Studies towards a concise synthesis of the duocarmycin SA subunit for antibody directed therapy

Farah Baig

A thesis submitted for the degree of
Doctor of Philosophy

UEA
University of East Anglia

University of East Anglia
School of Pharmacy

December 2009

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.

Farah Baig

Acknowledgements

Thank you to my supervisor, Dr. Mark Searcey.

Thank you to those I came across during the course of my research, at The School of Pharmacy, University of London, and University of East Anglia, for their help and support, and many excellent friendships that have formed.

Thank you to my family.

Thank you to BBSRC, and Dr Tim Norman at UCB Pharma for funding.

Abstract

Duocarmycin SA, a potent antitumour agent derived from *Streptomyces sp*, is an excellent candidate drug for development. High associated toxicity of this compound as well as inefficient synthetic routes, have limited clinical development. We aim to optimise the synthetic conditions rendering commercially viability, and use conjugation strategies for targeting and specificity, thereby reducing non-specific toxicity.

CI, the minimum pharmacophore has been synthesised as a model structure for initial conjugation studies. For synthesis of DSA, the key intermediate structure and active pharmacophore for duocarmycin SA, free radical allylation, Fischer cyclisation, azide chemistry, and organopalladium catalysis strategies for indole synthesis have been explored.

Synthesis of DSA involves construction of the first aromatic heterocyclic ring through Negishi coupling, utilising *in situ* palladium-ligand complex formation in the microwave. We optimised Negishi coupling by studies into palladium and ligand choice, and the use of microwave irradiation in comparison with solution phase methodology.

Installation of the full duocarmycin SA alkylation subunit structure will provide the core for generating analogues with enhanced activity. We intend to extend this further.
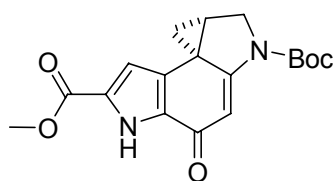
Boc - DSA

Abbreviations

Ac	acetyl
ADC	antibody-drug conjugate
ADEPT	antibody directed enzyme prodrug therapy
AML	acute myeloid leukaemia
anhyd.	anhydrous
aq.	aqueous
Bn	benzyl
BnBr	benzyl bromide
Boc	*tert*-butyloxycarbonyl
*t*BuLi	tert-butyl lithium
CBI	cyclopropabenzo[e]indolone
CI	cyclopropylindolone
CPI	cyclopropylpyrrolo[e]indolone
CTI	cyclopropylthioindolone
conc.	concentrated
DCM	dichoromethane
DHA	docosahexaenoic acid
DIBAL	diisobutylaluminium hydride
DIC	diisopropylcarbodiimide
DMA	N,N-dimethylacetamide
DMAP	dimethylaminopyridine
DMF	dimethyl formamide
DMSO	dimethylsulphoxide
DSA	alkylation subunit of duocarmycin SA
EGFR	epidermal growth factor receptor

EPR	enhanced permeation & retention
EtOAc	ethyl acetate
GlcNAc	N-acetyl-D-glucosamine
GlcUA	D-glucuronic acid
h	hour
HA	hyaluronic acid
HER2	human epidermal growth factor receptor type 2
HIV	human immunodeficiency virus
HPLC	high pressure liquid chromatography
HPMA	hydroxypropyl methacrylamide
IR	infra-red
M	molar concentration
mAb	monoclonal antibody
MeOH	methanol
min	minute
mmol	millimoles
mp	melting point
Ms	mesyl
MsCl	mesyl chloride
NaHMDS	sodium bis(trimethylsilyl)amide
NBS	N bromosuccinimide
NIS	N iodosuccimide
NMR	nuclear magnetic resonance
PDGF	platelet-derived growth factor
PE	petroleum ether
P(o-Tol)$_3$	tri-o-tolyl phosphine
PPA	polyphosphoric acid
PPh$_3$	triphenylphosphine

iPr$_2$NEt   *N* ethyldiisopropylamine

PUFA   polyunsaturated fatty acid

RES   reticuloendothelial system

R$_f$ retention factor

RT   room temperature

sat.   saturated

SM   starting material

TBAI   tetra-*N*-butyl ammonium iodide

TBDMS   tert-butyldimethylsilyl

TEA   triethylamine

TFA   trifluoroacetic acid

THF   tetrahydrofuran

TLC   thin layer chromatography

TMI   5, 6, 7- trimethoxyindole

TPPTS   tris(3-sulfonatophenyl) phosphane trisodium salt

TXPTS   tris(2,4-xylyl)3-sulfonatophenyl phosphane trisodium salt

UV   ultra violet

Contents

1. Introduction part 1: Cancer and targeting

1.1. The biology of cancer

Malignant disease or cancer is a heterogenous disease characterised by cell hyperproliferation, forming tumours which then metastasise to other parts of the body. Tumour growth has a destructive effect, causing both damage and physiological loss of function to the organs affected and to surrounding tissues. Chaotic growth results in high consumption of nutrients by diversion from surrounding tissues and organs. This is primarily facilitated by vascularisation of the cancerous tissue i.e. angiogenesis, ensuring a continued supply of nutrients and continued tumour growth.[1][2]

Cell hyperproliferation and disruption of apoptosis (programmed cell death) can be caused by a number of factors, including deregulation of cell signalling and loss of immune response. The cell cycle becomes disrupted by mutations to key genes such as bcl-2 which is responsible for prolonging cell survival by inhibiting apoptosis, and p53 which is responsible for generation of cell death signals.[1][2]

1.2. Environmental causes

The aetiology of malignant development is often unknown, however both genetic and environmental factors can contribute to the molecular causes of cancer. Known environmental triggers include exposure to tobacco, radiation, chemicals, and infectious organisms.[3] Smoking is strongly associated with not only lung cancer but also cancer of the mouth, larynx, oesophagus and bladder.[3][4] Alcohol is associated with upper respiratory and

gastrointestinal tract cancers, with a possible link to breast cancer.[2] Exposure to UV light is known to increase the risk of skin cancer.[3][4] Dietary factors have also been implicated however these are difficult to differentiate from epidemiological ones. The role of infectious disease in initiating cancer is becoming more apparent, particularly in the immunocompromised. It has been shown that HIV patients have a high incidence of Epstein Barr Virus-related lymphoma and Kaposi's sarcoma. Increasing incidence of cervical cancer may be due to the papilloma virus, hepatitis is associated with liver cancer, and *Heliobactor pylori* infection is recognised as a causative factor in gastric malignancies.[2][4]

1.3. The statistics

Cancer is a highly relevant and urgent problem in the twenty-first century. As medical advances have been able to deal with most other illnesses, mental health and malignant disease appear more prominent. In Western countries, approximately a third of the population will develop cancer at some point in their lifetime. Worldwide, there were 11 million new cases diagnosed in 2002, a quarter of which were in Europe.[5] One in eight deaths in the world is due to cancer, a mortality rate second only to cardiovascular disease.[6][7]

In the UK, approximately 300,000 new cases are diagnosed each year. Cancers of the breast, lung, bowel and prostate are the most prevalent. Potentially avoidable cancers such as malignant melanomas, uterine and kidney cancers are becoming increasingly common. Smoking related lung cancer contributes to the relatively high cancer mortality rate recorded in England, Scotland and Wales.[2]

1.4. Traditional treatment strategies

Cancer is essentially a moving target as increased mutations and natural selection of "the survival of the fittest" governs that the most resilient and toughest tumours survive and thrive. Initiation of treatment requires detection of the cancer, which can itself be a problem. Often, diagnosis can not be made early enough, as small tumour masses (when growth fraction is high) are undetectable with current technology. Alternatively, symptoms may manifest themselves only at a late stage of development. This has led to current strategies focussing on the removal of tumour masses and the prevention of metastasis. Typical therapies include surgery where tumour tissue is removed and chemotherapy particularly in metastatised cancers. Irradiation can be used as pre-treatment to shrink the tissue prior to surgery. The aim of any treatment should be to remove cancerous tissue without harming any normal tissue.

Historically, non-targeted therapy has been based on screening compound libraries or natural products against cancer cell lines, often without an in depth knowledge of the target or chemical and biological function of the therapies.[9] Chemotherapeutic agents employ various mechanisms of action to target tumour cells (Figure 1).

Many aspects of the cell's regular processes can be affected or hindered. Nucleotide manufacture inhibition adversely affects DNA replication, interfering with gene transcription and preventing the formation of the mitotic spindle, such that cell division fails. Enzymes of DNA synthesis can be inhibited; antimetabolites are usually structural analogues of naturally occurring metabolites which interfere with normal nucleic acid synthesis by false base substitution into metabolic pathways. Classic examples are

methotrexate and the fluoropyrimidines (5-fluorouracil). DNA function can be inhibited; alkylating agents e.g. cyclophosphamide, bind covalently to DNA bases. This cross-linking interferes with DNA synthesis. Inhibitors of topoisomerase I and II (etoposide) inhibit the unwinding of supercoiled DNA. Non-classical alkylators (cisplatin) cause intra-strand DNA cross-links. Cytotoxic antibiotics (doxorubicin) intercalate adjacent nucleotide base pairs on the same DNA strand. DNA transcription is targeted by agents such as actinomycin-D. Microtubule assembly is inhibited by plant alkaloids (vincristine) acting as tubulin binders. Taxanes work in a similar fashion. Effective combinations are achieved by using drugs that target DNA synthesis and microtubule formation together.

Side effects commonly experienced with current chemotherapy are a significant limiting factor. These side effects arise due to the non-specific action of cytotoxic drugs on all dividing cells. Common side effects are nausea, vomiting, alopecia, and myelosuppression. Rarely, toxicities idiosyncratic to specific drugs can present problems such as cardiotoxicity, neurotoxicity, nephrotoxicity, and even secondary malignancies. Relapse is also common due to hypoxic tissue reoxygenating. This tissue can then begin to hyperproliferate once again, so that the cancerous tissue continues to grow. This can render therapies inadequate. Thus, the combination of intolerable side effects and non-specific targeting results in reduced effectiveness of current chemotherapies.
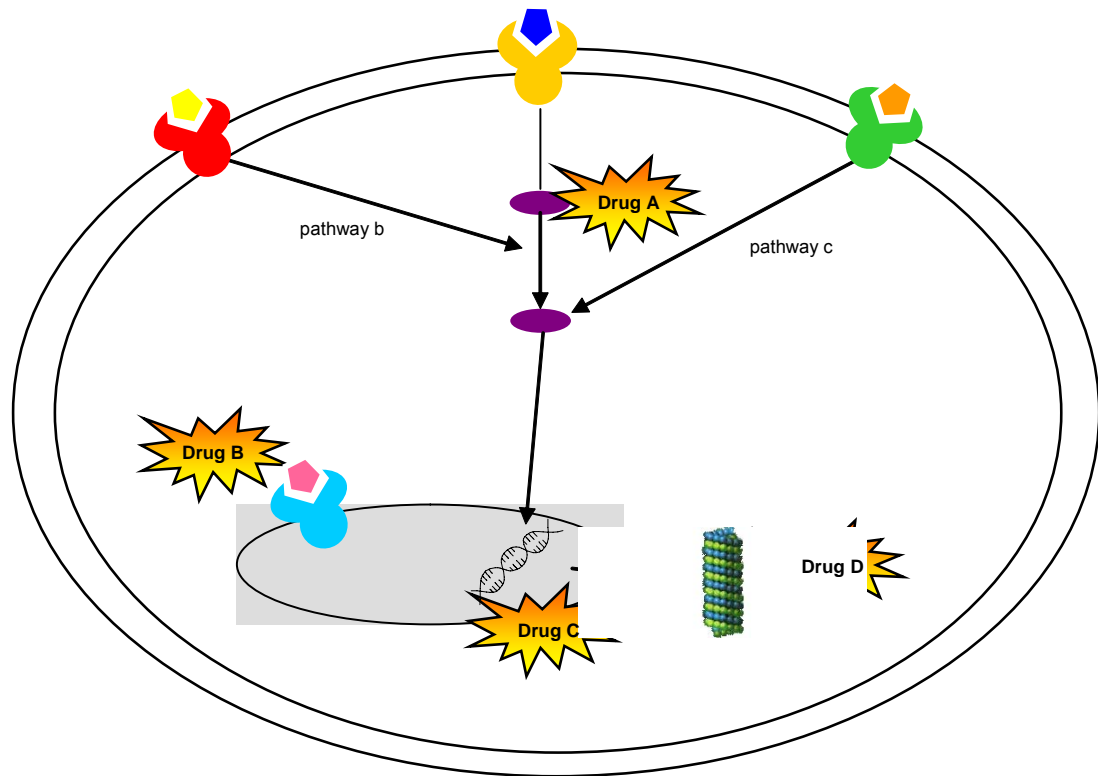
Figure 1 Subcellular effects of therapies. Abnormal transformations activate signal transduction pathways resulting in DNA synthesis, mitotic spindle assembly, cell division and viability. Targeting upstream parts of signal transduction pathways can leave susceptibility to resistance developing further downstream. Drug A targets an activated oncogene tyrosine kinase pathway, and is susceptible to resistance if pathway b or c is activated. Drug B targets a nuclear receptor, directly interacting with DNA. Drug C targets DNA directly. Drug D targets microtubules. All these are less susceptible to upstream counteraction.[8] [9]

1.5. Targeted therapy

The ultimate aim of any therapy is to either destroy tumour cells, or prevent them from proliferating further. By targeting therapy i.e. delivering the cytotoxic agent directly and specifically to the target site, side effects can be reduced, and selective action can be achieved. Targeted therapy and structure-based design is not a novel concept. For example, Elion and Hitching's work on nucleic acid synthesis led to the development of 6-mercaptopurine and 6-thioguanine which interfere with purine nucleotide

synthesis.[9] [10] [11] A discussion of ways to exploit characteristics unique to the target site and agents that can be utilised for targeting follows.

1.5.1 Exploiting the target

The tumour area's characteristics can be used as target sites. To arrest cell hyperproliferation, one approach is to directly target the cell cycle, either by preventing the formation of new cells or inducing cell death. Some recently developed targeted therapies act at the sub-cellular plasma membrane. These selectively and directly affect the signalling pathways which are responsible for abnormalities, rather than the end product of the pathway. These therapies have been found to be more effective in haemopoietic cancers than solid tumours due to ease of reaching the target area. However, their upstream location renders them more susceptible to resistance developing downstream. In this respect, classical agents acting directly at the end points of signalling pathways have an advantage.[9]

Protein synthesis or degradation can also be targeted to disrupt the cell cycle. The immunosuppressive activity of the macrolide antibiotic rapamycin derives from endothelial cell inhibition due to decreases in VEGF (vascular epithelial growth factor) production. It also inhibits mTOR (mammalian target of rapamycin), controlling cell growth and proliferation. mTOR has been found to be an upstream activator of hypoxia-inducible factor 1 (HIF-1) in tumour cells. Therefore, rapamycin's activity may be derived from inhibiting cellular responses to hypoxic stress.[12] Kinases, involved in protein modification, are also an intergral part of cell signalling. Tumour cell growth may result from cyclin overexpression, low levels or absence of cyclin-dependant kinase (cdk) inhibitor. In addition to regulating the cell cycle and other cellular processes by monitoring and facilitating the turnover of

proteins, the ubiquitin proteasome pathway is the principal mechanism for damaged or abnormal protein repair and degradation. Breakdown of this process may affect tumour growth and apoptosis. Proteasome inhibitors such as bortezomib increase cyclin-dependent kinase inhibitor levels causing G2-M cell cycle arrest followed by tumour cell apoptosis.[12]

Unregulated angiogenesis, i.e. the formation of new vasculature, is seen in pathologies such as cancer. Tumour cells more than 100 μm away from blood vessels become hypoxic, and acquire oxygen and nutrients by diffusion, ensuring viability and growth. If new blood vessels do not form, tumours are typically confined to 1-1.5 mm diameter.[13] As the tumour grows, vascularisation becomes increasingly chaotic, with hypervascularisation at the peripheries, and hypoxia and necrosis in the centre. Angiogenesis also provides a gateway for tumour cells to enter systemic circulation and metastasis.

Folkman first published the idea of using angiogenesis to target tumours thirty years ago.[14] Different strategies can be employed to inhibit tumour angiogenesis; interference of angiogenic factors or their receptors, inhibition of endothelial cell proliferation or adhesion, or of matrix metalloproteinases.

Amongst all growth factors, VEGF and bFGF (basic fibroblast growth factor) play a more prominent role in angiogenesis (Figure 2). Bevacizumab (Avastin), a VEGF-binding antibody, became the first anti-angiogenesis agent to be given FDA (Federal Drug Agency, regulates drug approval in the US) approval for colorectal cancer in 2004.[13] Thalidomide, infamous for its unexpected teratogenicity in pregnant women, has also found new applications in cancer therapy as an antiangiogenesis agent. Thalidomide down regulates VEGF expression and bFGF, suppressing tumour necrosis

factor-α (TNF-α) production and modulating cytokine responses. Its benefit is limited by cumulative and dose-dependent toxicities of neuropathy, fatigue, constipation and sedation.[15][16]
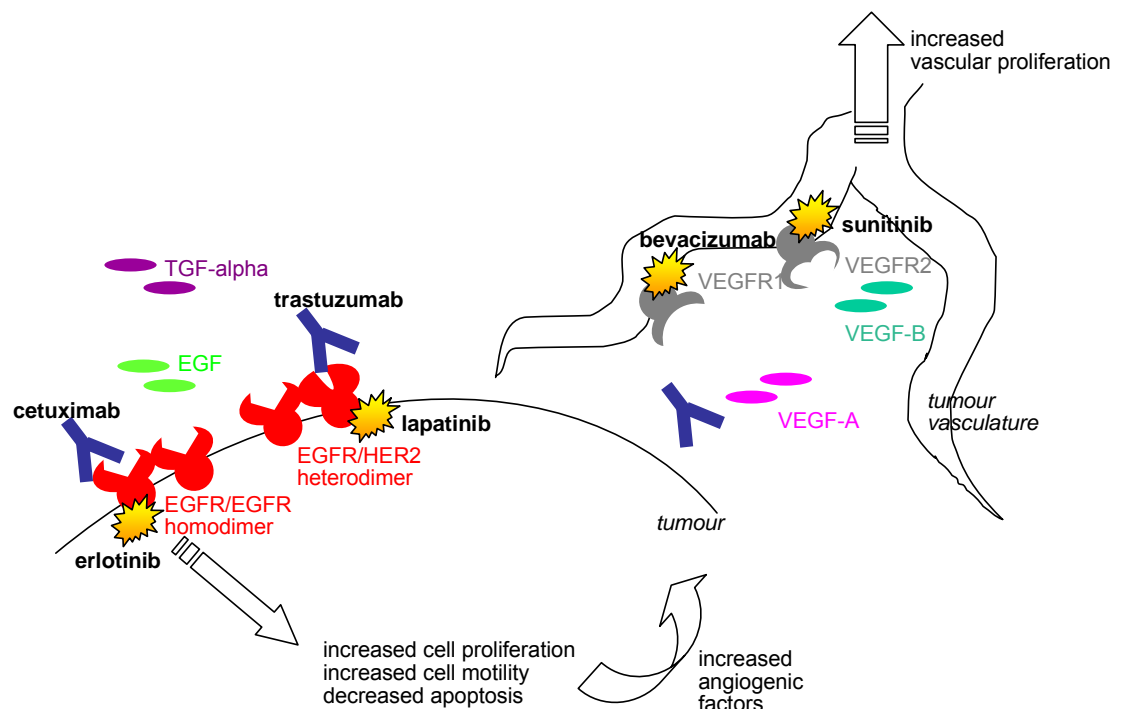
Figure 2 Antitumour agents targeting growth factors as targets of therapy (EGF=epidermal growth factor, TGFα=transforming growth factor-α).[17]

Hypoxia in tumours can also be utilised as a targeting characteristic. As tumour vasculature chaotically develops, oxygen delivery can become disrupted resulting in hypoxic cells of varying degrees depending on the distance from the vasculature. These hypoxic cells become problematic particularly in treatment response of solid tumours. Bioreductive drugs can be activated to their cytotoxic forms by enzymes that only exist in hypoxic environments.
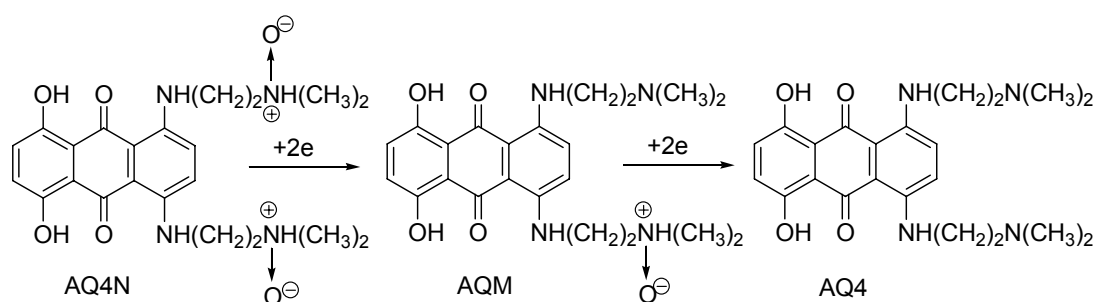
Scheme 1 AQ4N and its reduced active metabolites [18]

AQ4N (di-N-oxide of 1,4-bis[{2-(dimethylaminoethyl}-amino]5,8-dihydroxyanthracene-9,10-dione (AQ4) is an example of this approach (Scheme 1). The reduced active metabolite AQ4 essentially has topoisomerase II activity, binding to DNA with great affinity, stabilising the DNA helix due to its alkyl amino side chains. Interestingly, AQ4N is additionally metabolised and hence activated by certain cytochrome P450 (CYP) isoforms. Thus, the presence of CYP isoforms in tumours enables highly selective activity.[18]

Due to the multidimensional multifactorial nature of cancer, where a clear target is often not apparent, it is difficult to present one approach as superior to another. Inhibiting clear targets is not a guarantee of clinical efficacy. A target may be present in a malignancy, but in the clinical setting, patients are non-responsive to it. It may also be that merely a single target is not sufficient, and relative non-selectivity may be therapeutically advantageous. It is hoped that advances in understanding the biology of cancer, biomarkers, target behaviour, advances in genomics and proteomics will contribute to further advances in cancer therapy. Drug resistance can then also be understood and overcome, retaining activity of therapies.

1.5.2 Old drugs with new tricks – using targeting agents to enhance drug delivery

The emergence of new therapeutically efficacious anti-cancer agents from discovery and synthesis is slow. Therefore, existing cytotoxics need to be improved by controlling their activity, minimising side effects, resistance, and improving their delivery and targeting. One approach is by mediated targeting, by antibodies, ligands or other means. Prodrugs, drugs in their inactive form can be delivered to the target by a conjugated agent. The prodrug only becomes activated to its cytotoxic form on internalisation at the target site, either by release from the conjugated agent, or tumour environment specific activation.

1.5.2.1 Carrier systems

Conjugation can take the form of a carrier system, improving solubility profiles. Carriers can facilitate the drug's entry into the appropriate tissue or cellular compartment by enabling the traversing of membranes, or prolonging systemic circulation to allow sufficient uptake at the target site. Hydrophobic compounds can pass across the plasma membrane, but may have poor aqueous solubility so that they may not necessarily reach the membrane without assistance. In contrast, hydrophilic compounds have difficulty permeating across membranes, so the use of endocytosis or conjugation is required.

Surfactants can be used for drug delivery, particularly to improve the solubility of hydrophobic agents such as taxanes. Biodegradable water-soluble polymers such as copolymer hydroxypropyl methacrylamide (HPMA) are also being explored as targeted drug carriers. Biodegradable

spacers such as short peptides need to be inserted to ensure that the drug is released from the polymer.[19]

Microreservoir carrier systems such as liposomes rely on a relatively passive method of targeting based on anatomical characteristics of tumour tissue. Permeation of tumour tissue is via the enhanced permeation & retention (EPR) effects of leaky tumour vasculature, enabling accumulation and sustained release over time. Liposomes can be loaded with the cytotoxic drug, with targeting ligands attached at the liposome surface. Post-insertion means that any drug can be incorporated into ready made liposomes without requiring individual specific manufacture for each therapy. Liposomes have the advantage of achieving very high loads of drug to ligand ratios, increasing drug loading capacity relative to other carrier systems. They can become relatively large in size (usually 100 nm diameter) which, on one hand can result in slower clearance increasing the chance of target tissue uptake, and on the other hand can result in poor target tissue penetration. Slow release of the drug can be achieved also by altering the pharmacokinetic profile of the liposome. The anthracyclines doxorubicin and daunorubicin are established anticancer agents, however, cumulative cardiotoxicity limits their use. This has been reduced by encapsulating the drug in liposomes.[20] A liposomal formulation of doxorubicin (Doxil or Caelyx) is an established licensed product.[19]

1.5.2.2 The use of antibodies as targeting agents

Antibodies are endogenous proteins involved in the body's immune response, roughly Y shaped with three parts loosely connected by a flexible linkage (Figure 3). They have a dual role of binding to antigens at the variable region, and binding to effector molecules and receptors.

Antibodies have come under intense interest due to their antigen-specificity conferring good targeting properties as "naked" i.e. active entities in themselves, or as antibody-drug conjugate contructs. This could potentially fulfil the role of the "magic bullet," a term coined by Paul Ehrlich referring to agents that could specifically target and destroy tumour cells as well as microorganisms, establishing the ideas for much of current immunoconjugate development.[21] [22]
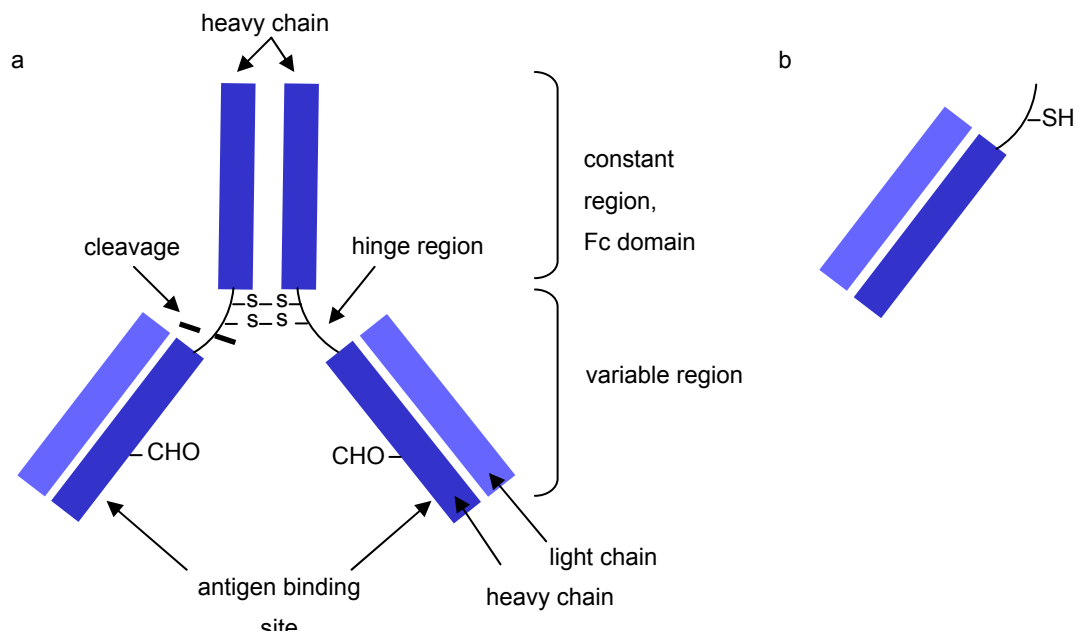

Figure 3 (a). antibody structure (b). Fab fragment

1.5.2.2.1 Antibodies as direct cytotoxic agents

Antibodies have been used as cytotoxic agents in their own right. Rituximab (Rituxan), an antibody targeting CD20 B cell antigens used in non-Hodgkins lymphoma, was the first to be approved in the US in 1997.[23]

Trastuzumab (Herceptin) used in breast cancer, became well known in the UK because of contraversies over high cost resulting in restricted availability on the NHS, despite the urgent demand for breast cancer therapies. In 1998, breast cancer caused over 11,000 deaths in England and Wales. This was the leading cause of death in women aged 35 to 54 years.[24] Herceptin is a humanised EGFR type 2 (HER2)-related monoclonal antibody targeting the oncogenic receptor tyrosine kinase (Her-2/Neu) on tumour cells. As HER-2 is found to be overexpressed in only 30% of invasive breast cancers,[9][17] it is often given in combination with other chemotherapy drugs such as the taxols. The main side effects are cardiotoxicity and infusion-related.[25]

The disadvantages of using antibodies are that manufacture is expensive and time-consuming, and stability and storage may be an issue. Immunogenicity of the early rodent antibodies was a major problem although this is being rectified with the use of antibodies with an increasingly human framework.[19] [26] However, idiopathic responses to antibodies can not be predicted.

Interest in immunoconjugates has also led to the development of radiolabelled monoclonal antibodies. As radiation can also destroy normal cells, it is essential that there is a high specificity for the target. β particles have millimetre penetration whereas α particles only penetrate within a range of a few cells and thus are only suitable for micrometastases or circulating tumour cells.[19] [27] [28] Anti-CD20 radioimmunoconjugates for treatment of lymphoma are in the most advanced stages of development. Zevalin, licensed in the UK, comprises of the antibody ibritumomab linked to the radioactive isotope ^{90}Y by covalent bonding of the chelator tiuxetan. Once the conjugate has been internalised, β-emission induces cellular damage via free radical formation in the target and neighbouring cells

(bystander effect).[29]   Other radioimmunotherapeutics such as with ^{131}I-tositumomab (Bexxar) are currently in clinical trials.[19]

1.5.2.2.2   Conjugation of antibodies to other moieties

The concept of antibody-drug conjugates (ADCs) is based upon conjugation of antibodies to drugs by means of a chemical linker, and to employ these ADCs as "homing missiles," reaching tumour specific antigens. This aims to improve the tumour-to-normal tissue selectivity of chemotherapy, reducing systemic toxicity.

One early type of this antibody-conjugate construct has been antibody directed enzyme prodrug therapy (ADEPT), a two step approach to targeting. Antibodies localise enzymes at the target site as they recognise tumour specific antigens on the cell surface. The prodrug is administered afterwards, and will only be activated to the cytotoxic drug by the enzyme localised at the tumour sites. Enzymes investigated in the past for ADEPT are β lactamases cleaving β lactam rings in paclitaxel and doxorubicin, cytosine deaminase activating 5-fluorouracil, and carboxypeptidase G2 cleaving terminal glutamic acid amides in nitrogen mustards.[19]

Figure 4 illustrates the pathway of the ADC to exert its cytotoxic activity at the target tumour site.[30]   The ADC is delivered intravenously, localises at tumour sites, and binds to a target antigen on the tumour cell surface. The ADC internalises, fuses with other vesicles and enters the endosome-lysosome pathway. Proteases in the endosome acidic environment digest the antibody and possibly the linker, releasing free drug. The drug then crosses the membrane to enter the cytoplasm, where it binds to its molecular target, resulting in cell cycle arrest and apoptosis. Some drug effluxes from the cell,

either by passive diffusion, active transport, or leakage from dying cells. If the effluxed drug is cell permeable, it may enter neighbouring cells, giving rise to the 'bystander cell killing effect.' The drug may also be metabolised in the cell. These metabolites may have differing cytotoxic abilities from the parent drug and propensity to be effluxed.[30]
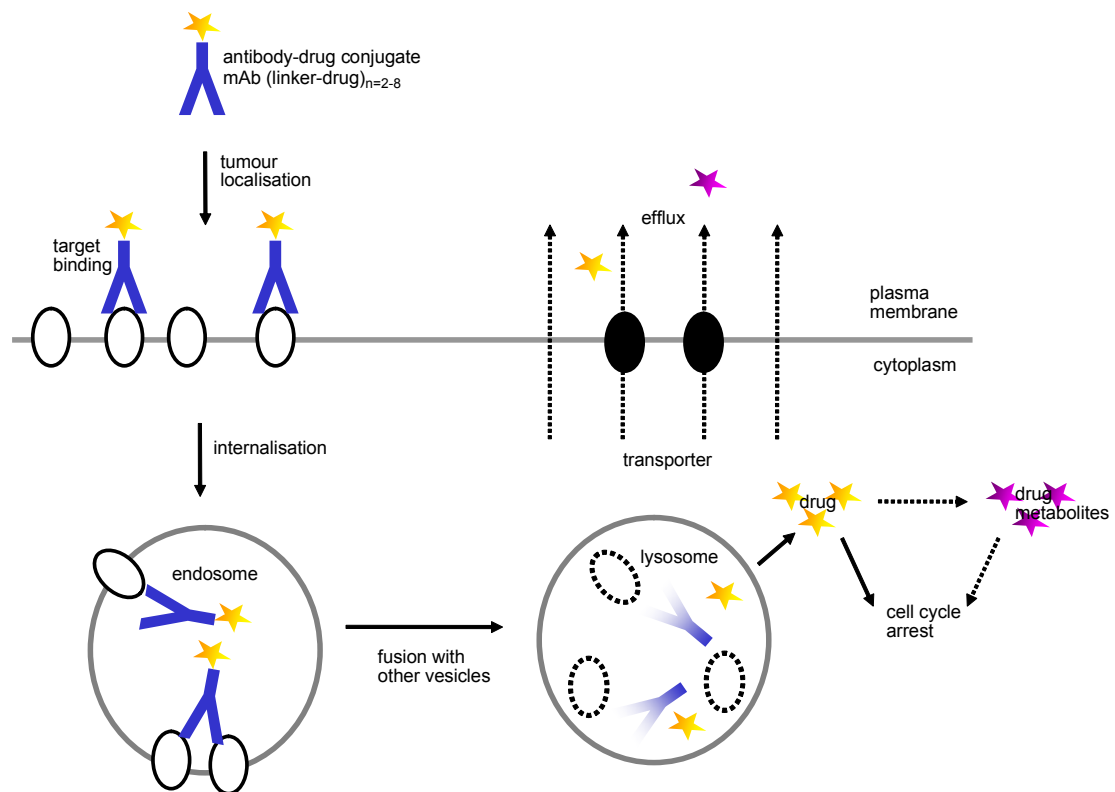
Figure 4 Cytotoxic activity of antibody-drug conjugate at tumour site.[30]

In 2000, gemtuzumab ozogamicin (Mylotarg, by Wyeth) emerged as the first licensed chemotherapeutic drug to directly target cancer cells using monoclonal antibodies. It is indicated for CD33-positive relapsed acute myeloid leukaemia (AML). CD33 is found on AML cell surfaces in about 80% of patients.[31] AML accounts for 80% of all acute leukaemia cases. Worldwide there are approximately 260,000 new AML cases each year, with 7,000 to 21,000 cases in Europe. However, Mylotarg was granted orphan

medicinal product status (a controlled licence for low prevalence illnesses for which medicines are not economically profitable) for AML in Europe but ultimately refused marketing authorisation in 2008 on the grounds that the open-label non-comparative studies carried out did not demonstrate clear benefits in AML. Many critical factors could not be determined such as how long remission would last or how Mylotarg affected survival rate. It was felt that the benefits in AML could not be established and did not outweigh the risks of the drug; severe and long-lasting bone marrow suppression, hepatotoxicity, veno-occlusive disease, and infusion related side effects such as chills, fever and low blood pressure.[31] Despite problems of therapeutic efficacy specific to this particular product, Mylotarg provides a good illustration of the scientific basis and scope of the ADC concept applicable to the development of any antibody-drug combination.

The recombinant humanised IgG4-κ antibody's (gemtuzumab, hP67.6) surface lysines are covalently bound to the cytotoxic enediyne antibiotic N-acetyl calicheamicin dimethyl hydrazide by means of the bifunctional acid labile hydrazone linker,(4-(4′-acetylphenoxy)) butanoic acid, AcBut and a sterically hindered disulfide bond (Figure 5), the latter present in all calicheamicin conjugates. The number of conjugated calicheamicin moieties can vary. Unconjugated antibody is also present in variable amounts. Mylotarg has a mean of 2-3 calicheamicin moles per mole of antibody, with 50% of ADCs with 4-6 moles of drug per mole of antibody, and 50% unconjugated antibody.[30] [32]

The antibody binds to the CD33 antigen on the target myeloid cell, and the resulting complex is internalised. It is believed that hydrazone cleavage is followed by reduction of the disulfide to release the cytotoxic agent.[33] Hydrazone cleavage occurs inside the acidic environment of the myeloid cell

lysosomes, releasing free calicheamicin which is activated by reductive cleavage of the disulfide bond. Methyl groups stabilise the disulfide to prevent premature calicheamicin release, for example by circulating reduced thiols such as glutathione.[34] The enediyne moiety is then able to bind to the DNA minor groove, rearranging to form diradicals which abstract hydrogen from the sugar-phosphate backbone inducing DNA double-strand breaks, ultimately leading to apoptosis.[35] However, any tumours overexpressing p-glycoprotein are resistant to gemtuzumab, which constitutes a major limitation.[33] [36]

Other ADCs currently in clinical development are SGN-35, made up of auristatin with a mean loading of around 4 drug moles per mole of antibody which has potential use in Hodgkins lymphoma.[30] It is designed to target CD30, utilising a valine-citrulline linker cleaved by proteases. The next generation Herceptin may be trastuzumab-DM1, comprising of maytansinoid linked to trastuzumab via a thioether which would target HER-2 receptors in breast cancer like its parent antibody.[37]
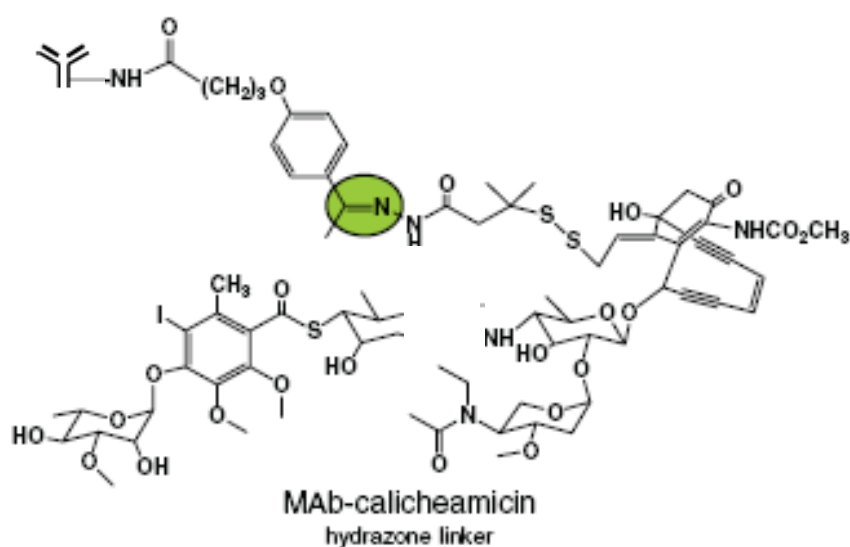

Figure 5 Gemtuzumab ozogamicin (Mylotarg) consists of calicheamicin linked to the anti CD33 antibody gentuzumab.[32] Acid-labile hydrazone linker shown in green.

1.5.2.3 Antibody-drug conjugate design

To design antibody-drug conjugates, choice of drug, antibody and linker requires careful selection.   Target selection is the critical first step in developing ADCs, which guides the method and design of the antibody. The antibody should selectively target tumour tissue, i.e. with minimum cross-reactivity with other tissue. Additional anti-tumour mechanisms can even be built in e.g. interfering with the target's biological function. The target antigen needs to be expressed ideally on most patients' tumour cells throughout the course of the disease, and absent from normal cells. Targeting is usually reliant on the highest antigen expression being on tumour cells relative to normal ones, and on antigens remaining at the tumour site and not entering the circulation. This can lead to clearance to the reticuloendothelial system (RES) and altered biodistribution to the liver, spleen and bone marrow. Thus, antigen shedding could either be tolerated or lead to toxicity in normal tissue. One solution may be to administer unconjugated antibody to bind to circulating shed antigen prior to treatment with conjugated drugs.[38]

1.5.3.2.1 Choice of antibody

Firstly, the antibody's antigen binding ability and specifity must not be compromised by the remainder of the conjugate. The antibody structure can be engineered, altering protein sequences to enhance binding.   Also, nonimmunogenic humanised antibodies are preferred to murine ones. Lower molecular weight mAb fragments display rapid clearance properties and improved tumour penetration.[39] This is due to the diffusion of larger sized antibodies being affected by high interstitial pressure, and high affinity

can result in "binding site barrier effects" where antibodies bind strongly to the first antigen encountered resulting in poor tumour penetration especially of solid tumours.[33] [39]

1.5.2.3.1 Highly cytotoxic drugs make good conjugates

The use of cytotoxic drugs with inherent ultrapotency is necessary (subnanomolar IC_{50} of free drug) as localisation of antibodies in human tumours is actually very inefficient; typically 0.0003 - 0.8 % of injected dose per gram of tumour.[30] Good candidates are calicheamicins, maytansinoids, auristatins and CC-1065 analogues. Drugs require amenable functional groups for linkage. A level of aqueous solubility and stability is also desirable. There is limited data on the effect of where the drug is conjugated to the antibody. Nonetheless, it is important that the drug is bound at a location remote from the antibody-antigen binding site, to prevent steric hindrance, and ultimately antigen binding affinity. Usually this is the Fc, constant region of the antibody, which does not participate in antigen binding (Figure 3). Binding interference is also avoided by carefully controlling the number of drug molecules attached to each antibody. Limiting drug loading i.e. drug-to-antibody ratio, also prevents ADC aggregation, solubility and pharmacokinetic problems. The drug also needs to be internalised in sufficient concentrations. This can be difficult to achieve if the level of antigens on the cell surface is low.

1.5.3.2.2 Linkers

The linker needs to maintain the stability of the conjugate during systemic circulation, but be sufficiently labile to release the active drug once internalised, at the target. The choice of linker is usually governed by any

amenable extracellular conditions which can be taken advantage of. Conditions that can be exploited are differences in pH, high cytosol glutathione concentrations, and lysosomal proteolytic enzymes, resulting in chemically or enzymatically labile, or even non-cleavable linkers.

Early conjugation was based on acid-labile hydrazones which are stable in the neutral pH of the bloodstream (pH 7.3 - 7.5), hydrolysing on internalisation into the acidic endosomes (pH 5.5-6.2) and lysosomes (pH 4.5-5.0). For example, (6-maleimidocaproyl)hydrazone has linked doxorubicin to the cysteine residues of the antibody. This conjugate ultimately failed in development due to low doxorubicin inherent potency, relatively short half life and insufficient clinical efficacy.[40] Mylotarg, as mentioned earlier, also employs a hydrazone linkage between calicheamicin and the antibody gemtuzumab (Figure 5). Work with inotuzumab ozogamicin using the same linkage as Mylotarg is ongoing at Wyeth and displays relatively better ADC stability.[40]

Disulfides linkers can be stable at physiological pH, with selective cleavage at the target tumour site where intracellular reduced glutathione levels are relatively higher (>1000 times) compared with blood plasma.[33] Hypoxic states of tumour cells due to poor blood flow, results in high activity of reductive enzymes and high glutathione concentrations. Work with mAb-ricin conjugates has used disulfide linkers. Developing hindered disulfides may confer some stability in the systemic circulation ensuring that the conjugate reaches the target whilst preserving susceptibility to intracellular reduction.[33] Self-immolative disulfide linkers have been designed where a glutathione-triggered cascade results in the release of the activated cytotoxic drug via thiolactone formation and ester bond cleavage (Scheme 2).[41]
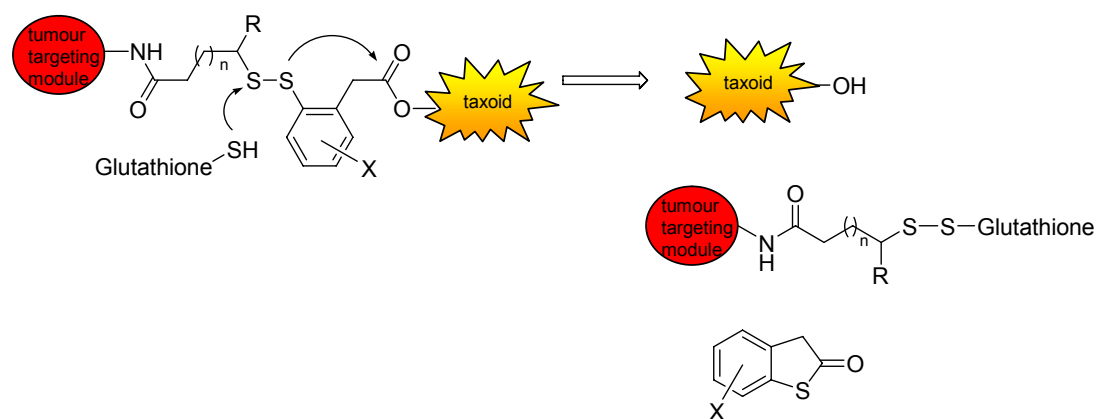
Scheme 2 Self-immolative disulfide linkers.[41]

Peptide-based linkers can confer better systemic stability as hydrolysis is reliant on tumour specific intracellular enzymes such as lysosomal proteases. Alternatively, specifically designed hydrolytic enzymes could be administered.[33] For example, Beeson and co-workers have designed radioimmunoconjugates consisting of ^{131}I-labelled cephalosporin linker conjugated to tositumomab, (against CD20 antigen) which is cleaved by β-lactamases.[42] Typically, the linker attaches via lysine residues on the antibody. Thus, the linker is responsible for attaching the drug to the antibody, controlling drug-retention half life in the blood circulation and drug release on reaching the target. Half-lives are in days rather than hours as is the case with hydrazones and disulfides.[33] Negatively charged residues on the ADC could impair cell membrane permeability, minimising systemic side effects.

Another example is Seattle Genetics' work on monomethyl auristatin E and F (auristatin derivatives binding to tubulin and inhibiting microtubule assembly) which are conjugated to the cysteine residues of the antibody via a valine citrulline linker (Figure 6). After enzymatic cleavage, 1,6-elimination of the electron-donating spacer 4-aminobenzyl group occurs to release free monomethyl auristatin E. This spacer is required to distance the steric bulk

of the drug which would otherwise prevent binding of enzymes to the amide bond.[43]
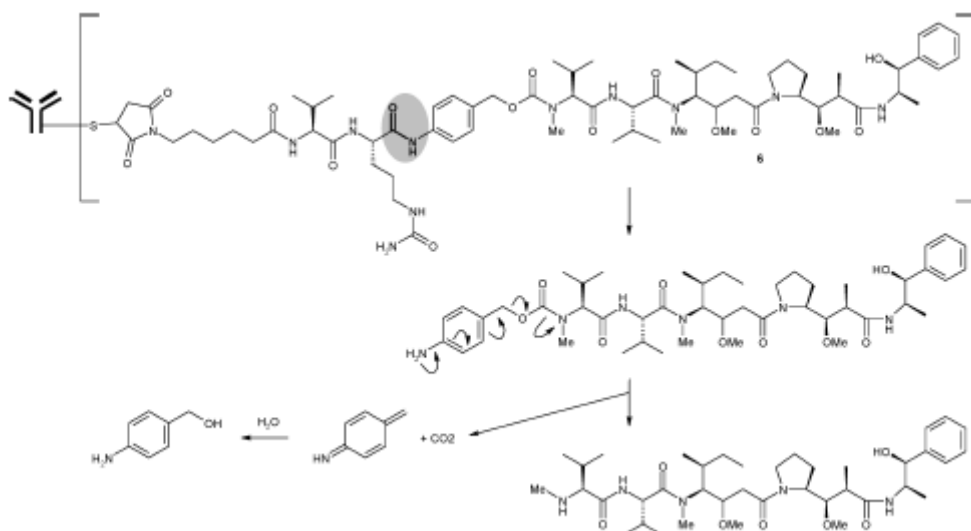

Figure 6 Proposed mechanism for release of monomethyl auristatin E from ADC 6 by cleavage of the valine-citrulline linker amide (shown in grey).[40]

A possible alternative method of targeting is pre-targeted therapy.[33] Here, localisation and delivery of the active drug by antibody-directed means is separated into two steps physically or by time. Antibody directed enzyme prodrug therapy (ADEPT) is a form of this type of pre-targeting. The mAb-enzyme conjugate localises at the tumour target followed by the administration of the cytotoxic prodrug. Selectivity is ensured as the prodrug only converts to the cytotoxic drug where the enzyme is present i.e. only at the tumour target site.[33]

To conclude, a new generation of chemotherapeutics are being developed which have improved targeting features, also minimising side effects. Targeting can be achieved by exploitation of tumour behaviour or environment, targeting of cellular processes, or by conjugation of the drug to targeting ligands. This design methodology is in contrast to the earlier use of

non specific cell cycle arrest and induction of cell death. Antibodies show great potential for improving targeting by conjugation to cytotoxic drugs.

1.6   Summary

At this juncture, the duocarmycins require introduction as excellent potential candidates for analogue generation due to their inherent cytotoxic activity in picomolar concentrations. Their structural diversity presents good opportunity for conjugation to moieties to improve targeting. An overview of these natural products, key issues for syntheses with reference to existing synthetic route methodology is dealt with in the following section. Our aim was to quickly reach an efficient and simple synthesis of the duocarmycin SA subunit, and use this to generate targeted duocarmycin analogues.

2. Introduction part 2: The Duocarmycins

The duocarmycins were initially isolated in the 1980s from *Streptomyces sp.* as potent antitumour antibiotics in Japan. This family of natural products includes (+) duocarmycin A and (+) duocarmycin SA (stable A). The duocarmycins are structurally similar to CC-1065, and to the more recently discovered yatakemycin, in that they all share the core dienone cyclopropane ring with subtle substituent differences (Figure 7). In this section the biological role and comparative merits of the active functional groups, and syntheses of key structures are discussed.
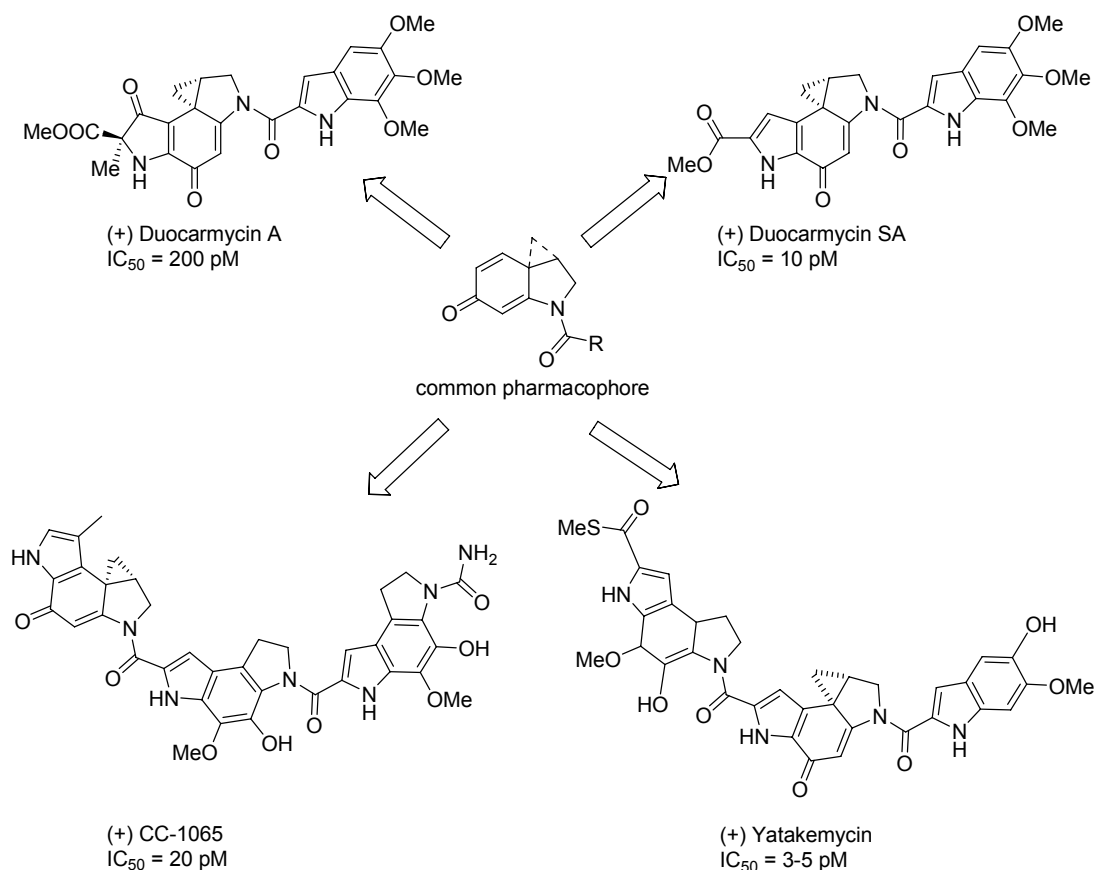

Figure 7 The duocarmycins, CC-1065, yatakemycin all share a common pharmacophore. IC50* values shown in L1210 murine cancer cell lines [44]

---

*IC_{50}=drug concentration at which 50% of cell growth is inhibited

2.1 Shape dependent binding

The structure of the duocarmycins results in a uniquely target-specific binding mechanism. Boger and co-workers have suggested that a conformational change or "twist" in planar structure occurs at the N-2 position upon binding, in order to fit into the DNA's helical structure (Figure 8).[47] This disrupts the vinylogous amide, activating nucleophilic attack on the cyclopropyl ring. Thus the duocarmycins only become activated at their biological target by shape-dependent catalysis.[45] This has profound implications for targeting as analogue design only requires encouragement to reach the target site.[46]

2.2 Biological activity

The duocarmycins' biological activity has been attributed to sequence-selective DNA minor groove alkylation by the dienone cyclopropane subunit. Reversible, stereoelectronic controlled N-3 addition of adenine to the least substituted cyclopropane carbon occurs at A-T rich base pairs in the minor groove of DNA (Figure 8), which ultimately damages DNA, inducing apoptosis.[47] [48] This activity makes the duocarmycins ideal anticancer candidates. The ease of reversibility depends on the relative reactivity and stability of the structure. DNA binding becomes less reversible and the adduct more stabilised by additional non-covalent interactions of the extended natural product structure.[48] Studies have shown that increases in pH, temperature and time favour retroalkylation i.e reversal of DNA alkylation, which is attributed to base-catalysed phenol deprotonation.[49] [50] The duocarmycins are highly potent, with cytotoxicity against tumourigenic cells in the range of 10 pM for duocarmycin SA. This potency is 10^3 times

higher than classical intercalating compounds such as adriamycin.[51] The duocarmycins have the additional advantage that they do not present the fatal delayed hepatotoxicity characteristic of CC-1065.[47][52]
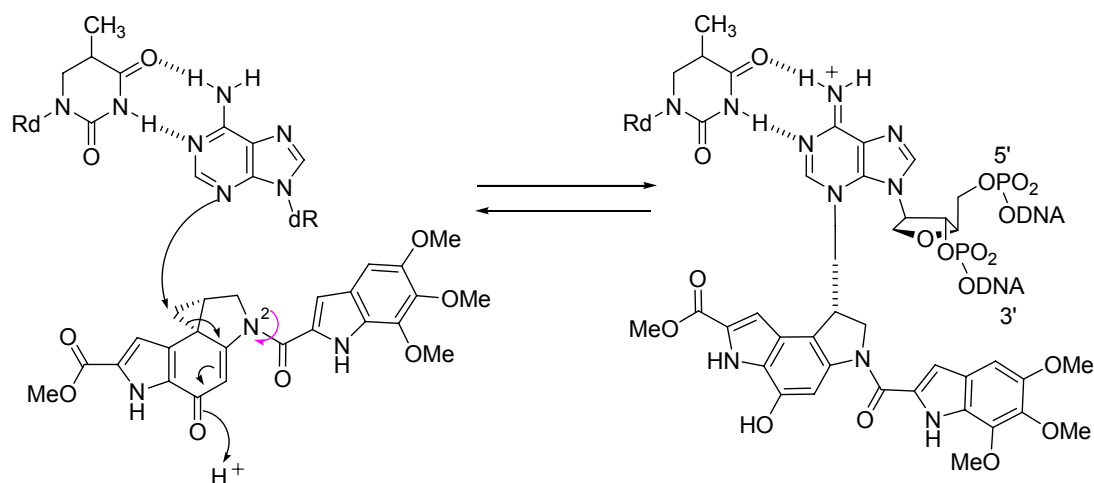

Figure 8 Mechanism of DNA binding by duocarmycin SA. Binding induced twist (in pink) between subunits at N-2 disrupts the amide as the N-2 electron lone pair/vinyl π orbital no longer overlap, destabilising the spirocyclopropane system. The cyclopropane ring undergoes nucleophilic attack by the adenine N-3 electron lone pair. This leads to drug binding at A-T base pairs in the DNA minor groove, causing DNA damage and induction of apoptosis.[46]

2.3 Importance of the cyclopropyl ring as the minimum potent pharmacophore

Figure 9 shows substituent and electronic effects on the biological activity of duocarmycin SA. Non-covalent forces also significantly contribute to the stabilisation of DNA alkylation. 5-Methoxyindole duocarmycin SA analogues have been found to retain all the DNA binding ability and biological activity of the full natural product, indicating that rigid length is important for activity as it provides stability, which reduces the rate of retroalkylation. The 6'- and 7'- methoxy, and N-1' positions point out of the minor groove with minimal DNA interaction and contribution to the extended structure that induces conformational change. These positions can

be points of substitution, however, any modifications must also take careful consideration of electronic effects.[46]
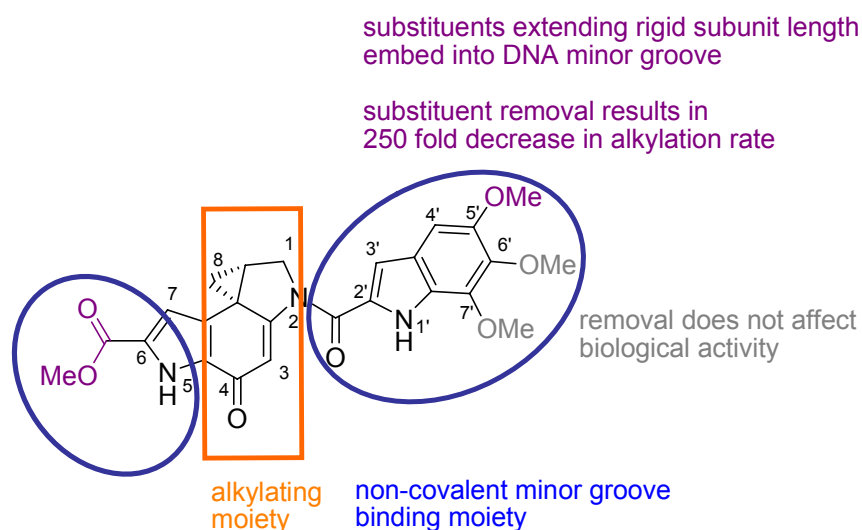

Figure 9 Duocarmycin SA is composed of the left hand alkylation (DSA) and right hand non-covalent binding subunits. The effects of substituents on biological activity are shown.[46]

The DSA structure (left hand subunit) is core to duocarmycin SA's DNA alkylating capabilities (Figure 10).[46] [53] Subsequent findings have indicated that the minimum pharmacophore CI (cyclopropylindolone) alkylates the same sites as duocarmycin-based agents regardless of the presence of the aromatic carbonyl, but with reduced potency and selectivity at the sites available for DNA alkylation.[47] [54] [55] This demonstrates the importance of the carbonyl group for potency. Figure 10 shows the DSA and CI structures.

CI
(cyclopropylindolone)

DSA
(duocarmycin SA left hand
alkylation subunit)
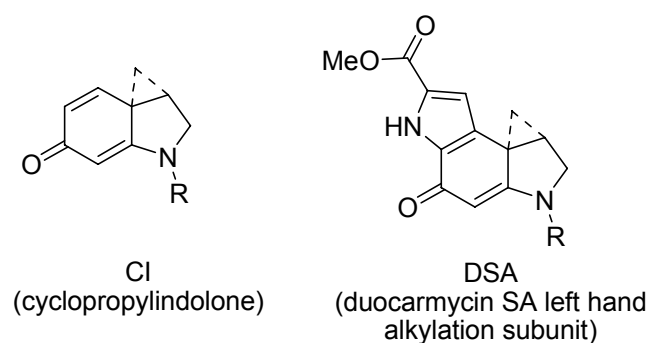
	CI	DSA
Relative IC$_{50}$ (pM, in L1210)	1 - 2	10
Relative DNA alkylation intensity	0.5 - 2.0	10

Figure 10 DSA and the minimum pharmacophore CI[55]

2.4 Relationship between solvolytic reactivity and cytotoxic potency

From structure-activity relationships, it has been found that cytotoxicity is directly related to chemical stability in aqueous acidic solutions. The acid-catalysed activation of the DNA alkylation reaction initially led to the expectation of a direct relationship between reactivity and cytotoxic activity, and that increasing electrophilic reactivity would enhance biological potency.[48] [56] In fact, studies demonstrated a direct relationship between solvolytic stability and biological potency.[48] An optimal balance is required between sufficient stability to reach the target and enough reactivity for DNA alkylation on getting there.[57] [58] DSA, the left hand alkylation subunit of duocarmycin SA, can be seen at the peak of this parabolic relationship, displaying the most ideal properties (Figure 11). CI (N-Boc-CI, cyclopropylindolone) is highly reactive or unstable, with lower associated cytotoxicity. In contrast, CPI (N-Boc-CPI, cyclopropylpyrrolo[e]indolone) is relatively more stable, with higher cytotoxicity. Subsequent work by Boger has found that replacing the pyrrole in the yatakemycin left hand subunit

with a thiophene has led to near optimal stability and potency. Thus, N-Boc-MeCTI (methylcyclopropathio[e]noindolone) analogues have been found to have 5-6 times greater stability than N-Boc-MeCPI (CC-1065 precursor subunit), and only 4-5 fold less potency than N-Boc-duocarmycin SA.[58] Studies on CBI (cyclopropabenzo[e]indolone) derivatives have shown a direct relationship between the electron-withdrawing properties of N-2 substituents and solvolytic reactivity; the most electron-withdrawing substituents provide the greatest chemical stability.[48] The C-6 methyl ester of CPI, and hence of duocarmycin SA lies deep in the minor groove. This increases rigid length and thus the degree of binding-induced twist.[59]
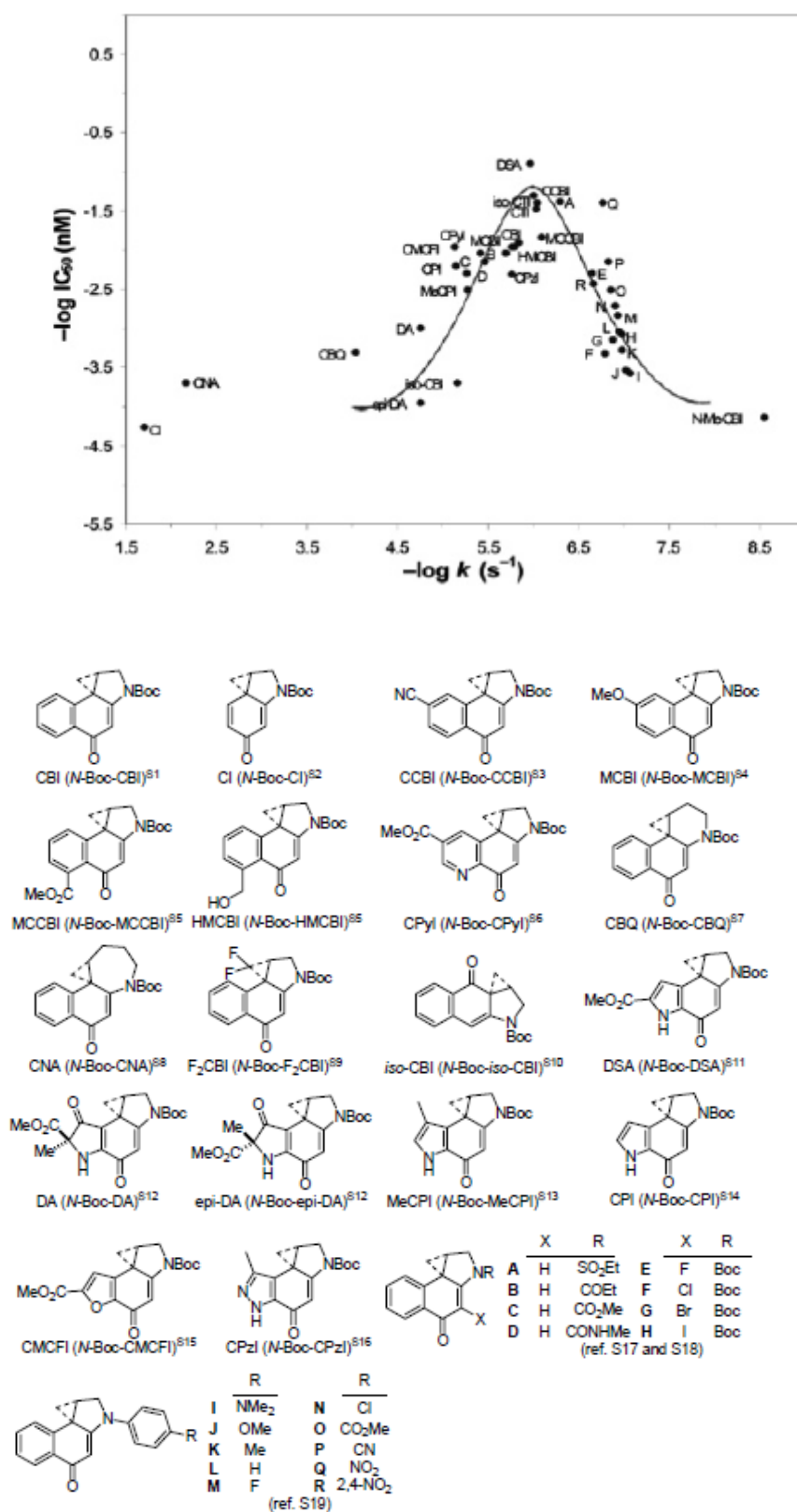
Figure 11 Relationship between reactivity (solvolysis) and cytotoxic potency for duocarmycin SA and related structures (L1210 cell lines), with key t structures.[58]

2.5 Synthesis overview

Our research focussed on the duocarmycin SA alkylation subunit due to the fact that it shows the highest cytotoxic activity of all the related structures. The key subunit, DSA (Figure 10) confers most of the activity of the duocarmycins and related natural products, as discussed above. Stability has been shown to arise from vinylogous resonance conjugation of the electron lone pairs on the nitrogen atom of the carboxamide unit. The carboxamide's electron withdrawing capabilities and extended aromatic subunit also contribute. By retaining the cyclopropane core but varying substituents, highly potent and specific analogues can thus be synthesised.

A key issue for synthesis of the core DSA subunit is the assembly of highly functionalised nitrogen-containing heterocycles. Thus, any synthesis of DSA will aim to form the featured indoles in a concise, efficient and elegant manner. A short overview of published duocarmycin SA syntheses follows, highlighting the methodology for indole formation.

The first reported total synthesis of duocarmycin SA was by Boger and co-workers.[47] [60] This synthesis starts with the stepwise construction of the indoles of the DSA subunit. Scheme 3 shows the key steps to indole formation. A *p*-quinone diimine **a** is malonated, reduced to diol **b**, then protected (as an acetonide, **c**) to ultimately form the C ring. Nucleophilic addition of the pyrrolidine enamine of pyruvaldehyde dimethyl acetal (to form **d**) is followed by acid hydrolysis to provide the C-6 adduct **e**, setting up the structure for A ring indole cyclisation. Mitsunobu activation (using Ph$_3$P-DEAD) forms the C ring **f-g**. This DSA subunit scaffold ultimately

forms duocarmycin SA. The synthesis was completed in 14 steps with an overall yield of 6%.[60]
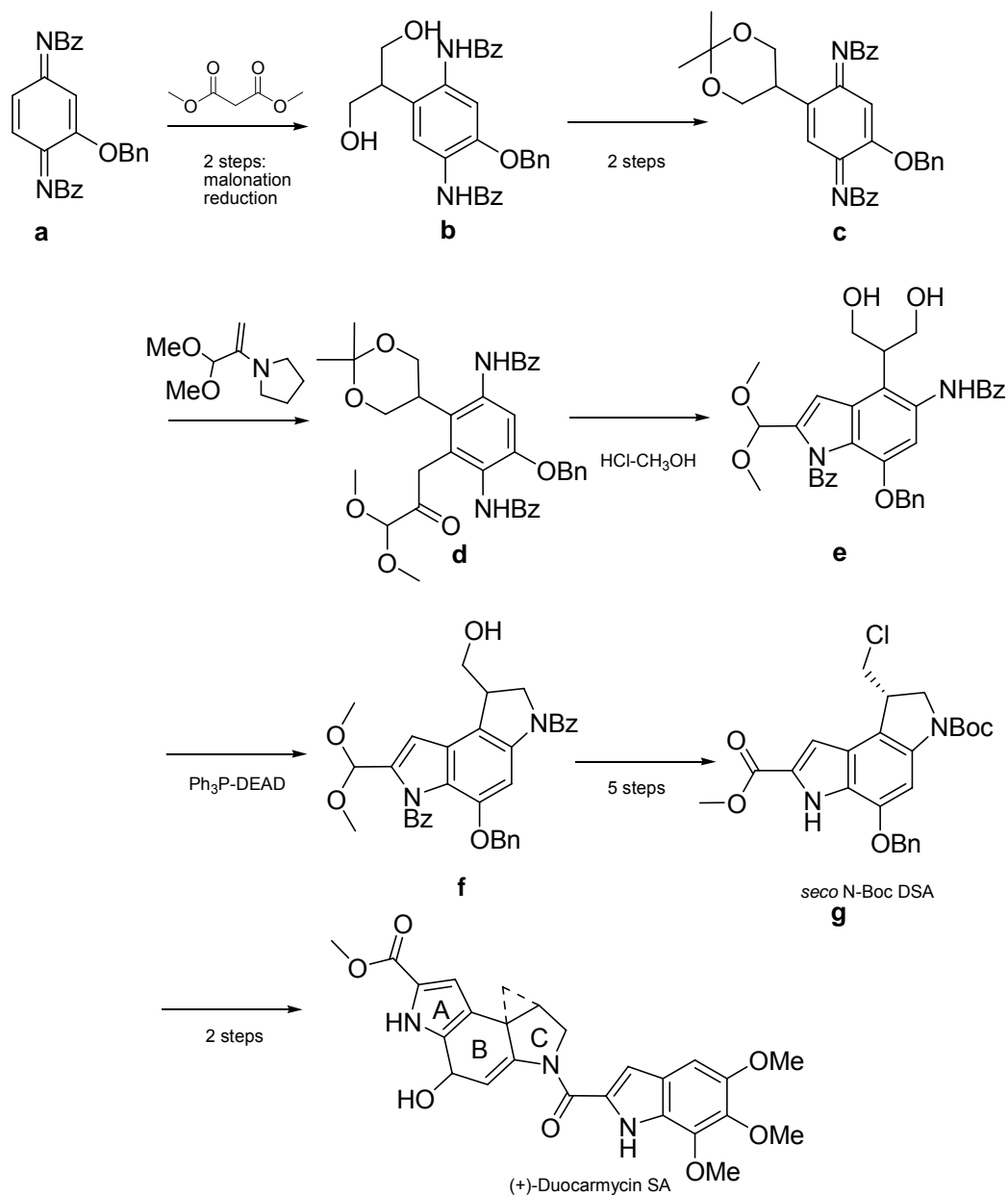
Scheme 3 Route to total synthesis of duocarmycin SA by Boger and co-workers, from p-quinone diimine via seco N-Boc DSA[47]

Organometallic chemistry has been often utilised for coupling and cyclisation to form indoles. One example is of Hiroya and co-workers' route to duocarmycin SA via a novel Pd catalysed method for synthesising the

indole-2-carboxylate (A ring indole)[61] as shown in Scheme 4. Negishi's coupling conditions (ZnBr$_2$, ^{i}Pr$_2$NEt, methyl propiolate, Pd(PPh$_3$)$_4$) were applied to couple an alkyne at the iodo substituent of **a**, followed by cyclisation of the mono-mesyl protected amine to form the indole **b**. This synthesis achieved duocarmycin SA in an overall yield of 3% over 17 steps.
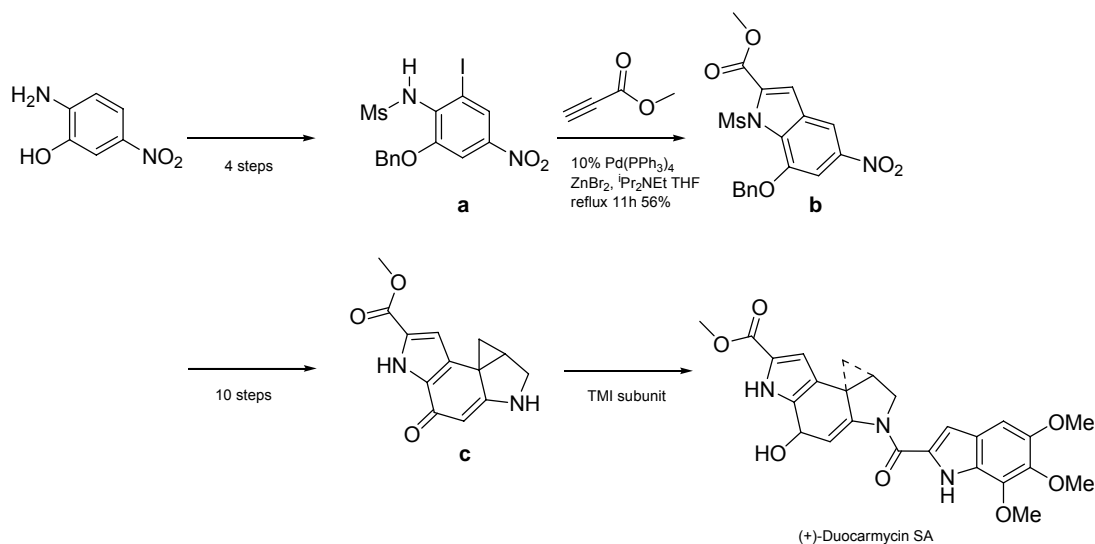

Scheme 4 Hiroya's synthesis of DSA using starting material 2-amino-5-nitrophenol, utilising Negishi coupling for indole formation, to ultimately form duocarmycin SA[61]

Tietze utilised the well known Fischer indole methodology in his concise route to *seco* duocarmycin SA as shown in Scheme 5. Diazotisation of the starting material **a** formed a diazonium salt, which was immediately reduced to provide the hydrazine. Treatment with methyl pyruvate formed the hydrazone **c**. Heating this precursor structure in PPA and xylene resulted in cyclisation to form the A ring indole **c**.[62]
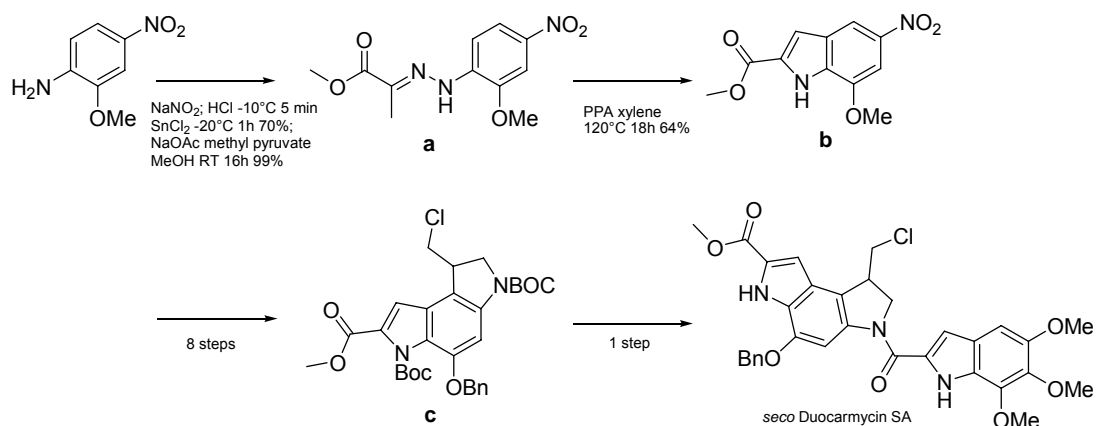
Scheme 5 Route to duocarmycin SA by Tietze using the starting material 2-methoxy-4-nitroaniline, utilising a hydrazone to achieve Fischer indole synthesis [62]

Natsume's approach was different in that the synthesis began with an intact A ring pyrrole, with the B (**b**) and C (**c**) rings sequentially built around it (Scheme 6).[47] Ester alkylation (L-malic acid ester derivative, *t*-BuOK, 86% to form **a**) provided the precursor structure for Lewis acid catalysed cyclisation to indole **b**. This synthesis was completed in 18 steps with an overall yield of 2%.
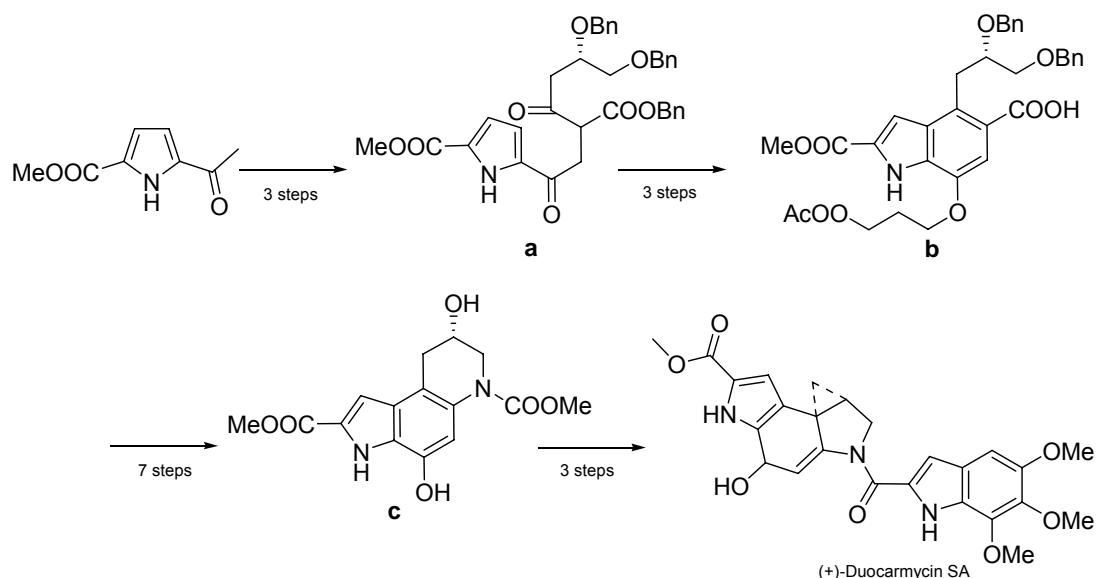
Scheme 6 Natsume's synthesis to duocarmycin SA initiating from a pyrrole structure[47]

2.5.1 Yatakemycin – the newest member

Yatakemycin has common structural features with both CC-1065 and the duocarmycins. It became the newest and most potent member when its structure was disclosed in 2003.[63] Despite its similarities, it also has distinct DNA binding subunits "sandwiching" or flanking the central alkylation subunit. In 2004, Boger's group reported a total synthesis,[64] but spectroscopic differences became evident between the synthesised proposed structure and the natural product. On the basis of this, an alternative structure bearing a left hand subunit thioester terminus (instead of a thioacetate), and right hand subunit 5-hydroxy, 6-methoxy substituents (instead of the reverse) were proposed.[65] Figure 12 shows the original disclosed and updated structures.
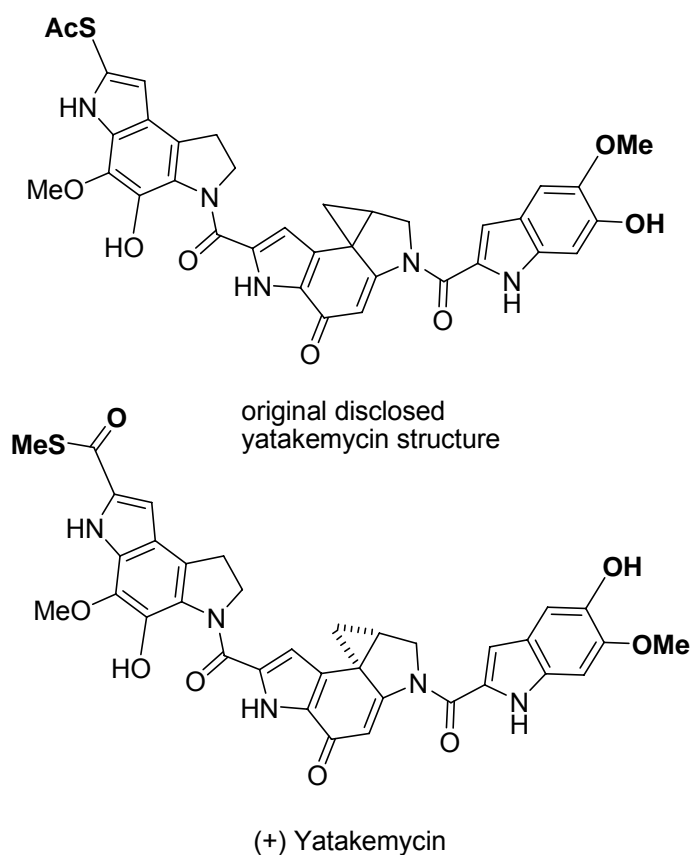

Figure 12 Original and reformulated yatakemycin structures[65] (differences highlighted in bold)

Boger's group then presented an asymmetric synthetic route to yatakemycin, which is applicable to duocarmycin SA.[65] This synthesis uses Hemetsberger methodology for indole synthesis; 3, 5 dinitrobenzyl alcohol is used as the starting material, from which an aldehyde **a** is condensed with methyl azidoacetate, followed by thermolysis to form **c**. This goes on to provide the central yatakemycin subunit, which is structurally analogous to DSA (**d**, Scheme 7).[65]
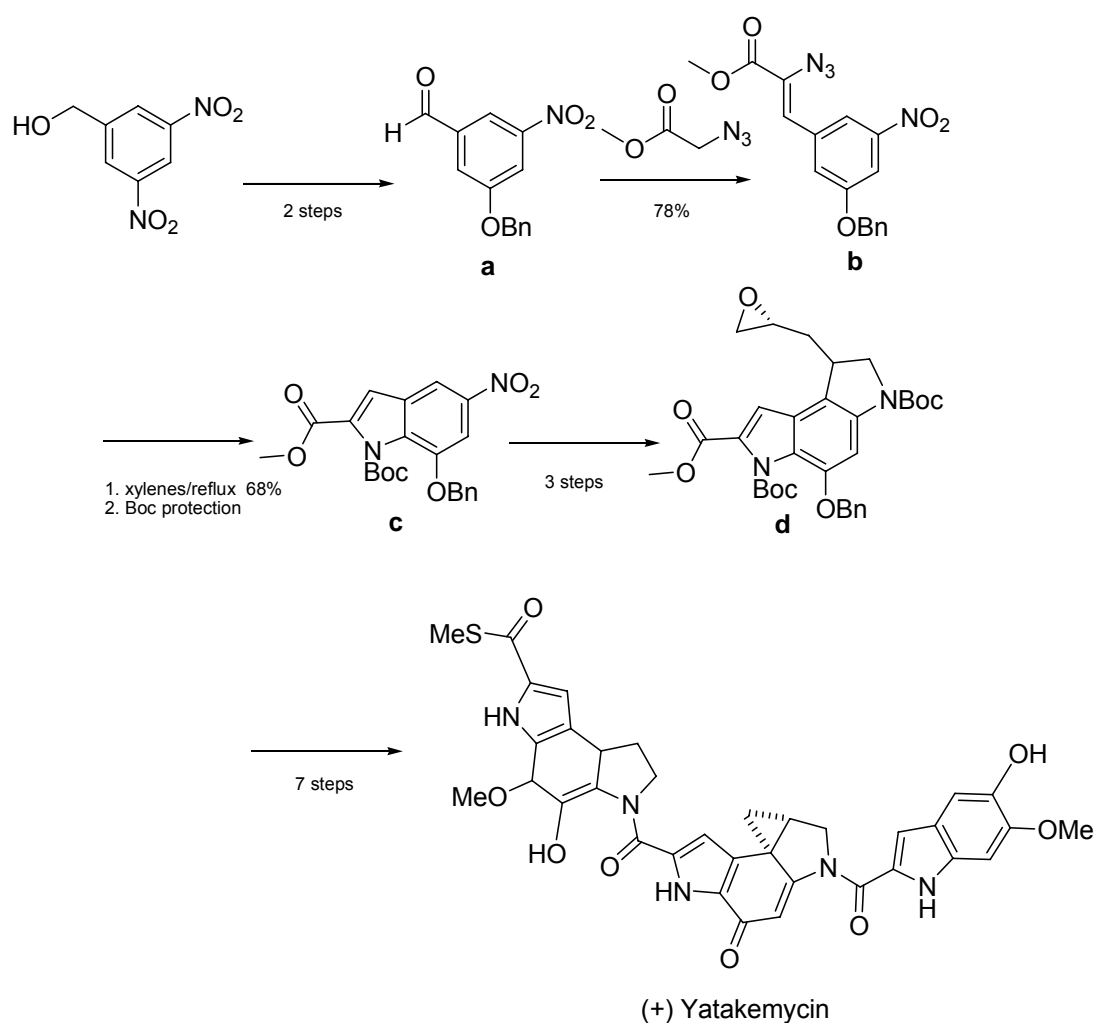
Scheme 7 Boger and co-workers' synthesis of the yatakemycin central subunit, starting from 3, 5 dinitrobenzyl alcohol utilising Hemetsberger methodology for indole formation analogous to the seco N-Boc-DSA intermediate[65]

Fukuyama's group published an alternative route to yatakemycin.[66] Here, synthesis of the middle segment was achieved by assembling the C ring **a** first (Scheme 8). Heck coupling of the aryl bromide with a dehydroalanine derivative forms the indole precursor structure (**b-c**). Cyclisation to form A ring indole **d** was achieved by intramolecular aryl amination using CuI; methodology based on organometallic catalytic couplings previously established in Fukuyama's published synthesis of duocarmycin SA.[67] The complete synthesis to (+)-yatakemycin was over 20 steps in 13% overall yield.[66]
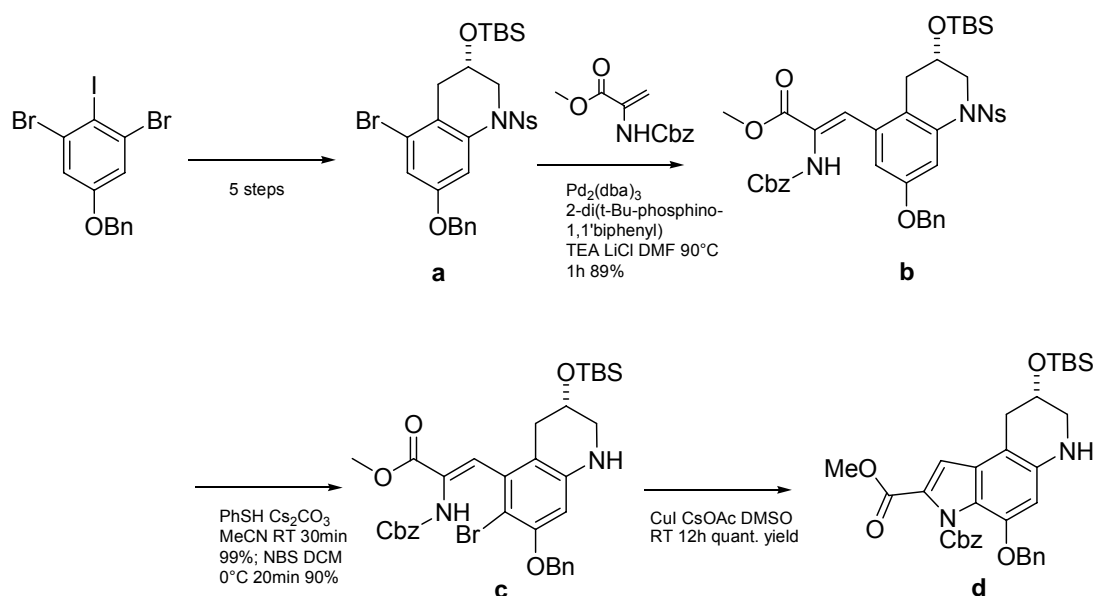
Scheme 8 Synthesis of the middle subunit of yatakemycin by Fukuyama and co-workers utilising Heck coupling/intramolecular aryl amination for indole formation[66]

2.5.2 Analogues

Efforts have been made to vary the structural design of the duocarmycins and related natural products to improve cytotoxicity and selectivity. There are numerous examples of structural analogues of duocarmycin SA in the literature, which have contributed to a better understanding of substituent

and subunit effects on DNA selectivity and alkylation.[44] [47] [48] A short overview of duocarmycin SA based analogues is presented below.

Boger and co-workers have synthesised bifunctional alkylating agents by coupling two DSA units together.[68] An existing N-Boc DSA was ring opened and coupled directly to another *seco* (ring open) compound. The resulting compound was spirocyclised to the duocarmycin dimer under basic conditions. With this approach, 4 diastereoisomers were synthesised (Table 1). The structure (+)(+)-duocarmycin SA_2 and (+)(-)-duocarmycin SA_2 displayed potencies of 3.0 pM and 5.0 pM respectively in L1210 cell lines. This shows a 2-3 fold increased cytotoxic activity compared to duocarmycin SA (10 pM). It is thought that these dimerised structures may demonstrate increased stability as well as improved cytotoxicity due to the presence of the right hand subunit also conferring DNA alkylation properties.[68]
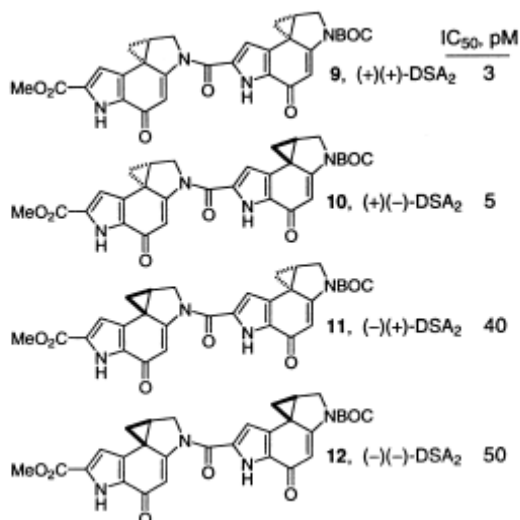

Table 1 In vitro cytotoxic activity in L1210 cell lines of bifunctional alkylating agents[68]

DSA based reversed analogues have been synthesised by Ellis and co-workers to study minor groove binding and affinity.[69] These analogues have the spirocyclised subunit as the right hand subunit in contrast to its usual left

hand position in the naturally occurring structures as shown in Scheme 9. Findings with these structures provide support for DNA alkylation being controlled by noncovalent binding, and not by alkylation subunit binding. In contrast, DNA alkylation is markedly reduced compared to the natural structures, due to disruption of the stabilising vinylogous amide preventing binding-induced catalysis.[69]
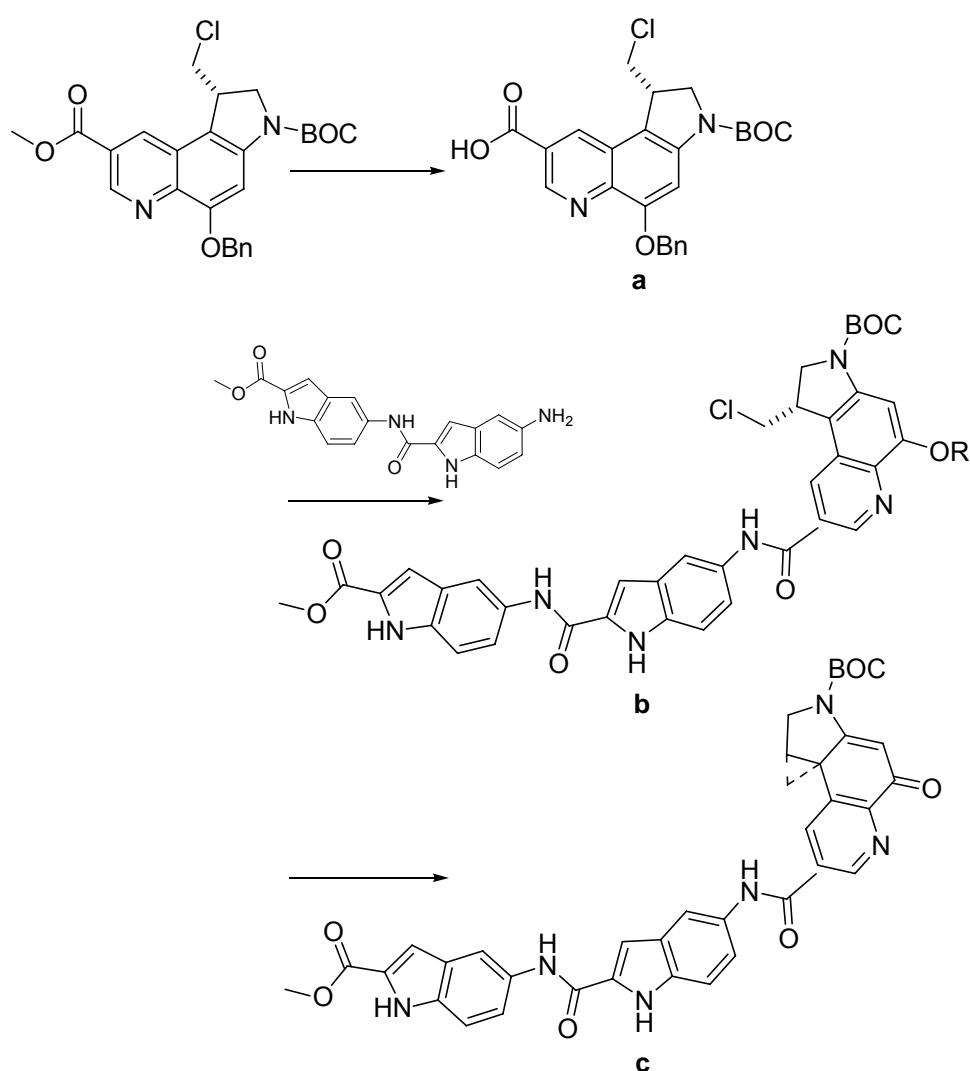
Scheme 9 Reversed duocarmycin SA analogues;[69] the spirocyclised subunit occupies a right hand position

Tercel and co-workers have synthesised *seco* duocarmycin SA analogues[70] with the aim of making a stable, low activity prodrug which would be

activated to its cytotoxic form at the target site. Synthesis commenced from 3-napthalic acid, with allylation (**a-b**) as the key precursor step prior to cyclisation (Scheme 10). Methyl azidoacetate was condensed with the aldehyde **c** for indole cyclisation using Hemetsberger methodology to form **d**. Replacement of the *seco* phenol by a nitro group showed the most reduced cytotoxicity.[70]
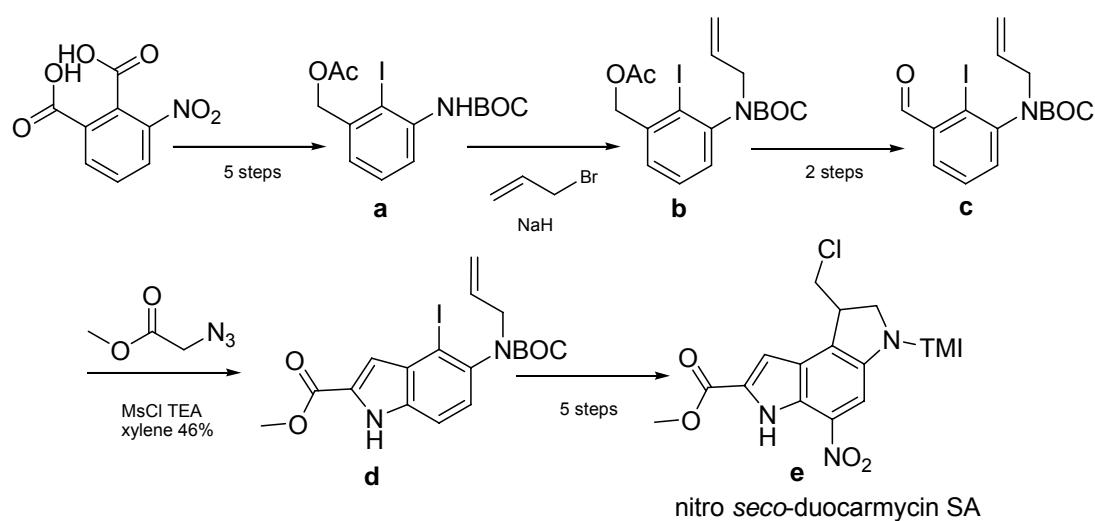
Scheme 10 Nitro-seco-duocarmycin SA synthesised by Tercel and co-workers[70] (TMI=trimethoxyindole)

Studies have shown that the (+)-(S) enantiomers of duocarmycin SA are more cytotoxic than their mirror images, and have different DNA binding orientations (3'-5' at AT sites).[71] Daniell and co-workers have synthesised achiral *seco* duocarmycin SA analogues showing that DNA binding and cytotoxicity can be retained if chirality is eliminated; achiral analogues had comparable cytotoxicities to their chiral counterparts.[71] The starting material 2-amino-4-chloro-5-nitrophenol was led through a series of steps, with the introduction of a diester (**b**) as the key precursor step prior to cyclisation. In their synthesis, indole formation (**c**) was achieved by intramolecular cyclisation of a dimethylcarboxylate using Pd(OAc)$_2$ as a catalyst to form a di-methoxy *seco* duocarmycin SA subunit (**d**, Scheme 11).[71]
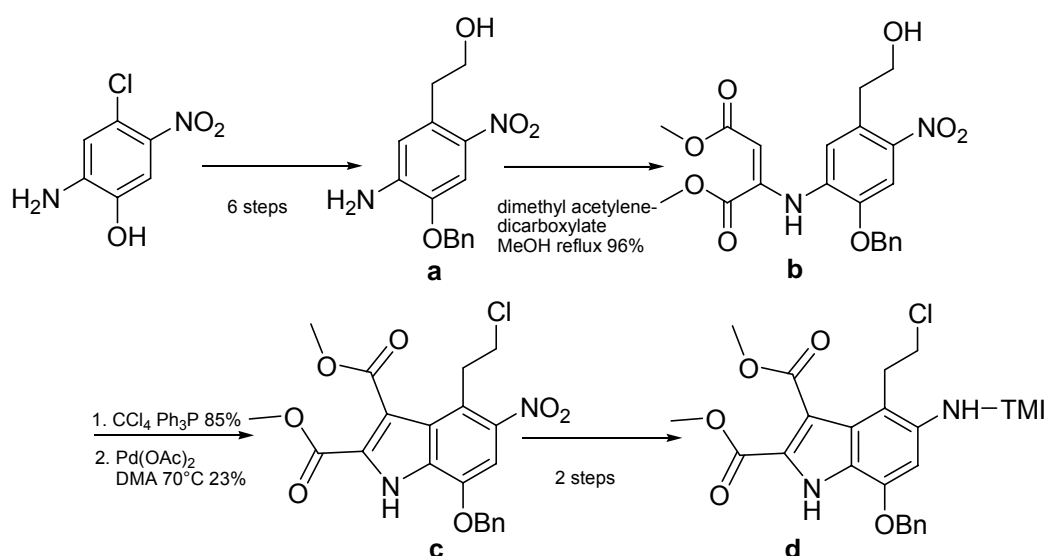
Scheme 11 Achiral duocarmycin SA analogues synthesised by Daniell and co-workers[71] (TMI=trimethoxyindole)

2.5.2.1 Analogue conjugates

The duocarmycins and related structures are inherently cytotoxic, however this activity needs to be harnessed in a manner that delivers the drug directly to the intended target. In addition to designing simple but highly cytotoxic analogues, another way to improve this activity is to conjugate the duocarmycin structure to another moiety with targeting characteristics, conferring even greater cytotoxicity but in a controlled fashion. This can minimise undesirable systemic effects by only allowing release of the drug due to target site-specific activation, or by acting as a "homing agent" improving selectivity by delivery to the target site. An overview of published duocarmycin SA based analogue conjugates is presented below.

Tietze has synthesised CBI prodrug analogues, which are activated at the target site by galactosidases and glucuronidases as shown in Figure 13. The *seco* drugs then rapidly cyclise *in situ* to the cyclopropyl moiety. Analogue

design incorporated tertiary amine R side chains enabling ammonium salt formation to improve water solubility.[44] β-glucuronidase cleavage has been utilised as high concentrations have been found in the extracellular space of solid necrotic tumours. *In vitro* tests on single cell proliferation in A549 human bronchial carcinoma cells showed that the glycosidic prodrugs had the same cytotoxicity in the presence of β-D-galactosidase as the *seco* drugs.[44] A 1-chloroethyl group was introduced into the prodrugs to increase steric hindrance and hence reduce direct DNA base alkylation. Although the *seco* drugs were found to have better stability and higher cytotoxicity, the methyl *seco* analogues exhibited a higher therapeutic index.[44]
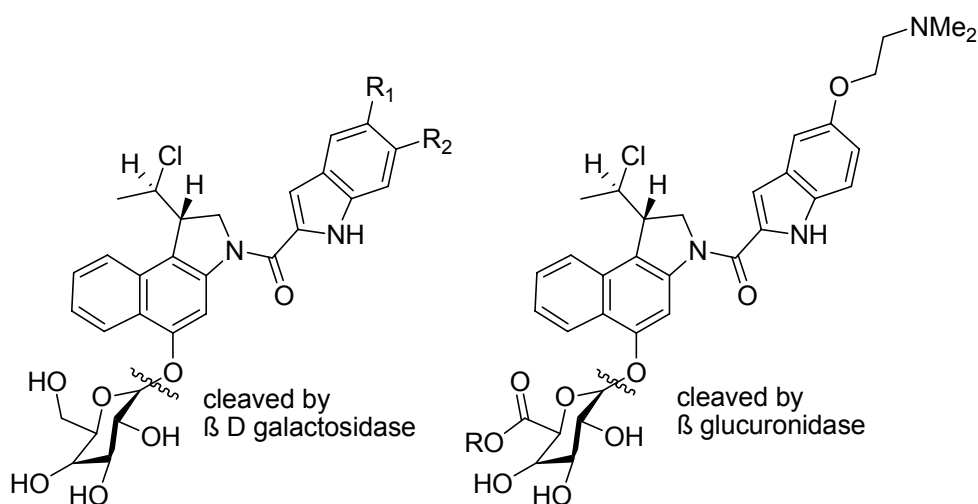
Figure 13 Tietze's glycosidic CBI analogues[44] [72] R=H or Me, R$_1$, R$_2$=tertiary amino side chain

Peptides have also been incorporated into prodrug design to improve targeting. Small peptides can have good tumour penetration, low immunogenicity, and target overexpressed receptors in a highly specific manner with great affinity.[44] Figure 14 shows Tietze and co-worker's pentagastrin conjugates of *seco* CBI based subunits.[73] However, cytotoxicity has not been found to be high enough to be sufficiently efficacious for *in vivo* application. It has been found that conjugation of the *seco* drug to the

pentagastrin N-terminus adversely affects pentagastrin activity. The authors suggest that longer gastrin peptide sequences may overcome this particular problem,[44] possibly due to the reduced proximity of the two components.
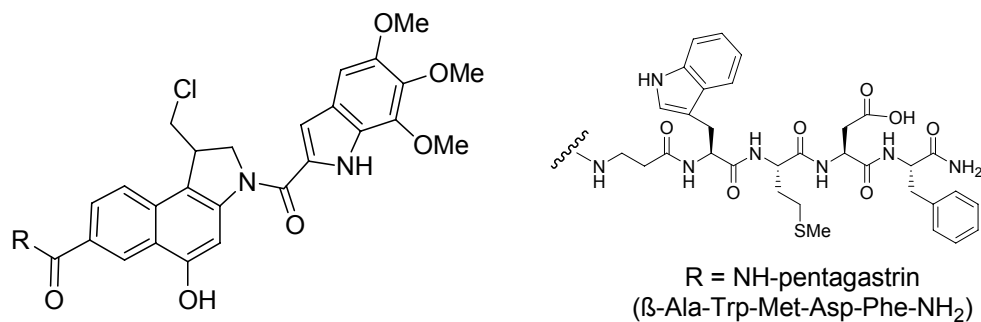

Figure 14 β-alanine modified pentagastrin CBI analogue conjugate[44]

2.5.2.2 Analogue conjugation to antibodies

As our research interests lay in the generation of duocarmycin SA analogues via conjugation to antibodies as a way to improve targeting and selectivity, an overview of published work in this area follows.

Anti-B4-DC1 is a bis indolyl-(*seco*)-CBI derivative and a potent CC-1065 analogue (IC$_{50}$ 0.02 nM) conjugated to an antibody (anti-B4) via a disulfide linker.[74] Conjugation was achieved by the attachment of sterically hindered sulfhydryl groups to thiol groups on the antibody. This compound was found to be 100-fold more potent than older conjugated drugs such as cyclophosphamide, doxorubicin at maximum tolerated doses, and more target-specific than the unconjugated form, with comparable potency to recent immunoconjugates based on maytansinoids. This analogue is highly selective for the target Namalwa (human lymphoma) cell line. However, poor aqueous buffer solubility made conjugation reactions difficult. This was improved by converting DC1 to a water-soluble phosphate prodrug DC4 (Figure 15).[74] [75]
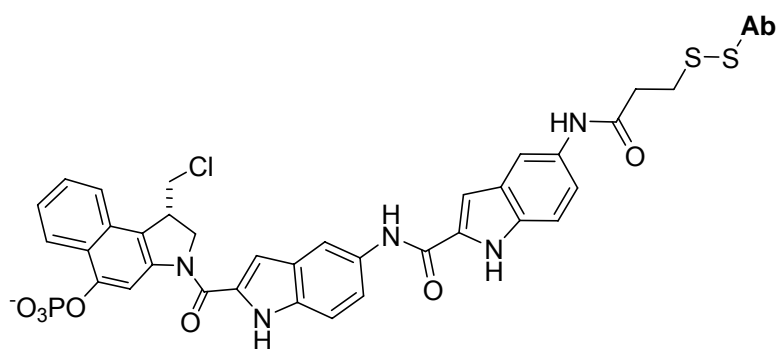
Figure 15 Anti-B4-DC4 analogue conjugate[75]

Seco-CBI analogues have previously been synthesised with dipeptide and hydrazone linkage to antibodies. The properties of the linker are important, providing stability to the conjugate whilst reaching the target, but also being sufficiently labile to free the cytotoxic drug at the target site. The hydrophobic planar structures and inherent instability of these analogues renders them liable to cause antibody-drug conjugate aggregation. Focus, such as work by Desbene and co-workers on doxorubicin ADEPT therapy introduced β-glucuronides as good linkers (Scheme 12),[76] [77] has now turned towards using glucuronide linkers as these confer improved solubility, are stable in plasma yet readily cleavable. It has been found that functionalising the CBI activating heteroatom (N or O) as a carbamate, amide or ether ensures the molecules are inactive whilst attached to the antibody. The antibody-drug conjugates (ADC) do not aggregate even when heavily loaded with hydrophobic drugs.[77]
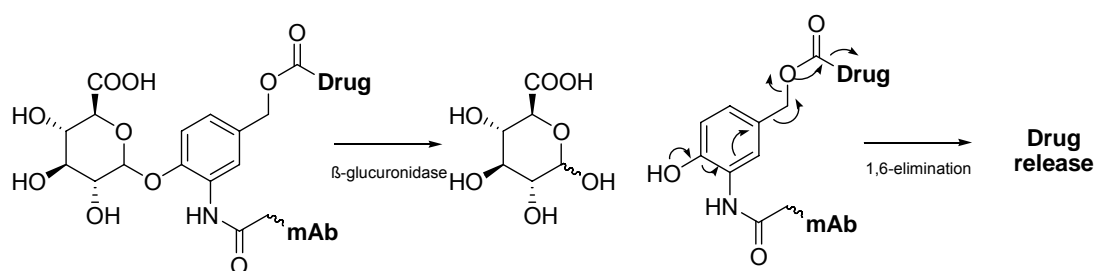
Scheme 12 Mechanism of drug release from antibody-drug conjugates with glucuronide linkers[77]

Another CBI-based ADC with peptide linkage has been designed by Mederex. This double prodrug strategy uses a *seco*-CBI derivative, a prodrug derivative of CC-1065, and provides improved therapeutic efficacy by conjugation to an antibody, improving targeting and minimising side effects. Two activating steps are required to release the cytotoxic to exert its activity (Figure 16) – amide bond cleavage to free the drug from the ADC, followed by cleavage of the carbamate initiating spirocyclisation.
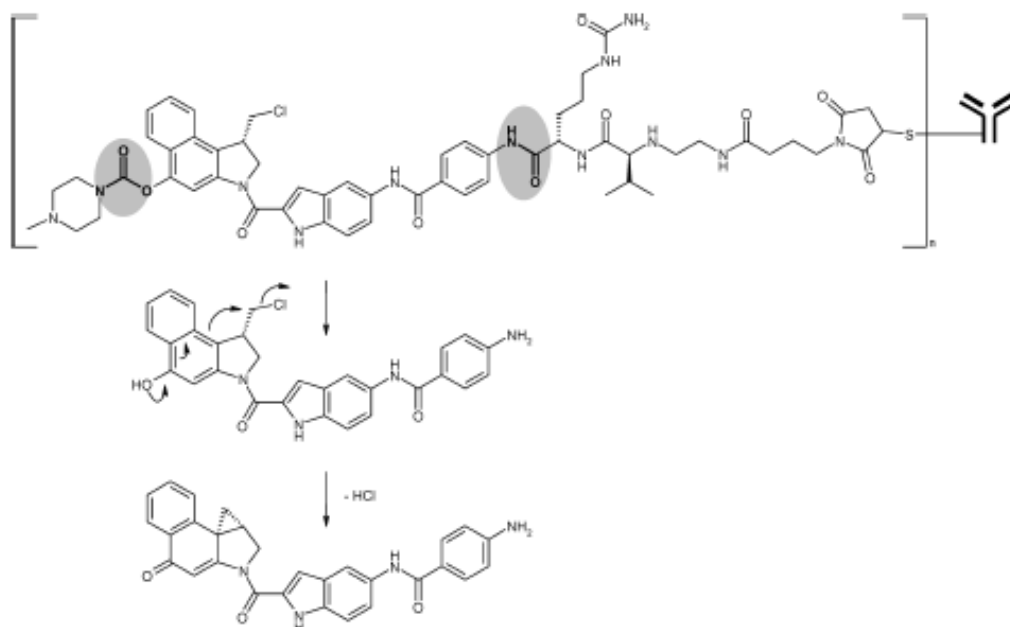

Figure 16 Proposed mechanism for release of CBI derivative from conjugate[40]

From the examples so far, it is evident that conjugation can create novel structures and modes of transport for the cytotoxic drug to reach its target, but that efficacious application in the clinic is still very much in development. Our aim is to synthesise a duocarmycin SA based analogue with optimised functionality, for subsequent linkage to antibodies. The N-2 position of DSA can be a point of ligand attachment by the use of amino acid sequences, or attaching alkynes to explore "click" chemistry. The C-4 position functionality could also be explored as an entry point for the introduction of ligands.

2.6. Duocarmycins are excellent candidate cytotoxic agents

In summary, the duocarmycins present attractive structures because of their DNA binding abilities and cytotoxicity. There is great potential for structural flexibility to generate analogues as well as for conjugation due to amino acid like properties (acid and amine at the ends of the structure). This latter feature is highly attractive for linkage to antibodies for example.

2.7. Summary with project aims

To date, duocarmycin SA has only been synthesised on relatively small scale. Thus a more efficient synthesis of duocarmycin SA is desirable as the first aim. Synthesis of duocarmycin SA with C-4 in the benzyl protected form would generate gram quantities that could be safely manipulated, only releasing the highly potent cytotoxic activity at the end of the synthesis. This structure could then be conjugated to other moieties, antibodies being of particular interest. Our efforts towards this goal are presented in the following chapters.

3. Synthesising carboxyindole (CI) as a model structure for the duocarmycin SA alkylation subunit, DSA

The carboxyindole (CI) subunit is the minimum pharmacophore required for DNA alkylation activity of duocarmycin SA as discussed in chapter 2 (Introduction part 2). Whilst retaining the biologically significant functionalities, it has a relatively simple structure in comparison to DSA, the pharmacophore common to this group of natural products. For these reasons, CI was chosen as a model structure, as it could be synthesised by a faster and more efficient route than DSA. We aimed to model our synthetic route on that established by Warpehoski,[56] and to optimise it further by improving reaction conditions and yields.

Our synthetic route is shown in Scheme 13. Commercially available starting material 4-chloro-3-nitrophenol was taken through a series of steps; benzyl protection of the phenol group, malonation to introduce the diester group at C-1, which was subsequently reduced using DIBAL. Reduction of the nitro group, followed by one pot mesylation/cyclisation provided the indoline heterocyclic structure. The mesyl group was displaced by a chloride to set up the core *seco* CI model subunit. This structure could then be used to attach to linkers and other targeting moieties. A discussion of this synthesis follows.
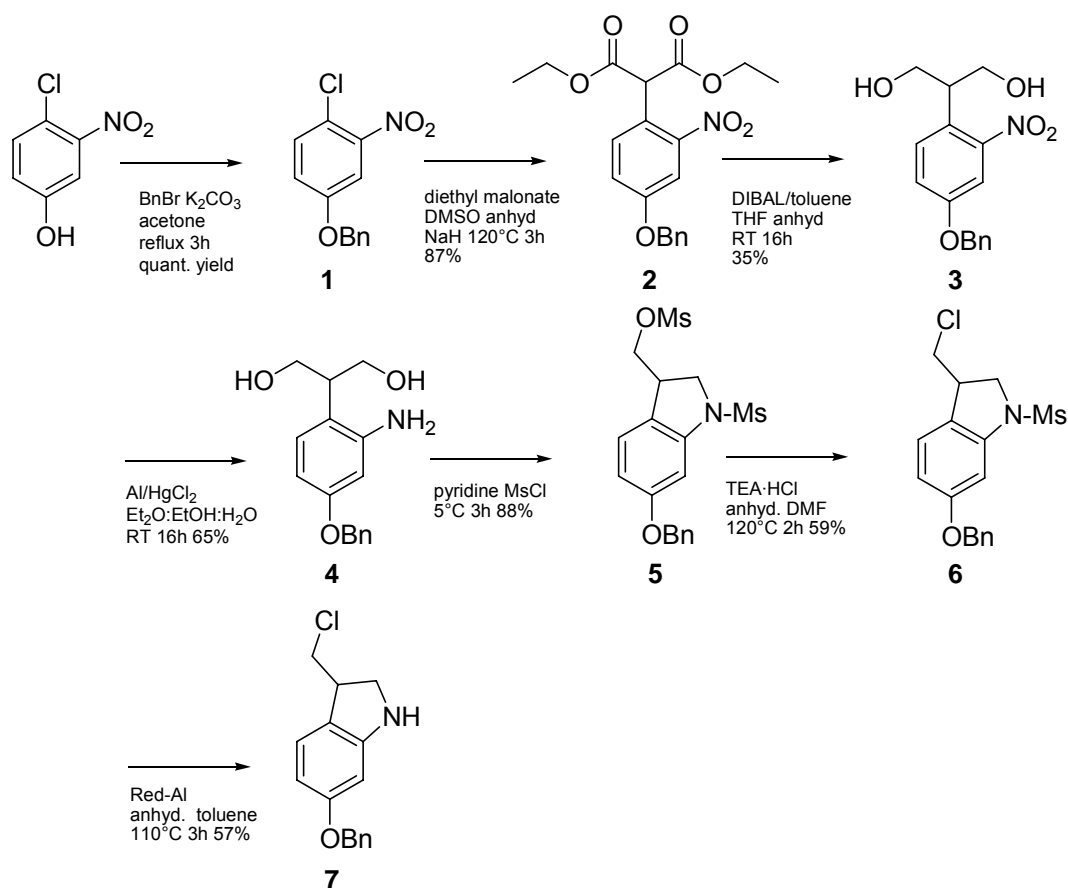
Scheme 13 Synthetic route leading to CI based model subunit

3.1.1 Benzyl protection of 4-chloro-3-nitrophenol

The first step in the synthesis of CI was protection of the biologically significant phenol group. Commercially available 4-chloro-3-nitrophenol was converted to a benzyl ether to form **1** in quantitative yield (Scheme 14).
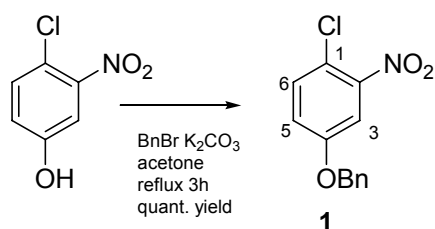
Scheme 14 Benzyl protection of 4-chloro-3-nitrophenol to form **1**

Benzylation occurs as shown in Scheme 15. The base (K_2CO_3) deprotonates the phenol group. Nucleophilic attack through an S_N2 mechanism occurs on the relatively electron-deficient carbon of benzyl bromide by the phenolate ion (**b**). Benzyl bromide is activated for this nucleophilic attack due to transition state orbital overlap stabilisation between the aromatic ring π-system and the orbitals involved in the transition state **c**. This ultimately liberates the bromide ion to form **1** (shown as **d**).
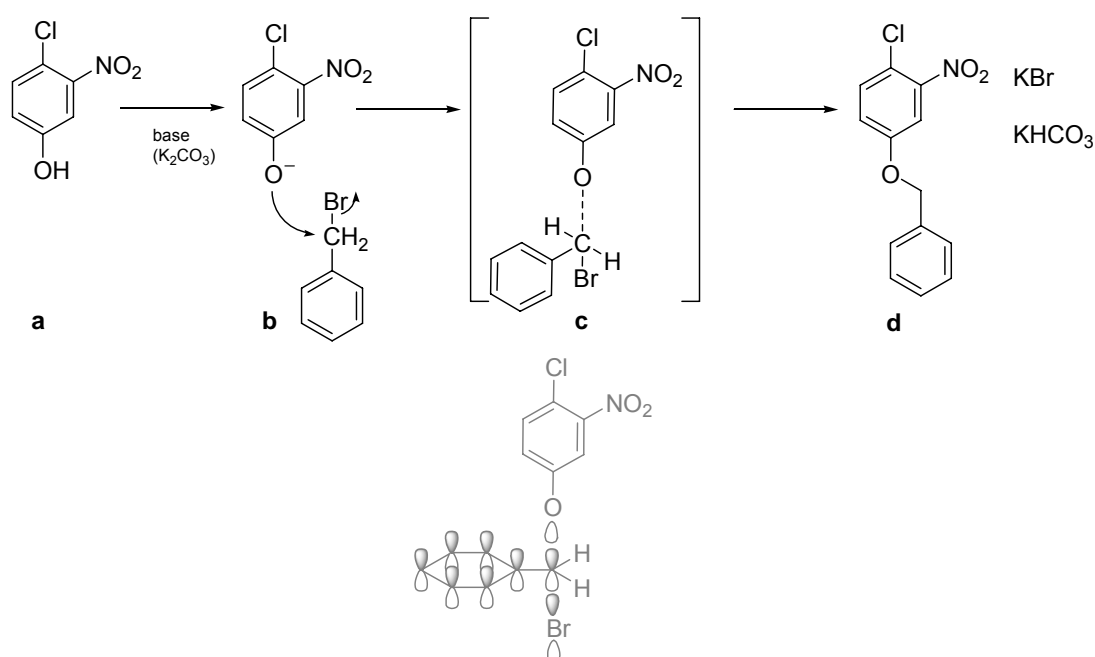
Scheme 15 Benzylation reaction mechanism to form **1** (transitional state orbital overlap shown in grey)

Benzylation was confirmed by the presence of ^{1}H NMR signals. A singlet at 5.10 ppm corresponded to two protons of the benzyl CH_2, and the multiplet at 7.42-7.44 ppm corresponded to the five benzyl aromatic ring protons. The protons on the phenol system represented an ABX system; the proton at C-5 appeared as a doublet due to *ortho* coupling with C-3 (J=2.6 Hz) and *meta* coupling with C-6 (J=9.4 Hz). The proton at C-6 was obscured within the multiplet of the benzyl aromatic protons, but was expected to be a doublet

with *ortho* coupling to C-5. The doublet at 7.48 ppm corresponded to the C-3 proton, which coupled to C-5 (J=3.2 Hz). This proton experiences a deshielded environment due to the neighbouring electron-withdrawing nitro group. Mass spectrometry confirmed the mass expected for this compound $C_{13}H_{10}ClNO_3Na$ (M+Na)+ 286.0241.

3.1.2 Alkyation of 1-chloro-2-nitro-4-phenylmethoxybenzene

The second step was malonate displacement at C-1. Dimethyl malonate and diethyl malonate were tested as candidate alkylating agents. Here, we report a significant improvement on Warpehoski's published method, which involved 24 hour reaction times, and recycling of unreacted starting material to yield only 77% of the product.[56] We reduced the reaction time to 3 hours and increased the yield to 87%.

Dimethyl malonate had previously successfully been used in our laboratory for a similar analogue (Scheme 16).[78] Displacement of halide substituents (fluoro and chloro) using dimethyl malonate was investigated using NaH (also anhydrous DMSO, heating at 120°C for 3 hours) to determine if this method was applicable to our route to CI.
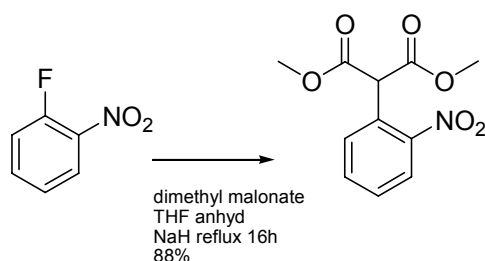
Scheme 16 Malonation of a fluoro compound using dimethyl malonate in 88% previously synthesised in-house

The use of dimethyl malonate for alkylation of the chlorobenzene compound consistently gave low yields (around 8%), in spite of efforts to optimise this by longer reaction times and increasing the molar ratio of dimethyl malonate. The use of newly purchased malonate also had no effect on the yield. Switching to diethyl malonate, for reasons that are not clear, resulted in successful alkylation to form **2** in 87% yield (Scheme 17). Adding malonate to NaH at a cooler temperature (approximately 0°C instead of below 30°C) followed by stirring at room temperature for 30 minutes before adding starting material was included. Activation of the malonate before the substrate appeared to both reduce the reaction time (3 hours) and increase the yield (87%).
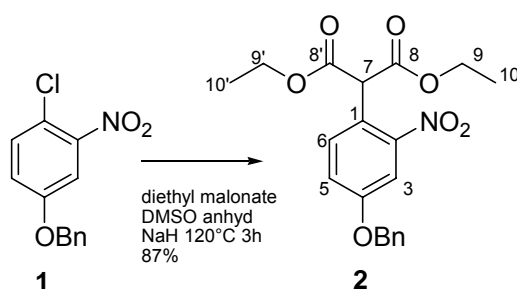
Scheme 17 Alkyation of 1-chloro-2-nitro-4-phenylmethoxybenzene **1** by diethyl malonate to form **2.**

Introduction of the malonate was confirmed by the presence of ^{1}H NMR signals at 1.26 ppm as a triplet (6 protons) and at 4.25 ppm (4 protons) as a quartet, which corresponded to two (CH_3) and (CH_2) groups of the malonate. The singlet at 5.21 ppm confirmed the presence of the central single dimalonate proton at C-7. Mass spectrometry confirmed $C_{20}H_{21}NO_7$ (M+H)$^+$ 388.1392. IR spectroscopy absorption at 1732 cm^{-1} due to carbonyl stretching and at 1225 cm^{-1} and 1023 cm^{-1} of the fingerprint region due to C-O, confirmed the presence of the malonate ester groups.

The malonation reaction mechanism is thought to proceed as follows (Scheme 18). Dimethyl malonate is activated by NaH; a hydride ion from NaH extracts a proton from the diethyl malonate central carbon (**a**), with keto-enol tautomerism stabilising the resultant anion (**b-c**). C-1 is activated for nucleophilic attack by the malonate carbanion due to the *ipso*-chloro and *ortho*-nitro substituents present on the aromatic ring (**d**). The transition state Meisenheimer complex **e** is stabilised by conjugation with the nitro group. Elimination of the chloride ion with regeneration of the conjugated system follows, to form the product **2** (shown as **g**).

It is not clear why the diethyl malonate would give better yields than the homologue dimethyl malonate. It may be that the methyl ester is less stable under these conditions than the ethyl ester, leading to side reactions involving, perhaps, hydrolysis. Although anhydrous solvents were utilised for these experiments, it is possible that residual moisture may have been present in the reagents and vessels.
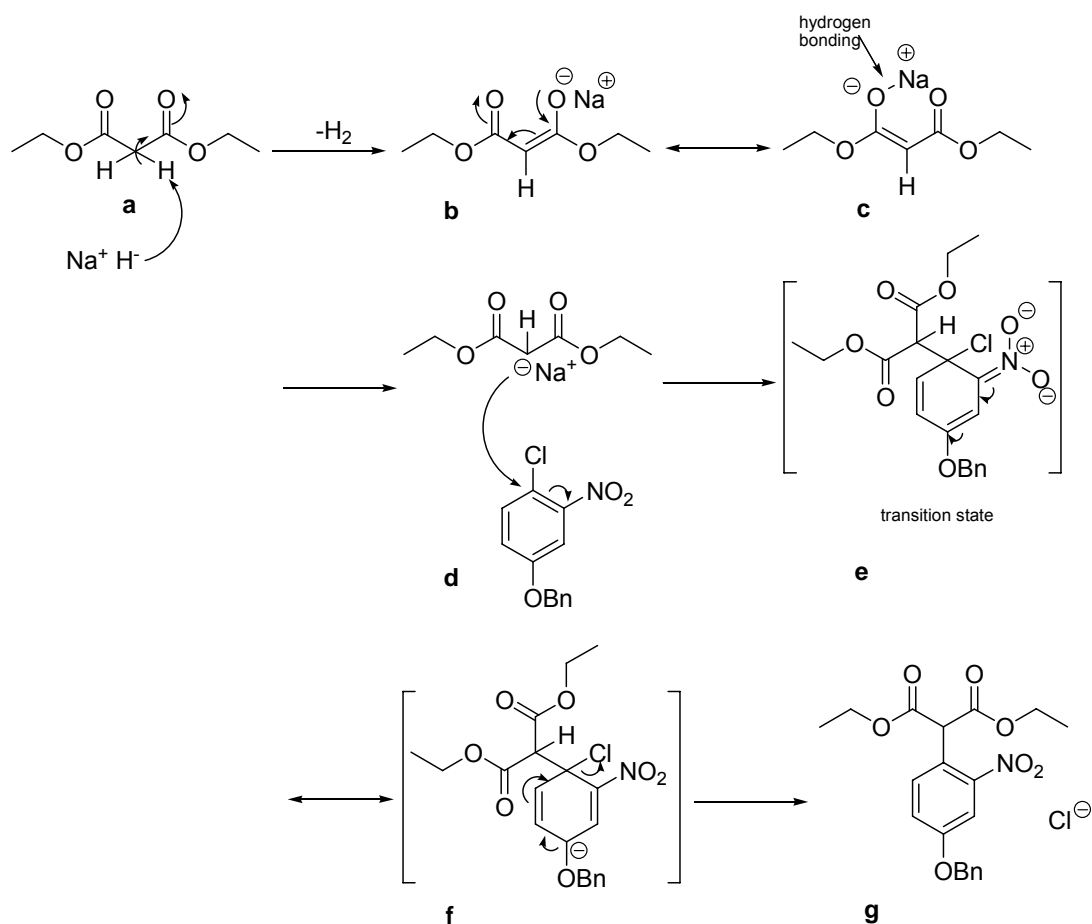
Scheme 18 Malonation reaction mechanism to form **2**

3.1.3 Diester reduction

At this juncture, Warpehoski's method requires reduction of the malonate group to a diol **3** using diisobutylaluminium hydride (DIBAL). In order to optimise this step, we compared yields of reduced malonate produced by DIBAL with other reducing agents including LiAlH$_4$ and NaBH$_4$. In our hands, use of DIBAL resulted in 35% yield (Scheme 19). This was comparable to Warpehoski's published method of 40% yield.[56]
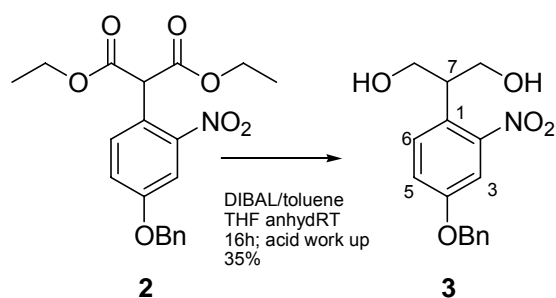
Scheme 19 Diester reduction to form **3**

DIBAL exists as a bridged dimer (Figure 17). In order to become a reducing agent, Lewis acid-base complex formation is required with the ester carbonyl. DIBAL has the advantages of relative ease of handling and control of the degree of reduction, i.e. reduction to aldehyde or complete reduction to alcohol (e.g. by temperature or molar equivalents used). However, it is not the most powerful reducing agent as it is more sterically hindered and thus has more difficulty transferring hydride ions.
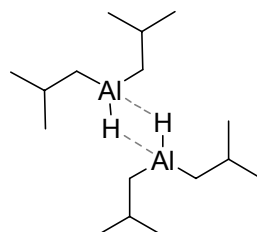
Figure 17 Diisobutylaluminium hydride (DIBAL) as a bridged dimer

Reduction may proceed as shown in Scheme 20. At least 4 equivalents of DIBAL are required to reduce a diester to a diol; we used 6. The DIBAL Al atom coordinates with the carbonyl oxygen (**a**). The critical step is hydride (H⁻) transfer (**b**) leading to a tetrahedral intermediate **c** which collapses to the aldehyde **e**. In the presence of a second equivalent of DIBAL (**f**), the resulting aldehyde intermediate (**h**) is further reduced as the aldehyde oxygen's electron lone pair now attack a second DIBAL Al atom (**i**). A second hydride transfer (**j**) results in a stable tetrahedral intermediate. This

stable intermediate collapses to form the alcohol only during acidic work up (**k**); we facilitated this by carefully pouring the crude reaction mixture into 1 M HCl. Reduction also occurs at the other ester, to ultimately form **3** (shown as **l**).

Reduction to diol was confirmed by ^{1}H NMR signals. The signals for the malonate diethyl groups were absent. The singlet for the central single C-7 proton was shifted upfield from 5.21 ppm in the malonate **2** to a multiplet at 3.5-3.6 ppm in the diol **3** as expected. This demonstrates its changing environment from the relative deshielding effect of the electron-withdrawing carbonyl groups in the esters, and upfield shift in the diol environment which is electron-donating. The signal at 4.0 ppm corresponded to four protons of the two CH_2 groups and appeared as a doublet due to coupling with the C-7 proton. A broad singlet representing two protons from the diol was not observed, as reported by Warpehoski.[56] This may due to the fact that in $CDCl_3$ the peak may be too broad to be easily observed due to labile OH protons which undero rapid exchange with each other. TLC monitoring showed that the reaction did not reach completion despite optimisation efforts such as prolonged reaction times, contributing to low yields. As aluminium hydrides are inherently unstable, degradation is a possibility and good storage is essential. Even lower yields were observed with older batches of DIBAL.
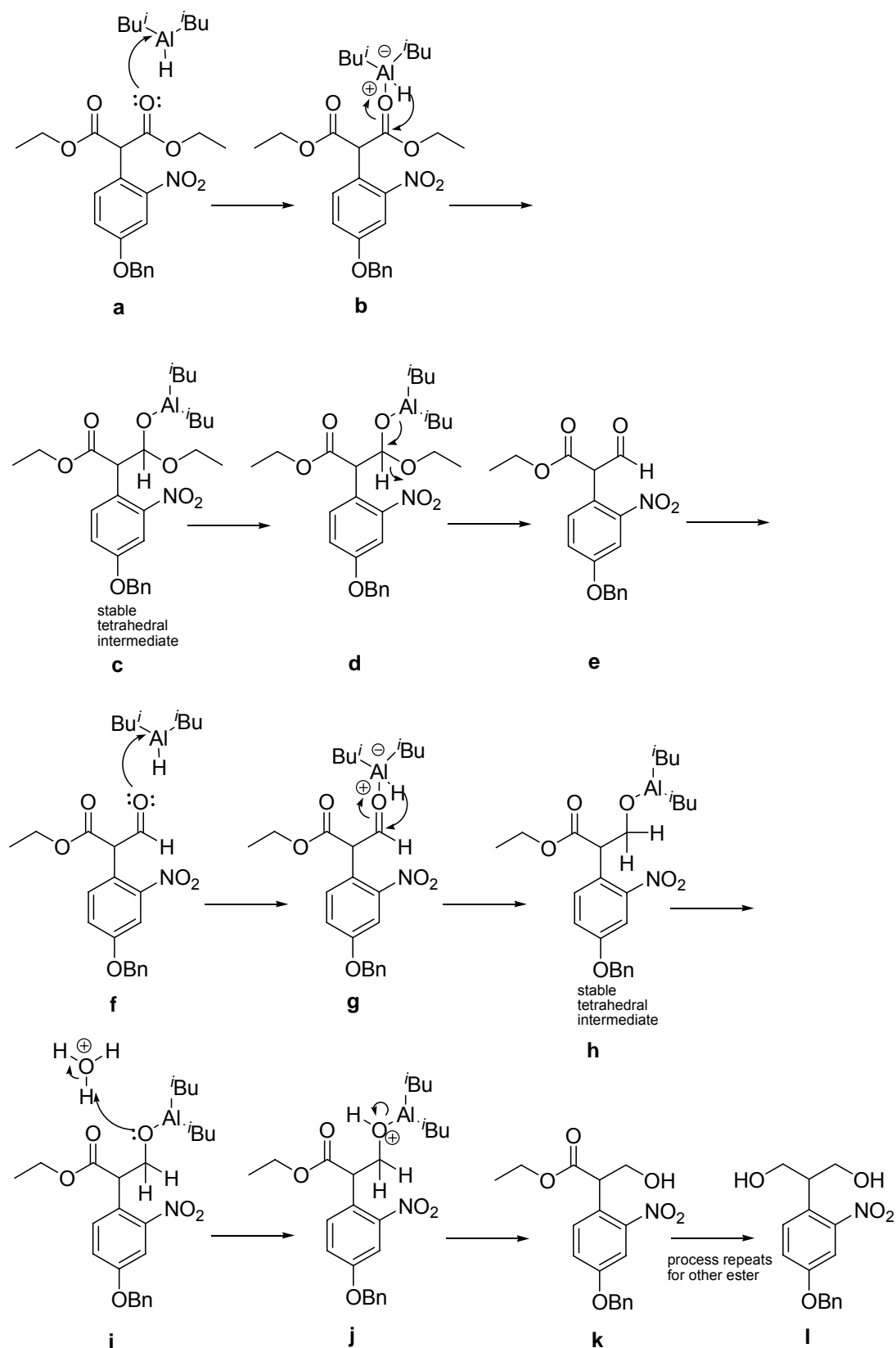
Scheme 20 Possible reaction mechanism for ester reduction by DIBAL to form **3**; 1 equivalent of DIBAL reduces the ester to aldehyde, the second equivalent of DIBAL reduces the aldehyde further to alcohol.[79]

Due to low yields of malonate product with DIBAL, various other reducing agents were investigated to optimise yields. The most common hydride transfer reducing agent, LiAlH₄ is powerful but relatively non-specific, with handling issues due to its reactivity with water and most other solvents. This did not provide a successful reaction as starting material decomposed. NaBH₄ is a less powerful reducing agent than LiAlH₄, but has better aqueous stability allowing greater variety in reaction conditions. However no reaction occurred with NaBH₄ as it reacts too slowly with esters. Esters by nature are not very reactive due to the delocalisation of electrons over the carbonyl and C-O bonds. A 1 M solution of DIBAL in THF was initially used as it was thought that retaining the same solvent for all parts of the reaction would be beneficial. An interesting observation is that reduction did not occur using this solution whereas 1 M solution of DIBAL in toluene provided a successful reduction.

Overall, after investigating various reducing agents, DIBAL provided the best yield for malonate reduction. Thus, synthesis was continued to the next step.

3.1.4   Nitro group reduction

The next step was reduction of the nitro group to form **4** in 65% yield using an Al-Hg amalgam method (Scheme 21).
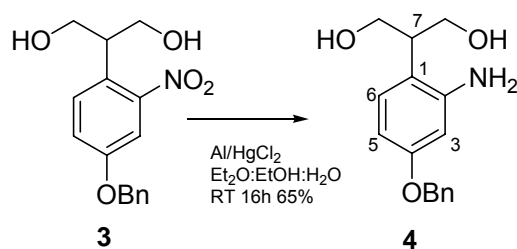
Scheme 21 Nitro group reduction to form **4**

The Al-Hg amalgam method, first described by Corey and co-workers,[80] has been previously successfully used by our group and others.[81] It was chosen as standard nitro group reduction conditions such as hydrogenation over Pd/C may result in preferential debenzylation. Thus, we achieved a 65% yield with the use of Al-Hg amalgam methodology which provides relatively mild, simple and cost-effective conditions for reduction. It is recognised that there are environmental considerations regarding Hg use, which at least in commercial scale up may lead to alternative conditions being sought. This yield is comparable to Warpehoski's 69% yield which used a relatively expensive PtO_2 method.[56]

The Al-Hg amalgam reduction is thought to work via a bimolecular reduction[82] where Hg oxidises Al, disrupting the latter's outer protective Al_2O_3 layer. This exposes more Al to oxidation for the subsequent reduction of the nitro group. The nitro group is reduced to form the amine by a gain in electrons where the aluminium is a source of electrons and the solvents (water and ethanol) provide the protons.

Reduction was confirmed by comparison of chemical shifts of ^1H NMR signals of the reduced amino product **4** to those of the nitro starting material **3**. The ^1H NMR multiplet for the central single diol proton at C-7 was upfield at 3.07-3.15 ppm compared to 3.5-3.6 ppm in **3**. This suggests a long

range effect from the electron-withdrawing nitro group on this proton in **3**. All aromatic protons also showed a more shielded, upfield shift due to the electron-donating amino substituent. The signal at 6.42 ppm due to the proton at C-5 was shifted from 7.18 ppm in **3**. The proton at C-3 was now observed at 6.45 ppm as a doublet (*J*=8.7 Hz) due to *meta* coupling with C-5, and the proton at C-6 appeared as a doublet at 6.93 ppm due to *ortho* coupling with C-5, whereas they appeared as part of a multiplet at 7.31-7.46 ppm in **3**. The proton at C-5 appeared at 6.42 ppm (compared to 7.18 ppm in **3**) as a doublet of doublets due to *ortho* coupling with C-6 and *meta* coupling with C-3.

3.1.5   One pot cyclisation

Mesylation and ring closure achieved one pot cyclisation to form the indoline **5** in 88% yield (Scheme 22), an improvement on the method reported by Warpehoski (59% yield).[56]
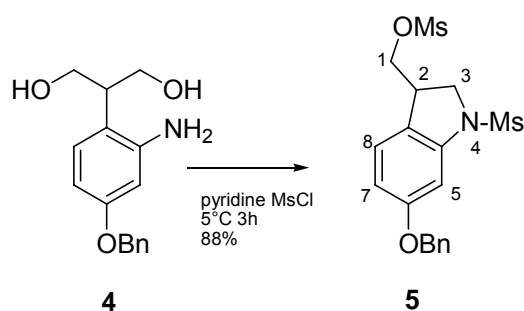
Scheme 22 One pot mesylation/cyclisation to form **5**

Mesylation, the first step in the one pot cyclisation is shown in Scheme 23. First, nucleophilic attack by the alcohol oxygen electron lone pairs occurs at the mesyl chloride sulphur atom, resulting in the elimination of chloride from the intermediate. The base, pyridine then extracts a proton (Scheme 23a).[83]   Attack by the amine group is unlikely; if mesylation of the nitrogen

did occur prior to cyclisation, this would form a sulphonamide which is much less reactive and thus unlikely to undergo the cyclisation reaction without strongly basic conditions. The nucleophilic electron lone pair on the amine then attacks the mesyl α carbon, resulting in a *5-exo-tet* intramolecular ring closure to give the cyclised indole product. On cyclisation, the secondary amine becomes activated for mesylation due to the extra positive-inductive effect of the alkane from the indoline ring. The indoline nitrogen lone pairs attack a sulfene in the same manner as before to become mesylated (Scheme 23b).

Cyclisation was confirmed by the presence of ^{1}H NMR signals. The singlet at 2.86 ppm corresponded to the three mesyl protons at N-4. The nitrogen is electron-donating resulting in slight chemical shielding for these protons compared to the mesyloxy group at C-1. The singlet at 2.97 ppm corresponded to the three mesyl protons at C-1. The multiplet at 3.91-4.09 ppm was due to two protons at 1-CH$_2$ and one proton at 2-H. The multiplet at 4.16-4.23 ppm (2 protons) corresponded to 3-CH$_2$. The doublet of doublets peak at 6.67 ppm was due to the proton at C-7, which has *ortho* coupling with C-8 (J=8.1 Hz) and *meta* coupling to C-5 (J=2.3 Hz). The doublet at 7.10 ppm was due to the C-5 proton *meta* coupling with C-7 (J=1.8 Hz). The C-8 proton is not clearly distinguishable as it overlaps with the benzyl aromatic proton multiplet at 7.29-7.48 ppm.
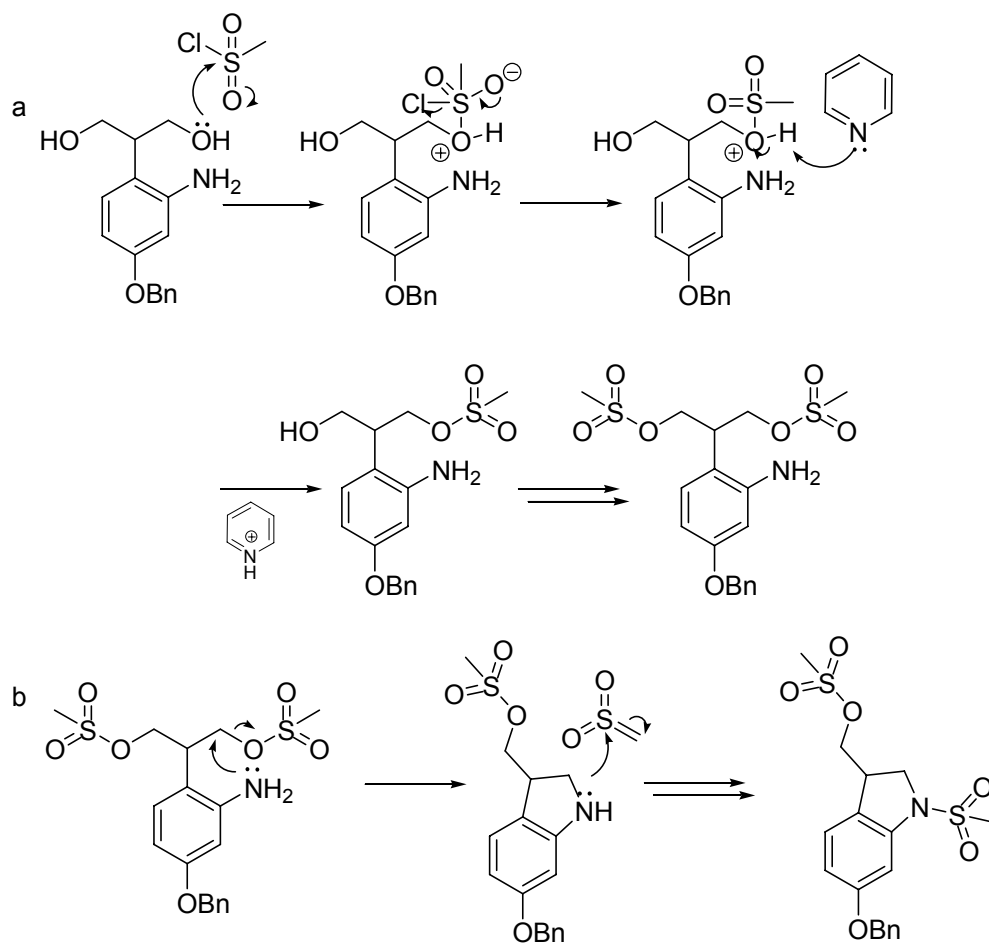
Scheme 23 One pot mesylation-cyclisation to form **5**; a. mesylation, b. 5-exo-tet intramolecular ring closure and indoline mesylation

3.1.6 Mesyl displacement by chloride

Cyclisation was followed by displacement of the mesyl group by a chloro substituent using triethylamine hydrochloride (TEA·HCl) to form **6** in 59% yield (Scheme 24).
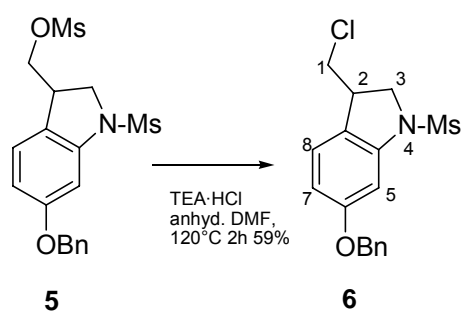
Scheme 24 Mesyl displacement by chloride, to form **6**

Scheme 25 shows the reaction mechanism for mesyl group displacement by a chloride. This is an S_N2 reaction in which the chloride ion of triethylamine hydrochloride attacks the carbon adjacent to the mesyl group, which acts as an excellent leaving group.
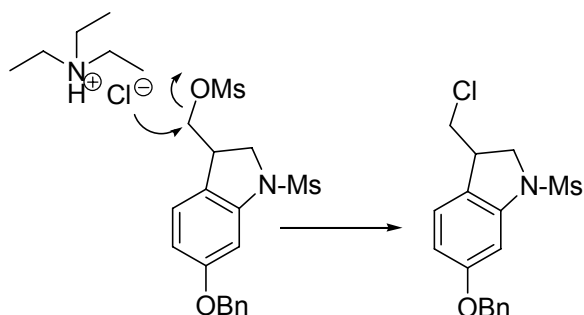
Scheme 25 Reaction mechanism for mesyl displacement by chloride to form **6**

Chloride substitution was confirmed by the absence of the C-1 O-mesyl peak by ^{1}H NMR spectroscopy. Here the restricted rotation at C-1 - C-2 results in each of the two 1-CH$_2$ protons appear as a distinct multiplets as the net effect of the subtly differing chemical environments. The multiplet at 3.56-3.60 ppm arose due to 1 proton at 1-CH, whilst the peak at 3.65-3.74 ppm resulted from the overlap of one 1-CH proton and the CH proton at C-2. The multiplets at 3.95-3.99 ppm and at 4.05-4.15 ppm corresponded to each of the two protons at 3-CH$_2$. All protons at C-1, C-2 and C-3 showed small upfield

chemical shifts with respect to the starting material **5** as the chloro substituent is relatively less electron withdrawing than the O-mesyl group. The aromatic protons at C-5, C-7 and C-8 were not affected by the new chloro substituent as NMR signals do not show any change in chemical shift. Mass spectrometry confirmed the expected mass $C_{17}H_{19}ClNO_3S$ (M+H)$^+$ 352.0765.

3.1.7   N-indoline deprotection

Deprotection of the indoline amine provided the structure for our model CI-based subunit **7**. The unprotected indoline amine would provide an access point for attaching other moieties, to improve targeting. Removal of the indoline N-mesyl protecting group proved particularly difficult. Initial efforts involved very simple conditions as described by Hiroya and co-workers using K_2CO_3 and MeOH, stirred at room temperature.[61] Refluxing in these conditions, and in hydrobromic acid (48%) in phenol[84] gave no reaction, with the return of starting material only. It is presumed that although these conditions have been described for mesyl group removal, the *N*-mesyl bond is too strong in the indoline environment for hydrolysis to occur. Successful removal of the mesyl group to form **7** was then achieved using sodium bis(2-methoxyethoxy)aluminium hydride (Red-Al) in 57% yield (Scheme 26).[56]
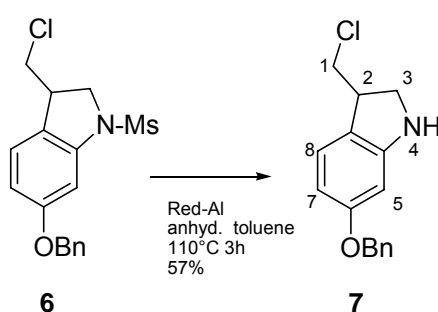
Scheme 26 N-indoline deprotection to form **7**

Deprotection was confirmed by the absence of a ^{1}H NMR peak for the N-mesyl CH$_3$ group at 4-position. The doublet of doublets at 6.97 ppm corresponded to the C-7 proton which couples to C-8 (*J*=8.8 Hz) and C-5, appears slightly more downfield than in **6**. This may be due to the absence of long range electron donating effects of the indoline mesyl group in **6**. The protons at C-5 and C-8 present as an overlapping doublet of doublets at 6.32-6.39 ppm. Yields were significantly reduced on scale-up, and for reasons that are not clear, this methodology was not reproducible. It may be that Red-Al or even the indoline itself is unstable and decomposed in storage, or during the reaction.

In comparison with other reducing agents (such as DIBAL), the electron donating alkoxy groups of Red-Al provide increased reactivity (Figure 18), and the extended side chains confer a better solubility profile.[85] It is probable that hydrolysis of the mesyl ester proceeds by coordination of the sulfonyl oxygen to the Al of Red-Al, in a manner similar to that outlined for the reduction of esters by DIBAL earlier (Scheme 20). Hydride transfer forms a tetrahedral intermediate, ultimately resulting in cleavage of the mesylate to give the free indoline N-H.
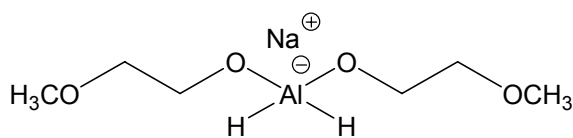
Figure 18 Sodium bis(2-methoxyethoxy)aluminium hydride (Red-Al® or Vitride®)

Deprotection of the indoline amine provided the structure for our model CI-based subunit **7**. The unprotected indoline amine would provide an access point for attaching other moieties, with the aim of improving targeting.

3.1.8 Investigations into attachment of other moieties to 7

Our aim was to use the free indoline N-H of our model CI-based structure **7** as a point of attachment. Any moiety could be accessible to **7** by amide bond formation. Two initial ideas were to couple alkynes, in order to explore click chemistry, and also to attach amino acids (Scheme 27).
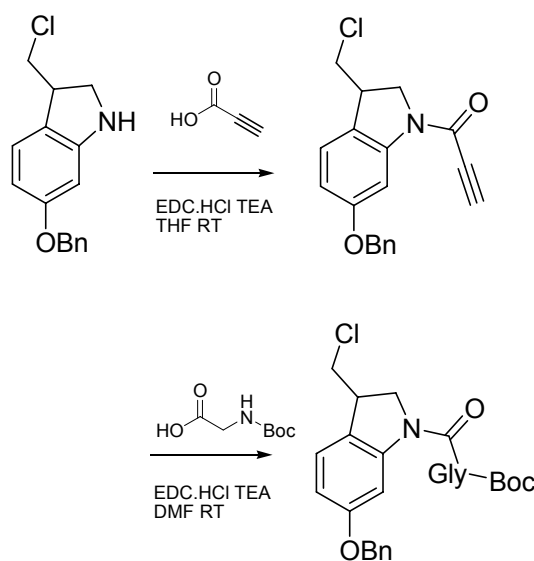
Scheme 27 Use of amide bond formation to couple alkynes,or amino acids to **7**

Initial studies on the attachment of both amino acids and propargyl groups to the CI subunit **7** were inconclusive. This was due to the lack of stability of the free amino compound **7** and the low yields obtained with Red-Al reduction on scale up. In retrospect, it may have been more useful to introduce a Boc group in place of the mesylate at an earlier stage in the synthesis, so that cleavage and subsequent coupling could be carried out without isolation of the amine intermediate. Although the coupling reactions were unsuccessful, reaching this stage prompted the study of potential right hand subunits for attachment to peptides and antibodies and the following section describes the synthesis of these compounds.

3.2 Synthesis of right hand duocarmycin SA based subunits

A small library of phthalamide based duocarmycin SA "right hand" subunits was designed to introduce amine linkers. These subunits could be attached via amide bond formation to **7**, or to any other subsequent duocarmycin SA "left hand" (DSA) subunits synthesised. The phthalamide group was coupled to the indole nitrogen of the indole carboxylate using NaH and tetrabutylammonium iodide (TBAI) as shown in Scheme 28. A description of the synthesis of this library of compounds follows.
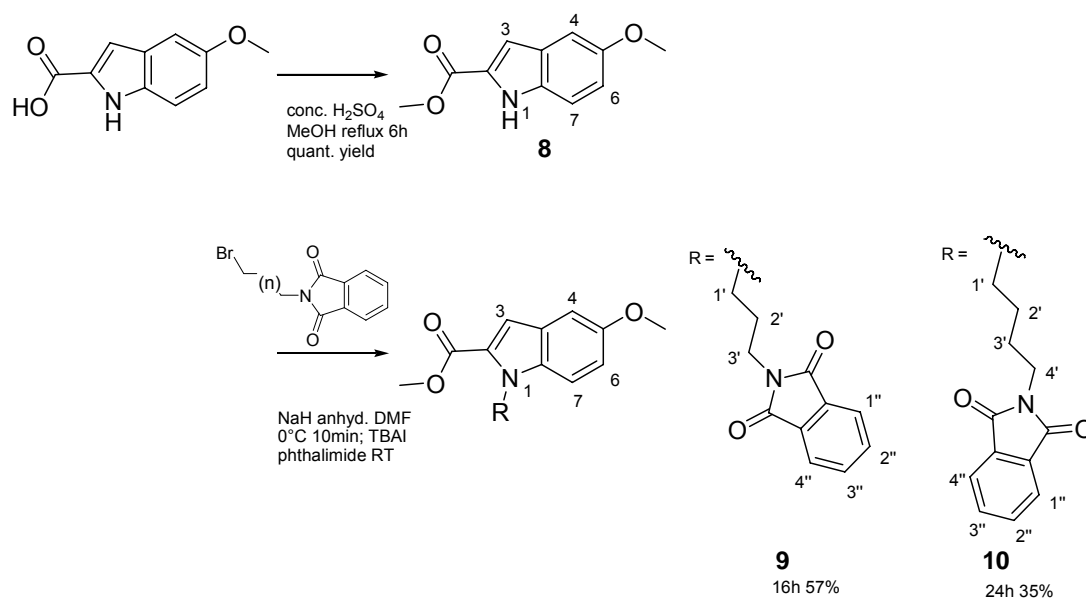
Scheme 28 Synthesis of 5-methoxy methyl ester phthalamide duocarmycin SA right hand subunits **9** and **10.**

For the 5-methoxy methyl ester phthalamide compounds (**9** and **10**), commercially available 5-methoxyindole 2-carboxylic acid was first refluxed in MeOH in concentrated H_2SO_4[86] to form the methyl ester product **8** in quantitative yield (Scheme 28).

Esterification occured under acidic conditions here, as shown in Scheme 29. The electron lone pair of the carbonyl becomes protonated by strong acid (**a**). This renders the carbonyl carbon highly susceptible to nucleophilic attack by alcohols (**b**), forming an unstable tetrahedral intermediate (**c**). The oxygen atoms become protonated in the presence of acid (**d**), with the loss of water pushing the reaction towards ester formation to form **8** (shown as **g**).
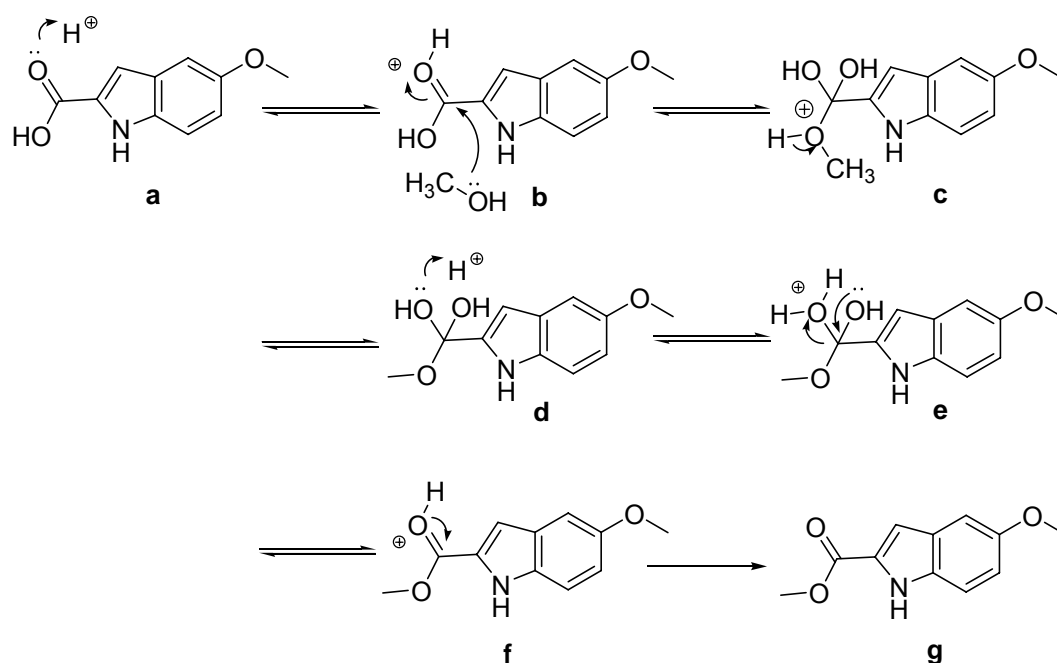
Scheme 29 Reaction mechanism for ester hydrolysis, to form **8**

Esterification was confirmed by ^{1}H NMR spectroscopy. The singlet at 3.85 ppm was due to the 5-methoxy group (3 protons). The singlet at 3.94 ppm corresponded to the 2-carboxylate ester CH$_3$; the neighbouring electron withdrawing ester group exerts a deshielding effect. The doublet of doublets peak at 7.00 ppm was due to C-6 (1 proton) *ortho* coupling with C-7 (*J*=8.8 Hz) and *meta* coupling with C-4 (*J*=2.4 Hz). The singlet at 7.08 ppm was due to one proton at C-3. The peak at 7.14 ppm was due the C-4 proton. It probably appears as a broad doublet due to *meta* coupling with C-6 and long range coupling with C-3. The doublet at 7.32 ppm was due to one proton at

C-7 showing *ortho* coupling with C-6 (*J*=8.8 Hz). The broad singlet at 8.92 ppm was due to the indole NH proton. Mass spectroscopy confirmed the expected mass $C_{11}H_{11}O_3N$ (M)$^+$ 205.0732.

3-Bromopropyl phthalimide was then coupled to the methyl ester **8** using NaH and TBAI[87] to form **9** in 57% yield (Scheme 28). The mechanism may proceed as shown in Scheme 30. NaH deprotonates the indole nitrogen, resulting in the formation of a sodium salt of the indole. TBAI is a source of I⁻, replacing Br for I, to make a more reactive electrophile. TBAI may also be acting as a phase transfer catalyst. The carbon α to the bromo substituent of the phthalimide then undergoes nucleophilic attack by the electron lone pair on the indole nitrogen, with the bromide acting as a good leaving group. This forms the coupled product **9**.
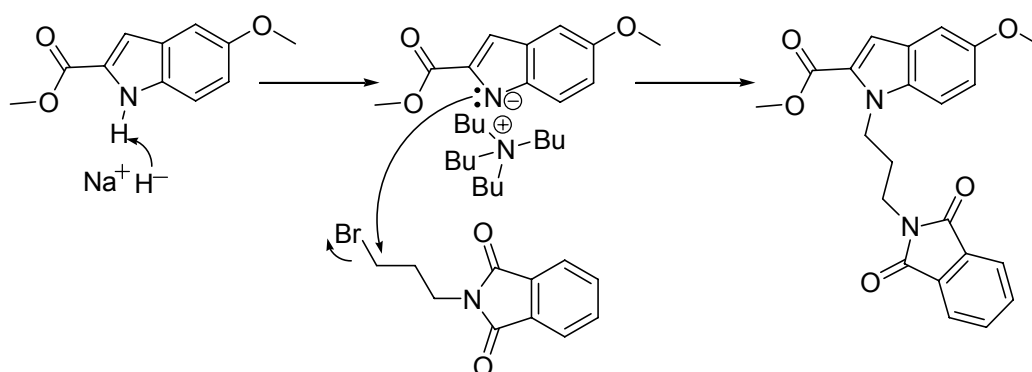
Scheme 30 Reaction mechanism for phthalimide coupling, to form **9**

The formation of **9** was confirmed by ¹H NMR and mass spectrometry. The propyl CH₂ groups showed distinctive signals; the peak at 2.20 ppm was a quintet due to 2'-CH₂ coupling with CH₂ groups on either side. The triplet at 3.77 ppm was due to 1'-CH₂ vicinal coupling with 2'-CH₂ (*J*=7.2 Hz). 3'-CH₂ was the most downfield of the propyl CH₂ groups at 4.61 ppm, due to the electron-withdrawing effect of the phthalimide carbonyls. This was a triplet

due to coupling with 2′-CH$_2$ (*J*=7.6 Hz). The 2-carboxylate ester CH$_3$ and the protons at C-4 snd C-7 show a slight upward chemical shift as they were now seen at 3.84 ppm (3.94 ppm in **8**), and 7.03 ppm (7.14 ppm in **8**) and 7.28 ppm (7.32 ppm in **8**) respectively, possibly due to a long range shielding effect from the propyl group. The proton at C-3 showed a slight downward shift relative to the starting material **8**. The aromatic protons of the phthalimide ring appeared as a multiplet at 7.72 ppm for the two protons at C-2″ and C-3″. The peak splits due to coupling with protons at C-1″ and C-4″. The multiplet at 7.84 ppm was due to the two protons at C-1″ and C-4″, coupling with C-2″ and C-3″. It appears relatively more deshielded than that for C-2″ and C-3″ due to the long range electron-withdrawing effect of the carbonyl groups. Mass spectroscopy confirmed the mass as expected for C$_{22}$H$_{24}$O$_5$N$_3$ (M+ NH$_4$)$^+$ as 410.1714.

The butyl compound **10** was synthesised in 35% yield (Scheme 28). This was confirmed by ^{1}H NMR spectroscopy. The 'central' butyl CH$_2$ groups 2′ and 3′ appeared at 1.73 ppm and 1.85 ppm respectively. They split as quintets due to coupling with CH$_2$ groups either side. The triplet at 3.71 ppm was due to 1′-CH$_2$ coupling with 2′-CH$_2$ (*J*=7.0 Hz). The triplet at 4.58 ppm was due to 4′-CH$_2$ being deshielded by the phthalimide carbonyl groups. The proton at C-3 at 7.20 ppm appeared as a doublet, possibly due to coupling with the C-4 proton. Phthalimide related signals (for protons C-1″ 2″ 3″ 4″) showed the same splitting patterns as for **9**. Mass spectrometry confirmed the expected mass as C$_{23}$H$_{22}$O$_5$N$_2$ (M)$^+$ at 406.1511.

The ethyl ester propyl phthalimide **11** was synthesised from commercially available ethyl indole-2-carboxylate in quantitative yield as shown in Scheme 31. The structure was confirmed by ^{1}H NMR spectroscopy. Signals for the ester group were a triplet at 1.33 ppm due to CH$_3$ (3 protons) coupling to the

neighbouring CH$_2$, and a quartet at 4.26 ppm due to CH$_2$ (2 protons) coupling to the neighbouring CH$_3$. The propyl group's signals appeared at 2.19 ppm as a quintet for 2'-CH$_2$, a triplet for 1'-CH$_2$ at 3.76 ppm, and a more deshielded triplet at 4.62 ppm for 3'-CH$_2$. The splitting pattern was due to coupling as outlined for **9**. The septet at 7.10 ppm corresponded to the proton at C-5, due *ortho* coupling to C-4 and C-6 (*J*=8.0 Hz) and *meta* coupling with C-7. The singlet at 7.26 ppm was due to the proton at C-3. The doublet of doublet of doublets at 7.30 ppm corresponded to C-6 (overlaps with C-3), which *ortho* coupled to C-5 and C-7. The doublet at 7.33 ppm corresponded to C-4. The doublet at 7.62 ppm corresponded to C-7, with *ortho* coupling to C-6 (*J*=8.4 Hz). The doublet of doublets at 7.69 ppm and 7.82 ppm corresponded to the phthalimide aromatic protons C-2 and C-3, C-1 and C-4 respectively. Phthalimide related signals (for protons C-1'' 2'' 3'' 4'') showed the same splitting patterns as for **9**. Mass spectrometry confirmed the expected mass MS (ES+) C$_{22}$H$_{24}$O$_4$N$_3$ (M + NH$_4$)$^+$ for 394.1764.
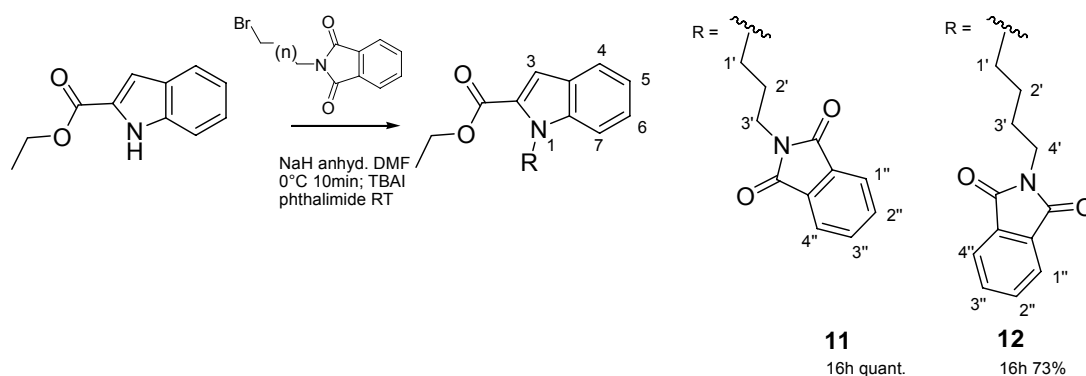
Scheme 31 Synthesis of ethyl ester phthalimide duocarmycin SA right hand subunits **11** and **12**

The ethyl ester butyl phthalimide **12** was synthesised in 73% yield (Scheme 31). The structure was confirmed by ^{1}H NMR spectroscopy. The butyl CH$_2$ group signals were the quintet at 1.71 ppm which corresponded to 2'-CH$_2$, the quintet at 1.84 ppm was due to 3'-CH$_2$, the triplet at 3.68 ppm was due to

1′-CH$_2$, and the triplet 4.58 ppm due to 4′-CH$_2$. Coupling occured in the same way as for **10**. The phthalimide aromatic protons C-2′′ and C-3′′ appeared at 7.67 ppm, and C-1′′ and C-4′′a appeared at 7.80 ppm. The splitting pattern was as for **9**. Mass spectrometry confirmed the expected mass as C$_{23}$H$_{26}$O$_4$N$_3$ (M + NH$_4$)$^+$ 408.1917.

3.3 Summary

In summary, a model CI-based subunit, **7** was synthesised (Scheme 13), with optimisation of some steps within the route, compared to that reported by Warpehoski.[56] Reaction conditions for malonation were developed; shorter reaction times (3 hours compared to 24 hours), and high yields (87%, compared to 77%) were achieved without recycling of starting material. Aromatic nitro group reduction was achieved in comparable yields (65%) to published methods (69%) via a cost-effective Al-Hg amalgam method. The key one pot mesylation/cyclisation step to form the indoline heterocycle was achieved in 88%, a significant improvement to published yields (59%). The difficult deprotection of the indoline nitrogen was achieved using Red-Al in 57% yield, providing the core structure for our model CI-based subunit **7**.

A phthalimide-based compound library (**9-12**) was synthesised in good yields (Scheme 28 & Scheme 31) to introduce amine terminus linkers on hydrolysis of the *N*-phthalimide, to which other moieties could be attached via amide bond formation to alter the DNA binding or interaction properties of duocarmycin SA subunit analogues. Ester hydrolysis at C-2 of the phthalimide compounds would also enable coupling via amide bond formation to our model CI-based subunit **7**, or more complex subunits.

Overall, we have achieved a synthetic route with overall greater efficiency, with significant reduction in time and monetary cost, in addition to increasing the yield of the CI subunit.

# 4. First route to the duocarmycin SA alkylation subunit, DSA

To date, syntheses for duocarmycin SA subunits have ranged from ten to twenty steps with only 3-12% yields.[47]   Our objective was to seek a concise and efficient route to DSA, the left hand alkylation subunit, as it is responsible for the biological activity of duocarmycin SA. This would be used to generate duocarmycin SA analogues, and linkage to other moieties.

Retrosynthetic analysis of the alkylation subunit of duocarmycin SA, known as Boc-DSA, indicated a sequential formation of two heterocycles. Key disconnections for our chosen synthetic route were the formation of the indole, disconnection of the five membered ring bond, and direct allylation (Scheme 32).

In this route we aimed to utilise direct allylation of the amino group of our substrate to introduce the six membered heterocycle, which would ultimately form the spirocyclised cyclopropane ring. Ek allylation had previously been successfully used in our laboratory to generate quinolone analogues (Scheme 33). We wanted to determine if this reaction could be applicable for our route.
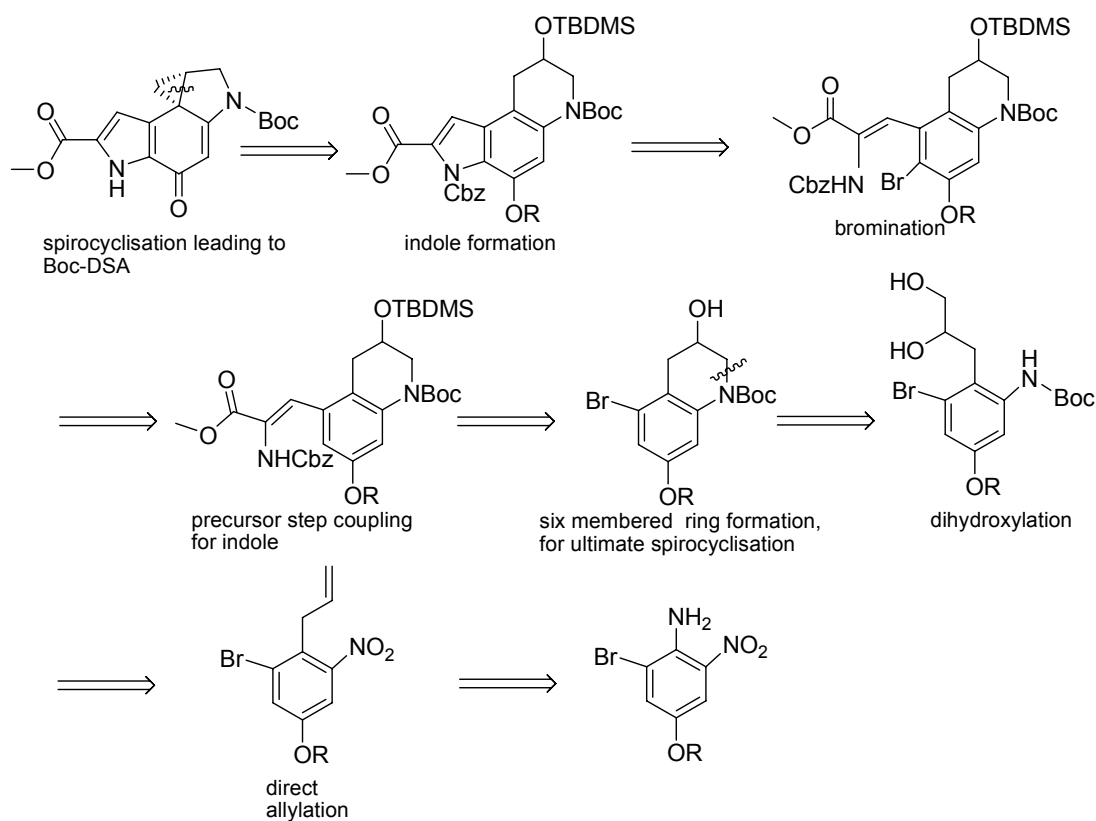
Scheme 32 Retrosynthetic analysis of Boc-DSA
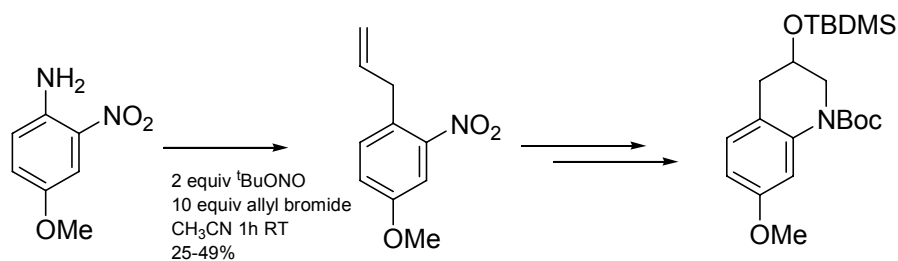
Scheme 33 Synthesis to quinolone analogue utilising Ek allylation in 49% yields previously achieved in-house in our laboratory

Scheme 34 summarises the initial steps of the intended synthetic route; phenol group protection via benzylation, bromination at C-2 (to allow indole formation later in the synthesis), and introduction of the allyl group at C-4 for ultimate formation of the spirocyclised indoline. A discussion of this synthetic route follows.
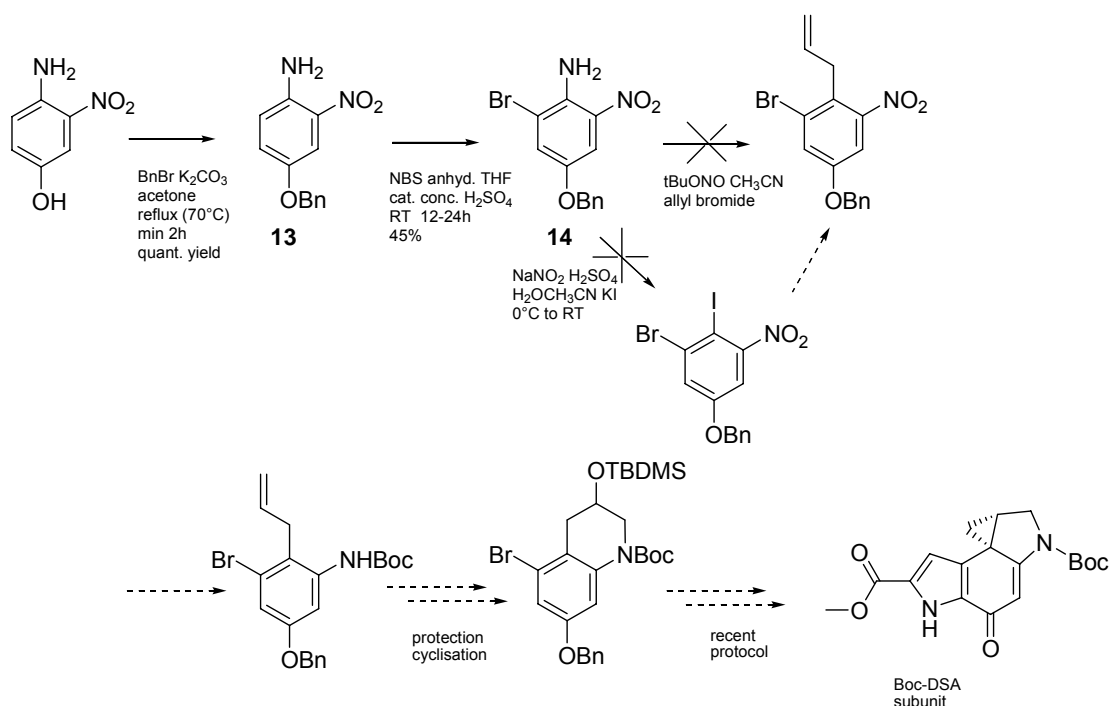
Scheme 34 Overview of the synthetic route to Boc-DSA with benzyl protecting group, employing direct allylation

4.1.1. Benzyl Protection of 4-amino-3-nitro-phenol

The synthetic route commenced using 4-amino-3-nitrophenol as the starting material. First, protection of the phenol group using benzyl bromide was achieved in quantitative yield to form **13** (Scheme 35).
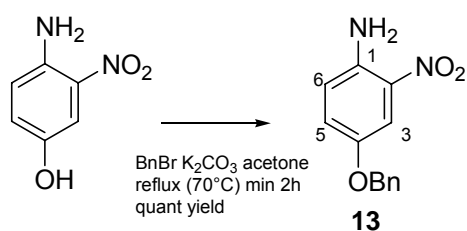
Scheme 35 Benzyl Protection of 4-amino-3-nitro-phenol to form **13**

Benzylation proceeds by S_N2 type attack on the reactive benzyl bromide by the phenolate ion (Scheme 36; see chapter 2, Scheme 15 for further description).
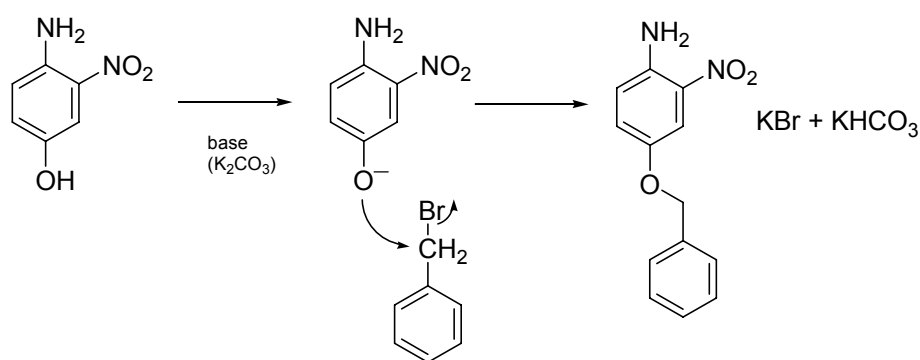
Scheme 36 Benzylation reaction mechanism to form **13**

Benzylation was confirmed by the presence of ^{1}H NMR signals. The singlet at 5.22 ppm corresponded to the two benzyl CH_2 protons. The broad singlet at 5.90 ppm was due to the amino group at C-1. The doublet peak at 6.78 ppm was due to the proton at C-6. The *J* value of 8.8 Hz indicates *ortho* coupling with C-5. This peak was the most upfield of all the aromatic protons due to the electron shielding effect of the *ortho* amino group. The doublet of doublets signal at 7.15 ppm corresponded to the proton at C-5, with *ortho* coupling to C-6 (*J*=9.0 Hz) and *meta* coupling with C-3 (*J*=3.0 Hz). The multiplet at 7.33-7.46 ppm corresponded to the five benzyl ring aromatic protons. The doublet at 7.67 ppm was due to the proton at C-3, which *ortho* couples to C-5 (*J*=2.8 Hz). This aromatic proton was deshielded due to the electron withdrawing effects of the *ortho* nitro group. Mass spectrometry confirmed the mass $C_{13}H_{11}O_3N_2$ (M)$^-$ 243.0776.

4.1.2. Bromination at C-2 position

Benzylation was followed by bromination at C-2 to form **14**, using N-bromosuccinimide (NBS) and catalytic H_2SO_4 in 45% yield (Scheme 37).
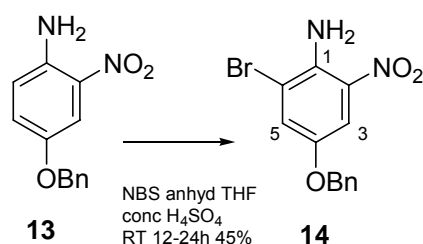
Scheme 37 Bromination at C-2 position to form **14**

The reaction proceeds in two steps (Scheme 38). First, NBS is activated through protonation in the presence of H_2SO_4, forming a species which effectively represents a bromonium ion (**a**-**b**). Regioselective bromination is dependent on the substituents on the aromatic ring. Benzyl substituents activate positions *ortho* and *para* to it by conjugative effects; in this case the *para* position is already occupied by the amino substituent. The nitro group deactivates *ortho* and *para* positions through its electron-withdrawing effects, directing towards positions *meta* to it which is C-2 in this case, as C-4 is substituted by the benzyl group. The direction of substitution is further augmented by the electron-donating effects of the amino substituent towards C-2. The bromonium ion undergoes nucleophilic attack by C-2 of **13** (**b**), and is substituted at this position, with the loss of a proton (**d**).
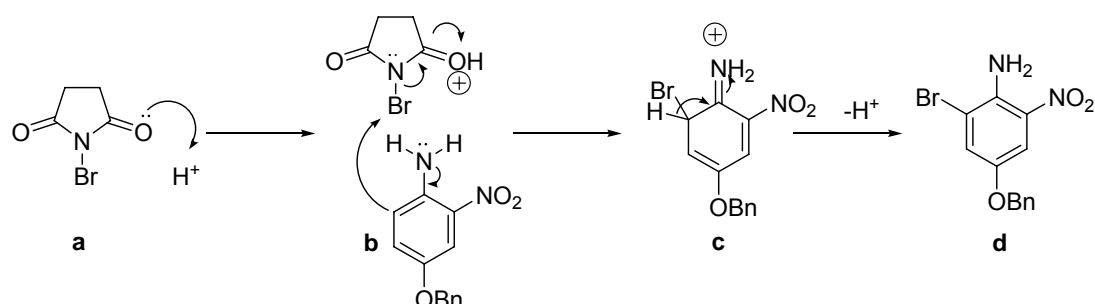
Scheme 38 Bromination reaction mechanism

Regioselective bromination to form **14** was confirmed by NMR and mass spectrometry. The ^1H NMR signal for C-2 was now absent due to the

presence of the bromo substituent, and aromatic protons at C-3 and C-5 showed a downfield chemical shift due to the electronegative effect of the bromide. The signal at 7.19 ppm due to the proton at C-3 was shifted downfield from 7.15 ppm in the starting material **13**. This peak now appeared as a doublet (doublet of doublets in **13**) as there is only *meta* coupling with C-5 (J=2.9 Hz). The doublet at 7.73 ppm corresponded to C-5 (downfield from 7.67 ppm in **13**), with *meta* coupling to C-3 (J=3.0 Hz). Mass spectrometry confirmed the mass $C_{13}H_{10}O_3N_2Br$ (M)$^-$ 320.9880.

4.1.3. Ek allylation

Bromination was followed by Ek allylation. The Ek method has the distinct advantage that allylation is direct, replacing the amino group without the intermediate conversion of aryl amine to aryl halide step seen in conventional methods, and without the use of metal catalysts.[88] [89] This allylated structure would be the precursor to cyclisation (Scheme 34).

Direct allylation with t-butylnitrite and allyl bromide following methodology described by Ek[89] (Scheme 39) only resulted in return of the starting material at the end of the reaction. Varying the reaction conditions such as increasing the temperature, adding more t-butylnitrite to push the reaction, and even adding Pd as a catalyst did not have any effect.
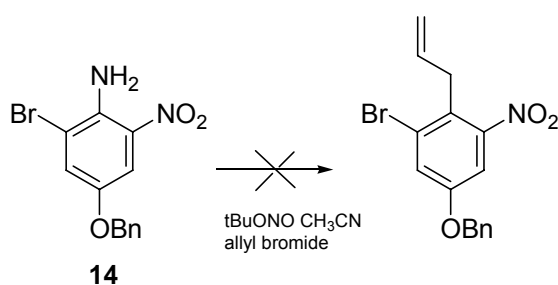
Scheme 39 Ek allylation of **14**

Direct allylation occurs via diazotisation, followed by free radical formation. *t*-Butyl nitrite is the source for generation of the reactive electrophile (Scheme 40). For successful formation of the product, this is attacked by the amino group to ultimately form the diazonium species. Dimerisation, followed by the loss of nitrogen generates the phenyl radical intermediate. This in turn is attacked by the allyl group, to ultimately form the allylated product. The free bromide radical could also continue to react with excess allyl bromide.[88] In this reaction, as starting material was being consumed during the reaction, one possible explanation may be that diazo formation occurs but that the aromatic substituents disfavour the subsequent allyl attachment.
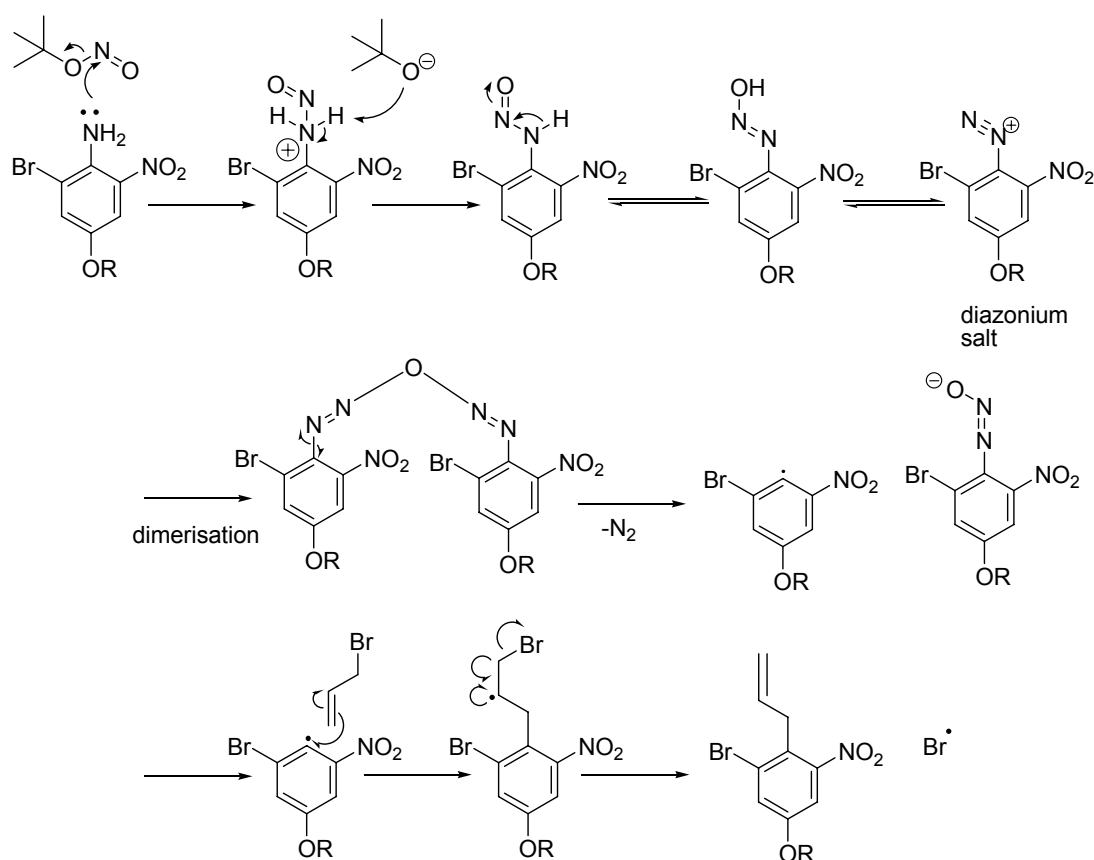
Scheme 40 Postulated direct allylation reaction mechanism. a. formation of the reactive electrophile from t-butyl nitrite. b. diazotisation followed by phenyl radical intermediate formation to give the allylated product (R=any protecting group e.g benzyl)

4.1.4. Displacement of C-4 amino group by iodo substituent

As direct allylation had not provided the desired product, a conventional two step route was investigated. First, displacement of the amino group by a halide (I) was carried out using $NaNO_2$ and KI (Scheme 41), followed by allylation.
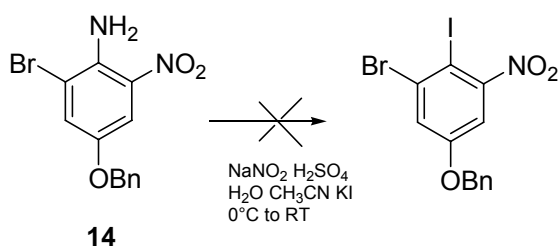
Scheme 41 Displacement of C-4 amino group of **14** by iodo substituent

These conditions did not yield amino group displacement – the presence of starting material at the end of the reaction suggested that diazotisation was not occurring. The –oxy *para* substituent may retard the reaction to form diazonium.

4.1.5. An alternative phenol protecting group: mesylation

The first route to Boc-duocarmycin SA described in Scheme 34 uses a benzyl group to protect the phenol. Halogenation resulted in almost 50% reduction in material in the second step of this route, which was not seen as an efficient start to the synthesis. Thus, we sought to change the protecting group to study conjugative effects within the aromatic ring on allylation. Scheme 42 shows phenol group protection using a mesyl group.
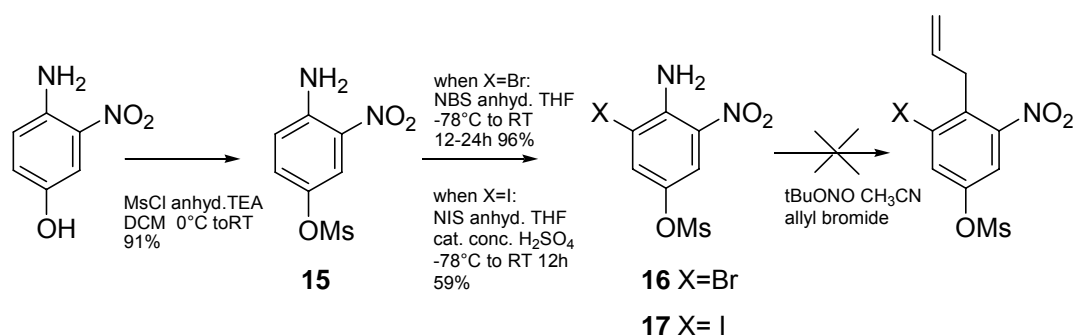
Scheme 42 Initial steps to Boc-DSA via Ek allylation, with mesyl group protection of phenol

4.1.6. Mesylation

The phenol group of the starting material 4-amino-3-nitrophenol was protected by a mesyl group, using mesyl chloride to form **15** in 91% yield (Scheme 43).
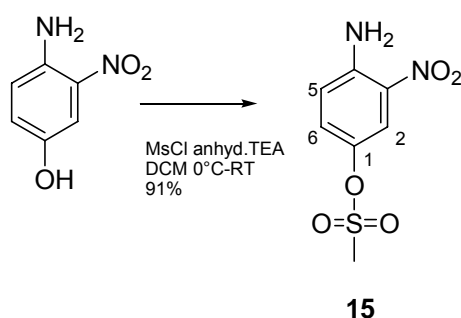
Scheme 43 Mesylation of phenol to form **15**

The mechanism for mesylation may occur as shown in Scheme 44. TEA extracts a proton from mesyl chloride (**a**). Electron rearrangement results in the elimination of chloride (**b**), to form a sulfene.[83] Nucleophilic attack by the phenol then occurs at the relatively electron deficient sulphur atom of the sulfene (**c**). Finally, the phenolate ion loses a proton to form the mesylated product (**15** shown as **d**).
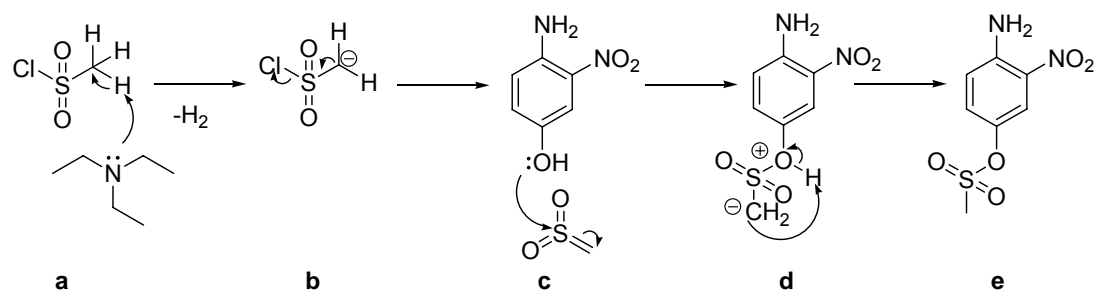
Scheme 44 Mesylation reaction mechanism

Mesylation to form **15** was confirmed by ^{1}H NMR spectroscopy. A singlet at 3.19 ppm was due to the presence of the mesyl group (3 protons). A broad singlet at 6.16 ppm was due to the two protons of the C-4 amino group. A doublet at 6.86 ppm corresponded to the proton at C-5, *ortho* coupling to C-6 (J=8.4 Hz). The doublet of doublets signal at 7.37 ppm corresponded to the proton at C-6. This splitting pattern was due to *ortho* coupling with C-5 (J=9.1 Hz) and *meta* coupling with C-2 (J=2.8 Hz). The doublet at 8.05 ppm was due to C-2 *meta* coupling with C-6 (J=2.8 Hz). This proton was highly deshielded because of the electron-withdrawing effect of the C-3 nitro group. Mass spectrometry confirmed the structure $C_7H_7O_5N_2S$ (M)$^-$ 231.0082.

4.1.7. Iodination and bromination of mesyl protected compound

Two candidate halides (I and Br) were investigated for halogenation of the C-3 position, to determine if the nature of the halogen would affect subsequent allylation. Electron-withdrawing substituents confer higher reactivity due to better conversion to the phenyl radical for allylation.[88] First, the C-3 position was iodinated using NIS and catalytic concentrated H_2SO_4 to form **16** in a yield of 59% (Scheme 45).
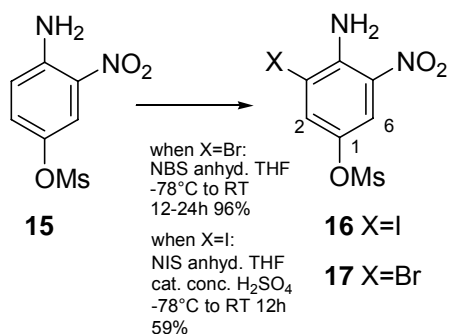
Scheme 45 Iodination & bromination of mesyloxy protected compound to form **16** & **17**

^{1}H NMR spectroscopy confirmed iodination to form **16**. No peak appeared for C-3 due to substitution, with both aromatic protons showing a downfield chemical shift due to the presence of the electron-withdrawing iodo substituent. The signal for C-2 at 7.89 ppm was now only a doublet due to *meta* coupling with C-6 (J= 2.8 Hz). This signal shifted downfield from 7.37 ppm in the starting material **15**. The proton at C-6 now appeared at 8.10 ppm, shifted downfield from 8.05 ppm in **15**. C-6 was a doublet due to *meta* coupling with C-2 (J=2.8 Hz).

The C-3 position was also brominated using NBS following the same methodology used in the benzyl protected synthesis (Scheme 37) to form **17** in an excellent yield of 96% (Scheme 45). This yield was greatly improved in comparison with 45% achieved with benzyl-protected phenol **14**. The combination of the two electron-withdrawing groups (nitro and mesyloxy) may deactivate all other positions on the aromatic ring, other than the electron rich position *ortho* to the amino group All ^{1}H NMR signals presented a downfield chemical shift due to the presence of the bromo substituent at C-3. The signal for the C-2 proton appeared at 7.76 ppm (previously at 7.37 ppm in **15**). The peak was now only a doublet, instead of doublet of doublets as there was now only *meta* coupling (J values between 2-3 Hz, as observed) with C-6. The signal for the C-6 proton now appeared

at 8.11 ppm, deshielded from 8.05 ppm in **15**. Mass spectrometry confirmed the mass $C_7H_6O_5N_2BrS$ (M)$^-$ 308.9190.

However, neither halide substituent (Br or I) led to successful direct allylation, using t-butylnitrite and allyl bromide (Scheme 42). The presence of only starting material by the end of both reactions indicates that the first step, diazotisation, may not be occurring.

Optimisation of the allylation step could have been investigated by the use of a model test substrate to establish and differentiate problems arising from reagent/technique employed, and substituent electronic effects. A net electron-withdrawing effect from the *ortho* nitro group may disfavour coupling. This seems to be the main difference in structure between **15**, and the compound previously successfully allylated via Ek methodology (Scheme 33).

4.1.8. Investigations into allylation via palladium catalysed couplings

Indirect strategies, such as conventional metal catalysed routes, were investigated as an alternative method for allylation. Conventional allylation usually employs halide substitution followed by allyl coupling. First, the amino group of **16** was substituted by an iodo substituent via diazotisation to form **18** (Scheme 46). A discussion of the subsequent palladium (Pd) catalysed couplings for attachment of the allyl group follows.
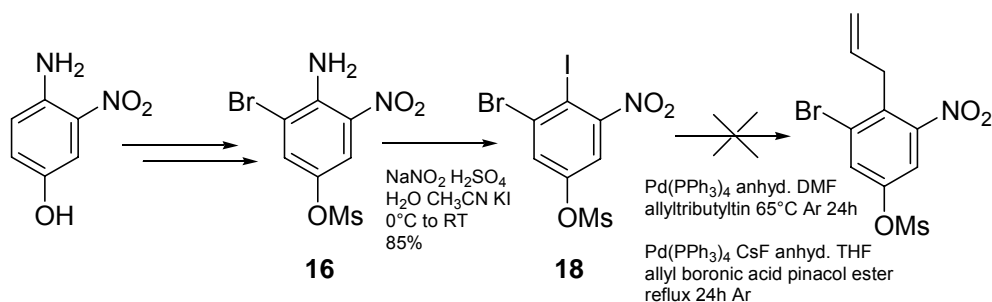
Scheme 46 Palladium coupling conditions investigated for allylation of **18**. Steps from 4-amino-3-nitrophenol to **16** are shown in Scheme 42.

4.1.8.1. Iodination via diazotisation

The amino group of **16** was displaced by an iodo group using NaNO$_2$ and KI, to form **18** in 85% yield (Scheme 47).
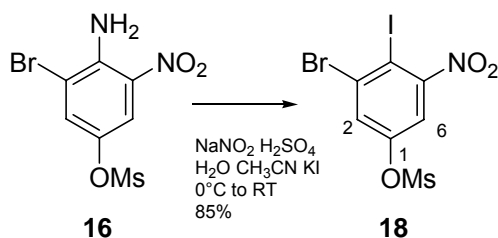
Scheme 47 Iodination via diazotisation to form **18**

The reaction mechanism is outlined in Scheme 48. The reactive electrophile (nitronium ion) is generated from NaNO$_2$ in the presence of H$_2$SO$_4$ (**a**). This electrophile undergoes attack by the electron lone pair of the amino substituent, to ultimately form the diazonium salt (**b**). N$_2$ is lost in a Sandmeyer-type S$_N$1 nucleophilic aromatic substitution reaction, with attack by KI at the electron deficient carbon (**c**). It is interesting that replacement of the electron-donating benzyl ether in **14** with the electron-withdrawing mesyl group in **16** activates this system. Whilst no reaction in the former case is observed, an excellent yield is achieved in the latter. There is clearly a fine balance between reactivity and stability in this reaction.

^{1}H NMR spectroscopy confirmed iodo displacement to form **18** as the broad C-4 amino singlet was now absent. Aromatic protons showed relative upfield chemical shifts, which may have been due to the iodo atom's relative larger size. The C-2 proton at 7.74 ppm was upfield from 7.76 ppm in the starting material **16**. The signal was a doublet due to *meta* coupling with C-6 (J=2.6 Hz). The C-6 proton appeared as a doublet due to *meta* coupling with C-2 at 7.45 ppm with an upfield shift from 8.11 ppm in **16**. Mass spectrometry confirmed $C_7H_4O_5NBrIS$ (M)$^-$ 419.8045.1.
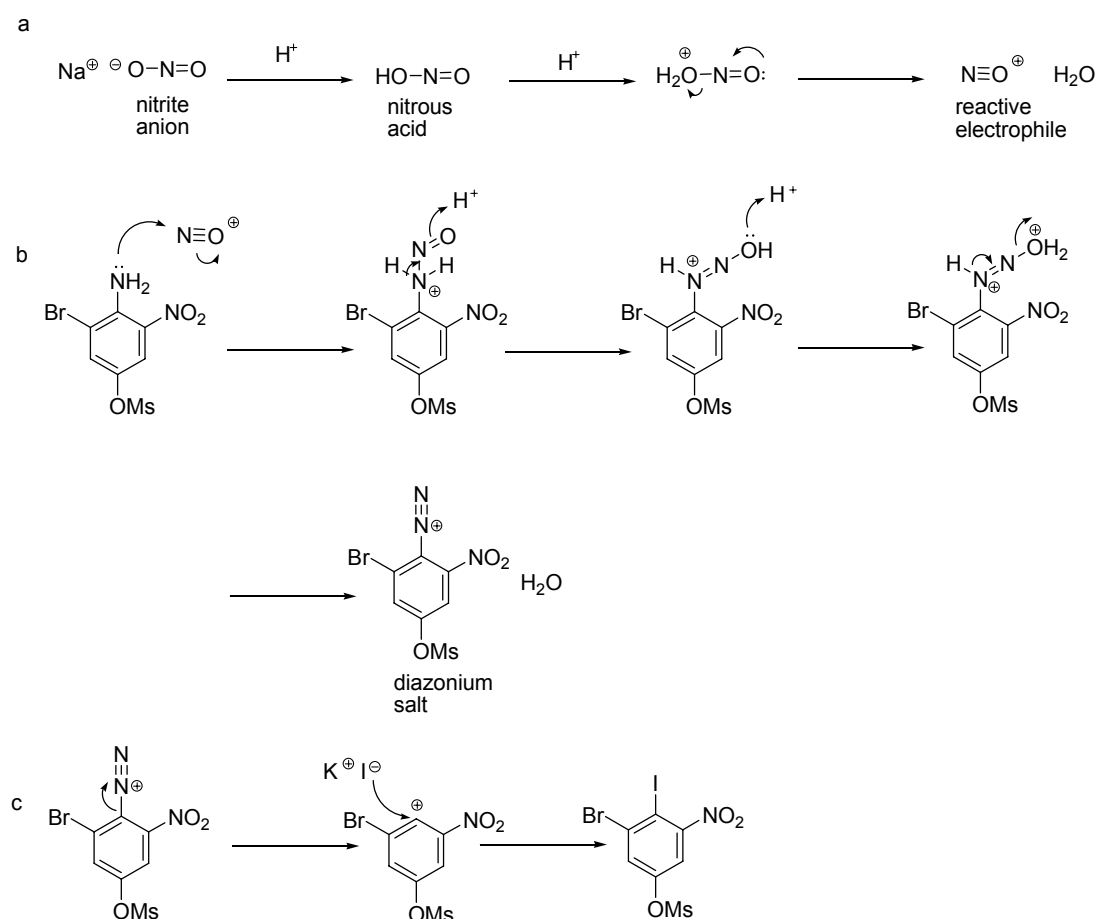
Scheme 48 Iodination reaction mechanism to form **18**; a. reactive electrophile formation from NaNO$_2$ b. diazonium salt formation c. Sandmeyer-type reaction to form aryl halide

4.1.8.2. Allylation of 18 via palladium catalysed couplings using Stille and Suuzuki conditions

Various Pd catalysed couplings were investigated for allylation. A discussion of these follows. Stille and Suzuki utilise tin and boron respectively for Pd catalysed coupling reactions.[90] [91] Couplings using these methods resulted in either no reaction or decomposition (Scheme 49, Table 2). One reason for this may be that these Pd-catalysed couplings require strictly anhydrous conditions, using dry glassware and maintaining an inert atmosphere. It was very difficult to maintain a fully intact reaction system throughout the course of the reaction. Build up of pressure interfered with the sealed system despite the use of a bubbler/balloon set-up; this may have affected the outcome. The presence of a *ortho* halide substituent, the bulky Br may have caused steric hindrance.
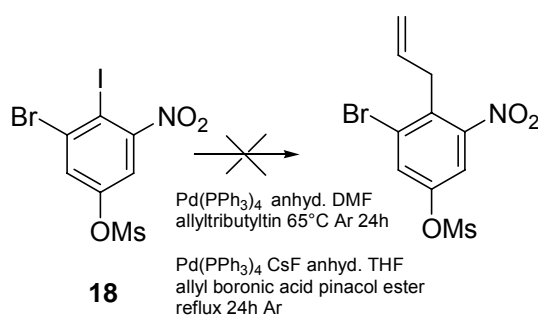
Scheme 49 Pd catalysed coupling conditions investigated for allylation of **18**

Coupling type	Catalyst	Allyl source	Conditions	Observations
Stille	Pd(PPh$_3$)$_4$ (0.20)	Allyltributyltin (1.18)	Under Ar, anhyd. DMF, 65°C	no reaction even after 72 h
Suzuki	Pd(PPh$_3$)$_4$ (0.10)	allyl boronic acid pinacol ester (3.60)	Under Ar, CsF, THF 30mins RT then reflux 24 h	decomposition

Table 2 Stille and Suzuki conditions investigated for allylation coupling reactions. Molar ratio to starting material in brackets.

4.1.8.3. Allylation of 14 via palladium catalysed coupling using Heck conditions

As both direct allylation via Ek methodology and indirect allylation methods had not been successful, we explored the use of Heck conditions to introduce the indole first, instead of the six membered heterocycle which ultimately spirocyclises. Attachment of the allyl functionality at C-3 was investigated as the precursor step for this purpose. Heck coupling is the reaction of aryl, benzyl and styryl halides with olefins at high temperature, in the presence of hindered amine bases and catalytic *in situ* generated Pd[(0)].

Yamada and co-workers reported Heck coupling with a carboxybenzyl, Cbz-protected amine substrate.[67] This was synthesised from alanine via a minimum of 5 steps.[92] We used a commercially available acetate, methyl 2-acetamidoacrylate,[93] to test the viability of this reaction as shown in Scheme 50.
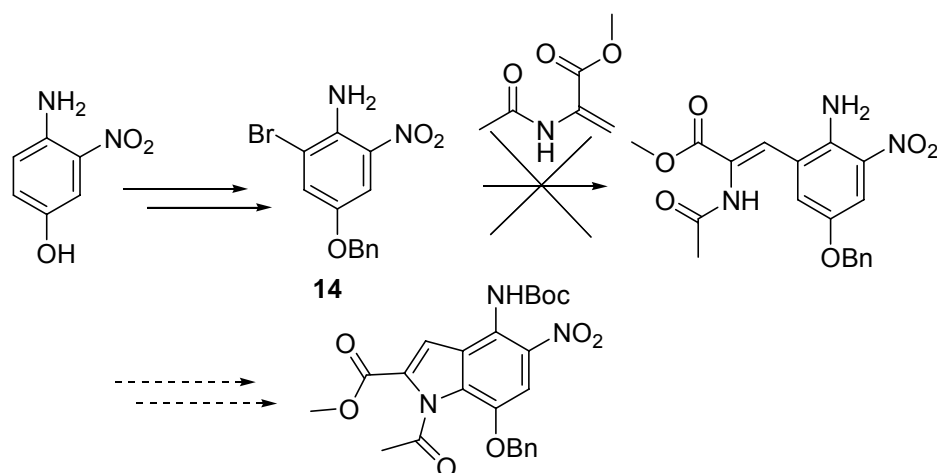
Scheme 50 Overview of synthetic route employing Heck coupling of **14** for indole cyclisation, leading to DSA

This reaction utilised Pd(OAc)[2], tri-*o*-tolyl phosphine, anhydrous TEA and CH[3]CN heated at 90°C (Table 3).[67] Here the Pd complex forms *in situ*, which

aims to improve the catalyst's stability, avoiding problems of degradation. However, TLC showed that starting material remained with decomposition evident after 12 hours.
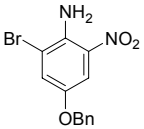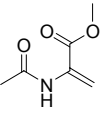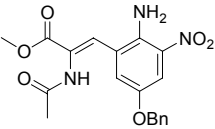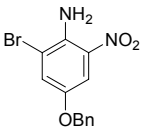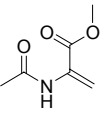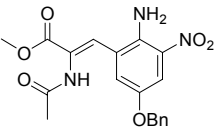
Aryl halide 14	Alkene	Conditions	Product	Yield
[structure: Br, NH_2, NO_2, OBn substituted benzene]	[structure: methyl 2-acetamidoacrylate]	1 mol% Pd(OAc)$_2$ 2.2 mol% P(o-tol)$_3$ 2 mol TEA under Ar, anhyd. CH$_3$CN, 90°C	[structure: Heck product]	Starting material only, with decomposition after 12 h
[structure: Br, NH_2, NO_2, OBn substituted benzene]	[structure: methyl 2-acetamidoacrylate]	10 mol% Pd(OAc)$_2$ 1.4 mol K$_3$PO$_4$ under Ar, DMA, 140°C	[structure: Heck product]	Starting material only, with decomposition after 72 h

Table 3 Heck coupling of **14** with and without ligand (alkene = methyl 2-acetamidoacrylate).

Many factors may have governed the outcome of this reaction. An examination of the reaction mechanism of the Heck catalytic cycle may highlight some of these factors (Figure 19). First, Pd$^{(II)}$ is reduced to Pd$^{(0)}$ by the ligand. Ratios of up to four equivalents of ligand to Pd catalyst may be necessary; we only used a 2:1 ratio which may have produced sub-optimal complex formation. After generation of Pd$^{(0)}$, oxidative addition is the rate determining step, where Pd inserts into the C-X bond. Next, the alkene inserts into the Pd-carbon bond during migratory insertion. In this reaction, the alkene, methyl 2-acetamidoacrylate, is overall slightly electron deficient due to the carbonyl groups. This is optimal as Heck coupling favours electron poor alkenes. Carbon-carbon bond rotation results in *cis* conformation to relieve strain. Elimination of the product is by β hydride

elimination. Base then regenerates the Pd catalyst to complete the cycle. As starting material was present at the end of the reaction, it is likely that problems occur at the beginning of the catalytic cycle. The Pd catalyst may not be able to insert into the C-X bond, or the aryl halide substituents hinder insertion of Pd$^{(0)}$ into the C-X bond. And so, the cycle cannot continue.
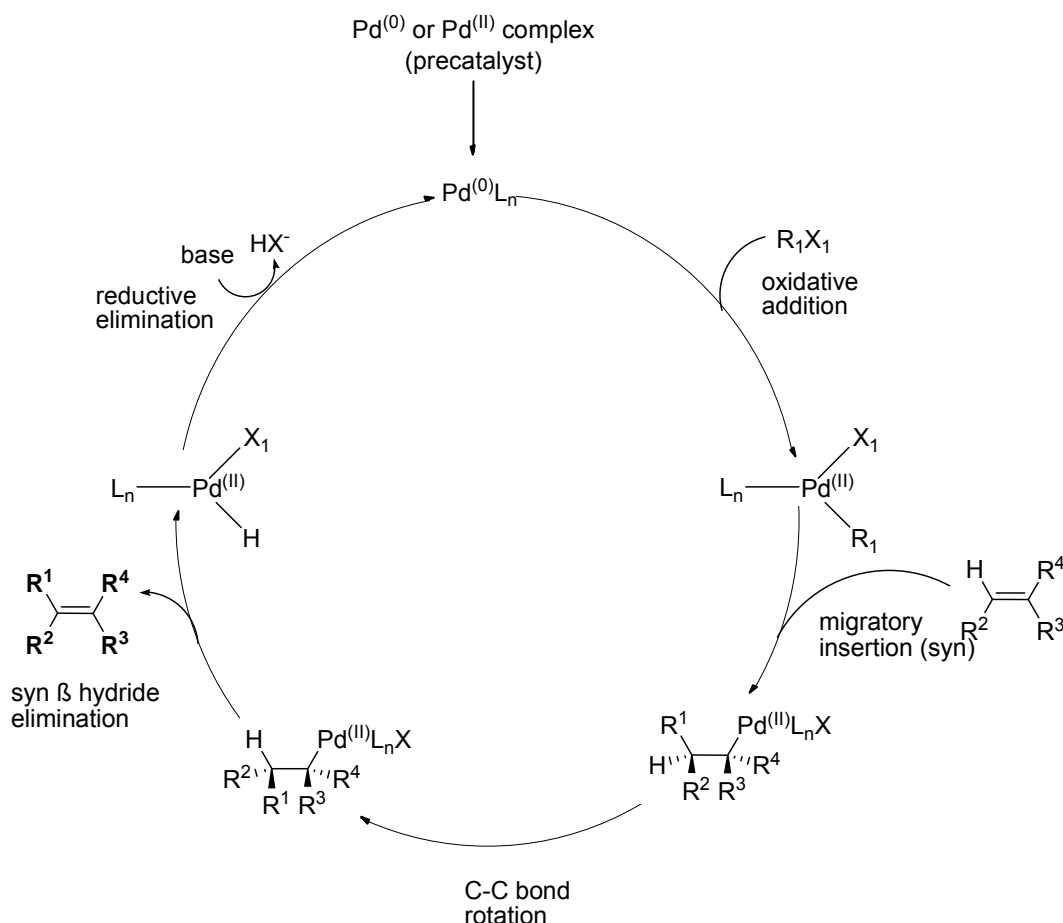
Figure 19 Classical Heck catalytic cycle[94]

In addition, phenyl ligands can become cleaved from the Pd catalyst at high temperatures (usually above 100°C), resulting in catalyst decomposition and coupling of the phenyl ligand to the alkene substrate. This effect is minimised with the use of P(o-tol)$_3$ as its steric bulk stabilises the Pd complex,[95] and subsequent formation of a highly thermally stable dimeric

palladacycle which is thought to be of Pd$^{(II)}$ character. However, the cleavage of phenyl ligands can not be completely prevented.

4.1.9. Ligand-free catalysis

Ligand-free catalysis was also investigated as a novel and cost-effective method using conditions described by Yao; refluxing Pd(OAc)$_2$ and K$_3$PO$_4$ in DMA (Table 3).[96] Yao proposes the formation of a transient palladacycle on reaction of Pd(OAc)$_2$ with the alkene as the first step. This palladacycle then undergoes oxidative addition by the aryl halide, generating Pd$^{(IV)}$ species. This converts to a Pd$^{(II)}$ species via a base-promoted intermediate releasing the product.[96] However, NMR indicated that only starting material remained by the end of the reaction. Again, a likely explanation is that an active Pd species may not be formed, preventing the catalytic cycle from continuing.

4.1.10. Microwave assisted coupling reactions

The use of microwave assisted organic synthesis has gained in popularity within the last decade. The advantages are short reaction times (seconds/minutes rather than hours/days), possible reduction in side products, increased yields and improved reproducibility. Disadvantages can be a limit on scale due to a maximum vessel size, potentially safety problems arising from vessel over-heating, and high equipment costs. Microwaves operate at an electromagnetic irradiation frequency of 2.45 GHz (corresponding to a wavelength of 12.24 cm). Reactions take place by solvents or reagents absorbing microwave energy and converting it to heat, which is more efficient and enables better temperature control. In contrast, conventional heating methods, typically oil baths, rely on the thermal

conductivity of reaction vessels, resulting in the reaction mixture temperature being lower than that of the vessel itself.[97] [98]

The use of the microwave for coupling reactions is well established for Heck, Stille, Sonogashira conditions, and Suzuki couplings.[97] [99] Thus, the microwave would provide a fast and simple way to study coupling conditions. Heck-type couplings were carried out in the microwave using $Pd(OAc)_2$, $P(o\text{-tol})_3$ and various bases; TEA,[100] K_3PO_4[96] and tributylamine,[101] as well as without the use of ligands (Table 4).

Aryl halide 14	Alkene	Conditions	Product	Yield
	1 mol	120°C, 15min, high abs → 1 mol% Pd(OAc)$_2$ 22 mol% P(o-tol)$_3$ 1 mol TEA CH$_3$CN		Incomplete reaction with decomposition
	1 mol	180°C, 35min high abs → 1 mol% Pd(OAc)$_2$ TEA CH$_3$CN		Starting material only
	1.25 mol	180°C, 35min high abs → 1 mol% Pd(OAc)$_2$ tributylamine DMF		Starting material only
	1.20 mol	140°C, 15min high abs → 1 mol% Pd(OAc)$_2$ 1.40 mol K$_3$PO$_4$ DMA		Starting material & decompostion

Table 4 Heck coupling in the microwave of **14** with and without ligand (alkene = methyl 2-acetamidoacrylate).

These conditions did not provide the desired allylated product. It may be that in this particular case of Heck coupling, a reactive Pd species cannot form in the absence of a ligand, so the catalytic cycle cannot continue. Also, catalyst/ligand ratios may not have been optimised, as discussed earlier. The electron-withdrawing electronic environment of the aryl halide, i.e the *ortho* C-5 nitro group, may be hindering the first step of Heck coupling, oxidative addition.

4.2. Route to DSA utilising Fischer indole cyclisation

Fischer cyclisation, the heating of ketone or aldehyde arylhydrazones in the presence of acid is a well known route to form indoles.[102] We investigated the potential of Fischer cyclisation as a facile step to indole formation, using Tietze's work on the synthesis of *seco-* duocarmycin SA.[62] Scheme 51 shows our synthetic route utilising Fischer methodology to form the first heterocycle; diazotisation to the hydrazone as the precursor structure to cyclisation, followed by indole formation.[62] A discussion of this follows.
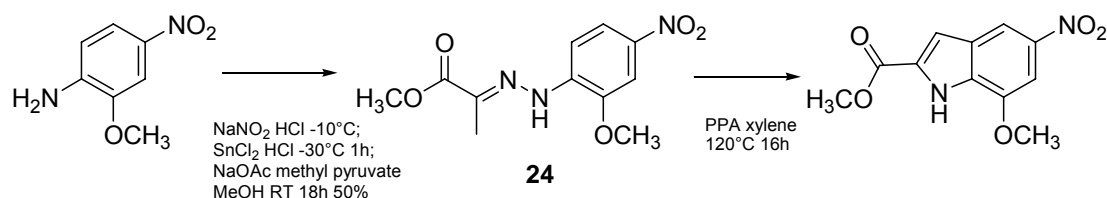
Scheme 51 Synthetic route utilising Fischer indole cyclisation

4.2.1. Diazotisation to hydrazone

To introduce the cyclising moiety for Fischer indole synthesis, commercially available 2-methoxy-4-nitroaniline was diazotised to the corresponding hydrazone[62] in 50% yield (Scheme 52). This yield was acceptable as it was maintained on scale up.
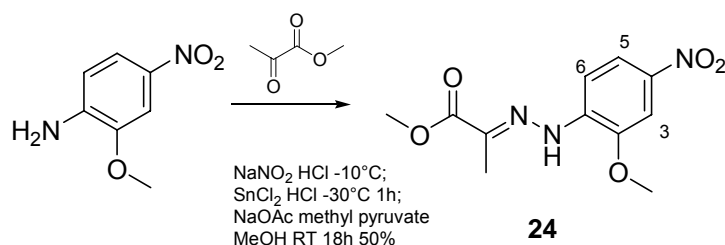
Scheme 52 Diazotisation of 2-methoxy-4-nitroaniline to form hydrazone 24

Scheme 53 shows the reaction mechanism for hydrazone formation. The reactive electrophilic nitrosonium ion is generated from $NaNO_2$ in the presence of HCl (Scheme 53a). This electrophile undergoes attack by the electron lone pair of the aromatic amino substituent, to form the diazonium ion. Tin(II) chloride in the presence of aqueous HCl forms a hydrate which reduces the diazonium ion (Scheme 53b). The methyl pyruvate keto group undergoes nucleophilic attack by the electron lone pairs of N-2 of the hydrazine (Scheme 53c). Protonation of the hydroxyl group in the acidic conditions results in the elimination of water, forming an imine type bond ie the hydrazone (Scheme 53c).
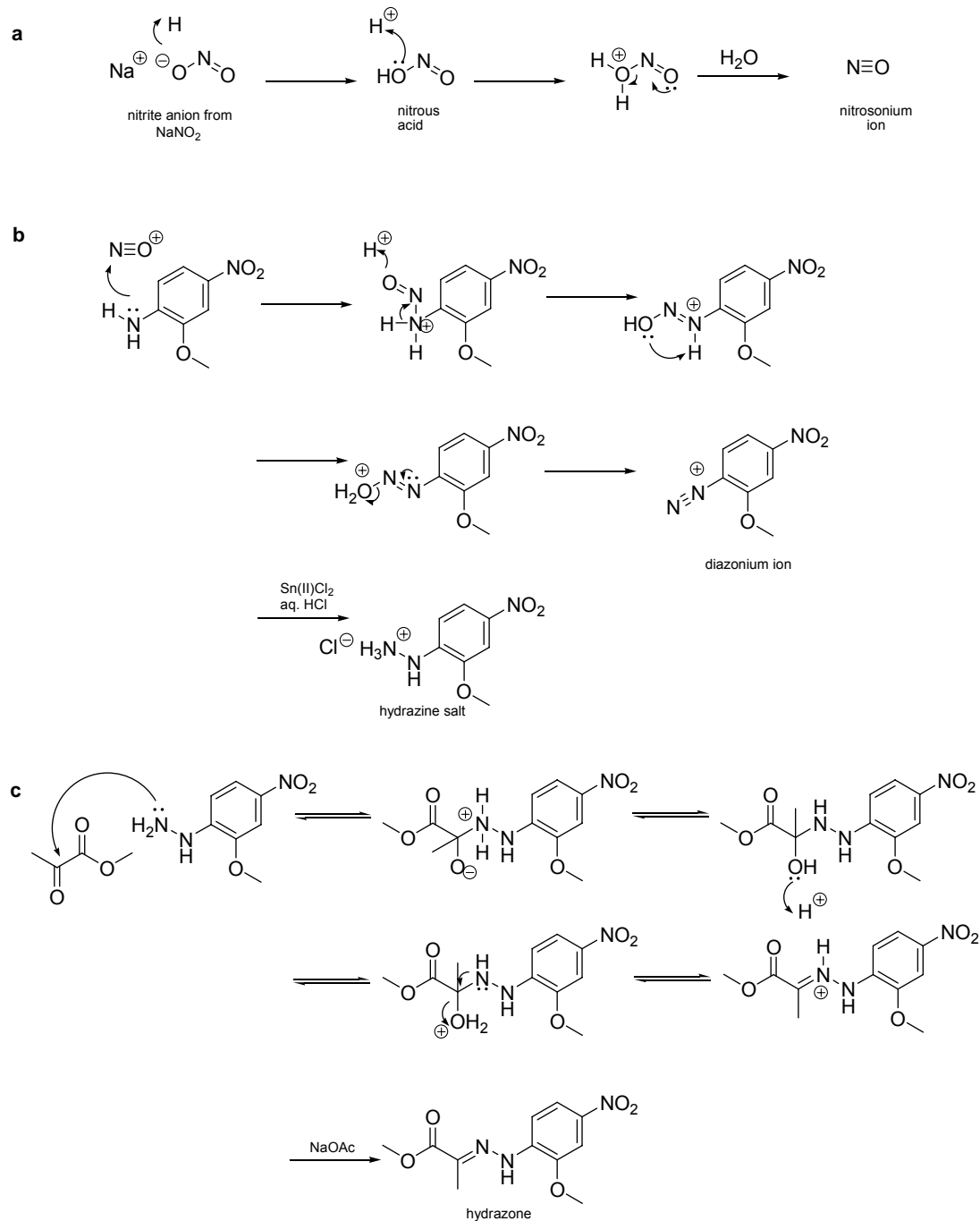
Scheme 53 Reaction mechanism for hydrazone formation 24 as the precursor step for Fischer indole cyclisation a. formation of the reactive nitrosonium ion from NaNO₂ b. reaction of amine substrate with nitrosonium ion via diazotisation to form hydrazine c. reaction of hydrazine with methyl pyruvate to form hydrazone

The structure of the hydrazone product **24** was confirmed by ^{1}H NMR and mass spectrometry and matched the literature compound.[62] The ^{1}H NMR singlet at 2.15ppm (3H) was due to CH₃. The two other methyl groups

(OCH$_3$) gave singlets at 3.86 and 3.98 ppm (3H each). However, it was difficult to assign which peak belonged to which OCH$_3$ in **24**. The aromatic OCH$_3$ may experience deshielding due to the ring current effect of the aromatic ring, and the carbonyl group may deshield the other OCH$_3$ group. The proton at C-6 gave a doublet at 7.59 ppm (1H, J=8.8 Hz) due to *ortho* coupling with C-5. The proton at C-3 appeared as a doublet at 7.73 ppm (1H, J=2.4 Hz) due to *meta* coupling with C-5. The proton at C-5 appeared as a doublet of doublets at 7.92 (1H, J=8.8, 2.4 Hz). This peak splitting pattern and J values confirmed that the C-5 proton is coupling with C-3 and C-6. The broad singlet peak at 8.30 ppm was due to NH. Mass spectrometry confirmed the mass $C_{11}H_{12}O_5N_3$ (M)$^-$ 266.0783. IR spectroscopy absorption at 3327 cm^{-1} was due to NH, 1713 cm^{-1} was due to carbonyl stretching although this could also have been due to the C=N bond, 1575 and 1338 cm^{-1} were due to aromatic NO_2, the split peak at 1092 cm^{-1} was assigned to C-O-C and 1023 cm^{-1} assigned to N-N. ^{13}C signals of note were the CH$_3$ at 10.45 ppm, and two OCH$_3$ groups at 51.82 and 56.18 ppm. The remainder of the peaks arose due to aromatic carbons (6 peaks in the 106-165 ppm range). The peak at 147.78 ppm was assigned as C=N, whilst 162.05 ppm was due to the carbonyl.

4.2.2. Fischer indole cyclisation

The hydrazone **24** underwent subsequent Fischer indole cyclisation by heating at 120° in PPA and xylene to provide the indole in 73% yield (Scheme 54).
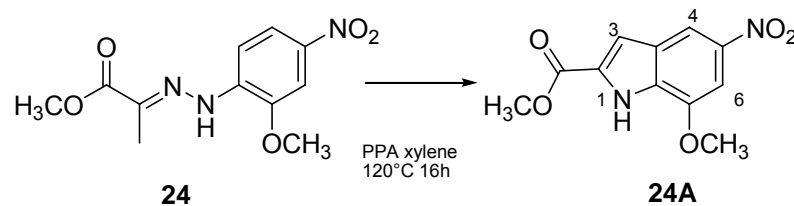
Scheme 54 Fischer indole cyclisation of **24** to form **24A**

The Fischer indole synthesis is composed of distinct steps as shown in Scheme 55.[102] First, the Lewis acid (ie the proton) coordinates to the imine nitrogen, with the hydrazone tautomerising to the corresponding hydrazine (**a**-**b**). [3,3]-sigmatropic rearrangement disrupts the aromatic ring to form **c**. Deprotonation causes re-aromatisation, followed by 5-*exo-trig* cyclisation (**d**-**e**) resulting in the formation of the indoline ring. Finally, the loss of ammonia provides the indole ring (**g**).

^{1}H NMR showed evidence of cyclisation and the presence of unidentifiable impurities. A hydrazone CH_3 peak was now absent. The indole peak was evident from a singlet at 7.27 ppm due to the proton at C-3. The aromatic protons C-4 and C-6 were not clearly distinguishable due to unidentifiable peaks presumably due to impurities. The proton at C-6 showed slight shielding; a doublet at 7.54 ppm appeared upfield from 7.73 ppm in the starting material **24**. The proton at C-4 should appear as a doublet as it only couples with C-6. This signal is deshielded due to the neighbouring indole at 8.27 ppm, from 7.92 ppm in **24**. The NH broad singlet peak also showed deshielding at 9.29 ppm, from 8.30 ppm in **24**. Mass spectra confirmed the presence of the cyclised structure **24A**, $C_{11}H_9O_5N_2$ (M)$^-$ 249.0519.

Thus it was evident from NMR spectra that the 73% yield achieved was not an accurate reflection of the true amount of indole product present due to the presence of unidentifiable impurities after purification. In retrospect, a

further purification would have been helpful to establish the true yield and to establish if the impurities present were significant.
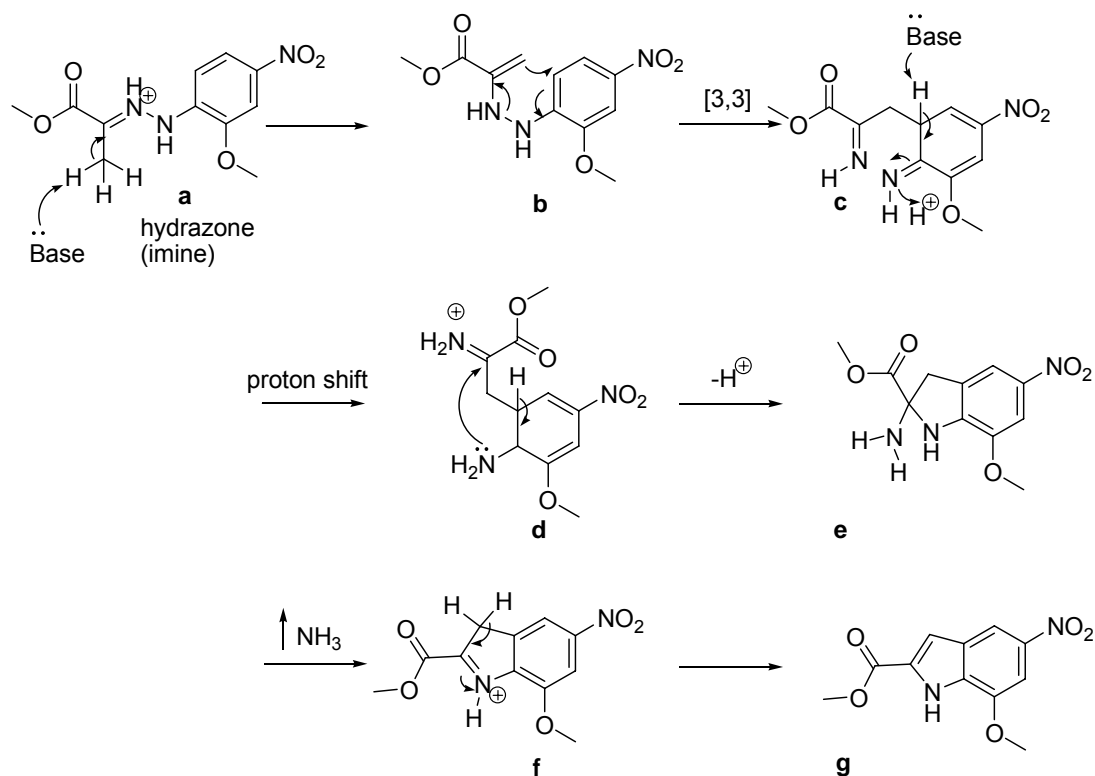
Scheme 55 Reaction mechanism for Fischer indole cyclisation from starting material **24** to form **24A**.[103]

Reaction efficiency could be improved by increasing the polyphosphoric acid (PPA) to starting material ratio (a 1.5:1 ratio was used in this reaction). Guy and co-workers have shown that a 4:1 ratio of PPA in xylene resulted in 78% yield when a structurally similar molecule was cyclised using Fischer methodology (Scheme 56).[104] Also, the use of xylene as a co-solvent has been shown to improve yields; a 9:1 PPA to starting material ratio without xylene was found to achieve only a 57% yield (Scheme 56).[105]
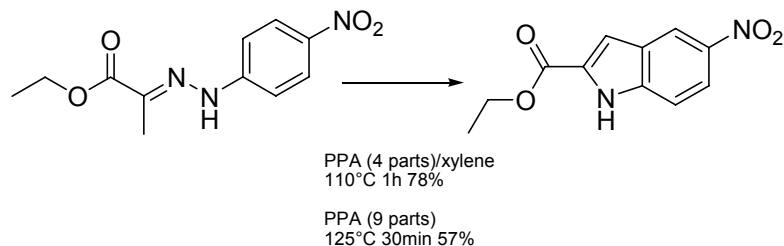
Scheme 56 Comparison of yields for Fischer indole synthesis of ethyl 2-(4-nitrophenylhydrazono)-propanoate at differing PPA ratios.[104] [105]

4.2.3. Protecting group manipulation - conversion of methyl ether to benzyl ester

Crude material from the Fischer indole cyclisation was carried forward as a test reaction. Tietze's conditions using AlCl₃ were used to cleave the methyl ether and convert to a benzyl group.[62] This would provide a protecting group which would be easier to manipulate later in the synthesis. Initial efforts at methyl ether hydrolysis did not result in any reaction, starting material only remained (Scheme 57). The indole may have been formed in insufficient quantities for methyl ether cleavage to occur due to the presence of impurities present in the previous step, or the impurities present may have interfered with this reaction. As the key Fischer indole cyclisation step would have required investigation to yield pure product before this step could be optimised, no further efforts were made on this step.
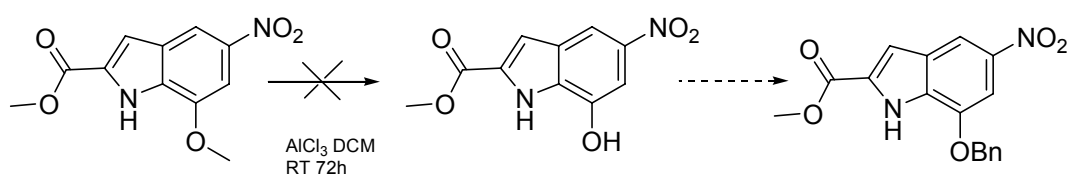
Scheme 57 Conversion of methyl ether to benzyl ester protecting group

4.3. Summary

In summary, the first synthetic route to DSA explored allylation methods as the precursor step to form the six membered heterocycle, as shown in the retrosynthetic analysis in Scheme 32. Allylation was investigated using direct methods (Ek methodology) and indirectly using Pd catalysed coupling reactions (Scheme 34). The use of the microwave for Pd catalysed couplings was also investigated (Table 4) to introduce the indole as the first heterocycle. Protecting groups were changed, specifically from benzyl to mesyl group phenol protection to determine the effects on allylation (Scheme 42). Our investigations indicated that allylation was not a feasible precursor step to DSA for our substrate, as discussed throughout this chapter.

The Fischer indole method was explored as it is a standard method for indole formation. The precursor hydrazone **24** was achieved in a facile step in yields of 50% which were consistently achievable. Cyclisation to indole was achieved using Fischer methodology (Scheme 54). However, the presence of variable amounts of inseparable impurities limited the use of this reaction for our purposes.

At this point, our attention turned to devising an alternative synthetic route to reach the core DSA alkylation subunit of duocarmycin SA.

5. Second route to the duocarmycin SA alkylation subunit, DSA

Our aim was to reach DSA, the core alkylation subunit of duocarmycin by a concise and efficient route. Our findings so far had shown that allylation as a precursor step strategy for indole formation was not feasible for our substrate (chapter 4). Thus we sought an alternative method to introduce the functionality that would lead to successful indole cyclisation. Key disconnections for our new route were essentially the stepwise formation of heterocycles; spirocyclisation of the cyclopropane, N-alkylation followed by *5-exo-trig* free radical ring closure to form the indoline, and left hand indole formation via Hemetsberger methodology (Scheme 58).
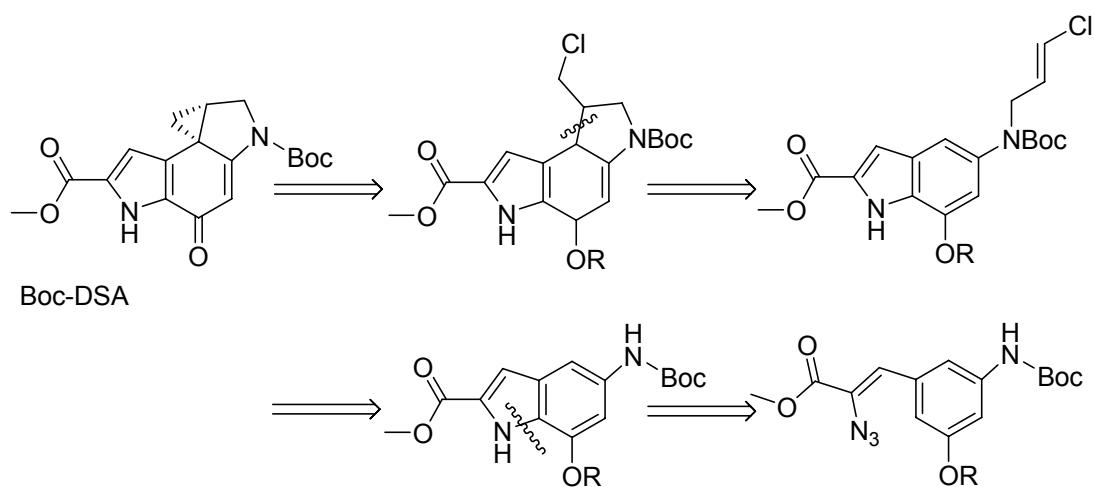
Scheme 58 Retrosynthetic analysis for second route to DSA utilising azide indole methodology

This route employed a different approach in that the left hand indole would be synthesised first using the Hemetsberger reaction as the key coupling precursor step. The Hemetsberger reaction is the condensation of a benzaldehyde with an azidoacetate, followed by pyrolysis to form an indole

(Scheme 59).[106] This methodology has been extensively used by Moody and co-workers in syntheses of many natural products, [107] [108] [109] [110] [111] most relevant of which is his work on CC-1065 analogues.[112]

Tichenor and co-workers detail a synthetic route using Hemetsberger indole cyclisation to reach duocarmycin SA.[65] An overview of this synthesis is described in chapter 2, introduction part 2. We thought this would be a feasible route to investigate, to quickly reach our target DSA structure. Different starting substrates were investigated to form the precursor aldehyde (Scheme 60 & Scheme 62). Scheme 59 shows the synthetic route which led to successful indole cyclisation. The aldehyde functionality was introduced using Swern oxidation, followed by one pot phenol formation/benzylation. Methyl azidoacetate was synthesised in-house, and condensed with the aldehyde to form the precursor to Hemetsberger indole cyclisation. A discussion of this synthesis follows.
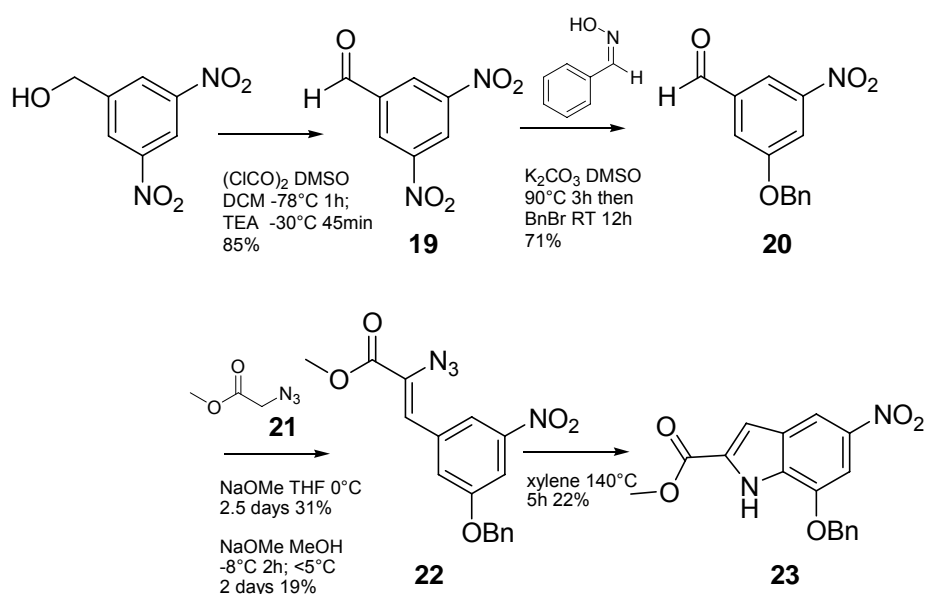
Scheme 59 Initial steps of the synthetic route to duocarmycin SA utilising the Hemetsberger azide route[65]

5.1.1. Aldehyde formation

An aldehyde functionality is required for condensation with an azide according to Hemetsberger indole cyclisation methodology. Three strategies, ortholithiation, reduction using DIBAL and Swern oxidation were investigated for this purpose. Swern oxidation was found to be an efficient and simple step to form the aldehyde. These strategies are discussed below.

Scheme 60 shows the use of the starting material 4-benzyloxy-2-bromo-6-nitro-phenylamine **14** to form the aldehyde. *Tert* butyl lithium (*t*BuLi) was utilised in this halogen-lithium exchange reaction with the bromo substituent, and conversion to aldehyde.[113]
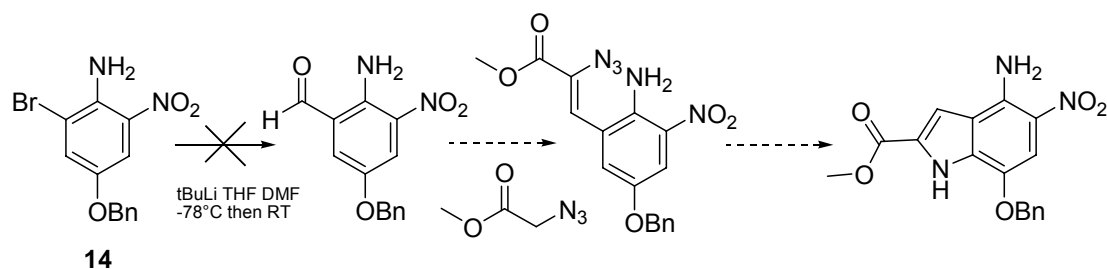
Scheme 60 Halogen-metal exchange to introduce the aldehyde substituent to **14** for Hemetsberger methodology

However, only starting material remained by the end of the reaction. Scheme 61 shows a possible mechanism. Usually nitrogen or oxygen based functional groups *ortho* to the halide can form a complex with the lithium (**a**). The lithium would then be "guided" to react with the bromo atom. The organolithium, *t*BuLi removes the halogen (bromide) from the aryl halide by nucleophilic attack, in a halogen-metal exchange (**a-b**). This forms a stable aryl lithium relative to the original highly basic *t*BuLi. Nucleophilic attack then occurs at the DMF carbonyl carbon, forming a tetrahedral intermediate (**b-c**). Protonation at the electron deficient oxygen would then ultimately lead to the aldehyde (**e**).

It may be that *t*BuLi was unable to form a complex with the *ortho* amino group for reasons that are not clear. Reactions are usually carried out at low temperatures to prevent lithiated structures attacking the starting material in a substitution reaction.[83] Tetrahydrofuran (THF) is a good solvent for lithiation as it remains liquid at low temperature. However, at higher temperatures it can form side products where lithium substitutes *ortho* to the THF cyclic ether, ultimately cleaving the heterocyclic ring.[83] In retrospect, the reaction may have been warmed to room temperature too quickly so that side reactions occurred. The most likely reason for this reaction not going to

completion is that the lithium base deprotonated the unprotected amine in preference to halogen-lithium exchange.
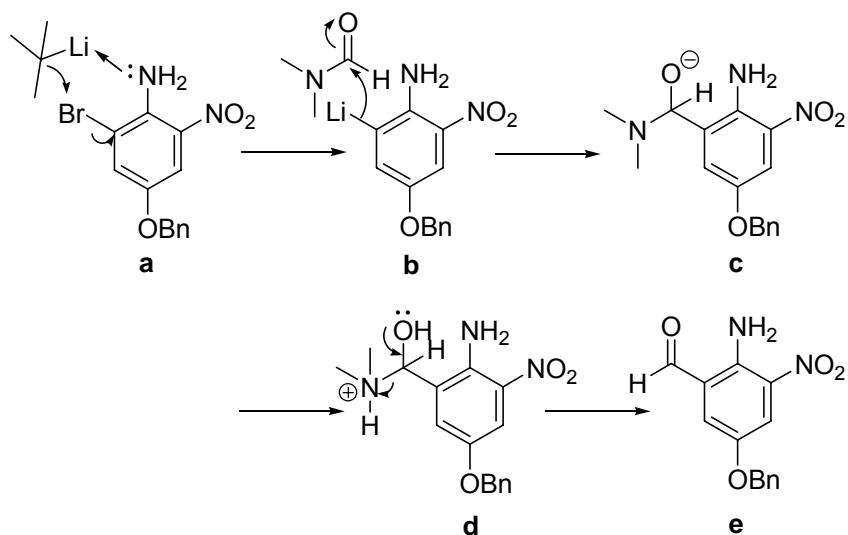
Scheme 61 Possible reaction mechanism for aldehyde formation via ortholithiation

As halogen-lithium exchange had not provided the desired aldehyde, reduction using DIBAL was investigated as an alternative method. The commercially available ester, methyl 3, 5-dinitrobenzoate was successfully reduced using DIBAL, which gave the desired product in 52% yield (Scheme 62).
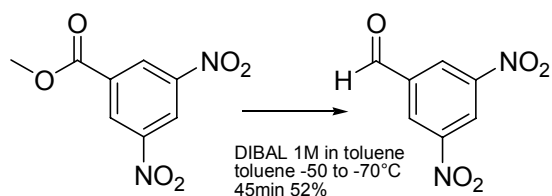
Scheme 62 Reduction of methyl 3,5-dinitrobenzoate using DIBAL to provide 3,5-dinitro benzaldehyde

The temperature had to be carefully controlled as excess DIBAL resulted in complete reduction to an alcohol. This was evident from an extremely polar

spot by TLC monitoring. On scale up, separation and purification from the starting material proved difficult. For these reasons, an alternative reaction to form the aldehyde was sought.

The use of Swern oxidation conditions for the preparation of the aldehyde was found to be a facile reaction. [65] [114] The aldehyde **19** was formed from commercially available 3,5-dinitrobenzyl alcohol in an excellent yield of 85% (Scheme 63).
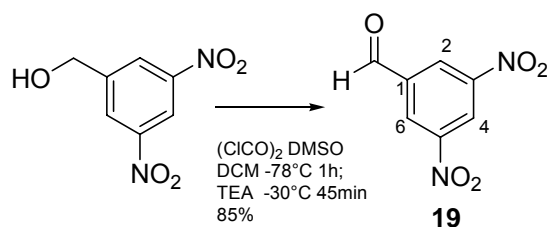
Scheme 63 Aldehyde formation via Swern oxidation of 3, 5-dinitro benzyl alcohol

The Swern oxidation proceeds through the following mechanism, as shown in Scheme 64. First, dimethyl sulphoxide (DMSO) is activated by oxalyl chloride. Oxalyl chloride undergoes nucleophilic attack at the carbonyl carbon by the sulfoxide ion (**a**), ultimately resulting in the release of the chloride group (**b**). This chloride ion attacks the electron deficient sulphur resulting in the release of the other chloride ion, carbon dioxide and carbon monoxide (**c**). This generates the reactive chlorosulfonium salt which then undergoes nucleophilic attack by the alcohol oxygen's electron lone pairs (**d**). Deprotonation of the carbon α to the sulfur generates the alkoxysulfonium ylide (**f**), ultimately rearranging to afford the aldehyde (**19** shown as **g**).

The formation of **19** was confirmed by the key aldehyde ^{1}H NMR singlet at 10.2 ppm. ^{13}C spectroscopy showed the carbonyl at 187.4 ppm. IR spectra clearly showed carbonyl absorbance at 1701 cm^{-1}, with nitro groups

absorbing at 1535 and 1342 cm^{-1}. Mass spectrometry confirmed the expected mass $C_7H_4O_5N_2$ (M+) 195.0039.
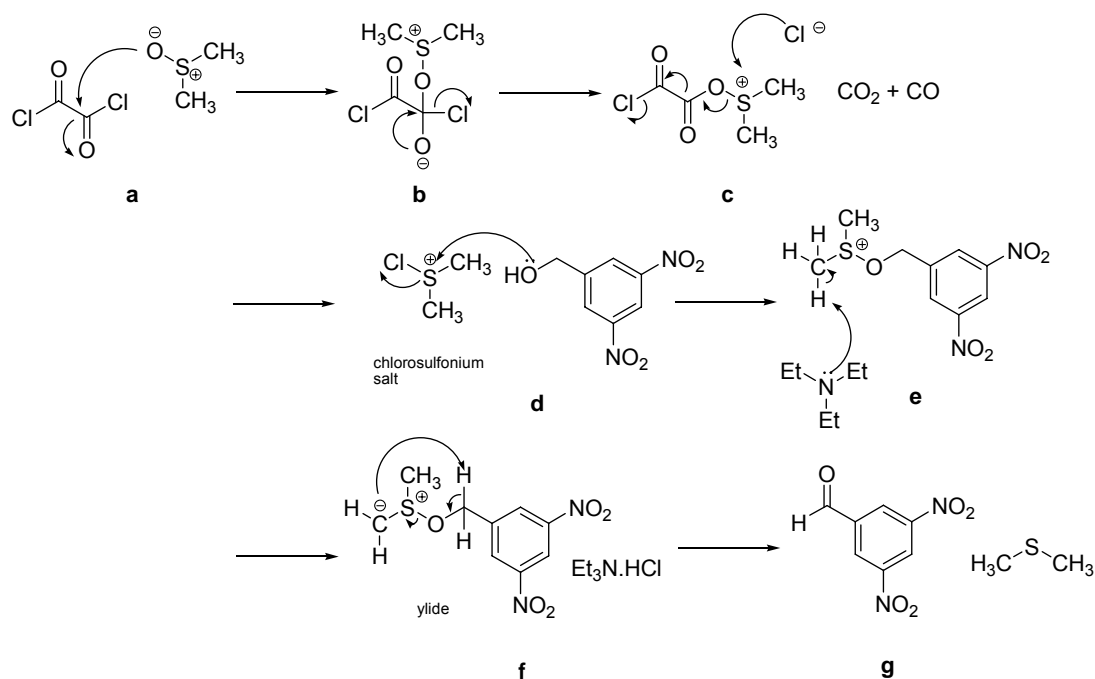
Scheme 64 Reaction mechanism for Swern oxidation to form **19**[83]

5.1.2. Phenol formation and benzyl protection

Introduction of the phenol functionality is important for biological activity of the duocarmycin SA subunit structure.   Phenol group protection by benzylation limits cytotoxicity until the end of the synthesis, and provides scope for functional group variation and ligand attachment at this position.

This functional group transformation was investigated as a two step process (Scheme 65);[65] nucleophilic displacement of one nitro group using benzaldehyde oxime formed the phenol in 30% yield after separation and purification.   The second step benzylation provided **20** using standard conditions (benzyl bromide, K$_2$CO$_3$) in 62% yield. This equates to 19% yield over two steps. Utilising a one pot method where oxime mediated nitro

group displacement and benzyl protection of the resulting phenol was carried out in one step resulted in a significantly improved 71% yield of the final product **20**.
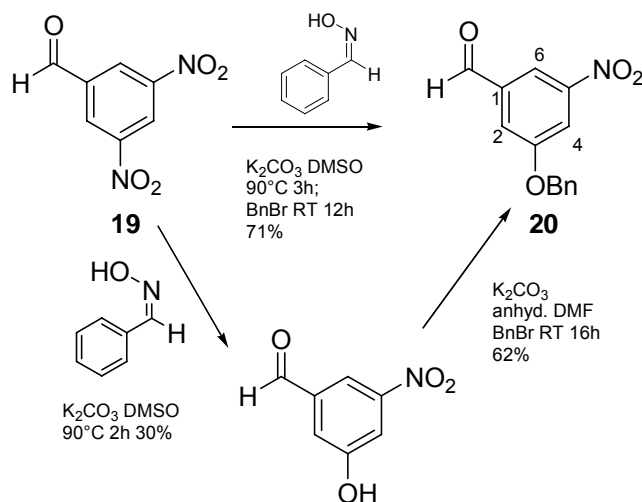
Scheme 65 Comparison of 2 step nitro group displacement to phenol then benzylation, versus one pot nitro phenol formation/benzylation to form **20**

The reaction mechanism for one pot benzylation may occur as shown in Scheme 66. First, the oxime becomes deprotonated in the presence of base (K$_2$CO$_3$). Nucleophilic attack by the negatively charged oxime oxygen results in displacement of the nitro group (**a**), followed by *in situ* elimination to form the phenoxide (**b**). Usual S$_N$2 type attack then occurs whereby the reactive benzyl bromide attaches to the phenolate ion (**d-e**).
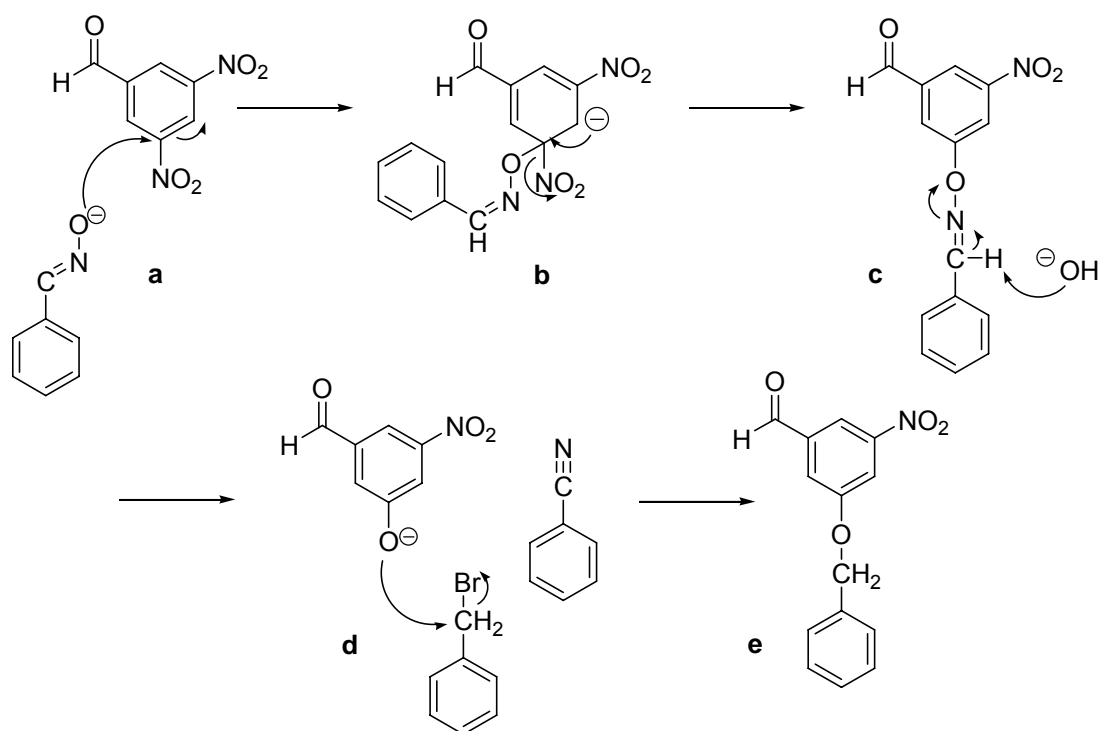
Scheme 66 One pot benzylation reaction mechanism to form **20**

The structure of **20** was confirmed by ^{1}H NMR and matches the literature reference; the CH$_2$ of the benzyl group appeared as a singlet (2H) at 5.19 ppm, and the five aromatic protons of the benzyl group as a multiplet at 7.31-7.44 ppm. Aromatic protons at C-2, C-4 and C-6 all have different chemical environments now due to the presence of the benzyl group (the two protons at C-2 and C-6 were in the same chemical environment in the starting material **19** at 8.97 ppm), and thus appeared as doublet of doublets due to *meta* coupling. C-4 was more upfield (8.04 ppm) relative to its position in the starting material **19** (9.23 ppm) due to the shielding effect of the neighbouring benzyloxy- group. C-6 was now the most deshielded proton at 8.28 ppm due to the *ortho* aldehyde and nitro groups. The aldehyde signal at 10.01 ppm showed slight upfield movement due to possible long range effects of the benzyl group. Mass spectrometry confirmed the expected mass C$_{14}$H$_{10}$O$_4$N (M-) 256.0615.

5.1.3. Synthesis of methyl azidoacetate

We investigated the synthesis of two candidate azidoacetates to explore azide coupling. Inspired by Farnier's work on indole cyclisation using ethyl azidoacetate,[115] [116] we investigated coupling with this commercially available azide. However, the desired coupling did not take place, as starting material remained by the end of the reaction.

Next, the synthesis of methyl azidoacetate was explored (Table 5). This involved the reaction of methyl bromoacetate with sodium azide under various conditions. Reactions using dimethyl formamide (DMF) at room temperature,[117] and refluxing in methanol,[118] did not provide the product. The Finkelstein method (in-house method in our laboratory [MS personal communication] Table 5, Scheme 67) in which the bromide atom is displaced first by an iodide atom and then an azido group, gave the correct NMR and IR spectra but subsequent coupling to the aldehyde was unsuccessful. Following Allen's method; refluxing in methanol (MeOH) for 4 hours, then a work up of MeOH removal *in vacuo* and diethyl ether/H_2O extraction[119] provided us with **21** in quantitative yield (Scheme 67) which proceeded to a successful coupling at the next step. It is not clear why using da Rosa's method[118] was unsuccessful as it had identical conditions but did not include the *in vacuo* removal of MeOH.

MeOH NaN₃
reflux 4h
quant. yield
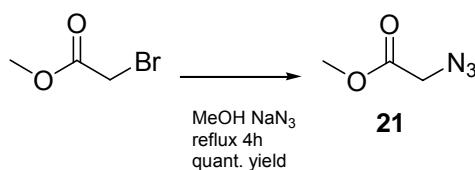
Scheme 67 Synthesis to form methyl azidoacetate **21**

Reference for method	NaN$_3$ molar ratio	Solvent	Conditions	Work up	Result
Sechi [117]	1.06	DMF	RT 2.5 h	Diethyl ether/H$_2$O extraction	Starting material only
In-house method in our laboratory (MS personal communication)	2	acetone	K$_2$CO$_3$ NaI RT 16 h	Filtered K$_2$CO$_3$, removed solvents *in vacuo*, EtOAc/H$_2$O extraction, removed solvents *in vacuo* carefully	Yellow oil in quantitative yield; No coupling
da Rosa[118]	1.22*	MeOH	reflux 4-12 h	EtOAc/H$_2$O extraction	Starting material only
Allen[119]	1.2*	MeOH	reflux 4 h	MeOH removed *in vacuo*, diethyl ether/H$_2$O extraction	Clear oil in quantitative yield; Successful coupling

Table 5 Methods used to synthesise methyl azidoacetate **21** (* mixed in water before adding).

Methyl azidoacetate is formed by the nucleophilic attack of the azido group at the relatively electro-deficient carbon α to the bromide atom. This displaces the bromide atom (Scheme 68).
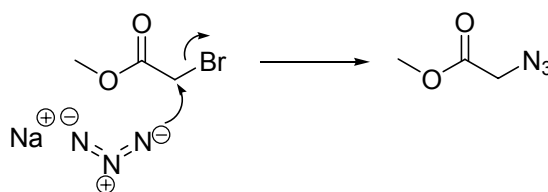
Scheme 68 Reaction mechanism to form methyl azidoacetate to form **21**

NMR spectroscopy confirmed the structure of **21**; ^{1}H NMR singlets at 3.64 ppm and 3.75 ppm corresponded to CH$_3$ and CH$_2$ respectively. IR spectra

showed absorbance at 2101 cm^{-1} due to the azide and at 1735 cm^{-1} due to the carbonyl.

5.1.4. Hemetsberger reaction step 1 – azide coupling

To achieve successful coupling of the azidoacetate **21** to the aldehyde **20**, a number of methods were investigated (Table 6). Optimisation of this step was achieved with successful coupling by two different methods, the highest yield of which was 31%. The results are discussed herein.

Reference for method	Base	Solvent	Conditions	Work up	TLC result
Tercel[70]	NaHMDS	THF	RT	EtOAc/ H$_2$O extraction	SM only
Tichenor[65]	NaOMe	THF	-78°C 3 h, 0°C 24 h	Extract H$_2$O/DCM, sat aq NaCl extraction	SM only
			-40°C 3 h, 0°C 16 h		SM only
			-25°C 3 h, 0°C 14 h		SM only
			0°C 2.5 days		Product 31%
			RT 3 h		Decomposition
Allen[119]	NaOMe	MeOH	-8°C 2 h, <5°C 2 days	Precipitation on ice	Product 19% & SM

Table 6 Conditions investigated for coupling methyl azidoacetate **21** to aldehyde **20** to form **22**, as the precursor step to Hemetsberger indole cyclisation (SM = starting material)

The use of the base sodium bis(trimethylsilyl)amide (NaHMDS) in THF at -78°C[70] was not sufficiently reactive as only starting material remained (Table 6). The coupling reaction was then carried out by the condensation of the aldehyde **20** and methyl azidoacetate **21** at -25°C in freshly prepared NaOMe, stirred for 3 hours at -25°C then at 0°C for 14 hours (Table 6).[65] Low temperatures (-78°C) resulted in no reaction despite extended reaction times.

Higher temperatures (room temperature) resulted in no product formation yet all the starting material had disappeared; the reaction may have been too rapid and resulted in decomposition. Azide coupling to form **22** was successfully achieved in 31% yield when the reaction was carried out at 0°C over 2.5 days (Scheme 69, Table 6). However, this yield was not consistently reproducible, particularly on scales above 0.1g.
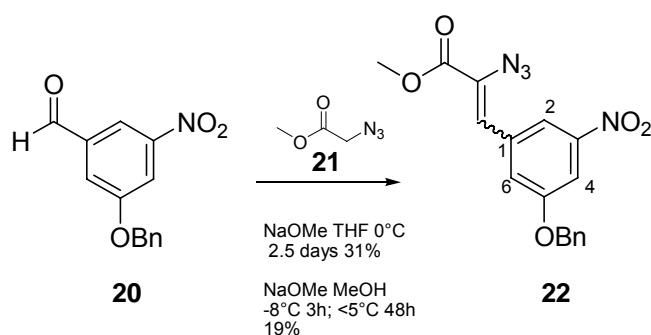
Scheme 69 Two methods for methyl aziodoacetate **21** coupling to aldehyde **20** to form **22**

Allen's method of stirring at -8°C for 2 hours, and below 5°C for 2 days thereafter,[119] provided the coupled product **22** in 19% yield (Scheme 69, Table 6). The reaction was incomplete as starting material also remained. Despite extended reaction times and increasing the molar ratio of sodium methoxide (NaOMe) and methyl azidoacetate **21**, this yield could not be optimised. This reaction had lower efficiency but was found to be reproducible.

The condensation reaction to form **22** may proceed by the mechanism shown in Scheme 70. The benzaldehyde condenses with the anion of methyl azidoacetate in an analogous way to aldol condensation. First, base (NaOMe) then deprotonates the carbon α to the carbonyl group of methyl azidoacetate, resulting in the formation of the anion (**a**). The aldehyde carbonyl carbon then undergoes nucleophilic attack by the carbanion to form the azido alcohol (**c-d**). Subsequent deprotonation of the second proton α to

the ester leads to elimination and loss of OH⁻ to generate the coupled product in a mix of *cis/trans* isomers (**d-e**).
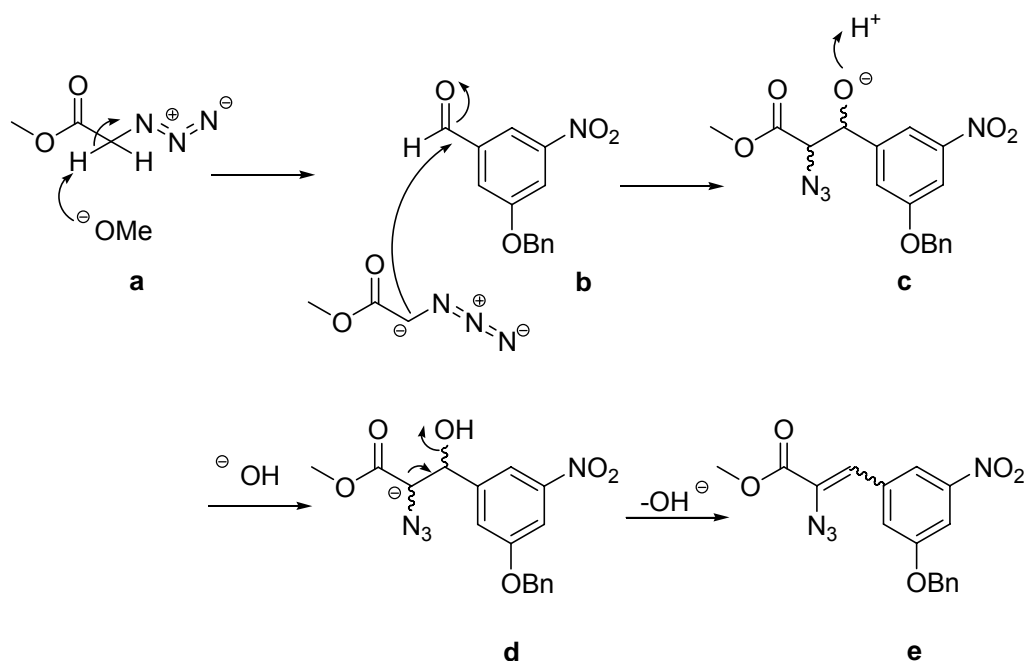
Scheme 70 Reaction mechanism for condensation of methyl azidoacetate with aldehyde to form **22**

This condensation step can be difficult as the methyl azidoacetate anion is unstable with the loss of nitrogen,[70] and may be the reason for low reaction efficiency seen here. To optimise this reaction, excess azidoacetate was used (4 molar equivalents used here) with as mild conditions as possible. Tercel and co-workers have also reported low yields for the coupling step (20-35%), which improved to only 45% when NaHMDS was used as the base instead of methoxide.[70] It may also be that the quality of MeOH used has a significant impact on the reaction; any presence of water would render the base ineffective.

^{1}H NMR spectra confirmed the structure of **22**, and is comparable to the literature reference.[65] The aldehyde peak from the starting material **20** is

absent. The ester CH_3 was present as a singlet at 3.93 ppm. A singlet at 6.84 ppm was due to the alkene CH. The proton at C-6 appears unaffected by the azide coupling at 7.78 ppm as there is no change in chemical shift from the starting material **20**. The proton at C-2 has shifted upfield at 7.81 ppm (from 8.28 ppm in **20**) and the proton at C-4 has shifted downfield at 8.20 ppm (from 8.04 ppm in **20**).

Thus, although low yielding, **22** was successfully synthesised as the condensation reaction of methyl azidoacetate **21** and the aldehyde **20** using two coupling methods.

5.1.5. Hemetsberger reaction part 2 – cyclisation to indole

Cyclisation to form the indole **23** following Hemetsberger methodology was achieved using a simple method published by Allen,[119] refluxing xylene at 140°C for 5 hours. This provided the desired product **23** in 22% yield (Scheme 71).
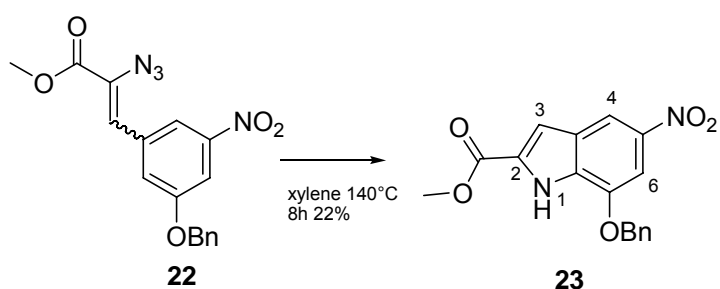
Scheme 71 Hemetsberger indole cyclisation to form **23**

TLC showed that the reaction was incomplete, with starting material remaining even after prolonged reaction times. ^{1}H NMR spectra show the appearance of a doublet at 7.32 ppm due to the new indole C-3 proton, which is now in the relatively more deshielded aromatic ring system

environment (compared to 6.84 ppm for the azide CH in the starting material **22** which is now absent). NH appeared as a broad singlet at 9.31 ppm confirming the formation of an indole.

The mechanism for indole cyclisation using Hemetsberger methodology is not completely confirmed. The observations of the original authors of this reaction, Hemetsberger and co-workers,[120] led them to suggest an indirect mechanism as outlined in Scheme 72. The azido precursor undergoes decomposition, by thermolysis (this can also occur by photolysis), to provide a nitrene, an electron deficient highly reactive species, and molecular nitrogen.[115 120 121] Hemetsberger suggested that the nitrene then attacks the neighbouring vinyl carbon, forming an azirine (also described as azacyclopropene). This is based on ^{1}H NMR spectroscopic presence of a singlet peak for the 2-H azirine proton at 3.30 ppm, and IR absorbance at 5.68 μ (equivalent to 1761 cm^{-1})* attributed to the existence of the azirine. This azirine was also isolated in a later paper.[122] The authors postulated that this azirine is in thermal equilibrium, and cleaved by thermolysis to form the nitrene which undergoes intramolecular insertion into the aromatic sp^2 C-H bond, cyclising to form the indole.[115 121] Support for this route comes from earlier work by Smolinsky where he showed that thermolysis of vinyl azides led to the formation of azirines,[123] and work carried out by Isomura and co-workers also demonstrates this.[124] Taniguchi's work on thermolysis of 2H-azirines showed that C-N bond fission occurs, giving rise to vinyl nitrenes which in turn can cyclise to indoles.[125] Hemetsberger and co-workers stated that it was not clear from their results if the mechanism proceeds directly to indole after nitrene formation, or indirectly via the azirine.[120] Subsequent authors, notably Farnier and co-workers, inferred the direct nitrene insertion

---

* λ (μ) = 10 000/wave no. (cm^{-1})

route.[115] As **22** had been used for cyclisation without any further purification of *cis/trans* isomers, thermolysis provided **23** and its regioisomer. Resolution of isomers was not attempted here due to the small scale involved, however selectivity in favour of **23** is established.[65]

In summary, the first step of this synthetic route, one pot nitro displacement and benzylation, was achieved in a good yield of 71%. Two coupling methods were investigated for condensation of methyl azidoacetate to an aldehyde to form the precursor Hemetsberger structure **22**. **22** was then successfully cyclised to form the indole **23**. As both key Hemetsberger steps were found to be low yielding in this instance, alternative methodology needed to be explored to generate larger quantities of indole, to progress the synthesis towards the target duocarmycin SA subunit structure.
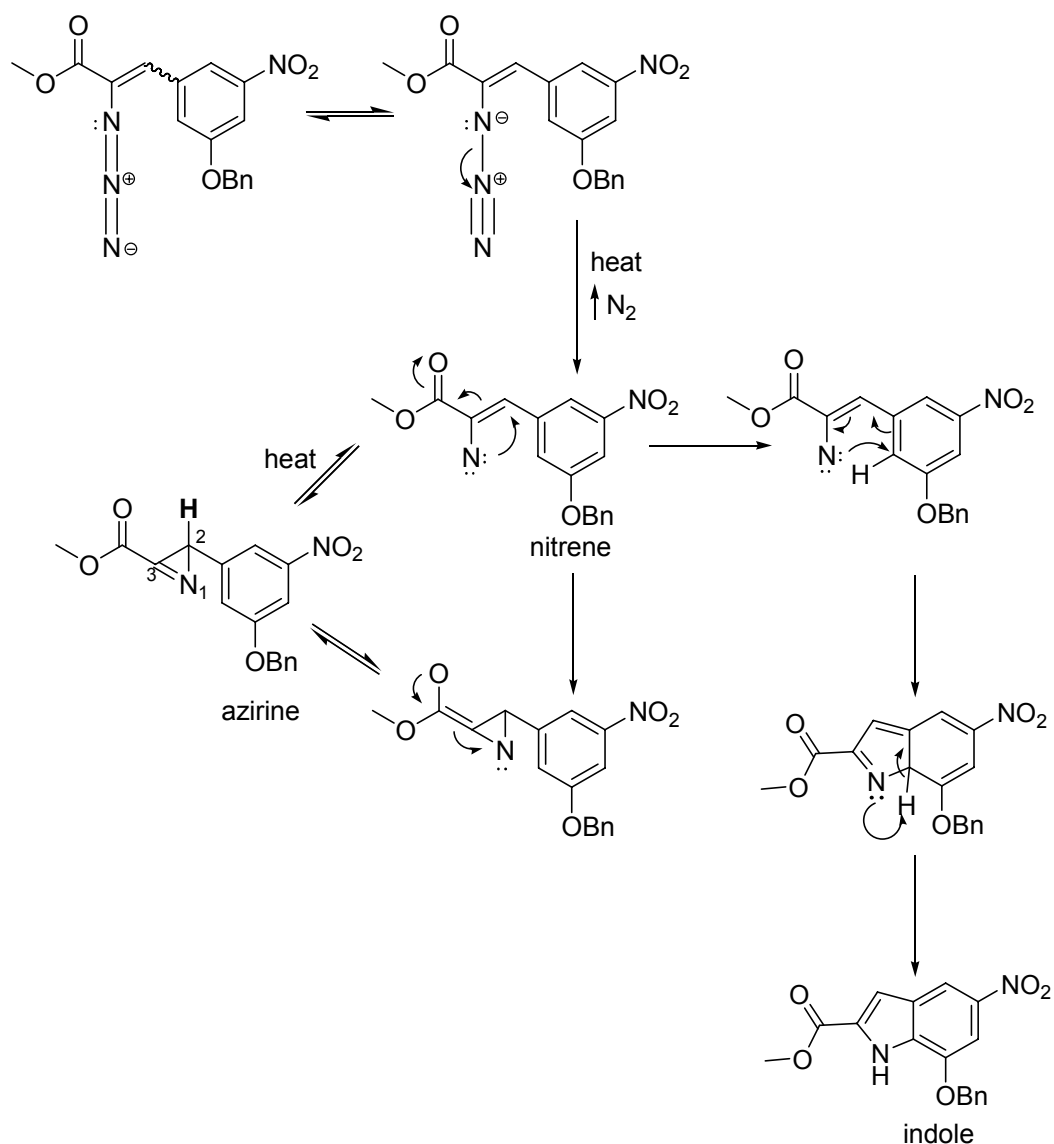
Scheme 72 Postulated reaction mechanism for Hemetsberger indole cyclisation[120] (2-H azirine proton shown in bold)

5.2. Summary

In summary, the second route to duocarmycin SA used Hemetsberger azide methodology to synthesise the left hand indole of our DSA structure (Scheme 59). Swern oxidation provided the precursor aldehyde **19** in high yields (85%). One pot nitro group displacement and phenol group protection via benzylation to form **20** was also achieved in good yields of 71%. Methyl azidoacetate **21** was synthesised in-house in excellent yields.

The precursor step of condensation of methyl azidoacetate **21** with the aldehyde **20** to form **22** was successful. Cyclisation to indole **23** via Hemetsberger methodology was also achieved. Efforts to optimise this coupling/cyclisation were made (Table 6), however low efficiency of methyl azidoacetate condensation and indole formation prevented us from pursuing this route further.

6. Third route to the duocarmycin SA alkylation subunit, DSA, utilising the Negishi/Hiroya approach

For assembly of the core subunit of duocarmycin SA, DSA, formation of the indole heterocycle is a key issue. A high yielding method is also required as this step features early in the synthetic route. The previous chapters outline a number of indole formation methods investigated to ultimately reach DSA. Allylation, both direct and metal catalysed was not found to be a feasible strategy (chapter 4). Hemetsberger azide methodology did provide the desired indole, however low yields prevented further development of this route (chapter 5). Low yields and difficulties in separation and purification also precluded the use of Fischer indole cyclisation (chapter 5).

With the aim of synthesising DSA by a concise and efficient route, we investigated the potential of the Negishi reaction for indole cyclisation, as described by Hiroya and co-workers.[61] Although, this synthesis was on a small scale, it offered a relatively direct way into the duocarmycin DSA.

Synthesis to Boc-DSA is shown in Scheme 73, modified from that described by Hiroya and co-workers.[61] 2-amino 5-nitrophenol was taken through a series of steps; benzyl protection of the phenol, iodination at C-3 as the precursor halide for Negishi coupling, mesylation to protect the amino group, and Negishi coupling utilising *in situ* Pd/ligand formation for indole formation. Synthesis with the cyclised structure proceeded further with nitro group reduction and iodination at C-4 as the steps leading to alkylation and indoline cyclisation, to ultimately form the DSA subunit. This would be

used as a core structure to attach targeting ligands. A discussion of this synthesis follows.
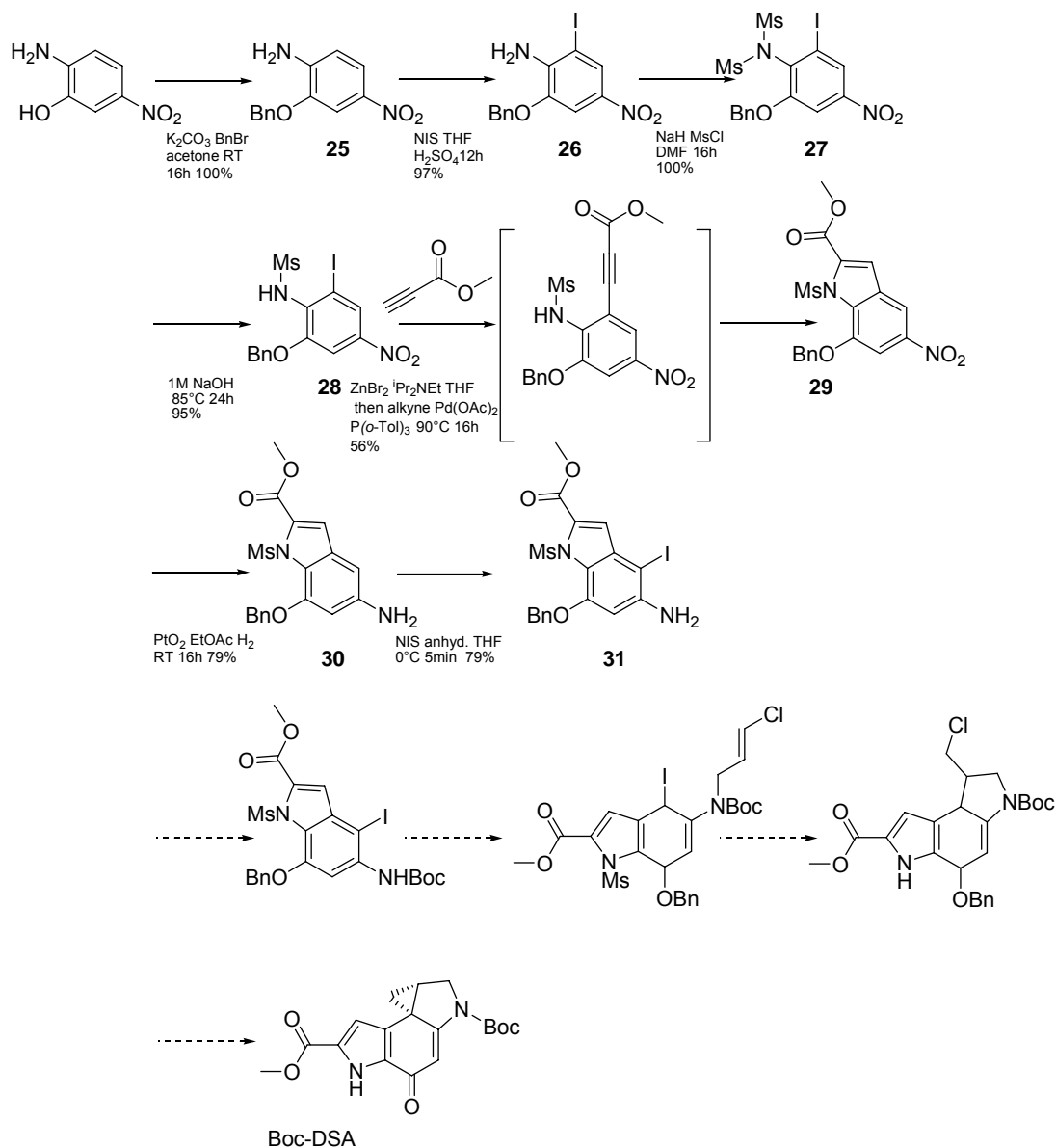
Scheme 73 Synthetic route leading to Boc-DSA based subunit, utilising Negishi coupling for indole formation

6.1.1 Benzyl Protection of 2-amino-5-nitro-phenol

The first step in this synthesis was protection of the phenol group of commercially available 2-amino-5-nitrophenol which proceeded to give a conversion to a benzyl ether **25** in 100% yield (Scheme 74).
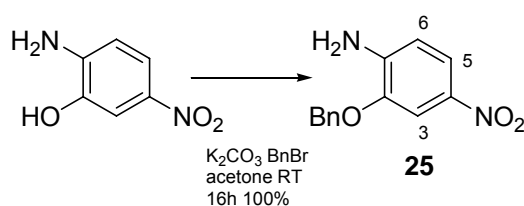
Scheme 74 Benzylation of 2-amino-5-nitro-phenol to form **25**

Benzylation proceeds by S_N2 type attack on the reactive benzyl bromide by the phenolate ion (chapter 3 Scheme 15 shows mechanism). The structure of **25** was confirmed by ^{1}H NMR and mass spectrometry, and matched the literature compound.[61] Benzylation was confirmed by the presence of ^{1}H NMR signals at 5.15 ppm (singlet, 2 protons) due to the CH_2 of the benzyl group, and the five aromatic protons of the benzyl group as a multiplet at 7.36-7.48 ppm. The proton at C-6 appeared as a doublet at 6.67 ppm due to *ortho* coupling with the neighbouring C-5 proton (J=8.7 Hz). It was relatively shielded compared to the other aromatic protons due to the neighbouring amino group at C-2. The proton at C-3 also appeared as a doublet at 7.78 ppm, this time due to long range *meta* coupling with the C-5 proton (J=2.4 Hz). The proton at C-5 appeared as a doublet of doublets at 7.83 ppm due to *ortho* (J=8.7 Hz) and *meta* coupling (J=2.4 Hz) with C-6 and C-3 respectively. The presence of the *ortho* nitro group deshields the C-5 proton. Mass spectrometry confirmed the mass $C_{13}H_{12}N_2O_3$ (M+) 244.0839.

6.1.2. Iodination at C-3

The following iodination step was regiospecific as the incoming halogen was directed *ortho* to the amino-group. The reaction utilised NIS with catalytic amounts of H_2SO_4 in THF to form **26**, proceeding in an excellent of yield 97%, following workup and purification (Scheme 75).
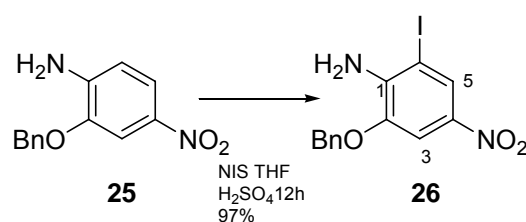
Scheme 75 Iodination at C-3 to form **26**

The reaction proceeds in two steps as shown in Scheme 76. First, NIS is activated through protonation in the presence of H_2SO_4, forming a species which effectively represents an iodonium ion (**a**). Regioselective iodination is dependant on the aromatic substituents. While the benzyloxy substituent is activating the position *ortho* and *para* to it through conjugative effects, the nitro-group is effectively deactivating through the opposite electron-withdrawing effect. The conjugative electron-donating effects of the amino substituent then direct substitution at C-6. The iodide then undergoes nucleophilic attack by C-6 of the aromatic ring (**b**). Aromaticity of the ring is maintained on the loss of the proton (**d**).
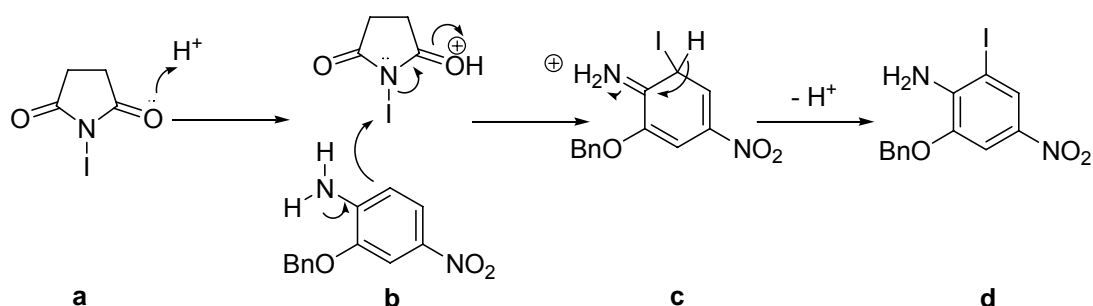
Scheme 76 Iodination reaction mechanism

Regioselective iodination was confirmed by NMR and mass spectrometry. The ^1H NMR signal for C-6 was now absent due to the presence of the iodide, and little change in chemical shift was seen for the proton signal at C-3 (7.74 ppm) as it is distant from the iodide at C-6. A peak at 8.29 ppm was

due to the proton at C-5, which is shifted more downfield compared to the starting material **25** (7.83 ppm). The peak splitting pattern was now a doublet as there was *meta* coupling only with C-3 (*J*=2.4 Hz). Mass spectrometry confirmed the mass $C_{13}H_{11}IN_2O_3$ (M+) 369.9811.

6.1.3. Dimesylation of the amino group

The next step was protection of the amino group. We report a facile method using cost effective reagents and simple work up of precipitation/filtration with higher yields (95% over 2 steps), compared to that described by Hiroya for mesylation (75% over 2 steps)[61] to form the mono-mesylated product **28**.

Hiroya utilises mesyl group protection[61] as it gives electronic characteristics conducive for the subsequent Negishi coupling and cyclisation; it reduces the nucleophilic character of the amino nitrogen, which could potentially complex with Pd during coupling. There is no discussion of alternative N-protecting groups in the original publication. Initial efforts at di-mesylation focussed on the methodology described using NaH, MsCl, THF at 0°C to room temperature, followed by hydrolysis to the mono-mesylated product using TBAF and THF at room temperature (Scheme 77). Employing this two-step method without purification at the intermediate dimesylated stage did not provide the mono-mesylated product. On separation of the two steps, it was evident from NMR that the first di-mesylation step was indeed working (88% yield), whereas hydrolysis to form the mono-mesylate was proving to be more problematic.
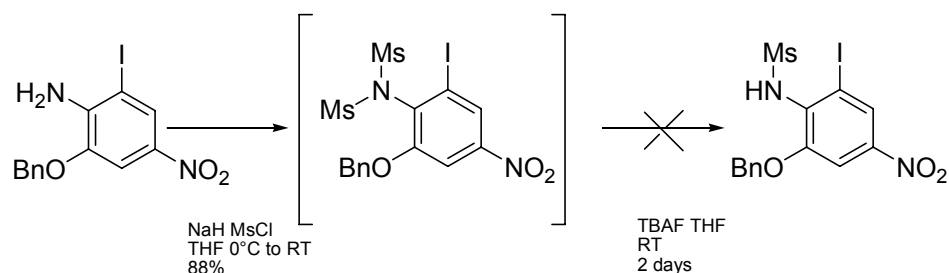
Scheme 77 One pot dimesylation/de-mesylation of **26** using Hiroya's method to form **27**[61]

An alternative facile method was found using NaH, MsCl in DMF, with a simple filtration work up and no further separation/purification,[126] to give the dimesylated product **27** in quantitative yield (Scheme 78).
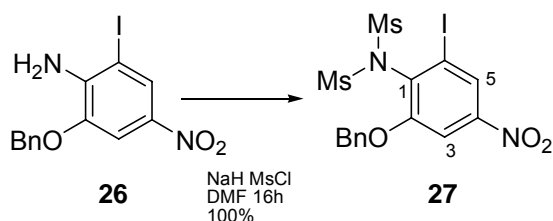
Scheme 78 Dimesylation to form **27**

The mechanism for mesylation may occur as shown in Scheme 79.[127] First, the amine of the starting material **26** undergoes deprotonation by NaH to generate an amide ion (**a**-**b**). The resulting amide ion acts as a nucleophile, attacking mesyl chloride (**b**). Electron rearrangement results in the elimination of chloride, to form a sulfene.[83] Nucleophilic attack by the aromatic amine occurs at the relatively electron deficient sulphur atom of the sulfene (**d**). The ammonium ion loses a proton to form the mono-mesylated product (**e**). The resulting sulphonamido group contains an acidic proton that is then removed by the excess NaH (**f**) and reacts with a second sulfene (**g**). The final proton is presumably extracted from mesyl chloride to generate further sulfene. Dimesylate will always be formed preferentially under these conditions due to the ease of deprotonation of the initial product

and reaction of the resulting anion. Hence, a two step process is required to achieve monomesylation.
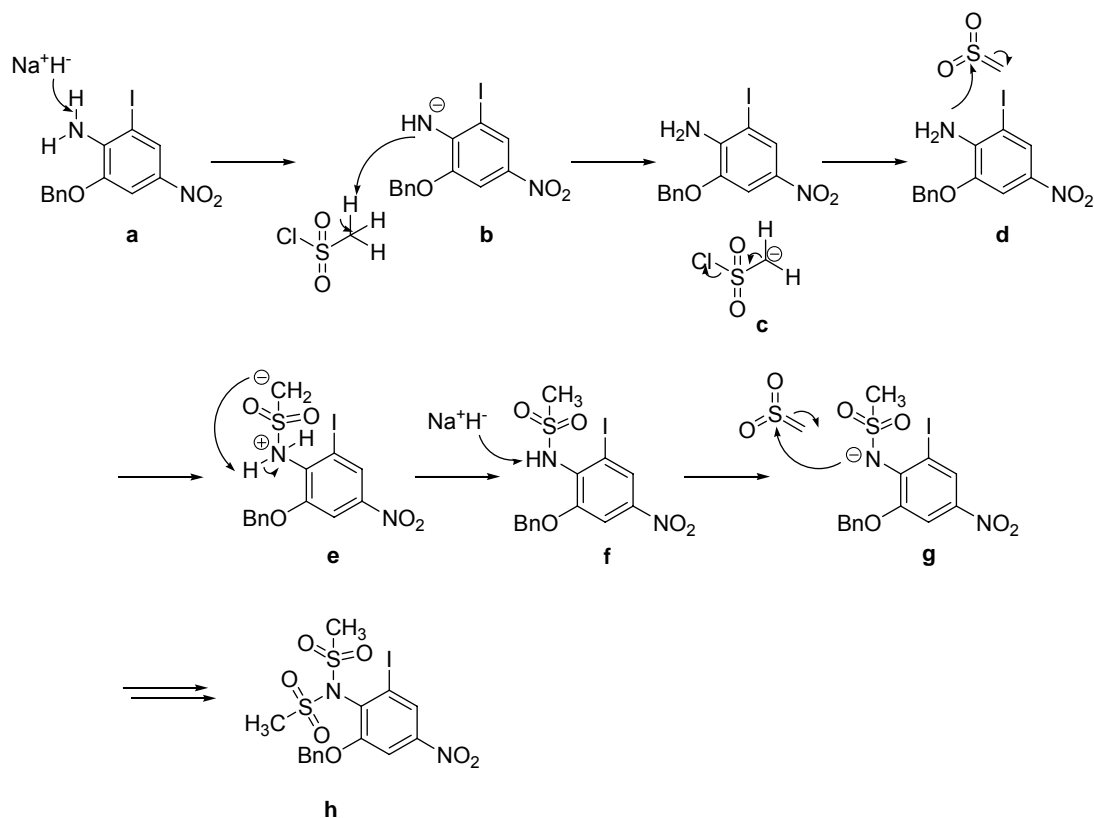
Scheme 79 Dimesylation reaction mechanism to form **27**

Dimesylation was confirmed by ^{1}H NMR; a singlet peak at 3.42 ppm was due to the two identical mesyl CH_3 groups (6 protons), and mass spectrometry confirmed the expected mass $C_{15}H_{15}IN_2O_7S_2$ (M+) 525.9363. IR absorptions at 1363 and 1160 cm^{-1} were attributable to S=O bond asymmetrical (1372-1335 cm^{-1}) and symmetrical (1195-1168 cm^{-1}) stretching.

Thus, confident that the dimesylate could be generated and characterised, efforts were directed towards de-mesylation to form mono-mesylate as a separate secondary step.

6.1.4. Demesylation

Subsequent formation of the mono-mesyl compound **28** was met with success, by heating at 85°C in 1 M NaOH, in 95% yield (Scheme 80).[126] This is a better dimesylation/de-mesylation method compared to that of Hiroya's[61] in terms of the simplicity of the work up and the improved yield. In our method, both reactions (dimesylation followed by demesylation) merely required precipitation and filtration to provide solids in improved yields. These could be used for subsequent steps without any further purification.
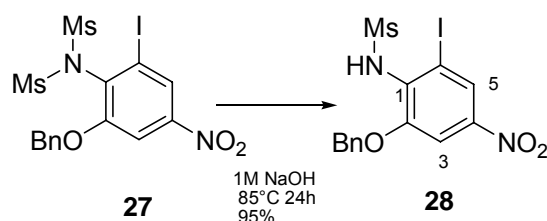
Scheme 80 Demesylation to form **28**

Demesylation occurs by the sulfonyl undergoing nucleophilic attack by a hydroxide ion as shown in Scheme 81. This leads to the cleavage of the S-N bond as the nitrogen acts as a leaving group. As the final step, the basic nitrogen gains a proton to form the mono-mesylated product. The presence of two mesyl groups result in delocalisation of the aniline nitrogen lone pair electrons over the N-S bonds and sulfoxide π system, and hence reduced stability. When one of the mesyl groups is removed by hydrolysis, the nitrogen lone pair is able to fully overlap with the remaining S=O, enhancing stability. Hence only one mesyl group can be removed in these conditions. It is possible that in harsher i.e more basic conditions, all mesyl groups would be cleaved, and this is why only mildly basic (1 M NaOH) conditions are used.
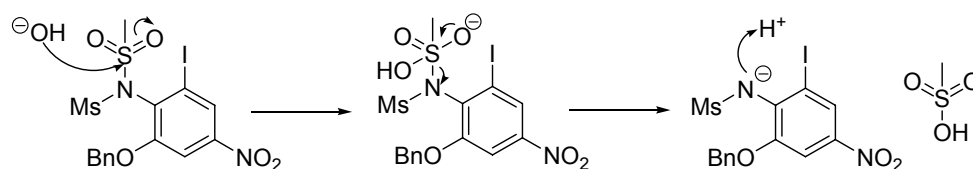
Scheme 81 Demesylation reaction mechanism

The structure of **28** was confirmed by ^{1}H NMR spectroscopy. A singlet at 3.02 ppm was now due to only one mesyl CH$_3$ group (3 protons), and the broad singlet at 6.39 ppm was due to sulphonamido proton. Mass spectrometry gave C$_{14}$H$_{17}$IN$_3$O$_5$S (M+NH$_4$)+ 465.9925, confirming the expected mass. IR absorption at 3254 cm^{-1} showed the presence of the NH bond.

6.1.5. Negishi cross-coupling for indole generation

Negishi coupling is the critical step for this synthesis, to generate the first heterocycle. Conventional solution phase and microwave methodology was tested. We report comparable yields to that of Hiroya of 56%,[61] using *in situ* Pd complex formation.

Negishi coupling is the Pd or Ni catalysed coupling of aryl-, alkenyl-, or alkynyl halides with organozincs, to form a new carbon-carbon bond[128] as shown in Scheme 82. It follows an oxidative-elimination pathway, although the structure of the active catalyst species and the exact role of ZnBr$_2$ are not fully understood.
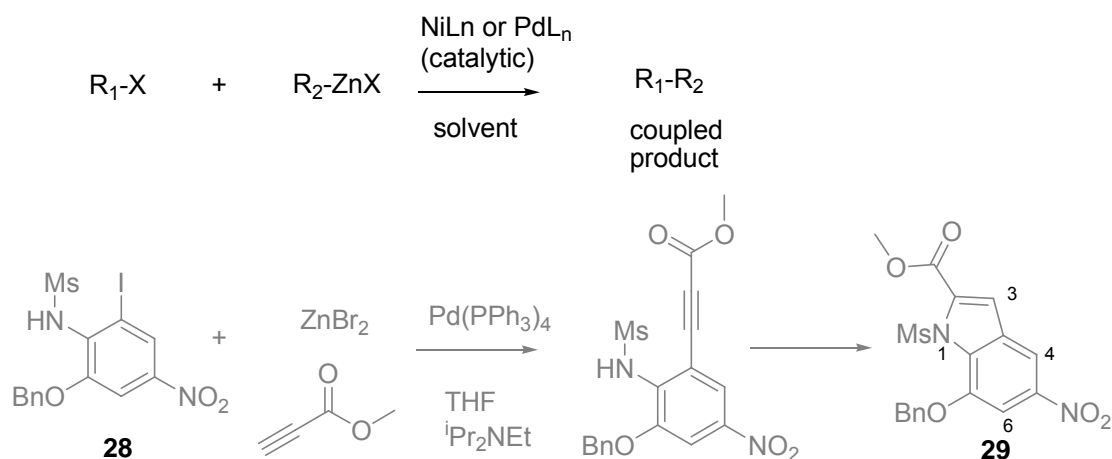
Scheme 82 General reaction for Negishi coupling, with reagents used in our case in grey, to form **29** (X=halide atom).

Negishi coupling was the pivotal step in this synthesis. Using Hiroya's method;[61] ZnBr₂, ᶦPr₂NEt, methyl propiolate, Pd(PPh₃)₄, THF and refluxing for 11 hours resulted in the methyl indole-2-carboxylate derivative **29** in yields of 29% (Scheme 82, Table 7) following a difficult separation. The cyclised compound **29** was accompanied by a co-eluting molecule that could only be separated by very careful column chromatography, requiring slow elution of the products, and hence large quantities of solvent. Increasing reaction times, increasing concentrations of ZnBr₂, ᶦPr₂NEt and catalyst did not improve yields.

The resulting indole was characterised by NMR and mass spectrometry. A ¹H NMR singlet at 3.43 ppm (3 protons) due to the mesyl CH₃ showed some deshielding i.e downfield chemical shift as a result of indole formation. A singlet at 3.94 ppm (3 protons) appeared due to the methyl group of the ester. The methylene of the benzyl group was still present as a singlet at 5.29 ppm (2 protons). The singlet peak for the indole proton at C-3 appeared very close to the chloroform peak at 7.27 ppm. The benzyl group appeared as two distinct multiplets at 7.46-7.44 ppm (3H, m, ArH) and 7.60-7.57 ppm

(2H, m, ArH). The aromatic protons at C-6 and C-4 appeared as doublets at 7.92 ppm and 8.19 ppm respectively, due to *meta* coupling with each other (*J*=1.8 Hz). Mass spectrometry confirmed the expected mass $C_{18}H_{16}N_2O_7S$ (M$^+$) 404.0675.

Hiroya and co-workers report a 56% yield for this reaction, but this was on a very small scale reaction of 57mg.[61] The authors used more material than obtained in the reported reaction in subsequent steps of the synthesis, so it seems likely that the yields were not improved on scale up and, as found here, were potentially even lower on a larger scale. Longer reaction times, increasing Pd or Pd/ligand loading did not improve yields. As a result, we investigated optimisation of the Negishi coupling by varying reaction conditions and the catalyst used.

6.1.5.1. The reaction mechanism for Negishi coupling

The reaction mechanism for Negishi coupling may proceed as outlined in Scheme 83. Pd$^{(0)}$ is generated from a Pd$^{(II)}$ complex, by reduction with the ligand substrate. This is followed by oxidative addition of Pd$^{(0)}$ to the halide (aryl halide in our case). The rate of nucleophilic attack on the aromatic ring is dependent on substituents on the aromatic ring. Electron-withdrawing groups accelerate oxidative addition, whereas electron-donating groups retard it.[129] Transmetallation then occurs, where the metal (ZnBr$_2$) transfers the R group (methyl propiolate in this reaction) to Pd, in exchange for the aryl halide (I in this reaction). Relatively fast reductive elimination requires a *cis* relationship between R^1 (aryl iodide in this reaction) and R^2 (methyl propiolate in this reaction),[130] providing the coupled product. *Cis/trans* isomerisation may occur at different rates depending on ligand types and

organometallic agent used.[131] The Pd catalyst is then regenerated, completing the catalytic cycle.
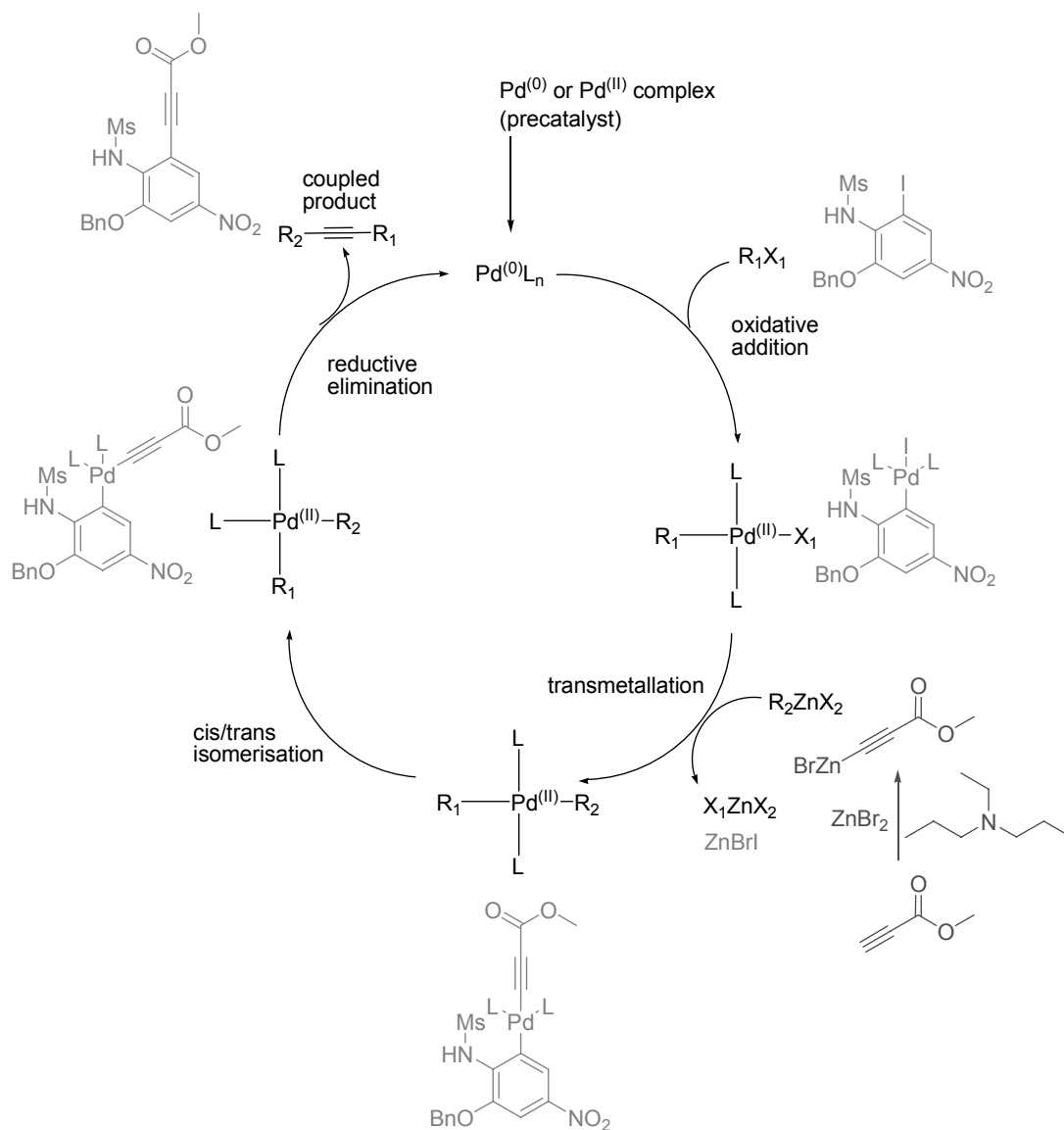
Scheme 83 General catalytic cycle for Negishi coupling,[94] using **28** to form **29** (in grey)

Transmetallation has been found to be the rate-determining step in the catalytic cycle; highly electropositive metals can inhibit subsequent reductive elimination by forming their own complexes with Pd.[130] In this case, ZnBr$_2$ is not sufficiently electropositive to have such an impact, and in general Zn containing organometallics show the highest reactivity.[132] Reductive

elimination can be affected by a number of factors such as the ligand used, countercations, and the carbon groups being coupled.[130] Steric hindrance by substituents on the alkyne may also play a part in the rate of Negishi coupling.

In the reaction described here, Negishi coupling is in fact the intermediary step, which is immediately followed by cyclisation to the indole product **29** (Scheme 82). The exact mechanism for indole cyclisation is not clear.

It is evident from our experiments that the difficult purification was due to the presence of the co-eluting uncyclised alkyne. This means that part of the reaction remains at the intermediate alkyne stage and the reaction does not progress fully to completion. This is mentioned by Hiroya,[61] and was identified as a possible impurity in product by TLC, NMR spectra on separation, and tentatively by IR. Progression to cyclisation was not possible for all material despite increasing reaction times and substrate concentrations. Thus, it appears that cyclisation is the problematic step in this reaction, and a likely reason for low yields (29%). This step could have been repeated using the alkyne mixture to see if further indole could be generated.

The similarity in polarity between the alkyne intermediate (as shown in Scheme 82) and the indole product **29** rendered separation by silica gel column chromatography particularly difficult. Despite using the slowest possible solvent systems (5% EtOAc in hexane), co-elution of **29** with the alkyne intermediate was unavoidable, affecting the final reported yield. Separation by automated column chromatography did not provide any superior separation. Increasing solvent system polarity to speed up fraction collection of **29** also resulted in co-elution. Thus, all of these factors

contributed to extremely poor yields. Purification was carried out by HPLC to characterise the intermediate and side products. However, as this purification was only carried out on a 3 mg scale, these products were not in sufficiently high quantitites to be characterised by NMR. Purification by HPLC on a larger scale would be needed to achieve this.

Overall, cyclisation to indole **29** was successfully achieved using Negishi coupling. Optimisation of this reaction was desired by varying conditions and catalysts used, to improve the 29% yield achieved so far. A discussion of optimisation methods follows.

6.1.5.2. Alternative palladium catalysed coupling methods

Palladium catalysis is often an efficient way to achieve many varied functional transformations. Alternative methods to couple the alkyne, methyl propiolate to the aryl iodide **28** were investigated.

Sonogashira's conditions were explored to achieve a fast and high yielding coupling, where Cu is used as the transition metal. The conditions were heating methyl propiolate, $Pd(PPh_3)Cl_2$ (2 mol%), CuI (4 mol%) and K_2CO_3 (20 mmol) in anhydrous THF.[133] The reported temperature of 65°C resulted in extensive decomposition. Lower temperatures of 45°C for 3 hours and 30°C for 5 hours resulted in no reaction and return of the starting material. Thus, either these conditions are not conducive for coupling, or the theoretical optimal temperature for the reaction results in decomposition, again preventing successful coupling. Prolonged reaction times at lower temperatures were not attempted. As there was no evidence for any reaction after 3 or 5 hours, it was unlikely to be successful. Eckert reports that electron-withdrawing substituents on the aromatic halide require shorter

reaction times (2 hours), whereas electron rich substituents confer longer reaction times (12 hours),[133] suggesting that this substrate should react rapidly. Also, Sonogashira's conditions have been widely reported to provide unacceptably low yields when electron-deficient alkynes are used.[134] [135] This may be due to anion stabilisation by the electron-deficient alkyne, rendering it unreactive during transmetallation. With this in mind, this method was not investigated further.

6.1.5.3. Optimising Negishi coupling in the microwave

The decision to retain Negishi conditions and optimise coupling using the microwave was due to the fact that there is much precedent with palladium couplings employing this medium. Considerable work has been published on Heck reactions by Larhed and co-workers,[97] [136] reactions using Stille and Sonogashira conditions, and even ligand-free aqueous Suzuki couplings.[97] Little work has been published on microwave-assisted Negishi couplings; some examples are the use of activated aryl bromides and heteroaryl chlorides with organozinc halides,[137] [138] [139] [140] enantioselective Negishi coupling[141] and Negishi coupling using Ni/C catalysts.[142] Additionally, heterocycle syntheses have been carried out in the microwave,[97] [98] including Fischer indoles[143] and indolines.[144] Although it is acknowledged that microwave irradiation would significantly reduce reaction times, scale up can be potentially limited.

The reagents were added as for the conventional reaction; **28**, $ZnBr_2$, $^iPr_2NEt$, methyl propiolate, $Pd(PPh_3)_4$ and THF were mixed at room temperature for 30 minutes under N_2, then submitted to microwave irradiation at 110°C (high absorbance) for 20 minutes. TLC showed that all starting material had been consumed so no further irradiation was required. Work up and separation

by silica gel chromatography provided the indole carboxylate **29** in 38% yield (Table 7). Increasing reaction times, increasing concentrations of ZnBr$_2$, ^{i}Pr$_2$NEt and catalyst did not improve yields, in fact they were observed to decrease.

Other reagents	Catalyst	Conditions	Yield
ZnBr$_2$ (3), ^{i}Pr$_2$NEt (6), methyl propiolate (2.75), anhyd. THF	Pd(PPh$_3$)$_4$ (2.5%)	reflux/11h	29%
	Pd(PPh$_3$)$_4$ (2.5%)	microwave 110°C/20mins/ high abs	38%
	Pd(PPh$_3$)$_4$ (2.5%)	5 mins pre-mix ZnBr$_2$ & ^{i}Pr$_2$NEt 90°C/3h	51%
	In situ complex: Pd(OAc)$_2$ (2.5%) & tri o tolyl phosphine (1%)	ZnBr$_2$ in THF, then 5 min pre-mix with ^{i}Pr$_2$NEt, then alkyne, SM, L, Pd added 90°C/16h	56%

Table 7 Different Negishi coupling methods investigated, with yield comparisons (molar equivalents in brackets, SM= starting material, L=ligand).

6.1.5.3.1. Pre-mix methodology

Negishi's earlier publications described pre-mixing ZnBr$_2$ and amine base before adding the remaining reagents.[134] Repeating the experiment with this initial pre-mix, increased the yield to 51% (Table 7) but required two columns for separation. However, for reasons that are not clear, this yield was not reproducible on scale up. Although Hiroya's published method used a 1:2 metal:base ratio,[61] others have recommended the use of a 1:4 ratio.[134][135] Increasing this ratio may also aid alkynylzinc derivative formation at the rate-determining transmetallation step of the catalytic cycle, although this was not attempted here.

6.1.5.3.2. Optimising reagents & *in situ* Pd complex formation

Attention was then turned towards optimisation by considering each reagent present in the Negishi catalytic cycle in turn. The catalyst is by nature an essential feature of the Negishi coupling. Although Ni catalysts are more cost effective and have higher reactivity, Pd catalysts have been found to provide better yields and greater variety in functional group compatibility.[132] Within the spectrum of Pd catalysts, $Pd^{(0)}$ complexes have been found to be relatively less stable than $Pd^{(II)}$.[130] Problems of air and moisture sensitivity are experienced more with commercially available $Pd^{(II)}$ catalysts such as $Pd(PPh_3)_4$,[145] leading to variable and irreproducible results. This may explain the behaviour seen with $Pd(PPh_3)_4$ in these experiments; yields increased from 22% to 38% when new batches of $Pd(PPh_3)_4$ were used during microwave-assisted experiments.

Using a method consisting of an initial pre-mix of $ZnBr_2$ and $^iPr_2NEt$, and forming an *in situ* Pd-ligand complex instead of using a pre-formed commercially available catalyst, **29** was synthesised in 56% yield. This was our highest yield attained on a preparative scale (Table 7). Satisfied with this yield at this point, the synthesis was continued.

For Negishi coupling, a commercially available $Pd(PPh_3)_4$ catalyst complex was used, as published by Hiroya and co-workers.[61] To eliminate problems of catalyst stability and degradation, a stable source of Pd was sought that could be converted to the active species via *in situ* complex formation which would improve yields. This would allow the study of ligand structure on reactivity, and confer greater versatility. For this, careful consideration of

both Pd and ligand sources was required, and parameters to be considered are discussed below.

Ligand behaviour and that of their transition metal complexes is governed by both electronic and steric effects, which are often inter-related. Ligand cone angle, a measure of ligand size, is defined as the solid angle formed with a metal at the vertex and hydrogen atoms at the cone perimeter (Figure 20).[146]
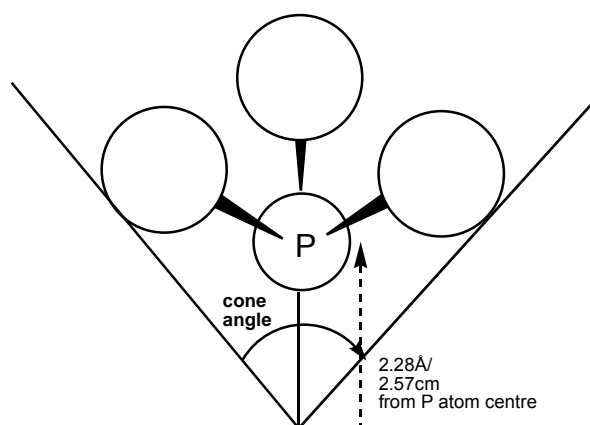
Figure 20 Cone angle is measured as the apex angle of a cylindrical cone, centred 2.28 Å/ 2.57 cm from the centre of the central phosphorous atom (outer spheres can be any atom/compound). Unsymmetrical ligands are measured by a model using the sum of each half-angle.[146]

Phosphines have been used extensively in Pd[(0)] catalysis. Triphenylphosphine is one of the most economical and effective in promoting desired catalysis (Figure 21).[130] Tolman[146] reported that increasing the substituent size of phosphine ligands increases the cone angle, bond lengths within the catalyst-ligand complex, and thus ligand dissociation. It also increases phosphine electron lone pair basicity. The resulting electron-rich palladium complexes are better able to undergo oxidative addition to less reactive substrates under mild conditions. Increased steric demand is required to promote dissociation of a ligand from

the Pd^0L_n resting state, which in turn is necessary prior to oxidative addition.[147] [148]

At the oxidative addition step, sterically demanding ligands stabilise $Pd^{(0)}$ complexes, but electron-donating ligands generate an electron-rich metal complex which undergoes faster oxidative addition reactions.[149] Oxidative addition is also more favoured when the ligand is more basic.[150] This has been found particularly in Heck and Suzuki reactions.[151] The most effective phosphine ligands have been those which are electron rich with large steric bulk; we chose to use tri-o-tolyl phosphine for these reasons (Figure 21). Electron-withdrawing substituents on the aromatic ring favour oxidative addition but may disfavour reductive elimination.[149] Thus, it is not possible to state a clear trend for optimal ligand properties; a balance is required.
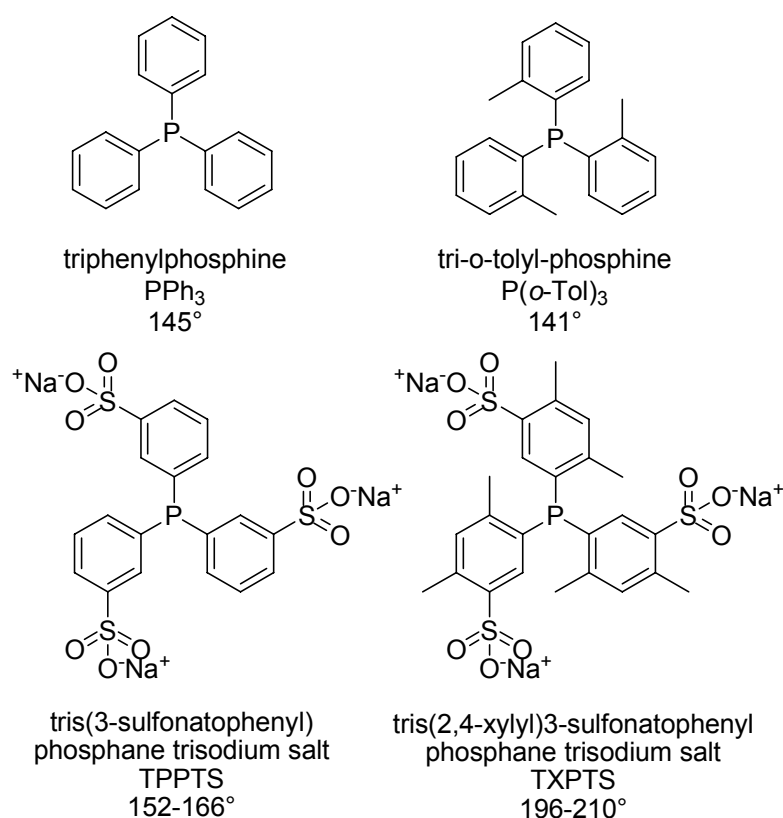
Figure 21 Phosphine ligands; PPh₃, P(o-Tol)₃, TPPTS and TXPTS with their respective cone angles[146 148]

Aqueous phase catalysts were also of interest as their use is particularly attractive for environmental and economical reasons. Hydrophilic catalysts are recoverable and thus the catalyst can be recycled. Initial development of water-soluble catalysts has been with sulfonated triphenylphosphane analogues, such as tris(3-sulfonatophenyl) phosphane trisodium salt (or triphenylphosphite trisulfonate, TPPTS) as shown in Figure 21, traditionally effective with the most reactive aryl halides corresponding to the strength of the C-X bond; iodides require the least energy to undergo oxidative addition to the Pd catalyst. In contrast, sterically demanding, electron-rich ligands are necessary for less reactive substrates.[148] These ligands have larger cone angles than PPh₃ thus greater steric demand, but their electronic nature is governed by the electron withdrawing sulfonate groups.[148]

Following on from the Pd and ligand factors discussed above, Pd(OAc)$_2$, a Pd(II) salt, was used with tri-o-tolyl phosphine as the ligand. Excess phosphine ligand is essential as Pd(OAc)$_2$ and excess triphenylphosphine have been found to form stable zerovalent complexes.[152] It is postulated that the phosphine ligand reduces Pd(OAc)$_2$ to a zerovalent Pd complex, which readily reacts with the aryl halide. The optimal Pd:L ratio has been found to be 1:4 when the ligands are phosphines.[152] For his work on intramolecular Heck arylation of 2-halo-*N*-allylanilines, Odle even suggests the use of Pd(OAc)$_2$: tri-o-tolyl phosphine in a 1:2 ratio.[153] Using less phosphine was found to generate a highly active but sensitive catalyst, giving irreproducible results.[145][152]

In terms of choice of halide for Negishi coupling, iodine is the most reactive halogen at the oxidative addition step (reactivity decreases with decreasing halogen size). Therefore, our choice of aryl halide **28** was optimised as only iodides and highly reactive bromides with electron-withdrawing substituents are effective at this step. The use of ZnBr$_2$ was also optimised, as Zn and its halides (notably chloride and bromide) have been demonstrated to provide the best yields and reaction rates.[132]

6.1.5.4. A study of *in situ* Pd/ligand complexes to optimise Negishi coupling and cyclisation

Use of the microwave would enable determination of the outcomes of varying reaction conditions rapidly, due to faster reaction times (minutes versus hours) in the microwave. Different Pd/ligand combinations were investigated to generate the most effective *in situ* Pd complex (Table 8).

Greater efficiency was desired in terms of increasing the amount of product **29** formed, and by improving the impurity profile.

Optimal molar ratios of 1:4 Pd to ligand were used, with a Pd to starting material ratio of 2.5% for the initial study. These Negishi coupling reactions were carried out in the microwave on a 0.25 g scale to obtain outcomes quickly. The reactions monitored by TLC, the crude reaction mixture worked up and subjected to HPLC analysis. Integration of UV absorbance peaks were compared as percentages (Table 8). A higher percentage would represent a high proportion of product, and a profile with less peaks would represent greater purity. The ideal would be to reach a profile as close as possible to a single peak of 100% area. These experiments showed that $Pd(OAc)_2$ with triphenylphosphine gave the highest result (40% of the crude mixture was **29**). A water soluble catalyst-ligand complex comprising of sodium tetrachloro palladate and TXPTS also showed an equally good result (41%).

As commercially available catalysts tend to be used in concentrations of 8-10%, a higher ratio of 8% Pd to starting material **28** was also investigated (see appendix for HPLC chromatograms). These results showed that poor yielding combinations at the lower concentration (2.5%) had better results at the higher concentration (8%); **29** was present in $Pd(OAc)_2$/ tri o tolyl phosphine as 62% (26% at 0.025 molar ratio) and in $Pd(PPh_3)Cl_2$ as 58% (32% at 0.025 molar ratio). $Pd(PPh_3)_4$ provided 53% of **29**, vastly improved from 13% at 0.025 molar ratio. This indicates that a clear pattern for superior catalyst/ligand combinations is not apparent at least from our initial studies.

Pd/ Complex	Ligand	Product Area (%)
Pd(OAc)₂	Tri o tolyl phosphine	26.5
Pd(OAc)₂	Triphenylphosphine	40.1
Pd(II) nitrate hydrate	Trixylylphosphite trisulfonate	37.0
Sodium tetrachloro palladate	Trixylylphosphite trisulfonate	41.4
Pd(PPh₃)Cl₂	-	31.8
PdCl₂	Tri o tolyl phosphine	25.0
PdCl₂	Triphenylphosphine	21.9
Pd(PPh₃)₄	-	13.4

Table 8 Microwave Negishi coupling to form **29** with various ligands, and yields obtained by HPLC analysis (see appendix for chromatograms)

Increasing Pd loading (from 2.5 mol% to 5 mol%), and increasing the molar ratio of both Pd and ligand (but retaining the 1:4 ratio of Pd:ligand) did not improve the results. As noted for other methods, reduction in yields were observed on scale-up and again it is not clear why this is the case, but may be due to inefficiencies in mixing or heating on a larger scale.

More ligands could be investigated, as well as further analysis of yields at higher loading concentrations.

The results from this initial study into Pd/ligand variation in the microwave could be improved and extended to provide a clearer trend of optimal Pd/ligand combinations. Although percentage of UV absorbance was used as measurement of comparative yield, the mass of **29** could be obtained by HPLC purification, providing a more meaningful comparison of actual yield.

HPLC analysis of solution phase experiments could provide a direct comparison, clearly highlighting any improvements to yield and purity obtained from conducting Negishi coupling in the microwave.

So far, an improved methodology has been established employing pre-mixing, *in situ* Pd/ligand complex formation, with microwave irradiation significantly reducing reaction times. Crude product can be placed directly on the HPLC for purification, without work up. The above described modifications to the study could provide a better picture, generating better optimised conditions for Negishi coupling.

6.1.6. Cleavage of N-mesyl group

On formation of **29**, cleavage of the N-mesyl protecting group was desired (Scheme 84) followed by benzyl group protection for easier manipulation later in the synthesis, to ultimately form DSA.
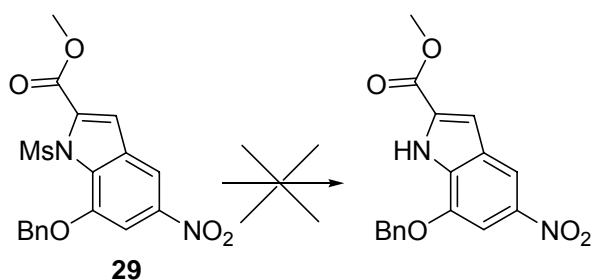
Scheme 84 N-mesyl group cleavage of **29**

Stirring at room temperature in K_2CO_3 and MeOH[61] did not provide any change to the starting material, even with prolonged reaction times, increasing the temperature, or adding more K_2CO_3. De-mesylation using 2 M NaOMe in MeOH, DCM, stirring at room temperature for 6 hours gave no clear result; starting material was consumed but no identifiable products

were formed. The N-mesyl bond may be too strong in this structure for hydrolysis. At this point, N-mesyl group cleavage was not possible.

6.1.7. Alternative amino protecting groups

As N-mesyl deprotection of the indole was found to be difficult, the synthetic route was re-examined to explore other amino protecting groups which would be readily cleaved on indole cyclisation (Scheme 85).
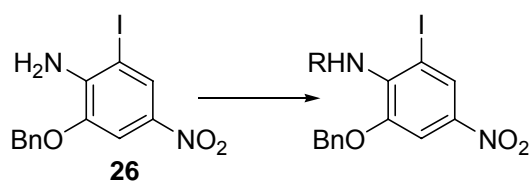
Scheme 85 Protection of amine substituent of **26** (R=Boc, triflate, tosylate)

Investigations into Boc protection seemed a natural place to start, as manipulation of this protecting group is usually relatively easy both in terms of addition and removal. Many subtly different reaction conditions were attempted (Table 9) using di-tert-butyl dicarbonate (Boc$_2$O). However, either two Boc groups became attached, or no reaction occurred. Addition of the first carbamate group activates the nitrogen for further reaction, in a similar fashion to the mesylation reaction (Scheme 79) with subsequent isolation of the di-Boc protected product.

Boc₂O molar ratio to **26**	Base	Solvent	Reaction conditions	Result
2.26	DMAP 0.2	THF	RT 24h	Di-Boc product
1.5	DMAP 1.0	THF	RT 1h	Unclear result
0.9	TEA 1.0	anhyd 1,4 dioxane	Reflux 18h	No reaction- SM only even after adding more Boc₂O
1.1	DMAP 0.1	CH₃CN	RT 16h	No reaction- SM only, even after adding more Boc₂O & DMAP
0.2	none	DCM	RT 16h	No reaction- SM only, even after adding more Boc₂O

Table 9 Different conditions investigated for Boc protection of the amino substituent of **26**

Although Boc protecting groups are usually hydrolysed by acid, this could also cleave the benzyl protecting group. Efforts to hydrolyse one Boc group from the di-Boc protected compound with K_2CO_3 (3.29 molar equivalents) in MeOH,[154] refluxed for 3 hours did not result in any reaction – both Boc protecting groups remained. Starting material **26** was consumed during the reaction, but no clear product was identifiable by NMR.

Other protecting groups were also investigated as shown in Table 10 to protect the free amine of **26** (Scheme 85). Triflate protection was studied[155] as this group is relatively easy to remove under mildly basic conditions (one example of triflate cleavage conditions is K_2CO_3, MeOH, H_2O at room temperature[156]). However in our case, NMR showed that on triflate formation, the benzyl protecting group was preferentially hydrolysed, and general decomposition was evident from TLC monitoring. Protection was also attempted using tosyl groups,[157] with the view that their reactive behaviour would be relatively close to that of mesyl groups. However, these conditions did not prove to be sufficiently reactive – starting material

remained after 16 hours. As a consequence, studies of alternative protecting groups were not pursued further.

Protecting group	Base	Solvent	Conditions	Result
BOC				See Table 10
Trifluoromethane sulfonic anhydride 1.1mol equiv	TEA 1.21mol equiv	DCM	0°C to RT 1.5h	Degradation, benzyl group hydrolysed
Trifluoromethane sulfonic anhydride 1.1mol equiv	pyridine 1.21mol equiv	DCM	0°C to RT 1.5h	Degradation, benzyl group hydrolysed
p-Toluene sulphonyl chloride	pyridine	THF	RT 16h	No reaction- SM only

Table 10 Other protecting groups investigated to protect the amino substituent of **26** (see Scheme 85).

6.1.8. Nitro group reduction

Synthesis was continued with the indole *N*-mesyl group still attached. This group would be pointing out of the groove on DNA binding and so would not be expected to affect DNA binding. Thus, the next step was reduction of the nitro group, achieved in 79% yield using PtO$_2$ (Scheme 86).
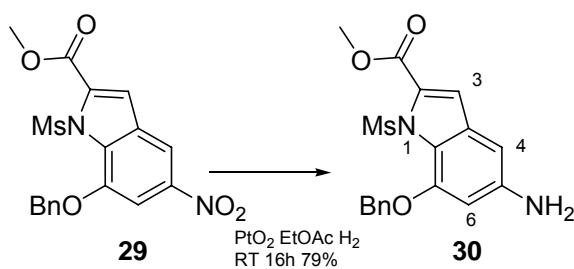
Scheme 86 Nitro group reduction to form **30**

Standard reduction methods were deemed to be too harsh resulting in degradation or unwanted reactions with other substituents on the molecule in preference to the intended reduction. For example, hydrogenation in the presence of Pd would be expected to cleave the benzyl group first, as these are standard conditions for the removal of this protecting group.[156]

Al-Hg amalgam methodology established in our lab was used as it is a mild reduction method (also discussed in chapter 3).[80] Two day reaction times provided the crude product **30** in 94% yield of mass recovered. When the crude product **30** was purified by silica gel chromatography, a slightly less polar spot than the highly polar reduced product was evident on TLC. This appeared as early as 4 hours into the reaction. NMR indicated that the mesyl protecting group was being cleaved as the major product (37%) during this reaction. Anderson and Chapman note that during the reduction of nitroamines to diamines using Al/Hg amalgam, there is a sharp rise in pH as water is used as the proton source, which may contribute to incomplete reaction to the hydroxylamine intermediate.[82] Thus, it may be that in our case the increasing basicity of the reaction mixture results in hydrolysis of the mesyl group. This could have been investigated by monitoring the pH during the reaction. Another way to improve yields and/or rate of reaction may be with the use of ultrasonication. Amalgam fragmentation results in a homogeneous dispersion, with turbulent flow then increasing contact between the amalgam and substrate.[81] [158] Metal amalgam techniques have been documented for N-sulfonyl cleavage,[159] which is the most likely reason for our observations.

This mesyl group cleavage seemed to be of use rather than a hindrance, as earlier attempts at de-mesylating had not been fruitful, as described above. The intact reduced mesylated pure product **30** was only present in 11%

yields, which left over half of the reaction mixture unaccounted for. Due to poor yields with the Al/Hg amalgam method and unpredictable cleavage of the mesyl group, alternative reduction methods were investigated (Table 11).

Reducing agent	Other reagents	Conditions	Yield
Al/Hg amalgam	$Et_2O/EtOH/H_2O$	RT 16h	94% crude; 11% pure 37% demesylated
In	EtOH, NH_4Cl/H_2O	Reflux 2.5h	Degradation probable
Pd/C 5%	Under H_2 EtOAc	RT 3h	Degradation probable
Pd/C 5%	Under H_2 EtOH	RT 2 days	Degradation probable
PtO_2	Under H_2 EtOAc	RT 16h	79%

Table 11 Various methods investigated for the reduction of the nitro group to amine **30**

The use of indium (In) is a recently described mild, ecologically friendly and relatively cost effective reduction method. Indium is also stable to air and water, unlike other reducing agents.[160] Selective reduction in the presence of other groups can often be difficult as they may be cleaved during the reaction, however this is less likely with In, mesylates remain unaffected.[160] Reduction using ethanol and aqueous NH_4Cl as reported by Banik and co-workers was investigated.[161] Consumption of starting material was observed by 3 hours but without products being identifiable (Table 11), probably due to decomposition. This reaction could have been investigated at lower temperatures to determine if heat was having an adverse effect.

Hydrogenation in the presence of Pd/C 5% is a standard method of reduction. However, this method is also known to cleave benzyl protecting groups, which may occur prior to nitro group reduction. Using these standard conditions in EtOAc, initial monitoring showed no reaction. After

2 days, starting material had been consumed but with no identifiable product (Table 11). Changing the solvent to ethanol gave the same undesired result. Hence, decomposition occurred with these conditions.

Hydrogenation in the presence of PtO_2 or Adam's catalyst, in EtOAc over 16 hours at room temperature[61] provided the product **30** in 79% yield (Scheme 86, Table 11).

^{1}H NMR spectra confirmed reduction to form **30**. All signals showed an upfield chemical shift due to the relative electron-donating effect of the amino substituent, particularly in comparison to the electron-withdrawing nitro group. Protons at C-4 and C-6 showed the most shielding; C-6 now appeared as a doublet at 6.41 ppm (7.92 ppm in the starting material **29**) and C-4 now appeared as a doublet at 6.45 ppm (8.19 ppm in **29**). Mass spectrometry $C_{18}H_{19}N_2O_5S$ (M+H)+ 375.1009 confirmed the structure.
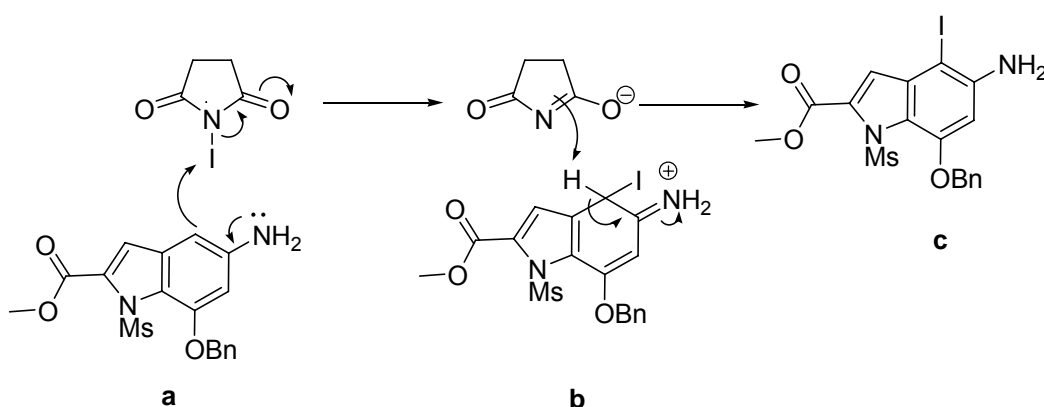
6.1.9. Iodination at C-4

Iodination at C-4 was successful using NIS at 0°C for 5 minutes[61] to provide the desired product **31** in 79% yield (Scheme 87).
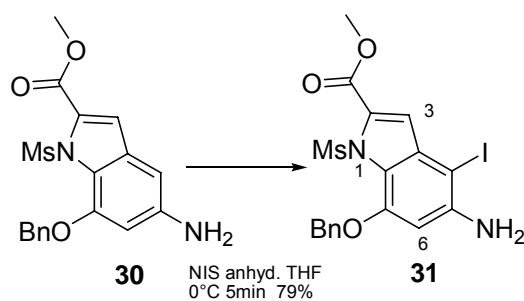
Scheme 87 Iodination at C-4 to form **31**

^{1}H NMR spectroscopy confirmed iodination. The peak for the C-4 proton was absent due to displacement by an iodo atom. The proton at C-3 showed a small long range deshielding effect due to the iodo substituent at 7.06 ppm as a singlet (7.01 ppm in the starting material **30**). Mass spectrometry $C_{18}H_{21}O_5N_3IS$ (M+NH4)+ 518.0241 confirmed the structure.

The reaction proceeds in a manner similar to that described earlier in this chapter (Scheme 75). In this case, iodination is only possible at C-4 or C-3. The presence of the amino group at C-5 directs iodination at C-4, the ether oxygen also directs to the *para* position while the benzyl group at C-7 causes steric hindrance preventing iodination at C-6. Scheme 88 outlines a plausible reaction mechanism for iodination of **30** at C-4. NIS is the source of the iodonium ion, which undergoes nucleophilic attack by C-4 of the aromatic ring (**a**). Aromaticity of the ring is maintained on the loss of the proton. This iodination proceeds without the need for NIS activation by an acid. The most likely reason for this is that NIS contains trace amounts of HI which acts as a catalyst. The substituents of **30** may also confer sufficient reactivity for nucleophilic attack to occur even in the absence of the addition of acid.

Scheme 88 Possible reaction mechanism for regioselective iodination at C-4 to form **31**

6.2. Summary

Negishi coupling was successfully achieved to form **29** and optimised using *in situ* Pd/ligand complex formation in 56% yield (Table 7), in comparable yields to Hiroya's published method (56%).[61] Scaleability and optimisation of this synthetic route has not proved to be facile. It is clear that there is great scope in investigating the critical cross-coupling step further, and progressing the synthesis overall.

Work on this synthesis by this author draws to a close here. The final steps to reach the target Boc-DSA structure would be to Boc-protect the free amine, alkylation using 1,3-dichloropropene at C-5, followed by *5-exo-trig* ring closure (Scheme 73). This would provide a core Boc-DSA structure to use as a template for further substituent modifications to generate derivatives. Key areas to vary would be at the indole amino group, and the aromatic phenol group to improve DNA binding properties. These points could also serve as handles to attach other targeting ligands such as antibodies.

7. Final Conclusion

We have optimised the route to the minimum active pharmacophore of duocarmycin SA, CI **7**, (Scheme 13). As compared to the conditions reported by Warpehoski,[56] introduction of the second step diester was improved by shorter reaction times (3 hours compared to 24 hours), and higher yields (87% compared to 77%) without recycling of the starting material. Aromatic nitro group reduction was achieved in comparable yields (65% compared to 69%) via a cost-effective Al-Hg amalgam method. The key one pot mesylation/cyclisation was achieved in 88%, a significant improvement to published yields (59%). Red-Al successfully deprotected the indoline nitrogen in 57% yield, providing the core structure for our model CI-based subunit, **7**. Thus, we have met our first aim of achieving a synthetic route to CI with greater overall efficiency.

In the first synthetic route to DSA, the alkylation subunit DSA has been the key intermediate target structure as shown in the retrosynthetic analysis in Scheme 32. Allylation was investigated for the precursor step to heterocycle formation, using direct methods (Ek methodology[89]) and indirectly using Pd catalysed coupling reactions, in addition to comparisons of solution phase versus microwave processes. Our work indicated that allylation was not a feasible precursor step to DSA for our substrate.

Fischer indole cyclisation was also explored. The precursor hydrazone **24** can be synthesised with ease, however, indole formation was low yielding, with difficulties in separation and purification.

In the second route to DSA, Hemetsberger azide methodology[106] was used to synthesise the left hand indole of our target DSA structure (Scheme 59). Swern oxidation provided the precursor aldehyde **19** in high yields (85%). One pot nitro group displacement and phenol group protection via benzylation[64] to form **20** was also achieved in good yields of 71%. Two methods were successfully explored for condensation of methyl azidoacetate **21** to the aldehyde **20** to form **22**. Subsequent indole cyclisation to form **23** using Hemetsberger's conditions was also achieved. However, low efficiency of methyl azidoacetate condensation and indole formation prevented us from pursuing this route further.

In the third route to DSA, Negishi coupling was optimised to form the indole using *in situ* Pd/ligand complex formation in 56% yield (Table 8), which is comparable to Hiroya's published method (56%).[61]

A phthalimide-based library of compounds **9**-**12** was also synthesised in good yields to introduce amide linkers (Scheme 28 & Scheme 31). The *N*-phthalimide position can be used to attach other moieties, with the aim of improving targeting, DNA binding and/or interaction properties of DSA analogues.

The third route to DSA utilising the optimised Negishi cross-coupling and cyclisation can provide the core Boc-DSA structure by a concise and scaleable route (Scheme 73). This would be used as a template for further substituent modifications. Candidate positions on the alkylation subunit structure are at the indole amino group and the aromatic phenol group to improve DNA binding properties. These points can also serve as handles to attach other targeting ligands such as antibodies.

8. Experimental

8.1. Physical Characterisation & Spectroscopic Techniques

8.1.1. NMR

Spectra were obtained using Bruker Avance 400, Varian Unity Plus 400, or Varian Gemini 300 spectrometers. Spectra were processed using Bruker Topspin & MestRe-C 4.9 software. Chemical shifts were referenced to the specified deuterated solvent. Chemical shifts recorded in parts per million (ppm). Peak splitting was reported as singlet (s), doublet (d), triplet (t), quartet (q), quintet (quin), septet (sept), multiplet (m), broad (br), doublet of doublets (dd). Coupling constants were reported in Hertz (Hz).

8.1.2. Mass Spectra

Low resolution mass spectra were obtained by electrospray analysis in positive or negative modes using Quadrople Mass Spectrometer, or Shimadzu LCMS 2010EV. Accurate mass spectra were obtained from EPSRC National Mass Spectrometry Service Centre (Swansea, UK).

8.1.3. Infra Red

Spectra were obtained from neat samples using Nicolet Smart Golden Gate Spectrometer (Avatar 360 FT-IR E.S.P.), and Perkin-Elmer Spectrum BX FT-IR. Spectra were processed by Spectrum v5.0 software.

8.1.4. Melting Point

Melting points were obtained using Bibby Stuart Scientific SMP3, and Mel-Temp electrothermal melting point apparatus.

8.2. Chromatographic Techniques

8.2.1. TLC

Merck precoated silica gel 60 F_{254} TLC plates were used. Flash chromatography for chromatographic separation/purifications: silica gel (particle size 40-63 μm, BDH Laboratory Supplies, Poole, Dorset).

8.3. Reagent, solvent and apparatus preparation

All chemicals used were reagent grade. All glassware was oven dried prior to use. Solvents where specified as anhydrous bought as such, and assumed to conform to manufacturing standards. All water used was distilled.

8.3.1. Microwave

Microwave reactions were carried out using Emrys Creator, Personal Chemistry microwave. 0.5-2 ml capacity Biotage microwave vials used for reactions, with aluminium crimp caps with PTFE lined rubber seals.

8.3.2. HPLC

HPLC analysis performed on Agilent 1200 Series LC. Column: 4.6 x 150 mm, 5 μm, Agilent Eclipse γ DV-C18. Mobile phase: A: 0.005% TFA in H_2O, B: 0.005% TFA in MeOH. Gradient: 5→95% B over 16 min. Flow rate: 1 ml/min. Column temperature: 40.0°C. Injection volume: 20 μl. Run time: 16 min. UV

detection λ: 254 nm. Data processed using Agilent ChemStation for LC 3D systems, rev B.02.01-SR1 [260].

8.4. Experimental procedures & characterisation

1-chloro-2-nitro-4-(benzyloyl)benzene [1][56]
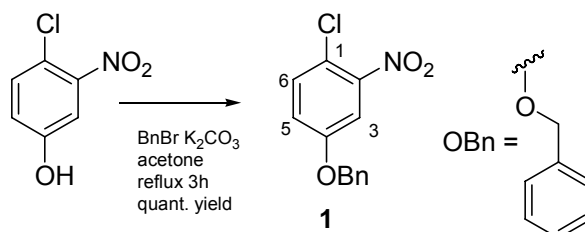
K$_2$CO$_3$ (0.62 g, 4.47 mmol) and benzyl bromide (BnBr, 0.42 ml, 3.57 mmol) were successively added to a solution of 4-chloro-3-nitrophenol (0.5 g, 2.88 mmol) in acetone (10 ml). The mixture was allowed to reflux for 3 h. After cooling, the crude mixture was filtered, and the filtrate was concentrated *in vacuo*. The resulting crude solid was purified by silica gel chromatography [20% EtOAc/hexane] to provide **1** (0.78 g, quantitative yield) as a bright yellow solid; R$_f$ 0.53 (20% EtOAc/hexane); mp 53.4-55.6°C; IR ν_{max} (film, cm^{-1}) 1525 (NO$_2$); ^{1}H-NMR (400 MHz, CDCl$_3$) δ 5.10 (2H, s, CH$_2$), 7.12 (1H, dd, *J*=9.4, 2.6 Hz, 5-H), 7.42-7.44 (6H, m, ArH + 6-H), 7.48 (1H, d, *J*=3.2 Hz, 3-H); ^{13}C-NMR (400 MHz, CDCl$_3$) δ 15.2, 71.1, 111.7, 118.7, 120.8, 127.8, 128.8, 129.0, 132.7, 135.5, 157.7; MS (ES+) calculated for C$_{13}$H$_{10}$ClNO$_3$Na (M+Na)+ 286.0247, found 286.0241.

Diethyl [4-(benzyloxy)-2-nitrophenyl] malonate [2][56]
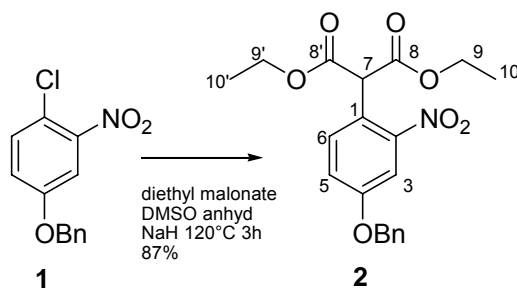
NaH (60% in mineral oil, 0.23 g, 5.69 mmol) was dispersed in anhyd. DMSO (5 ml) under N_2, and cooled in an ice bath (0-5°C approx.). Diethyl malonate (0.86 ml, 5.60 mmol) was added dropwise to this suspension (caution: H_2 evolution) and warmed to room temperature with stirring for 30 min. **1** (0.5 g, 1.90 mmol) was added and the reaction mixture stirred at 120°C for 3 h. The reaction was quenched by carefully pouring into a beaker of ice/water solution (50 ml) and stirred. Once the ice had melted, the crude mixture was extracted with DCM and water. The aqueous layer was washed four times with DCM. The combined organic layers were dried over $MgSO_4$, and concentrated *in vacuo*. The residue was purified by silica gel chromatography [10% EtOAc/hexane] to provide **2** (0.73 g, 87%) as a yellow oil; R_f 0.34 (20% EtOAc/hexane); IR ν_{max} (film, cm^{-1}) 1732 (C=O), 1506 (NO_2); ^{1}H-NMR (400 MHz, $CDCl_3$) δ 1.26 (6H, t, *J*=8.2 Hz, 2(CH_3)), 4.25 (4H, q, *J*=6.6 Hz, 2(CH_2)), 5.12 (2H, s, CH_2), 5.21 (1H, s, 1'-H), 7.23 (1H, dd, *J*=6.0, 2.4 Hz, 5-H), 7.35-7.42 (6H, m, ArH + 6-H), 7.65 (1H, d, *J*=2.8 Hz, 3-H); ^{13}C-NMR (400 MHz, $CDCl_3$) δ 14.2, 54.1, 62.4, 70.9, 111.3, 120.5, 120.6, 127.8, 128.7, 129.0, 132.4, 135.8, 149.5, 159.0, 167.8; MS (ES+) calculated for $C_{20}H_{21}NO_7$ (M+H)$^+$ 388.1391, found 388.1392.

2-[4-(benzyloxy)-2-nitrophenyl]propane-1,3-diol [3][56]
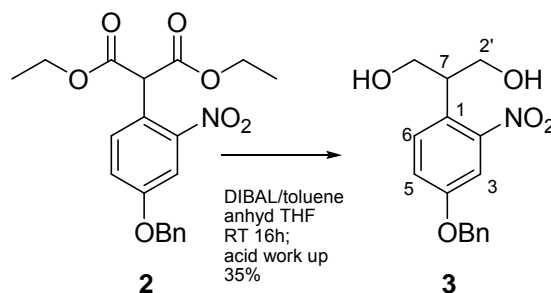
Under N_2, to a solution of diisobutylaluminium hydride (DIBAL, 1 M in toluene, 30 ml, 31 mmol) in anhyd. THF (20 ml), a solution of **2** (2 g, 5.17 mmol) was carefully added dropwise, and stirred at room temperature for 16 h. To quench the reaction, the reaction mixture was carefully and slowly poured into a beaker of 1 M HCl (150 ml), with vigorous manual stirring (caution: foaming and gas evolution). The crude mixture was extracted with EtOAc twice. The combined organic layers were dried over $MgSO_4$, and concentrated *in vacuo*. The residue was purified by silica gel chromatography [increasing gradient 50-70% EtOAc/hexane] to provide **3** (0.55 g, 35%) as a red solid; mp 101-105°C; ^{1}H-NMR (400 MHz, $CDCl_3$) δ 3.54 (1H, quin, *J*= 6.0 Hz, CH), 4.01 (4H, d, *J*=5.2 Hz, 2(CH_2)), 5.10 (2H, s, CH_2), 7.18 (1H, dd, *J*= 8.4 Hz, 2.8 Hz, 5-H), 7.31-7.46 (7H, m, ArH + 3-H + 6-H);

## 2-[2-amino-4-(benzyloxy)phenyl]propane-1,3-diol [4][56]
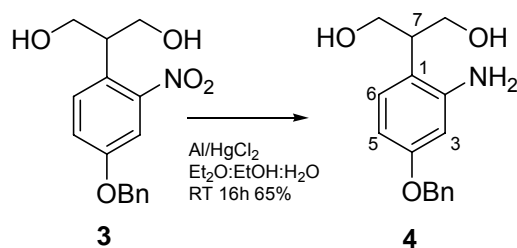
3 (0.1 g, 0.33 mmol) was dissolved in an $Et_2O/EtOH/H_2O$ (22:7:1, 10 ml) solution. Al foil (0.09 g, 3.3 mmol) was cut into strips and rolled into coils. Each coil was briefly dipped in Et_2O, immersed in 2% aq. $HgCl_2$ solution (~20 secs), briefly dipped in Et_2O again, and added to the solution containing **3**. This was stirred at room temperature for 16 h. The crude mixture was then filtered through Celite, the solid residue washed with EtOH and the filtrate concentrated *in vacuo*. EtOAc was added to the crude fuchsia solid. The resulting suspension was filtered to yield **4** (0.06 g, 65%) as a beige solid, which was used in the next step without further purification; mp 143.5-148.8°C; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 3.05-3.16 (1H, m, CH), 3.86-4.01 (4H, m, 2(CH_2)), 5.00 (2H, s, CH_2), 6.42 (1H, dd, *J*=3.0, 8.6 Hz, 5-H), 6.45 (1H, d, *J*=2.7 Hz, 3-H), 6.93 (1H, d, *J*=8.7 Hz, 6-H), 7.31-7.42 (5H, m, ArH);

6-(benzyloxy)-2,3-dihydro-1-(methylsulfonyl)-1H-indole-3-methanol,

methanesulfonate [5][56]
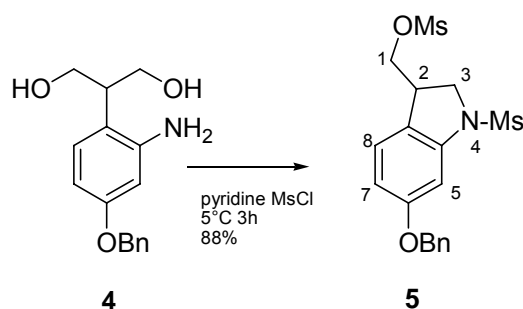
Under N_2, **4** (0.2 g, 0.73 mmol) was dissolved in pyridine (3 ml), cooled to 5°C, and MsCl (0.20 ml, 2.56 mmol) added dropwise. The reaction mixture was stirred at 5°C for 3 h. The reaction was quenched with ice, diluted with EtOAc, and extracted. The organic layer was washed with cold 3 M HCl, then 1 M HCl (until the wash remained acidic), 5% $NaHCO_3$ and sat. aq. NaCl. The combined organic layers were dried over $MgSO_4$, and concentrated *in vacuo*. The residue was purified by silica gel chromatography [80% EtOAc/hexane] to provide **5** (0.27 g, 88%) as a beige solid; mp 130.1-135.6°C; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 2.86 (3H, s, NSO_2CH_3), 2.97 (3H, s, OSO_2CH_3), 3.91-4.09 (3H, m, 1-CH_2 + 2-H), 4.16-4.23 (2H, m, 3-CH_2), 5.02 (2H, s, CH_2), 6.67 (1H, dd, *J*=2.3, 8.1 Hz, 7-H), 7.10 (1H, d, *J*=1.8 Hz, 5-H), 7.29-7.48 (6H, m, ArH + 8-H);

6-(benzyloxy)-2,3-dihydro-1-(chloro)-1H-indole-3-methanol, methanesulfonate [6]
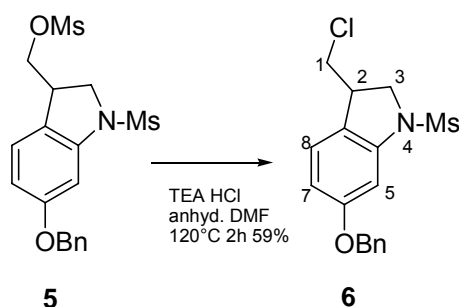
5 (0.05 g, 0.12 mM) was dissolved in anhyd. DMF (1 ml), and TEA.HCl (0.03 g, 0.24 mmol) added. This was stirred at 120°C for 2 h. The crude mixture was diluted with EtOAc, washed with water and sat. aq. NaCl. The combined organic layers were dried over $MgSO_4$, and concentrated *in vacuo* to provide 6 (0.02 g, 59%) as a cream solid, which was used in the next step without further purification; ^{1}H-NMR (400 MHz, CDCl$_3$) δ 2.85 (3H, s, NSO$_2$CH$_3$), 3.56-3.60 (1H, m, 1-CH), 3.65-3.74 (2H, m, 1-CH + 2-H), 3.95-3.99 (1H, m, 3-CH), 4.05-4.15 (1H, m, 3-CH), 5.06 (2H, s, CH$_2$), 6.67 (1H, dd, *J*=8.4, 2.4 Hz, 7-H), 7.10 (1H, d, *J*=2.4 Hz, 5-H), 7.12 (1H, d, *J*=8.4 Hz, 8-H), 7.30-7.44 (5H, m, ArH); ^{13}C-NMR (400 MHz, CDCl$_3$) δ 34.9 42.1, 46.9, 54.9, 70.5, 101.2, 110.5, 123.0, 125.8, 127.8, 128.3, 128.8, 136.7, 143.6, 160.2; MS (ES+) calculated for C$_{17}$H$_{19}$ClNO$_3$S (M+H)+ 352.0769, found 352.0765;

6-(benzyloxy)-2,3-dihydro-1-(chloro)-1H-indole [7]
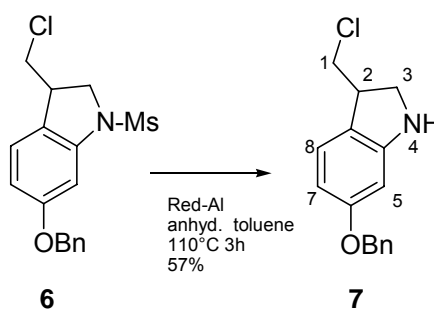
6 (0.1 g, 0.28 mmol) was dissolved in anhyd. toluene (1.4 ml), to which sodium bis(2-methoxyethoxy)aluminium hydride (Red-Al, 3.4 M in toluene, 0.42 ml, 1.42 mmol) was slowly added whilst stirring, and heated to reflux (110°C) for 3 h. The reaction mixture was cooled to 0°C by placing in an ice bath, and poured into ice cold sat. aq. NaCl. The crude reaction mixture was extracted three times with diethyl ether. The combined organic layers were dried over $MgSO_4$, and concentrated *in vacuo*. The residue was purified by silica gel chromatography [10% EtOAc/hexane] to provide **7** (0.04 g, 57%) as a brown oil; ^{1}H-NMR (400 MHz, $CDCl_3$) δ 3.12 (2H, t, *J*=8.4 Hz, 1-CH), 3.35-3.30 (1H, m, 2-CH), 3.45 (1H, m, 1-CH), 3.74-3.69 (2H, m, 3-CH), 5.02 (2H, s, CH_2), 6.32-6.39 (2H, overlapping dd, 5-H + 8-H), 6.97 (1H, dd, *J*=36.0, 8.8 Hz, 7-H), 7.37-7.44 (5H, m, ArH);

5-methoxy-indole-2-carboxylate [8]
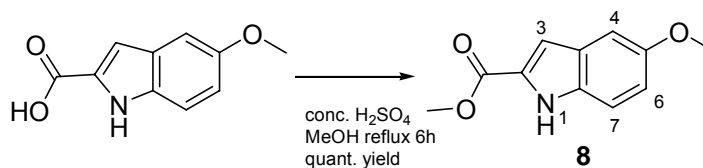
Concentrated H_2SO_4 (0.1 ml) was added to a solution of 5-methoxyindole 2-carboxylic acid (0.1 g, 0.52 mmol) in MeOH (3.4 ml). This was heated to reflux (approx 65°C) for 6 h. The reaction mixture was allowed to cool and concentrated *in vacuo*. The residue was dissolved in DCM (20 ml), washed twice with sat. aq. $NaHCO_3$ (10 ml), and the organic layers dried over $MgSO_4$, and concentrated *in vacuo* to yield **8** (0.13 g, quant. yield) as a cream solid, which was used in the next step without further purification. R_f 0.48 [20% EtOAc/hexane]; mp 176.1-182.1°C; IR ν_{max} (film, cm^{-1}) 3327 (NH), 1681 (C=O); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 3.85 (3H, s, 5-OCH$_3$), 3.94 (3H, s, CH$_3$), 7.00 (1H, dd, *J*=8.8, 2.4 Hz, 6-H), 7.08 (1H, s, 3-H), 7.14 (1H, m, 4-H), 7.32 (1H, d, *J*=8.8 Hz, 7-H), 8.92 (1H, br s, NH); ^{13}C NMR ((CD$_3$)$_2$CO, 400 MHz) δ 51.2, 55.0, 102.2, 107.7, 113.4, 116.8, 127.9, 128.0, 133.1, 155.0, 162.0; MS (ES+) calculated for C$_{11}$H$_{11}$O$_3$N (M)$^+$ 205.0733, found 205.0732;

**Methyl 5-methoxy-1-(4-phthalimidopropyl) indole-2-carboxylate [9]**
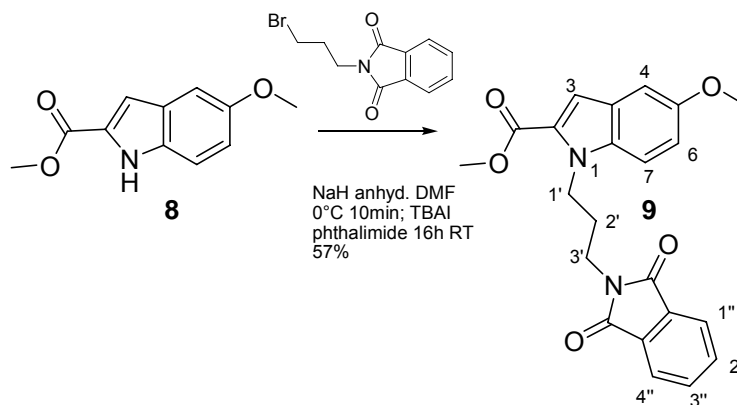
NaH (60% in mineral oil, 0.03 g, 0.73 mmol) was dispersed in anhyd. DMF (1 ml) at 0°C. **8** (0.1 g, 0.49 mmol) was added to this suspension and stirred for 10 min. Tetra-*N*-butyl ammonium iodide (TBAI, 0.05 g, 0.14 mmol) and 3-bromopropyl phthalimide (0.17 g, 0.60 mmol) were added, and the reaction mixture left to stir at room temperature for 16 h. The mixture was then poured into H_2O (10 ml), and the aqueous layer extracted three times with EtOAc. The combined organic layers were washed with H_2O, sat. aq. NaCl, dried over $MgSO_4$, and concentrated *in vacuo*. The residue was purified by silica gel chromatography [20% EtOAc/hexane] to provide **9** (0.11 g, 57%) as a yellow solid; R_f 0.33 [20% EtOAc/hexane]; mp 149.9-152.8°C; IR ν_{max} (film, cm^{-1}) 1704 (C=O); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 2.20 (2H, quin, *J*=7.2 Hz, 2'-CH$_2$), 3.77 (2H, t, *J*=7.2 Hz, 1'-CH$_2$), 3.83 (3H, s, 5-OCH$_3$), 3.84 (3H, s, CH$_3$), 4.61 (2H, t, *J*=7.6 Hz, 3'-CH$_2$), 6.99 (1H, dd, *J*=9.2, 2.4 Hz, 6-H), 7.03 (1H, d, *J*=2.4 Hz, 4-H), 7.18 (1H, s, 3-H), 7.28 (1H, overlapping CDCl$_3$, 7-H), 7.72 (2H, m, 2''-H + 3''-H), 7.84 (2H, m, 1''-H + 4''-H); ^{13}C NMR ((CD$_3$)$_2$CO, 400 MHz) δ 35.57, 42.4, 51.1, 55.1, 102.5, 110.0, 111.7, 116.8, 123.0, 126.5, 132.6, 134.2, 134.7, 155.1, 162.0, 168.1; MS (ES+) calculated for C$_{22}$H$_{24}$O$_5$N$_3$ (M+ NH$_4$)$^+$ 410.1710, found 410.1714;

**Methyl 5-methoxy-1-(4-phthalimidobutyl)indole-2-carboxylate [10]**
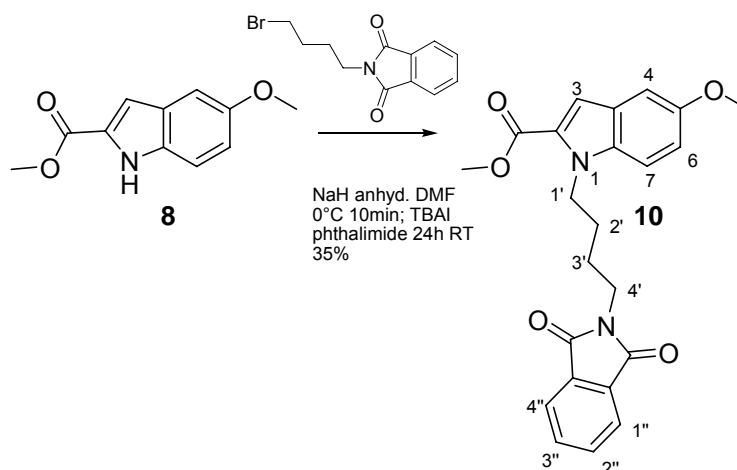
NaH (60% in mineral oil, 0.03 g, 0.73 mmol) was dispersed in anhyd. DMF (1 ml) at 0°C. 8 (0.1 g, 0.49 mmol) was added to this suspension and stirred for 10 min. Tetra-n-butyl ammonium iodide (TBAI, 0.05 g, 0.14 mmol) and 3-bromobutyl phthalimide (0.18 g, 0.63 mmol) were added, and the reaction mixture left to stir at room temperature for 24 h. The mixture was then poured into H_2O (10 ml), and the aqueous layer extracted three times with EtOAc. The combined organic layers were washed with H_2O, sat. aq. NaCl, dried over $MgSO_4$, and concentrated *in vacuo*. The residue was purified by silica gel chromatography [20% EtOAc/hexane] to provide **10** (0.07 g, 35%) as a cream solid; R_f 0.33 [20% EtOAc/hexane]; mp 131.8-134.4°C; IR ν_{max} (film, cm^{-1}) 1698 (C=O); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.73 (2H, quin, *J*=3.2 Hz, 2'-CH$_2$), 1.85 (2H, quin, *J*=6.8 Hz, 3'-CH$_2$), 3.71 (2H, t, *J*=7.0 Hz, 1'-CH$_2$), 3.84 (3H, s, 5-OCH$_3$), 3.89 (3H, s, CH$_3$), 4.58 (2H, t, *J*=7.4 Hz, 4'-CH$_2$), 7.00 (1H, dd, *J*=9.0, 2.6 Hz, 6-H), 7.04 (1H, d, *J*=2.4 Hz, 4-H), 7.20 (1H, d, *J*=0.8 Hz, 3-H), 7.32 (1H, overlapping CDCl$_3$, 7-H), 7.71 (2H, m, 2''-H + 3''H), 7.83 (2H, m, 1''-H + 4''-H); ^{13}C NMR ((CD$_3$)$_2$CO, 400 MHz) δ 25.9 28.1, 37.4, 44.1, 51.1, 55.1, 102.4, 109.9, 111.9, 116.7, 123.0, 132.4, 134.2, 162.1, 168.1; MS (ES+) calculated for C$_{23}$H$_{22}$O$_5$N$_2$ (M)$^+$ 406.1523, found 406.1511

Ethyl 5-methoxy-1-(4-phthalimidopropyl)indole-2-carboxylate [11]
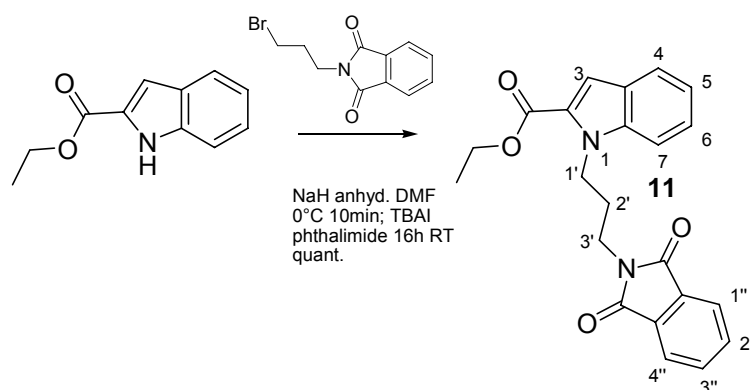
NaH (60% in mineral oil, 0.95 g, 24 mmol) was dispersed in anhyd. DMF (15 ml) at 0°C. Ethyl indole-2-carboxylate (3 g, 16 mmol) was added to this suspension and stirred for 10 min. Tetra-*N*-butyl ammonium iodide (TBAI, 1.46 g, 3.96 mmol) and 3-bromopropyl phthalimide (5.53 g, 21 mmol) were added, and the reaction mixture left to stir at room temperature for 16 h. The mixture was then poured into H_2O, and the aqueous layer extracted three times with EtOAc. The combined organic layers were washed with H_2O, sat. aq. NaCl, dried over $MgSO_4$, and concentrated *in vacuo*. The residue was purified by silica gel chromatography [20% EtOAc/hexane] to provide **11** (6.12 g, quant. yield) as a cream solid; R_f 0.42 [20% EtOAc/hexane]; mp 106.6-110.4°C; IR ν_{max} (film, cm^{-1}) 1700 (C=O); ^1H NMR (CDCl$_3$, 400 MHz) δ 1.33 (3H, t, *J*=7.0 Hz, CH$_3$), 2.19 (2H, quin, *J*=7.2 Hz, 2'-CH$_2$), 3.76 (3H, t, *J*=7.2 Hz, 1'-CH$_2$), 4.26 (2H, q, *J*=7.2 Hz, CH$_2$) 4.62 (2H, t, *J*=7.6 Hz, 3'-CH$_2$), 7.10 (1H, ddd, *J*=8.0, 6.8, 1.6 Hz, 5-H), 7.26 (1H, s, 3-H), 7.30 (1H, dd, *J*=6.8, 1.2 Hz, 6-H), 7.33 (1H, d, *J*=8.2 Hz, 4-H), 7.62 (1H, d, *J*=8.4 Hz, 7-H), 7.69 (2H, m, 2''-H + 3''-H), 7.82 (2H, m, 1''-H + 4''-H); ^{13}C NMR ((CD$_3$)$_2$CO, 400 MHz) δ 25.9, 28.1, 37.4, 44.1, 51.1, 55.1, 102.4, 109.9, 111.9, 116.7, 123.0, 126.4, 132.4, 134.2, 162.1, 168.1; MS (ES+) calculated for C$_{22}$H$_{24}$O$_4$N$_3$ (M + NH$_4$)$^+$ 394.1761, found 394.1764.

**Ethyl 5-methoxy-1-(4-phthalimidobutyl)indole-2-carboxylate [12]**
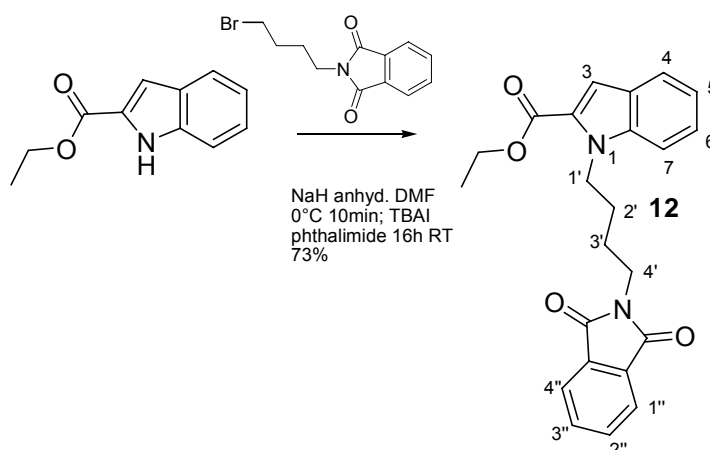
NaH (60% in mineral oil, 0.73 g, 18.2 mmol) was dispersed in anhyd. DMF (10 ml) at 0°C. Ethyl indole-2-carboxylate (2.3 g, 12.2 mmol) was added to this suspension and stirred for 10 min. Tetra-*N*-butyl ammonium iodide (TBAI, 1.12 g, 3.0 mmol) and 3-bromobutyl phthalimide (4.46 g, 15.8 mmol) were added, and the reaction mixture left to stir at room temperature for 16 h. The mixture was then poured into H_2O, and the aqueous layer extracted three times with EtOAc. The combined organic layers were washed with H_2O, sat. aq. NaCl, dried over $MgSO_4$, and concentrated *in vacuo*. The residue was loaded onto silica and purified by silica gel chromatography [20% EtOAc/hexane] to provide **12** (3.44 g, 73%) as a cream solid; R_f 0.45 [20% EtOAc/hexane]; mp 120.8-126.4°C; IR ν_{max} (film, cm^{-1}) 1698 (C=O); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.37 (3H, t, *J*=7.0 Hz, CH$_3$), 1.71 (2H, quin, *J*=7.2 Hz, 2'-CH$_2$), 1.84 (2H, quin, *J*=6.8 Hz, 3'-CH$_2$), 3.68 (2H, t, *J*=7.0 Hz, 1'-CH$_2$), 4.32 (2H, q, *J*=7.0 Hz, CH$_2$), 4.58 (2H, t, *J*=7.2 Hz, 4'-CH$_2$), 7.09 (1H, ddd, *J*=8.4, 6.8, 1.5 Hz, 5-H), 7.27 (1H, s, 3-H), 7.30 (1H, dd, *J*=6.8, 1.2 Hz, 6-H), 7.37 (1H, d, *J*=8.8 Hz, 4-H), 7.63 (1H, d, *J*=8.0 Hz, 7-H), 7.67 (2H, m, 2''-H + 3''-H), 7.80 (2H, m, 1''-H + 4''-H); ^{13}C NMR ((CD$_3$)$_2$CO, 400 MHz) δ 13.9, 25.9, 28.1, 37.4, 44.0, 60.4, 110.4, 110.9, 120.7, 122.6, 123.0, 125.0, 126.2, 127.6, 132.4, 134.2, 139.3, 162.1,

168.1; MS (ES+) calculated for $C_{23}H_{26}O_4N_3$ (M + NH_4)$^+$ 408.1918, found 408.1917.

## 4-benzyloxy-2-nitro-phenylamine [13]
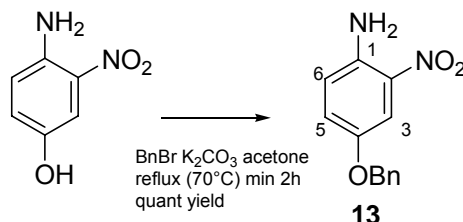
4-amino-3-nitrophenol (5 g, 30 mmol) was dissolved in acetone (approx. 50 ml). K_2CO_3 (22.4 g, 160 mmol) and benzyl bromide (BnBr, 4.63 ml, 40 mmol) were added. The reaction mixture was stirred and heated to reflux until the reaction was complete by TLC (minimum 2 h). The solution was filtered and the solvent removed from the filtrate *in vacuo*. The product **13** was recrystallised from hexane/EtOAc as a red solid (5.23 g, quant.): R_f 0.60 [30% EtOAc/hexane]; mp 135.0-145.1°C; IR ν_{max} (film, cm^{-1}) 3473 (NH_2), 3349 (NH_2), 1506 (NO_2), 1202 (C-O); ^1H NMR (CDCl$_3$, 400 MHz) δ 5.04 (2H, s, CH_2), 5.90 (2H, br s, NH_2), 6.78 (1H, d, *J*=8.8 Hz, 6-H), 7.15 (1H, dd, *J*=9.0, 3.0 Hz, 5-H), 7.33-7.46 (5H, m, ArH), 7.67 (1H, d, *J*=2.8 Hz, 3-H); ^{13}C NMR (CDCl$_3$, 400 MHz) δ 70.9, 107.9, 115.8, 120.3, 127.2, 127.5, 127.9, 128.4, 128.9, 136.5, 140.1; MS (ES-) calculated for $C_{13}H_{11}O_3N_2$ (M)$^-$ 243.0775, found 243.0776.

## 4-benzyloxy-2-bromo-6-nitro-phenylamine [14]
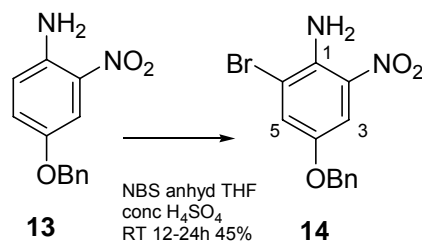
13 (7 g, 29 mmol) was dissolved in 50 ml anhyd. THF. A few drops (approx 2 ml) of conc. H_2SO_4 were added to the solution. *N*-bromosuccinimide (NBS, 5.11 g, 29 mmol) was dissolved in anhyd. THF (40 ml, with slight heating to aid dissolution) then added dropwise. The reaction mixture was left to stir at room temperature for 2 h. Further NBS (1.28 g, 7 mmol), dissolved in anhyd. THF was added and the solution left to stir at room temperature (12-24h) until the reaction was complete by TLC. The solvent was concentrated *in vacuo* and the crude product purified by silica gel chromatography [10% EtOAc/hexane] to afford **14** as (3.66 g, 45%) an orange-red solid: R_f 0.80 [10% EtOAc/hexane]; mp 116.9-125.3°C; IR v_{max} (film, cm^{-1}) 3477 (NH_2), 3366 (NH_2), 1502 (NO_2), 1328 (NO_2), 1247, 1194 (C-O), 1023 (C-Br); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 5.04 (2H, s, CH_2), 6.40 (2H, br s, NH_2), 7.36-7.42 (5H, m, ArH), 7.19 (1H, d, *J*=2.9 Hz, 3-H), 7.73 (1H, d, *J*=3.0 Hz, 5-H); ^{13}C NMR (CDCl$_3$, 400 MHz) δ 71.2, 94.6, 109.0, 113.0, 127.9, 128.6, 128.9, 130.0, 136.0, 138.0, 148.9; MS (ES-) calculated for $C_{13}H_{10}O_3N_2$ ^{79}Br (M)$^-$ 320.9880, found 320.9880

Methanesulfonic acid 4-amino-3-nitro-phenyl ester [15]
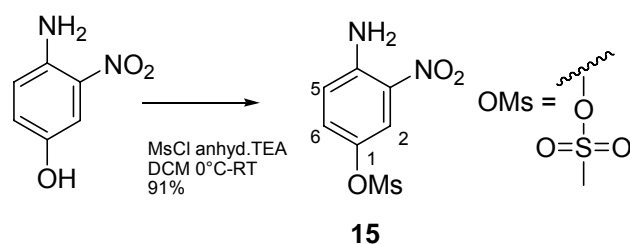
4-amino-3-nitrophenol (2 g, 13 mmol) was suspended in DCM (20 ml, yellow solution-only partial dissolution). This was cooled to 0°C and anhyd. TEA (1.90 ml, 1.38 g, 14 mmol) added (dark orange-brown solution formed). Mesyl chloride (MsCl, 1.06 ml, 14 mmol) was added dropwise whilst maintaining the temperature at 0°C (caution: exothermic). The resulting solution was left to stir at room temperature overnight. Further anhyd. TEA (0.47 ml, 3.4 mmol) was added and the solution left to stir at room temperature until the reaction was complete by TLC. The reaction mixture was loaded onto silica and the solvent was concentrated *in vacuo*. The crude product was purified by silica gel chromatography [10% MeOH/DCM] to afford **15** (2.93 g, 91%) as a yellow solid. R_f 0.40 [10% MeOH/DCM]; mp 146.4-160.8°C; IR ν_{max} (film, cm^{-1}) 3493 (NH$_2$), 3375 (NH$_2$), 1510 (NO$_2$), 1340 (NO$_2$), 1152 (S=O); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 3.19 (3H, s, CH$_3$), 6.16 (2H, br s, NH$_2$), 6.86 (1H, d, *J*=8.4 Hz, 5-H), 7.37 (1H, dd, *J*=9.1, 2.8 Hz, 6-H), 8.05 (1H, d, *J*=2.8 Hz, 2-H); ^{13}C NMR (CD$_3$)$_2$CO, 400 MHz) δ 36.5, 118.9, 120.6, 130.9, 138.5, 145.2, 162.1; MS (ES-) calculated for C$_7$H$_7$O$_5$N$_2$S (M)$^-$ 231.0081, found 231.0082.

Methanesulfonic acid 4-amino-3-iodo-5-nitro-phenyl ester [16]
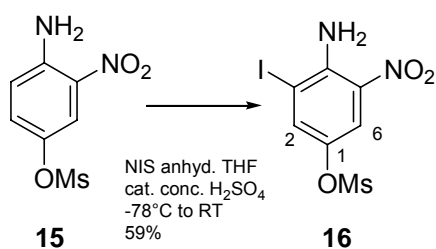
15 (0.9 g, 3.6 mmol) was dissolved in anhyd. THF (10 ml) and cooled to -78°C. Concentrated H_2SO_4 (4 drops) was added, followed by dropwise addition of NIS (0.9 g, 4.4 mmol) dissolved in anhyd. THF (40 ml). The reaction mixture was left to stir at -78°C for 5 mins, then at room temperature for 12 h. Further NIS solution (0.23 g, 1.0 mmol, in 10 ml anhyd. THF) was added dropwise; the reaction was complete by TLC after 30 mins. The solvent was concentrated *in vacuo* and the crude product purified by silica gel chromatography [0.5% MeOH/DCM] to afford **16** (0.81 g, 59%) as a solid; R_f 0.66 [1% MeOH/DCM]; mp 143.7-145.6°C; IR ν_{max} (film, cm^{-1}) 3489 (NH_2), 3371 (NH_2), 1510 (NO_2), 1359 (NO_2), 1152 (S=O); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 3.14 (3H, s, CH$_3$), 6.68 (2H, br s, NH$_2$), 7.89 (1H, d, *J*=2.8 Hz, 2-H), 8.10 (1H, d, *J*=2.8 Hz, 6-H); ^{13}C NMR ((CD$_3$)$_2$CO, 400 MHz) δ 36.6, 118.9, 120.6, 130.8, 138.5, 145.2, 162.1;

Methanesulfonic acid 4-amino-3-bromo-5-nitro-phenyl ester [17]
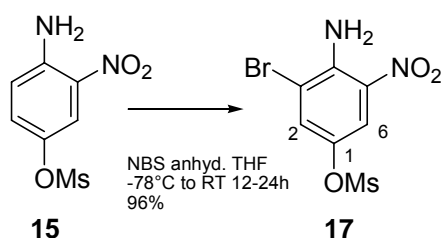
15 (1 g, 4.0 mmol) was dissolved in anhyd. THF (12 ml) and cooled to -78°C. NBS (0.79 g, 4.4 mmol) was dissolved in anhyd. THF (40 ml, required slight heating to aid dissolution), and then added dropwise. The reaction mixture was left to stir at room temperature for 16 h. Further NBS (0.18 g, 1.0 mmol, dissolved in anhyd. THF) was added dropwise and the reaction mixture left to stir, until the reaction was complete by TLC. The solvent was concentrated *in vacuo* and the crude product purified by silica gel chromatography [10% MeOH/DCM] to afford 17 (1.32 g, 96%) as a solid. R_f 0.83 [10% MeOH/DCM]; mp 126.7-140.5°C; IR ν_{max} (film, cm^{-1}) 3481 (NH$_2$), 3363 (NH$_2$), 1504 (NO$_2$), 1353 (NO$_2$), 1164 (S=O), 1073 (C-Br); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 3.23 (3H, s, CH$_3$), 6.72 (2H, br s, NH$_2$), 7.76 (1H, d, *J*=2.6 Hz, 2-H), 8.11 (1H, d, *J*=2.8 Hz, 6-H); ^{13}C NMR (CDCl$_3$, 400 MHz) δ 37.9, 94.6, 112.6, 119.5, 133.9, 137.3, 141.8; MS (ES-) calculated for C$_7$H$_6$O$_5$N$_2$^{79}BrS (M)$^-$ 308.9186, found 308.9190.

## Methanesulfonic acid 3-bromo-4-iodo-5-nitro-phenyl ester [18]
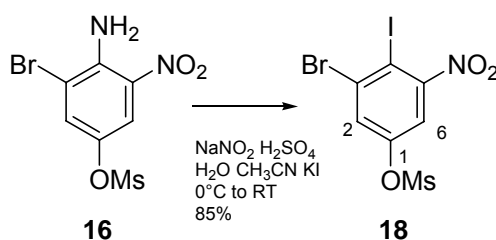
A round bottom flask was filled with 1 g ice and 1 ml conc. H_2SO_4, and stirred at 0°C. **16** (1 g, 3.1 mmol) was dissolved in anhyd. CH_3CN (3 ml) and added dropwise to the ice/H_2SO_4 mixture. $NaNO_2$ (0.38 g, 5.5 mmol) was dissolved in cold H_2O (5 ml) and added dropwise. This was left to stir at 0°C (approx. 15 min, until the reaction mixture turned from a suspension to a clear yellow solution), and then allowed to warm to room temperature. The reaction mixture was poured into a solution of KI (1.77 g, 10.6 mmol) dissolved in 1 ml H_2O, and stirred vigorously at room temperature for 16 h until the reaction was complete by TLC. The reaction was diluted with chloroform and washed with H_2O twice. The organic layers were then washed with sat. aq. $NaHCO_3$, sat. aq. sodium thiosulphate and sat. aq. NaCl consecutively, dried with $MgSO_4$, filtered and the solvent was concentrated *in vacuo*. The crude oil was purified with silica gel chromatography [10% MeOH/DCM] to afford **18** (1.13 g, 85%) as a yellow solid; R_f 0.88 [1% MeOH/DCM]; mp 126.6-135.6°C; IR v_{max} (film, cm^1) 1532 (NO_2), 1348 (NO_2), 1162 (S=O), 954 (C-Br); ^1H NMR ($CDCl_3$, 400 MHz) δ 3.20 (3H, s, CH_3), 7.45 (1H, d, J=2.8 Hz, 6-H), 7.74 (1H, d, J=2.8 Hz, 2-H); ^{13}C NMR ((CD_3)$_2$CO, 400 MHz) δ 37.5, 93.3, 117.6, 129.4, 133.3, 150.1, 162.1; MS (ES-) calculated for $C_7H_4O_5N^{79}BrIS$ (M)$^-$ 419.8044, found 419.8045

3,5-dinitrobenzaldehyde [19]
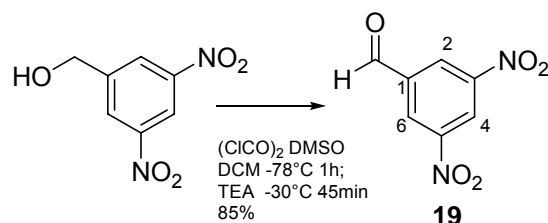
DMSO (10.75 ml, 151 mmol) with DCM (~2 ml) was cooled to -78°C. At -78°C, (ClCO)$_2$ (6.57 ml, 75 mmol) was added and the reaction mixture stirred for 10 min. 3, 5 dinitrobenzyl alcohol (10 g, 50 mmol) was dissolved in DCM (approx. 40 ml) and added dropwise. The reaction mixture was then stirred for 1 h. TEA (35 ml, 25 mmol) was added dropwise and stirred for 15 min at -78°C and at -30°C thereafter. After 1 h, further (ClCO)$_2$ (6.57 ml, 75 mmol) was added and the reaction mixture continued to stir. After 30 min, the reaction mixture was extracted at room temperature with DCM/H$_2$O, the organic layers were dried over MgSO$_4$, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography [30% EtOAc/hexane] to afford **19** (8.45 g, 85%) as a cream solid; mp 82.0-84.0°C; R$_f$ 0.27 [20% EtOAc/hexane]; IR v_{max} (film, cm^{-1}) 1701 (C=O), 1535 (NO$_2$), 1342 (NO$_2$); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 8.97 (2H, d, *J*=2.0 Hz, 2-H + 6-H), 9.23 (1H, t, *J*=2.0 Hz, 4-H), 10.18 (1H, s, CHO); ^{13}C NMR (CDCl$_3$, 400 MHz) δ 123.5, 128.9, 138.7, 149.4, 187.4; MS (ES+) calculated for C$_7$H$_4$O$_5$N$_2$ (M)$^+$ 195.0036, found 195.0039

3-(benzyloxy)-5-nitrobenzaldehyde [20][65]
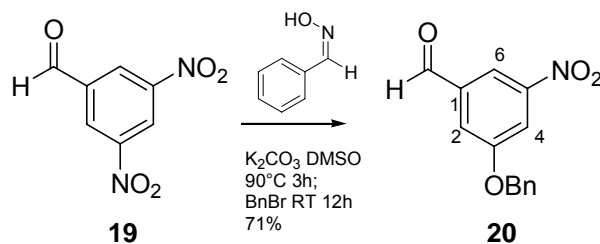
To 5 ml DMSO, benzaldehyde oxime (92.6 mg, 0.76 mmol) and K_2CO_3 (0.21 g, 1.53 mmol) were added, and heated to 90°C. **19** (0.1 g, 0.51 mmol) was dissolved in DMSO, added to the reaction mixture (dark purple solution formed), and vigorously stirred at 90°C for 3 h. The reaction mixture was allowed to cool to room temperature, BnBr (97 μl, 0.82 mmol) added and left to stir for 12 h. The reaction mixture was diluted with EtOAc, and extracted with 5% HCl, H_2O, sat. aq. NaHCO$_3$, and sat. aq. NaCl. The combined organic layers were dried over MgSO$_4$ and concentrated *in vacuo*. The crude product was purified using silica gel chromatography [20% EtOAc/hexane] to afford **20** (93.5 mg, 71%) as a pale yellow solid; mp 67.1-69.9°C; R$_f$ 0.61 [50% DCM/hexane); IR ν_{max} (film, cm^{-1}) 1705 (C=O), 1529 (NO$_2$), 1347 (NO$_2$), 1276 (C-O-C); ^{1}H NMR (CDCl$_3$, 300 MHz) δ 5.19 (2H, s, CH$_2$), 7.31-7.44 (5H, m, ArH), 7.77 (1H, dd, *J*=2.4, 1.8 Hz, 2-H), 8.04 (1H, dd, *J*=4.0, 2.0 Hz, 4-H), 8.28 (1H, dd, *J*=2.0, 1.2 Hz, 6-H), 10.01 (1H, s, CHO); ^{13}C NMR (CDCl$_3$, 400 MHz) δ 71.3, 115.6, 117.5, 120.3, 127.9, 128.9, 129.1, 135.2, 138.4, 160.0, 189.8; MS (ES-) calculated for C$_{14}$H$_{10}$O$_4$N (M)$^-$ 256.0615, found 256.0615

Methyl azidoacetate [21][119]
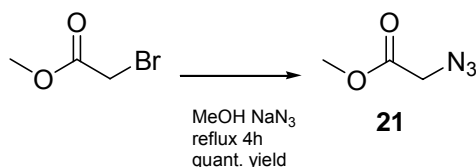
Methyl bromoacetate (3.09 ml, 33 mmol) was dissolved in MeOH (5 ml). NaN$_3$ (2.55 g, 39 mmol) was added with stirring and the mixture allowed to reflux (65-68°C) for 4 h. After cooling, MeOH present in the reaction mixture was removed *in vacuo*. H$_2$O was added, and the solution extracted with diethyl ether (4x50 ml). The combined organic layers were dried over MgSO$_4$, filtered and the solvent concentrated *in vacuo* carefully (danger of distillation as low boiling point) to afford **21** (quant. yield) as a pale yellow liquid, which was used in the next step without further purification; IR ν_{max} (film, cm^{-1}) 2101 (N$_3$) 1736 (C=O) cm^{-1}; ^{1}H NMR (CDCl$_3$, 400 MHz) δ 3.64 (3H, s, CH$_3$), 3.75 (2H, s, CH$_2$);

### Methyl 2-azido-3-(3-benzyloxy-5-nitrophenyl)acrylate [22][65]
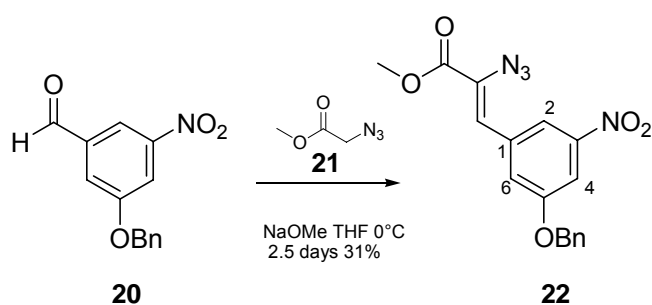
NaOMe was freshly prepared (25%w/v, 0.04 g Na in 1 ml MeOH, 1.56 mmol) and stirred at room temperature until Na was fully dissolved, then cooled to -8°C. **20** (0.1 g, 0.39 mmol) and **21** (0.18 g, 1.56 mmol) were dissolved in 1 ml MeOH, and added dropwise to the cold NaOMe solution. The reaction

mixture was left to stir at -8°C for 3 h, then below 5°C for 48 h. The reaction mixture was then poured over ice (2 g) and stirred manually. The resulting suspension was filtered and washed with H_2O. This gave **22** (42.4 mg, 31% yield) as a white solid; R_f 0.80 (50% DCM/hexane); mp 132.8-138.2°C; ^{1}H NMR (CDCl$_3$, 300 MHz) δ 3.93 (3H, s, CH$_3$), 5.19 (2H, s, CH$_2$), 6.84 (1H, s, CH), 7.36-7.47 (5H, m, ArH), 7.78 (1H, t, *J*=2.3 Hz, 6-H), 7.81 (1H, m, 2-H), 8.20 (1H, m, 4-H); ^{13}C NMR (CDCl$_3$, 300 MHz) δ 52.7, 70.9, 113.8, 116.8, 121.9, 127.1, 127.6, 128.6, 128.8, 132.7, 135.3, 149.3, 159.3, 164.9.

## Methyl 7-benzyloxy-5-nitroindole-2-carboxylate [23][65]
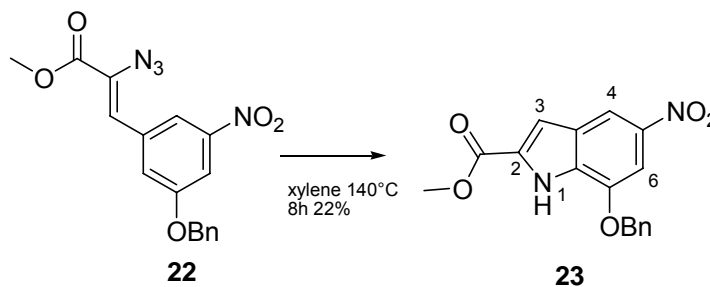
22 (0.1 g, 0.28 mmol) was dissolved in xylene (5 ml) and added dropwise to xylene (5 ml) heated to reflux (140°C) for 8 h. The solvent was removed *in vacuo*. The crude mixture was loaded onto silica and purified by silica gel chromatography [10-20% EtOAc-hexane] to provide **23** (20.7 mg, 22%) as a yellow solid; ^{1}H NMR (CDCl$_3$, 400 MHz) δ 3.94 (3H, s, CH$_3$), 5.26 (2H, s, CH$_2$), 7.32 (1H, d, *J*=2.0 Hz, 3-H), 7.39-7.49 (5H, m, ArH), 7.69 (1H, s, 6-H), 8.33 (1H, s, 4-H), 9.31 (1H, br s, NH).

## Methyl 2-[(2-methoxy-4-nitrophenyl)hydrazono]propanoate [24][62]
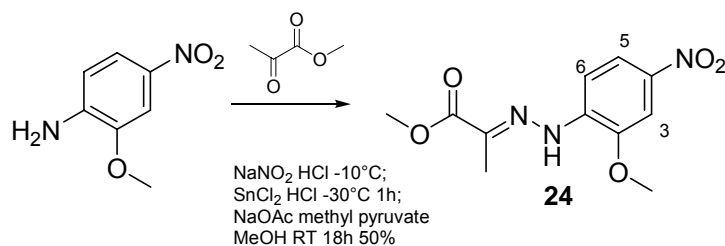
A suspension of 2-methoxy-4-nitroaniline (3.5 g, 20.8 mmol) in conc. HCl (105 ml) was cooled to -10°C (peach colour). A solution of $NaNO_2$ (1.51 g, 20.8 mmol) in H_2O (243 ml) was added dropwise to the suspension. The resulting yellow solution was added to a Sn(II)chloride (14 g, 62.4 mmol) solution in conc. HCl (70 ml) and stirred for 1 h at -30°C. The reaction mixture was allowed to warm to room temperature and the precipitate filtered. The filtrate was used to wash the precipitate. The precipitate was then dissolved in a minimal amount of MeOH, and NaOAc (2.08 g, 25.0 mmol) and methyl pyruvate (1.82 ml, 20.8 mmol) added. The reaction mixture was stirred vigorously at room temperature for 18 h. The resulting yellow precipitate was filtered and dried *in vacuo* to provide **24** (2.76 g, 50%) as a yellow solid; R_f 0.27 [30% EtOAc/hexane]; mp 163.7-167.0°C; IR ν_{max} (film, cm^{-1}) 3327 (NH$_2$), 1713 (C=O), 1575 (NO$_2$), 1338 (NO$_2$), 1092 (C-O-C), 1023 (N-N); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 2.15 (3H, s, CH$_3$), 3.86 (3H, s, 2-OCH$_3$), 3.98 (3H, s, OCH$_3$), 7.59 (1H, d, *J*=8.8 Hz, 6-H), 7.73 (1H, d, *J*=2.4 Hz, 3-H), 7.92 (1H, dd, *J*=8.8, 2.4 Hz, 5-H), 8.30 (1H, br s, NH); ^{13}C NMR ((CD$_3$)$_2$CO, 400 MHz) δ 10.5, 51.8, 56.2, 106.1, 111.9, 118.3, 138.3, 139.2, 145.8, 162.1, 165.1; MS (ES-) calculated for $C_{11}H_{12}O_5N_3$ (M)$^-$ 266.0782, found 266.0783.

Methyl 7-methoxy-5-nitro-1H-indole-2-carboxylate [24A][62]
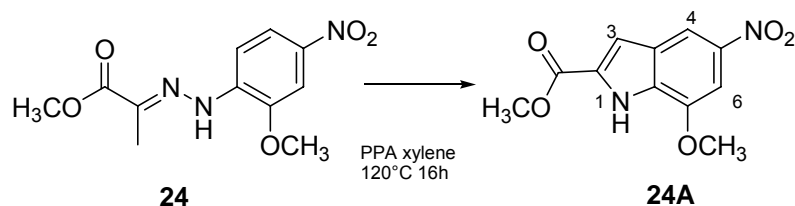
24 (0.1g, 0.38 mmol) was added to a stirred suspension of polyphosphoric acid (PPA, 0.15g) in xylene (20ml), and heated at 120°C for 16 h. The solvent was decanted and the residue thoroughly washed with xylene and DCM. The combined organic layers were then washed with water, dried over $MgSO_4$, and the solvents removed *in vacuo*. The residue was purified by silica gel chromatography [50% EtOAc/hexane] to provide **24A** (0.094 g, 73%) as a bright yellow solid; R_f 0.51 (50% EtOAc/hexane); mp 170.5-190.0°C; IR ν_{max} (film, cm^{-1}) 3310 (NH), 1712 (C=O), 1521 (NO$_2$), 1342 (NO$_2$), 1094 (C-O-C); ^{1}H NMR (CDCl$_3$, 400 MHz) δ 3.94 (3H, s, 2-OCH$_3$), 4.00 (3H, s, OCH$_3$) 7.27 (1H, s, 3-H), 7.54 (1H, d, *J*=9.6 Hz, 6-H), 8.27 (1H, s, 4-H), 9.29 (1H, br s, NH); ^{13}C NMR ((CD$_3$)$_2$CO, 400 MHz) δ 51.8, 56.2, 98.9,106.1, 110.9, 111.9, 112.6, 118.3, 162.0; MS (ES-) calculated for C$_{11}$H$_9$O$_5$N$_2$ (M)$^-$ 249.0517, found 249.0519.

2-Benzyloxy-4-nitroaniline [25][61]
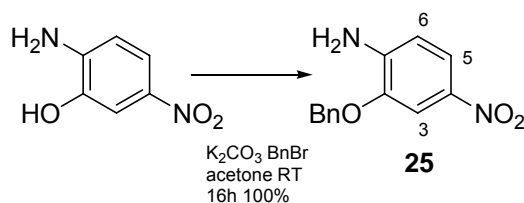
K_2CO_3 (18 g, 130 mmol) and benzyl bromide (BnBr, 7.85 ml, 77.9 mmol) were successively added to a solution of 2-amino-5-nitrophenol (10 g, 64.9 mmol) in acetone (250 ml). The mixture was stirred for 16 h at room temperature. After cooling, the crude mixture was filtered, and the filtrate was concentrated *in vacuo*. The resulting solid was recrystallised from EtOAc to provide **25** (16.6 g, 100%) as a yellow solid; R_f 0.44 (30% EtOAc/hexane); mp 149.5-153.4°C; IR ν_{max} (film, cm^{-1}) 3482 (NH$_2$), 3357 (NH$_2$), 1618 (NO$_2$), 1478 (NO$_2$), 1176 (C-O); ^{1}H-NMR (300 MHz, CDCl$_3$) δ 5.15 (2H, s, CH$_2$), 6.67 (1H, d, J=8.7 Hz, 6-H), 7.36-7.48 (5H, m, ArH), 7.78 (1H, d, J=2.4 Hz, 3-H), 7.83 (1H, dd, J=8.7, 2.4Hz, 5-H); ^{13}C-NMR (400 MHz, CDCl$_3$) δ 71.1, 107.4, 119.6, 128.1, 128.8, 129.0, 135.9, 138.7, 143.7, 144.7; MS (ES+) calculated for $C_{13}H_{12}N_2O_3$ (M+) 244.0842, found 244.0839.

2-Benzyloxy-6-iodo-4-nitroaniline [26][61]
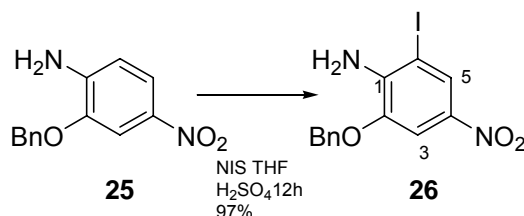
N-iodosuccinimide (NIS, 2.90 g, 12.9 mmol) and catalytic amounts of H_2SO_4 (2 drops) were successively added to a solution of **25** (3 g, 11.7 mmol) in THF (30 ml), and stirred for 16 h at room temperature. The crude mixture was

extracted with EtOAc and H$_2$O. The organic layer was dried over MgSO$_4$, and concentrated *in vacuo*. The residue was purified by silica gel chromatography [20% EtOAc/hexane] to provide **26** (4.34 g, 97%) as a yellow solid; R$_f$ 0.53 (30% EtOAc/hexane); mp 107.6-109.0°C; IR ν_{max} (film, cm^{-1}) 3475 (NH$_2$), 3355 (NH$_2$), 1602 (NO$_2$), 1494 (NO$_2$), 1275 (C-O); ^{1}H-NMR (300 MHz, CDCl$_3$) δ 5.03 (2H, br s, NH$_2$), 5.16 (2H, s, CH$_2$), 7.28-7.53 (5H, m, ArH), 7.74 (1H, d, *J*=2.4 Hz 3-H), 8.29 (1H, d, *J*=2.4 Hz, 5-H); ^{13}C-NMR (400 MHz, CDCl$_3$) δ 71.3, 106.5, 128.0, 128.1, 128.7, 128.8, 135.1, 138.6, 143.1, 143.9; MS (ES+) calculated for C$_{13}$H$_{11}$IN$_2$O$_3$ (M+) 369.9809, found 369.9811

## Dimethanesulfonyl-2-Benzyloxy-6-iodo-4-nitroaniline [27]
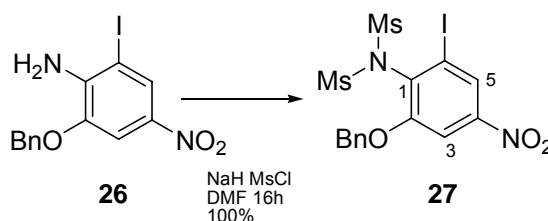
NaH (60% in mineral oil, 0.04 g, 1.09 mmol) was added slowly to a solution of **26** (0.1 g, 0.26 mmol) in anhyd. DMF (1 ml). After stirring for 1 h at room temperature, the mixture was cooled in an ice bath, and MsCl (62.8 μl, 0.79 mmol) added dropwise. The mixture was stirred for a further 16 h at room temperature. H$_2$O (2 ml) and sat. aq. Na$_2$CO$_3$ (0.3 ml) were added to the mixture. The reaction mixture was filtered, and the precipitated solid washed with H$_2$O and cold diethyl ether/hexane (3:7). This provided **27** (0.14 g, 100%) as a beige solid, which was used in the next step without further purification; R$_f$ 0.43 (30% EtOAc/hexane); mp 168.7-170.1°C; IR ν_{max} (film, cm^{-1}) 1528 (NO$_2$), 1363 (NO$_2$), 1160 (S=O); ^{1}H-NMR (400 MHz, CDCl$_3$) δ 3.42 (6H, s, 2CH$_3$), 5.20 (2H, s, CH$_2$), 7.33-7.52 (5H, m, ArH), 7.91 (1H, d, *J*=2.4 Hz, 3-H), 8.37 (1H, d, *J*=2.0 Hz, 5-H); ^{13}C-NMR (400 MHz, CDCl$_3$) δ 45.7, 72.8,

108.4, 125.8, 127.1, 128.3, 128.6, 129.1, 129.4; MS (ES+) calculated for $C_{15}H_{15}IN_2O_7S_2$ (M+) 525.9360, found 525.9363.

## N-methanesulfonyl-2-Benzyloxy-6-iodo-4-nitroaniline [28][61]
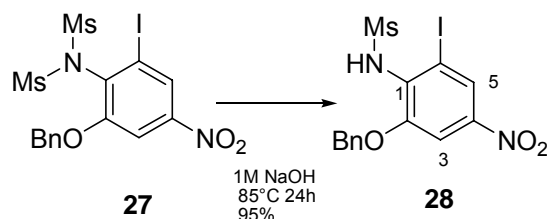
1 M NaOH (30 ml) was added to **27** (1.25 g, 2.37 mmol) and heated at 85°C for 24 h. The mixture was allowed to cool. 5 M HCl (approx 5 ml) was added dropwise until the mixture became acidic (pH 1-2). The suspension was filtered, and the precipitate washed with H_2O and cold diethyl ether. This provided **28** (1.01 g, 95%) as a cream solid, which was used in the next step without further purification; R_f 0.25 (30% EtOAc/hexane); mp 198.2-204.6°C; IR ν_{max} (film, cm^{-1}) 3254 (NH), 1523 (NO$_2$), 1324 (NO$_2$), 1148 (S=O); ^{1}H-NMR (300 MHz, CDCl$_3$) δ 3.02 (3H, s, CH$_3$), 5.20 (2H, s, CH$_2$), 6.39 (1H, br s, NH) 7.39-7.52 (5H, m, ArH), 7.91 (1H, d, J=2.4 Hz, 3-H), 8.36 (1H, d, J=2.4 Hz, 5-H); ^{13}C -NMR (400 MHz, CDCl$_3$) δ 43.7, 72.5, 94.6, 108.3, 126.8, 129.0, 129.2, 129.5, 134.1, 134.6, 152.7; MS (ES+) calculated for $C_{14}H_{17}IN_3O_5S$ (M+NH$_4$)+ 465.9925, found 465.9925.

Methyl 7-benzyloxy-1-methanesulfonyl-5-nitroindole-2-carboxylate [29][61]
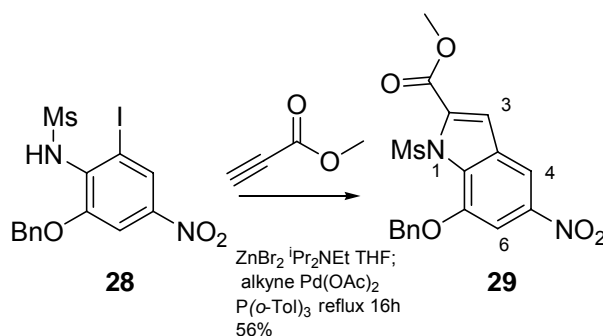
N-ethyldiisopropylamine (^{i}Pr$_2$NEt, 0.59 ml, 3.43 mmol) was added to a solution of ZnBr$_2$ (0.37 g, 1.63 mmol) in THF (6 ml), and stirred for 5 min at room temperature. Methyl propiolate (0.14 ml, 1.54 mmol), **28** (0.25 g, 0.56 mmol), tri-o-tolylphosphine (P(o-Tol)$_3$, 17 mg, 0.06 mmol) and Pd(OAc)$_2$ (3.1 mg, 0.01 mmol) were added successively, and the mixture heated at 90°C for 16 h. Saturated aq. NH$_4$Cl was added to the mixture, and extracted with EtOAc. The organic layer was washed with sat. aq. NaCl, dried over MgSO$_4$, and concentrated *in vacuo*. The residue was purified by silica gel chromatography [10% EtOAc/hexane] to provide **29** (0.13 g, 56%) as a pale yellow solid; IR ν_{max} (film, cm^{-1}) 1731 (C=O), 1601 (NO$_2$), 1336 (NO$_2$), 1108 (S=O); ^{1}H-NMR (300 MHz, CDCl$_3$) δ 3.43 (3H, s, 1-CH$_3$), 3.94 (3H, s, 2-CH$_3$), 5.30 (2H, s, CH$_2$), 7.27 (1H, s, 3-H), 7.44-7.46 (3H, m, ArH), 7.57-7.60 (2H, m, ArH), 7.92 (1H, d, J=1.8 Hz, 6-H), 8.19 (1H, d, J=1.8 Hz, 4-H); ^{13}C-NMR (400 MHz, CDCl$_3$) δ 43.1, 50.7, 53.1, 72.4, 105.2, 112.1, 117.9, 126.6, 128.7, 129.0, 129.3, 134.8; MS (ES+) calculated for C$_{18}$H$_{16}$N$_2$O$_7$S (M+) 404.0673, found 404.0675.

General method for microwave Negishi coupling reactions & HPLC analysis to form methyl 7-benzyloxy-1-methanesulfonyl-5-nitroindole-2-carboxylate [29]
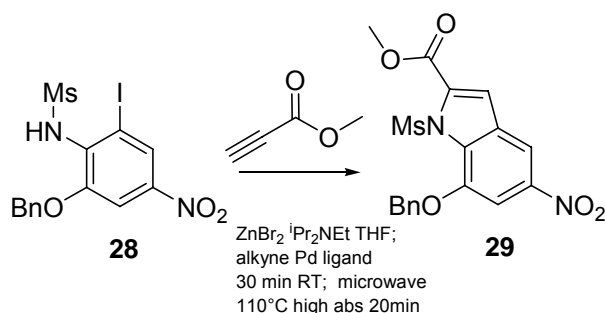

Under N_2, $ZnBr_2$ (0.15 g, 0.65 mmol) was stirred in THF (2.3 ml) at room temperature until fully dissolved in a microwave vial. $^iPr_2NEt$ (0.24 ml, 1.37 mmol) was added and stirred for 5 min at room temperature. Methyl propiolate (0.14 ml, 1.54 mmol), **28** (0.25 g, 0.56 mmol), Pd catalyst (for ready-made catalyst - Pd catalyst:starting material molar ratio 0.025:1 or 0.08:1, for *in situ* complex - Pd:ligand:starting material molar ratio 0.025:0.1:1, or 0.08:0.32:1) were added successively, and the mixture stirred at room temperature for 30 min.

The microwave vial was sealed, and the reaction mixture pre-mixed for 30 sec in the microwave, followed by irradiation at 110°C, high absorbance for 20 min. Microwave irradiation for a further 20 min was given if incomplete reaction after monitoring by TLC [20% EtOAc/hexane]. Saturated aq. NH_4Cl was added to the mixture, and extracted with EtOAc. The organic layer was washed with sat. aq. NaCl solution, dried over $MgSO_4$, and concentrated *in vacuo*.

For HPLC analysis, the crude residue was reconstituted in CH_3CN (HPLC grade) to a concentration of 1 mg/ml. HPLC analysis was performed using a

5-95% gradient of mobile phase B (0.005% TFA in MeOH) over 16 min (see "Reagent, solvent and apparatus preparation" above).

Methyl 7-benzyloxy-1-methanesulfonyl-5-(amino) indole-2-carboxylate [30][61]
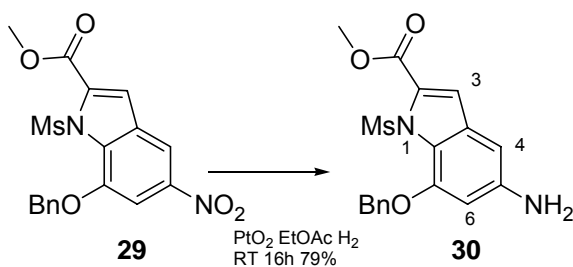
PtO$_2$ (6.1 mg, 0.03 mmol) was added to a solution of **29** (50 mg, 0.13 mmol) in EtOAc (2.5 ml). The flask was purged with H$_2$. The reaction mixture was then left to stir under H$_2$ at room temperature for 16 h. PtO$_2$ was removed by filtration of the crude reaction mixture through Celite®. The filtered bed was washed with EtOAc, and the filtrate concentrated *in vacuo*. The residue was purified by silica gel chromatography [40% EtOAc/hexane] to provide **30** (39.4 mg, 79%) as a yellow oil; ^{1}H-NMR (400 MHz, CDCl$_3$) δ 3.35 (3H, s, SO$_2$CH$_3$), 3.86 (3H, s, CO$_2$CH$_3$), 5.11 (2H, s, CH$_2$), 6.41 (1H, d, *J*=3.0 Hz, 6-H), 6.45 (1H, d, *J*=3.0 Hz, 4-H), 7.01 (1H, s, 3-H), 7.28-7.38 (3H, m, ArH), 7.48 (2H, m, ArH), 7.53 (2H, d, *J*=8.0 Hz, ArH); MS (ES+) calculated for C$_{18}$H$_{19}$N$_2$O$_5$S (M+H)+ 375.1010, found 375.1009.
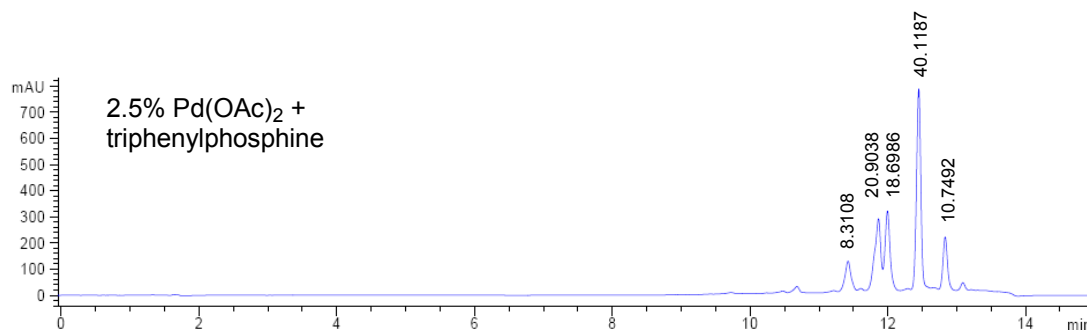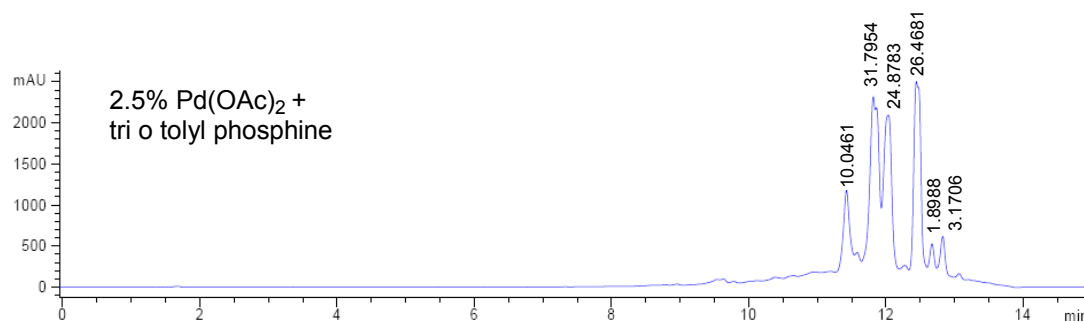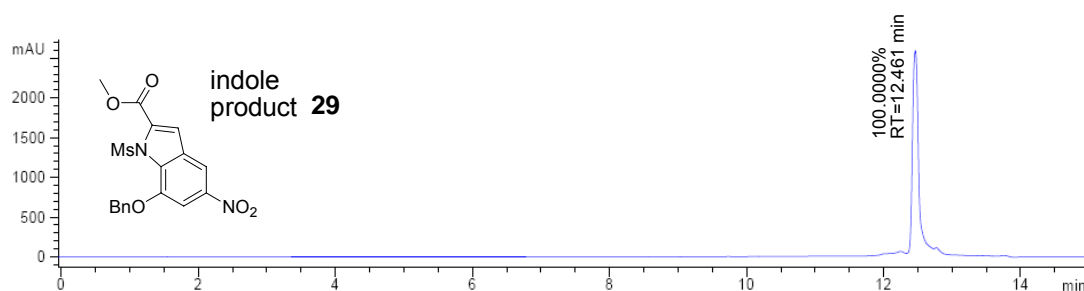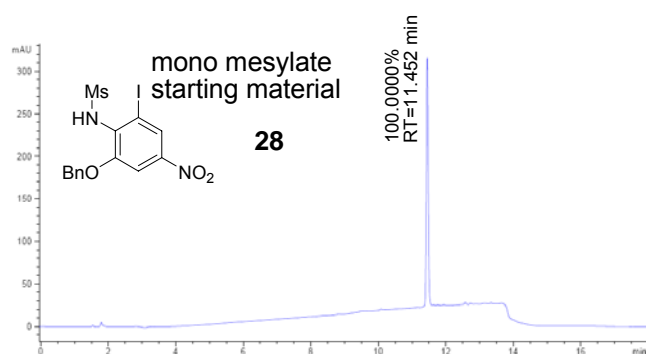
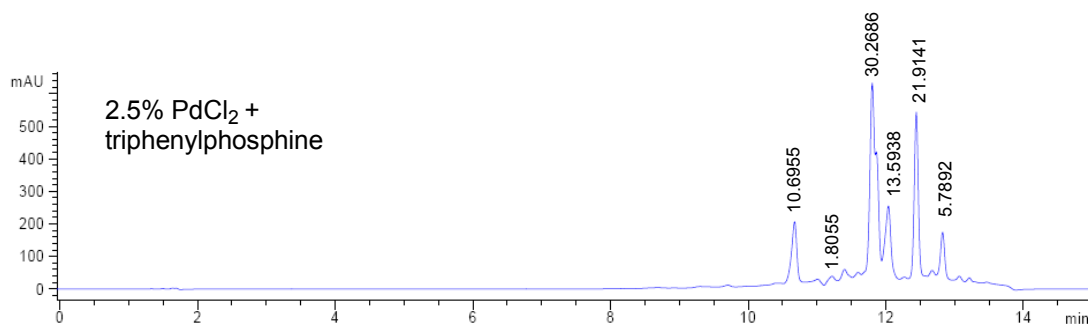**Methyl-7-benzyloxy-4-iodo-1-methanesulfonyl-5-(amino)indole-2 carboxylate [31][61]**
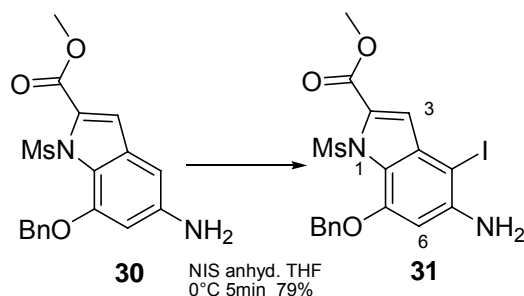
30 (58.9 mg, 0.16 mmol) was dissolved in anhyd. THF (6 ml) and cooled to 0°C. NIS (27.8 mg, 0.12 mmol) was added, and the reaction mixture stirred at 0°C for 5 min. Saturated aq. Na_2SO_3 solution was added to the mixture and extracted with EtOAc. The organic layer was washed with sat. aq. NaCl solution. The combined organic layers were dried over $MgSO_4$ and the filtrate concentrated *in vacuo*. The residue was purified by silica gel chromatography [40% EtOAc/hexane] to provide **31** (quant. yield) as a yellow oil; ^{1}H-NMR (400 MHz, $CDCl_3$) δ 3.37 (3H, s, SO_2CH_3), 3.91 (3H, s, CO_2CH_3), 5.14 (2H, s, CH_2), 6.55 (1H, s, 6-H), 7.06 (1H, s, 3-H), 7.37-7.45 (3H, m, ArH), 7.52-7.56 (2H, m, ArH); MS (ES+) calculated for $C_{18}H_{21}O_5N_3IS$ (M+NH_4)+ 518.0235 found 518.0241.

Appendix

HPLC chromatograms for studies on ligand structure effects on the formation of *in situ* Pd/ligand complexes (chapter 6). UV absorbance peaks annotated with area under curve (%).

2.5% Pd(II) nitrate hydrate + trixylylphosphite trisulfonate

2.5% Sodium tetrachloro palladate + trixylylphosphite trisulfonate

2.5% PdCl$_2$(PPh$_3$)

2.5% PdCl$_2$ + tri o tolyl phosphine

2.5% PdCl$_2$ + triphenylphosphine
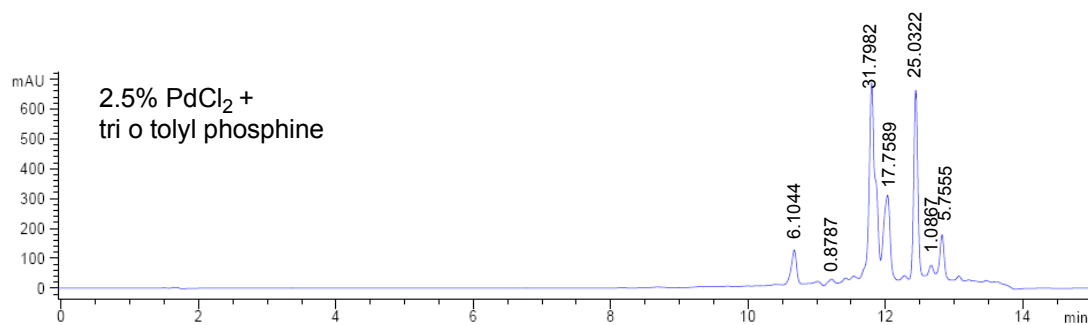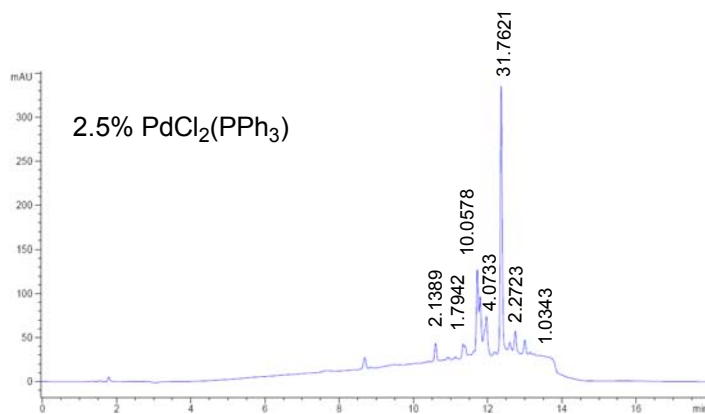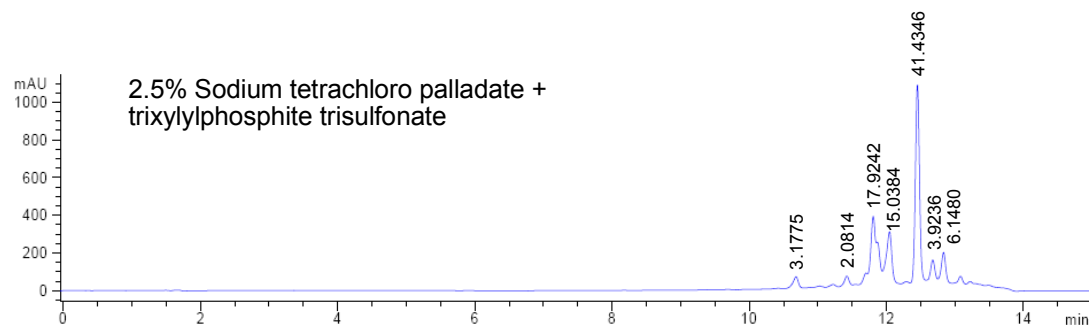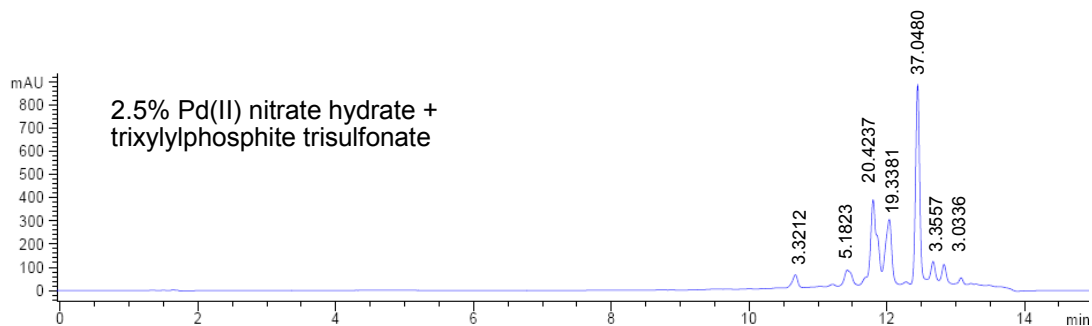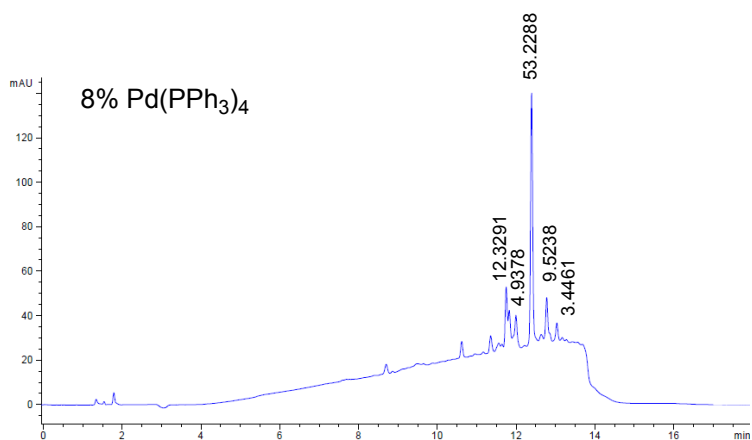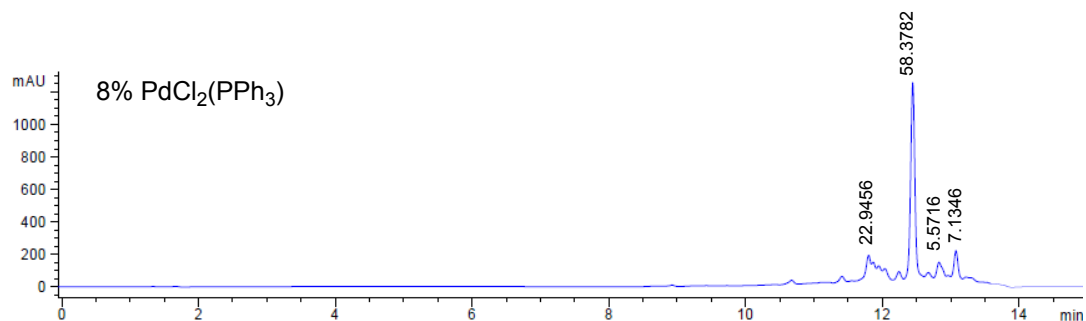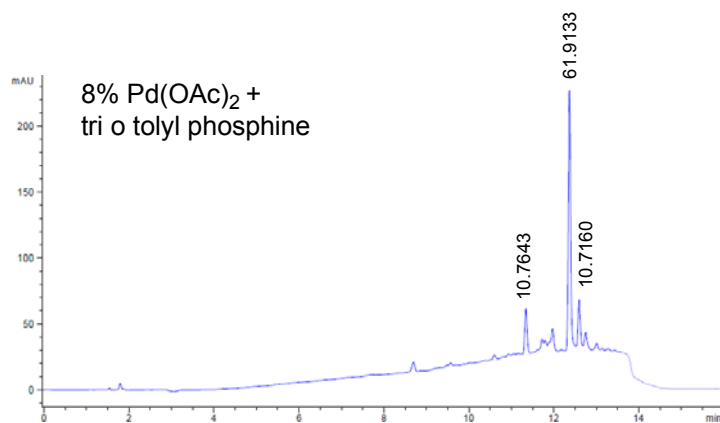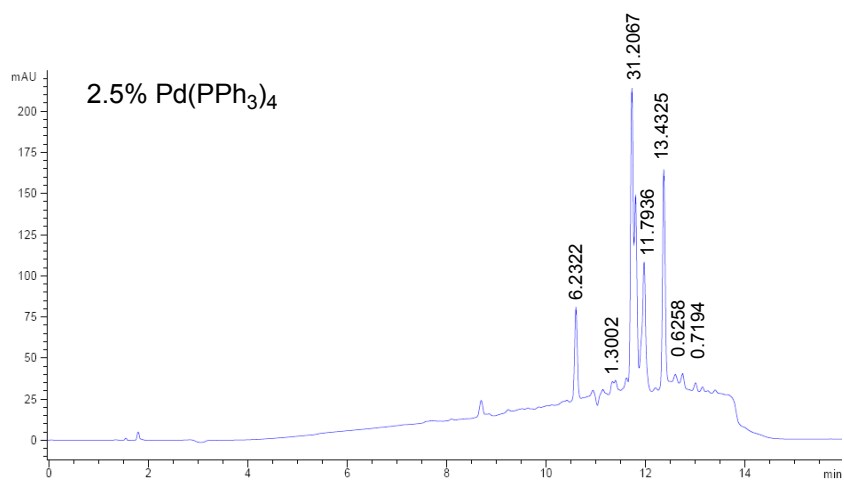
References

[1] King, R.J.B. & Robins, M.W., *Cancer Biology*, Prentice Hall, Essex, 2nd edn., 2000, ch 1, pp 1-8

[2] Kumar, P. & Clark, M., *Clinical Medicine*, WB Saunders, 4th edn., 1999, ch 7, pp 415-445

[3] National Cancer Institute, *Cancer and the environment*, www.cancer.gov/cancertopics/understandingcancer/environment, last updated July **2006**, accessed 31/10/09

[4] Montesano, R. & Hall, J., *Eur J Cancer* **2001,** 37, S67–S87

[5] *Cancer Research UK*, http://info.cancerresearchuk.org/cancerstats/ incidence, last updated November **2008**, accessed 29/04/09

[6] *World Health Organisation*, Core Health Indicators, **2002**, www.who.int, accessed 27/04/09

[7] *American Chemical Society*, **2007**, Globals facts & figures, www.cancer.org, accessed 29/04/09

[8] Jordan, M.A. & Wilson, L. *Nat Rev Cancer* **2004,** 4, 253-265

[9] Hait, W.N. *Cancer Res* **2009,** 69(4), 1263-1267

[10] Elion, G.B. & Hitchings, G.H. *J Bio Chem* **1950**, 185, 651-655

[11] Ling, Y.H., Chan, J.H., Beattie, K.L. & Nelson, A.,,, *J. Mol Pharmacol* **1992,** 42 (5), 802-807

[12] Rosa, D.D., Ismael, G., Dal Lago, L. & Awada, A. *Cancer Treatment Rev* **2008**, 34, 61-80

[13] Mauriz, J.L. & Gonzalez-Gallego, J. *J Pharm Sci* **2008**, 97 (10), 4129-4154

[14] Folkman, J. *N Engl J Med* **1971,** 285, 1182–1186

[15] Dy, G.K. & Adjei, A.A. *Cancer* **2008**, 113 (7) suppl, 1857-1887

[16] List, A., Kurtin, S., Roe, D.J., Buresh, A., Mahadevan, D., Fuchs, D., Rimsza, L., Heaton, R., Knight, R. & Zeldis, J.B. *New Engl J Med* **2005,** 35 (2), 549-557

[17] Press, M.F. & Lenz, H. *Drugs* **2007**, 67 (14), 2045-2075

[18] Patterson, L.H. & McKeown, S.R. *Brit J Cancer* **2000,** 83 (12), 1589-1593

[19] Allen, T.M. Nat Rev Cancer 2002, 2, 750-763

[20] Ismael, G.F.V., Dornelles Rosa, D., Mano, M.S. & Awada, A. *Cancer Treatment Rev* **2008**, 34, 81-91

[21] Gensini, G.F., Conti, A.A. & Lippi, D. *J Infection* **2007**, 54, 221-224

[22] Brumlik, M.J., Daniel, B.J., Waehler, R., Curiel, D.T., Giles, F.J. & Curiel T.J. *Expert Opin Drug Del* **2008**, 5 (1), 87-100

[23] Milenic, D.E., Brady, E.D. & Brechbiel, M.W. *Nat Rev Drug Discovery* **2004**, 488-498

[24] *NICE* **2002**, TA34 Breast cancer - trastuzumab: guidance, www.nice.org/guidance, accessed 10/05/09

[25] *British National Formulary* **2008**, 56, protein kinase inhibitors, 8.1.5

[26] Chari, R.V.J. *Adv Drug Del Rev* **1998**, 31, 89-104

[27] Ercan, M.T. & Caglar, M. *Curr Pharm Des* **2000**, 6, 1085–1121

[28] Goldenberg, D.M. *J Nucl Med* **2002**, 43, 693–713

[29] Zevalin description **2009**, www.zevalin.com/HealthCarePro/description.htm, accessed 14/05/09

[30] Carter, P.J. & Senter, P.D. *Cancer J* **2008**, 14 (3), 154-169

[31] EMA, Refusal assessment report for Mylotarg, **2008**, EMEA/H/C/000705, www.ema.eu, accessed 10/05/09

[32] Wyeth **2008**, Mylotarg (gemtuzumab ozogamicin for injection) prescribing information, www.wyeth.com, accessed 10/05/09

[33] Wu, A.M. & Senter, P.D. *Nat Biotech* **2005**, 23 (9), 1137-1146

[34] Hamann, P.R., Hinman, L.M., Beyer, C.F., Lindh, D., Upeslacis, J., Shochat, D., Mountain, A. *Bioconj Chem* **2005**, 16, 354-360

[35] Walker, S., Landovitz, R., Dong Ding, W., Ellestad, G. A., Kahne, D. *Proc. Nati. Acad. Sci. USA* **1992**, 89, 4608-4612

[36] Linenberger, M.L. *Blood* **2001**, 98, 988-994

[37] Phillips, G.D.L., Li, G., Dugger, D.L., Crocker, L.M., Parsons, K.L., Mai, E., Blattler, W.A., Lambert, J.M., Chari, R.V.J., Lutz, R.J., Wong, W.L.T.,

Jacobson, F.S., Koeppen, H., Schwall, R.H., Kenkare-Mitra, S.R., Spencer, S.D. & Sliwkowski, M.X. *Cancer Res* **2008**, 68 (22), 9280-9290

[38] Ricart, A.D. & Tolcher, A.W. *Nat Clin Prac Oncol* **2007**, 4 (4), 245-255

[39] Adams, G.P., Schier, R., McCall, A.M., Simmons, H.H, Horak, E.M., Alpaugh, R.K., Marks, J.D. & Weiner, L.M. *Cancer Re*s **2001**, 61, 4750-4755

[40] Ducry, L. & Stump, B. *Bioconjugate Chem* **2010**, 21, 5-13

[41] Ojima, I. *Acc Chem Res* **2008**, 41 (1), 108-119

[42] Beeson, C., Butrynski, J.E., Hart, M.J., Nourigat, C., Matthews, D.C., Press, O.W., Senter, P.D. & Bernstein, I.D. *Bioconjug Chem* **2003**, 14, 927-933

[43] Toki, B.E., Cerveny, C.G., Wahl, A.F., Senter, P.D. *J Org Chem*, **2002**, 1866-1872

[44] Tietze, L.F. & Krewer, B. *Anti Canc Agents Med Chem* **2009,** 9, 304-325

[45] Tichenor, M.S., MacMillan, K.S., Trzupek, J.D., Rayl, T.J., Hwang, I. & Boger, D.L. *J Am Chem Soc* **2007,** 129, 10858-10869

[46] Searcey, M. *Curr Pharm Des* **2002**, 8, 1375-1389

[47] Boger, D.L., Boyce C.W., Garbaccio, R.M. & Goldberg, J. *Chem Rev* **1997**, 97, 787-828

[48] Boger, D.L. & Johnson, D.S. *Angew Chem Int Ed Engl* **1996,** 35, 1438-1474

[49] Boger, D.L., Johnson, D.S., Yun, W. *J Am Chem Soc* **1994**, 116, 1635-1656

[50] Boger, D.L., Yun, W. *J Am Chem Soc* **1993**, 115, 9872-9873

[51] Ganton, M.D. & Kerr, M.A. *J Org Chem* **2007,** 72, 574-582

[52] McGovren, J.P., Clarke, G.L., Pratt, E.A. & DeKoning, T.F. *J Antibiot* **1984**, 37 (1), 63-70

[53] Hurley, L.H., Lee, C.S., McGovren, J.P., Warpehoski, M.A., Mitchel, M.A., Kelly, R.C. & Aristoff, P.A. *Biochemistry* **1988,** 27, 3886

[54] Boger, D.L., Zarrinmayeh, H., Munk, S.A., Kitos, P.A. & Suntornwat, O. *Proc Natl Acad Sci* **1991**, 88, 1431-1435

[55] Boger, D.L., Munk, S.A. & Zarrinmayeh, H. *J Am Chem Soc* **1991**, 113, 3980-3983

[56] Warpehoski, M.A., Gebhard, I., Kelly, R.C., Krueger, W.C., Li, L.H., McGovren J.P., Prairie, M.D., Wicnienski, N. & Wierenga, W. *J Med Chem* **1988**, 31, 590-603

[57] Parrish, J.P., Hughes, T.V., Hwang, I. & Boger, D.L. *J Am Chem Soc* **2004**, 126 (1), 80-81

[58] Tichenor, M.S., MacMillan K.S., Stover, J.S., Wolkenberg, S.E., Pavani, M.G., Zanella, L., Zaid, A.N., Spalluto, G., Rayl, T.J., Hwang, I., Baraldi, P.G. & Boger, D.L. *J Am Chem Soc* **2007**, 129, 14092-14099

[59] Boger, D.L., Santillan, A., Searcey, M., Brunette, S.R., Wolkenberg, S.E., Hedrick, M.P. & Jin, Q. *J Org Chem* **2000**, 65, 4101-4111

[60] Boger, D.L. & Zarrinmayeh, H.J. *J Org Chem* **1990**, 55, 1379-1390

[61] Hiroya, K., Matsumoto, S. & Sakamoto, T. *Org Lett* **2004**, 6 (17), 2953-2956

[62] Tietze, L.F., Haunert, F., Feuerstein, T. & Herzig, T *Eur J Org Chem* **2003**, 562-566

[63] Igarashi, Y., Futamata, K., Fujita, T., Sekine, A., Senda, H., Naoki, H. & Furumai, T. *J Antibiot* **2003,** 56, 107-113

[64] Tichenor, M.S., Kastrinsky, D.B. & Boger, D.L. *J Am Chem Soc* **2004**, 126, 8396-8398

[65] Tichenor, M.S., Trzupek, J.D., Kastrinsky, D.B., Shiga, F., Hwang, I. & Boger, D.L. *J Am Chem Soc* **2006**, 128, 15683-1596

[66] Okano, K., Tokuyama, H. & Fukuyama, T. *J Am Chem Soc* **2006**, 123, 7136-7137

[67] Yamada, K., Kurokawa, T., Tokuyama, H. & Fukuyama T. *J Am Chem Soc* **2003**, 125, 6630-6631

[68] Boger, D.L., Searcey, M., Tse, W.C. & Jin, Q. *Bioorg Med Chem Lett* **2000**, 5 (14), 2577-2579

[69] Ellis, D., Wolkenberg, S.E. & Boger, D.L. *J Am Chem Soc* **2001**, 123, 9299-9306

[70] Tercel, M., Gieseg, M.A., Denny, W.A. & Wilson, W.R. *J Org Chem* **1999**, 64, 5946-5953

[71] Daniell, K., Stewart, M., Madsen, E., Le, M., Handl, H., Brooks, N., Kiakos, K., Hartley, J.A. & Lee, M. *Bioorg Med Chem Lett* **2005**, 15, 177-180

[72] Tietze, L.F., Major,F., Schuberth, I., Spiegl, D.A., Krewer, B., Maksimenka, K., Bringmann, G. & Magull, J. *Chem Eur J* **2007**, 13, 4396-4409

[73] Tietze, L.F., Panknin, O., Krewer, B., Major, F. & Schuberth, I. *Int J Mol Sci* **2008**, 9, 821-837

[74] Chari, R.V.J, Jacket, K.A., Bourret, L.A., Derr, S.M., Tadayoni, B.M., Kristin M. Mattocks, K.M., Shah, S.A., Liu, C., Blattler, W.A. & Goldmacher, V.S. *Cancer Res* **1995**, 55, 4079-4084

[75] Chari, R.V.J. *Acc Chem Res* **2008,** 41 (1), 98-107

[76] Desbene, S., Van, H.D., Michel, S., Koch, M., Tillequin, F., Fournier, G., Farjaudon, N. & Monneret, C. *Anticancer Drug Des* **1998**, 13, 955

[77] Jeffrey, S.C., Nguyen, M. T., Moser, R.F., Meyer, D.L., Miyamoto, J.B. & Senter, P.D. *Bioorg Med Chem Lett* **2007**, 2278-2280

[78] Searcey, M. **2008,** private communication

[79] Clayden, J., Greeves, N., Warren, S. & Wothers, P., *Organic Chemistry*, OUP, Oxford, 2005 (last reprint), ch 24, p 621

[80] Corey, E.J., Andersen, N.H., Carlson, R.M., Paust, J., Vedejs, E., Vlattas, I & Winter, R.E.K. *J Am Chem Soc* **1968**, 90, 3247

[81] Fitch, R.W. & Luzzio, F.A. *Tet Lett* **1994**, 35 (33), 6013-6016

[82] Anderson, J.C. & Chapman, H.A. *Synthesis* **2006**, 19, 3309-3315

[83] Solomons, T.W.G. & Foyle, C.B. *Organic Chemistry*, Wiley, Asia, 2011 pp 520

[84] Snyder, H.R. & Hecker, R.E. *J Am Chem Soc* **1952**, 74, 2006-2009

[85] Malek, J. & Cerny, M. *Synthesis* **1972,** 217-234

[86] Jorgensen, M.R., Olsen, C.A., Mellor, I.R., Usherwood, P.N.R., Witt, M., Franzyk, H. & Jaroszewski. J.W. *J Med Chem* **2005,** 48 56-70

[87] Jennings, L.D., Foreman, K.W., Rush, T.S., Tsao, D.H.H., Mosyak, L., Kincaid, S.L., Sukhdeo, M.N., Sutherland, A.G., Ding, W., Kenny, C.H., Sabus, C.L., Liu, H., Dushin, E.G., Moghazeh, S.L., Labthavikul, P., Petersen, P.J., Tuckman M. & Ruzin, A.V. *Bioorg Med Chem* **2004,** 12, 5115-5131

[88] Ek, F., Axelsson, O., Wistrand, L.G. & Frejd, T. *J Org Chem* **2002,** 67, 6376-6381

[89] Ek, F., Wistrand, L.G. & Frejd, T. *J Org Chem* **2003**, 68, 1911-1918

[90] Sohn, J., Li, Z., Deng, L., Safi, A., Pirrung, M.C. & Rudolph, J. *J Med Chem* **2003,** 46, 2580-2588

[91] Kotha, S., Behera, M. & Shah, V.R. *Synlett* **2005**, 12, 1877-1880

[92] Kolar, A. & Olsen, R. *Synthesis* **1977**, 457-459

[93] Dygos, J., Yonan, E.E., Scaros, M.G., Goodmonson, O.J., Getman D.P., Periana, R.A. & Beck, G.R. *Synthesis* **1992,** 741-743

[94] Kurti, L. & Czako, B. Strategic applications of named reactions in organic synthesis, Elsevier, London, 2005

[95] Crisp, G.T. *Chem Soc Rev* **1998**, 27, 427-436

[96] Yao, Q., Kinney, E.P. & Yang, Z. *J Org Chem* **2003**, 68, 7528-7531

[97] Kappe, C.O. *Angew Chem Int Ed* **2004,** 43, 6250 –6284

[98] Lidstrom, P., Tierney, J., Wathey, J. & Westman, J. *Tetrahedron* **2001**, 57, 9225-9283

[99] Leadbeater, N.E. & Marco, M. *Org Lett* **2002**, 4, 2973 – 2976

[100] Stadler, A., Yousefi, B., Dallinger, D., Walla, P., Van der Eycken, E., Kaval, N. & Kappe, C.O. *Org Process Res Dev* **2003**, 7, 707-716

[101] Larhed, M. & Hallberg, A. *J Org Chem* **1996**, 61, 26

[102] Robinson, B. *Chem Rev* **1962**, 374-401

[103] Clayden, J., Greeves, N., Warren, S. & Wothers, P., *Organic Chemistry*, OUP, Oxford, 2005 (last reprint), ch 36, p 950; ch 44, p 1204

[104] Guy, A. & Guette, J.P. *Synthesis* **1980** 222-223

[105] Parmenter, S.M., Cook, A.G. & Dixon, W.B. *J Am Chem Soc* **1968**, 80, 4621

[106] Hemetsberger, H., Knittel, D. & Weidmann, H. *Monatsh Chem* **1969**, 100, 1599-1603

[107] MacKenzie, A.R., Moody, C.J. & Rees, C.W. *J Chem Soc Chem Commun* **1983**, 1372-1373

[108] Moody, C.J. & Ward, J.G. J *Chem Soc Perkin Trans I* **1984,** 2903-2909

[109] MacKenzie, A.R., Moody, C.J. & Rees, C.W. *Tet* **1986**, 42 (12), 3259-3286

[110] Martin, T. & Moody, C.J. *J Chem Soc Perkin Trans I* **1988,** 241-246

[111] Jones, G.B. & Moody, C.J. *J Chem Soc Chem* **1989**, 186-187

[112] Bolton, R., Moody, C.J., Pass, M., Rees, C.W. & Tojo, G. *J Chem Soc Perkin Trans* I **1988,** 2491-2499

[113] Bolton, R., Moody, C.J., Rees, C.W. & Tojo, G. *J Chem Soc Chem Commun* **1985**, 1775-1776

[114] Mancuso, A. & Swern, D. *Synthesis* **1981**, 165-185

[115] Farnier, M., Soth, S. & Fournari, P. *Can J Chem* **1976**, 1066-1073

[116] Henn, L., Hickey, B., Moody, C.J. & Rees, C.W. *J Chem Soc Perkin Trans I* **1984**, 2189-2196

[117] Sechi, M., Derudas, M., *J Med Chem* **2004**, 47, 5298-5310

[118] da Rosa, A.F., Rebelo, R.A. & Nascimento, M.G. *J Braz Chem Soc* **2003**, 14 (1), 11-15

[119] Allen, M.S., Hamaker, L.K., La Loggia, A.J. & Cook, J.M. *Synth Commun* **1992**, 22 (14), 2077-2102

[120] Hemetsberger, H., Knittel, D. & Weidmann, H. *Montsh. Chem* **1970**, 101, 161-165

[121] Moody, C.J. Oxidation by nitrene insertion in *Comprehensive Organic Synthesis* **1991,** ed. Trost & Fleming

[122] Knittel, D. *Synthesis* **1985**, 186-188

[123] Smolinsky, G. *J Org Chem* **1962**, 3557-3559

[124] Isomura, K., Okada, M. & Taniguchi, H. *Tet Lett* **1969**, 46, 4073-4076

[125] Taniguchi, H., Isomura, K. & Tanaka, T. *Heterocycles* **1977**, 6 (9), 1563-1568

[126] Su, B., Landini, S., Davis, D.D., Brueggemeier, R.W. *J Med Chem* **2007,** 50, 1635-1644

[127] Fischer, E & Lipschitz, W., *Ber*. **1915**, 48, 360

[128] Negishi, E. & King, A. O. & Okukado, N, *J. Org. Chem* **1977**, 42, 1821-1823

[129] Trost, B.M. 'Organopalladium compounds in organic synthesis and catalysis' in *Comprehensive Organometallic Chemistry* **1982** vol 8, chap 57, ed Wilkinson G, Oxford

[130] Negishi, E., Takahashi, T. & Akiyoshi, K. *Chem Ind* **1988**, 33, 381-407

[131] Casares, J.A., Espinet, P., Fuentes, B. & Salas, G. *J Am Chem Soc* **2007**, 129, 3508-3509

[132] Negishi, E. *Acc Chem Res* **1982,** 15, 340-348

[133] Eckert, T. & Ipaktschi, *Synthetic Comm* **1998**, 28 (2), 327-335

[134] Anastasia, L. & Negishi, E. *Org Lett* **2001**, 20, 3111-3113

[135] Sakamoto, T., Shiga, F., Yasuhara, A., Uchiyama, D., Kondo, Y. & Yamanaka, H. *Synthesis* **1992** , 746-748

[136] Larhed, M., Moberg, C. & Hallberg, A. *Acc Chem Res* **2002**, 35, 717 – 727

[137] Stanetty, P., Schnürch, M. & Mihovilovic, M.D. *Synlett* **2003** 1862-1864

[138] Ohberg, L. & Westman, *Synlett* **2001,** 1893-1896

[139] Walla, P. & Kappe, C.O. *Chem Commun* **2004,** 564-565

[140] Mutule, I. & Suna, E. *Tet Lett* **2004** 45, 3909-3912

[141] Genov, M., Almorin, A. & Espinet, P. *Tetrahedron:Asymmetry* **2007,** 18, 625-627

[142] Lipshutz, B.H., Frieman, B.A., Lee, C.T., Lower, A., Nihan, D.M. & Taft, B.R. *Chem Asian J* **2006**, 1 (3), 417-429

[143] Majetich, G. & Hicks, R. *Radial Phys Chem* **1995,** 45 (4), 567-579

[144] Olofsson, K., Kim, S.K., Larhed, M., Curran, D.P. & Hallberg, A. *J Org Chem* **1999,** 64, 4539-4541

[145] O'Donnell, M., Zhou, C., Mi, A., Chen, N. & Kyle, J.A. *Tetrahedron Lett* **1995**, 24, 4205-4208

[146] Tolman, C.A. *Chem Rev* **1977**, 77 (3), 313-346

[147] Barrios-Landeros, F. & Hartwig, J.F. *J Am Chem Soc* **2005**, 127 (19), 6944-6945

[148] Shaughnessy, K.H. Eur *J Org Chem* **2006**, 8, 1827-1835

[149] Christmann, U. & Vilar, R. *Angew Chem Int Ed* **2005**, 44 (3), 366-374

[150] Amatore, C. et al. *Organometallics* **1995**, 14, 1818-1826

[151] Valentine, D.H. & Hillhouse, J.H. *Synthesis* **2003**, 16, 2437–2460

[152] Amatore, C., Emmanuelle Carre, E., Jutand, A. & M'Barkii, M.A. *Organometallics* **1992**, 11, 3009-3013

[153] Odle, R., Blevins, B., Ratcliff, M. & Hegedus, L.S. *J Org Chem* **1980**, 45, 2710-2713

[154] Ortuzar-Kerr, N. 2006, Thesis (PhD), The School of Pharmacy, University of London

[155] Faucher, N., Martres, P., Laroze, A., Pineau, O., Potvain, F. & Grillot, D. *Bioorg Med Chem Lett* **2008**, 18, 710-715

[156] Wuts, P.G.M. & Greene, T.W. *Greene's Protective Groups in Organic Synthesis* **2007**, 4th ed, Wily

[157] McCarroll, A.J., Bradshaw, T.D., Westwell, A.D., Matthews, C.S. & Stevens, M.F.G. *J Med Chem* **2007**, 50, 1707-1710

[158] Luzzio, F.A. & O'Hara, L.C. *Synth Comm* **1990**, 20(20), 3223-3234

[159] Birkinshaw, T. N. & Holmes, A. B. *Tetrahedron Letters* **1987,** 28, 813-816

[160] Pitt, M.R. *J Chem Soc Perkin Trans I* **2001**, 955-977

[161] Banik, B.K., Banik, I. & Becker, F.F. *Org Synth* **2004**, 81, 188-193