The likely mechanism of amide bond formation using HBTU is depicted in scheme 3.3. A tertiary amine base, such as DIPEA is used to deprotonate the carboxylic acid. The resulting carboxylate anion attacks the electrophilic sp^2 hybridised carbon of the guanidinium structure. Subsequent collapse of the tetrahedral intermediate, results in the formation of the isouronium cation, and the oxybenzotriazole anion. These two species then react to form oxybenzotriazole ester, with the release of tetramethylurea. Attack of this activated ester by the amine gives the desired amide bond after deprotonation.



Scheme 3.3 Mechanism of amide bond formation using HBTU. Also shown are the different isomers of HBTU **33** and **34**.

Uronium salts are used in an equimolar quantity to the carboxylic acid, and are typically premixed before addition to the amine. This is to reduce the occurrence of competing guanylation of the amine, which can lead to capping of the resin (scheme 3.4).[145] When using HBTU, it is also common to also use HOBt as an additive. This increases the

efficiency of the oxybenzotriazole ester formation. When coupling amino acids, this is also said to act to supress potential racemization. As the isouronium ester, *N*-protected amino acids can be particularly prone to base-induced racemization through either enolization or oxazolone formation (Scheme 3.5).[154] This is due to the increased acidity of the alpha proton.



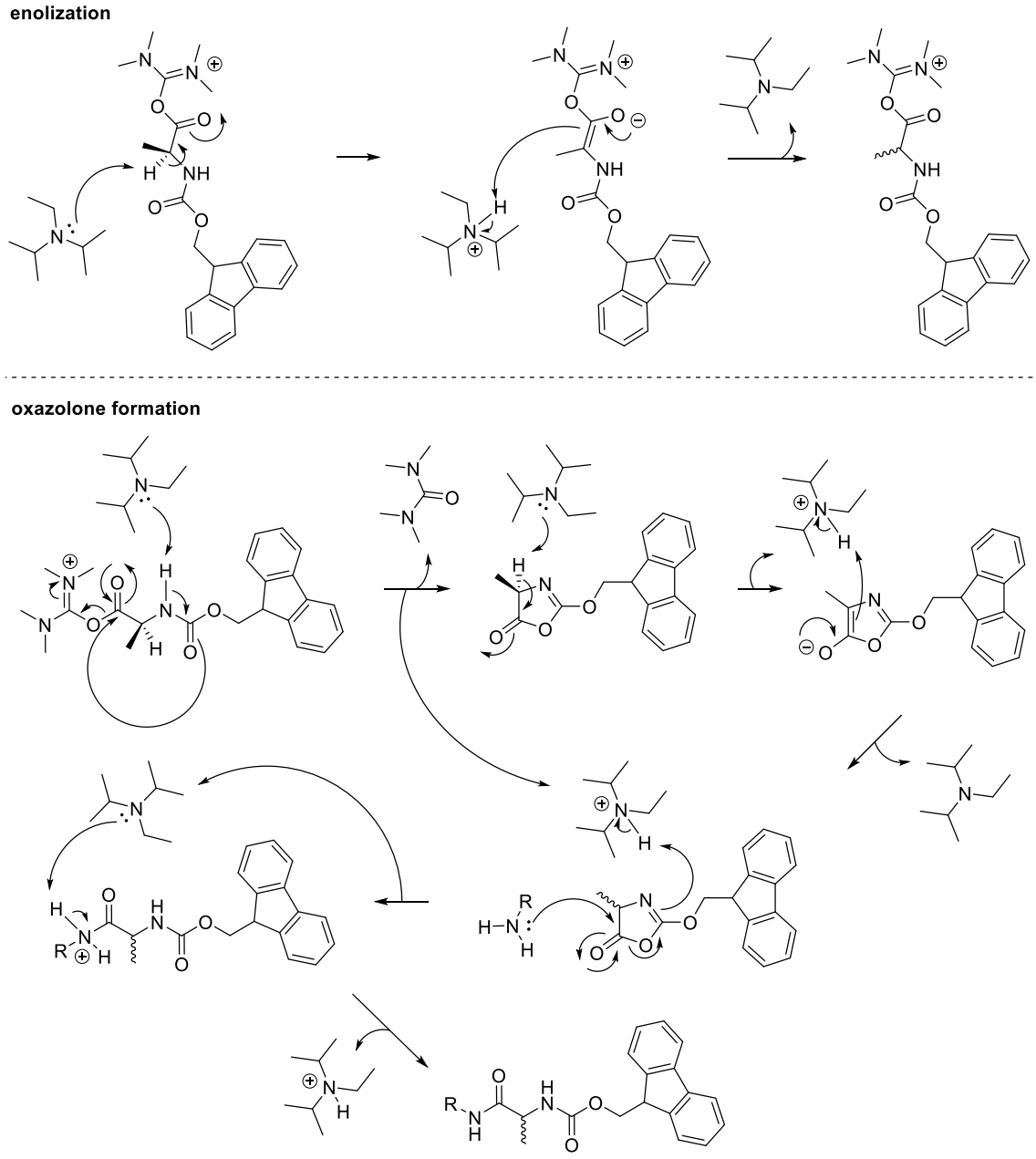
Scheme 3.4 Potential capping of resin by guanylation.

For coupling of **11**, this is not a specific concern, but there is still the potential to benefit from faster formation of the oxybenzotriazole ester with the use of HOBt as an additive.

Solid phase couplings are usually performed with a large access of the active ester, in order to drive reactions to completion. Near quantitative conversions are important to prevent the accumulation of resin bound deletion products. However, given the precious nature of **11**, it was decided that couplings should be first attempted with the smallest excess possible.

Therefore, 1.1 equiv. of **11** was treated with an equimolar quantity of HBTU, and a twofold excess of both HOBt and DIPEA, in DMF. After 30 seconds the mixture was added to the resin. The resin was shaken for 2 hours. At this point a sample of the resin was taken, and subjected to Kaiser testing.

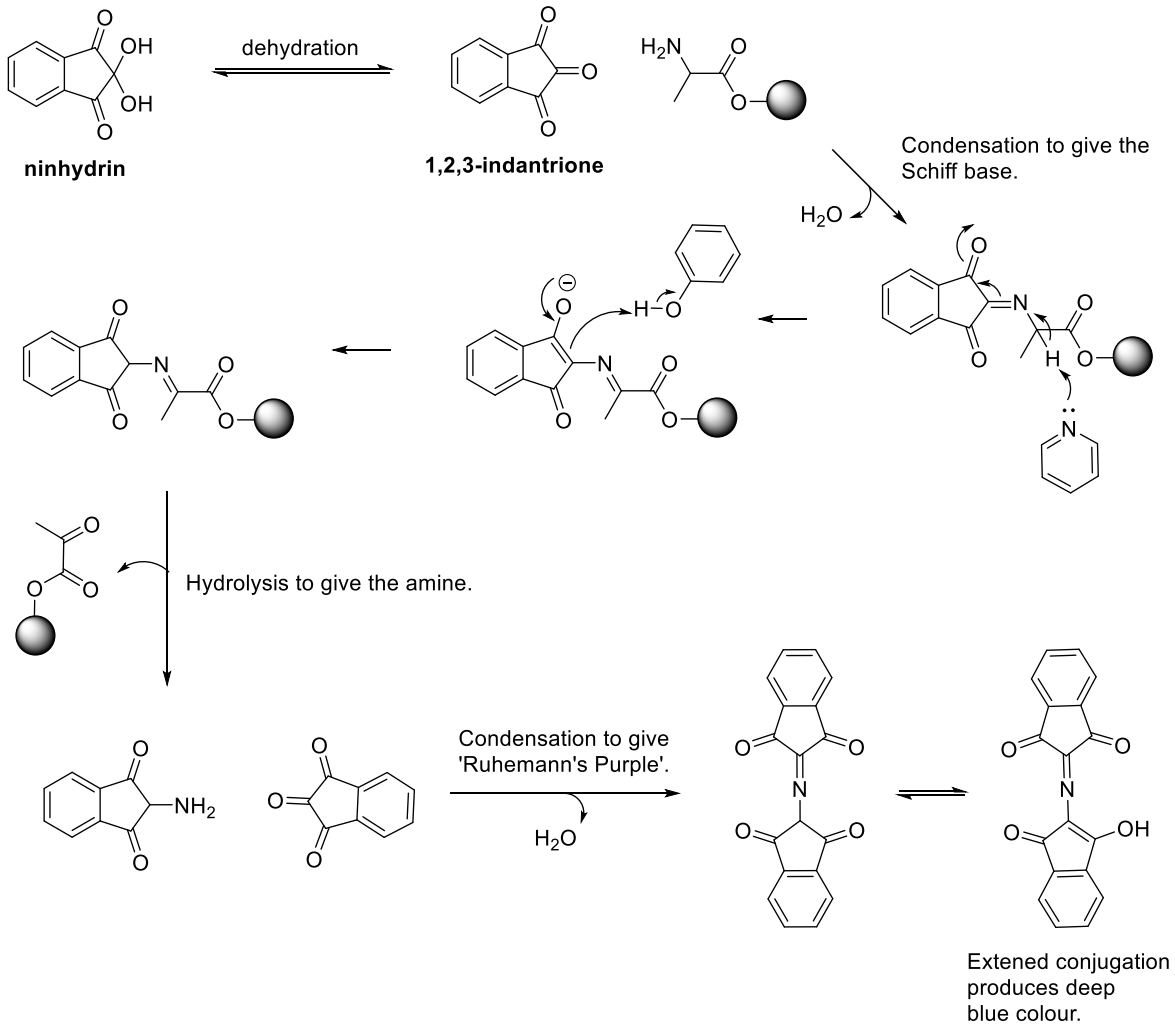
The Kaiser test is a qualitative test used to detect the presence of primary amines.[155] A small sample of the washed resin is treated with a drop of three solutions. These are, 5% w/v ninhydrin in ethanol, 80 % w/v phenol in ethanol, and a 2:98 mixture of aqueous 0.001 M KCN and neat pyridine. The sample is heated to 120 ^{o}C for 5 mins. Formation of a deep blue colour indicates the presence of primary amines.



Scheme 3.5 Potential base catalysed racemization, of the isouronium ester form of Fmoc-protected amino acids, by enolization or oxazolone formation during couplings.

The Kaiser test is based on the reaction between ninhydrin and primary amines, which results in the formation of the chromophoric compound 'Ruhemann's Purple'. The mechanism of this reaction has been the subject of some controversy. However it is now

generally accepted that the most likely mechanism (scheme 3.6),[156] involves the dehydration of ninhydrin to give 1,2,3-indantrione. This condenses with the amine to give the Schiff base. Decomposition gives an intermediary amine derivative of ninhydrin, which condenses with another molecule of 1,2,3-indantrione to give ‘Ruhemann's Purple’.



Scheme 3.6 Mechanism of the Kaiser test.

Unexpectedly, the Kaiser test performed after 2 hours, was negative, suggesting complete coupling of **11**, despite the modest excess of the reagent used. This was a surprise, as it was anticipated that 1.1 equiv. of **11**, would not be sufficient to affect a quantitative coupling. The plan in the event of a positive Kaiser test would have been to perform additional 2 hour couplings with 0.5 equiv. of **11** until a negative Kaiser test was observed. This was deemed to be a good reagent conserving strategy for driving couplings to completion. Fortunately this did not appear to be necessary.

Subsequently the indoline nitrogen was Fmoc-deprotected with piperidine as before, and coupled to a final alanine residue. This was achieved using Fmoc-Ala-OH (5 *equiv.*), HBTU (5 equiv.), and DIPEA (10 equiv.), in DMF. As before the reagents were premixed for 30 secs prior to addition to the resin. Since a large excess of amino acid was used a more standard 45 min reaction time was employed. Kasier testing was not employed here, as the formation of 'Ruhemann's Purple', does not occur between the reaction of ninhydrin and secondary amines. Other resin tests exist which are capable of detecting free secondary amines. These are typically used to monitor couplings of proline. A good example is the chloranil test.[157] However this was not employed here due to the lack of reagent availability. Therefore, it was decided to assume that the single coupling with a large excess of amino acid was likely sufficient.

After removal of the *N*-terminal Fmoc group of the alanine residue, the resin was prepared for cleavage by extensive washing with DMF, followed by DCM, and drying under a stream of nitrogen. Cleavage was affected by treatment of the resin with a solution of 95 % TFA, 2.5 % TIPS, and 2.5 % water. After shaking for 2 hours, the cleavage mixture was filtered and concentrated by rotary evaporation, followed by precipitation with the addition of cold Et_2O.

The crude product was analysed by reverse phase analytical HPLC. Initial results were disappointing. Although an obvious product dominated the HPLC trace, with a strong peak at 9.0 min, several significant side products were also observed (figure 3.1a). It seemed unlikely that these could represent deletion products. Firstly there were more side products than could be predicted to have resulted from incomplete couplings. Furthermore, Kaiser testing had suggested that the first coupling was quantitative. Incomplete Fmoc-deprotection also has the potential to result in deletion products, however considering the reliability of this step, it too seemed unlikely to explain the observed impurities.

HPLC analysis was monitored by UV absorbance at the 254 nm, and 214 nm wavelengths. The side products showed greater intensity at 254 nm, and this strongly suggested that they contained aromatic character derived from the duocarmycin residue. It was possible that the large excess of activated alanine was able to react, not only with the indoline nitrogen, but also lead to acylation of nucleophilic sites on the indole scaffold of the duocarmycin residue. However, this seemed unlikely with activation of the alanine by HBTU. Even if an acid chloride of alanine had been used, such a reaction would be likely to require Friedel-Craft like conditions.[158] Furthermore, the indole side chain of tryptophan is not reported to be susceptible to acylation during peptide couplings.

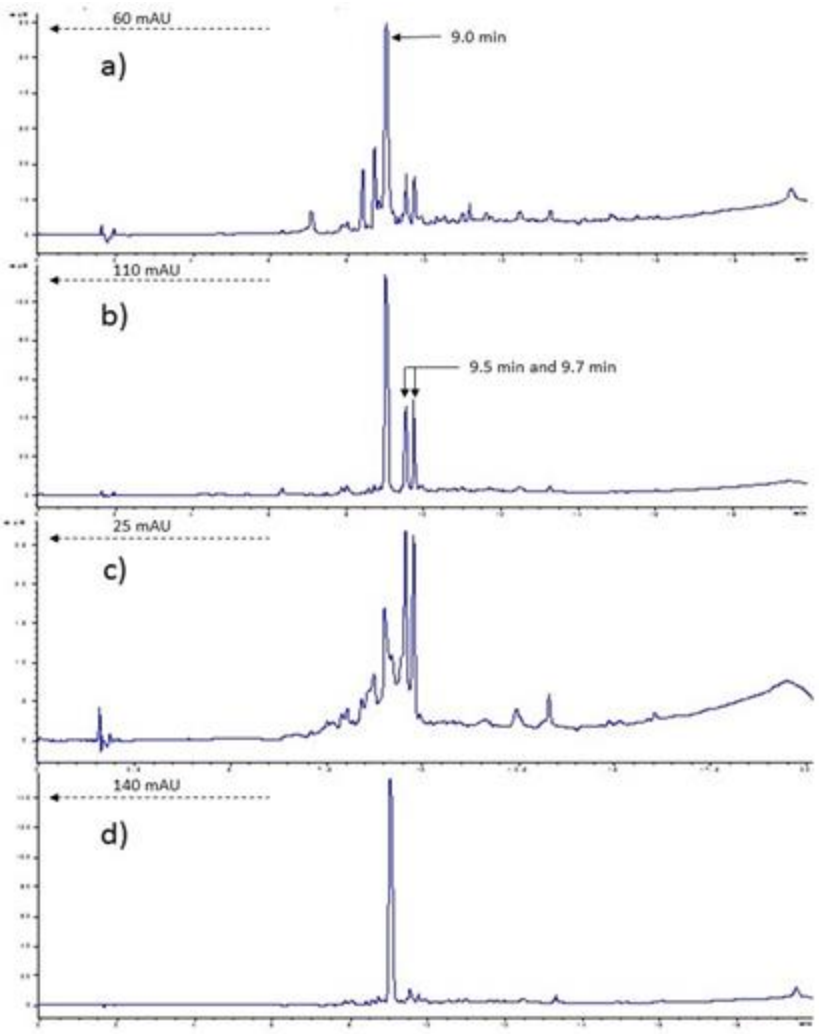


Figure 3.1 HPLC analysis of crude **32** after cleavage under varying conditions. a) 95 % TFA, 2.5 % TIPS, 2.5 % H_2O. b) 50 % TFA, 50 % DCM. c) 95 % TFA, 5 % DCM. d) 47.5 % TFA, 47.5 % DCM, 2.5 % TIPS, 2.5 % H_2O. 10 mg of dried resin was cleaved under either conditions a, b ,c, or d, with 5 mL of the respective cleavage cocktail for 2 hours. The cleavage mixture was filtered and evaporated to dryness. The crude was dissolved in 1 mL of MeOH and analysed by HPLC at 254 nm. Agilent Eclipse XDB-C18 column, 4.8 x 150 mm, 5 µm. Solvent A: [Water and 0.05 % TFA], Solvent B: [ACN and 0.05 % TFA]. Gradient: 0% [B] to 95 % [B], from 0 min to 15 mins, 95 % [B] to 0 % [B] from 15 to 20 mins. Monitored UV 254 nm. Flow rate 1 mL/min. Column temperature 40 ^{o}C.

As this synthesis had been conducted with racemic **11**, it was briefly considered that perhaps the impurities were the result of separation of diastereomeric deletion products. The pattern of impurities certainly could be described as two pairs either side of the main product. However, the main product would also be a mixture of diastereoismers, and no separation of the peak at 9 min had been observed.

These explanations seemed unlikely, and therefore attention turned to the cleavage conditions. The cleavage cocktail contained 95 % neat TFA, and it was possible that the duocarmycin residue was not stable under these conditions. Therefore, small sample of

11 was treated with neat TFA. Monitoring by TLC showed decomposition of this reagent. This suggested, that the unexpected impurities were the result of side reactions during cleavage. As a result, the synthesis of **32** was repeated as before. On completion, the dried resin was portioned into 10 mg samples, and exposed to varying cleavage conditions. The crude products were again analysed by reverse phase analytical HPLC (figure 3.1).

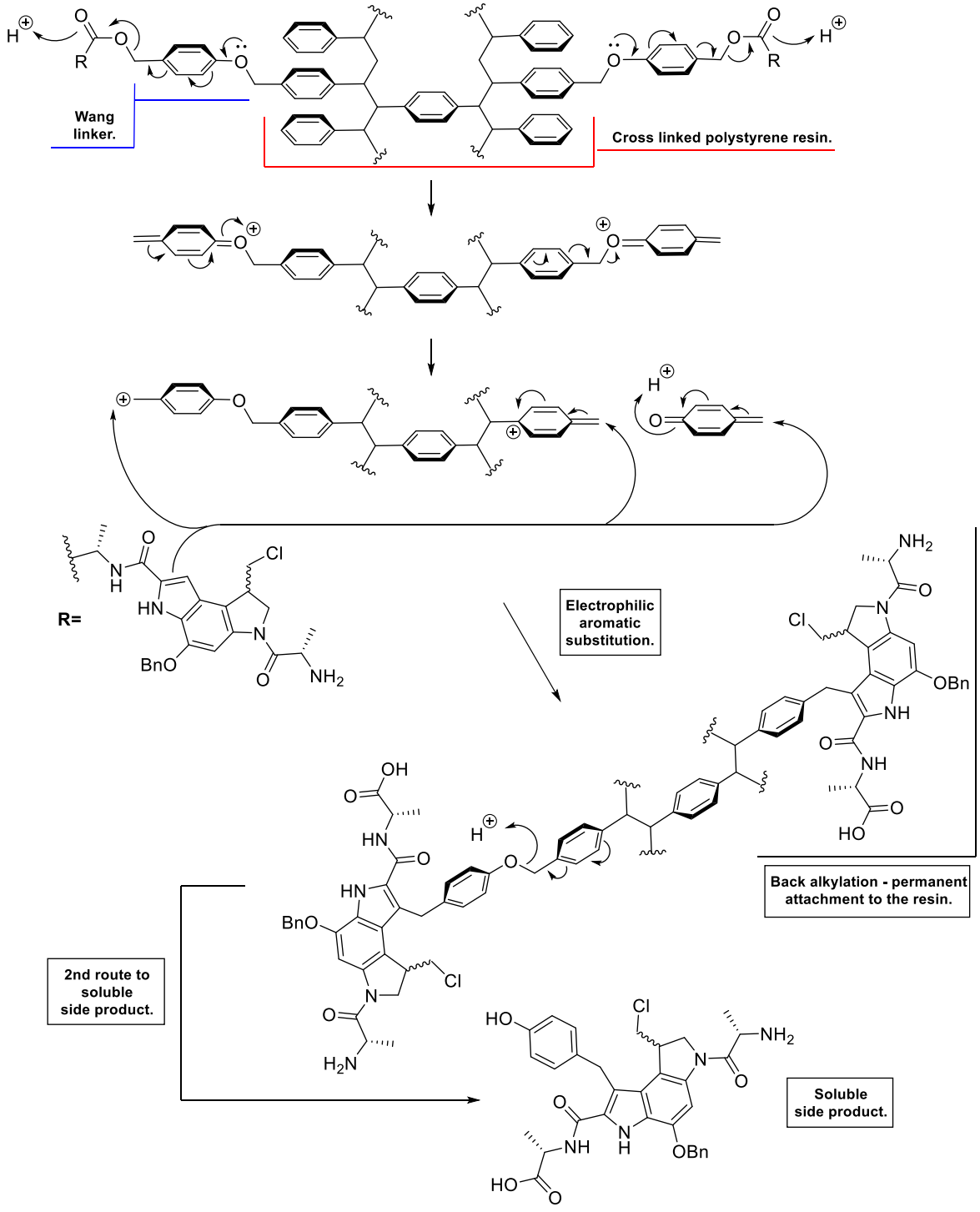
Reducing the TFA concentration to 50 % by dilution with DCM, and the omission of the scavengers (TIPS and H_2O), saw a dramatic improvement (figure 3.1b). The side products that eluted before the main product at 9 min, which had been the most intense impurities under the original conditions, were no longer present. However the two impurities at 9.5 and 9.7 min appeared to have grown in intensity when compared to the main product at 9 min. It seemed that the inclusion of scavengers protected against the formation of these two impurities; but the question arose, was it the reduction in TFA concentration, or the omission of scavengers which had protected against formation of the other side products? It was at least conceivable, that the combination of TFA as proton donor, and the trialkylsilane, as potential hydride donor, could lead to reduction of the indole scaffold to an indoline.[159] Indeed, the use of TES as a scavenger in the cleavage of tryptophan containing peptides is known to cause reduction of the indole side chain.[160] However, TIPS is cited as a good alternative when this problem is encountered, as it is significantly less prone to producing this side reaction.

In order to ascertain whether reducing the TFA concentration had derived a benefit. The next sample of resin was subjected to cleavage with 95 % TFA and 5 % DCM (figure 3.1c). Here, it can clearly been seen that the high concentration of TFA has had a detrimental effect. The main product peak previously observed at 9 mins, is no longer the majority product, in fact it is arguably no longer observed at all. Furthermore, the two impurities at 9.5 and 9.7 min now dominate the HPLC trace.

The logical extension of these findings was to combine a reduced concentration of TFA with the inclusion of the scavengers. Thus, the next sample of resin was cleaved using 47.5 % TFA, 47.5 % DCM, 2.5 % TIPS, and 2.5 % H_2O. This combination appeared to produce a synergistic benefit (figure 3.1d). Under these conditions the main product at 9 min, was the only discernible peak. This was confirmed as the desired product by accurate mass analysis.

Scavengers, are included in cleavage cocktails to protect potentially nucleophilic residues from alkylation by cationic species derived from cleavage. In the case of peptides containing trifunctional amino acids, this can include the *tert*-butyl cations derived from

simultaneous side chain deprotection. This is of course not applicable to the alanine residues present in **32**. However, another source of carbocations is the linker itself.



Scheme 3.7 Acidic cleavage of **32** from Wang resin. Production of *p*-quinone methide, and resin bound cations. Potential alkylation of indole scaffold leading to soluble and permanently resin bound impurities.

Scheme 3.7 depicts the acidic cleavage of **32** from Wang resin. The Wang linker is attached to the polystyrene resins *via* a benzyl ether, and is linked to the product by a benzyl ester. As discussed previously, it is the electron-donating effect of the *para*-benzyl ether, which increases the acid sensitively of the benzyl ester, and allows for complete cleavage by treatment with TFA. However, an undesired consequence is the potential for subsequent cleavage of the linker remnant itself. This liberates the highly reactive p-quinone methide species, which is then available to alkylate sensitive residues;[161] in this case, the indole scaffold of the duocarmycin structure, by electrophilic aromatic substitution. This would result in soluble side products, and is likely responsible for the impurities seen at 9.5 and 9.7 min, which are reduced by the addition of scavengers.

It is likely that the side product at 9.5 min represents alkylation at the most nucleophilic 3 position of the indole, depicted in scheme 3.7, and the side product at 9.7 mins, represents the additional alkylation of this product at the 6 position (standard indole numbering) or at the indole nitrogen itself. Attack of the aromatic ring of the benzyl ether protecting the phenol group seems less likely. A second route to this impurity can also be envisioned, when *p*-quinone methide in not released. Here, the resin bound carbocation generated reacts first with the indole scaffold, and this product is then cleaved from the resin at the benzyl ether.

It can be seen in figure 3.1, that addition of TIPS and H_2O protects against the formation of these side products, as the intensity of their peaks grows with the omission of these scavengers (figure 3.1a vs figure 3.1b, and figure 3.1c). Here H_2O acts as a completing nucleophile, and TIPS as a hydride donor. The synergy in protection observed between reduction in TFA concentration and the addition of scavengers (figure 3.1a vs figure 3.1d), most likely results from the reduced liberation of *p*-quinone methide. Thus, the smaller quantity of this soluble reactive species, is more effectively quenched by the concentration of scavengers. The resin bound-carbocation resulting from cleavage of the benzyl ester still exists in the same quantity. However, this species can also be quenched by the polystyrene resin itself, leading to increased crosslinking of the resin. This effect has been used to explain the reduced swelling capacity of Wang resins post TFA cleavage.[161]

As can be seen in scheme 3.7, liberation of *p*-quinone methide also produces a second carbocation on the polystyrene resin itself. This too can be envisaged to react with the sensitive indole scaffold. The consequence here would be a so called ‘back alkylation’, which permanently sequesters the product to the resin, reducing yields. Closer inspection of figure 3.1 provides some tentative evidence for this effect. The test cleavages were identical except for the composition of the cleavage cocktail, and were analysed by HPLC

after dissolution of the crude in the same volume of MeOH. Thus the intensity of the HPLC peaks gives a qualitative indication of recovery. Under the most destructive cleavage conditions of no scavengers and high TFA concentration, where back alkylation would be greatest (figure 3.1c), the highest intensity peaks are those of the soluble alkylated side products, at around 25 mAU. At reduced TFA concentration and with the presence of scavengers (figure 3.1d) the only discernible peak is the desired product with much greater intensity of around 140 mAU. This suggests a reduction in recovery of the product, which is consistent with loss due to back alkylation. Furthermore, comparison of figure 3.1a, and figure 3.1c, demonstrates that at high TFA concentration but with the presences of scavengers, which would protect against back alkylation, the desired product still dominates with an intensity of around 60 mAU.

Further support for these suspected side reactions during cleavage, can be found in the literature, concerning the cleavage of tryptophan containing peptides from Wang resin. Side protects resulting from the alkylation of the indole side chain of tryptophan by liberated p-quinone methide, have been isolated and characterised.[162, 163] Furthermore, Wang resin is reported to produce low yields of tryptophan containing peptides, even after soluble side products have been taken into account, and this has been attributed to losses incurred by back alkylation to the resin.[162, 164] Considering that it is the more nucleophilic 3- position of the pyrrole ring of the duocarmycin indole which is unsubstituted, compared to the 2-position in the tryptophan indole, syntheses incorporating **11**, may be particularly susceptible to such side reactions. Furthermore, the 6-position of the duocarmycin indole is further activated over the same position in the tryptophan indole, by the *ortho*-benzyl ether. Although this position is sterically shielded by the benzyl group.

It is reported that for tryptophan containing peptides such side reactions can be reduced by the use of Boc side chain protection.[165, 166] This reduces the nucleophilicity of the indole. It is suggested that removal of this group during cleavage by TFA is retarded at the carbamic acid intermediate,[166, 167] which is lost during work up, offering protection throughout cleavage. It may be that **11** would be an improved solid phase building block, if it retained Boc-protection of the indole nitrogen.

It is also clear from the comparison of figure 3.1a, b, and c, reducing the TFA concentration alone is also beneficial. Considering the observed instability of **11** itself to neat TFA, this likely represents the decomposition of the duocarmycin residue at high concentrations of TFA. It is not obvious what causes decomposition and it is likely to occur *via* multiple routes. It is possible for example that some benzyl ether protection is lost. The occurrence of 3-trifluoroacetylindoles has also been reported by the reaction of TFA with

indole-2-carboxylic acids, although this was under refluxing conditions.[168] Furthermore the authors state that acylation occurs after decarboxylation, and this should not be possible once **11** is incorporated into a polyamide structure. However, they do report that the reaction produces many other by-products which were not isolated.

This first set of experiments, had clearly demonstrated that **11**, could be incorporated into polyamide structures using Fmoc solid phase synthesis. Pleasingly, large excesses of **11** did not appear to be necessary to affect quantitative couplings to resin bound amino acids. It is clear however that cleavage is not a trivial step, and care must be taken to avoid degradation of the product. Reducing TFA concentration and the use of appropriate scavengers effectively suppresses these side reactions, at least for short sequences which do not contain side chain protected trifunctional amino acids.

3.4 The first attempt at a small library of amino acid-duocarmycin conjugates.

It was decided to attempt to use the solid phase synthesis and optimised cleavage conditions employed in the first experiments, to access a small library of amino acid-duocarmycin conjugates, and assess the effects of different amino acids on the antiproliferative activity of the resulting analogues. This work predated the scale up synthesis, and thus was conducted using racemic **11** from the pilot synthesis.

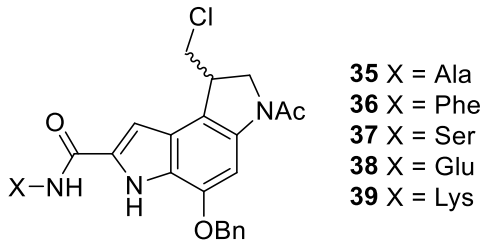
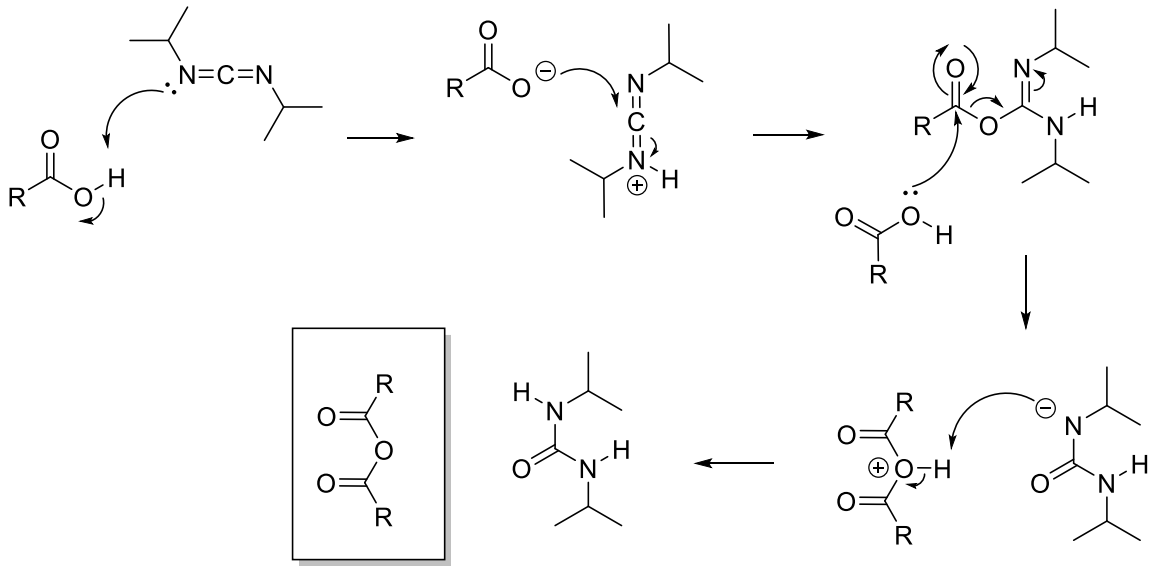


Figure 3.2 Targets of the first attempt at a library of amino acid-duocarmycin conjugates.

For this series of analogues the decision was made to acetylate the indoline nitrogen as opposed to coupling a final amino acid residue. This would allow comparison of biological activity with previously described, truncated duocarmycin analogues.[169] Alanine, phenyl alanine, serine, glutamic acid, and lysine were chosen as the *C*-terminal amino acids (figure 3.2). These were chosen to represent non-polar, polar, and charged side chains. It was considered particularly interesting to see if the positively charged side chain of lysine, could both increase water solubility, and benefit minor grove binding, by association with the polyanionic backbone of DNA.

The lab already had a large supply of unsubstituted Wang resin, and bulk quantities of the required Fmoc-protected amino acids. Therefore, for economic reasons, it was decided not to purchase preloaded resins, and instead load the first amino acid residues manually. Loading of the first amino acid residue, requires esterification to the hydroxyl group of the Wang linker. This was achieved using the symmetrical anhydride method.[145] For the preparation of each resin, the Fmoc-protected amino acid was converted to its anhydride by treatment with DIC in anhydrous DCM. Typically a few drops of anhydrous DMF were also required to aid dissolution of the amino acid. The mixture was stirred under N_2 for 1 hour at 0 ^{o}C. During this time the Wang resin was swelled with anhydrous DCM for 30 min, followed by anhydrous DMF for a further 30 min. The anhydride was concentrated by rotary evaporation, and the residue dissolved in anhydrous DMF, before being added to the swelled resin. Addition of 10 mol % DMAP, was also added to the swelled resin to promote esterification. The resin was shaken under N_2 for 1 hour. At this point the resin was washed thoroughly with anhydrous DMF, before being treated with a large excess of acetic anhydride, again in anhydrous DMF under N_2 for 1 hour. The acetic anhydride treatment was use to acetylate unreacted hydroxyl groups. This was necessary because the substitution level of the resin with respect to the Wang linker was higher than desired. Therefore, sub-stoichiometric quantities of the amino acid anhydrides were used. The lower substitution allowed larger quantities of resin to be used which were more convenient to weigh. Following this treatment the resins were washed thoroughly with DMF followed by DCM. The washed resins were then dried under a steam of N_2, followed by desiccation under vacuum overnight.



Scheme 3.8 Mechanism of anhydride formation using DIC.

Scheme 3.8 depicts the mechanism of anhydride formation. In brief, DIC promotes the condensation of two molecules of the amino acid. First the carboxyl group of one amino acid reacts with DIC to form the O-acylisourea intermediate. This is attacked by the carboxylic acid group of a second amino acid to liberate the symmetrical anhydride. The reaction of the symmetrical anhydride with the hydroxyl group of the Wang linker is analogous to the formation of the *t*-butyl carbamates, described in section 2.2.5, with DMAP acting as an acyl transfer agent.

Before the loaded resins could be used, an accurate estimation of the resin loading with respect to the amino acids was required. This was especially important due to the small excesses of **11** being used for the following couplings. Fortunately, this can be determined by monitoring of the Fmoc-deprotection reaction of a sample of the resin. The relationship between UV absorption and concentration of the fulvene-piperidine adduct is well-characterised, and thus can be used to calculate the resin loading. [145]

For each resin the following procedure was followed. Two accurately weighed 5 mg samples of the thoroughly dried resin were treated with 3 mL of 20 % piperidine in DMF. The samples were allowed to react for 3 hours with occasional agitation. The long reaction time was to ensure complete deprotection, as the dried resin was not swelled prior to addition of the piperidine solution. UV absorbance at 290 nm was then measured for each solution against a blank of 20 % piperidine in DMF, in 1 cm quartz cuvettes. It is important that the UV absorbance is within the linear range, and typically this required a series of 2 fold serial dilutions, until an absorbance reading below 1 was observed. Resin loading was then calculated from the mean absorbance using the following equation,[145] which is based on the Beer-Lambert law ($A = \varepsilon cl$), where A is absorbance at 290 nm, ε is the molar extinction coefficient (M^{-1} cm^{-1}), c is the concentration (M), and l is the path length (cm). The 1.75 constant is derived from the extinction coefficient of the fulvene-piperidine adduct (5253 M^{-1} cm^{-1}) and a reaction volume of 3 mL, with a path length of 1 cm. The dilution factor corrects for the dilutions needed to give an absorbance reading below 1 when using 5 mg samples.

$$Loading\ (mmol\ per\ g) = \frac{[Absorbance][dilution\ factor]}{[mg\ of\ sample]\ 1.75}$$

To validate the method, it was first used to estimate the resin loading of the commercially available preloaded resin used in the first solid phase experiment. The results match the manufacturer's stated value.

The loaded resins were then used to conduct the solid phase synthesis of the small library on a 0.038 mmol scale. Couplings of **11** were conducted as before, and after Fmoc-removal, the indoline nitrogen was acetylated by treatment with acetyl chloride and DIPEA in DMF. The analogues were cleaved using the optimised conditions described in section 3.2, and concentrated to dryness. HPLC analysis of the crude products confirmed the successful cleavage, with all exhibiting just one major product peak. Figure 3.3 shows an example HPLC trace of crude lysine analogue **39** after cleavage.

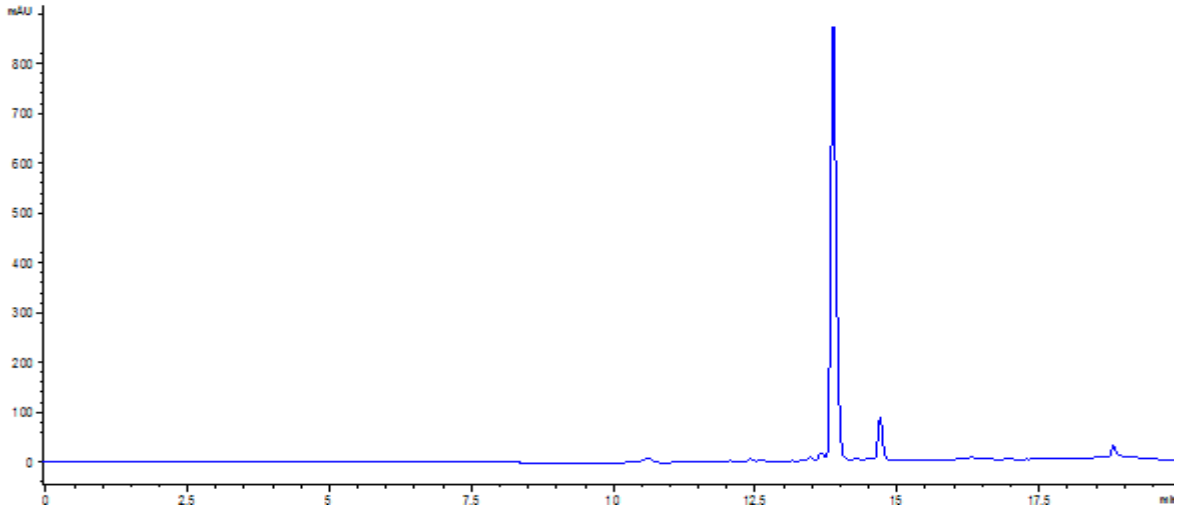
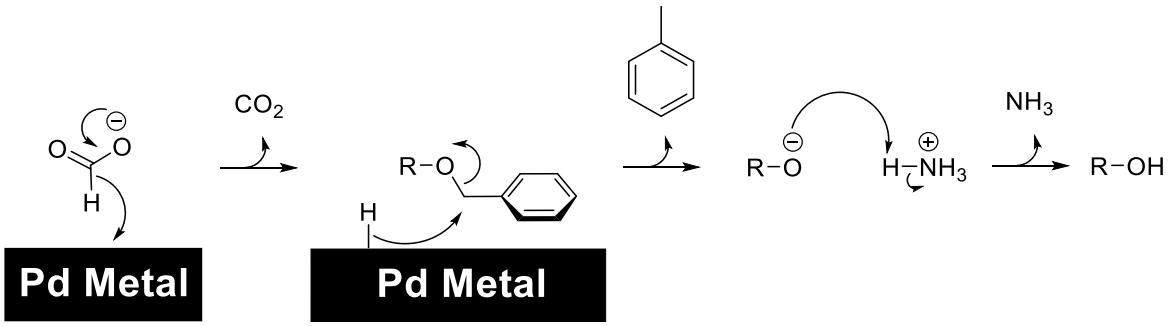


Figure 3.3 Crude HPLC trace of **39** after cleavage. HPLC conditions: Agilent Eclipse XDB-C18 column, 4.8 x 150 mm, 5 µm. Solvent A: [Water and 0.05 % TFA], Solvent B: [MeOH and 0.05 % TFA]. Gradient: 0% [B] to 95 % [B], from 0 min to 15 mins, 95 % [B] to 0 % [B] from 15 to 20 mins. Monitored UV 254 nm. Flow rate 1 mL/min. Column temperature 40 ^{o}C.

Mass spec analysis confirmed the desired products. This showed that the reduced TFA concentration of the optimised cleavage conditions was still sufficiently acidic to affect removal of *tert*-butyl side chain protection of the serine, lysine, and glutamic acid analogues. It also confirmed that the *tert*-butyl carbocations released during cleavage of these analogues were not alkylating the duocarmycin scaffolds to any significant extent. This suggests that additional scavengers are not required for coupling of *t*-butyl protected trifunctional amino acids, at least not for analogues containing single amino acid residues. The crude products were subjected directly to benzyl-deprotection to provide the active seco forms of the duocarmycin analogues for biological evaluation.

This was achieved *via* catalytic transfer hydrogenation, using Pd/C and aqueous ammonium formate,[170] in either MeOH, or a mixture of MeOH and THF depending on the solubility of the analogue. This procedure is typical for the benzyl deprotection of duocarmycin analogues, and has been shown to allow the selective removal of the benzyl ether, without completing dehalogenation of the substituted indoline ring, or indole reduction.[96]

The reaction takes place on the surface of the heterogeneous Pd catalyst using ammonium formate as a hydrogen donor,[171] and most likely proceeds *via* a mechanism consistent with that depicted in scheme 3.9. Here, the formate ion first dissociates from NH_4 and is adsorbed on to the Pd surface. This results in its decomposition to CO_2 and H-Pd. When the substrate is also adsorbed on to the Pd surface, the hydride is transferred to the benzyl ether, producing toluene, and the phenoxide, which in turn can be protonated by the ammonium, H_2O, or MeOH. In the absence of the substrate, the hydride can react with the ammonium, H_2O, or MeOH to release hydrogen gas.



Scheme 3.9 Likely mechanism of benzyl ether cleavage, *via* heterogeneous Pd catalysed transfer hydrogenation.

The Pd/C was easily removed by filtration through a short column of diatomaceous earth. Analytical HPLC analysis revealed good conversion to one major product for all analogues. Figure 3.4 shows an example HPLC trace of crude Lysine analogue **39** after benzyl-deprotection.

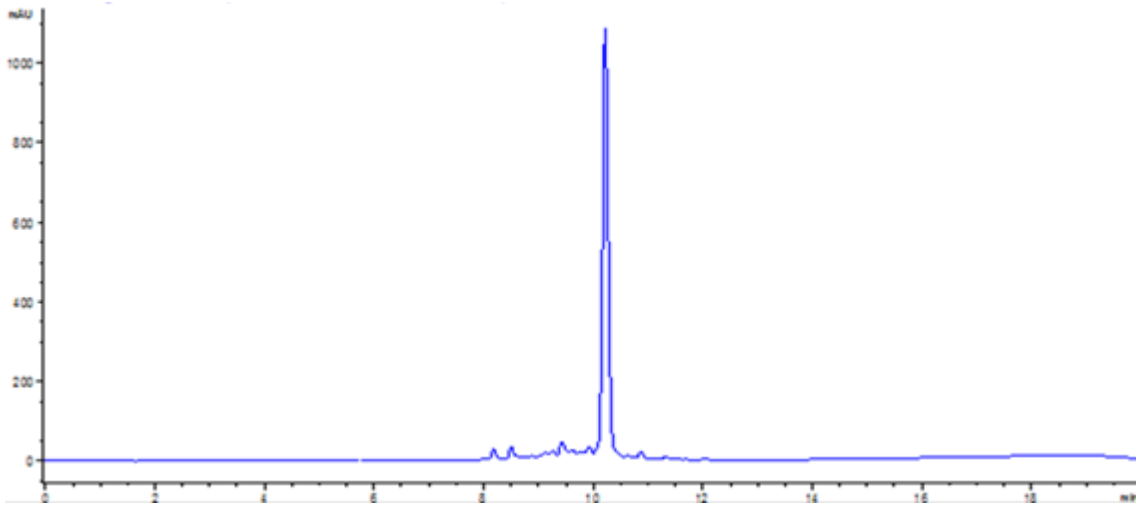


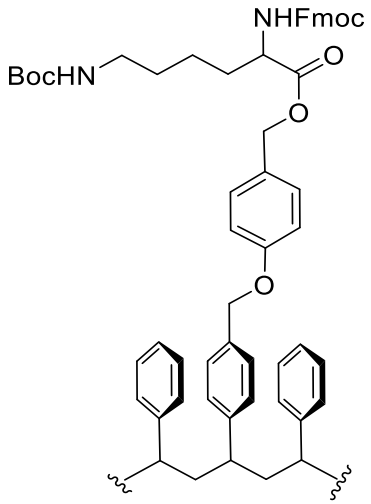
Figure 3.4 Crude HPLC trace of **39** after benzyl-deprotection. HPLC conditions: Agilent Eclipse XDB-C18 column, 4.8 x 150 mm, 5 µm. Solvent A: [Water and 0.05 % TFA], Solvent B: [MeOH and 0.05 % TFA]. Gradient: 0% [B] to 95 % [B], from 0 min to 15 mins, 95 % [B] to 0 % [B] from 15 to 20 mins. Monitored UV 254 nm. Flow rate 1 mL/min. Column temperature 40 ^{o}C.

Mass spectrometry confirmed these were the desired phenols. However, the crude products were clearly over 100 % yield in mass, and had a glossy appearance. This seemed most likely the result of excess ammonium formate contamination. As a result some further purification would be required. The analogues, particularly those with polar side chains, did not run well on TLC plates, and required mobile phases with large concentrations of MeOH to move the products from the baseline. Therefore purification was attempted using reverse phase preparative HPLC. However only trace quantities of the products were recovered. It seemed likely that the particularly poor recoveries were the result of un-optimised preparative HPLC conditions. However, intrinsically low yields could not be ruled out. It was questioned whether, despite the optimised cleavage conditions, the use of Wang resin could still be having a determinately effect on yield *via* the suspected back alkylation. Therefore it was decided that before the library would be resynthesised, a brief exploration of different resins should be conducted.

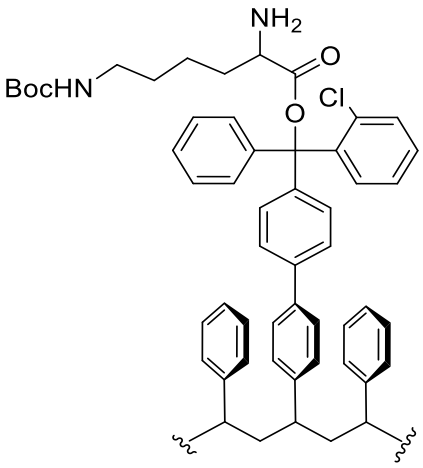
3.5 Resin screening.

Although it is true that all linkers cleaved by acidic conditions will result in a resin bound cation, the propensity for this species to participate in the back alkylation of sensitive nucleophilic residues is not necessarily equal. The Wang linker is particularly prone to this side reaction, most likely due to the exposed nature of the cation. An alternative, often recommended for the synthesis of peptides containing tryptophan, is the 2-chlorotrityl linker.[172, 173] The reduced occurrence of back alkylation to the indole side chain of tryptophan, has been attributed to a steric shielding effect provided by the triphenylmethyl structure that encapsulates the cleavage site.[145] This shielding effect has also been shown to effectively supress diketopiperazine formation, another common side reaction which can lead to near quantitative losses in the synthesis of sensitive peptide sequences.[174]

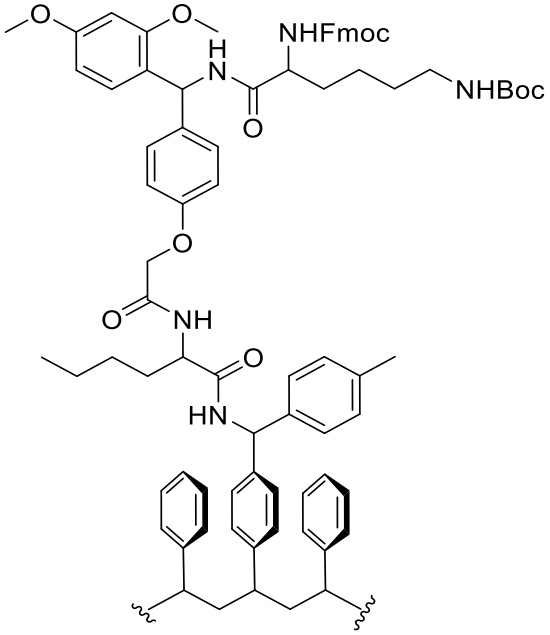
It was therefore decided to resynthesize the lysine analogue **39**, on a number of different resins, and compare the recovery. In order to reduce the influence of handling errors, this would be achieved by conducting the synthesis for each resin, on the same scale, using identical reaction conditions, and analysing the crudes by analytical reverse phase HPLC, after cleavage. Although there would be no internal standard, and therefore estimation of yield would not be possible, by preparing the HPLC samples in an identical manner, the area of the product peak could serve as a qualitative comparison of recovery between each resin.



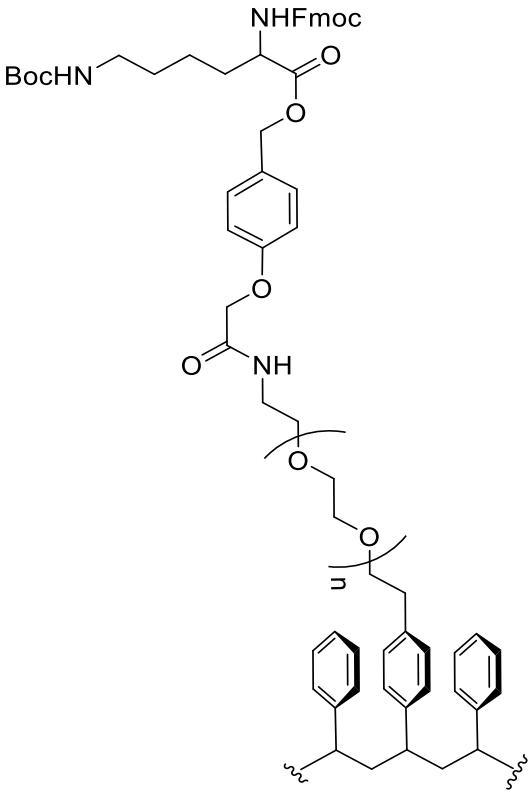
Wang resin



2-Chlorotrityl resin



Rink amide MHBA resin



NovaSyn TGA resin

Figure 3.5 Structures of the different lysine substituted resins.

In addition to the original Wang resin, and the suspected superior candidate 2-chlorotrityl resin, the screen would also include Rink amide MBHA resin,[175] and NovaSyn® TGA resin.[176] The structures of the lysine substituted resins are shown in figure 3.5.

NovaSyn® TGA resin also contains the Wang linker, however instead of being attached directly to the polystyrene polymer by a benzyl ether, it is spaced by a polyethylene glycol chain. Such polyethylene glycol functionalised resins, are said to provide a more favourable matrix environment, for solid phase synthesis.[177] The polyethylene spacers help to improve the solvation of the growing peptide chains, and protect against aggregation which can reduce the coupling efficiencies of susceptible sequences. Association of the hydrophobic duocarmycin structure with polystyrene resins might be a foreseeable issue when considering the synthesis of more complex analogues in the future. It was also speculated that the improved mass transfer kinetics of the matrix environment of polyethylene glycol functionalised resins, may facilitate the escape of products from the matrix before they could react with the cleavage cation.

Rink amide MHBA resin, contains the Rink amide linker, attached *via* a norleucine spacer to a polystyrene based 4-methylbenzhydrylamine resin. This spacer protects against alternative detachment of the Rink amide linker at the ether during cleavage. This can occur when the Rink amide linker is attached directly to polystyrene, in an analogous fashion to release of p-quinone methide observed with Wang resin. Unlike the other linkers, cleavage provides the product as a terminal amide, as opposed to a carboxylic acid. This was deemed to have potential useful applications in the future. Furthermore, the two phenyl groups surrounding the cleavage site may provide some protection against back alkylation. In fact, it could be said to offer a level of steric protection which is intermediary between the well shielded cleavage site of the 2-chlorotrityl linker, and the very exposed cleavage site of the Wang linker.

For the screening, the synthesis was conducted on a 0.038 mmol scale, using the reaction conditions already described in section 3.3. The unnatural enantiomer of **11** was used as this was deemed less valuable. After cleavage, the crudes were evaporated to dryness and dissolved in 1 mL of MeOH. 100 µL of the resulting solutions were diluted with 900 µL of MeOH and analysed by reverse phase HPLC. A comparison of product peak area for each resin is shown as a bar chart in figure 3.6.

It can be seen that recovery was greatest with 2-chlorotrityl resin. The two resins containing the Wang linker produced the lowest recovery, and were essentially identical. Rink amide MHBA fell somewhere in between. This is consistent with the degree of steric shielding for each resin. Furthermore, it would seem that the improved mass transfer kinetics of the NovaSyn® TGA resin provided no detectable protection against back alkylation with the Wang linker. In fact, the polyethylene glycol spacer may be detrimental

as it increases the distance of the carbocation from the polystyrene core, which itself can act to scavenge this reactive species.

Although it is clear that 2-chlorotrityl resin provides the greatest recovery of product, we cannot be sure from these results, how large the difference is, as the relationship between concentration and absorbance at 254 nm may not be linear. Furthermore, the product from Rink amide cleavage is different to the other resins, being the amide rather than the carboxylic acid. It is not known what effect this minor change has on the absorbance character of the compound. However, as it is the shared aromatic structure that will be responsible for the majority of absorbance at 254 nm, this is unlikely to have a significant impact.

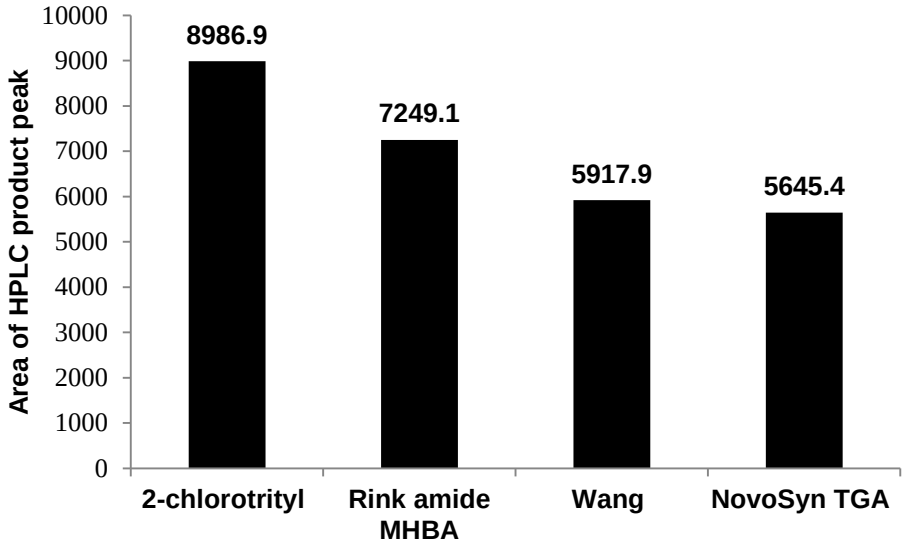


Figure 3.6 Bar chart comparing the HPLC product peak area between different resins (see text). HPLC conditions: 10 µL injection. Agilent Eclipse XDB-C18 column, 4.8 x 150 mm, 5 µm. Solvent A: [Water and 0.05 % TFA], Solvent B: [MeOH and 0.05 % TFA]. Gradient: 0% [B] to 95 % [B], from 0 min to 15 mins, 95 % [B] to 0 % [B] from 15 to 20 mins. Monitored UV 254 nm. Flow rate 1 mL/min. Column temperature 40 ^{o}C.

These results at least confirmed the suspicion that 2-chlorotrityl resin is a more appropriate choice than Wang resin for syntheses incorporating **11**. The increased recovery is most likely due to the reduced occurrence of back alkylation of the resin with the indole scaffold on cleavage. 2-Chlorotrityl resin holds further advantages. Firstly, it is cleaved in only one place. This means it does not produce a soluble linker derived carbocation, which can alkylate the indole scaffold to produce soluble side products. It is possible this also reduces the overall burden on the scavengers. In addition, the triphenylmethyl structure produces a highly stabilised carbocation, and as such 2-chlorotrityl resin can be cleaved under far milder conditions. For example cleavage when

simultaneous *t*-butyl side chain protection is not required, can be affected with a TFA concentration as low as 1%.

Due to the above advantages, and the results of the screening, 2-chlorotrityl resin would be chosen for future work, when a carboxylic acid end group was desired.

3.6 Coupling agent screening.

Following the resin screening, it was decided to use a similar test to compare a range of coupling agents. These were chosen on the basis of availability. In this test the lysine analogue was synthesised on a 0.016 mmol scale using 2-chlorotrityl resin as the solid support. Again the unnatural enantiomer of **11** was used. Reaction conditions were identical, and as already described, except for the reagents used to couple **11** to the resin bound lysine residue. The different coupling conditions tested were as follows:

- **11** (1.1 equiv.), HBTU (1.1 equiv.), HOBt (2 equiv.), DIPEA (2.2 equiv.), DMF, 2 hours.

- **11** (1.1 equiv.), HATU (1.1 equiv.), HOBt (2 equiv.), DIPEA (2.2 equiv.), DMF, 2 hours.

- **11** (1.1 equiv.), PyBop (1.1 equiv.), HOBt (2 equiv.), DIPEA (2.2 equiv.), DMF, 2 hours.

- **11** (1.1 equiv.), EDCI (1.1 equiv.), HOBt (2 equiv.), DIPEA (2.2 equiv.), DMF, 2 hours.

- **11** (1.1 equiv.), DIC (1.1 equiv.), HOBt (2 equiv.), DIPEA (2.2 equiv.), DMF, 2 hours.

After cleavage, the crudes were evaporated to dryness and dissolved in 1 mL of MeOH. 100 µL of the resulting solutions were diluted with 900 µL of MeOH and analysed by reverse phase HPLC. Again the area of the product peak was used as a qualitative comparator of performance. The results are displayed as a bar chart in figure 3.7.

As can be seen HATU was clearly the best performing coupling reagent, HBTU and PyBOP appeared fairly similar, and EDCI and DIC were the worst. HATU is often reported to give better yields than the closely related HBTU. However, the results in this case were surprising. In the first solid phase experiment (section 3.2) a negative Kaiser test suggested that the coupling of **11** with HBTU, admittedly to a different amino acid, was

complete after 2 hours. Therefore, it was not expected that the performance of HBTU could be bettered. There are a number of possible explanations for the improved coupling observed with HATU.

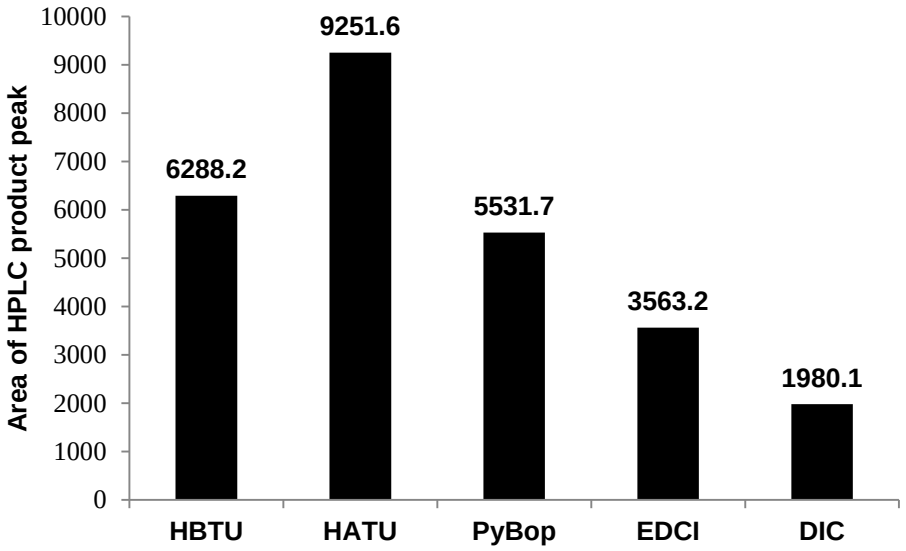


Figure 3.7 Bar chart comparing the HPLC product peak area between different coupling conditions (see text). HPLC conditions: 10 µL injection. Agilent Eclipse XDB-C18 column, 4.8 x 150 mm, 5 µm. Solvent A: [Water and 0.05 % TFA], Solvent B: [MeOH and 0.05 % TFA]. Gradient: 0% [B] to 95 % [B], from 0 min to 15 mins, 95 % [B] to 0 % [B] from 15 to 20 mins. Monitored UV 254 nm. Flow rate 1 mL/min. Column temperature 40 ^{o}C.

It may be that coupling of **11** to lysine is less efficient than to alanine, and therefore HATU was able to perform better. Another possibility is that Kaiser testing after coupling of **11** is unreliable, and gives a false negative result. This could have occurred for example, if hydrophobic or aromatic stacking interactions between different Fmoc-protected duocarmycin residues, precluded entry of the ninhydrin to the inner core of the matrix, where the amine of the unreacted resin bound amino acid would be situated. A third possibility is that despite premixing before addition to the resin, HBTU reacts slower than HATU to form the activated ester, and therefore was able to undergo some completing guanylation of the amine. This would also produce a negative Kasier test but leave room for HATU to perform better. However, the last two explanations seem unlikely, as if the Kaiser test in the first experiment was a false negative, we would have expected to see at least the deletion product arising from cleavage of the uncoupled or guanylated alanine, which was not observed.

Nevertheless, the screen clearly suggests HATU to be superior for couplings of **11**. HATU is a very closely related uronium salt to HBTU. In fact, it differs by only one atom; a pyridine ring in place of a benzene ring (figure 3.8). The increased coupling efficiency

often reported with HATU is attributed to a neighbouring group effect produced by this extra nitrogen atom, which stabilises the transition state of amide bond formation (figure 3.8).[154] Another, and not mutually exclusive explanation, is that the electron-withdrawing effect of the nitrogen atom improves the stability of the HOAt moiety as a leaving group, increasing reactivity over HOBt esters.

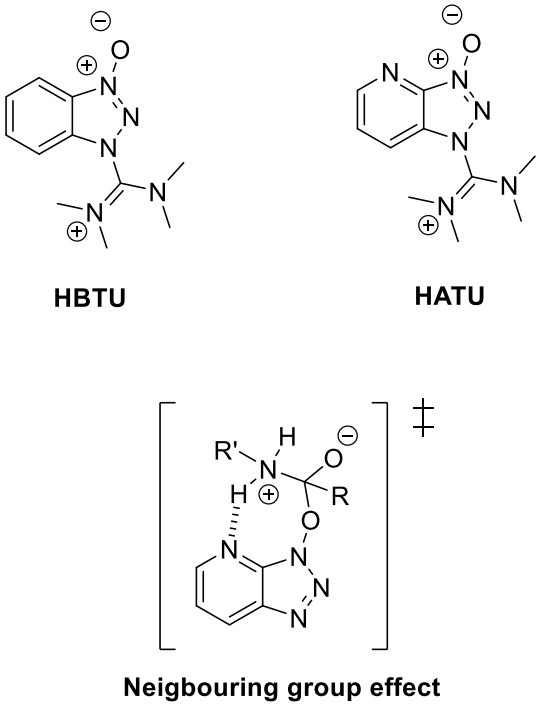


Figure 3.8 Structure of HBTU and HATU. Neighbouring group effect, during amide coupling of HOAt ester.

The structures of PyBop, EDCI, and DIC are shown in figure 3.9.[154] PyBOP is phosphonium salt, and serves as another means of introducing the HOBt ester, *via* the acyloxyphosphonium intermediate. EDCI, and DIC are both carbodiimides which react with carboxylic acids to form their respective *O*-acylisoureas. In the presence of HOBt as in this case, the *O*-acylisoureas serve as intermediates towards the HOBt ester.

It is actually illogical to use HOBt as an additive with HATU. This is because it would be completing with the HOAt released from HATU, and therefore potentially hinder the formation of the desired more reactive HOAt ester. Thus, it is possible that HATU would perform even better than the screen has suggested if this additive was not used. Even so HATU clearly outperformed the other coupling reagents, and was therefore chosen for future work. However HOBt would not be used as an additive in the future.

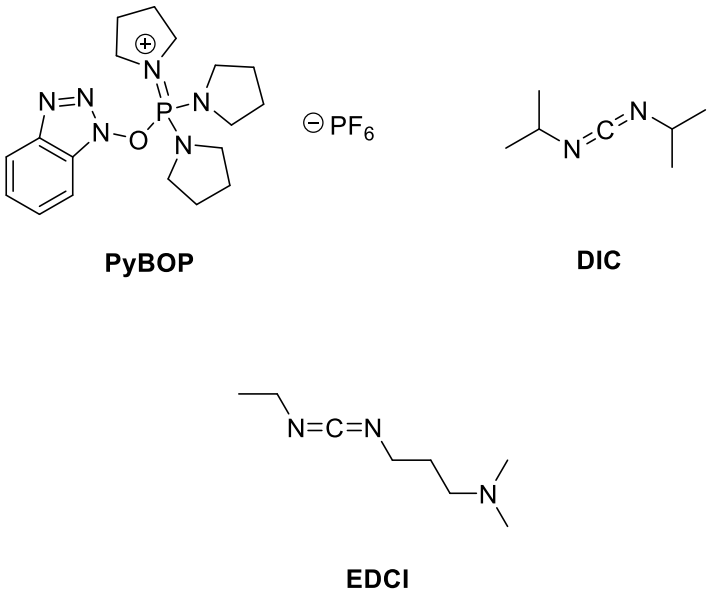


Figure 3.9 Stuctures of PyBOP, EDCI, and DIC.

In retrospect, a better way to compare the coupling performance of these reagents, would have been to stop the synthesis after coupling of **11**. Then the resins could have been dried, and the loading of **11**, determined by monitoring Fmoc-deprotection, as described in section 3.3. However this would have been significantly more time consuming, and the simple HPLC test, provided the desired information.

3.7 On resin benzyl deprotection tests.

Before reattempting the synthesis of the small library using the improved resin and coupling agent choice, it was decided to briefly explore whether benzyl deprotection could be performed on the solid phase immediately prior to cleavage. This would provide compounds for assay without the need for post solid phase reactions.

The conditions previously used to remove benzyl protection after cleavage would not be appropriate, as the heterogeneous catalyst would not be able to access the matrix of the swelled resin. Therefore attempts were made to affect benzyl ether cleavage, under homogenous conditions, using soluble Pd species, with organosilanes as the hydrogen source. Such conditions have been reported for the reductive cleavage of benzyl ethers in solution.[178]

To test the conditions, the lysine analogue **39** was again resynthesized, using the previously described conditions, but on 2-chlorotrityl resin and with HATU as the coupling reagent. After acetylation of the indoline nitrogen, the resin was dried. Samples of the dried resin were then swelled in anhydrous DCM and degassed with a steam of N_2, before

being subjected to different deprotection conditions. As the HPLC retention time of both the benzyl protected and the free phenol form of the lysine analogue were known, the success of on resin benzyl deprotection could be assessed by HPLC analysis of the crudes after release from the solid support.

In the end only two conditions were tested, and neither was successful. In both cases only the benzyl protected product was detected by HPLC after cleavage. The conditions tested were as follows:

- $Pd(OAc)_2$ (30 mol %), Et_2SiH (25 equiv.), Et_3N (25 equiv), anhydrous DCM, N_2, overnight.

- $Pd(PPh_3)_4$ (25 mol %), Ph_3SiH (25 equiv), anhydrous DCM, N_2, overnight.

This was a far from exhaustive search, and it is possible that a method to affect on resin benzyl deprotection could be found if more time was invested. It would be a useful addition to the process, but as it is not essential, it was decided to focus efforts on analogue generation.

It is also possible that the structure of **11** could be improved by the replacement of the benzyl ether, with an alternative protecting group. For example, the use of the more acid liable *para*-methoxybenzyl ether, would most likely allowed simultaneous deprotection during TFA mediated cleavage of the products from the resin. This would of course preclude cleavage conditions that utilise ultralow concentrations of TFA. However, cleavage with up 47.5 % TFA has been shown to be tolerable, so it is a feasible option. It would certainly offer flexibility, allowing cleavage of the protected products in 1 % TFA, or the deprotected products at higher concentrations.

However, the *para*-methoxybenzyl ether protecting group would not be compatible with the current synthesis of the **11**. The activating effect of the methoxy group, would likely be susceptible to iodination, by NIS. Furthermore, it would probably be cleaved during the late stage Boc-deprotection step. Although, it could conceivable be reintroduced after Fmoc-protection. In fact, if the group was to be cleaved, and then reintroduced, then proceeding iodination of the original *para*-methoxybenzyl ether might be acceptable.

It certainly seems that the replacement of the benzyl ether, with a *para*-methoxybenzyl group, would be worth exploring in the future. However, it was not within the scope of these studies, as large scale synthesis of **11**, had provided a multigram quantity of the benzyl-protected building block, and it was deemed efforts were better focused on continuing to explore its incorporation in to solid phase synthesises.

3.8 Repeat of the solid phase synthesis of the small library of amino acid-duocarmycin conjugates.

The library of amino acid-duocarmycin analogues was resynthesized, on 2-chlorotrityl resin, using the natural enantiomer of **11**, and employing HATU as the coupling reagent. The solid phase synthesis of each analogue was conducted on a 0.038 mmol scale, starting from commercially available preloaded resins. In addition to the five original analogues a sixth β–alanine analogue was also synthesised. Commercially preloaded 2-chlorotrityl resins are supplied as the free α-amine, thus do not require initial Fmoc-deprotection. They are supplied in this manner, as 2-chlorotrityl resins loaded with an Fmoc-protected amino acid have reduced stability on storage.[145] This may be attributable to mechanical stress imposed by the proximity of the large aromatic Fmoc group to the trityl structure.

The increased acid sensitivity of the 2-chlorotrityl linker was exploited to further reduce the TFA concentration during cleavage. For analogues not requiring *tert*-butyl side chain deprotection, cleavage was affected using 1 % TFA, and 10 % TIPS in DCM, while the TFA concentration was increased to between 10 % and 20 % with extended reaction times for those that did. Considerable variability in side chain deprotection rates was observed. For example, removal of the *tert*-butyl carbamate of the lysine analogue was complete with 2 hours at a TFA concentration of 10 %. However, the more stable *tert*-butyl ether of the serine analogue, required treatment with a TFA concentration of 20 % overnight. It is not clear whether a benefit is derived by reducing the TFA concentration below 47.5 %, when extended reaction times are required at lower concentrations. It maybe that 47.5 % is the optimal compromise between TFA concentration and reaction time for such analogues.

The scavenger composition of the cleavage cocktails was also changed from the original optimised conditions, which employed 2.5 % TIPS, and 2.5 % H_2O. This change was made because it was found that H_2O no longer served as an effective scavenger when TFA concentration was reduced below 47.5 %. When the lysine analogue was first cleaved using 10 % TFA in DCM with 2.5 % TIPS, and 2.5 % H_2O, significant formation of a side product suspected to have resulted from alkylation of the indole with the *tert*-butyl cation, was observed. Increasing the concentration of TIPS to 10 % effectively suppressed the formation of this side product. It is likely the reduction in the effectiveness of H_2O as a scavenger results from its reduced miscibility with the larger volumes of DCM now being used. Interestingly, if 10 % H_2O was also included in the cleavage cocktail, this

supressed side chain deprotection of the lysine. This might be because the immiscible water layer acted as a proton sink, reducing the acidity of the DCM layer. It is likely that the concentration of TIPS could be reduced by including other DCM miscible scavengers such as EDT. However, this was not explored, due to the noxious nature of this compound.

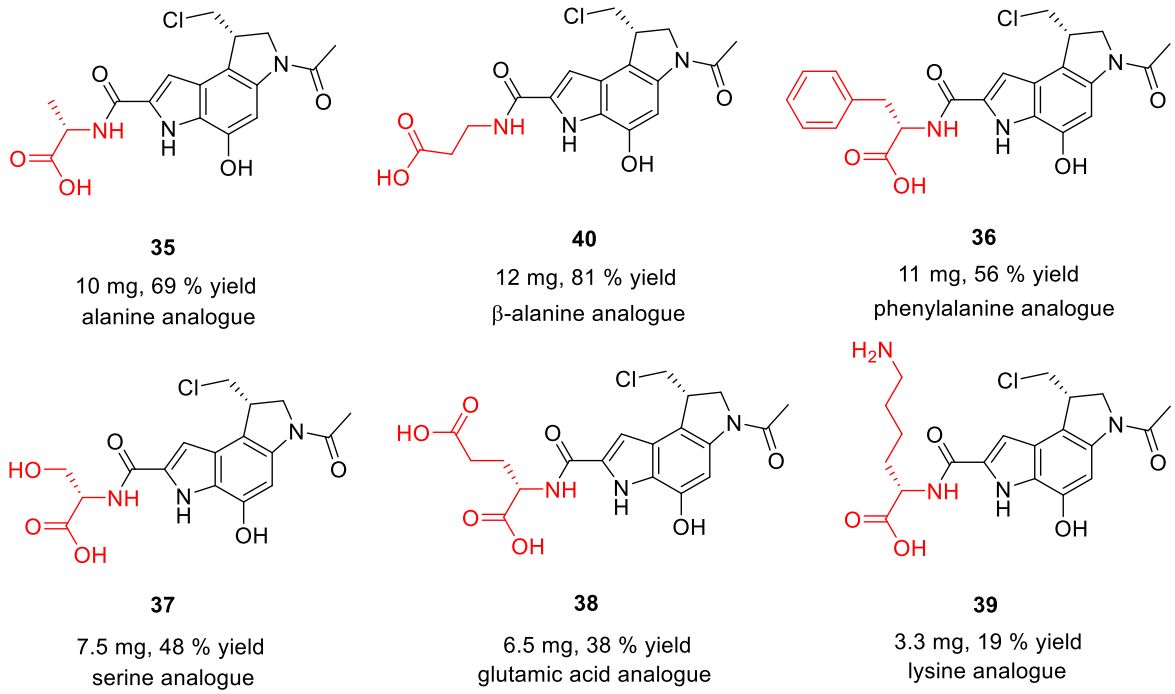


Figure 3.10 Structure, recovery, and yield of the analogues made in this library.

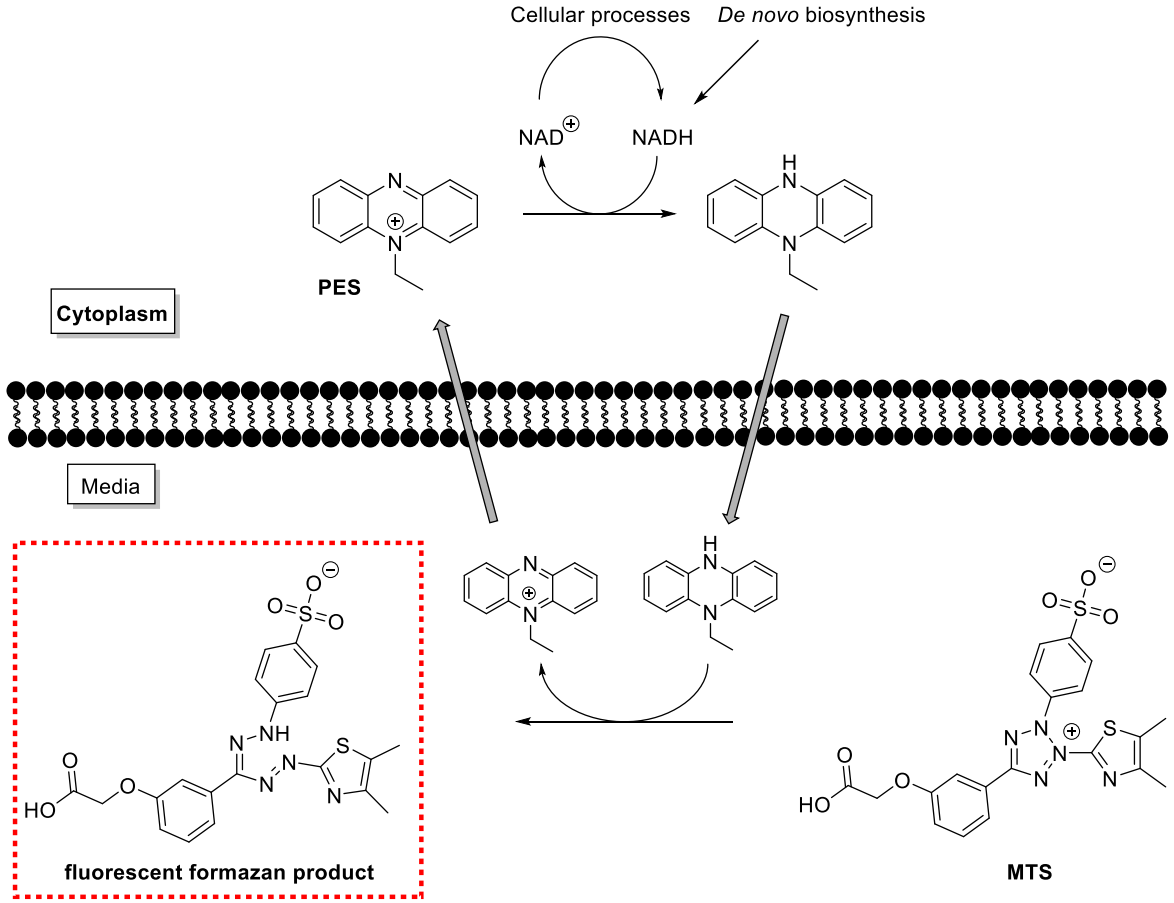
Following cleavage the crude products were subjected to benzyl deprotection by catalytic transfer hydrogenation as previously described. Isolation from the excess ammonium formate, was affected by either reverse phase preparative HPLC, or silica gel flash chromatography, depending on the polarity of the amino acid side chain. Suitable mobile phase conditions could not be found to elute polar analogues on normal phase silica gel. In an attempt to mediate against losses during preparative HPLC the crudes were injected in a minimal volume of DMSO, and the injection method set to insert an air plug either side of the sample. This technique prevents mixing with the mobile phase, until the sample has reached the column head.

All analogues were recovered in a usable quantity, ranging from 3.3 mg to 12 mg, representing isolated yields of between 19 % and 81 % (see figure 3.10). The best yields were for those analogues purified by flash chromatography (**35**, **36**, **40**). This may not represent intrinsically better yielding reactions for these analogues, but reflect that suboptimal recovery was being experienced with the preparative HPLC instrument. The

desired structures were confirmed by a combination of accurate mass spectrometry, ^{1}H NMR and DEPT-edited-HSQC experiment.

3.9 Assessment of the antiproliferation activity of 35, 36, 37, 38, 39, and 40.

The antiproliferative activity of the analogues against an HL-60 cell line was assessed using a MTS assay.[179] MTS is a tetrazolium compound that can be reduced by incubation with metabolising cells to give a fluorescent formazan product. This signal can be measured, and is proportional to the number of viable cells. The cell permeability of MTS is poor, and it is therefore used in conjunction with an intermediary electron acceptor, in this case PES. PES is reduced in the cytoplasm by NADH, and in turn reduces MTS in the media (scheme 3.10). In non-metabolizing cells, NADH is quickly depleted, and thus MTS reduction is inhibited.

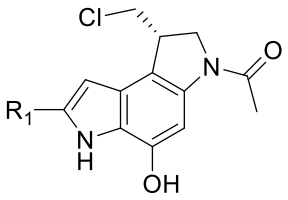


Scheme 3.10 Reduction of MTS by viable cells, mediated by PES.

In the antiproliferation assay, a fixed starting concentration of cells is incubated with different concentrations of the compound under investigation, for a fixed time; in this case 72 hours. The MTS/PES reagent is then added. After incubation for a further 4 hours, the formazan signal is read. Plotting this against the log of compound concentration produces a dose response curve, from which an IC_{50} value can be calculated for the inhibition of proliferation, and this serves as an indirect measure of cytotoxcity.

HL-60 cells are a naturally immortal cell line, originally derived from leukocytes isolated from blood donated by a female patient diagnosed with acute myeloid leukemia.[180] Being derived from blood, HL-60 cells are naturally adapted to allow *in vitro* culture as single-cell suspensions. This allows HL-60 cell cultures to be conveniently passaged by dilution. Passaging is essential to maintain the cells in a consistent growth phase, and avoid variation between assays that might arise from senescence. When appropriately cultured, HL-60 cultures have a doubling time of around 24 hours.[181] The HL-60 cell line was chosen here for reasons of availability and ease of culturing.

The resulting IC_{50} for each analogue is shown in table 3.1, along with that of the commercially available cyctotoxic DNA intercalating agent doxorubicin, which was also tested as a positive control.

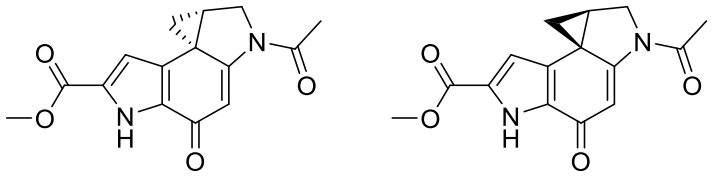


	R_1	IC_{50} (μM)	[95 % CI]
35	CO-Ala-OH	85	[45 - 165]
40	CO-β-Ala-OH	>300	
36	CO-Phe-OH	32	[13-78]
37	CO-Ser-OH	40	[12-115]
38	CO-Glu-OH	286	[101-814]
39	CO-Lys-OH	>300	
Doxorubicin (+ve control)		0.125	

Table 3.1 Mean IC_{50} values with 95 % confidence intervals returned by the MTS assay. The assay was performed in triplicate HL-60 cell line. Detailed protocol can be found in the experimental chapter.

It was expected, that all of the analogues would have similar activity to previously reported truncated duocarmycin analogues. For example the *N*-acetyl duocarmcyin analogue **41** (fig 3.11), possessing the *C*-terminal ester of the natural product, has been reported to

have low nM activity, with IC_{50} values of 1 nM and 45 nM, for the natural and unnatural enantiomers respectively.[182] Admittedly these values are reported in a different cell line (murine L1210 cells). However, there is no reported evidence of HL-60 cells possessing an inherent resistance to duocamycin analogues.



(+)-**41**
IC50 (L1210) 1 nM

(-)-**41**
IC50 (L1210) 45 nM

Figure 3.11 Structure of **41**.

It is also true that the low nM IC_{50} values are reported for the intact spirocyclised alkylation subunit. However, it is generally accepted that halide seco forms of the alkylation subunit are equipotent to their spirocyclised counterparts, and that spriocyclisation in cells is not rate limiting.[94]

Therefore, it would appear that addition of a *C*-terminal amino acid to the *N*-acetyl alkylation subunit results in at least a 1000 fold decrease in cytotoxicity. The results also tentatively suggest that the structure of the side chain may not be trivial, and apparent variation in activity was observed between the tested analogues, ranging from 30 μM to greater than 300 μM. However, with many of the confidence intervals overlapping these differences have not been shown to be significant in this study.

The suspicion at this point was that the large reduction in activity most likely reflected a decrease in cell permeability. Common to all the analogues tested is the α-carboxyl group. These carboxylic acids will be largely ionized at physiological pH, thus inhibiting passive diffusion through the hydrophobic core of the cell membrane. Differences in cell permeability might also explain some of the apparent side chain effects. For example, the most potent analogue **36** is that possessing the non-polar aromatic side chain of phenyl alanine. This could increase the lipophilicity of this analogue promoting passive diffusion. In contrast, two of the least potent analogues **38** and **39**, contain additional ionisable groups in their side chains; the negativity charged carboxylate of glutamic acid, and the positively charged amine of lysine, this could further inhibit passive diffusion.

However the correlation between hydrophilicity and reduced activity, is not completely consistent, as is highlighted by the increased activity of the more polar serine analogue

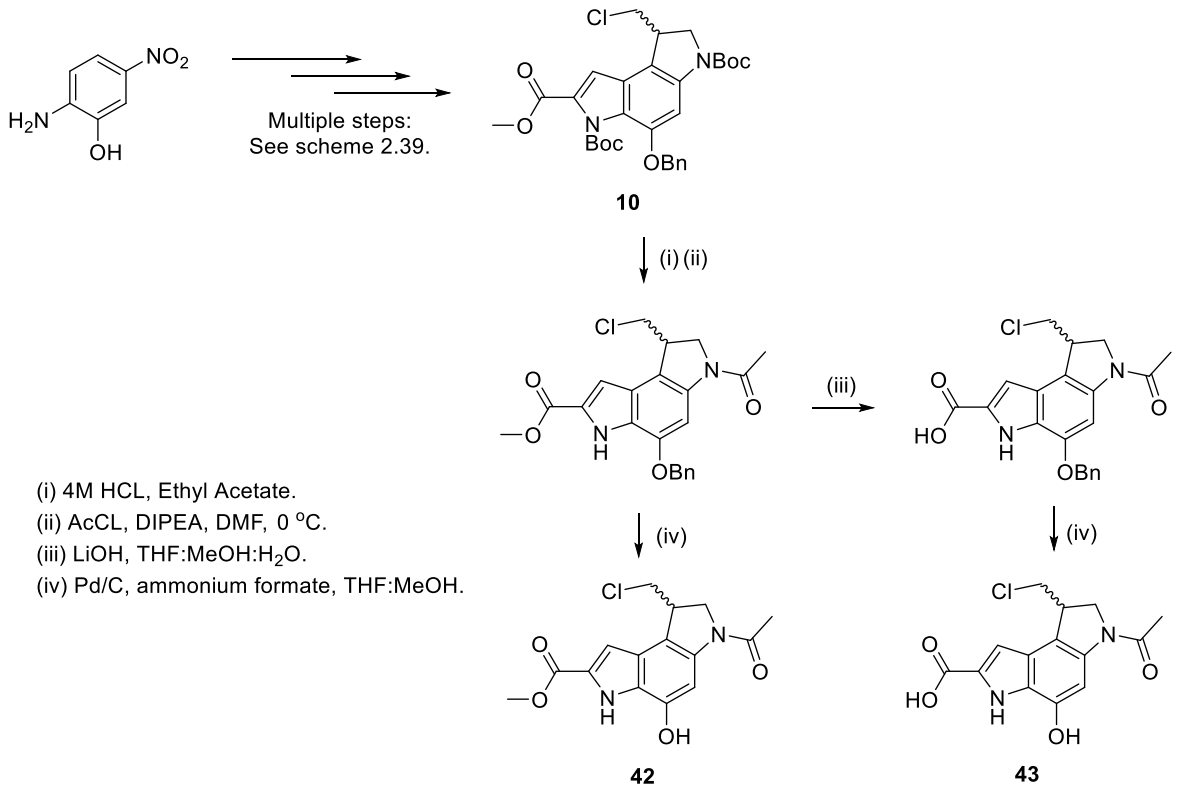
(**37**) relative to the alanine analogue (**35**). Furthermore, the linear β-alanine analogue (**40**) appears to be considerably less active than its branched alanine counterpart. Clearly if the observed differences in activity are real, more than one factor would have to be contributing. It seems likely that the different sidechains, could have both positive or negative effects on DNA binding, as well as cell permeability, and it could be the relative combination of these factors which disrupts the trend between hydrophilicity and reduced activity.

As discussed, despite the possible differences in activity observed between the analogues, they all exhibited a reduction in cytotoxicity of over a 1000 fold, when compared to **41**, and this most likely represents decreased cell permeability due to ionisation of the common α-carboxyl group. Of course the pK_a of this group will not be identical for each analogue. However, they are likely to be similar, and thus small differences in the precise molar fraction of unionised acid, are unlikely to contribute significantly to differences in cell permeability. The pK_a of the α-carboxyl group is around 2, for all natural amino acids.[183] This is considerably lower than most carboxylic acids (for example acetic acid has a pK_a of around 4.75).[183] The low pK_a values of the α-carboxyl group of amino acids is attributable to the presence of the α-amino group. In the case of these analogues this group is masked by the amide bond. Therefore, it is likely that the pK_a of the α-carboxyl group of these analogues is higher than that of the corresponding free amino acids. The lysine analogue may be an exception, and is likely to have the lowest pK_a, considering it is still able to form a zwitterion with the amino group of its side chain. This might be contributing to the apparent inactivity of this species in our assay. However, assuming a physiological pH of around 7.4, and that the pK_a of all the α-carboxyl groups, are likely to be below 6.15, it seems likely that all the analogues would be over 95 % ionized.

It is conceivable that this negative charge could also have a disruptive effect on minor grove binding, considering the potential charge repulsion with the polyanionic backbone of DNA. However, as will be seen from later results this appears not to be the case.

3.10 Synthesis and antiproliferative activity of controls 42, and 43.

In order to assess the impact of a *C*-terminal carboxyl group in isolation from the effects of the amino acid residue, it was decided to synthesise control compounds **42**, and **43**; these being the seco *N*-acetyl alkylation subunit with the natural *C*-terminal ester, and its carboxyl counterpart respectively (scheme 3.11).



Scheme 3.11 Synthesis and structure of **42** and **43**.

Compounds **42**, and **43**, were accessed *via* the same synthetic route used to synthesize **11**, up until the common di-Boc protected intermediate **10**. Exhaustive Boc-deprotection and acetylation with acetyl chloride gave the benzyl protected precursor to **42**. Benzyl ether cleavage by catalytic transfer hydrogenation as before, either immediately, or following methyl ester hydrolysis afforded **42** and **43** respectively (scheme 3.11). The desired structures were confirmed by a combination of accurate mass spectrometry, ^{1}H NMR and DEPT-edited-HSQC experiment, and further characterised by ^{13}C NMR.

Antiproliferative activity of the racemates was assessed by MTS assay as before in the HL-60 cell line. While the intact ester **42** returned a mean IC_{50} of 25 nM (95 % CI: 11-37 nM, triplicate), in line with its spirocyclized counterpart **41**, and over 1000 fold more potent

than the amino acid analogues, the carboxylic acid **43**, showed no activity at all even at the highest concentration tested (500 μM). This clearly demonstrated that the presence of the ionisable *C*-terminal carboxyl group was having a detrimental effect on cytotoxicity, most likely by inhibiting passive diffusion through the cell membrane.

Given the complete lack of activity shown by **43**, it now seemed surprising that the amino acid-duocarmycin conjugates had shown any activity at all. Thus it would appear, in comparison to **43**, the amino acid structures seemed to be offering some mitigation against the negative impact on cellular activity imposed by the free carboxylic acid. As discussed previously in relation to the variation in activity observed between the different amino acid analogues, it seems unlikely that potential differences in the pK_a could account for the difference in activity between **43** and the amino acid-duocarmycin conjugates.

It is possible that passive diffusion of the unionised fraction of the amino acid analogues is faster than that of **43** due to the increased hydrocarbon scaffold provided by the amino acid. However, it is also intriguing to speculate that perhaps the amino acid analogues are able to benefit, at least partially, from active transport processes used by cells to control intracellular amino acid homeostasis. It is possible that recognition by such transport proteins, may be providing greater cell permeability when compared to **43**. This might also offer insight into why the β-alanine analogue appeared considerably less active than the alanine analogue. For example, perhaps HL-60 cells do not possess the same capacity for the active transport of β-alanine, as they might for alanine. However this is entirely speculative, although some amino acid related drugs have been shown to utilise transporters whose endogenous function is to mediate the uptake of amino acids. For example the phenylalanine derived nitrogen mustard melphalan, has been shown to cross cell membranes *via* recognition of the LAT-1 amino acid transporter.[184]

3.11 Synthesis and antiproliferative activity of the first extended amino acid-duocarmycin conjugate.

The complete inactivity of **43** in the MTS assay, clearly demonstrated that a *C*-terminal carboxyl group had the potential to abate entirely the cytotoxicity of the duocarmycin alkylation subunit. However, the surprising activity of the amino acid-duocarmycin conjugates, although still over 1000 fold lower than that of the ester **42**, was encouraging. Clearly the additional structure of the amino acid scaffold, was providing some mitigation against the detrimental effect of the carboxyl group. As discussed, this may in part be due to recognition by amino acid transporters, but could also represent faster passive diffusion due to the increased hydrocarbon structure. This appeared evident in the observation that

the phenylalanine analogue containing the most hydrophobic sidechain appeared to be the most active of the amino acid conjugates. It was therefore intriguing to question how much activity would be recovered by replacing *N*-terminal acetylation, with the coupling of an additional hydrophobic group.

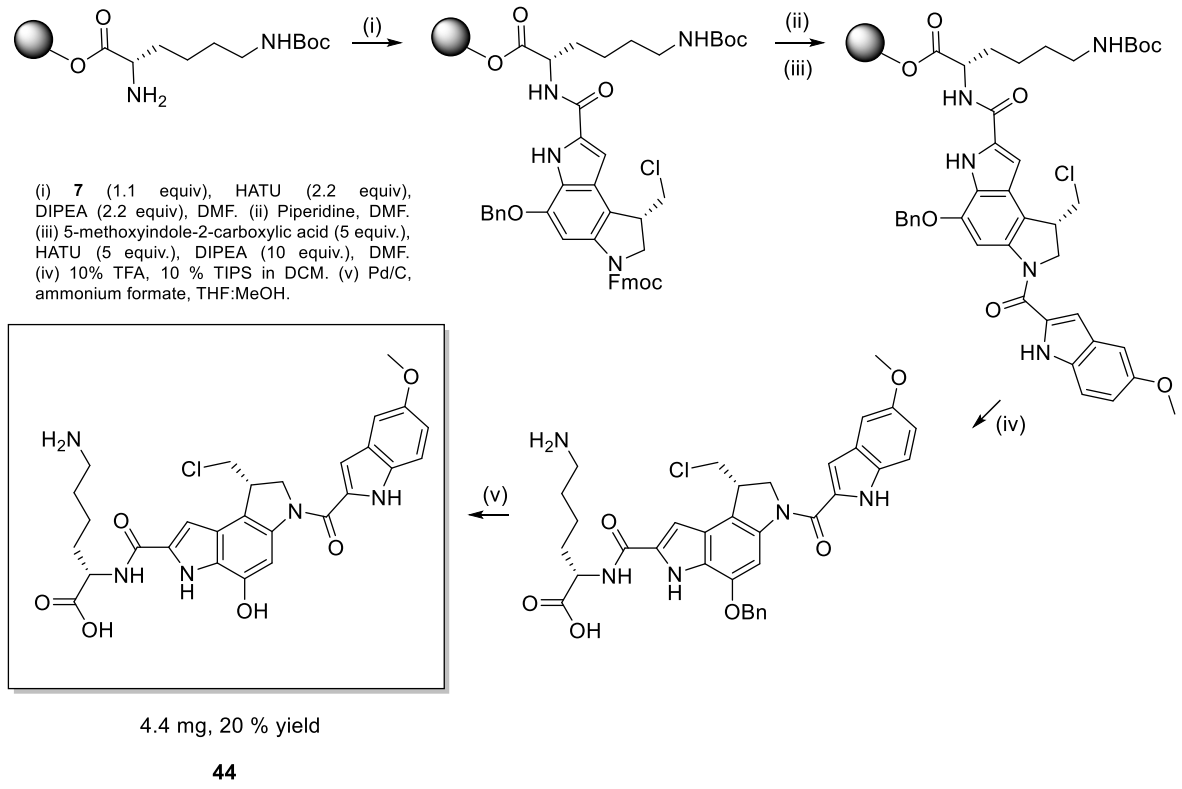
The group chosen was 5-methoxyindole. Coupling of this group *via* the commercially available 5-methoxyindole-2-carboxylic acid, would provide an amino acid-duocarmycin conjugate which was more similar in structure to the full natural product, which possesses an *N*-terminal trimethoxyindole unit. Such an analogue would be expected to benefit from faster passive diffusion, and increased affinity for the minor groove.

It was decided to synthesis the 5-methoxyindole equivalent of the zwitterionic lysine analogue **39**. This had been one of the least potent of the original analogues, and it was therefore particularly interesting to see how much activity could be recovered by *N*-terminal extension. Furthermore, it was still anticipated that if cell permeability was improved by the additional indole subunit, that the cationic side chain of lysine might provide further favourable electrostatic interactions with the polyanionic backbone of DNA.

As such, the solid phase methodology was employed to synthesize the extended amino acid-duocarmycin conjugate **44** (Scheme 3.12). As before, the synthesis was conducted on a 0.038 mmol scale, beginning from a commercially preloaded lysine 2-chlorotrityl resin. Synthesis proceeded as already described, with the replacement of acetylation of the indoline nitrogen with the coupling of 5-methoxyindole-2-carboxylic acid, using HATU. Cleavage was affected using 10 % TFA, 10 % TIPS, in DCM for 2 hours, to give one product.

Benzyl deprotection was affected as previously described by catalytic transfer hydrogenation. However, it is noteworthy to highlight that this was not as straightforward as with previous analogues. During initial attempts it was found that the extended lysine analogue was unexpectedly susceptible to completing dehalogenation. The particular susceptibility of the lysine extended analogue to this side reaction is not known. It was observed that benzyl deprotection was much slower than with previous analogues, and the later treatment with additional equivalents of Pd/C and ammonium formate to drive the reaction to completion, promoted the initiation of the side reaction. Furthermore, performing the reaction with lower quantities of both the catalyst and hydrogen donor resulted in no reaction. It was questioned whether the amine side chain could be poisoning the catalyst, or promoting dehalogenation. However, attempts to perform benzyl deprotection prior to removal of Boc-protection of the sidechain made no improvement. This was achieved by cleavage of the Boc-protected product, using 1 % TFA. In the end

limiting the reaction time, and separating the desired product from unreacted starting material by reverse phase HPLC, gave the best recovery (4.4 mg, 20 % yield). The desired structure was confirmed by a combination of accurate mass spectrometry, ^{1}H NMR and DEPT-edited-HSQC experiment.



Scheme 3.12 Synthesis of the extended lysine analogue **44**.

Antiproliferative activity was assessed by MTS assay as before in the HL-60 cell line, returning a mean IC_{50} value of 374 nM (95 % CI: 200-701 nM, triplicate). Therefore, addition of the methoxyindole subunit had increased the potency of the extended lysine analogue **44** by over 1000 fold, compared to the truncated counterpart **39**, the IC_{50} of which was estimated to fall between 300 µM and 500 µM, but could not be calculated due to lack of activity at the available concentrations returning an incomplete dose response curve.

Clearly the methoxyindole unit was very effective at limiting the detrimental impact of the zwitterionic *C*-terminal lysine residue. This is consistent with expected faster passive diffusion of unionised fractions promoted by the increased hydrophobic structure. However, improved activity is also likely to have resulted from more efficient DNA alkylation. This being a result of both improved non-covalent affinity for the minor groove,

and more effective activation of the cyclopropane caused by the increased ridged length of the compound (see chapter one).

Despite the dramatic improvement, **44**, was still over 14 times less active than the truncated alkylation subunit possessing the *C*-terminal ester, **42**. Furthermore, it was over 1000 fold least active then the full natural product, to which it bears closer structural homology, with duocarmycin SA reportedly returning IC_{50} values of as low as 0.006 nM (murine L1210).[185]

It was therefore clear that the *C*-terminal amino acid was still having a detrimental effect on cytotoxicity. This may still be through limited cell permeability, but might also be the result of disruption to minor groove binding, or perhaps more likely a combination of both factors.

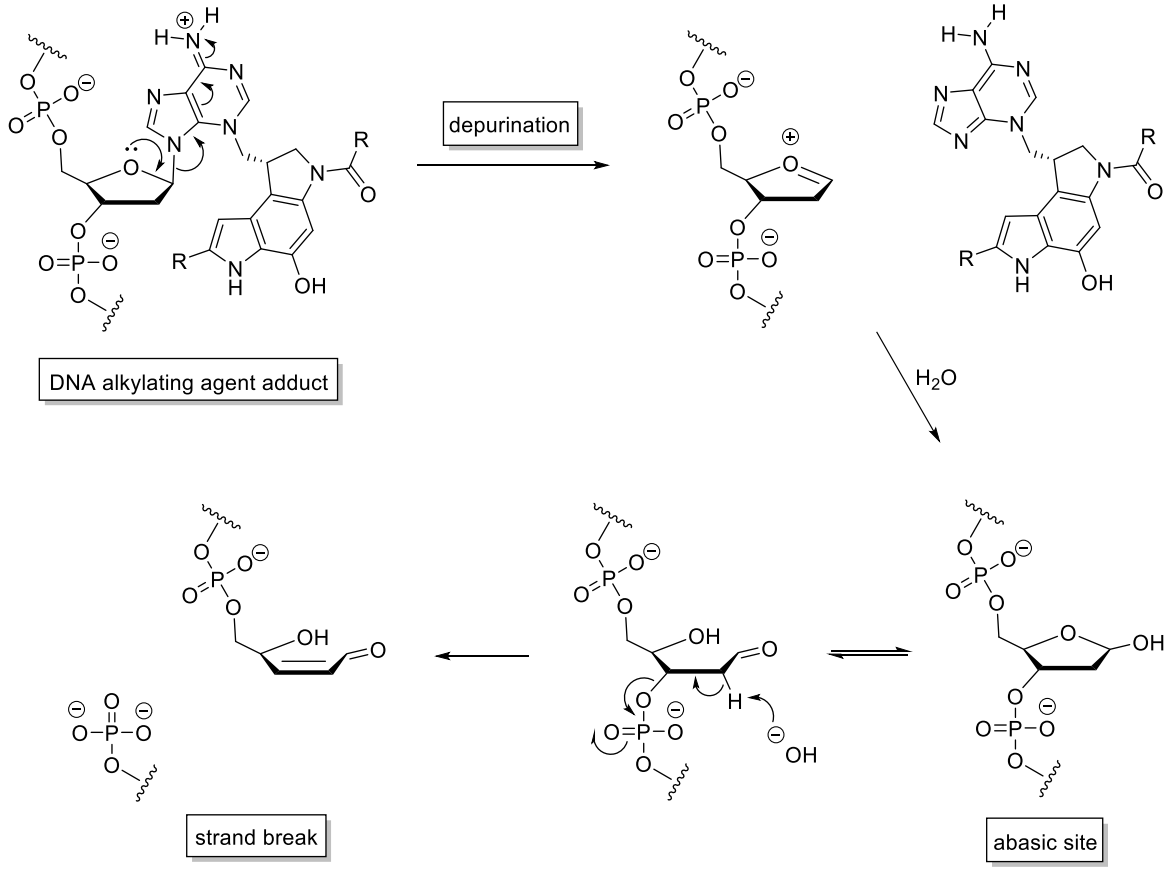
3.12 Cell free DNA alkylation Assay.

The comparably low activity of these analogues in the MTS assay seemed most likely to be the result of poor cell permeability. However, it was also possible that the *C*-terminal carboxyl group or the amino acid structure itself could have a detrimental impact on DNA alkylation by disrupting minor grove recognition. It was therefore apparent that it would be desirable to ascertain if these compounds could indeed alkylate DNA as expected.

DNA footprinting techniques were deemed most useful to this end, as they would give information not just on the ability of these compounds to alkylate DNA, but also highlight any sequence selectivity they may exhibit.[186] This work was carried out by Prof. Keith R Fox from the University of Southampton, an expert in the sequence specific recognition of DNA by small molecules.

This assay takes advantage of the destabilising effect of DNA alkylation,[187] which results in a strand break upon thermal treatment at the site of covalent attachment of the alkylating agent. The mechanism of strand cleavage is outlined in scheme 3.13. Briefly, alkylation of the nucleobase results in a positively charged adduct. This species destabilises the glycosidic bond between the deoxyribose sugar and the alkylated base. Thermal treatment accelerates spontaneous cleavage of this bond producing an abasic site in the DNA backbone. The resulting oxocarbenium ion reacts with water to give the cyclic acetal. This species exist in equilibrium with the ring opened aldehyde. The acidic hydrogen vicinal to the carbonyl can be abstracted initiating strand cleavage through the β-elimination of the phosphate group at the 3’ end of the abasic site.

In the thermal cleavage assay, the compounds under investigation are incubated with singly end-labelled DNA. In this case the 3' end is radiolabeled with ^{32}P. Following incubation the DNA is isolated and thermally treated to induce strand breaks at the site of alkylation. Separation of the DNA fragments based on their relative sized is achieved by denaturing polyacrylamide gel electrophoresis. The gel is then dried and the bands visualised using a phosphorimaging screen, which detects the radioactive signal. As the sequence of DNA is known, the relative position of the bands against a Maxam and Gilbert control ladder, allows identification of the sites of alkylation.



Scheme 3.13 Thermal cleavage at the site of DNA alkylation.

The analogues were tested in the thermal cleavage assay using the MS1 DNA fragment.[188] This is a sequence designed by the Fox group as a general tool for studying the sequence selectivity of minor groove binders, and contains all 136 possible tetranucleotide sequences. The exact sequence of the MS1 fragment can be seen later in figure 3.17.

Figure 3.12 shows the DNA cleavage by the *N*-acetyl analogues. The compounds were incubated with 1.5 µL (10 nM) of the radiolabelled MS1 fragment at concentrations of 50 µM and 5 µM, overnight at 37 ^{O}C, in 10 mM Tris-HCl (pH 7.5), containing 10 mM NaCl.

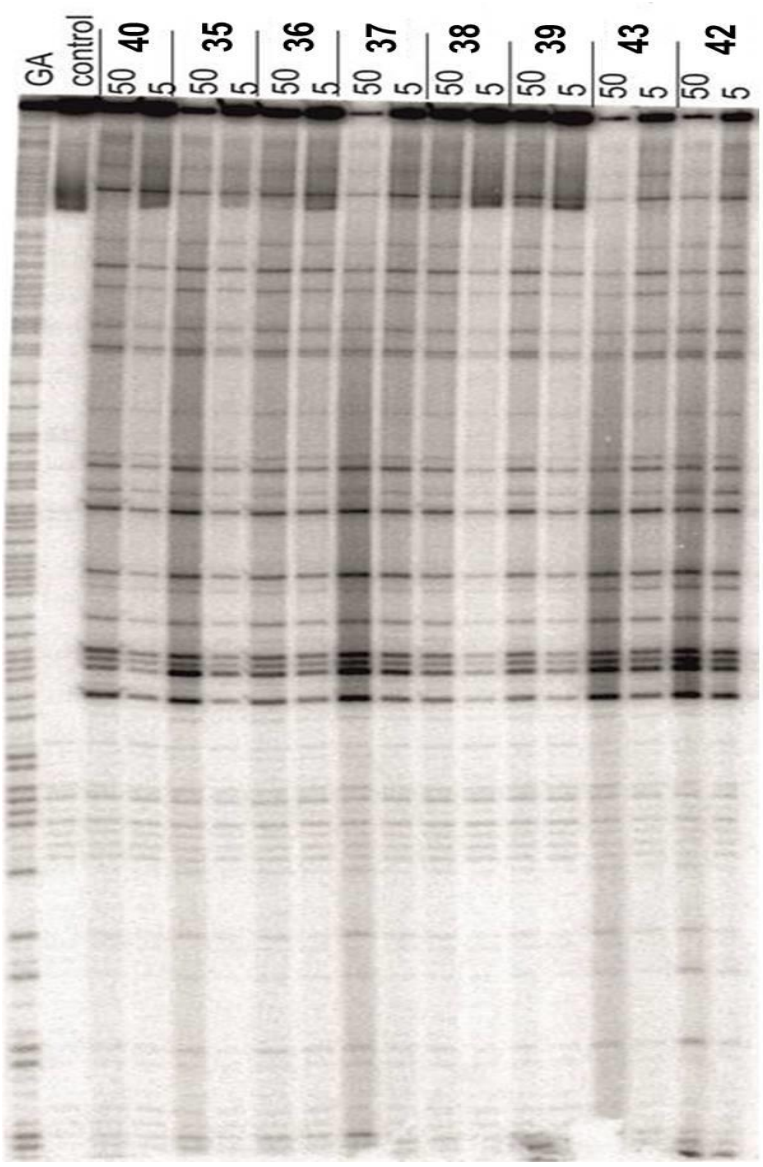
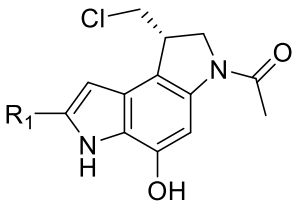


Figure 3.12 DNA cleavage by the *N*-acetyl analogues at incubation concentrations of 50 μM and 5 μM. Lane GA = G+A Maxam and Gilbert ladder. Lane Control = negative control (non-cleaved DNA).

Sample Key:



*(racemic C-CH_2Cl)

	R_1
35	CO-Ala-OH
40	CO-β-Ala-OH
36	CO-Phe-OH
37	CO-Ser-OH
38	CO-Glu-OH
39	CO-Lys-OH
42 *	$COOCH_3$
43 *	COOH

The gel clearly shows that all of the analogues alkylate DNA under these conditions. There is no apparent difference in sequence selectivity exhibited between the truncated analogues. All appear to alkylate only A as expected, and show a preference only for either a 5' T or A neighbouring the alkylation site. This is consistent with previously reported truncated duocarmycin analogues.[189]

The apparent requirement for a neighbouring 5' A or T base has been attributed to offering greater access to the alkylation site due to allowing deeper penetration of the minor groove. The additional amine of a flanking G base projects into the minor groove and thus sterically precludes sufficient minor groove penetration, reducing access to the alkylation site. In addition the narrower minor groove width of AT rich sequences has also been suggested to lead to improved non-covalent binding through stronger van der Waals interactions. Stronger non-covalent binding is advantageous due to the reversible nature of the alkylation event.

As all of the *N*-acetyl analogues including the ester and acid control compounds, **42** and **43**, exhibit essentially identical preference for alkylation sites, it would appear that *C*-terminal amino acid functionality has no effect on the sequence selectivity of the alkylation event. However it would seem that with the exception of the serine analogue, the *C*-terminal amino acid group, does appear to have a negative impact on the efficiency of DNA alkylation. This is evident in the differences in the amount of uncleaved DNA remaining at the top of each lane after electrophoresis. It is clear from visual inspection of the gel, that with the exception of the serine analogue, all analogues containing *C*-terminal amino acid functionality have considerable more uncleaved DNA remaining relative to the ester control **42**.

It would be possible to accurately quantify the differences in the remaining uncleaved DNA by measuring the intensity of the radioactive signal. Unfortunately, this was not possible as the observation was made after the results were returned by the Fox group, and this processing could not be performed in house. However, for illustrative purposes, a crude quantitation has been attempted. For this, the image of the gel was expanded, and boxes drawn around the bands of uncleaved DNA remaining in each 50 μM lane. The height of each box was then compared to that of the uncleaved control lane giving an estimation of the percentage of DNA cleaved by each analogue. The results are presented as a bar chart in figure 3.13.

The results of this analysis suggest some interesting differences in alkylation efficiency. However, to confirm these differences as real, it would be necessary to conduct several replicates, and ideally to perform more rigorous kinetic based alkylation assays. Nevertheless, the observed differences will be discussed.

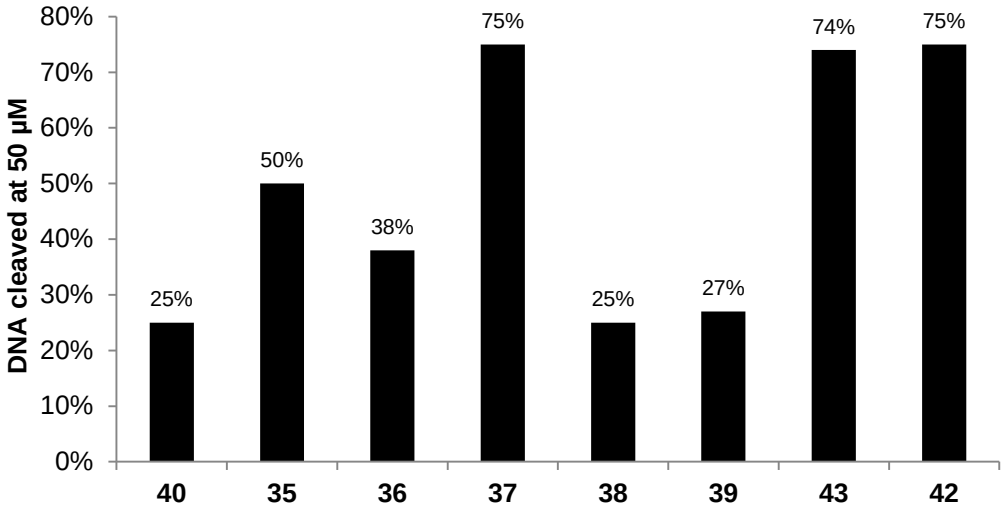
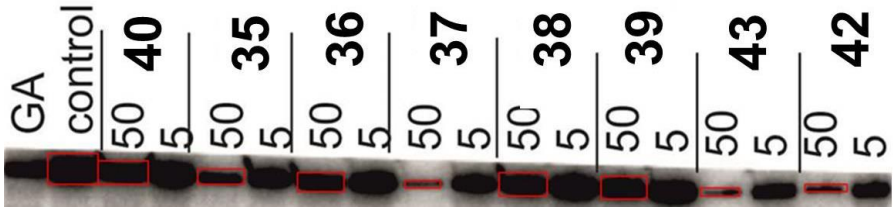


Figure 3.13 Bar chart of the percentage of DNA cleaved relative to the control lane at 50 µM compound incubation concentration. Estimation of uncleaved DNA band height achieved using the box method (see below image). Boxes were drawn using Microsoft PowerPoint, and heights measured using the 'autoshape' size function.



As the cleavage induced by the carboxylic acid control **43**, appears to match that of the ester **42**, the detrimental impact of the amino acids is unlikely to be a consequence of the negatively charged carboxylic acid group. Thus the feared potential electrostatic repulsion with the polyanionic backbone of DNA appears not to be the dominant problem. Therefore, the amino acid structures themselves might be inhibiting the alkylation event.

As discussed, the nature of the side chain may have the potential to mitigate the negative effect of the *C*-terminal amino acid on DNA alkylation. This is most apparent in the serine analogue **37**, which appears to be the only amino acid analogue to have cleaved the same percentage of DNA as the ester control **42**. The serine analogue was also one of the best performing of the amino acid analogues in the MTS assay, breaking the observed trend between hydrophobicity and cellular activity, returning a lower IC_{50} than the alanine

analogue. It is possible that the hydroxyl group is offering an advantage towards DNA alkylation. This will be discussed in more detail later.

The results indicate that the remaining *N*-acetyl amino acid-duocarmycin conjugates all alkylate DNA with reduced efficiency compared to the ester and acid controls, **42**, and **43**. This might be the result of effects on non-covalent minor groove binding. The detrimental impact could result from a number of potential effects, and would likely be the result of a combination of contributing factors.

The most simple explanation may be that the amino acid structure disrupts the planar nature of the compounds and this reduces minor groove binding affinity. In such cases the potential reversal of the alkylation event may be more noticeable, thus reducing statistically the quantity of available compound alkylated to the DNA at any one point in time. It is also possible that dissociation from the hydrophobic environment of the minor groove to the more polar solvent environment may be further favoured by the increased hydrophilic nature of the compounds, particularly for the lysine and glutamic acid analogues which contain a second charge on their side chains. This would seem consistent with a recent report by Boger *et al.* that describes a series of increasingly water soluble duocarmycin SA derivatives. These analogues contained varying degrees of PEG substitution at the methoxy groups of the trimethoxyindole subunit. A convincing linear relationship between hydrophilicity and reduced DNA alkylation efficiency was demonstrated.[190]

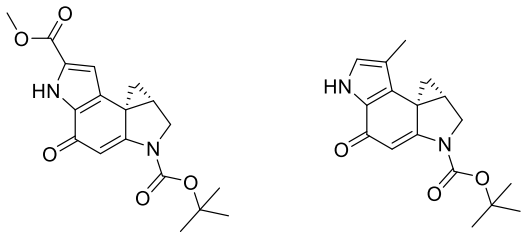


Figure 3.14 Structure of N-Boc-DSA, and N-Boc-CPI.

The *C*-terminal amino acid structures may also impose steric demands which might decrease alkylation efficiency. It has been reported that a simple truncated Boc derivative of the alkylation subunit of CC-1065 (*N*-Boc-CPI) shows reduced DNA alkylation efficiency when compared to that of duocarmycin (*N*-Boc-DSA). This has been attributed to a steric effect of the extra indole methyl group of the CPI subunit (see figure 3.14).[189] It is possible that the amino acid structures sterically inhibit the depth of minor grove penetration, or perhaps in some cases, can shield the alkylation event itself.

There would certainly appear to be a tentative relationship between increased side chain size and reduced alkylation activity. This is apparent in the observation that the alanine analogue appears to alkylate DNA more efficiently then the phenylalanine analogue, which in turn is better than the lysine and glutamic acid analogues. However the serine and β-alanine analogues both break this trend. The serine analogue has a larger side chain than alanine, but is as good as the ester control **42** at alkylating DNA. Again this could be the result of a positive effect specific to the presence of the hydroxyl group, and as mentioned previously, this will be discussed later. The β-alanine analogue has no side chain and yet performs as badly as the lysine and glutamic acid analogues. This is consistent with its apparent reduced activity in the MTS assay when compared to the alanine analogue. Clearly the size of the side chain could not be the only factor.

The apparent reduced alkylation efficiency of the β-alanine analogue compared to the alanine analogue may be a consequence of the greater flexibility of the linear structure of β-alanine, compared to the branched structure of alanine. For example this increased flexibility may incur a greater entropic cost to binding to the minor groove. This would also be true for the flexible side chains of the lysine and glutamic acid analogues. The greater entropic cost might be in part countered by the entropically favourable displacement of ordered water molecules known to be associated with the minor groove.[191, 192] However, given the comparable size of the β-alanine and alanine analogues, the increased entropic penalty could be significant.

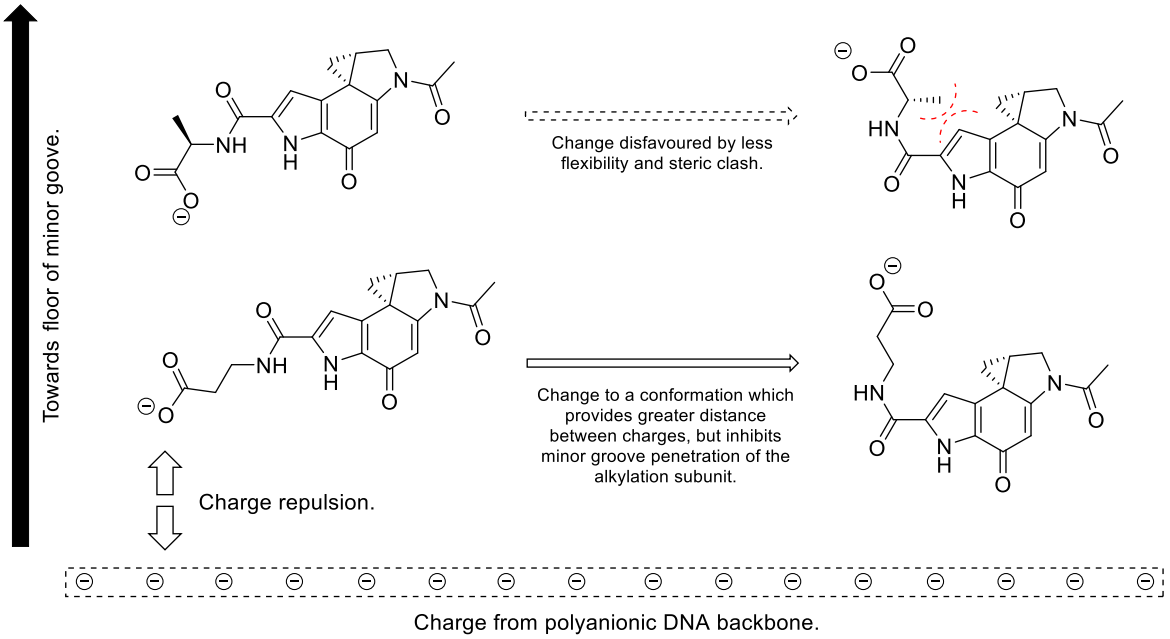


Figure 3.15 Schematic representation of a potential reason for the observed reduced alkylation efficiency between the alanine and β-alanine analogues.

Another potential negative consequence of the greater flexibility of the β-alanine structure can also be envisioned. In order to produce the greatest distance between the negatively charged *C*-terminal carboxyl group and the polyanionic phosphodiester backbone of DNA, the linear aliphatic chain of β-alanine, may bend to position the carboxyl group towards the inside floor of the minor groove. As a result this may produce a greater steric constraint on minor groove penetration of the alkylation subunit, than is produced by the branched less flexible structure of alanine (see figure 3.15). In such a configuration the β-alanine analogue appears to share more spatial homology with the glutamic acid analogue than the alanine analogue, and this is consistent with the observed alkylation activities.

As discussed, it is clear that the presence of the *C*-terminal amino acid structures are not altering the sequence selectivity of the alkylation event. However, it is possible that other non-covalent binding positions might be favoured thus reducing alkylation efficiency. For example, it was originally expected that the lysine analogue would be one of the best performing analogues, as it was hoped that the positively charged side chain might increase non-covalent binding affinity, by forming a favourable electrostatic interaction with the polyanionic phosphodiester backbone. This does not appear to be the case, and suggests that when the alkylation event occurs, the side chain of lysine may not be projecting out of the minor groove to associate with the phosphodiester backbone, but may lay across the floor of the minor groove, or be positioned towards the interior of the minor groove. Therefore it is possible that the lysine analogue prefers to bind with its side chain projecting out of the minor groove, but that in this binding orientation, the cyclopropane is no longer in a favourable position to alkylate the N3 of adenine. Similar such alternative binding orientations may exist for the other analogues.

It is most likely that the true reason for the reduced alkylation efficiency of the truncated amino acid-ducarmycin analogues (if real) is a complicated combination of different interplaying factors, which might include some of the ideas discussed above. The reduced activity also seen in the MTS assay, is likely a further more complicated interplay between the factors contributing to reduced DNA alkylation efficiency, coupled with factors contributing to reduced cell permeability.

The lysine extended analogue **44** containing the *N*-terminal methoxyindole subunit was also tested in the thermal cleavage assay. Consistent with the results of the MTS assay, the extended structure performed considerable better than the truncated *N*-acetyl analogues. Figure 3.16a, shows the DNA cleavage gel of **44**, over a 100 fold concentration range (1 μM to 0.01 μM). Run in comparison on the same gel were the *N*-

acetyl *C*-terminal ester and acid compounds **42**, and **43**, at the previous 5 μM and 50 μM concentrations.

The first observation is the far superior DNA alkylation efficiency. Alkylation was observable down to at least the 0.01 μM concentration. The increased alkylation activity was further highlighted by altering the incubation temperature. The gel in figure 3.16a, shows cleavage after incubation at 37 ^{o}C as before. However, figure 3.16b, shows the cleavage gel of all the compounds when incubated at a concentration of 10 μM at only 25 ^{o}C. Under these conditions, only the extended analogue **44**, shows any observable alkylation bands.

These results are entirely consistent with the reduced reactivity in thermal cleavage assays of truncated duocarmycin analogues such as N-Boc-DSA compared to the full natural product, which has be shown to alkylate DNA at lower incubation temperatures.[189] The trimethoxyindole subunit of the natural product, and the methoxyindole subunit of **44** likely improve DNA alkylation in the same manner. This is generally accepted to occur *via* two processes. Firstly, the methoxyindole subunit improves hydrophobic non-covalent minor groove binding affinity. Secondly, the increased linear length induces greater *in situ* activation of the cyclopropane, as described in chapter one.

It is likely that the improved activity of **44** compared with its truncated counterpart **39** in the MTS assay, is the result of both improved DNA alkylation efficiency, and superior cell permeability; both as a consequence of the methoxyindole subunit. However, as **44** clearly out performs the truncated *C*-terminal ester compound **42** in the thermal cleavage assay, but returned a higher IC_{50} in the MTS assay, this compound is likely to still suffer from some reduction in cellular uptake, when compared to the *C*-terminal ester.

In addition to the improved alkylation activity, **44** also displays greater discrimination in alkylation sites when compared to the truncated compounds. This change in sequence selectivity is again consistent with what would be expected from the extended structure which is more similar to the full natural product.[189] The sequence selectivity of alkylation was quantified by measuring the intensity of the radioactive signal of each band. The results have been plotted as a proportion of the total cleaved DNA, against the sequence of the MS1 fragment in figure 3.17. Filled bars corresponded to cleavage by **44**, and open bars to that of **42**.

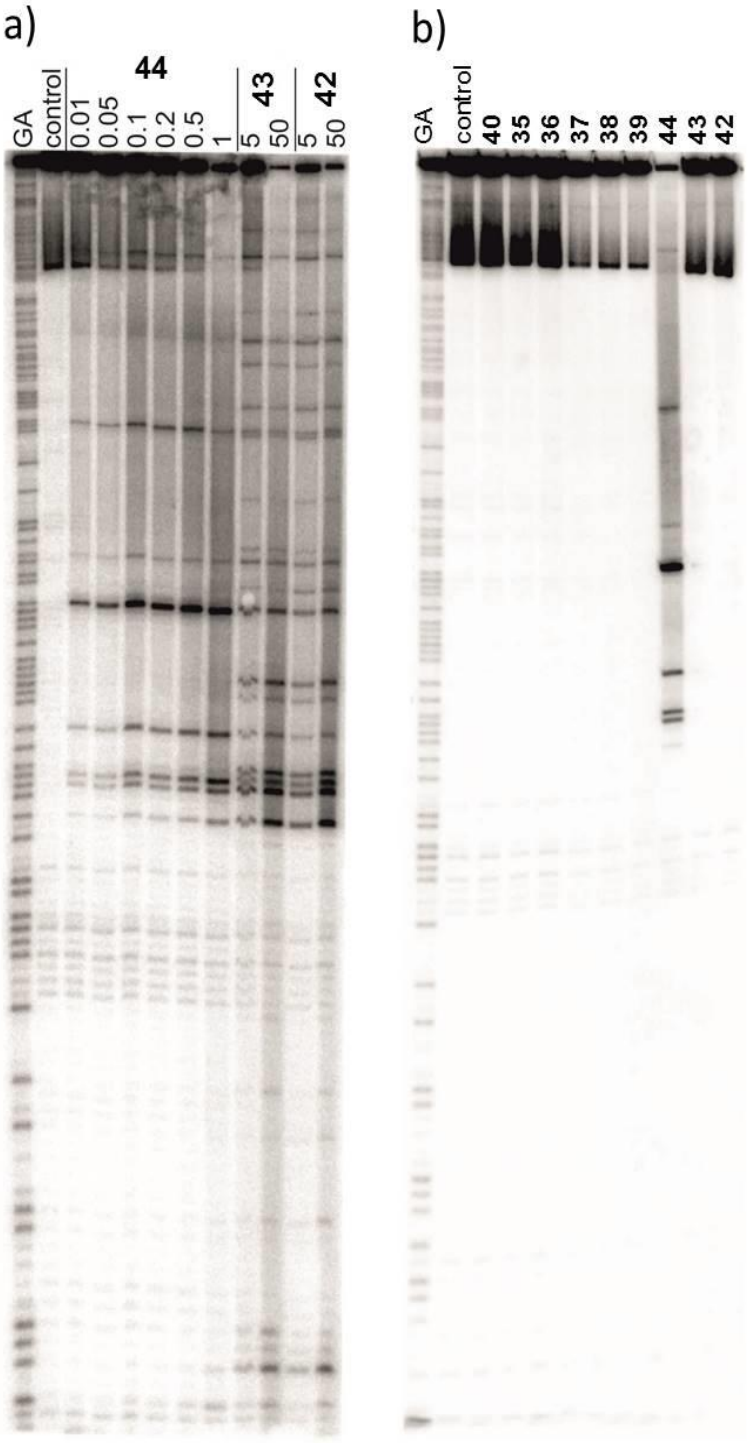


Figure 3.16 (a) DNA thermal cleavage gel of **44**, **42**, and **43** at various concentrations (μM). Incubation at 37 ^{o}C. (b) DNA thermal cleavage gel of all the analogues at 10 μM. Incubation temperature 25 ^{o}C.

As discussed, while **42** demonstrates little selection other than the requirement for a 5' A or T base neighbouring the alkylated A, **44** exhibits a clear preference for two sites within the MS1 sequence; one corresponding to 40 % of the total alkylation and the other to 26 %. The remaining 36 % of the total alkylation observed is spread relatively equally over six additional minor sites, but does range from 1 % to 13 %.

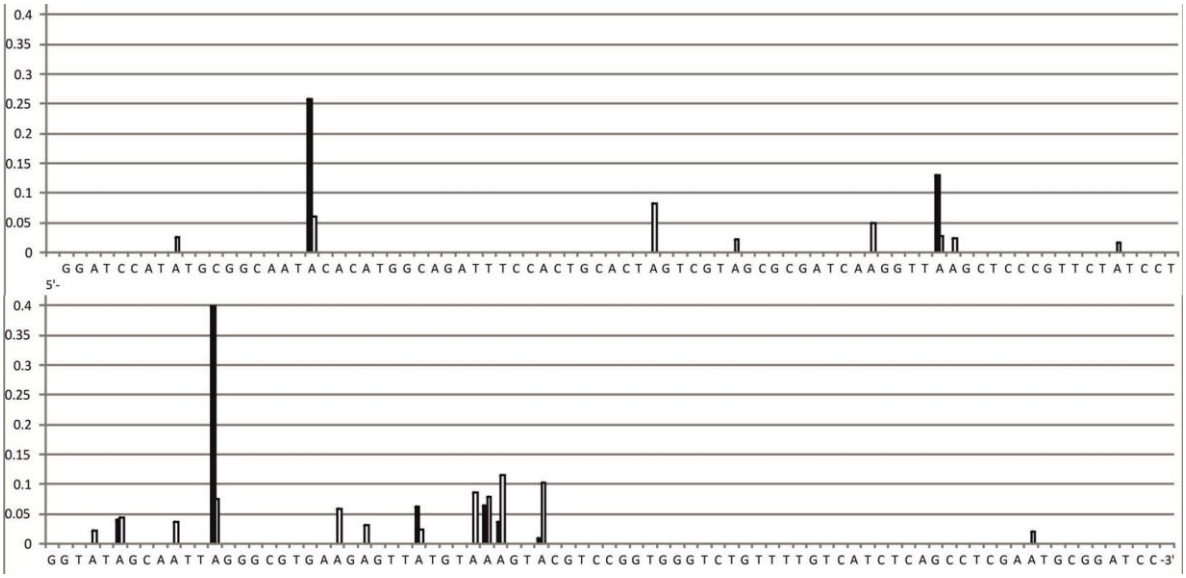


Figure 3.17 Intensity of each cleavage band, plotted as a proportion of total cleavage against the sequence of the MS1 DNA fragment. Open bars = **42**. Filled bars = **44**.

With the exception of the anomalous 1 % binding site, there would appear to be an absolute requirement for at least two 5' A or T bases neighbouring the alkylated A. There is also a clear preference for longer AT rich sequences. The most favoured alkylation site (40 %) represents the longest AT rich run, comprising of four flanking 5' A or T bases (5'-AATTA). In addition the second most favoured alkylation site (26 %) contains three 5' A or T bases (5'-AATA). It would also appear that a flanking 5' T base is favoured over A. This is apparent in the observation that in only two out of the eight sites alkylated by **44**, did the first neighbouring 5' base constitute an A. Furthermore, in the 13 % alkylation site, **44** has alkylated the A neighbouring the T base, when a longer 5' AT rich sequence could have been achieved by alkylating the furthest 3' A of the tetranucleotide AT rich sequence (5'-TTAA). Additionally, the only alkylation site not to contain more than one 5' flanking A or T bases (the 1% site), happens to consist of a TA step (5'-TA).

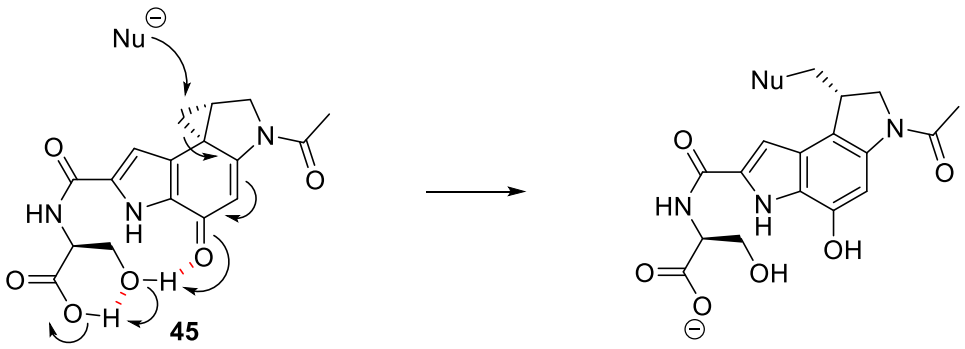
These results are broadly in agreement with the sequence selectivity of duocarmycin SA, where in a different DNA fragment the major alkylation site was also found to constitute 5'-AATTA.[189] In line with duocarmycin SA the preference for longer AT rich sequences is most likely a consequence of the increased length of **44**, compared to the truncated *N*-acetyl analogues. Here longer sequences are required to benefit from the increased

hydrophobic binding affinity provided by the narrower groove topology of AT rich regions. In addition longer G free sequences are required to prevent the steric inhibition of groove penetration, by this base's minor groove occupying amine.

3.13 Potential explanation for the possible superior activity of the serine analogue 37.

The apparent increased activity of the serine analogue **37** (if real) is interesting. It is the only truncated amino acid analogue to have exhibited alkylation activity comparable to the truncated ester control **42**. In addition, it was one of the better performing analogues in the MTS assay, returning a lower IC_{50} than the alanine analogue, despite possessing a more polar side chain, thus appearing to break the tentative trend between hydrophilicity and reduced activity. These results (if real) would suggest that the presence of the hydroxyl group of the serine side-chain may be providing a favourable effect on the efficiency of DNA alkylation.

This observation led to the consideration that the hydroxyl group may be increasing the electrophilicity of the cyclopropane by participating in an intramolecular hydrogen bond chain between the carbonyl of the cyclohexadienone system, and the *C*-terminal carboxylic acid, as outlined in scheme 3.14.



Scheme 3.14 Potential hydrogen bonding promoting the alkylation by spirocylised form of **37**, (**45**).

If this theory is correct, it would be predicted that the spirocylised form of **37** (**45**), would be more reactive towards solvolysis at neutral pH than the spirocylised form of the alanine analogue, **35**. The reactivity of these two species could be compared by monitoring the solvolysis reaction by UV spectroscopy. In order to conduct this experiment the seco-compounds would need to be converted into their spirocylised counterparts. Unfortunately, preliminary attempts to spirocyclise **37** were unsuccessful. The reaction

proved difficult to monitor, and the product could not be isolated in sufficient purity to confirm its structure.

To investigate the plausibility of the idea, Density Functional Theory (DFT) calculations were employed. In chemistry, DFT is a powerful computational method of studying the energies of molecules. In the most basic terms, DFT can be considered as a method for producing approximate solutions to the Schrodinger equation for multielectron systems. In practical terms, DFT calculations can be run quite simply by chemists using commercially available software, and so provides a method to determine possible conformations of molecules, and to model reaction pathways.[193]

The feasibility for **45** to adopt a conformation consistent with the proposed hydrogen bonded structure, was examined by gas-phase DFT calculations employing the B3LYP functional[194-196] and 6-31G(d) basis set[197] implemented by Gaussian09.[193] Structure optimisation calculations identified five significant conformations of **45** (figure 3.18a-e). All were confirmed by frequency calculations to represent true energy minima. Frequency calculations can be used to distinguish a true minimum from a transition state. Of the five optimised structures, 'conformation a' would appear to be consistent with the proposed hydrogen bonding. This is apparent from the short interatomic separations between the oxygen atom of the carboxylic acid and the serine side-chain oxygen (2.75 Å), and from there to the oxygen atom in the cyclohexadienone ring (3.75 Å). In addition, a slight lengthening of the C=O bond of the cyclohexadienone was observed when compared to an alternative minimum 'conformation b', in which hydrogen bonding was not present (1.24 Å vs 1.23 Å).

It is clear that in the gas-phase, 'conformation a' is not the global minimum. However the energy differences between the identified conformations are small [most stable conformer 'conformation b' is only 6.1 kcal mol^{-1} (25.5 kJ mol^{-1}) more stable], and this would suggest that 'conformation a' is realistically accessible. Furthermore, it is possible that in the environment of the minor groove, the order of the relative energies might change, and this could be in favour of 'conformation a'.

Having identified 'conformation a', it was decided to study the solvolytic opening of the cyclopropane using methanol as a model nucleophile. A transition state (see figure 3.19) consistent with ring opening was identified, and confirmed by a frequency calculation that indicated one imaginary (i.e. negative) vibrational frequency. Once a transition state has been identified, calculations termed the IRC (intrinsic reaction coordinate), and reverse IRC can be used to determine the reaction pathway. The reaction pathway shown in figure 3.19 has a calculated activation energy of 53.9 kcal mol^{-1} (225.6 kJ mol^{-1}), determined

from the energy difference between the transition state and the sum of the energies of the individually optimised reactants (**45** in 'conformation a', and methanol). The methanol adduct at the end of the calculated reaction path initially seemed only slightly lower in energy (1.4 kcal mol^{-1}, 5.8 kJ mol^{-1}) than the transition state itself, but a further optimisation of the structure gave a new conformer (see figure 3.19) that was 7.3 kcal $mole^{-1}$ (30.7 kJ mol^{-1}) more stable than the transition state. Finally, loss of the proton from the $MeO(H)^{+}CH_2$- section of the structure ultimately forms a carboxylate which was also confirmed by a DFT calculation and is also illustrated in figure 3.19. Interestingly, the hydrogen bonding involving the serine OH is retained in this final structure, together with additional hydrogen bonding to the indole NH.

These results suggest that it is possible for **45** to adopt a conformation that allows hydrogen bonding between the hydroxyl and the cyclohexadienone. A reaction pathway for ring opening of the cyclopropane consistent with activation by the hydroxyl group has been identified, and offers support to the plausibility of this theory. It would be interesting to compare the activation energy of an alternative pathway calculated from a transition sate which does not involve the hydrogen bonding conformation of **45**. However, so far IRC calculations have not been successful for transition states of different conformations preventing this comparison. This does not mean they do not exist but that the correct transition state for an alternative pathway has not yet been identified.

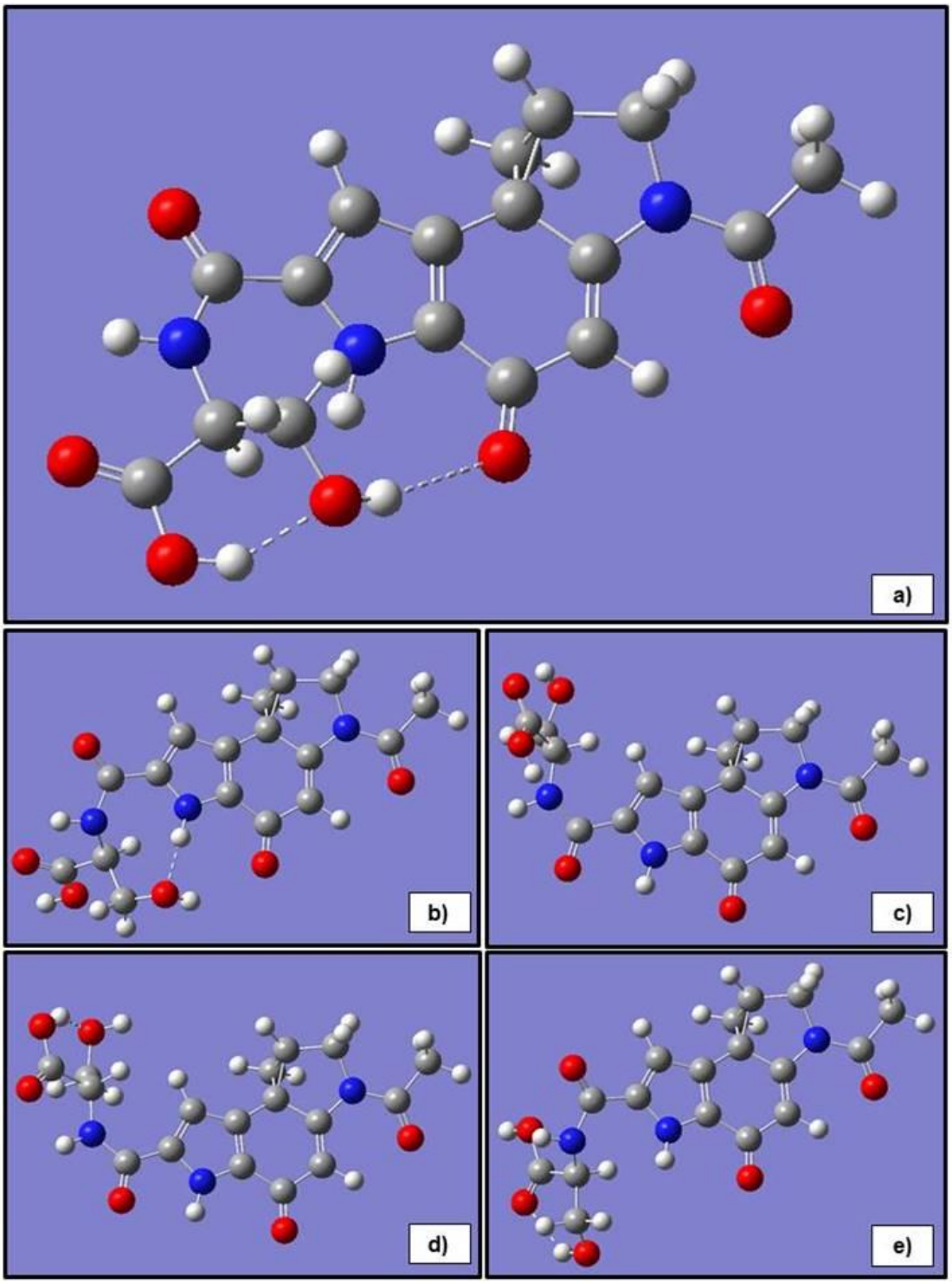


Figure 3.18 Conformers of **45** from structure optimisation calculations. (a) Conformation consistent with proposed hydrogen bonding. (b) Confirmation 6.1 kcal mol^{-1} (25.5 kJ mol^{-1}) more stable than conformation a. (c) Confirmation 2.8 kcal mol^{-1} (11.7 kJ mol^{-1}) more stable than conformation a. (d) Confirmation 1.9 kcal mol^{-1} (8.0 kJ mol^{-1}) more stable than conformation a. (e) Confirmation 0.8 kcal mol^{-1} (3.2 kJ mol^{-1}) more stable than conformation a.

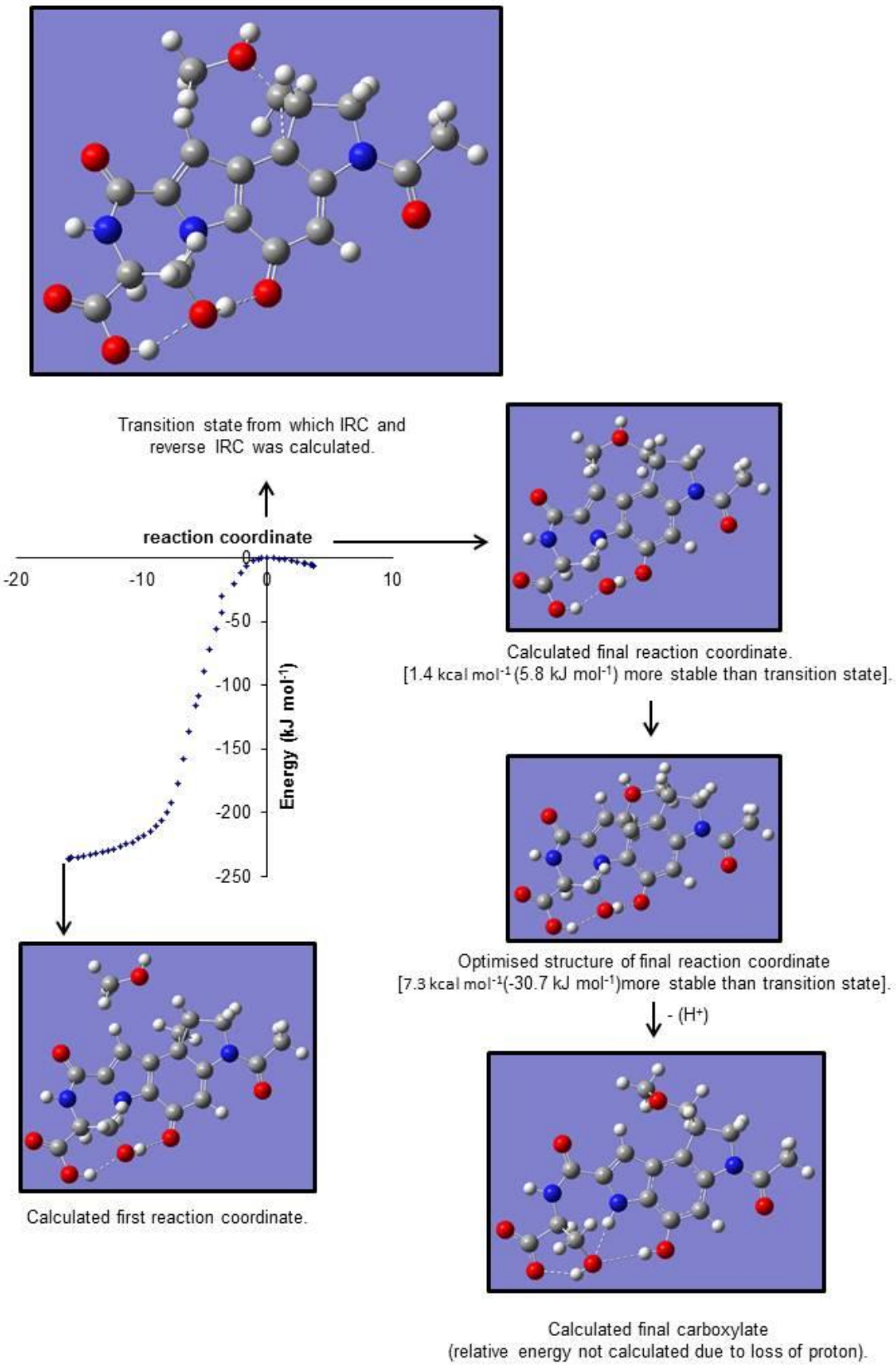
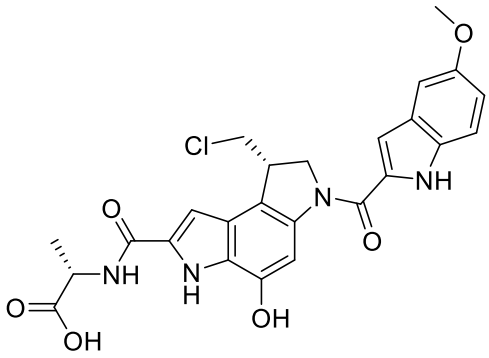


Figure 3.19 Calculated reaction pathway for methanolysis and key structures.

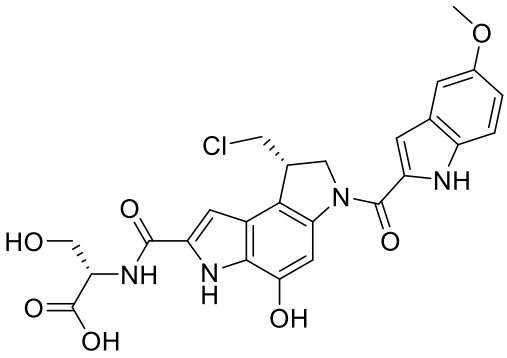
3.14 Synthesis and antiproliferative activity of the serine and alanine extended amino acid-duocarmycin conjugates.

Given the interesting suspected increased activity of the truncated serine analogue when compared to the truncated alanine analogue, the decision was made to use the solid phase methodology to synthesise the extended methoxyindole counterparts of these two compounds, **46** and **47** (figure 3.20). The synthesis of each analogue proceeded as previously described on a 0.038 mM scale. No issues were experienced during benzyl deprotection, which had been troublesome during the synthesis of the extended lysine analogue. The desired structures, **46** and **47** were confirmed by a combination of accurate mass spectrometry, ^{1}H NMR and DEPT-edited-HSQC experiment.



7 mg, 35 % yield

46



3.3 mg, 16 % yield

47

Figure 3.20 Structure of the extended alanine and serine analogues, **46**, and **47**.

Antiproliferative activity was assessed by MTS assay as before in the HL-60 cell line, returning a mean IC_{50} value of 38 nM (95 % CI: 23-62 nM, triplicate), and 153 nM (95 % CI: 75-312 nM, triplicate), for **46** and **47** respectively. Consistent with the results from the lysine extended analogue **44**, inclusion of the methoxyindole unit increased the antiproliferative activity of both **46**, and **47** by over 1000 fold when compared to their respective truncated counterparts. Interestingly, in their extended forms the serine analogue no longer exhibits superior activity to the alanine analogue, and in this instance the confidence intervals suggest the result may be significant. However, this is not inconsistent with the theory put forward in section 3.12 to explain the apparent increased activity of the truncated serine analogue (although the confidence intervals overlapped in this case). Inclusion of the methoxyindole unit means activation of the cyclopropane can now be expected to be controlled by more efficient disruption of the stabilising vinylogous amide upon binding to the minor groove (see chapter one). Thus, any activating potential

of the hydroxyl group of the serine side-chain may be inconsequential. In fact, it could even be detrimental, by increasing the reactivity of the cyclopropane to solvolytic opening, or reaction with other biological nucleophiles before reaching its DNA target.

With the theorised advantage of the serine side-chain abated by the presence of the methoxyindole unit, a more consistent trend between both increased hydrophilicity and side-chain length and reduced activity is now observed for the extended analogues (lysine extended analogue **44** 374 nM [95 % CI: 200-701 nM, triplicate], serine extended analogue **47** 153 nM [95 % CI: 75-312 nM, triplicate], alanine extended analogue **46** 38 nM [95 % CI: 23-62 nM, triplicate]). Although, the difference between the lysine and serine extended analogues may not be significant.

Although the possible increased activity of the truncated serine analogue is interesting, it may be of limited practical value, as it does not appear to produce a synergistic effect with the methoxyindole unit, on activation.

3.15 Masking of the *C*-terminal carboxylic acid.

As discussed, although inclusion of the methoxyindole unit dramatically improved antiproliferative activity, and showed that low nM activity was achievable with a *C*-terminal amino acid, this is still over 1000 fold less active than duocarmycin SA. Given the complete lack of activity of **43**, this was still suspected to be a consequence of a negative effect on cell permeability, imposed by the carboxylic acid. In an attempt to confirm this, two additional versions of the alanine extended analogue were synthesised varying the nature of the *C*-terminal end (figure 3.21).

48 contained a neutral end. This was achieved by conducting the synthesis on Rink amide MHBA resin, which affords a terminal amide after cleavage. In **49**, the carboxylic acid was masked by a group which would be positively charged at physiological pH. This was achieved by activating the cleaved carboxylate with HATU and DIPEA, followed by treatment with an excess of 3-(dimethylamino)-1-propylamine. The desired structures were confirmed by a combination of accurate mass spectrometry, ^{1}H NMR and DEPT-edited-HSQC experiment.

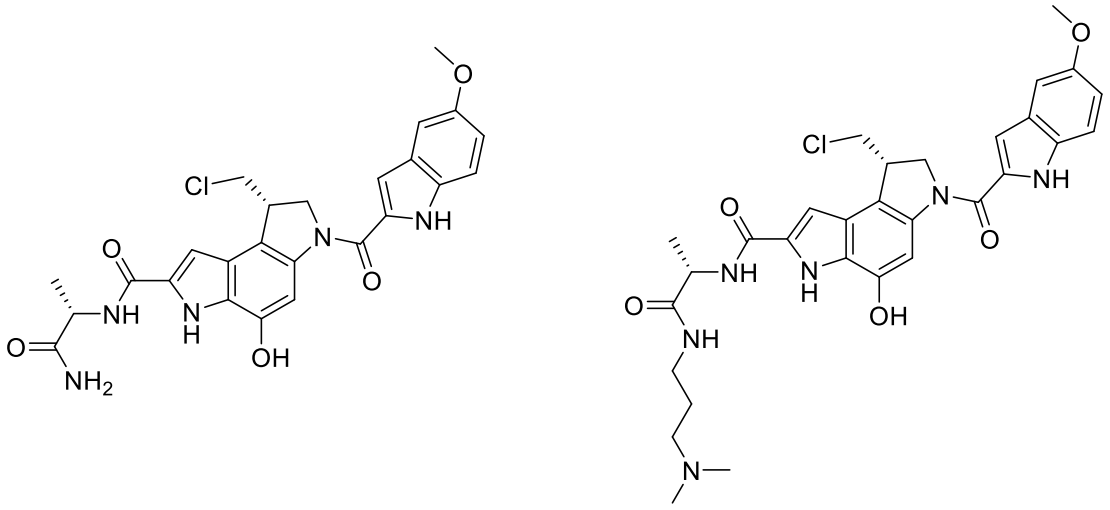


Figure 3.21 Structure of **48** and **49**.

Antiproliferative activity was assessed by MTS assay as before in the HL-60 cell line. Inclusion of the dimethylamine tail appeared to be detrimental, producing roughly a twofold drop in activity over the free carboxylic acid (**49** mean IC_{50} 64 nM [95 % CI: 36-117 nM, triplicate]), although confidence intervals overlap so this may not be a significant difference. This compound is of course still charged, which is likely to inhibit cellular uptake. However, it was expected that the positive charge of the amine would be more beneficial than the negative charge of the carboxylic acid. It is possible, as suggested for the lysine analogue, that electrostatic attraction to the polyanionic backbone of DNA, may promote non-covalent minor groove binding in an orientation which does not place the cyclopropane in the most favourable position for alkylation.

The neutral amide analogue **48** showed only a marginal improvement in IC_{50} (mean 28 nM [95 % CI: 12-68 nM, triplicate]) compared to the free carboxylic acid **46**. However again the confidence intervals overlap suggesting the difference may not be significant. Indeed, this compound was expected to have low pM activity comparable to the natural product; yet the observed activity is only comparable to that of the truncated ester control **42.** It would therefore appear that for the extended analogues at least, it is not the *C*-terminal carboxylic acid which is producing the largest detrimental effect on antiproliferative activity. The thermal cleavage assays of the truncated analogues, provided tentative evidence suggesting that the amino acid structures might have the potential to reduce the efficiency of DNA alkylation. However, given that the extended lysine analogue **44** was more efficient than the truncated ester control **42** at alkylating DNA, it seems that the

reduced antiproliferative activity of the neutral amide extended analogue **48**, is likely to still be a product of reduced cellular uptake. Although, it is not obvious what properties of the neutral non-polar *C*-terminal structure of **48** could be affecting cell permeability. It might be simply the increased size when compared to the truncated ester analogue **42**. Thus the improved uptake of **42** promoted by its reduced size, might allow comparable antiproliferative activity with **48**, despite it expected reduced alkylation efficiency.

However, this reasoning is unlikely to explain the reduced antiproliferative activity of **48**, when compared to the natural products. Duocarmycin SA is only marginally smaller than **48**, and yatakemycin and CC-1065 are considerably larger. It is possible that cell permeability of the extended amino acid analogues is actually comparable to the natural products, and that the differences in activity here, represent negative impacts on alkylation efficiency produced by the amino acid structures. The extended lysine analogue was more efficient at alkylating DNA than the truncated ester control **42**, but it could still be less efficient than the full natural products.

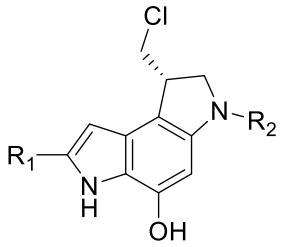
The above would seem the most likely explanation for the observed results, although it is possible to envision other contributing factors. It may be that the presence of the *C*-terminal amino acids can affect the distribution of the compounds once inside the cell. For example, recognition of the amino acid structures could lead to partial participation in amino acid homeostasis pathways, potentially resulting in efflux from the cell, or compartmentalisation to organelles other than the nucleus. If such processes took place it would reduce amount of the compounds available to alkylate DNA, reducing antiproliferative activity.

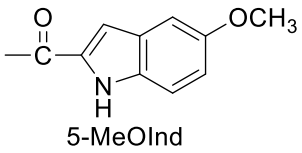
Whatever the true reason, it is clear that inclusion of *C*-terminal amino acids has an impact on the antiproliferative activity of the duocarmycin alkylation subunit both in truncated and extended forms. However, low nM activity is still possible. The ionisable *C*-terminal carboxylic acid, may be a significant negative contributor for truncated analogues, but appears not to be significant when considering the reduced activity of extended analogues possessing the methoxyindole subunit.

3.16 Conclusions from chapter three.

It has been demonstrated that **11**, serves as a suitable agent, for incorporation of the duocarmycin SA alkylation subunit into polyamide structures *via* Fmoc based solid phase synthesis. Large excesses of this reagent are not necessary for short sequences when coupling to amino acids, but could be employed if difficulties were experienced for more

complicated targets. The results indicate that **11** should be considered as a sensitive residue, and complications analogous to those reported for synthesises incorporating tryptophan can be experienced. As such, the choice of solid support, and cleavage conditions are not trivial. 2-Chlorotrityl resin appears to be superior to Wang resin, and also allows greater flexibility in cleavage conditions. Minimising the concentration of TFA used to affect cleavage is advantageous, and appropriate scavengers should be included. It is possible that **11** could be improved by replacement of the benzyl ether with a *para*-methoxybenzyl ether, allowing simultaneous phenol deprotection during cleavage from the resin. Boc-protection of the indole nitrogen might also be beneficial, reducing the sensitivity of the residue during cleavage.





	R_1	R_2	IC_{50} (μM)	[95 % CI]
35	CO-Ala-OH	Ac	85	[45 - 165]
40	CO-β-Ala-OH	Ac	>300	
36	CO-Phe-OH	Ac	32	[13-78]
37	CO-Ser-OH	Ac	40	[12-115]
38	CO-Glu-OH	Ac	286	[101-814]
39	CO-Lys-OH	Ac	>300	
44	CO-Lys-OH	5-MeOInd	0.374	[0.200-0.701]
47	CO-Ser-OH	5-MeOInd	0.153	[0.075-0.312]
46	CO-Ala-OH	5-MeOInd	0.038	[0.023-0.062]
49	CO-Ala-$NH(CH_2)_2N(CH_3)_2$	5-MeOInd	0.064	[0.036-0.117]
48	CO-Ala-NH_2	5-MeOInd	0.028	[0.012-0.068]
42*	$COOCH_3$	Ac	0.025	[0.011-0.037]
43*	COOH	Ac	>300	

* racemic

Table 3.2 Summary table of IC_{50} valves for all analogues discussed during chapter three.

It is clear that the presence of a *C*-terminal amino acid is detrimental to the antiproliferative activity of the duocarmycin alkylation subunit, when compared to the natural products. This appears to be due to a combination of poor cellular uptake, and disruption of minor groove binding affinity. However, low nM activity is possible by inclusion of the methoxyindole subunit. The activity of such extended analogues cannot be further improved by masking of the carboxylic acid by a terminal amide, or a

dimethylamine group. The nature of the side-chain may have an effect of activity, with short non polar groups appearing to be broadly favoured. Allthough this may not be significant in some cases (see confindence intervals table 3.2). In the truncated analogue **37** the hydroxyl group of the serine side-chain might be beneficial, and this may be due to specific activation of the cyclopropane through intramolecular hydrogen bonding. However, this has not been proven beyond doubt. Furthermore, although interesting, this effect has little practical value, as it does not appear to provide a synergistic benefit with addition of the methoxyindole unit. Thus more rigous exploration of this potential effect may not be of valve.

These results would indicate that **11** could be a useful building block for the solid phase synthesis of peptide based linkers towards the development of novel antibody drug conjugates. Low nM activity could be sufficient; however it may also be possible to recover the low pM activity exhibited by the natural products through incorporation of multiple alkylation subunits. Furthermore, it could also be argued that if the reduced activity of the amino acid duocarmycin conjugates is the result of reduced cell uptake, then this could in fact be seen as advantageous for antibody drug conjugates; serving as a protective mechanism against off-target cytotoxicity should premature cleavage from the antibody prior to internalisation occur. There is clearly some potential for the *C*-terminal amino acid to affect activity. For example, although most of the minor side chain effects have not been proven to be significant in this study, there are some tentative trends. Furthermore, some of the differences are more striking, (for example the complete inactivity of the β-alanine analogue **40** at the tested concentrations). This at least suggests that when designing peptide based linkers for novel antibody drug conjugates of duocarmycin, the nature of the *C*-terminal amino acid my not be trivial, and this warrants further investigation.

4

Chapter Four

*Initial investigations towards a second application of **11**: The solid phase synthesis of duocarmycin- distamycin hybrid compounds, highlighting the potential for direct incorporation of the duocarmycin alkylation subunit into sequence selective hairpin polyamides.*

4.1 Aims.

The work described in this chapter was carried out in order to address the following aims:

- To begin preliminary investigations to explore the utility of **11** to serve as a 'building block' for the solid phase incorporation of the duocarmycin alkylation subunit in to sequence selective minor grove binding polyamide structures.

Such compounds could have the potential to target the cytotoxicity of the duocarmycins to gene mutations specific to malignant cells.

4.2 Introduction to distamycin A and the hairpin polyamides.

The work reported in chapter three, clearly demonstrates that **11** can serve as a building block for the direct solid phase synthesis of polyamide structures incorporating the alkylation subunit of the duocarmycin SA. This work focused on coupling **11** to natural amino acids, with the view to the future development of novel antibody drug conjugates through peptide based linkages. Such an approach could improve the therapeutic index of the duocarmycins by harnessing the exquisite cellular targeting power of antibodies, perhaps removing the necessity for prodrug strategies. Another approach to targeting of the duocarmycin alkylation subunit could be to extend its structure in such a way as to refine the sequence selectivity of minor groove recognition. For example, if it were possible to bind selectively to a gene mutation specific to a malignant cell line, this could facilitate the targeting of cytotoxicity.

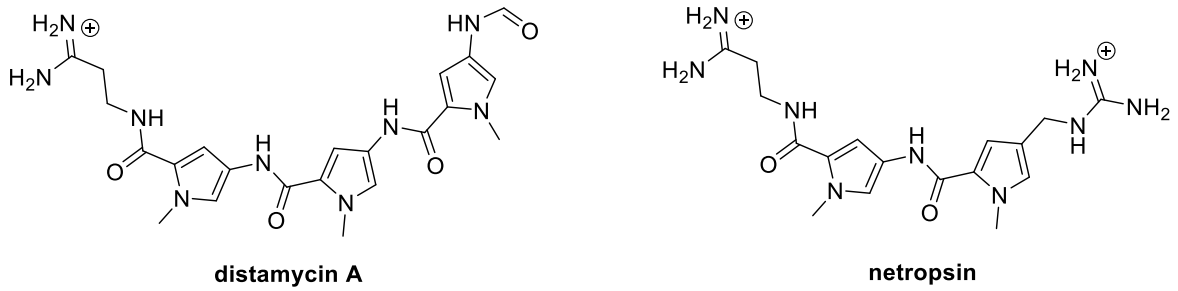
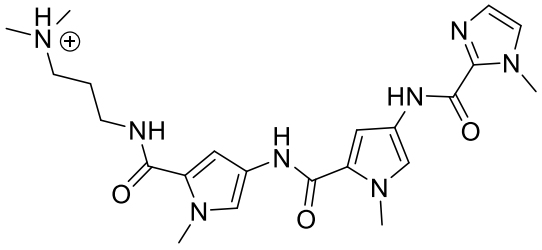


Figure 4.1 Structure of distamycin A and netropsin.

One of the most successful synthetic approaches towards sequence selective minor groove recognition has been the development of the hairpin polyamides. This advance can be traced back to biological evaluation and synthetic exploration of the natural products distamycin A[198] and netropsin[199] (figure 4.1). These natural lexitropsins are heterocyclic polyamides, comprising of multiple *N*-methylpyrrole residues (Py), and bind in the minor groove of DNA, favouring AT rich sequences.[200] Like the non-covalent binding of duocarmycin SA, the narrower groove topology of AT rich sequences is favoured due to stronger hydrophobic binding interactions between the planar aromatic structures of the lexitropsins and the walls of the minor groove. The crescent shape of these molecules also complements the curvature of the helical groove. Furthermore it was shown that the NH groups of the polyamide chains form favourable hydrogen bonding interactions with the N3 of A and O2 of T, while the amino group of G imposes steric inhibition of this close association.[201] The complexes are further stabilised by electrostatic interactions between

the cationic amidinium and guanidinium tails and the polyanionic phosphate backbone of DNA.[202]

The observed hydrogen bonding led to the independent proposals by both the Dickerson and Lown groups that incorporation of 1-methylimidazole (Im), in place of Py residues may permit binding to GC sequences.[203, 204] This was based on the rationale that the additional imidazole nitrogen would both alleviate the steric clash between the Py and the amino group of G, and potentially form a favourable hydrogen bond with it. Indeed this was shown to be the case for Im containing analogues of both distamycin A and netropsin. Selectivity for GC sequences was greatest for distamycin A analogues, and this was attributed to less nonspecific electrostatic affinity, resulting from possessing only a single amidinium tail, as opposed to the two cationic tails present in netropsin derivatives.[204-206] This discovery essential produced a preliminary binary code for sequence recognition, with Py selecting for A·T or T·A base pairs, and Im selecting for G·C or C·G base pairs.



50

Figure 4.2 Structure of **50**.

Up until this point binding of the lexitropsins had only been observed as a 1:1 complex with DNA. However, the report by Wemmer *et al.* that distamycin A could also bind in the minor groove as an antiparallel dimer,[207] would prove to mark a turning point in the development of sequence recognition *via* polyamide structures. At the time of this discovery, the Dervan group had also been working on Im variants of distamycin related analogues. Their novel polyamide **50** containing an ImPyPy chain (figure 4.2) had been designed based on a 1:1 binding model to select for the sequence 5'-XZZ-3' (where X = G or C, and Z = A or T). However, this proved not to be the case, and instead **50** was observed to bind to the 5 base sequence 5'-ZGZCZ-3' (Z = A or T).[208] Dervan *et al*. suggested that **50** may also be binding as an antiparallel 2:1 dimer analogous to the alternative binding model described for distamycin A. As such, the heterocycles of each chain could be envisioned to form cofacial pairs, with the polyamide edge of each facing towards the floor of the minor groove.

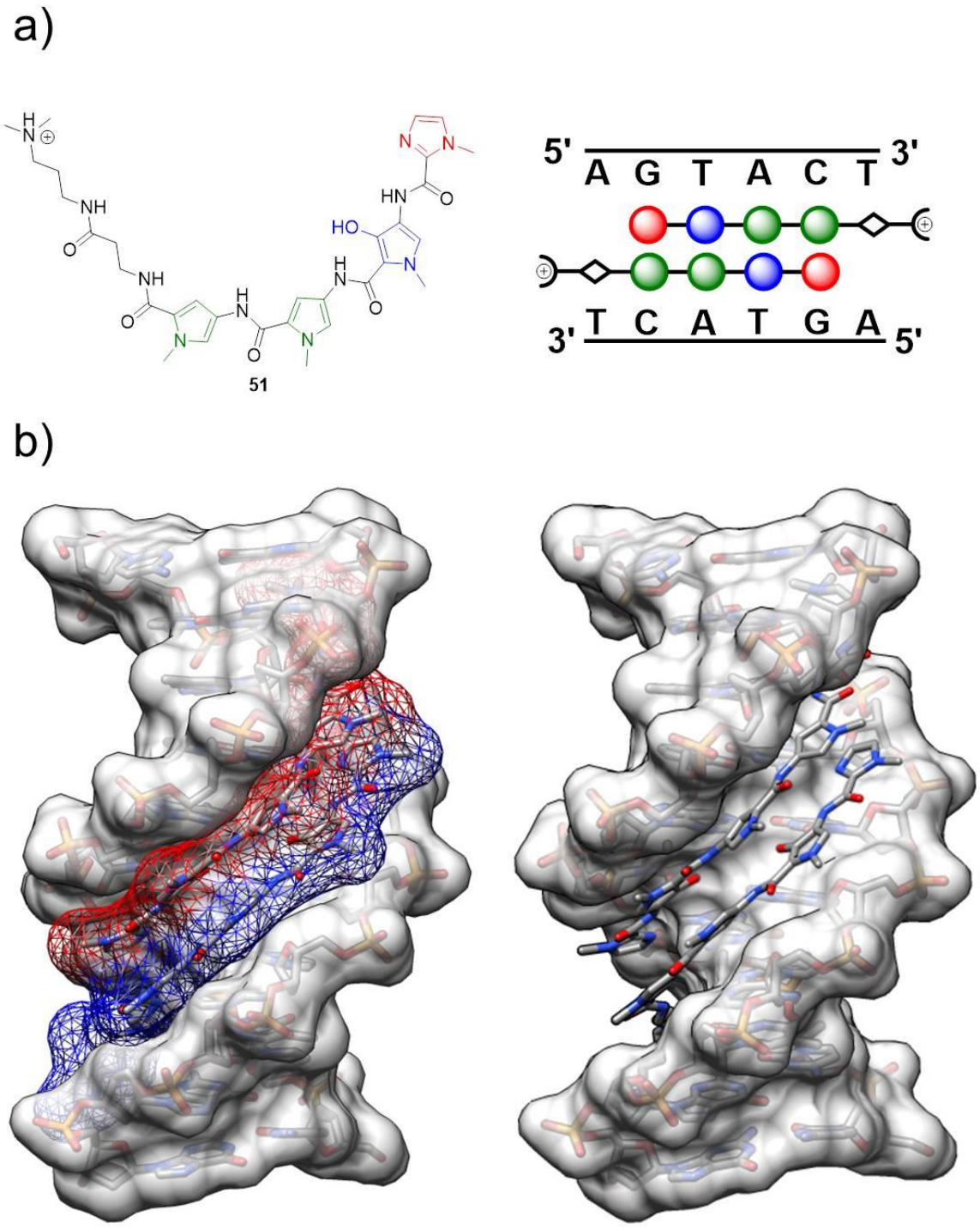


Figure 4.3 (a) Left: Colour coded structure of **51**. Py = Green, Hp = Blue, Im = Red. Right: schematic representation of minor binding of **51** depicting pairing rules. (b) Crystal structures of 2:1 complex of **51** and DNA. Left: with ligand surface (red surface upper polyamide, blue surface lower polyamide). Right: Crystal structures of 2:1 complex of **51** and DNA without ligand surface. Image generated using 'Chimera 1.10.1', coordinates from protein data bank (ID:407d).[211]

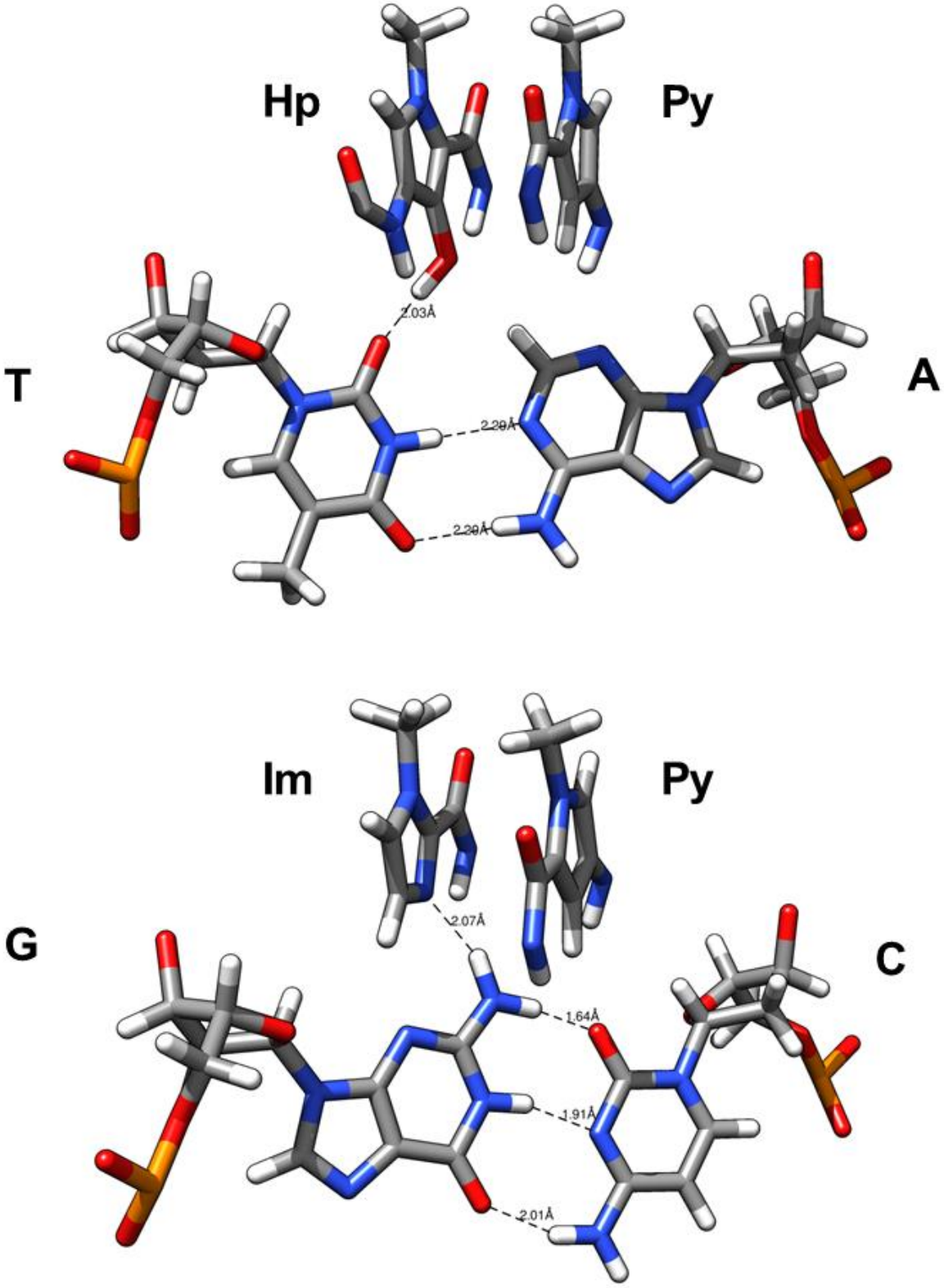


Figure 4.4 Expansion of the crystal of 2:1 complex of **51** and DNA, highlighting the physical basis of Dervan's pairing rules. Top: Hp·Py selecting for a T·A base pair. Hydroxyl group sits in cleft and hydrogen bonds to O2 of T. Bottom: Im·Py selecting for G·C base pair. Im permitted by G, and hydrogen bonds to amine of G. Image generated using 'Chimera 1.10.1', coordinates from protein data bank (ID:407d)[211] Hydrogens added by 'Chimera 1.10.1.

In this way, a chain partnership of Im•Py selected for a G•C base pair base, and the opposite pairing of Py•Im selected for a C•G base pair, while Py•Py was degenerate selecting for both an A•T and T•A base pair. Collaboration with Wemmer *et al.* would confirm this suggestion by solving the NMR structure for a 2:1 dimer of **50** with DNA.[209, 210]

The recognition code would later be completed by the design of a third heterocyclic residue, *N*-methyl-3-hydroxypyrrole (Hp).[211] Here, Hp•Py partners select for T•A base pairs, and Py•Hp partners select for A•T base pairs. The rationale for this specificity was both steric and electronic. Here the hydroxyl group of Hp produces a steric constraint to binding towards the A side of the T•A base pair, while the cleft on the T side can accommodate this group. This selective association is further promoted by a favourable hydrogen bond between the hydroxyl group and the O2 of T. The physical basis for selectivity would later be confirmed from the crystal structure of an antiparallel 2:1 dimer of the novel polyamide **51**, which contain the polyamide sequence PyPyHpIm, with the oligonucleotide sequence 5'-CC**AGTACT**GG-3' (figure 4.3, and figure 4.4).[212]

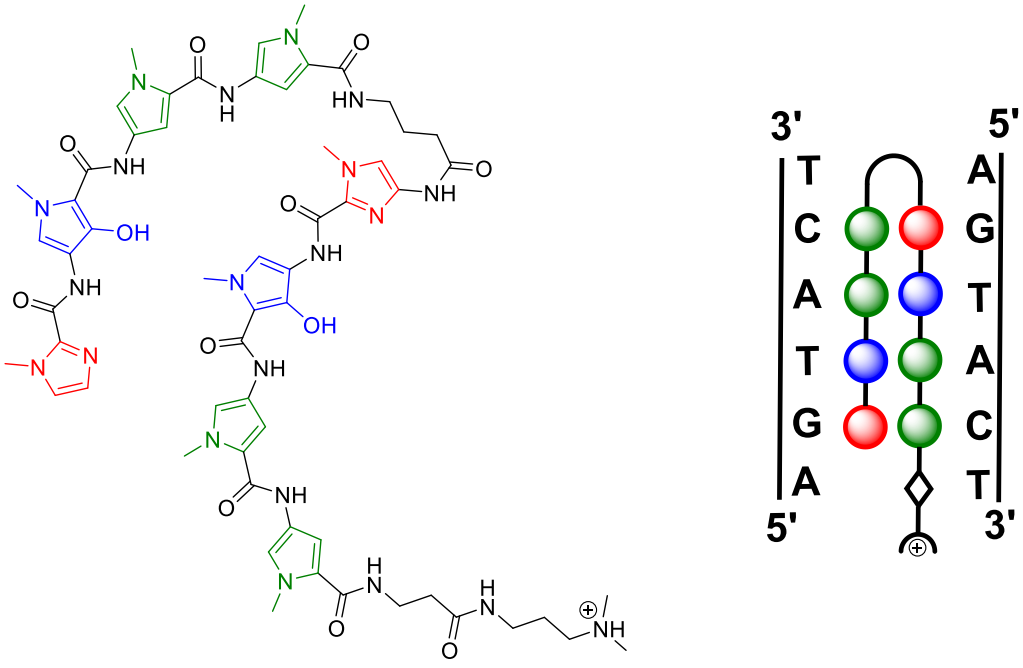


Figure 4.5 Left: Colour coded structure of a Hairpin polyamide version of **51**. Right: Schematic representation of minor groove binding.

The Dervan group was quick to explore the covalent linkage of antiparallel dimers. Several strategies have been employed including bridging the nitrogen atoms of central Py resides with aliphatic linkers, the so called H-pin approach,[213] and the linking of both terminals to produce cyclic compounds.[214] However, owning to the ease of synthesis, by far the most popular linkage has been the U-pin strategy.[215] Here a γ-aminobutyric acid

residue links the *N*-terminus of one chain with the *C*-terminus of the other. This linking motif was found to be optimal for promoting the desired turn, and antiparallel binding. Such compounds are usually referred to as 'hairpin polyamides' (figure 4.5).[215]

These hairpin structures typically exhibit improved binding specificity and affinity when compared to the equivalent unpaired dimers. Covalent linkage prevents competing off target 1:1 binding of unlinked polyamide chains, enforcing the pairing rules. Binding affinity is improved despite the loss of one of the cationic tails, and this most likely represents the decreased entropic cost of binding.[215]

One of the major limitations of the simple hairpins has been the length of sequence it is possible to target. The natural arc of the polyamide chains is too curved and thus is not completely complementary to the helical turn of the minor groove. As a result, binding affinity and specificity begins to decreases when the polyamide exceeds five heterocyclic pairs.[216, 217]

Several strategies have been employed to compensate for this. For example the inclusion of internal aliphatic β-alanine residues have been used to relax the curvature of the polyamide chains, allowing the successful targeting of longer sequences as antiparallel dimers.[218] Another approach has evolved from the exploration of substitution at the α-position of the γ-aminobutyric acid turn residue. This led to the development of tandem hairpin polyamides linked both turn to tail, and turn to turn.[219, 220] Such compounds have been designed and shown to recognise up to 10 base pair sequences, with high affinity (figure 4.6).[220, 221]

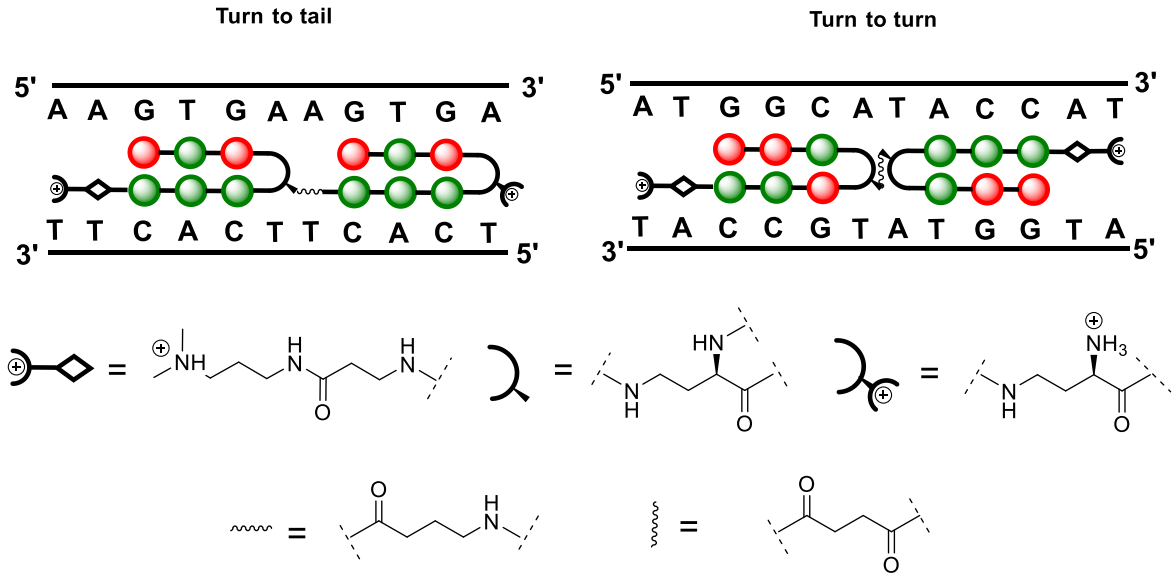


Figure 4.6 Schematic representation of the tandem hairpin approaches.

The steric influence of chiral amino substituents at the α-position of the γ-aminobutyric acid turn residue, has also been shown to further promote the desired 5'-3' binding orientation, as well as inhibit competing unfolded minor groove binding.[222, 223] Furthermore, β-alanine/heterocycle pairings have seen success in targeting difficult sequences, allowing more flexibility in the orientation of the selecting ring structure.[224]

A truly generalisable set of rules for sequence selective minor groove recognition by hairpin polyamides is still an active area of research. For example, an alternative T selective residue to replace the Hp heterocycle is desirable. Hp has proven to be more context dependent than the Im residue, and also less stable. Furthermore, Hp containing polyamides often exhibit reduced binding affinity.[225] The optimisation of long sequence recognition, cell permeability and nuclear uptake are also important challenges.

However, despite the remaining issues, simple hairpin polyamides have already seen success in the sequence selective inhibition of DNA-Protein interactions. Both suppression and activation of gene transcription through the rational design of hairpin polyamides has been demonstrated. One of the first examples was the competitive displacement of the TFIIIA transcription factor, which was shown to suppress related gene transcription both *in vitro* and in cellular assays.[226] More recently, a hairpin polyamide designed to target the difficult 5'-CGCG-3' sequence, demonstrated low nM allosteric inhibition of a major groove binding methyltransferase *in vitro*.[227]

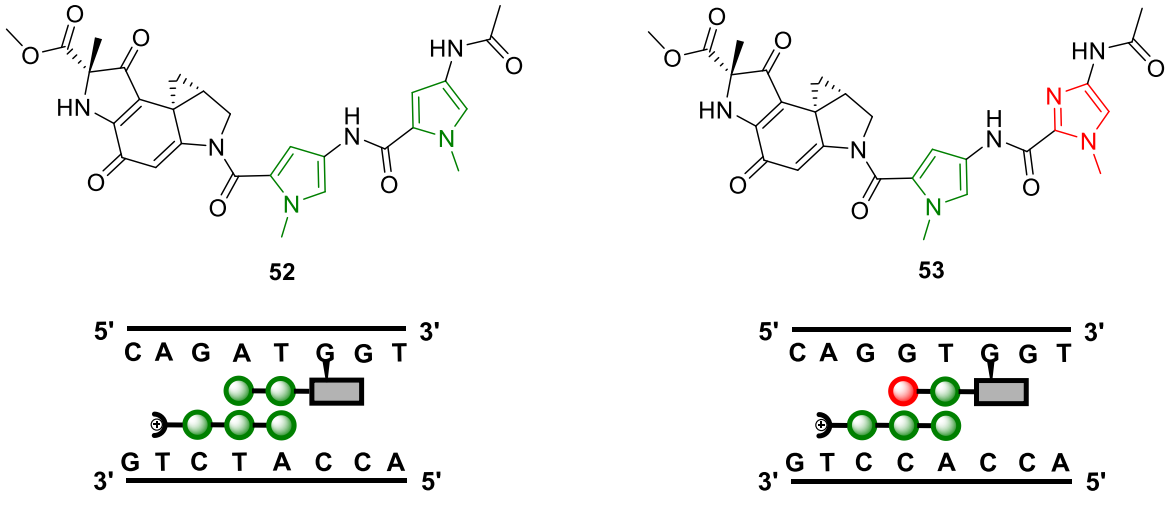


Figure 4.7 Structures of **52** and **53** with schematic representations of the cooperative minor groove binding of each with distamycin A.

Combining the sequence selective minor groove recognition of the hairpin polyamides, with the alkylation activity of the duocarmycins, is not a novel concept. The core structures of both families exhibit planar aromatic ring systems, affording tight hydrophobic

interactions with the walls of the minor groove, and thus position one edge of the respective ligands towards the protruding face of the internal base pairs. This homology in binding mode was quickly recognised to have the potential to produce interesting hybrid compounds.

In 1996, Sugiyama *et al.* discovered that duocarmycin A could be induced to alkylate the G of GC rich sequences, *via* the co-incubation of distamycin A. They demonstrated that this was the result of cooperative minor groove binding through the formation of a cofacial heterodimer.[228] In response to this finding, Sugiyama *et al.* began to explore replacement of the indole subunit of duocarmycin A with short polyamide sequences. Conjugates **52**, and **53**, were shown to selectively alkylate DNA in the presence of distamycin A, in a way which was consistent with the Im/Py pairing rules. Here, **52** which contained the PyPy sequence, preferentially alkylated the 3' G of a 5'-AGATG-3' site, while **53**, containing the ImPy sequence, alkylated the 3' G of a 5'-AGGTG-3' site (see figure 4.7).[229]

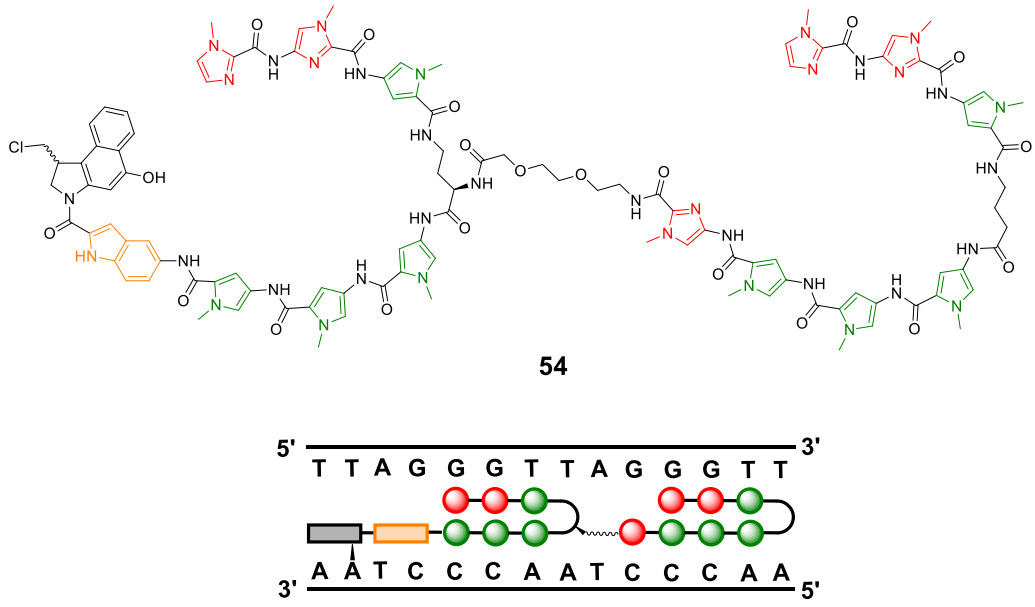
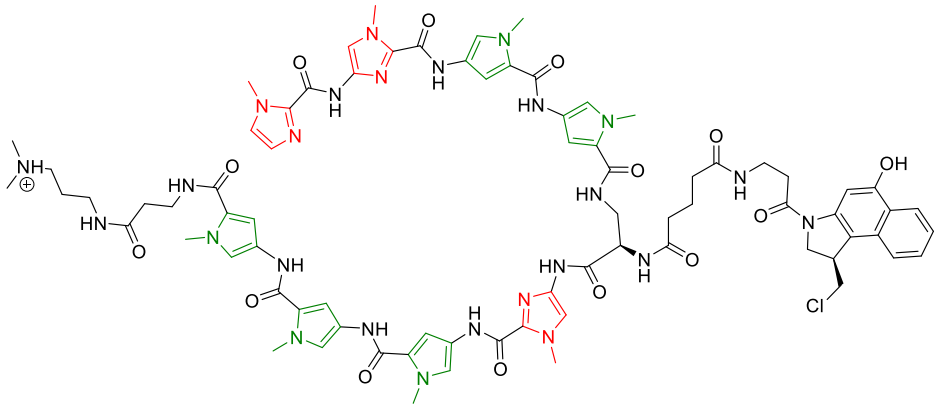


Figure 4.8 Structure of **54** and a schematic representation of binding to the human telomere repeat sequence sequence.

Ultimately, these results would prompt the Sugiyama group to explore the conjugation of various duocarmycin related alkylation subunits to the *C*-terminal tail of hairpin polyamide structures through the indoline nitrogen. This has included the selective alkylation of a nine base sequence (5'-ACAAATCCA-3'), *via* the use of an internal β-alanine/β-alanine pairing to relax the curvature of the polyamide chains.[230] Interestingly, they report increased alkylation efficiencies when duocarmycin subunits are spaced from the polyamide *C*-terminals by vinyl, or indole linkers.[230, 231] Molecular dynamic simulations

suggested that in the absence of such a linker, the cyclopropane of these compounds is oriented in a suboptimal position between two base pairs. This could reflect the fact, that in this approach the duocarmycin subunit is orientated in an unnatural 5' to 3' direction. More recently, Sugiyama *et al.* have demonstrated the impressive, and apparent specific alkylation of the human telomere repeat sequence through the synthesis of the tandem hairpin seco-CBI conjugate **54** (figure 4.8).[232]



55

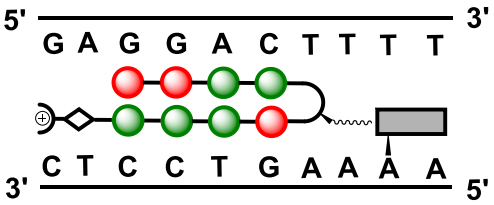


Figure 4.9 Structure of **55** and schematic representation of binding.

An alternative approach has been explored by Dervan *et al.* (figure 4.9). Here, the seco-CBI alkylation subunit is linked through the turn residue of the hairpin structure by an aliphatic linker. Such compounds have been shown to selectivity alkylate A bases neighbouring the preferred binding site of the hairpin structure. It was also observed that the chirality of the CBI subunit could affect binding preference. For example, conjugate **55**, containing the R enantiomer of seco-CBI, was shown to bind a matched site containing a flanking 5'-TTT-3' sequence, alkylating the third A of the complementary strand (figure 4.9). The same compound exhibited no alkylation of a second matched binding site which contained a flanking 5'-ATA-3' sequence. Alkylation of the sense stand of this second matched site could be induced by incorporation of the *S* enantiomer of seco-CBI.[233]

Being oligomeric in nature, the hairpin polyamides are ideally suited to construction through solid phase synthesis. Indeed exploration of this class of molecule has been greatly accelerated by its implementation. The Dervan group were first to employ solid

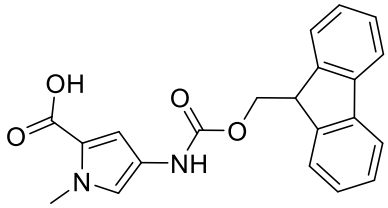
phase synthesis to this effect.[234] They have exploited both Boc- and Fmoc-protection strategies. In the most common approach, the core polyamide structure is synthesised from a solid supported β-alanine residue.[235] Cleavage from both Wang and PAM resins is affected by aminolysis, allowing simultaneous introduction of the dimethylamine tail. The *C*-terminal β-alanine residue, has been shown to be AT selective. However, the use of Kaiser oxime resin to introduce shorter GC selective tails has also been explored.[236]

Despite the extensive application of solid phase synthesis to access the core structures of hairpin polyamide targets, conjugation of duocarmycin related subunits has to date been restricted to post cleavage modifications. In the case of the Sugiyama approach this strategy is obligatory due to the requirement to reveal the *C*-terminal acid before amide coupling to the indoline nitrogen of the duocarmycin species can be undertaken. The Dervan group also couple CBI *via* the indoline nitrogen, and achieve this following release and isolation of the precursory polyamide structure from the solid support. In this case CBI is linked at the turn residue. It is therefore possible to envision an orthogonal protecting strategy that would allow this coupling to be implemented prior to cleavage, perhaps *via* a diacid linker. However, this would complicate the methodology, disrupting the operational simplicity of sequential *N*-terminal amide couplings, and thus require reconfiguration of the automated synthesis routinely used by this group.

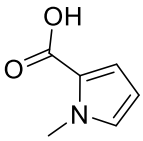
Given the continued interest and exciting potential of bifunctional hairpin polyamides, our new solid phase 'building block' **11**, would seem ideally placed to provide a substrate for the direct and inline solid phase incorporation of the duocarmycin alkylation subunit, potentially at any position in the chain sequence. As the alkylation subunit of duocarmycin SA is known to be AT selective, it may be possible to replace a PyPy position without effecting sequence selectivity of the parent polyamide. However the rigid length of a duocarmycin alkylation subunit, is marginally shorter than that of a PyPy sequence. It remains to be seen whether this will prove beneficial for mid sequence incorporation by relaxing the overall curvature of the chain structures, or deleteriously by negatively affecting minor groove recognition.

4.3 Synthesis of the Fmoc-protected pyrrole amino acid monomer 56 and unsubstituted pyrrole acid 57.

In order to conduct preliminary investigations towards the direct solid phase incorporation of **11** into hairpin polyamide sequences, access to the desired Fmoc-protected pyrrole amino acid **56** and unsubstituted pyrrole acid **57** was required (figure 4.10).



56



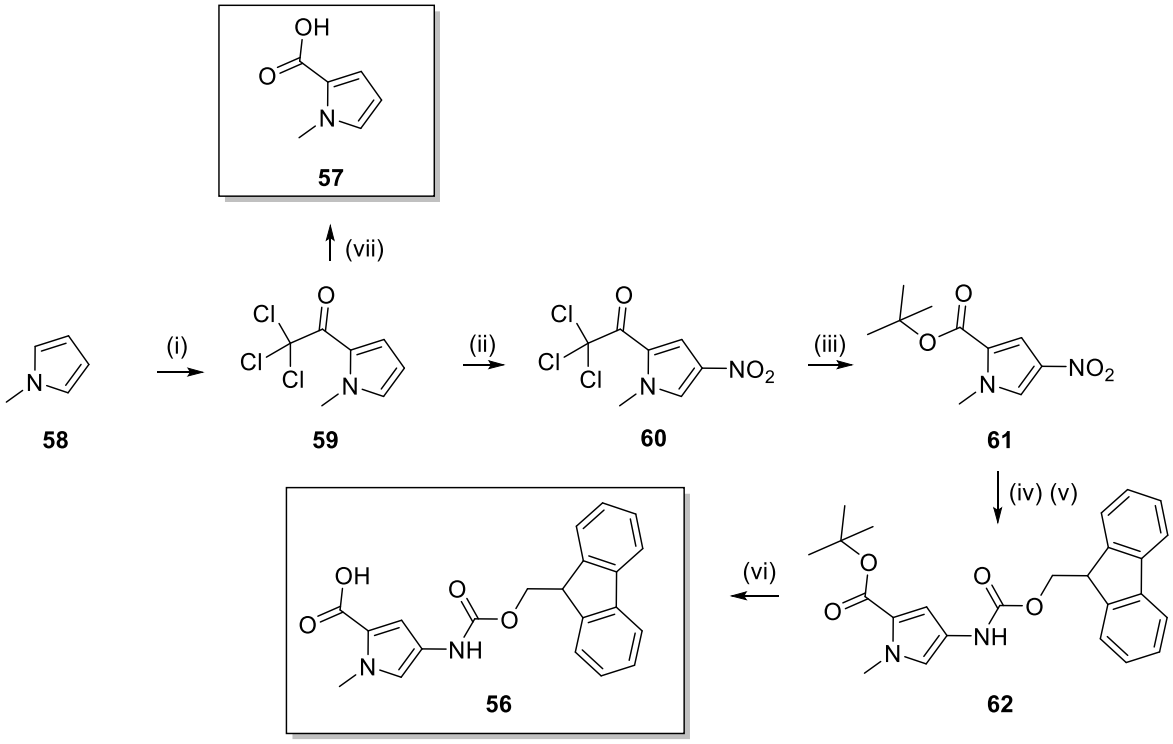
57

Figure 4.10 Structure of **56** and **57**.

Both building blocks were obtained through synthesis from commercially available *N*-methylpyrrole as previously described by Dervan *et al.* (scheme 4.1).[235] Minor changes from the published procedures were implemented when necessary, which will be highlighted in the following discussion.

Synthesis of **56** began by the direct acylation of *N*-methylpyrrole with trichloroacetyl chloride affording the acyl substituted pyrrole **59**. Subsequent nitration with fuming HNO_3 and catalytic H_2SO_4 in Ac_2O provided the nitro pyrrole **60**. Conversion of the trichloroacetyl substituent to a *t*-butyl ester was achieved by treatment with NaO*t*Bu in refluxing *t*-BuOH. Interestingly, initial attempts to substitute KO*t*Bu for NaO*t*Bu, due to reagent availability, resulted in no reaction. The apparent necessity for the Na^+ counterion over K^+ is surprising. This presumably represents a steric effect resulting from the increased atomic radius of K^+.

In the published procedure, subsequent nitro group reduction is achieved by treatment with Pd/C in DMF under a 500 psi positive pressure of hydrogen. This is followed by removal of the catalyst by filtration and immediate protection of the amine by addition of Fmoc-Cl and DIPEA. Attempts to repeat this step using a balloon of hydrogen proved unsuccessful. The reduction was slow and could not be driven to completion. Treatment of the partially reduced **61** did not yield the desired product. It was suspected that the prolonged reaction times (days) resulted in degradation of the amine.



Scheme 4.1 Synthesis of **56** and **57**. (i) trichloroacetyl chloride, DCM, 0 ^{o}C. (ii) HNO_3, H_2SO_4, AcO_2, - 40 ^{o}C. (iii) NaOtBu, HOtBu, reflux. (iv) 10 % Pd/C, Ammonium formate 25 % w/v aq, THF. (v) Fmoc-Cl, DCM 0 ^{o}C. (vi) $TiCl_4$, DCM, 0 ^{o}C. (vii) NaOH, THF:Water.

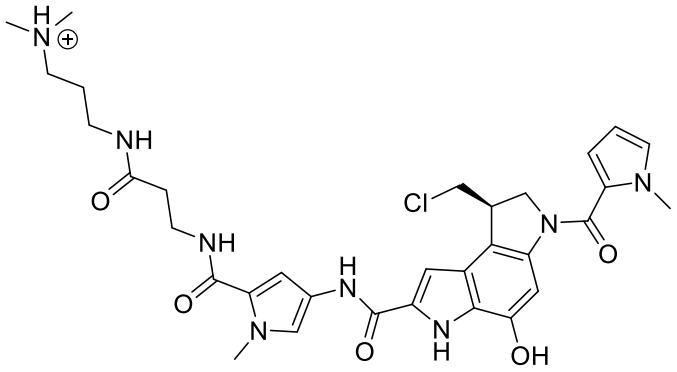
As a result ammonium formate was utilised as an alternative hydrogen source. The reduction was conducted in THF, and driven to completion within 4 hours by repeated addition of Pd/C as a slurry in aqueous ammonium formate every 30 mins. On consumption of the starting material as monitored by TLC, the reaction mixture was filtered through a pad of diatomaceous earth, directly into a stirring solution of Fmoc-Cl in DCM. Subsequent aqueous work up and column chromatography afforded the desired carbamate **62** in an acceptable yield of 84 %.

Selective hydrolysis of the *t*-butyl ester to afford the desired carboxylic acid was affected as per the published procedure by treatment with $TiCl_4$. The previous authors report completing decarboxylation if this step is attempted by treatment with TFA.

The unsubstituted pyrrole acid **57** was accessed *via* direct hydrolysis of the common trichloroacetyl pyrrole intermediate **59**, through treatment with aqueous NaOH in THF.

4.4 Initial investigations towards the solid phase synthesis of distamycin-duocarmycin hybrid compounds.

To explore the use of **11** as a building block for the solid phase synthesis of distamycin-duocarmycin hybrid structures, a simple target compound was envisioned (**63**, figure 4.11). This consisted of the duocarmycin alkylation subunit sandwiched between two pyrrole residues. The *C*-terminal of the target compound would possess the β-alanine-dimethylaminopropylamine tail commonly employed by the Dervan group in the synthesis of hairpin polyamides. The *N*-terminal pyrrole residue would be unsubstituted removing the need for final Fmoc deprotection and acetylation steps. In order to conserve supplies of the more valuable natural enantiomer of **11** these studies would be conducted using its unnatural enantiomer.



63

Figure 4.11 Structure of **63**.

4.4.1 First attempt at the solid phase synthesis of 63 using 2-ClTrt resin and aminolytic cleavage.

In the initial synthetic strategy, cleavage from the resin by aminolysis, as employed by the Dervan group, was planned to allow simultaneous introduction of the dimethylamino tail. The synthesis was conducted on a 0.078 mmol scale from commercially available preloaded 2-ClTrt β-alanine resin. General solid phase methodology and Fmoc deprotection steps were performed as previously described in chapter 3. HATU and DIPEA were employed for carboxylic acid activation during all amide bond formations again as previously described in chapter 3. In the case of pyrrole couplings, 5 equiv. of either **56**, or **57** were utilised with reaction times of 5 hours. The final terminal pyrrole

coupling was repeated, as this was anticipated to be the most inefficient reaction. To conserve reagent, couplings of the duocarmycin alkylation subunit were conducted using only 1.5 equiv. of **11**, with overnight reaction times.

Aminolysis was conducted as described by Dervan *et al.*[235] The washed resin was treated with neat *N,N*-dimethylaminopropylamine and incubated at 60 ^{o}C overnight. However, analytical HPLC analysis, did not detect any cleavage products. It was suspected that the 2-ClTrt structure provided a steric shield of the cleavage site, inhibiting aminolysis.

4.4.2 Second attempt at the solid phase synthesis of 63 using PAM resin and aminolytic cleavage.

Due to the suspected effect of 2-ClTrt resin towards aminolytic cleavage, the synthesis was repeated but beginning from a commercially preloaded Boc-β-alanine-PAM resin. This was one of the resins originally employed by the Dervan group for synthesises of polyamides cleaved by aminolysis.[234] The structure of Boc-β-alanine-PAM resin is shown in figure 4.12 alongside β-alanine-2-ClTrt resin. It can be seen that the cleavage site is considerably less sterically hindered for nucleophilic attack.

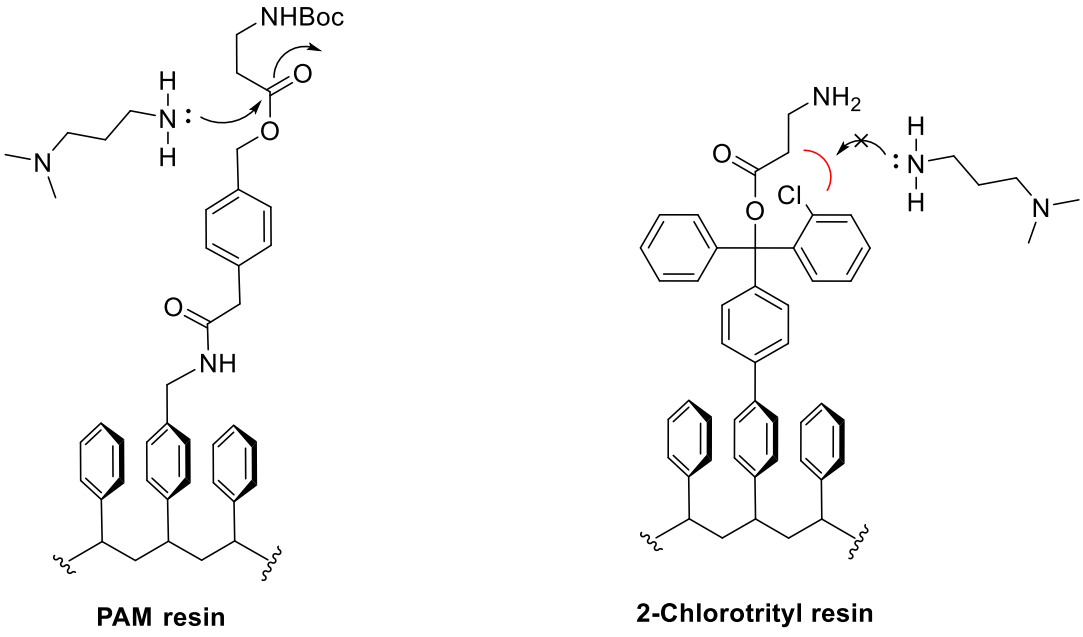


Figure 4.12 Structure of Boc-β-Alanine PAM resin, and β-Alanine 2-ClTrt resin. Steric hindrance of aminolytic cleavage with 2-ClTrt resin.

The synthesis was again conducted on a 0.078 mmol scale, and was identical to that described in section 4.3.1, with the exception of an additional initial Boc deprotection step of the immobilised β-alanine. This was achieved using 95 % TFA in DCM. PAM resins are typically used for Boc solid phase synthesis and so the linkage is stable towards TFA treatment.

Aminolysis was repeated as described in section 4.3.1. Analytical HPLC analysis of the cleavage mixture detected several products, with two dominating the HPLC trace (See figure 4.13).

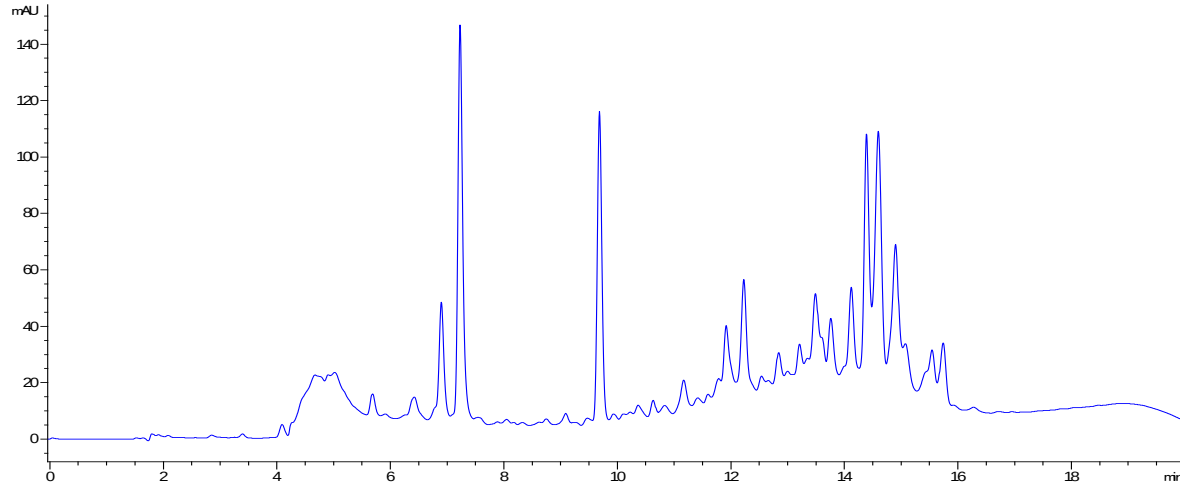
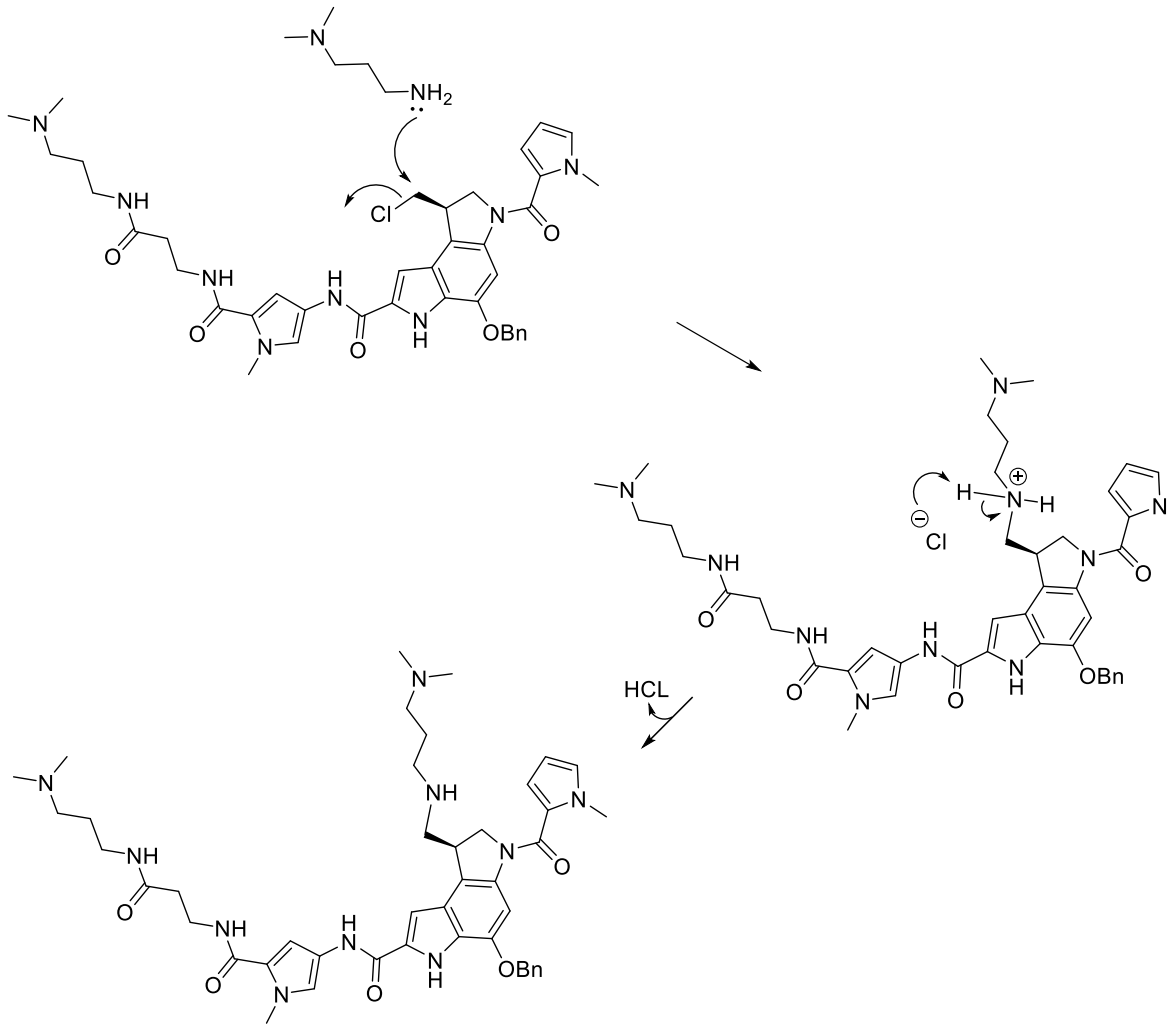


Figure 4.13 HPLC trace of cleavage mixture after aminolysis. Agilent Eclipse XDB-C18 column, 4.8 x 150 mm, 5 µm. Solvent A: [Water and 0.05 % TFA], Solvent B: [MeOH and 0.05 % TFA]. Gradient: 0% [B] to 95 % [B], from 0 min to 15 mins, 95 % [B] to 0 % [B] from 15 to 20 mins. Monitored UV 254 nm. Flow rate 1 mL/min. Column temperature 40 ^{o}C.

Although it was of course possible that the observed multiple peaks represented true side products of the synthesis, degradation during cleavage was again suspected. For example it is possible to envision displacement of the Cl atom under the aminolytic cleavage conditions, especially when considering the large excess of amine and heat energy available (see scheme 4.2).

The products were not isolated to confirm this, as it was considered more time efficient to repeat the synthesis again on 2-ClTrt resin which would allow milder cleavage using 1% TFA, followed by introduction of the dimethylaminopropylamine tail as a post cleavage modification. This strategy had seen success in the synthesis of analogue **49**, when exploring the effect on MTS activity of masking the free carboxylic acid group in chapter 3 (see section 3.14).



Scheme 4.2 Possible degradation route under aminolytic cleavage conditions.

4.4.3 Third attempt at the solid phase synthesis of 63, 2-ClTrt resin with introduction of the dimethylaminopropylamine tail as a post cleavage modification.

The synthesis was again conducted on 0.078 mmol scale, and was identical to that described in section 4.3.1. Cleavage was affected by treatment with 1 % TFA, 10% TIPS in DCM, to give the free acid. HPLC analysis showed one major product (see figure 4.14a blue trace). The crude acid was dissolved in DMF and treated with equimolar quantity of HATU and a 2-fold excess of DIPEA. After 30 secs 5 equivalence of *N*,*N*-dimethylaminopropylamine was added. HPLC analysis after one hour showed consumption of the starting material and formation of a new product (figure 4.14a red trace). As the product would not precipitate from the reaction mixture on the addition of water, it was subjected directly to preparative HPLC. The isolated product was immediately benzyl deprotected by catalytic transfer hydrogenation as previously

described in chapter 3. HPLC analysis confirmed complete deprotection after 45 min, figure 4.14a green trace. The deprotected product was isolated from excess ammonium formate, by preparative HPLC (figure 4.14b), to afford 4.7 mg of **63**, representing an overall yield of 9 %. Accurate mass spectrometry returned a mass spectrum consistent with the desired product.

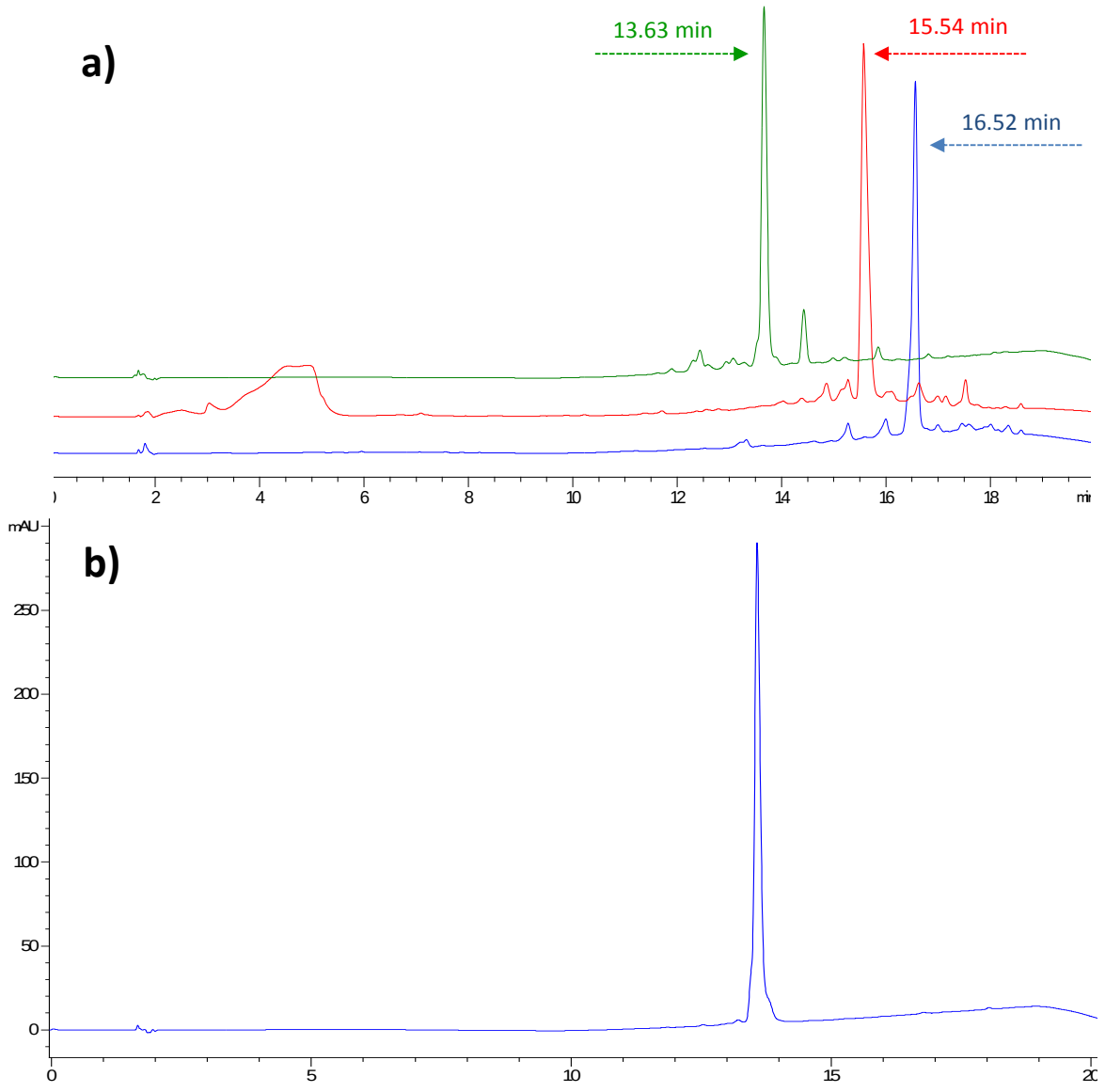


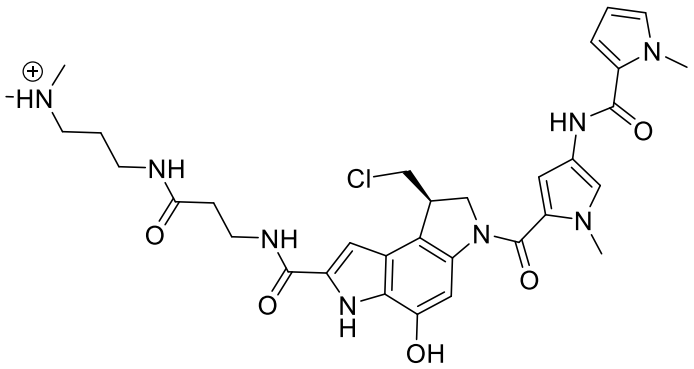
Figure 4.14 HPLC analysis during the synthesis of **63** (a) Blue trace: HPLC analysis of crude cleavage product. Red trace: HPLC analysis after 1 hour of *N*,*N*-dimethylaminopropylamine coupling. Green trace: HPLC analysis after 45 min of benzyl deprotection conditions. (b) HPLC analysis of **63** after preparative HPLC purification. Agilent Eclipse XDB-C18 column, 4.8 x 150 mm, 5 µm. Solvent A: [Water and 0.05 % TFA], Solvent B: [MeOH and 0.05 % TFA]. Gradient: 0% [B] to 95 % [B], from 0 min to 15 mins, 95 % [B] to 0 % [B] from 15 to 20 mins. Monitored UV 254 nm. Flow rate 1 mL/min. Column temperature 40 ^{o}C.

4.4.4 Introduction of the duocarmycin alkylation subunit at different chain positions.

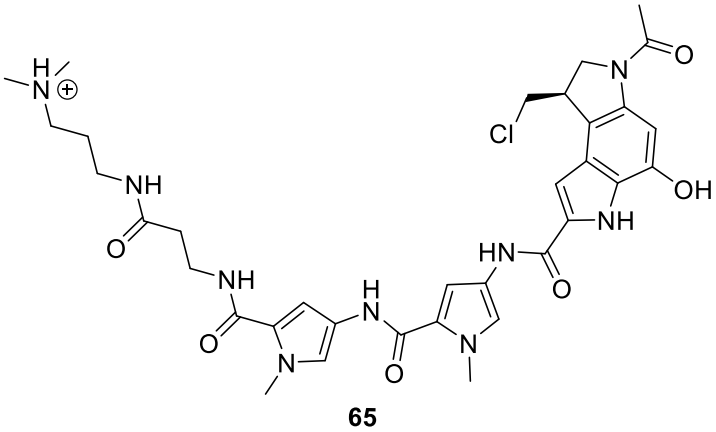
The optimised methodology was successfully applied to the synthesis of hybrids **64**, and **65** (figure 4.15). These analogues represented introduction of the duocarmycin alkylation subunit at the remaining possible chain positions, of this novel bispyrrole family.

Synthesis of **64** and **65** was conducted on a smaller 0.043 mmol scale. In the case of **65**, an additional final Fmoc deprotection and acetyl capping step was required, and this was achieved using AcCl as previous described in chapter 3.

2.2 mg of **65** and 2 mg of **64** were isolated representing overall yields of 8 % and 7 % respectively. Accurate mass spectrometry returned mass spectra consistent with the desired products.



64



65

Figure 4.15 Structure of **64** and **65**.

4.5 Antiproliferative activity of 63, and 64 and 65.

The antiproliferative activity of the three distamycin-duocarmycin hybrids against an HL-60 cell line was assessed using a MTS assay as previously described in chapter 3. The results are shown in table 4.1.

	Chain sequence	IC_{50} (μM)	[95 % CI]
63	Py-DSA-Py	1.2	[0.61-2.40]
65	Py-Py-DSA	3.6	[2.11-6.15]
64	DSA-Py-Py	>100	

Table 4.1 IC_{50} values returned by the MTS assay. HL-60 cell line. Detailed protocol can be found in the experimental chapter. Chain Sequence key: Py = *N*-methylpyrrole residue, DSA = seco-duocarmycin alkylation subunit (unnatural enantiomer). **64** = no activity at the top concentration tested (100 μM)

These preliminary results indicate that the sequence position of the duocarmycin alkylation subunit is not trivial with respect to antiproliferative activity. Low μM IC_{50} values were returned for analogues **63** and **65**, representing mid sequence and *N*-terminal incorporation. In contrast no activity was observed, at the top concentration tested (100 μM), for analogue **64** in which the duocarmycin alkylation subunit was incorporated at the *C*-terminal position.

This was an intriguing result as **64** was anticipated to possess the greatest activity. The prediction was based on this analogue containing the greatest rigid length extending from the nitrogen of the vinylogous amide. Thus, it might be predicted that this structure would produce the greatest binding induced activation of the cyclopropane (see chapter one). The lack of activity of **64** suggests that incorporation at this chain position has a detrimental effect on DNA binding, at least with respect to positioning of the cyclopropane. For example it is possible to imagine a binding model, in which the pyrrole pair bind the minor groove, but in doing so, position the duocarmycin subunit projecting out of the minor groove. This could perhaps be promoted by closer association of the cationic tail, with the polyanionic backbone of DNA, and hydrogen bonding of the carbonyl of the linking amide (see figure 4.16). Such binding would be likely to prevent nucleophilic attack of the cyclopropane by the nucleic bases.

The results also indicate there may be a small benefit to mid sequence incorporation, with compound **63** possessing a 3 fold increase in activity over **65**. However, overlapping confidence intervals suggest this may not be significant. Nevertheless, if a real effect, this may reflect subtle differences in the positioning of the cyclopropane on binding, or could

result from greater activation of the cyclopropane due to the increase rigid length extending from the *N*-terminal of the alkylation subunit inherent in the structure of **63**. It is not known whether these compounds are binding the minor groove as 1:1 complexes or as 2:1 antiparallel homodimers. Therefore it is also possible that the differences in activity could represent different preferences to bind in one mode over the other.

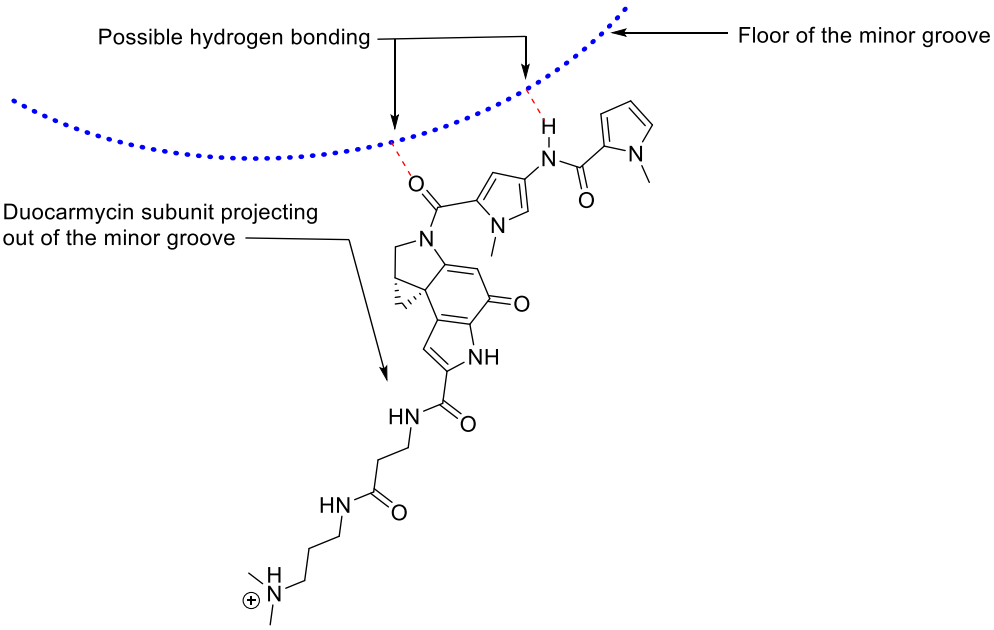


Figure 4.16 Schematic representation of possible binding model of the ring close form of **64** which could account for lack of activity.

4.6 Conclusions from chapter four and suggestions for future work continuing from these preliminary results.

Due to time constraints the above results represent the extent of progress in the exploration of this second solid phase application of our novel building block **11**. These preliminary findings suggest that **11** is a convenient substrate for the direct and in line incorporation of the duocarmycin alkylation subunit in to heterocyclic polyamide sequences, and should be applicable to the solid phase construction of longer bifunctional hairpin polyamide structures. The use of **11** in this way offers the additional advantage of allowing versatility in the sequence position in which the alkylation subunit is incorporated. It should be noted that commonly employed aminolytic cleavage strategies appear not to be suitable for release of such structures from the solid support, due to sensitivity of the duocarmycin residue. However, where a dimethylamino tail is desired, this can be introduced through post cleavage modification. The MTS assay results indicate that the chosen sequence position of the alkylation subunit, at least with respect to the unnatural enantiomer, can have a pronounce effect on antiproliferative activity. There appears to be

a mild preference for mid sequence incorporation over *N*-terminal placement (but this may not be significant), while *C*-terminal positioning may be detrimental towards activity.

The obvious next step in any future work in this area, would be to resynthesise compounds **63**, **64**, and **65**, using the natural enantiomer of **11**. It would be prudent to conduct these synthesises on a slightly larger scale, so as to allow rigorous confirmation of the structure of these novel compounds through the use of a 2-dimensional carbon-proton correlation NMR experiments such as DEPT-edited HSQC. It would be interesting to ascertain if the relationship between antiproliferative activity and sequence position of the alkylation subunit, observed with the original analogues, is consistent after the change in stereochemistry, as well confirming the expected increase in activity which might be predicted from use of the natural enantiomer based on previous literature on the duocarmycin family.

Following on from these experiments, the logical extension of this work would be to begin to explore the incorporation of the duocarmycin alkylation subunit in to full hairpin polyamide structures. It would be particularly interesting to investigate the effect on sequence selectivity of incorporation of the alkylation subunit at different chain positions of a well characterised existing hairpin polyamide sequence. This work might highlight potential sequence positions where the alkylation subunit can be substituted for native heterocycles without compromising minor groove sequence recognition. Such results could make a significant contribution towards the development of selective targeting of cytotoxicity through bifunctional hairpin structures which target specific mutations in the genome of tumour cells; thus potentially significantly improving the therapeutic window of the duocarmycin family while ablating the necessity for prodrug strategies or additional targeted delivery vehicles. Such an approach might benefit from allowing the construction of bespoke bifunctional hairpin polyamides, for the targeting of different malignancies based on their specific genetics.

These approaches are likely to require recognition of long sequences in order to produce the necessary genome wide selectivity. As discussed in section 4.1, progress in this area has already been made, with the advent of tandem hairpin structures, or the use of mid sequence aliphatic residues to relax the curvature of long hairpin structures. It maybe that the increased length of the duocarmycin residue when compared to a PyPy sequence, may also prove to serve as a mechanism to relax the over-curvature of longer hairpin structures, when incorporated in mid sequence chain positions.

In addition to allowing versatility in the chain position of the alkylation subunit, the use of **11** as a solid phase building block, also offers the potential for the convenient introduction

of multiple alkylation subunits. Continuing from the above suggested work, it would be interesting to investigate the effect of incorporating a duocarmycin residue in both the parallel and antiparallel chains of a single hairpin structure. Perhaps such compounds would yield sequence selective DNA crosslinking agents. If so these compounds would likely possess increased potency with respect to antiproliferative activity.

These ideas represent a significant extension in the use of **11** to date. However, it is the author's hope that the work contained within this thesis, including the development of the building block itself, and the initial investigations towards its use in different solid phase applications, have provided a suitable grounding from which more ambitious projects, such as those suggested, could be launched in the future.

5

Chapter Five

Experimental

5.1 General considerations.

5.1.1 Reagent, and Solvent Preparation.

All chemicals were reagent grade and purchased from Sigma Aldrich, Novabiochem, VWR, AGTC, Fluorochem, or Fisher Scientific. HPLC mobile phases were prepared using HPLC grade solvents. Solvents where specified as anhydrous were bought as such and assumed to conform to the manufacturers standards. All water used was distilled. All DMF for solid phase synthesis was purchased as peptide grade from AGTC.

5.1.2 Physical Characterisation and Spectroscopic Techniques.

^{1}H and ^{13}C-NMR spectra were recorded in Fourier Transform mode on a Bruker B-ACS 60 Ultrashield 400 plus spectrometer, operating at a nominal ^{1}H NMR frequency of 400 MHz, using the specified deuterated solvent. All spectra were processed using Topspin 3.0 software. The chemical shifts for both ^{1}H and ^{13}C-spectra were recorded in ppm and were referenced to the residual solvent peak. Multiplicities in the NMR spectra are described as: s = singlet, d = doublet, dd = doublet of doublets, t = triplet, q = quartet, m = multiplet, br = broad, appt = apparent; coupling constants are reported in Hz. Accurate mass spectra were recorded at the EPSRC National Mass Spectrometry Service Centre, Swansea. Infrared spectra were recorded as neat samples using a Perkin-Elmer Spectrum BX FT-IR and manipulated using Spectrum v5.3 Software.

5.1.3 General Chromatographic Techniques.

Thin layer chromatography: was performed on Merck aluminium plates coated with 0.2 mm silica gel-60 F_{254}. After elution, the TLC plates were visualised under UV light.

Flash chromatography: Unless otherwise stated normal phase flash chromatography was performed in glass columns on silica gel for column chromatography (particle size 60 µm), using hand bellows to apply positive pressure.

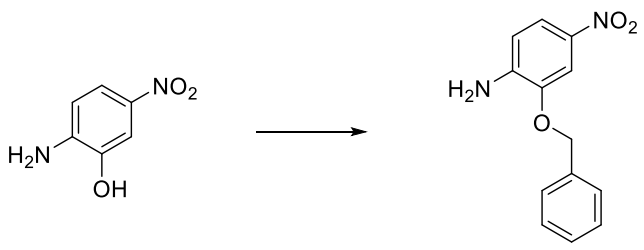
General RP-Flash chromatography: Unless otherwise stated RP-Flash chromatography was performed on a Biotage Isolera 4, using a pre-packed Biotage SNAP 12 g C18 column, and a flow rate of 20 mL/min. Solvent A = 95 % H_2O 5 % MeOH + 0.05% TFA and Solvent B = 95 % MeOH 5 % H_2O + 0.05% TFA. Gradient 0 % B → 100 % B over 15 minutes 100 % B → 0 % B over 5 mins. Detection wavelength 254 nm. Samples were dry loaded by adsorption on celite.

General analytical RP-HPLC: Unless otherwise stated analytical RP-HPLC was performed on an Agilent 1200 using an Agilent eclipse XDB-C18 column, 4.6 x 150mm, 5 µm and a flow rate of 1 mL/min. Solvent A = H_2O + 0.05% TFA and Solvent B = MeOH + 0.05% TFA. Gradient 5% B → 95% B over 15 minutes 95% B → 5 % B over 5 mins. Detection wavelength 254 nm.

General preparative RP-HPLC: Unless otherwise stated preparative RP-HPLC was performed on an Agilent 1260 infinity using an Agilent eclipse XDB-C18 column, 21.2 x 150 mm, 5 µm and a flow rate of 20 mL/min and a flow rate of 20 mL/min. Solvent A = 95 % H_2O 5 % MeOH + 0.05% TFA and Solvent B = 95 % MeOH 5 % H_2O + 0.05% TFA. Gradient 0 % B → 100 % B over 15 minutes 100 % B → 0 % B over 5 mins. Detection wavelength 254 nm.

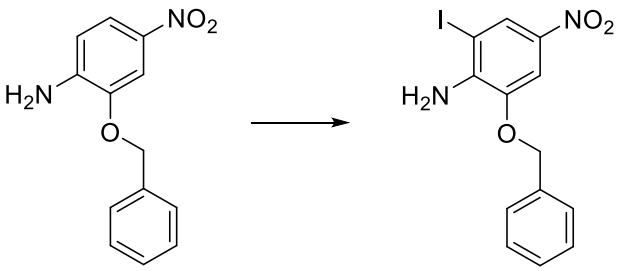
5.2 Organic Synthesis.

5.2.1 Synthesis of 11.



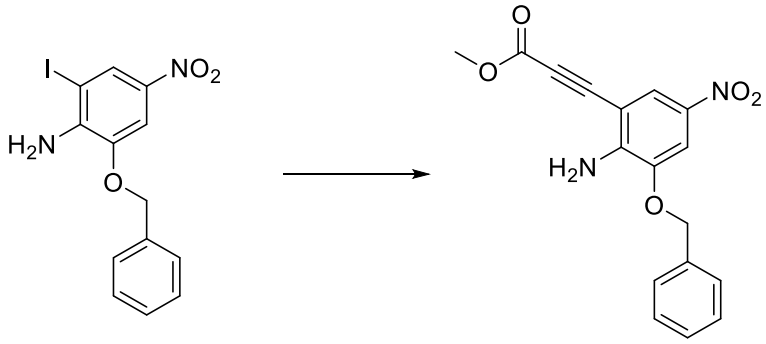
21 **22**

(22): BnBr (21.00 mL, 178 mmol) was added dropwise to a stirring suspension of 2-amino-5-nitrophenol **21** (25.00 g, 162 mmol) and K_2CO_3 (49.30 g, 357 mmol) in DMF (250 mL) at room temperature. After 20 hours, the reaction mixture was poured over crushed ice. The precipitate was collected by filtration and triturated with cold water prior to drying at 40 ^{O}C under vacuum overnight. The reaction was repeated on the same scale and the two batches combined to afford 78.05 g of **22** as a yellow/brown amorphous solid (98.5 % average yield over the 2 batches). R_f 0.17 (20 % EtOAc in hexane); mp 147-149 ^{o}C, lit mp 144-145 ^{o}C.[110] ^{1}H NMR ($CDCl_3$, 400 MHz) δ 7.83 (1H, dd, *J*=2.4, 8.7), 7.77 (1H, d, *J*=2.4), 7.37-7.46 (5H, m), 6.66 (1H, d, *J*=8.7), 5.15 (2H, s), 4.60 (2H, brs). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 144.6, 143.6, 138.7, 135.9, 128.9, 128.7, 128.0, 119.5, 112.1, 107.4, 71.0. IR (neat) ν_{max} 3483, 3359, 3225, 3188, 3075, 2939, 2876, 1622, 1579, 1519, 1480, 1455, 1386, 1282, 1222, 1176, 1091, 1007, 950, 914, 870, 853, 818, 797, 755, 744, 727, 697, 643, 623 cm^{-1}. HRMS (ES+) calculated for $C_{13}H_{13}N_2O_3$ $(M+H)^+$ 245.0921 found 245.0923. NMR consistent with the literature.[110]



22 23

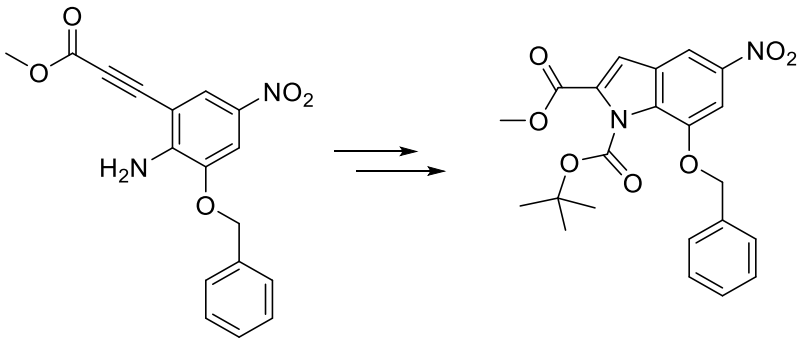
(23): Conc.H_2SO_4 (800 µL, 15.15 mmol) was added to a stirring solution of **22** (37.00 g, 151 mmol) in DMF (555 mL), followed by portionwise addition of NIS (51.10 g, 227 mmol) at room temperature. After 4 hours, the reaction mixture was poured over crushed ice. The precipitate was collected by filtration and triturated with cold water, followed by cold hexane, prior to drying at 40 ^{O}C under vacuum overnight. The reaction was repeated with 39.10 g of **23** and the two batches combined to afford 105.65 g as a bright yellow amorphous soild (91.5 % average yield over the 2 bathes). R_f 0.31 (20 % EtOAc in hexene); mp 103-106 ^{o}C, lit mp 105-106 ^{o}C.[110] ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.29 (1H, d, *J*=2.3), 7.74 (1H, d, *J*=2.3), 7.38-7.44 (5H, m) 5.16 (2H, s), 5.02 (2H brs). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 144.1, 143.3, 138.9, 135.4, 129.0, 128.9, 128.3, 128.1, 106.7, 178.5, 71.5. IR (neat) ν_{max} 3476, 3379, 3359, 3091, 3056, 3030, 2357, 2333, 1602, 1568, 1497, 1451, 1425, 1386, 1282, 1237, 1099, 1037, 1025, 869, 849, 819, 740, 726, 692 cm^{-1}. HRMS (ES+) calculated for $C_{13}H_{12}IN_2O_3$ $(M+H)^+$ 370.9887 found 370.9890. NMR consistent with the literature.[110]



23 26

(26): 23 (40.80 g, 110 mmol) was dissolved in anhydrous DMF (1225 mL). The resulting solution was degassed with a stream of N_2 for 30 mins prior to addition of methyl propiolate (37.10 mL, 441 mmol), $Pd(PPh_3)_2CL_2$ (3.87 g, 5.51 mmol), $ZnBr_2$ (99.00 g, 441 mmol), and DIPEA (77.00 ml, 441 mmol) at room temperature . The reaction mixture was

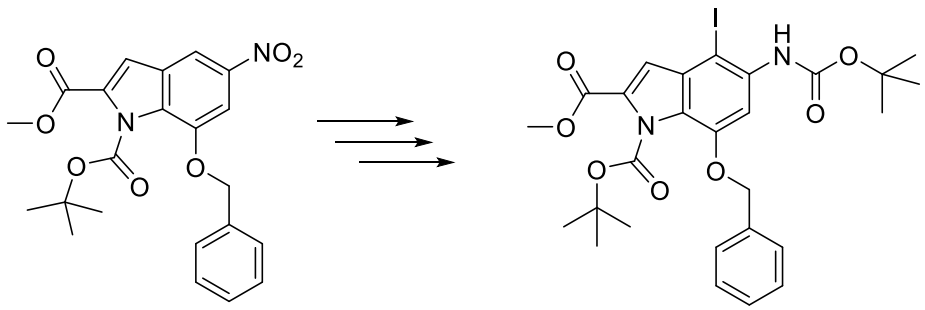
then heated to 66 ^{o}C and stirred overnight under N_2. After cooling to room temperature the reaction mixture was poured over crushed ice, and the resulting chocolate colour precipitate collected by filtration. The reaction was repeated with 51.00 g of **23**, and the precipitates were combined prior to adsorption onto 250 g of silica. Elution through a 1 kg silica plug with 50% ethyl acetate and hexane afforded 62.00 g of **26** as an orange amorphous solid (77 % yield). R_f 0.16 (20 % EtOAc in hexane); mp 136-139 ^{o}C. ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.06 (1H, d, *J*=2.4), 7.76 (1H, d, *J*=2.4), 7.38-7.45 (5H, m), 5.32 (2H, brs), 5.17 (2H, s), 3.86 (3H, s). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 154.1, 146.7, 144.5, 137.8, 135.2, 129.1, 129.0, 128.1, 123.0, 108.3, 101.0, 87.1, 81.1, 71.5, 53.1. IR (neat) ν_{max} 3499, 3391, 3351, 3087, 3063, 3030, 2951, 2204, 1698, 1611, 1455, 1430, 1393, 1325, 1299, 1237, 1215, 1148, 1093, 1040, 1028, 1001, 886, 859, 755, 740, 731, 694, 657, 612 cm^{-1}. HRMS (ES+) calculated for $C_{17}H_{15}N_2O_5$ $(M+H)^+$ 327.0975 found 327.0979.



26 **16**

(16): 26 (60.00 g, 184 mmol) in anhydrous THF (858 mL) was treated with 1M TBAF in THF solution (368.00 mL, 368 mmol) and refluxed at 66 ^{o}C for 1 hour. After cooling to room temperature the THF was removed by rotary evaporation under reduced pressure. The residue was dissolved in ethyl acetate (1000 mL) and washed 3 times with water (1000 mL). Concentration of the ethyl acetate followed by co-evaporation of the residue with DCM afforded crude **15** as a dark purple foam. The foam was dissolved in DCM (1000 mL) and treated with Boc_2O (80.00 g, 368 mmol), and DMAP (22.46 g, 184 mmol) at room temperature for 1.5 hours. Removal of the DCM gave a dark foam which was purified by silica gel chromatography using an Isco automated flash chromatography system. The crude was dry loaded on to a 1.5 kg pre-packed silica column adsorbed on to 200 g of silica. A linear gradient of 0 to 30 % ethyl acetate in hexane was run over 23 column volumes and then held at 30 % ethyl acetate until complete elution of the product. Removal of the solvent afford 31.00 g of **16** as an orange amorphous solid (39 % yield of 2 steps). R_f 0.33 (20 % EtOAc in hexane); mp 168-171 ^{o}C, reported as an oil in the literature.[90, 96] ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.26 (1H, d, *J*=1.9), 7.67 (1H, d, *J*=1.9), 7.49-

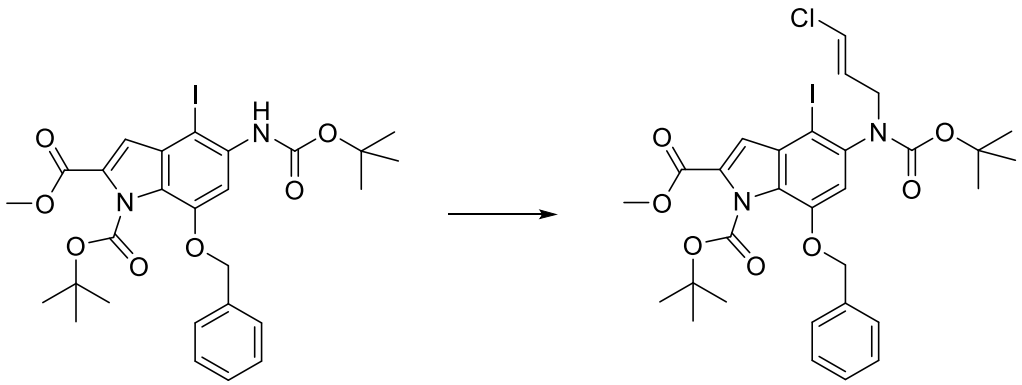
7.45 (2H, m), 7.41-7.34 (3H, m), 7.33 (1H, s) 5.33 (2H, s), 3.94 (3H, s), 1.47 (9H, s). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 160.5, 149.3, 145.6, 143.6, 135.2, 130.2, 128.9, 128.7, 128.2, 126.4, 112.7, 112.5, 102.2, 86.5, 71.3, 52.5, 27.9, 27.3. IR (neat) ν_{max} 3127, 3099, 3050, 2981, 2949, 1765, 1722, 1586, 1512, 1437, 1388, 1372, 1325, 1252, 1223, 1151, 1115, 1073, 982, 875, 840, 822, 801, 778, 766, 742, 729, 697, 606 cm^{-1}. HRMS (ES+) calculated for $C_{22}H_{23}N_2O_7$ $(M+H)^+$ 427.1500 found 427.1499. NMR consistent with the literature.[90, 96]



16 **29**

(29): 16 (15.00 g, 35.2 mmol) was dissolved in THF (293 ml) and treated with zinc powder (34.50 g, 528 mmol), NH_4CL (18.82 g, 352 mmol), Boc_2O (23.03 g, 106 mmol), DMAP (430 mg, 3.52 mmol), and water (58.60 mL). The resulting suspension was stirred vigorously at room temperature overnight. After removal of the zinc by filtration, the THF was evaporated and the residue taken up in ether (500 mL). The ether was washed 3 times with water (250 mL) and died over $MgSO_4$. Co-evaporation with DCM gave crude **18** as a light brown foam. The reaction was repeated on the same scale and the crudes combined and dissolved in DMF (352 mL). Conc. H_2SO_4 (0.375 mL, 7.04 mmol) was added followed by portionwise addition of NIS (23.75 g, 106 mmol) at room temperature. After 3 hours the reaction mixture was diluted with Et_2O (1000 mL), and washed once with 50 % saturated brine in water (1000 mL), twice with water (1000 mL), and once with saturated brine (1000 mL). The first wash was back extracted 3 times with Et_2O (500 mL), which was subsequently combined and washed twice with saturated brine (1000 mL). All the Et_2O was combined and concentrated to give a dark red foam which was purified by silica gel chromatography using an Isco automated flash chromatography system. The crude was dry loaded on to a 750 g pre-packed silica column adsorbed on to 170 g of celite. A linear gradient of 0 to 20 % ethyl acetate in hexane was run over 16 column volumes. Removal of the solvent afford 26.00 g of **29** as an off white foam which dried to an amorphous solid (59 % yield over 3 steps). R_f 0.37 (20 % EtOAc in hexane); mp 158-161 ^oC, the literature does not report a mp.[90] ^1H NMR ($CDCl_3$, 400 MHz) δ 7.79 (1H, brs), 7.49-7.46 (2H, m), 7.30-7.38 (3H, m), 7.09 (1H, s), 6.77 (1H, brs), 5.24 (2H, s), 3.91 (3H,

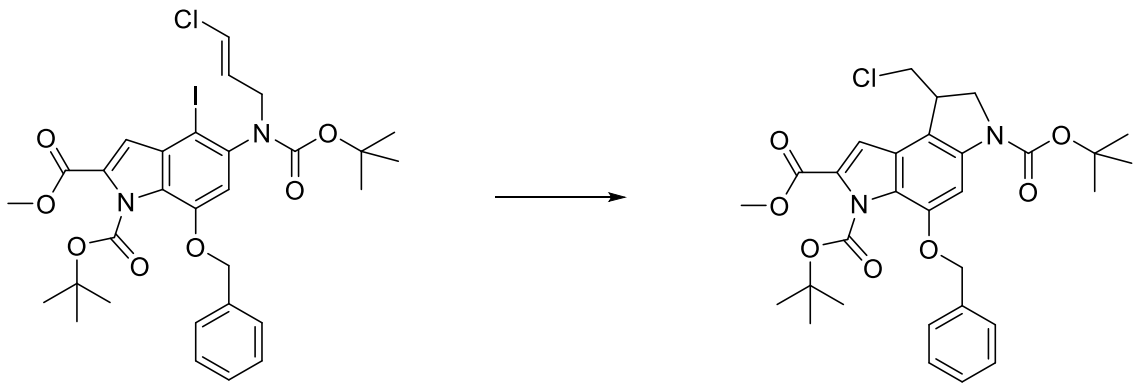
s), 1.54 (9H, s), 1.41 (9H, s). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 160.9, 153.1, 149.9, 146.6, 136.0, 134.7, 131.5, 128.7, 128.6, 128.4, 127.8, 123.7, 114.6, 102.6, 85.6, 81.0, 71.1, 52.3, 28.5, 27.3. IR (neat) ν_{max} 3355, 2984, 2933, 1763, 1725, 1716, 1615, 1575, 1541, 1505, 1449, 1393, 1361, 1338, 1310, 1256, 1221, 1152, 1080, 980, 908, 878, 843, 817, 758, 723, 693 cm^{-1}. HRMS (ES+) calculated for $C_{27}H_{32}O_7N_2I$ $(M+H)^+$ 623.1249 found 623.1246. NMR consistent with the literature.[90]



29 **30**

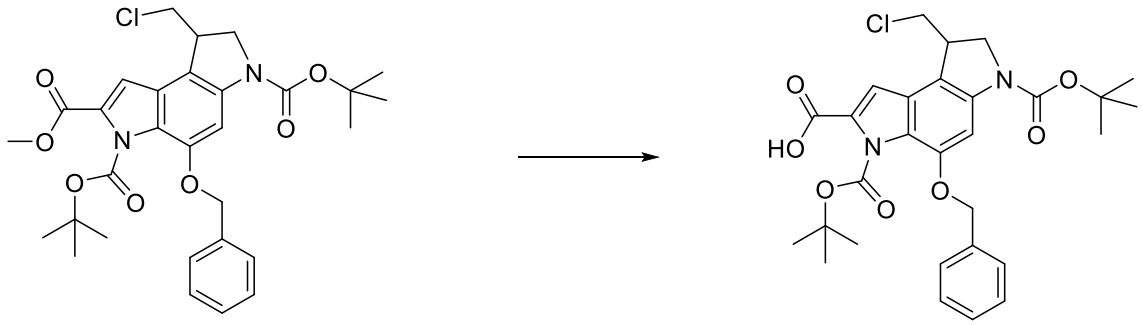
(30): 29 (26.00 g, 41.8 mmol) was dissolved in DMF (418 mL) and treated with t-BuOK (9.37 g, 84 mmol) and technical grade (90 %) 1,3-dichloropropene as a mixture of cis and trans isomers (12.90 mL, 125 mmol). After stirring for 1.5 hours with the vessel submerged in a room temperature water bath, the reaction mixture was cooled to 0 ^oC and quenched with saturated aqueous NH_4CL (20 mL). The mixture was diluted with Et_2O (1000 mL), and washed twice with of 50 % saturated brine in water (1000 mL), and once with saturated brine (1000 mL). The Et_2O was dried over $MgSO_4$, concentrated, and co-evaporated with DCM 6 times to afford a brown foam which was purified by silica gel chromatography using an Isco automated flash chromatography system. The crude was dry loaded on to a 220 g pre-packed silica column adsorbed on silica. A linear gradient of 0 to 10 % ethyl acetate in hexane was run over 16 column volumes. Removal of the solvent afforded 18.00 g of **30** as a light brown foam which dried to an amorphous solid (62 % yield - mixture of *E*/*Z* isomers). R_f 0.31 (20 % EtOAc in hexane); mp 104-106 ^oC, reported as an oil in the literature.[90] ^{1}H NMR ($CDCl_3$, 400 MHz, mixture of *E*/*Z* isomers) δ 7.28-7.44 (5H, m) 7.18 (1H, s), 6.65-6.47 (1H, m), 5.80-6.00 (2H, m]), 5.17-5.28 (2H, m), 4.46 & 4.18 (1H, m), 4.33 & 3.73 (1H, m), 3.93 (3H, s), 1.53 (9H, s), 1.29 & 1.27 (9H, s). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 160.9, 154.2, 150.0, 145.7, 138.7, 135.9, 132.2, 128.9, 128.4, 128.0, 127.5, 125.4, 121.8, 120.7, 115.3, 109.7 86.0, 83.9, 80.6, 70.7, 52.4, 49.5, 46.2, 28.4, 27.3. IR (neat) ν_{max} 2976, 2921, 1775, 1731, 1702, 1694, 1571, 1535, 1467,

1454, 1435, 1391, 1372, 1299, 1251, 1227, 1150, 1118, 1077, 978, 932, 885, 842, 829, 782, 764, 739, 731, 699 cm^{-1}. HRMS (ES+) calculated for $C_{30}H_{35}O_7N_2ClI$ $(M+H)^+$ 697.1172 found 697.1174. NMR consistent with the literature.[90]



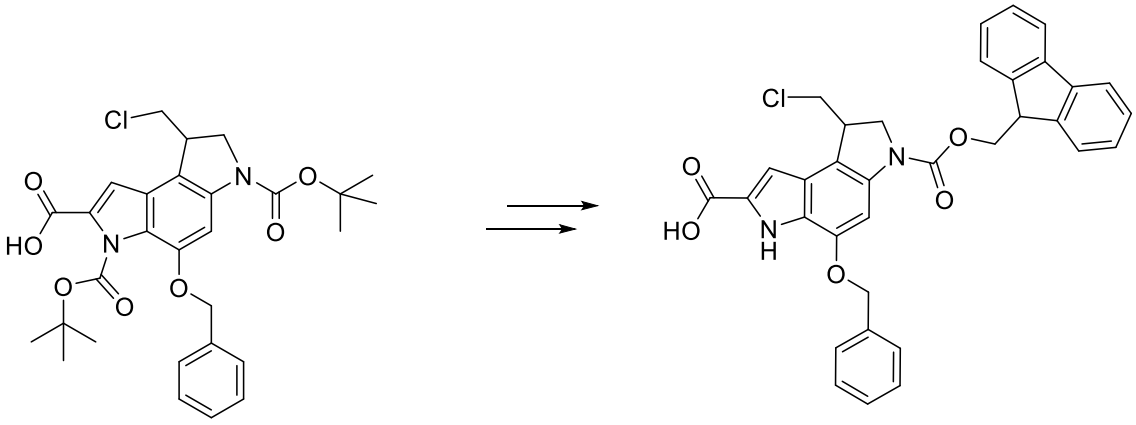
30 **10**

(10): 30 (9.00 g, 12.9 mmol) was dissolved in anhydrous toluene and degassed with a stream of N_2 for 45 mins prior to addition of AIBN (0.53 g, 3.23 mmol) and TTMSS (4.38 mL, 14.20 mmol). The resulting solution was refluxed at 90 ^oC under N_2. After 1 hour the reaction mixture was allowed to cool to room temperature before being concentrated and subjected directly to silica gel column chromatography using an Isco automated flash chromatography system. A 120 g pre-packed silica column was used, and 0 % ethyl acetate in hexane run for 5 column volumes rising to 10 % linearly over the subsequent 5 column volumes, holding at 10 % until complete elution of the product. The reaction was repeated on the same scale and the products combined, affording 10.28 g of **10** as white foam which dried to an amorphous solid (70 % yield). R_f 0.30 (20 % EtOAc in hexane); mp 115-118 ^oC, reported as an oil in the literature.[96] ^1H NMR (DMSO-d_6, 400 MHz) δ 7.69 (1H brs), 7.47-7.29 (6H, m), 5.27 (2H, s), 4.13 (1H, t, *J*=9.7), 4.06-3.89 (4H, m), 3.87 (3H, s), 1.48 (9H, s), 1.39 (9H, s). ^{13}C NMR (DMSO-d_6, 100 MHz) δ 160.4, 151.4, 149.5, 145.1, 136.2, 128.4, 128.0, 127.9, 123.5, 113.2, 108.5, 97.4, 85.0, 80.3, 69.7, 52.3, 52.2, 47.6, 40.7 (obscured by DMSO peak observed by HSQC), 28.0, 26.8, 22.0. IR (neat) ν_{max} 3002, 2977, 2921, 2357, 1782, 1720, 1698, 1593, 1538, 1494, 1494, 1439, 1417, 1379, 1343, 1241, 1214, 1141, 1089, 1022, 988, 918, 899, 836, 765, 745, 712, 699, 691, 664 cm^{-1}. HRMS (ES+) calculated for $C_{30}H_{36}O_7N_2Cl$ $(M+H)^+$ 571.2206 found 571.2201. NMR consistent with the literature (literature NMR recorded in a different solvent [$CDCl_3$]).[96]



10 **31**

(31): 10 (10.28 g, 18.00 mmol) was dissolved in a mixture of THF (167 ml) and MeOH (111 ml) and treated with a saturated aqueous solution of LiOH (56.00 mL) dropwise. After 3 hours the THF and MeOH was removed under reduced pressure, and the residue diluted with water (100 mL). Acidification with 5 M HCl promoted the precipitation **31** as a white solid which was collected by filtration. The solid was recovered by dissolution in ethyl acetate and co-evaporation with DCM afforded 10.00 g of **31** as a light green foam which dried to an amorphous solid (100 % yield). R_f 0.48 (10 % MeOH in CH_2Cl_2); mp 174-178 ^oC. ^1H NMR (DMSO-d_6, 400 MHz) δ 13.43 (1H, brs), 7.69 (1H, brs), 7.49-7.29 (6H, m), 5.26 (2H, s), 4.13 (1H, t, *J*=9.7), 4.05-3.87 (4H, m), 1.49 (9H, s), 1.37 (9H, s). ^{13}C NMR (DMSO-d_6, 100 MHz) δ 161.5, 151.5, 149.1, 136.2, 129.5, 128.4, 128.0, 127.6, 123.6, 123.4, 122.4, 107.7 97.1, 84.6, 80.0, 69.7, 52.2, 47.6, 40.8 (obscured by DMSO peak observed by HSQC), 28.1, 26.8, 22.0. IR (neat) ν_{max} 2976, 2929, 2361, 2328, 1770, 1694, 1683, 1593, 1538, 1495, 1418, 1393, 1368, 1251, 1142, 1085, 1013, 978, 908, 942, 792, 745, 695, 668 cm^{-1}. HRMS (ES+) calculated for $C_{29}H_{34}O_7N_2Cl$ $(M+H)^+$ 557.2049 found 557.2044.



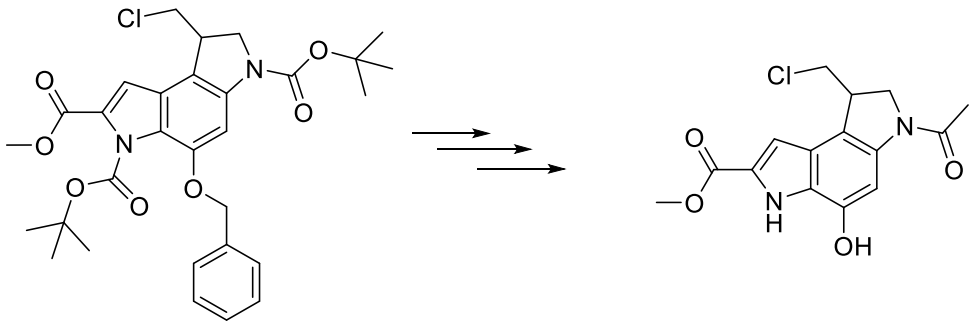
10 **11**

(11): 10 (10.00 g, 17.95 mmol) was dissolved in 4 M HCl in dioxane (180.00 mL) and stirred at room temperature overnight. Following removal of the dioxane under reduced pressure, the residue was dissolved in THF (269 mL). The resulting solution was cooled to

0 ^{o}C, before being treated with $NaHCO_3$ (4.52 g, 53.9 mmol) in water (90 mL), followed by Fmoc-Cl (4.64 g, 17.95 mmol) dropwise in THF (100 mL). After 5 min the reaction mixture was quenched with MeOH (2 mL), and the THF and MeOH removed under reduced pressure. The remaining mixture was acidified with 2 M HCl, and extracted 3 times with 2-MeTHF, and dried over $MgSO_4$. Crude **11** was purified by silica gel chromatography using an Isco automated flash chromatography system. The crude was dry loaded on to a 220 g pre-packed silica column adsorbed on to 18 g of silica. A linear gradient of 0 to 5 % MeOH in DCM was run. Removal of the solvent afforded 8.30 g of racemic **11** as a light green/brown foam which dried to an amorphous solid (80 % yield over 2 steps). R_f 0.42 (10 % MeOH in CH_2Cl_2); 125-128 ^{o}C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.97 (1H, brs), 11.90 (1H, s), 7.90 (2H, d, *J*=6.7), 7.74-7.68 (2H, m), 7.67-7.57 (2H, m), 7.53-7.23 (8H, m), 7.20 (1H, d, *J*=1.8), 5.35-5.84 (2H, brs, [rotameric coalescence observed at 333K, δ 5.17, 2H, s]), 4.74-4.31 (3H, m, [rotameric coalescence observed at 333K, δ 4.55, 2H, app quin, δ 4.39, 1H, t, *J*=6.6]), 4.23-4.14 (1H, m), 4.10-3.94 (3H, m), 3.93-3.84 (1H, m). ^{13}C NMR (DMSO- d_6, 100 MHz) δ 162.5, 152.0, 145.6, 143.8, 140.8, 136.8, 129.9, 128.2, 127.8, 127.6, 127.4, 127.2, 125.6, 125.1, 124.1, 120.2, 112.7, 105.8, 95.4, 69.5, 66.6, 51.9, 47.6, 46.7, 41.0, 34.4. IR (neat) ν_{max} 2950, 2367, 2320, 1694, 1682, 1593, 1538, 1441, 1404, 1318, 1247, 1218, 1171, 1131, 1085, 1028, 966, 903, 827, 737, 696, 667, 621 cm^{-1}. HRMS (ES-) calculated for $C_{34}H_{26}O_5N_2Cl$ $(M-H)^-$ 577.1536 found 577.1527.

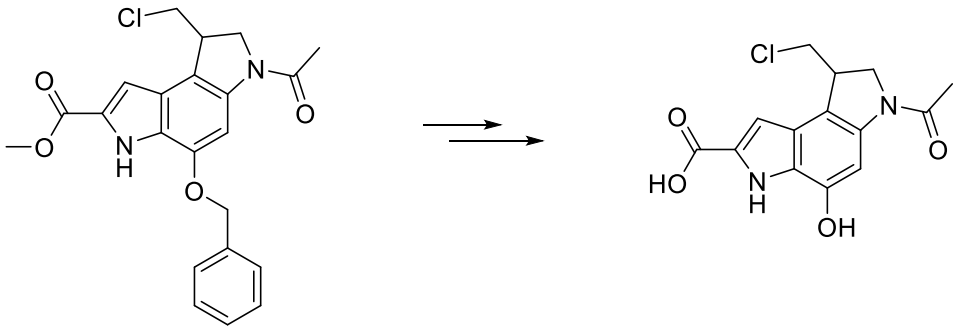
Preparative chiral resolution of **11** was achieved using super critical fluid chromatography. Separation was affected using a Chiralpak AD-H column (250 x 30 mm, 5 micron), and an isocratic flow of 50 % CO_2, and 50 % IPA containing 0.1 % TFA, at 45 mL per min. The back pressure was regulated at 10 MPa, and column temperature controlled at 40 ^{o}C. A racemate of **11** (9.8 g) was dissolved in THF:MeOH 1:1 (100 mL), and 1.25 mL (125 mg) injected every 9 mins. Fractions were monitored by UV (220 nm), collected, combined and dried to afford 2.82 g of peak 1 (5.5 min), and 3.1 g of peak 2 (7 min), both as cream amorphous solids (α = 1.27). Peak 1 $[\alpha]_D$ –20 (*c* 0.05, DMF); mp 204-207 ^{o}C. Peak 2 $[\alpha]_D$ +20 (*c* 0.05, DMF); mp 204-207 ^{o}C.

5.2.2 Synthesis of control compounds 42 and 43



10 **42**

(42): 10 (50.0 mg, 0.087 mmol) was dissolved in 4 M HCL in EtOAc (5 mL) containing TIPS (500 µL) and the solution was stirred overnight at room temperature. After removal of the solvent under reduced pressure, the residue was taken up in DMF (7 mL), and cooled to 0 ^{o}C. The solution was treated with DIPEA (30 µL, 0.17 mmol), and AcCl (6 µL, 0.087 mmol) and stirred under N_2. After 2 hours the reaction mixture was poured over crashed ice and the product was collected as a beige precipitate. This was dissolved in a 1:1 mixture of THF and MeOH (2 mL) and added to a suspension of 10 % Pd/C (20 mg) in 25 % aqueous ammonium formate (300 µL) under N_2. After 1 hour, the reaction mixture was filtered through celite. Flash chromatography (silica gel, 7 x 1 cm, 5 % MeOH in DCM) afforded 17.0 mg **42** as a white amorphous solid (60 % yield over 3 steps). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 11.55 (1H, brs), 9.72 (1H, s), 7.75 (1H, s), 7.22 (1H, app d, *J*=2.10), 4.31 (1H, t, *J*=11.6), 4.10-3.96 (3H, m), 3.91-3.87 (1H, m), 3.85 (3H, s), 2.15 (3H, s). ^{13}C NMR (DMSO-d_6, 100 MHz) δ 167.4, 161.4, 143.6, 138.0, 127.7, 125.4, 124.0, 111.7, 106.0, 99.7, 53.2, 51.8, 47.7, 41.5, 24.1. HRMS (ES+) calculated for $C_{15}H_{16}ClN_2O_4$ $(M+H)^+$ 323.0793 found 323.0797.



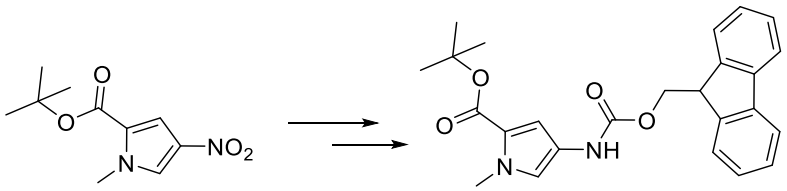
66 **43**

(43): 66 (57.0 mg, 0.14 mmol) was dissolved in a 3:2:1 mixture of THF:MeOH:H_2O (6 mL), and treated with LiOH.H_2O (110.0 mg, 2.62 mmol) overnight at room temperature. The

organic solvents were removed under reduced pressure, and the residue diluted with 1 M HCl (10 mL). The mixture was cooled to 4 ^{o}C for 72 hrs, and the product was collected as a beige precipitate by centrifugation. This was dissolved in a 1:1 mixture of THF and MeOH (2 mL) and added to a suspension of 10 % Pd/C (20 mg) in 25 % aqueous ammonium formate (300 µL) under N_2. After 1 hour the reaction mixture was filtered through celite and the crude purified by preparative HPLC. Lyophilization afforded 6.5 mg of **43** as a tan amorphous solid (15 % yield over 2 steps). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.92 (1H, brs), 11.33 (1H, s), 9.62 (1H, s), 7.72 (1H, s), 7.13 (1H, s), 4.37-4.25 (1H, m), 4.12-3.95 (3H, m), 3.92-3.81 (1H, m), 2.15 (3H, s). ^{13}C NMR (DMSO-d_6, 100 MHz) δ 167.4, 162.5, 143.5, 137.9, 129.1, 125.1, 124.1, 111.6, 105.6, 99.5, 53.2, 47.7, 41.5, 24.1. HRMS (ES+) calculated for $C_{14}H_{14}ClN_2O_4$ $(M+H)^+$ 309.0637 found 309.0637.

5.2.3 Synthesis of 56 and 57.

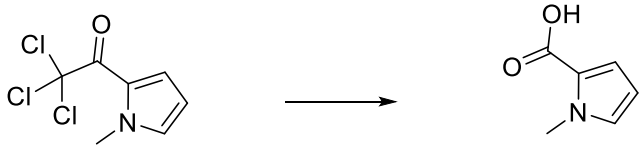
Synthesis of the pyrrole building blocks **56** and **57** (scheme 4.1, chapter four) was conducted as prevoiusly described by Dervan *et al.*, with the exception of the convertion of intermediate **61** to intermediate **62**. Which was conducted as descriped below along with the final steps in each synthesis.



61 **62**

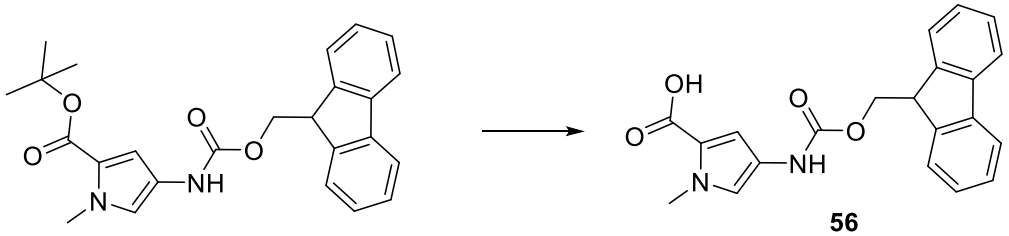
(62): 61 (1.00 g, 4.4 mmol) was dissolved in THF (60 mL). The resulting solution was treated with 10 % Pd/C (176 mg), as a slurry in 25 % w/v aq. ammonium formate (12 mL), and the suspension stirred vigorously at room temperature. Further additions of 10 % Pd/C (50 mg) as a slurry in 25 % w/v aq. ammonium formate (1 mL), were made every 30 mins until complete consumption of the starting material was observed, as monitored by TLC. After 4 hours the reaction mixture was filtered through a short column of diatomaceous earth directly in to a stirring solution of Fmoc-Cl (1.143 g, 4.4 mmol) in anhydrous DCM (180 mL), at 0 ^{o}C. The mixture was left to stir at 0 ^{o}C for 30 mins, before being wash 3 times with water. The organic layer was concentrated by rotary evaporation and subjected directly to flash chromatography (silica gel 20 x 3 cm, EtOAc in DCM, 0 % to 1 %) affording 1.57 g of **62** as a crunchy white foam which dried to an amorphous solid (84 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 9.42 (1H, s), 7.92-7.88 (2H, m), 7.73-7.67

(2H, m), 7.45-7.39 (2H, m), 7.36-7.31 (2H, m), 7.01 (1H, appt brs), 6.61 (1H, appt brs), 4.43 (2H, d, *J*= 6.3), 4.27 (1H, t, *J*=6.3), 3.76 (3H, s), 1.48 (9H, s). Consistent with the literature.[235]



59 **57**

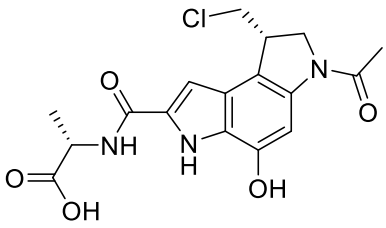
(57): 59 (3.00 g, 13.25 mmol) was dissolved in THF (4 mL). The resulting solution was treated with a 4.5 M aqueous solution of NaOH (1.60 mL) and left to stir at room temperature overnight. After removal of the THF under reduced pressure, the residue was diluted with 8 mL of H_2O. The solution was cooled to 0^oC. Acidification by the dropwise addition of concentrated HCl promoted the precipitation **57** which was collected by filtration. The precipitate was dried under vacuum, affording 0.81 g of **62** as a light yellow amorphous solid (49 % yield). ^1H NMR (DMSO-d_6, 400 MHz) δ 12.12 (1H, brs), 7.03 (1H, t, *J*=2.1) 6.78 (1H, dd, *J*= 3.9, 1.8), 6.05 (1H, dd, *J*=3.9, 2.5), 3.83 (3H, s). Consistent with the literature.[237]



62 **56**

(56): 62 (1.54 g, 3.68 mmol) was dissolved in anhydrous DCM (30 mL). The resulting solution was cooled to 0^oC before being treated with the dropwise addition of a 1 M solution of $TiCl_4$ in anhydrous DCM (8.00 mL) under N_2. After 45 mins of stirring at 0^oC the reaction mixture was quenched by the addition of a cold 1 M aqueous solution of HCl (100 mL). The resulting precipitate was collected by filtration, triturated in cold water, and dried under vacuum, affording 1.19 g of **56** as a white amorphous solid (89 % yield). ^1H NMR (DMSO-d_6, 400 MHz) δ 12.15 (1H, brs), 9.43 (1H, s), 7.90 (2H, appt d, *J*=7.4), 7.72 (2H, appt d, *J*=7.4), 7.42 (2H, appt t, *J*= 7.4), 7.34 (2H, appt t, *J*=7.4), 7.04 (1H, appt brs), 6.62 (1H, appt brs), 4.45 (2H, d, *J*= 6.3), 4.27 (1H, t, *J*=6.3), 3.78 (3H, s). Consistent with the literature.[235]

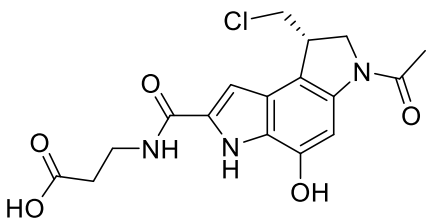
5.3 Solid phase synthesis.



35

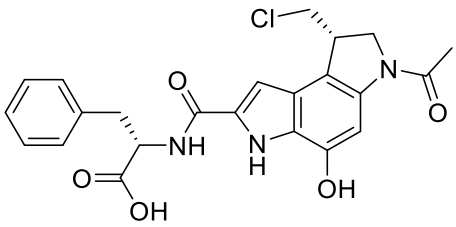
(35): H-Ala-2ClTrt resin (53.0 mg, 0.039 mmol Ala, [manufacturer's resin loading 0.73 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL) and 3 times with anhydrous DMF (10 mL). The resin was placed under an atmosphere of N_2 and treated with anhydrous DMF (2 mL), DIPEA (75 µL, 0.43 mmol), and AcCL (16 µL, 0.225 mmol). After 1 hour of shaking the resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 1% TFA, 10% TIPS in DCM (10 mL). After 2 hours of shaking the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in THF:MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude cleavage product was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (300 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Flash chromatography (silica gel, 7 x 1 cm, 0 % to 30 % MeOH in EtOAc) and trituration with hexane, afforded 10.0 mg of **35** as a beige amorphous solid (69 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 11.23 (1H, s), 9.70, (1H, s), 8.57 (1H, d, *J*=7.4), 7.70 (1H, s), 7.21 (1H, appt d, *J*=2.1), 4.44 (1H, appt quin, *J*=7.4), 4.33 (1H, appt t, *J*=11.7], 4.11-4.06 (1H, m), 4.04-3.96 (2H, m), 3.89-3.82 (1H, m), 2.15 (3H, s), 1.41 (3H, d, *J*=7.4). ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100

MHz) δ 104.0 (CH), 100.2 (CH), 53.3 (CH_2), 50.4 (CH), 47.7 (CH_2), 41.6 (CH), 24.2 (CH_3), 20.0 (CH_3). HRMS (ES+) calculated for $C_{17}H_{19}ClN_3O_5$ $(M+H)^+$ 380.1008 found 380.1004.



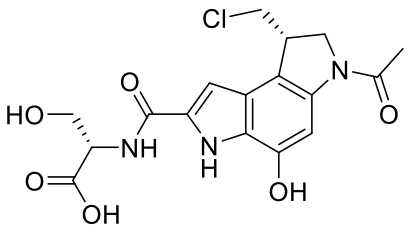
40

(40): H-β-Ala-2ClTrt resin (53.0 mg, 0.039 mmol β-Ala, [manufacturer's resin loading 0.73 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL) and 3 times with anhydrous DMF (10 mL). The resin was placed under an atmosphere of N_2 and treated with anhydrous DMF (2 mL), DIPEA (75 µL, 0.43 mmol), and AcCL (16 µL, 0.225 mmol). After 1 hour of shaking the resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 1% TFA, 10% TIPS in DCM (10 mL). After 2 hours of shaking the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude cleavage product was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (300 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Flash chromatography (silica gel, 7 x 1 cm, 0 % to 30 % MeOH in EtOAc) and trituration with hexane afforded 12.0 mg of **40** as a beige amorphous solid (81 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.29 (1H, brs), 11.18 (1H, s), 9.69 (1H, s), 8.45 (1H, brs), 7.69 (1H, s), 7.11 (1H, s), 4.35-4.28 (1H, m), 4.07-3.95 (3H, m), 3.87-3.81 (1H, m), 3.48 (2H, obscured by H_2O peak observed by HSQC and COSY), 2.53 (2H, obscured by DMSO peak observed by HSQC and COSY), 2.15 (3H, s). ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100 MHz) δ 101.4 (CH), 98.8 (CH), 53.3 (CH_2), 47.4 (CH_2), 41.8 (CH), 35.2 (CH_2), 33.9 (CH_2), 24.2 (CH_3). HRMS (ES+) calculated for $C_{17}H_{19}ClN_3O_5$ $(M+H)^+$ 380.1008 found 380.1009.



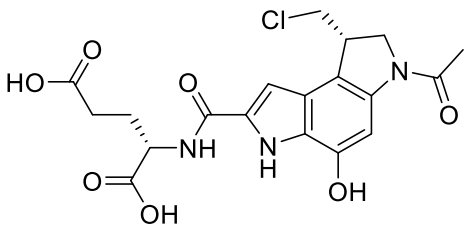
36

(36): H-Phe-2ClTrt resin (53.0 mg, 0.039 mmol Phe, [manufacturer's resin loading 0.73 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL) and 3 times with anhydrous DMF (10 mL). The resin was placed under an atmosphere of N_2 and treated with anhydrous DMF (2 mL), DIPEA (75 µL, 0.43 mmol), and AcCL (16 µL, 0.225 mmol). After 1 hour of shaking the resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 1% TFA, 10% TIPS in DCM (10 mL). After 2 hours of shaking the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in THF:MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude cleavage product was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (300 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Flash chromatography (silica gel, 7 x 1 cm, 0 % to 10 % MeOH in EtOAc) and trituration with hexane, afforded 11.0 mg of **36** as a beige amorphous solid (56 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.87 (1H, brs), 11.20 (1H, s), 9.70 (1H, s), 8.65 (1H, d, *J*=8.1), 7.70 (1H, s) 7.32-7.25 (4H, m), 7.21-7.16 (2H, m), 4.72-4.65 (1H, m), 4.36-4.27 (1H, m), 4.10-3.97 (3H, m), 3.89-3.82 (1H, m), 3.24-3.16 (1H, m), 3.06-2.98 (1H, m), 2.15 (3H, s) . ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100 MHz) δ 128.7 (CH), 128 (CH), 126.2 (CH), 102.0 (CH), 98.6 (CH), 53.5 (CH), 52.9 (CH_2), 47.1 (CH_2), 41.5 (CH), 36.5 (CH_2), 23.8 (CH_3). HRMS (ES+) calculated for $C_{23}H_{23}O_5N_3Cl$ $(M+H)^+$ 456.1321 found 456.1317.



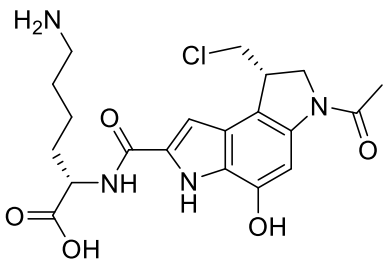
37

(37): H-Ser(tBu)-2ClTrt resin (51.0 mg, 0.039 mmol Ser, [manufacturer's resin loading 0.76 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL) and 3 times with anhydrous DMF (10 mL). The resin was placed under an atmosphere of N_2 and treated with anhydrous DMF (2 mL), DIPEA (75 µL, 0.43 mmol), and AcCl (16 µL, 0.225 mmol). After 1 hour of shaking the resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 20% TFA, 10% TIPS in DCM (10 mL). After shaking overnight the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in THF:MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude cleavage product was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (300 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Preparative HPLC (see general Prep HPLC method) and lyophilization, afforded 7.5 mg of **37** as a beige amorphous solid (48 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.76 (1H, brs), 11.35 (1H, s), 9.72 (1H, s), 8.49 (1H, d, *J*= 8.3), 7.71 (1H, s), 7.24 (1H, s), 5.01 (1H, brs), 4.58-4.48 (1H, m), 4.37-4.28 (1H, m), 4.14-3.97 (3H, m), 3.90-3.76 (3H, m), 2.15 (3H, s). ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100 MHz) δ 102.0 (CH), 98.3 (CH), 61.0 (CH_2), 55.0 (CH), 53.0 (CH_2), 47.2 (CH_2), 41.6 (CH), 23.8 (CH_3). HRMS (ES+) calculated for $C_{17}H_{19}ClN_3O_6$ $(M+H)^+$ 396.0957 found 396.0956.



38

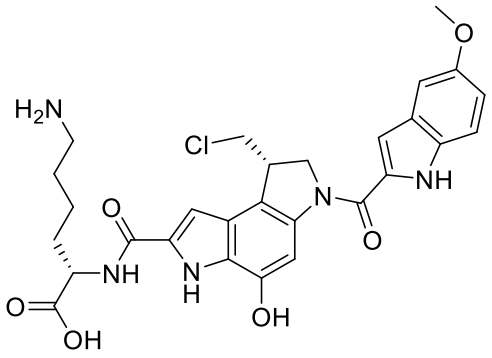
(38): H-Glu(OtBu)-2ClTrt resin (59.0 mg, 0.039 mmol Glu, [manufacturer's resin loading 0.65 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL) and 3 times with anhydrous DMF (10 mL). The resin was placed under an atmosphere of N_2 and treated with anhydrous DMF (2 mL), DIPEA (75 µL, 0.43 mmol), and AcCL (16 µL, 0.225 mmol). After 1 hour of shaking the resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 10% TFA, 10% TIPS in DCM (10 mL). After shaking for 6 hours the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in THF:MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude cleavage product was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (300 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Preparative HPLC (see general Prep HPLC method) and lyophilization, afforded 6.5 mg of **38** as a beige amorphous solid (38 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.51 (2H, brs), 11.25 (1H, s), 9.73 (1H, s), 8.52 (1H, brs), 7.70 (1H, s), 7.21 (1H, s), 4.49-4.30 (2H, m), 4.18-3.93 (3H, m), 3.91-3.80 (1H, m), 2.43-2.35 (2H, m), 2.21-2.04 (4H, m), 1.97-1.86 (1H, m). ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100 MHz) δ 102.2 (CH), 98.8 (CH), 53.1 (CH_2), 51.3 (CH), 47.2 (CH_2), 41.7 (CH), 30.1 (CH_2), 26.1 (CH_2), 23.9 (CH_3). HRMS (ES+) calculated for $C_{19}H_{21}ClN_3O_7$ 438.1063 $(M+H)^+$ found 438.1053.



39

(39): H-Lys(Boc)-2ClTrt resin (53.0 mg, 0.039 mmol Lys, [manufacturer's resin loading 0.73 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL) and 3 times with anhydrous DMF (10 mL). The resin was placed under an atmosphere of N_2 and treated with anhydrous DMF (2 mL), DIPEA (75 µL, 0.43 mmol), and AcCL (16 µL, 0.225 mmol). After 1 hour of shaking the resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 10% TFA, 10% TIPS in DCM (10 mL). After shaking for 2 hours the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in THF:MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude cleavage product was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (300 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Preparative HPLC (see general Prep HPLC method) and lyophilization, afforded 3.3 mg of **39** as a beige amorphous solid (19 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.79 (1H brs), 11.28 (1H, s), 9.79 (1H, s), 8.54 (1H, d, *J*=8.3), 7.73-7.69 (2H, brs), 7.71 (1H, s), 7.22 (1H, s), 4.46-4.38 (1H, m), 4.37-4.28 (1H, m), 4.14-3.95 (3H, m), 3.91-3.83 (1H, m), 2.82-2.76 (2H, m), 2.16 (3H, s), 1.86-1.74 (2H m), 1.60-1.54 (2H, m), 1.47-1.41 (2H m). ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100 MHz) δ 102.2 (CH), 98.8 (CH), 53.0 (CH_2), 51.6 (CH), 47.2 (CH_2), 41.5

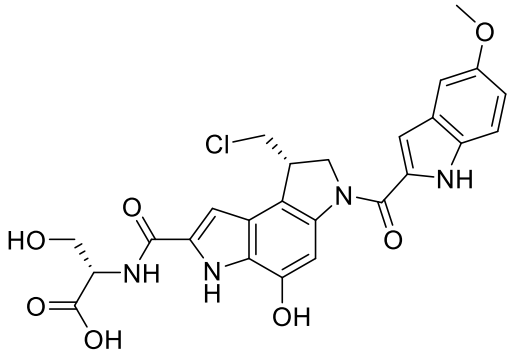
(CH), 38.4 (CH_2), 30.3 (CH_2), 26.3 (CH_2), 24.0 (CH_3), 22.5 (CH_2). HRMS (ES+) calculated for $C_{20}H_{26}ClN_4O_5$ 437.1586 $(M+H)^+$ found 437.1593.



44

(44): H-Lys(Boc)-2ClTrt resin (53.0 mg, 0.039 mmol Lys, [manufacturer's resin loading 0.73 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). 5-methoxyindole-2-carboxylic acid (38.0 mg, 0.199 mmol) was disolved in 2 mL of DMF and treated with HATU (73.0 mg, 0.191 mmol) and DIPEA (70 µL, 0.401 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 10% TFA, 10% TIPS in DCM (10 mL). After shaking for 2 hours the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in THF:MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude cleavage product was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (30 mg) in a 25 % aqueous solution ammonium formate (500 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Preparative HPLC (see general Prep HPLC method) and lyophilization, afforded 4.4 mg of **44** as a beige amorphous solid (20 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 11.56, (1H, s), 11.38 (1H, brs), 9.89 (1H, brs), 8.59 (1H brs), 7.76 (1H, brs), 7.59 (2H brs), 7.38 (1H, d, *J*=8.9), 7.26 (1H, s), 7.15 (1H, d, *J*= 2.2), 7.02 (1H, d, *J*=1.4), 6.89 (1H, dd, *J*=8.9, 2.2),

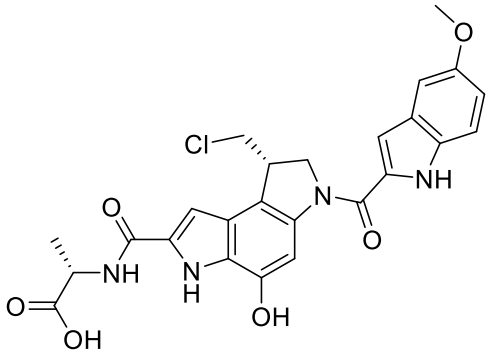
4.82-4.67 (1H, m), 4.51-4.39 (2H, m), 4.18-4.03 (2H, m), 3.99-3.91 (1H, m), 3.78 (3H, s), 2.84-2.75 (2H, m), 1.92-1.74 (2H, m), 1.64-1.55 (2H, m), 1.49-1.42 (2H, m). ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100 MHz) δ 114.8 (CH), 112.8 (CH), 104.5 (CH), 102.4 (CH), 101.8 (CH), 99.5 (CH), 54.9 (CH_3), 54.4 (CH_2), 51.8 (CH), 47.1 (CH_2), 42.0 (CH), 38.4 (CH_2), 3.3 (CH_2), 26.4 (CH_2), 22.4 (CH_2). HRMS (ES+) calculated for $C_{28}H_{31}ClN_5O_6$ 568.1957 $(M+H)^+$ found 568.1949.



47

(47): H-Ser(tBu)-2ClTrt resin (56.0 mg, 0.039 mmol Lys, [manufacturer's resin loading 0.76 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). 5-methoxyindole-2-carboxylic acid (38.0 mg, 0.199 mmol) was disolved in 2 mL of DMF and treated with HATU (73.0 mg, 0.191 mmol) and DIPEA (70 µL, 0.401 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 20% TFA, 10% TIPS in DCM (10 mL). After shaking overnight the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in THF:MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude cleavage product was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (300 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Preparative HPLC (see general Prep HPLC method) and lyophilization, afforded 3.3 mg of

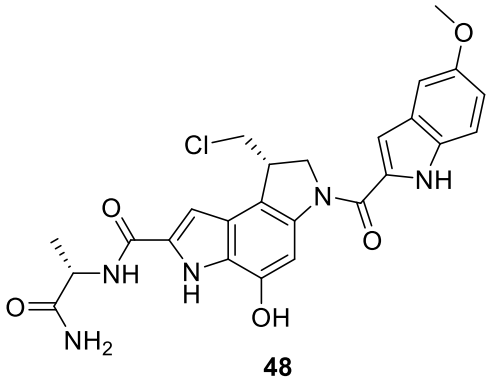
47 as a beige amorphous solid (16 % yield). ^{1}H NMR (DMSO- d_6, 400 MHz) δ 11.54 (1H, s), 11.43 (1H, s), 9.80 (1H, s), 8.52 (1H, d, *J*=7.4), 7.77 (1H, brs), 7.38 (1H, d, *J*= 8.6), 7.29 (1H, d, *J*=1.4), 7.15 (1H, d, *J*=2.3), 7.02 (1H, d, *J*=1.3), 7.89 (1H, dd, *J*= 8.6, 2.3), 4.79-4.71 (1H, m), 4.57-4.50 (1H, m), 4.47-4.37 (1H, m), 4.17-4.04 (2H, m), 3.99-3.89 (1H, m), 3.82-3.76 (2H, m), 3.78 (3H, s). ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100 MHz) δ 114.6 (CH), 112.8 (CH), 104.2 (CH), 102.4 (CH), 101.7 (CH), 99.5 (CH), 61.1 (CH_2), 55.0 (CH_3), 54.8 (CH), 54.4 (CH_2), 47.0 (CH_2), 42.0 (CH). HRMS (ES-) calculated for $C_{25}H_{22}ClN_4O_7$ 525.1182 $(M-H)^-$ found 525.1187.



46

(46): H-Ala-2ClTrt resin (58.0 mg, 0.039 mmol Ala, [manufacturer's resin loading 0.72 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). 5-methoxyindole-2-carboxylic acid (38.0 mg, 0.199 mmol) was disolved in 2 mL of DMF and treated with HATU (73.0 mg, 0.191 mmol) and DIPEA (70 µL, 0.401 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 1% TFA, 10% TIPS in DCM (10 mL). After shaking for 2 hours the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in THF:MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude cleavage product was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (300 µL) under N_2.

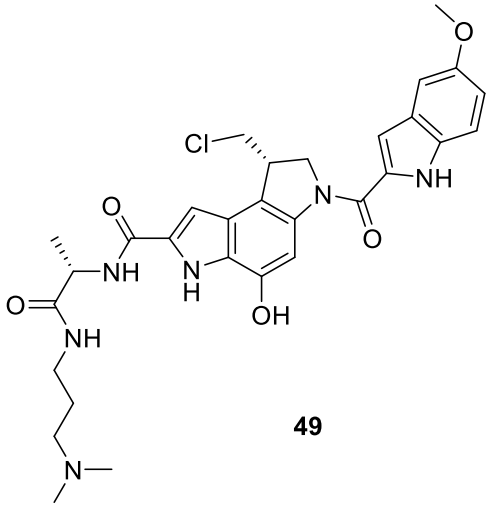
After 1 hour the Pd/C was removed by filtering through a plug of celite. Reverse phase flash chromtography (see general reverse phase flash chromotography method) and lyophilization, afforded 7.0 mg of **46** as a beige amorphous solid (35 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.67 (1H, Brs), 11.5 (1H, s), 11.32 (1H, s), 9.82 (1H, s), 8.62, (1H, d, *J*=7.0), 7.76 (1H, brs), 7.38 (1H, d, *J*=8.9), 7.27 (1H, s), 7.15 (1H, s), 7.02 (1H, s), 6.89 (1H, d, *J*=8.9), 4.81-4.71 (1H, m), 4.50-4.39 (2H, m), 4.16-4.06 (2H, m), 3.98-3.91 (1H, m), 3.78 (3H, s), 1.42 (3H, d, *J*=7.3). ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100 MHz) δ 114.7 (CH), 112.7 (CH), 104.4 (CH), 102.2 (CH), 101.8 (CH), 99.5 (CH), 54.9 (CH_3), 54.4 (CH_2), 47.5 (CH), 46.9 (CH_2), 42.0 (CH), 16.9 (CH_3). HRMS (ES-) calculated for $C_{25}H_{22}ClN_4O_6$ 509.1233 $(M-H)^-$ found 509.1234.



48

(48): Rink amide MBHA resin (107.0 mg, 0.039 mmol, [manufacturer's resin loading 0.36 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins, and treatment with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). The resin was washed 6 times with DMF (10 mL). Fmoc-Ala-OH (121.0 mg, 0.39 mmol), was disolved in 2 mL of DMF and treated with HBTU (133.0 mg, 0.39 mmol), HOBt.H_2O (54.0 mg, 0.39 mmol) and DIPEA (135 µL, 0.78 mmol). After 30 secs the solution was added to the resin and shaken for 45 mins. The coupling was repeated and the resin washed 6 times with DMF (10 mL) prior to Fmoc deprotection with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was wash a further 6 times with DMF (10 mL). **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). 5-methoxyindole-2-carboxylic acid (38.0 mg, 0.199 mmol) was disolved in 2 mL of DMF and treated with HATU (73.0 mg, 0.191 mmol) and DIPEA (70 µL, 0.401 mmol). After 10 secs the resulting solution was added to the resin and the

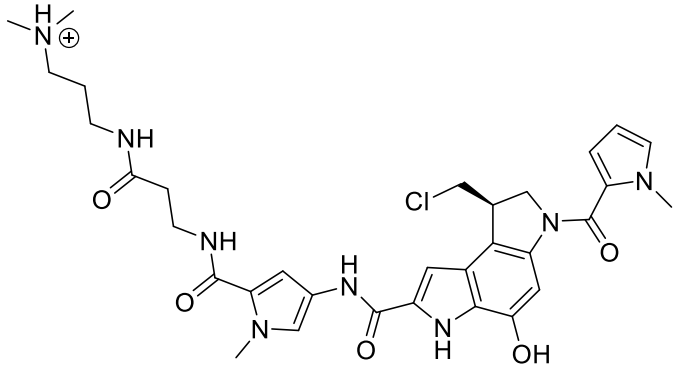
mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 47 % TFA, 47 % DCM, 3 % TIPS and 3 % H_2O (10 mL). After shaking for 2 hours the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. The crude cleavage product was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (300 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Flash chromatography (silica gel, 7 x 1 cm, 10 % MeOH in DCM) and trituration with hexane, afforded 6.8 mg of **48** as a beige amorphous solid (34 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 11.54 (1H, s), 11.35 (1H, s), 9.82 (1H, s), 8.46 (1H, d, *J*=7.6), 7.76 (1H, brs), 7.47 (1H, brs), 7.38 (1H, d, *J*= 8.9), 7.27 (1H, d, *J*= 2.0), 7.15 (1H, d, *J*=2.3), 7.04 (1H, brs), 7.02 (1H, s), 6.89 (1H, dd, *J*= 8.9, 2.3), 4.80-4.72 (1H, m), 4.51-4.40 (2H, m), 4.16-4.06 (2H, m), 3.98-3.89 (1H, m), 3.78 (3H, s), 1.35 (3H, d, *J*=7.1). ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100 MHz) δ 114.7 (CH), 112.8 (CH), 104.4 (CH), 102.3 (CH), 101.8 (CH), 99.3 (CH), 54.9 (CH_3), 54.4 (CH_2), 48.0 (CH), 47.1 (CH_2), 41.9 (CH) 18.0 (CH_3). HRMS (ES+) calculated for $C_{25}H_{25}ClN_5O_5$ 510.1539 $(M+H)^+$ found 510.1533.



49

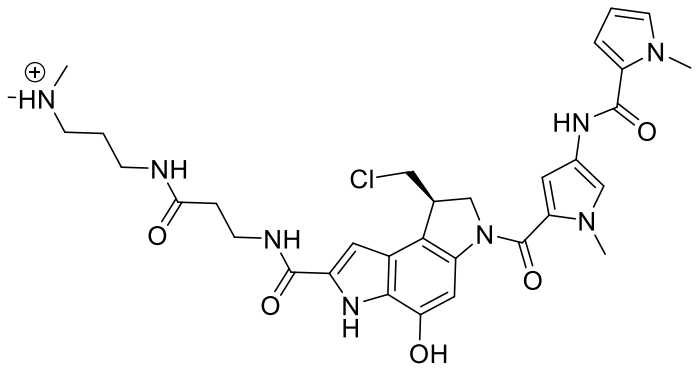
(49): H-Ala-2ClTrt resin (58.0 mg, 0.039 mmol Ala, [manufacturer's resin loading 0.72 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (25.0 mg, 0.043 mmol) was disolved in 2 mL of DMF and treated with HATU (13.0 mg, 0.043 mmol) and DIPEA (16 µL, 0.086 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). 5-methoxyindole-2-

carboxylic acid (38.0 mg, 0.199 mmol) was disolved in 2 mL of DMF and treated with HATU (73.0 mg, 0.191 mmol) and DIPEA (70 µL, 0.401 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and 6 times with DCM (10 mL). Cleavage was affected by addition of a solution of 1% TFA, 10% TIPS in DCM (10 mL). After shaking for 2 hours the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in THF:MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude was disolved in DMF (1 mL), and treated with HATU (16.0 mg, 0.042 mmol), and DIPEA (20 µL, 0.12 mmol). After 10 secs the resulting solution was treated with 3-(Dimethylamino)-1-propylamine (30 µL, 0.24 mmol) and stirred at room temperture for 2 hours, prior to preciptation with cold H_2O (15 mL), and collection by centrifugation. The precipate was was disolved in THF:MeOH (2 mL) and treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (300 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Reverse phase flash chromtography (see general reverse phase flash chromotography method) and lyophilization, afforded 4.0 mg of **49** as a beige amorphous solid (17 % yield). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 11.54 (1H, s), 11.37 (1H, s), 9.91 (1H, s), 9.33 (1H, brs), 8.56 (1H, appt t, *J*=7.3), 8.21 (1H, appt q, *J*=5.6), 7.79 (1H, brs), 7.39 (1H, d, *J*=8.7), 7.28 (1H, s), 7.16 (1H, d, *J*=2.3), 7.03 (1H, s), 6.90 (1H, dd, *J*= 8.7, 2.3), 4.81-4.72 (1H, m), 4.47-4.39 (2H, m), 4.16-4.08 (2H, m), 3.99-3.91 (1H, m), 3.78 (3H, s), 3.21-3.12 (2H, m), 3.08-2.99 (2H, m), 2.76 (6H, s), 1.84-1.75 (2H, m), 1.38 (3H, d, *J*=7.1). ^{13}C NMR (observed by DEPT-ed-HSQC) (DMSO-d_6, 100 MHz) δ 114.8 (CH), 112.8 (CH), 104.4 (CH), 102.6 (CH), 101.8 (CH), 99.4 (CH), 54.9 (CH_3), 54.5 (CH_2), 54.3 (CH_2), 48.7 (CH), 47.1 (CH_2), 42.2 (CH_3), 42.0 (CH), 35.2 (CH_2), 24.1 (CH_2), 17.6 (CH_3). HRMS (ES+) calculated for $C_{30}H_{36}ClN_6O_5$ 595.2430 $(M+H)^+$ found 595.2418.



63

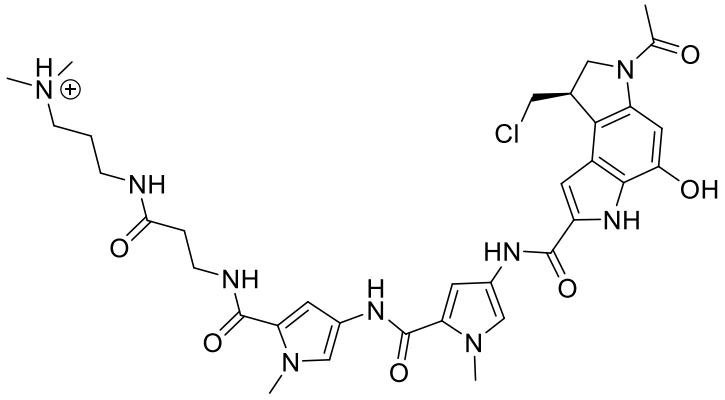
(63): H-β-Ala-2ClTrt resin (106.0 mg, 0.078 mmol β-Ala, [manufacturer's resin loading 0.73 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **56** (141.0 mg, 0.39 mmol) was disolved in 2 mL of DMF and treated with HATU (148.0 mg, 0.39 mmol) and DIPEA (135 µL, 0.78 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken for 6 hours. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). **11** (68.0 mg, 0.117 mmol) was disolved in 2 mL of DMF and treated with HATU (44.0 mg, 0.117 mmol) and DIPEA (40 µL, 0.229 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). **57** (48.0 mg, 0.39 mmol) was disolved in 2 mL of DMF and treated with HATU (148.0 mg, 0.39 mmol) and DIPEA (135 µL, 0.78 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken for 5 hours. This coupling was then repeated on the same scale with freash reagents for a further 5 hours. The resin was washed 6 times with DMF (10 mL) and 6 times with DCM (10 mL). Cleavage was affected by addition of a solution of 1% TFA, 10% TIPS in DCM (10 mL). After shaking for 2 hours the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude was disolved in DMF (1 mL), and treated with HATU (30.0 mg, 0.078 mmol), and DIPEA (30 µL, 0.156 mmol). After 10 secs the resulting solution was treated with 3-(Dimethylamino)-1-propylamine (49 µL, 0.39 mmol) and stirred at room temperture. After 2 hours the reaction mixture was subjected directly to reverse phase preparative HPLC (see general Prep HPLC method). Fractions containing the the desired product, were concentrated, and redisolved in MeOH (1.5 mL). The solution was treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (500 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Reverse phase preparative HPLC (see general Prep HPLC method)and lyophilization, afforded 4.7 mg of **63** as a tan amorphous solid (9 % yield). HRMS (ES+) calculated for $C_{32}H_{40}O_5N_8Cl$ 651.2805 $(M+H)^+$ found 651.2795.



64

64: H-β-Ala-2ClTrt resin (59.0 mg, 0.043 mmol β-Ala, [manufacture's resin loading 0.73 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **11** (38.0 mg, 0.066 mmol) was disolved in 2 mL of DMF and treated with HATU (24.0 mg, 0.066 mmol) and DIPEA (23 µL, 0.13 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). **56** (79.0 mg, 0.21 mmol) was disolved in 2 mL of DMF and treated with HATU (82.0 mg, 0.21 mmol) and DIPEA (75 µL, 0.44 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken for 5 hours. This coupling was then repeated on the same scale with freash reagents for a further 5 hours. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). **57** (28.0 mg, 0.21 mmol) was disolved in 2 mL of DMF and treated with HATU (82.0 mg, 0.21 mmol) and DIPEA (75 µL, 0.44 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken for 5 hours. The resin was washed 6 times with DMF (10 mL) and 6 times with DCM (10 mL). Cleavage was affected by addition of a solution of 1% TFA, 10% TIPS in DCM (10 mL). After shaking for 2 hours the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude was disolved in DMF (1 mL), and treated with HATU (17.0 mg, 0.043

mmol), and DIPEA (16.5 µL, 0.086 mmol). After 10 secs the resulting solution was treated with 3-(Dimethylamino)-1-propylamine (27 µL, 0.22 mmol) and stirred at room temperture. After 2 hours the reaction mixture was subjected directly to reverse phase preparative HPLC (see general Prep HPLC method). Fractions containing the the desired product, were concentrated, and redisolved in MeOH (1.5 mL). The solution was treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (500 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Reverse phase flash preparative HPLC (see general Prep HPLC method)and lyophilization, afforded 2.2 mg of **64** as a tan amorphous solid (8 % yield). HRMS (ES+) calculated for $C_{32}H_{40}O_5N_8Cl$ 651.2805 $(M+H)^+$ found 651.2796.



65

65: H-β-Ala-2ClTrt resin (59.0 mg, 0.043 mmol β-Ala, [manufacture's resin loading 0.73 mmol/g]) was prepared for coupling by swelling in DCM for 30 mins followed by DMF for a further 30 mins. **56** (79.0 mg, 0.21 mmol) was disolved in 2 mL of DMF and treated with HATU (82.0 mg, 0.21 mmol) and DIPEA (75 µL, 0.44 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken for 5 hours. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). **56** (79.0 mg, 0.21 mmol) was disolved in 2 mL of DMF and treated with HATU (82 mg, 0.21 mmol) and DIPEA (75 µL, 0.44 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken for 5 hours. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection affected with piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL). **11** (38.0 mg, 0.066 mmol) was disolved in 2 mL of DMF and treated with HATU (24 mg, 0.066 mmol) and DIPEA (23 µL, 0.13 mmol). After 10 secs the resulting solution was added to the resin and the mixture shaken overnight. The resin was washed 6 times with DMF (10 mL) and removal of the Fmoc protection of the indoline nitrogen affected with

piperidine in DMF (3 mL 40 % 10 mins, 3 mL 20 % 5 mins twice). Following Fmoc deprotection the resin was washed 6 times with DMF (10 mL) and 3 times with anhydrous DMF (10 mL). The resin was placed under an atomsphere of N_2 and treated with anhydrous DMF (2 mL), DIPEA (75 µL, 0.43 mmol), and AcCL (16 µL, 0.225 mmol). After 1 hour of shaking the resin was washed 6 times with DMF (10 mL) and 6 times with DCM. Cleavage was affected by addition of a solution of 1% TFA, 10% TIPS in DCM (10 mL). After shaking for 2 hours the cleavage mixture was flitered. The resin was rinsed 3 times with DCM (3 mL) and the combined filtrates were concentrated to dryness by rotary evaporation under vacuum. To ensure full recovery of the product the resin was soaked in MeOH (10 mL), and after filtering this was combined with the rest of the cleavage product and again evaporated to dryness. The crude was disolved in DMF (1 mL), and treated with HATU (17.0 mg, 0.043 mmol), and DIPEA (16.5 µL, 0.086 mmol). After 10 secs the resulting solution was treated with 3-(Dimethylamino)-1-propylamine (27 µL, 0.22 mmol) and stirred at room temperture. After 2 hours the reaction mixture was subjected directly to reverse phase preparative HPLC (see general Prep HPLC method). Fractions containing the the desired product, were concentrated, and redisolved in MeOH (1.5 mL). The solution was treated with a slurry of 10 % Pd/C (20 mg) in a 25 % aqueous solution ammonium formate (500 µL) under N_2. After 1 hour the Pd/C was removed by filtering through a plug of celite. Reverse phase flash preparative HPLC (see general Prep HPLC method) and lyophilization, afforded 2.0 mg of **65** as a tan amorphous solid (7 % yield). HRMS (ES+) calculated for $C_{34}H_{43}O_6N_9Cl$ 708.3019 $(M+H)^+$ found 708.3014.

5.4 Biological assays.

5.4.1 MTS assay.

Cell culture: The HL60 cell line was purchased from ECACC (Porton Down, UK). Cells were cultured in RPMI 1640 medium supplemented with 10 % foetal calf serum and 2 mM L-glutamine. HL-60 cells were passaged twice weekly and maintained between 1-9 x 10^5 cells/ml at 37 ^oC and 5 % CO_2.

Antiproliferative assay: Antiproliferative activity was determined by MTS assay using the CellTiter 96 Aq_{ueous} One Solution Cell Proliferation Assay (Promega) and following the manufacturer's instructions. Briefly, HL-60 cells (3 x 10^4/100 µl) were seeded in 96-well plates and left untreated or treated with DMSO (vehicle control), duocarmycins, or doxorubicin hydrochloride at 8 concentrations (see below) in triplicate for 72 hr at 37 °C with 5 % CO_2. Following this, MTS assay reagent was added for 4 hrs and the absorbance

measured at 490 nm using the Polarstar Optima microplate reader (BMG Labtech). IC_{50} values were calculated using GraphPad Prism Version 5.0 software.

Concentrations tested: Compounds **35**, **36**, **37**, **38**, **39**, **40**, **42**, **43**, **44**, and **47**, (500 µM, 250 µM, 100 µM, 10 µM, 1 µM, 0.1 µM, 0.01 µM, 0.001 µM). Compounds **46**, **48**, **49**, **63**, **64, 65**, and doxorubicin (100 µM, 10 µM, 1 µM, 0.1 µM, 0.01 µM, 0.001 µM, 0.0001 µM, 0.00001 µM)

5.4.2 DNA alkylation (thermal cleavage assay).

DNA cleavage**:** The MS1 DNA fragment was prepared as previously described,[188] by cleaving the parent plasmid with HindIII and SacI and labelling the 3'-end of the HindIII site with α-^{32}P[dATP] using Klenow fragment (exo-). 1.5 µL of each compound (diluted in 10 mM Tris-HCl pH 7.5, containing 10 mM NaCl) was incubated with 1.5 µL of the radiolabelled DNA and incubated overnight at 37 °C. The samples were then mixed with an equal volume of formamide containing 10 mM EDTA and ligand specific cleavage was induced by heating at 100 °C for 3 minutes. Samples were loaded onto 8% denaturing polyacrylamide gels containing 8M urea. The dried gel was exposed to a phosphorimager screen and analysed using a Typhoon phosphorimager.

5.5 DFT calculations.

5.5.1 General information for DFT calculations.

DFT geometry optimisations (in the gas phase) were performed on the UEA Grace high performance computational cluster using 12 dedicated nodes to run Gaussian09 employing the B3YLP density functional and the 6-31G(d) basis set. Optimised structures were confirmed as true minima by the absence of virtual (i.e. negative) vibrational modes in frequency calculations and transition states were confirmed by the presence of one negative vibrational frequency which was inspected (Gaussview05) to confirm that it corresponded to the approach of a methanol molecule to the spirocyclic cyclopropane. Activation energies were calculated by subtracting the sum of the individual energies of the optimised conformations of the reactants (compound 45 and methanol) from the energy of the transition state.

The transition state calculations were performed using the command line:

b3lyp/6-31G* opt=(TS,NOEIGENTEST,CALCALL,TIGHT)

For geometry optimisations, the corresponding command line was:

b3lyp/6-31G* opt scf=tight test gfinput iop(6/7=3)

For frequency calculations, the corresponding command line was:

b3lyp/6-31G* freq scf=tight test gfinput iop(6/7=3)

For single point calculations, the corresponding command line was:

b3lyp/6-31G* sp scf=tight test gfinput iop(6/7=3)

The reaction pathway presented in Figure 3.19 was calculated from the checkpoint file generated by a frequency calculation, using the commands:

b3lyp/6-31G* IRC=(RCFC,FORWARD,MAXPOINTS=800,MAXCYCLE=800,STEPSIZE=30) GUESS=READ GEOM=CHECK

and

b3lyp/6-31G* IRC=(RCFC,REVERSE,MAXPOINTS=800,MAXCYCLE=800, STEPSIZE=30) GUESS=READ GEOM=CHECK sometimes in combination with an allocation of additional memory (%mem=1gb).

5.5.2 Selected bond lengths, angles and atomic coordinates for single point and optimised structures.

Figure 3.18, CONFORMATION a): –1273.6307249 Ha.

Atom	NA	NB	NC	Bond	Angle	Dihedral	X	Y	Z
C							3.0931269	2.1013942	-1.0235569
C	1			1.54			2.3224542	1.0516837	-0.1999654
C	2	1		1.50	114.2		2.9551139	-0.3028422	-0.1865095
N	3	2	1	1.41	107.6	51.5	4.3026674	-0.1392725	0.1854903
C	4	3	2	1.48	111.2	18.8	4.5203933	1.1667622	0.8529961
C	1	2	3	1.49	60.4	-95.0	3.3234712	2.0053439	0.4447055
C	2	1	6	1.47	120.0	121.2	0.8525363	1.0661213	-0.1821251

C	7	2	1	1.39	116.3	129.1	0.2227701	-0.1387116	-0.4600261
C	8	7	2	1.45	127.8	9.4	0.8308079	-1.4548304	-0.5823534
C	3	2	1	1.36	124.3	-131.7	2.2866199	-1.4653514	-0.426185
C	7	2	1	1.42	136.7	-47.0	-0.166843	2.0337794	-0.0120132
C	11	7	2	1.40	107.4	177.0	-1.3924872	1.3831677	-0.1772668
N	8	7	2	1.36	108.6	-177.5	-1.1201231	0.0627699	-0.4573155
C	4	3	2	1.40	125.0	-169.7	5.2793265	-1.1355719	0.1161763
C	14	4	3	1.52	116.0	-176.4	6.6523314	-0.7617757	0.6491832
O	9	8	7	1.24	120.5	173.2	0.1216918	-2.4662591	-0.7119776
C	12	11	7	1.48	125.8	174.6	-2.7318011	1.9892895	0.0266472
N	17	12	11	1.38	120.2	-163.8	-3.8463747	1.1832475	0.1607245
C	18	17	12	1.44	131.1	-6.6	-4.0088936	-0.2443136	0.019111
C	19	18	17	1.56	114.4	73.9	-3.4219661	-1.0791236	1.1967967
O	20	19	18	1.43	110.0	-164.7	-3.3077089	-2.4533839	0.8138669
O	17	12	11	1.23	119.5	13.7	-2.8343162	3.2082341	0.1203148
O	14	4	3	1.22	122.0	4.0	5.0472807	-2.2366301	-0.3520174
C	19	18	17	1.54	108.7	-163.8	-5.5114585	-0.5580264	-0.1241718
O	24	19	18	1.21	123.1	3.9	-6.3607316	0.298217	-0.1813477
O	24	19	18	1.34	114.2	-176.2	-5.7968849	-1.8664368	-0.18351
H	10	3	2	1.08	121.5	179.7	2.7907699	-2.4203473	-0.4101745
H	6	1	2	1.09	120.7	-111.0	2.9927355	2.7957195	1.1107969
H	5	4	3	1.10	110.2	100.2	4.5568589	1.0349094	1.9424099
H	5	4	3	1.09	110.6	-140.0	5.4651304	1.6116424	0.5278457

H	15	14	4	1.10	112.0	57.0	6.6115483	-0.4292473	1.6926182
H	15	14	4	1.09	107.5	176.9	7.2829862	-1.6484689	0.5796757
H	15	14	4	1.10	111.5	-63.8	7.1061023	0.0442326	0.0606636
H	1	6	5	1.09	118.0	-160.2	2.4834925	2.9204183	-1.392079
H	1	6	5	1.09	119.4	-11.7	3.8631028	1.7541299	-1.7083506
H	13	8	7	1.01	120.9	173.0	-1.7530815	-0.6660511	-0.7403314
H	11	7	2	1.08	129.0	-5.2	-0.0751378	3.0816815	0.2365104
H	18	17	12	1.02	113.7	-174.4	-4.7210581	1.7002526	0.2049921
H	19	18	17	1.10	111.7	-49.6	-3.5668398	-0.6129885	-0.9209048
H	26	24	19	0.99	110.0	6.2	-4.9817957	-2.397455	-0.0232775
H	20	19	18	1.09	110.7	-41.0	-2.4609211	-0.6717835	1.5167763
H	20	19	18	1.09	109.8	77.9	-4.1032228	-1.0419732	2.0517773
H	21	20	19	0.97	109.8	90.2	-2.431516	-2.6137855	0.4174663

Figure 3.18, CONFORMATION b): –1273.6404232 Ha.

Atom	NA	NB	NC	Bond	Angle	Dihedral	X	Y	Z
C							-0.3026237	1.913281	0.2823304
C	1			1.42			0.7471881	0.9694529	0.1849251
C	2	1		1.39	107.1		0.2054633	-0.2786238	0.4888543
N	3	2	1	1.36	108.5	-0.4	-1.11292	-0.1190626	0.7706576
C	4	3	2	1.37	109.4	0.5	-1.448193	1.2008596	0.6472823
C	2	1	5	1.47	135.6	173.5	2.1489534	0.9993502	-0.257482
C	6	2	1	1.49	115.6	175.8	2.9085551	-0.2745668	-0.0785996
C	7	6	2	1.36	124.0	15.5	2.3259804	-1.479008	0.1621888

C	3	2	1	1.46	126.8	173.7	0.8779628	-1.5680568	0.3934124
C	6	2	1	1.53	129.7	32.0	3.1725557	2.1284904	-0.1804178
C	10	6	2	1.52	106.7	145.1	4.519583	1.4785837	0.0744852
N	7	6	2	1.41	107.7	-162.1	4.2826615	0.0326273	-0.1507733
C	10	6	2	1.49	61.3	-105.1	2.5575889	1.8096894	-1.5000998
O	9	3	2	1.24	121.6	-171.1	0.2597815	-2.6346729	0.5307587
C	12	7	6	1.39	125.1	-170.2	5.3106915	-0.9065032	-0.2249641
C	15	12	7	1.52	116.0	-175.8	6.7252408	-0.3500853	-0.195588
C	5	4	3	1.47	121.4	178.1	-2.8328335	1.6645914	0.8536967
N	17	5	4	1.40	116.6	-20.2	-3.8389665	0.6981408	0.7467916
C	18	17	5	1.46	126.5	-39.1	-3.8961874	-0.3954414	-0.2213846
C	19	18	17	1.56	113.7	124.2	-4.0264405	-1.8063166	0.4328856
O	20	19	18	1.42	113.9	-65.6	-2.8904339	-2.1939613	1.1975499
O	15	12	7	1.22	122.2	4.6	5.0919905	-2.1027243	-0.3169708
O	17	5	4	1.22	122.3	157.0	-3.0966457	2.8209689	1.1556924
C	19	18	17	1.53	108.7	-115.8	-5.1121463	-0.1853865	-1.1252857
O	24	19	18	1.21	124.9	-17.3	-6.0439668	0.5441407	-0.8740357
O	24	19	18	1.35	111.6	164.8	-5.0467374	-0.9564481	-2.2327298
H	8	7	6	1.08	121.6	177.0	2.9087976	-2.38516	0.238906
H	10	6	2	1.09	120.9	5.5	2.9404594	3.0577502	0.3298531
H	11	10	6	1.10	111.0	-107.3	4.8522491	1.653304	1.1062819
H	11	10	6	1.09	112.1	131.4	5.2921603	1.852135	-0.6040742
H	16	15	12	1.10	112.0	56.7	6.9087851	0.2578317	0.6975365

H	16	15	12	1.09	107.5	176.7	7.4109469	-1.1977891	-0.200231
H	16	15	12	1.10	111.5	-64.0	6.9287219	0.2769315	-1.0716387
H	13	10	6	1.09	118.0	103.8	1.7873383	2.4755055	-1.876821
H	13	10	6	1.09	119.4	-108.0	3.157401	1.3190172	-2.2631394
H	4	3	2	1.02	124.9	-171.9	-1.7266545	-0.851698	1.1402582
H	1	5	4	1.08	124.6	179.3	-0.2733303	2.9792793	0.1031473
H	18	17	5	1.02	110.0	-179.2	-4.7501814	1.1276708	0.8928868
H	19	18	17	1.09	110.3	1.9	-3.0062546	-0.3858416	-0.8563509
H	26	24	19	0.98	106.7	177.7	-5.8683278	-0.7955009	-2.7355258
H	20	19	18	1.09	108.0	52.2	-4.8706575	-1.7806756	1.1290542
H	20	19	18	1.10	109.0	170.0	-4.2421813	-2.5437746	-0.3485907
H	21	20	19	0.98	111.3	-91.0	-2.2446349	-2.6812106	0.6521617

Figure 3.18, CONFORMATION c): –1273.6351730 Ha.

Atom	NA	NB	NC	Bond	Angle	Dihedral	X	Y	Z
C							-0.8194998	0.021947	0.317245
C	1			1.41			0.5659305	0.2789649	0.2367248
C	2	1		1.39	107.3		0.7256464	1.6557413	0.104075
N	3	2	1	1.36	108.3	-0.1	-0.5034954	2.228542	0.0992532
C	4	3	2	1.37	110.0	-0.4	-1.4725035	1.2618926	0.2172448
C	2	1	5	1.47	134.5	175.2	1.7760468	-0.5461246	0.3616886
C	6	2	1	1.49	115.4	172.3	3.0557457	0.1403527	0.0195973
C	7	6	2	1.36	123.8	14.9	3.1822358	1.4873487	-0.099022
C	3	2	1	1.47	125.9	-179.8	2.0059378	2.3657122	-0.018714

C	6	2	1	1.53	129.9	28.4	1.9838704	-2.0279258	0.0659464
C	10	6	2	1.52	106.6	145.1	3.3944989	-2.1682484	-0.4762191
N	7	6	2	1.41	107.6	-162.3	4.0270185	-0.855142	-0.2069908
C	10	6	2	1.49	61.3	-104.9	1.8943443	-1.5881031	1.4876148
O	9	3	2	1.23	121.5	-178.8	2.0546132	3.5931224	-0.1102268
C	12	7	6	1.40	125.0	-170.2	5.3937914	-0.6191609	-0.3612073
C	15	12	7	1.52	116.0	-175.9	6.2464742	-1.8331753	-0.692165
C	5	4	3	1.48	115.2	177.4	-2.8577024	1.766807	0.2856425
N	17	5	4	1.39	121.0	164.8	-3.9435121	0.9145546	0.0905345
C	18	17	5	1.46	128.0	7.7	-3.9318964	-0.5350756	-0.0411512
C	19	18	17	1.54	115.0	86.2	-3.9827027	-1.3115613	1.2924633
O	20	19	18	1.41	111.1	-166.2	-3.6552326	-2.6701581	1.1028183
O	15	12	7	1.22	122.1	4.5	5.8778375	0.4918682	-0.2299465
O	17	5	4	1.23	119.9	-13.2	-3.0657816	2.9615546	0.4798867
C	19	18	17	1.54	110.5	-147.8	-5.0570105	-0.9997342	-0.9811238
O	24	19	18	1.21	122.4	-154.7	-5.5424637	-2.1087395	-0.9147223
O	24	19	18	1.34	115.9	27.0	-5.4492353	-0.112868	-1.9076881
H	8	7	6	1.08	121.6	178.3	4.1328443	1.9496051	-0.3215642
H	10	6	2	1.09	121.2	5.4	1.1960392	-2.6283563	-0.3778284
H	11	10	6	1.10	111.0	-107.3	3.3832475	-2.371644	-1.5552362
H	11	10	6	1.09	112.1	131.4	3.9476219	-2.9712871	0.0196215
H	16	15	12	1.10	112.0	56.8	5.9042889	-2.3392077	-1.6019599
H	16	15	12	1.09	107.5	176.8	7.2693661	-1.4850535	-0.8381465

H	16	15	12	1.10	111.5	-63.9	6.2334312	-2.5674735	0.1218433
H	13	10	6	1.09	117.9	104.1	0.9708757	-1.7851856	2.0233225
H	13	10	6	1.09	119.4	-107.8	2.7916821	-1.5832661	2.1018482
H	4	3	2	1.01	125.9	177.3	-0.6959561	3.221744	0.0537429
H	1	5	4	1.08	127.0	176.8	-1.2786121	-0.9442204	0.470359
H	18	17	5	1.02	110.4	162.9	-4.7969961	1.3411621	0.4406571
H	19	18	17	1.09	107.4	-33.8	-3.0047916	-0.8087567	-0.5537026
H	26	24	19	0.98	107.8	-2.4	-4.9376279	0.7105325	-1.7688255
H	20	19	18	1.09	108.7	-48.4	-3.2372688	-0.88963	1.9729222
H	20	19	18	1.10	109.5	69.2	-4.9734155	-1.1817395	1.7565204
H	21	20	19	0.98	105.3	-62.0	-4.3419472	-3.019086	0.5040375

Figure 3.18, CONFORMATION d): –1273.6337649 Ha.

Atom	NA	NB	NC	Bond	Angle	Dihedral	X	Y	Z
C							-0.7936302	0.0132285	0.0944784
C	1			1.41			0.5974079	0.2609305	0.1294885
C	2	1		1.39	107.1		0.767325	1.6422569	0.1360358
N	3	2	1	1.36	108.2	-0.3	-0.459799	2.2202489	0.0986528
C	4	3	2	1.37	110.3	0.0	-1.4367871	1.2603087	0.0669098
C	2	1	5	1.47	134.8	175.7	1.7976027	-0.5813092	0.24617
C	6	2	1	1.49	115.4	171.7	3.0975505	0.1246474	0.0436087
C	7	6	2	1.36	123.7	15.1	3.2328009	1.4742136	0.0632152
C	3	2	1	1.47	125.8	-179.6	2.0534382	2.3503349	0.1600423
C	6	2	1	1.52	129.8	28.0	2.0192828	-2.0282034	-0.179227

C	10	6	2	1.52	106.7	144.8	3.4571517	-2.1272773	-0.6554469
N	7	6	2	1.41	107.6	-162.2	4.077838	-0.8541657	-0.2213352
C	10	6	2	1.49	61.3	-105.3	1.8534127	-1.7250751	1.2722758
O	9	3	2	1.23	121.7	-178.6	2.1151532	3.5795371	0.1902114
C	12	7	6	1.40	125.0	-170.0	5.4519511	-0.6167155	-0.2703518
C	15	12	7	1.52	116.0	-175.9	6.3193541	-1.7997579	-0.6689525
C	5	4	3	1.49	114.4	-178.1	-2.8203577	1.7941006	-0.0200029
N	17	5	4	1.37	121.5	-178.6	-3.9107502	0.9604562	-0.0323995
C	18	17	5	1.44	130.2	-7.8	-3.9835218	-0.4772208	-0.120168
C	19	18	17	1.55	111.9	84.9	-3.8327473	-1.1493756	1.2673527
O	20	19	18	1.44	111.3	-174.3	-3.774553	-2.581051	1.1501489
O	15	12	7	1.22	122.1	4.6	5.9304103	0.471462	-0.0009492
O	17	5	4	1.23	118.5	0.8	-2.9764483	3.0142856	-0.0708945
C	19	18	17	1.55	107.9	-153.0	-5.3411431	-0.8466843	-0.7623955
O	24	19	18	1.21	122.5	8.4	-6.0804741	-0.0183958	-1.2374561
O	24	19	18	1.34	115.1	-171.9	-5.6341287	-2.1539145	-0.7586404
H	8	7	6	1.08	121.7	178.3	4.1948246	1.9497609	-0.0606224
H	10	6	2	1.09	121.2	5.3	1.2577887	-2.5778736	-0.7238044
H	11	10	6	1.10	110.9	-107.1	3.5024645	-2.2216057	-1.7486482
H	11	10	6	1.09	112.1	131.7	3.979498	-2.9813853	-0.2144508
H	16	15	12	1.10	112.0	56.9	6.0316827	-2.2100229	-1.6435908
H	16	15	12	1.09	107.5	176.8	7.3502352	-1.4480733	-0.7179798
H	16	15	12	1.10	111.5	-63.9	6.2551364	-2.6109502	0.0658066

H	13	10	6	1.09	118.1	104.0	0.9026766	-1.9632908	1.7394507
H	13	10	6	1.09	119.3	-108.0	2.7159121	-1.7878365	1.9315525
H	4	3	2	1.01	126.8	-178.5	-0.6626239	3.2129409	0.0725004
H	1	5	4	1.08	126.7	-179.5	-1.2656994	-0.9577935	0.0735179
H	18	17	5	1.02	113.1	-169.2	-4.7764406	1.4382992	-0.2694653
H	19	18	17	1.10	111.0	-36.7	-3.2040536	-0.8711158	-0.7891215
H	26	24	19	0.99	109.5	9.9	-4.9841491	-2.6359519	-0.194528
H	20	19	18	1.09	109.9	-50.8	-2.945626	-0.7604395	1.7758619
H	20	19	18	1.09	110.1	68.2	-4.7050584	-0.9287089	1.8877271
H	21	20	19	0.97	108.8	80.9	-2.8718078	-2.8375268	0.9038103

Figure 3.18, CONFORMATION e): –1273.6319410 Ha.

Atom	NA	NB	NC	Bond	Angle	Dihedral	X	Y	Z
C							2.5343897	1.7167848	-1.4892072
C	1			1.54			2.1986686	0.878682	-0.2435649
C	2	1		1.49	114.5		3.0376048	-0.3441971	-0.077355
N	3	2	1	1.41	107.6	51.4	4.3885029	0.048364	-0.1739012
C	4	3	2	1.48	111.1	19.1	4.5378666	1.5051265	0.0549315
C	1	2	3	1.49	60.4	-95.0	3.1484316	2.0698715	-0.1771438
C	3	2	1	1.36	123.8	-131.6	2.5327613	-1.5766134	0.1864579
C	7	3	2	1.47	120.7	-6.0	1.0982568	-1.7587539	0.4566472
C	8	7	3	1.46	114.4	-3.0	0.3237465	-0.5163112	0.4806781
C	9	8	7	1.39	126.3	3.0	0.8086701	0.7599238	0.2193939
N	9	8	7	1.36	125.2	-175.1	-0.9991422	-0.4197035	0.802804

C	11	9	8	1.38	109.1	177.1	-1.3738419	0.9078177	0.7812797
C	12	11	9	1.39	107.9	1.1	-0.2618764	1.6613971	0.4111179
C	12	11	9	1.48	125.8	-175.8	-2.699574	1.437157	1.1573527
O	14	12	11	1.22	120.9	151.4	-2.8207472	2.5697623	1.6055403
O	8	7	3	1.23	124.0	-179.3	0.5773798	-2.8467218	0.714751
C	4	3	2	1.39	125.0	-169.9	5.4703677	-0.8261708	-0.2704056
O	17	4	3	1.22	122.1	4.6	5.3222621	-2.033013	-0.3610911
C	17	4	3	1.52	116.0	-175.9	6.848788	-0.1848577	-0.2680981
N	14	12	11	1.40	118.4	-25.0	-3.8041448	0.5901148	1.0620882
C	20	14	12	1.46	126.3	-40.0	-4.064924	-0.3717316	-0.0033914
C	21	20	14	1.53	112.7	-91.3	-4.8968049	0.2232507	-1.1467348
O	22	21	20	1.34	112.8	35.7	-4.605195	1.5133579	-1.3875476
C	21	20	14	1.55	111.4	144.2	-4.743201	-1.648054	0.543901
O	24	21	20	1.41	112.5	-172.0	-4.8420758	-2.6623883	-0.4262368
O	22	21	20	1.22	124.5	-144.9	-5.7282476	-0.3921748	-1.7872441
H	7	3	2	1.08	121.7	178.6	3.1708406	-2.4433452	0.2787286
H	6	1	2	1.09	120.7	-111.0	2.8669597	2.9822065	0.338906
H	5	4	3	1.10	110.3	100.1	4.8737113	1.696762	1.0827437
H	5	4	3	1.09	110.6	-140.1	5.2752522	1.9293306	-0.6329637
H	19	17	4	1.10	112.0	56.5	7.0125218	0.4352193	0.6205026
H	19	17	4	1.09	107.5	176.5	7.5843091	-0.9895825	-0.2848434
H	19	17	4	1.10	111.5	-64.2	6.997513	0.4506642	-1.149037
H	1	6	5	1.09	117.9	-160.0	1.7188402	2.3347636	-1.8517942

H	1	6	5	1.09	119.4	-11.9	3.1519433	1.2663947	-2.2628089
H	11	9	8	1.01	122.4	6.4	-1.5319966	-1.2075763	1.1469423
H	13	12	11	1.08	124.0	179.7	-0.268434	2.7377286	0.3088128
H	20	14	12	1.02	108.9	177.4	-4.6274541	1.0841326	1.3942459
H	21	20	14	1.09	108.9	25.6	-3.112247	-0.6682162	-0.4540629
H	23	22	21	0.98	106.8	178.5	-5.1627788	1.7926617	-2.1392713
H	24	21	20	1.09	108.1	-53.7	-4.1267713	-2.03608	1.3603798
H	24	21	20	1.10	109.1	62.7	-5.7237942	-1.3806481	0.9721568
H	25	24	21	0.97	105.9	-63.2	-5.4179134	-2.3063486	-1.1260996

Transition state (Figure 3.19), –1389.2591938 Ha.

Atom	NA	NB	NC	Bond	Angle	Dihedral	X	Y	Z
N							3.8695984	-1.0567495	0.2915422
C	1			1.44			2.4598999	-0.9973829	0.0265181
C	2	1		1.41	107.9		2.0140535	0.2996038	0.3417571
C	3	2	1	1.49	111.2	7.8	3.1610815	1.2044939	0.6543608
C	1	2	3	1.47	108.9	8.7	4.2767348	0.175661	0.9839147
C	2	1	5	1.40	127.2	-170.2	1.6404865	-2.036111	-0.4397349
C	6	2	1	1.43	119.2	-178.7	0.2523874	-1.7796217	-0.6737218
C	7	6	2	1.44	113.3	4.3	-0.1123409	-0.3986189	-0.4772434
C	8	7	6	1.41	127.9	-11.3	0.6566607	0.6222336	0.1072388
N	8	7	6	1.36	122.7	-178.2	-1.3868295	0.049195	-0.6134383
C	10	8	7	1.37	110.6	165.8	-1.5187963	1.3047203	-0.0740637
C	11	10	8	1.40	107.3	3.3	-0.2530309	1.7056122	0.3622074

C	11	10	8	1.48	124.7	-164.6	-2.8124905	1.954769	0.2425999
O	13	11	10	1.23	120.5	-167.7	-2.8602355	3.1545203	0.510923
O	7	6	2	1.28	127.8	-173.2	-0.6569831	-2.620834	-0.9789387
C	4	3	2	1.54	108.9	99.1	3.4860397	2.048573	-0.5927005
O	16	4	3	1.61	110.6	116.0	3.2268306	3.6081947	-0.3022629
C	17	16	4	1.49	116.9	-104.7	2.0032878	4.208498	-0.894832
C	1	2	6	1.38	127.0	1.9	4.7062401	-2.1396975	0.1053145
O	19	1	2	1.22	123.1	5.4	4.3415521	-3.1846471	-0.4136931
C	19	1	2	1.53	116.0	-175.3	6.1483805	-1.9597682	0.5747345
N	13	11	10	1.38	118.5	16.6	-3.9460576	1.167039	0.318569
C	22	13	11	1.45	130.4	-7.5	-4.1413342	-0.2308473	-0.0034838
C	23	22	13	1.53	109.6	-168.1	-5.6412418	-0.5486788	-0.0280699
O	24	23	22	1.34	113.3	-178.7	-5.9064056	-1.8372975	-0.2925381
C	23	22	13	1.57	116.5	68.5	-3.4317134	-1.2698231	0.9339065
O	26	23	22	1.42	109.2	-157.6	-3.2736782	-2.4960728	0.2351305
O	24	23	22	1.21	123.9	1.8	-6.5078124	0.2728643	0.1624302
H	6	2	1	1.08	120.8	4.1	2.0345575	-3.031596	-0.5801364
H	4	3	2	1.10	116.4	-138.6	3.0257071	1.8916739	1.5038107
H	5	1	19	1.10	110.8	-76.2	4.301522	0.0220828	2.0706455
H	5	1	19	1.09	111.4	43.9	5.2734092	0.5006575	0.6684318
H	21	19	1	1.10	112.4	55.4	6.2109087	-1.666483	1.629149
H	21	19	1	1.09	107.2	175.3	6.653723	-2.916673	0.4402166
H	21	19	1	1.10	111.8	-65.9	6.6767194	-1.2001074	-0.0150948

H	16	4	3	1.09	113.5	2.0	2.8177127	1.8591864	-1.430338
H	16	4	3	1.09	115.6	-128.6	4.5328006	2.0434661	-0.907656
H	10	8	7	1.01	120.0	-14.7	-2.0750469	-0.5254534	-1.0689184
H	12	11	10	1.08	123.1	-172.1	-0.0853875	2.6050995	0.9413504
H	22	13	11	1.02	114.4	-178.1	-4.8021606	1.680806	0.5086505
H	23	22	13	1.10	111.1	-54.8	-3.8231934	-0.4483857	-1.0344167
H	25	24	23	0.99	109.8	5.2	-5.0613627	-2.3502255	-0.3378278
H	26	23	22	1.09	110.6	-33.7	-2.4810483	-0.8756207	1.2945475
H	26	23	22	1.10	109.3	84.6	-4.0648906	-1.4599564	1.8080126
H	27	26	23	1.00	110.9	95.3	-2.3678855	-2.5538546	-0.1823553
H	18	17	16	1.09	105.8	47.3	1.1917654	3.5148689	-0.6663574
H	18	17	16	1.09	106.2	165.7	1.8595333	5.1648121	-0.3933813
H	18	17	16	1.09	109.2	-72.9	2.1441741	4.3214503	-1.9710696
H	17	16	4	0.98	111.8	126.7	4.0094945	4.1523419	-0.5145944

Calculated conformation at the first reaction coordinate (Figure 3.19), –1389.3489081 Ha.

Atom	NA	NB	NC	Bond	Angle	Dihedral	X	Y	Z
N							3.8963894	-1.1533909	0.3278373
C	1			1.41			2.5609765	-1.0657853	-0.1023492
C	2	1		1.50	107.6		2.1491689	0.3684039	-0.0093271
C	3	2	1	1.53	106.4	-9.1	3.2717284	1.0930691	0.726606
C	1	2	3	1.48	111.2	19.0	4.2817352	0.0357261	1.1255915
C	2	1	5	1.36	127.8	-158.9	1.7410943	-2.0819666	-0.495904
C	6	2	1	1.46	119.6	171.6	0.3267126	-1.8158137	-0.7343511

C	7	6	2	1.45	114.7	1.7	-0.0678927	-0.4297116	-0.5555832
C	8	7	6	1.39	127.7	-7.5	0.6995437	0.614892	-0.0544079
N	8	7	6	1.36	122.2	-171.8	-1.3659995	-0.0401169	-0.576726
C	10	8	7	1.37	110.1	164.8	-1.4883271	1.2237126	-0.0566334
C	11	10	8	1.40	107.2	2.4	-0.1975849	1.6681784	0.2574128
C	11	10	8	1.49	124.8	-169.0	-2.7774593	1.8771849	0.2901649
O	13	11	10	1.23	120.1	-170.0	-2.8059004	3.0696641	0.5754388
O	7	6	2	1.25	124.8	-172.0	-0.53455	-2.6909992	-0.9563828
C	4	3	2	1.49	61.7	106.4	3.1333393	1.3099752	-0.7367783
O	16	4	3	3.20	87.8	126.9	3.0911227	4.4563104	-0.1443159
C	17	16	4	1.42	112.2	-131.2	2.0143486	5.1139999	-0.8057903
C	1	2	6	1.40	125.0	17.1	4.7366724	-2.2481148	0.1157538
O	19	1	2	1.22	122.0	0.1	4.3632988	-3.2434128	-0.4804568
C	19	1	2	1.52	116.1	179.3	6.1516424	-2.1178733	0.6546266
N	13	11	10	1.38	119.1	13.6	-3.9258383	1.1072265	0.3589729
C	22	13	11	1.45	131.0	-7.0	-4.1600743	-0.2781741	0.0176831
C	23	22	13	1.54	109.5	-166.7	-5.6733772	-0.5443523	-0.0325087
O	24	23	22	1.34	113.6	-178.4	-5.9858487	-1.8196971	-0.310898
C	23	22	13	1.56	115.6	70.5	-3.5042658	-1.3302739	0.9726239
O	26	23	22	1.43	109.4	-159.3	-3.3929371	-2.5808759	0.2962274
O	24	23	22	1.21	123.6	1.8	-6.5083935	0.3071884	0.1554688
H	6	2	1	1.08	121.7	-2.2	2.0873926	-3.1038393	-0.5426724
H	4	3	2	1.08	121.3	-144.5	3.0622582	1.9090213	1.4095456

H	5	1	19	1.10	110.1	-77.1	4.2233025	-0.1934825	2.1980046
H	5	1	19	1.09	110.6	42.7	5.3054691	0.3400586	0.8920694
H	21	19	1	1.10	112.2	55.4	6.1670627	-1.8850088	1.7251182
H	21	19	1	1.09	107.5	175.5	6.6550218	-3.0702536	0.4858737
H	21	19	1	1.10	111.3	-65.4	6.7048214	-1.3274402	0.1335706
H	16	4	3	1.08	116.6	106.0	2.7286899	2.2640105	-1.0496608
H	16	4	3	1.09	119.5	-106.7	3.8517757	0.8566137	-1.4164208
H	10	8	7	1.01	121.2	-17.3	-2.0805848	-0.630403	-0.9654031
H	12	11	10	1.08	124.0	-178.1	0.0186461	2.6237727	0.7140283
H	22	13	11	1.02	113.9	-176.9	-4.7693909	1.6407321	0.5534573
H	23	22	13	1.10	111.4	-53.1	-3.8295764	-0.5004541	-1.0096826
H	25	24	23	0.99	110.1	5.4	-5.1674826	-2.3689785	-0.3471274
H	26	23	22	1.09	110.8	-35.4	-2.5383008	-0.9755314	1.3348872
H	26	23	22	1.09	109.6	83.2	-4.1481449	-1.4858855	1.8440387
H	27	26	23	0.98	110.5	94.3	-2.5061152	-2.6704057	-0.117335
H	18	17	16	1.09	106.8	39.5	1.1488085	4.4500142	-0.7381081
H	18	17	16	1.10	112.3	157.9	1.7529209	6.068865	-0.3281285
H	18	17	16	1.10	112.4	-79.1	2.2253789	5.296594	-1.8696961
H	17	16	4	0.97	123.3	97.3	3.8739658	5.022088	-0.218577

Calculated conformation at the final reaction coordinate (Figure 3.19), –1389.2614077 Ha.

Atom	NA	NB	NC	Bond	Angle	Dihedral	X	Y	Z
N							3.8794531	-1.0293802	0.2735111
C	1			1.44			2.4714605	-0.9483332	0.0010539

C	2	1		1.41	108.3		2.0367016	0.3574962	0.303858
C	3	2	1	1.50	110.8	5.5	3.1963646	1.2335227	0.6705503
C	1	2	3	1.47	108.8	10.4	4.2926983	0.1804039	0.9927264
C	2	1	5	1.40	126.2	-168.0	1.65364	-1.9941666	-0.453287
C	6	2	1	1.43	119.3	180.0	0.264739	-1.7477687	-0.6769281
C	7	6	2	1.44	112.9	4.0	-0.0998117	-0.369537	-0.4778359
C	8	7	6	1.41	128.3	-11.2	0.6663925	0.6619758	0.0978638
N	8	7	6	1.36	122.0	-177.9	-1.3825371	0.0606525	-0.600152
C	10	8	7	1.37	110.6	165.5	-1.5262846	1.311397	-0.05928
C	11	10	8	1.40	107.3	3.1	-0.2616241	1.7288241	0.3670495
C	11	10	8	1.48	124.6	-165.5	-2.826416	1.9483308	0.2539843
O	13	11	10	1.23	120.6	-167.5	-2.8871525	3.1483704	0.5198774
O	7	6	2	1.28	127.9	-173.2	-0.6446781	-2.5951126	-0.9712262
C	4	3	2	1.54	113.9	100.7	3.6665589	2.1581471	-0.4696555
O	16	4	3	1.56	109.8	98.2	3.0576569	3.5772962	-0.2835737
C	17	16	4	1.49	120.0	-98.6	1.8423586	3.9730542	-1.05113
C	1	2	6	1.38	127.4	2.9	4.7078796	-2.1197591	0.0931721
O	19	1	2	1.22	123.3	5.8	4.3434369	-3.1598588	-0.4355096
C	19	1	2	1.53	116.0	-175.1	6.1469345	-1.9539272	0.5794771
N	13	11	10	1.38	118.6	16.6	-3.9531597	1.1504306	0.3246073
C	22	13	11	1.45	130.5	-7.4	-4.1355359	-0.249433	0.0024707
C	23	22	13	1.53	109.6	-168.2	-5.6325597	-0.5798669	-0.0252427
O	24	23	22	1.34	113.3	-178.6	-5.8858435	-1.870901	-0.2895648

C	23	22	13	1.57	116.5	68.3	-3.4191013	-1.2829039	0.9409944
O	26	23	22	1.42	109.2	-157.6	-3.2491199	-2.5072778	0.2426717
O	24	23	22	1.21	123.9	1.8	-6.5069578	0.2340705	0.1629006
H	6	2	1	1.08	120.6	3.2	2.0537772	-2.9885935	-0.583715
H	4	3	2	1.10	115.8	-135.7	3.0492867	1.8801456	1.5490564
H	5	1	19	1.10	110.9	-75.6	4.3003736	0.0071356	2.0771875
H	5	1	19	1.09	111.4	44.2	5.2994295	0.4949649	0.6992711
H	21	19	1	1.10	112.5	55.6	6.2017005	-1.6675442	1.6363544
H	21	19	1	1.09	107.2	175.4	6.6448883	-2.9146766	0.4449709
H	21	19	1	1.10	111.8	-65.8	6.688653	-1.19623	-0.0007507
H	16	4	3	1.09	113.8	-19.4	3.3420564	1.8361939	-1.4609176
H	16	4	3	1.09	114.3	-148.9	4.7365157	2.3820706	-0.4547436
H	10	8	7	1.01	119.7	-15.4	-2.0630186	-0.5236762	-1.0547625
H	12	11	10	1.08	122.9	-172.0	-0.1034032	2.6299267	0.9465512
H	22	13	11	1.02	114.4	-177.9	-4.8145968	1.6567338	0.5104659
H	23	22	13	1.10	111.1	-54.9	-3.8129042	-0.4647775	-1.0276771
H	25	24	23	0.99	109.7	5.2	-5.0355899	-2.3757776	-0.3334545
H	26	23	22	1.09	110.6	-33.7	-2.4720851	-0.8805819	1.3022125
H	26	23	22	1.10	109.3	84.6	-4.0519995	-1.4777387	1.8143543
H	27	26	23	1.00	110.7	95.1	-2.3416763	-2.5531734	-0.1756929
H	18	17	16	1.09	106.0	53.3	1.0978912	3.1969039	-0.8580287
H	18	17	16	1.09	105.4	171.2	1.5325512	4.9286738	-0.629348
H	18	17	16	1.09	108.7	-67.9	2.1042973	4.0469158	-2.107724

H	17	16	4	0.97	110.0	132.1	3.7524534	4.2537534	-0.3830247

Optimised structure starting from the conformation at the final reaction coordinate (Figure 3.19), –1389.2708845 Ha.

Atom	NA	NB	NC	Bond	Angle	Dihedral	X	Y	Z
N							3.9048052	-0.7898829	0.4219128
C	1			1.44			2.5100217	-0.7399496	0.0724818
C	2	1		1.41	107.9		2.0150332	0.525728	0.44823
C	3	2	1	1.50	110.3	5.4	3.1322616	1.396349	0.9465071
C	1	2	3	1.47	108.5	13.5	4.2022531	0.3420004	1.3149026
C	2	1	5	1.40	126.8	-164.0	1.7475337	-1.7732423	-0.4880389
C	6	2	1	1.43	119.2	177.0	0.3478735	-1.56553	-0.7131909
C	7	6	2	1.44	113.2	5.1	-0.0797562	-0.2219391	-0.4208722
C	8	7	6	1.41	128.3	-10.8	0.636402	0.7982614	0.231995
N	8	7	6	1.36	122.3	-178.6	-1.3759687	0.1703272	-0.5424219
C	10	8	7	1.37	110.3	166.9	-1.5645337	1.3885948	0.0645409
C	11	10	8	1.39	107.6	3.4	-0.3230159	1.8252558	0.5266091
C	11	10	8	1.48	124.4	-165.5	-2.8917833	1.9646161	0.3909511
O	13	11	10	1.23	120.7	-166.6	-3.0032771	3.1476205	0.7056906
O	7	6	2	1.27	127.6	-172.4	-0.5162259	-2.4259858	-1.0834975
C	4	3	2	1.53	111.1	94.3	3.6889412	2.2728677	-0.1744374
O	16	4	3	1.54	108.7	49.5	2.5120985	2.971577	-0.8857905
C	17	16	4	1.49	118.6	-98.6	1.9992003	2.3889499	-2.1612404
C	1	2	6	1.38	126.9	6.0	4.7870731	-1.8229881	0.1677052

O	19	1	2	1.22	122.9	9.0	4.5060349	-2.7772707	-0.5427474
C	19	1	2	1.53	116.1	-171.9	6.1651581	-1.7035029	0.8136553
N	13	11	10	1.38	118.4	17.6	-3.9873202	1.1200461	0.4192522
C	22	13	11	1.45	130.6	-8.1	-4.1179332	-0.267661	0.0290312
C	23	22	13	1.53	109.6	-168.0	-5.6024077	-0.649406	-0.0233345
O	24	23	22	1.34	113.3	-178.6	-5.8088871	-1.9345903	-0.3512245
C	23	22	13	1.57	116.5	68.6	-3.3683125	-1.3185729	0.9198817
O	26	23	22	1.42	109.2	-157.5	-3.1498465	-2.5014426	0.1630513
O	24	23	22	1.21	123.9	1.8	-6.5058718	0.1229969	0.1976993
H	6	2	1	1.08	120.8	1.5	2.185404	-2.7410396	-0.6837645
H	4	3	2	1.10	115.1	-141.0	2.8816579	2.0293766	1.8061724
H	5	1	19	1.10	111.1	-79.7	4.0695295	0.0599587	2.3680452
H	5	1	19	1.09	111.3	40.9	5.2307712	0.6936657	1.1834757
H	21	19	1	1.10	112.3	54.7	6.1051242	-1.561905	1.8989626
H	21	19	1	1.09	107.2	174.5	6.7021006	-2.6289752	0.6030409
H	21	19	1	1.10	111.9	-66.6	6.7396335	-0.8656256	0.3992167
H	16	4	3	1.09	112.5	-66.0	4.180237	1.6906536	-0.9557693
H	16	4	3	1.10	116.0	164.4	4.3115599	3.1171095	0.1403216
H	10	8	7	1.01	120.1	-13.8	-2.0334925	-0.408737	-1.0349874
H	12	11	10	1.08	123.6	-176.9	-0.1803534	2.7377075	1.0899609
H	22	13	11	1.02	114.2	-178.2	-4.8677704	1.5859038	0.6214236
H	23	22	13	1.10	111.2	-54.7	-3.7847355	-0.4220974	-1.0087734
H	25	24	23	0.99	109.8	5.2	-4.9420561	-2.4070295	-0.4146089

H	26	23	22	1.09	110.7	-33.6	-2.438853	-0.900824	1.3079058
H	26	23	22	1.10	109.4	84.8	-3.9975676	-1.5826916	1.7772406
H	27	26	23	1.00	110.8	94.3	-2.2475709	-2.4890271	-0.2610004
H	18	17	16	1.09	105.2	52.1	1.8063861	1.3427692	-1.940107
H	18	17	16	1.09	105.1	169.3	1.0713231	2.9211872	-2.3631919
H	18	17	16	1.09	108.3	-69.5	2.7556009	2.5362927	-2.9342622
H	17	16	4	0.98	110.8	131.7	2.6862042	3.927668	-0.988212

Calculated conformation for the carboxylate (Figure 3.19), –1388.8679690 Ha.

Atom	NA	NB	NC	Bond	Angle	Dihedral	X	Y	Z
N							-4.0754136	-0.8315955	-0.1594032
C	1			1.43			-2.6517803	-0.9064552	-0.0645443
C	2	1		1.39	109.8		-2.0957808	0.3415021	-0.3334885
C	3	2	1	1.51	111.0	1.8	-3.1733736	1.3703561	-0.5922034
C	1	2	3	1.48	109.3	5.9	-4.4613803	0.5066291	-0.6478951
C	2	1	5	1.41	126.4	-172.5	-1.8901858	-2.0599678	0.2209308
C	6	2	1	1.39	118.5	178.3	-0.5040102	-1.9514549	0.2382345
C	7	6	2	1.41	118.3	1.0	0.0681742	-0.6911332	-0.0084931
C	3	2	1	1.41	118.5	-178.9	-0.6918261	0.4714617	-0.3135223
N	8	7	6	1.37	128.2	-179.7	1.3949313	-0.3564211	-0.0465002
C	10	8	7	1.39	109.6	177.4	1.5273442	0.9856391	-0.3816801
C	11	10	8	1.38	108.1	0.6	0.2592564	1.514491	-0.5461914
C	11	10	8	1.50	127.4	-177.8	2.8027352	1.7549538	-0.5798642
O	13	11	10	1.23	118.2	-178.7	2.7221208	2.956403	-0.8531282

O	7	6	2	1.37	119.2	-179.4	0.2505378	-3.0631202	0.5016624
C	4	3	2	1.54	114.3	112.7	-3.2754005	2.4635441	0.482882
O	16	4	3	1.41	108.7	75.1	-2.200452	3.3615073	0.3181787
C	17	16	4	1.41	113.5	-170.3	-2.0691308	4.2804905	1.3792555
C	1	2	3	1.37	126.9	-178.2	-4.9705116	-1.8441515	0.0697693
O	19	1	2	1.23	123.7	2.4	-4.6463043	-2.9708016	0.4381304
C	19	1	2	1.53	115.8	-178.0	-6.4406298	-1.4942942	-0.1525124
N	13	11	10	1.37	119.7	2.9	4.0032616	1.1110037	-0.4876074
C	22	13	11	1.45	128.0	-13.3	4.2673058	-0.2094993	0.0417057
C	23	22	13	1.56	112.8	-155.3	5.7374333	-0.3765631	0.5485491
O	24	23	22	1.27	111.2	-178.0	5.9465884	-1.5541314	0.9681888
C	23	22	13	1.55	115.0	81.4	3.9874408	-1.3744794	-0.9388995
O	26	23	22	1.43	107.7	-150.2	3.6073306	-2.51691	-0.1625909
O	24	23	22	1.24	117.7	1.5	6.491267	0.6106783	0.4748883
H	6	2	1	1.08	120.9	-0.6	-2.3692682	-3.0098852	0.4049964
H	4	3	2	1.10	111.9	-126.2	-3.0235656	1.8884676	-1.5482537
H	5	1	19	1.10	110.1	-66.8	-4.8434849	0.4379632	-1.6745095
H	5	1	19	1.10	110.5	51.8	-5.2608383	0.9263892	-0.0270377
H	21	19	1	1.10	112.3	55.9	-6.6283364	-1.1091155	-1.1617904
H	21	19	1	1.09	107.3	176.0	-7.0182795	-2.40756	-0.0034948
H	21	19	1	1.10	111.7	-64.8	-6.788561	-0.7349054	0.5586324
H	16	4	3	1.10	109.8	-46.4	-3.2557567	2.0039812	1.4857373
H	16	4	3	1.11	110.0	-164.6	-4.2365327	3.0040001	0.382686

H	10	8	7	1.01	123.4	-6.7	2.137441	-0.9943552	0.2032584
H	12	11	10	1.08	124.4	179.5	0.0563265	2.5420051	-0.8016674
H	22	13	11	1.02	116.4	-170.5	4.8354295	1.7014433	-0.4809166
H	23	22	13	1.10	109.7	-40.5	3.6645907	-0.3742519	0.9443828
H	26	23	22	1.09	112.6	-31.2	3.1721633	-1.1519631	-1.6332222
H	26	23	22	1.10	110.1	89.2	4.8849042	-1.5934331	-1.5317297
H	18	17	16	1.10	111.6	59.8	-1.9025307	3.7691224	2.3404288
H	18	17	16	1.09	107.1	178.9	-1.2016976	4.9064094	1.1560918
H	18	17	16	1.10	111.8	-61.7	-2.961605	4.9223089	1.4772848
H	27	26	23	1.00	99.7	-52.3	4.3700494	-2.5471059	0.4772035
H	15	7	6	0.98	111.2	166.0	1.1873106	-2.9158427	0.2716644

Reaction pathway shown in Figure 3.19, obtained from IRC and reverse IRC calculations:

FORWARD path complete.

Energies reported relative to the TS energy of -1389.259194

Summary of reaction path following

	Energy	RxCoord
1	0.00000	0.00000
2	-0.00009	0.46471
3	-0.00021	0.69872
4	-0.00079	1.16341

5	-0.00218	1.64670
6	-0.00451	2.13250
7	-0.00768	2.61835
8	-0.01155	3.10410
9	-0.01608	3.58981
10	-0.02128	4.07550
11	-0.02720	4.56122
12	-0.03384	5.04696
13	-0.04114	5.53278
14	-0.04399	5.71226
15	-0.05194	6.19872
16	-0.05994	6.68469
17	-0.06725	7.16557
18	-0.07308	7.64136
19	-0.07602	7.98440
20	-0.07833	8.39081
21	-0.08017	8.85196
22	-0.08166	9.32615
23	-0.08290	9.80822
24	-0.08393	10.28632
25	-0.08481	10.76872
26	-0.08557	11.25160
27	-0.08624	11.73305

28	-0.08683	12.21675
29	-0.08736	12.69747
30	-0.08784	13.18210
31	-0.08827	13.66464
32	-0.08866	14.14933
33	-0.08902	14.63139
34	-0.08934	15.11521
35	-0.08963	15.59317
36	-0.08969	15.69849
37	-0.08973	15.78218
38	-0.08975	15.81361

Total number of points: 37

Total number of gradient calculations: 1221

Total number of Hessian calculations: 0

REVERSE path complete.

Energies reported relative to the TS energy of -1389.259194

Summary of reaction path following

	Energy	RxCoord
1	-0.00220	-3.71856
2	-0.00217	-3.68625
3	-0.00211	-3.61646
4	-0.00199	-3.48048
5	-0.00158	-2.99724
6	-0.00147	-2.87414
7	-0.00109	-2.39425
8	-0.00075	-1.92114
9	-0.00046	-1.43652
10	-0.00022	-0.95460
11	-0.00006	-0.47125
12	0.00000	0.00000

Total number of points: 11

Total number of gradient calculations: 1162

Total number of Hessian calculations: 0

6

Chapter Six

References

1. Cancer Research UK. http://www.cancerresearchuk.org/about-cancer/what-is-cancer (August 2015).

2. Casás-Selves, M.; DeGregori, J. How cancer shapes evolution, and how evolution shapes cancer. *Evolution* **2011,** 4, 624-634.

3. Vincent, T. L.; Gatenby, R. A. An evolutionary model for initiation, promotion, and progression in carcinogenesis. *International Journal of Oncology* **2008,** 32, 729-737.

4. Ruddon, R. W. *Cancer Biology*. fourth ed.; Oxford University Press: New York, 2007.

5. Hanahan, D.; Weinberg, R. A. The Hallmarks of Cancer. *Cell* **2000,** 100, 57-70.

6. Karin, M.; Greten, F. R. NF-kB: linking inflammation and immunity to cancer development and progression. *Nat Rev Immunol* **2005,** 5, 749-759.

7. Hanahan, D.; Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **2011,** 144, 646-74.

8. Shay, J. W.; Bacchetti, S. A survey of telomerase activity in human cancer. *Eur J Cancer* **1997,** 33, 787-91.

9. Bearss, D. J.; Hurley, L. H.; Von Hoff, D. D. Telomere maintenance mechanisms as a target for drug development. *Oncogene* **2000,** 19, 6632-41.

10. Bergers, G.; Benjamin, L. E. Tumorigenesis and the angiogenic switch. *Nat Rev Cancer* **2003,** 3, 401-10.

11. Spano, D.; Heck, C.; De Antonellis, P.; Christofori, G.; Zollo, M. Molecular networks that regulate cancer metastasis. *Semin Cancer Biol* **2012,** 22, 234-49.

12. Sudhakar, A. History of Cancer, Ancient and Modern Treatment Methods. *Journal of cancer science & therapy* **2009,** 1, 1-4.

13. DeVita, V. T., Jr.; Chu, E. A history of cancer chemotherapy. *Cancer Res* **2008,** 68, 8643-53.

14. Joint Formulary Committee. *British National Formulary*. sixty ninth ed.; BMJ Group and Pharmaceutical Press: London, 2015.

15. Agudelo, D.; Bourassa, P.; Berube, G.; Tajmir-Riahi, H. A. Intercalation of antitumor drug doxorubicin and its analogue by DNA duplex: structural features and biological implications. *Int J Biol Macromol* **2014,** 66, 144-50.

16. Emadi, A.; Jones, R. J.; Brodsky, R. A. Cyclophosphamide and cancer: golden anniversary. *Nat Rev Clin Oncol* **2009,** 6, 638-47.

17. Boddy, A. V.; Yule, S. M. Metabolism and pharmacokinetics of oxazaphosphorines. *Clin Pharmacokinet* **2000,** 38, 291-304.

18. Moen, M. D.; McKeage, K.; Plosker, G. L.; Siddiqui, M. A. Imatinib: a review of its use in chronic myeloid leukaemia. *Drugs* **2007,** 67, 299-320.

19. Deininger, M. W. N.; Goldman, J. M.; Melo, J. V. The molecular biology of chronic myeloid leukemia. *Blood* **2000,** 96, 3343-3356.

20. Druker, B. J. Imatinib as a paradigm of targeted therapies. *Adv Cancer Res* **2004,** 91, 1-30.

21. Dasgupta, A.; Kumar, L. Chronic myeloid leukaemia: A paradigm shift in management. *Natl Med J India* **2014,** 27, 316-23.

22. Hudis, C. A. Trastuzumab--mechanism of action and use in clinical practice. *N Engl J Med* **2007,** 357, 39-51.

23. Moja, L.; Tagliabue, L.; Balduzzi, S.; Parmelli, E.; Pistotti, V.; Guarneri, V.; D'Amico, R. Trastuzumab containing regimens for early breast cancer. *Cochrane Database Syst Rev* **2012,** 4, Cd006243.

24. Pinto, A. C.; Ades, F.; de Azambuja, E.; Piccart-Gebhart, M. Trastuzumab for patients with HER2 positive breast cancer: delivery, duration and combination therapies. *Breast* **2013,** 22 Suppl 2, S152-5.

25. Huang, M.; Shen, A.; Ding, J.; Geng, M. Molecularly targeted cancer therapy: some lessons from the past decade. *Trends in Pharmacological Sciences* 35, 41-50.

26. Shaikh, A. J. Molecular targeting agents in cancer therapy: science and society. *Asian Pac J Cancer Prev* **2012,** 13, 1705-8.

27. Hawkes, N. High cost of cancer treatment doesn't reflect benefits, say specialists. *BMJ* **2011,** 343.

28. Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the 30 Years from 1981 to 2010. *Journal of natural products* **2012,** 75, 311-335.

29. Oki, T. New anthracycline antibiotics. *Jpn J Antibiot* **1977,** 30 Suppl, 70-84.

30. Arcamone, F.; Cassinelli, G.; Fantini, G.; Grein, A.; Orezzi, P.; Pol, C.; Spalla, C. Adriamycin, 14-hydroxydaunomycin, a new antitumor antibiotic from S. peucetius var. caesius. *Biotechnology and Bioengineering* **1969,** 11, 1101-1110.

31. Kunitomo, N.; Mori, S. Studies on pterygium. Report IV. A treatment of the pterygium by mitomycin C installation. *Acta Soc Ophthalmol Jpn* **1963,** 67, 601–7.

32. Wakaki, S.; Marumo, H.; Tomioka, K.; Shimizu, G.; Kato, E.; Kamada, H.; Kudo, S.; Fujimoto, Y. Isolation of new fractions of antitumor mitomycins. *Antibiot Chemother (Northfield)* **1958,** 8, 228-40.

33. Bass, P. D.; Gubler, D. A.; Judd, T. C.; Williamsa, R. M. The Mitomycinoid Alkaloids: Mechanism of Action, Biosynthesis, Total Syntheses and Synthetic Approaches. *Chemical reviews* **2013,** 113, 6816-6863.

34. Maiese, W. M.; Lechevalier, M. P.; Lechevalier, H. A.; Korshalla, J.; Kuck, N.; Fantini, A.; Wildey, M. J.; Thomas, J.; Greenstein, M. Calicheamicins, a novel family of antitumor antibiotics: taxonomy, fermentation and biological properties. *J Antibiot (Tokyo)* **1989,** 42, 558-63.

35. Smith, A. L.; Nicolaou, K. C. The enediyne antibiotics. *J Med Chem* **1996,** 39, 2103-17.

36. Ghosh, N.; Sheldrake, H. M.; Searcey, M.; Pors, K. Chemical and biological explorations of the family of CC-1065 and the duocarmycin natural products. *Curr Top Med Chem* **2009,** 9, 1494-524.

37. Chidester, C. G.; Krueger, W. C.; Mizsak, S. A.; Duchamp, D. J.; Martin, D. G. The structure of CC-1065, a potent antitumor agent and its binding to DNA. *Journal of the American Chemical Society* **1981,** 103, 7629-7635.

38. Hanka, L. J.; Dietz, A.; Gerpheide, S. A.; Kuentzel, S. L.; Martin, D. G. CC-1065 (NSC-298223), a new antitumor antibiotic. Production, in vitro biological activity, microbiological assays and taxonomy of the producing microorganism. *J Antibiot (Tokyo)* **1978,** 31, 1211-7.

39. Boger, D. L.; Johnson, D. S. CC-1065 and the duocarmycins: unraveling the keys to a new class of naturally derived DNA alkylating agents. *Proc Natl Acad Sci U S A* **1995,** 92, 3642-9.

40. MacMillan, K. S.; Boger, D. L. Fundamental Relationships Between Structure, Reactivity, and Biological Activity for the Duocarmycins and CC-1065. *Journal of medicinal chemistry* **2009,** 52, 5771-5780.

41. Wolkenberg, S. E.; Boger, D. L. Mechanisms of in Situ Activation for DNA-Targeting Antitumor Agents. *Chemical Reviews* **2002,** 102, 2477-2496.

42. Boger, D. L.; Turnbull, P. Synthesis and Evaluation of a Carbocyclic Analogue of the CC-1065 and Duocarmycin Alkylation Subunits: Role of the Vinylogous Amide and Implications on DNA Alkylation Catalysis. *The Journal of Organic Chemistry* **1998,** 63, 8004-8011.

43. Eis, P. S.; Smith, J. A.; Rydzewski, J. M.; Case, D. A.; Boger, D. L.; Chazin, W. J. High resolution solution structure of a DNA duplex alkylated by the antitumor agent duocarmycin SA. *Journal of Molecular Biology* **1997,** 272, 237-252.

44. Boger, D. L.; Garbaccio, R. M. Shape-Dependent Catalysis: Insights into the Source of Catalysis for the CC-1065 and Duocarmycin DNA Alkylation Reaction. *Accounts of Chemical Research* **1999,** 32, 1043-1052.

45. Boger, D. L.; Santillán, A.; Searcey, M.; Jin, Q. Critical Role of the Linking Amide in CC-1065 and the Duocarmycins: Implications on the Source of DNA Alkylation Catalysis. *Journal of the American Chemical Society* **1998,** 120, 11554-11557.

46. Boger, D. L.; Mesini, P.; Tarby, C. M. Chemical and Structural Comparison of N-BOC-CBQ and N-BOC-CBI: Identification and Structural Origin of an Unappreciated but Productive Stability of the CC-1065 and Duocarmycin SA Alkylation Subunits. *Journal of the American Chemical Society* **1994,** 116, 6461-6462.

47. Boger, D. L.; Turnbull, P. Synthesis and Evaluation of CC-1065 and Duocarmycin Analogs Incorporating the 1,2,3,4,11,11a-Hexahydrocyclopropa[c]naphtho[2,1-b]azepin-6-one (CNA) Alkylation Subunit: Structural Features that Govern Reactivity and Reaction Regioselectivity. *The Journal of Organic Chemistry* **1997,** 62, 5849-5863.

48. Boger, D. L.; Santillán, A.; Searcey, M.; Brunette, S. R.; Wolkenberg, S. E.; Hedrick, M. P.; Jin, Q. Synthesis and Evaluation of 1,2,8,8a-Tetrahydrocyclopropa[c]pyrrolo[3,2-e]indol-4(5H)-one, the Parent Alkylation Subunit of CC-1065 and the Duocarmycins: Impact of the Alkylation Subunit Substituents and Its Implications for DNA Alkylation Catalysis. *The Journal of Organic Chemistry* **2000,** 65, 4101-4111.

49. Boger, D. L.; Hughes, T. V.; Hedrick, M. P. Synthesis, Chemical Properties, and Biological Evaluation of CC-1065 and Duocarmycin Analogues Incorporating the 5-Methoxycarbonyl-1,2,9,9a-tetrahydrocyclopropa[c]benz[e]indol-4-one Alkylation Subunit. *The Journal of Organic Chemistry* **2001,** 66, 2207-2216.

50. Boger, D. L.; McKie, J. A.; Cai, H.; Cacciari, B.; Baraldi, P. G. Synthesis and Properties of Substituted CBI Analogs of CC-1065 and the Duocarmycins Incorporating the 7-Methoxy-1,2,9,9a-tetrahydrocyclopropa[c]benz[e]indol-4-one (MCBI) Alkylation Subunit: Magnitude of Electronic Effects on the Functional Reactivity. *The Journal of Organic Chemistry* **1996,** 61, 1710-1729.

51. Boger, D. L.; Han, N.; Tarby, C. M.; Boyce, C. W.; Cai, H.; Jin, Q.; Kitos, P. A. Synthesis, Chemical Properties, and Preliminary Evaluation of Substituted CBI Analogs of CC-1065 and the Duocarmycins Incorporating the 7-Cyano-1,2,9,9a-tetrahydrocyclopropa[c]benz[e]indol-4-one

Alkylation Subunit: Hammett Quantitation of the Magnitude of Electronic Effects on Functional Reactivity. *The Journal of Organic Chemistry* **1996,** 61, 4894-4912.

52. Parrish, J. P.; Hughes, T. V.; Hwang, I.; Boger, D. L. Establishing the Parabolic Relationship between Reactivity and Activity for Derivatives and Analogues of the Duocarmycin and CC-1065 Alkylation Subunits. *Journal of the American Chemical Society* **2004,** 126, 80-81.

53. Searcey, M. Duocarmycins--natures prodrugs? *Curr Pharm Des* **2002,** 8, 1375-89.

54. Bhuyan, B. K.; Newell, K. A.; Crampton, S. L.; Von Hoff, D. D. CC-1065 (NSC 298223), a most potent antitumor agent: kinetics of inhibition of growth, DNA synthesis, and cell survival. *Cancer Res* **1982,** 42, 3532-7.

55. McGovren, J. P.; Clarke, G. L.; Pratt, E. A.; DeKoning, T. F. Preliminary toxicity studies with the DNA-binding antibiotic, CC-1065. *J Antibiot (Tokyo)* **1984,** 37, 63-70.

56. Shamdas, G. J.; Alberts, D. S.; Modiano, M.; Wiggins, C.; Power, J.; Kasunic, D. A.; Elfring, G. L.; Earhart, R. H. Phase I study of adozelesin (U-73,975) in patients with solid tumors. *Anticancer Drugs* **1994,** 5, 10-4.

57. Fleming, G. F.; Ratain, M. J.; O'Brien, S. M.; Schilsky, R. L.; Hoffman, P. C.; Richards, J. M.; Vogelzang, N. J.; Kasunic, D. A.; Earhart, R. H. Phase I study of adozelesin administered by 24-hour continuous intravenous infusion. *J Natl Cancer Inst* **1994,** 86, 368-72.

58. Burris, H. A.; Dieras, V. C.; Tunca, M.; Earhart, R. H.; Eckardt, J. R.; Rodriguez, G. I.; Shaffer, D. S.; Fields, S. M.; Campbell, E.; Schaaf, L.; Kasunic, D.; Von Hoff, D. D. Phase I study with the DNA sequence-specific agent adozelesin. *Anticancer Drugs* **1997,** 8, 588-96.

59. Foster, B. J.; LoRusso, P. M.; Poplin, E.; Zalupski, M.; Valdivieso, M.; Wozniak, A.; Flaherty, L.; Kasunic, D. A.; Earhart, R. H.; Baker, L. H. Phase I trial of Adozelesin using the treatment schedule of daily x5 every 3 weeks. *Invest New Drugs* **1996,** 13, 321-6.

60. Cristofanilli, M.; Bryan, W. J.; Miller, L. L.; Chang, A. Y.; Gradishar, W. J.; Kufe, D. W.; Hortobagyi, G. N. Phase II study of adozelesin in untreated metastatic breast cancer. *Anticancer Drugs* **1998,** 9, 779-82.

61. Pitot, H. C.; Reid, J. M.; Sloan, J. A.; Ames, M. M.; Adjei, A. A.; Rubin, J.; Bagniewski, P. G.; Atherton, P.; Rayson, D.; Goldberg, R. M.; Erlichman, C. A Phase I study of bizelesin (NSC 615291) in patients with advanced solid tumors. *Clin Cancer Res* **2002,** 8, 712-7.

62. Schwartz, G. H.; Patnaik, A.; Hammond, L. A.; Rizzo, J.; Berg, K.; Von Hoff, D. D.; Rowinsky, E. K. A phase I study of bizelesin, a highly potent and selective DNA-interactive agent, in patients with advanced solid malignancies. *Ann Oncol* **2003,** 14, 775-82.

63. Pavlidis, N.; Aamdal, S.; Awada, A.; Calvert, H.; Fumoleau, P.; Sorio, R.; Punt, C.; Verweij, J.; van Oosterom, A.; Morant, R.; Wanders, J.; Hanauske, A. R. Carzelesin phase II study in advanced breast, ovarian, colorectal, gastric, head and neck cancer, non-Hodgkin's lymphoma and malignant melanoma: a study of the EORTC early clinical studies group (ECSG). *Cancer Chemother Pharmacol* **2000,** 46, 167-71.

64. van Tellingen, O.; Punt, C. J.; Awada, A.; Wagener, D. J.; Piccart, M. J.; Groot, Y.; Schaaf, L. J.; Henrar, R. E.; Nooijen, W. J.; Beijnen, J. H. A clinical pharmacokinetics study of carzelesin given by short-term intravenous infusion in a phase I study. *Cancer Chemother Pharmacol* **1998,** 41, 377-84.

65. Wolff, I.; Bench, K.; Beijnen, J. H.; Bruntsch, U.; Cavalli, F.; de Jong, J.; Groot, Y.; van Tellingen, O.; Wanders, J.; Sessa, C. Phase I clinical and pharmacokinetic study of carzelesin (U-80244) given daily for five consecutive days. *Clin Cancer Res* **1996,** 2, 1717-23.

66. Awada, A.; Punt, C. J.; Piccart, M. J.; Van Tellingen, O.; Van Manen, L.; Kerger, J.; Groot, Y.; Wanders, J.; Verweij, J.; Wagener, D. J. Phase I study of Carzelesin (U-80,244) given (4-weekly) by intravenous bolus schedule. *Br J Cancer* **1999,** 79, 1454-61.

67. Li, L. H.; DeKoning, T. F.; Kelly, R. C.; Krueger, W. C.; McGovren, J. P.; Padbury, G. E.; Petzold, G. L.; Wallace, T. L.; Ouding, R. J.; Prairie, M. D.; et al. Cytotoxicity and antitumor activity of carzelesin, a prodrug cyclopropylpyrroloindole analogue. *Cancer Res* **1992,** 52, 4904-13.

68. Lajiness, J. P.; Robertson, W. M.; Dunwiddie, I.; Broward, M. A.; Vielhauer, G. A.; Weir, S. J.; Boger, D. L. Design, Synthesis, and Evaluation of Duocarmycin O-Amino Phenol Prodrugs Subject to Tunable Reductive Activation. *Journal of Medicinal Chemistry* **2010,** 53, 7731-7738.

69. Jin, W.; Trzupek, J. D.; Rayl, T. J.; Broward, M. A.; Vielhauer, G. A.; Weir, S. J.; Hwang, I.; Boger, D. L. A Unique Class of Duocarmycin and CC-1065 Analogues Subject to Reductive Activation. *Journal of the American Chemical Society* **2007,** 129, 15391-15397.

70. Wolfe, A. L.; Duncan, K. K.; Parelkar, N. K.; Brown, D.; Vielhauer, G. A.; Boger, D. L. Efficacious Cyclic N-Acyl O-Amino Phenol Duocarmycin Prodrugs. *Journal of medicinal chemistry* **2013,** 56, 4104-4115.

71. Tercel, M.; Atwell, G. J.; Yang, S.; Stevenson, R. J.; Botting, K. J.; Boyd, M.; Smith, E.; Anderson, R. F.; Denny, W. A.; Wilson, W. R.; Pruijn, F. B. Hypoxia-Activated Prodrugs: Substituent Effects on the Properties of Nitro seco-1,2,9,9a-Tetrahydrocyclopropa[c]benz[e]indol-4-one (nitroCBI) Prodrugs of DNA Minor Groove Alkylating Agents. *Journal of Medicinal Chemistry* **2009,** 52, 7258-7272.

72. Stevenson, R. J.; Denny, W. A.; Tercel, M.; Pruijn, F. B.; Ashoorzadeh, A. Nitro seco Analogues of the Duocarmycins Containing Sulfonate Leaving Groups as Hypoxia-Activated Prodrugs for Cancer Therapy. *Journal of Medicinal Chemistry* **2012,** 55, 2780-2802.

73. Wilson, W. R.; Stribbling, S. M.; Pruijn, F. B.; Syddall, S. P.; Patterson, A. V.; Liyanage, H. D.; Smith, E.; Botting, K. J.; Tercel, M. Nitro-chloromethylbenzindolines: hypoxia-activated prodrugs of potent adenine N3 DNA minor groove alkylators. *Mol Cancer Ther* **2009,** 8, 2903-13.

74. Sutherland, M.; Gill, J. H.; Loadman, P. M.; Laye, J. P.; Sheldrake, H. M.; Illingworth, N. A.; Alandas, M. N.; Cooper, P. A.; Searcey, M.; Pors, K.; Shnyder, S. D.; Patterson, L. H. Antitumor activity of a duocarmycin analogue rationalized to be metabolically activated by cytochrome P450 1A1 in human transitional cell carcinoma of the bladder. *Mol Cancer Ther* **2013,** 12, 27-37.

75. Pors, K.; Loadman, P. M.; Shnyder, S. D.; Sutherland, M.; Sheldrake, H. M.; Guino, M.; Kiakos, K.; Hartley, J. A.; Searcey, M.; Patterson, L. H. Modification of the duocarmycin pharmacophore enables CYP1A1 targeting for biological activity. *Chem Commun (Camb)* **2011,** 47, 12062-4.

76. Chari, R. V.; Miller, M. L.; Widdison, W. C. Antibody-drug conjugates: an emerging concept in cancer therapy. *Angew Chem Int Ed Engl* **2014,** 53, 3796-827.

77. Petersen, B. H.; DeHerdt, S. V.; Schneck, D. W.; Bumol, T. F. The human immune response to KS1/4-desacetylvinblastine (LY256787) and KS1/4-desacetylvinblastine hydrazide (LY203728) in single and multiple dose clinical studies. *Cancer Research* **1991,** 51, 2286-2290.

78. Almagro, J. C.; Fransson, J. Humanization of antibodies. *Front Biosci* **2008,** 13, 1619-33.

79. Chari, R. V.; Jackel, K. A.; Bourret, L. A.; Derr, S. M.; Tadayoni, B. M.; Mattocks, K. M.; Shah, S. A.; Liu, C.; Blattler, W. A.; Goldmacher, V. S. Enhancement of the selectivity and antitumor efficacy of a CC-1065 analogue through immunoconjugate formation. *Cancer Res* **1995,** 55, 4079-84.

80. Zhao, R. Y.; Erickson, H. K.; Leece, B. A.; Reid, E. E.; Goldmacher, V. S.; Lambert, J. M.; Chari, R. V. J. Synthesis and Biological Evaluation of Antibody Conjugates of Phosphate Prodrugs

of Cytotoxic DNA Alkylators for the Targeted Treatment of Cancer. *Journal of Medicinal Chemistry* **2012,** 55, 766-782.

81. Derwin, D.; Passmore, D. Abstract 2575: Activation of antibody drug conjugate MDX-1203 by human carboxylesterase. *Proceedings of the 101st Annual Meeting of the American Association for Cancer Research* **2010,** 70, Washington.

82. Bristol-Myers Squibb. Presentation Slides on Mdx-1203 http://c.ymcdn.com/sites/casss.site-ym.com/resource/resmgr/Mass_Spec_Speaker_Slides/2010_MS_SrinivasanMohan.pdf. (Aug 2015).

83. Bristol-Myers Squibb. Study of MDX-1203 in Subjects With Advanced/Recurrent Clear Cell Renal Cell Carcinoma (ccRCC) or Relapsed/Refractory B-Cell Non-Hodgkin's Lymphoma (B-NHL) https://clinicaltrials.gov/ct2/show/NCT00944905. (August 2015).

84. Elgersma, R. C.; Coumans, R. G. E.; Huijbregts, T.; Menge, W. M. P. B.; Joosten, J. A. F.; Spijker, H. J.; de Groot, F. M. H.; van der Lee, M. M. C.; Ubink, R.; van den Dobbelsteen, D. J.; Egging, D. F.; Dokter, W. H. A.; Verheijden, G. F. M.; Lemmens, J. M.; Timmers, C. M.; Beusker, P. H. Design, Synthesis, and Evaluation of Linker-Duocarmycin Payloads: Toward Selection of HER2-Targeting Antibody–Drug Conjugate SYD985. *Molecular Pharmaceutics* **2015,** 12, 1813-1835.

85. van der Lee, M. M.; Groothuis, P. G.; Ubink, R.; van der Vleuten, M. A.; van Achterberg, T. A.; Loosveld, E. M.; Damming, D.; Jacobs, D. C.; Rouwette, M.; Egging, D. F.; van den Dobbelsteen, D.; Beusker, P. H.; Goedings, P.; Verheijden, G. F.; Lemmens, J. M.; Timmers, M.; Dokter, W. H. The Preclinical Profile of the Duocarmycin-Based HER2-Targeting ADC SYD985 Predicts for Clinical Benefit in Low HER2-Expressing Breast Cancers. *Mol Cancer Ther* **2015,** 14, 692-703.

86. Meyer, Y.; Richard, J. A.; Delest, B.; Noack, P.; Renard, P. Y.; Romieu, A. A comparative study of the self-immolation of para-aminobenzylalcohol and hemithioaminal-based linkers in the context of protease-sensitive fluorogenic probes. *Org Biomol Chem* **2010,** 8, 1777-80.

87. Blencowe, C. A.; Russell, A. T.; Greco, F.; Hayes, W.; Thornthwaite, D. W. Self-immolative linkers in polymeric delivery systems. *Polymer Chemistry* **2011,** 2, 773-790.

88. Boger, D. L.; Searcey, M.; Tse, W. C.; Jin, Q. Bifunctional alkylating agents derived from duocarmycin SA: potent antitumor activity with altered sequence selectivity. *Bioorganic & Medicinal Chemistry Letters* **2000,** 10, 495-498.

89. Chan, W.; White, P. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*. OUP Oxford: 2000.

90. Tichenor, M. S.; Trzupek, J. D.; Kastrinsky, D. B.; Shiga, F.; Hwang, I.; Boger, D. L. Asymmetric Total Synthesis of (+)- and ent-(−)-Yatakemycin and Duocarmycin SA: Evaluation of Yatakemycin Key Partial Structures and Its Unnatural Enantiomer. *Journal of the American Chemical Society* **2006,** 128, 15683-15696.

91. Patel, V. F.; Andis, S. L.; Enkema, J. K.; Johnson, D. A.; Kennedy, J. H.; Mohamadi, F.; Schultz, R. M.; Soose, D. J.; Spees, M. M. Total Synthesis of Seco (+)- and ent-(−)-Oxaduocarmycin SA: Construction of the (Chloromethyl)indoline Alkylating Subunit by a Novel Intramolecular Aryl Radical Cyclization onto a Vinyl Chloride. *The Journal of Organic Chemistry* **1997,** 62, 8868-8874.

92. MacMillan, K. S.; Nguyen, T.; Hwang, I.; Boger, D. L. Total Synthesis and Evaluation of iso-Duocarmycin SA and iso-Yatakemycin. *Journal of the American Chemical Society* **2009,** 131, 1187-1194.

93. Boger, D. L.; Boyce, C. W.; Garbaccio, R. M.; Searcey, M. Synthesis of CC-1065 and duocarmycin analogs via intramolecular aryl radical cyclization of a tethered vinyl chloride. *Tetrahedron Letters* **1998,** 39, 2227-2230.

94. Boger, D. L.; Ishizaki, T.; Zarrinmayeh, H.; Kitos, P. A.; Suntornwat, O. A potent, simple derivative of an analog of the CC-1065 alkylation subunit. *Bioorganic & Medicinal Chemistry Letters* **1991,** 1, 55-58.

95. Wuts, P. G. M. *Greene's Protective Groups in Organic Synthesis, 5th Edition*. Wiley: 2014.

96. Tietze, Lutz F.; Haunert, F.; Feuerstein, T.; Herzig, T. A Concise and Efficient Synthesis of seco-Duocarmycin SA. *European Journal of Organic Chemistry* **2003,** 2003, 562-566.

97. Brisbois, R. G.; Wanke, R. A. Iodine Monochloride. In *Encyclopedia of Reagents for Organic Synthesis*, John Wiley & Sons, Ltd: 2001.

98. Chaikovskii, V. K.; Filimonov, V. D.; Skorokhodov, V. I.; Ogorodnikov, V. D. Superactivity and dual reactivity of the system N-iodosuccinimide-H2SO4 in the iodination of deactivated arenes. *Russian Journal of Organic Chemistry* **2007,** 43, 1278-1281.

99. Sonogashira, K.; Tohda, Y.; Hagihara, N. A convenient synthesis of acetylenes: catalytic substitutions of acetylenic hydrogen with bromoalkenes, iodoarenes and bromopyridines. *Tetrahedron Letters* **1975,** 16, 4467-4470.

100. Omura, K.; Swern, D. Oxidation of alcohols by “activated” dimethyl sulfoxide. a preparative, steric and mechanistic study. *Tetrahedron* **1978,** 34, 1651-1660.

101. Kurti, L.; Czako, B. *Strategic Applications of Named Reactions in Organic Synthesis*. Elsevier Science: 2005.

102. Linden, J. J. M. v. d.; Hilberink, P. W.; Kronenburg, C. M. P.; Kemperman, G. J. Investigation of the Moffatt−Swern Oxidation in a Continuous Flow Microreactor System. *Organic Process Research & Development* **2008,** 12, 911-920.

103. Keirs, D.; Overton, K. Conversion of amines into imines by swern oxidation. *Journal of the Chemical Society, Chemical Communications* **1987**, 1660-1661.

104. Mori, N.; Togo, H. Facile oxidative conversion of alcohols to esters using molecular iodine. *Tetrahedron* **2005,** 61, 5915-5925.

105. McClintock, S. P.; Forster, N.; Herges, R.; Haley, M. M. Synthesis of α-Ketoester- and α-Hydroxyester-Substituted Isoindazoles via the Thermodynamic Coarctate Cyclization of Ester-Terminated Azo-Ene-Yne Systems. *The Journal of Organic Chemistry* **2009,** 74, 6631-6636.

106. Chinchilla, R.; Najera, C. Recent advances in Sonogashira reactions. *Chemical Society Reviews* **2011,** 40, 5084-5121.

107. Radhakrishnan, U.; Stang, P. J. Palladium-Catalyzed Arylation of Enynes and Electron-Deficient Alkynes Using Diaryliodonium Salts. *Organic Letters* **2001,** 3, 859-860.

108. King, A. O.; Okukado, N.; Negishi, E.-i. Highly general stereo-, regio-, and chemo-selective synthesis of terminal and internal conjugated enynes by the Pd-catalysed reaction of alkynylzinc reagents with alkenyl halides. *Journal of the Chemical Society, Chemical Communications* **1977**, 683-684.

109. Anastasia, L.; Negishi, E.-i. Highly Satisfactory Procedures for the Pd-Catalyzed Cross Coupling of Aryl Electrophiles with in Situ Generated Alkynylzinc Derivatives†. *Organic Letters* **2001,** 3, 3111-3113.

110. Hiroya, K.; Matsumoto, S.; Sakamoto, T. New Synthetic Method for Indole-2-carboxylate and Its Application to the Total Synthesis of Duocarmycin SA. *Organic Letters* **2004,** 6, 2953-2956.

111. Hiroya, K.; Matsumoto, S.; Ashikawa, M.; Kida, H.; Sakamoto, T. The optimization for cyclization reaction of 2-(2-carbomethoxyethynyl)aniline derivatives and formal synthesis of pyrroloquinoline quinone and its analogue utilizing a sequential coupling-cyclization reaction. *Tetrahedron* **2005,** 61, 12330-12338.

112. Yasuhara, A.; Kanamori, Y.; Kaneko, M.; Numata, A.; Kondo, Y.; Sakamoto, T. Convenient synthesis of 2-substituted indoles from 2-ethynylanilines with tetrabutylammonium fluoride. *Journal of the Chemical Society, Perkin Transactions 1* **1999**, 529-534.

113. Li, H.-Y.; Sun, H.; DiMagno, S. G. Tetrabutylammonium Fluoride. In *Encyclopedia of Reagents for Organic Synthesis*, John Wiley & Sons, Ltd: 2001.

114. Kondo, Y.; Kojima, S.; Sakamoto, T. General and Facile Synthesis of Indoles with Oxygen-Bearing Substituents at the Benzene Moiety. *The Journal of Organic Chemistry* **1997,** 62, 6507-6511.

115. Hiroya, K.; Jouka, R.; Kameda, M.; Yasuhara, A.; Sakamoto, T. Cyclization reactions of 2-alkynylbenzyl alcohol and 2-alkynylbenzylamine derivatives promoted by tetrabutylammonium fluoride. *Tetrahedron* **2001,** 57, 9697-9710.

116. Ezquerra, J.; Pedregal, C.; Lamas, C.; Barluenga, J.; Pérez, M.; García-Martín, M. A.; González, J. M. Efficient Reagents for the Synthesis of 5-, 7-, and 5,7-Substituted Indoles Starting from Aromatic Amines: Scope and Limitations. *The Journal of Organic Chemistry* **1996,** 61, 5804-5812.

117. Hay, A. S. Oxidative Coupling of Acetylenes. II1. *The Journal of Organic Chemistry* **1962,** 27, 3320-3321.

118. Okuma, K.; Seto, J.-i.; Sakaguchi, K.-i.; Ozaki, S.; Nagahora, N.; Shioji, K. Palladium-free zinc-mediated hydroamination of alkynes: efficient synthesis of indoles from 2-akynylaniline derivatives. *Tetrahedron Letters* **2009,** 50, 2943-2945.

119. Hashmi, A. S. K.; Hutchings, G. J. Gold Catalysis. *Angewandte Chemie International Edition* **2006,** 45, 7896-7936.

120. Arcadi, A.; Bianchi, G.; Marinelli, F. Gold(III)-Catalyzed Annulation of 2-Alkynylanilines: A Mild and Efficient Synthesis of Indoles and 3-Haloindoles. *Synthesis* **2004,** 2004, 610-618.

121. Hudlicky, M. *Reductions in Organic Chemistry*. Ellis Horwood Limited: 1986.

122. Desai, S.; Sun, W.; Gabriel, J.; Canney, D. J. The synthesis and preliminary evaluation of substituted chromones, coumarins, chromanones, and benzophenones as retinoic acid receptor ligands. *Heterocycl. Commun.* **2008,** 14, 129-136.

123. Kamm, O. β-Phenylhydroxylamine. *Org. Synth.* **1925,** IV, 57-8.

124. Clayden, J.; Greeves, N.; Warren, S. *Organic Chemistry*. OUP Oxford: 2012.

125. Beckwith, A. L. J.; Schiesser, C. H. Selectivity and synthetic applications of radical reactionsRegio- and stereo-selectivity of alkenyl radical ring closure: A theoretical study. *Tetrahedron* **1985,** 41, 3925-3941.

126. Hinman, R. L.; Whipple, E. B. The Protonation of Indoles: Position of Protonation. *Journal of the American Chemical Society* **1962,** 84, 2534-2539.

127. Höhne, G. W. H.; Hemminger, W.; Flammersheim, H. J. *Differential Scanning Calorimetry: An Introduction for Practitioners*. Springer Berlin Heidelberg: 2013.

128. Iwasawa, T.; Tokunaga, M.; Obora, Y.; Tsuji, Y. Homogeneous Palladium Catalyst Suppressing Pd Black Formation in Air Oxidation of Alcohols. *Journal of the American Chemical Society* **2004,** 126, 6554-6555.

129. Amatore, C.; Bahsoun, A. A.; Jutand, A.; Meyer, G.; Ndedi, N.; Ricard, L. Mechanism of the Stille Reaction Catalyzed by Palladium Ligated to Arsine Ligand: PhPdI(AsPh3)(DMF) Is the Species Reacting with Vinylstannane in DMF. *Journal of the American Chemical Society* **2003,** 125, 4212-4222.

130. Jacquemard, U.; Bénéteau, V.; Lefoix, M.; Routier, S.; Mérour, J.-Y.; Coudert, G. Mild and selective deprotection of carbamates with Bu4NF. *Tetrahedron* **2004,** 60, 10039-10047.

131. Bretherick, L. *Handbook of Reactive Chemical Hazards*. 4th ed.; Butterworth-Heinemann Ltd: Boston, 1990.

132. SciFinder databases has over 13340 entries for reactions where NaH is a reagent, and the solvent is DMF. (29/08/2015). In.

133. UK Chemical Reaction Hazards Forum. Sodium Hydride / DMF process stopped http://www.crhf.org.uk/incident101.html. (27 August).

134. Laird, T. SPECIAL FEATURE SECTION: SAFETY OF CHEMICAL PROCESSES. *Organic Process Research & Development* **2002,** 6, 876-876.

135. SciFinder database has over 15303 entries for reactions where NaH is a reagent, and the solvent is THF. (29/08/2015). In.

136. Kiran, E.; Sengers, J. M. H. L. *Supercritical Fluids: Fundamentals for Application*. Springer Netherlands: 2013.

137. Rajendran, A. Design of preparative-supercritical fluid chromatography. *Journal of Chromatography A* **2012,** 1250, 227-249.

138. Kalíková, K.; Šlechtová, T.; Vozka, J.; Tesařová, E. Supercritical fluid chromatography as a tool for enantioselective separation; A review. *Analytica Chimica Acta* **2014,** 821, 1-33.

139. MacMillan, K. S.; Boger, D. L. An Additional Spirocyclization for Duocarmycin SA. *Journal of the American Chemical Society* **2008,** 130, 16521-16523.

140. Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society* **1963,** 85, 2149-2154.

141. Hermkens, P. H. H.; Ottenheijm, H. C. J.; Rees, D. Solid-phase organic reactions: A review of the recent literature. *Tetrahedron* **1996,** 52, 4527-4554.

142. Mitchell, A. R. Bruce Merrifield and solid-phase peptide synthesis: A historical assessment. *Peptide Science* **2008,** 90, 175-184.

143. Merrifield, R. B. Solid Phase Peptide Synthesis. II. The Synthesis of Bradykinin. *Journal of the American Chemical Society* **1964,** 86, 304-305.

144. The Nobel Prize in Chemistry 1984 was awarded to Bruce Merrifield *"for his development of methodology for chemical synthesis on a solid matrix"*. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1984/. (30/08/2015).

145. Chan, W.; White, P. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*. OUP: Oxford, 2000.

146. Mäde, V.; Els-Heindl, S.; Beck-Sickinger, A. G. Automated solid-phase peptide synthesis to obtain therapeutic peptides. *Beilstein Journal of Organic Chemistry* **2014,** 10, 1197-1212.

147. Vaino, A. R.; Janda, K. D. Solid-Phase Organic Synthesis: A Critical Understanding of the Resin. *Journal of Combinatorial Chemistry* **2000,** 2, 579-596.

148. Davenport, K. G. Hydrogen Fluoride. In *Encyclopedia of Reagents for Organic Synthesis*, John Wiley & Sons, Ltd: 2001.

149. Chang, C.-D.; Meienhofer, J. SOLID-PHASE PEPTIDE SYNTHESIS USING MILD BASE CLEAVAGE OF NαFLUORENYLMETHYLOXYCARBONYLAMINO ACIDS, EXEMPLIFIED BY A SYNTHESIS OF DIHYDROSOMATOSTATIN. *International Journal of Peptide and Protein Research* **1978,** 11, 246-249.

150. Stawikowski, M.; Fields, G. B. Introduction to Peptide Synthesis. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* **2002,** CHAPTER, Unit-18.1.

151. Wang, S.-S. p-Alkoxybenzyl Alcohol Resin and p-Alkoxybenzyloxycarbonylhydrazide Resin for Solid Phase Synthesis of Protected Peptide Fragments. *Journal of the American Chemical Society* **1973,** 95, 1328-1333.

152. Carpino, L. A. 1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive. *Journal of the American Chemical Society* **1993,** 115, 4397-4398.

153. Carpino, L. A.; Imazumi, H.; El-Faham, A.; Ferrer, F. J.; Zhang, C.; Lee, Y.; Foxman, B. M.; Henklein, P.; Hanay, C.; Mügge, C.; Wenschuh, H.; Klose, J.; Beyermann, M.; Bienert, M. The Uronium/Guanidinium Peptide Coupling Reagents: Finally the True Uronium Salts. *Angewandte Chemie International Edition* **2002,** 41, 441-445.

154. Al-Warhi, T. I.; Al-Hazimi, H. M. A.; El-Faham, A. Recent development in peptide coupling reagents. *Journal of Saudi Chemical Society* **2012,** 16, 97-116.

155. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Analytical Biochemistry* **1970,** 34, 595-598.

156. Bottom, C. B.; Hanna, S. S.; Siehr, D. J. Mechanism of the ninhydrin reaction. *Biochemical Education* **1978,** 6, 4-5.

157. Vojkovsky, T. Detection of secondary amines on solid phase. *Pept Res* **1995,** 8, 236-7.

158. Ketcha, D. M.; Gribble, G. W. A convenient synthesis of 3-acylindoles via Friedel Crafts acylation of 1-(phenylsulfonyl)indole. A new route to pyridocarbazole-5,11-quinones and ellipticine. *The Journal of Organic Chemistry* **1985,** 50, 5451-5457.

159. Lanzilotti, A. E.; Littell, R.; Fanshawe, W. J.; McKenzie, T. C.; Lovell, F. M. Stereoselective reduction of some indoles with triethylsilane-trifluoroacetic acid. *The Journal of Organic Chemistry* **1979,** 44, 4809-4813.

160. Pearson, D. A.; Blanchette, M.; Baker, M. L.; Guindon, C. A. Trialkylsilanes as scavengers for the trifluoroacetic acid deblocking of protecting groups in peptide synthesis. *Tetrahedron Letters* **1989,** 30, 2739-2742.

161. Stanger, K. J.; Krchňák, V. Incorporation of the Wang Linker upon Cleavage from Polystyrene-based Resin to Form O-(4-Hydroxy)benzyl Derivatives. *Journal of Combinatorial Chemistry* **2006,** 8, 652-654.

162. Riniker, B.; Kamber, B.; In Jung, G.; Bayer, E. In *Byproducts of Trp-peptides synthesized on a p-benzyloxybenzyl alcohol polystyrene resin.*, Peptides 1988 (Proceedings of the 20th European Peptide Symposium), Berlin, 1989; Walter de Gruyter: Berlin, 1989; pp 115-117.

163. Rovero, P.; Pegoraro, S.; Viganò, S.; Bonelli, F.; Triolo, A. Solid support-dependent alkylation of tryptophan residues in SPPS using a 2-methoxybenzyl alcohol-based linker. *Letters in Peptide Science* **1994,** 1, 149-155.

164. Atherton, E.; Cameron, L. R.; Sheppard, R. C. Peptide synthesis. Part 10. Use of pentafluorophenyl esters of fluorenylmethoxycarbonylamino acids in solid phase peptide synthesis. *Tetrahedron* **1988,** 44, 843-857.

165. Fields, C. G.; Fields, G. B. Minimization of tryptophan alkylation following 9-fluorenylmethoxycarbonyl solid-phase peptide synthesis. *Tetrahedron Letters* **1993,** 34, 6661-6664.

166. Riniker, B.; Flörsheimer, A.; Fretz, H.; Sieber, P.; Kamber, B. A General Strategy for the Synthesis of Large Peptides: r1~he Combined Solid-Phase and Solution Approach. *Tetrahedron* **1993,** 49, 9307-9320.

167. Franzén, H.; Grehn, L.; Ragnarsson, U. Synthesis, properties, and use of Nin-boc-tryptophan derivatives. *Journal of the Chemical Society, Chemical Communications* **1984**, 1699-1700.

168. Kost, A. N.; Budylin, V. A.; Romanova, N. N.; Matveeva, E. D. Reaction of indoles with trifluoroacetic acid. *Chemistry of Heterocyclic Compounds* **1981,** 17, 921-923.

169. Boger, D. L.; Hertzog, D. L.; Bollinger, B.; Johnson, D. S.; Cai, H.; Goldberg, J.; Turnbull, P. Duocarmycin SA shortened, simplified, and extended agents: a systematic examination of the role of the DNA binding subunit. *J. Am. Chem. Soc.* **1997,** 119, 4977-4986.

170. Bieg, T.; Szeja, W. Removal of O-Benzyl Protective Groups by Catalytic Transfer Hydrogenation. *Synthesis* **1985,** 1985, 76-77.

171. Rajagopal, S.; Spatola, A. F. Mechanism of Palladium-Catalyzed Transfer Hydrogenolysis of Aryl Chlorides by Formate Salts. *The Journal of Organic Chemistry* **1995,** 60, 1347-1355.

172. Barlos, K.; Gatos, D.; Kallitsis, J.; Papaphotiu, G.; Sotiriu, P.; Wenqing, Y.; Schäfer, W. Darstellung geschützter peptid-fragmente unter einsatz substituierter triphenylmethyl-harze. *Tetrahedron Letters* **1989,** 30, 3943-3946.

173. Barlos, K.; Gatos, D.; Kapolos, S.; Papaphotiu, G.; Schäfer, W.; Wenqing, Y. Veresterung von partiell geschützten peptid-fragmenten mit harzen. Einsatz von 2-chlortritylchlorid zur synthese von Leu15 -gastrin I. *Tetrahedron Letters* **1989,** 30, 3947-3950.

174. Rovero, P.; Viganò, S.; Pegoraro, S.; Quartara, L. Synthesis of the bradykinin B1 antagonist [desArg10]HOE 140 on 2-chlorotrityl resin. *Letters in Peptide Science* **1996,** 2, 319-323.

175. Rink amide MBHA (http://www.merckmillipore.com/GB/en/product/Rink-Amide-MBHA-resin-(100-200-mesh),MDA_CHEM-855003). (30/08/15).

176. NovaSyn® TGA (http://www.merckmillipore.com/GB/en/product/NovaSynTGA-resin-(90-%c2%b5m),MDA_CHEM-855005). (30/08/15).

177. Wang, Z.; Yang, R.; Zhu, J.; Zhu, X. PEG-related polymer resins as synthetic supports. *Science China Chemistry* **2010,** 53, 1844-1852.

178. Coleman, R. S.; Shah, J. A. Chemoselective cleavage of benzyl ethers, esters, and carbamates in the presence of other easily reducible groups. *Synthesis* **1999**, 1399-1400.

179. Cory, A. H.; Owen, T. C.; Barltrop, J. A.; Cory, J. G. Use of an aqueous soluble tetrazolium/formazan assay for cell growth assays in culture. *Cancer Commun* **1991,** 3, 207-12.

180. Collins, S. J.; Gallo, R. C.; Gallagher, R. E. Continuous growth and differentiation of human myeloid leukaemic cells in suspension culture. *Nature* **1977,** 270, 347-9.

181. Birnie, G. D. The HL60 cell line: a model system for studying human myeloid cell differentiation. *The British Journal of Cancer. Supplement* **1988,** 9, 41-45.

182. Boger, D. L.; Bollinger, B.; Hertzog, D. L.; Johnson, D. S.; Cai, H.; Mésini, P.; Garbaccio, R. M.; Jin, Q.; Kitos, P. A. Reversed and Sandwiched Analogs of Duocarmycin SA: Establishment of the Origin of the Sequence-Selective Alkylation of DNA and New Insights into the Source of Catalysis. *Journal of the American Chemical Society* **1997,** 119, 4987-4998.

183. Lide, D. R. *CRC Handbook of Chemistry and Physics, 85th Edition*. Taylor & Francis: 2004.

184. Uchino, H.; Kanai, Y.; Kim, D. K.; Wempe, M. F.; Chairoungdua, A.; Morimoto, E.; Anders, M. W.; Endou, H. Transport of amino acid-related compounds mediated by L-type amino acid transporter 1 (LAT1): insights into the mechanisms of substrate recognition. *Mol Pharmacol* **2002,** 61, 729-37.

185. Robertson, W. M.; Kastrinsky, D. B.; Hwang, I.; Boger, D. L. Synthesis and evaluation of a series of C5'-substituted duocarmycin SA analogs. *Bioorg Med Chem Lett* **2010,** 20, 2722-5.

186. Hartley, J. A.; Wyatt, M. D. Determination of the DNA Sequence Specificity of Alkylation Damage Using Cleavage-Based Assays. In *Drug-DNA Interaction Protocols*, 1997; Vol. 90, pp 147-156.

187. Gates, K. S. An Overview of Chemical Processes That Damage Cellular DNA: Spontaneous Hydrolysis, Alkylation, and Reactions with Radicals. *Chemical research in toxicology* **2009,** 22, 1747-1760.

188. Lavesa, M.; Fox, K. R. Preferred binding sites for [N-MeCYs(3), N-MeCys(7)]TANDEM determined using a universal footprinting substrate. *Anal Biochem* **2001,** 293, 246-50.

189. Boger, D. L.; Johnson, D. S.; Yun, W. (+)- and ent-(-)-Duocarmycin SA and (+)- and ent-(-)-N-BOC-DSA DNA Alkylation Properties.Alkylation Site Models That Accommodate the Offset AT-Rich Adenine N3 Alkylation Selectivity of the Enantiomeric Agents. *Journal of the American Chemical Society* **1994,** 116, 1635-1656.

190. Wolfe, A. L.; Duncan, K. K.; Lajiness, J. P.; Zhu, K.; Duerfeldt, A. S.; Boger, D. L. A Fundamental Relationship between Hydrophobic Properties and Biological Activity for the Duocarmycin Class of DNA-Alkylating Antitumor Drugs: Hydrophobic-Binding-Driven Bonding. *Journal of Medicinal Chemistry* **2013,** 56, 6845-6857.

191. Drew, H. R.; Dickerson, R. E. Structure of a B-DNA dodecamer: III. Geometry of hydration. *Journal of Molecular Biology* **1981,** 151, 535-556.

192. Liepinsh, E.; Otting, G.; Wüthrich, K. NMR observation of individual molecules of hydration water bound to DNA duplexes: direct evidence for a spine of hydration water present in aqueous solution. *Nucleic Acids Research* **1992,** 20, 6549-6553.

193. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.;

Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, revision A.1*, Gaussian, Inc.: Wallingford, CT, USA, 2009.

194. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993,** 98, 5648-5652.

195. Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **1988,** 37, 785-789.

196. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *The Journal of Physical Chemistry* **1994,** 98, 11623-11627.

197. Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *The Journal of Chemical Physics* **1971,** 54, 724-728.

198. Arcamone, F.; Penco, S.; Orezzi, P.; Nicolella, V.; Pirelli, A. Structure and Synthesis of Distamycin A. *Nature* **1964,** 203, 1064-1065.

199. Finlay, A. C.; Hochstein, F. A.; Sobin, B. A.; Murphy, F. X. Netropsin, a New Antibiotic Produced by a Streptomyces. *Journal of the American Chemical Society* **1951,** 73, 341-343.

200. Walker, W. L.; Kopka, M. L.; Goodsell, D. S. Progress in the design of DNA sequence-specific lexitropsins. *Biopolymers* **1997,** 44, 323-34.

201. Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proceedings of the National Academy of Sciences* **1985,** 82, 1376-1380.

202. Wellenzohn, B.; Flader, W.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R. Significance of ligand tails for interaction with the minor groove of B-DNA. *Biophysical Journal* **2001,** 81, 1588-1599.

203. Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. Binding of an antitumor drug to DNA: Netropsin and C-G-C-G-A-A-T-T-BrC-G-C-G. *Journal of Molecular Biology* **1985,** 183, 553-563.

204. Lown, J. W.; Krowicki, K.; Balzarini, J.; De Clercq, E. Structure-activity relationship of novel oligopeptide antiviral and antitumor agents related to netropsin and distamycin. *Journal of Medicinal Chemistry* **1986,** 29, 1210-1214.

205. Kissinger, K.; Krowicki, K.; Dabrowiak, J. C.; Lown, J. W. Molecular recognition between oligopeptides and nucleic acids. Monocationic imidazole lexitropsins that display enhanced GC sequence dependent DNA binding. *Biochemistry* **1987,** 26, 5590-5595.

206. Burckhardt, G.; Luck, G.; Zimmer, C.; Störl, J.; Krowicki, K.; Lown, J. W. Variation of DNA sequence specificity of DNA-oligopeptide bindings ligands related to netropsin: imidazole-containing lexitropsins. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1989,** 1009, 11-18.

207. Pelton, J. G.; Wemmer, D. E. Structural characterization of a 2:1 distamycin A.d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc Natl Acad Sci U S A* **1989,** 86, 5723-7.

208. Wade, W. S.; Mrksich, M.; Dervan, P. B. Design of peptides that bind in the minor groove of DNA at 5'-(A,T)G(A,T)C(A,T)-3' sequences by a dimeric side-by-side motif. *Journal of the American Chemical Society* **1992,** 114, 8783-8794.

209. Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proceedings of the National Academy of Sciences* **1992,** 89, 7586-7590.

210. Wade, W. S.; Mrksich, M.; Dervan, P. B. Binding affinities of synthetic peptides, pyridine-2-carboxamidonetropsin and 1-methylimidazole-2-carboxamidonetropsin, that form 2:1 complexes in the minor groove of double-helical DNA. *Biochemistry* **1993,** 32, 11385-11389.

211. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **1998,** 391, 468-471.

212. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. A Structural Basis for Recognition of A·T and T·A Base Pairs in the Minor Groove of B-DNA. *Science* **1998,** 282, 111-115.

213. Mrksich, M.; Dervan, P. B. Design of a Covalent Peptide Heterodimer for Sequence-Specific Recognition in the Minor Groove of Double-Helical DNA. *Journal of the American Chemical Society* **1994,** 116, 3663-3664.

214. Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B. Cycle Polyamide Motif for Recognition of the Minor Groove of DNA. *Journal of the American Chemical Society* **1999,** 121, 1121-1129.

215. Mrksich, M.; Parks, M. E.; Dervan, P. B. Hairpin Peptide Motif. A New Class of Oligopeptides for Sequence-Specific Recognition in the Minor Groove of Double-Helical DNA. *Journal of the American Chemical Society* **1994,** 116, 7983-7988.

216. Kelly, J. J.; Baird, E. E.; Dervan, P. B. Binding site size limit of the 2:1 pyrrole-imidazole polyamide-DNA motif. *Proc Natl Acad Sci U S A* **1996,** 93, 6981-5.

217. Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. Structural basis for G.C recognition in the DNA minor groove. *Nat Struct Biol* **1998,** 5, 104-9.

218. Swalley, S. E.; Baird, E. E.; Dervan, P. B. A Pyrrole-Imidazole Polyamide Motif for Recognition of Eleven Base Pair Sequences in the Minor Groove of DNA. *Chemistry – A European Journal* **1997,** 3, 1600-1607.

219. Herman, D. M.; Baird, E. E.; Dervan, P. B. Tandem Hairpin Motif for Recognition in the Minor Groove of DNA by Pyrrole–Imidazole Polyamides. *Chemistry – A European Journal* **1999,** 5, 975-983.

220. Kers, I.; Dervan, P. B. Search for the optimal linker in tandem hairpin polyamides. *Bioorg Med Chem* **2002,** 10, 3339-49.

221. Weyermann, P.; Dervan, P. B. Recognition of Ten Base Pairs of DNA by Head-to-Head Hairpin Dimers. *Journal of the American Chemical Society* **2002,** 124, 6872-6878.

222. de Clairac, R. P. L.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. NMR Characterization of Hairpin Polyamide Complexes with the Minor Groove of DNA. *Journal of the American Chemical Society* **1997,** 119, 7909-7916.

223. Herman, D. M.; Baird, E. E.; Dervan, P. B. Stereochemical Control of the DNA Binding Affinity, Sequence Specificity, and Orientation Preference of Chiral Hairpin Polyamides in the Minor Groove. *Journal of the American Chemical Society* **1998,** 120, 1382-1391.

224. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. Aliphatic/Aromatic Amino Acid Pairings for Polyamide Recognition in the Minor Groove of DNA. *Journal of the American Chemical Society* **1998,** 120, 6219-6226.

225. Nguyen, D. H.; Szewczyk, J. W.; Baird, E. E.; Dervan, P. B. Alternative heterocycles for DNA recognition: an N-methylpyrazole/N-methylpyrrole pair specifies for A.T/T.A base pairs. *Bioorg Med Chem* **2001,** 9, 7-17.

226. Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. Regulation of gene expression by small molecules. *Nature* **1997,** 387, 202-5.

227. Kang, J. S.; Meier, J. L.; Dervan, P. B. Design of Sequence-Specific DNA Binding Molecules for DNA Methyltransferase Inhibition. *Journal of the American Chemical Society* **2014,** 136, 3687-3694.

228. Sugiyama, H.; Lian, C.; Isomura, M.; Saito, I.; Wang, A. H.-J. Distamycin A modulates the sequence specificity of DNA alkylation by duocarmycin A. *Proceedings of the National Academy of Sciences of the United States of America* **1996,** 93, 14405-14410.

229. Tao, Z.-F.; Fujiwara, T.; Saito, I.; Sugiyama, H. Sequence-Specific DNA Alkylation by Hybrid Molecules between Segment A of Duocarmycin A and Pyrrole/Imidazole Diamide. *Angewandte Chemie International Edition* **1999,** 38, 650-653.

230. Bando, T.; Sasaki, S.; Minoshima, M.; Dohno, C.; Shinohara, K.; Narita, A.; Sugiyama, H. Efficient DNA alkylation by a pyrrole-imidazole CBI conjugate with an indole linker: sequence-specific alkylation with nine-base-pair recognition. *Bioconjug Chem* **2006,** 17, 715-20.

231. Tao, Z.-F.; Saito, I.; Sugiyama, H. Highly Cooperative DNA Dialkylation by the Homodimer of Imidazole−Pyrrole Diamide−CPI Conjugate with Vinyl Linker. *Journal of the American Chemical Society* **2000,** 122, 1602-1608.

232. Yamamoto, M.; Bando, T.; Kawamoto, Y.; Taylor, R. D.; Hashiya, K.; Sugiyama, H. Specific alkylation of human telomere repeat sequences by a tandem-hairpin motif of pyrrole-imidazole polyamides with indole-seco-CBI. *Bioconjug Chem* **2014,** 25, 552-9.

233. Chang, A. Y.; Dervan, P. B. Strand Selective Cleavage of DNA by Diastereomers of Hairpin Polyamide-seco-CBI Conjugates. *Journal of the American Chemical Society* **2000,** 122, 4856-4864.

234. Baird, E. E.; Dervan, P. B. Solid Phase Synthesis of Polyamides Containing Imidazole and Pyrrole Amino Acids. *Journal of the American Chemical Society* **1996,** 118, 6141-6146.

235. Wurtz, N. R.; Turner, J. M.; Baird, E. E.; Dervan, P. B. Fmoc Solid Phase Synthesis of Polyamides Containing Pyrrole and Imidazole Amino Acids. *Organic Letters* **2001,** 3, 1201-1203.

236. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B. Solid-Phase Synthesis of DNA Binding Polyamides on Oxime Resin. *Bioorganic & Medicinal Chemistry* **2002,** 10, 2767-2774.

237. Ikeda, M.; Nakagawa, H.; Ban, S.; Tsumoto, H.; Suzuki, T.; Miyata, N. Development of a DNA-binding TEMPO derivative for evaluation of nuclear oxidative stress and its application in living cells. *Free Radical Biology and Medicine* **2010,** 49, 1792-1797.