Design and Synthesis of Duocarmycin Analogues for the Treatment of Cancer

Lily Rebecca Cassidy

Supervisor: Professor Mark Searcey, Dr Andrew Beekman

A thesis submitted for the degree of Doctor of Philosophy

2024

University of East Anglia

SCHOOL OF PHARMACY

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 and acknowledgment is given, this work is original and has been carried out by the author alone.

Lily Rebecca Cassidy

Abstract

The duocarmycins are a family of natural compounds and their synthetic analogues that are highly potent DNA alkylating agents. The parent compounds are among the most potent cytotoxic compounds discovered in nature and therefore offer great potential for cancer treatment. However, high toxicity and low selectivity mean there has been limited progress towards the clinic as chemotherapeutic agents. Research into finding ways to achieve tumour selective cytotoxicity is of great importance. One approach is with antibody drug conjugates, or ADCs. Here the cytotoxic drug (the payload) is bound by a linker to the antibody and in doing this the payload can be targeted directly to the tumours. A duocarmycin-based ADC called SYD985 or trastuzumab duocarmazine has recently finished phase III clinical trials for the treatment of HER-2 metastatic breast cancer, for which approval is currently being sought.[1] This shows an exciting new field to be explored and this project will look at preparing dimeric payloads that could potentially be used as a warhead for an ADC.

The duocarmycin family consists of different alkylating subunits, including the duocarmycin SA unit, DSA, one of the most potent subunits and of great interest in the design of analogues. This thesis gives a review over chemotherapeutic agents and the new focus of treatments shifting towards more targeted therapies. The project initially follows work previously done by Searcey *et al*. where the DSA alkylating unit was synthesised and will then go on to discuss an alternative route which is able to avoid some of the challenges from the initial synthesis. The synthesis of a set of novel dimers is then described, and the importance of stereochemistry is addressed. DNA stapling with the duocarmycins is something that has not previously been investigated, so opens this project to a new field of study. Finally, the ability for the compounds to alkylate DNA as well as their anti-proliferate activity was investigated and all the dimers synthesised were shown to be active against HL-60 cells, a human leukemia cell line.

Contents

Acknowledgements

I would like to firstly thank my supervisors Professor Mark Searcey and Dr Andrew Beekman for their support and guidance throughout my studies. I am so grateful to Mark for giving me this opportunity to study such an interesting class of compounds, and for giving me the freedom to take my project in directions that I thought were exciting. Andrew thank you for help throughout the years but especially thank you for your feedback and encouragement of my presenting and public speaking. You have given me the confidence and experience to happily stand up in conferences and talk about the exciting things I have accomplished throughout my PhD.

It has been a pleasure to work alongside both the Searcey and Beekman research groups in a professional and social environment. I am hugely grateful to have been part of such a fun, sociable and supportive work environment, and I'm certain I have gained friends for life.

I would like to say a few special thanks; Dr Ryan Tinson and Dr Bethany Hood for your help and advice with my synthesis during my PhD. Dr Marco Cominetti with all your experience with the duocarmycins as well as your ability to fix the 'unfixable' machines in the lab that are so crucial in allowing everyone to do their research! I would like to thank Dr Zoë Goddard for your help, in learning new techniques in biology and how to work with cells for my different assays. And finally, Vicky Kamperi for your help with the Gold nanoparticle assay, and for being a great friend in and out of the labs.

Outside of research I have been lucky to have such supportive friends and family. Thank you to my mum and dad for the years of love and guidance. My brother James, who has now also just finished his PhD in Sweden, you are a huge inspiration but also my biggest cheerleader. My friends for always being there, despite not knowing much about science, have always been excited and interested in what I am up too! Lastly a special thank you to my partner Jonny and my cat Blue for your everyday love and support, you have made finishing and writing this thesis a happy time to look back on.

Abbreviations

ADC	Antibody-drug conjugates
A	Adenine can also be used for Adenosine
AT	Adenine and thymine
AuNPs	Gold nanoparticles
CBI	1,2,9,9a-Tetrahydrocyclopropa[*c*]benz[*e*]indol-4-one
C	Cytosine
DCM	Dichloromethane
DMF	Dimethylformamide
DMP	Dess-martin periodinane
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
Dox	Doxorubicin
DSA	Duocarmycin SA alkylation subunit
DSI	Duocarmycin SA-indole
EtOAc	Ethyl acetate
FDA	The Food and Drug Administration
Fmoc	9-Fluorenylmethoxycarbonyl
Fab	Variable region of an antibody
Fc	Constant region of an antibody
G	Guanine
HER2	Human epidermal growth factor receptor 2
HOBt	1-Hydroxybenzotriazole hydrate
IR	Infrared
MeOH	Methanol
MTS	3-(4,5-Dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium
NitroCBIs	Nitrochloromethylbenzindolines
NMR	Nuclear magnetic resonance
ppm	Parts per million
PBS	Phosphate buffered saline
PDC	Peptide drug conjugate

PEG	Polyethylene glycol
R&D	Research and development
T	Thymine
TLC	Thin layer chromatography
TTMSS	Tris(trimethylsilyl)silane
TIPS	Triisopropylsilane
UV	Ultraviolet
Uv vis	UV/Vis spectrophotometer
5-FU	5-Fluorouracil

Chapter 1 – Introduction

Chapter 1 – Introduction

The aim of this chapter is to introduce the different approaches to cancer treatment and show the importance of finding new ways to target this group of diseases. The reader will see the shift in focus within modern cancer treatments towards more targeted therapies, with particular focus on antibody drug conjugates, ADCs. There will be a discussion on different payloads and why duocarmycin analogues make exciting candidates, lastly touching on dimers and higher order structures to enhance potency and selectivity.

1.1. Cancer

Cancer is a leading cause of death globally. The World Health Organization estimates that in 2020 there were almost 10 million deaths as a result of this group of diseases.[1] The number of deaths is predicted to rise significantly over the coming years, with 4 million new cases a year being predicted in the UK alone by 2030 (**Figure 1.1**).[2] Finding new cancer treatments is therefore of huge importance. Furthermore, an ageing population contributes to the increased chance of cancer, with ageing associated to an increased likeliness of DNA mutations in healthy cells.[3] For this reason there is an intense focus on finding ways to treat and prevent cancer; in fact oncology is the largest area of focus within the pharmaceutical industry.[4]

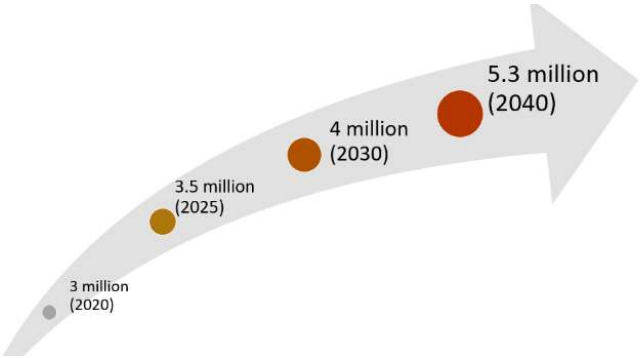


Figure 1.1: *Predicted rise in cancer cases 2020 to 2040 in the UK.* [2]

Cancer is the name given to a group of diseases where changes in DNA cause cells to replicate uncontrollably. There are different ways to try to stop or kill cancer cells and the method of choice depends on the type of cancer that is present: it is estimated that there are more than 100 types of cancer.[5] **Figure 1.2** shows some of the most common

cancer types in 2020 worldwide and their percentages relative to one another according to figures reported by The World Health Organization.[1]

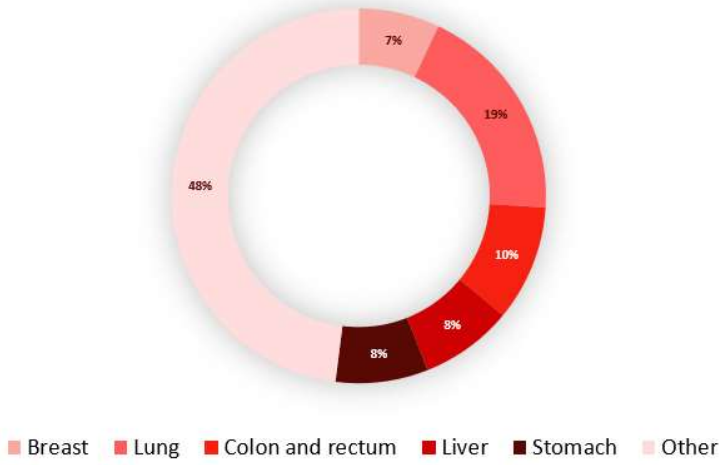


Figure 1.2: *Most common new cases of cancer in 2020.* [1]

There are different approaches to cancer treatment, including surgery, immunotherapy, radiotherapy, chemotherapy, and more targeted approaches. The first line of treatment for most solid tumours is surgery and then treatment is often followed by a combination of different therapies depending on the type of cancer and how advanced the disease is.[6]

1.2. Classical Cancer Chemotherapy

Chemotherapy is one approach to treat cancer, which uses cytotoxic drugs to kill rapidly dividing cancer cells. For cancers to grow and spread, the DNA must replicate. Drugs that target or interrupt any stage of DNA replication can therefore kill cancer cells. Different drugs have different mechanisms but most chemotherapeutic agents target DNA replication.[7] There are four main categories of cancer chemotherapy agents that suggest the mode of action; topoisomerase inhibitors, antimitotics, antimetabolites and alkylating agents. There are also antitumour antibiotics and more recently a general group called targeting agents/approaches.

1.2.1 Topoisomerase inhibitors

Topoisomerase inhibitors work by inhibiting a family of enzymes called topoisomerases that are responsible for regulating supercoiling in DNA.[8] Controlling supercoiling is essential for DNA replication, and if not controlled will result in cell death.[9] DNA topoisomerase I and II are enzymes responsible for resolution of supercoils, therefore finding compounds that are able to inhibit these can be an effective way to inhibit replication.[10]

Camptothecins are a class of natural anticancer agents that act on topoisomerase I. Most of the topoisomerase I inhibitors are derivatives or analogues of the natural plant product camptothecin (**Figure 1.3**).[11]

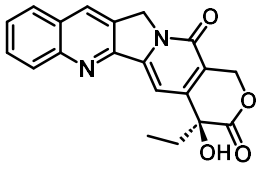


Figure 1.3: *Structure of camptothecin.*

Topoisomerase I and DNA form a complex during DNA replication which causes a relaxation of the supercoiling.[12] Although the exact mechanism of action is not clear, camptothecin traps topoisomerase I in this complex with DNA, preventing replication and leading to cell death.[13] Since the discovery of camptothecin, different analogues have been synthesised to try and overcome issues with its use, such as poor water solubility, instability in plasma and high toxicity.[14]

One of the most well-known topoisomerase II inhibitors is doxorubicin, this anthracycline is used to treat a range of different cancers.[15] Interestingly, doxorubicin also falls into another category of cancer chemotherapy, which is 'antitumour antibiotics.' Doxorubicin is a natural compound that was derived from *Streptomyces peucetius* bacteria.[15] Its structure (**Figure 1.4**) allows it to act as a non-covalent DNA binding agent.[16] It has planar aromatic rings which allow it to intercalate between the base pairs of DNA, positioning the side chains in one of the grooves.[16] It is then also able to interact with the topoisomerase II enzyme and lead to its inhibition.[17]

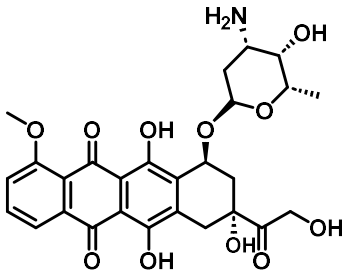


Figure 1.4: *Structure of doxorubicin.*

Doxorubicin is dose limited in its use. It has the added complication that, through its effects on an enzyme called superoxide dismutase in the heart, it is known to cause irreversible heart failure and cardiotoxicity.[18]

1.2.2 Antimitotics

Antimitotics inhibit mitosis. These complex plant alkaloids include, for example, vinca alkaloids and taxanes. Both the vinca alkaloids and taxanes work by disrupting the formation and disassembly of microtubules, which are important in mitosis.[19] They work by different mechanisms, with vinca alkaloids disrupting the formation of the microtubules whereas taxanes stabilise them.[20,21] Most processes of inhibition have a catastrophic effect on the cell.

Vinca alkaloids have been used as anticancer agents for over 30 years, with one of the most well-known agents, vinblastine, being used for the treatment of late-stage testicular cancer, breast cancer and Hodgkins lymphoma (**Figure 1.5**).[22,23] They all have the same mechanism of action. They bind to tubulin, which is a protein used in the assembly of microtubules. In this way the cell is arrested in this stage of the cell cycle and apoptosis is induced.[24]

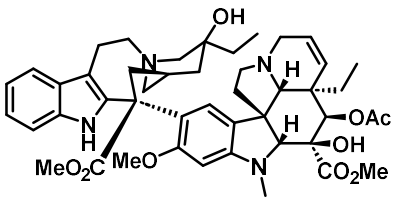


Figure 1.5*: Structure of vinblastine.*

Paclitaxel (Taxol) was first discovered in 1969 and its structure published in 1971 when the National Cancer Institute set up a program to screen plant extracts for potential anticancer activity.[25] By 2009 over 550 of these taxane diterpenoids had been isolated, all from different yew trees.[26] The structural backbone of these compounds is very similar for many of the compounds; however, there can be variations and in 2005 there were 9 different skeleton arrangements recorded.[27] **Figure 1.6** shows one of these skeletal arrangements and the complete structures of two of the most successful taxanes, paclitaxel and docetaxel (Taxotere) highlighting key differences in red.[27]

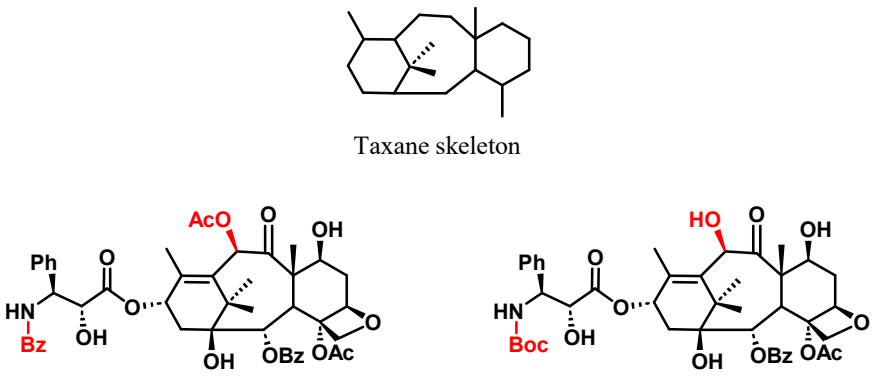


Figure 1.6: *The backbone of paclitaxel and docetaxel, highlighting key functional group differences.*

These drugs are classed as microtubule-stabilising agents and they bind irreversibly to tubulin. In doing this, the microtubules are unable to disassemble and the cell is unable to complete mitosis, resulting in apoptosis.[28]

Despite poor solubility and dose-limiting toxicity, paclitaxel is one of the most widely prescribed anticancer drugs, used to treat a range of cancers.[29] Docetaxel is a semi-synthetic analogue of paclitaxel which has a higher potency, requiring less drug to have a therapeutic effect. There are minor differences, which are highlighted red in the **Figure 1.6**, showing that docetaxel has a tert-Butyloxycarbonyl (Boc) group instead of a benzoyl group and a hydroxyl group instead of an ester group. This change affects the pharmacokinetics and means that docetaxel is more water soluble and has a higher affinity for its target, increasing its therapeutic effect.[30]

1.2.3 Antimetabolites

A metabolite is a substance formed during metabolism, or a product necessary for metabolism.[31] An antimetabolite is a product that interferes or inhibits normal metabolic processes.[31]

Antimetabolites affect the synthesis phase of the cell cycle, where DNA is replicated and a complete copy is synthesised.[32] Most antimetabolites have a similar structure to nucleotides or folates which allow them to effectively target DNA or RNA synthesis.[33] Antimetabolites work by either replacing the DNA bases with other bases or by inhibiting the enzymes involved in nucleotide synthesis. Within antimetabolites there are three subcategories;

1. Antifolates / folate antagonists

Folate antagonists act by blocking the binding of folic acid, which inhibits folate-dependent enzymes.[34] In doing this, the key enzymes involved in DNA and RNA synthesis are stopped, which can stop tumour progression.

2. Purine analogues

These analogues prevent the synthesis of adenine and guanine, necessary for the production of DNA and RNA. They are able to do this as they are structurally very similar to natural metabolic purines, and can therefore compete with these during DNA replication, therefore inhibiting them.[35]

3. Pyrimidine analogues

Pyrimidine analogues work via the same method as the purine analogues. Pyrimidines are again structurally very similar to those natural pyrimidines required for metabolic functions, here they are able to block the synthesis of nucleotides cytosine, thymine and uracil.

Two of the most widely used and well-known antimetabolites are methotrexate and 5-flurouracil (5-FU). Methotrexate is among the first antimetabolites to be discovered, having been used in clinic since 1983.[36] Methotrexate is a folate antagonist. It is designed to mimic the structure of folic acid, a key co-factor in the biosynthesis of

thymidine and purines. **Figure 1.7** shows methotrexate next to folic acid, with the key functional group differences highlighted in red font.[34] Methotrexate works by inhibiting the enzyme dihydrofolate reductase (DHFR), which is folate-dependent. DHFR is responsible for synthesis of purines and pyrimidines, therefore hugely important in DNA and RNA synthesis.

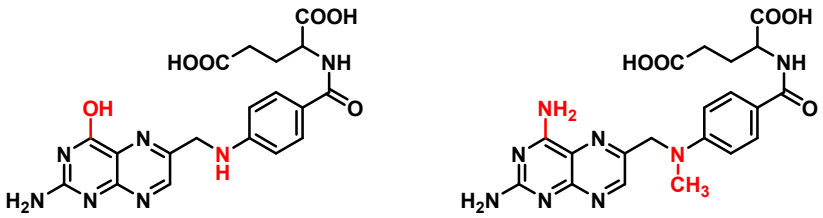


Figure 1.7: *Structures of Folic Acid and Methotrexate, highlighting key functional group differences in red.*

The pyrimidine analogue 5-fluorouracil (5-FU) is shown in **Figure 1.8**. This uracil nucleotide has been used as a chemotherapeutic drug since 1958.[37]

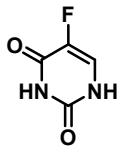


Figure 1.8: *Structure of 5-flurouracil.*

5-FU is one of the most widely used chemotherapeutics, used to treat a range of cancers including breast, colon, pancreas, ovarian as well as several other cancers.[38] It primarily works by inhibiting cellular thymidylate synthase (TS) preventing DNA synthesis, but also by incorporating into growing DNA strands via conversion to the nucleotide.[39]

1.2.4 Alkylating agents

Cell survival relies on DNA replication. Alkylating agents are able to destroy cells by damaging DNA and as a result prevent mitosis.[40] Alkylating agents work in different ways; they can cause DNA strand breaks, abnormal base pairing, or they can crosslink

the bases in the DNA double helix. All of these interactions with DNA will cause apoptosis as the strands are unable to replicate and tumour growth therefore stops.[41]

Alkylating agents were discovered during World War 1 through the use of mustard gas, also known as sulphur mustard. They were used in the war as a chemical warfare agent. Scientists noted significant effects on soldiers with volume of lymphoid tissues shrinking after exposure.[42] Research into this group of compounds as a potential cancer treatment was therefore pursued. Among the sulphur mustards, important nitrogenous variants were considered, the nitrogen mustards. The most well known nitrogen mustard, mechlorethamine, is among this family of compounds, and was the first chemotherapeutic drug to be used on patients.[43]

This group of compounds are able to crosslink DNA. Mechlorethamine is bifunctional due to having two chloroalkyl groups (**Figure 1.9**). These groups form an aziridine-ring intermediate which is then able to undergo DNA alkylation due to the highly strained rings. These key structural features allow it to form crosslinks as there are two sites that can alkylate at the guanine-N7 twice on DNA strands.[44] Their mechanism of action and the concept of DNA crosslinking will be discussed further in the introduction to **Chapter 4**.

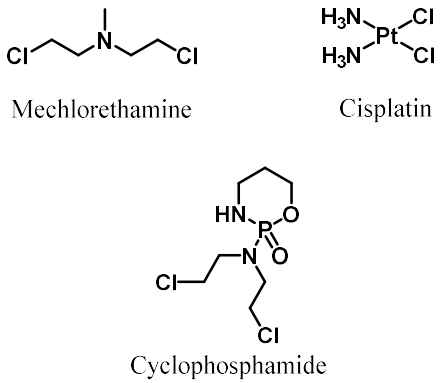


Figure 1.9: *The structures of three alkylating units: Mechlorethamine, cyclophosphamide and cisplatin.*

The nitrogen mustards, although successful, had issues with selectivity and toxicity.[45] Cyclophosphamide was developed by Arnold and co-workers in an attempt to increase selectivity (**Figure 1.9**).[46] Cyclophosphamide is among the most used of the alkylating agents, used not just to treat a range of cancers but also as an immunosuppressant.[47]

Its mechanism of action is similar to that of the nitrogen mustards in that it is inactive until it cyclises to form the active compound, which in the case of cyclophosphamide is the 4-hydroxycyclophosphamide. Unlike the nitrogen mustards, cyclophosphamide is only activated *in vivo* and requires activation by hepatic microsomal enzymes.[48]

Other well-known examples of drugs in this class are platinum drugs such as cisplatin and carboplatin. Cisplatin (**Figure 1.9**) also crosslinks DNA. Similarly to the other two groups, cisplatin is able to form crosslinks, mainly instrastand but also interstrand crosslinks.[49] DNA crosslinking is an interesting concept and could be applied to other drugs to try and improve potency and selectivity. The mechanism of action in which these compounds are able to alkylate DNA and form these crosslinks is discussed further in **Chapters 3** and **4** where mechanisms of interstrand vs intrastrand crosslinks are compared. Cisplatin is among the most commonly used antitumor agents, however, due to side effects including severe nausea, many different analogues have been developed including; carboplatin, oxaliplatin, JM216, and many others.[50]

1.2.5 Issues with classical chemotherapeutics

A common problem with cytotoxics is their side effects, which mainly derive from their lack of selectivity in targeting all types of proliferating cells in the body.

Similarly to the topoisomerase inhibitors, antimitotic drugs also have significant side effects, and although they are highly effective and widely used, finding alternatives that may avoid some of these side effects is important. New analogues may be able to overcome drug resistance but the issue of off-target effects and toxicity are still prevalent.

Often all these categories of drugs are used in combination with each other to work synergistically to target different pathways. Cisplatin is often used with the antimetabolite 5-FU.[51] Antimetabolites are also often used in combination with antimitotics, for example it is recommended that docetaxel (see **Section 1.2.3**.) is used alongside gemcitabine (an antimetabolite) for the treatment of metastatic breast cancer.[52] In fact, most of these treatments are used as combination treatments but this still fails to address some of the issues faced with selectivity. Almost all of these drugs

and combinations have significant issues with toxicity, drug resistance and side effects. Finding new analogues and new compounds is hugely important but to really avoid these issues, a different strategy is needed.

1.3. Targeted Therapies in Cancer

There is a need for new approaches in cancer chemotherapy to avoid the side effects associated with cytotoxics, and this is being reflected in the pharmaceutical industry with a greater focus on targeted therapies. This concept has been around for a very long time, with Paul Ehrlich winning a Nobel Prize in 1908 for his theory of a 'magic bullet' describing the delivery of a toxic payload directly to a target disease.[53] Developments in science and a better understanding of cancer biology, means that there is now a clearer view of the unique or close to unique environments associated with a tumour that can now be exploited to treat cancer.

Targeted therapies avoid the many side effects associated with off-target effects. In theory healthy cells should be left unharmed, also allowing an increased therapeutic window. In practice, this is difficult and although targeted therapies are significantly more advanced than classical chemotherapy methods, there are still issues to be overcome. One major issue is that of resistance to single agent targeted treatments, especially in cancers that are more advanced.[54] Targeted therapies are still a relatively new field in oncology, and there are limited cancers that can be treated with these approaches. This shows the importance of finding new targets as well as new payloads.

There are now two main approaches to modern cancer chemotherapy: targeting particular cellular pathways or targeting a cytotoxic agent to the tumour.

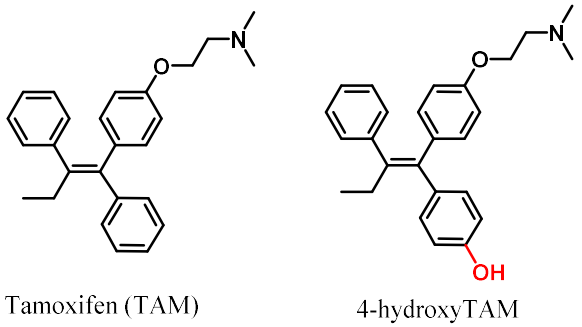
1.3.1 Targeting a specific pathway

Tamoxifen is a hormonal drug to treat breast cancer and was one of the first targeted therapies to be discovered and used. Tamoxifen has been used since the late 1960s and is still one of the most widely prescribed anticancer drugs in the world.[55]

The endocrine system is responsible for making hormones and the growth of certain cancers (such as breast cancer) can be influenced by these hormones. Anti-hormones

are used in cases like these where the tumour is hormone-dependent. By blocking the action of the hormone responsible for the growth of the tumour, the cancer can be slowed significantly or stopped.

Tamoxifen is a selective oestrogen receptor modulator (SERM), which once metabolised into its active form has a high affinity for the oestrogen receptor and in this way blocks it.[56] Tamoxifen is biologically inactive or has relatively low activity until after it has been metabolised in the liver, mainly by the cytochrome P450 CYP2D6, into 4-hydroxy-tamoxifen (**Figure 1.10**).[57] This is known as a prodrug approach. Figure 10 shows the structure of tamoxifen and its active form after metabolism in the liver.



Tamoxifen (TAM) 4-hydroxyTAM

Figure 1.10: Structure of tamoxifen and 4-hydroxyTAM.

Since the discovery of tamoxifen, scientific understanding of human cancers and the different signalling pathways has evolved significantly allowing new targets for anticancer treatments to be pursued.

Bcr-Abl is a gene sequence associated with different human leukemias.[58] The discovery of imatinib (Gleevec), a drug that targets this bcr-Abl kinase, changed the course of cancer chemotherapy towards more targeted approaches.[59]

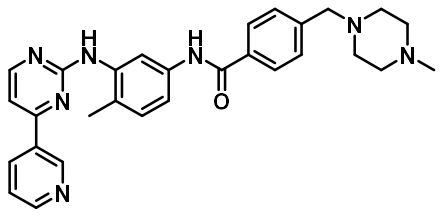


Figure 1.11*:* Structure imatinib (Gleevec).

The structure of imatinib (**Figure 1.11**) allows it to act as a tyrosine kinase enzyme inhibitor.[60] It binds specifically to bcr-abl and because it binds competitively it inhibits ATP.[61] Substrates cannot enter the kinase site and therefore the tumour cell cannot proliferate as the downstream signalling pathway has been blocked. **Figure 1.12** shows a simplified scheme of how imatinib acts as a competitive inhibitor of ATP to inhibit kinase activity.[62] It has been used to treat several diseases but most commonly chronic myelogenous leukemia and acute lymphocytic leukemia.[63] This drug is a chronic therapy and is not used to cure disease but more to control disease progression.

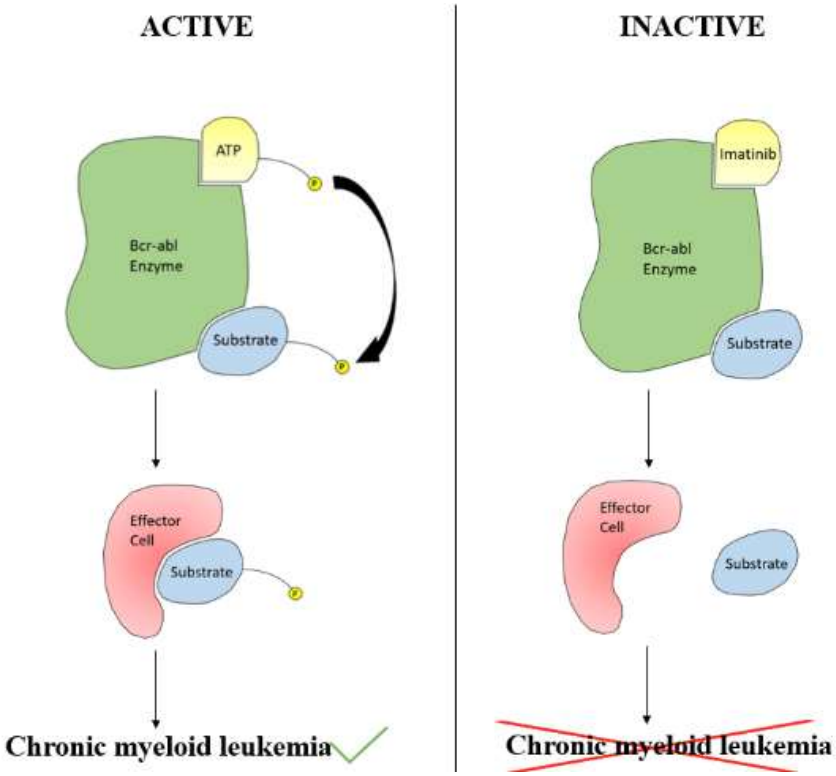


Figure 1.12: *Activity of bcr-abl enzyme by ATP activation vs Gleevec inhibition.*

Other examples of the pathway targeting approach include angiogenesis inhibitors. These are small drug molecules that work by blocking a process called angiogenesis, which is responsible for making new blood vessels. This is a necessary process for tumours to grow.[64] Angiogenesis inhibition stops the tumour from being able to grow as the blood vessels that supply essential nutrients and oxygen are blocked, in this way the tumour is indirectly targeted and starved of the necessary ingredients for growth.

Examples of angiogenesis inhibitors include bevacizumab (Avastin), approved in 2004 to treat colon cancer.[65] Thalidomide, originally used to treat morning sickness in pregnant women, was also reported to be an angiogenesis inhibitor and was used to treat multiple myeloma since 1999.[66] There are several other examples and a larger number in preclinical and clinical trials. Angiogenesis inhibitors have fewer side effects and a lower risk of drug resistance than other compounds.[67]

1.3.2 Peptide drug conjugates – targeting cytotoxics to the tumour

A second approach to targeted therapy involves targeting a cytotoxic agent directly to the tumour. One example of doing this is through the use of peptides.

Peptides can themselves be used as a cytotoxic agent or can be bound to a cytotoxic agent as a targeting moiety. Peptide drug conjugates (PDC) harness the potency of cytotoxic agents by binding them to specific peptide sequences. A PDC is made up of three components; the targeting peptide, the linker and the cytotoxic payload, (**Figure 1.13)**. Some tumours overexpress proteins, for example, epidermal growth factor receptor (EGFR) or human epidermal growth factor receptor 2 (Her2). Peptides are selected that target these, allowing specific binding to protein receptors and therefore targeting the payload directly to the tumour. The linker is used to bind the payload to the targeting peptide. The linker is important as it must show stability when in circulation and not prematurely release the payload. The main class of linkers include; cleavable, non-cleavable, pH-sensitive, enzyme sensitive and redox sensitive.[68] The choice of payload can vary, however, in PDCs the main examples used are; doxorubicin, gemcitabine, daunorubicin, paclitaxel, camptothecin.[69]

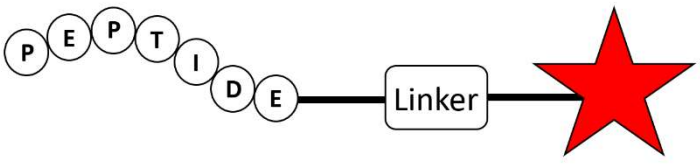


Figure 1.13: *Simplified Structure of a Peptide Drug Conjugate.*

Two PDCs have been clinically approved.[70] The first to be approved in 2020, ^{177}Lu-DOTATATE, is used for the treatment of gastroenteropancreatic neuroendocrine

tumours (GEP-NETs).[71] This is a radioactive peptide drug-conjugate as it combines the radionucleotide Lu-177 to a targeting peptide specific to the somatostatin (SST) receptor. The homing peptide is a somatostatin analogue called DOTA-TATE **(Figure 1.14)**.[72] This is a cyclic peptide formed by a disulphide bridge, this makes it more stable to degradation in the cell. The radionucleotide is shown in red and it is conjugated via an amide linker to the somatostatin peptide shown in black. The peptide radionucleotide delivers ionising radiation directly in the tumour environment causing DNA single and double strand breakage and therefore cell death.[73]

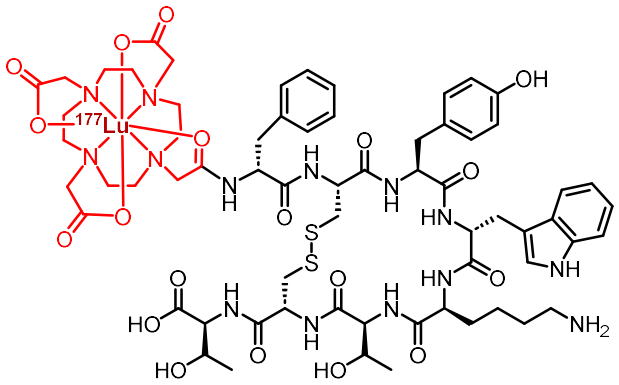


***Figure 1.14**: Structure of PDC ^{177}Lu-DOTATATE.*

Several more PDCs are in clinical trials, including one in phase III clinical trials with a paclitaxel payload for the treatment of breast cancer.[74]

Peptides have advantages over other therapies, including other targeting therapies. Peptide drug conjugates allow specific control of the number of payloads attached to the targeting agent. The small size of peptides, which range from 5 to 25 amino acids in length, means that tumour penetration is improved.[75] However, there are several drawbacks to using peptides, including their lack of stability *in vivo*. The half-life of a peptide in circulation is determined by a class of enzymes called exopeptidases, which can break the peptide down. PDCs show promise for the future, however there is a lot more optimisation needed for more to reach the clinic.

1.3.3 Antibody drug conjugates – targeting cytotoxics to the tumour

Another approach to targeting the cytotoxic drug directly to the tumour is to use an antibody. Antibody-drug conjugates (ADCs) were first described in the 1980s and

success in clinical trials was seen for the first time in 2000.[76] Over the last decade there has been a significant increase in the number of ADCs making it to market for cancer treatments, especially in this last year; with eleven on the market by June 2023, and fifteen by November 2023.[77,76] With over 100 ADCs in preclinical and clinical trials, they represent an important class of products for future cancer therapies.[78]

The small therapeutic window of chemotherapeutic drugs and their high side effects has severely limited their usefulness in the clinic. In order for untargeted cancer drugs to have an effect on the patient, the dose must be sufficiently high, and in practice is close to the maximum tolerated dose (MTD), resulting in off-target toxicity and hence unwanted side effects. ADCs are able to increase this MTD and lower the efficacious dose (MED) as they are more specific to the tumours and have less off-target effects, hence an increase in the therapeutic window (**Figure 1.15**). [79]

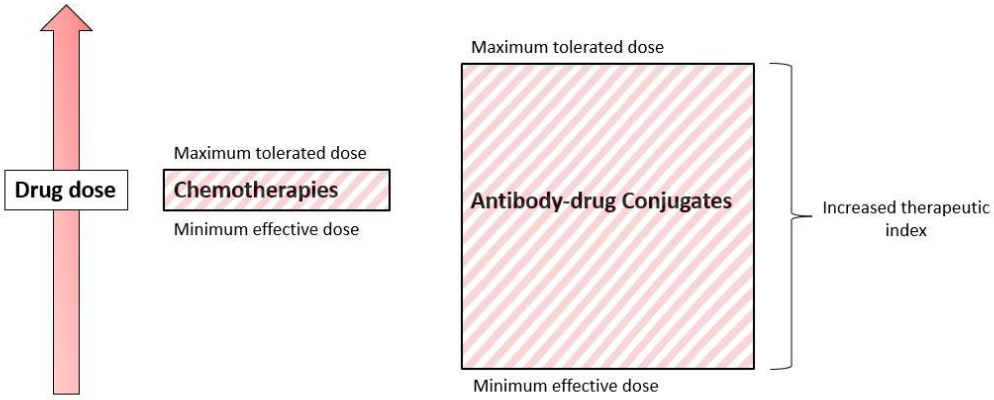


Figure 1.15: *An illustration of the increased therapeutic index that ADCs offer.*

ADCs are comprised of monoclonal antibodies (mAbs) bound via a linker to a highly potent drug. **Figure 1.16** shows the generic arrangement of an ADC.[80] On the surface of some tumours some antigens may be overexpressed that are not present on healthy cells.[81] ADCs exploit this unique environment by using an antibody specific to the tumour antigens. In this way the antibody acts as the targeting device and is able to direct a highly toxic payload to the tumour environment without affecting the healthy cells. The linker, shown in green in this schematic, is important as it enables the drug and antibody to remain bound while in circulation and therefore 'safe' until it reaches the targeted site.

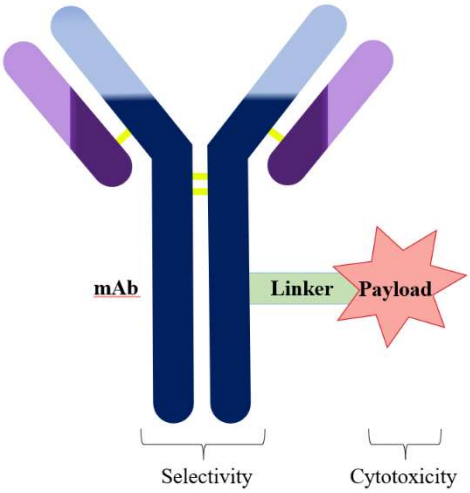


Figure 1.16: *The simplified structure of an ADC.*

The antibody is roughly Y shaped. All antibodies are made up of two heavy (blue) and two light (purple) polypeptide chains and are referred to as immunoglobulin G antibodies (IgG).[82] IgG antibodies are linked together by disulphide bonds (represented in yellow), with two disulphide bonds between the heavy chains and one between the heavy and light chains. The top of the antigen, the V region (shown as light purple and light blue), is the site responsible for antigen binding. This part of the antibody can vary and is often referred to as the Fab region (the fragment antigen binding region).[83] The stem of the two heavy chains (dark blue) are less variable and are involved in interacting with other molecules. This region is often referred to as the Fc region (the fragment crystallisable region or the constant region).[83]

All ADCs have the same general mechanism of targeting cells and killing them, illustrated in **Figure 1.17**.[84] The mechanism of action of the ADC can be broken down into seven steps; The first step is the ADC in the bloodstream and in circulation. The second step is the binding step, here the monoclonal antibody, recognises and binds to specific antigens present on the tumour surface. The ADC is then internalised while bound to the antigen in a process called receptor-mediated endocytosis.[85] The ADC-antigen complex is internalised in an endosome, where it is transported within the cell. This then matures to a lysosome which has conditions that are able to break down the ADC complex.[86] The ADC is then 'activated' as the linker either cleaves or breaks down to release the toxic payload. The free drug is then able to exert its effect, which

will vary depending on the mechanism of the warhead. Finally, the cell will undergo apoptosis.

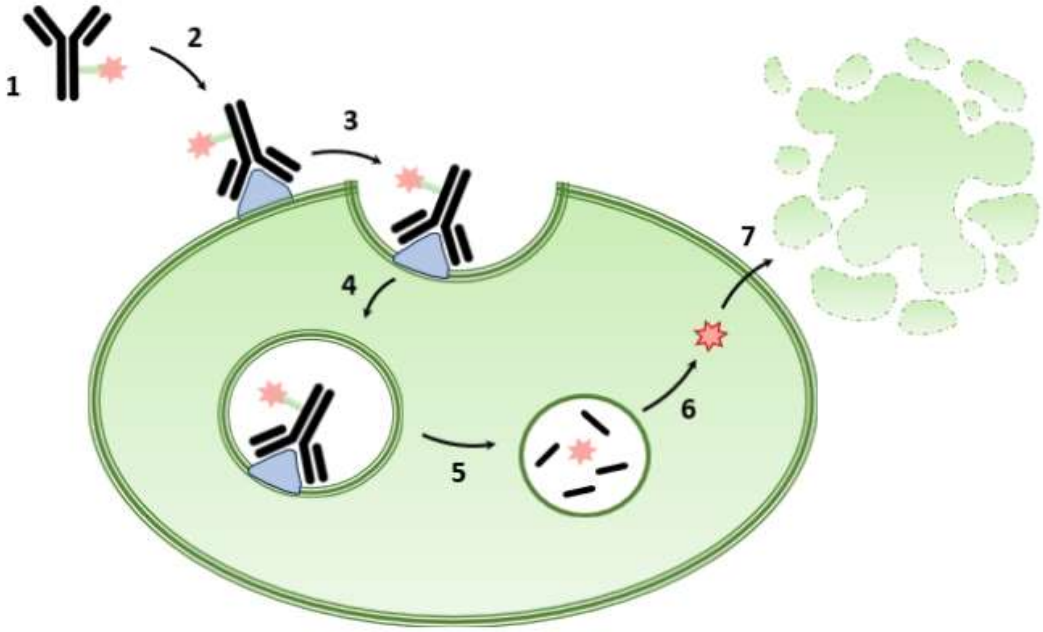


Figure 1.17: *Mechanism of an ADC on a cancer cell broken down into seven steps.*

1.3.3.1 Payloads in ADCs

Payloads that have been too cytotoxic as prodrugs or as chemotherapeutic agents on their own have been prioritised for ADCs, as their potencies can now be harnessed in a safe, effective way. Payloads for these compounds need to be potent cytotoxic compounds so they can effectively destroy the cancer cells once delivered.

In 2000 the first ADC approved was gemtuzumab ozogamicin (Mylotarg) (**Figure 1.18**). This was later withdrawn after it showed increased fatality in the treatment of myeloid leukaemia.[87] This ADC was made up of an anti-CD33 mAb bound via a cleavable hydrazone linker (shown in black) to a calicheamicin derivative (red).[88] The calicheamicin derivative is *N*-acetyl gamma calicheamicin dimethyl hydrazine, this is a minor groove DNA binder and induces double-stranded breaks, cell cycle arrest and hence apoptosis.[89]

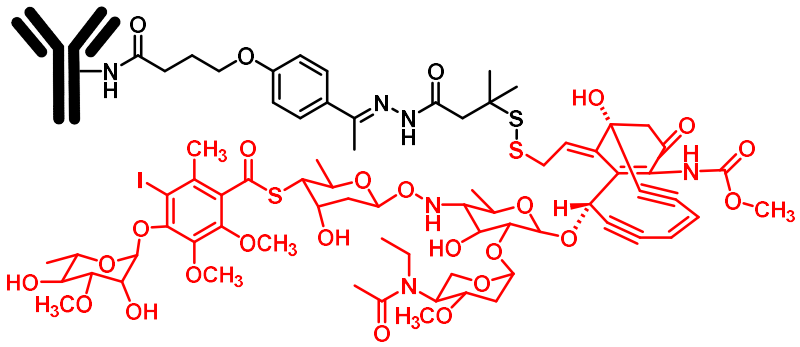


Figure 1.18: *Gemtuzumab Ozogamicin (Mylotarg).*

Despite the first ADC being withdrawn, there were continued efforts to develop this class of compounds. Since trastuzumab emtansine and brentuximab vedotin reached the clinic, this is one of the largest areas of research within oncology, as these demonstrate advantages in increased therapeutic index that classical treatments fail to offer.[90]

The payloads used in ADCs generally fall into two categories; DNA-damaging agents and microtubule-targeting payloads.[91] This thesis will focus on DNA alkylating agents (in particular the duocarmycins), despite there being more microtubule targeted payloads in clinical trials. There is a shift of focus with ADCs towards solid tumour treatment, which could result in DNA-damaging agents being of more interest due to potency.[92] The payloads that disrupt DNA synthesis that are promising for ADCs are pyrrolobenzodiazepine (PBD) dimers, pyridinobenzodiazapines (PDDs), indolinobenzodiazapines (IGNs) and the duocarmycins.[79]

1.3.3.2 Linkers in ADCs

Linkers in ADCs can be cleavable and non-cleavable. There are three features of an effective linker; most importantly that it is stable when in circulation, but the linker must at the same time give effective release of the active payload when in the tumour environment. Furthermore, the linker needs to have high water solubility, this helps stop ADC aggregation.[93]

Cleavable linkers release the payload once exposed to certain conditions, so that the payload can have its potent effect. There are three main categories of cleavable linker and their mechanism of release is dependent on different conditions; enzyme cleavability, pH sensitivity and glutathione sensitivity.

Hydrazone linkers (as found in gemtuzumab ozogamicin) are pH sensitive and are stable in neutral and alkaline conditions, however in more acidic environments they are very labile. In this way the design of these ADCs takes advantage of low pH in the endosome and lysosome to trigger the hydrolysis of the acid-labile hydrazone linkers. Disulfide-linkers are also stable while in circulation in the plasma, however once internalised there is elevated glutathione in cells which then cleaves the linker to release the payload.[94]

Enzyme cleavable linkers are generally preferred over reducible disulfides and acid-sensitive linkers as they are more stable in plasma.[95] This stability is due to the linkers being cleaved by specific enzymes (lysosomal proteases) that are overexpressed in the target cells.[93]

Gemtuzumab ozogamicin, shown previously (**Figure 1.18**), is an example of an ADC with a cleavable linker. The cleavable linker is shown in black, and this binds together the targeting antibody gemtuzumab and the calicheamicin shown in red.

Non-cleavable linkers are linkers where the mAb degrades after it is internalised in the cell and this releases the active payload into the lysosome. Non-cleavable linkers can be divided into two groups; thioether and maleimidocaproyl (MC). ADCs with non-cleavable linkers are not at risk of being cleaved in circulation as they require the entire antibody to degrade before the payload is released, and for this reason they have much lower risk of systemic toxicity.

So far there are only two non-cleavable ADCs on the market: trastuzumab emtansine (Kadcyla) and belantamab mafodotin .[96] **Figure 1.19** shows the structure of Kadcyla, the first non-cleavable ADC, and used for the treatment of HER2 positive breast cancer.[97] Kadcyla is an ADC made up of a payload emtansine (red), a MCC non-cleavable linker (black) and the anti-Her 2 antibody.

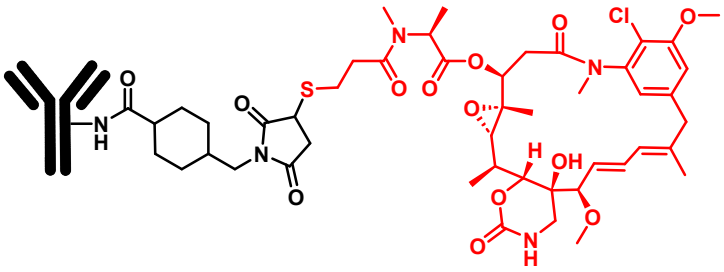


Figure 1.19: *Structure of Trastuzumab emtansine (Kadcyla).*

When comparing cleavable with non-cleavable linkers, there are benefits and downsides to each. Non-cleavable linkers have higher plasma stability, so these ADCs tend to have an increased therapeutic index.[98] However, cleavable linkers provide increased bystander effects (due to early release), i.e., they are able to kill cancer cells surrounding the target.[98] Also, the payload is released from cleavable linkers on its own, without the linker still attached, meaning it often has more activity as its structure is not changed or compromised.[95]

1.3.3.3 Drug Antibody Ratios, DARs

The linker is also important as it determines where and how many payloads attach to the antibody. The potency and lifetime of the ADC is usually determined by this drug-antibody ratio (DAR). The drug to antibody ratio is the average number of drugs that are bonded to the antibody and ranges from DAR 0 to DAR 8. A DAR>4 can lead to aggregation, as well as high DAR ADCs having faster clearance.[99]

Linkers, payloads and control of the DARs have evolved significantly since the first ADC, in fact this class of compounds is evolving progressively from the first, second third and next generation.

1.3.3.4 Next-Generation ADCs

First-generation ADCs such as gemtuzumab ozogamicin showed the potential of this class of compound However, they had some significant shortcomings such as; potency, ADC internalisation, tumour localisation, linker stability, target antigen

expression and immune response.[100] Significant improvements were needed for these drugs to be effective and safe to administer.

Second-generation ADCs used more potent payloads and addressed some of the other shortcomings of the first generation. The payloads of second-generation ADCs were between 100 and 1000 fold more potent in comparison to those of the first generation.[101] Struggles of second generation were issues with the final DAR ratios as good coupling techniques had not yet been developed. These issues including off-target effects contributed to the drugs having a narrow therapeutic index. There was also still limited tumour penetration and high resistance towards these drugs.

Third-generation ADCs have improved linkers and conjugation chemistries, which has benefitted the therapeutic index of these drugs. The DAR is more controlled in third and next-generation ADCs, with an average DAR of 2 or 4.[87] Because of these improvements, more potent payloads can be used as they can be delivered safely, and payloads are now being seen in the pM range.[102] Using more potent compounds or compounds with different mechanisms of action can help to combat intratumour heterogeneity which is seen in many cancers. Next-generation ADCs tend to use the following compounds as payloads; cytokines, lysosome inhibitors, PBD dimers and duocarmycins.[103]

Tiberghein described the use of PBD dimers, which is the use of dimeric payloads and dual warheads.[104] Studies have shown that dual-dual payloads in an ADC have a better clinical outcome than using two different types of drugs.[105] This is something that will be discussed further in **Chapter 3**.

Next generation duocarmycin-based ADCs include a seco-CBI-dimer e.g. CD22-SN36248 and a DCM-PBD or DCM-PDD dimer. Pcm5B14-DCM is another duocarmycin-based ADC which has shown promise. DCM-SA, adozelesin and KW-2189 show activity against cancer cell lines showing drug resistance which is one reason why they are an excellent group of compounds to investigate further.[106] Duocarmycins have also been shown to work with both cleavable and non-cleavable linkers so there are more options and avenues to explore. [96]

1.4. The Duocarmycins

The duocarmycins make excellent candidates for cancer treatment due to their potency, however, due to a lack of selectivity, they have not yet made it to market as effective cancer treatments. If this potency can be harnessed so they are safe in the presence of healthy cells and only become activated in the presence of tumour cells, then they have a much higher chance of becoming druggable. Studies have been carried out to try and harness the potency of these compounds by using them as prodrugs, however their small therapeutic window and high toxicity has limited this progress.

The duocarmycins are a family of natural anticancer antibiotics (and their synthetic analogues) that have highly potent antitumour activity in cancer cell lines. The first member of this family, CC-1065, was isolated in the 1970s from a *Streptomyces* species. More of these compounds were then discovered to make up a family of natural products, all showing potent anti-tumour activity.[107] There are four key members of the duocarmycin family, CC-1065, duocarmycin A, duocarmycin SA and yatakemycin. Their structures are shown in **Figure 1.20**. They consist of different alkylating subunits, including the duocarmycin SA unit, DSA, one of the most potent subunits and of great interest in the design of analogues.[108]

CC-1065 was the first of this group of compounds to be discovered, by The Upjohn Company who also characterised it.[109] This compound showed high potency *in vitro* but only modest activity *in vivo*.[110] Duocarmycin A was isolated by the Japanese company Kyowa Hakko Kogyo in 1988.[111] This was more potent than CC-1065 and did not show delayed fatal toxicity.[112]

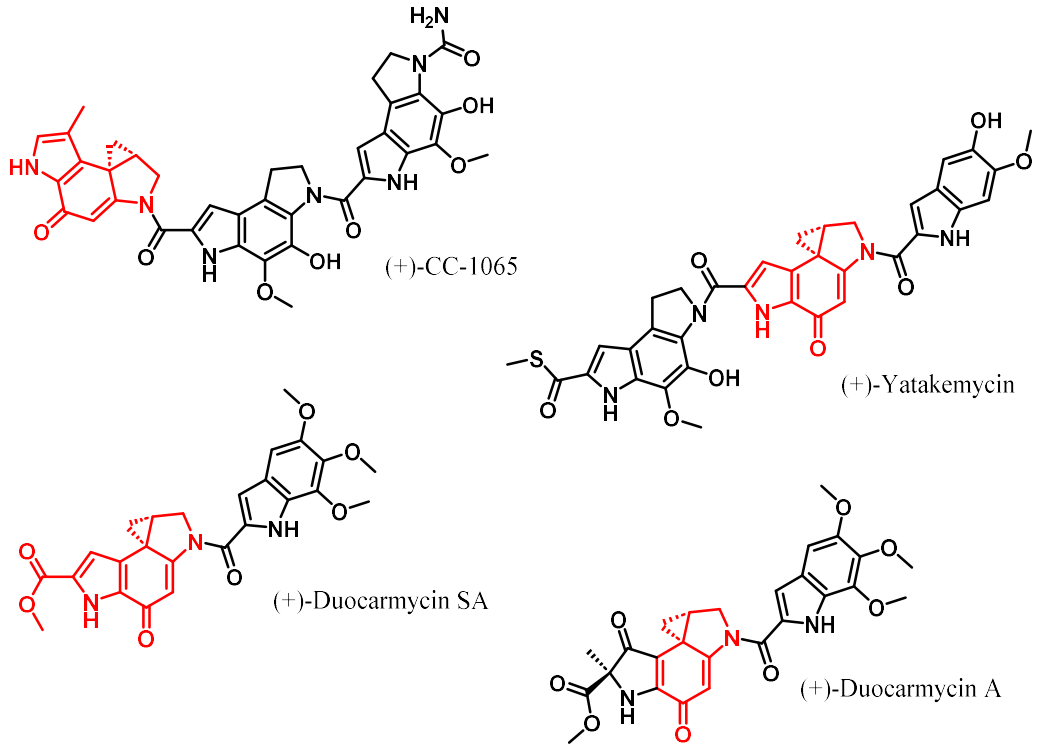


Figure 1.20: *Structures of CC-1065, duocarmycin SA, duocarmycin A and yatakemycin.*

In 1990 Duocarmycin SA was discovered and was found to be more potent than both its predecessors.[111] It showed improved stability and potency, and like duocarmycin A did not show delayed fatal toxicity. [111] Yatakemycin, the most recent member of this group of compounds to be discovered in 2003, has an unusual 'sandwiched' structure.[113] Unlike the other members in the duocarmycin family it has the alkylating unit between a left- and right-hand binding unit. Note that the left binding unit is similar to the binding unit of the first member of this group to be discovered, CC-1065. It can also be seen from the structures that yatakemycin and duocarmycin SA have the same alkylation unit, these are the two most potent members of this family of compounds.

Yatakemycin has a slightly improved potency despite the alkylating subunit being the same. Studies by Boger on around 50 analogues showed that it is the arrangement of these three subunits in the "sandwiched" order that gives yatakemycin the increased rate and potency of DNA alkylation.[114] Despite there being some differences in potencies within this group of compounds, all of the duocarmycin family are able to

alkylate DNA with picomolar IC_{50} values in the L1210 murine tumour cell line, showing that most have the potential to be used in cancer treatments.[115]

1.4.1 Mechanism of action of the duocarmycins

In order to investigate the mechanism of action of the duocarmycins, Boger has synthesised several alkylating units (**Figure 1.21**).[116] These compounds are relatively unreactive until they reach their biological target, Boger assessed the stability of these units at pH 3 (at pH 7 some were too stable, so the rate of hydrolysis could not be measured). It was found that stability corresponds to biological activity. Unstable units or units that were too stable had little biological activity, because stability appears to be correlated with its ability to reach its site of action in the cell. The alkylating unit of duocarmycin SA, DSA, is one of the most potent subunits.[117] The cyclopropabenzindol-4-one (CBI) alkylation unit, a synthetic analogue not found in nature, was shown to be a good alternative to that of DSA, as it is almost as potent and stable as duocarmycin SA. It is four times more stable and more potent than the alkylating unit of CC-1065.[118]

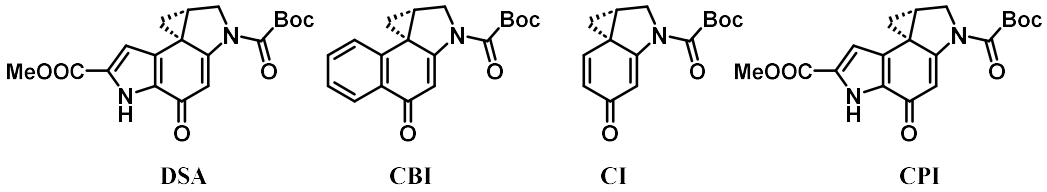


Figure 1.21: *Common alkylating sub-units.*

Mechanistically, all members of the duocarmycins alkylate DNA in a similar way. They bind selectively to AT-rich regions in the minor groove of DNA.[115] On binding, the planar curved shape of the compounds is twisted slightly to adapt to fit the helical shape of DNA. This binding-induced conformational change causes the chemistry of the duocarmycins to be 'switched on' so they are more reactive alkylators.[119] Boger *et al* carried out studies to investigate this binding-induced activation by comparing DSI and duocarmycin SA. DSI differs from DSA as there is a simple indole instead of the trimethoxyindole-binding subunit in the compound (**Figure 1.22**). The red arrow in **Figure 22** highlights where this binding-induced twisting occurs, and hence how the extended conjugation is disrupted. The most notable of several findings was that there

was a bigger twist with duocarmycin SA between the alkylating subunit and the trimethoxyindole DNA binding unit, correlating in ~20-fold faster alkylation rate of DNA.[120] The conjugated amide stabilises the α,β-unsaturated ketone making the cyclopropane less reactive.[121] It is this twisting that disrupts the vinylogous amide stabilisation and hence the larger the twist the more activated the cyclopropane ring becomes.[122]

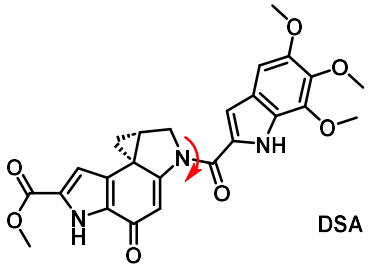
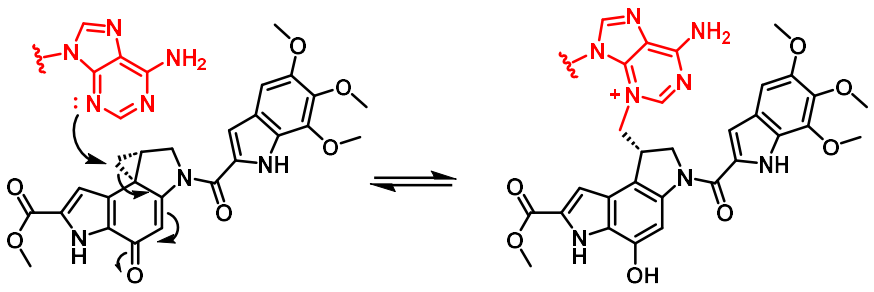


***Figure 1.22**: Structure of DSA, with arrow showing where the binding induced conformational twist occurs.*

On binding to the DNA, alkylation with the adenine base occurs at the N3 position with the least substituted carbon of the highly strained cyclopropane ring. The mechanism for the alkylation between an adenine base in DNA and duocarmycin SA can be seen in **Scheme 1.1** where the adenine base is shown in red and the duocarmycin compound in black.[121] The nitrogen lone pair attacks the highly reactive cyclopropane strained ring and a driving force for the reaction is this strain release and aromatisation of benzene.



***Scheme 1.1:** The alkylation mechanism of DSA to adenine base in DNA.*

The mechanism for alkylation of DNA is reversible. This is true for most members of the duocarmycin class, except for CC-1065, which alkylates DNA irreversibly.[123] Despite the reaction being reversible, DNA repair mechanisms do not work after

alkylation due to a distortion of the DNA.[124] As a result of this, alkylation leads to cell death (apoptosis).

Unlabelled duplex DNA can be used to follow the rate of this reversible reaction.[125] Different duocarmycins have different alkylation selectivity.[108] The stereochemistry of the duocarmycins also determines how these molecules alkylate DNA. It was found that natural CC-1065 and DSA alkylate the 3'-terminal adenine of an AT-rich sequence, whereas unnatural enantiomers bind in the reverse orientation and alkylate at the 5'-end.[126] This is represented in **Figure 1.23** where the natural (+)-duocarmycin SA is shown bound to a DNA helix.

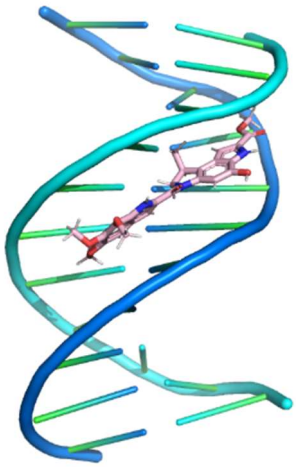


Figure 1.23: *Shows (+)-duocarmycin SA bound to adenine in the minor groove of DNA.*

Yatakemycin differs slightly in that both the natural and unnatural enantiomers of yatakemycin alkylate the central adenine in an AT-rich sequence.[108] In nature the duocarmycins are produced are shown to be 100 to1000 fold better at alkylating DNA than the unnatural enantiomer.[127]

1.4.2 The seco form – duocarmycins

In most synthetic routes to the duocarmycin alkylating subunit, the precursor to the ring closed, cyclopropane form is the 'seco' form, or the ring opened analogue. This often has the structure shown in **Figure 1.24**, where '**X**' is a good leaving group and R=H. Any group that can donate a pair of electrons and cause spirocyclization can be used in place of the hydroxyl.[128] Tercel *et al* reports that the DSA in the phenol form

(R=H) is the most potent member of this class.[128] In nature the seco form is also found, which spontaneously cyclises in a biological system to form the cyclopropane group prior to DNA alkylation. This form can also be protected at the OH (for example in this work with an OBn) which prevents spirocyclisation and makes duocarmycin precursors much safer to handle. Deprotection can then be carried out on a small scale when the compound is ready for biological testing.

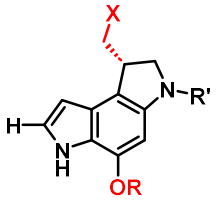
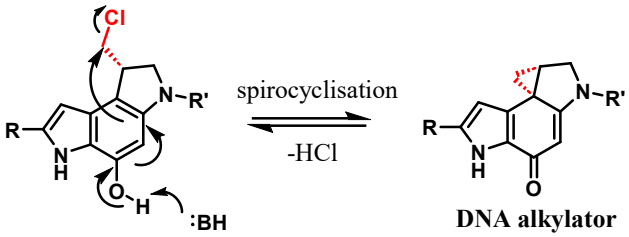


Figure 1.24: *The structure of protected seco-DSA.*

The seco form can become activated in different ways, this can depend on the 'R' group. Ways to do this include, in basic conditions if the 'R' group was a phenol then spirocyclization follows deprotonation (**Scheme 1.2**).[129] The mechanism of cyclisation is shown in **Scheme 1.2** where X=Cl and R=H. Studies have shown that seco phenolic derivatives have no difference in toxicity than when the cyclopropane is already present giving evidence for fast cyclisation.[128] It has been shown that the alkylating ability and potency correlates to the electron-withdrawing ability of this 'X' group.[129]



Scheme 1.2: *Mechanism of spirocyclisation of phenolic DSA.*

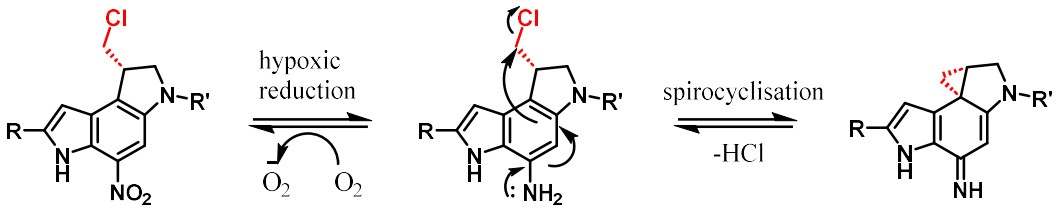
Studies of the seco for of the duocarmycin alkylating subunits have allowed the development of prodrug strategies.

1.4.3 Prodrug strategies

A prodrug is defined as a compound that is pharmacologically inert/inactive that is converted *in vivo* to the active drug which can then have a therapeutic effect.[130] The duocarmycins as stand-alone agents have failed to reach the clinic due to their toxicity and small therapeutic window (refer to **Figure 1.15**). A number of strategies, principally prodrug and targeting strategies, have been developed to try and improve this therapeutic window. The referenced review by Pors *et al* in 2020 provides an overview of all the prodrug approaches that have been applied to the duocarmycins to try and improve their success in the clinic.[127] Some of these approaches will be discussed however for a more detailed review see Pors *et al*.

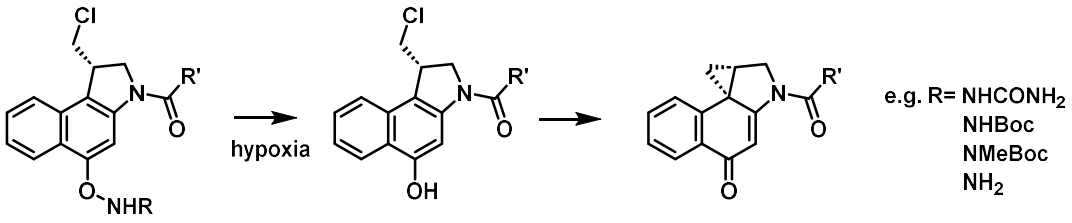
1.4.3.1 Hypoxia targeted prodrugs

Environments such as hypoxia can be exploited with these prodrug approaches. As a solid tumour grows away from the proximity of the blood supply, it develops a layer of hypoxic cells, which survive in a low oxygen environment.[131] These cells, which eventually die as the tumour grows further, can potentially reoxygenate when the tumour mass is decreased through radiotherapy or chemotherapy, as they are inherently resistant, and can cause the tumour to reestablish leading to relapse. Referring to **Figure 1.24**, where the phenol group is replaced by a nitro group, the seco form is then inactive. In the presence of hypoxic tumour environments (**Scheme 3**), which are inherently a reducing environment, there will be conversion to an amine.[132] The amine is then able to donate a lone pair of electrons and spirocyclisation occurs forming the cyclopropane ring. While this approach has been shown to work in model systems, these chemotherapeutic agents are not selective enough and are too cytotoxic *in vivo*.



Scheme 1.3: *Hypoxic reduction of prodrug in decreased oxygen environment of tumour.*

Boger *et al* show how you can have a stable phenol protecting group on the prodrug and then activate it using reductive activation (**Scheme 1.4**) ensuring that the compound is only active *in vivo* or at the final stages of synthesis.[133]



***Scheme 1.4**: A simplified scheme of the prodrug requiring reductive activation.*

Here the prodrug is cleaved by reductive activation at the weak N-O bond and the compound is then able to spirocyclise once in the seco form to the final compound shown, the active CBI compound. Boger reported a library of *N*-acyl *O*-aminophenol prodrugs of CBI-indole$_2$ that were all able to undergo nucleophilic cleavage of this N-O bond. This study highlighted the ability of these compounds to act as a prodrug with varying stabilities based on the R group to be activated in hypoxic tumour environments and remain stable in normal environments.[134]

1.4.3.2 Carbamate and carbonate prodrugs

These prodrugs work by incorporating either a carbamate or a carbonate to the parent drug which masks the potency or toxicity until the drug is in the desired environment. A carbamate is made up of a carbonyl group bound to an amino group and has the general structure OC(O)NHR. A carbonate is made up of a carbonyl group bonded to two oxygen atoms with the general structure OC(O)OR. Once these carbamates or carbonates are in a tumour environment, they are cleaved *in vivo*, often using an enzyme to release the parent drug which can go on to have its cytotoxic effect.

In 2012 Boger reported the synthesis of a prodrug of seco-CBI-indole$_2$ that utilises a cyclic carbamate that was stable and cleaved by hydrolysis. There are several examples of duocarmycins utilising this approach, but one that has been most successful is KW-2189. Another successful example was carzelesin. The structures for both these cyclopropylpyrroloindole compounds are shown in **Figure 1.25** with the masking groups highlighted in red. [127]

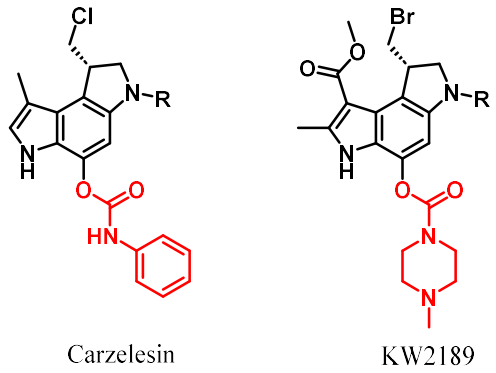


Figure 1.25: *Structures of Carzelesin and KW2189.*

Carzelesin is based on CC-1065 and is a carbamate prodrug and KW-2189 is a duocarmycin B2 carbamate-based prodrug.[135] They vary in how they are both activated. Carzelesin is first hydrolysed before it closes the ring to form the active spirocyclised compound.[136] KW-2189 is activated by enzyme hydrolysis and although based on CC-1065 (which in trials showed delayed death) does not show this same side effect. This was then suspected to be because of a differing right hand binding unit.[137] Both these drugs were tested in phase II clinical trials. However, hematologic toxicity was seen when patients were administered the required dose, so the trials were not continued.[129,138]

1.4.3.3 Glycoside prodrugs

These prodrugs work by incorporating a sugar moiety via a glycosidic bond to the parent drug which masks the potency or toxicity until the drug is in the desired environment.

Tietze reported work that targets drugs using antibody-directed enzyme prodrug therapy (ADEPT), where he uses glycosidic prodrugs bound to a targeting antibody in a way making these compounds even safer.[139] He also discusses using these drugs as a monotherapy without the antibody, as the prodrug approach means that they only become activated in conditions specific to tumour environments.

Duocarmycin SA based compounds were synthesised it in the seco form with different sugar compounds bound to where the phenol would be. In doing this, the active drug

will be released once cleaved using glycohydrolases releasing the seco drug which can then spirocyclise to become active.

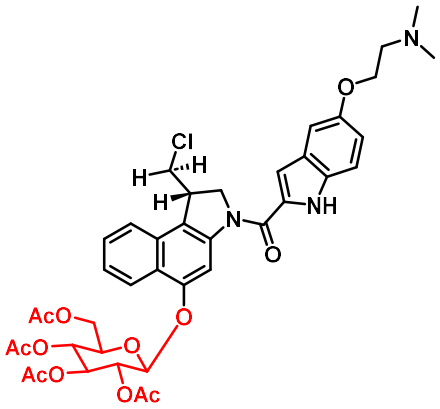
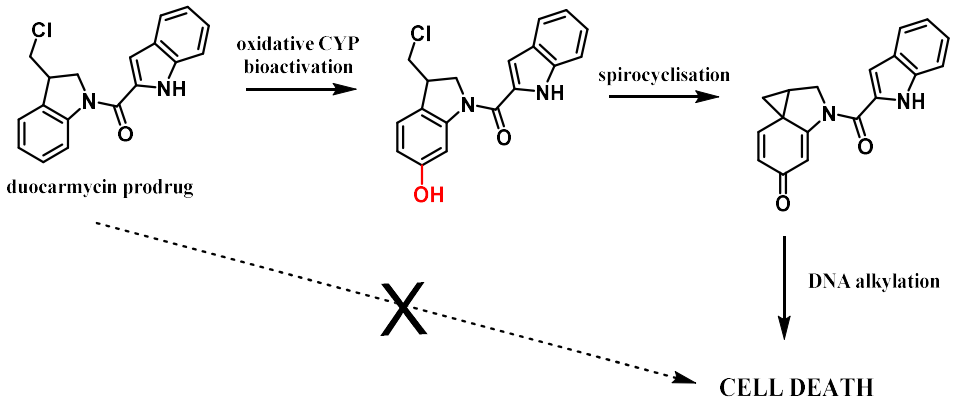


***Figure 1.26**: (D-man) duocarmycin SA sugar conjugate prodrug developed by Tietze.*

Tietze reports a library of different prodrugs all with differing sugar moieties and discusses their stability and their activity with and without enzyme activation. The compound he found to be most active is shown in **Figure 1.26** where the protecting sugar is highlighted in red.[140] This is a prodrug containing an α-mannoside moiety with 94% stability over 24 hours and an IC_{50} of 0.6 nM.[140]

1.4.3.4 Bio oxidative duocarmycins

Searcey *et al* investigated cytochrome P450 bioactivation of duocarmycin prodrugs. They describe the synthesis of compounds that exploit cytochrome P450 enzymes that are responsible for genetic mutations.[141] They aimed to synthesise duocarmycin analogues that are missing the key 'warhead', the cyclopropane ring, similarly to how the other methods mask the toxicity. They aimed to make one that could then undergo regioselective aryl oxidation by CYP1A1 (a cytochrome P450 enzyme overexpressed in tumour environments) give the seco-duocarmycin, which is able to spirocyclise.[141]



Scheme 1.5: *Duocarmycin prodrug undergoing oxidative bioactivation using CYP resulting in spirocyclisation, DNA alkylation and cell death.*

Searcey and Pors show a scheme which highlights the principles behind this approach **(Scheme 1.5)**.[141] The precursor is unable to cause cell death due to the cyclopropane ring not being present. However, once activated by oxidation using the CYP enzyme, spirocyclisation can occur and the duocarmycin analogue can perform its cytotoxic effect.

1.5. Targeting Duocarmycins – ADCs and PDCs

So far, prodrug approaches to try and get the duocarmycins to market have failed, because of their high toxicity and low selectivity. Efforts to synthesise less toxic analogues have also been tried but failed as myelotoxicity has meant all analogues have failed in clinical trials.[115]

Targeting strategies provide the best solution to try and get these highly potent compounds to market, and examples of peptide conjugates and antibody drug conjugates that have shown promise will be discussed further in this section.

1.5.1 Peptide-duocarmycin conjugate

In 2020 a peptide-duocarmycin conjugate was synthesised and published by the Searcey research group.[142] The Thomas-Friedenreich antigen (TFα) is expressed in 90% of cancer cells.[143] Because it is expressed in many human carcinomas, it makes a desirable target for peptide binding. The peptide sequence they aimed to synthesise

was HGRFILPWWYAFSPS, also known as a TFα-peptide. This sequence peptide has been shown to bind to TFα and therefore inhibit TFα accessibility. The payload was based on duocarmycin, which has appropriate protection for the application in solid phase synthesis.[144]

Solid-phase synthesis is a technique used to build a structure by binding together subunits onto an insoluble material. It is often used to build biological molecules including peptides, nucleic acids and oligosaccharides.[36] This method has advantages over classic solution phase synthesis as larger more complicated structures can be synthesised in smaller building blocks and then joined together sequentially. This type of synthesis is desirable if looking to explore different linear polymers, as once a synthetic route has been found for one of the building blocks, these can be coupled in a step-by-step reaction.

Another advantage to solid phase synthesis is efficiency. Because the growing chain is insoluble, solvents and excess unreacted materials can simply be filtered off. It allows multistep synthesis in one vessel consecutively and avoids repeated purification processes.[37] This is especially effective when designing higher order structures like peptides and peptide conjugates.

The duocarmycin SA alkylating unit that the Searcey group aimed to incorporate into the peptide was an amino acid ester so it was suitable compound for solid phase synthesis once suitably protected. There are different approaches and different protection methods. One approach is to use Fmoc as a protecting group. This is preferred due to the milder deprotection conditions that can be utilised, which avoids decomposition of the DSA.[144]

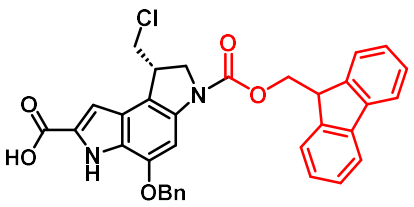
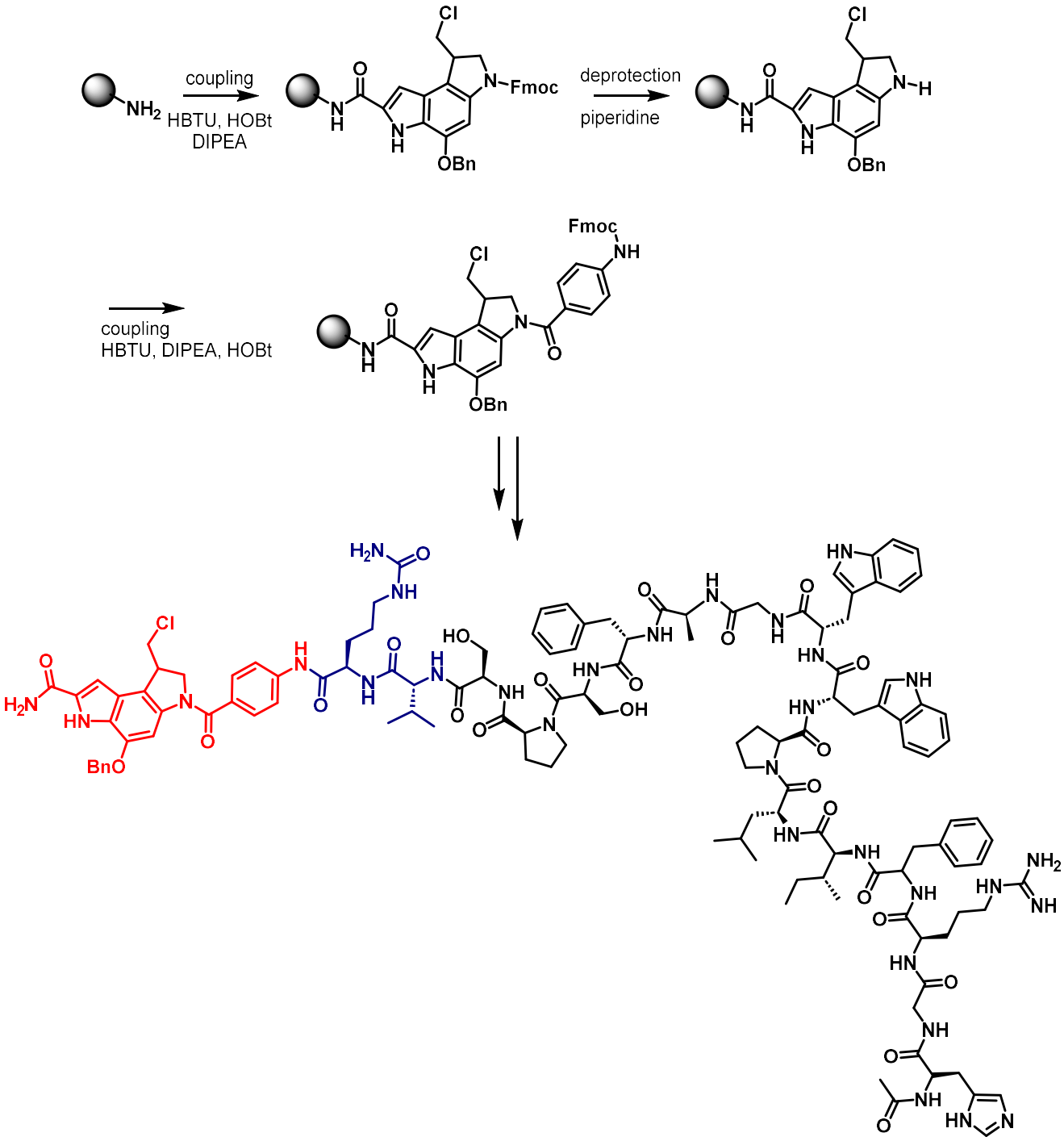


Figure 1.27: *Alkylating unit of DSA with Fmoc on terminal amine.*

Fmoc is 9-Fluorenylmethoxycarbonyl and is shown in red in **Figure 1.27** where it is bound to the DSA alkylating unit. The Fmoc protection on the amine allows the extension of the linear peptide through control of the amide formation steps. After each coupling the protecting group is removed with piperidine. The final compound is then cleaved from the resin using TFA and all protecting groups are removed as well. The design of the PDC incorporated a cathepsin B cleavable sequence, as studies have shown that the use of a cleavable sequence with duocarmycin payloads is favourable. **Scheme 1.6** summarises the solid phase peptide synthesis to incorporate the active payload, the cleavable linker and the targeting peptide by a sequence of couplings and deprotections. The duocarmycin payload is shown in red (this has a para-aminobenzoic acid unit as a spacing unit), the cleavable linker is shown in blue and the TFα binding peptide is shown in black.



Scheme 1.6: *Peptide drug conjugate built on a solid-phase.*

The DSA alkylating subunit in the seco form was successfully incorporated, making the handling of these compounds much safer. Following deprotection of the benzyl group the PDC was tested against seven different cancer cell lines using a noncancerous human cell line as the negative control. Five of the cancer cell lines: breast (MCF-7 and SKBR3), colorectal (HCT116), fibrosarcoma (HT1080), and

promyelocytic leukemia (HL-60) all showed activity in low nM range.[142] A competitive binding assay, which showed that there was a reduced potency on cells in the presence of jacalin (a known TFα binder), showed they bind competitively to the TFα.[142]

These exciting results show the potential for the seco-DSA alkylating unit to be incorporated safely for the production of not just PDCs but also ADCs.

1.5.2 ADCs with duocarmycin payloads.

A review by Wang on the progress of duocarmycin-based ADCs published in 2021 illustrates the strengths and weaknesses of the duocarmycin based ADCs.[106] As the duocarmycins are one of the most potent anticancer agents known, possessing low pM IC50 values in numerous cancer cell lines, they make excellent candidates for payloads to an ADC. There has been success with a duocarmycin based ADC called SYD985 currently under regulatory review, having just passed phase III clinical trials for the treatment of HER-2 metastatic breast cancer.[96] SYD985 has shown promise at killing various cancer cells expressing HER-2. The FDA gave SYD985 (**Figure 1.28**) a Fast-Track Designation status.[106] In the structure of SYD985, there is a short polyethylene glycol (PEG), solubility is an issue with these compounds and their hydrophobic nature can lead to issues of aggregation in the ADCs formed. Incorporation of PEG structures into the linker can help to mitigate these effects. SYD985 has a DAR of 2.8 and uses a protease cleavable linker. The linker is a maleimidocaproyl like linker. There is also a duocarmycin based ADC called MGC018 (vobramitamab duocarmazine) in phase II/III clinical trials for the treatment of prostate cancer.[96]

Seco-DUBA, the analogue of duocarmycin SA used in the ADC is a prodrug. While bound to the antibody the duocarmycin analogue isn't cyclised, its in the seco form. Once cleaved spirocyclisation occurs and the active drug is released.

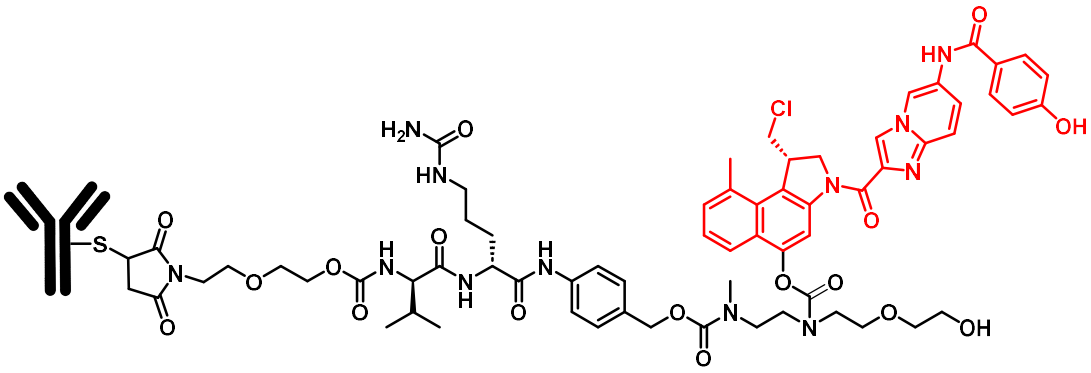


Figure 1.28: *Chemical structure of SYD985.*

Despite SYD985 being fast tracked, it is still a second generation ADC, and there will be new and more promising ideas to be seen in preclinical ADCs with duocarmycin conjugates. Having discussed first, second and third generation ADCs, there is the potential to have significant improvements on DARs, linker technologies and antibody targeting.

Linker design to an antibody is not something this project will look into. It will focus more on developing a highly potent payload. The concept of dimers and higher order structures will be looked at in much greater detail in chapters three and four.

1.6. Aims of Study

Targeted therapies are the focus of modern medicine. Current cancer treatments have many limitations, mostly due to affecting metabolically active and quickly multiplying cells as well as cancer cells. ADCs allow highly potent payloads to be targeted directly to the cancer without affecting healthy cells and hence increase the therapeutic index.

This project aims to synthesise duocarmycin analogues that could potentially be used as a payload for an antibody drug conjugate. The focus will be on synthesising the highly potent DSA alkylating unit in the seco-form and then use this to make dimers. This work will be based upon the previous synthesis described by Searcey, Boger and co-workers. Dimers in ADCs are having success, shown with two CBI dimers in clinical trials.[145] Boc-protection of the DSA amino acid ester will mean that solution

phase synthesis can be applied to this and explore a range of different structures, with focus on preparing dimers. It would be helpful to have access to both stereoisomers so once the racemic mixture has been synthesised, finding a way to separate the enantiomers or explore synthetic routes to make the DSA enantiopure is also something this project explores. This aim of this project is to show alternative routes to synthesise the DSA alkylating unit and discuss improvements over previously followed routes.

The hypothesis is that in using a dimer is that the payload will be more potent and more selective. In order to optimise the delivery, the use of different linkers within the dimer was explored, as well as varying the separation distance between two alkylation units. Once a range of analogues and dimers have been prepared using the DSA alkylating unit then the compounds will be tested for biological activity and DNA crosslinking studies will be performed. This project can have impact, especially as there is one duocarmycin based ADC having just finished phase III clinical trials and one in phase II/III.[96,146] If these show success, there is likely to be a much greater focus on this family of compounds.

References

1 World Heath Organization, Cancer, https://www.who.int/news-room/fact-sheets/detail/cancer, (accessed 9 February 2024).

2 M. Quaresma, M. P. Coleman and B. Rachet, 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971–2011: a population-based study, *The Lancet*, 2015, **385**, 1206–1218.

3 J. P. De Magalhães, How ageing processes influence cancer, *Nat. Rev. Cancer*, 2013, **13**, 357–365.

4 EvaluatePharma, EvaluatePharma World Preview 2019, Outlook to 2024 12th Edition, 2019, 13–14.

5 G. M. Cooper, The Development and Causes of Cancer, *Cell Mol. Approach*, 2000, 743.

6 L. Wyld, R. A. Audisio and G. J. Poston, The evolution of cancer surgery and future perspectives, *Nat. Rev. Clin. Oncol.*, 2015, **12**, 115–124.

7 H. Kitao, M. Iimori, Y. Kataoka, T. Wakasa, E. Tokunaga, H. Saeki, E. Oki and Y. Maehara, DNA replication stress and cancer chemotherapy, *Cancer Sci.*, 2018, **109**, 264–271.

8 R. P. Hertzberg, M. J. Caranfa and S. M. Hecht, On the mechanism of topoisomerase I inhibition by camptothecin: evidence for binding to an enzyme-DNA complex, *Biochemistry*, 1989, **28**, 4629–4638.

9 D. A. Koster, K. Palle, E. S. M. Bot, M.-A. Bjornsti and N. H. Dekker, Antitumour drugs impede DNA uncoiling by topoisomerase I, *Nature*, 2007, **448**, 213–217.

10 A. K. McClendon, A. C. Rodriguez and N. Osheroff, Human topoisomerase iiα rapidly relaxes positively supercoiled dna: implications for enzyme action ahead of replication forks *, *J. Biol. Chem.*, 2005, **280**, 39337–39345.

11 Y. Pommier, Topoisomerase I inhibitors: camptothecins and beyond, *Nat. Rev. Cancer*, 2006, **6**, 789–802.

12 C. J. Thomas, N. J. Rahier and S. M. Hecht, Camptothecin: current perspectives, *Bioorg. Med. Chem.*, 2004, **12**, 1585–1604.

13 X. Fan, X. Lin, Q. Ruan, J. Wang, Y. Yang, M. Sheng, W. Zhou, G. Kai and X. Hao, Research progress on the biosynthesis and metabolic engineering of the anti-

cancer drug camptothecin in Camptotheca acuminate, *Ind. Crops Prod.*, 2022, **186**, 115270.

14 P. Botella and E. Rivero-Buceta, Safe approaches for camptothecin delivery: Structural analogues and nanomedicines, *J. Controlled Release*, 2017, **247**, 28–54.

15 S. Rivankar, An overview of doxorubicin formulations in cancer therapy, *J. Cancer Res. Ther.*, 2014, **10**, 853.

16 S. M. Cutts, A. Nudelman, A. Rephaeli and D. R. Phillips, The Power and Potential of Doxorubicin-DNA Adducts, *IUBMB Life*, 2005, **57**, 73–81.

17 C. F. Thorn, C. Oshiro, S. Marsh, T. Hernandez-Boussard, H. McLeod, T. E. Klein and R. B. Altman, Doxorubicin pathways: pharmacodynamics and adverse effects, *Pharmacogenet. Genomics*, 2011, **21**, 440.

18 M. R. Bristow, J. W. Mason, M. E. Billingham and J. R. Daniels, Dose-effect and structure-function relationships in doxorubicin cardiomyopathy, *Am. Heart J.*, 1981, **102**, 709–718.

19 W. Krause, Resistance to anti-tubulin agents: From vinca alkaloids to epothilones, *Cancer Drug Resist.*, 2019, **2**, 82.

20 S. Lobert, B. Vulevic and J. J. Correia, Interaction of Vinca Alkaloids with Tubulin: A Comparison of Vinblastine, Vincristine, and Vinorelbine, *Biochemistry*, 1996, **35**, 6806–6814.

21 S. M. Maloney, C. A. Hoover, L. V. Morejon-Lasso and J. R. Prosperi, Mechanisms of Taxane Resistance, *Cancers*, 2020, **12**, 3323.

22 Y. Zhang, Y. Xue, G. Li, H. Yuan and T. Luo, Enantioselective synthesis of Iboga alkaloids and vinblastine via rearrangements of quaternary ammoniums, *Chem. Sci.*, 2016, **7**, 5530–5536.

23 A. Duflos, A. Kruczynski and J.-M. Barret, Novel Aspects of Natural and Modified Vinca Alkaloids, *Curr. Med. Chem. - Anti-Cancer Agents*, 2002, **2**, 55–70.

24 M. Moudi, R. Go, C. Yong, S. Yien and M. Nazre, *Vinca Alkaloids*, 1231, vol. 4.

25 D. Guenard, F. Gueritte-Voegelein and P. Potier, Taxol and taxotere: discovery, chemistry, and structure-activity relationships, *Acc. Chem. Res.*, 1993, **26**, 160–167.

26 Y.-F. Wang, Q.-W. Shi, M. Dong, H. Kiyota, Y.-C. Gu and B. Cong, Natural Taxanes: Developments Since 1828, *Chem. Rev.*, 2011, **111**, 7652–7709.

27 Q.-W. Shi and H. Kiyota, New Natural Taxane Diterpenoids from Taxus Species since 1999, *Chem. Biodivers.*, 2005, **2**, 1597–1623.

28 B. T. McGrogan, B. Gilmartin, D. N. Carney and A. McCann, Taxanes, microtubules and chemoresistant breast cancer, *Biochim. Biophys. Acta BBA - Rev. Cancer*, 2008, **1785**, 96–132.

29 L. Min, J.-C. Han, W. Zhang, C.-C. Gu, Y.-P. Zou and C.-C. Li, Strategies and Lessons Learned from Total Synthesis of Taxol, *Chem. Rev.*, 2023, **123**, 4934–4971.

30 K. Tkaczuk and J. Yared, Update on taxane development: new analogs and new formulations, *Drug Des. Devel. Ther.*, 2012, 371.

31 G. J. Peters, Novel Developments in the Use of Antimetabolites, *Nucleosides Nucleotides Nucleic Acids*, 2014, **33**, 358–374.

32 C. Avendaño and J. C. Menéndez, in *Medicinal Chemistry of Anticancer Drugs*, Elsevier, 2008, pp. 9–52.

33 E. Y. Chu and H. H. Kong, in *Dermatologic Principles and Practice in Oncology*, John Wiley & Sons, Ltd, 2013, pp. 160–169.

34 I. S. Kovalev, G. V. Zyryanov, S. Santra, A. Majee, M. V. Varaksin and V. N. Charushin, Folic Acid Antimetabolites (Antifolates): A Brief Review on Synthetic Strategies and Application Opportunities, *Molecules*, 2022, **27**, 6229.

35 W. B. Parker, Enzymology of Purine and Pyrimidine Antimetabolites Used in the Treatment of Cancer, *Chem. Rev.*, 2009, **109**, 2880–2893.

36 S. B. Kaye, New antimetabolites in cancer chemotherapy and their clinical impact, *Br. J. Cancer*, 1998, **78**, 1–7.

37 F. J. Ansfield, J. M. Schroeder and A. R. Curreri, Five years clinical experience with 5-fluorouracil, *JAMA*, 1962, **181**, 295–299.

38 R. Thirumaran, G. C. Prendergast and P. B. Gilman, in *Cancer Immunotherapy*, eds. G. C. Prendergast and E. M. Jaffee, Academic Press, Burlington, 2007, pp. 101–116.

39 M. Mehrmohamadi, S. H. Jeong and J. W. Locasale, Molecular features that predict the response to antimetabolite chemotherapies, *Cancer Metab.*, 2017, **5**, 8.

40 Z. H. Siddik, in *The Cancer Handbook*, John Wiley & Sons, Ltd, 2005.

41 C. Marosi, Complications of chemotherapy in neuro-oncology, *Handb. Clin. Neurol.*, 2012, **105**, 873–885.

42 L. F. Craver, The Nitrogen Mustards: Clinical Use, *Radiology*, 1948, **50**, 486–493.

43 D. Fu, J. A. Calvo and L. D. Samson, Balancing repair and tolerance of DNA damage caused by alkylating agents, *Nat. Rev. Cancer*, 2012, **12**, 104–120.

44 M. S. Highley, B. Landuyt, H. Prenen, P. G. Harper and E. A. D. Bruijn, The Nitrogen Mustards, *Pharmacol. Rev.*, 2022, **74**, 552–599.

45 Y. Chen, Y. Jia, W. Song and L. Zhang, Therapeutic Potential of Nitrogen Mustard Based Hybrid Molecules, *Front. Pharmacol.*, 2018, **9**, 1–12.

46 R. Papac, N. L. Petrakis, F. Amini and D. A. Wood, Comparative clinical evaluation of two alkylating agents: mannitol mustard and cyclophosphamide (cytoxan), *J. Am. Med. Assoc.*, 1960, **172**, 1387–1391.

47 M. J. Moore, Clinical Pharmacokinetics of Cyclophosphamide, *Clin. Pharmacokinet.*, 1991, **20**, 194–208.

48 R. A. Fleming, An Overview of Cyclophosphamide and Ifosfamide Pharmacology, *Pharmacother. J. Hum. Pharmacol. Drug Ther.*, 1997, **17**, 146S-154S.

49 J. B. Vermorken, W. W. ten Bokkel Huinink, E. A. Eisenhauer, G. Favalli, D. Belpomme, P. F. Conte and S. B. Kaye, Carboplatin versus cisplatin, *Ann. Oncol.*, 1993, **4**, S41–S48.

50 P. J. O'Dwyer, J. P. Stevenson and S. W. Johnson, in *Cisplatin*, John Wiley & Sons, Ltd, 1999, pp. 29–69.

51 G. J. Peters, C. L. van der Wilt, C. J. van Moorsel, J. R. Kroep, A. M. Bergman and S. P. Ackland, Basis for effective combination cancer chemotherapy with antimetabolites, *Pharmacol. Ther.*, 2000, **87**, 227–253.

52 A. D. Seidman, Gemcitabine and docetaxel in metastatic breast cancer, *Oncol. Williston Park N*, 2004, **18**, 13–16.

53 P. Hoppenz, S. Els-Heindl and A. G. Beck-Sickinger, Peptide-Drug Conjugates and Their Targets in Advanced Cancer Therapies, *Front. Chem.2020, **8**, 671, 1-24*

54 H. Jin, L. Wang and R. Bernards, Rational combinations of targeted cancer therapies: background, advances and challenges, *Nat. Rev. Drug Discov.*, 2023, **22**, 213–234.

55 A. Howell and S. J. Howell, Tamoxifen evolution, *Br. J. Cancer*, 2023, **128**, 421–425.

56 V. C. Jordan, Tamoxifen: a most unlikely pioneering medicine, *Nat. Rev. Drug Discov.*, 2003, **2**, 205–213.

57 C. K. Osborne, Tamoxifen in the Treatment of Breast Cancer, *N. Engl. J. Med.*, 1998, **339**, 1609–1618.

58 F. Pane, M. Intrieri, C. Quintarelli, B. Izzo, G. C. Muccioli and F. Salvatore, BCR/ABL genes and leukemic phenotype: from molecular mechanisms to clinical correlations, *Oncogene*, 2002, **21**, 8652–8667.

59 M. Zoubir, T. Tursz, C. Menard, L. Zitvogel and N. Chaput, Imatinib Mesylate (Gleevec©): Targeted Therapy Against Cancer with Immune Properties, *Endocr. Metab. Immune Disord. - Drug Targets*, **10**, 1–7.

60 R. Capdeville, E. Buchdunger, J. Zimmermann and A. Matter, Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug, *Nat. Rev. Drug Discov.*, 2002, **1**, 493–502.

61 F. Stegmeier, M. Warmuth, W. R. Sellers and M. Dorsch, Targeted Cancer Therapies in the Twenty-First Century: Lessons From Imatinib, *Clin. Pharmacol. Ther.*, 2010, **87**, 543–552.

62 I. Tamascar and J. Ramanarayanan, Targeted treatment of chronic myeloid leukemia: role of imatinib, *OncoTargets Ther.*, 2009, **2**, 63–71.

63 O. Uziel, E. Fenig, J. Nordenberg, E. Beery, H. Reshef, J. Sandbank, M. Birenbaum, M. Bakhanashvili, R. Yerushalmi, D. Luria and M. Lahav, Imatinib mesylate (Gleevec) downregulates telomerase activity and inhibits proliferation in telomerase-expressing cell lines, *Br. J. Cancer*, 2005, **92**, 1881–1891.

64 K. M. Cook and W. D. Figg, Angiogenesis Inhibitors: Current Strategies and Future Prospects, *CA. Cancer J. Clin.*, 2010, **60**, 222–243.

65 F. a. L. M. Eskens, Angiogenesis inhibitors in clinical development; where are we now and where are we going?, *Br. J. Cancer*, 2004, **90**, 1–7.

66 N. Vargesson, Thalidomide-induced teratogenesis: History and mechanisms, *Birth Defects Res.*, 2015, **105**, 140–156.

67 J. Folkman, Endogenous angiogenesis inhibitors, *APMIS*, 2004, **112**, 496–507.

68 L. Gong, H. Zhao, Y. Liu, H. Wu, C. Liu, S. Chang, L. Chen, M. Jin, Q. Wang, Z. Gao and W. Huang, Research advances in peptide–drug conjugates, *Acta Pharm. Sin. B*, 2023, **13**, 3659–3677.

69 E. I. Vrettos, G. Mező and A. G. Tzakos, On the design principles of peptide–drug conjugates for targeted drug delivery to the malignant tumor site, *Beilstein J. Org. Chem.*, 2018, **14**, 930–954.

70 C. Fu, L. Yu, Y. Miao, X. Liu, Z. Yu and M. Wei, Peptide–drug conjugates (PDCs): a novel trend of research and development on targeted therapy, hype or hope?, *Acta Pharm. Sin. B*, 2023, **13**, 498–516.

71 B. M. Cooper, J. Iegre, D. H. O' Donovan, M. O. ¨ Lwegård Halvarsson and D. R. Spring, Peptides as a platform for targeted therapeutics for cancer: peptide-drug conjugates (PDCs), *Chem Soc Rev*, 2021, **50**, 1480–1494.

72 V. M. Miranda, Medicinal inorganic chemistry: an updated review on the status of metallodrugs and prominent metallodrug candidates, *Rev. Inorg. Chem.*, 2022, **42**, 29–52.

73 W. Delbart, J. Karabet, G. Marin, S. Penninckx, J. Derrien, G. E. Ghanem, P. Flamen and Z. Wimana, Understanding the Radiobiological Mechanisms Induced by 177Lu-DOTATATE in Comparison to External Beam Radiation Therapy, *Int. J. Mol. Sci.*, 2022, **23**, 12369.

74 Y.-S. Zhu, K. Tang and J. Lv, Peptide–drug conjugate-based novel molecular drug delivery system in cancer, *Trends Pharmacol. Sci.*, 2021, **42**, 857–869.

75 M. Alas, A. Saghaeidehkordi and K. Kaur, Peptide–Drug Conjugates with Different Linkers for Cancer Therapy, *J. Med. Chem.*, 2021, **64**, 216–232.

76 P. Gogia, H. Ashraf, S. Bhasin and Y. Xu, Antibody–Drug Conjugates: A Review of Approved Drugs and Their Clinical Level of Evidence, *Cancers*, 2023, **15**, 3886.

77 D.-Y. Ruan, H.-X. Wu, Q. Meng and R.-H. Xu, Development of antibody-drug conjugates in cancer: overview and prospects, *Cancer Commun.*, 2023, 1–20.

78 A. Mullard, FDA approves 100th monoclonal antibody product, *Nat. Rev. Drug Discov.*

79 I. Pysz, P. Jackson and D. E. Thurston, in *RSC Drug Discovery Series*, 2019.

80 A. Ungaro, M. Tucci, A. Audisio, L. Di Prima, C. Pisano, F. Turco, M. D. Delcuratolo, M. Di Maio, G. V. Scagliotti and C. Buttigliero, Antibody-Drug Conjugates in Urothelial Carcinoma: A New Therapeutic Opportunity Moves from Bench to Bedside, *Cells*, 2022, **11**, 803.

81 V. Chudasama, A. Maruani and S. Caddick, Recent advances in the construction of antibody–drug conjugates, *Nat. Chem.*, 2016, **8**, 114–119.

82 J. Charles A Janeway, P. Travers, M. Walport and M. J. Shlomchik, in *Immunobiology: The Immune System in Health and Disease. 5th edition*, Garland Science, 2001.

83 M. L. Chiu, D. R. Goulet, A. Teplyakov and G. L. Gilliland, Antibody Structure and Function: The Basis for Engineering Therapeutics, *Antibodies*, 2019, **8**, 55.

84 F. A. N. Biteghe, N. Mungra, N. E. T. Chalomie, J. D. L. C. Ndong, J. Engohang-Ndong, G. Vignaux, E. Padayachee, K. Naran and S. Barth, Advances in epidermal growth factor receptor specific immunotherapy: lessons to be learned from armed antibodies, *Oncotarget*, 2020, **11**, 3531–3557.

85 C. Theocharopoulos, P.-P. Lialios, H. Gogas and D. C. Ziogas, An overview of antibody–drug conjugates in oncological practice, *Ther. Adv. Med. Oncol.*, 2020, **12**, 1758835920962997.

86 J. T. W. Tong, P. W. R. Harris, M. A. Brimble and I. Kavianinia, An Insight into FDA Approved Antibody-Drug Conjugates for Cancer Therapy, *Molecules*, 2021, **26**, 5847.

87 A. Beck, L. Goetsch, C. Dumontet and N. Corvaïa, Strategies and challenges for the next generation of antibody–drug conjugates, *Nat. Rev. Drug Discov.*, 2017, **16**, 315–337.

88 CHMP, Committee for Medicinal Products for Human Use (CHMP) Assessment report.

89 N. K. Damle and P. Frost, Antibody-targeted chemotherapy with immunoconjugates of calicheamicin, *Curr. Opin. Pharmacol.*, 2003, **3**, 386–390.

90 P. D. Senter and E. L. Sievers, The discovery and development of brentuximab vedotin for use in relapsed Hodgkin lymphoma and systemic anaplastic large cell lymphoma, *Nat. Biotechnol.*, 2012, **30**, 631–637.

91 S. Ponziani, G. Di Vittorio, G. Pitari, A. M. Cimini, M. Ardini, R. Gentile, S. Iacobelli, G. Sala, E. Capone, D. J. Flavell, R. Ippoliti and F. Giansanti, Antibody-Drug Conjugates: The New Frontier of Chemotherapy, *Int. J. Mol. Sci.*, 2020, **21**, 5510.

92 Y. Fu and M. Ho, DNA damaging agent-based antibody-drug conjugates for cancer therapy, *Antib. Ther.*, 2018, **1**, 43.

93 J. D. Bargh, A. Isidro-Llobet, J. S. Parker and D. R. Spring, Chem Soc Rev Cleavable linkers in antibody-drug conjugates, *Chem Soc Rev*, **48**, 4361.

94 J. R. McCombs and S. C. Owen, Antibody Drug Conjugates: Design and Selection of Linker, Payload and Conjugation Chemistry, *AAPS J.*, 2015, **17**, 339–351.

95 J. D. Bargh, S. J. Walsh, A. Isidro-Llobet, S. Omarjee, J. S. Carroll and D. R. Spring, Sulfatase-cleavable linkers for antibody-drug conjugates †, , DOI:10.1039/c9sc06410a.

96 C. Dumontet, J. M. Reichert, P. D. Senter, J. M. Lambert and A. Beck, Antibody–drug conjugates come of age in oncology, *Nat. Rev. Drug Discov.*, 2023, **22**, 641–661.

97 S. Kitson, L. Mansi, V. Cuccurullo and A. Ciarmiello, Targeted Therapy Towards Cancer-A Perspective, *Anticancer Agents Med. Chem.*, 2017, **17**, 311.

98 J. D. Bargh, A. Isidro-Llobet, J. S. Parker and D. R. Spring, Cleavable linkers in antibody–drug conjugates, *Chem. Soc. Rev.*, 2019, **48**, 4361–4374.

99 X. Sun, J. F. Ponte, N. C. Yoder, R. Laleau, J. Coccia, L. Lanieri, Q. Qiu, R. Wu, E. Hong, M. Bogalhas, L. Wang, L. Dong, Y. Setiady, E. K. Maloney, O. Ab, X. Zhang, J. Pinkas, T. A. Keating, R. Chari, H. K. Erickson and J. M. Lambert, Effects of Drug–Antibody Ratio on Pharmacokinetics, Biodistribution, Efficacy, and Tolerability of Antibody–Maytansinoid Conjugates, *Bioconjug. Chem.*, 2017, **28**, 1371–1381.

100 M. Abdollahpour-Alitappeh, M. Lotfinia, T. Gharibi, J. Mardaneh, B. Farhadihosseinabadi, P. Larki, B. Faghfourian, K. S. Sepehr, K. Abbaszadeh-Goudarzi, G. Abbaszadeh-Goudarzi, B. Johari, M. R. Zali and N. Bagheri, Antibody–drug conjugates (ADCs) for cancer therapy: Strategies, challenges, and successes, *J. Cell. Physiol.*, 2018, **234**, 5628–5642.

101 S. Sau, H. O. Alsaab, S. K. Kashaw, K. Tatiparti and A. K. Iyer, Advances in antibody–drug conjugates: A new era of targeted cancer therapy, *Drug Discov. Today*, 2017, **22**, 1547–1556.

102 D. M. Goldenberg and R. M. Sharkey, Sacituzumab govitecan, a novel, third-generation, antibody-drug conjugate (ADC) for cancer therapy, *Expert Opin. Biol. Ther.*, 2020, **20**, 871–885.

103 F. Mack, M. Ritchie and P. Sapra, The Next Generation of Antibody Drug Conjugates, *Semin. Oncol.*, 2014, **41**, 637–652.

104 Next-Generation Antibody-Drug Conjugates (ADCs) for Cancer Therapy, *ACS Med. Chem. Lett.*, 2016, **7**, 972–973.

105 C. M. Yamazaki, A. Yamaguchi, Y. Anami, W. Xiong, Y. Otani, J. Lee, N. T. Ueno, N. Zhang, Z. An and K. Tsuchikama, Antibody-drug conjugates with dual payloads for combating breast tumor heterogeneity and drug resistance, *Nat. Commun.*, 2021, **12**, 3528.

106 H. P. Yao, H. Zhao, R. Hudson, X. M. Tong and M. H. Wang, Duocarmycin-based antibody–drug conjugates as an emerging biotherapeutic entity for targeted cancer therapy: Pharmaceutical strategy and clinical progress, *Drug Discov. Today*, 2021, **26**, 1857–1874.

107 A. Kumar, J. White, R. James Christie, N. Dimasi and C. Gao, in *Annual Reports in Medicinal Chemistry*, Academic Press Inc., 2017, vol. 50, pp. 441–480.

108 M. S. Tichenor, J. D. Trzupek, D. B. Kastrinsky, F. Shiga, I. Hwang and D. L. Boger, Asymmetric Total Synthesis of (+)-and ent-(-)-Yatakemycin and Duocarmycin SA: Evaluation of Yatakemycin Key Partial Structures and Its Unnatural Enantiomer, *J AM CHEM SOC*, 2006, **128**, 15683–15696.

109 D. D. Von Hoff, K. A. Newell and S. L. Crampton, *CC-1065 (NSC 298223), a Most Potent Antitumor Agent: Kinetics of Inhibition of Growth, DNA Synthesis, and Cell Survival*, 1982, vol. 42.

110 R. C. Elgersma, R. G. E. Coumans, T. Huijbregts, W. M. P. B. Menge, J. A. F. Joosten, H. J. Spijker, F. M. H. De Groot, M. M. C. Van Der Lee, R. Ubink, D. J. Van Den Dobbelsteen, D. F. Egging, W. H. A. Dokter, G. F. M. Verheijden, J. M. Lemmens, C. M. Timmers and P. H. Beusker, Design, Synthesis, and Evaluation of Linker-Duocarmycin Payloads: Toward Selection of HER2-Targeting Antibody−Drug Conjugate SYD985, *Mol. Pharm.*, 2015, **12**, 1813–1835.

111 D. L. Boger, C. W. Boyce, R. M. Garbaccio and J. A. Goldberg, CC-1065 and the duocarmycins: Synthetic studies, *Chem. Rev.*, 1997, **97**, 787–828.

112 Y. Wang, H. Yuan, S. C. Wright, H. Wang and J. W. Larrick, Synthesis and cytotoxicity of a biotinylated CC-1065 analogue, *BMC Chem. Biol.*, 2002, **2**, 1–19.

113 J. P. Parrish, D. B. Kastrinsky, S. E. Wolkenberg, Y. Igarashi and D. L. Boger, DNA alkylation properties of yatakemycin, *J. Am. Chem. Soc.*, 2003, **125**, 10971–10976.

114 M. S. Tichenor, K. S. MacMillan, J. D. Trzupek, T. J. Rayl, I. Hwang and D. L. Boger, Systematic exploration of the structural features of yatakemycin

impacting DNA alkylation and biological activity, *J. Am. Chem. Soc.*, 2007, **129**, 10858–10869.

115 G. A. Vielhauer, M. Swink, N. K. Parelkar, J. P. Lajiness, A. L. Wolfe and D. Boger, Evaluation of a reductively activated duocarmycin prodrug against murine and human solid cancers, *Cancer Biol. Ther.*, 2013, **14**, 527–536.

116 A. M. Beekman, M. M. D. Cominetti and M. Searcey, in *Cytotoxic Payloads for AntibodyDrug Conjugates*, 2019, pp. 185–191.

117 K. S. MacMillan and D. L. Boger, Fundamental relationships between structure, reactivity, and biological activity for the duocarmycins and CC-1065, *J. Med. Chem.*, 2009, **52**, 5771–5780.

118 A. L. Wolfe, K. K. Duncan, N. K. Parelkar, S. J. Weir, G. A. Vielhauer and D. L. Boger, A Novel, Unusually Efficacious Duocarmycin Carbamate Prodrug That Releases No Residual Byproduct, *J. Med. Chem.*, 2012, **55**, 5878–5886.

119 R. J. Stevenson, W. A. Denny, M. Tercel, F. B. Pruijn and A. Ashoorzadeh, Nitro seco Analogues of the Duocarmycins Containing Sulfonate Leaving Groups as Hypoxia-Activated Prodrugs for Cancer Therapy, *J Med Chem*, 2012, **55**, 2780–2801.

120 J. R. Schnell, R. R. Ketchem, D. L. Boger and W. J. Chazin, Binding-Induced Activation of DNA Alkylation by Duocarmycin SA: Insights from the Structure of an Indole Derivative−DNA Adduct, *J. Am. Chem. Soc.*, 1999, **121**, 5645–5652.

121 D. L. Boger and R. M. Garbaccio, Shape-dependent catalysis: Insights into the source of catalysis for the CC-1065 and duocarmycin DNA alkylation reaction, *Acc. Chem. Res.*, 1999, **32**, 1043–1052.

122 D. L. Boger, M. Searcey, W. C. Tse and Q. Jin, Bifunctional alkylating agents derived from duocarmycin SA: Potent antitumor activity with altered sequence selectivity, *Bioorg. Med. Chem. Lett.*, 2000, **10**, 495–498.

123 D. L. Boger, D. S. Johnson and W. Yun, (+)- and ent-(-)-Duocarmycin SA and (+)- and ent-(-)-N-BOC-DSA DNA Alkylation Properties.Alkylation Site Models That Accommodate the Offset AT-Rich Adenine N3 Alkylation Selectivity of the Enantiomeric Agents, *J. Am. Chem. Soc.*, 1994, **116**, 1635–1656.

124 W. Wrasidlo, D. S. Johnson and D. L. Boger, Induction of endonucleolytic DNA fragmentation and apoptosis by the duocarmycins, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 651–656.

125 D. L. Boger and W. Yun, Reversibility of the duocarmycin A and SA DNA alkylation reaction, *J. Am. Chem. Soc.*, 1993, **115**, 9872–9873.

126 D. L. Boger, H. W. Schmitt, B. E. Fink and M. P. Hedrick, Parallel Synthesis and Evaluation of 132 (+)-1,2,9,9a-Tetrahydrocyclopropa[c]benz[e]indol-4-one (CBI) Analogues of CC-1065 and the Duocarmycins Defining the Contribution of the DNA-Binding Domain, *J. Org. Chem.*, 2001, **66**, 6654–6661.

127 Z. Jukes, G. Ribeiro Morais, P. M. Loadman and K. Pors, How can the potential of the duocarmycins be unlocked for cancer therapy?, *Drug Discov. Today*, 2021, **26**, 577–584.

128 M. Tercel, M. A. Gieseg, W. A. Denny and W. R. Wilson, Synthesis and Cytotoxicity of Amino- *s eco* -DSA: An Amino Analogue of the DNA Alkylating Agent Duocarmycin SA, *J. Org. Chem.*, 1999, **64**, 5946–5953.

129 N. Ghosh, H. Sheldrake, M. Searcey and K. Pors, Chemical and Biological Explorations of the Family of CC-1065 and the Duocarmycin Natural Products, *Curr. Top. Med. Chem.*, 2009, **9**, 1494–1524.

130 H.-K. Han and G. L. Amidon, Targeted prodrug design to optimize drug delivery, *AAPS PharmSci*, 2000, **2**, 6.

131 D. C. Singleton, A. Macann and W. R. Wilson, Therapeutic targeting of the hypoxic tumour microenvironment, *Nat. Rev. Clin. Oncol.*, 2021, **18**, 751–772.

132 M. Tercel, G. J. Atwell, S. Yang, R. J. Stevenson, K. J. Botting, M. Boyd, E. Smith, R. F. Anderson, W. A. Denny, W. R. Wilson and F. B. Pruijn, Hypoxia-Activated Prodrugs: Substituent Effects on the Properties of Nitro *seco* -1,2,9,9a-Tetrahydrocyclopropa[*c*]benz[*e*]indol-4-one (nitroCBI) Prodrugs of DNA Minor Groove Alkylating Agents, *J. Med. Chem.*, 2009, **52**, 7258–7272.

133 W. Jin, J. D. Trzupek, T. J. Rayl, M. A. Broward, G. A. Vielhauer, S. J. Weir, I. Hwang and D. L. Boger, A unique class of duocarmycin and CC-1065 analogues subject to reductive activation, *J. Am. Chem. Soc.*, 2007, **129**, 15391–15397.

134 J. P. Lajiness, W. M. Robertson, I. Dunwiddie, M. A. Broward, G. A. Vielhauer, S. J. Weir and D. L. Boger, Design, synthesis, and evaluation of

duocarmycin O-amino phenol prodrugs subject to tunable reductive activation, *J. Med. Chem.*, 2010, **53**, 7731–7738.

135 Y. Wang, L. Li, Z. Tian, W. Jiang and J. W. Larrick, Synthesis and antitumor activity of CBI-bearing ester and carbamate prodrugs of CC-1065 analogue, *Bioorg. Med. Chem.*, 2006, **14**, 7854–7861.

136 L. H. Li, T. F. DeKoning, R. C. Kelly, W. C. Krueger, J. P. McGovren, G. E. Padbury, G. L. Petzold, T. L. Wallace, R. J. Ouding, M. D. Prairie and I. Gebhard, Cytotoxicity and Antitumor Activity of Carzelesin, a Prodrug Cyclopropylpyrroloindole Analogue1, *Cancer Res.*, 1992, **52**, 4904–4913.

137 E. Kobayashi, A. Okamoto, M. Asada, M. Okabe, S. Nagamura, A. Asai, H. Saito, K. Gomi and T. Hirata, Characteristics of Antitumor Activity of KW-2189, a Novel Water-soluble Derivative of Duocarmycin, against Murine and Human Tumors, *Cancer Res.*, 1994, **54**, 2404–2410.

138 S. R. Alberts, V. J. Suman, H. C. Pitot, J. K. Camoriano and J. Rubin, Use of Kw-2189, a Dna Minor Groove-Binding Agent, in Patients with Hepatocellular Carcinoma: A North Central Cancer Treatment Group (ncctg) Phase Ii Clinical Trial, *J. Gastrointest. Cancer*, 2007, **38**, 10–14.

139 L. F. Tietze, K. Schmuck, H. J. Schuster, M. Müller and I. Schuberth, Synthesis and Biological Evaluation of Prodrugs Based on the Natural Antibiotic Duocarmycin for Use in ADEPT and PMT, *Chem. – Eur. J.*, 2011, **17**, 1922–1929.

140 L. F. Tietze, H. J. Schuster, B. Krewer and I. Schuberth, Synthesis and Biological Studies of Different Duocarmycin Based Glycosidic Prodrugs for Their Use in the Antibody-Directed Enzyme Prodrug Therapy, *J. Med. Chem.*, 2009, **52**, 537–543.

141 N. Ortuzar, K. Karu, D. Presa, G. R. Morais, H. M. Sheldrake, S. D. Shnyder, F. M. Barnieh, P. M. Loadman, L. H. Patterson, K. Pors and M. Searcey, Probing cytochrome P450 (CYP) bioactivation with chloromethylindoline bioprecursors derived from the duocarmycin family of compounds, *Bioorg. Med. Chem.*, 2021, **40**, 116167.

142 O. C. Cartwright, A. M. Beekman, M. M. D. Cominetti, D. A. Russell and M. Searcey, A Peptide-Duocarmycin Conjugate Targeting the Thomsen-Friedenreich Antigen Has Potent and Selective Antitumor Activity, *Bioconjug. Chem.*, 2020, **31**, 1745–1749.

143 O. Kurtenkov, Profiling of Naturally Occurring Antibodies to the Thomsen-Friedenreich Antigen in Health and Cancer: The Diversity and Clinical Potential, *BioMed Res. Int.*, 2020, **2020**, 1–12.

144 M. J. Stephenson, L. A. Howell, M. A. O'Connell, K. R. Fox, C. Adcock, J. Kingston, H. Sheldrake, K. Pors, S. P. Collingwood and M. Searcey, Solid-Phase Synthesis of Duocarmycin Analogues and the Effect of C-Terminal Substitution on Biological Activity, *J. Org. Chem.*, 2015, **80**, 9454–9457.

145 D. Su, J. Chen, E. Cosino, J. Dela Cruz-Chuh, H. Davis, G. Del Rosario, I. Figueroa, L. Goon, J. He, A. V. Kamath, S. Kaur, K. R. Kozak, J. Lau, D. Lee, M. V. Lee, D. Leipold, L. Liu, P. Liu, G. L. Lu, C. Nelson, C. Ng, T. H. Pillow, P. Polakis, A. G. Polson, R. K. Rowntree, O. Saad, B. Safina, N. J. Stagg, M. Tercel, R. Vandlen, B. S. Vollmar, J. Wai, T. Wang, B. Wei, K. Xu, J. Xue, Z. Xu, G. Yan, H. Yao, S. F. Yu, D. Zhang, F. Zhong and P. S. Dragovich, Antibody-drug conjugates derived from cytotoxic seco-CBI-Dimer payloads are highly efficacious in xenograft models and form protein adducts in Vivo, *Bioconjug. Chem.*, 2019, **30**, 1356–1370.

146 M. Nadal-Serrano, B. Morancho, S. Escrivá-de-Romaní, C. Bernadó Morales, A. Luque, M. Escorihuela, M. Espinosa Bravo, V. Peg, F. A. Dijcks, W. H. A. Dokter, J. Cortés, C. Saura and J. Arribas, The Second Generation Antibody-Drug Conjugate SYD985 Overcomes Resistances to T-DM1, *Cancers*.

Chapter 2– Synthesis of DSA

Chapter 2 – Synthesis of DSA

The aims of this chapter are to demonstrate an improved and alternative route to synthesize the DSA alkylating unit. Initially, the chapter provides a brief overview of the synthetic approaches that have influenced the work presented in this thesis. For more comprehensive reviews of duocarmycin synthesis, see published references. In 1997, Boger published a review of synthetic studies of CC-1065 and the duocarmycins, which laid the foundation for subsequent research.[1] Since then, numerous analogues within this family have been synthesized. A novel member named yatakemycin was originally reported in error and later corrected by Boger and Tichenor who describe the total synthesis of yatakemycin.[2] Providing a more recent perspective, in 2022, Felber published a comprehensive overview of the development of the duocarmycins over the last 40 years.[3]

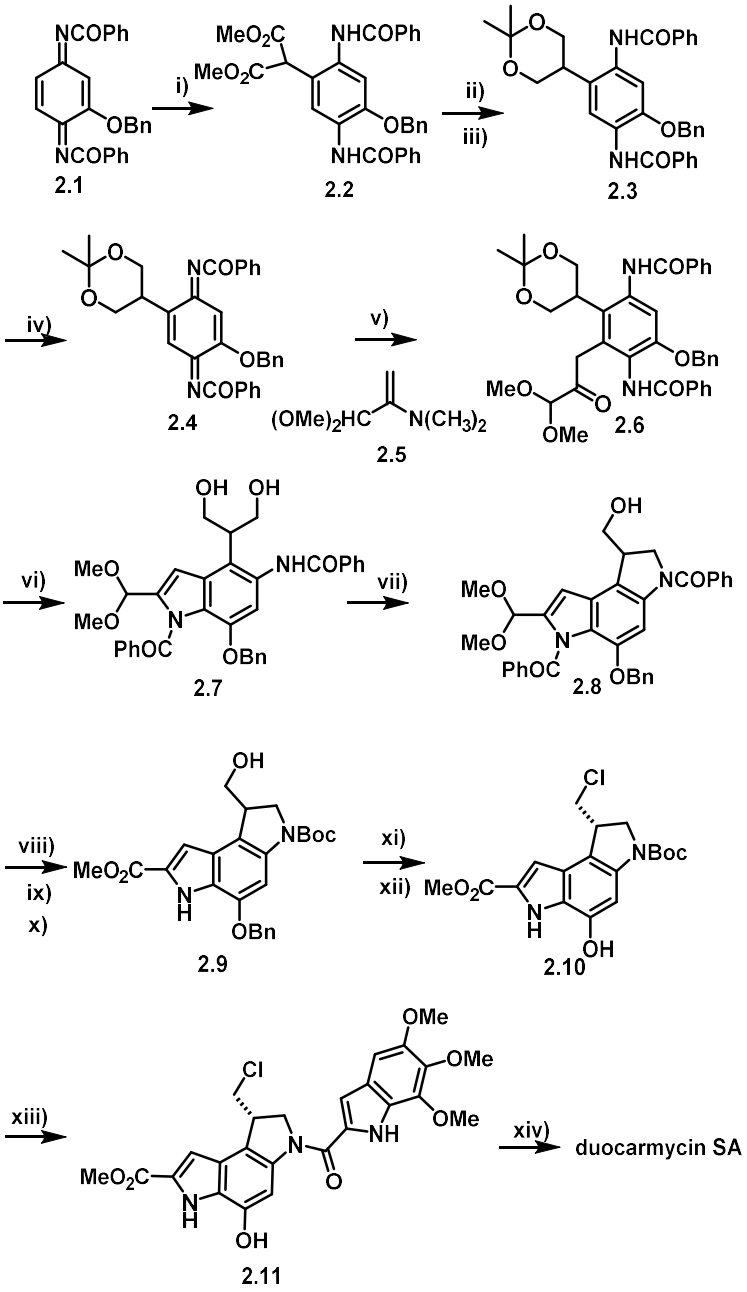
A route to the Fmoc-protected DSA group has previously been described by the Searcey group. This chapter discusses this route and the difficulties associated with it. An alternative route that avoids previous bottlenecks will then be discussed before comparing both routes in the conclusion. Both routes generate the DSA as a racemic mixture, and **Chapter 3** will go on to discuss the importance of using enantiopure material in the development of dimeric DSA structures.

2.1. Synthetic approaches to duocarmycin SA and DSA

The first reported synthesis of a member of the duocarmycin family was of (+)-CC-1065 in 1988 (**Chapter 1, Figure 1.20**).[4] The indole ring systems of the alkylating units tend to be synthetically challenging and analogues have been made that are more tractable for synthesis such as the cyclopropabenzaindole (CBI) unit (**Chapter 1**, **Figure 1.21**). Since then extensive studies have been carried out on the other members of this family as well as analogues to investigate structure activity relationships.[5]

2.1.1 First total synthesis of (+)-DSA

In 1992 Boger reported the first total synthesis of (+)- duocarmycin SA. The synthesis gave a racemic mixture where both enantiomers were separated using chiral HPLC (**Scheme 2.1**).[6]



Scheme 2.1: *The first total synthesis of (+)-duocarmycin SA by Boger et al.*

i) $CH(CO_2Me)_2$, $NaOCH_3$, THF, 64%, ii) $NaBH_4$, EtOH, 71%, iii) $Me_2C(OMe)_2$, TsOH, DMF, 99%, iv) $PB(OAc)_4$, $CHCl_3$, 100%, v) pyruvaldehyde dimethyl acetal, THF followed immediately by pH 4 phosphate buffer, 61%, vi) HCl, MeOH, 91%, vii) DEAD-Ph_3P, THF, 100%, viii) NH_2NH_2, EtOH followed by Boc_2O, THF, 67%, ix) DMSO-pH4 phosphate buffer- dioxane, 91%, x) MnO_2, NaCN, MeOH, AcOH 89%, xi) 10% Pd/C, HCO_2NH_4, 77%, resolution xii) $PPh_3.CCl_4$, CH_2Cl_2, 81%, xiii) first HCl, EtOAc followed by, 5,6,7-trimethoxyindole-2-carboxylic acid, EDCl, $NaHCO_3$, 61%, xiv) NaH, THF.DMF, 87%.

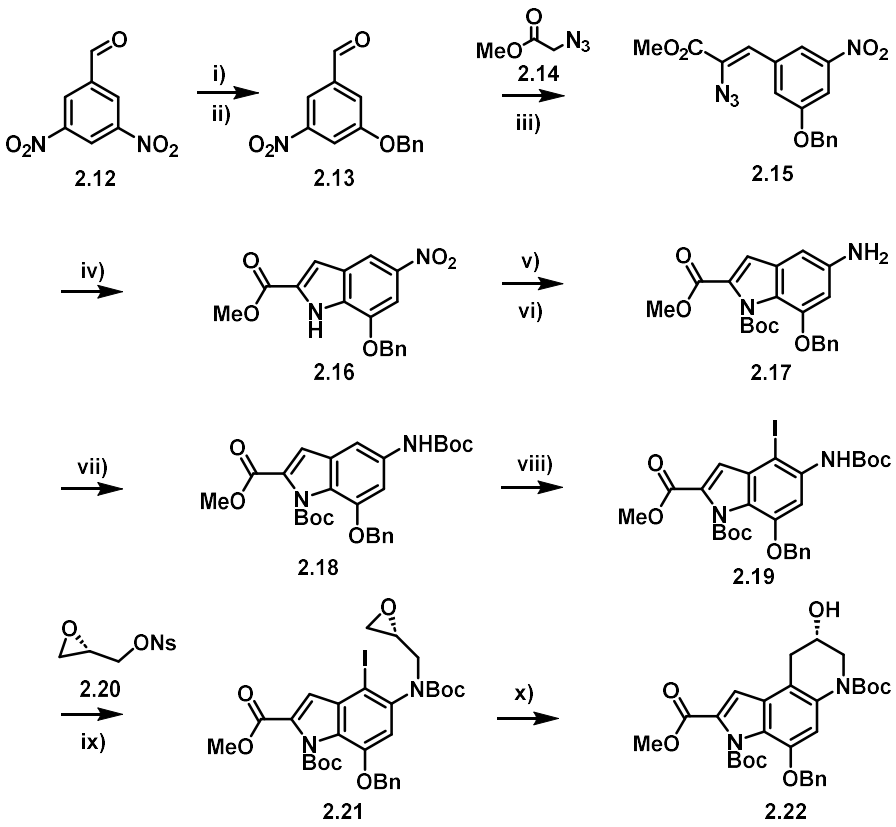
The first step involved the nucleophilic addition of dimethyl malonate to a quinonediimine (**2.1**), which occurred regioselectively to generate compound **2.2**. Reduction of the dimethyl ester and protection of the resulting diol as the acetonide was achieved using 2,2-dimethoxypropane ($Me_2C(OMe)_2$) with catalytic p-

toluenesulfonic acid. Oxidation using lead tetraacetate regenerated a quinonediimide (compound **2.4**) and allowed a second nucleophilic addition to this species, which after treatment with pH 4.0 buffer generated ketone **2.6**. Cyclisation to the indole was carried out under acidic conditions alongside regeneration of the diol to give compound **2.7**. This diol was now set up for an intra-nuclear Mitsunobu reaction to cyclise and generate the indoline structure (**2.8**), which was achieved using DEAD-Ph_3P. The next step involved manipulation of the protecting groups, moving from the benzoyl groups, which are required on the quinonediimines to help regioselectivity, to a single Boc group on the indoline nitrogen. This was achieved by treatment with hydrazine followed by Boc anhydride. Acid catalysed hydrolysis of the dimethyl acetal gave the aldehyde which was then oxidised with MnO_2 and NaCN in MeOH and AcOH to give the methyl ester, **2.9**. This was followed by conversion of the hydroxyl group to the chloride using PPh_3 and CCl_4, giving compound **2.10**. Boc deprotection was carried out using HCl and then the coupling of 5,6,7-trimethoxyindole-2-carboxylic acid using EDCl, $NaHCO_3$ in DMF, followed by chiral resolution resulted in compound **2.11** with a yield of 61%. The final reaction was an intramolecular ring closure using NaH in a mixture of THF and DMF with a high yield of 87% giving the final duocarmycin SA product (**Chapter 1, Figure 1.20**).

Since this first reported synthesis, extensive research has been carried out on this family of compounds with notable contributions from the following groups; Boger, Searcey, Sugiyama, Lee, Denny, Tercel, Saito and Tietze (reviewed by Felber and Thorn-Seshold).[3] This project explored the alkylating subunit of duocarmycin SA. A summary of the different approaches used to make the subunit by different groups was provided by Schmidt et al. who reported the enantioselective total synthesis of (+)-duocarmycin SA.[7]

2.1.2 Bogers enantioselective synthesis of (+)-DSA

All these synthetic routes failed to synthesise duocarmycins enantioselectively. As discussed in **Chapter 1**, stereochemistry determines how these compounds bind to DNA and affect the potencies.[8] In 2006 the Boger group synthesised duocarmycin SA, as well as yatakemycin in an asymmetric synthesis.[2] As well as this, in 2018 Schmidt published an enantioselective total synthesis of (+)- duocarmycin SA.[7]



Scheme 2.2: *Enantioselective synthesis of duocarmycin SA.*

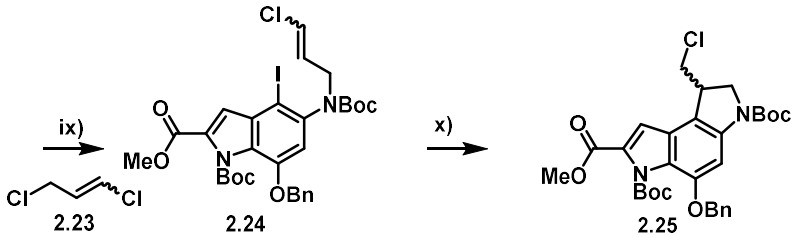
i) benzaldehyde oxime, K_2CO_3, DMF, ii) BnBr, DMF, 86%, iii) methyl azidoacetate, NaOMe, MeOH, 78%, iv) Xylenes 140 °C, 68%, v) Boc_2O, 98%, vi) Zn, NH_4Cl, 98%, vii) Boc_2O, 95%, viii) NIS, toluene, AcOH, 91%, ix) (S)-glycidyl-3-nosylate, NaH, DMF, 96%, x) i-PrMgCl, -42 °C, followed by CuI-Bu_3P, -78 °C, 69%

Both the Boger and Schmidt routes to synthesis have been investigated in our laboratory, although the results have been equivocal, with some steps proving to be irreproducible in our hands. However, the Boger synthesis was reproducible and elements of it were utilised in the later, successful, racemic synthesis described in this work. Key to this in an enantioselective sense, is the introduction of a chiral epoxide, which is subject to Grignard conditions and transmetallation with copper. This reaction generated the enantiopure product within the Boger group.

This synthesis will be discussed in this chapter and individual reactions examined in more detail. More generally, the synthesis started as shown in **Scheme 2.2** with aldehyde **2.12**. Nucleophilic substitution using benzaldehyde oxime followed by BnBr allowed one nitro group to be converted into a benzyl group giving **2.13**. The next step was a condensation reaction using methyl azido acetate and NaOMe acting as a base.

The product (**2.14**) was dissolved in xylene and heated to reflux allowing cyclisation to form the indole ring (**2.16**). There was tert-butyloxycarbonyl (Boc) protection of the amine group followed by reduction and protection of the nitro group. Regioselective iodination was achieved using *N*-iodosuccinimide in toluene with AcOH. The epoxide (*S*)-glycidyl-3-nosylate was introduced using NaH in DMF. Finally, treatment with *i*-PrMgCl at -42 °C, followed by CuI-Bu_3P at -78 °C left the ring opened benzyl protected enantiopure alkylating subunit (**2.22**).

This paper by Boger also described the synthesis of the racemic alkylating subunit of duocarmycin SA (**Scheme 2.3**).[2] This is from iodinated alkylating subunit **2.19.**



Scheme 2.3*: Alternative route by Boger to synthesise racemic alkylating subunit of duocarmycin SA.*

ix) 1,3-dichloropropene, NaH, 86%, x) AIBN, Bu_3SnH, 87%

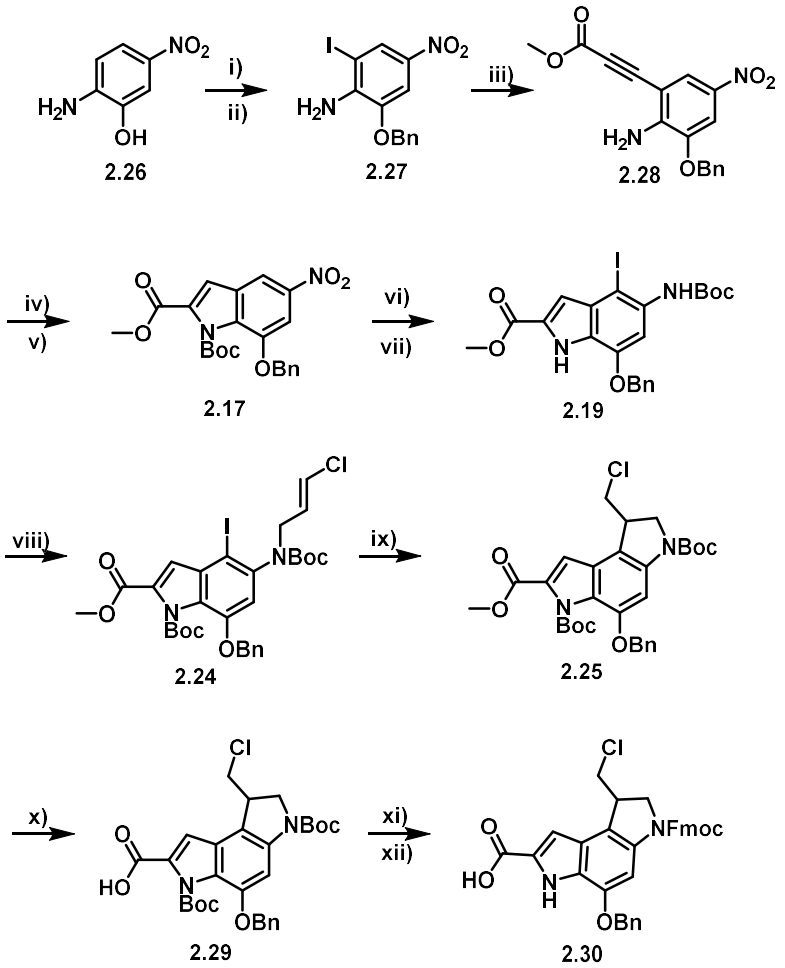
This free radical cyclisation to form the core DSA subunit has been the mainstay of racemic duocarmycin analogue synthesis since its first disclosure in 1997 and its application to CBI analogues.[9,10] Tietze and co-workers described the first application of this compound using tris (trimethylsilyl)silane (TTMSS) in place of the tributyl tin hydride as the reagent in synthesis. This 5-exo trig radical cyclisation has been successful and utilised in many synthetic routes to the duocarmycins.

2.1.3 Searcey synthesis of (+)-DSA for use in solid phase synthesis

The routes to synthesise duocarmycin SA typically involve long complex multistep synthetic routes. Searcey and co-workers aimed to develop synthesis of duocarmycin SA and analogues of this compound in a more efficient way. They took inspiration from building peptides in a sequential way on solid phase and applied this to the different subunits (alkylating and binding units) of duocarmycin SA. In using this approach, they could not only make duocarmycin SA in a more straightforward

synthesis but also generate a library of other analogues by breaking down the compound into different subunits.

In 2015, the Searcey laboratory described the synthesis of the (+)-DSA subunit that was protected with fluorenylmethoxycarbonyl (Fmoc), so is suitable for the application of solid-phase synthesis.[11] Solid phase synthesis is further discussed in **Chapter 3**. This synthetic route produced a racemic mixture, which was ultimately separated into the enantiomers by supercritical fluid chromatography (SFC).[11] The route shown in **Scheme 2.4** will initially be followed in this chapter to make the natural DSA subunit.



Scheme 2.4: *The synthetic route followed by Searcey and co-workers to obtain (+)-DSA unit with Fmoc protection.*

i)BnBr, K_2CO_3, DMF, 98.5%, ii) NIS, H_2SO_4, DMF, 92%, iii) $ZnBr_2$, $Pd(PPh_3)_2Cl_2$, DIPEA, methyl propiolate, DMF, 66 ^oC, N_2, 77%, iv) TBAF, THF, 66 ^oC, v) Boc_2O, DMAP, CH_2Cl_2,, 39%, vi) Zn, NH_4Cl, Boc_2O, DMAP, THF/H_2O vii) NIS, H_2SO_4, DMF, 59%, viii) potassium tert-butoxide, 1,3-dichloropropene, DMF, 62%, ix) AIBN, Tris(trimethylsilyl)silane (TTMSS), toluene, 90 ^oC, N_2, 70% x) LiOH, THF/MeOH/H_2O, 100%, xi) 4M HCl in dioxane, xii) Fmoc-Cl, $NaHCO_3$, THF/H_2O, 80%

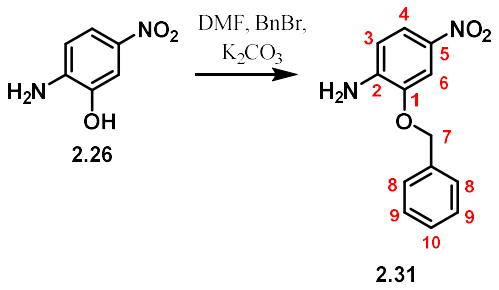
In this chapter, compounds **2.25** and the free acid derivative **2.29** shown in **Scheme 2.4** were the target compounds following the synthesis previously reported by the Searcey group. This project will not use solid phase methods, although there will be a discussion of this in **Chapter 3**.

As duocarmycins show promise as targeted therapies, finding new synthetic routes to these, or improving yields and ease of synthesis is important.

2.2. Initial synthetic route

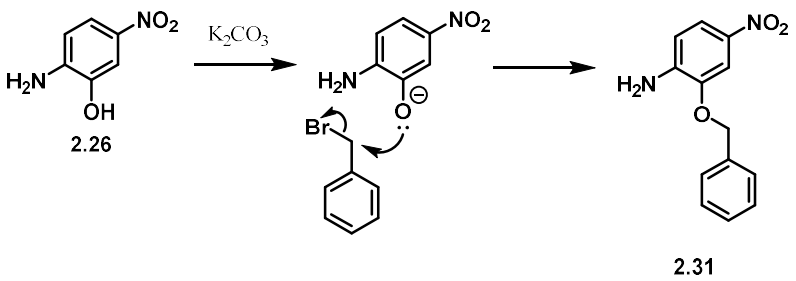
The total synthesis for Fmoc protected DSA was described by Searcey *et al* (**Scheme 2.4**). This synthesis was initially followed with minor changes to alternative reaction conditions in attempts to overcome bottlenecks in the synthesis and improve the overall yield. The overall scheme follows this route to the desired final product, **2.25**. The synthesis also shows some extra steps as products were worked up and purified, unlike the original synthesis, where several steps were 'telescoped'.

2.2.1 Benzyl protection of starting material to give compound 2.31



Scheme 2.5: *Benzyl protection of starting material* ***2.26*** *to give compound* ***2.31.***

The first step in the synthetic route utilised 2-amino-5-nitrophenol (**2.26**) as the readily available starting material for this synthesis (**Scheme 2.5**). The reaction to introduce this protecting group, the benzyl ether, proceeded via nucleophilic substitution reaction. The reaction was carried out in DMF at room temperature by treatment of the starting material with BnBr and K_2CO_3. Here the K_2CO_3 acts as base to deprotonate the phenol group, hence generating a more reactive phenoxide ion. As shown in **Scheme 2.6**, the lone pair on the oxygen attacks the delta positive carbon, to form this carbon-oxygen bond in an SN2 nucleophilic substitution reaction.

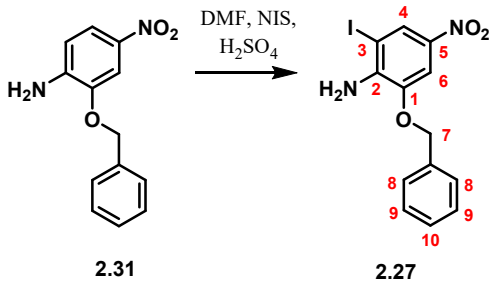


Scheme 2.6: *Mechanism of benzyl protection.*

The product was precipitated from the reaction by pouring over ice. To remove excess DMF the product was washed with cold water and the filtered product was dried *in vacuo* to obtain a near quantitative yield of 96%.

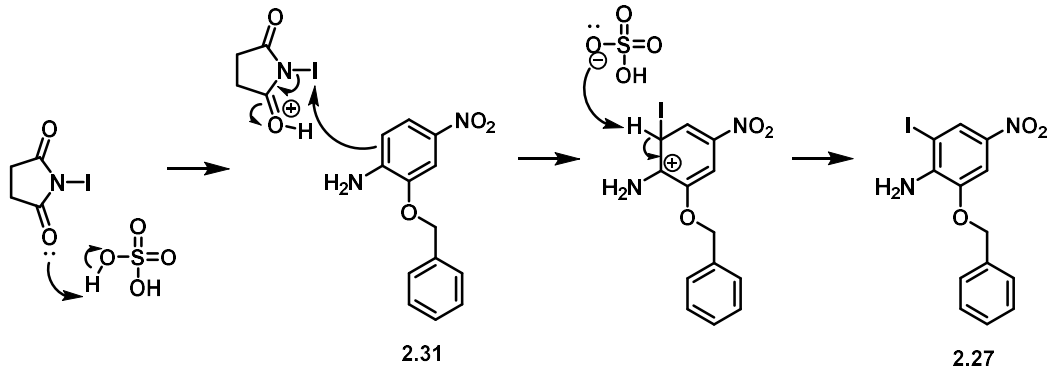
All data collected agrees with published procedure and confirmed the successful synthesis of the benzylated product.[11] ^{1}H NMR correctly accounted for all protons. There are in theory seven environments, but only five are seen as the benzyl group is observed as a multiplet in the region 7.36-7.46 ppm with the correct integration of five accounting for protons on C8, 9 and 10 (**Scheme 2.5**). A doublet of doublets is observed at 7.84 ppm; this integrates for one and accounts for the proton at the C4 position. This was supported by the *J*-values obtained (8.7 and 2.4 Hz). The doublet integrating for one at 7.78 ppm is a result of the aromatic proton at C6 (supported by showing a low J = 2.4 Hz value showing the meta coupling). The proton on the C3 position of the aromatic ring is seen as a doublet integrating for one at 6.67 ppm, supported by the high *J*-value of 8.6 Hz due to its ortho coupling to the C4. The singlet at 5.15 ppm that integrates for two is characteristic for the protons on the carbon adjacent to the oxygen on the benzyl group, C7. The ^{13}C NMR data agrees with that of the published procedure and correctly accounts for all the eleven carbon environments.[11] The IR data is also supported by that in literature and characteristic peaks confirm the correct product has been synthesised.[11] Peaks at 3482 cm^{-1} and 3358 cm^{-1} are evidence of an N-H functional group. The two peaks at 1578 cm^{-1} and 1386 cm^{-1} are characteristic of the N-O stretches in a nitro group.

2.2.2 Iodination using N-iodosuccinimide to give compound 2.27



Scheme 2.7: *Iodination of* ***2.31*** *to* ***2.27*** *using N-iodosuccinimide.*

The next step was an iodination at the C3 position (**Scheme 2.7**). This regioselective iodination was achieved by treating starting material **2.31** with *N*-iodosuccinimide in DMF. The reaction was catalysed by sulfuric acid (**Scheme 2.8**). This regioselective iodination was complete after 4 hours at room temperature. A solid precipitate was formed when the reaction was poured over crushed ice. The bright yellow solid was then filtered and washed with water removing excess DMF, followed by washing with hexane (purple washings often seen, showing excess iodine removed). The product was left to dry and the reaction afforded a yield of 78%.



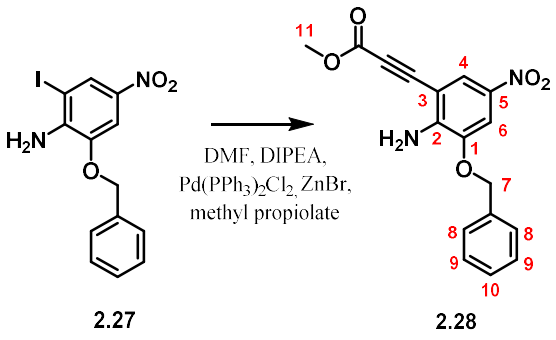
Scheme 2.8: *Mechanism of iodination.*

Successful iodination was shown by characterisation using ^{1}H NMR. When comparing to the previous NMR there was a similar shift in the five aromatic protons from the benzyl group, and similarly the singlet integrating for two because the protons at C7 (**Scheme 2.7**) had a very similar shift at 5.18 ppm. Notably, the previous peak at 6.67 ppm disappeared, as this proton has now been exchanged for an iodine. The proton adjacent to this iodine on the C4 shows an expected shift downfield from 7.78 ppm to

8.31 ppm. This shift is expected as it is now adjacent to a more electron withdrawing group. Further support of this assignment is that there is only one *J*-coupling now at 2.4 Hz which is a result of the meta coupling to the proton at position C6. The proton at C6 shows a similar shift at 7.76 ppm which is expected due to there being no changes to functional groups within two bond distances. ^{13}C NMR accounts correctly for all 11 carbon environments and the shifts correlate to that in the procedure described in the literature.[11]

When compared with the starting material, it can be seen that there is no significant difference in the IR. Alkyl halides in general have vibrational frequencies in the region 850-515 cm^{-1}. As this is the fingerprint region, it is difficult to see a clear appearance of this peak. There are slight shifts, but the same characteristic peaks confirm the correct product functional groups with a peak at 3475 cm^{-1} and 3378 cm^{-1} indicating the presence of an N-H functional group. The two peaks at 1495 cm^{-1} and 1276 cm^{-1} are characteristic of the N-O stretches in a nitro group.

2.2.3 Negishi cross coupling to give compound 2.28

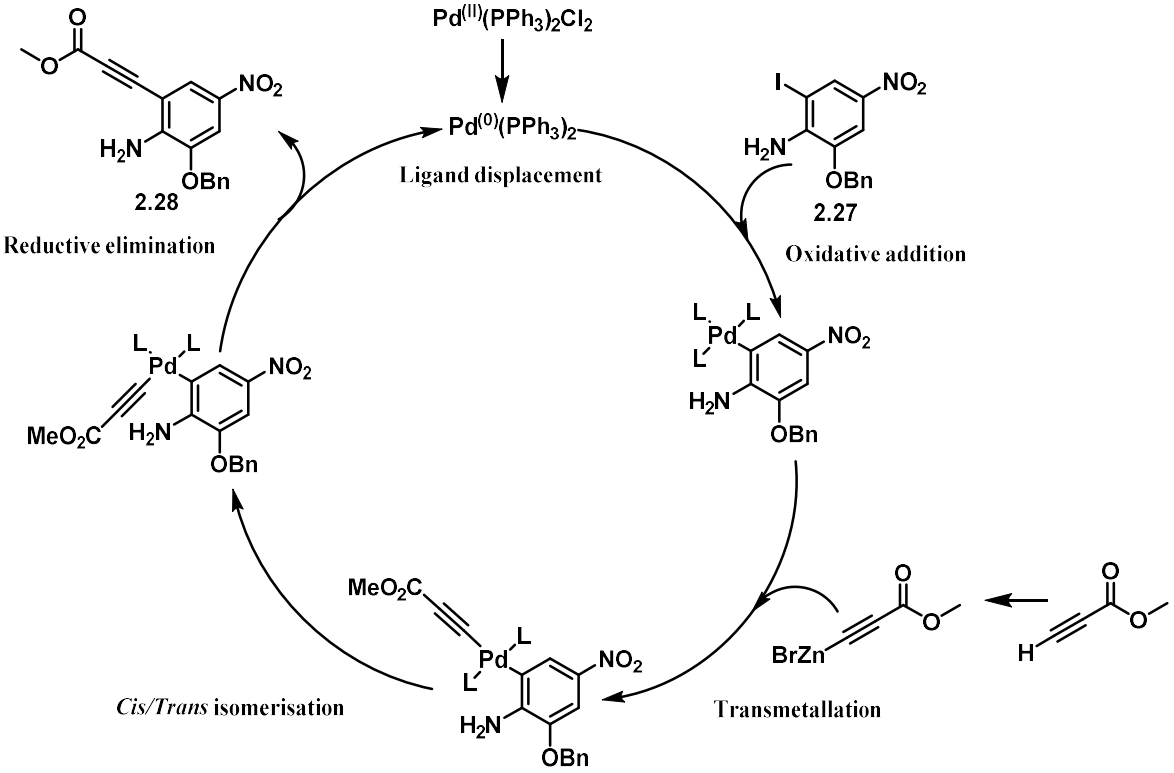


2.27 2.28

Scheme 2.9: *Negishi cross coupling of* ***2.27*** *to add the alkyl group and give compound* ***2.28.***

The introduction of the alkyne group was achieved using a Negishi cross coupling (**Scheme 2.9**). Initially, this reaction was repeated and carried out with parallel experiments as 10 g was the largest amount that gave acceptable yields. The starting iodinated compound was dissolved in anhydrous DMF and degassed with nitrogen for 1 hour. Methyl propiolate, $Pd(PPh_3)_2Cl_2$, $ZnBr_2$ and DIPEA were then added before stirring the reaction overnight at 66 ºC. The reaction was complete after 24 hours poured over crushed ice ready to work up.

This Negishi cross coupling resembles that of a Sonogashira reaction as it is a palladium catalysed cross coupling between an aryl halide and an alkyne; however, it does not use a copper co-catalyst in transmetallation characteristic of this type of reaction.[12] In 2001 Negishi explored varying these reactions to investigate the use of alkynylzinc compounds to improve reactions with electron deficient alkynes such as the methyl propiolate used in this reaction.[13] This Negishi style approach was used by Sakamoto in their paper to synthesise duocarmycin SA.[13]



Scheme 2.10: *Catalytic cycle showing the Negishi cross coupling.*

The catalytic cycle shows how some of the different reagents are used during the reaction (**Scheme 2.10**). The first step involves the palladium catalyst undergoing ligand displacement to go from Pd in the oxidation state 2 to Pd in the oxidation state 0. Oxidative addition then occurs between the palladium catalyst and the iodoaryl compound. The third step is transmetallation with zinc bromide. Cis/trans isomerisation around the palladium complex then allows the final step of reductive elimination to occur to give the final compound.

This reaction presented several challenges. After cooling to room temperature, the reaction was worked up by precipitating the product. The solution was poured over crushed ice, and the precipitate was filtered. The reaction afforded a thick sticky crude mixture that was hard to filter and hard to purify. There were several different attempts to try and improve this workup.

The product was wet loaded directly on to the silica column, but this caused streaking on the column and prolonged elution. Wet loading was tried again onto a silica plug, with the hope to remove the initial baseline impurity and subsequently make the final column easier. However, due to the viscosity of the mixture this was not efficient timewise. Efforts were made to address the baseline impurity by dry loading onto a silica plug. Despite using a substantial amount of silica to convert the viscous oil into a more manageable powder, difficulties persisted when solvents were introduced, leading to similar issues in the elution process.

Given that the reaction took place in DMF, additional attempts were made to enhance purification. Washing the mixture more extensively with water and triturating using cold ethanol were explored, yielding some improvement. Ultimately, the optimal method followed a published procedure, but with the modification of washing the product with water and subsequent trituration with cold ethanol, which effectively mitigated the inherent stickiness and tar-like properties of the compound.

The reaction was explored across various scales, ranging from 20 g, resulting in a notably poor yield, to 100 mg, which exhibited improved yield but still displayed inconsistency. Given the extensive synthetic route involved, a compromise was established at a 5 g scale, and parallel reactions were conducted. While individual columns marginally enhanced yields compared to combined efforts, the extended workup time rendered the process impractical for negligible material gains. Consequently, reactions were combined during the purification step to optimize time efficiency.

The low yields could be for several reasons. One issue could be the use of methyl propiolate; this is an electron-deficient alkyne, which according to Sakamoto is a well-known issue in these types of reaction.[13] As well as this, the substituents on the

aromatic ring are known to affect yields. Studies by Sakamoto showed that the couplings gave moderate yields if there were either no functional groups or weakly electron donating groups. However, if the groups were electron donating there were poor yields and sometimes the reaction wouldn't occur at all.[13] Here there is one electron donating group, the NH_2, which may contribute to the low yields; however, there are two electron withdrawing groups, the NO_2 and the benzyl ether group, which in theory should promote the reaction.

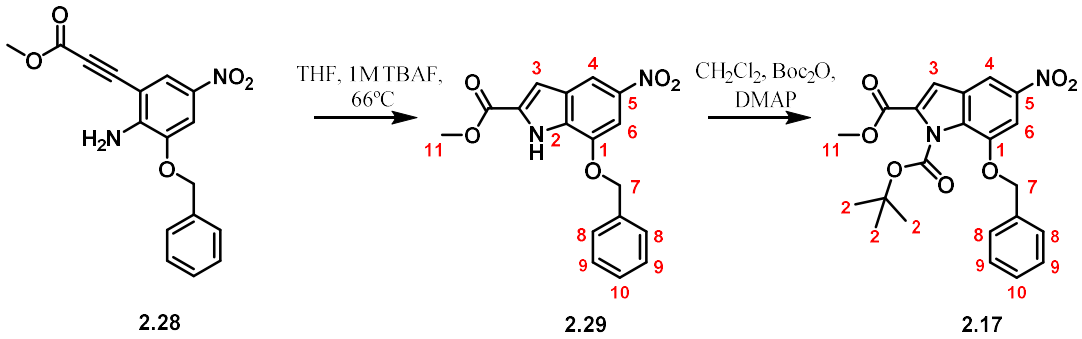
The repetition of this step, even when adhering to the improvements in workup, yielded inconsistent results. Other members of the research group also encountered difficulties with the same step. Despite the yield variations, the correct product was isolated as an orange solid, and its structure and purity was confirmed using ^{1}H NMR, ^{13}C NMR and IR. The ^{1}H NMR was recorded in chloroform and was in agreement with literature procedure correctly accounting for all fourteen hydrogens in the expected environments. The benzyl group is shown as a multiplet in the region ranging from 7.39-7.35 ppm correctly accounting for all five protons. There are two aromatic protons at positions C4 and C6 that have integrated for one proton, a doublet was observed at 8.06 ppm and the other doublet at 7.76 ppm. Assigning these, C4 is likely to be the doublet at 8.06 ppm. It has shifted from 8.31 ppm up field, which is expected as it is no longer ortho to the electron withdrawing iodo group. The doublet at 7.76 ppm is likely to be the single proton at position C6; the previous shift for this position was very similar and this environment has not changed. This doublet still shows a low *J*-coupling value at 2.3 Hz that is characteristic of the meta coupling to the position C4. The broad singlet observed at 5.31 ppm correctly accounts for the two protons on the NH_2 group. The singlet at 5.17 ppm accounts for the protons on the carbon chain of the benzyl group at position C7; this shift is characteristic for these two protons. The singlet at 3.86 ppm is a result of the addition of methyl propiolate and shows the three protons at region C11. There are 15 different carbon environments on the compound, and all were shown on the ^{13}C NMR. The shifts conform with that in the literature data.[11]

There are slight changes in the IR, but the same characteristic peaks confirm the correct product functional groups with a peak at 3498 cm^{-1} and 3347 cm^{-1} indicating the presence of an N-H functional group. The two peaks at 1506 cm^{-1} and 1298 cm^{-1}

are characteristic of the N-O stretches in a nitro group. Notable differences in the IR between the starting material and final product highlight successful reaction, as there is now a carbonyl group C=O at 1697 cm^{-1}. There is also an alkyne group in the final product which gives expected shifts of 2203 cm^{-1}. The disappearance of the alkyl halide group was not observed, likely due to its presence in the fingerprint region, making changes harder to observe.

Inconsistent and low yields so early in a long synthetic route causes bottlenecks, so finding ways to try and avoid them is important. Designing new routes to improve the overall synthetic yield will be discussed in section **2.3.**

2.2.4 TBAF mediated cyclisation and Boc protection of the indole to give compound 2.17



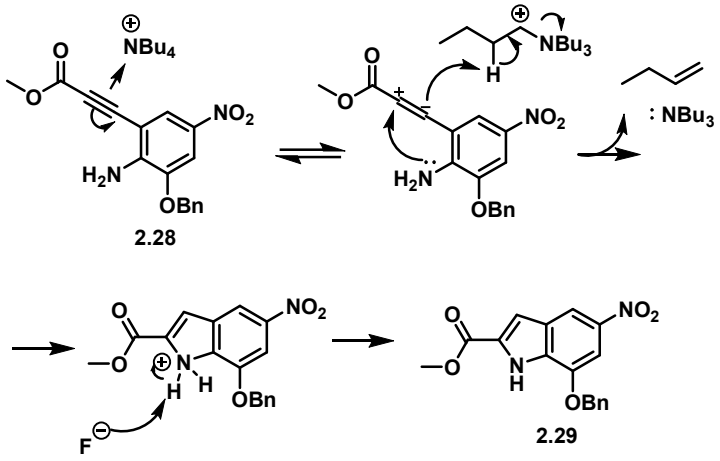
Scheme 2.11: *Cyclisation and subsequent Boc protection of the indole.*

The next step was the TBAF mediated cyclisation (**Scheme 2.11**). Compound **2.28** was dissolved in anhydrous THF and treated with 1M Tetra-*n*-butylammonium fluoride (TBAF) solution in THF. The solution was refluxed at 66ºC for 1 h. After cooling, the solvent was then removed under reduced pressure. The product was then dissolved in ethyl acetate and washed with water, followed by brine. A crude NMR was then run to check that the compound was fully cyclised and had no significant impurities. If the NMR and TLC indicated the cyclisation reaction had gone to completion, the crude was dissolved in CH_2Cl_2 and treated with di-*tert*-butyl dicarbonate (Boc_2O) and 4-dimethylaminopyridine (DMAP) (**Scheme 2.11**). The reaction was stirred at room temperature for 2 h. The solvent was removed under reduced pressure before dry loading onto silica for column chromatography. The

product was then purified using automated flash chromatography with a gradient of 0% to 10% ethyl acetate in hexane. This afforded the final product, **2.17,** as a bright yellow solid with a yield of 46%.

The cyclisation and Boc protection in the published procedure was carried out in one step. When trying to replicate this, there were difficulties, as it appeared that that only half the product was Boc protected and the other half was not. Reactions were followed closely with TLC. The cyclised product (**2.29**) and the cyclised and Boc protected product (**2.17**) had similar R_f values, posing challenges in column chromatography purification. Different purification conditions to separate the mixture of cyclised (**2.29**) and cyclised and Boc protected compounds (**2.17**) were trialled but were unsuccessful. Attempts to enhance reaction completion by adding more starting materials were unsuccessful. Therefore, the mixed material was re-Boc protected to give the final desired product, now allowing for the straightforward isolation of the pure product by column chromatography.

To try and avoid these issues and improve the low yields, the reaction was worked up and checked after cyclisation and if necessary purified after the cyclisation. Over the two steps after optimisation, there was a yield of 46% obtained, which compared with a yield of 33% when avoiding the more challenging purifications.

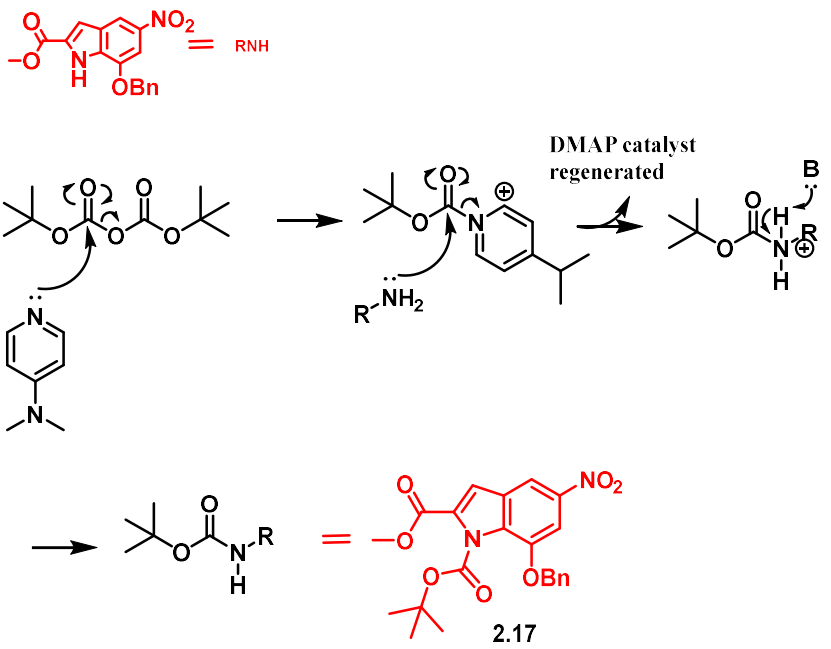


Scheme 2.12: *Proposed mechanism for TBAF promoted ring closure.*

The mechanism for this cyclisation was proposed by Sakamoto.[14,15] Results supported evidence of both the $(NBu_4)+$ and the fluoride ion being involved. They show several

mechanistic approaches in the article and from this the proposed route for this reaction may be the one seen in **Scheme 2.12**. Here it can be seen that the alkyne in **2.28** becomes activated by coordination to the tetra butyl ammonium ion. This forms a charged double bond that is more susceptible to nucleophilic attack from the lone pair on the amino group. Protonation of the double bond is through an elimination reaction on the tetrabutylammonium ion. The fluoride ion stabilises the positive charge ammonium group by acting as a base to deprotonate the indole nitrogen, forming **2.29**.

The mechanism for the Boc protection can be seen in **Scheme 2.13**. This shows the reaction between Boc anhydride with the amine. DMAP acts as a catalytic base and activates the Boc anhydride. This acylated DMAP is susceptible to nucleophilic attack from the lone pair on the amine. Catalytic DMAP is regenerated by proton transfer, and the formation of the final Boc protected amine.

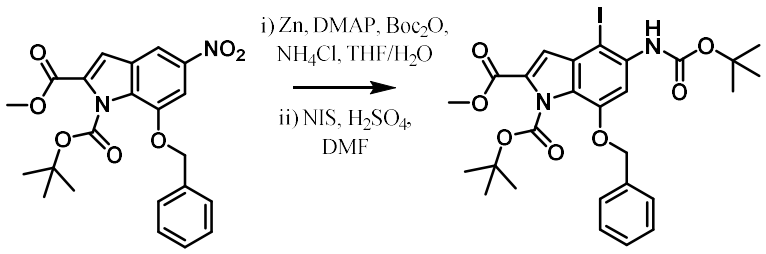


Scheme 2.13: *Mechanism of DMAP catalysed Boc protection of the indole.*

Successful cyclisation and Boc protection was supported by literature data.[11] The ^{1}H NMR showed a doublet at 8.26 ppm that integrates for one proton and corresponds to C6 (**Scheme 2.11**). The doublet at 7.68 ppm that integrates for one proton corresponds to C4. The peak seen upfield to these at 7.34 ppm is a singlet that corresponds to C3. It is shifted upfield in comparison to these other two environments (C4 and C6) as it

is not adjacent to any electron-withdrawing groups. The aromatic benzyl protons are observed as multiplets and have the correct total integration of five protons in the expected regions for aromaticity. The protons on C7 as part of the benzyl group are again in the expected region and seen at 5.32 ppm and correctly seen as a singlet that integrates for two protons. The singlet at 3.94 ppm that integrates for three protons is a result of the methyl group. The Boc group was in the expected region at 1.47 ppm with the expected splitting pattern and integration as a singlet that integrates for nine protons. ^{13}C NMR was run in chloroform and shows seventeen of the eighteen environments expected. The values obtained correspond closely to that seen in literature. However, at 27.3 ppm, only one environment is seen, whereas literature reports two very close together in this region (27.3 ppm and 27.9 ppm), therefore they may have overlapped. The IR disappearance of peaks at 3498 cm^{-1} and 3347 cm^{-1} that were present due to the NH_2 group is evidence for the cyclisation and Boc protection. The peak at 2203 cm^{-1} is also not present which was showing the presence of an alkyne group. Expected peaks for the carbonyl as well as the two peaks observed for the nitro group are still there, and are in support of literature values, showing that the correct compound was isolated.

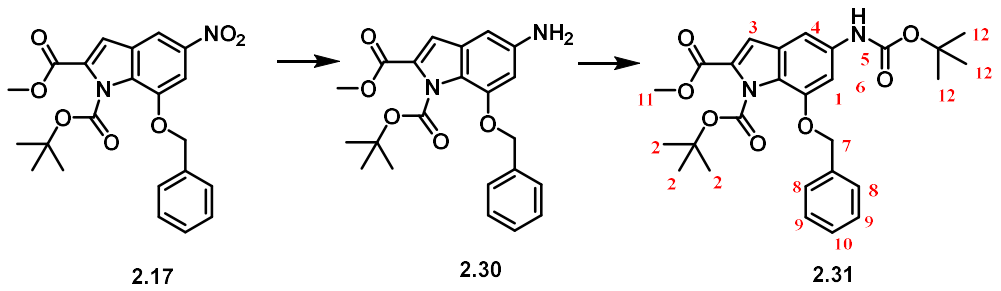
2.2.5 Reduction of nitro group followed by iodination to give compound 2.19



2.17 2.19

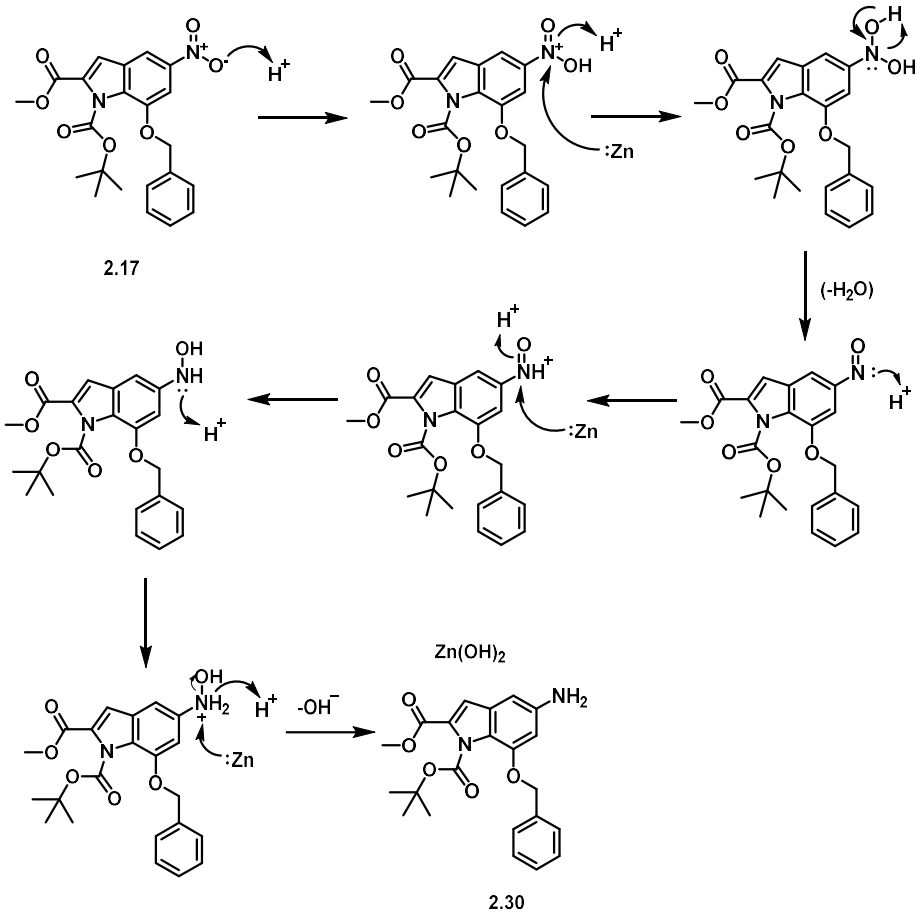
Scheme 2.14: *Reduction of compound **2.17** followed immediately by iodination to give compound **2.19.***

The published procedure reports going directly from the nitro indole to the Boc-protected iodo-compound **2.19** without purification of intermediates. In this work, enhanced yields were obtained when the intermediate reduced and protected indole, **2.31,** was purified prior to iodination.



***Scheme 2.15**: Reduction and Boc protection of the nitro group.*

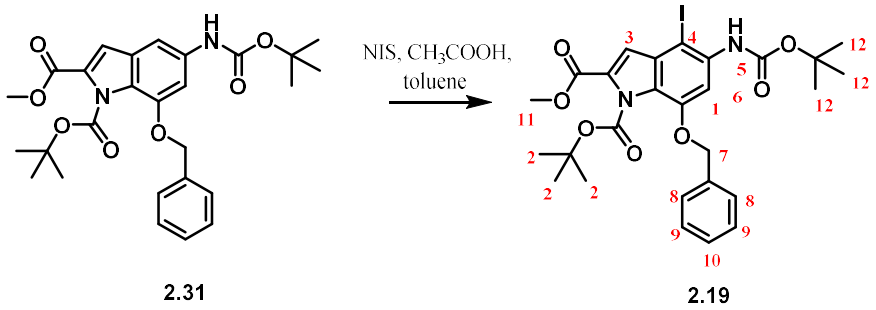
This first conversion involved a one pot reaction. Compound **2.17** was dissolved in THF and treated with zinc powder, ammonium chloride, Boc anhydride, DMAP and water. This reaction was left to stir overnight, and the excess zinc was filtered off. The reaction was concentrated in vacuo before dissolving the mixture in ethyl acetate and washing with water and brine. The organic layer was concentrated and dry loaded onto silica where it was subjected to flash chromatography.



***Scheme 2.16:** Suggested possible mechanism for nitro reduction.*

Zinc reacts with ammonium chloride to form reactive hydrogen. This hydrogen reduces the nitro group. A number of mechanisms for reductions of a nitro group have been postulated, and one route is shown in **Scheme 2.16**.

The Boc protection follows the same mechanism as shown previously in **Scheme 2.13**, to form final compound **2.31** with a yield of 73%. The ^{1}H NMR data accounted for all the protons. The splitting patterns and shifts are similar to the starting material with only minor changes; however, the more notable change and evidence for the Boc protection is the appearance of another singlet that integrates for nine protons in the expected region. These are seen as two intense singlets at 1.51 ppm and 1.44 ppm, and it is likely that C2 is the more downfield peak at 1.51 ppm because the carbamate is attached directly to the aromatic system. IR data shows minor differences compared to the starting material. This is expected due to the only change being an additional functional group which is already present in the starting material. The other key expected peaks as a result of the functional groups are seen and are in support of the correct compound being synthesised.



2.31 2.19

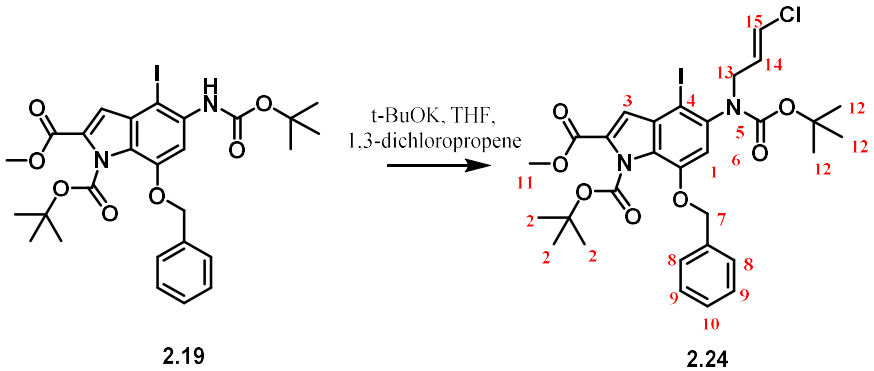
Scheme 2.17**: Iodination of **2.31** with NIS to form product **2.19.

The following iodination was carried out following two separate procedures (**Scheme 2.17**).[2,11] Initially, purified compound **2.31** was dissolved in THF, and treated with PTSA followed by *N*-iodosuccinimide. The reaction was worked up after 3 h by addition of ethyl acetate, and the solution was washed with water followed by brine. The viscous mixture was dry loaded onto silica and a manual column of 10% EtOAc:hexane was run. The product was obtained as a pure white solid with a disappointing yield of 56%.

To try and increase yields and improve ease of workup an alternative route was found.[2] The product **2.31** was dissolved in toluene and treated with acetic acid and NIS. This was stirred for three hours and loaded directly onto a flash column. It was purified using a gradient of 0 to 15% EtOAc in hexane and afforded the pure product as a white solid with a yield of 73%.

Both routes gave the correct compound in agreement with the literature. The ^{1}H NMR shows a broad singlet at 7.76 ppm due to the proton at C3. It is shifted more downfield due to the presence of the adjacent electron withdrawing iodo group. The singlet at 7.07 ppm is a result of C6 and correctly integrates for one proton. A broad singlet at 6.74 ppm is a result of the NH group in position C5. The aromatic benzyl protons are observed as multiplets and have the correct total integration of five protons in the expected regions for aromaticity. The protons on C7 as part of the benzyl group are again in the expected region and seen at 5.22 ppm and is correctly seen as a singlet that integrates for two protons. The singlet at 3.89 ppm that integrates for three protons is a result of the methyl group. The two strong singlets at 1.51 ppm and 1.38 ppm, are a result of the two Boc groups, correctly integrating for nine protons each. It is likely that C2 is the more downfield peak at 1.51 ppm, as noted previously. All twenty carbon environments were accounted for and values in agreement with literature data.[11] Further support for iodination were the IR results which also corresponded to known data.[11] However, the same key functional groups still show they are present due to key peaks seen carbonyl at 1763 cm^{-1} and 1716 cm^{-1}, N-H at 3357 cm^{-1}

2.2.6 Alkylation using 1,3-dichloropropene to give compound 2.24

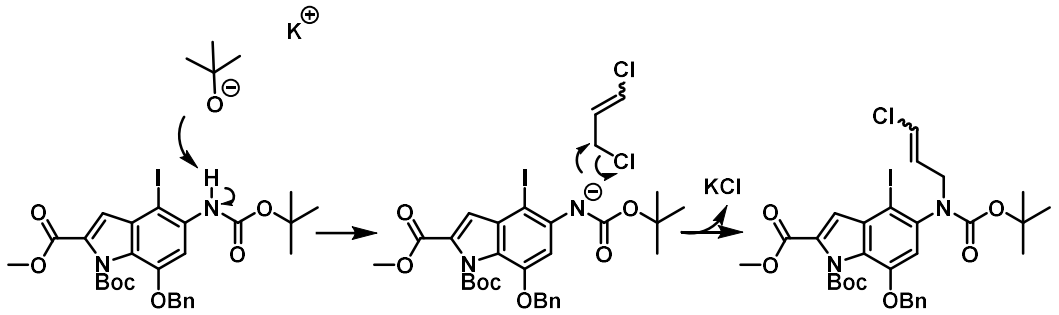


2.19 2.24

Scheme 2.18: *Alkylation using 1,3-dichloropropene.*

The published procedure was followed with minor variations (**Scheme 2.18**).[11] The starting material **2.19** was dissolved in anhydrous THF and treated with t-BuOK and 1,3-dichloropropene. After the reaction had stirred for 1.5 hours, it was cooled in an ice bath and treated dropwise with a solution of ammonium chloride. EtOAc was added to the mixture and several washes with water followed by brine. The solution was dried *in vacuo* and was co-evaporated with CH_2Cl_2 to form crude mixture. This was dry loaded onto silica and purified using flash chromatography, on a gradient of 0 to 15% ethyl acetate in hexane, giving final product as a light brown foam with a yield of **82**%.

This mechanism is likely to proceed via S_N2 mechanism (**Scheme 2.19**). The potassium tert-butoxide is acting as a strong base to deprotonate the amide. Once deprotonated the negatively charged amide is able to react with the delta positive carbon adjacent to electron withdrawing chloride group. The chloride is able to act as a good leaving group and forms a salt with the potassium. This can be removed after the reaction in the workup.

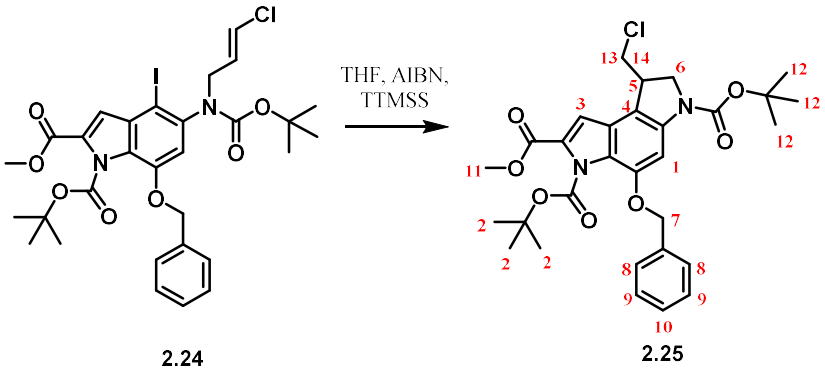


Scheme 2.19: *Introduction of the vinyl chloride to give* ***2.24.***

Despite affording a high yield and clean product, the NMR data was broad and hard to interpret, as it was a mixture of the *E/Z* isomers. This was due to the 1,3-dichloropropene being a mixture of the *cis* and *trans* isomers. Despite this, clear peaks were seen and all data correlated to that in the literature. The ^{1}H NMR correctly showed all protons with notable new peaks compared to starting material NMR. The appearance of new peaks in the alkene region showed that product **2.24** was correctly isolated. The multiplet between 5.19-5.30 ppm integrated for two protons and is showing environment C13. Two further multiplets at 4.31-4.49 ppm and 4.11-4.20 ppm each integrated for one proton and were a result of C15 and C14 respectively.

The stability of the product became a concern during an extended time. The compound decomposed when stored for an excess of one month, even at low temperature. As such, a strategy was implemented to immediately move small bathes through to the next step and store these at -80 °C prior to the reaction. These precautions helped and were important as maintaining product stability was important so close to the end of the synthesis.

2.2.7 Radical cyclisation using AIBN

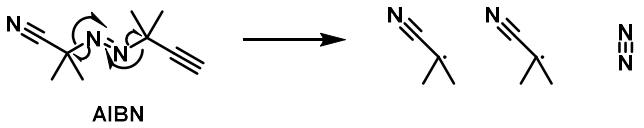


Scheme 2.20: *Radical cyclisation using AIBN.*

The conditions for the radical cyclisation is shown in **Scheme 2.20**. Product **2.24** was dissolved in anhydrous THF and degassed for 1 h in nitrogen. AIBN and tris(trimethylsilyl)silane (TTMSS) were then added, before the solution was refluxed until the reaction appeared to be complete by TLC, ranging from 1.5 to 4 h. After the solution was cooled to room temperature, the solvent was removed under reduced pressure and the compound was wet loaded and purified using flash chromatography. A gradient of 0 to 4% ethyl acetate in hexane gave the final compound as a white foam, which solidified to a light brown powder.

This is a free radical reaction and an intramolecular cyclisation. This reaction uses AIBN, which initiates the radical reactions. There are three stages in a radical reaction; initiation (**Scheme 2.21**), propagation (**Scheme 2.22**, **Scheme 2.23** and **Scheme 2.24**) and termination (**Scheme 2.25**). These three stages are summarised in the following figures with brief descriptions as to what happens at each stage.

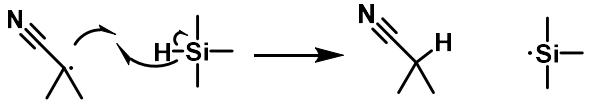
First is initiation, where the radical species are first formed. This homolytic cleavage usually requires a catalyst, such as heat, UV radiation, a metal catalyst etc. In the case of this radical cyclisation, AIBN is activated by heat as the reaction is heated at reflux.



Scheme 2.21: *Initiation - generation of the radicals from AIBN.*

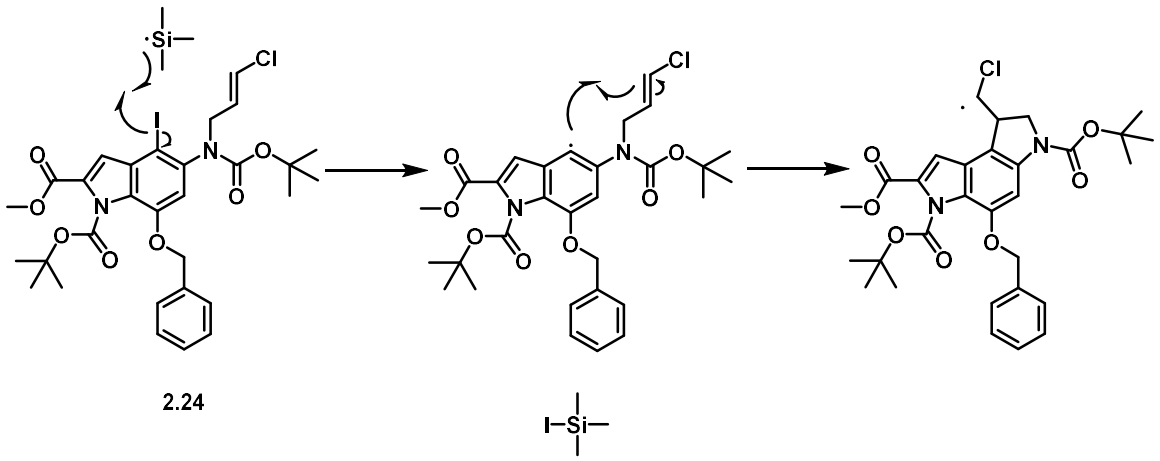
AIBN acts as the initiator. **Scheme 2.21** shows that AIBN decomposes, forming two radicals and eliminating a byproduct of nitrogen gas. Two identical free radicals are formed, which are then able to go onto the next step of a radical reaction, known as propagation.

This propagation reaction is carried out using trimethylsilyl silane solution (TTMSS) in toluene at reflux. Once the AIBN has generated the two free radicals, they react with the TTMSS to form a new radical **(Scheme 2.22)**.



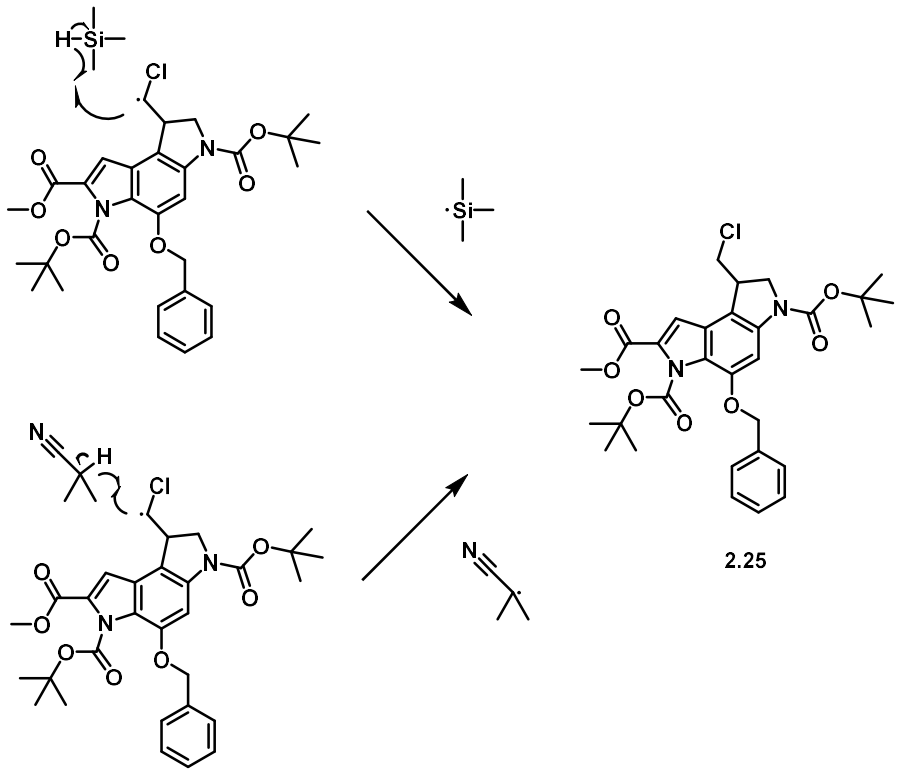
Scheme 2.22: *Propagation cascade reaction - formation of the TTMSS radical.*

This radical is then able to react with compound **2.24** in another propagation reaction, **Scheme 2.23** highlights the mechanism for the 5-exo-trig cyclisation and how this compound forms the duocarmycin SA alkylating subunit.



Scheme 2.23: *Propagation showing the 5-exo-trig radical cyclisation.*

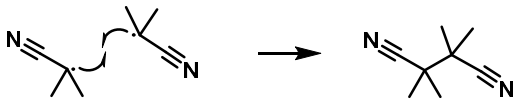
The step to generate the final cyclised compound **2.25** can proceed via two routes, summarised in **Scheme 2.24.**



Scheme 2.24: *Protonation of cyclised structure to give final compound* ***2.25.***

The TTMSS radical is regenerated, and the final step of the reaction continues without further generations of radicals.

The third and final step in a radical reaction is termination (**Scheme 2.25**). Here, the two radicals are able to recombine and form a stable bond.



Scheme 2.25: *Termination - radical recombination*

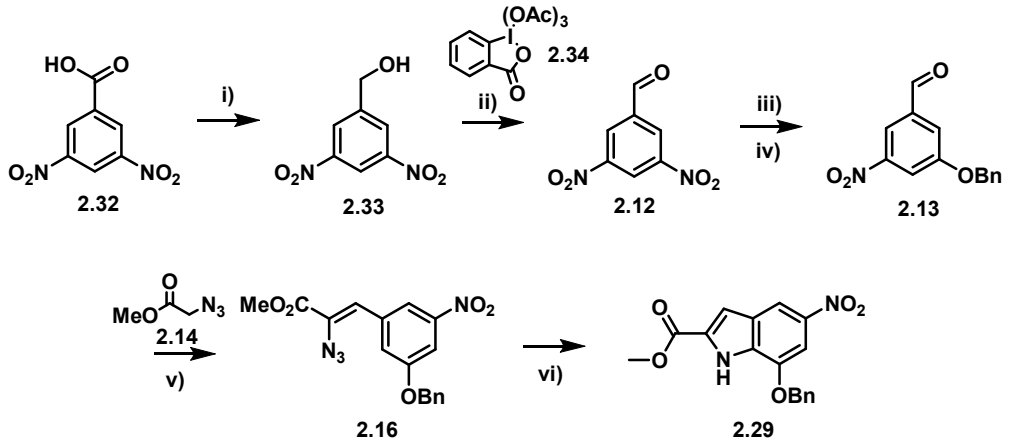
This reaction was successful, and a range of yields was achieved from 39% to 82%. Evidence of the success was supported by that of literature data.[11] ^{1}H NMR were run in DMSO. There is a multiplet in the range 7.29-7.46 ppm that is characteristic of aromatic protons and integrates for seven protons. C8, C9 and C10 aromatic protons of the benzyl group account for five of these protons. The other two protons are those of aromatic C1 and C3. The C7 benzyl protons, seen at 5.27 ppm, are shown as a singlet integrating for two protons. The triplet that integrates for one proton and is seen at 4.13 ppm is the result of C5. The multiplet at 3.89-4.05 ppm that integrates for four protons is showing two protons from C13 and two protons from C6. The singlet integrating for three at 3.87 ppm is characteristic of the methyl at position C11. The same two characteristic Boc protons are seen as two singlets at 1.48 and 1.39 ppm, both integrating for nine protons. The ^{13}C NMR was also run in DMSO and also accounts for all twenty-three environments. There were no significant changes in the IR from the starting material, as the functional groups stay the same, but the peaks correspond to that of literature data and highlight the key functional groups.

2.2.8 Bottlenecks in route and overall yield

This synthetic route to the target compound gave an overall yield of 2.7% but suffered from several bottlenecks. Of particular note is the Negishi step (**Scheme 2.9**), which gave inconsistent yields, a challenging workup and purification. *(This overall yield calculation is assuming a 25% yield for this inconsistent step).* To develop a new route, it was decided to investigate an alternative approach to the formation of the indole ring system.

2.3 Improvements to synthesis – a new route

To try and avoid these bottlenecks, a new route to cyclised compound **2.29** was investigated. Inspiration was taken from a previous route to synthesise (+)-duocarmycin SA enantioselectively.[2] The route for this synthesis can be seen in at the start of this chapter (**Scheme 2.2**). The starting aldehyde was the first target. Several routes were explored to find a suitable method for synthesising the starting aldehyde, compound **2.12.** 3,5-dinitrobenzoic acid (**2.32**) was used as the starting material as it is readily available.



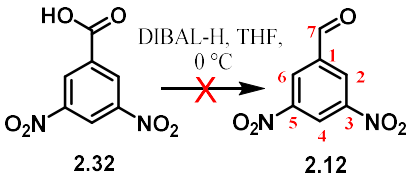
***Scheme 2.26**: Synthetic route avoiding bottleneck*

i)BH_3·THF, 0 °C to rt, 18 h, ii) Dess-Martin periodinane, DCM rt, 24 h, iii)K_2CO_3, benzaldehyde oxime, DMF, 90 °C, 1.25 h, iv) BnBr, DMF, 25 °C, 2h, v) methyl azidoacetate, NaOMe, -41 °C, 48 h, vi) xylenes, 140 °C

2.3.1 Synthesis of starting material (2.12)

2.3.1.1 Partial reduction using DIBAL

The reduction of the acid (**2.32**) to the aldehyde (**2.12**) was first attempted using diisobutylaluminum hydride (DIBAL) to partially reduce the carboxylic acid to an aldehyde (**Scheme 2.27**).

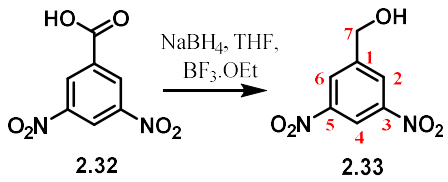


***Scheme 2.27**: Partial reduction of **2.32** to **2.12**.*

DIBAL-H, compound **2.32** and THF were stirred at 0 °C for 2 h until all the starting material appeared to be consumed by TLC analysis. The reaction was quenched with ethyl acetate, $KHSO_4$ and water, then worked up by extraction with ethyl acetate. The H NMR showed a mixture of products with no aldehyde present as an expected peak around 10 ppm was absent, characteristic of a proton adjacent to a carbonyl.

2.3.1.2 Reduction using sodium borohydride to give the alcohol compound 2.33

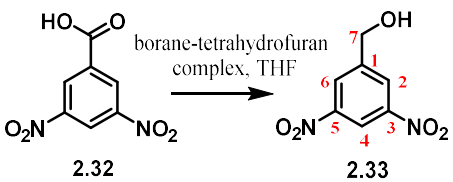
A more suitable method to form an aldehyde from a carboxylic acid is to reduce to the alcohol and oxidise the alcohol to the aldehyde, as it is difficult to partially reduce carboxylic acids. Two methods were followed to try and determine which provides the best yield and ease of synthesis.[16,17] The first reduction tried used boron trifluoride and sodium borohydride (**Scheme 2.28**).



Scheme 2.28: Reduction of **2.32** to **2.33** using sodium borohydride and boron trifluoride.

The preparation of **2.33** followed the procedure previously described by Caron.[17] Firstly sodium borohydride was suspended in dry THF and cooled to 0 °C before a solution of compound **2.33** in THF and boron trifluoride etherate ($BF_3{\cdot}OEt$) was added and warmed to 25 °C for 3 hours. The solution was quenched with 1 M HCl and extracted with CH_2Cl_2. ^{1}H NMR analysis showed that this reaction was successful and the data for the isolated product correlated with the literature data.[17] The reaction afforded a yield of 40%.

Another method used to try the reduction followed the work of Barker and co-workers who used a borane complex to reduce the carboxylic acid (**Scheme 2.29**).[16]



Scheme 2.29: Reduction of **2.32** to **2.33** using a borane-THF complex.

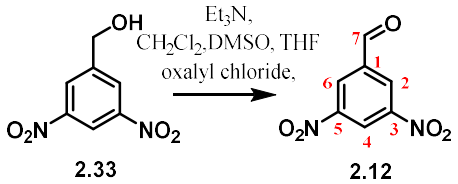
Minor changes were made to the procedure to try to increase the yield and purity obtained. Compound **2.32** was dissolved in anhydrous THF and cooled to 0 °C before treating with the borane-tetrahydrofuran complex. This was stirred for 3 h before warming to 25 °C overnight. Differing from the literature, further addition of the borane complex was needed throughout the reaction as TLC revealed starting material still present. The reaction was quenched with water after TLC revealed all starting material was consumed. The reaction was concentrated *in vacuo* and precipitated solid was collected by filtration. The aqueous washings were then extracted with EtOAc. The literature reports that the product can be used without further purification.[16] However, some impurities were removed by trituration with toluene. This removed the volatile impurities under reduced pressure. The yield obtained with these minor changes afforded a 99.4% yield compared with a literature yield of 89%.[16] The data obtained for this reaction supported that of the literature.[16] ^{1}H NMR in deuterated methanol showed correctly all three hydrogen environments, with the correct integrations and environments shown for the aromatic hydrogens at ~8 ppm. The aliphatic proton adjacent to the alcohol group integrated for two protons as a singlet in the expected region around at 4.8 ppm. The ^{13}C NMR also showed five peaks for the five carbon environments in the product. IR data further supported the rest showing the key functional groups in the expected regions with a broad peak at 3248 cm^{-1} indicative of an OH group, and a peak at 1519 cm^{-1} showing presence of the nitro groups.

2.3.2 Attempts to synthesise dinitro benzaldehyde compound 2.12

There were several attempts to synthesise the aldehyde, which are discussed in the following sections.

2.3.2.1 Swern oxidation to synthesise 2.12

The next step is oxidation of the alcohol to aldehyde. One approach to this is a Swern oxidation using oxalyl chloride in CH_2Cl_2, DMSO, THF and triethylamine (**Scheme 2.30**).[18]

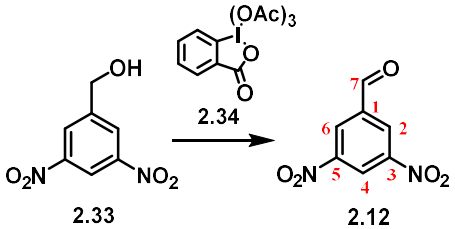


***Scheme 2.30*: Oxidation of **2.33** to **2.12** by Swern oxidation.**

Oxalyl chloride, anhydrous CH_2Cl_2 and DMSO were cooled to to -61 °C and treated with compound **2.33** dissolved in THF. It was stirred under nitrogen for 30 minutes before treating with triethylamine. The reaction was then warmed to room temperature and quenched with H_2O, washed with $NaHCO_3$ followed by 1M HCl and then brine before the organic layer was concentrated under reduced pressure. ^1H NMR of the crude product showed the reaction was successful, due to the expected peaks including the aldehyde proton at 10.2 ppm. The product was not purified but the crude yield obtained for this was 38%. It is important when doing a multi-step synthesis to have high yields, so an alternative method was explored to try and improve this low yield.

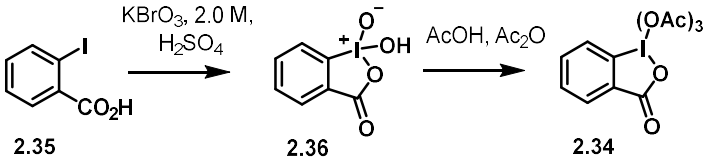
2.3.2.2 Dess-Martin periodinane for oxidation

The oxidation of 3,5-dinitrobenzyl alcohol can also be carried out using Dess-Martin periodinane (DMP), compound **2.34**, to make the aldehyde **2.12** (**Scheme 2.31**).[19]



***Scheme 2.31**:* Oxidation of **2.33** with DMP to give aldehyde **2.13**

In order to follow this procedure, the oxidising agent needed to be synthesised. The oxidising agent DMP was synthesised following a published procedure.[20] This two-step synthesis is summarised in **Scheme 2.32**.



Scheme 2.32: Synthesis of Dess-Martin Periodinane, **2.34**.

The intermediate compound and the final compound are known to be heat and shock sensitive so all work involving these was carried out behind a blast shield. Care was also taken on scaling up of the DMP as the internal temperatures could rise very suddenly if heating too quickly in the last step.

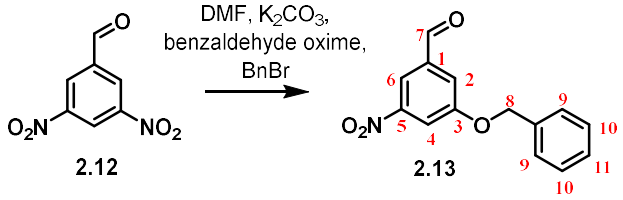
Potassium bromate and 2.0 M sulfuric acid were heated to 60 °C and treated with 2-iodobenzoic acid (**2.35**). The reaction was stirred for 2.5 hours, with the reaction visibly complete when orange bromine gas stopped being produced. The reaction mixture was then cooled to 2 °C in an ice-water bath, and the solution was filtered collecting the white solid. The solid was washed and importantly kept wet due to risk of explosion. This afforded compound **2.36**, hydroxy-1,2-benziodoxol-3(1H)-one -1-oxide. This solid was treated with glacial acetic acid and acetic anhydride under nitrogen, and slowly heated to 85 °C over one hour. The literature reports that heating should be increased gradually over 30 minutes, but during one of the attempts at synthesising DMP, the internal temperature when the hot plate reached 60 °C rose to over 110 °C and as a result the final product was destroyed. Once the internal temperature was at 85 °C and all the solid had dissolved, the stirring and heating was stopped. The solution was then left for 24 hours and colourless crystals formed. These crystals were collected by vacuum filtration under nitrogen affording compound **2.34**, 1,1,1-triacetoxy-1,1-dihydro-1,2-benziodoxol-3(1H)-one. Poor solubility meant data for this was hard to obtain; however, the product was obtained as white crystals that were used in the next step without purification with a yield of 40% compared with yield of 74% reported in the literature.[20]

The oxidation of compound **2.33** with DMP proceeded as expected (**Scheme 2.31**). Compound **2.33** was dissolved in CH_2Cl_2 treated with 1.3 equiv of DMP, and stirred at room temperature for 24 h. The solvent was removed under reduced pressure to

afford the crude material. The reaction worked well, with only some baseline impurities that were removed using a silica plug with CH_2Cl_2 as the solvent. This oxidation reaction gave a high yield of **92%.** Despite no published procedure for this method, the data strongly correlated with that obtained from the procedure of the Swern oxidation and showed an improved yield on that synthesis, which reported an 89% yield.[18] The product obtained was pure as shown by ^{1}H NMR. All three proton environments with the correct integration were observed. There is a significant difference in the NMR of the starting material, compound **2.33,** in comparison with product **2.12**, as a peak at 10.2 ppm is observed. This is characteristic of a proton adjacent to a carbonyl and shows the oxidation has been successful. ^{13}C NMR also shows the desired aldehyde has been made as there are five environments. This was supported by the IR, which when compared to the starting material has the disappearance of the broad peak at around 3248 cm^{-1} characteristic of an O-H and shows the appearance of sharp peak at 1699 cm^{-1} characteristic of a conjugated aldehyde.

2.3.3 Synthesis of benzaldehyde, compound 2.13

With **2.12** in hand, the alternative route to the DSA alkylating subunit could be explored (**Scheme 2.33**).



Scheme 2.33: Benzylation of nitro group to yield **2.13**

S*yn*-benzaldehyde oxime was dissolved in DMF and treated with K_2CO_3. This solution was heated to 90 °C before adding a solution of 3,5-dinitrobenzaldehyde in DMF. The literature procedure states it is stirred for 1 h and then cooled before continuing to the second part.[2] TLC analysis showed that there was still a significant amount of starting material present after 1 h, so the reaction was given longer to go to completion before proceeding. In doing this, unfavourable side reactions occurred and after the benzylation with benzyl bromide (BnBr) at 25 °C for 2 hours the desired product was not obtained. TLC showed that there were many impurities in the crude

mixture. These were isolated using flash chromatography in 50% CH_2Cl_2-hexane however, none of them were the desired product.

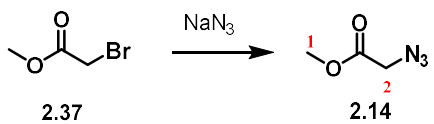
The second step of the synthesis was carried out despite not all the starting material, compound **2.12,** being consumed beforehand (as shown by TLC). Purification for this reaction proved challenging due to several impurities with similar polarities. After flash chromatography in 50% CH_2Cl_2-hexane, minor impurities could be removed with cold diethyl ether, leaving the product as a colourless solid. The product was clean and the data obtained agreed with the published procedure.[11] All eleven protons were accounted for, the five protons in the O-benzyl group (C9, 10 and 11) can be seen in the ^{1}H NMR in the aromatic region as well as the other three aromatic hydrogen environments (C2, 4 and 6) on the other benzene ring (**Scheme 2.33**). The NMR differs notably from the starting material as there is a singlet with an integration of two in the region ~5 ppm that is characteristic of addition of an O-benzyl group as the protons on the carbon adjacent to the oxygen (C8). Twelve carbon environments are seen, which support the structure formed, and the IR data shows the characteristic peaks of the conjugated aldehyde at 1700 cm^{-1} and the nitro group at 1529 cm^{-1}. A high yield of 75% was obtained; this is slightly lower than of the yield of 86% reported in the literature, however, it was still satisfactory.[2]

2.3.4 Synthesis of methyl 2-azido-3-(3-benzyloxy-5-nitrophenyl)acrylate compound 2.16

To carry out the following step, the starting material methyl 2-azidoacetate (compound **2.14**) had to be synthesised, as this was not commercially available.

2.3.4.1 Synthesising methyl 2-azidoacetate, 2.14

The method for this synthesis was established previously within the Searcey group, with significant changes to the published procedure (**Scheme 2.34**).[21]



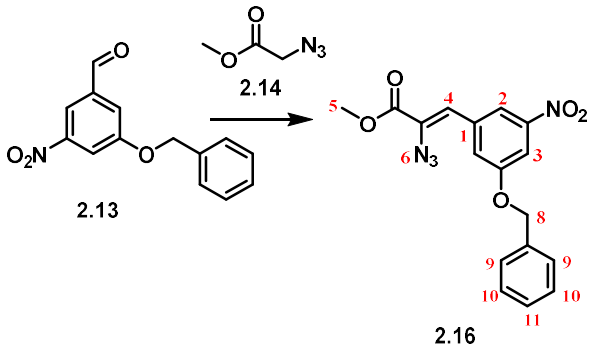
2.37 2.14

Scheme 2.34: Synthesis of methyl 2-azidoacetate, **2.14**

Methyl bromoacetate (**2.37**) and sodium azide were refluxed in acetone under nitrogen for 12 hours. The acetone was then removed under reduced pressure and the excess sodium azide was filtered and washed with diethyl ether. The diethyl ether containing product **2.14** was then removed under reduced pressure affording pure methyl 2-azidoacetate **2.14** as a pale yellow oil with a yield of 84%. Two singlets were seen around 3 ppm, one integrated for three protons and the other for two, confirming the structure. The ^{13}C NMR accounts for all three carbons and the IR further supports the correct synthesis with a characteristic azido peak at 2102 cm^{-1} and a carbonyl peak at 1743 cm^{-1}.

2.3.4.2 Synthesising the azidoacrylate compound 2.16

The preparation of the azidoacrylate, **2.16** was carried out following a previously reported procedure (**Scheme 2.35**).[2] Sodium methoxide was prepared beforehand by dissolving sodium metal in anhydrous methanol. This solution was then cooled to - 41 °C, differing from the temperature in the literature of -25 °C, using dry ice and acetonitrile. This was then treated with a solution of the aldehyde **2.13** and methyl azido acetate, compound **2.14** in THF (**Scheme 2.35**).

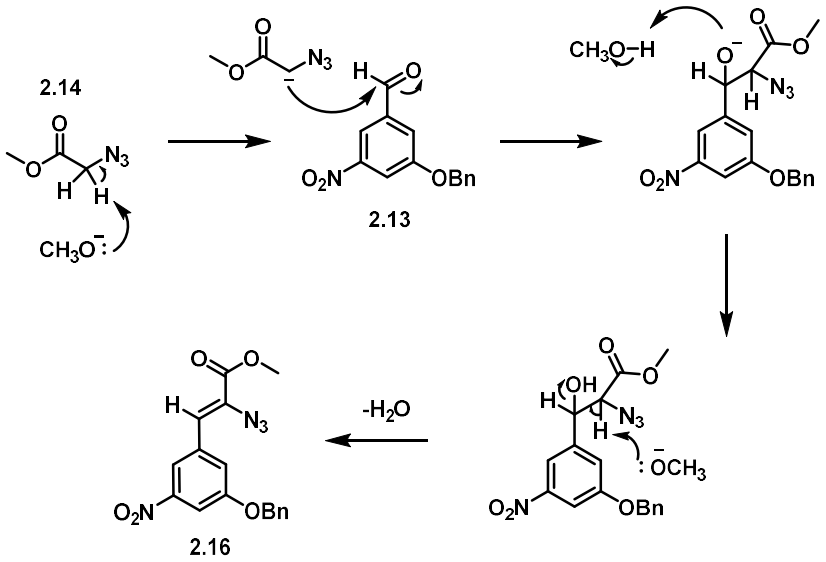


Scheme 2.35: Preparation of azidoacrylate, **2.16**

The reaction was kept at this temperature for 3 h and then at 0 °C for 14 h. The reaction mixture was diluted with CH_2Cl_2 and washed with water and brine. The organic layers were dried, and the solvent was removed under reduced pressure to obtain crude product, which was purified by flash chromatography to obtain yellow crystals with a yield of 73%. The quality was good as shown by lack of impurity peaks in the NMR. The ^{1}H NMR correctly showed fourteen hydrogens. A multiplet is seen in the aromatic region, which is a result of the aromatic protons on the *O*-benzyl group. This spectrum

differs from the starting material as there is now a singlet at 3.93 ppm that integrates for three protons. This is a result of the methyl group that has been added from the methyl 2-azido acetate as well as a singlet that integrates for one proton at 6.83 ppm responsible for the proton on the alkene. The ^{13}C NMR accounts for all fifteen environments, supporting the successful synthesis of the target. The IR differs from that of the starting material, compound **2.13,** as there is now the characteristic azido peak at 2132 cm^{-1} as well as the other functional groups in the compound.

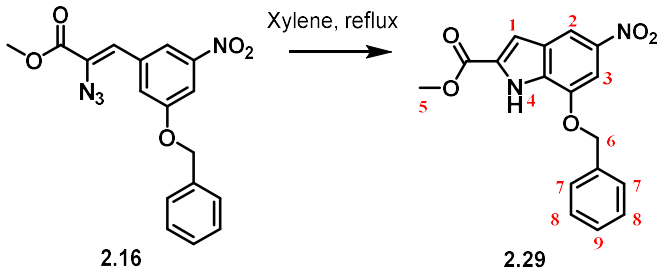
This reaction proceeds via a Knoevenagel condensation, shown in **Scheme 2.36**. The base sodium methoxide deprotonates methyl 2-azidoacetate, **2.14**. This proceeds via nucleophilic addition to the aldehyde. Protonation of the subsequent alkoxide is followed by elimination of water leaving the desired product **2.16**.



Scheme 2.36: *Mechanism of Knoevenagel condensation forming* ***2.16***

2.3.5 Cyclisation to form indole ring structure, compound 2.29

The cyclisation of **2.16** to **2.29** was attempted first in xylene, following the same procedure carried out by Boger and co-workers (**Scheme 2.37**).[2] The reaction was followed by TLC analysis. Although the reaction worked well when followed by ^{1}H NMR, problems were associated with the purification of the products.



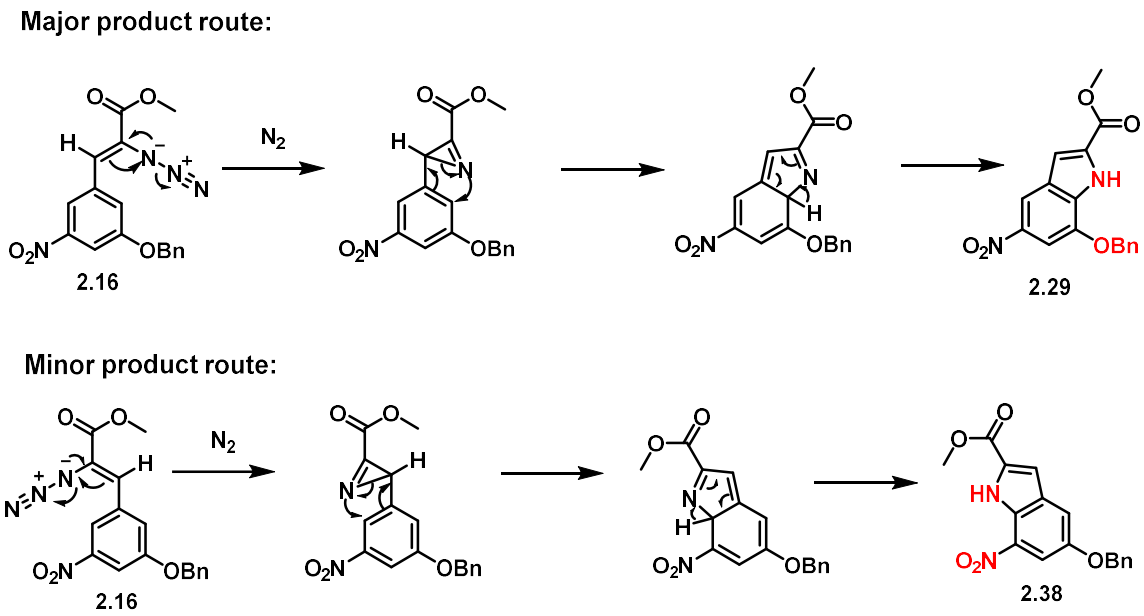
Scheme 2.37: *Synthesis of* ***2.29*** *via cyclisation in xylene.*

Two isomers are formed during cyclisation (**Scheme 2.38**) that have very similar polarities so are difficult to separate by flash chromatography. This reaction proceeds via a Hemetsberger-Knittel indole cyclisation. This product can cyclise in two ways and these give differing isomers with differing chemical properties. The major product is the desired compound **2.29,** which was determined as the major product due to the intensity in the TLC spot when visualised under UV light compared to the lower spot, which was due to minor product, compound **2.38.**

Different methods to separate these have been explored including flash chromatography with a gradient 0 to 15% EtOAc in hexane, gradient 0 to 10% EtOAc in hexane, and isocratic at 10% EtOAc in hexane, and all failed to separate. The method that had most success in getting some separation was the gradient to 10% of EtOAc but the conditions are not ideal. Other routes to purify were also explored such as recrystallisation in hexane and recrystallisation in hexane/EtOAc. Neither of these had success.

An alternative method was attempted using microwave conditions, which followed work carried out on similar compounds.[22] The starting material was heated in hexane at 200 °C for 15 min. It was carried out on a small scale of approximately 25 mg as a proof of concept. This worked extremely well with a yellow precipitate from the reaction, which was filtered and ^{1}H NMR showed that it preferentially synthesised only the desired isomer rather than both. However, this reaction was not scalable, as the microwave was unable to reach the temperature of 200 °C when using 100 mg. The resulting compound made was similar to reaction in xylene when the reflux was

interrupted after 2 h. This could mean that during the cyclisation there is an intermediate that can be isolated. The mechanism can be seen in **Scheme 2.38**.



Scheme 2.38: *Hemetsberger-Knittel indole cyclisation to afford compound* ***2.29.***

To try and mimic the reaction carried out under microwave conditions, finding a straight chained alkane with a high boiling point was investigated. Undecane has a boiling point of 196 °C so the starting azido compound, **2.16,** was refluxed at approximately 215 °C to see if this was selective in cyclising just the major isomer, **2.29**. Unfortunately, both isomers were produced, and lack of solubility meant that some starting material was still present.

An alternative approach was to see if in subsequent steps, it would be easier to separate the two isomers. As a consequence, crude compound **2.29** was used in the next step, similar to the reaction in the alternative route to Boc-protect the indole nitrogen (See section **2.2.4** and mechanism in **Scheme 2.13**.) However, this led to the two products becoming close in polarity, with only one product seen on TLC, but with two clear regioisomers in the ^{1}H NMR.

The products did not become easier to purify, so efforts went back into the original route but trying new separation techniques. For future synthesis, the xylene reflux will be followed as this works. The initial purification strategy using a slow gradient of EtOAc in hexane resulted in impurities still being present. Tweaking of the solvents

resulted in 5% acetone in hexane successfully purifying the crude product, obtaining a yield of 43% of **2.29.** This compound matched the material obtained from the previous route.

The original route could then be followed as there were no significant challenges in the subsequent steps, and achieved the final product, **2.25**, with an overall reaction yield of 3.9%.

2.4 Conclusion of Chapter 2

Both routes led to isolation of the final product, **2.25**. Each route had specific challenges, so it is important to have a comparison to decide which route to follow when scaling up the synthesis. **Table 2.1** summarises some of the simplified comparisons of the two routes.

Comparison	**Route 1**	**Route 2**
Synthesis	Yields were consistently higher when worked up and purified at each step.	Required more steps than the original route and the synthesis of two starting materials **2.14** and **2.34**
Cost	Negligible	Negligible
Overall yield	2.7%	3.9%
Purification	Issue with Negishi (**Scheme 2.9**)	Issue with cyclisation step (**Scheme 2.37**)
Overall scalability	Could scale up (equipment limiting)	Scaling up provided lower yields
Route for future synthesis	Continued	Did not use

Table 2.1: *Comparison of the two routes used to synthesise DSA alkylating subunit,* ***2.25***

The route followed initially had been perfected and scaled in an industry setting, so changes were needed to scale it down for use with laboratory facilities. A common approach was to run several smaller scale reactions parallel to one another for both routes, which meant reaction setup and workups took longer than carrying out one

larger scale reaction. The difference in cost between these routes in negligible, as the cost of the various raw materials for each route is similar overall.

The first route required more purifications and workups between steps than initially expected in order to achieve acceptable and sufficiently consistent yields. These additional reactions and purification steps added to the time taken to synthesise the final product. However, none of the starting materials needed to be synthesised as they were all commercially available, which saved time. In comparison, the second route required synthesis of methyl-2-azido acetate, **2.14.** A new route to synthesise this was discovered and provided a better yield of product (3.9%). However, the synthesis of the other starting material, Dess-martin periodinane, DMP, compound **2.34,** provided inconsistent yields and the synthesis was reportedly explosive, so all work and storage was conducted behind a blast shield.

Both routes had reaction bottlenecks. The first route resulted in poor yields at the Negishi step, synthesis of compound **2.28**. The second route also resulted in poor yields when scaling reactions up. For example, the synthesis of the azidoacrylate, compound **2.16,** carried out at -41 ºC, provided very poor yields when carried out on a scale larger than 2 g, and therefore it was necessary to run parallel experiments while keeping the reactions temperature controlled. Comparing the overall yields for each route is difficult due to the inconsistent yields in this step. An average calculation was done assuming a consistent yield of 25% for the Negishi cross-coupling reaction, which showed that the second route was on average 1.4 times more efficient than the first route.

Overall, the second route may provide a more consistent and slightly higher yield. However, the first route, established previously in the Searcey lab, was used to bring the DSA alkylating subunit forward due to its better performance at a larger scale, even despite bottlenecks in the synthesis and one particularly challenging purification.

References

1 D. L. Boger, C. W. Boyce, R. M. Garbaccio and J. A. Goldberg, CC-1065 and the duocarmycins: Synthetic studies, *Chem. Rev.*, 1997, **97**, 787–828.

2 M. S. Tichenor, J. D. Trzupek, D. B. Kastrinsky, F. Shiga, I. Hwang and D. L. Boger, Asymmetric Total Synthesis of (+)-and ent-(-)-Yatakemycin and Duocarmycin SA: Evaluation of Yatakemycin Key Partial Structures and Its Unnatural Enantiomer, *J AM CHEM SOC*, 2006, **128**, 15683–15696.

3 J. G. Felber and O. Thorn-Seshold, 40 Years of Duocarmycins: A Graphical Structure/Function Review of Their Chemical Evolution, from SAR to Prodrugs and ADCs, *JACS Au*, 2022, **2**, 2636–2644.

4 T. Imaizumi, Y. Yamashita, Y. Nakazawa, K. Okano, J. Sakata and H. Tokuyama, Total Synthesis of (+)-CC-1065 Utilizing Ring Expansion Reaction of Benzocyclobutenone Oxime Sulfonate, *Org. Lett.*, 2019, **21**, 6185–6189.

5 A. M. Beekman, M. M. D. Cominetti and M. Searcey, in *Cytotoxic Payloads for AntibodyDrug Conjugates*, 2019, pp. 185–191.

6 D. L. Boger and K. Machiya, Total synthesis of (+)-duocarmycin SA, *J. Am. Chem. Soc.*, 1992, **114**, 10056–10058.

7 M. A. Schmidt, E. M. Simmons, C. S. Wei, H. Park and M. D. Eastgate, An Enantioselective Total Synthesis of (+)-Duocarmycin SA, *J. Org. Chem.*, 2018, **83**, 3928–3940.

8 D. L. Boger, H. W. Schmitt, B. E. Fink and M. P. Hedrick, Parallel Synthesis and Evaluation of 132 (+)-1,2,9,9a-Tetrahydrocyclopropa[c]benz[e]indol-4-one (CBI) Analogues of CC-1065 and the Duocarmycins Defining the Contribution of the DNA-Binding Domain, *J. Org. Chem.*, 2001, **66**, 6654–6661.

9 V. F. Patel, S. L. Andis, J. K. Enkema, D. A. Johnson, J. H. Kennedy, F. Mohamadi, R. M. Schultz, D. J. Soose and M. M. Spees, Total Synthesis of Seco (+)- and ent-(−)-Oxaduocarmycin SA: Construction of the (Chloromethyl)indoline Alkylating Subunit by a Novel Intramolecular Aryl Radical Cyclization onto a Vinyl Chloride, *J. Org. Chem.*, 1997, **62**, 8868–8874.

10 D. L. Boger, C. W. Boyce, R. M. Garbaccio and M. Searcey, Synthesis of CC-1065 and duocarmycin analogs via intramolecular aryl radical cyclization of a tethered vinyl chloride, *Tetrahedron Lett.*, 1998, **39**, 2227–2230.

11 M. J. Stephenson, L. A. Howell, M. A. O, K. R. Fox, C. Adcock, J. Kingston, H. Sheldrake, K. Pors, S. P. Collingwood and M. Searcey, Solid-Phase Synthesis of Duocarmycin Analogues and the Effect of C-Terminal Substitution on Biological Activity,

12 D. Astruc, The 2010 Chemistry Nobel Prize to R.F. Heck, E. Negishi, and A. Suzuki for palladium-catalyzed cross-coupling reactions, *Anal. Bioanal. Chem.*, 2011, **399**, 1811–1814.

13 K. Hiroya, S. Matsumoto and T. Sakamoto, New Synthetic Method for Indole-2-carboxylate and Its Application to the Total Synthesis of Duocarmycin SA, *Org. Lett.*, 2004, **6**, 2953–2956.

14 A. Yasuhara, Y. Kanamori, M. Kaneko, A. Numata, Y. Kondo and T. Sakamoto, Convenient synthesis of 2-substituted indoles from 2-ethynylanilines with tetrabutylammonium fluoride, *J. Chem. Soc. Perkin 1*, 1999, 529–534.

15 K. Hiroya, R. Jouka, M. Kameda, A. Yasuhara and T. Sakamoto, Cyclization reactions of 2-alkynylbenzyl alcohol and 2-alkynylbenzylamine derivatives promoted by tetrabutylammonium fluoride, *Tetrahedron*, 2001, **57**, 9697–9710.

16 D. Barker, A. L. Lehmann, A. Mai, G. S. Khan and E. Ng, Synthesis of non-symmetrical 3,5-diamidobenzyl amines, ethers and sulfides, *Tetrahedron Lett.*, 2008, **49**, 1660–1664.

17 J. Keillor and K. Caron, Can. Pat. Appl., 2746891, 19 Jan 2012.

18 WO 00/66544, .

19 H. Muratake, I. Abe and M. Natsume, Total synthesis of an antitumor antibiotic, (±)-duocarmycin SA, *Tetrahedron Lett.*, 1994, **35**, 2573–2576.

20 R. K. Boeckman, P. Shao and J. Mullins, The Dess-Martin periodinane: 1,1,1-triacetoxy-1,1-dihydro-1,2-benziodoxol-3(1h)-one, *Org. Synth.*, 2004, **10**, 696.

21 Y. Tian, Z. Liu, J. Liu, B. Huang, D. Kang, H. Zhang, E. De Clercq, D. Daelemans, C. Pannecouque, K. H. Lee, C. H. Chen, P. Zhan and X. Liu, Targeting the entrance channel of NNIBP: Discovery of diarylnicotinamide 1,4-disubstituted 1,2,3-triazoles as novel HIV-1 NNRTIs with high potency against wild-type and E138K mutant virus, *Eur. J. Med. Chem.*, 2018, **151**, 339–350.

22 F. Lehmann, M. Holm and S. Laufer, Rapid and easy access to indoles via microwave-assisted Hemetsberger–Knittel synthesis, *Tetrahedron Lett.*, 2009, **50**, 1708–1709.

Chapter 3 – Synthesis of Duocarmycin Dimers

Chapter 3 – Synthesis of duocarmycin dimers

The aims of this chapter are to introduce the concept of dimers, their utility in cancer treatment, and to highlight the importance of finding novel high-potency analogues to treat this group of diseases. The introduction discusses dimeric payloads for cancer treatments that are under development, followed by a discussion on the synthesis of dimeric payloads using the alkylating subunit synthesised in **Chapter 2**. This chapter then highlights the importance of stereochemistry in dimer synthesis and discusses methods to separate racemic mixtures, with a focus on chiral column chromatography. Lastly, this chapter briefly discusses potential methods of dimer synthesis via solid phase techniques and highlights the superiority of solution phase approaches in the synthesis of these compounds. The concept of crosslinking using dimers and dual warheads will be discussed in greater detail in **Chapter 4**.

3.1 Introduction to dimers

Dimers are made up of two compounds bound together. They are used as anti-cancer agents, both in their own right, and by binding them to different targeting groups, such as proteins or antibody drug conjugates (ADCs).[1] The basic principles are the same for both categories: when using a dimeric compound, it is possible to bind to two distinct binding sites on a specific target, which may increase not only potency but also selectivity.[2] This brief introduction to dimers focuses on dimeric compounds used as payloads in the application of ADCs. For a more in-depth review on duocarmycins in ADCs with focus on dimeric payloads, see Chapters 9 and 10 in *Cytotoxic Payloads for Antibody-Drug Conjugates*.[3]

3.1.1 Cyclopropapyrroloindole (CPI) Dimers

Cyclopropapyrroloindole (CPI) units are used as the alkylating subunit of the duocarmycin family member CC-1065 (**Chapter 1, Figures 1.20** and **1.21**). Despite CC-1065 having delayed fatal toxicity, analogues of this compound incorporating the alkylating subunit have been studied for the application as anticancer agents with some success.[4,5] There is a considerable interest currently in the development of DNA sequence specific or selective agents for genetic targeting for the control of gene expression, for application in diagnosis or ultimately in therapy. In this context CC-

1065 is one of the most impressive lead compounds isolated in trace quantities from the culture of Streptomyces zelensis at Upjohn in 1978. The unique structure was confirmed by single molecule X-ray structure determination in 1981. However CC-1065 cannot be used in humans because it was found that it caused delayed deaths in experimental animals. In the search for compounds with better antitumor selectivity and DNA sequence specificity many CC-1065 analogues have been synthesized in an attempt to avoid the undesired side effects while retaining its potency against tumor cells. Two successful attempts in the modification in the active moiety of the parent natural product 1,2,8,8a-tetrahydro-7-methylcyclopropa[3,2-e]indole-4-one (CPI) and 1,2,9,9a-tetrahydrocyclopropa[c]benz[e]indole-4-one (CBI) have been made. We review here recent progress with the analogs of CPI and CBI and their conjugates both by solution and solid phase, also the progress and development of CPI and CBI conjugates with polyamides (information reading molecules in the minor groove of DNA). Since CPI-CPI dimers are significantly more potent than CC-1065 in vitro and in vivo, a large number of CBI-CBI dimers with varying linkers lengths and positions have been reviewed.[4] Thurston and Jackson developed a series of CPI dimers (**Figure 3.1**), and reported that dimers with three, five and eight methylenes between the subunits, highlighted in red, showed increased potency over the parent compound. Compounds with three methylenes (**3.1**) had the highest potency against L1210 leukaemia cells (IC_{50} = 0.004 nM vs 0.043 nM of CC-1065).[3]

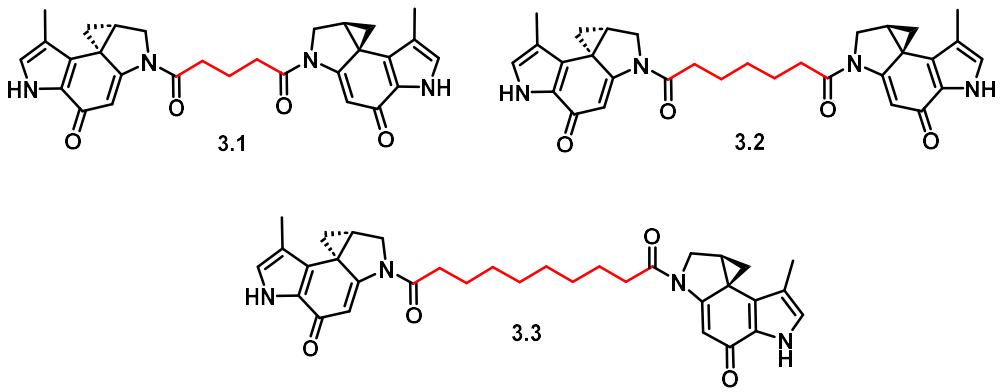
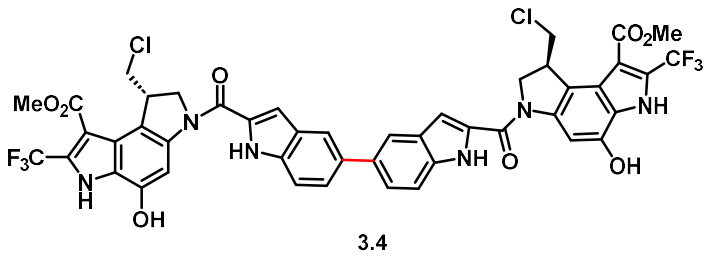


Figure 3.1: *CPI bis-alkylators with varying flexible (poly)methylene linkers highlighted in red.*

A range of CPI dimers were also developed to perform structure-activity relationships. The most potent dimer had a potency of 0.005 nM against HeLaS3 (human uterine cervix carcinoma cells) **Figure 3.2**.[3] The heightened potency is owed to the single carbon-carbon bond between the two indole rings (**Figure 3.2**, red). The other dimers

in the series with lower activity had alternative linker groups of varying length. An important conclusion was that for cytotoxicity in this family of compounds, the length of the linker had a greater effect than the type of linker.[6]



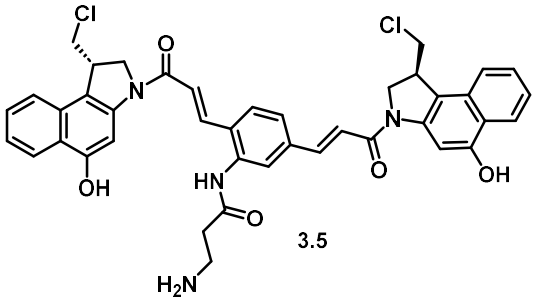
3.4

Figure 3.2: *CPI dimer by Fukuda, with a C-C direct link (red) showing highest potency.*

Since then, research into these compounds has focussed on coupling them to an antibody for targeting. These CPI dimers as ADCs include a next-generation dimeric CPI-based ADC used to treat myeloid leukemia (AML) by binding to CD-33 and CD-123.[7]

3.1.2 Cyclobenzindole (CBI) dimers

Duocarmycin based *seco*-CBI dimers have been shown to be highly potent *in vitro*.[8] This potency has made *seco*-CBI dimers of great interest as payloads for ADCs (refer to **Chapter 1, Figure 1.21** for CBI structure**)**. Su *et al.* published the first example of ADCs bearing dimeric *seco*-CBI structures. **Figure 3.3** provides an example of these dimers.[8]



3.5

Figure 3.3: *seco-CBI dimer payload.*

When designing dimers, it is important to note that position and length of linker affects the potency of the compound. Lown *et al*. investigated this structure-activity relationship in *seco*-CBI dimers.[9] It was found that for the compounds bound carbon to carbon C7-C7 (**3.6**), the longer the chain, the less potent the compound (**Figure 3.4**).[9] Conversely, for the C7-N3 dimers, the longest linker (a chain of six carbons, **3.7**) proved to be the most potent.[9] This shows that both the linkage position within the monomer and length of the linker have a significant effect on the potency of the dimers.

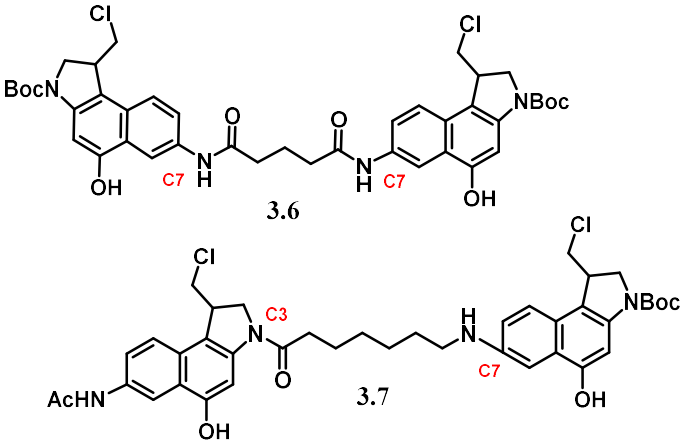


Figure 3.4: *CBI units, with varying linker position and linker length, position highlighted in red.*

3.1.3 CBI-PBD dimers

Compounds with dual 'warheads' that show differences in binding are also being investigated as chemotherapeutic treatments.[10,11] This is particularly true in the case of hybrid payloads in next generation ADCs. The pyrrolobenzodiazepines (PBDs) are tricyclic natural products that bind to the minor groove in DNA and selectively alkylate guanine bases.[12] PBDs have been developed as homodimers. For example the ADC SG3249 consists of a PBD dimer (SG3199, compound **3.8**) linked to a maleimidocaproyl valine-alanine dipeptide linker; **Figure 3.5** shows the payload structure).[13]

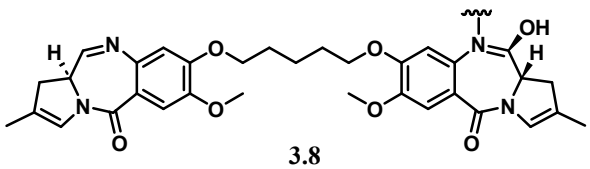
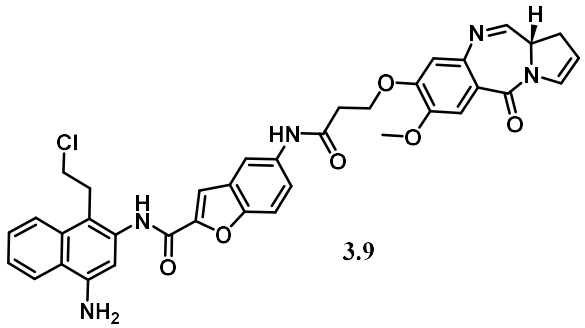


Figure 3.5: *The PBD dimeric payload in ADC SG3249*

Thurston *et al.* also synthesised hybrids of two distinct DNA alkylating agents, CBI-PBD and *seco*-CBI, to investigate how they interact with DNA.[13] These duocarmycin based dimers are having success as payloads for ADCs in the treatment of cancer, as CBI dimers and CBI-PBD hybrids have both progressed into clinical trials.[14]



3.9

***Figure 3.6**: structure of seco-amino-CBI-PBD*

Purnell *et al.* synthesised several bisalkylating units, including compound **3.9** shown in **Figure 3.6**.[15] Compound **3.9** was shown to be active against several human cancer cell lines (IC_{50} of 0.026 μM in P815 cells), in addition to being able to form interstrand crosslinks. This crosslinking is thought to confer an increased potency, as compound 3.9 shows a greater potency over both parent monomers.[15] The CBI section of the structure alkylates at the N3 position of adenine and the PBD section alkylates at the N2 position of guanine, allowing these dimers to crosslink DNA. Tercel *et al.* described the synthesis of ten compounds, containing one *seco*-CBI payload and one PBD payload.[16] Their work investigated the linker length of these dimers, and assessed sequence selectivity using thermal cleavage assays. Their work showed that the CBI unit determined the alkylation site. Both its enantiomeric form and the linker length influenced cytotoxicity greatly with over a 35,000-fold difference in cytotoxicity in different cell lines.[16]

3.1.4 DSA dimers

As next generation ADCs allow the use of more potent 'warheads', compounds with improved activity are needed. As described in **Chapter 1**, the DSA alkylating subunit (**2.25**) is the most potent in this family, so this is the best target among the duocarmycins for design of a more effective 'warhead'.

In this thesis, for simplicity, the terms 'natural' and 'unnatural' will be used to describe the enantiomers of the DSA. 'Natural' will be used to describe the enantiomer found in nature (compound **2.25 (a)** in **Figure 3.7**, and 'unnatural' will describe the enantiomer that has only been synthesised in the laboratory (compound **2.25 (b)** in **Figure 3.7**.

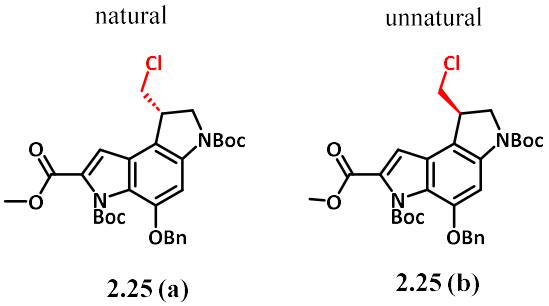


Figure 3.7: *The different enantiomers that the racemic starting material,* ***2.25*** *can be separated into, highlighting in red the difference between the 'natural' and the 'unnatural' isomers*

Boger *et al*. reported the synthesis of all four diastereomers available from using both enantiomers of the DSA alkylating unit, attached directly with no linker, as shown in **Figure 3.8**.[17]

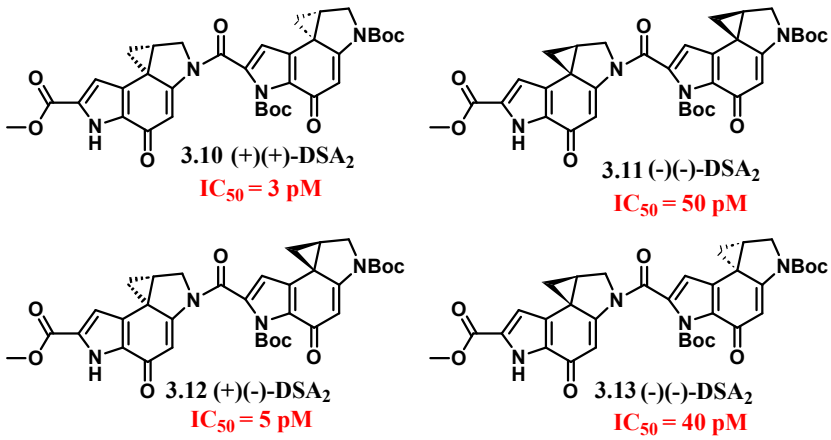


Figure 3.8: *The four possible diastereomers of DSA alkylating unit as dimers synthesised by Boger, highlighting their relative* IC_{50} *in red*

All four diastereomers showed picomolar IC_{50} values in L1210 cell lines, and their relative potencies are highlighted in **Figure 3.8** in red. The left-hand subunit controlled where the first site of DNA alkylation occurred. The units where the natural enantiomer (**2.25 (a)**) was first were approximately ten times more potent than those

where the unnatural enantiomer was first. They were also more potent than the parent compound duocarmycin SA, which has a potency in the same cell line of 10 pM.[17]

The ability of these compounds to crosslink DNA was also investigated. When the cyclopropanes were oriented in opposite directions, i.e., (+)(-)-DSA_2 and (-)(+)-DSA_2, interstrand crosslinking of DNA was observed.[17] Units with the same enantiomers, i.e., (+)(+)-DSA_2 and (-)(-)-DSA_2, were only able to form intrastrand DNA cross links. Their ability to alkylate DNA in different ways, depending on their orientation, may be crucial to their biological activity and this project will explore this idea further.

One aim of this project was to synthesise duocarmycin dimers, similar to those in **Figure 3.8**, but in which the alkylating subunits are separated by varying linkers. Finding appropriate ways to synthesise dimers and other higher order structures is of significance when hoping to later explore their structure activity relationships (SAR).

3.2 Synthesis – trail dimer studies with racemic material

The synthesis of a dimer involved bringing together two subunits of the DSA derivative, **compound 2.25 in Chapter 2,** synthesised previously. To address solubility issues, and to investigate adding a linker, β-alanine (**3.14**) and polyethylene glycol (PEG) linkers (**3.16**) were chosen (**Figure 3.9)**. β-alanine, being unreactive, was expected to maintain compound potency, while PEG-based linker groups aimed to improve solubility. This approach allowed investigation of DNA binding and crosslinking phenomena, with potential applications in SARs by varying linker lengths.

A control monomer, compound **3.15,** was synthesized for later activity comparison with the dimers. **Figure 3.9** outlines the proposed structures of the three compounds (**3.14, 3.15, 3.16**) targeted for synthesis in this project.

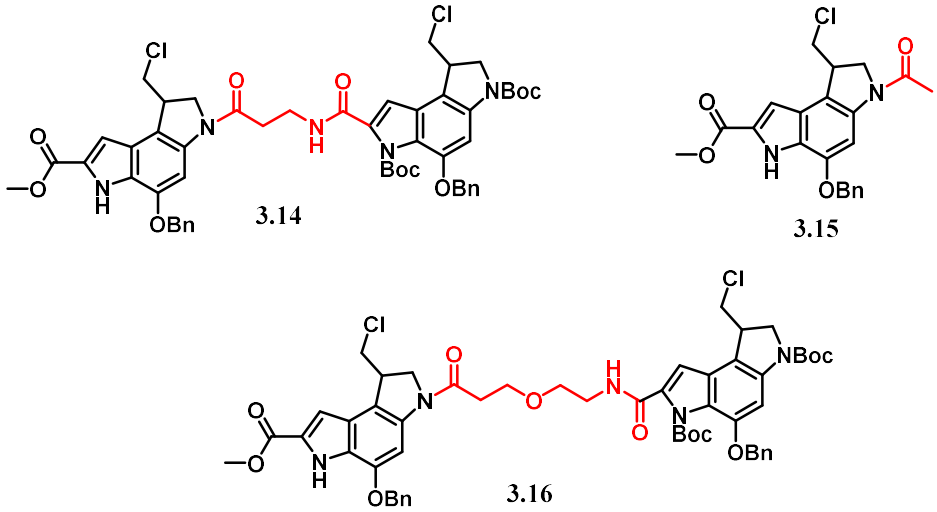
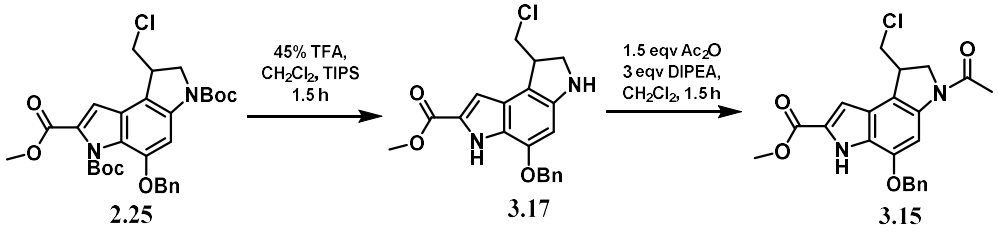


Figure 3.9: *The two proposed racemic dimers (**3.14, 3.16**) and the control monomer (**3.15**)*

3.2.1. Synthesis of the monomer DSA-Ac (3.15)

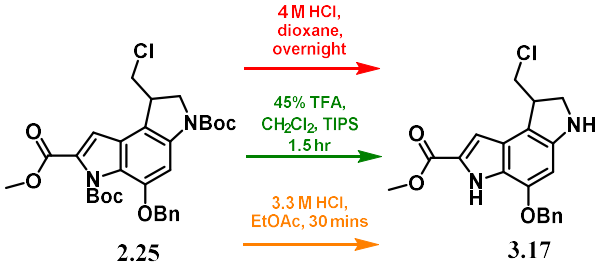
To carry out subsequent couplings, the free amine needed to be present to form both the left half of the dimers and the control monomer. The first step in the synthesis of the monomer, and of all the dimers, was to find appropriate conditions to remove the Boc protection group, compounds **2.25** to **3.17** (**Scheme 3.1**).



Scheme 3.1: *Reaction scheme for the synthesis of the acetylated monomer **3.15.***

There were several challenges during the Boc deprotection procedure. The different methods are summarised in **Scheme 3.2**. Initially, a literature-recommended protocol employing 4M hydrochloric acid in dioxane overnight was pursued.[18] Despite several attempts, including modifications such as reduced reaction times and monitoring closely by TLC and analytical high-performance liquid chromatography (HPLC), it was evident that the reactions were overly harsh, resulting in degradation. Degradation

was observed as a black baseline impurity on TLC and the analytical HPLC analysis revealed a complex mixture of peaks.



Scheme 3.2: *An overview of the different strategies to remove the Boc protecting group.*

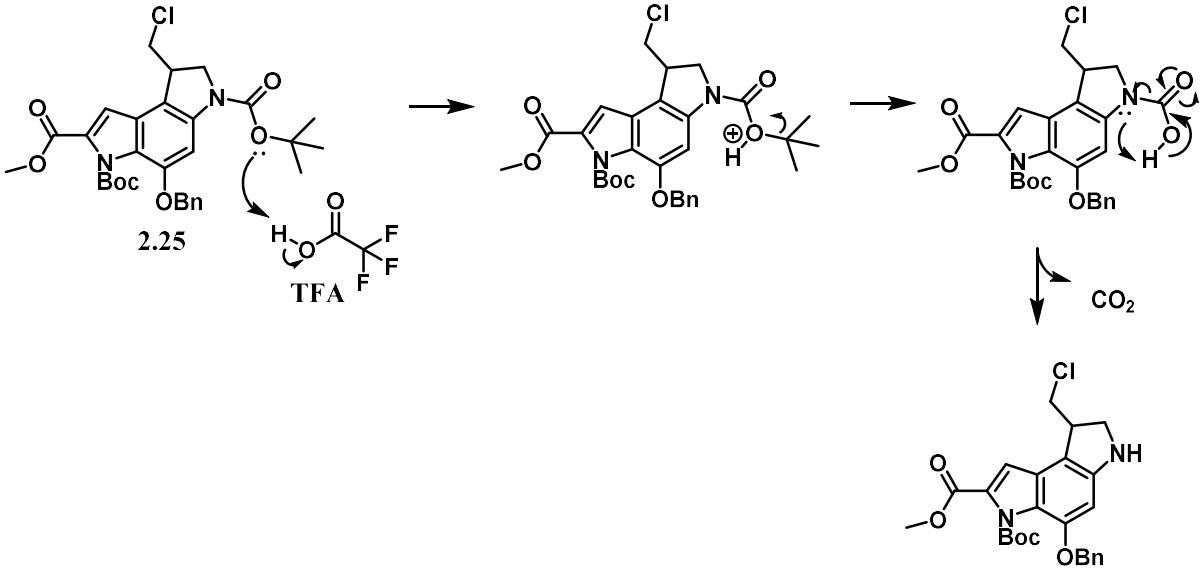
In response to the observed issues, alternative reaction conditions were investigated. A less concentrated solution of 3.3 M HCl in ethyl acetate was used. TLC analysis indicated a more promising result; however, some degradation was noted as the reaction progressed. Subsequently, the reaction was repeated and terminated after a shorter duration of 30 minutes. Analysis by NMR and liquid chromatography–mass spectrometry (LC-MS) confirmed the presence of the desired product. The NMR spectrum showed the disappearance of strong singlets, corresponding to the Boc protecting group at the upfield section, 1.3-1.6 ppm. The LC-MS, with an expected mass of 370.11 and an observed m/z ratio of 371.1, corresponded to the $[M+H]^+$. The presence of the free amine group posed challenges during purification, as poor separation was observed, so an alternative route that avoided purification was investigated.

A commonly used alternative approach for the deprotection of Boc groups uses trifluoroacetic acid (TFA). In fluorenylmethoxycarbonyl (Fmoc) peptide synthesis, TFA combined with triisopropylsilane (TIPS) in CH_2Cl_2 has been used to cleave compounds from the resin. In the solid phase synthesis of duocarmycin analogues, it was shown that when handling the duocarmycins, a lower concentration of TFA (47.5%) is required to cleave when compared with normal peptide synthesis, which routinely uses a concentration of 85% TFA.[18] For this reason, a concentration of 45% TFA in 52.5% CH_2Cl_2 and 2.5 % TIPS was used. TIPS is crucial in solid phase peptide synthesis (SPPS), as it acts as a scavenger for unreacted carbocations and prevents them from producing side-reactions with the peptide.[19] It was reasonable to maintain

the use of a scavenger, due to the generation of carbocations, and it can easily be removed due to its volatility.

The reaction progress was monitored, again, through analytical HPLC and TLC. The reaction was terminated after 1.5 hours, revealing a seemingly complete conversion of starting material to product. The reaction was worked up by removing all solvent, including TFA, *in vacuo*, and the product was triturated with cold diethyl ether. The final product was confirmed through NMR analysis of the crude product. This showed the expected peaks and, importantly, showed that the two characteristic Boc group singlets, both integrating for nine protons, were no longer present when compared with the NMR of the starting material. The purity was also confirmed with HPLC, where a single peak was observed. The accurate mass was validated by LCMS, showing the same m/z as before, corresponding to the $[M+H]^+$.

The mechanism for the Boc deprotection is shown in **Scheme 3.3** and is not selective for the Boc shown, with both Boc groups being removed; however, for simplicity, the scheme shows only the removal of the Boc group of interest.

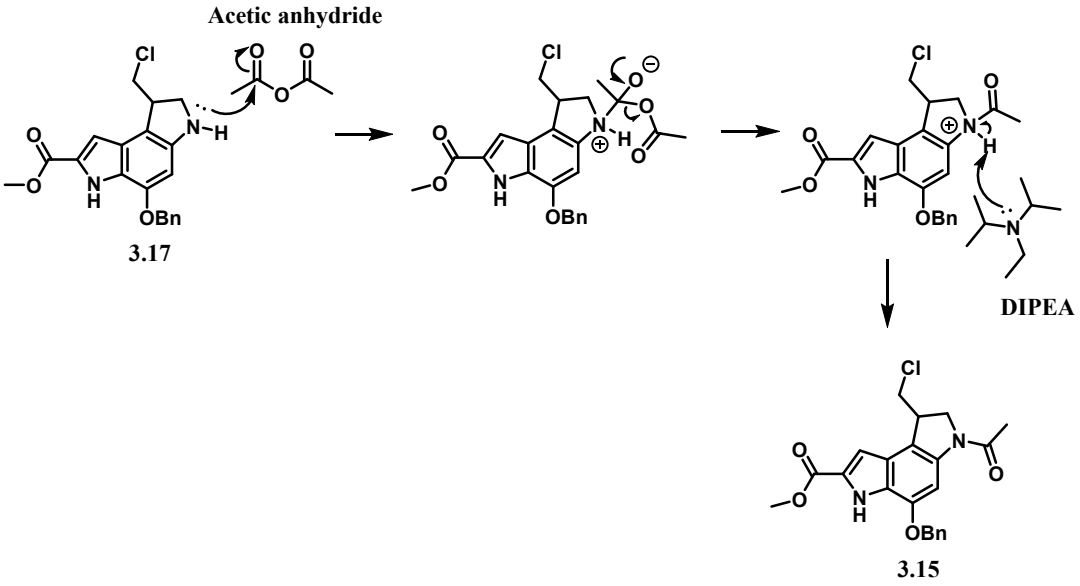


Scheme 3.3: *Mechanism for Boc deprotection using TFA.*

This product was used immediately in the acylation step without further purification. The Boc deprotected DSA was dissolved in CH_2Cl_2 and treated with 1.5 equiv. of acetic anhydride and 3 equiv. of *N,N*-Diisopropylethylamine (DIPEA). The reaction was complete after 1.5 h, confirmed by TLC and analytical HPLC. As the reaction

progressed, solid product was precipitating out of the CH_2Cl_2. The solvent was removed under reduced pressure and the product was subjected to flash column chromatography. Different eluents were assessed for purification by flash chromatography, the most successful being isocratic conditions of 50:50 EtOAc in hexane. This produced a white amorphous solid with a yield of 58% over the two steps.

The mechanism of the acylation starts with nucleophilic attack of the lone pair of electrons on the nitrogen of **3.17** towards the δ^+ carbon on the carbonyl of the acetic anhydride (**Scheme 3.4**). This nucleophilic addition is followed by an elimination to form the protonated product. The DIPEA is used as a base in this reaction to deprotonate the positively charged amine, to give the final acetylated product, compound **3.15**.



Scheme 3.4: *Mechanism of acetylation to form compound* ***3.15***

The success of this reaction was confirmed using ^{1}H NMR, ^{13}C NMR, LC-MS, and HRMS. ^{1}H NMR showed the disappearance of both Boc groups in comparison to the starting material. Additionally, a new single peak integrating for three protons was observed at 2.16 ppm, evidencing the methyl group from the acetylation. The LC-MS, with an expected mass of 412.12 and an observed m/z ratio of 413.1, corresponded to the $[M+H]^+$. This was also confirmed on HRMS, which gave an *m/z* ratio of 413.1288. This data, alongside a single observed peak on the analytical HPLC, further confirmed that the correct product had been synthesised. The purified monomer was then stored,

ready to be deprotected and brought forward for biological testing later in the project (**Chapter 4**).

3.2.2. The importance of stereochemistry in dimer synthesis

The synthesis of the racemic dimers is a model system for the final synthesis of the stereoselective compounds. Use of a model was important to optimise the coupling conditions before using the precious enantiopure starting material. Understanding stereochemistry, i.e., the 3D arrangement of chiral atoms within a larger compound, is hugely important. As the DNA alkylation units of duocarmycin SA (DSA) in the *seco* form contain a stereocentre, generating dimers generates a set of isomers. The four compounds fall into two pairs of enantiomers and each compound will also have two diastereomers (**Figure 3.10**).

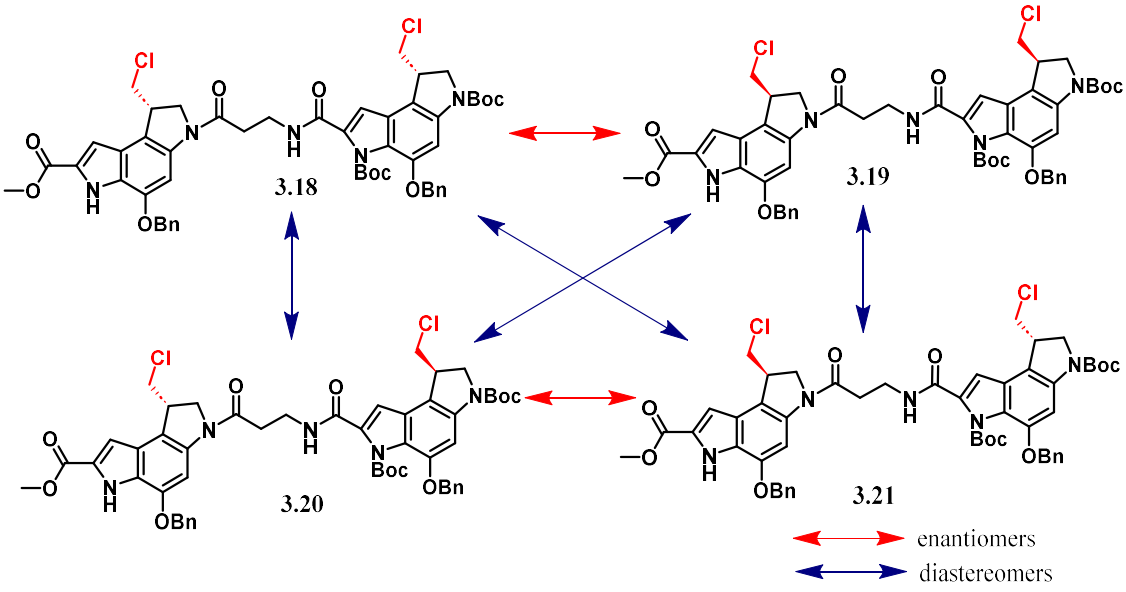


Figure 3.10: *The DSA- β-alanine -DSA dimer with the four possible arrangements*

The chiral enantiomers are non-superimposable mirror images of each other. They have the same chemical formula but a different arrangement of atoms around a chiral centre.[20] They possess the same physical and chemical properties, but can have different activity when acting on biological systems.[20] Diastereomers have the same chemical formula and the same connectivity. However, they have a different spatial arrangement of atoms and are not a non-superimposable mirror image of each other.

Consequently, their chemical and physical properties will vary, albeit sometimes slightly. They are also likely to have different biological activities.[21]

Figure 3.11 shows the four possible arrangements of the DSA-PEG dimer. When synthesising these compounds, it is expected that there may be some difficulties in purification, and likely two very similar but still different environments observed as a result of the two pairs of enantiomers formed from using racemic starting materials.

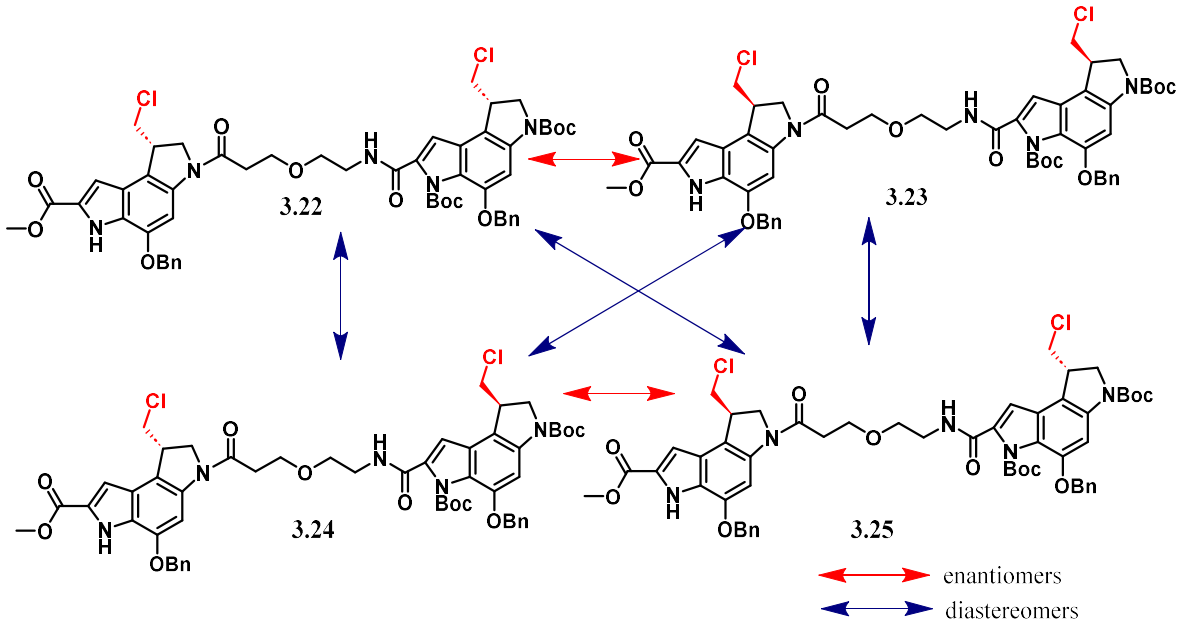


Figure 3.11: *The DSA-Peg 1- DSA dimer with four possible arrangements*

3.2.3. Synthesis of the left-hand side of DSA- β-alanine -DSA dimer

Duocarmycin alkylation subunits possess a structure that can be manipulated, due to the presence of carboxylic acid, (compound **2.29** in red) and terminal amine groups, (compound **3.17** in red) that can be coupled through an amide bond and bind directly to each other or other units (**Figure 3.12)**. There were several steps involved in synthesising the dimer with a β-alanine linker. The dimer has a left-hand and a right-hand unit and both need to be prepared prior to the final couplings. The Boc group had to be removed for the left-hand unit, and the ester needed to be hydrolysed for the right-hand unit. The individual units making up the final dimer are summarised in **Figure 3.12**, with the functional groups needed for the couplings highlighted in red.

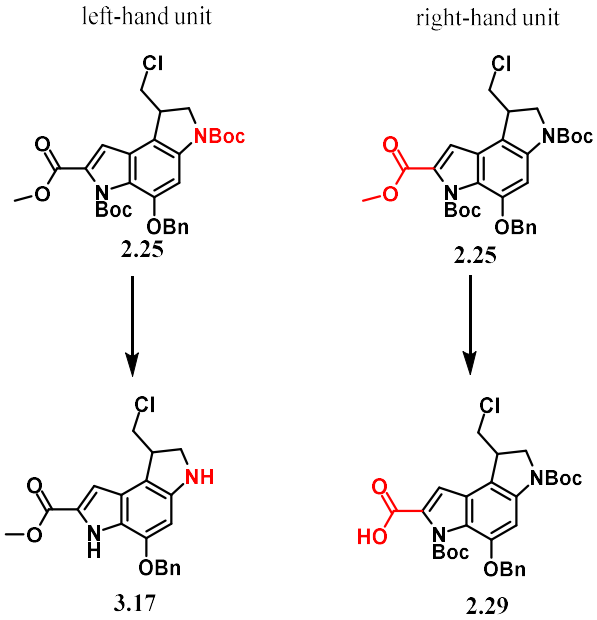
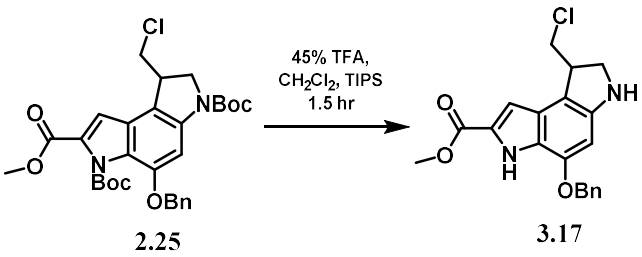


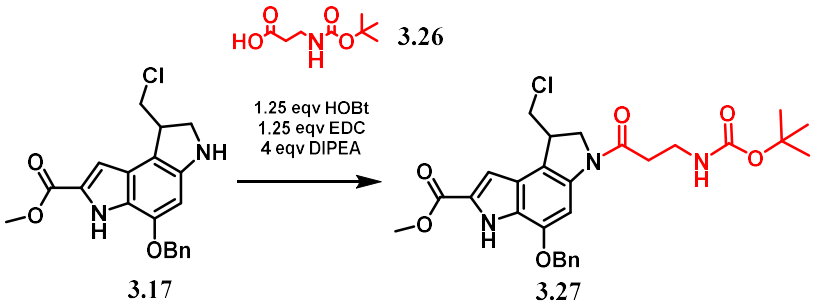
Figure 3.12: *Necessary starting materials to synthesise DSA based dimers.*

The first step was the Boc deprotection, which used the same conditions as when making the monomer unit (**Scheme 3.5**).



Scheme 3.5: *Conditions for Boc deprotection of* ***2.25*** *to give* ***3.17.***

For the first series of dimers, compound **3.17** was then reacted with Boc-protected β-alanine, **3.26** (**Scheme 3.6**).



Scheme 3.6: *Conditions for coupling of β-alanine linker,* ***3.26*** *to* ***3.17*** *to give* ***3.27.***

Several coupling methods were trialled to generate this product. The initial conditions were in DMF with 1 equiv. of **3.17**, 1.1 equiv. of Boc-β-alanine, 1.2 equiv. 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide **(**EDC), 1.2 equiv. hydroxybenzotriazole (HOBt) and 3 equiv. of *N*-methyl morpholine. However, these conditions failed to drive the reaction to completion, even after extended reaction times and led to a complicated work-up, due to the insolubility of the final product. The use of extended reaction times also generated unidentified side-products that further complicated purification. Finally, substitution of the base with 4 equiv. DIPEA, shortening the reaction time, 1.5 to 3.5 h, and immediate purification gave the required product, compound **3.27**.

Difficulties in purification came from the insolubility of the target compound in many solvents. Initially the reactions were extracted using CH_2Cl_2 or EtOAc and the organic layer was washed several times with water to remove any DMF. This, however, yielded very little crude product, and after purification there was too much product lost to determine if the reactions had been successful. The best method to work-up was to add water to the reaction mixture, precipitating the product, which was then collected following centrifugation. The product was triturated with hexane and collected. Purification attempts included normal phase flash chromatography with varying conditions, reverse-phase flash chromatography and preparative HPLC in acetonitrile, which gave a clean product. However, due to being able to inject only 10 microlitres every 30 minutes, this was inefficient for the purposes of this project. For this reason, the optimised conditions were two consecutive normal phase columns. First, to remove other impurities, was a 50:50 EtOAc:hexane isocratic column, followed by a 100% EtOAc flush to collect the product. The product was then concentrated and subjected to another column, this time, of a gradient from 0 to 3% MeOH in CH_2Cl_2, affording pure material in a 42% yield.

The last difficulty was product insolubility, which meant that yields were low, and data was hard to obtain. This was unfortunately unavoidable. NMRs were best carried out in acetone or methanol. Obtaining ^{13}C NMR was not possible due to product precipitating out of the solvent over the run time. The product was checked on analytical HPLC to show one clean peak (**Figure 3.13**), and the correct mass was

identified using LCMS. The LC-MS, with an expected mass of 541.20 and an observed m/z ratio of 542.2, corresponded to the $[M+H]^+$.

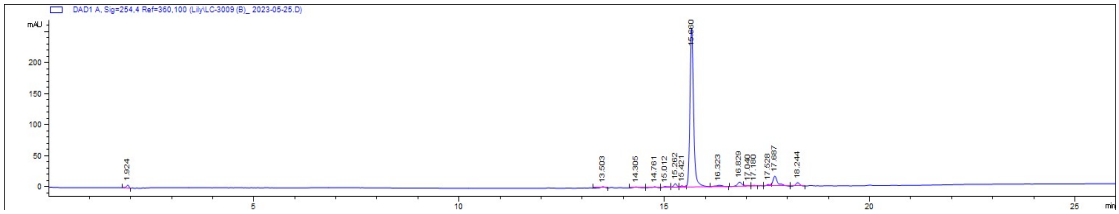
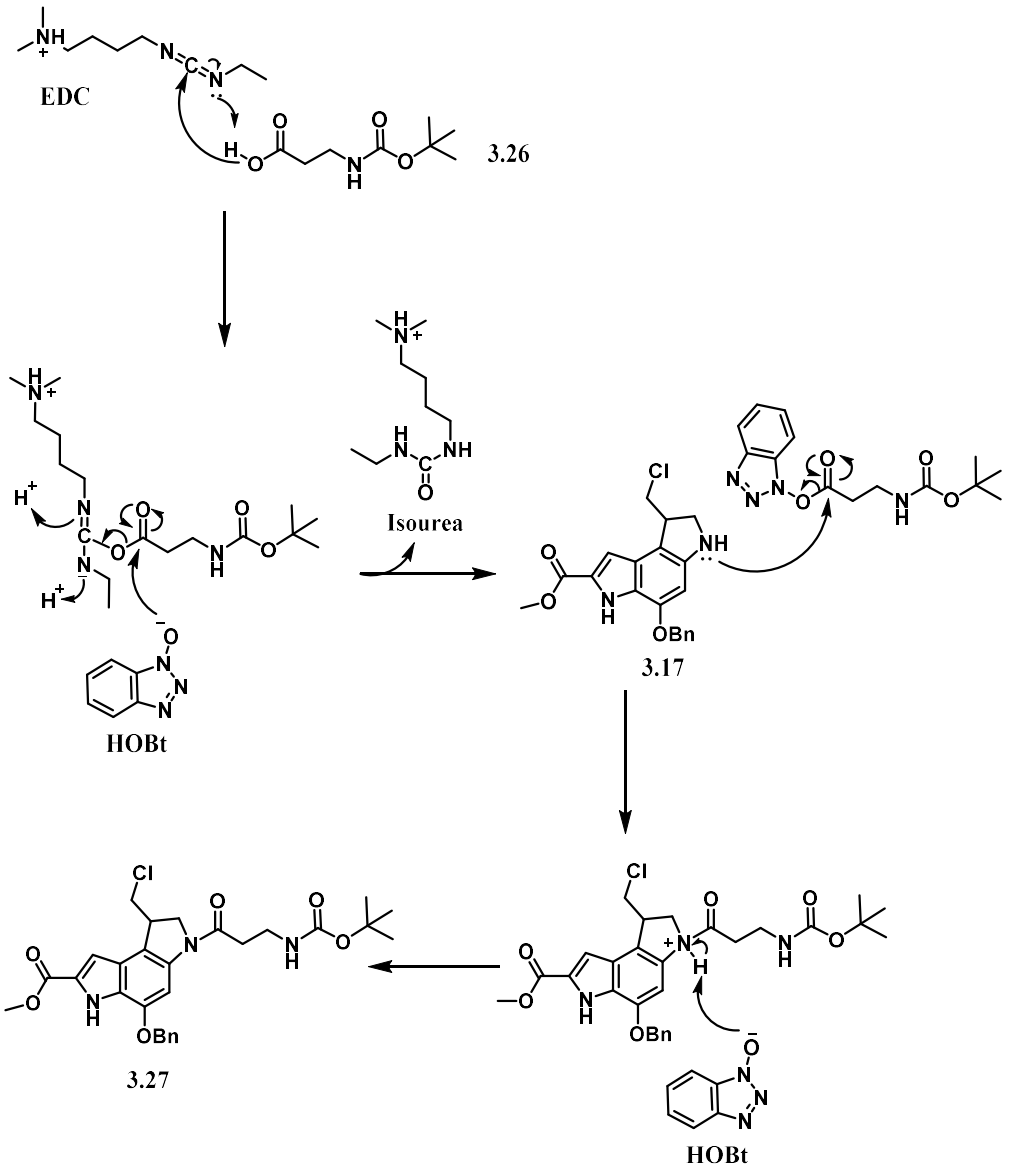


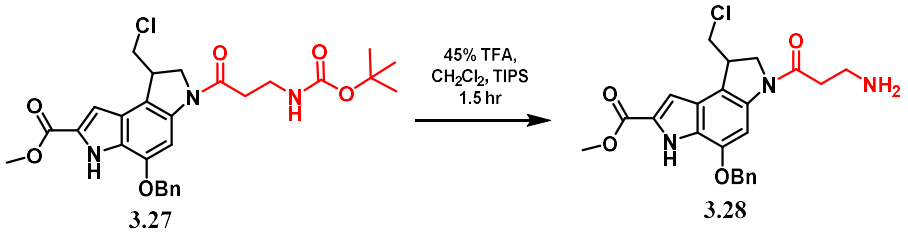
Figure 3.13: *Analytical HPLC trace of **3.27** after 2 purifications by flash chromatography. Recorded at 214 nm at 75°C (Agilent eclipse XDB-C18 column (4.6 x 150 mm, 5 µm) and a flow rate of 1 mL/min. Spectra were run with a solvent gradient of 0-100% B over 20 min. Solvent A: H_2O, 0.05% TFA, solvent B: MeCN, 0.05% TFA)*

The coupling reagents used for the formation of the amide were EDC and DIPEA. EDC is often used, as the urea by-product is easily removed. DIPEA is used as the base in this reaction to neutralise acidic by-products. The mechanism for the amide coupling can be seen in **Scheme 3.7.** EDC can activate the carboxylic acid, shown in the first step. HOBt acts as a catalyst in this reaction and can react to give the activated ester. Then the amide bond formation occurs between the DSA free amine, compound **3.17** and HOBt ester. The lone pair attacks the δ^+ carbon in a nucleophilic attack, and the following step eliminates the HOBt. The final product, compound **3.27,** is released after proton transfer.



Scheme 3.7: *Mechanism coupling of β-alanine linker,* ***3.26*** *to* ***3.17*** *using EDC and HOBt to give* ***3.27.***

Finally, the Boc needed to be deprotected to liberate the free amine. This was done using the previously established conditions of 45% TFA (**Scheme 3.8**). This reaction worked well and provided a one-to-one conversion of starting material to final product, as shown on analytical HPLC.



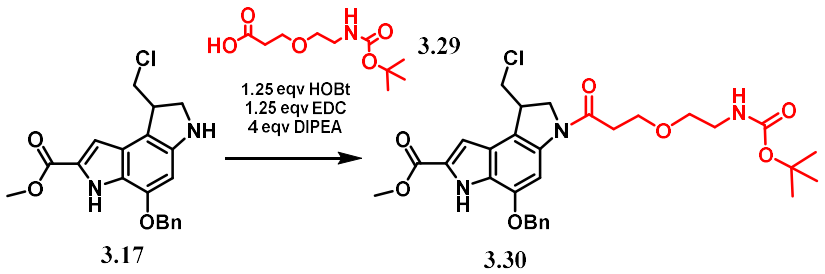
Scheme 3.8: *Boc deprotection of β-alanine linker in* ***3.27*** *to give* ***3.28***

The product was confirmed using LC-MS and a ^{1}H NMR of the crude run in acetone. The LC-MS, with an expected mass of 441.2 and an observed m/z ratio of 441.2, corresponded to the $[M]^{+}$. The crude NMR was difficult to interpret, possibly due to issues in solubility; however, it was clear that the Boc protons that are usually visible as a singlet integrating for nine protons, were no longer present. The crude product from this deprotection was carried forward to the final coupling.

3.2.4. Synthesis of the left-hand side of the DSA-PEG-DSA dimer

With the racemic mixture of the starting materials, we also explored conditions to generate the DSA-PEG-1-DSA dimer, compound **3.16**. The reactions were very similar to that of the β-alanine dimer; however, some minor differences in work up and purifications were required (**Scheme 3.9**).

The Boc-deprotection of **2.25** was carried out as before, using TFA in CH_2Cl_2 for 1.5 hours and monitoring by TLC and HPLC for completion to give compound **3.17** ready for the coupling.



Scheme 3.9*: Conditions for coupling of PEG-1 linker*

The same coupling conditions used for the synthesis of the DSA β-alanine dimer, compound **3.14,** were used to make the DSA-PEG-1-Boc, compound **3.29**: 1 equiv.

of the Boc-deprotected DSA, 1.1 equiv. of the Boc-protected PEG-1 linker, 1.25 equiv. of EDC, 1.25 equiv. of HOBt, and 4 equiv. of DIPEA. This reaction was, again, monitored closely by TLC and HPLC and was complete after 3 hours.

The DSA-PEG compound, **3.30,** was much more soluble in organic solvents; therefore, instead of precipitating it and centrifuging to collect the product, an extraction was used. After the reaction was complete, the mixture was diluted with water and extracted with EtOAc three times. The organic layers were combined and concentrated before being subjected directly to column chromatography.

The product was purified using two columns. A gradient of 20-50% acetone in hexane was followed by flushing with methanol. The methanol flush was then concentrated and subjected to another column with a gradient of 0 to 3% MeOH in CH_2Cl_2.

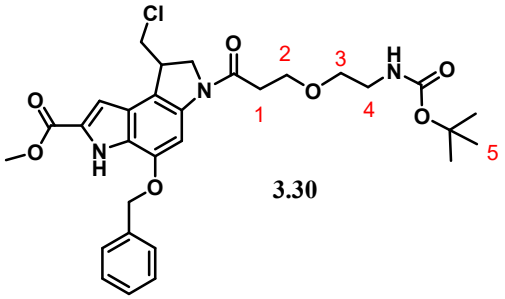
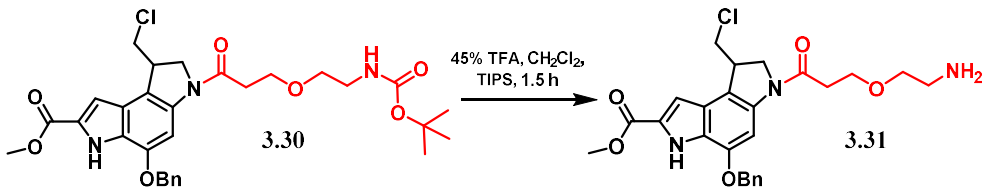


Figure 3.14: *DSA-Peg-1-Bocprotected,* ***3.30***

The solubility was much improved with this linker and meant that data could be obtained from NMR. The ^{1}H NMR was run in acetone-d^6. When compared to the starting material, there were notable changes, which supported evidence of the linker being present. All seventeen environments were accounted for, with the correct number of protons and expected splitting patterns. The ester methyl group was observed as a singlet at 3.86 ppm, integrating for 3H. The benzyl protons were noted as a singlet at 5.26 ppm, integrating for two protons. All four peaks corresponding to the PEG linker were observed and assigned as the following: (**Figure 3.14**); environment 1 at 2.78 ppm (multiplet, 2H), environment 2 at 3.83 ppm (triplet, 2H), environment 3 at 3.54 ppm (triplet, 2H), environment 4 at 3.28 ppm (doublet of triplets, 2H). The Boc group on the linker was then observed at 1.38 ppm as a singlet,

integrating for 9H. The product was further confirmed using LC-MS, with an expected mass of 585.22 and an observed *m/z* ratio of 608.1, corresponded to the $[M+Na]^+$.



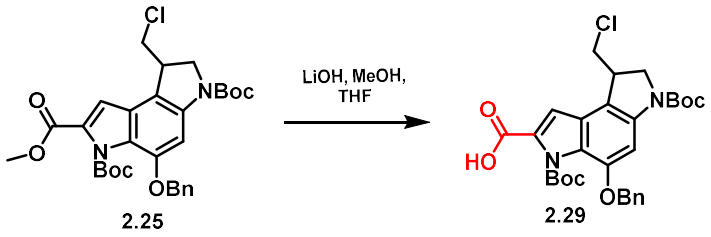
Scheme 3.10: *Boc deprotection of DSA-Peg-1 linker*

The Boc group was removed using TFA, TIPS and CH_2Cl_2 (**Scheme 3.10)** and the solvent was removed under reduced pressure. Cold diethyl ether was used to triturate the residue, giving a pale green solid.

3.2.5. Synthesis of the right-hand side of the dimers

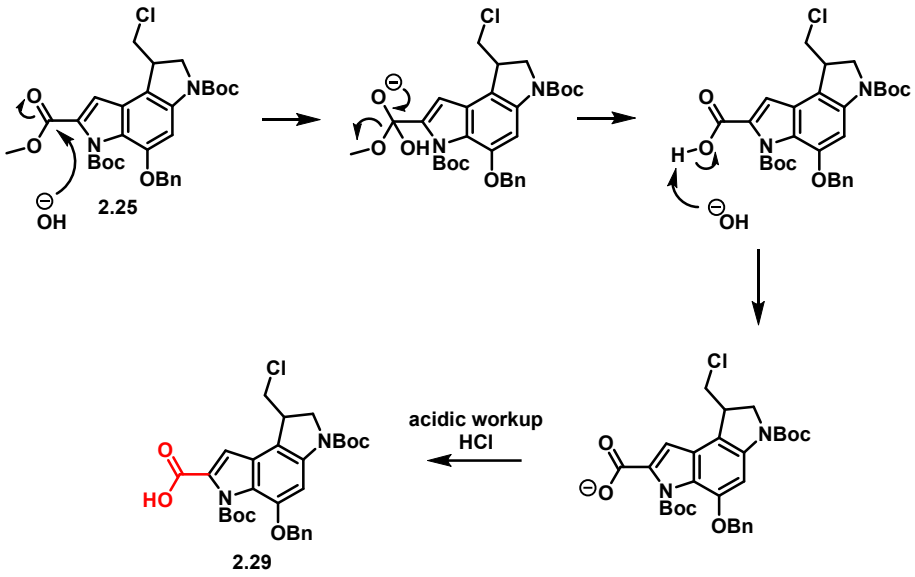
To make the right-hand side of the dimer, a reaction to generate the carboxylic acid from the ester was needed (**Scheme 3.11**). Conditions for the ester hydrolysis were determined using the route followed to generate the alkylating subunit.[12]

The DSA starting material, **2.25** was dissolved in THF and MeOH and treated dropwise with a solution of saturated aqueous LiOH. After 3 h, or when TLC showed the reaction was complete, the solvents were removed under reduced pressure. The residue was diluted with water and treated with 6 M HCl. Acidification was carried out in an ice bath. The solution was checked with litmus paper and acidification promoted precipitation. Finally, the solid was collected by filtration and dried, affording a pale green solid.



Scheme 3.11: *Conditions for ester hydrolysis to give* ***2.29***

The mechanism for this methyl ester hydrolysis is shown in **Scheme 3.12**. The first step of a base-catalysed ester hydrolysis is the hydroxide ion attacking the δ^+ carbon of the carbonyl group. There is then an electron transfer and an elimination of the methoxide ion. This leaves the carboxylic acid, the reaction is in basic conditions due to the LiOH, so the carboxylic acid is easily deprotonated. The workup therefore must be acidic to allow the protonation and the final compound, **2.29,** to be present. This protonation precipitates the final compound out of the water and allows it to be collected by vacuum filtration.



Scheme 3.12: *The base catalysed ester hydrolysis of* ***2.25*** *to give* ***2.29***

This reaction worked well, and no further purification was required. It produced a pale green solid at a consistent yield of 97% and 99.9% purity, according to HPLC integration (**Figure 3.15).** ^{1}H NMR, ^{13}C NMR and LC-MS were all used to confirm that the compound was correct. All the expected shifts were observed, with a notable difference when compared with the starting material as there was the loss of the methyl group, which was seen as a singlet that integrated for three protons at 3.87 ppm. The product was further confirmed using LC-MS, with an expected mass of 556.2 and an observed *m/z* ratio of 557.2, corresponded to the $[M+H]^+$.

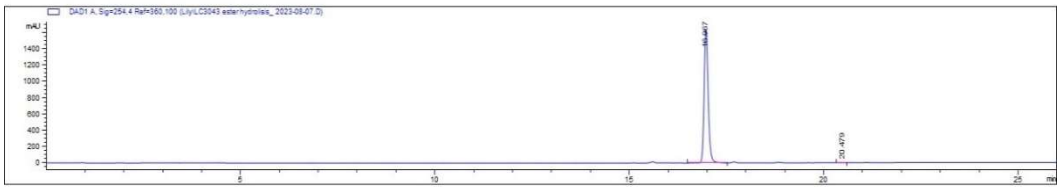
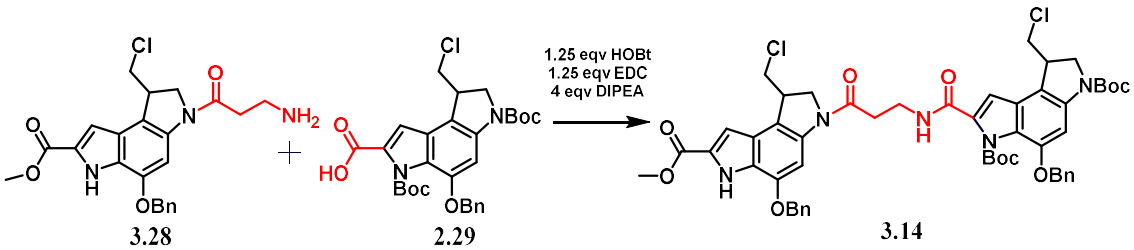


Figure 3.15: *Analytical HPLC trace for ester hydrolysed,* ***2.29****, showing 99.9% purity. Recorded at 214 nm at 75°C (Agilent eclipse XDB-C18 column (4.6 x 150 mm, 5 µm) and a flow rate of 1 mL/min. Spectra were run with a solvent gradient of 0-100% B over 20 min. Solvent A: H2O, 0.05% TFA, solvent B: MeCN, 0.05% TFA)*

3.2.6. Final coupling to synthesise the dimers, 3.14 and 3.16

To develop suitable conditions to synthesise the dimers, the racemic starting material was used to establish the appropriate conditions. However, separation of this mixture was not possible with non-chiral purification techniques, and, although late-stage separation in a chiral synthesis can be desirable, in this case the properties of the products, and the relative lack of solubility in many solvents, would make them unsuitable for such a late-stage approach. Therefore, finding a way to separate the racemic starting material, compound **2.25,** was of great importance in being able to synthesise these dimers and have a better understanding of how their varying stereochemistry may affect their biological activity.

Now both the left- (**3.28**) and right-hand **(2.29**) side of the dimer had been synthesised they were ready to be coupled together (**Scheme 3.13**).



Scheme 3.13: *Final coupling of the two synthesised subunits,* ***3.28*** *and* ***2.29*** *to give the final DSA β-alanine-DSA,* ***3.14***

The coupling conditions established in the first part of this synthesis were utilised: 1 equiv. of the DSA with linker **(3.28),** 1.1 equiv. of ester hydrolysed **(2.29),** 1.25 equiv. EDC, 1.25 equiv. HOBt, 4 equiv. of DIPEA. Initially, this was left stirring overnight.

The longer the reaction was left, the more side products were observed by TLC. As a consequence, the reaction was followed closely by TLC and analytical HPLC. The reaction time was capped at a maximum of 4 hours to try and reduce decomposition or unwanted side products.

The crude impurities could not be separated using flash chromatography, both normal and reverse phase. Therefore, attempts were made using preparative HPLC and semi-preparative HPLC. Preparative HPLC allowed two compounds, very close in polarity, to be separated. The HPLC was run in acetonitrile and the peaks came off at 15.50 minutes and 16.39 minutes. The product mass was seen in both fractions with the LC-MS, showing an expected mass of 979.33 and an observed m/z ratio of 490.0, which corresponded to the $[M+2H]^+$. After preparative HPLC, the same fractions were combined, and solvent removed. However, once checked using TLC and HPLC for purity, these products were still a mixture with the adjacent spot. This is a result of the diastereoisomers that had been formed. Two sets of enantiomers were formed in this synthesis, as described in the introduction, and as outlined in **Figure 3.10**, each enantiomeric pair will have the same physical properties; therefore, two environments are seen for the four diastereoisomers. The TLC plate was visualised under UV light. An image of this TLC after the products had been purified by preparative HPLC is shown in **Figure 3.16** to highlight the similarities of polarity**.**

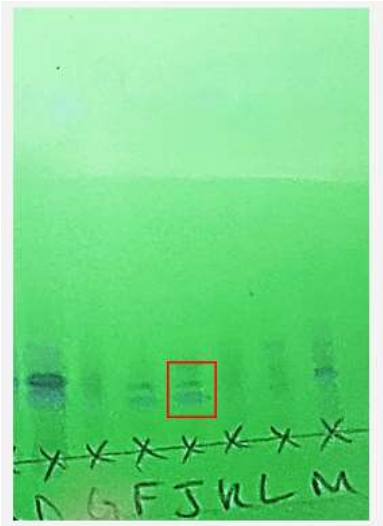
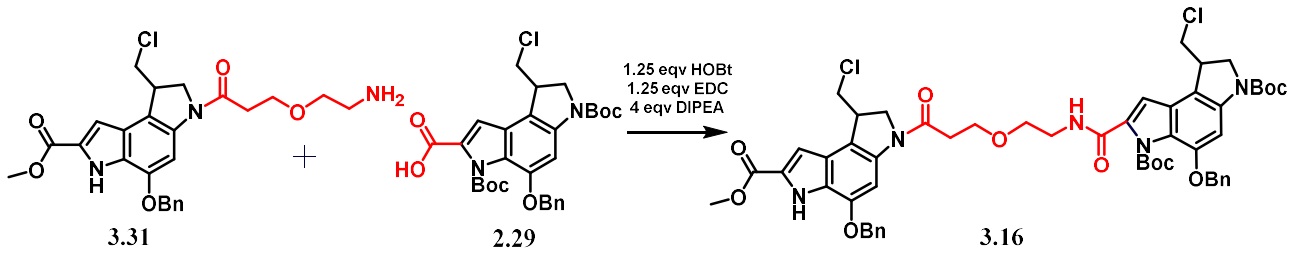


Figure 3.16: *An Image of the mixed fraction of the β-alanine dimer after purification by prep HPLC visualised on TLC under UV light*

Despite fractions containing multiple compound spots, the correct mass was found for the desired dimers. As expected, these two spots were observed because racemic starting material was used to synthesise these dimers. Their varying stereochemistry gave them different properties, although they are chemically quite similar. This TLC demonstrates that there is a diastereomeric mixture.

The same conditions were employed in an attempt to synthesise the racemic PEG dimer (**Scheme 3.14**).



Scheme 3.14: *Final amide coupling to give the DSA-Peg1-DSA dimer,* ***3.16***

Impurities were separated using preparative HPLC. Similar to the synthesis of the model beta alanine dimer, two spots were observed on TLC plate, accounted for by the two pairs of enantiomers of the four diastereoisomers. The product mass was seen in both fractions with the LC-MS, showing an expected mass of 1023.26 and an observed m/z ratio of 1023, which corresponded to the $[M+H]^+$. The expected masses of both dimers and the compounds having such similar polarities is evidence that these diasteoeoisomers were formed as expected. This mixture could not be separated; however, it was proof that the methods to synthesise these dimers were suitable, so efforts were now focussed into separating the original starting compound, **2.25,** into its enantiomers to form enantiopure dimers.

3.3 Enantiopurity – resolution of enantiomers

Racemic mixtures are not desirable for many reasons. They may have different chemical, physical and biological properties. In particular, the biological activity of compounds containing a chiral centre, or multiple chiral centres, will differ, as they bind to chiral biological molecules, such as DNA and proteins, which will lead to

unacceptable variation in activity and specificity. Purity must also be controlled to get a drug to market, as the FDA and regulatory bodies in other countries have strict rules on this. Successful separation also allows a library of compounds to be tested. For all these reasons, the separation of the racemic mixture of compound **2.25** was necessary.

3.3.1 Achieving enantiopurity through synthesis

There are different ways of achieving enantiopurity. Synthesis to produce the enantiopure material can achieve a theoretical yield of 100%, whereas separating them after synthesis of a racemic mixture achieves a maximum theoretical yield of 50%. As mentioned in **Chapter 2**, there are reported methods of synthesising the duocarmycins enantioselectively; however, these have not been reproducible. Therefore, finding alternative ways to achieve chiral separation of these compounds is of interest.

3.3.2 Achieving enantiopurity through purification

The idea of separating chiral compounds has been around for a long time. It was reported in 1815 when Jean-Baptiste Biot described the ability of compounds to rotate plane polarised light.[22] Racemic mixtures can be separated into their constituent enantiomers in different ways, and new and improved methods are constantly being developed. Some of these methods include chromatographic techniques, liquid-liquid extraction, crystallisation methods and chemical resolution.[23]

In the solid-phase synthesis approach to DSA, the successful separation of the racemic mixture of the DSA subunit with Fmoc-protection was achieved using supercritical fluid chromatography.[18] This approach was not available for this project and a different approach was sought.

A more cost-effective alternative method was identified and employed for the first time for this project and related projects. This involved chiral flash column chromatography. In normal-phase flash column chromatography, the stationary phase is silica or a silica gel which is polar. The purification or separation of compounds works by using a mobile phase, which is usually an organic solvent or mixture of solvents. The separation of mixtures is based on polarity, as compounds with a higher affinity for the stationary phase will elute slower than non-polar compounds with a

lower affinity for the stationary phase. Enantiomers have the same physical properties and therefore the same retention time, so separating them using silica is not possible. However, chiral chromatography uses a material for the stationary phase that interacts with the compounds based on their chirality. There are several different approaches to this.[24,25]

In this case, a column was used that separates the enantiomers on the basis of their interaction with a chiral sugar-based stationary phase.[26] The stationary phase is a chiral selector bonded to virgin silica. This project used chiral columns supplied by Interchim which this project uses supplies four chiral columns, the differing compounds are; Amylose tris-(3.5-dimethylphenyl carbamate), amylose tris-(3-chlorophenylcarbamate), cellulose tris-(3.5-dichlorophenylcarbamate), and cellulose tris-(3.5 dimethylphenylcarbamate).[27] This project utilises the IA Chiral column bearing amylose tris-(3.5-dimethylphenyl carbamate) immobilized on silica gel (**Figure 3.17**).[28]

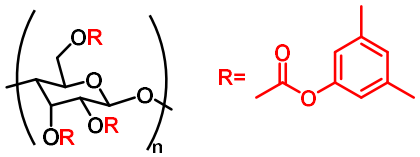


Figure 3.17: *structure of amylose tris-(3.5-dimethylphenylcarbamate)*

Separation of the enantiomers was first achieved using the 5 g puriFlash® Chiral IA 20 µm and a solvent system of an isocratic mixture of 35% isopropyl alcohol (IPA) in hexane. A maximum of 10 mg could be loaded onto this column. Using a slow gradient, each run could take several hours to complete. Having shown that the column was able to separate the enantiomers on a small scale, the approach was scaled up to allow larger amounts to be purified. Using a 25 g chiral column, different amounts were loaded to determine the maximum load that still achieved separation of the enantiomers. A maximum of 100 mg dry loaded in a minimum amount of silica could be loaded and still achieve separation. A typical run would produce a spectrum similar to that shown in **Figure 3.18 (**see **Figure 3.7** for compound structures). The graph shows absorbance, mAU against column volumes, CV.

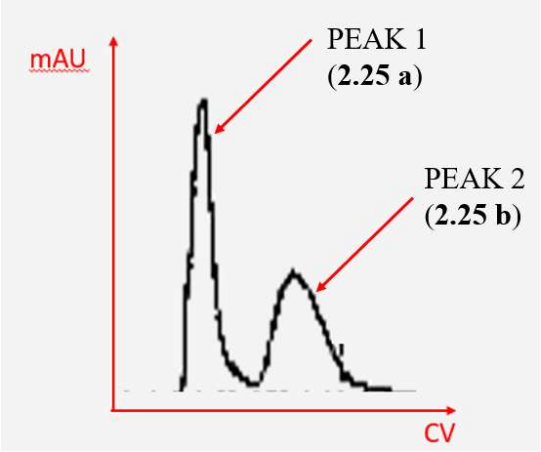


Figure 3.18: *UV trace after chiral separation of racemic* ***2.25*** *in 35% isopropyl alcohol (IPA) in hexane to give its enantiomers,* ***2.25a*** *and* ***2.25b.*** *Graph showing absorbance, mAU against column volumes, CV.*

Two tests were used to confirm that the combined fractions after purification were enantiopure. Firstly, a small amount of the combined and concentrated fractions were injected back onto the chiral column. The presence of just one peak confirmed purity to the limit of the detector. Secondly, polarimetry was used to check the optical rotation of the product and to determine which fractions corresponded to which enantiomer. There is a negative rotation of the *seco* form of the natural isomer, and a positive rotation of the *seco* form of the unnatural isomer. The data corresponded to this and the results from the polarimeter where 100 mg of the Boc-protected ester was dissolved in CH_2Cl_2 are summarised in **Table 3.1**. The table also shows the diastereomeric ratios of the two compounds, calculated by a comparison of the areas under each peak

Trace	Compound	Optical rotation	Diastereomeric ratios
PEAK 1	**2.25 a**	**-0.10**	**78**
PEAK 2	**2.25 b**	**0.11**	**22**

Table 3.1: *A summary of the data highlighting the peaks obtained from chiral separation and the values obtained from polarimetry*

3.4 Synthesis – enantiopure dimer synthesis

For the duocarmycins, the stereochemistry of the alkylation subunit determines their binding orientation to DNA. Once activated by removal of the protecting groups, the enantiomer (+)-duocarmycin SA binds preferentially to AT-rich regions in the orientation 3' to 5', the unnatural enantiomer (-)-duocarmycin SA also binds to AT rich regions but with the orientation 5' to 3'.[29] When trying to design an ultra-potent dimer, it is important to understand the order in which the units are bound. Previous work showed it was the first subunit, that controlled binding and, therefore, potency (section **3.1.4**, the studies done by Boger *et al)*.[17] It is not clear, however, that this will continue to be the case with dimers in which a linker is introduced, although it would suggest that the two most potent dimers would be the ones starting with the natural subunit.

As discussed in the introduction of this chapter, Boger's studies also showed that having a group with the same stereochemistry vs different influenced whether these compounds were able to form intrastrand crosslinks or interstrand crosslinks with DNA.[17] As this project aims to synthesise a highly potent payload, the dimers of interest will start with the natural enantiomer first. DNA binding can also affect the potency so having access to both the natural-natural (*S, S*) dimers and the natural-unnatural (*S*, *R*) dimers would be of interest when testing these compounds in a biological setting.

As the unnatural *R*-enantiomer will be required in smaller amounts for the target compounds, trial synthesis of the non-racemic dimers focussed on this enantiomer and the synthesis of the unnatural-unnatural (*R*, *R*) dimers. All six of the enantiopure dimers that this project aims to synthesise are summarised in **Figure 3.19.** The next chapter will discuss the removal of the benzyl protecting group in readiness for biological testing.

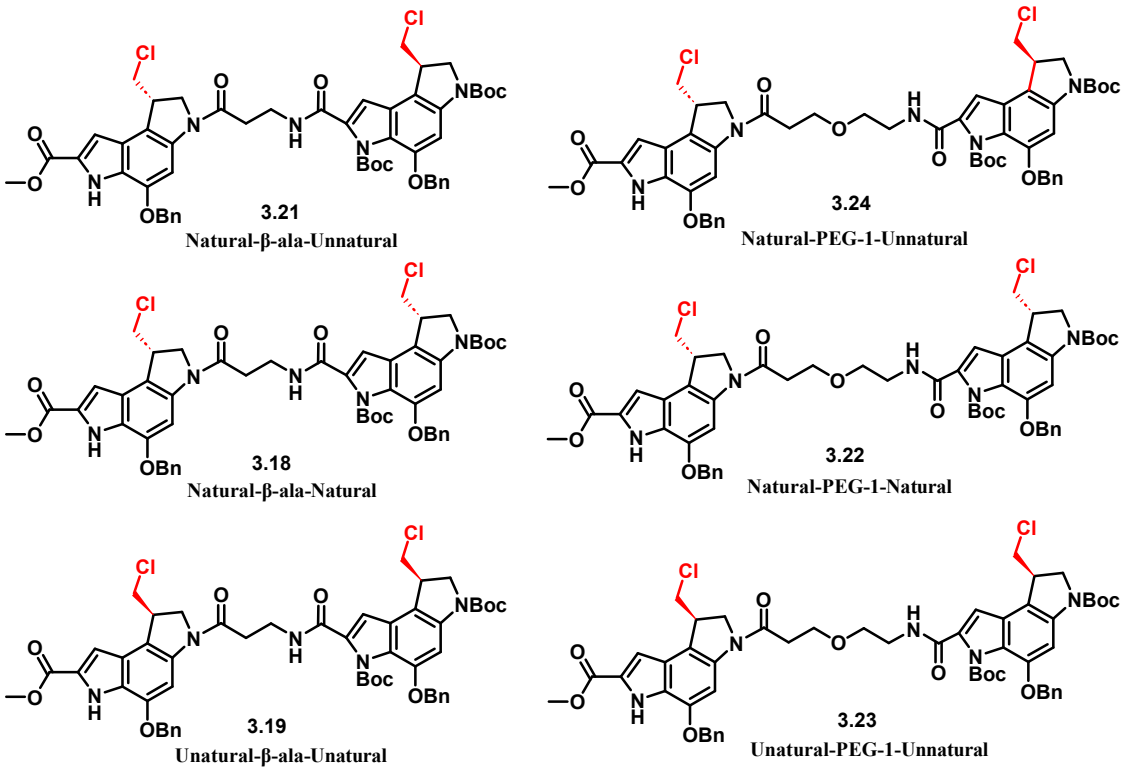


Figure 3.19: *The six enantiopure dimers this project aims to synthesise*

The reactions were followed as described in the trial synthesis earlier in this chapter. However, increased equivalents of coupling agents drove the reactions to completion and there were often fewer impurities. This meant that the final coupling to make the DSA dimers did not need to be purified by preparative HPLC. The reactions were worked up as before, and the yields, purities and mass spectral data for these compounds are shown in the following table (**Table 3.2**). **Figure 3.20** shows the HPLC traces for compounds **3.16, 3.17, 3.18, 3.20, 3.22** and **3.21.** The purifications varied depending on the compounds and yields and data will be described in the experimental. The mass of the product was confirmed by high resolution mass spectrometry for the natural-natural dimers (**3.16** and **3.20**) and LC-MS for each individual dimer (See **Chapter 5** Experimental details).

Compound Number	Linker	Stereochemistry	Yield (mg, %)	Purity (HPLC)	Mass spec
3.18	β–Ala	Natural-natural	2.6 mg 7.45%	97%	HRMS
3.20	β–Ala	Natural-unnatural	4.6 mg 13.39 %	87%	LC-MS
3.19	β–Ala	Unnatural-unnatural	15.9 mg 46%	86%	LC-MS
3.22	PEG-1	Natural-natural	9.7 mg 27 %	83%	HRMS
3.24	PEG-1	Natural-unnatural	8.5 mg 24%	85%	LC-MS
3.23	PEG-1	Unnatural-unnatural	10.6 mg 44%	88%	LC-MS

Table 3.2: *A summary of the results from the dimer synthesis of compounds* ***3.18, 3.19, 3.20, 3.22, 3.24*** *and* ***3.23*** *showing percentage yield and purity*

Table 3.2 shows that the yields from the deprotection varied across the different compounds, although this might be expected when working on such small amounts of material. The purities on the whole were in excess of 80% according to the HPLC. NMR data for these compounds could only be obtained in DMSO, due to solubility. Recovery from DMSO solvent is poor, so NMR attempts would be made after the benzyl group had been removed at the final step. The purity and masses were confirmed by observing a single peak on HPLC, the correct m/z ratio on LC-MS and the correct m/z ratio on HRMS.

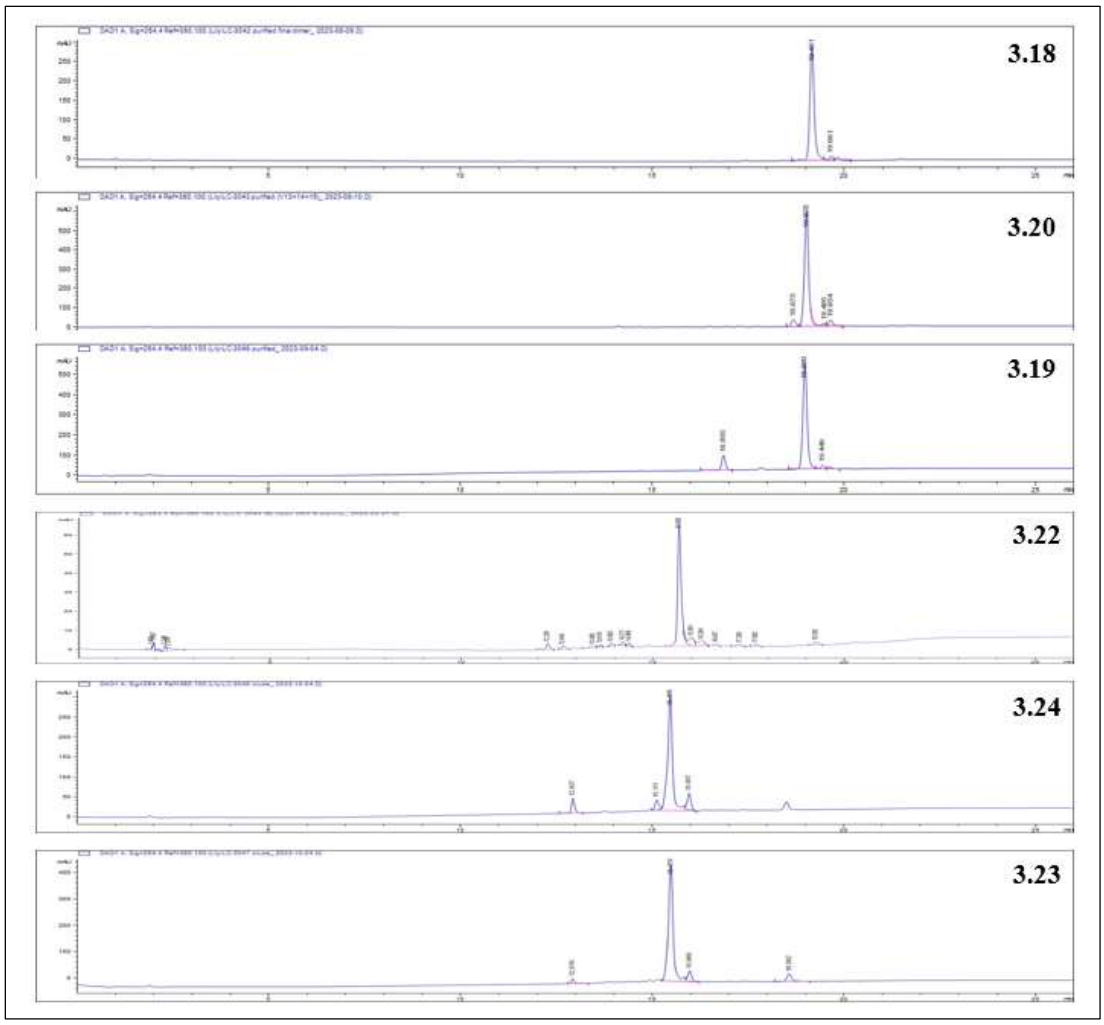


Figure 3.20: *Analytical HPLC traces of compounds* ***3.18, 3.19, 3.20, 3.22, 3.23, 3.24***

Recorded at 214 nm at 75°C (Agilent eclipse XDB-C18 column (4.6 x 150 mm, 5 µm) and a flow rate of 1 mL/min. Spectra were run with a solvent gradient of 0-100% B over 20 min. Solvent A: H2O, 0.05% TFA, solvent B: MeCN, 0.05% TFA)

3.5 Solid Phase synthesis for dimers

Solid-phase synthesis has been used previously for the synthesis of duocarmycin analogues.[18] The concept of solid-phase synthesis was introduced in **Chapter 1**, which describes solid-phase synthesis as a technique used to build a structure by binding together subunits onto an insoluble material.[30] This approach, in theory, has several advantages over solution-phase synthesis. For this project, the advantages include easy application, due to the subunit having access to one carboxylic acid and one amine. This allows consecutive couplings and, therefore, the ability to make dimers and higher order structures.[31] Additionally, it could offer ease of purification.[32] As the

growing compound is insoluble, solvents and excess unreacted materials are washed away, allowing the synthesis to be carried out in one vessel, avoiding repeated purifications between reaction steps.

Solid-phase synthesis requires a solid support, which acts as the initial building block. This is often a resin bead commonly made up of crosslinked polystyrene, or crosslinked polyamide-based resins and composite polystyrene-polyethylene glycol-based resins.[33] The resin type will depend on the type of coupling planned and the compounds used. The first step in solid phase synthesis is to load the *N*-protected *C*-terminal amino acid. The type of resin used will determine the end product having either a *C*-terminal acid or an amide. The linker is important as it determines what conditions can be used for the couplings and what conditions are needed to cleave the compound from the resin. This is a key feature to be considered if the compounds being designed are unstable to acid or base. Many resins can be purchased with the linker pre-attached. In this case the DSA payload will have a terminal amide. Resins used for the synthesis of compounds such as this include Rink-amide resins, Sieber resins and PAL resins (**Figure 3.21**).[34]

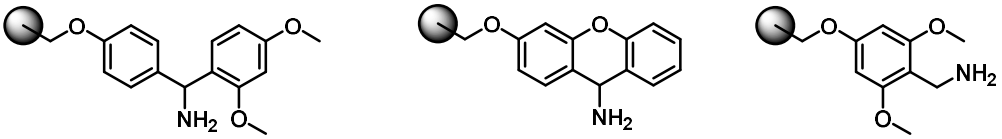


Figure 3.21: *Resins for the use in amide peptide synthesis*

This section aims to investigate the use of solid-phase synthesis to make duocarmycin-based dimers. The following two sub-sections will describe the synthesis of dimers shown in **Figure 3.22**, comparing the synthesis using solution phase vs solid phase methods to decide which is the preferred approach for this project. These compounds were previously described as noted above. Previous research in the group investigated the activity of the duocarmycins with different amino acids attached, following solid phase synthesis.[18] From these findings, it was decided that a rink-amide MBHA resin with an alanine would be the optimum starting point. Compounds with an alanine attached maintained their potency, while other amino acids such as lysine led to

inactive compounds.[18]An alanine will be first coupled to the resin bead and then the dimer will be built and acetylated before cleavage.

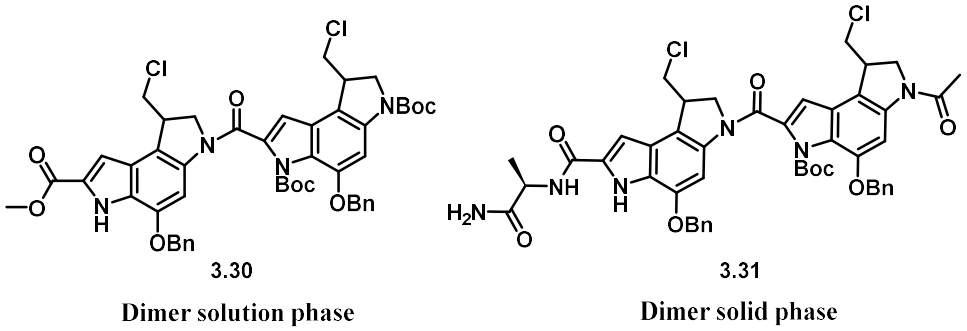
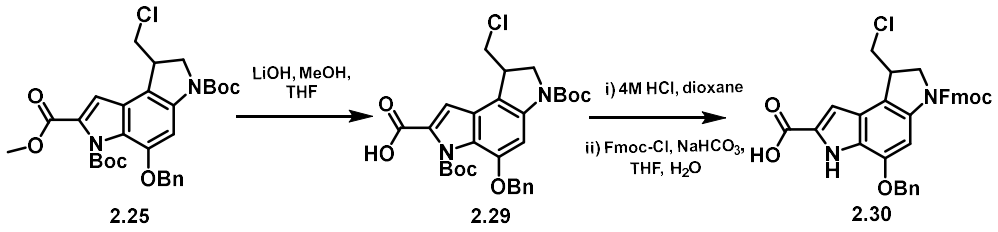


***Figure 3.22**: Dimers for comparison of solution-phase vs solid-phase*

3.5.1 Synthesis of Fmoc-protected subunit

Chapter 1 discussed the importance of Fmoc-protection in solid-phase synthesis, a strategy previously used to generate duocarmycin analogues by solid-phase synthesis.[18] The terminal amine needs to be protected using an Fmoc group, which is carried out by reacting the amine with fluorenyl methoxycarbonyl chloride (Fmoc-Cl). Protection is necessary to prevent unwanted side reactions: without the Fmoc protecting group the individual subunits could couple to one another rather than to the resin or to the next subunit, as they have an amino group and a carboxylic acid. The synthesis to generate this subunit is described by *Stephenson*, where the alkylating subunit is Fmoc-protected in a two-step reaction (**Scheme 3.15).**[18]



***Scheme 3.15:** Synthesis of DSA alkylating subunit to make it applicable in solid-phase synthesis*

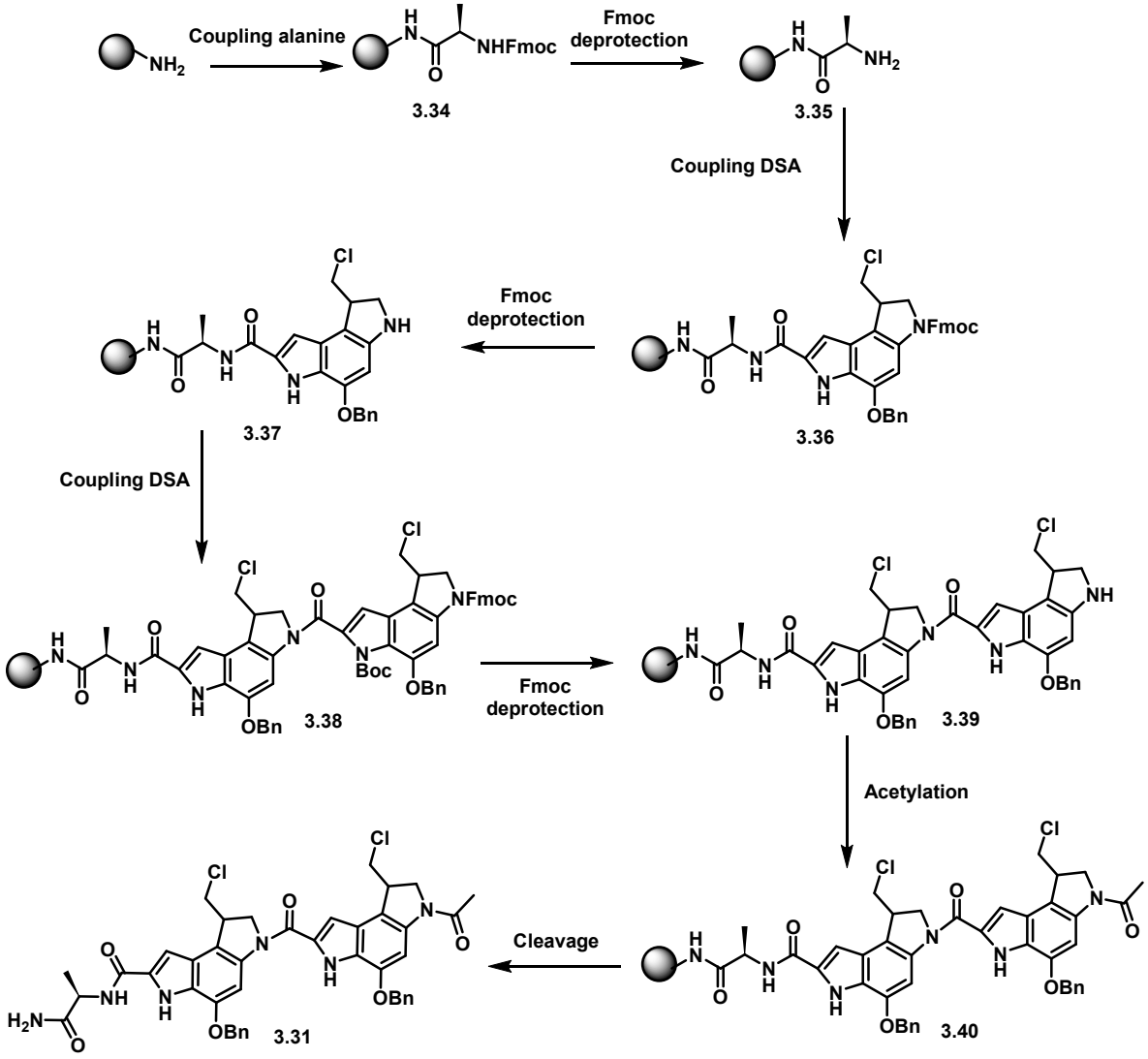
The synthesis of the starting subunit was described in **Chapter 2**. The ester hydrolysis was already described earlier in this chapter (**Scheme 3.11**). The next steps followed the procedure described in the literature to synthesise the Fmoc-protected final subunit.[18]

This is a two-step procedure, the first of which involves the deprotection of the Boc group using 4M HCl in dioxane. This was left overnight to stir before the dioxane was removed under reduced pressure. The mixture was then cooled to 0 ºC, dissolved in THF, and treated dropwise with $NaHCO_3$ in water with Fmoc-Cl dissolved in THF. This was stirred for 5 min and then quenched using MeOH. The solvents were then removed under reduced pressure, and the remaining solution was acidified using 2 M HCl. This solution was extracted using EtOAc and then the organic layer was combined and concentrated before the product was subject directly to flash column chromatography. Purification was conducted using a gradient of 0 to 5% MeOH in CH_2Cl_2, and the yield for this two-step process was 37%. This final compound was sensitive to degradation so was stored in the freezer in dark conditions. NMR data was in agreement with that of literature data.[18] Notable differences were observed in the ^{1}H NMR when compared with the starting material: there was no longer a methyl group at 3.87 ppm seen as a singlet, integrating for three hydrogens. Both Boc groups were removed and a Fmoc group used in place of one. This was confirmed on ^{1}H NMR as disappearance of both singlets, integrating for nine hydrogens, was observed at 1.48 and 1.39 ppm. And the new peaks as a result of the Fmoc were observed as an increase of eight protons in the aromatic region, two protons at 4.55 ppm, and a multiplet integrating for one hydrogen at 4.18.

3.5.2 Solid-phase synthesis to make the dimer DSA-DSA

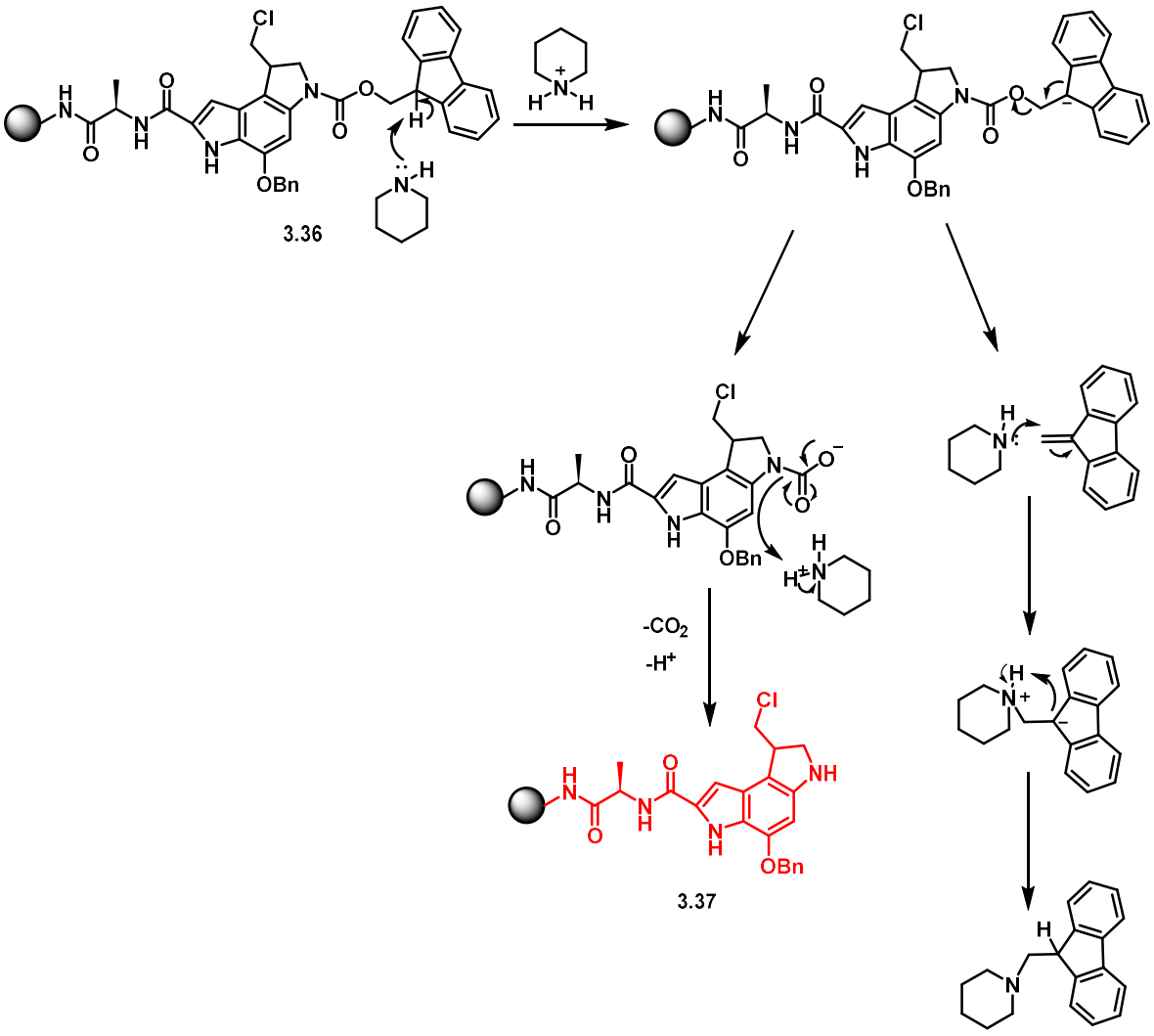
As this will be a linear synthesis, differing from solution phase chemistry to make dimers, there is no difference in preparing the left-hand subunit and the right-hand subunit – they both utilise the Fmoc-protected subunit directly on the resin. The overall reaction scheme can be seen in **Scheme 3.16**, showing each step of the solid-phase synthesis to obtain the final dimer. The first step is to swell the resin – this project used the rink-amide MBHA resin. Swelling is important as 99% of the coupling sites are not on the surface but are inside the resin bead, and therefore should be swollen in DMF for 20-30 mins prior.[33] The resin was then deprotected using 20% piperidine in DMF. This was repeated twice, with the solutions drained and the resin washed in between. Afterwards, the resin was washed thoroughly with DMF. The next step was then the coupling of the alanine. This was carried out using Fmoc-alanine with the coupling agents hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU)

and DIPEA. The mixture was added to the resin and was shaken for 45 minutes, washed and repeated. After the alanine coupling, Fmoc-deprotection involved using a 20% piperidine solution in DMF. This was shaken for 10 minutes and repeated twice with washings in between. The DSA was then coupled using HATU and DIPEA. This was added to the resin and shaken overnight – the extended reaction times for coupling were reported to be necessary for good conversion. After washing, Fmoc deprotection was repeated. The second DSA subunit was then coupled using HATU and DIPEA, and, again, the reaction was shaken overnight. Final Fmoc-deprotection was followed by thorough washing, before the acetylation. AcCl and DIPEA were added to the resin and the mixture was shaken for 1 h. This process was repeated. The cleavage of the dimer from the resin was then carried out using 47.5% TFA, 46.5% DCM, 2.5% TIPS, 2.5% H_2O cleavage for 2 h. The collected washings were concentrated to give a black solid.



Scheme 3.16: *Attempted formation of a duocarmycin dimer using solid phase synthesis*

The Fmoc groups are removed using piperidine (see mechanism in **Scheme 3.17**). Removing the Fmoc with piperidine is suitable as it does not affect the acid-labile linker to the resin. The lone pair of the nitrogen on the piperidine acts as a nucleophile to remove the acidic proton shown on the ring system in Fmoc. The other side products formed in the subsequent electron transfers are soluble in DMF so can be washed away prior to the next coupling.



Scheme 3.17: *Mechanism of Fmoc deprotection*

The product from this attempted synthesis could not be identified. It did not give the correct mass spectrum or NMR spectrum for the target compound. After several attempts to synthesise this compound, the solid-phase route was deemed not reproducible for this project.

3.5.3 Solution phase synthesis to make the dimer DSA-DSA

The synthesis of the directly linked dimers, similar to those previously described, **(**sections **3.2.3 and 3.2.4)** was also carried out in solution. This reaction was trialled with racemic starting material. The overall scheme to synthesise the DSA-DSA subunit in solution-phase is summarised in **Figure 3.23**. TFA was used to deprotect the Boc group and generate the free amine for the left-hand alkylating subunit. LiOH

was used to hydrolyse the ester and synthesise the right-hand subunit. These two subunits were coupled to one another directly, using the coupling conditions established previously which were: 1 equiv. of Boc deprotected (**3.15),** 1.1 equiv. of ester hydrolysed **(2.29),** 1.25 equiv. EDC. 1.25 equiv. HOBt, 4 equiv. of DIPEA.

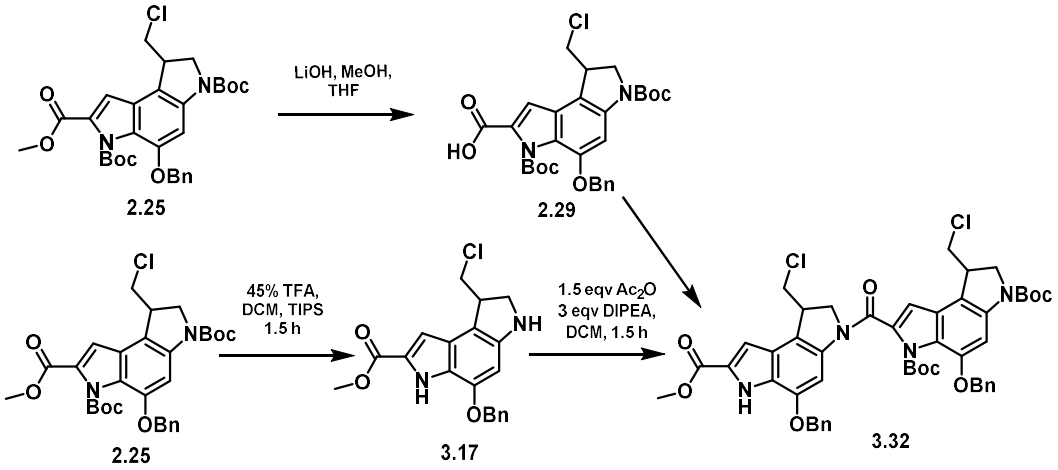


Figure 3.23: *Scheme showing solution phase synthesis to make DSA-DSA,* ***3.32***

The final coupling worked well, and the correct product was isolated and confirmed using LC-MS and HRMS. The product was purified using flash chromatography with a gradient of 0 to 15% EtOAc in hexane. Pure product was obtained in a 6% overall yield. There was one spot observed on a TLC plate and one peak observed on HPLC, despite this synthesis providing all 4 diastereoisomers. This is likely to be due to the four compounds having the same or very similar polarities so they appear as one compound. This theory could be confirmed using chiral chromatography and then followed by optical rotation to confirm the presence of four differing compounds.

The product was confirmed using LC-MS, showing an expected mass of 909.86 and an observed m/z ratio of 304.4, corresponded to the $[M+3H]^+$. Data was also obtained on the HRMS which gave an m/z ratio of 932.2989 corresponding to the $[M+Na]^+$. The NMR was difficult to interpret due to broad overlapping peaks, perhaps due to the presence of multiple diastereoisomers.

3.6. Conclusion of Chapter 3

This chapter describes finding appropriate conditions to synthesise and purify six novel DSA-DSA dimers with varying linker units. The project also demonstrated effective chiral separation using chiral flash column chromatography, which is a new method for the duocarmycins and relatively inexpensive. Having access to both enantiomers and having developed a method to synthesise dimers meant that a set of six novel enantiopure dimers could be synthesised. As well as this, a duocarmycin alkylating monomer was synthesised and purified that will allow direct comparison with the dimers in **Chapter 4.**

This chapter investigated the feasibility of synthesising dimers through solid-phase methodology. Despite the inherent advantages associated with solid-phase synthesis, it proved unsuitable for dimers due to the absence of product formation. The equivalents needed are high and, in our hands, the minimum amount needed for efficient coupling is around 15 mg, whereas in solution much less can be used. In this case, solution-phase chemistry also had the advantages that we could closely follow the reactions using standard liquid or thin layer chromatography methods, which is advantageous when working on such a small scale. In this project, we also generated low yields of the Fmoc-protected DSA subunit, which, at this late stage of the synthesis was a disadvantage. Starting the synthesis of the dimers two steps earlier, from the di-Boc-protected compound (**2.25**), was more efficient.

References

1 H. Tang, Y. Liu, Z. Yu, M. Sun, L. Lin, W. Liu, Q. Han, M. Wei and Y. Jin, The Analysis of Key Factors Related to ADCs Structural Design, *Front. Pharmacol.*, 2019, **10**, 1–11.

2 M. K. Hadden and B. S. J. Blagg, *Dimeric Approaches to Anti-Cancer Chemotherapeutics*, .

3 D. E. Thurston and P. J. M. Jackson, *Cytotoxic Payloads for Antibody–Drug Conjugates*, Royal Society of Chemistry, 2019.

4 S. K. Sharma, G. Jia and J. W. Lown, Novel Cyclopropylindole Conjugates and Dimers Synthesis and Anti-Cancer Evaluation, *Curr. Med. Chem. - Anti-Cancer Agents*, 2001, **1**, 27–45.

5 Tietze, F. Lutz and B. Krewer, Novel Analogues of CC-1065 and the Duocarmycins for the Use in Targeted Tumour Therapies, *Anticancer Agents Med. Chem.*, 2009, **9**, 304–325.

6 Y. Fukuda, S. Seto, H. Furuta, H. Ebisu, Y. Oomori and S. Terashima, Novel Seco Cyclopropa[c]pyrrolo[3,2-e]indole Bisalkylators Bearing a 3,3‘-Arylenebisacryloyl Group as a Linker, *J. Med. Chem.*, 2001, **44**, 1396–1406.

7 Y.-C. Han, J. Kahler, N. Piché-Nicholas, W. Hu, S. Thibault, F. Jiang, M. Leal, M. Katragadda, A. Maderna, R. Dushin, N. Prashad, M. B. Charati, T. Clark, L. N. Tumey, X. Tan, A. Giannakou, E. Rosfjord, H.-P. Gerber, L. Tchistiakova, F. Loganzo, C. J. O’Donnell and P. Sapra, Development of Highly Optimized Antibody–Drug Conjugates against CD33 and CD123 for Acute Myeloid Leukemia, *Clin. Cancer Res.*, 2021, **27**, 622–631.

8 D. Su, J. Chen, E. Cosino, J. Dela Cruz-Chuh, H. Davis, G. Del Rosario, I. Figueroa, L. Goon, J. He, A. V. Kamath, S. Kaur, K. R. Kozak, J. Lau, D. Lee, M. V. Lee, D. Leipold, L. Liu, P. Liu, G. L. Lu, C. Nelson, C. Ng, T. H. Pillow, P. Polakis, A. G. Polson, R. K. Rowntree, O. Saad, B. Safina, N. J. Stagg, M. Tercel, R. Vandlen, B. S. Vollmar, J. Wai, T. Wang, B. Wei, K. Xu, J. Xue, Z. Xu, G. Yan, H. Yao, S. F. Yu, D. Zhang, F. Zhong and P. S. Dragovich, Antibody-drug conjugates derived from cytotoxic seco-CBI-Dimer payloads are highly efficacious in xenograft models and form protein adducts in Vivo, *Bioconjug. Chem.*, 2019, **30**, 1356–1370.

9 G. Jia and J. W. Lown, Design , Synthesis and Cytotoxicity Evaluation of 1-Chloromethyl- 5-hydroxy-1 , 2-dihydro-3 H -benz [e] indole (seco -CBI) Dimers, 2000, **8**, 1607-1617

10 K. W. Świderska, A. Szlachcic, Ł. Opaliński, M. Zakrzewska and J. Otlewski, FGF2 Dual Warhead Conjugate with Monomethyl Auristatin E and α-Amanitin Displays a Cytotoxic Effect towards Cancer Cells Overproducing FGF Receptor 1, *Int. J. Mol. Sci.*, 2018, **19**, 2098.

11 M. A. Krzyscik, Ł. Opaliński and J. Otlewski, Novel Method for Preparation of Site-Specific, Stoichiometric-Controlled Dual Warhead Conjugate of FGF2 via Dimerization Employing Sortase A-Mediated Ligation, *Mol. Pharm.*, 2019, **16**, 3588–3599.

12 S. C. Wilson, H. Philip, S. Forrow, J. A. Hartley, T. Adams, J. kelland and D. E. Thurston, Design, Synthesis, and Evaluation of a Novel Sequence-Selective Epoxide-Containing DNA Cross-Linking Agent Based on the Pyrrolo[2,1-c][1,4]benzodiazepine System | Journal of Medicinal Chemistry, *J Med Chem*, 1999, **42**, 4028–4041.

13 P. J. M. Jackson, K. M. Rahman and D. E. Thurston, The use of molecular dynamics simulations to evaluate the DNA sequence-selectivity of G–A cross-linking PBD–duocarmycin dimers, *Bioorg. Med. Chem. Lett.*, 2017, **27**, 102–108.

14 T. H. Pillow and T. Moana, in *RSC Drug Discovery Series*, Chapter 11: Duocarmycin-PBD Dimers as Antibody-Drug Conjugate (ADC) Payloads 2019.

15 B. Purnell, A. Sato, A. O'Kelley, C. Price, K. Summerville, S. Hudson, C. O'Hare, K. Kiakos, T. Asao, M. Lee and J. A. Hartley, DNA interstrand crosslinking agents: Synthesis, DNA interactions, and cytotoxicity of dimeric achiral seco-amino-CBI and conjugates of achiral seco-amino-CBI with pyrrolobenzodiazepine (PBD), *Bioorg. Med. Chem. Lett.*, 2006, **16**, 5677–5681.

16 M. Tercel, S. M. Stribbling, H. Sheppard, B. G. Siim, K. Wu, S. M. Pullen, K. J. Botting, W. R. Wilson and W. A. Denny, Unsymmetrical DNA Cross-Linking Agents: Combination of the CBI and PBD Pharmacophores, *J. Med. Chem.*, 2003, **46**, 2132–2151.

17 D. L. Boger, M. Searcey, W. C. Tse and Q. Jin, Bifunctional alkylating agents derived from duocarmycin SA: Potent antitumor activity with altered sequence selectivity, *Bioorg. Med. Chem. Lett.*, 2000, **10**, 495–498.

18 M. J. Stephenson, L. A. Howell, M. A. O'Connell, K. R. Fox, C. Adcock, J. Kingston, H. Sheldrake, K. Pors, S. P. Collingwood and M. Searcey, Solid-Phase Synthesis of Duocarmycin Analogues and the Effect of C-Terminal Substitution on Biological Activity, *J. Org. Chem.*, 2015, **80**, 9454–9457.

19 F. J. Ferrer-Gago, L. Q. Koh and D. P. Lane, Functionalized Resins for the Synthesis of Peptide Alcohols, *Chem. – Eur. J.*, 2020, **26**, 379–383.

20 M. J. Cope, SAC 92. Understanding chirality: how molecules that are mirror images of each other can act differently in the body, *Anal. Proc.*, 1993, **30**, 498–500.

21 P. Reveglia, A. Cimmino, M. Masi, P. Nocera, N. Berova, G. Ellestad and A. Evidente, Pimarane diterpenes: Natural source, stereochemical configuration, and biological activity, *Chirality*, 2018, **30**, 1115–1134.

22 D. J. Cordato, L. E. Mather and G. K. Herkes, Stereochemistry in clinical medicine: a neurological perspective, *J. Clin. Neurosci.*, 2003, **10**, 649–654.

23 J. Sui, N. Wang, J. Wang, X. Huang, T. Wang, L. Zhou and H. Hao, Strategies for chiral separation: from racemate to enantiomer, *Chem. Sci.*, 2023, **14**, 11955–12003.

24 T. E. Beesley and R. P. W. Scott, *Chiral Chromatography*, John Wiley & Sons, 1999.

25 A. Tarafder and L. Miller, Chiral chromatography method screening strategies: Past, present and future, *J. Chromatogr. A*, 2021, **1638**, 461878.

26 P. Bako, G. Keglevich, Z. Rapi and L. Toke, The Enantiomeric Differentiation Ability of Chiral Crown Ethers Based on Carbohydrates, *Curr. Org. Chem.*, 2012, **16**, 297–304.

27 Purification of enantiomers with chiral puriFlash® columns, https://blog.interchim.com/purification-of-enantiomers-with-chiral-puriflash-columns/, (accessed 4 March 2024).

28 C. Yamamoto, E. Yashima and Y. Okamoto, Structural Analysis of Amylose Tris(3,5-dimethylphenylcarbamate) by NMR Relevant to Its Chiral Recognition Mechanism in HPLC, *J. Am. Chem. Soc.*, 2002, **124**, 12583–12589.

29 D. L. Boger, D. L. Hertzog, B. Bollinger, D. S. Johnson, H. Cai, J. Goldberg and P. Turnbull, Duocarmycin SA shortened, simplified, and extended agents: A systematic examination of the role of the DNA binding subunit, *J. Am. Chem. Soc.*, 1997, **119**, 4977–4986.

30 R. B. Merrifield, Automated Synthesis of Peptides, *Science*, 1965, **150**, 178–185.

31 K. Rose and J. Vizzavona, Stepwise Solid-Phase Synthesis of Polyamides as Linkers, *J. Am. Chem. Soc.*, 1999, **121**, 7034–7038.

32 P. Seneci, *Solid-Phase Synthesis and Combinatorial Technologies*, John Wiley & Sons, 2003.

33 M. Amblard, J.-A. Fehrentz, J. Martinez and G. Subra, Methods and protocols of modern solid phase peptide synthesis, *Mol. Biotechnol.*, 2006, **33**, 239–254.

34 P. T. Shelton and K. J. Jensen, in *Peptide Synthesis and Applications*, eds. K. J. Jensen, P. Tofteng Shelton and S. L. Pedersen, Humana Press, Totowa, NJ, 2013, pp. 23–41.

Chapter 4 – DNA Crosslinking and Cancer Therapy

Chapter 4 – DNA Crosslinking and Cancer Therapy

This chapter explores how the compounds synthesised in this project bind to DNA and investigates how this may affect potency. As discussed previously, the stereochemistry of the duocarmycins determines binding, and binding determines potency. Having found a route to synthesise a set of novel dimers, and having also found a way to separate the racemic mixture, the next steps of this project focused on a series of biological assays. This chapter discusses the different ways these compounds may bind to DNA and investigates how varying the stereochemistry within the dimers affects their structure-activity relationships.

4.1 DNA crosslinking

DNA crosslinking can occur through **inter**strand crosslinking and **intra**strand crosslinking. There is also the option for dual 'warheads' to cause duplex-duplex crosslinks.

4.1.1 Introduction to DNA crosslinking

Figure 4.1 provides a schematic to compare different methods of crosslinking. It highlights how a compound (blue) may bind to the same strand of DNA, to different strands of the same DNA molecule, or form crosslinks between two separate DNA molecules.

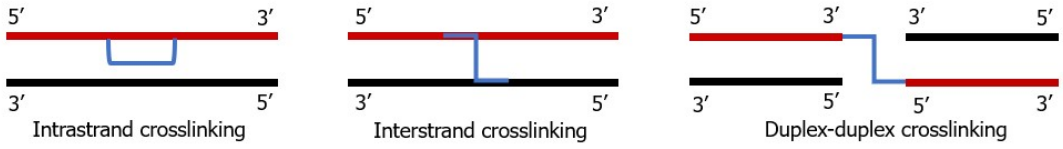


***Figure 4.1**: A simplified schematic showing the differences between different DNA crosslinking mechanisms.*

Intrastrand crosslinking is where binding occurs at two sites on the same DNA strand.[1] Interstrand crosslinking is where opposite strands of the same DNA helix are connected by nucleotides bound to a chemical bridge.[2] Interstrand crosslinks are believed to have the most potent effect, and are able to inhibit both replication and transcription, leading to cell death.[3] These crosslinks can form through several different methods, including normal cellular processes (e.g. oxidation of lipids),

irradiation of DNA in the presence of intercalating drugs, and also by the action of chemotherapeutic agents.[4] This thesis focuses on bifunctional alkylating compounds that can be applied as chemotherapeutic agents. Examples of compounds that are able to form these covalent bonds between purine bases on opposing strands include mechlorethamine, cisplatin and mitomycin C.[5]

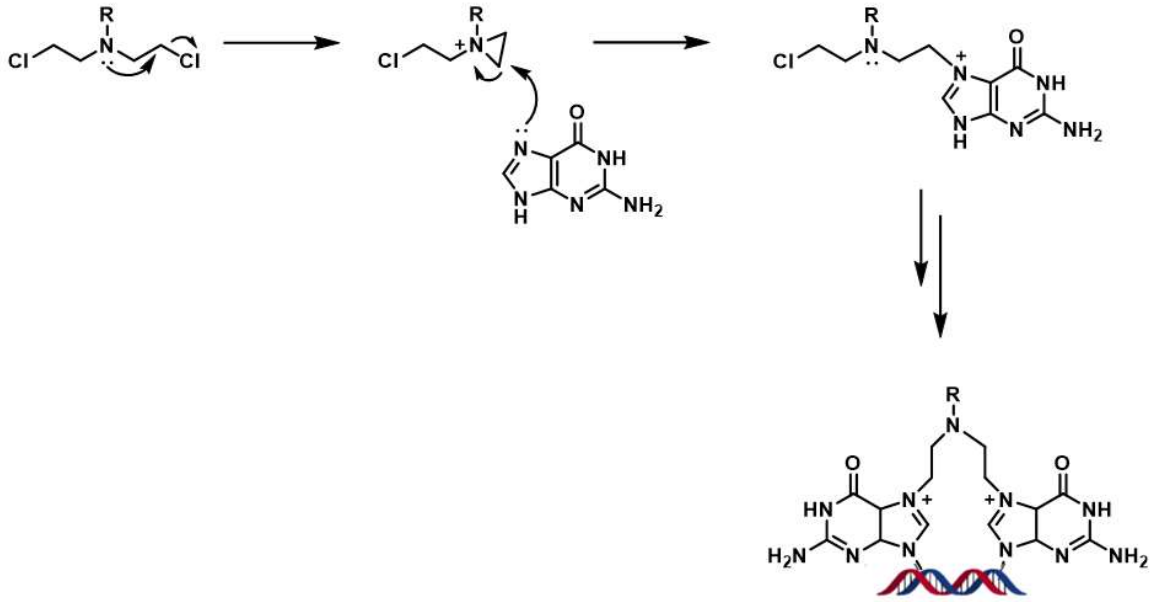
Another interesting concept is interstrand crosslinking between two separate DNA helices. Some bifunctional compounds induce this duplex-duplex crosslinking.[6] This concept will be explored in more detail later in the chapter using a gold nanoparticle assay to show the ability of known bis-acridines to form these crosslinks. These known crosslinking agents will be compared to compounds synthesised in this project to determine whether these are also able to bind to DNA in this way.

4.1.2 DNA crosslinking and its value in cancer treatment

Compounds able to crosslink DNA are promising agents for cancer therapeutics. The uncontrolled replication of cancer cells means they are particularly vulnerable to agents that disrupt DNA replication. When crosslinking compounds bind to DNA, they inhibit the separation of strands necessary for replication and disrupt the accurate copying of base pair sequences. Consequently, the cell is unable to divide and replicate, ultimately leading to cell death. DNA crosslinking agents are therefore highly desirable, particularly if they can be directed specifically toward cancer cells. This has been a successful approach, and studies on these compounds started as early as the late 1940s.[7] DNA crosslinking agents are still used as an effective way to treat cancer, and in 2018, there were reportedly 4400 clinical trials utilising these agents.[3]

An example of a family of DNA intrastrand crosslinking agents are the nitrogen mustards. These group of compounds were introduced in **Chapter 1** (**Figure 1.9**), which discussed that these were the first cancer chemotherapeutic agents used, after being developed from mustard gas used in the first world war.[8] These compounds are DNA alkylators and their high potency is due to their ability to form crosslinks with DNA. **Scheme 4.1** highlights the mechanism by which these compounds cyclise and form crosslinks with DNA. Mechlorethamine, where R=CH_3, is among this family of compounds and was the first chemotherapeutic drug to be used on patients.[9] The

mechanism shows how these compounds are able to cyclise and form a highly strained ring. Guanine at the N7 position is able to react and form a covalent bond.[10] This highly strained ring is a good driving force for the reaction; however, by also having a positively charged nitrogen group, this makes the aziridinum ion even more susceptible to ring opening. The nitrogen mustards have two reactive functional groups so can cyclise and alkylate another guanine base, hence the final compound is able to bind twice to the strands of DNA in the helix.



***Scheme 4.1**: A figure showing the mechanism of the cyclisation of nitrogen mustards and how this reacts with DNA to form intrastrand crosslinks*

The nitrogen mustards are shown to bind favourably to 5'-GXC-3' where X is any of the 4 DNA bases.[10] As the DNA is now bound in the complex, it is unable to replicate since the DNA strands are held together by a covalent bond, leading to cell death.

A compound that can perform both intrastrand and interstrand crosslinking is cisplatin. Despite being an FDA approved drug since 1978, the mode of action of cisplatin is still not fully understood.[11] Cisplatin has two chloride groups in the cis position, which act as leaving groups. While in circulation, the cisplatin remains in its square planar complex, due to high concentrations of chloride ions in blood plasma.[11] However, once inside the cell, this concentration decreases and the chloride ligands are replaced by water. The complex is then able to bind to DNA. As this compound is bifunctional, this happens twice and can form inter- or intra- strand crosslinks (**Figure 4.2**).

Cisplatin mostly acts on adjacent guanines in N7 position and often forms 1,2-intrastrand crosslinks but also 1,3 interstrand crosslinks. When cisplatin forms interstrand crosslinks at 5'GC there is severe distortion of DNA structure. [12]

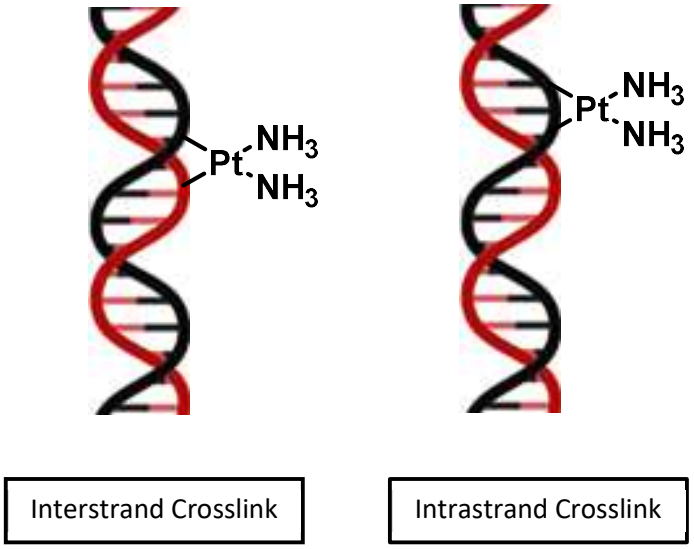


***Figure 4.2**: Illustration of the ability of cisplatin to form both interstrand and intrastand crosslinks with DNA*

This neutral square planar complex has a *cis* arrangement, and studies have shown this arrangement produces only 2% of interstrand adducts. Studies on the *trans* isomer have shown there are approximately 4% of interstrand adducts formed.[4] Although there is a majority of intrastrand crosslinks, the cause of the cytotoxicity is unknown, and some studies have shown that the interstrand crosslinks result in high potency.[13]

4.1.3 DNA binding with the duocarmycins

The alkylation mechanism of the duocarmycins was introduced in **Chapter 1** (**Scheme 1.1**) and this was shown mechanistically and schematically. The discussion of monomeric analogues emphasised that stereochemistry determined binding. When investigating dimeric payloads and designing assays to explore their ability to alkylate DNA and explore crosslinking, the monomers must fully be understood first.

Boger *et al.* carried out comparative studies on the two enantiomers of the duocarmycin compounds as a way to understand the structural reasons for sequence selectivity.[14] Duocarmycin SA alkylates the 3'-terminal adenine of an AT rich

sequence, whereas unnatural enantiomers bind in the reverse orientation and alkylate at the 5'-end.[15] Studies on different sequences found that for duocarmycin SA, preferential alkylation sites were: 5'-AA**A** > 5'-TT**A** > 5'-TA**A** > 5'-AT**A**. This was true for both the natural and unnatural enantiomers, but the binding orientations were reversed so that the adenine-N3 was able to bind to the least substituted cyclopropane carbon.[14] **Figure 4.3.4** shows the stick and space models taken from the referenced paper.[16] This diagram highlights the difference in binding between the natural (+)-duocarmycin SA and unnatural (-)-duocarmycin SA in the same minor groove region.

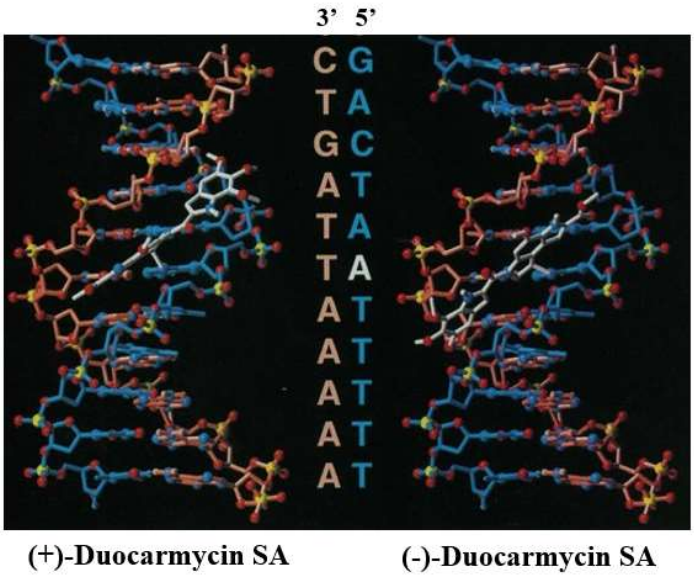


Figure 4.3: *Stick and space models showing the binding and alkylation direction of by (+)-duocarmycin SA and (-)-duocarmycin SA with w794 DNA.*

It is important to note that it is not just the conformation of the different enantiomers that determines selectivity. The duocarmycins are planar and curved, and they twist slightly as they fit into the helical shape of DNA.[17] The activation of the compounds is induced when they are in the minor groove as there is a binding-induced conformational change.[18] This concept was introduced in **Chapter 1,** which describes the difference between the different duocarmycin analogues and how their shape and hence the conformational change affects their potency.

Boger *et al.* published several reports showing how the arrangement of duocarmycin units and analogues affects their potency. For example, they developed 50 analogues to determine whether having the alkylating subunit in the centre of binding units in a

'sandwiched' arrangement increased the binding induced conformational twist, which, therefore, increased the rate and potency of DNA alkylation.[19] Boger and Garbaccio also investigated the angle of the twist in varying duocarmycin structures using X-ray structures.[20] They were investigating the angle of twist of the compounds between the alkylating group and its *N*-substituent once bound to DNA to determine how disruption to the compounds was affected by their structure. While looking at CC-1065, duocarmycin A and duocarmycin SA, it was found that the latter had the greatest conformational twist, with an angle estimated at 22.1°, and that the angle of the twist corresponded to the potency of the compound.[20] The shape of the compounds is therefore also important, and binding two alkylating subunits together to form a dimer with varying linkers will have some effect on this binding-induced confirmational change, and consequently their potency.

The sequence around the AT rich region has also been shown to affect the binding affinity of these compounds. Deep penetration into the minor groove is needed for alkylation. AT rather than GC sequences either side of the binding adenine are predicted to improve non-covalent binding through van der Waals interactions.[20] Non-covalent binding is important due because alkylation of the compounds with DNA is a reversible reaction.

Binding, however, does not always follow an exact pattern and, although duocarmycin analogues with varying stereochemistry tend to bind in opposite directions, they may bind in the same direction. **Figure 4.4** shows duocarmycin SA bound in the same binding site but with both the natural and the unnatural enantiomer, generated using PyMOL software.[21] Despite binding in the same site (AATTA), there is a difference in potency, and the cause of this is highlighted in red. The alkylating units are pointing in the opposite directions, resulting in alkylation sites on opposite DNA strands and differing by one base pair.

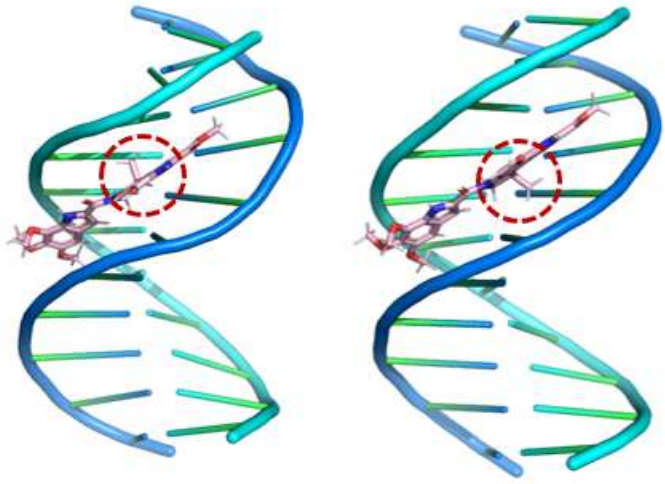


Figure 4.4: *Binding in DNA, same binding site opposite strands, image generated using PyMOL software.*

These are just some of the factors of monomeric duocarmycins that influence binding and potency. When two subunits are bound together to form a dimer, new considerations must be understood. The introduction of dimeric duocarmycin payloads was touched on in **Chapter 3**. To summarise, studies showed that it is the left-hand subunit that controls the first site of DNA alkylation.[22] The natural enantiomer is, on average, 10 times more potent than the unnatural, and this is reflected in the dimers bearing the natural enantiomer as the left-hand subunit being more potent than those with the unnatural enantiomer.[22,23]

Duocarmycin SA alkylating subunits bound directly to one another were shown to be able to form intra- and interstrand crosslinks.[22] Interestingly, the stereochemistry determined the type of crosslink. Cyclopropanes in opposite directions showed interstrand crosslinking, while units with the same enantiomers were only able to form intrastrand DNA crosslinks.[22]

There are a number of other examples of hetero-bifunctional or hybrid bifunctional alkylating agents utilising the alkylation units of the duocarmycins, for example, pyrrolobenzodiazepine (PBD) – cyclopropapyyroloindole (CPI) dimers. By incorporating a PBD unit, these compounds bind to guanine in GC rich regions, by using dual ‘warheads’, they are able to crosslink to G and A bases.[24]

There are many different properties and factors that may slightly alter where and how well these compounds bind to DNA. Studies have shown that the length and distance of cross-link affects the stability of the DNA.[25] Having variations in linker length, and potentially crosslinking distance, may also affect potency. This project explores some preliminary studies on how the dimers synthesised may vary in potency and utilises a number of biological assays to understand how they bind to DNA. Four assays were carried out with the compounds: MTS assay, gold nanoparticle assay and two gel electrophoresis assays. Potency was calculated using the MTS cell killing assay to determine IC_{50} values. The gold nanoparticle assay will show if certain dimers are able to form duplex-duplex crosslinks. Gel electrophoresis will possibly give some more information on how these compounds bind to and crosslink DNA. Before doing any of the biological assays, the compounds must be activated by removing protecting groups.

4.2. Removing protecting groups for biological evaluation

All the dimers and control monomers had been synthesised; however, to activate these compounds, the protecting groups needed to be removed. This was carried out using palladium on carbon (Pd/C) and ammonium formate. After removing protecting groups these compounds are known to be cytotoxic, therefore these compounds were handled with great care. The project aimed to deprotect all compounds that were synthesised in **Chapter 3**. **Figure 4.5** shows these compounds numbered **4.1 – 4.7** with the key functional group changes and differences highlighted in red.

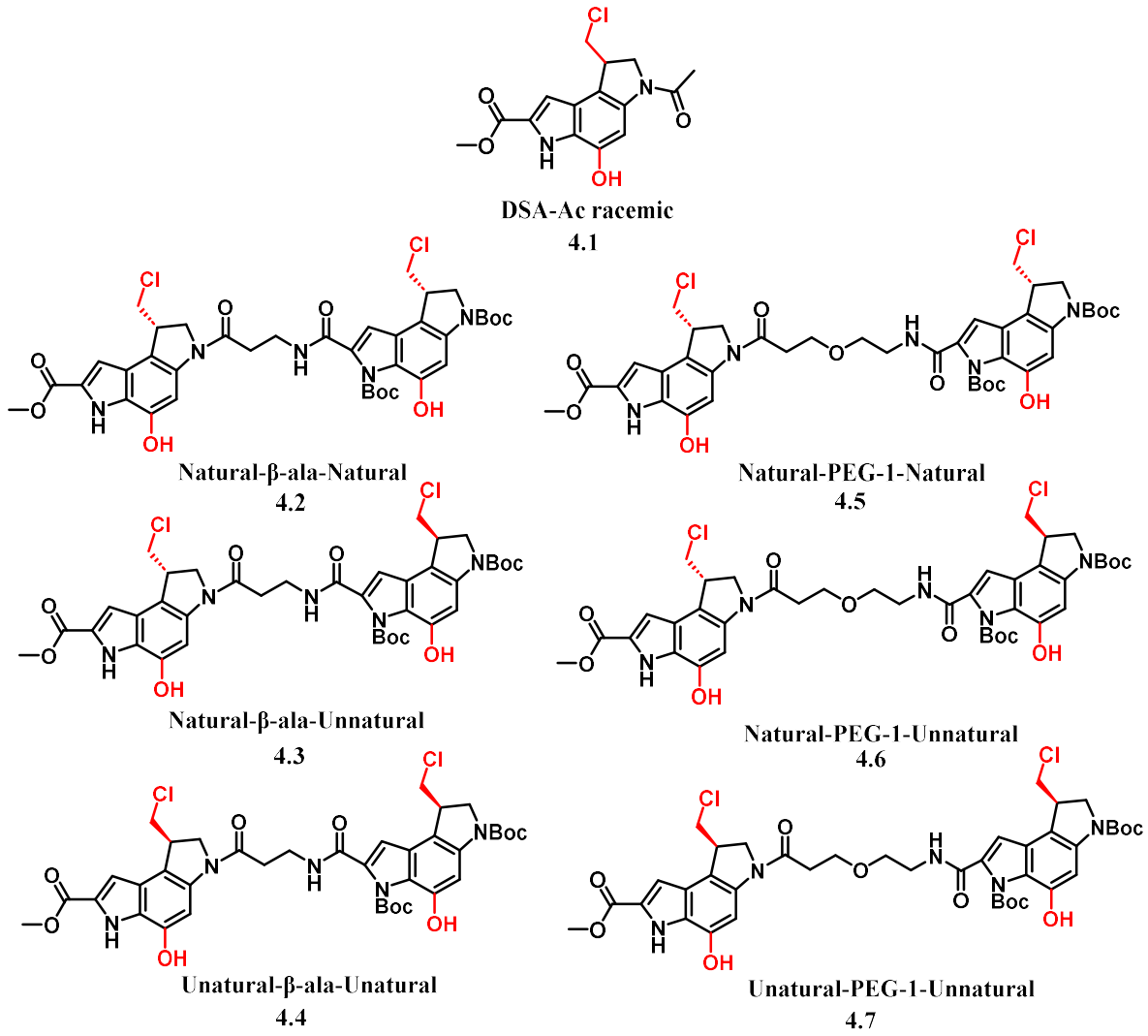


Figure 4.5: *Structures of the seven toxic compounds this project aimed to synthesise.*

The same conditions to deprotect were used for all compounds. The general scheme for this is shown in **Figure 4.6**. The debenzylated compounds were confirmed with a single HPLC peak and a correct HRMS and LC-MS. NMR investigations were also carried out and will be discussed; however, due to their toxicity, only small amounts were synthesised, so full characterisation by NMR was not possible.

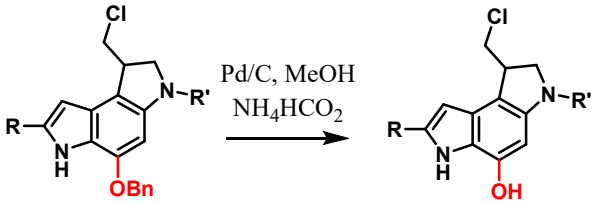


Figure 4.6: *Conditions for benzyl deprotection on generalised subunit.*

This debenzylation reaction uses a palladium on carbon catalyst and ammonium formate to reduce the benzyl group and generate a free alcohol on the DSA subunit. The reaction works by a catalytic transfer hydrogenation. The ammonium formate decomposes in the presence of the palladium catalyst, forming byproducts including hydrogen gas. A hydride transfer generates the benzyl deprotected product. The reaction allows selective removal of the benzyl ether.

The conditions for all debenzylations were the same. The benzyl protected DSA analogues were dissolved in MeOH and degassed with nitrogen. Pd/C was then added before further degassing. A solution of 25% ammonium formate in distilled water was then added and the reaction was vigorously stirred. THF was often added after the ammonium formate addition, due to product insolubility. The reactions were all monitored closely using HPLC and LC-MS with a UV trace attachment. Reactions were worked up between 1.5 and 3 hours after the reactions were started. The workups were carried out by filtering the palladium residue, then adding water to the reaction mixture, before extracting three times with EtOAc. The combined organic layers were then concentrated under reduced pressure before being subjected to different purification techniques.

The yields, purities, and mass spectral data for these compounds are shown in the following table, **Table 4.1**. **Figure 4.7** shows the HPLC traces for compounds **4.1 – 4.7**. The purifications varied depending on the compounds and yields and data will be described in the experimental. The mass of the product was confirmed by high-resolution mass spectrometry, see **Chapter 5** Experimental for details.

Compound Number	Linker	Stereochemistry	Yield (mg, %)	Purity (HPLC)	Mass spec
4.1	N/A	Racemic monomer	ND	93%	HRMS
4.2	β-Ala	Natural-natural	3.2 mg 99%	81%	HRMS
4.3	β-Ala	Natural-unnatural	2.8 mg 64%	91%	HRMS
4.4	β-Ala	Unnatural-unnatural	1.4 mg 39%	80%	HRMS
4.5	PEG-1	Natural-natural	3.3 mg 68%	85%	HRMS
4.6	PEG-1	Natural-unnatural	6.4 mg 98%	75%	HRMS
4.7	PEG-1	Unnatural-unnatural	1.8 mg 74%	99%	HRMS

Table 4.1: *A summary of the results from the debenzylations of compounds* ***4.1-4.7,*** *showing percentage yield and purity.*

As can be seen from **Table 4.1**, the yields from the deprotection varied across the different compounds, although this might be expected when working on such small amounts of material. The toxic nature of the resulting compounds, which may show very high cytotoxic potential, mean it is essential to keep amounts deprotected to a minimum. The purities, overall, were in excess of 80% according to the HPLC data, with the exception of the PEG-linked natural-unnatural dimer, which, while the target compound was clearly the major product, proved difficult to get to higher levels of purity.

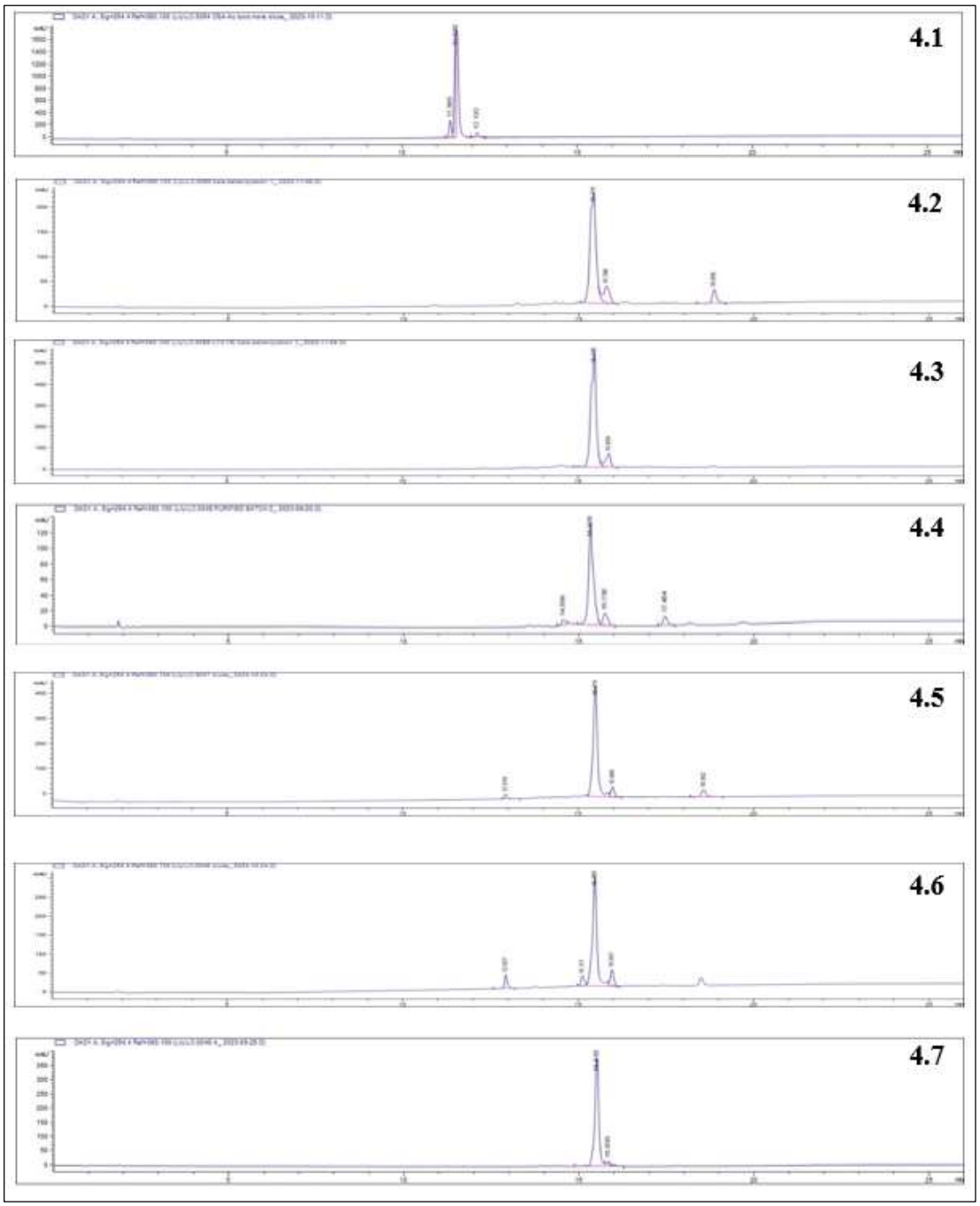


Figure 4.7: *Analytical HPLC trace recorded at 214 nm showing compounds* ***4.1-4.7*** *at 75°C (Agilent eclipse XDB-C18 column (4.6 x 150 mm, 5 µm) and a flow rate of 1 mL/min. Spectra were run with a solvent gradient of 0-100% B over 20 min. Solvent A: H2O, 0.05% TFA, solvent B: MeCN, 0.05% TFA)*

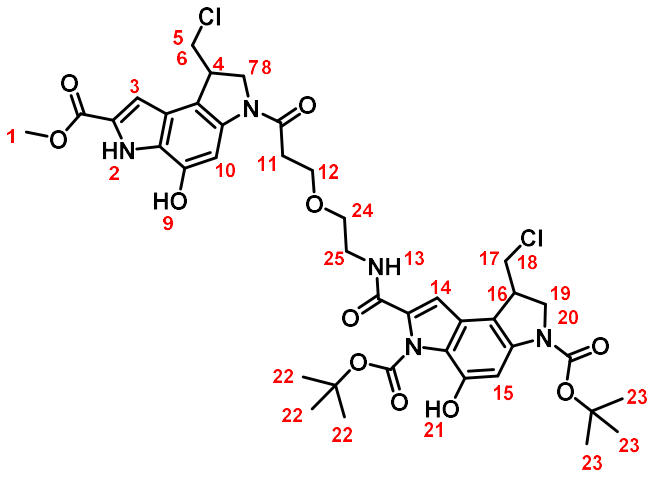
4.2.1. NMR investigations

To further characterise these products, proton and carbon NMR were investigated. As biological assays could be carried out with the compounds dissolved in deuterated

DMSO of known concentration, it was possible to subject the compounds to NMR in this solvent. Success in these investigations was limited to the small quantities available and only the PEG dimer gave some limited success.

Interpretation of the proton NMR was challenging due to the two DSA compounds with very similar environments having overlapping peaks. Attempts were made on the 800 MHz NMR with experiments running over a range of times. To get clearly defined peaks, the ^{1}H NMR, COSY, HSQC and HMBC were run over 18 hours. Despite this, and with expert input from Dr Ignacio Delso, an NMR specialist, the resolution could not be improved with the quantities of sample provided, therefore full compound assignment and characterisation was not possible.

The most success was seen with the natural-PEG-1-natural compound, **4.5**, and the findings for these experiments are summarised in **Table 4.2**. The different environments are numbered and shown in red.



Environment	Shift and assignment / ppm
C1	3.85 ppm, singlet, 3H
C2	11.56 ppm, singlet, 1H
C3	7.23 ppm, singlet, 1H = **3** or **15**
C4	4.81 ppm, doublet of doublets, 2H
C5	*Too much overlap*
C6	*Too much overlap*
C7	*Too much overlap*
C8	*Too much overlap*
C9	11.19 ppm, singlet 1H = **9** or **21**
C10	7.08 ppm, singlet, 1H = **10**
C11	*Not assigned*
C12	*Not assigned*
C13	8.22 ppm, triplet, 1H
C14	*Too much overlap*
C15	7.23 ppm, singlet, 1H = **3** or **15**
C16	*Too much overlap*
C17	*Too much overlap*
C18	*Too much overlap*
C19	*Too much overlap*
C20	*Too much overlap*
C21	11.19 ppm, singlet 1H **9** or **21**
C22	1.42 ppm, singlet, 9H
C23	1.48 ppm, singlet 9H
C24	*Not assigned*
C25	*Not assigned*

Table 4.2: *Assignments of protons of compound* ***4.5*** *to the different environments, as highlighted in* ***Figure 4.8***

All four pairs of diastereotopic CH_2 were accounted for; **C11**, **C12**, **C14** and **C25**, however, these could not be assigned. The environments were observed at: 4.32 ppm as a triplet integrating for two protons, 3.47 ppm as a multiplet integrating for two protons, 3.70 ppm as a multiplet integrating for two protons and 2.96 ppm as a doublet of doublets integrating for two protons.

With purity in excess of 80%, alongside the correct LC-MS and HRMS data, these compounds were then subjected to biological testing with some confidence that the correct compound structures were being tested. Despite not being able to fully characterise the NMR due to similar environments and limited product, there is clear evidence that both groups have coupled together with the correct linker.

4.3. Anti-proliferative activity – MTS assay

The MTS assay is used to assess cell viability, cell proliferation and cytotoxicity. MTS, 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, is a yellow dye that is used in cell proliferation assays due to its change in colour once metabolised by cells. Cells that are proliferating have a mitochondrial dehydrogenase enzymes, which are able to reduce the yellow MTS into a purple coloured formazan product (**Figure 4.9**).[26]

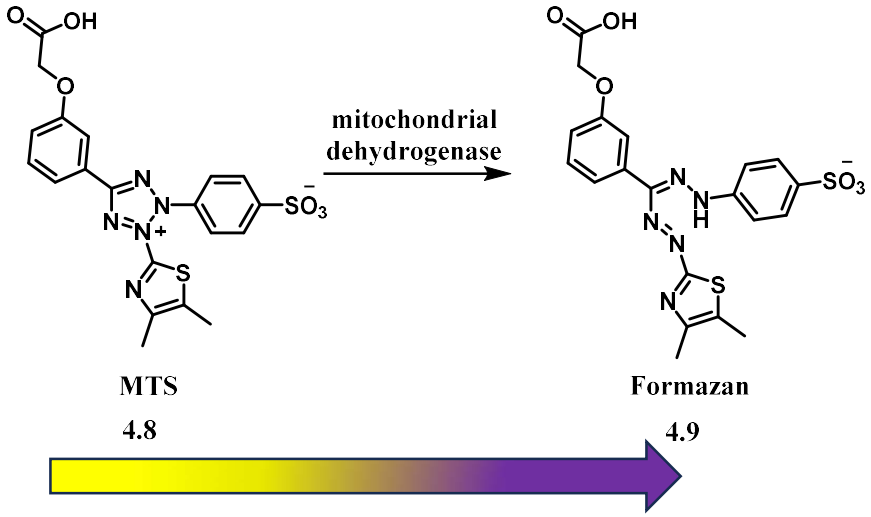


Figure 4.9: *Figure showing the change in chemical structure of the MTS once metabolised by mitochondrial dehydrogenase into the formazan from a yellow solution to a purple solution dependent on concentration*

This colour change is the key feature of this cell assay. The absorbance is measured at 492 nm and can be correlated with cell proliferation. Positive and negative controls are used as a standard to read against the samples when the absorbance is measured. The concentration of the coloured formazan product is directly proportional to the number of viable cells in that well, and, therefore, a gradient can be run to determine potency of a compound.

The assay used a human leukemia cell line, HL-60. This is a cell line that has been used previously to study the duocarmycins.[27,28] The positive control for these experiments was the antitumour agent doxorubicin, as this is known to cause cell death in this cell line. The negative control was a blank addition of DMSO. The plates were read on a microplate reader and the data was analysed to generate IC_{50} values for each compound. The experiments were done in triplicate, this meaning that experiments

were done at least three times on different days with different stock solutions and different cell passages. An example of one of the plates after the cells had been treated with the MTS dye is shown in **Figure 4.10**.

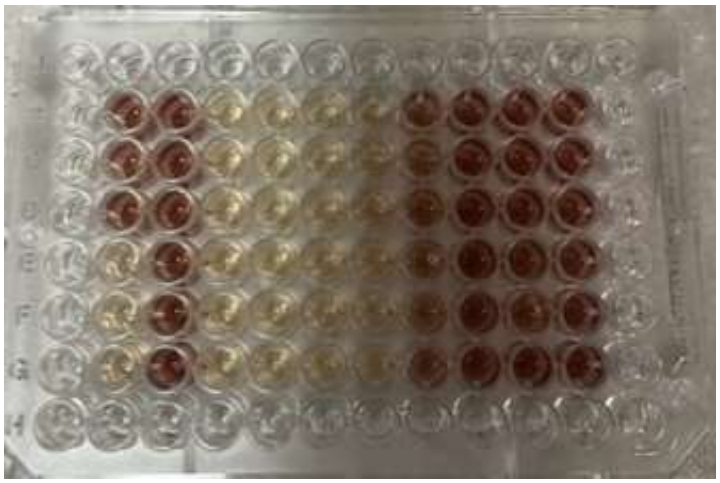


Figure 4.10: *A plate showing MTS results for compounds* ***4.5*** *(top) and compound* ***4.6*** *(bottom) in HL-60 cell line. DMSO control top, doxorubicin control bottom*

Table 4.3 summarises the compounds tested and their allocated compound number, as well as the results from the assay.

Compound number	Stereochemistry	Linker	IC_{50}/nM
4.1	Racemic	Monomer	984
4.2	Natural-natural	β-Ala	239
4.3	Natural-unnatural	β-Ala	26
4.4	Unnatural-unnatural	β-Ala	1131
4.5	Natural-natural	PEG	558
4.6	Natural-unnatural	PEG	664
4.7	Unnatural-unnatural	PEG	2090

Table 4.3: *Compound 4.1-4.7 and their results from the MTS assay*

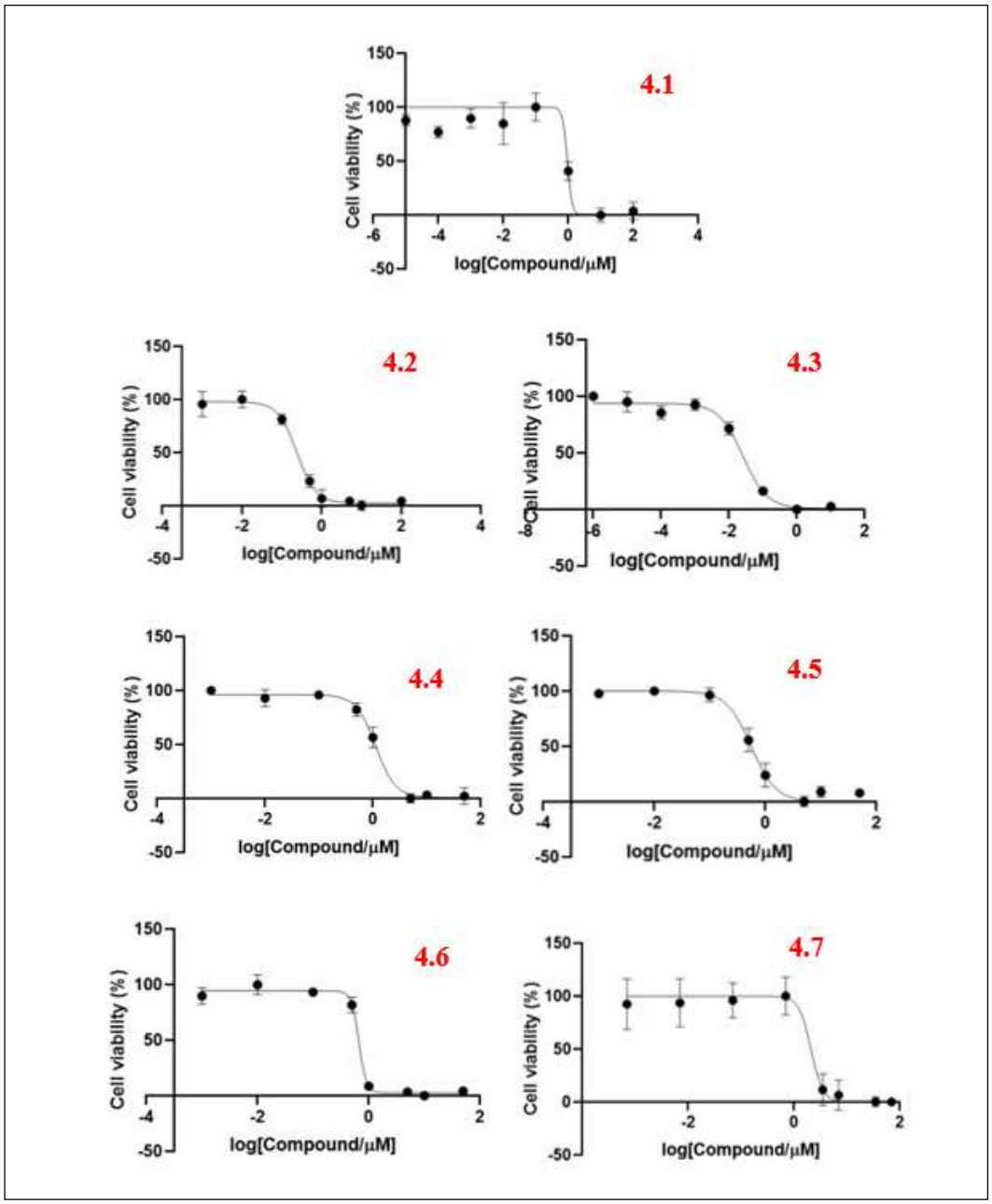


Figure 4.11: *Data obtained from MTS assay from compounds* ***4.1-4.7****. Values obtained using a BMG Labtech POLARstarOPTIMA microplate reader, reading at a 492 nm wavelength and plotted on GraphPad Prism Version 6.0, using a four-parameter logistic nonlinear regression model.*

There were six dimers designed and a monomer for direct comparison. From previous studies on the duocarmycins, it was hypothesised that the dimers starting with the natural DSA subunit would be more potent than both the monomer and the dimers starting with the unnatural subunit. It was also hypothesised that the monomer will be more potent than the dimers bearing the unnatural subunit on the left-hand side.

The assay showed that all seven compounds had anti-proliferative activity in the HL60 cell line, ranging from 2000 to 26 nM in activity. The most potent compounds in this study were the compounds in which the natural subunit is on the left-hand side of the dimer (compounds **4.2** and **4.5**), as expected from previous studies. These compounds are on average six times more potent than the dimers with the unnatural stereochemistry subunit in this position. As expected, the racemic monomer was also more active than the dimers with the unnatural left-hand subunit.

The dimers with the *β*-alanine linker were generally more potent than those with the PEG linker. This may be an indication that the more hydrophobic nature of dimers with the *β*-alanine linker mimics the hydrophobic nature of most minor groove binding ligands more efficiently than the ether linkage. The narrow deep minor groove is efficiently bound by extended aromatic compounds, such as duocarmycin SA and distamycin A, in A-T rich regions in a thermodynamically favoured process where the spine of hydration along the groove is displaced which may be mirrored by the more hydrophobic linker.

Interestingly, of the two PEG dimers, it was the one bearing both natural enantiomers (**4.5,** IC_{50} = 558 nm) that was the most potent. This varied from the dimer with the *β*-alanine, where the dimer with the natural and unnatural subunit, **4.3** (IC_{50} = 26 nm), showed the highest potency against the cell line. The most potent compound in this series, compound **4.3**, showing low nanomolar potency, was the natural - *β*-alanine - unnatural dimer. This toxicity could be due to the ability of the compound to form DNA crosslinks, although this was not, as noted, replicated in the PEG series. The following assays studied the ability of the compounds to crosslink DNA, both forming duplex to duplex crosslinks and also intermolecular crosslinks in a single DNA strand.

4.4. Duplex to Duplex crosslinking assay – A gold nanoparticles assay

The dimers synthesised in this project have two independent reactive groups. This is a key feature of a compound that may be able to crosslink DNA via intermolecular crosslinks in a single DNA strand or via DNA to DNA crosslinking. The first assay in this project to investigate DNA duplex to duplex crosslinking used gold nanoparticles.

4.4.1 Introduction to nanoparticles and nanomedicines

A nanoparticle is a material that is measures 1-100 nm in size.[29] These small particles can be made of different materials. Gold is one of the least reactive metals and has intrinsic properties that mean, as a nanoparticle, it has been studied by physicists, biologist and chemists. Gold nanoparticles, AuNPs, have been used in many areas of science and healthcare, including targeting drugs for therapy, bioimaging, catalysts, and in various biological assays.[30] The high surface area to volume ratio of nanoparticles means they have increased surface reactivity and also that they can interact with surrounding environments. These physical and chemical properties change as the size of the nanoparticles increases. In aqueous solutions, gold nanoparticles below 100 nm have an intense red colour; this colour changes when the particles aggregate and become larger, when they start to change to a blue or purple colour.[31]

Gold nanoparticles can be functionalised with various molecules, allowing various applications.[32] Marin *et al*. exploited this property to investigate duplex to duplex interactions in bis-acridines.[33] It was shown that these compounds could bind to two separate DNA stands and demonstrate AuNP aggregation. This aggregation of the AuNPs was followed using UV-Vis absorbance spectroscopy, where there is a colour change on aggregation of the gold nanoparticles if crosslinking is occuring. A simplified illustration for this assay is shown in **Figure 4.12.**

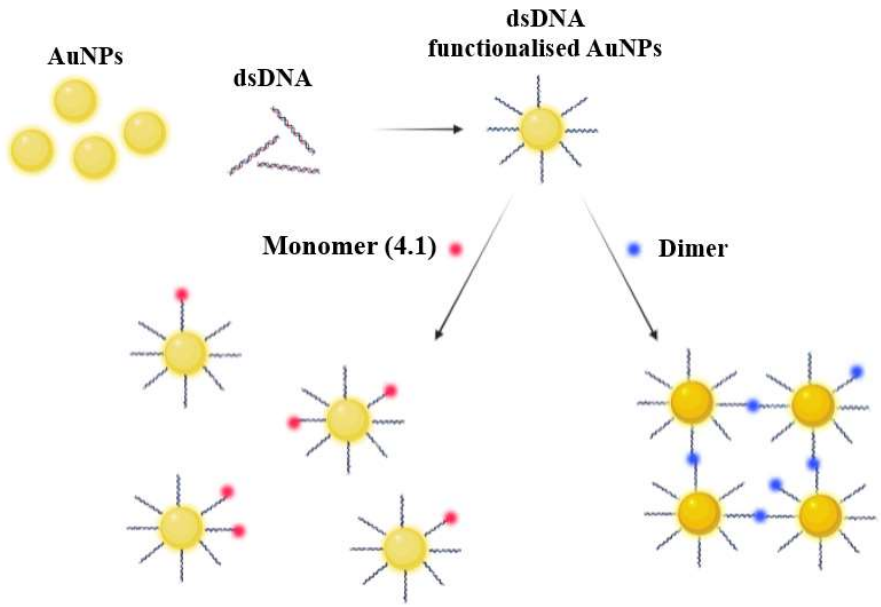


Figure 4.12: *An illustration of gold nanoparticle aggregation with a dimer vs non-aggregation with a monomer*

Figure 4.13 provides an example of a positive trace from UV-Vis absorption spectroscopy, demonstrating aggregation.[33] The initial spectrum trace, labelled 1, shows the absorbance of the starting mixture where the gold nanoparticles are not aggregated. Peak 2 shows the expected trace for a compound that is able to crosslink duplex-duplex DNA strands. Peak 3 shows the expected trace for a compound that is not able to crosslink, here the peak stays at the same wavelength and doesn't have the same characteristic shift right towards the higher wavelength.

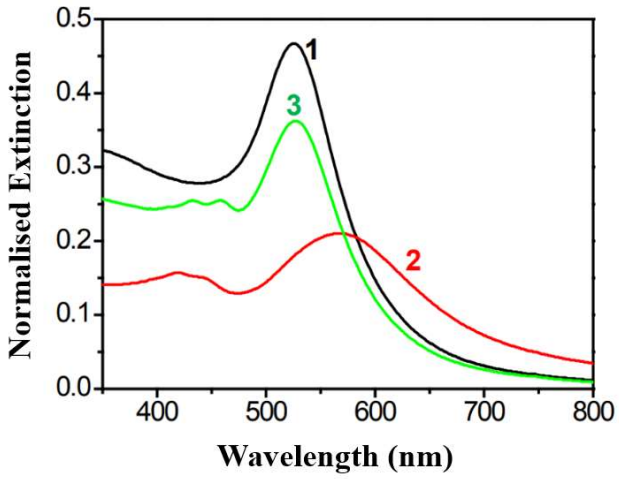


Figure 4.13: *A graph showing the expected traces for a positive and negative result of duplex-duplex crosslinking of DNA as seen by UV-Vis.*[30]

This assay was used to investigate the duplex-duplex crosslinking ability of the duocarmycin dimers, with a simple diacridine-4-carboxamide, synthesised by Victoria Kamperi, a PhD student in the laboratory, as a positive control.[33] The assay can be broken down into several steps before the crosslinking assay could be done and these are summarised in the following sections.

4.4.2 Preparation of the gold nanoparticles, AuNPs

The first step was the synthesis of the gold nanoparticles, using a synthesis developed by Enüstün and Turkevich via a citrate reduction.[34] The nanoparticles were prepared using solutions of gold (III) chloride trihydrate in water and trisodium citrate in water, which were combined and heated. After cooling, the characteristic red colour of the nanoparticles was observed and these were collected after filtration and stored cold for future use.

4.4.3 Functionalisation of the AuNPs with thiolated DNA

The functionalisation of the AuNPs followed a procedure that had been previously reported by Zhang.[35] Functionalisation was confirmed using a NaCl test, in which NaCl was added to the AuNPs, and would only cause aggregation to non-functionalised nanoparticles. This can be observed visually and also using UV-Vis spectroscopy. The functionalised AuNPs would have no change with NaCl, however, if they had not been correctly functionalised with the DNA, then aggregation and the characteristic colour change would occur.

4.4.4 Preparation of the double-stranded DNA

In this assay there were two different double-stranded DNA strands investigated. The DNA has one strand containing a thiol group to enable binding to the AuNPs and forming a strong covalent bond.[36] The preliminary studies of this assay were carried out using a DNA sequence that was already available in the lab. The sequences of the oligonucleotides were as follows:

DNA sequence 1:

5’ - [ThiC6] - CTACGTGGACCTGGAGAGAGGAAGGAGACTGCCTG – 3’

3' – GATGCACCTGGACCTCTCTCCTTCCTCTGACGGAC – 5'

These sequences are rich in guanine and cytosine; there are, however, some adenine and thymine tracts available. As discussed previously, the duocarmycins bind preferentially into AT-rich regions. For this reason, another thiolated DNA sequence was designed that contained AT sequences at the end furthest from the thiolated end, away from the AuNPs and potentially accessible to the duocarmycin dimers. The following two single-stranded oligonucleotides were designed:

DNA sequence 2:

5' – [ThiC6] – CGACGCGCGCGAGCTTAA – 3'

3' – GCTGCGCGCGCTCGAATT – 5'

These complementary sequences were annealed to give 35 and 18 base pair duplex DNA sequences, respectively. **Figure 4.14** shows how this annealing process works: the oligomers are combined in a buffer solution and heated to 90 °C to denature any complexes the single strands may have formed. They are then allowed to cool slowly to room temperature overnight which allows hybridisation as new hydrogen bonds form between the complementary strands.

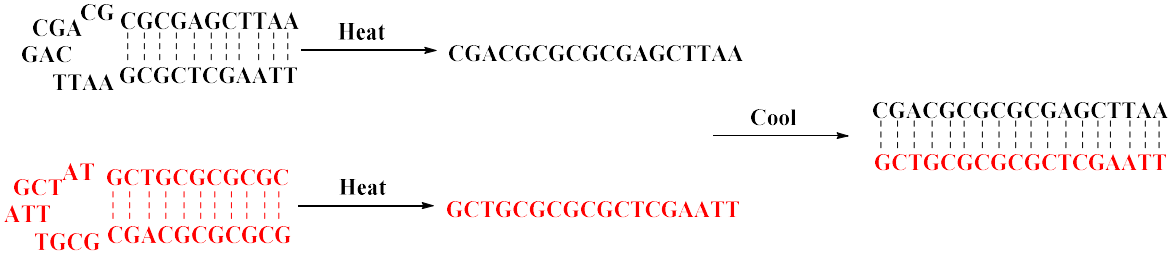
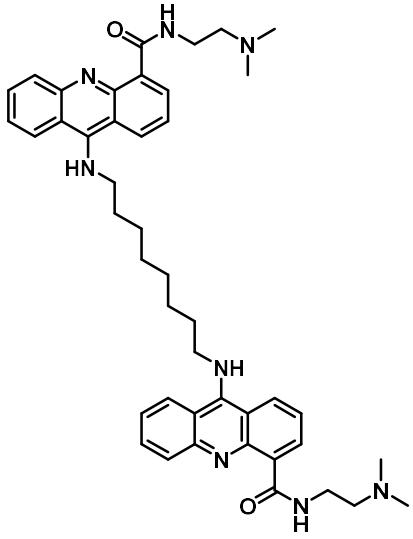


Figure 4.14: *A visual representation of how the annealing of two single-stranded oligonucleotides form a double-stranded DNA*

4.4.5 Duplex – duplex crosslinking Assay

The dimer was assessed alongside positive and negative controls. The negative control in this assay was the monomer **4.1**. This compound is not able to form these duplex-duplex crosslinks, as there is only a single unit able to form a bond with DNA. The positive control for this experiment was the acridine bis-intercalator, shown in **Figure 4.15**, which had shown a positive result previously in this assay. These compounds are

cytotoxic due to their flat planar structure and they are able to intercalate between base pairs in DNA.[37] 9-Aminoacridine has been investigated for use as an anticancer chemotherapeutic due to its mode of action, which not only includes intercalation causing a stability change and unwinding of DNA, but also its ability to initiate reactive oxygen species and topoisomerase inhibition.[38] A bifunctional agent can increase the potency of this compound by binding to two sides in the DNA duplex.[37] By varying the linkers, it was shown that acridines are able to bind to different DNA complexes and form duplex-duplex interstrand crosslinks.[33,39]



4.10

Figure 4.15: *The bis-acridine that is a known duplex-duplex DNA crosslinker*

An initial study was carried out with the unnatural-PEG-1-unnatural dimer **4.7**, to see if there was any evidence of it crosslinking DNA sequence 1 that was available already within the laboratory. No evidence for crosslinking was seen with the sequence.

With DNA sequence 2, the natural-PEG-natural compound **4.5** was investigated. The PEG dimer was chosen due to its solubility and slightly longer linker. By increasing the distance between the two alkylating units, it was hypothesised that DNA duplex to duplex crosslinking would be enhanced. An initial reading of DMSO in the gold nanoparticles was carried out and then the compounds were added, titrated in increasing concentration. The UV-Vis absorption spectrum of dsDNA functionalised gold nanoparticles was recorded after increasing concentrations of negative control (DMSO), positive control (known acridine crosslinker), and the duocarmycin dimer.

The experiments were run for 10 titrations and aggregation could be detected by a characteristic red shift, right, and a decrease in absorbance. Four crosslinking assays are summarised in **Table 4.4**.

Name	Compound	Sequence	Observation
Preliminary study	**4.7**	DNA sequence 1	No aggregation
Negative control (monomer)	**4.1**	DNA sequence 2	No aggregation
Positive control (acridine)	**4.10**	DNA sequence 2	Aggregation
Dimer result	**4.5**	DNA sequence 2	No aggregation

Table 4.4: *A summary of the observations found by the AuNP assay by the different compounds on the custom oligomer sequences*

The graphs in **Figure 4.16** show that in the regions 380-490 nm and 300-400 nm there is an increase in intensity recorded. This is because the compounds being titrated absorb light at these wavelengths and are being added in increasing concentrations.

This assay successfully synthesised functionalised gold nanoparticles with two different DNA strands and tested them on a series of compounds. The assay correctly showed the expected shift for the negative control. The addition of the monomer had no effect on the colour of the AuNPs, shown as there was no change in the peak shift on the UV-Vis extinction spectrum. The positive control, acridine, correctly showed the shift in the spectrum peak down and towards the right, towards the red shift; this was the colorimetric change of the gold nanoparticles aggregating.

The preliminary study of the unnatural-PEG-unnatural dimer with the initial DNA sequence showed that this compound was unable to form duplex-duplex crosslinks, as the spectrum shows no shift in wavelength.

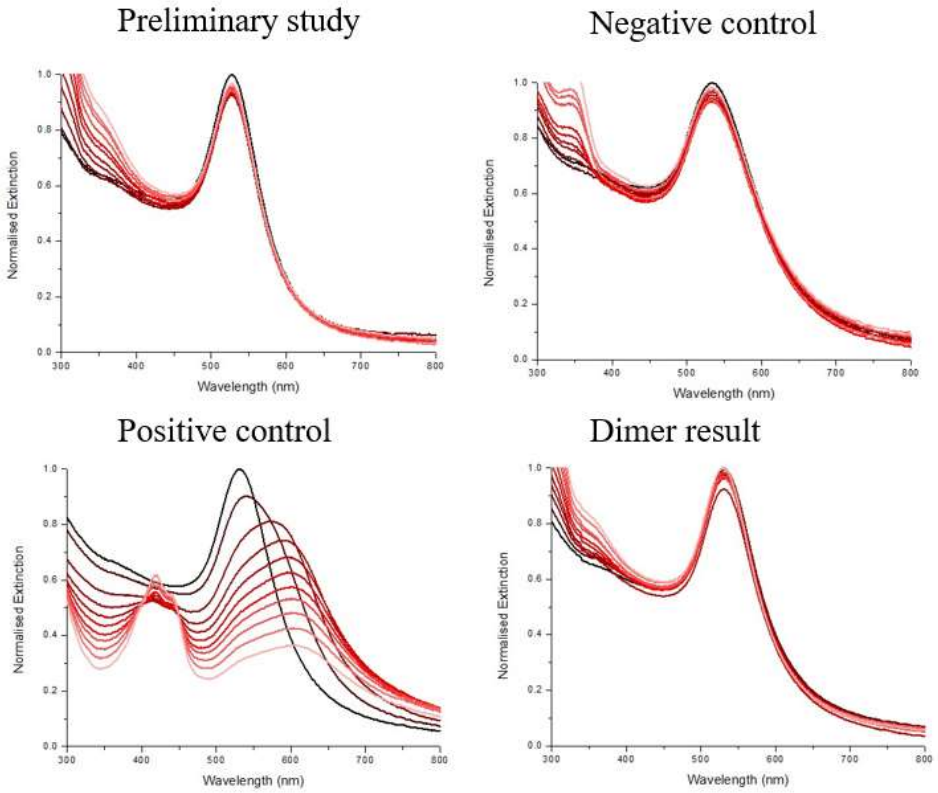


Figure 4.16: *UV-Vis extinction spectrum of dsDNA functionalised gold nanoparticles after increasing concentrations of negative control, no compound, positive control, known acridine crosslinker, and PEG dimers with varying stereochemistries.*

The rationale in this project was that a DNA sequence rich in adenine and thymine and the natural-PEG-natural dimer would be able to form duplex-duplex crosslinks. However, on titration of this dimer to the AuNPs, there was no evidence of aggregation. This suggests that these compounds are not able to form duplex to duplex crosslinks in the conditions of the AuNP assay. To confirm this observation, we designed and investigated a further, gel-based, assay using small, designed oligonucleotides.

4.5. Gel Electrophoresis to investigate DNA crosslinking

Gel electrophoresis is a technique where molecules such as DNA can be separated based on size and charge.[40] All types of gel electrophoresis have roughly have the same general procedure, which is outlined in **Figure 4.17** The samples are prepared before loading onto the gel using a loading dye. Electrodes are attached to the gel tank. The anode is attached to the end closest to where the samples are loaded and the cathode is attached to the opposite end. Negatively charged DNA travels down the plate towards the cathode, and there is a separation based on DNA fragment size, as lighter chains will travel faster and further. The loading dye will help determine when the gel run is complete and can be analysed under a specific wavelength to see the fluorescently tagged compounds being run.

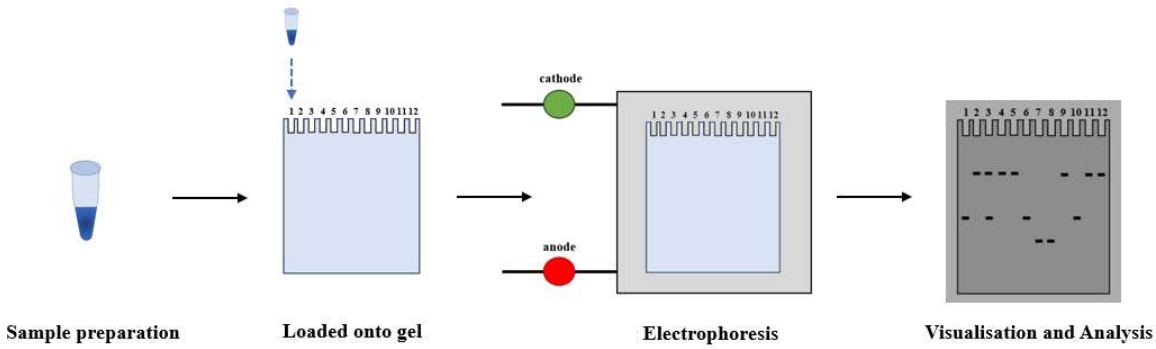


***Figure 4.17**: The general method for gel electrophoresis.*

Although there are similar general procedures, the types of gel, buffers, run conditions, and what is being tested can all vary significantly. Therefore, trying to find appropriate conditions to run compounds on the gel and get a clear result may take some adjustment.

In this project, a 10% polyacrylamide gel was used. This gel should be able to show clearly if there are separations in the double-strand DNA fragments. There were two types of assays which we aimed to investigate using gel electrophoresis: a thermal cleavage assay and a crosslinking assay. For each of these assays a different sequence of oligomers was needed. Four custom single-stranded oligonucleotides were ordered: the thermal cleavage assay where the AT sequence was central to the strand, and a

DNA crosslinking assay, where the AT was at the terminus. One strand of each complementary oligonucleotides was requested with a 6-Fam group at the 5' position.

4.5.1 Thermal cleavage assay – the central AT sequence

One way to prove compounds are able to bind to DNA is using a thermal cleavage assay. The binding of the DSA monomer or dimer to the DNA strand adds negligible weight therefore the band shift is not something that can be seen. However, if the duocarmycins are able to bind to the DNA strand and this strand is heated, then in theory the DNA strand will cleave where the duocarmycin has bound. This cleavage could theoretically be visualised on the gel, as there will be a band that travels further on the gel, as it should be smaller after cleavage. The hope is that both the monomer and the dimers should be successful in this assay, as they are able to bind to DNA, whilst the negative control in this assay would be the compounds incubated with the DNA without heating, as this should allow them to remain bound and not cleave the DNA.

The custom oligomers were designed with an AT sequence in the middle and one of the strands fluorescently labelled so visualisation of the bands could be done after the electrophoresis.
5' - [6-FAM] - CGACGCGTTAACGTTAACGCGC – 3'
3' – GCTGCGCAATTGCAATTGCGCG – 5'

The single-stranded oligomers were annealed overnight and then used the following day in the assay.

The compounds tested in the thermal cleavage assay were the monomer unit **4.1**, and the dimer unit, **4.5**. These were tested in several experiments, with concentrations ranging from 50 μM to 5 μM. After incubation, the compounds that were being tested for the cleavage were then heated at 90 °C for 3 minutes, before allowing them to cool. After cooling, all samples were stained with a blue tracking dye and loaded onto the gel. The gels showed the expected result of a band travelling faster through the gel after heating, **Figure 4.18**. However, no concentration dependence was observed,

suggesting that the band observed was not due to DNA cleavage, but was perhaps due to the denaturing of the DNA under the conditions of the experiment.

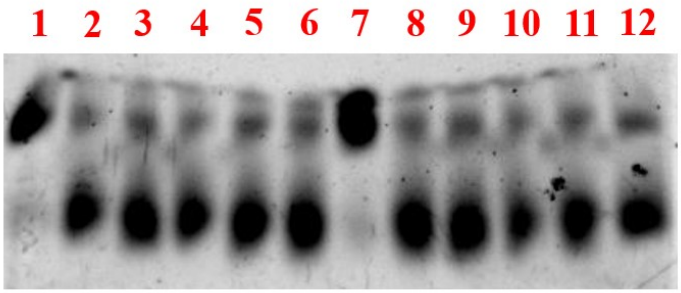


Figure 4.18: *Thermal cleavage assay of monomer* ***4.1*** *and dimer* ***4.5.*** *Wells 1-12, far left to far right. 1) monomer control 5 μM. 2) monomer heat 5 μM. 3) monomer heat 0.5 μM. 4) monomer heat 0.05 μM. 5) monomer heat 0.005 μM. 6) monomer heat 0.0005 μM. 7) dimer control 5 μM. 8) dimer heat 5 μM. 9) dimer heat 0.5 μM. 10) dimer heat 0.05 μM, 11) dimer heat 0.005 μM. 12) dimer heat 0.0005 μM*

Further tests, with much lower concentrations of the ligands, showed a similar effect, which further supported the idea that this was due to DNA denaturation, rather than cleavage. To test this hypothesis, a further experiment with a control lane of just the labelled oligonucleotide that had not been annealed was carried out. The results confirmed the two bands being observed were due to a mixture of the double- and single-stranded sequences.

To overcome this issue, the assay was repeated and, after heating the samples at 90 °C for 3 minutes, the samples were allowed to cool slowly overnight before staining and loading the gels. Despite several attempts, this was unsuccesful in showing any form of cleavage (**Figure 4.19**).

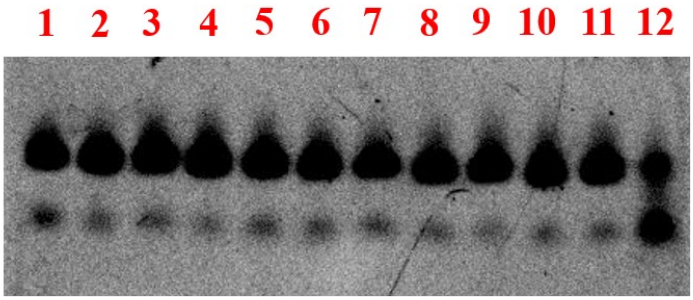


Figure 4.19: *Thermal cleavage assay with condition changes. Wells 1-12, far left to far right. 1) control DMSO and heated. 2) control monomer 5 μM. 3) monomer heat 0.5 μM. 4) monomer heat 5 μM. 5) monomer heat 10 μM. 6) monomer heat 50 μM. 7) control DMSO heat. 8) control dimer 5 μM. 9) dimer heat 0.5 μM. 10) dimer heat 5 μM, 11) dimer heat 10 μM. 12) dimer heat 50 μM*

These gels were re-attempted with this method using the PEG dimer with two varying stereochemistries, compound **4.5** and compound **4.6**. The samples were run with a DMSO control, and a range of concentrations were tried, however, there was no evidence of cleavage using this method. No further work was carried out on this assay.

4.5.2 Crosslinking assay – the terminal AT sequence

By having the AT sequence at the end of the strand, it was hypothesised that the dimer could bind to two separate duplex-duplex strands and form an end-to-end crosslink. If the dimer was able to bind to two differing strands, this could be visualised using gel electrophoresis, as this would generate a higher molecular weight oligonucleotide strand. This could be compared with the monomer or DMSO control, both unable to form crosslinks, and if the duplex was formed, it would be roughly twice the size and would travel less far along the gel. The two custom strand complementary oligomers ordered from Eurogentech had the following sequences:

5’ - [6-Fam] - CGACGCGCGCGAGCTTAA – 3’

3’ – GCTGCGCGCGCTCGAATT – 5’

Stock solutions of 100 μM were prepared for each oligonucleotide in buffer following the manufacturer’s guidelines: 10 mM Tris-HCl, pH 7.5, 10 mM NaCl.

No difference between the control samples and dimer samples was observed in the gels, indicating that either the dimer was unable to form these end-to-end crosslinks, or that this assay was unable to show this. Additionally, there were two bands observed for each well, even for the control with DMSO and the control monomer. This indicated that there was an issue with the DNA sequence, which seemed to be denatured under the conditions of the gel electrophoresis. The results of this assay can be seen in **Figure 4.20**, which shows in the first well the single-stranded fluorescently-labelled oligomer. The two bands were clearly visible, and it was observed that the band that travelled further was, in fact, a result of this labelled single stranded oligomer, indicating that the other mixtures the DNA had not fully annealed or been denatured in the experiment.

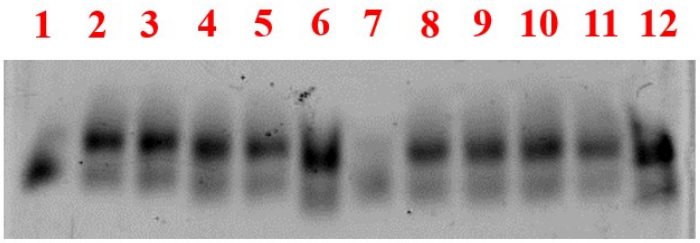


Figure 4.20: *Crosslinking assay results - numerical left to right; 1 and 7) control labelled DNA. 2 and 8) control monomer 50* µM. *3 and 9) dimer 0.05* µM. *4 and 10) dimer 0.5* µM. *5 and 11) dimer 5* µM. *6 and 12) dimer 50* µM.

There was no evidence of the dimer being able to form end to end crosslinks of two DNA strands therefore no further work was done on this assay.

4.6. Conclusions and Future Work

This project successfully synthesised a set of novel compounds and then tested these in three different biological assays.

This project aimed to investigate whether binding together two duocarmycin subunits would make a potent payload potentially for an ADC. This was achieved, supported by the data obtained from the MTS assay. The MTS assay showed that all seven compounds synthesised in this project were able to effectively induce cell death in HL-60 cells. This assay showed consistent results in potency of dimers with varying stereochemistry when compared to a similar study.[22] It was determined that dimers with the *β*-alanine linker were more potent than those bearing the PEG dimer. Although these dimers were more potent, synthetically, they gave lower yields, potentially due to a lack of solubility when compared to the PEG dimers. Despite later assays not being able to prove DNA crosslinking, it is likely that compound (**4.3**) is able to form interstrand crosslinks, leading to the highly potent low nanomolar antiproliferative activity observed compared with the other dimers and the monomer.

The AuNP assay was developed to show duplex to duplex crosslinking and the assay was run successfully, as supported by the results with a control acridine. The negative control correctly showed no aggregation of AuNPs, as expected with a monomer. There was no evidence to support that the dimers synthesised were able to form

duplex-duplex crosslinks. One possible reason for this may be the linkers were too short, an interesting test would be to make the dimers using alternative linkers, maybe a PEG-3 or -4, that may allow the crosslinking to occur. Another interesting idea may include binding an acridine to a duocarmycin to see if these are able to form duplex-to-duplex crosslinks.

Gel electrophoresis was unable to show DNA binding, cleavage or crosslinking with the procedures run. Given more time, this project could investigate the use of different sequences of the oligomers, different kinds of gel, and different times for annealing and incubation for the compounds with these custom DNA strands. Investing longer and different strands of DNA, and altering potential binding sites, could be an interesting concept to study if crosslinking could be proven. The length of cross-link on duplex stability could be linked to values obtained in the MTS assay, hence, help to further understand how these compounds are binding and how it relates to their potency.

References

1 P. J. McHugh, V. J. Spanswick and J. A. Hartley, Repair of DNA interstrand crosslinks: molecular mechanisms and clinical relevance, *Lancet Oncol.*, 2001, **2**, 483–490.

2 A. J. Deans and S. C. West, DNA interstrand crosslink repair and cancer, *Nat. Rev. Cancer*, 2011, **11**, 467–480.

3 H. B. Rycenga and D. T. Long, The evolving role of DNA inter-strand crosslinks in chemotherapy, *Curr. Opin. Pharmacol.*, 2018, **41**, 20–26.

4 D. Noll, A. Noronha, C. Wilds and P. Miller, Preparation of interstrand cross-linked DNA oligonucleotide duplexes, *Front. Biosci. J. Virtual Libr.*, 2004, **9**, 421–37.

5 P. B. Hopkins, J. T. Millard, J. Woo, M. F. Weidner, J. J. Kirchner, S. Th. Sigurdsson and S. Raucher, Sequence preferences of DNA interstrand cross-linking agents: Importance of minimal DNA structural reorganization in the cross-linking reactions of mechlorethamine, cisplatin and mitomycin C, *Tetrahedron*, 1991, **47**, 2475–2489.

6 N. H. Hopcroft, A. L. Brogden, M. Searcey and C. J. Cardin, X-ray crystallographic study of DNA duplex cross-linking: simultaneous binding to two d(CGTACG)2 molecules by a bis(9-aminoacridine-4-carboxamide) derivative, *Nucleic Acids Res.*, 2006, **34**, 6663–6672.

7 K. W. Kohn, in *Molecular Aspects of Anti-Cancer Drug Action*, eds. S. Neidle and M. J. Waring, Macmillan Education UK, London, 1983, pp. 315–361.

8 L. F. Craver, The Nitrogen Mustards: Clinical Use, *Radiology*, 1948, **50**, 486–493.

9 D. Fu, J. A. Calvo and L. D. Samson, Balancing repair and tolerance of DNA damage caused by alkylating agents, *Nat. Rev. Cancer*, 2012, **12**, 104–120.

10 J. O. Ojwang, D. A. Grueneberg and E. L. Loechler, Synthesis of a Duplex Oligonucleotide Containing a Nitrogen Mustard Interstrand DNA-DNA Cross-Link1, *Cancer Res.*, 1989, **49**, 6529–6537.

11 Y. Mantri, S. J. Lippard and M.-H. Baik, Bifunctional Binding of Cisplatin to DNA: Why Does Cisplatin Form 1,2-Intrastrand Cross-Links with AG But Not with GA?, *J. Am. Chem. Soc.*, 2007, **129**, 5023–5030.

12 K. Stehlikova, DNA bending and unwinding due to the major 1,2-GG intrastrand cross-link formed by antitumor cis-diamminedichloroplatinum(II) are flanking-base independent, *Nucleic Acids Res.*, 2002, **30**, 2894–2898.

13 J.-M. Malinge, M.-J. Giraud-Panis and M. Leng, Interstrand cross-links of cisplatin induce striking distortions in DNA, *J. Inorg. Biochem.*, 1999, **77**, 23–29.

14 D. L. Boger and D. S. Johnson, CC-1065 and the Duocarmycins: Understanding their Biological Function through Mechanistic Studies, *Angew. Chem. Int. Ed. Engl.*, 1996, **35**, 1438–1474.

15 D. L. Boger, H. W. Schmitt, B. E. Fink and M. P. Hedrick, Parallel Synthesis and Evaluation of 132 (+)-1,2,9,9a-Tetrahydrocyclopropa[c]benz[e]indol-4-one (CBI) Analogues of CC-1065 and the Duocarmycins Defining the Contribution of the DNA-Binding Domain, *J. Org. Chem.*, 2001, **66**, 6654–6661.

16 CC-1065 and the duocarmycins, https://www.pnas.org/doi/10.1073/pnas.92.9.3642, (accessed 2 April 2024).

17 R. J. Stevenson, W. A. Denny, M. Tercel, F. B. Pruijn and A. Ashoorzadeh, Nitro seco Analogues of the Duocarmycins Containing Sulfonate Leaving Groups as Hypoxia-Activated Prodrugs for Cancer Therapy, *J Med Chem*, 2012, **55**, 2780–2801.

18 D. L. Boger, C. W. Boyce, R. M. Garbaccio and J. A. Goldberg, CC-1065 and the duocarmycins: Synthetic studies, *Chem. Rev.*, 1997, **97**, 787–828.

19 M. S. Tichenor, K. S. MacMillan, J. D. Trzupek, T. J. Rayl, I. Hwang and D. L. Boger, Systematic exploration of the structural features of yatakemycin impacting DNA alkylation and biological activity, *J. Am. Chem. Soc.*, 2007, **129**, 10858–10869.

20 D. L. Boger and R. M. Garbaccio, Catalysis of the CC-1065 and duocarmycin DNA alkylation reaction: DNA binding induced conformational change in the agent results in activation, *Bioorg. Med. Chem.*, 1997, **5**, 263–276.

21 J. A. Smith, G. Bifulco, D. A. Case, D. L. Boger, L. Gomez-Paloma and W. J. Chazin, The structural basis for in situ activation of DNA alkylation by duocarmycin SA11Edited by I. Tinoco, *J. Mol. Biol.*, 2000, **300**, 1195–1204.

22 D. L. Boger, M. Searcey, W. C. Tse and Q. Jin, Bifunctional alkylating agents derived from duocarmycin SA: Potent antitumor activity with altered sequence selectivity, *Bioorg. Med. Chem. Lett.*, 2000, **10**, 495–498.

23 N. Ghosh, H. Sheldrake, M. Searcey and K. Pors, Chemical and Biological Explorations of the Family of CC-1065 and the Duocarmycin Natural Products, *Curr. Top. Med. Chem.*, 2009, **9**, 1494–1524.

24 P. J. M. Jackson, K. M. Rahman and D. E. Thurston, The use of molecular dynamics simulations to evaluate the DNA sequence-selectivity of G–A cross-linking PBD–duocarmycin dimers, *Bioorg. Med. Chem. Lett.*, 2017, **27**, 102–108.

25 S. K. Sharma, G. Jia and J. W. Lown, Novel Cyclopropylindole Conjugates and Dimers Synthesis and Anti-Cancer Evaluation, *Curr. Med. Chem. - Anti-Cancer Agents*, 2001, **1**, 27–45.

26 S. A. O'Toole, B. L. Sheppard, E. P. J. McGuinness, N. C. Gleeson, M. Yoneda and J. Bonnar, The MTS assay as an indicator of chemosensitivity/resistance in malignant gynaecological tumours, *Cancer Detect. Prev.*, 2003, **27**, 47–54.

27 W. A. Chen, T. G. Williams, L. So, N. Drew, J. Fang, P. Ochoa, N. Nguyen, Y. Jawhar, J. Otiji, P. J. Duerksen-Hughes, M. E. Reeves, C. A. Casiano, H. Jin, S. Dovat, J. Yang, K. E. Boyle and O. L. Francis-Boyle, Duocarmycin SA Reduces Proliferation and Increases Apoptosis in Acute Myeloid Leukemia Cells In Vitro, *Int. J. Mol. Sci.*, , DOI:10.3390/ijms25084342.

28 O. C. Cartwright, A. M. Beekman, M. M. D. Cominetti, D. A. Russell and M. Searcey, A Peptide-Duocarmycin Conjugate Targeting the Thomsen-Friedenreich Antigen Has Potent and Selective Antitumor Activity, *Bioconjug. Chem.*, 2020, **31**, 1745–1749.

29 M. E. Davis, Z. (Georgia) Chen and D. M. Shin, Nanoparticle therapeutics: an emerging treatment modality for cancer, *Nat. Rev. Drug Discov.*, 2008, **7**, 771–782.

30 J. D. E. T. Wilton-Ely, The surface functionalisation of gold nanoparticles with metal complexes, *Dalton Trans.*, 2007, 25–29.

31 K. Sztandera, M. Gorzkiewicz and B. Klajnert-Maculewicz, Gold Nanoparticles in Cancer Treatment, *Mol. Pharm.*, 2019, **16**, 1–23.

32 J. R. Nicol, D. Dixon and J. A. Coulter, Gold nanoparticle surface functionalization: a necessary requirement in the development of novel nanotherapeutics, *Nanomed.*, 2015, **10**, 1315–1326.

33 M. J. Marín, B. D. Rackham, A. N. Round, L. A. Howell, D. A. Russell and M. Searcey, A rapid screen for molecules that form duplex to duplex crosslinks in DNA †, *Chem Commun*, 2013, **49**, 9113.

34 B. V. Enustun and John. Turkevich, **coagulation of colloidal gold**, *J. Am. Chem. Soc.*, 1963, **85**, 3317–3328.

35 X. Zhang, M. R. Servos and J. Liu, Instantaneous and quantitative functionalization of gold nanoparticles with thiolated DNA using a pH-assisted and surfactant-free route, *J. Am. Chem. Soc.*, 2012, **134**, 7266–7269.

36 J. R. Reimers, M. J. Ford, S. M. Marcuccio, J. Ulstrup and N. S. Hush, Competition of van der Waals and chemical forces on gold–sulfur surfaces and nanoparticles, *Nat. Rev. Chem.*, 2017, **1**, 1–13.

37 N. Capelle, J. Barbet, P. Dessen, S. Blanquet, B. P. Roques and J. B. L. Pecq, Deoxyribonucleic acid bifunctional intercalators, https://pubs.acs.org/doi/pdf/10.1021/bi00582a023, (accessed 20 March 2024).

38 P. Prasher and M. Sharma, Medicinal chemistry of acridine and its analogues, *MedChemComm*, 2018, **9**, 1589.

39 J. Markovits, C. Garbay-Jaureguiberry, B. P. Roques and J.-B. Le PECQ, Acridine dimers: influence of the intercalating ring and of the linking-chain nature on the equilibrium and kinetic DNA-binding parameters, *Eur. J. Biochem.*, 1989, **180**, 359–366.

40 P. Y. Lee, J. Costumbrado, C.-Y. Hsu and Y. H. Kim, Agarose Gel Electrophoresis for the Separation of DNA Fragments, *JoVE J. Vis. Exp.*, 2012, e3923.

Chapter 5 – Experimental

Chapter 5 – Experimental

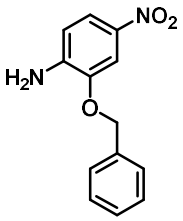
5.1 Materials and General considerations

All solvents and chemicals were reagent grade and purchased from appropriate suppliers. The reagents used were used from the following named suppliers: Sigma Aldrich, Alfa Aesar, Acros Organics, BHD Biochemicals, Fluorochem, Thermo Fisher. All water used was distilled. All reactions requiring anhydrous conditions were performed under an atmosphere of nitrogen. All ^{1}H and ^{13}C NMR spectra were recorded in Fourier transform mode operating at a ^{1}H NMR frequency of 400 MHz and a ^{13}C NMR frequency of 100 MHz, using the specified deuterated solvent. ^{1}H and ^{13}C NMR spectra were obtained using a Bruker spectrometer. The chemical shifts for both ^{1}H and ^{13}C spectra were recorded in ppm relative to internal tetramethylsilane (0.00 ppm) and were referenced to the residual solvent peak. CD_3COCD_3 was referenced to 2.05 ppm, and $CDCl_3$ was referenced to 7.26 ppm. Abbreviations of multiplicities in the NMR spectra are described as; s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad, dd = doublet of doublets, dt=doublet of triplets, td= triplet of doublets. Coupling constants are reported in Hertz. Infrared spectra were obtained on a Perkin Elmer Spectrum Two. Organic solutions which had been in contact with water were dried over sodium sulphate unless otherwise specified. Thin-layer chromatography was performed on aluminium plates coated with 0.20 mm silica gel 60 with florescent indicator UV_{254}. After elution, the TLC plates were visualized under UV light. Flash chromatographic separations were performed on silica gel for column chromatography. Detection wavelength was 254 nm and 280 nm. Reactions done under microwave conditions used a CEM Discover SP machine at Dynamic mode, with low stirring and heating at 200 °C for 15 minutes. Analytical RP-HPLC was performed using an Agilent 1200 HPLC, fitted with an Agilent eclipse XDB-C18 column (4.6 x 150 mm, 5 μm) and a flow rate of 1 mL/min. Analyses were run with a solvent gradient of 0-100% B over 20 min. Solvent A: H_2O, 0.05% TFA, solvent B: MeCN, 0.05% TFA. Detection wavelengths were 214 nm and 254 nm. Low resolution masses were recorded using an Agilent 1200 Series equipped with a 6460 triple quadrupole. High resolution masses were recorded using Waters Synapt XS QTOF under electrospray ionisation. UV-vis spectra were recorded on an Agilent Cary 60 spectrometer, using quartz cuvettes with 1 cm pathlength. For the MTS assay, the

plates were measured using the BMG Labtech POLARstarOPTIMA microplate reader at a wavelength of at 492 nm. IC_{50} values were calculated using GraphPad Prism Version 6.0 software, using a four-parameter logistic nonlinear regression model. For gel electrophoresis, samples were stained by a mixture of two dyes (bromophenol blue and xylene cyanol FF) from ThermoFisher prior to loading. Gels were visualised in an Image Quant LAS 4000.

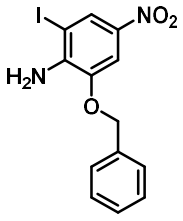
5.2 Synthesis of DSA Subunit Method 1

5.2.1. Preparation of 4-Nitro-2-(phenylmethoxy)benzenamine (2.31)



2.31 was prepared following a published procedure.[1] 2-Amino-5-nitrophenol, **2.26** (25.05 g, 0.163 mol) was dissolved in DMF (250 mL) and treated with K_2CO_3 (49.3 g, 0.356 mol) and dropwise with BnBr (21.0 mL, 0.178 mmol). The reaction was stirred overnight at room temperature, before being poured over crushed ice. The yellow/brown precipitate was collected by filtration. This was then triturated with cold water and dried under vacuum overnight. Compound **2.31** was used without further purification (37.95, 0.155 mol, 96 %). 96 % For 25: IR (ATR) 3482, 3358, 1821, 1578, 1518, 1478, 1386, 1272, 1223, 1006, 697cm^{-1}. ^{1}H NMR (CDCl3, 400 MHz) δ 7.84 (dd, *J* =8.7, 2.4 ,1H), 7.78 (d, *J* =2.4, 1H), 7.36-7.46 (m, 5H), 6.67 (d, J =8.6 ,1H), 5.15 (s, 2H). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 144.7, 143.6, 138.8, 135.9, 128.9, 128.7, 128.0, 119.5, 112.1, 107.4, 71.0

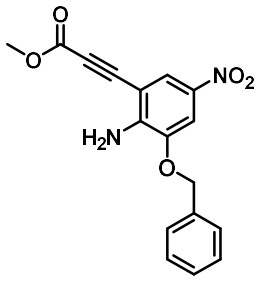
5.2.2. Preparation of 2-Iodo-4-nitro-6-(phenylmethoxy)benzenamine (2.27)



Compound **2.27** was prepared following a published procedure.[1] **2.31** (16.18 g, 66 mmol) was dissolved in DMF (240 mL) and treated with H_2SO_4 (350 µL, 67 mmol). NIS (22.38 g, 99 mmol) was added portion wise. The reaction was stirred for 4 h at room temperature before being poured over crushed ice. The bright yellow precipitate that crashed out was filtered and then triturated with water followed by cold hexane. Compound **2.27** was dried under vacuum and was used without further purification (19.22 g, 52 mmol, 78 %). For **2.27**: IR (ATR) 3475, 3378, 3357, 1603, 1495, 1276, 1023, 692 cm^{-1}. ^{1}H NMR (CDCl3, 400 MHz) δ 8.31 (d, *J* = 2.4, 1H), 7.76 (d, *J*=2.3, 1H), 7.42-7.46 (m,

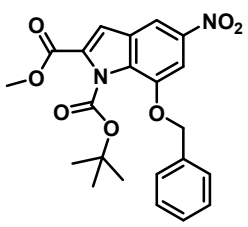
5H), 5.18 (s, 2H).). 13C NMR (CDCl3, 100 MHz) δ 144.1, 142.7, 138.9, 135.4, 129.0, 128.9, 128.3, 128.1, 107.1, 78.5, 72.0

5.2.3 Preparation of Methyl 3-[2-Amino-3-(phenylmethoxy)-5-nitrophenyl]-2-propynoic Acid (2.28).



2.28 was prepared following a published procedure, minor changes were made and this reaction was repeated on a maximum of a 10 g scale due to limitations of the laboratory equipment.[1] Product **2.27** (10 g) was dissolved in anhydrous DMF (250 mL). The solution was degassed with nitrogen for 1 hour before addition of methyl propiolate (9.8 mL, 117 mmol), $Pd(PPh_3)_2Cl_2$ (1.00 g, 1.42 mmol), $ZnBr_2$ (26.00 g, 115 mmol) and DIPEA (20 mL, 113 mmol). The reaction was stirred overnight under N_2 at 66 ºC. After cooling to room temperature, the reaction was poured over crushed ice, the precipitate was then filtered. This was triturated with water followed by cold ethanol. The dark brown mixture was then dry loaded and purified by manual column chromatography with 30% Ethyl acetate and hexane to afford product **2.28** as an orange solid (2.2 g, 0.027 mol, 25%): IR (ATR) 3498, 3347, 2203, 1697, 1610, 1506, 1298, 1000, 740, 695 cm^{-1}. ^1H NMR (CDCl3, 400 MHz) δ 8.06 (d, *J*=2.3, 1H), 7.76 (d, *J*=2.4, 1H), 7.39-7.25 (m, 5H), 5.31 (brs, 2H), 5.17 (s, 2H), 3.86 (s, 3H). (CDCl3, 100 MHz) δ 154.0, 146.7, 144.5, 137.6, 135.1, 129.0, 128.9, 128.1, 123, 108.2, 100.9, 87.1, 81.1, 71.4, 53.1.

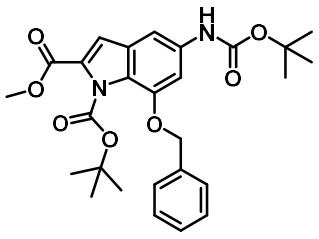
5.2.4. Preparation of 1-(1,1-Dimethylethyl)-2-methyl 5-Nitro-6-(phenylmethoxy) indole-1,2-dicarboxylate (2.17).



Compound **2.17** was prepared following a published procedure.[1] Compound **2.28** (1.75 g, 5.37 mmol) was dissolved in anhydrous THF (25 mL) and treated with TBAF solution (1M in THF, 11.16 mL, 11.16 mmol). The solution was refluxed at 66 ºC for 1.5 h. After cooling, the solvent was then removed under reduced pressure. The product was then dissolved in ethyl acetate (25 mL) and washed 3 times with water, followed by brine. A crude NMR was then run in chloroform, if there were no major impurities then the crude was carried on to the next step, if not then it was purified before Boc protection. The crude was dissolved in DCM (30 mL) and treated with Boc_2O (2.44 g, 11.2 mmol) and DMAP (0.68 g, 5.6 mmol). The reaction was

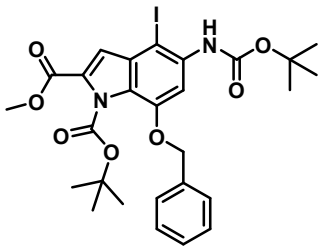
stirred at room temperature for 2 h. The solvent was then removed under reduced pressure before dry loading onto silica for column chromatography. A gradient of 0% to 10% ethyl acetate in hexane was run to afford **2.17** as a bright yellow solid (1.0592 g, 2.48 mmol, 46%). For **2.17**: IR (ATR) 2981, 1765, 1721, 1511, 1325, 1225,1152, 742, 697 cm^{-1}. ^{1}H NMR (CDCl3, 400 MHz) δ 8.26 (d, *J* = 1.9, 1H), 7.68 (d, *J*= 1.9, 1H), 7.45-7.48 (m, 2H), 7.34-7.41 (m, 3H), 7.34 (s, 2H), 5.32 (s, 2H) 3.94 (s, 3H) 1.47 (s, 9H). ^{13}C NMR (CDCl3, 100 MHz) δ 160.0, 149.3, 145.6, 143.6, 135.2, 130.1, 128.9, 128.7, 128.2, 126.4, 112.7, 112.5, 102.2, 86.5, 71.30, 52.6, 27.3

5.2.5. Preparation of (2-Methyl 1-(2-methyl-2-propanyl) 7-(benzyloxy)-5-({[(2-methyl-2-propanyl)oxy]carbonyl}amino)-1H-indole-1,2-dicarboxylate) (2.31)



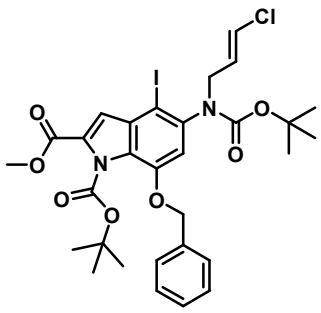
Compound **2.31** was prepared following a published procedure.[1] Compound **2.17** (2.45 g, 5.74 mmol) was dissolved in THF (48 mL) and treated with zinc powder (5.64 g, 86 mmol), NH_4Cl (3.08 g, 57.6 mmol), Boc_2O (3.76 g, 17.2 mmol), DMAP (77 mg, 0.63 mmol) and water (9.6 mL). The reaction was stirred overnight before filtering off the excess zinc. The solution was then concentrated under reduced pressure. The mixture was dissolved in EtOAc and washed with water (25 mL) three times, followed by brine. The product was then concentrated and dry loaded onto silica before being subject to column chromatography. A manual column 20% EtOAc in hexane was run to obtain pure white foam**, 2.31** (2.0911 g, 4.21 mmol, 73%). For **2.31**: IR (ATR) 3345, 2979, 1764, 1716, 1436, 1238, 1149, 842 cm^{-1}. ^{1}H NMR (CDCl3, 400 MHz) δ 7.42-7.46 (m, 2H), 7.30-7.38 (m, 5H) 7.10 (s, 1H), 7.45-7.48 (m, 2H), 7.34-7.41 (m, 3H), 7.34 (s, 2H), 5.21 (s, 2H) 3.89 (s, 3H) 1.51 (s, 9H), 1.44 (s, 9H). ^{13}C NMR (CDCl3, 100 MHz) δ – 160.5, 149.29, 145.64, 143.62, 135.20, 130.2, 130.1, 128.9, 128.7, 128.2, 126.4, 112.7, 112.5, 102.2, 86.5, 71.3, 53.0, 31.7 , 27.2, 22.8

5.2.6. Preparation of 1-(1,1-Dimethylethyl)-2-methyl 5-[[(1,1-dimethylethoxy)-carbonyl]amino]-4-iodoo-7-(phenylmethoxy)indole-1,2-dicarboxylate (2.19)



2.19 was prepared following a published procedure.[2] Starting material **2.31** (1.05 g, 2.1 mmol) was dissolved in toluene (57 ml) and treated with acetic acid (0.47 ml) and NIS (0.94 g, 4.18 mmol), this was then stirred for three hours and loaded directly onto a flash column, by injecting the toluene. It was purified using a gradient of 0 to 15% EtOAc in hexane and afforded the pure product as a white solid, **2.19** (0.9546 g, 1.53 mmol, 73%): IR (ATR) 3356, 2984, 2935, 1763, 1716, 1449, 1361, 1222, 1149, 1079, 980, 758 cm^{-1}. ^{1}H NMR (CDCl3, 400 MHz) δ 7.76 (brs, 1H), 7.42-7.46 (m, 2H), 7.27-7.35 (m, 3H), 7.07 (s, 1H), 6.74 (brs, 1H), 5.22 (s, 2H), 3.89 (s, 3H), 1.51 (s, 9H), 1.38 (s, 9H).^{13}C NMR (CDCl3, 100 MHz) δ 160.9, 153.12, 149.9, 146.57, 136.0, 135.0, 132.0, 128.7, 128.6, 128.4, 127.8, 123.70, 114.5, 102.6, 85.6, 81.0, 71.0, 52.3, 28.5, 27.3

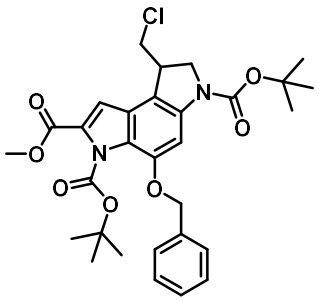
5.2.7. Preparation of 1-(1-,1-Dimethylethyl)-2-methyl 5-[(3-Chloro-2-propenyl) [(1,1-dimethylethoxy)carbonylamino]-4-iodo-7-(phenylmethoxy)indole-1,2-dicarboxylate (2.24)



2.24 was prepared following a published procedure, with minor changes.[1] Compound **2.19** (0.9518 g, 1.53 mmol) was dissolved in THF (5 mL) and treated with t-BuOK (70 mg, 0.62 mmol) and 1,3-dichloropropene (a mix of cis and trans isomer) (0.1 mL, 0.9 mmol a mixture of cis and trans isomers). The reaction was stirred at room temperature for 1.5 h. The reaction was worked up by cooling to 0 ºC and then quenched by treating the reaction dropwise with saturated aqueous ammonium chloride (1 mL). EtOAc (15 mL) was added to the solution which was washed 3 times with water followed by brine. The organic was dried in vacuo and was co-evaporated with CH_2Cl_2 to form a crude mixture. This was then dry loaded onto silica and purified using flash chromatography, on a gradient of 0 to 15% ethyl acetate in hexane giving final product as a light brown foam, **2.24** (0.88 g, 1.25 mmol, 82%): ^{1}H NMR ($CDCl_3$, 400 MHz, mixture of E/Z isomers); δ7.30-7.43 (m, 5H), 7.18 (s, 1H), 6.48-6.52, 2H), 5.82-5.98 (m, 2H), 5.19-5.30 (m, 2H), 4.31-4.49 (m, 1H), 4.11- 4.20 (m, 1H), 3.93 (s, 3H), 1.53

(s, 9H), 1.26 and 1.28 (s, 9H). IR (ATR) 2979, 1770, 1697, 1570, 1366, 1220, 1149, 1079, 841, 734 cm^{-1}.

5.2.8. Preparation of 6-Bis(1,1-dimethylethyl)-2-methyl 8-(chloromethyl)-7,8-dihydro-4-(phenylmethoxy)benzo[1,2-b:4,3-b′]dipyrrole-2,3,6-tricarboxylate (2.25).

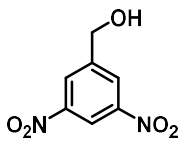


2.25 was prepared following a published procedure, with minor changes.[1] Product **2.24** (0.84 g, 1.2 mmol) was dissolved in anhydrous toluene (12 ml) and degassed for 1 h in nitrogen. AIBN (0.05 g, 0.3 mmol) and TTMSS (0.41 mL, 1.3 mmol). The resulting solution was refluxed at 90 ºC under nitrogen for 2 h. The mixture was then concentrated and loaded onto silica before column chromatography, gradient of 0% to 4% ethyl acetate in hexane over 15 column volumes. Purification gave a white foam, **2.25** (0.4928 g, 0.86 mmol, 72%): IR (ATR) 2978, 1766, 1719, 1694, 1494, 1418, 1345, 1368, 1221, 1136, 1083, 842, 695 cm^{-1}. ^{1}H NMR (DMSO-d6, 400 MHz) δ 7.29-7.46 (m, 7H), 5.27 (s, 2H), 4.13 (t, J=9.8, 1H), 3.89-4.05 (m, 4H), 3.87 (s, 3H), 1.48 (s, 9H), 1.39 (s,9H). ^{13}C NMR (DMSO-d6, 100 MHz) δ 160.9, 151.9, 150.0, 145.6, 136.6, 128.9, 128.5, 128.3, 124.0, 113.2, 109.0, 97.4, 85.5, 80.3, 70.2, 52.8, z52.7, 48.1, 28.5, 27.2, 22.5

- *The proton differs from the literature as 13C peak at 40.7 was not seen, as it was obscured by the DMSO peak*

5.3 Synthesis of DSA Subunit Method 2

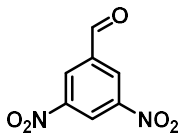
5.3.1 Preparation of 3,5-dinitrobenzyl alcohol (2.33)



2.33 was prepared following a published procedure.[3] Small changes were made. The procedure was as follows: To a solution of 3,5-dinitrobenzoic acid, **2.32** (2.13 g, 10 mmol) in anhydrous THF (15 mL) at 0 °C, borane-tetrahydrofuran complex (15 mL of IM in THF, 1.50 mo l) was added over 1 h. The resulting heterogeneous mixture was stirred at 0 °C for 3 h and at 25 °C overnight with an additional 4 mL borane-tetrahydrofuran complex. After 12 h a further 4 mL borane-tetrahydrofuran complex added, TLC showed complete

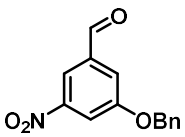
conversion after 2 h. The resulting solution was quenched with H_2O and concentrated *in vacuo* until solids were present. The solids were filtered, washed with H_2O and dissolved in EtOAc. The aqueous filtrate was extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with aqueous $NaHCO_3$ and saturated NaCl. The organic layer was dried with sodium sulphate and solvent removed under reduced pressure to afford crude material.The compound was then purified by trituration with toluene to afford **2.33** (1.973 g, 9.96 mmol, 99%) as a beige solid. IR (ATR) 3248 (O-H), 3101 (C-H), 1519 (N-O) cm^{-1}. ^1H NMR (400 MHz, CD_3OD) δ 8.88 (m, 1H), 8.55-8.68 (m, 2H), 4.83 (s 2H) – there was impurity of water at 4.85. ^{13}C NMR (100 MHz, CD_3OD) δ 179.25, 176.31, 155.26, 146.03, 91.09

5.3.2 Preparation of 3,5-dinitrobenzaldehyde (2.12).



Product **2.12** was prepared following procedure previously done in the research group. 3,5-dinitrobenzylalcohol, **2.33**, (0.33 g, 1.68 mmol) in 11 mL CH_2Cl_2 was treated with Dess-Martin periodinane **(**0.94 g) and stirred overnight at room temp. The solution was concentrated in vacuo, and the residue was purified by flash chromatography (100% CH_2Cl_2) to give methyl 5-benzyloxy-5-nitrobenzaldehyde **(2.12)** as pale-yellow crystals (0.3017 g, 1.54 mol, 91%). IR (ATR) 3075 (C-H), 1699 (C=O), 1538 (N-O) cm^{-1}. ^1H NMR ($CDCl_3$, 400 MHz): 10.22 (s, 1H), 9.30 (t, *J*=2.22 *Hz* 1H), 9.05 (d, *J*=2.0 *Hz*, 2H); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 187.1, 150.10, 138.31, 129.05, 122.59

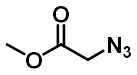
5.3.3 Preparation of 3-Benzyloxy-5-nitrobenzaldehyde (2.13).



Product **2.13** was prepared following published procedure with minor changes.[2] A solution of *syn*-benzaldehyde oxime (0.611 g, 5.03 mmol) in DMF (5.7 mL) was treated with K_2CO_3 (1.43, 10.3 mmol) and heated to 90 °C. After 10 minutes of heating, a solution of 3,5-dinitrobenzaldehyde, **2.12,** (0.505 g, 2.57 mmol) in DMF (5.7 mL). This solution was stirred at 90 °C for 1 h. The reaction mixture was cooled to 25 °C, BnBr (0.71 mL, 5.88 mmol) was added and stirred for 2 h. The mixture was diluted with Et_2O, extracted with 1M aqueous HCl and then washed with H_2O, saturated aqueous $NaHCO_3$, saturated aqueous NaCl and dried with sodium sulphate. The solvent was

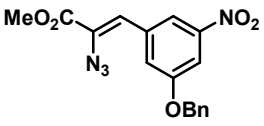
removed in vacuo and purified by flash chromatography (35% DCM−hexanes) to provide **2.13** (0.494 g, 1.92 mmol, 75%) as a pale yellow solid: IR (ATR) 1700 (C=O), 1529 (N-O) cm^{-1}.^{1}H NMR (CDCl3, 400 MHz) δ 10.03 (s, 1H), 8.30 (dd, *J* = 2.1, 1.3 Hz, 1H), 8.07 (t, *J* = 2.3 Hz, 1H), 7.79 (dd, *J* = 2.5, 1.2 Hz, 1H), 7.47−7.35 (m, 5H), 5.22 (s, 2H); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 190.01, 159.64, 150.12, 138.69, 135.45, 129.33, 129.14, 128.09, 120.53,117.67, 115.84, 71.61

5.3.4 Preparation of methyl 2-azidoacetate (2.14)



To a flask, methyl bromoacetate, **2.37** (20.61 g, 0.135 mol), NaN_3 (33.04 g, 0.51 mol, 4 eq) and 150 mL acetone were added. The mixture was refluxed for 12 h under N_2. The solvent was removed in vacuo, and the product was extracted with diethyl ether three times and the ppt was filtered off. The organic phase was then removed in vacuo to give methyl 2-azidoacetate, **2.14** as a pale-yellow oil (12.9977 g, 0.113 mol, 84%). The isolated methyl 2-azidoacetate should be used without purification. IR (ATR) 2102 ($N^-=N^+=N^-$), 1743 (C=O), 1203 (C-O) cm^{-1}. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 3.89 (s, 2H), 3.81 (s, 3H). ^{13}C NMR ($CDCl_3$, 100 MHz): δ 169.4, 53.2, 50.4

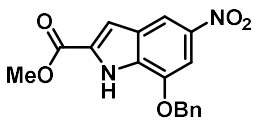
5.3.5 Preparation of methyl 2-azido-3-(3-benzyloxy-5-nitrophenyl)acrylate (2.16)



Product **2.16** was prepared following a previously reported procedure.[2] Sodium methoxide was prepared beforehand by adding sodium metal (1.48 g, 64.4 mmol) to anhydrous MeOH (36 mL) at 0 °C and stirring until the sodium fully dissolved. The solution was cooled using acetonitrile and dry ice to –41 °C and then treated dropwise with a solution of 3- benzyloxy-5-nitrobenzaldehyde, **2.13** (3.18 g, 212.3 mmol) and methyl azido acetate, **23** (5.72 g, 49.7 mmol) in THF (11.5 mL). The reaction was stirred at –41 °C for an additional 3 h, and then at 0 °C for 14 h. The reaction mixture was diluted with CH_2Cl_2, washed with H_2O, saturated aqueous NaCl and dried with sodium sulphate. The solvent was removed in vacuo compound **2.16** used without further purification (1.90 g, 5.35 mmol 43%) as a beige solid: IR (ATR) 2132 ($N^-=N^+=N^-$), 1709 (C=O), 1527 (N-O) cm^{-1}. ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.19 (t, *J* = 1.5 Hz, 1H), 7.81 (t, *J* = 1.8 Hz 1H), 7.77 (t, *J* = 2.2 Hz, 1H), 7.36−7.46 (m, 5H), 6.83 (s, 1H), 5.17 (s, 2H),

3.93 (s, 3H); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 163.6, 158.9, 149.5, 136.0, 135.2, 128.6 (2C), 128.3, 127.9, 127.4 (2C), 122.6, 121.8, 117.8, 109.6, 70.7, 53.1

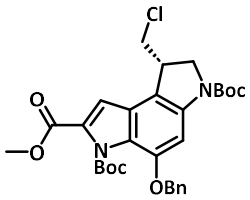
5.3.6 Preparation of methyl 7-Benzyloxy-5-nitroindole-2-carboxylate (2.29)



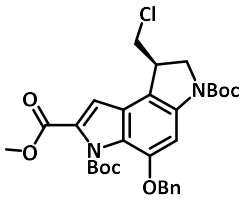
Product **2.29** was prepared following previously reported procedure referenced.[2] A solution of **2.16** (1.36 g, 3.83 mmol) in anhydrous xylenes (340 mL) refluxed at 140 °C for 7 h. The solution was then cooled, and solvent removed in vacuo. The product was purified using a manual column of 5% acetone in hexane to obtain the product as pale-yellow crystals (0.54 g, 1.65 mmol, 43%). For **2.29**: IR (ATR) 3296 (N-H), 1704 (C=O), 1521 (N-O) cm^{-1}. ^{1}H NMR ($CDCl_3$, 400 MHz) δ 9.33 (br s, 1H), 8.36 (dd, *J* = 1.8, 0.7 Hz, 1H), 7.72 (d, *J* = 1.9 Hz, 1H), 7.42-7.52 (m, 5H), 7.35 (d, *J* = 2.3 Hz, 1H), 5.29 (s, 2H), 3.96 (s, 3H). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 161.2, 145.0, 143.0, 135.1, 130.6, 129.5, 128.7 (2C), 128.6, 128.1 (2C), 126.6, 113.1, 110.8, 100.3, 70.9, 52.2

5.4 Separation of enantiomers

Separation was achieved using a chiral column modus 25 g puriflash. Samples were dry loaded in a minimum amount of silica and run continuously in isocratic conditions of 35% IPA in hexane. The peaks were collected based on UV trace and checked using optical rotation recorded using a Polarimeter. For 100 mg of the Boc protected ester dissolved in DCM.



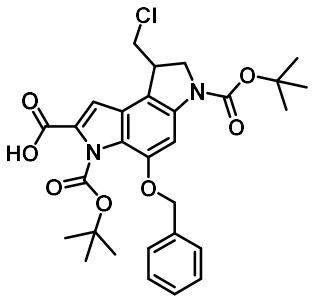
Natural = -0.10
2.25 (a)



Unnatural = 0.11
2.25 (b)

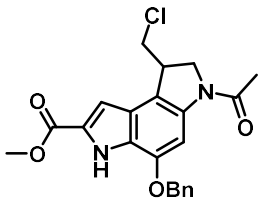
5.5 Synthesis of protected dimers and control monomer

5.5.1 Ester hydrolysis: Preparation of 8-(Chloromethyl)-7,8-dihydro-4-(phenylmethoxy)benzo-[1,2-b:4,3-b′]dipyrrole-3,6-dicarboxylic Acid 3,6-Bis(dimethylethyl)ester (2.29)



This reaction followed a literature procedure, with minor changes to workup.[1] 2.25 (0.030 g, 0.053 mmol) was dissolved in a MeOH (530 µL) and mixture of THF (322 µL). This solution was treated dropwise with a saturated aqueous solution of LiOH (165 µL). The reaction was stirred under N_2 for 3 h before removing the solvent under reduced pressure. The product was diluted with distilled water and acidified using 6 M HCl. Acidification was done in and ice bath, solution was checked using litmus paper, acidification promoted a pale green solid. The solid was collected by filtration and dried, affording a pale green solid. The solid was used without further purification, **2.29** (0.0284 g, 0.0051 mol, 97%). Purity confirmed using HPLC run in acetonitrile, single peak observed at 16.967 ppm showing 99.9% purity. Confirmed using LCMS with an expected mass of 556.2 and an observed m/z ratio of 557.2, corresponded to the $[M+H]^+$. ^{1}H NMR (MeOH, 400 MHz): 7.73 (br s, 1H), 3.36-7.52 (m, 7H), 5.26 (s, 2H), 4.12 (t, 1H), 3.85-4.09 (m, 4H), 1.45 (s, 9H), 1.38 (s, 9H). ^{13}C NMR (MeOH, 400 MHz): 170.4, 161.5, 151.54, 149.8, 145.1, 136.6, 129.6, 128.4, 128.0, 123.6, 123.3, 108.1, 84.6, 70.0, 60.2, 55.7, 52.3, 49.6, 47.6, 28.0, 26.8, 21.0,

5.5.2. Synthesis of DSA-Ac protected, 3.15

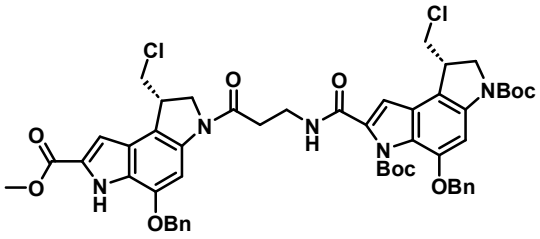


Starting material **2.25** (25 mg, 0.044 mmol) was dissolved in DCM (525 µL) and treated with TFA (450 µL) and TIPS (25 µL). This reaction was stirred vigorously for 1.5 hours before solvent was removed under reduced pressure. The crude was triturated with cold ether giving a green solid, **3.17**. This solid was used dissolved in DCM (500 microlitres) and treated with acetic anhydride (6.3 µL), and DIPEA (22.87 microlitres). The reaction was stirred for 1.5 hours. The solvent was removed under reduced pressure and subject directly to flash chromatography. Isocratic conditions of 50:50 EtOAc in hexane obtained a white amorphous solid, **3.15** (10.4 mg, 0.025 mmol, 57%): ^{1}H NMR ($CDCl_3$ 400 MHz) δ 8.17 (1H, NH) 7.4-7.58 (m, 6H), 7.14 (1H,s) 5.28 (s, 2H), 4.13 (t, 1H), 3.89-4.05 (m, 4H), 3.96 (s, 3H), 2.32 (s, 3H). Product purity was

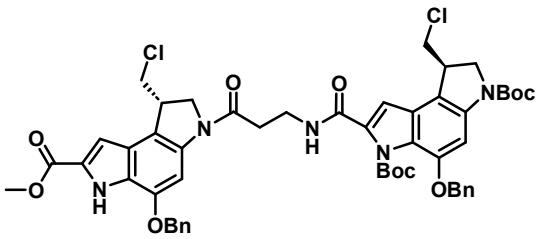
confirmed using single peak on HPLC, correct mass on the LCMS $[M+H]^+ = 413.1$ and HRMS $[M+H]^+ = 413.1288$

5.5.3. Synthesis of DSA-β-ala-DSA Protected

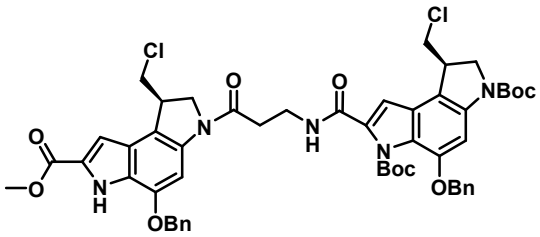
The three different β--alanine dimers were synthesised using the same conditions.



Natural-β-ala-Natural
3.18



Natural-β-ala-Unnatural
3.20



Unatural-β-ala-Unatural
3.19

The procedure for the dimer (**3.18**, **3.19** or **3.20**) synthesis is as follows: Starting material **2.25** dissolved in DCM and treated with TFA and TIPS solution (45% TFA in 52.5% DCM and 2.5 % TIPS). This was stirred for 1.5 hours and checked by HPLC for completion. Reaction worked up by removal of solvent under reduced pressure and trituration with cold ether. No purification required. This deprotected analogue, **3.17** was dissolved in DMF and treated with β-alanine Boc, **3.26** (1.2 eqv) and coupling agents, EDC (1.25 eqv), HOBT (1.25 eqv) and DIPEA (4 eqv). The progress was checked by HPLC and TLC, reaction worked up after ~1.5-3 hours. Crude was crashed out using water and the precipitated product collected by centrifuging off the supernatant. Product was purified using flash chromatography twice (first to remove excess impurities, second to isolate the pure compound). An isocratic run used 50:50 EtOAc in hexane and the product was collected in a 100% EtOAc flush at the end. This was purified using a gradient of 0 to 3% MeOH in DCM obtaining **3.27**. This product was dissolved in DCM and treated with TFA and TIPS solution (45% TFA in 52.5% DCM and 2.5 % TIPS). This was stirred for 1.5 h and checked by HPLC for

completion. The reaction was worked up by removal of solvent under reduced pressure and trituration with cold ether. No purification was required. This compound, **3.28**, could then be coupled with the ester hydrolysed product, **2.29** (synthesised as described in section **5.51**). The final couplings were carried out by dissolving both compounds in DMF and treating with coupling agents, EDC (1.25 equiv), HOBT (1.25 equiv) and DIPEA (4 equiv). The progress was checked by HPLC and TLC, reaction worked up after ~1.5-3 hours. The differences in purifications, yields and purity are summarised in the following sections.

5.5.3.1 Synthesis of Natural-β-ala-Natural, 3.18

This reaction proceeded as expected, two columns were needed to purify the final product, first 0 to 20% EtOAc in hexane. This was flushed with 100% EtOAc, concentrated, and followed with an isocratic column of 3% MeOH in DCM, yielding 2.56 mg of product and a yield of 8%. The mass was confirmed on LC-MS and HRMS. The product purity was confirmed using HPLC, giving a dimer purity of 97%
LRMS (ESI-QQQ): $C_{52}H_{55}Cl_2N_5O_{10}$, calculated $[M+Na]^+ = 1002.32$, found = (1002.3)
HRMS (ESI): $C_{52}H_{55}Cl_2N_5O_{10}$, calculated $[M+H]^+ = 980.3319$, found = $[M+Na]^+ =$ 1002.3149 and with Boc removed m/z ratio correctly showing M/Z =824.2181

5.5.3.2 Synthesis of Natural-β-ala-unnatural, 3.20

This reaction proceeded as expected, the final product was purified by flash chromatography using a gradient of 0 to 50% EtOAc in hexane to obtain 4.6 mg of product. The reaction had a yield of 13%. The mass was confirmed using LC-MS and correlated the HPLC peak to product **3.18** with correct HRMS. The product purity was confirmed using and gave this dimer a purity of 87%. LRMS (ESI-QQQ): $C_{52}H_{55}Cl_2N_5O_{10}$, calculated $[M+Na]^+ = 1002.32$, found = (1002.4)

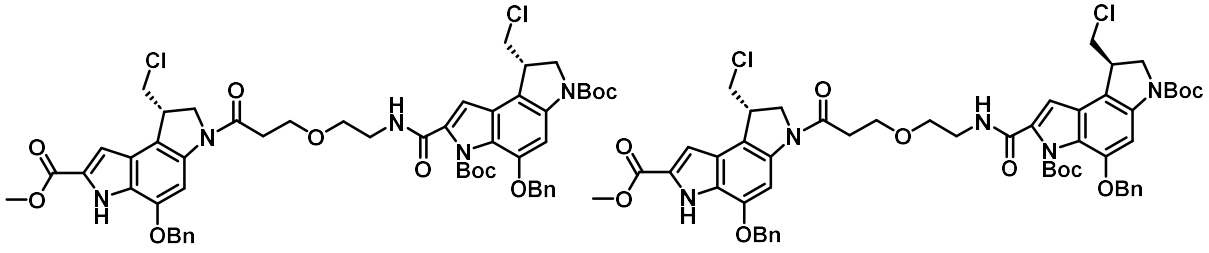
5.5.3.3 Synthesis of Unnatural-β-ala-unnatural, 3.19

This reaction proceeded as expected, the final product was purified by flash chromatography with a gradient of 0 to 50% EtOAc in hexane and yielded 15.9 mg of product with a yield of 46%. The mass was confirmed on LC-MS and correlated the HPLC peak to product **3.18** with correct HRMS. The product purity was confirmed

using HPLC and gave this dimer a purity of 86%. LRMS (ESI-QQQ): $C_{52}H_{55}Cl_2N_5O_{10}$, calculated, $[M+Na]^+ = 1002.32$, found = 1002.3

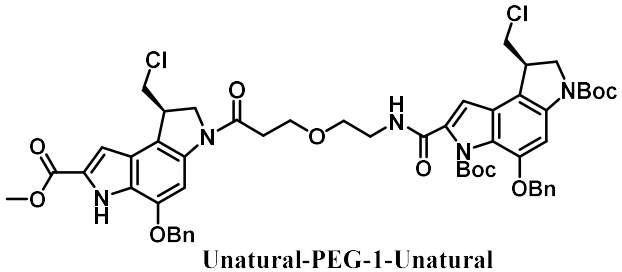
5.5.4. Synthesis of DSA-Peg 1-DSA Protected

The three different PEG dimers were synthesised using the same conditions.



Natural-PEG-1-Natural
3.22

Natural-PEG-1-Unnatural
3.24



Unatural-PEG-1-Unatural
3.23

The general procedure for the dimers (**3.22**, **3.23** or **3.24**) is as follows: Compound **2.25** dissolved in DCM and treated with TFA and TIPS solution (45% TFA in 52.5% DCM and 2.5 % TIPS). This was stirred for 1.5 hours and checked by HPLC for completion. Reaction worked up by removal of solvent under reduced pressure and trituration with cold ether. No purification required. This Boc deprotected analogue, **3.17** was dissolved in DMF and treated with PEG-1-Boc, **3.26** (1.2 eqv) and coupling agents, EDC (1.25 eqv), HOBT (1.25 eqv) and DIPEA (4 eqv). The progress was checked by HPLC and TLC, reaction worked up after ~1.5-3 hours. The mixture was diluted with water and extracted with EtOAc three times. The organic was then combined and concentrated before being subject directly to column chromatography. Product was purified using flash chromatography twice (First to remove excess impurities, second to isolate the pure compound). A gradient was run of 20 to 50% EtOAc in Hexane, the product was then collected in MeOH flush at the end. This concentrated product was then purified using a gradient of 0 to 3% MeOH in DCM to give **3.29**. This product was dissolved in DCM and treated with TFA and TIPS solution

(45% TFA in 52.5% DCM and 2.5% TIPS). This was stirred for 1.5 hours and checked by HPLC for completion. Reaction worked up by removal of solvent under reduced pressure and trituration with cold ether. No purification required. This compound, **3.30**, could then be coupled with the ester hydrolysed product, **2.29** (synthesised as described in 5.51). The final couplings were done by dissolving both compounds in DMF and treating with coupling agents, EDC (1.25 eqv), HOBT (1.25 eqv) and DIPEA (4 eqv). The progress was checked by HPLC and TLC, reaction worked up after ~1.5-3 hours. The differences in purifications, yields and purity are summarised in the following sections.

5.5.4.1 Synthesis of Natural-PEG-Natural, 3.22

This reaction proceeded as expected, purification was done using flash chromatography by wet loading in DCM onto the column with an isocratic gradient of 50% EtOAc in hexane giving 9.71 mg of product with a yield of 27%. The product purity was confirmed using HPLC and gave this dimer a purity of 83% . LRMS (ESI-QQQ): $C_{54}H_{59}Cl_2N_5O_{11,}$ calculated $[M+2H]^+$ = 512.69 , found = 512.7. HRMS (ESI): $C_{54}H_{59}Cl_2N_5O_{11}$, calculated $[M+Na]^+$ =1046.35, found = 1046.9

5.5.4.2 Synthesis of Natural-PEG-unnatural, 3.24

This reaction proceeded as expected, purification was done using flash chromatography by wet loading in DCM onto the column with an isocratic gradient of 50% EtOAc in hexane. This purification gave 8.5 mg of product with a yield of 24%. The mass was confirmed using LC-MS and correlated the HPLC peak to product **3.22** with correct HRMS. The product purity was confirmed using HPLC and gave this dimer a purity of 85%. LRMS (ESI-QQQ): $C_{54}H_{59}Cl_2N_5O_{11,}$ calculated $[M+Na]^+$ = 1046.32, found = 1046.9

5.5.4.3 Synthesis of Unnatural-PEG-unnatural, 3.23

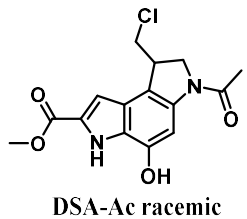
This reaction proceeded as expected, purification was done using flash chromatography by wet loading in DCM onto the column with an isocratic gradient of 50% EtOAc in hexane. This purification gave 10.57 mg of product with a yield of 44%. The mass was confirmed using LC-MS and correlated the HPLC peak to product

3.22 with correct HRMS. The product purity was confirmed using HPLC and gave this dimer a purity of 88%. LRMS (ESI-QQQ): $C_{54}H_{59}Cl_2N_5O_{11}$, calculated $[M+Na]^+$ = 1046.35, found = 1046.9

5.6 Synthesis of deprotected dimers

The benzylated DSA analogues were dissolved in CH_3OH, degassed with N_2 for 10 min. Pd/C added to the solution and it was degassed for a further 10 mins. Ammonium formate solution (25% in distilled water) was added and the rection was stirred vigorously. If there was product insolubility, then an appropriate volume of THF was added. The reactions were all monitored closely using HPLC and LCMS with a UV trace attachment. Reactions were worked up between 1.5-3 hours after the reactions were started. The workups were done by filtering off the palladium, then adding water to the reaction mixture before extracting three times with EtOAc. The combined organic layers were then concentrated under reduced pressure before being subject to different purification techniques. The purifications varied depending on the compounds and yields and data will be described in more detail in the following sections.

5.6.1. Debenzylation of DSA-Ac monomer, 4.1

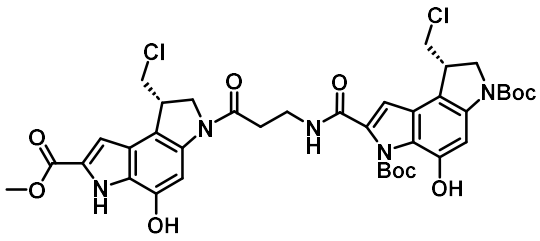


DSA-Ac racemic
4.1

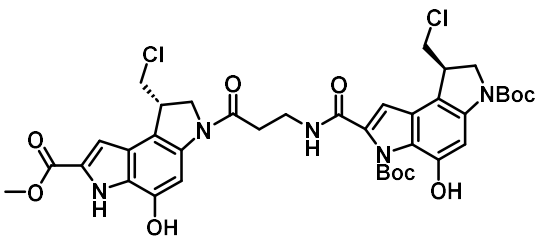
Compound **2.25** (26.8 mg) dissolved in MeOH (1 mL) and degassed with N_2 for 10 minutes, treated with Pd/C (16 mg) and further degassed for another 10 minutes. The reaction was treated with ammonium formate solution (0.5 mL 25% in distilled water). Filtered off Pd/C and solution was diluted with water and extracted 3 times with EtOAc. Solvent was removed under reduced pressure and purified using semi prep HPLC in acetonitrile. Due to the injections being limited to 30 microlitres at a time only part of the crude mixture was purified due to each injection taking 25 minutes. Therefore no yield for reaction could be calculated but 3.09 mg of clean product obtained. Purity confirmed using HPLC to show one peak, and product confirmed using HRMS and LCMS. HRMS (ESI): $C_{15}H_{15}ClN_2O$, calculated $[M+H]^+$ = 323.0752, found = 323.0713. LRMS (ESI-QQQ): $C_{15}H_{15}ClN_2O$, calculated $[M+H]^+$ = 323.08, found = 323.1

5.6.2. Debenzylation of DSA-β–alanine-DSA

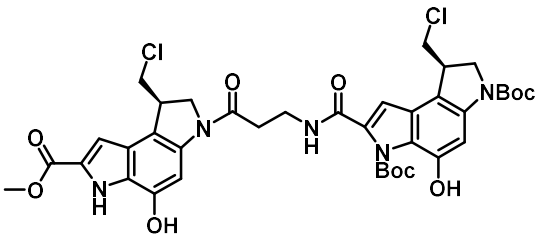
The three different β-alanine dimers were debenzylated using the same conditions as described. The procedure can be seen summarised in **5.6.1**. The differences in the purifications, yields and purity after reaction are summarised in the following sections.



Natural-β-ala-Natural
4.2



Natural-β-ala-Unnatural
4.3



Unatural-β-ala-Unatural
4.4

5.6.2.1 Enantiopure Natural DSA – β-alanine – Natural DSA TOXIC, 4.2

This reaction proceeded as expected and obtained 3.18 mg of product without final purification and a yield of 99%. The product purity was confirmed using HPLC showing a purity of 81%. HRMS (ESI): $C_{38}H_{43}Cl_2N_5O$, calculated $[M+2H]^+$ = 802.2621, found = 802.2743. LRMS (ESI-QQQ): $C_{38}H_{43}Cl_2N_5O$, calculated $[M+H]^+$ = 800.25, found = 800.3

5.6.2.2 Enantiopure Natural DSA – β-alanine – Unnatural DSA TOXIC, 4.3

This reaction proceeded as expected and obtained 2.75 mg of product without final purification and a yield of 64%. The product purity was confirmed using HPLC which showed a purity of 91.23%. HRMS (ESI): $C_{38}H_{43}Cl_2N_5O$, calculated $[M+2H]^+$ = 802.2621, found = 802.2643

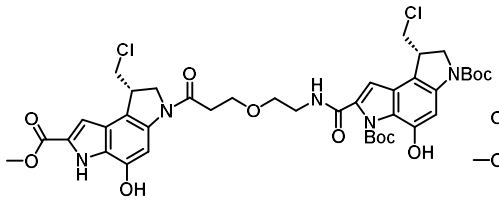
LRMS (ESI-QQQ): $C_{38}H_{43}Cl_2N_5O$, calculated $[M+H]^+$ = 800.25, found = 800.3

5.6.2.3 Synthesis enantiopure Unnatural- DSA – β-alanine – Unnatural DSA TOXIC, 4.4

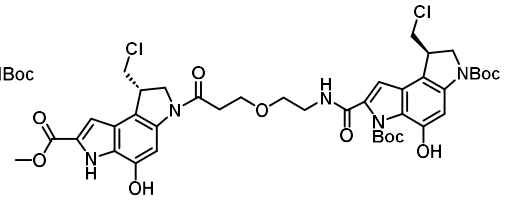
This reaction proceeded as expected and was purified using semi preparative HPLC which obtained 1.4 mg of product. The reaction had a yield of 39%. The product purity was confirmed using HPLC and gave this dimer a purity of 80%. The mass was confirmed on LCMS and HRMS. HRMS (ESI): $C_{38}H_{43}Cl_2N_5O$, calculated $[M+2H]^+$ = 802.2621, found = 802.2602. LRMS (ESI-QQQ): $C_{38}H_{43}Cl_2N_5O$, calculated $[M+H]^+$ = 800.25, found = 800.3

5.6.3. Synthesis of DSA-PEG-DSA toxic

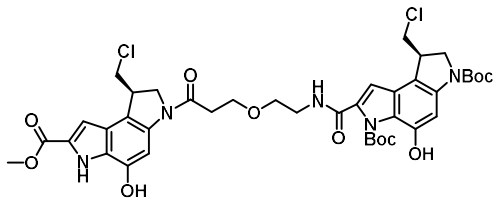
The three different PEG dimers were debenzylated using the same conditions. The general procedure can be described. The procedure can be seen summarised in **5.6.1**. The differences in the purifications, yields and purity after reaction are summarised in the following sections.



Natural-PEG-1-Natural
4.5



Natural-PEG-1-Unnatural
4.6



Unatural-PEG-1-Unatural
4.7

5.6.3.1 Synthesis enantiopure Natural DSA – PEG – Natural DSA TOXIC, 4.5

This reaction proceeded as expected, requiring no final purification, obtaining 3.31 mg of product with a yield of 68%. The product purity was confirmed using HPLC, and showed a purity of 86%. The mass was confirmed on LCMS and HRMS.
HRMS (ESI): $C_{40}H_{47}Cl_2N_5O_{11}$ calculated $[M+2H]^+$ = 846.2884, found = 846.2897
LRMS (ESI-QQQ): $C_{40}H_{47}Cl_2N_5O_{11}$, calculated $[M+3H]^+$ = 282.09, found = 282.1

5.6.3.2 Synthesis enantiopure Natural DSA – PEG – Unnatural DSA TOXIC, 4.6

This reaction proceeded as expected, requiring no final purification, obtaining 6.4 mg of product, with a yield of 98%. The product purity was confirmed using HPLC, and showed a purity of 75%. The mass was confirmed on LCMS and HRMS. HRMS (ESI): $C_{40}H_{47}Cl_2N_5O_{11}$ $[M+2H]^+$ = 846.2884, found = 846.2901. LRMS (ESI-QQQ): $C_{40}H_{47}Cl_2N_5O_{11}$, calculated $[M+3H]^+$ = 282.09, found = 282.1

5.6.3.3 Synthesis enantiopure Unnatural DSA – PEG – Unnatural DSA, 4.7

This reaction proceeded as expected, the crude was purified using semi preparative HPLC to obtain 1.79 mg of product. The reaction had a yield of 74%. The product purity was confirmed using HPLC which showed a purity of 99%. HRMS (ESI): $C_{40}H_{47}Cl_2N_5O_{11,}$ calculated$[M+2H+Na]^+$ = 868.2703, found = 868.2684. LRMS (ESI-QQQ): $C_{40}H_{47}Cl_2N_5O_{11}$ calculated $[M+3H]^+$ = 282.09, found = 282.3

5.7 MTS assay

The cells were seeded at 3 X 10^5 cells /mL in a 96-well plate and diluted from a stock solution using RPM1 media and incubated for 24 h at 37 °C. The cells were either treated with DMSO (negative control), treated with doxorubicin (positive control), or compounds **4.1-4.7** at a range of concentrations (**Table .1**). These were pipetted in triplicate and incubated for 72 h at 37 °C. The cells were then treated with the MTS reagent, and incubated for a further 4 hours at 37 °C. The plates were then read using the BMG Labtech POLARstarOPTIMA microplate reader, which reads at a 492 nm wavelength. The data values were collected for each well and the IC50 values for each compound could be calculated. The values were submitted on GraphPad Prism Version 6.0 software, using a four-parameter logistic nonlinear regression model. The results are from three different experiments carried out on different days with different stock solutions and different cell passages.

Compound	Stereochemistry	IC_{50}
4.1	Racemic monomer	**0.9839**
4.2	Natural- β-Ala-natural	**0.2387**
4.3	Natural- β-Ala unnatural	**0.02600**
4.4	Unnatural- β-Ala - unnatural	**1.131**
4.5	Natural-PEG-natural	**0.5589**
4.6	Natural-PEG-unnatural	**0.6642**
4.7	Unnatural-PEG-unnatural	**2.09**

Table 5.1: *Data from the MTS assays ran with relevant concentrations and data from analysis.*

5.8 Gold nanoparticles

5.8.1. Synthesis of gold nanoparticles

The synthesis of the citrate stabilised AuNPs was carried out using a method developed by Enüstün and Turkevich.[4] A solution of gold (III) trihydrate was prepared (12.5 mg, 32 µmol, in 100 mL distilled water). This was combined with a solution of trisodium citrate tribasic dihydrate dihydrate (50 mg, 168 µmol, in 50 mL distilled water). These combined solutions were stirred vigorously at 60 °C for 5 minutes. The solution was then heated to 85 °C and stirred for a further 2.5 hours. After cooling to room temperature, the now red nanoparticle solution was filtered through a 0.22 µm Miller GP syringe driven filter unit and this stock solution was stored in the fridge for future use.

5.8.2 Preparation of the custom strand thiolated DNA, dsDNA

There were two sequences used for the AuNP assay, the first was sequence was used a sequence readily available in the lab due to a PhD student following work done Marin and co-workers.[5] These two custom oligomers had been ordered from Sigma UK. The sequences are shown below with one strand having a thiol group at the 5' position:

5' - [ThiC6] - CTACGTGGACCTGGAGAGAGGAAGGAGACTGCCTG – 3'

3' – GATGCACCTGGACCTCTCTCCTTCCTCTGACGGAC – 5'

The other oligomers were ordered from Eurogentec and were custom strands containing an AT region at the end furthest from the custom thiolated end. The sequence of these oligomers is shown below:

5' – [ThiC6] – CGACGCGCGCGAGCTTAA – 3'

3' – GCTGCGCGCGCTCGAATT – 5'

Stock solutions of these oligomers were prepared separately at a concentration of 50 μM in 5 mM HEPES and 10 mM trisodium citrate buffer. The stock solution for the first sequences from Sigma UK were already prepared within the lab. The sequences from Eurogentec followed manufacturer's instructions to make the stock solutions. These were stored in a -80 °C freezer until they were ready to be annealed and used in the assay. To anneal these solutions 25 μL of each complimentary oligomer was combined in an Eppendorf and heated added to a water bath at 100 °C. This solution was then allowed to cool slowly to room temperature overnight and then kept in the fridge until used in the assay.

5.8.3. Synthesis of gold nanoparticles functionalised with thiolated DNA, dsDNA

The method used to functionalise the gold nanoparticles with the thiolated DNA followed work done by Zang and co-workers.[6] To 1.5 mL of gold nanoparticles the annealed DNA was added (31.5 μL of a 25 μM solution in trisodium citrate (10 mM, pH 3.0). This solution was thoroughly mixed for 1 minute. Then a trisodium citrate solution (30 μL of a 500 μM solution, pH 3.0) was added to the AuNPs and mixed thoroughly before letting the solution stand for 10 minutes. Complete functionalisation was confirmed using NaCl. Here NaCl solution (52.5 μL of a 2 M solution) was added

to the nanoparticle solution and immediately mixed before mixing for a further 20 minutes on a rotary mixer with a stirring bead. If the solution remained red then a further NaCl test was done (150 µL of a 2 M solution) and the mixing was repeated and added to the rotary mixer for 40 minutes. If this solution remained red this indicated that all the gold nanoparticles were fully functionalised as there was no aggregation. The solution was then centrifuged at 8,000 rpm for 30 minutes in an AllegraTM X-22R centrifuge, Beckman Coulter. The nanoparticles precipitated at the bottom of the Eppendorf. The supernatant was removed and a PB buffer was added (1.5 mL of 10 mM, pH 7.4), the solution was then centrifuged again, and the supernatant removed. This washing was repeated 3 times before storing the particles in a new PB buffer (1.5 mL of 10 mM, pH 7.4) an d stored in the fridge.

5.8.4. Titration of the functionalised AuNPs with the compounds being tested

The titration of the dimers and monomers was then tested with the functionalised nanoparticles. Stock solutions of DSA monomer **4.1** (negative control, acridine **4.10** (positive control), unnatural PEG **4.7** and natural PEG **4.5** were prepared at a concentration of 1 mM. Before any measurements were taken a baseline reading of the buffer was taken on the UV-Vis spectrometer. 570 µL of dsDNA functionalised AuNPs was added to a cuvette and an initial UV-Vis extinction recording was taken. Then a series of 10 titrations (1 µL of 1 mM stock solution) was added and shaken for 1 minute by hand before recording an extinction measurement for every titration. This procedure was repeated for all the compounds and the data recorded on a USB so the appropriate curves could be plotted.

5.9 Gel electrophoresis

5.9.1. Preparation of the buffer

The buffer was made up using TRIS HCl (100 mL milliQ water, 157.6 mg, 10 mM), to this NaCl (58.44 mg) added. Solution made up to pH 7.5 using pH probe.

5.9.2. Annealing custom labelled single stranded oligonucleotides

Two different DNA sequences were needed to test for DNA binding and DNA crosslinking. Four single stranded oligonucleotides were ordered two with the AT

sequence in middle and AT sequence at the end. The two complementary strands for the AT sequence in the middle:

5’ - [6-FAM] - CGACGCGTTAACGTTAACGCGC – 3’
3’ – GCTGCGCAATTGCAATTGCGCG – 5’

The two complementary strands for the AT sequence at the end:

5’ - [6-Fam] - CGACGCGCGCGAGCTTAA – 3’
3’ – GCTGCGCGCGCTCGAATT – 5’

Stock solutions of 100 μM were prepared for each oligonucleotide in buffer (10 mM Tris-HCl (pH 7.5), containing 10 mM NaCl following manufacturers (Eurogentech) guidelines. These were stored in a -80 °C freezer until they were ready to be annealed and used in the assay. To anneal these solutions 10 μL of each complimentary oligomer was combined to 980 μL of the buffer in an Eppendorf and heated added to a water bath at 100 °C. This solution was then allowed to cool slowly to room temperature overnight and then kept in the fridge until used in the assay.

5.9.3. Thermal cleavage assay

The thermal cleavage assay was attempted using several concentrations of monomer and dimer. After the oligomer sequence with the AT sequence in the middle had been annealed, separate eppendorfs of 100 μL of annealed oligomerand 1 μL of each compound in desired concentration (diluted in DMSO). Incubation was done for 2 hours at 37 °C. The compounds that were being tested for the cleavage were then heated at 90 °C for 3 minutes before allowing them to cool. The samples were then stained with a DNA Gel loading dye by ThermoFisher (mixture of two dyes bromophenol blue and xylene cyanol FF) (1 μL) so they could be visualised during the electrophoresis and loading. The samples were loaded onto a pre-prepared 10% TBE polyacrylamide gel. The gel was loaded in the electrophoresis mini tank with a 10% diluted TBE Buffer (Tris-borate EDTA) in milli-Q water. The gels were transported in milli-Q water before reading in the Image Quant LAS 4000.

5.9.4. DNA crosslinking assay

The thermal cleavage assay was attempted using several concentrations of monomer and dimer. After the oligomer sequence with the AT sequence at the end had been annealed, separate eppendorfs of 100 μL of annealed oligomerand 1 μL of each compound in desired concentration (diluted in DMSO). Incubation was done for 2 hours at 37 °C. The samples were then stained with a DNA Gel loading dye by ThermoFisher (mixture of two dyes bromophenol blue and xylene cyanol FF) (1 μL) so they could be visualised during the electrophoresis and loading. The samples were loaded onto a pre-prepared 10% TBE polyacrylamide gel. The gel was loaded in the electrophoresis mini tank with a 10% diluted TBE Buffer (Tris-borate EDTA) in milli-Q water. The gels were transported in milli-Q water before reading in the Image Quant LAS 4000.

References

1 M. J. Stephenson, L. A. Howell, M. A. O'Connell, K. R. Fox, C. Adcock, J. Kingston, H. Sheldrake, K. Pors, S. P. Collingwood and M. Searcey, Solid-Phase Synthesis of Duocarmycin Analogues and the Effect of C-Terminal Substitution on Biological Activity, *J. Org. Chem.*, 2015, **80**, 9454–9457.

2 M. S. Tichenor, J. D. Trzupek, D. B. Kastrinsky, F. Shiga, I. Hwang and D. L. Boger, Asymmetric Total Synthesis of (+)-and ent-(-)-Yatakemycin and Duocarmycin SA: Evaluation of Yatakemycin Key Partial Structures and Its Unnatural Enantiomer, *J. Am. Chem. Soc*., 2006, **128**, 15683-15696

3 D. Barker, A. L. Lehmann, A. Mai, G. S. Khan and E. Ng, Synthesis of non-symmetrical 3,5-diamidobenzyl amines, ethers and sulfides, *Tetrahedron Lett.*, 2008, **49**, 1660-1664

4 B. V. Enustun and John. Turkevich, Coagulation of Colloidal Gold, *J. Am. Chem. Soc.*, 1963, **85**, 3317–3328.

5 M. J. Marín, B. D. Rackham, A. N. Round, L. A. Howell, D. A. Russell and M. Searcey, A rapid screen for molecules that form duplex to duplex crosslinks in DNA †, *Chem Commun*, 2013, **49**, 9113.

6 X. Zhang, M. R. Servos and J. Liu, Instantaneous and quantitative functionalization of gold nanoparticles with thiolated DNA using a pH-assisted and surfactant-free route, *J. Am. Chem. Soc.*, 2012, **134**, 7266–7269.