# DESIGN AND DNA BINDING OF MODEL ANTI-CANCER DRUGS

Submitted by
ELIZABETH ANKERS
B. Sc. (Hons) – La Trobe University

A thesis submitted in total fulfilment
of the requirements for the degree of
Doctor of Philosophy

School of Molecular Science
Faculty of Science, Technology and Engineering

La Trobe University
Bundoora, Victoria 3086
Australia

October, 2011

A table from pages 195 and 196 of this thesis had to be removed due to copyright issues.

# TABLE OF CONTENTS

## SYMBOLS AND ABBREVIATIONS

| | |
|---|---|
| 1D | One-Dimensional |
| 2D | Two-Dimensional |
| $^{31}P$ | Phosphorus (NMR) |
| 3D | Three-dimensional |
| 6-311G | Basis set |
| 6-31G | Basis set |
| 6-mer | Six base pair long oligonucleotide |
| 8-mer | Eight base pair long oligonucleotide |
| Å | Angstroms |
| a.u. | Atomic Units |
| *ab initio* | From first principles |
| AM1 | Austin Model 1 semi-empirical method |
| AMPAC | Semi-empirical quantum chemistry program |
| AO | Atomic orbital |
| aq | Aqueous |
| Asp | Aspartic acid |
| C2-DACA | *N*-(2-aminoethyl)acridine-4-carboxamide |
| C4-DACA | *N*-(4-aminobutyl)acridine-4-carboxamide.oxalate |
| CBS-Q | Complete basis set |
| COSY | Correlation Spectroscopy |
| ct DNA | Calf Thymus DNA |
| CVFF | Consistent Valence Forcefield |
| d6-DMSO | Deuterated dimethylsulfoxide |
| DACA | *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide |
| DEPT | Distortion Enhancement by Polarisation Transfer |
| $\Delta H_T$ | Full thermal correction |
| DMSO | Dimethylsulfoxide |
| ds | Double stranded |
| ε | Extinction coefficient |
| *E. coli* | *Escherichia coli* |
| $E_0$ | Overall reaction energy at 0 K |
| $E_A$ | Activation enthalpy |
| EDTA | Ethylenediaminetetraacetic acid |
| $E_e$ | Electronic reaction energy |
| ESI | Electrospray Ionization (Mass Spectrometry) |
| FDA | The Food and Drug Administration |
| FLT | Full length transcript |
| G3/G4 | Gaussian-*n* methods |
| HB101 | *E. coli* competent cells of strain HB10 |
| HCT-8 | Human ileocecal colorectal adenocarcinoma cell line |
| $H_f$ | Enthalpy of formation |
| HF | Hartree-Fock |
| HPLC | High performance liquid chromatography |
| $H_{rxn}$ | Enthalpy of reaction |
| HSQC | Heteronuclear Single Quantum Correlation |
| $ID_{50}$ | Median infective dose |
| ILS | Increased Life Span |

| | |
|---|---|
| i-PrOH | Isopropanol |
| J | Coupling constant |
| *lac* UV5 | Mutated bacterial vector promoter region |
| LB Broth | Luria-Bertani broth |
| m/z | Mass to charge ratio |
| MacGAMESS | *Ab initio* quantum chemistry program |
| MacMolPlot | Molecular builder for MacGAMESS |
| MAR07 | Synthetic daunomycin derivative |
| MIDI | Basis set |
| MilliQ | Millipore ultrapure water |
| MINI | Basis set |
| MM | Master mix |
| mmHg | Millimetres of mercury |
| MP2 | Moller-Plesset pertubation theory method |
| NOE | Nuclear Overhauser Effect |
| NOESY | Nuclear Overhauser Effect Spectroscopy |
| NTP | All four ribonuclease 5'-triphosphates |
| ρ | Rho (protein factor) |
| P388 | Murine lymphocytic leukaemia cell line |
| PBS | Phosphate buffered saline |
| pCC1 | Cloning vector constructed from pSP64 |
| pdb | Protein Data Bank file |
| ppm | Parts per million |
| pRW1 | Restriction fragment from pSP64 |
| pSP64 | Standard cloning vector |
| Sarcoma-180 | Murine solid sarcoma cell line |
| SGI | Silicon Graphics Inc. |
| SN-07 | Barminomycin |
| ss | Single stranded |
| STO-3G | Basis set |
| $t_{1/2}$ | Half-life |
| TAE buffer | Tris-acetate-EDTA buffer |
| TBE buffer | Tris-borate-EDTA buffer |
| Tc buffer | Transcription buffer |
| TE buffer | Tris-EDTA buffer |
| TEMED | Tetramethylethylenediamine |
| TLC | Thin Layer Chromatography |
| TOCSY | Total Correlation Spectroscopy |
| Topo | Topoisomerase |
| Tris | 2-Amino-2-hydroxymethyl-propane-1,3-diol |
| TS | Transition state |
| $\mu M_{bp}$ | Micromolar base pair |
| WebMO | Web interface for Gaussian program |

**SUMMARY**

This thesis describes the design and measurement of the DNA binding properties in the presence of formaldehyde of several acridine derivatives with primary amino-containing side chains. An aim of this work was to emulate the activation of Adriamycin by formaldehyde and consequent binding to DNA, using simpler, easier to synthesise compounds.

Two derivatives of the anti-tumour agent *N*-[2-(dimethylamino)ethyl] acridine-4-carboxamide (designated C2-DACA and C2-DACA dimer in this thesis) were studied. The length of the side chain was varied, and the tertiary amine was modified to a primary amine. This was to facilitate interaction with formaldehyde and subsequent covalent binding of the formaldehyde-activated drug to DNA.

The DNA-binding abilities of a derivative *N*-(2-aminoethyl)acridine-4-carboxamide.oxalate, C2-DACA, and its dimer were investigated using radio-labelled DNA electrophoretic crosslinking assays, and binding studies based on absorbance, but did not show DNA binding. Molecular modelling showed that the side chain of C2-DACA was too short and under considerable steric hindrance once intercalated and covalently bound to DNA. The length of the side chain proved to be insufficient to reach the preferred binding site. Molecular modelling also showed that extending the chain by 2 or 3 carbons would allow favourable interactions between the intercalating acridine derivative and DNA, and allow the side chain to covalently bind to the DNA. The compound arising from the molecular modelling *N*-(4-aminobutyl) acridine-4-carboxamide, C4-DACA, was synthesised in a similar manner to C2-DACA. DNA binding experiments including an *in vitro* transcription assay showed that formaldehyde-activated C4-DACA bound to DNA.

The energetics of Schiff base and aminal formation were calculated using quantum mechanic molecular orbital methods. Formaldehyde was found to be the most energetically favourable aldehyde in both enthalpy and entropy. The overall binding reaction of a primary amine reacting with formaldehyde, and subsequently with guanine to produce an aminal was exothermic. Most of this energy came from the Schiff base formation, with an additional smaller amount from the subsequent reaction of the imine with the exocyclic amine of guanine.

## STATEMENT OF AUTHORSHIP

Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis submitted for the award of any other degree or diploma.

No other person's work has been used without due acknowledgement in the main text of the thesis.

This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

……. 21 October 2011

ELIZABETH ANKERS

## ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisors, Prof. Robert Brownlee and Prof. Don Phillips for their continual support and guidance. Big thanks go to Dr David Wilson for his expert advice with the Gaussian 09 calculations. I'd like to thank the NMR/computer modelling crew of Dr Jason Dang, Andrew Box, Dr Daniel Oehme and the various honours students for making the time in the lab so enjoyable. Viva la Friday lab lunch! I'm also grateful for the intensive month of help from Michel Rajzmann during the AMPAC theoretical calculations.

Secondly, I'd like to thank Dr Les Deady for his guidance with the synthesis, and for synthesising all of the drugs used in this thesis. I'm also grateful to Dr Belinda Abbott for letting me loose in her lab, and to her students Elisse Browne and Lori Ferrins for their advice, assistance and encouragement.

I also thank Dr Suzi Cutts, Dr Benny Evison and Dr Kate Coldwell for their assistance and advice during my time in the DRP lab, with an extra special thanks to Benny for performing the *in vitro* transcription assay.

Thanks and love goes to my family for their continual support and encouragement through the long haul, and for providing a place to retreat to and recharge.

To all my friends, old and new, for their friendship and advice throughout the years, I say a very special thankyou. I must admit that it felt like this PhD would never end, and I'm sure I would have given up long ago without all your help and encouragement.

Everything that has a beginning has an end. I see the end coming.

The Oracle, The Matrix Revolutions

## LIST OF PUBLICATIONS

### Conferences and Presentations

1. Bioactive Discovery in the New Millennium, Lorne February 2002
   "Studies Towards New Anti-Cancer Drugs"
   Ankers, E. A.; Brownlee, R. T. C.; Phillips, D. R.

2. ANZMAG Barossa Valley, February 2004
   "Development of New Anti-Cancer Drugs"
   Ankers, E. A.; Brownlee, R. T. C.; Phillips, D. R.

3. ANZMAG Murramarang Resort, February 2006
   "Anti-Cancer Drug Design"
   Ankers, E. A.; Brownlee, R. T. C.; Phillips, D. R.

# CHAPTER 1. GENERAL INTRODUCTION

The aim of this work was to find model acridine based compounds that intercalate with and covalently bind to DNA. Such compounds would mimic the binding, activation, and cytotoxic response of model chemotherapeutic drugs. The most effective drugs currently used in chemotherapy are expensive and difficult to synthesise due to their enantioselectivity and copious synthetic steps. While the world searches for improved cures for cancer, we can also try to discover or create some cheaper and more target-specific alternatives to these expensive drugs.

One of the most effective chemotherapy drugs in clinical use is Adriamycin (doxorubicin) [1]. Crystal structures of Adriamycin intercalated in DNA show that the side chain lies in the minor groove with the axis of the anthracycline moiety intercalating perpendicular to the base pair axis. This side chain was found to be activated by formaldehyde *in vitro* [2-7] and to form a covalent link with the free $NH_2$ of guanine which extends into the minor groove of DNA [8].

This discovery inspired the study of a family of acridine based compounds which had promising activity with cancer cell lines in culture [9, 10]. The acridine moiety intercalates perpendicular to the base pairs, similar to the anthracyclines. The majority of crystal structures of acridine derivatives intercalated with DNA have the side group in the major groove, but there is also potential for binding via the minor groove. In this work molecular modelling and energetic studies were used to assess drug-DNA interactions in order to identify suitable drug structures. This study was directed towards elucidating energetic and steric factors involved in the covalent binding of model compounds to DNA. Multiple biochemical methods were utilised to discover if the acridine based drug intercalated into DNA or bound covalently to DNA in the presence of formaldehyde, and to elucidate the optimum conditions for drug-DNA formation.

Nuclear magnetic resonance spectroscopy was employed to give 3D structural information of the drug-DNA complex. Molecular modelling software was utilised to obtain the structure of a 6-mer of DNA in the presence of the acridine based drug and formaldehyde.

This chapter presents a review of the structure and mode of action of DNA intercalating compounds and describes intercalators that also bind covalently to DNA.

## 1.1. General Aspects

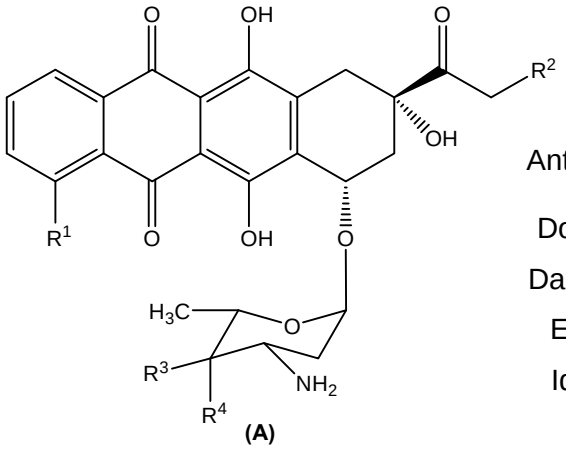
The nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), are the chemical carriers of a cells genetic information. Coded within DNA is the information needed to determine the nature of the cell, control cell growth and division, and instruct biosynthesis of all proteins and enzymes required for normal cellular function. It is this that makes DNA and its role in biological functions so interesting, and exceedingly important. There are many different ways in which this process can be interfered with, and where mutation of the DNA occurs, leads to the possibility of uncontrolled cell growth (i.e. cancer).

Topoisomerases are vital cellular enzymes. They prevent DNA strands from becoming tangled by cutting the DNA, which causes the DNA to wind and/or unwind. Topoisomerases adjust the topological structure of DNA during transcription, replication and chromosome compaction, and are regarded as housekeeping genes [11]. The result of inhibition of topoisomerases is cell death, thus making topoisomerases a potential target for chemotherapy. Topoisomerase inhibitors are among the most active anticancer agents in use today.

Topoisomerases are located in the nucleus of cells and function to prevent entanglement of DNA strands during DNA replication, RNA synthesis from DNA (transcription), exchanges of DNA segments between chromosomes (recombination) and elimination of erroneous DNA sequences (repair) [11]. Inhibition of topoisomerases disrupts these essential processes, and consequently leads to cell death. DNA topoisomerases catalyse the regulation of the topology of double-stranded DNA [11]. The two types of topoisomerases in mammalian cells are classified as topoisomerase (Topo) I and II. Topo I induces single strand breaks in DNA, whereas Topo II induces double strand breaks, and both cause topological changes in DNA [11].

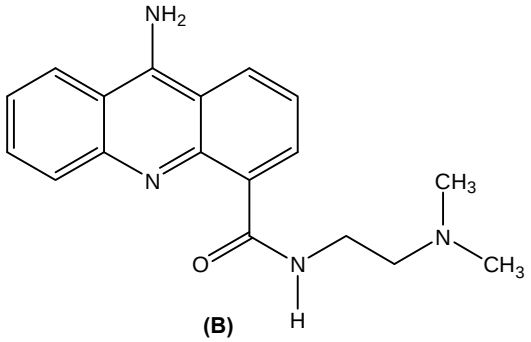
The activity of various enzymes requires the DNA to be in a specific 3D structure for replication or transcription to occur. Low molecular weight ligands that intercalate, reversibly and irreversibly, can inhibit the replication and

expression of DNA and RNA. Some examples of reversible ligand intercalators are anthracyclines (**A**), acridines (**B**), and actinomycin D (**C**), as shown in **Figure 1.1**.

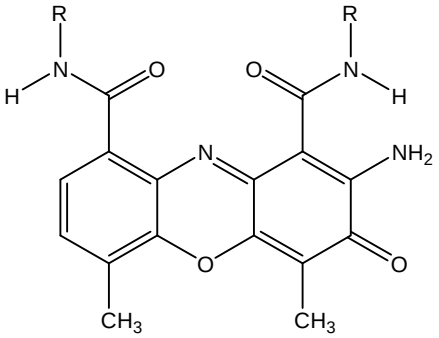



(A)

| Anthracycline | R1 | R2 | R3 | R4 |
|---|---|---|---|---|
| Doxorubicin | $OCH_3$ | OH | H | OH |
| Daunorubicin | $OCH_3$ | H | H | OH |
| Epirubicin | $OCH_3$ | OH | OH | H |
| Idarubicin | H | H | H | OH |




(B)




R = cyclo(Thr-D-Val-Pro-Ser-MeVal)

(C)

**Figure 1.1** Chemical structures of anthracyclines (**A**), *N*-[2-(dimethylamino)ethyl]-9-aminoacridine-4-carboxamide (**B**), and actinomycin D (**C**).

Several of these ligands bind selectively to specific DNA sequences, whereas others may only favour GC rich areas. An example of this specificity is in the clinical anticancer agent mitoxantrone, which intercalates specifically at CpA and CpG sequences [12, 13]. The incorporation of such sites into small oligonucleotides of DNA has allowed the interactions of such ligands to be studied in detail using NMR [14-18].

The structure and dynamics of the DNA strand are integral in determining how a ligand will interact. Thus, the changes that occur when a ligand intercalates are very important. For a ligand to intercalate, the DNA must undertake major

local conformational changes, such as unwinding, unfolding, and in some cases, disruption of the base pair hydrogen bonding [19-21].

An example of large structural changes on binding is seen with the anthracycline antibiotic nogalamycin (**Figure 1.2**), which is produced by *Streptomyces nogalater*. It contains two sugar moieties, and has an overall structure similar to that of a dumbbell (**Figure 1.3**). One end of nogalamycin contains both a methyl ester and a nogalose sugar, while the other end has a positively charged bicyclo amino sugar fused at the 1 and 2 positions (**Figure 1.2**) [19]. The flat anthracycline core intercalates between the DNA base pairs, accompanied by extensive unwinding of the DNA. The planarity of the B-DNA base pairs is well preserved.

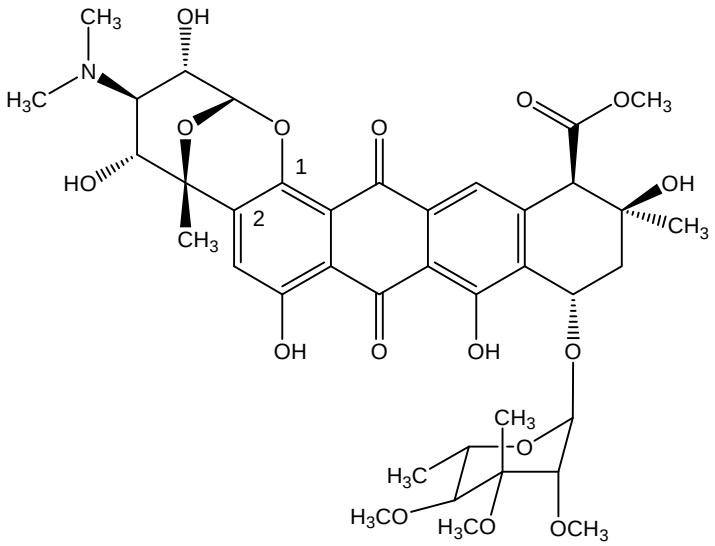



**Figure 1.2** The chemical structure of nogalamycin.

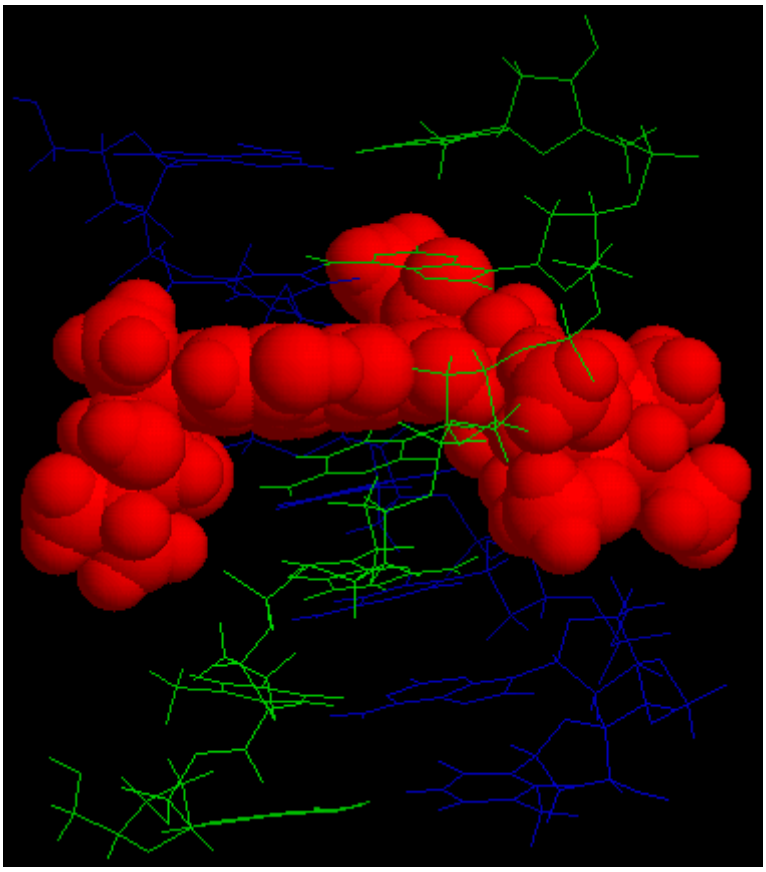


**Figure 1.3** Nogalamycin intercalated into DNA [15].

Nogalamycin is thought to require a local unwinding and denaturing of the local DNA base pairs in order to enable it to thread itself through the helix and intercalate, and once intercalated, inhibits both DNA replication and transcription [21]. Because of its dumbbell shape, nogalamycin essentially locks itself into the DNA, with interactions in both grooves imparting a high binding affinity and also extremely slow dissociation kinetics, requiring a major structural disruption of the DNA for both association and dissociation events to occur [15, 20]. The major structural disruption of the DNA required for dissociation is thought to be the reason for the high binding affinity, leading to its high potent biological activity and high level of cytotoxicity compared to other members of the anthracycline family [15].

Studies into the interactions between chemotherapeutic drugs and DNA have provided a greater understanding of their structure and dynamics, and ultimately of how the drugs operate. Understanding the localised spatial requirements for favourable interactions of small drugs in the DNA binding receptor may lead to the discovery of more medicinally effective

chemotherapeutics with enhanced binding abilities and reduced toxicological effects.

## 1.2. Anthracyclines (Adriamycin and Daunomycin)

Chemotherapy of cancer started after the second World War, and although its concept was simple (i.e. killing cancer cells), the chemotherapeutic arsenal for the treatment of cancer is still rather limited [11]. The first anticancer drugs were obtained from plants and fungi. The pigmented compound actinomycin A was isolated from *Streptomyces* cultures, and was found to exhibit antitumour activity [11]. Consequently, various antitumour antibiotics were isolated from fermentation broth of various *Streptomyces* species [11]. This included doxorubicin (Adriamycin), daunorubicin (daunomycin), nogalamycin, actinomycin D, respinomycin, mitomycin C and bleomycin [11, 21-23].

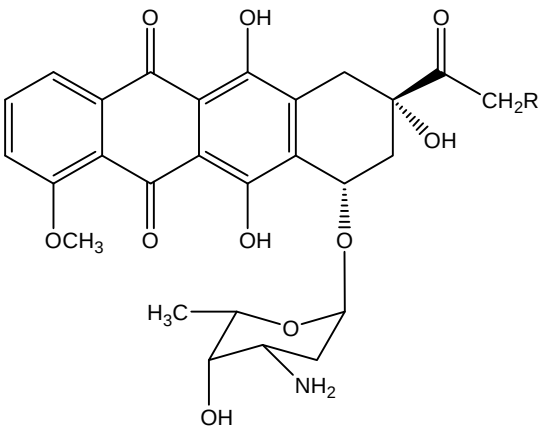



**Figure 1.4** Structure of Adriamycin [23] and daunomycin [24].

DNA intercalating agents are important types of anticancer drugs. The anthracycline antibiotics, of which daunomycin and Adriamycin are the parent compounds (**Figure 1.4**), are widely used in cancer chemotherapy [25]. They were isolated from *Streptomyces peucetius* in Italy and France in the 1960s [11]. Daunomycin is mainly used to treat acute lymphocytic and myelocytic leukaemias [11, 19]. Adriamycin is much more versatile and shows clinical activity against carcinomas of the breast, lung (small-cell lung cancer), ovary and bladder, sarcomas of bone and soft tissues, various childhood malignancies

(neuroblastoma, Wilms' tumour, Ewing sarcoma), acute leukaemias, Hodgkin and non-Hodgkin lymphomas and multiple myeloma [19, 26].

Daunomycin and Adriamycin are possibly the best characterised DNA intercalators. The mechanism of action of anthracyclines within the cell remains in debate even though it has been discussed extensively [27-29]. There has also been substantial investigation into the effect of anthracyclines inhibiting Topo I and II [28, 30-32]. Although these drugs have been extremely useful in the treatment of cancer, there have been problems associated with the unwanted side-effects of cardiotoxicity and drug resistance [26, 33]. Consequently, there has been extensive research into finding anthracyclines with improved pharmacological properties compared to the parent compounds.

To date there has only been two anthracycline analogues synthesised which are comparable to their parent compounds of daunomycin and Adriamycin. These are epirubicin and idarubicin (**Figure 1.5**).

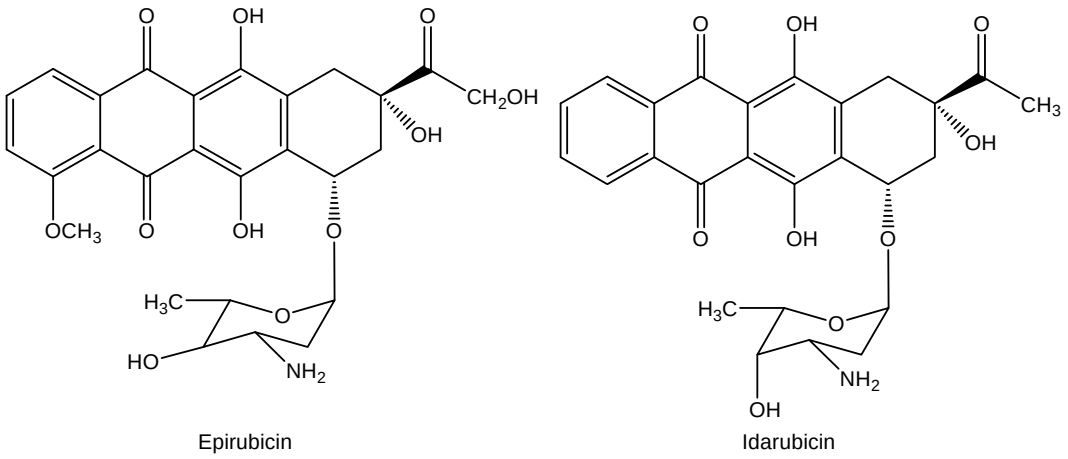



**Figure 1.5** Structures of Epirubicin and Idarubicin.

The only difference between Adriamycin and epirubicin is in the amino sugar moiety, where the hydroxyl group at the 4' position is in the L-configuration [11]. Epirubicin is used predominantly in the treatment of solid tumours, especially breast cancer, much the same as the parent compound Adriamycin. The potency and efficiency of epirubicin is less than that of Adriamycin. It also has the same drug resistance pattern as Adriamycin, so may not be utilised for the treatment of drug resistant tumours.

Idarubicin is used in the treatment of acute leukaemias, though its abilities are yet to be fully evaluated [11]. It differs from daunomycin only by the lack of a methoxy group in the four-ring anthracycline structure. Idarubicin is yellow in colour due to the lack of the methoxy group, is able to cross the blood brain barrier, and may also be given orally (unlike the acid labile parent compounds) [11]. Although idarubicin exhibits the typical anthracycline toxicity, oral use reduces the severity [11].

### 1.3. Acridines and Acridones

Anticancer drugs like Adriamycin and daunomycin are still obtained in the same manner as when originally discovered (i.e. extracted from cultures). Synthesising anthracyclines from scratch is a very costly task in both time and money. Their complex substituents require multiple protection and deprotection steps, while some reactions may not afford chiral centres with high enantioselectivity. The number of steps in a scheme for such elaborate molecules makes the complete synthesis a major task. This is one reason why the search for anticancer drugs has headed toward smaller, more easily synthesised molecules. The broad class of acridines and acridones are a starting point for chemotherapeutics that are more economically favourable to synthesise.

A major problem in cancer chemotherapy today is the treatment of solid tumours. Ligands do not respond any where near as well to chemotherapy of solid tumours as they do for disseminated diseases such as leukaemias and lymphomas [34]. There are several reasons for this, including a higher “intrinsic resistance” of many carcinoma cell lines to common agents compared to leukaemia cell lines, the problems of hypoxic, noncycling cells in solid tumours (“environmental resistance”), and enhanced genetic instability leading to “acquired resistance” [34]. Another important factor is the requirement for drugs to penetrate poorly vascularised solid-tumour masses [34].

Amsacrine, **Figure 1.6**, is a non-chiral antineoplastic with a low molecular weight, and is used to treat acute adult leukaemia [35]. Amsacrine has an electron rich side chain that is capable of forming electron-transfer complexes [36].

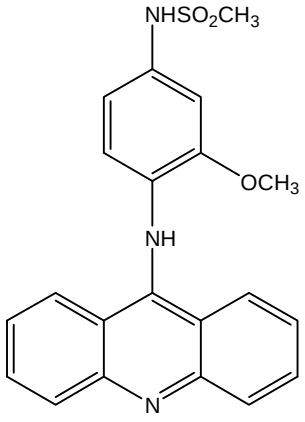



**Figure 1.6** Structure of amsacrine.

The amsacrine analogue *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide (DACA) shown in **Figure 1.7**, has a greatly extended spectrum of activity against solid tumours compared to that of amsacrine, which may be due to its more weakly basic acridine nucleus [34].

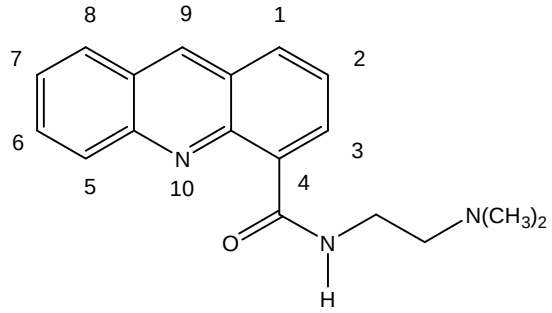



**Figure 1.7** The structure of *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide (DACA) and substituent position numbering.

DACA can be synthesised with relative ease, quickly and inexpensively. DACA is a Topo I and II poison [37-54]. In 1995, DACA became the first synthetic dual inhibitor to commence clinical trials [54]. These agents increase the concentration of Topo-generated DNA breaks that are normally short-lived intermediates in enzyme catalysis [39], and inevitably initiate programmed cell death pathways. DACA is able to circumvent several types of drug resistance, appears to cross the blood-brain barrier, is water soluble and is easily formulated [36, 45-51, 54-56]. The diluted drug is stable at room temperature (pH 7) for at least 6 months [50, 55]. DACA was found to have high anti-leukemic and solid tumour activity, with an $ID_{50}$ against murine leukaemia cell models (L1210 cells)

of 15 nM, and against HCT-8 human colon adenocarcinoma cells in culture of 55 nM [10, 36, 40, 55, 56]. It also proved highly active against the P388 leukaemia model *in vivo*, giving an increased life span (ILS) value in mice of around 100% at the optimal dose of 4.5 mg/kg [10, 55], and has curative activity against Lewis lung carcinoma [36, 48, 51, 54-56].

An absolute requirement for anti-tumour activity in the acridine-4-carboxamide class appears to be a suitably placed side chain. Studies of acridine substitution patterns among the 9-aminoacridine-4-carboxamides (e.g. **Figure 1.8**) revealed that pharmacological properties were enhanced with substituents in the 5-position [10].

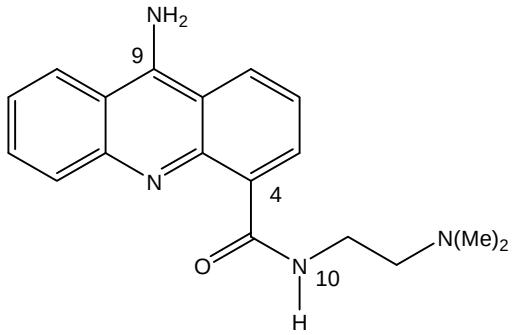



**Figure 1.8** An example of a 9-aminoacridine-4-carboxamide.

In culture, analogues with 5,7-disubstitution (numbering shown in **Figure 1.7**) show both the higher cytotoxic potency of 5-monosubstituted analogues, and the broad-spectrum cell line effectiveness shown by DACA and 7-monosubstituted analogues [45, 54]. Other necessities are that the side chain is at the acridine 4 position, the carboxamide has an unsubstituted NH group, and there are two methylene groups between the carboxamide NH and the terminal protonated N,N-dimethylamino group [10]. In the 9-amino-DACA derivative, the carboxamide group forms an internal hydrogen bond with the protonated N10 atom of the acridine ring. The side chain lies in the DNA major groove with its protonated N,N-dimethylamino group forming hydrogen bonding interactions with the O6 and N7 atoms of guanine G2 [57].

The related compound mitoxantrone is a synthetic anthracenedione antibiotic, with a typical quinone-semiquinone structure, but lacks the fourth ring of the anthracyclines, including the amino sugar [11, 13, 58-60]. Instead, it has two substituted basic aminoalkylamino groups (**Figure 1.9**).

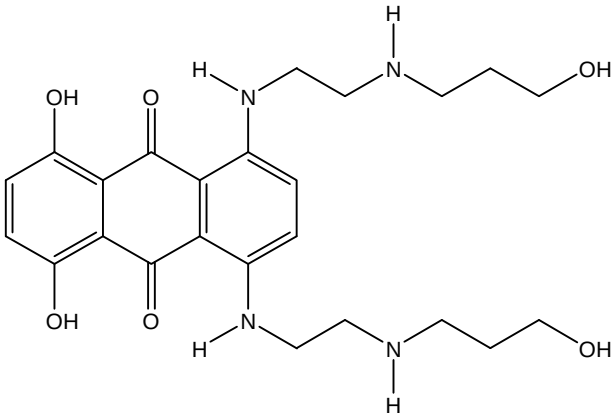



**Figure 1.9** Structure of mitoxantrone.

In comparison to the anthracyclines, the major dose-limiting factor for mitoxantrone is leucopoenia (lowering of leukocytes in the blood), with side-effects including myelosuppression (bone marrow suppression) and cardiotoxicity [58, 61]. Mitoxantrone has an improved clinical tolerability, and on a molar basis is more cytotoxic than anthracyclines [61]. Clinical use showed, however, that there was no clear advantage of using mitoxantrone over the anthracyclines, except in the case of prostate cancer [11]. It shows promise in the treatment of acute non-lymphocytic leukaemia, melanoma, advanced breast cancer, non-small cell lung cancer, and non-Hodgkin's lymphomas [13, 60]. The blue coloured drug is delivered intravenously, and "poisons" Topo II, causing breaks in DNA strands [62, 63].

Mitoxantrone can also be activated by formaldehyde to form drug-DNA adducts. However, excessively high levels of drug are required in order to generate adducts, which makes this an unlikely biological mechanism of action. An analogue of mitoxantrone with potential to form adducts more efficiently is the 2-aza-anthracenedione pixantrone (**Figure 1.10**) which is currently undergoing Phase III clinical trials for the treatment of aggressive non-Hodgkin's lymphoma [64].

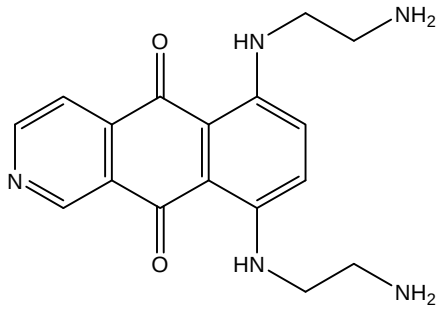



**Figure 1.10** Structure of pixantrone.

The cardiotoxicity of mitoxantrone was attributed to the 5,8-substituents (hydroxyl groups) [61, 64, 65]. Notable differences in structure between mitoxantrone and pixantrone are that the side chains in the latter are shorter, and that both ends possess a free amine. This minor change in functional group would allow for formaldehyde activation and consequent reaction with DNA. Pixantrone is efficiently activated by formaldehyde to generate covalent drug-DNA adducts with the ability to stabilise the DNA double helix [61]. Formaldehyde is an absolute requirement for pixantrone-DNA adduct formation, and has a 10 to 100 fold greater tendency to form drug-DNA adducts when compared with equimolar formaldehyde and mitoxantrone concentrations [61]. Preclinical studies confirmed that pixantrone had a greater activity than that of doxorubicin or mitoxantrone against haematologic neoplasias in a mouse model [64]. A Phase I study established that pixantrone was well tolerated and showed encouraging antitumor activity in heavily pretreated patients with relapsed/refractory malignant lymphoma [64]. A Phase II study on the activity of pixantrone as a single-agent chemotherapeutic against multiply relapsed and aggressive non-Hodgkin's lymphoma elucidated myelosuppressive toxicity after the second dose [64].

## 1.4. Covalent Binding to DNA

Several ligands have the ability to bind to DNA, forming a covalent aminal linkage (N-C-N) between the drug and DNA, which results in an enhanced cytotoxic activity in cell studies. The capacity to form an imine linkage can be inbuilt in the ligand (e.g. barminomycin shown in **Figure 1.11**), or can be introduced in the laboratory by the use of aldehydes, where Adriamycin is activated by formaldehyde (**Scheme 1.1**) [66].

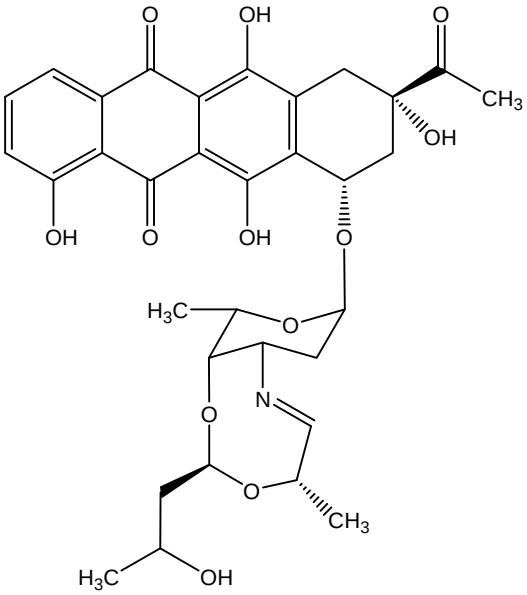
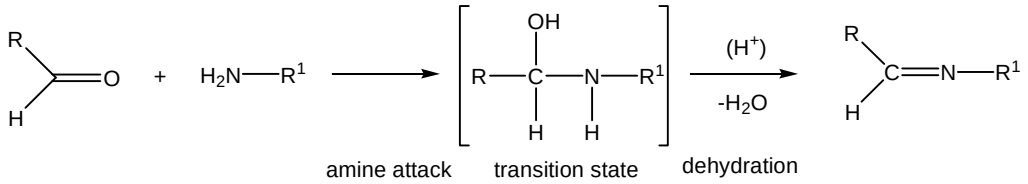


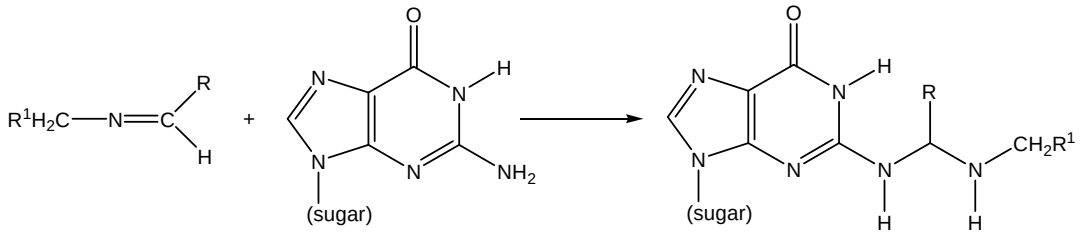
**Figure 1.11** Structure of barminomycin.

Schiff base formation, shown in **Scheme 1.1**, undergoes a transition in rate-determining step from rate-determining dehydration of the carbinolamine addition product at neutral pH, to rate-determining amine attack under acidic conditions [67]. At low pH values, the rate of dehydration becomes rapid, whereas the rate of amine attack is retarded (due to conversion of the amine to its conjugate acid) [67]. Thus, a compromise is required between the rates of dehydration and amine attack.




**Scheme 1.1** Schiff base formation of a primary alkylamine ($R^1$ = remainder of amino-containing drug).

The activated form of the drug can then intercalate to the DNA (or more likely, is already intercalated prior to activation) and covalently binds through the $NH_2$ of guanine in the minor groove (**Scheme 1.2**).




**Scheme 1.2** Imine formation is the reaction of the activated Schiff base with guanine.

### 1.4.1. Cisplatin

There are a number of metal containing drugs with the ability to covalently bind to DNA and stabilise the double stranded form of DNA. Cisplatin is still among the most widely used of the antitumour drugs [14, 68, 69]. It is well established that the anticancer activity of cisplatin derives from the formation of platinum-DNA adducts that inhibit DNA replication and thereby cell division [14, 68-73]. The unexpected and interesting biological activity of platinum complexes was first noted in a study that was designed to test the effects of electric fields on growing cells [69, 74]. One of the electrolysis products from the platinum electrode suppressed bacterial cell division without killing the bacteria [69, 74]. Cisplatin was first tested in 1968 for anticancer activity in a mouse-tumour model system, where it completely inhibited the development of the solid Sarcoma-180 tumour [74]. However, the trans isomer did not suppress bacterial growth except at high concentrations [69, 74]. The first four known active platinum complexes are shown in **Figure 1.12** [74].

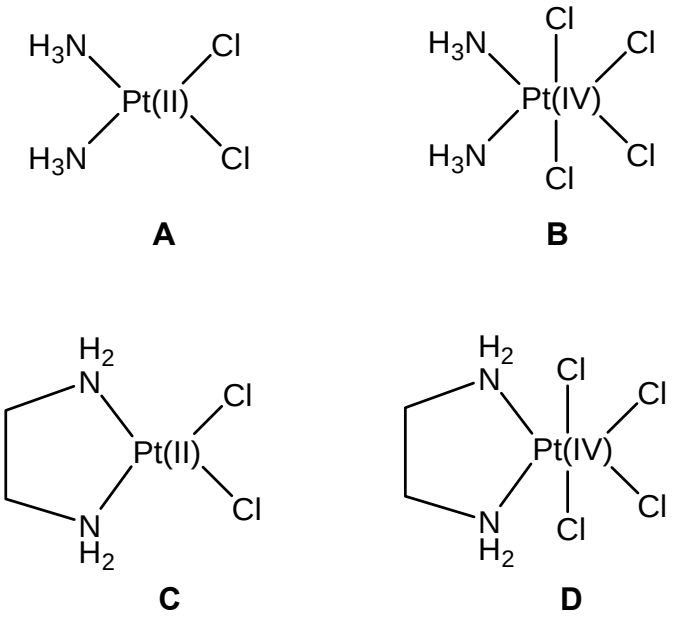



**Figure 1.12** First four known active platinum complexes: **A**. *cis*-dichlorodiammine-platinum(II); **B**. *cis*-tetrachlorodiammineplatinum(IV); **C**. dichloroethylene-diammineplatinum(II); **D**. tetrachloroethylenediammine-platinum(IV) [74].

Early clinical trials showed some notable tumour regressions, and by 1973 it was shown that cisplatin had interesting activity in terminal and previously unresponsive patients [74]. Cisplatin was given FDA approval in 1979 for use as an antitumour agent [69].

Cisplatin is most effective in the treatment of metastatic testicular tumours, but is also included in various combination chemotherapeutic regiments for ovarian, head, neck, bladder, cervical, lung and other neoplasms [68, 69]. A major side effect is kidney toxicity [69, 74, 75], which is greatly reduced by the simple procedure of hydrating the patient. This affords little to no loss of anticancer activity, with a further advance of administering the drug as a slow infusion over 6 to 8 hours [74]. Other problems involved with treatment are myelosuppression, nausea and vomiting, tinnitus and/or hearing loss in the high-frequency range, and peripheral neuropathy [75].

Despite the limitations of acquired resistance and toxic side effects [68], the effectiveness of cisplatin as a chemotherapeutic encouraged research into its molecular interactions with DNA [68, 72]. Animal studies showed that cisplatin would act either additively or synergistically with a number of other anticancer

drugs, affording a substantial improvement in the treatment of animal tumours [74]. The combination of cisplatin and Adriamycin showed a very significant increase of activity both in response rate and duration of remissions for advanced metastatic ovarian cancer [74]. However, a major limitation of cisplatin chemotherapy is acquired resistance, especially for ovarian cancer [68].

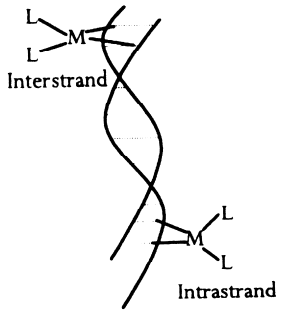


**Figure 1.13** Inter- and intra-strand cross-links.

The low (approx 4 mM) chloride ion concentration within cells allows hydrolysis of cisplatin to the aqua species [76]. In the *in vivo* reactions, major adducts arise from intrastrand cross-links (shown in **Figure 1.13**) between the hydrated cisplatin species and two adjacent guanine residues, whereas major *in vitro* adducts arise between an adenine residue and an adjacent guanine residue [69, 70]. Minor adducts arise from intrastrand cross-links between two guanine residues separated by at least one base, and from interstrand cross-links between two guanine residues [69, 70]. Cisplatin binds to the N7 atom of guanine, the N7 and N1 atoms of adenine, and the N3 atom of cytosine [69]. The cisplatin bridge resides in the minor, rather than the major, groove of the duplex [14]. The two major adducts which have been characterised for cisplatin are d(GpG) and d(ApG) 1,2 intrastrand cross-links, representing about 65% and 25% of the bound platinum, respectively [68, 71-73], The other adducts are d(GXG) 1,3 intrastrand cross-links, X being a base residue, d(GG) 1,2 interstrand cross-links, monofunctional adducts, and protein-DNA crosslinks [68, 71].

### 1.4.2. Adriamycin and Daunomycin Covalent DNA Adducts

While Adriamycin can intercalate with DNA to induce disruption of the DNA replication process, it is a much better drug when activated by formaldehyde where it forms covalent adducts with DNA almost exclusively at GpC sequences [2, 7, 66, 77]. Earlier work displayed that cross-linking was dramatically enhanced by the presence of $Fe^{2+}$ or $Fe^{3+}$ ions, and since $Fe^{2+}$ is rapidly oxidised to $Fe^{3+}$, it was thought that the $Fe^{3+}Ad_3$ species was an intermediate [2, 77, 78]. It has also been shown that adduct formation has the absolute requirement for the exocyclic amine of guanine [2, 66]. High resolution X-ray crystallography (1.5 Å) revealed

that the anthracycline section intercalates between the DNA base pairs, with the amino sugar in the minor groove [79].

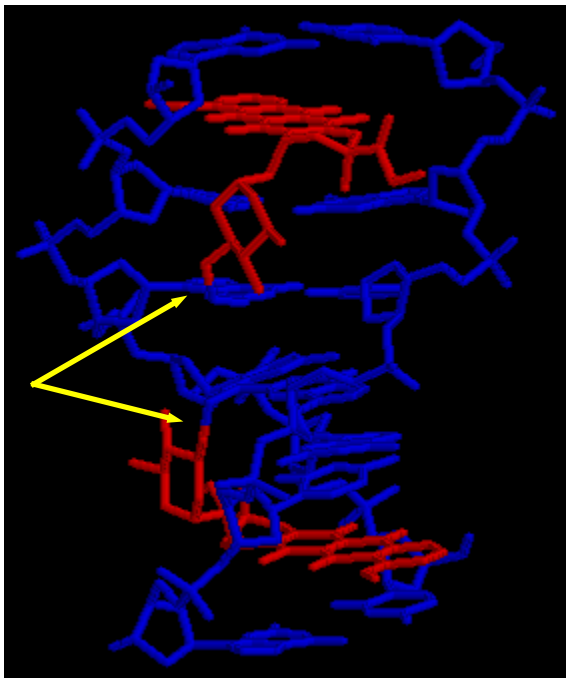


**Figure 1.14** X-ray structure showing two adducts with a covalent methylene bridge between Adriamycin (red) and DNA (blue) [8].

An X-ray structure of the Adriamycin-DNA complex showed a covalent methylene bridge was formed between the N3'of Adriamycin (free amine of the sugar) and the N2 of guanine (indicated by the yellow arrows in **Figure 1.14**) [66, 79]. Guanine has the only free amine in the minor groove whereas that of adenine lies in the major groove. This methylene bridge was attributed at that time to trace amounts of formaldehyde in the buffer solution reacting with the free amine of the drug. However, NMR studies of the Adriamycin-DNA adduct formed on $d(ATGCAT)_2$ in the presence of formaldehyde confirmed that a single covalent adduct was formed with the DNA [66]. An example of two formaldehyde-mediated adducts is shown in **Figure 1.14**. Epirubicin and daunomycin are believed to be cytotoxic to tumour cells by catalysing the production of formaldehyde through redox cycling and using the formaldehyde for covalent attachment to DNA at guanine [8].

As in the nogalamycin/DNA structure, the Adriamycin intercalates in the B form DNA, accompanied by unwinding of the DNA. The adjacent bases are

opened up, but there is minimal distortion in the 3D base pair structure. This lack of base pair distortion is seen directly above and below the red intercalated anthracycline in **Figure 1.14**.

### 1.4.3. Mitoxantrone and Barminomycin

Mitoxantrone can also be activated by formaldehyde *in vitro*, producing DNA adducts which function as interstrand crosslinks. Cells of myeloid origin have increased levels of formaldehyde [80], possibly hinting as to why mitoxantrone is effective against myeloid tumours. The attraction of neutrophils to sites of inflammation of advanced solid tumours causes the release of hydrogen peroxide accompanying a respiratory burst [13, 81]. Hydrogen peroxide produces oxygen and hydroxyl radicals which oxidise Tris (*in vitro*) and the biologically available polyamines such as spermine, to generate formaldehyde [4, 13]. This affords an increased level of formaldehyde at the site of these solid tumours, and may also result in the activation of mitoxantrone. Therefore, tumours which do not contain increased levels of formaldehyde may be resistant to mitoxantrone, because the drug cannot be activated sufficiently in these cells [13, 59].

The X-ray structure of another ligand which covalently binds to DNA is MAR07, a synthetic daunomycin derivative that contains an additional sugar, attached to the O4' of daunosamine (**Figure 1.15**). While crystallising the derivative, Gao et al. (1991) observed the drug readily formed a cross-link to DNA [33]. The adduct formed readily with high yield during the crystallisation steps, likely due to the trace amounts of formaldehyde in the solvent [33]. The drug was bound via a methylene bridge (HCHO) between the N3' of daunosamine and the N2 of 2-aminoadenine of the DNA [33].

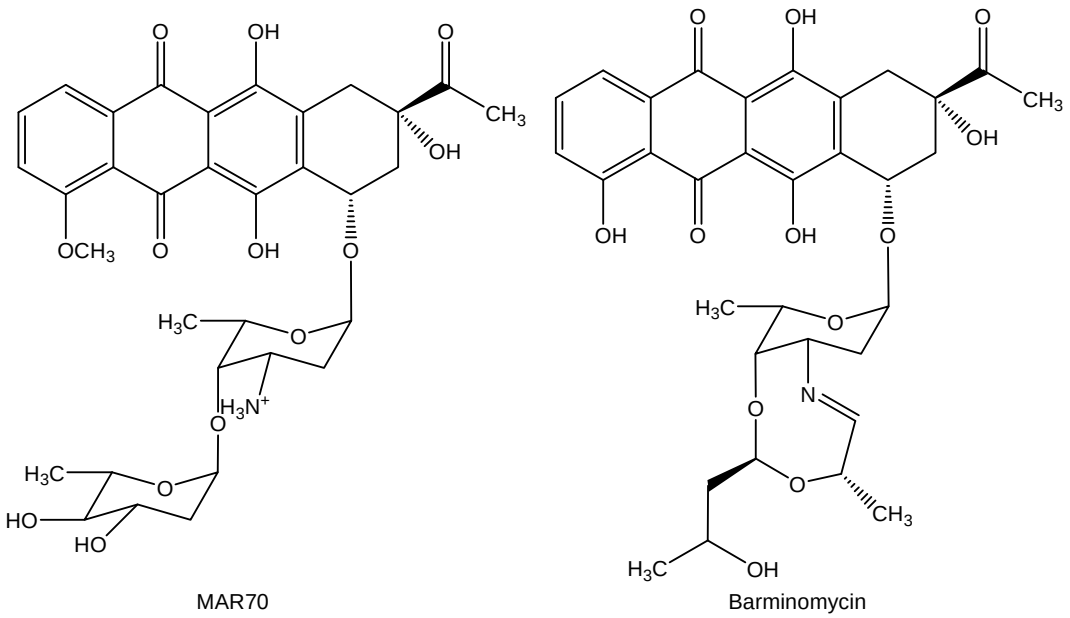



**Figure 1.15** Structure of MAR70 (and barminomycin).

The macromolecular antibiotic SN-07 was extracted from *Actinomadura roseoviolacea* var. *miuraensis* nov. var. in 1985 [82, 83]. An organism from a soil sample in Japan was grown in oatmeal starch medium, and strain 07 was isolated from the supernatant. The pinkish red powder was found to have activity against lymphocytic leukaemia P388, and was mutagenic in both prokaryotic and eukaryotic cells without the need for metabolic activation [82, 83]. The active chromophore (SN-07 chromophore) was isolated from SN-07, and was shown to be the anthracycline antibiotic barminomycin (**Figure 1.15**) [84]. Barminomycin intercalates in the minor groove of DNA, and binds to the amine group of an adjacent guanine, forming a covalent bond from the carbon of the double bond on the sugar ring to the DNA [85].

### 1.4.4. Anti-Cancer Drugs Containing Free Amine Groups

It was observed that most of the naturally occurring drugs previously mentioned contain a free amino group, whereas the majority of the synthetic drugs do not. The results from these examples led to the suggestion that the incorporation of a free amine group within the ligand may increase its ability to form a covalent bond with DNA. The ligand may need the ability to form a Schiff base with endogenous levels of formaldehyde, and then react with the $NH_2$ of

guanine. As there have already been derivatives of acridine which exhibit antitumour activity [10, 34], the terminal N-dimethyl side chain of DACA was modified for the work in this report to the free amine. This was designated *N*-(2-aminoethyl)acridine-4-carboxamide (C2-DACA).

## 1.5. DNA Structure

Some readers of this thesis will have a strong background in chemistry and molecular modelling, with perhaps only a limited knowledge of molecular biology. For this reason a brief introduction has been included below to the key concepts involved in DNA structure, since much of the drug-DNA interaction presented in this thesis requires an understanding of steric aspects of B-DNA.

Deoxyribonucleic acid and ribonucleic acid are the nucleic acids that carry all of a cell's genetic information. The individual building blocks consist of nucleotides (heterocyclic purine or pyrimidine) held in specific sequences along a long phosphate-sugar backbone. In both DNA and RNA there are four different heterocyclic amine bases. In DNA, the two purines are guanine and adenine, and the two pyrimidines are cytosine and thymine. The only difference between DNA and RNA bases is that the latter has uracil in place of thymine. Two strands of complementary DNA are held together by the hydrogen bonds between specific nucleotides, whereas RNA remains as a single strand. Guanine pairs with cytosine, and adenine with thymine (**Figure 1.16**).

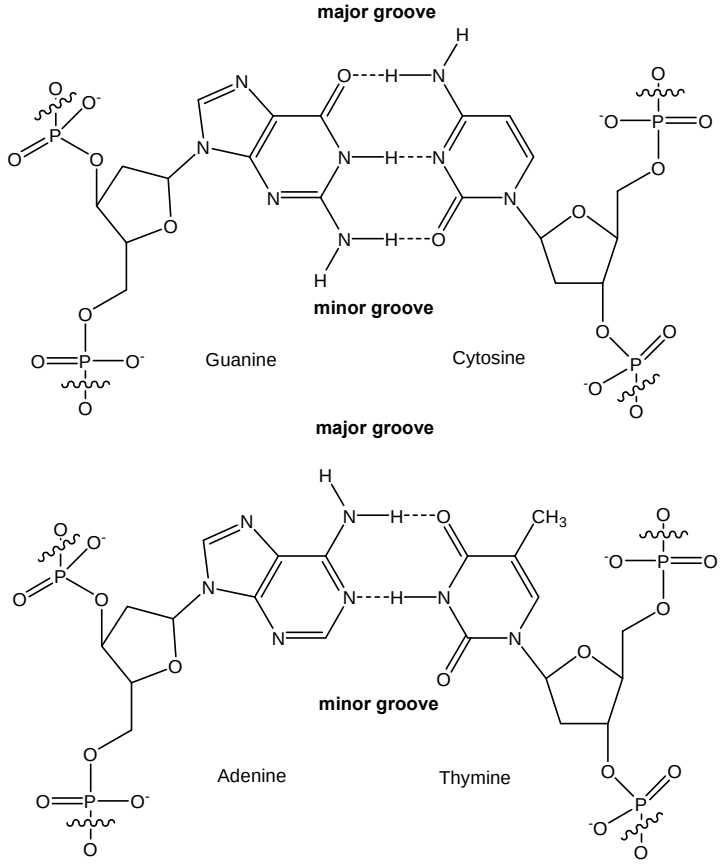



**Figure 1.16** Base pairing in DNA. The two strands are complimentary

The Watson and Crick model for DNA was published in 1953, and was determined by use of X-ray crystallography [86]. Their model was later found to be the B form of DNA. This structure (**Figure 1.17**) has two helical chains, represented as ribbons, coiled about a central axis, with the phosphate diester groups joining β-D-deoxyribofuranose residues with 3'-5' linkages [86]. The horizontal rods represent the pairs of bases holding the chains together through hydrogen bonds (**Figure 1.16**). The helix makes a turn every 3.4 nm, with a distance between two neighbouring base pairs of 0.34 nm, giving about 10 pairs per turn [87]. The intertwined strands make two grooves of different widths, referred to as the major groove and the minor groove, which may facilitate binding with specific proteins or drugs since only specific functional groups

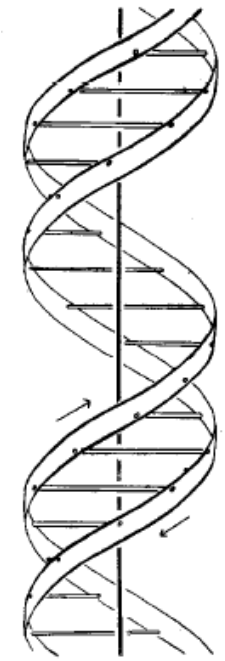


**Figure 1.17** Watson and Crick model of DNA.

(e.g. amino, carboxyl etc) protrude into specific grooves (see **Figure 1.16**). DNA is in the B form in aqueous conditions such as experiments performed *in vitro* [88].

Fibre diffraction analyses showed that the secondary structure of DNA is polymorphic (A, B, C, D, and Z-forms) [89, 90], with conformational species such as right-handed, left-handed, and achiral allomorphs. Solutions with higher salt concentrations or with alcohol added may change the DNA structure to an A form or, in DNA with alternating purine-pyrimidine sequences, to the Z form (**Figure 1.18**) [88]. A-DNA is still right-handed, but every 2.3 nm makes a turn, giving 11 bases per turn [91]. This induces the major groove to become very deep and the minor groove quite shallow.

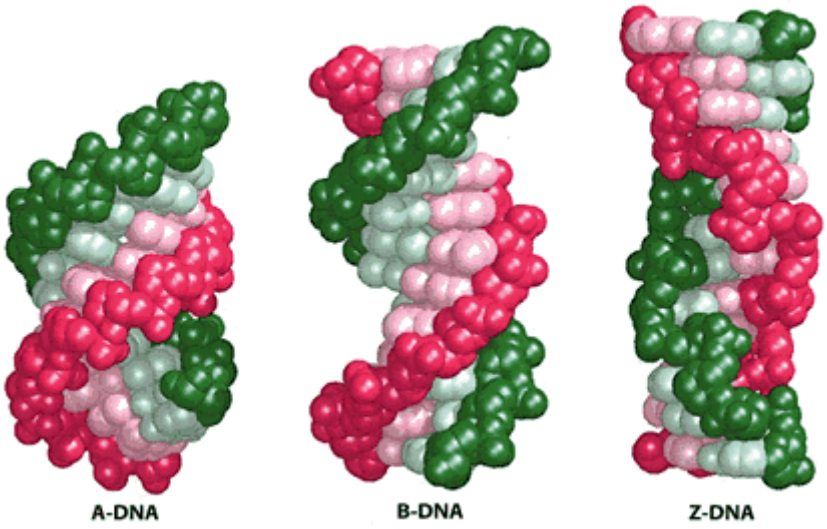



**Figure 1.18** Space-filling models of A-DNA, B-DNA and Z-DNA, each 12 base pairs in length [92].

The dodecamer 5'-d(CGCGAATTCGCG)$_2$ revealed that the local helix parameters, twist, tilt and roll, are much more strongly influenced by base sequence than by crystal packing or any other external forces [93]. Free oxygen atoms of adjacent phosphate groups in B-DNA are found at least 6.6 Å apart and individually hydrated, whereas they are as close as 5.3 Å in A-DNA and 4.4 Å in Z-DNA, and bridged by water molecules [88]. A suggested reasoning behind the transition of B to A and B to Z is that A- and Z-DNA have a more economical hydration compared with B-DNA [88]. This implies that A- and Z-DNA are stable under a water deficit, whereas B-DNA is stable under water excess, making B-

DNA the most likely form present in the nucleus of the cell, and also in aqueous conditions.

The structure that the DNA adopts relies heavily upon the base pair sequence. Water hydration is base-specific and is an important element in stabilising regions of DNA with particular base sequences, and in influencing conformational transitions of DNA with non-random sequence, such as the B to A transitions of DNA fibres with high A+T content, and the B to Z transition of poly [d(G-C)] [88].

## 1.6. Aims and Overall Approach for the Project

The aim of the project was to emulate the way in which Adriamycin becomes activated by formaldehyde and consequently binds to DNA, and to employ much simpler compounds to enable the formation of similar drug-DNA adducts. The binding ability of the compounds was to be determined using biochemical techniques and spectroscopic methods. If there were evidence of covalent binding, the binding site would be determined using a short oligonucleotide (6-8 base pairs) by NMR and computer modelling. The ultimate goal was to determine the complete structure using these methods. In parallel, quantum mechanic calculations were to be undertaken to quantify the energetics of Schiff base and drug-DNA adduct formation.

The compound shown in **Figure 1.19**, *N*-(2-aminoethyl)acridine-4-carboxamide.oxalate (C2-DACA) was chosen for initial testing for several reasons. The parent compound *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide (DACA), which contains a terminal $NMe_2$ group, exhibits antitumour activity, is relatively easy to synthesise, requires reasonably inexpensive starting materials, has no need for enantioselective reactions (unlike daunomycin and Adriamycin, it contains no chiral centres), and has the potential to mimic Adriamycin (in terms of formaldehyde activation and binding to DNA). It might be perceived that the side chain was too short for covalent binding (as was shown in this thesis) but there is reasonable flexibility in the intercalation position of the acridines which could possibly allow for the side chain to bind covalently.

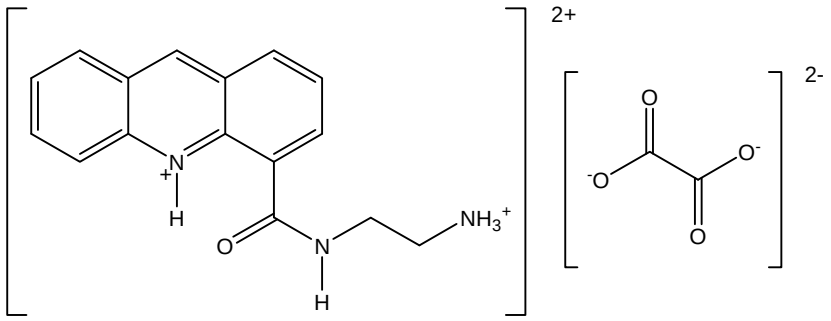


**Figure 1.19** Structure of C2-DACA.

Biological studies proved that C2-DACA does not mimic Adriamycin. This led to the computer modelling (using Insight II software on an SGI Iris workstation) of C2-DACA intercalated into a short DNA sequence (6 base pairs). The modelling revealed that the side arm of C2-DACA was not long enough to undergo the desired reaction (intercalation and subsequent binding to guanine). Further modelling of longer side arms showed that an extra 2-3 carbons in the chain would greatly reduce steric hindrance of the DNA at the intercalation site.

Two main approaches were decided upon:

- React C2-DACA with acrolein derivatives (Michael addition) to create a longer chain, then test for conjugation to DNA
- Create a new drug with four carbons in the side chain (instead of only two), which would facilitate the formation of an aminal linkage with DNA.

The first approach was dismissed due to the probability of formation of unwanted acrolein-DNA adducts [94]. Thus the side chain of C2-DACA was increased by using diaminobutane instead of diaminoethane in the synthesis of C4-DACA. The compound shown in **Figure 1.20**, *N*-(4-aminobutyl)acridine-4-carboxamide.oxalate (C4-DACA), was synthesised in collaboration with Dr L Deady (La Trobe University).

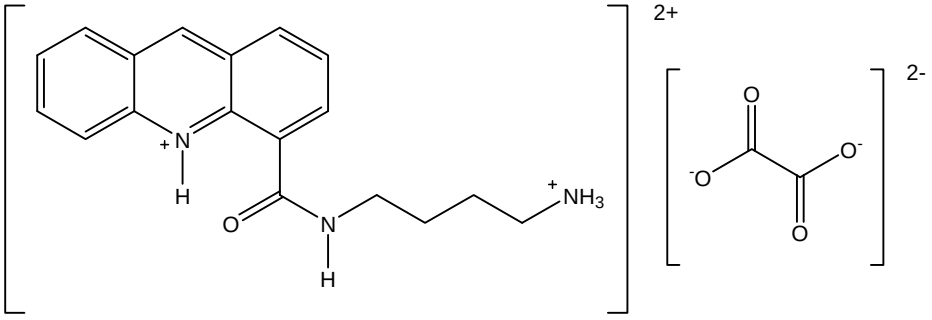


**Figure 1.20** Structure of C4-DACA.

C4-DACA was assessed for its ability to form drug-DNA adducts in the presence of formaldehyde. Tests included absorbance based binding studies, stability of adducts with DNA, cross-linking assays, transcription assays, together with other biochemical studies.

It was expected that the free amine on the extended side arm would enable binding to DNA, and this means that many other anti-cancer drugs could be modified similarly to yield more active DNA binding agents. This would mean that chemotherapy could be a less traumatic experience, and that the growth of tumours may be slowed, stopped, or even reversed.

The ultimate goals of these studies were to determine the conditions at which this reaction would take place by using NMR studies of C4-DACA bound covalently to DNA. This would allow a 3D model of the bound and intercalated drug to be created, using specialist software such as Accelrys Discovery Studio and Amber software on a Linux based operating system. The outcome of this work would also help us model new and better anti-cancer drugs in the future.

# CHAPTER 2. DRUG SYNTHESES AND CHARACTERISATION

## 2.1. Introduction

This chapter describes the synthesis of the compounds that were used in the DNA binding studies. That includes C4-DACA, C2-DACA and the C2-DACA dimer. The first step in all of these syntheses was to start with the reduction of 9-oxoacridan-4-carboxylic acid to oxoacridan-4-carboxylic acid. This could then be converted to the three target compounds. As discussed previously, these compounds have no chiral centres and thus the synthesis should be quite straightforward. The first intermediate, acridine-4-carboxylic acid, is a known lachrymator and sternutator [9].

Two syntheses of C4-DACA were performed. A small scale synthesis of C4-DACA was undertaken by myself, however it quickly became apparent that I was allergic by lachrymatory effect and skin allergy to the intermediate. Due to this, all DACA derivatives used within this thesis were kindly synthesised by Dr Les Deady, of the Department of Chemistry, La Trobe University.

In order to fully characterise each drug unequivocally, multiple 2D nuclear magnetic resonance spectroscopy was employed. $^{1}H$ and $^{13}C$ NMR gives information on the chemical environment of each equivalent proton or carbon in a substance by a certain predictable chemical shift. This is done by "pulsing" a sample within a magnetic field with radio frequency radiation, and recording and amplifying the emitted radiation decay. Different information can be obtained by using certain pulse sequences.

The Distortion Enhancement by Polarisation Transfer or DEPT spectrum determines the presence of primary, secondary or tertiary carbon atoms by varying the flip angle. A DEPT135 will have negative CH and $CH_3$ peaks and positive $CH_2$ peaks.

Heteronuclear NMR is helpful due to the very low abundance of naturally occurring $^{13}C$ when only small concentrations of sample are available. The Heteronuclear Single Quantum Correlation or HSQC spectrum correlates information between a carbon atom and the protons that are directly attached as magnetisation is transferred from protons to the hetero nucleus.

Homonuclear spectrums (usually $^{1}H$) can give information about the connectivity of a compound. Spectra are symmetrical about a diagonal plane, with cross peaks arising from magnetisation exchanged between two nuclei during the mixing time (time between pulses). COrrelation SpectroscopY or COSY spectra show coupling (magnetisation transfer) between protons that are separated by three bonds (direct through-bond neighbours). Total COrrelation SpectroscopY or TOCSY spectra are similar to COSY spectra, but depending on the mixing time, coupling is seen through more bonds. A short mixing time results in a COSY-like spectrum whereas a long mixing time will increase the range of polarisation through multiple bonds. A particularly useful tool is the Nuclear Overhauser Effect SpectroscopY or NOESY spectrum. This method allows nuclei that are spatially close together (within 5 Å) to undergo cross relaxation. NOESY spectra are particularly useful in assigning the backbone of DNA via the close proximity of the sugars and the bases. It is possible to find the aromatic protons of the 5' DNA base then jump from aromatic to H1' to aromatic all the way down the backbone to the 3' DNA base.

## 2.2. Materials and Synthesis

The starting material, 9-oxoacridan-4-carboxylic acid, 1,1'-carbonyldiimidazole, 1,2-diaminoethane and 1,4-diaminobutane were purchased from Sigma Aldrich. The remaining chemicals and reagents were of analytical grade. NMR spectra were recorded on a Bruker Avance-300 spectrometer ($^{1}H$ at 300.1317 MHz and $^{13}C$ at 75.47 MHz), a Bruker Avance-400 spectrometer ($^{1}H$ at 400.100 MHz and $^{13}C$ at 100.5877 MHz) and a Bruker Avance-500 spectrometer ($^{1}H$ at 500.1900 MHz and $^{13}C$ at 125.7729 MHz). All the chemical shifts in the NMR were recorded as δ values in parts per million (ppm) and coupling constants (*J*) were recorded in Hz. Electrospray ionization (ESI) mass spectrometry was carried out using a Bruker Daltronics (Germany) Esquire$^{6000}$ ion trap mass spectrometer. The samples were analysed with a flow rate of 4 μL/min and a mass range of 50 – 3000 m/z. A scan rate of 5500 m/z/second was used with the temperature set at 140°C. The numbering system used in these compounds is shown in **Figure 2.1**. This numbering scheme is also used in the NMR assignments.

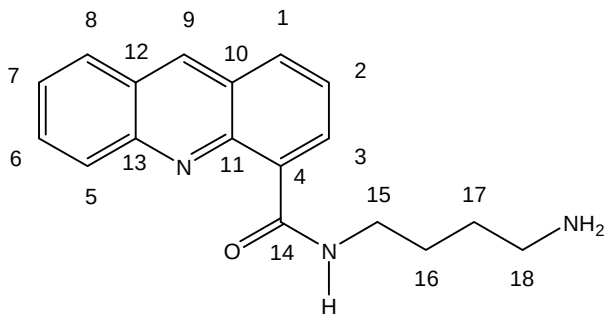
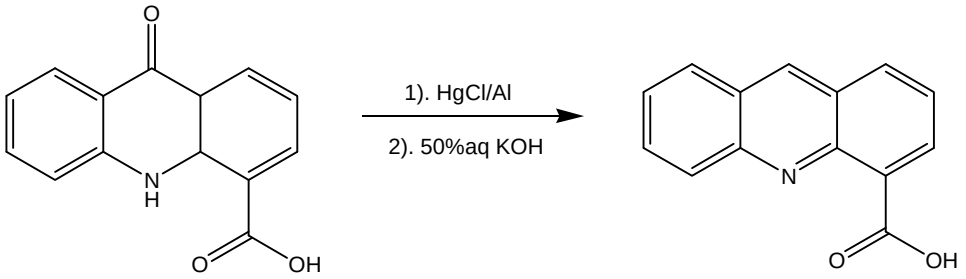



**Figure 2.1** Numbering system employed for C4-DACA assignment.

## 2.3. Reduction of 9-oxoacridan-4-carboxylic acid to acridine-4-carboxylic acid




**Scheme 2.1** Reduction of 9-oxoacridan-4-carboxylic acid to acridine-4-carboxylic acid.

Method A from Atwell et al (1987) [34] was scaled down by a factor of 6. The reaction scheme is shown in **Scheme 2.1**. Portions (0.1-0.2 g each) of aluminium foil (total 0.9 g, 0.0333 mol) were amalgamated with mercury by immersing them for 1 min in a solution of mercuric chloride (3 g, 0.0111 mol) in EtOH (30 mL).

The resulting amalgam was washed with EtOH and added portionwise to a refluxing solution of 9-oxoacridan-4-carboxylic acid (1 g, 0.0042 mol) and KOH (0.27 g, 0.0048 mol) in 50% aqueous EtOH (30 mL). When the reaction was complete, as indicated by TLC, 1M KOH (8 mL) was added, the hot mixture was filtered, and the precipitate washed with hot 50% aqueous EtOH (25 mL). The filtrate was acidified with 12M HCl and treated with $FeCl_3$ (1.8 g). The solution was refluxed until homogenous, and clarified by filtration. The product was precipitated with KOAc, filtered and washed with water.

The crude product was dissolved in 1M KOH, filtered and diluted with hot EtOH. HOAc was added dropwise and the resulting precipitate was filtered. The precipitate was recrystallised in aqueous acetone to give acridine-4-carboxylic acid as a pale yellow, granular solid in 32% yield (0.298 g, 0.0013 mol).

The aromatic section of a 300 MHz NMR spectrum of the acid is shown in **Figure 2.2** and shows the acridine protons at low field. The acridine protons show the expected splitting pattern for this compound with the H9 singlet appearing at 9.5 ppm, and the other protons appearing as four doublets and three triplets (actually doublets of doublets with similar couplings).

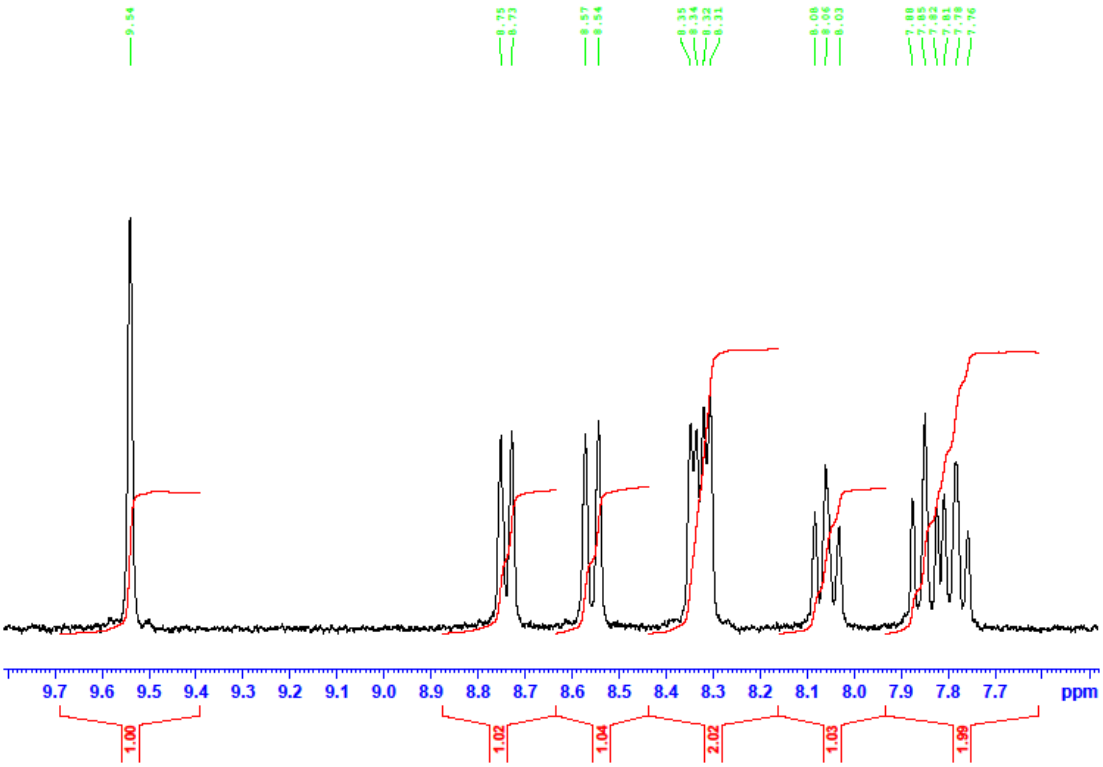
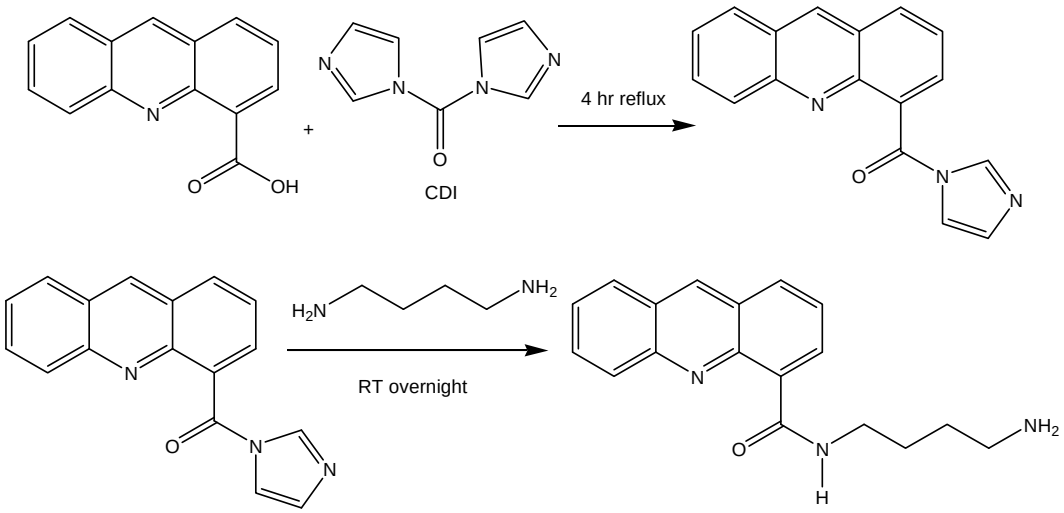



**Figure 2.2** Expanded aromatic section of a 300 MHz $^{1}$H spectrum of acridine-4-carboxylic acid in d6-DMSO showing the splitting pattern of the acridine aromatic protons.

## 2.4. Conversion of acridine-4-carboxylic acid to *N*-(4-aminobutyl)acridine-4-carboxamide (C4-DACA)




**Scheme 2.2** Conversion of acridine-4-carboxylic acid to *N*-(4-aminobutyl) acridine-4-carboxamide (C4-DACA).

A solution of acridine-4-carboxylic acid from **2.3** (0.2 g, 0.0009 mol) and 1,1'-carbonyldiimidazole (0.6 g, 0.0037 mol) in dioxane (5 mL) was refluxed with moisture exclusion for 4 h.

The solvent was removed and the residue was dissolved in 5 mL dichloromethane. This was added dropwise to a stirred solution of 1,4-diaminobutane (0.7 g, 0.0079 mol) in dichloromethane (10 mL) and stirred overnight at room temperature.

The solution was washed with 10% $Na_2CO_3$ (5 mL), and water (2 x 5 mL), dried with $MgSO_4$ and solvent was removed *in vacuo* ($^1$H spectrum shown in **Figure 2.3**). The oxalate was prepared by the slow addition of a solution of oxalic acid in EtOH. The resulting precipitate was recrystallised in i-PrOH to afford *N*-(4-aminobutyl)acridine-4-carboxamide.oxalate salt (C4-DACA) as a pale yellow solid (**Figure 2.4**).

## 2.5. Results

The final product, C4-DACA, was identified using $^{1}H$ and $^{13}C$ NMR. The $^{1}H$ NMR data of the solid in 98% d6-DMSO at about 1 mg concentration before conversion to the oxalate salt is shown in **Figure 2.3**. The spectrum shows the aromatic peaks as four "doublets" and three "triplets" as expected for the acridine ring. The protons on the side chain appear at 3.52, 2.62, 1.72 and 1.56 ppm with appropriate splittings. The peak at 3.52 ppm is adjacent to the large residual water peak and shows coupling to the adjacent NH at 11.43 ppm (not shown). The side chain peaks are also broadened as a consequence of dynamic effects. All spectra were calibrated to DMSO (2.49 or 39.5 ppm) or H9 of C4-DACA (9.326 ppm). This solid was recrystallised from i-PrOH and was fully assigned by various NMR methods, and electrospray ionisation mass spectrometry.

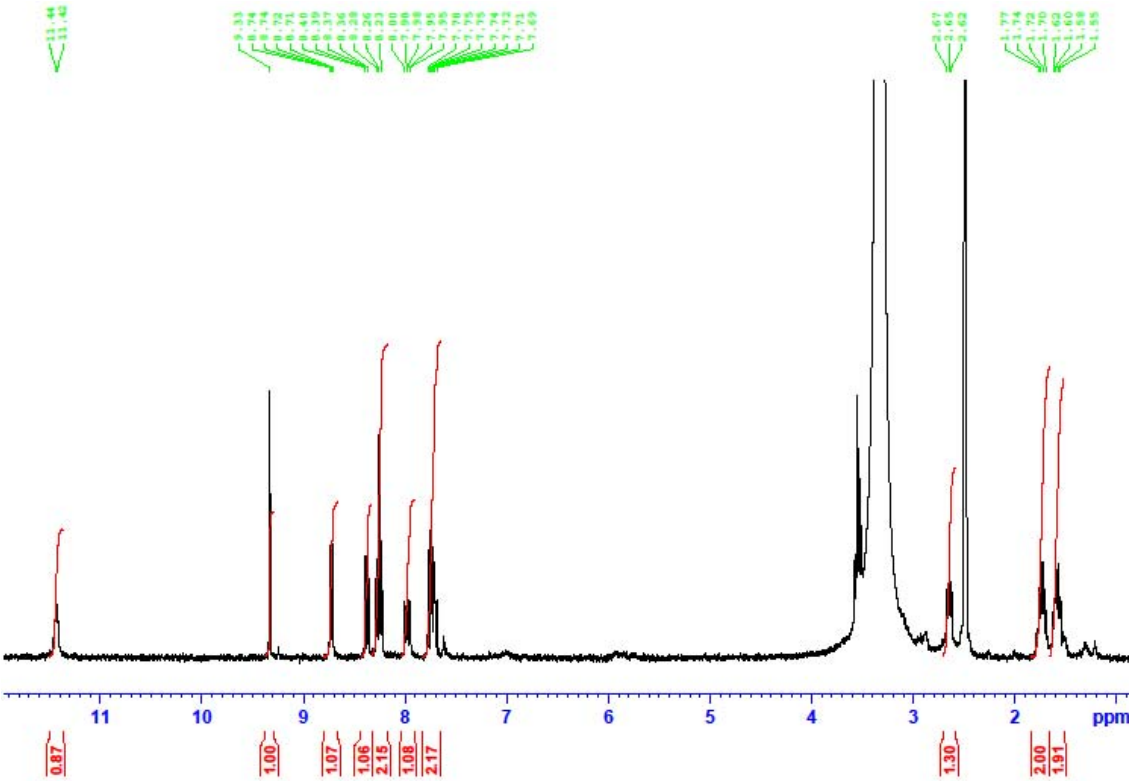



**Figure 2.3** 300 MHz $^{1}H$ NMR in d6-DMSO of the C4-DACA solid before recrystallisation.

### 2.5.1 Characterisation of C4-DACA

The full assignment of C4-DACA was required not only to identify the product unequivocally, but also for ease of assignment for the drug-DNA NMR studies. All spectra were calibrated to DMSO (2.49 or 39.5 ppm) or H9 of C4-DACA (9.35 ppm).

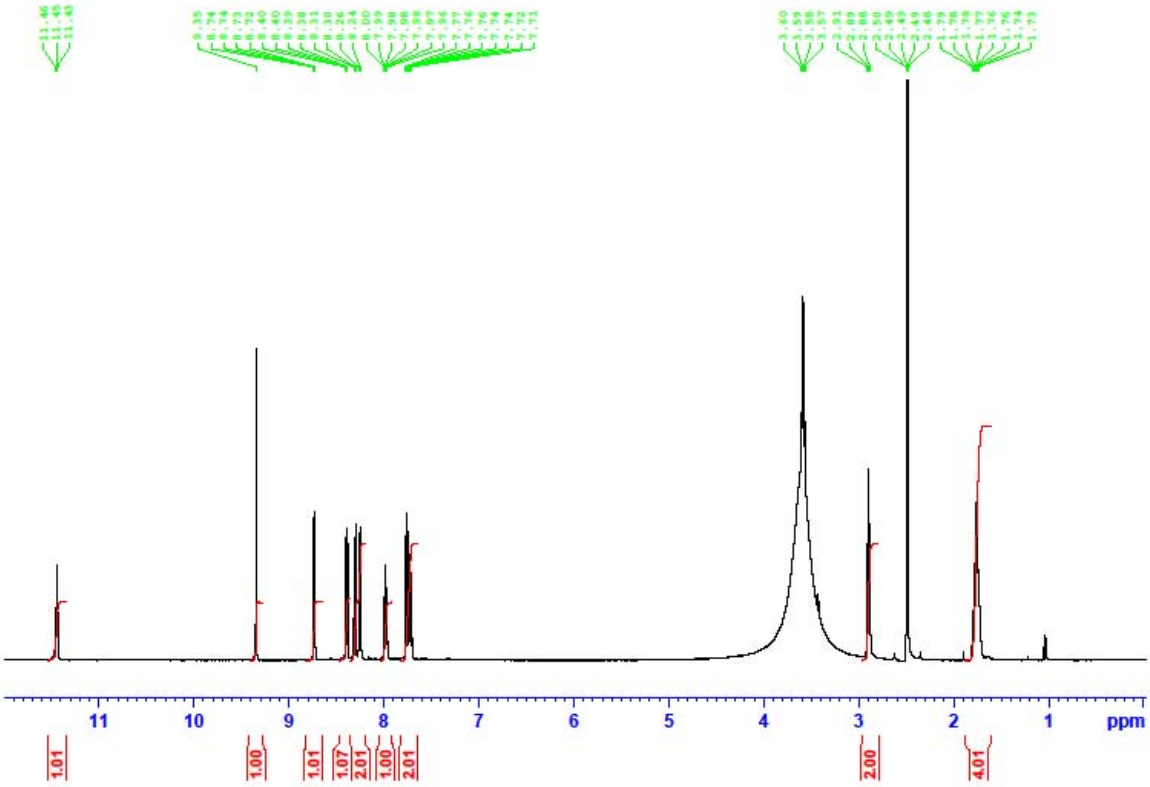



**Figure 2.4** 500 MHz $^{1}$H NMR spectrum in d6-DMSO of C4-DACA at 300 K.

The 500 MHz $^{1}$H spectrum of C4-DACA is given in **Figure 2.4** and shows the expected splitting patterns. The protons on the side chain are at ~1.77 (overlapping multiplet), 2.89 and 3.59 ppm with appropriate splitting. The "quartet" under the small water peak at 3.59 ppm showed coupling to the adjacent NH at 11.45 ppm (shown in the magnitude COSY, **Figure 2.11**), indicating that it was the H15 proton. The aromatic section of this spectrum is shown in **Figure 2.5** with four "doublets" and three "triplets" for the acridine protons, as expected.

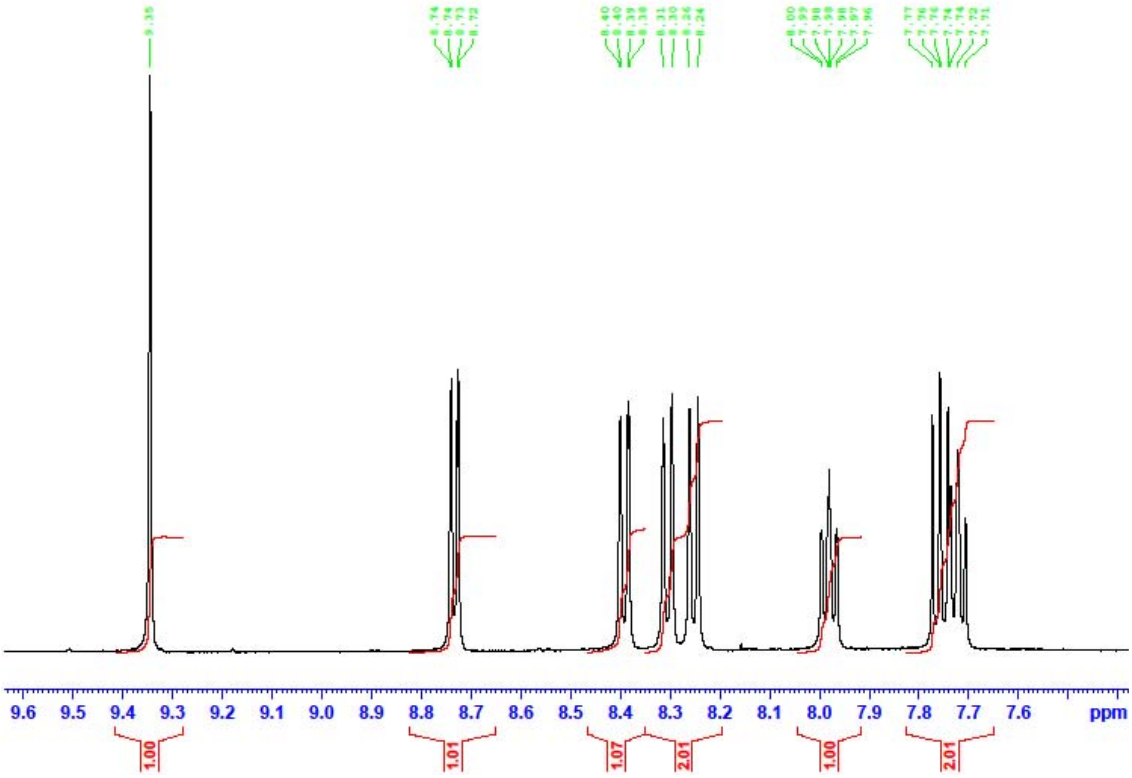



**Figure 2.5** Aromatic section of a 500 MHz $^{1}$H NMR spectrum in d6-DMSO of C4-DACA at 300 K.

The $^{13}$C spectrum in **Figure 2.6** shows that there are 19 carbon peaks as expected (C4-DACA contains 18 carbons, and oxalate contains 2 equivalent carbons) with appropriate integration. Two of the side chain carbons appear very close to the DMSO multiplet as indicated in **Figure 2.7** by the red arrows.

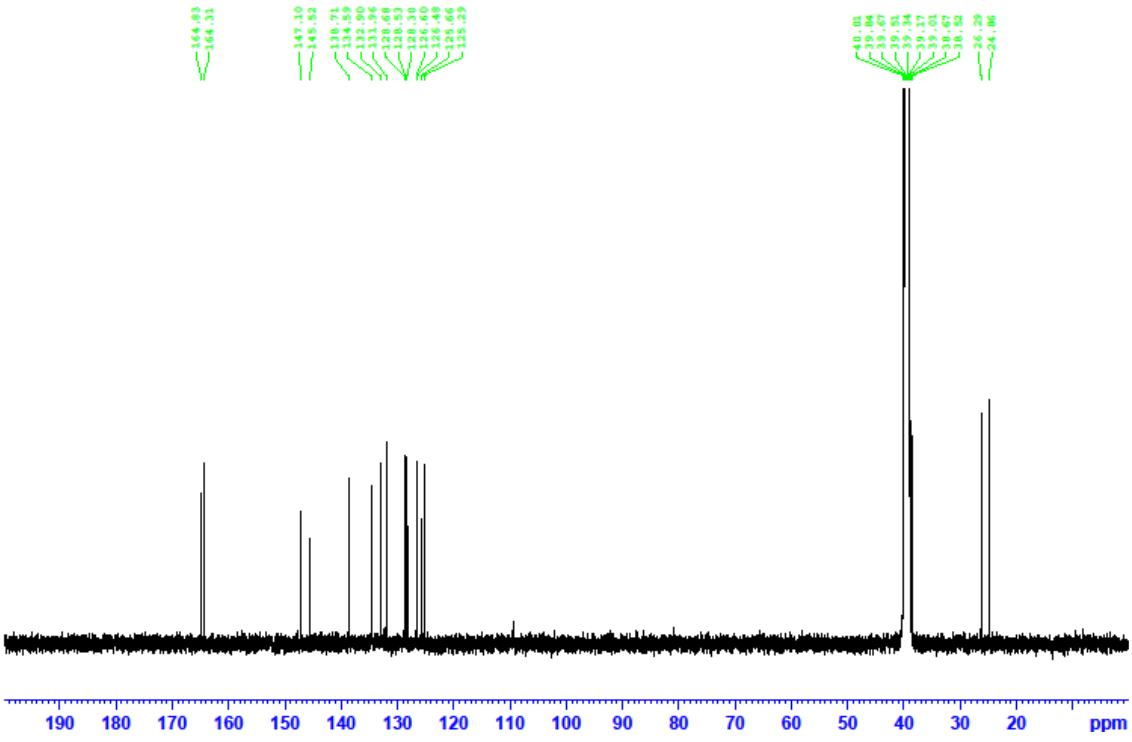



**Figure 2.6** 500 MHz $^{13}$C NMR spectrum in d6-DMSO of C4-DACA at 300 K.

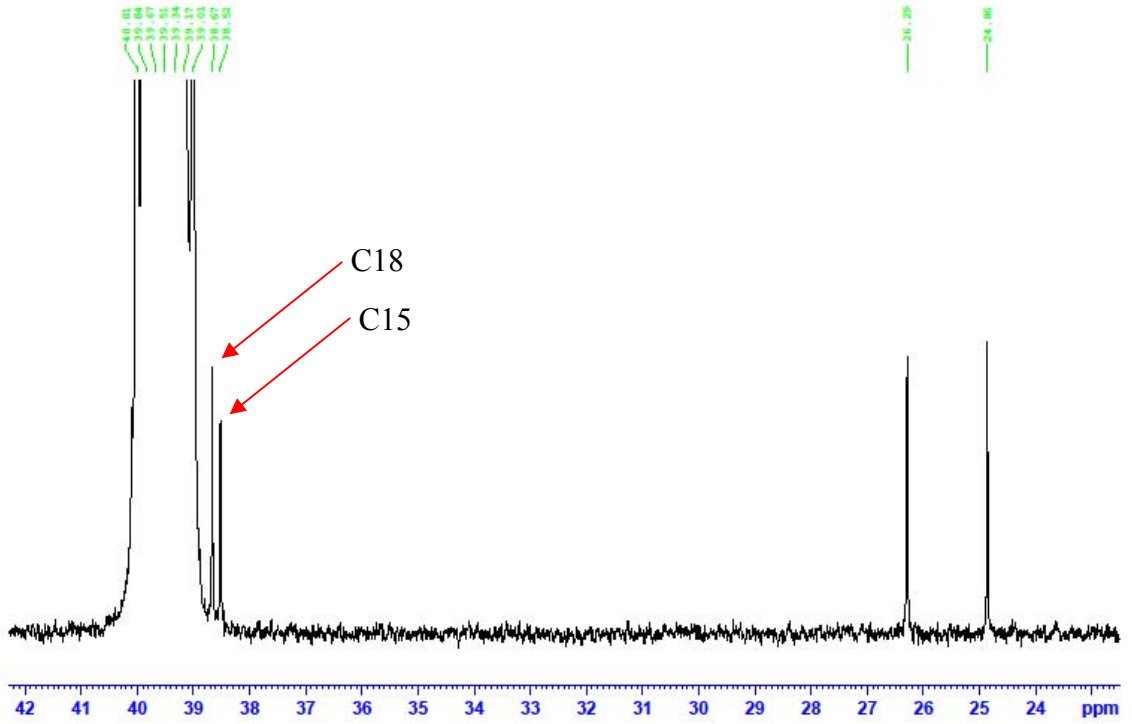



**Figure 2.7** 500 MHz $^{13}$C NMR spectrum in d6-DMSO of the side chain carbons of C4-DACA at 300 K.

The 500 MHz DEPT135 spectrum in **Figure 2.9** indicates that all of the carbon peaks from 120 to 140 ppm are CH carbons, and those from 20 to 38.5 ppm are $CH_2$ carbons, as expected.

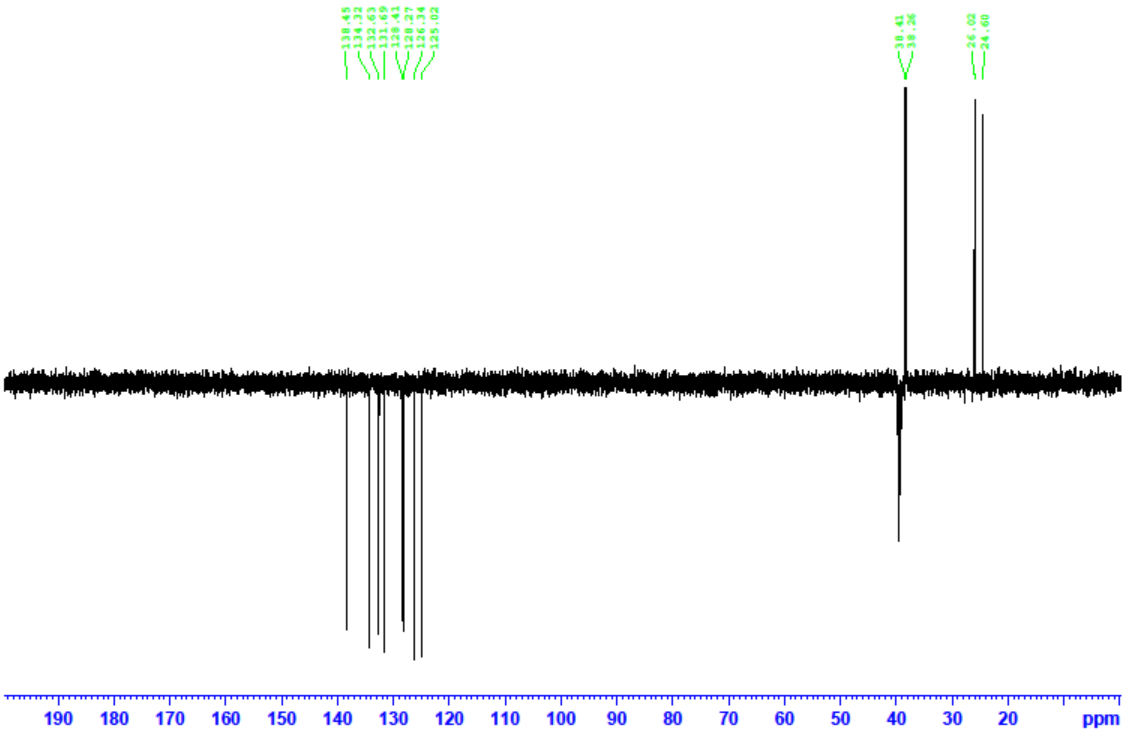



**Figure 2.8** The 500 MHz DEPT135 NMR spectrum in d6-DMSO of C4-DACA at 300 K.

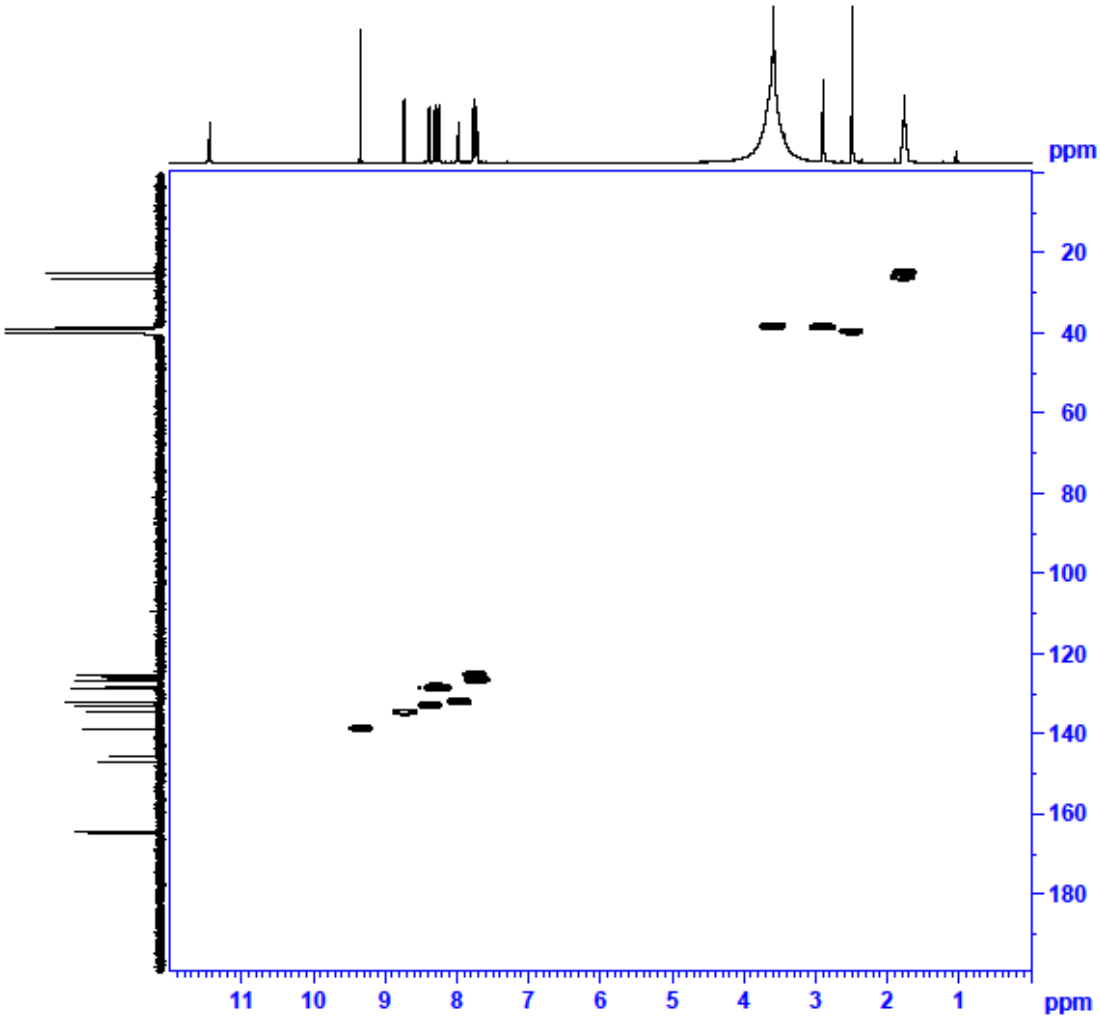



**Figure 2.9** 500 MHz HSQC NMR spectrum in d6-DMSO of C4-DACA at 300 K.

The 500 MHz HSQC spectrum in **Figure 2.9** allows the assignment of the C15 to C18 carbons, and the separation of the overlapping acridine protons. The overlapping multiplet at 7.7 ppm in the proton spectrum correlates to two separate carbons at 125.3 and 126.6 ppm. These carbon shifts are also consistent with model $^{13}C$ predictions using the ChemDraw NMR software.

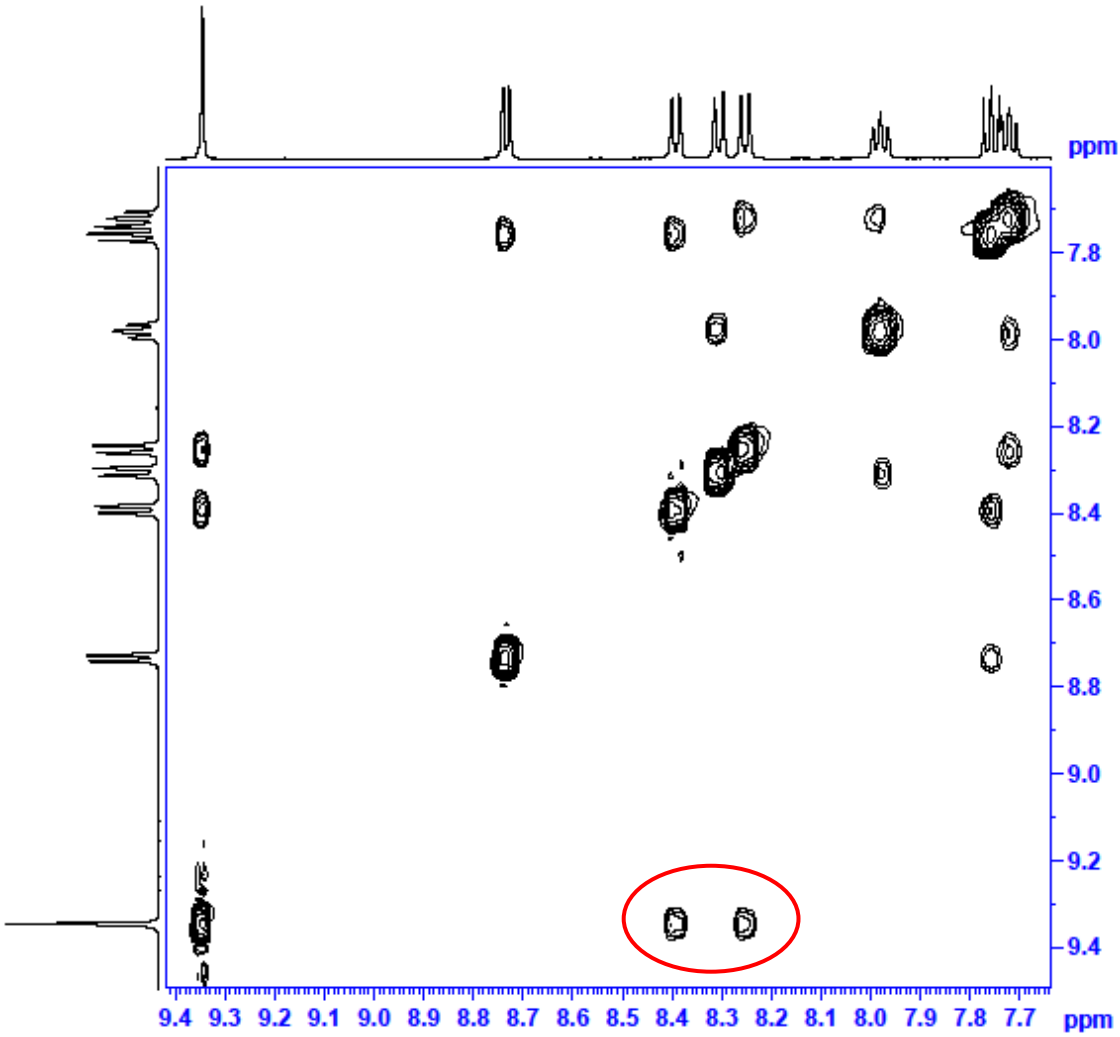



**Figure 2.10** Aromatic section of a 500 MHz 900 ms mixing time NOESY NMR spectrum in d6-DMSO of C4-DACA at 300 K. This shows the NOE peaks from H9 at 9.35 ppm to H1 at 8.40 ppm and H8 at 8.25 ppm, indicated by the red oval.

The aromatic section of a 500 MHz NOESY spectrum with 900 ms mixing time is presented in **Figure 2.10** and shows the NOE coupling from the H9 "singlet" at 9.35 ppm to the H1 "doublet" at 8.40 ppm and the H8 "doublet" at 8.25 ppm, indicated by the red oval.

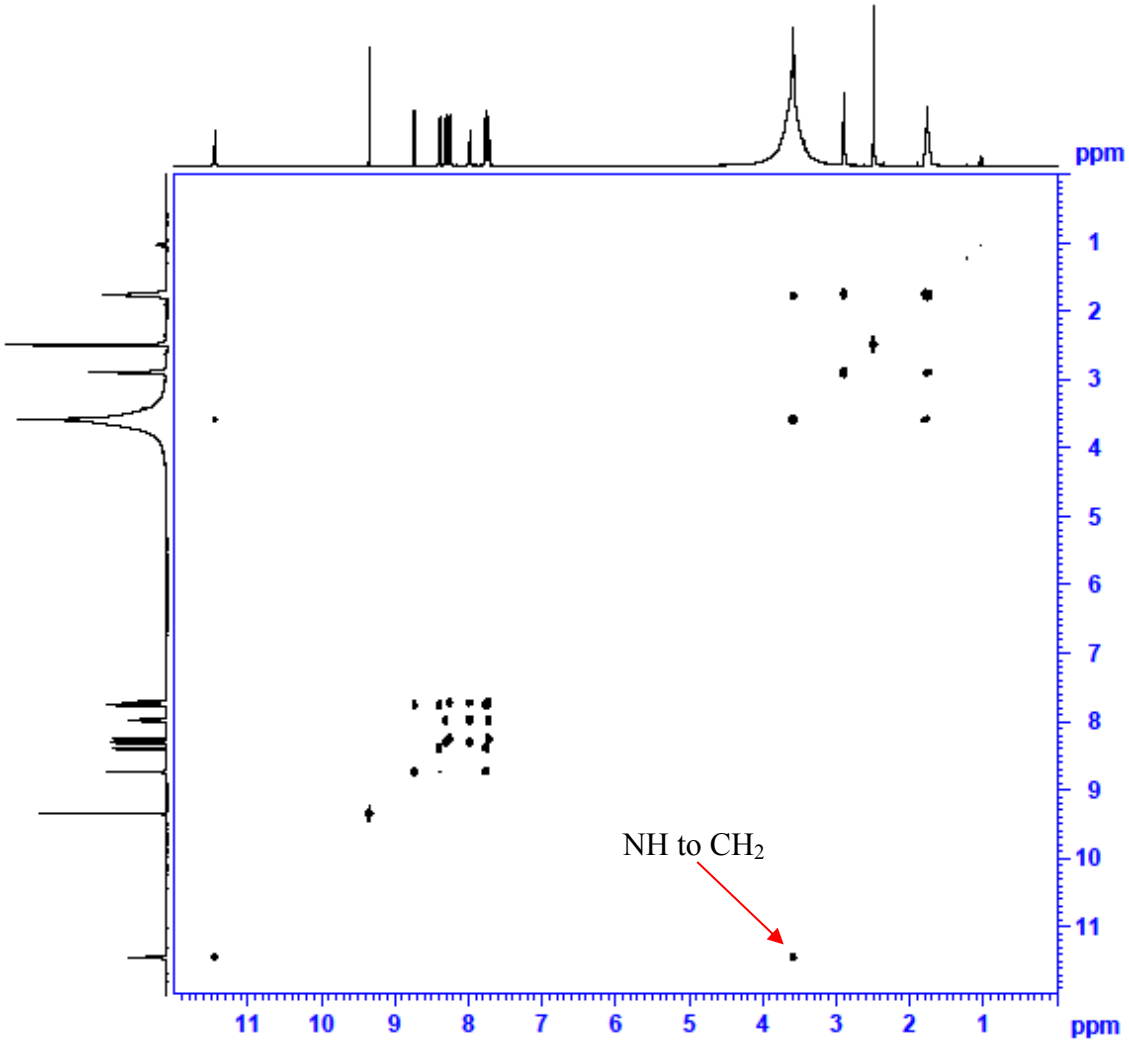



**Figure 2.11** 500 MHz magnitude COSY NMR spectrum in d6-DMSO of C4-DACA at 300 K. This highlights the NOE peak between the NH proton and adjacent $CH_2$ protons of the side chain, indicated by the red arrow.

The 500 MHz magnitude COSY spectrum seen in **Figure 2.11** shows the NOE peaks between the NH peak and the adjacent $CH_2$ allowing the identification of the H15 protons at 3.59 ppm. An expanded view of this spectrum is shown in **Figure 2.12** which afforded the assignment of the acridine protons.

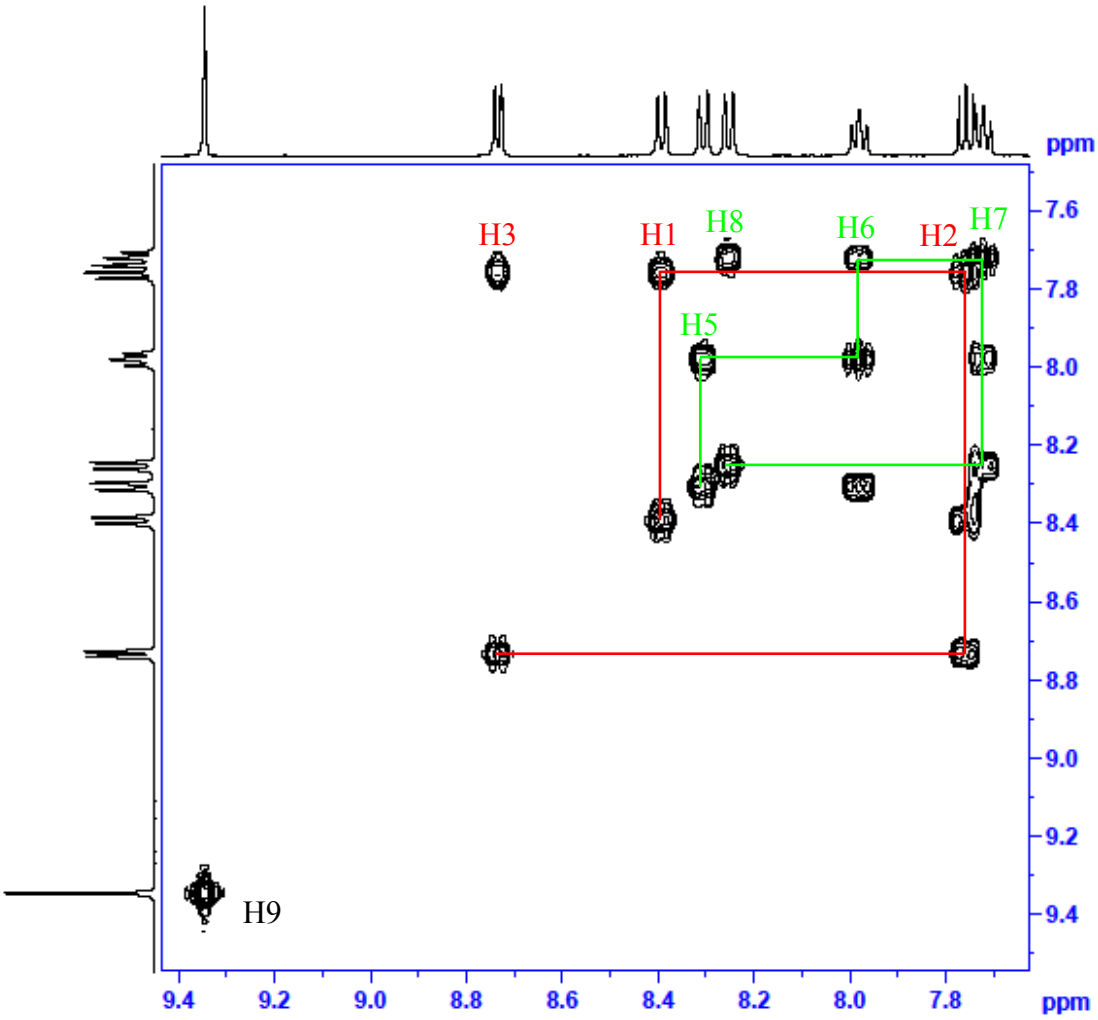



**Figure 2.12** Aromatic section of a 500 MHz magnitude COSY NMR spectrum in d6-DMSO of C4-DACA at 300 K.

The 500 MHz NOESY spectrum with 900 ms mixing time presented in **Figure 2.10** shows NOE peaks from the H9 proton to the H1 and the H8 protons as a result of their proximity. This information was used as a starting point for assigning the aromatic section of the acridine ring using the 500 MHz magnitude COSY spectrum shown in **Figure 2.12**. The coupled protons of the acridine ring are in a group of three protons (highlighted in red) and a group of four protons (highlighted in green) which together with the coupling information allowed the complete assignment. Once the group of four and three were identified, the doublet at 8.40 ppm must be H1 and the doublet at 8.25 ppm must be H8. This allows unequivocal identification of all other aromatic protons.

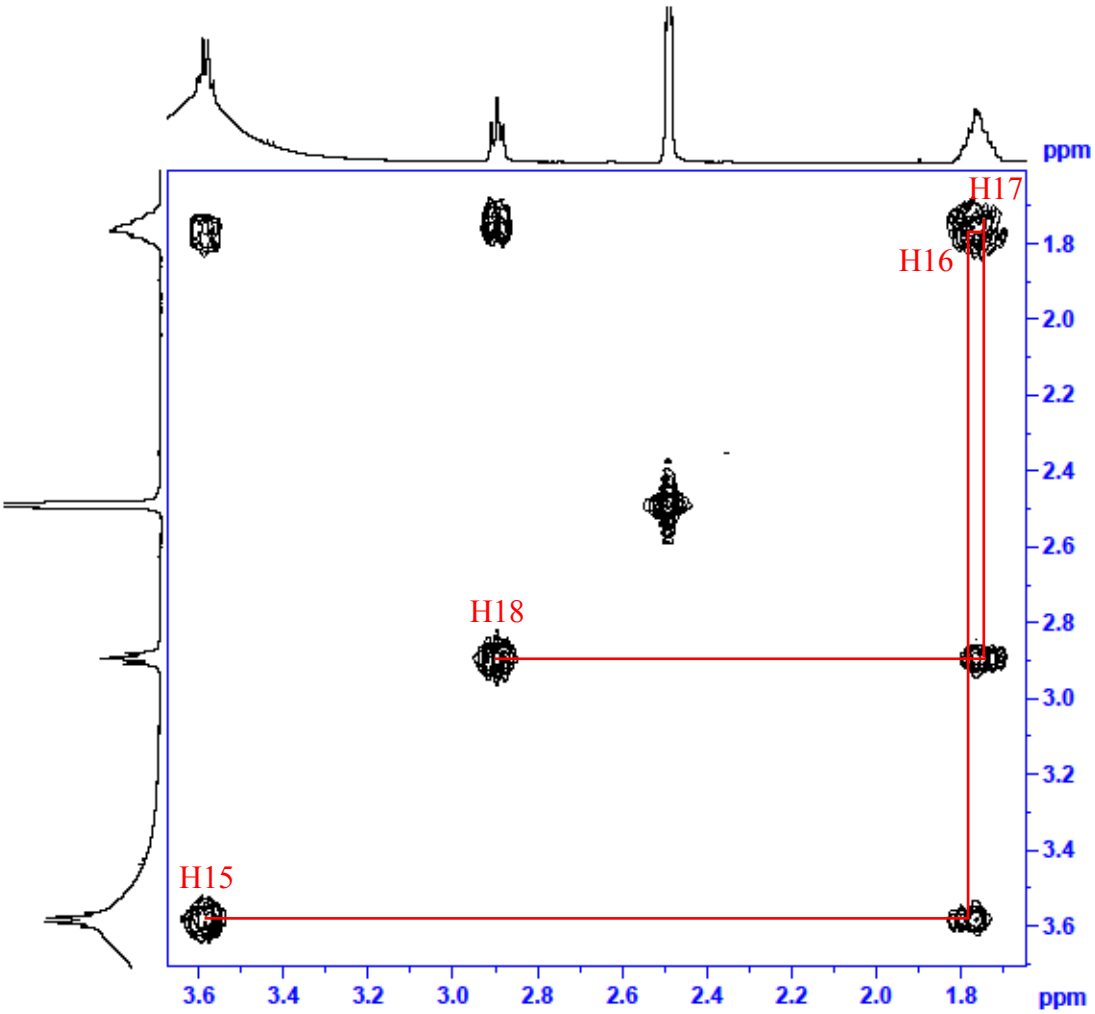



**Figure 2.13** The side chain section of a 500 MHz magnitude COSY NMR spectrum in d6-DMSO of C4-DACA at 300 K.

The expanded high field section of the 500 MHz magnitude COSY spectrum presented in **Figure 2.13** allowed the assignment of the protons of the side chain. The starting point is given from the coupling information between the NH proton and $CH_2$ protons from **Figure 2.11**. The multiplet coupled to the NH proton is the H15 protons at 3.59 ppm with the coupling along the chain indicated by the red line. The H15 protons give an NOE peak correlating to the H16 protons at 1.78 ppm, which in turn sees the H17 protons at 1.74 ppm, which finally couples to the H18 protons at 2.89 ppm.

**Table 2.1** $^{1}H$ and $^{13}C$ chemical-shift values in ppm of C4-DACA in d6-DMSO at 300 K and 500 MHz. Carbon peaks assigned using ChemDraw and compared to literature compounds are indicated by an asterisk (*).

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
|---|---|---|---|---|---|---|---|---|---|---|
| **H-1** | 8.40 | 7.76 | 8.74 | N/A | 8.31 | 7.98 | 7.72 | 8.25 | 9.35 | N/A |
| **C-13** | 132.9 | 125.3 | 134.6 | 128.3* | 128.7 | 131.9 | 126.6 | 128.5 | 138.7 | 126.5* |

| | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **NH** |
|---|---|---|---|---|---|---|---|---|---|
| **H-1** | N/A | N/A | N/A | N/A | 3.59 | 1.78 | 1.74 | 2.89 | 11.45 |
| **C-13** | 147.1* | 125.7* | 145.5* | 164.8[1]* | 38.52 | 26.29 | 24.86 | 38.67 | N/A |

[1] The two $^{13}C$ peaks at 164.3 and 164.8 ppm were ambiguous (C14 or the oxalate salt).

### 2.5.2 C4-DACA in $H_2O/D_2O$ NMR Sample

Since it was hoped to react C4-DACA with DNA in the presence of formaldehyde, the C4-DACA NMR sample was studied in a 10% $D_2O$/90% $H_2O$ solution at low concentration. This NMR sample consisted of 450 µl of 2 mM C4-DACA (in MilliQ $H_2O$) and 50 µl $D_2O$ (final concentration of 1.8 mM). The high water peak was reduced by use of a standard presaturation (watergate) pulse sequence. The full carbon assignment was not necessary in this solvent as the DNA related experimentation would be essentially proton based.

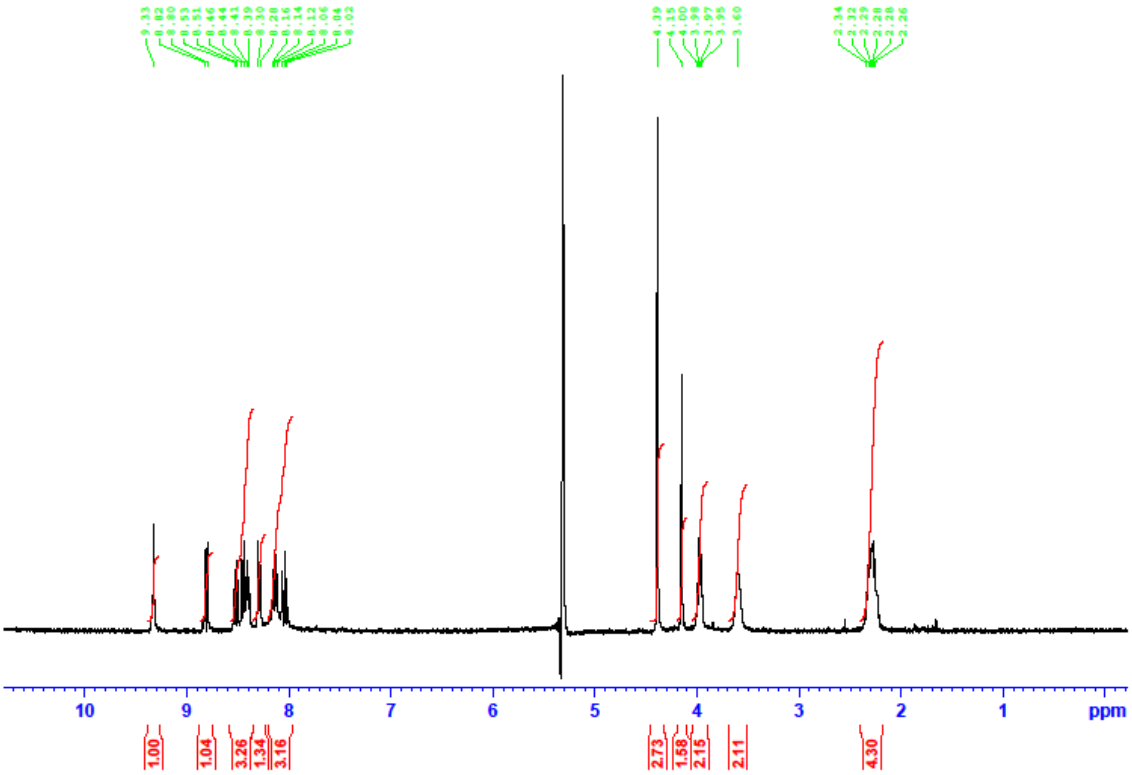



**Figure 2.14** 400 MHz $^{1}$H NMR spectrum with water suppression (watergate) in $D_2O$ of C4-DACA at 300 K.

The 400 MHz $^{1}$H NMR spectrum employing the watergate sequence was used to reduce the large water peak and is shown in **Figure 2.14**. Compared with the spectrum in d6-DMSO, the side chain protons are considerably broadened and coupling information lost as a result of different orientations of the side chain. The multiplet at 2.86 ppm is the overlapping H16 and H17 protons, while the H18 protons are seen at 3.60 ppm. The H15 protons at 3.97 ppm were coupled to the low field NH proton at 11.42 ppm (not shown).

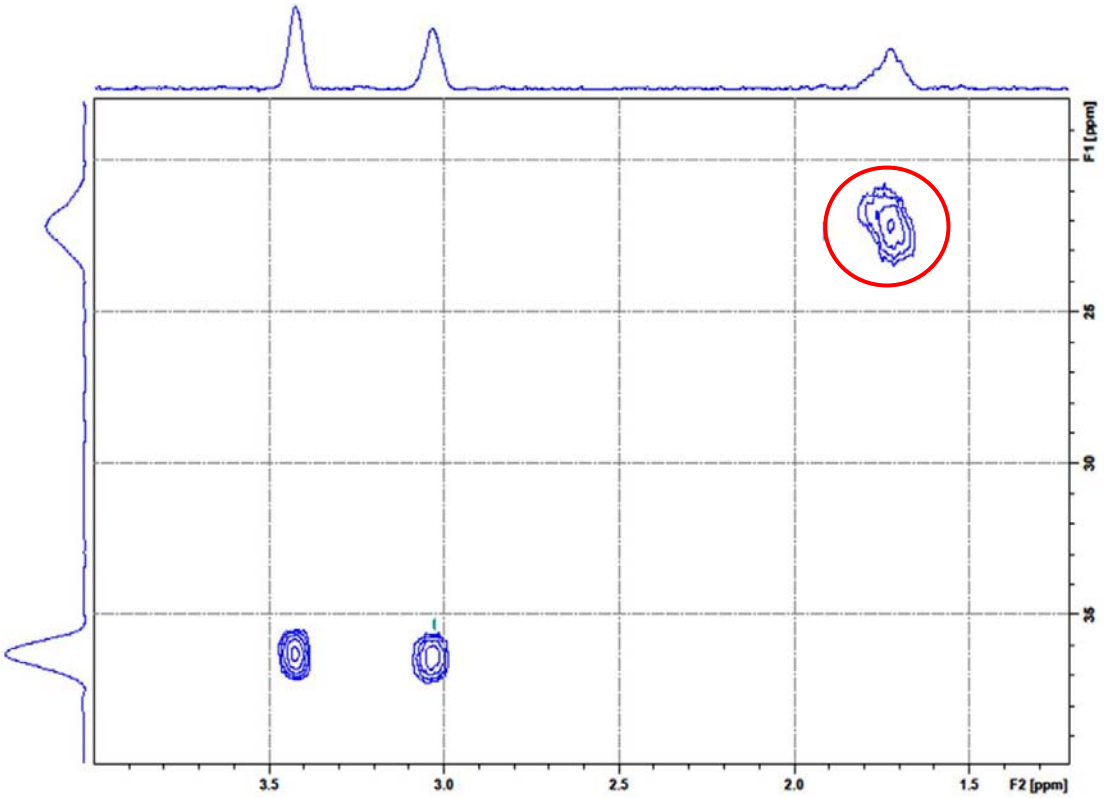



**Figure 2.15** Expanded view of the high field section of a 400 MHz HSQC NMR spectrum of C4-DACA in $D_2O$ at 300 K. The red circle indicates the peak where the H16 and H17 side chain protons overlap.

The HSQC spectrum from **Figure 2.15** allows the separation of the C16 and C17 carbons at 21.6 and 22.4 ppm, the C18 carbon at 36.5 ppm, and the C15 carbon at 36.3 ppm. The aromatic section of the HSQC spectrum presented in **Figure 2.16** allows the acridine protons to be separated and identified. These carbon shifts were consistent with model $^{13}C$ prediction using the HIPPO NMR software.

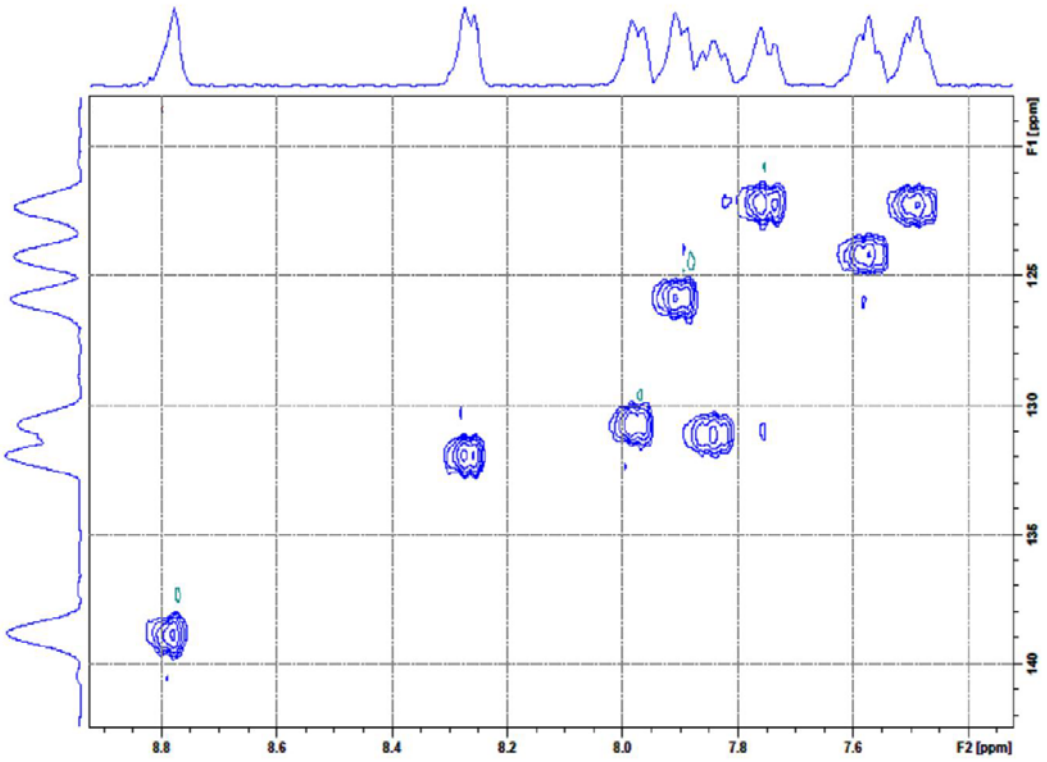



**Figure 2.16** Expanded view of the aromatic portion of the 400 MHz HSQC NMR spectrum of C4-DACA in $D_2O$ at 300 K. This shows how the protons overlap but the carbons allow separation and identification.

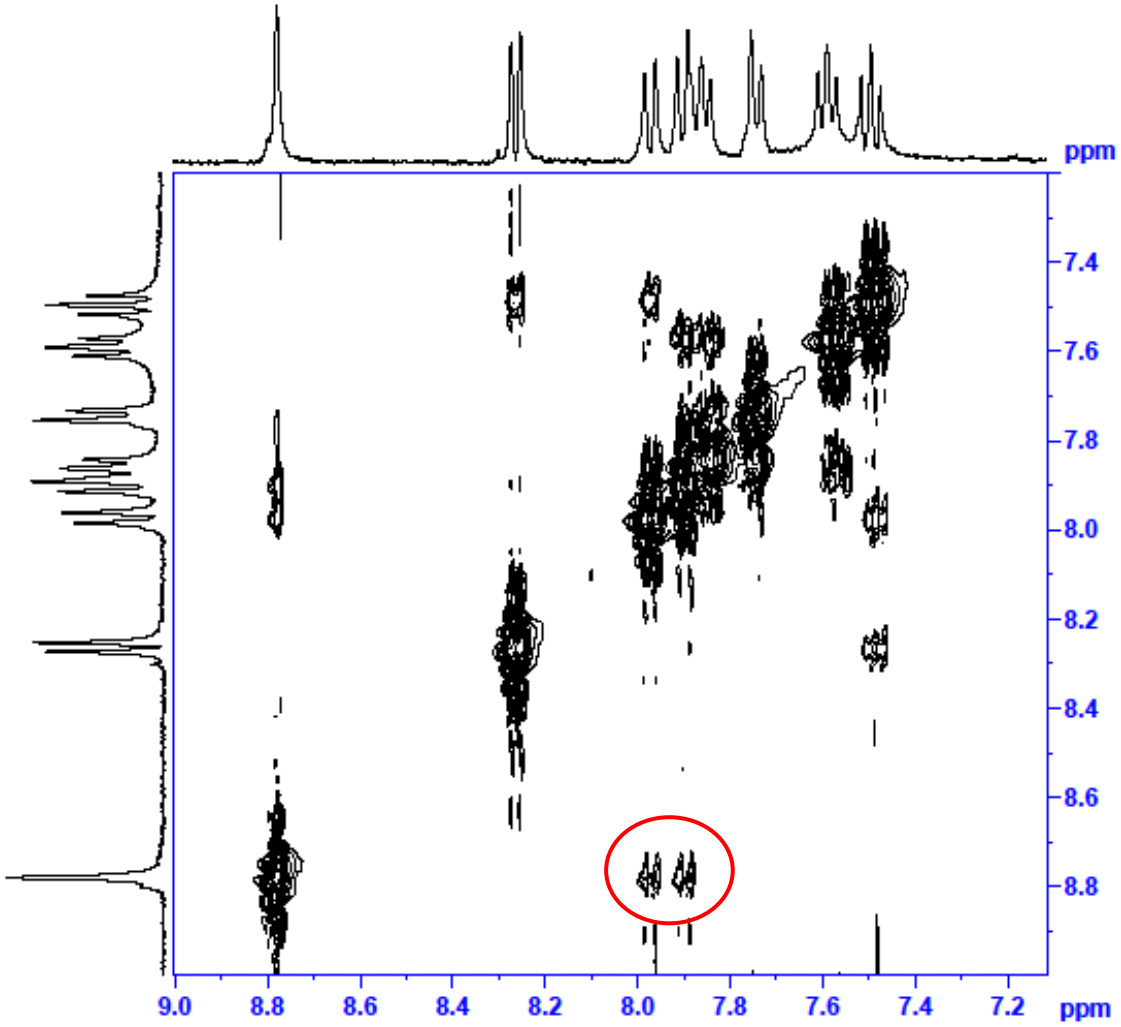



**Figure 2.17** Aromatic section of a 400 MHz 900 ms mixing time NOESY NMR spectrum of C4-DACA in $D_2O$ at 300 K. This shows the NOE peaks from the H9 "singlet" at 9.32 ppm to the H1 "doublet" at 8.52 ppm and the H8 "doublet" at 8.45 ppm, indicated by the red oval.

The coupling information from the 900 ms mixing time NOESY spectrum presented in **Figure 2.17** allowed the assignment of the H9 "singlet" at 9.32 ppm. The H9 proton gave NOE peaks to the H1 "doublet" at 8.52 ppm and the H8 "doublet" at 8.45 ppm, indicated by the red oval.

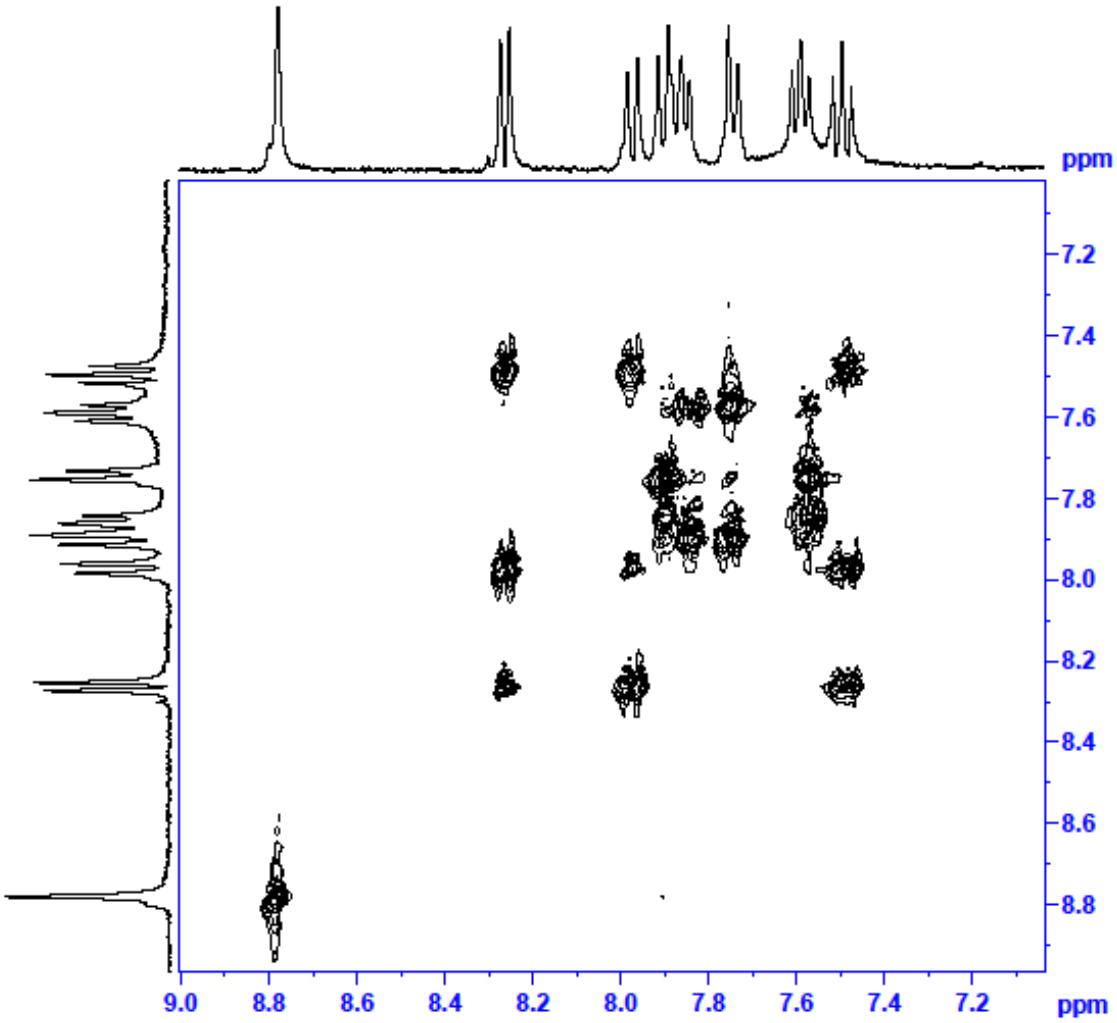



**Figure 2.18** Aromatic section of a 400 MHz 120 ms mixing time TOCSY NMR spectrum of C4-DACA in $D_2O$ at 300 K.

The absolute assignment of C4-DACA in $D_2O$ is presented in **Table 2.2**. This assignment was achieved by using data from a 120 ms mixing time TOCSY (**Figure 2.18**), which gave COSY type peaks, and from a 900 ms NOESY (**Figure 2.17**). The NOESY spectrum showed NOE cross peaks from the H9 proton at 9.32 ppm to the H1 proton at 8.52 ppm and the H8 proton at 8.45 ppm as a result of their proximity. The 120 ms TOCSY spectrum allowed the division of the coupled protons in the acridine ring into a group of three protons and a group of four protons which together with the coupling allowed the complete assignment. Once the group of three and four peaks are identified, the crosspeak from the H9 proton to the doublet at 8.51 ppm within the group of three must be the H1 proton, and

the NOE peak from the H9 proton to the doublet at 8.44 ppm within the group of four must be the H8 proton. This allowed identification of all other protons except H6 and H7 which were identified from a magnitude COSY spectrum (not shown). A standard HSQC (**Figure 2.15**) assisted in the assignment of the side chain protons. The numbering system employed in assigning C4-DACA is shown in **Figure 2.1**.

**Table 2.2** $^{1}H$ chemical-shift values in ppm of C4-DACA in $H_2O/D_2O$ at 300 K and 400 MHz.

| H1 | H2 | H3 | H5 | H6 | H7 | H8 | H9 | H15 | H16/17 | H18 | NH |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 8.52 | 8.04 | 8.81 | 8.29 | 8.13 | 8.41 | 8.45 | 9.32 | 3.97 | 2.28 | 3.60 | 11.42 |

### 2.5.3 Electrospray Ionisation (ESI) Mass Spectrum of C4-DACA

The ESI mass spectrum of C4-DACA, **Figure 2.20**, shows the molecular ion at 294.1 m/z as expected which corresponds to the structure shown in **Figure 2.19**.

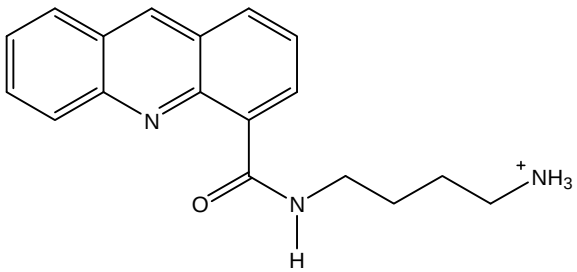



**Figure 2.19** Structure of C4-DACA molecular ion (MW 294.16 g/mol).

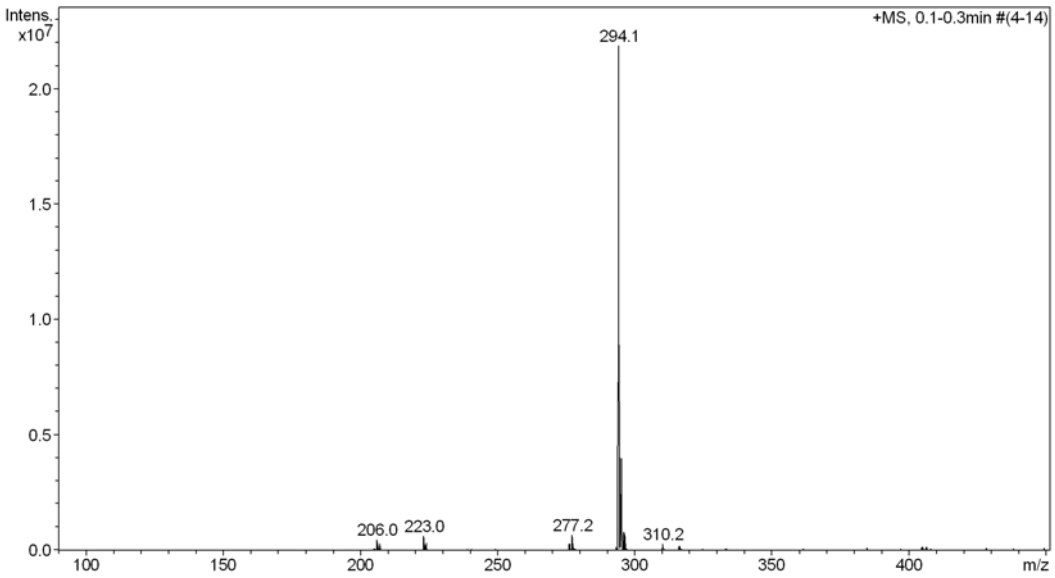



**Figure 2.20** ESI mass spectrum of C4-DACA.

### 2.5.4. Synthesis and Characterisation of C2-DACA

The synthesis of C2-DACA was identical to the synthesis outlined in **2.4** of C4-DACA except that the activated acid was added dropwise to a stirred solution of 1,2-diaminoethane (0.5 g, 0.0079 mol) in dichloromethane (10 mL) to achieve the shorter side chain. All spectra were calibrated to DMSO (2.49 or 39.5 ppm) or H9 of C2-DACA (9.33 ppm). The numbering system employed for the assignment of C2-DACA is shown in **Figure 2.21**.

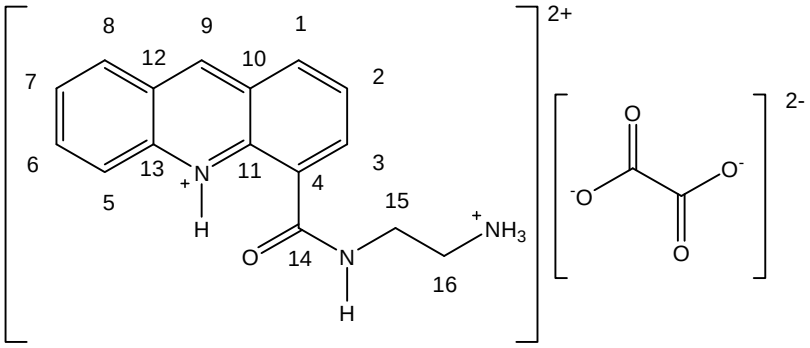



**Figure 2.21** Structure of C2-DACA.oxalate (MW = 355.12).

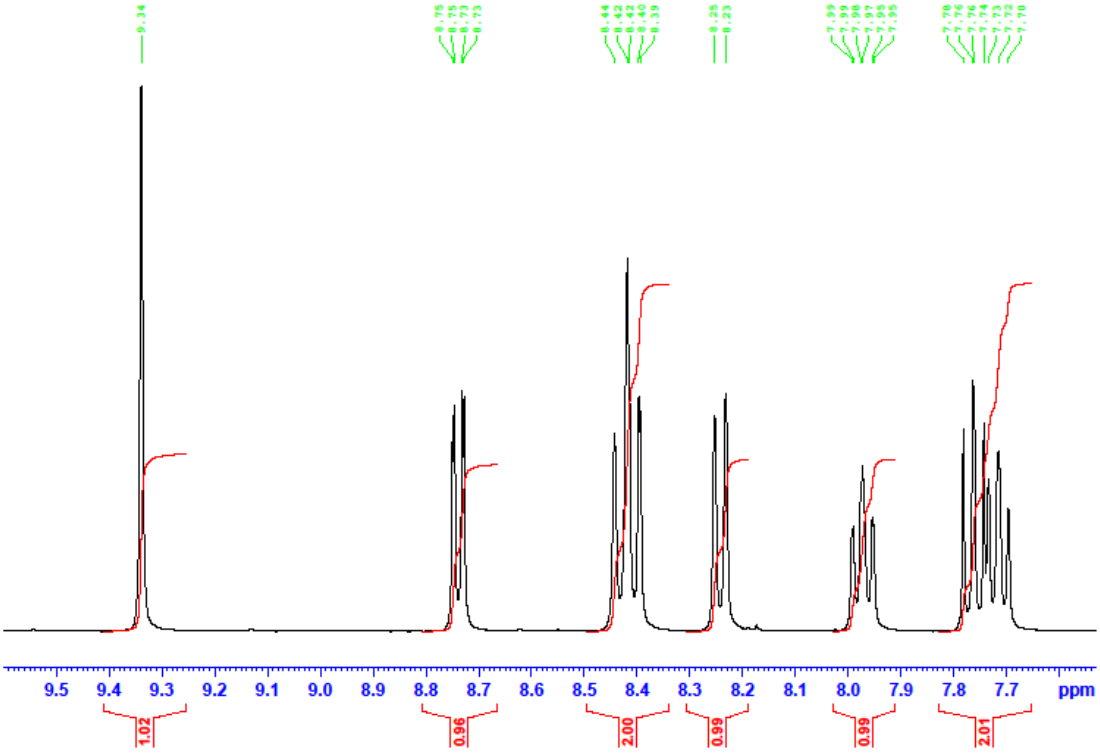



**Figure 2.22** Expanded aromatic section of a 400 MHz $^{1}$H NMR spectrum of C2-DACA in d6-DMSO showing the splitting pattern of the acridine protons.

The aromatic section of a 400 MHz $^{1}$H spectrum of C2-DACA is shown in **Figure 2.22**. There are four "doublets" and three "triplets" as expected for the acridine protons, similar to that of C4-DACA but overlapping in different sections. An expansion of the high field side chain section presented in **Figure 2.23** revealed coupling from the "quartet" at 3.81 ppm to the adjacent NH at 11.52 ppm indicating that the peak was the H15 proton. The H16 and H15 protons of the side chain are at 3.16 and 3.81 ppm respectively with appropriate splitting.

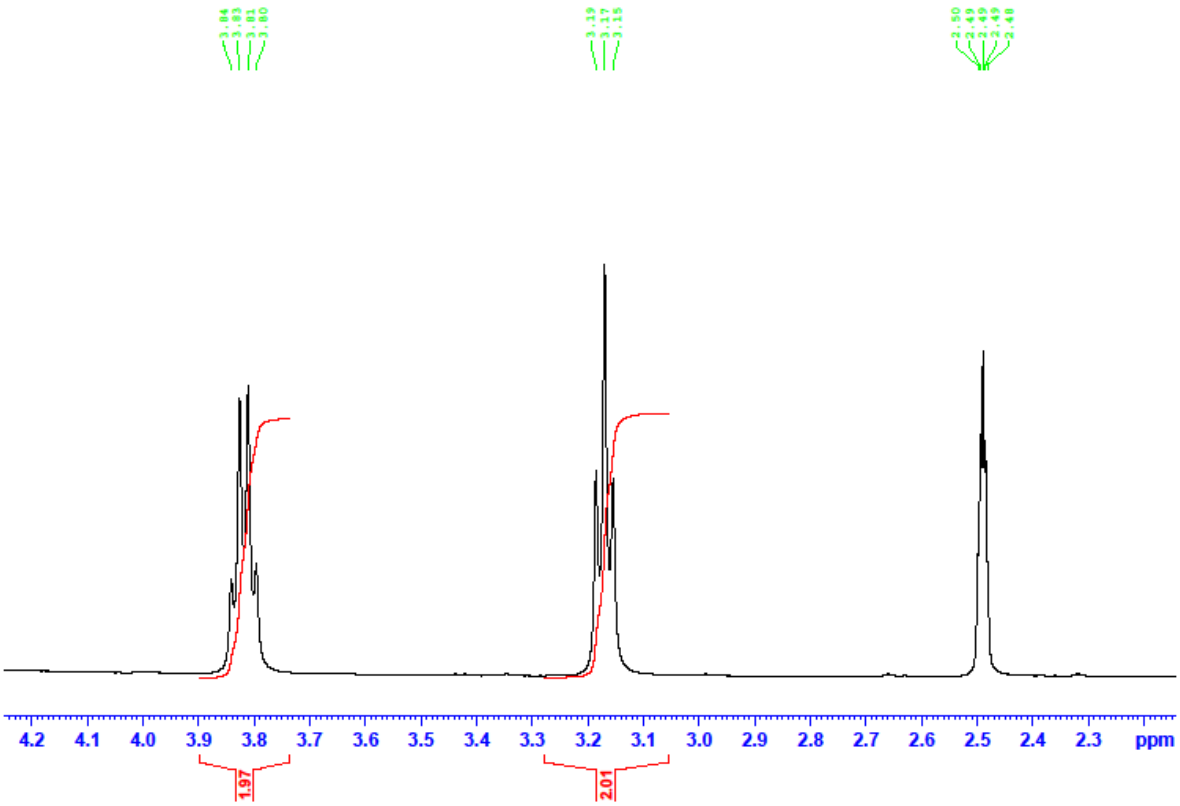



**Figure 2.23** Expanded high field section of a 400 MHz $^{1}$H NMR spectrum of C2-DACA in d6-DMSO showing the splitting pattern of the side chain protons. The "quartet" at 3.81 ppm showed coupling to the adjacent NH proton hence was assigned as the H15 proton.

The carbon spectra in **Figure 2.24** and **2.25** show that there are 17 carbon peaks as expected (C2-DACA contains 16 carbons, oxalate contains 2 equivalent carbons) with appropriate integration. The C15 carbon of the side chain appears under the DMSO multiplet, and is highlighted by the red arrow in **Figure 2.25**.

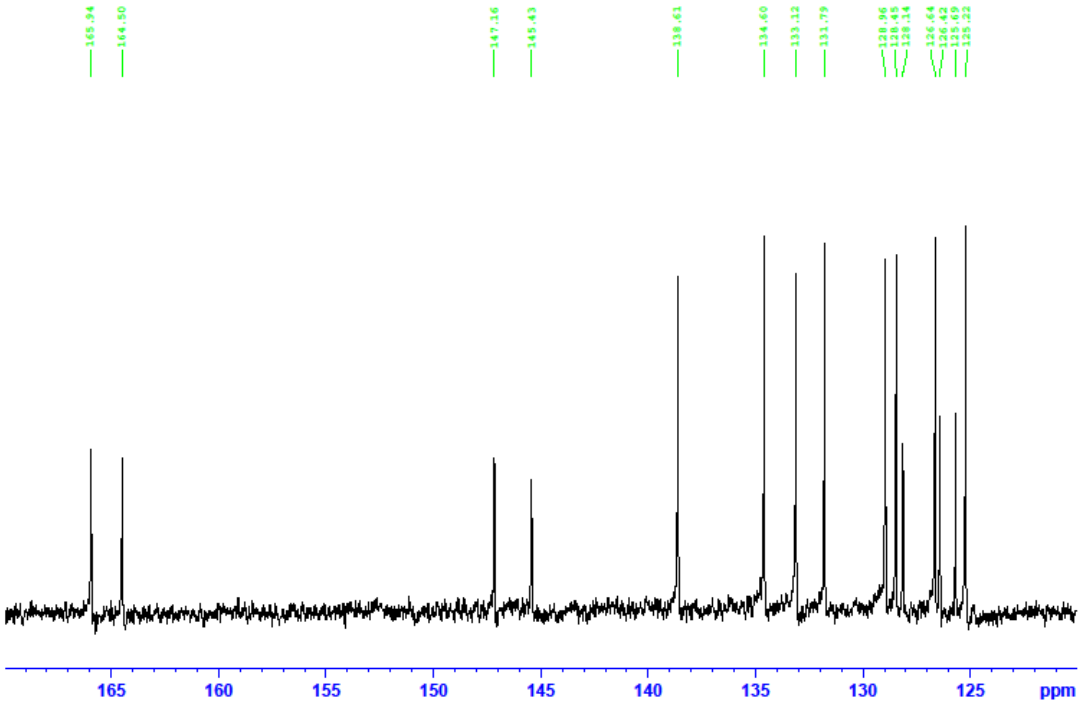



**Figure 2.24** 400 MHz $^{13}$C NMR spectrum of C2-DACA in d6-DMSO showing the acridine and carboxy carbons.

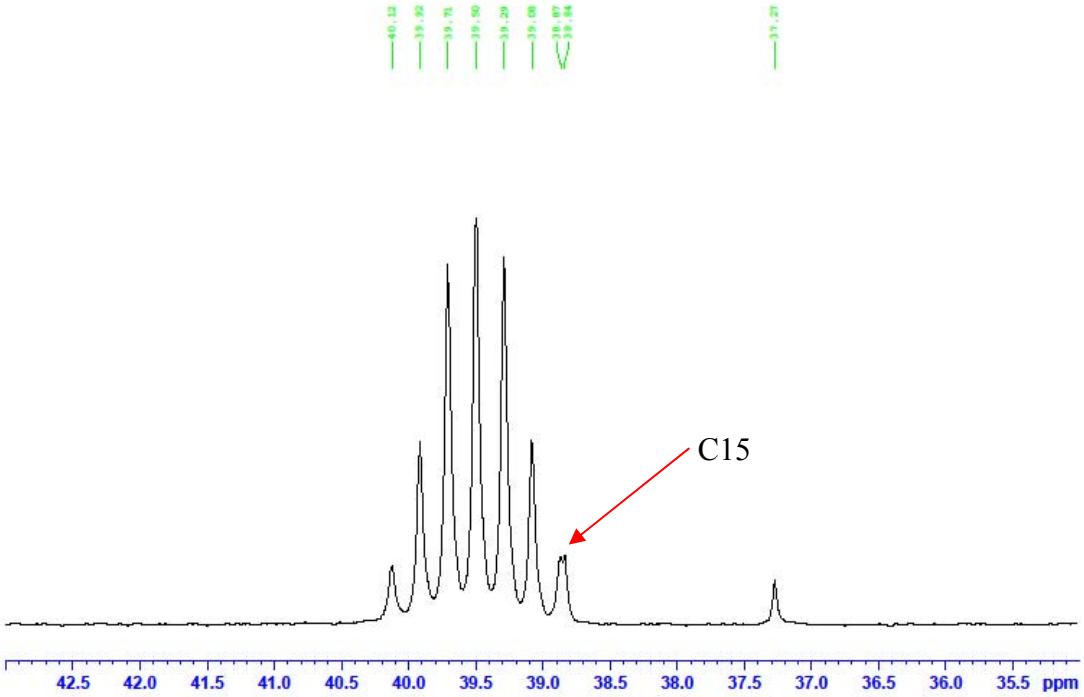



**Figure 2.25** Expanded section of a 400 MHz $^{13}$C NMR spectrum of C2-DACA in d6-DMSO showing the side chain carbons. The C15 carbon is hidden under the DMSO peak at 38.84 ppm, and the C16 carbon is at 37.27 ppm.

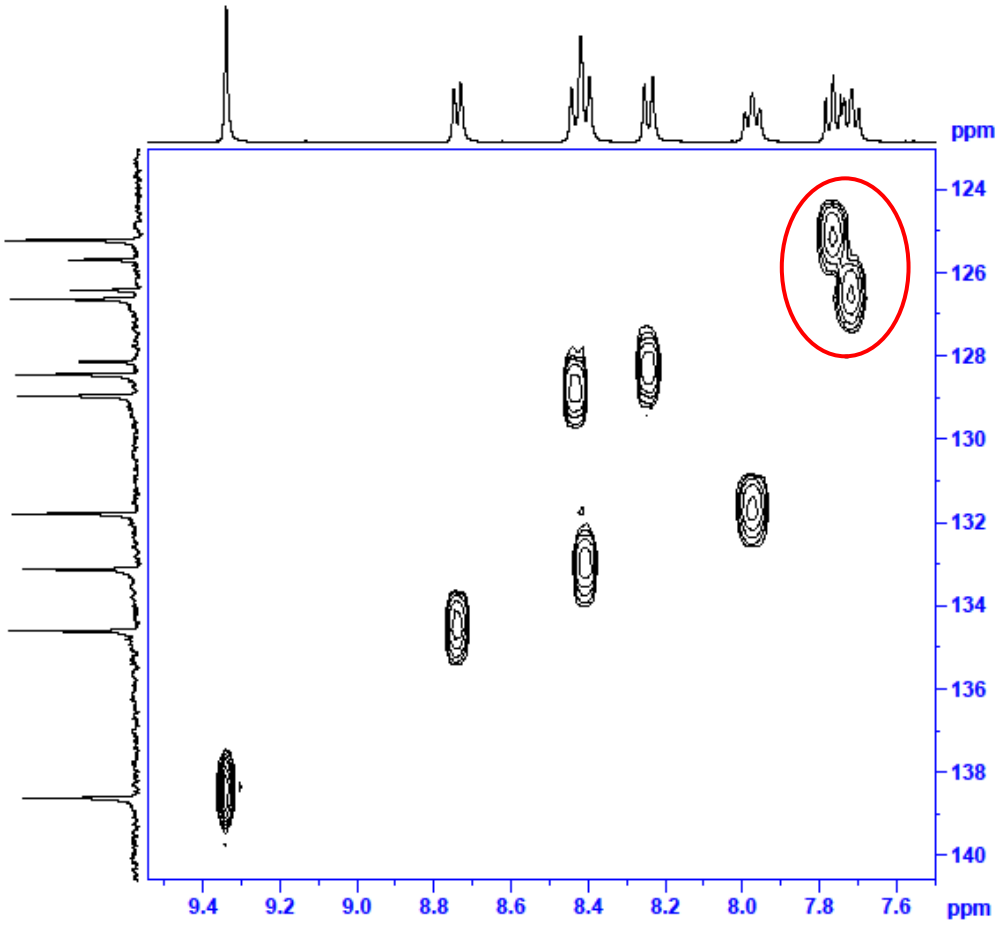



**Figure 2.26** Expanded aromatic section of a 400 MHz HSQC NMR spectrum of C2-DACA in d6-DMSO. This shows that the peaks at 7.7 ppm and 8.4 ppm are actually two separate overlapping protons on individual carbons, indicated by the red oval.

The HSQC spectrum presented in **Figure 2.26** allowed the identification of C15 at 38.84 ppm and C16 at 37.27 ppm, and also allowed the acridine protons to be separated. The overlapping multiplet at 7.7 ppm in the proton spectrum is attached to two separate carbons at 125.4 and 126.7 ppm, whereas the "triplet" at 8.4 ppm corresponds to the two carbon peaks at 129.0 and 133.2 ppm. These carbon shifts are also consistent with model $^{13}C$ prediction using the ChemDraw NMR software.

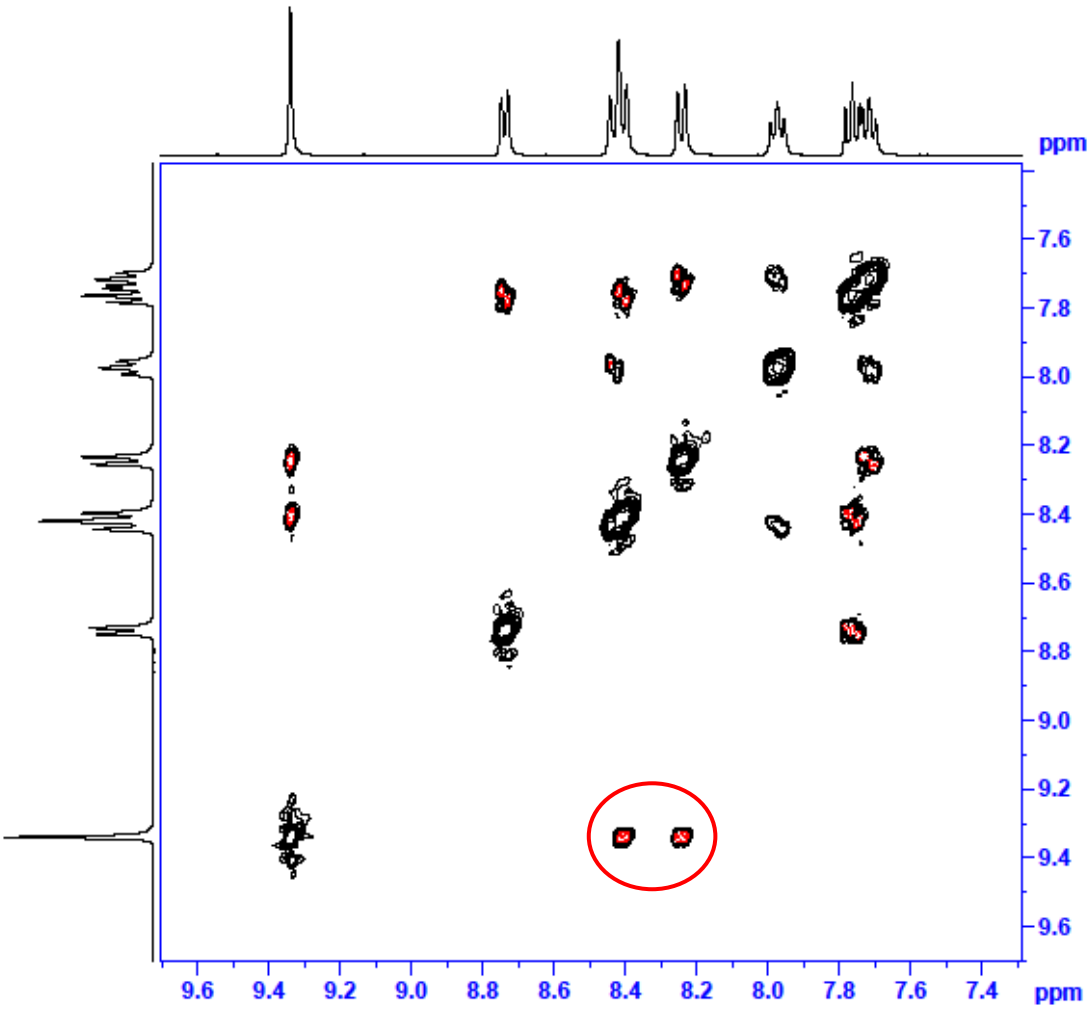



**Figure 2.27** Expanded aromatic section of a 400 MHz NOESY NMR spectrum of C2-DACA in d6-DMSO. This shows the NOE peaks from the H9 proton at 9.30 ppm to the H1 proton at 8.37 ppm and to the H8 proton at 8.21 ppm, indicated by the red oval.

The NOESY spectrum presented in **Figure 2.27** showed a correlation from the H9 proton at 9.30 ppm to the H1 proton at 8.37 ppm and also to the H8 proton at 8.21 ppm as a result of their proximity. This information was used as a starting point for assigning the protons of the acridine ring from the magnitude COSY spectrum in **Figure 2.28**. The coupled protons of the acridine ring are in a group of three protons (highlighted in red) and a group of four protons (highlighted in green) which together with the coupling allowed the complete assignment. Once the group of four and three have been identified, the part of the multiplet at 8.37

ppm must be the H1 proton and the doublet at 8.21 ppm must be the H8 proton. This allowed unequivocal identification of all other acridine protons.

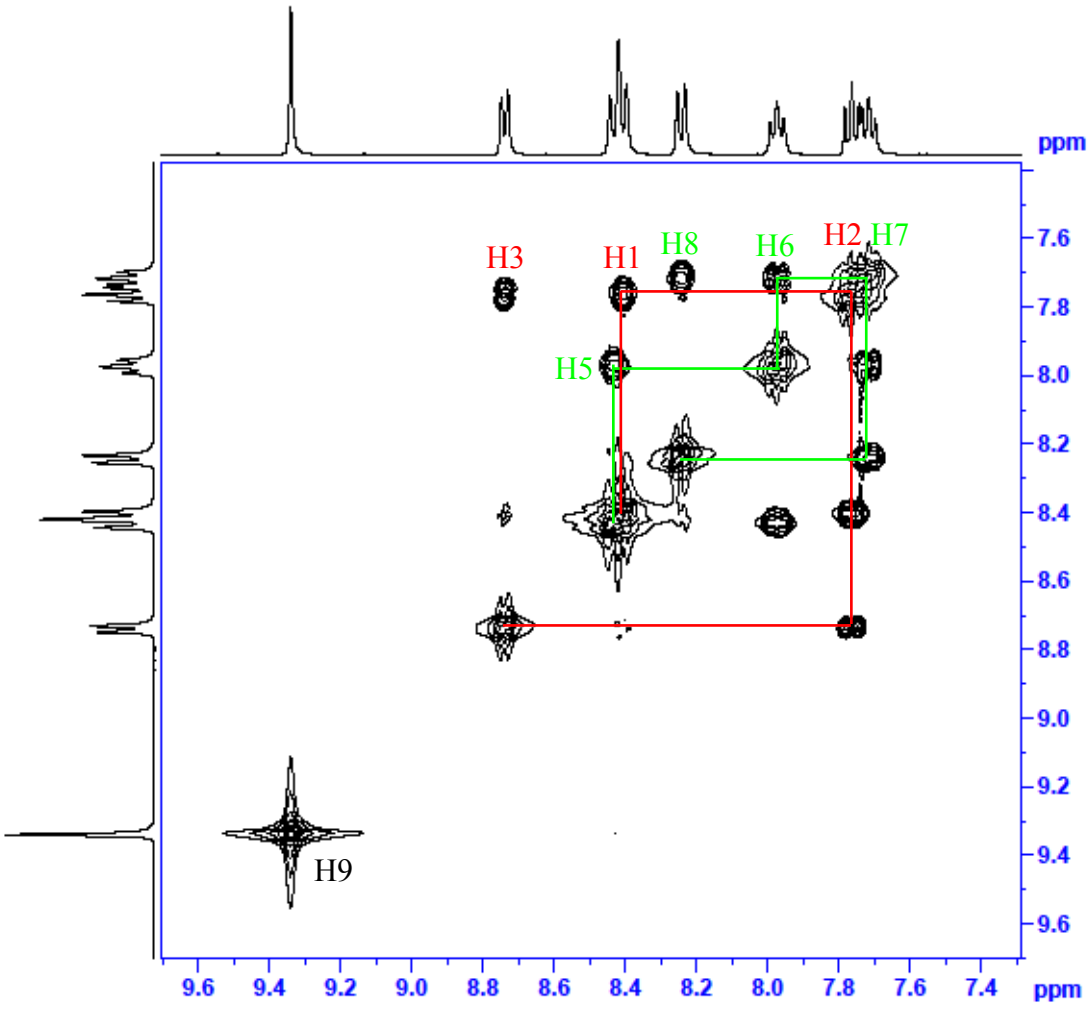



**Figure 2.28** Expanded aromatic section of a 400 MHz magnitude COSY NMR spectrum of C2-DACA in d6-DMSO. This shows the groups of three protons (H1 to H3) and four protons (H5 to H8) and defines the H6 and H7 protons via the crosspeaks from the H5 proton to the H6 proton, and also the crosspeaks of the H8 proton to the H7 proton.

### 2.5.5. NMR Assignments of C2-DACA

The 1D and 2D NMR methods summarised in **Table 2.3** showed that there were two separate $CH_2$ side chain peaks in the proton spectrum (3.13 and 3.78 ppm) and in the carbon spectrum (37.4 and 38.9 ppm), consistent with the structure in **Figure 2.21** as expected. The 2D HSQC spectrum also gave crosspeaks where expected (H/C16 at 3.13, 38.9 ppm and H/C15 at 3.78, 37.4 ppm).

**Table 2.3** $^{1}H$ and $^{13}C$ chemical-shift values in ppm of C2-DACA in d6-DMSO at 300 K and 400 MHz. Carbon peaks assigned using ChemDraw and compared to literature compounds are indicated by an asterisk (*).

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
|---|---|---|---|---|---|---|---|---|---|---|
| **H-1** | 8.39 | 7.75 | 8.73 | N/A | 8.42 | 7.96 | 7.70 | 8.23 | 9.33 | N/A |
| **C-13** | 133.2 | 125.4 | 134.7 | 128.1* | 129.0 | 131.9 | 126.7 | 128.6 | 138.7 | 126.4* |

| | **11** | **12** | **13** | **14** | **15** | **16** | **NH** |
|---|---|---|---|---|---|---|---|
| **H-1** | N/A | N/A | N/A | N/A | 3.81 | 3.16 | 11.52 |
| **C-13** | 147.2* | 125.7* | 145.4* | 164.5[1]* | 37.3 | 38.9 | N/A |

[1] The two $^{13}C$ peaks at 164.5 and 165.9 ppm were ambiguous (C14 or the oxalate salt).

### 2.5.6. Electrospray Ionisation (ESI) Mass Spectrum of C2-DACA

The ESI mass spectrum of C2-DACA, **Figure 2.30**, shows the molecular ion at 266.1 m/z as expected which corresponds to the structure shown in **Figure 2.29**.

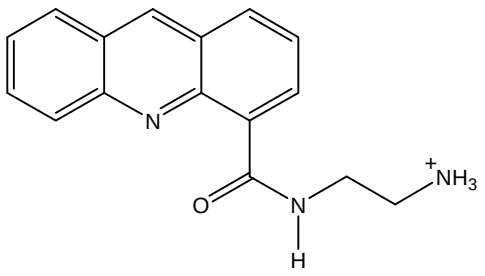



**Figure 2.29** Structure of C2-DACA molecular ion (MW 266.13 g/mol).

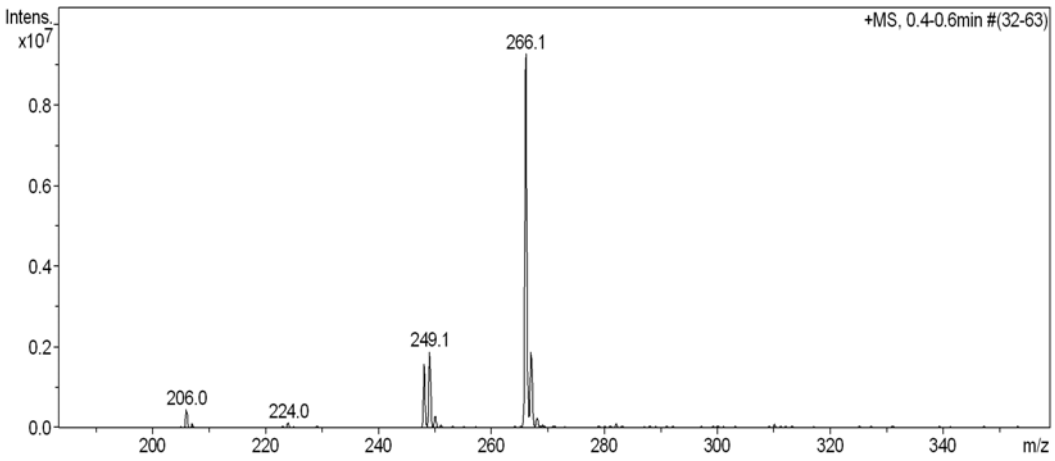



**Figure 2.30** ESI mass spectrum of C2-DACA.

### 2.5.7. Synthesis and Characterisation of C2-DACA dimer

The synthesis of the C2-DACA dimer was identical to the synthesis outlined in **2.4** of C2-DACA. The only difference was that the activated acid was added quickly (not dropwise) to the 1,2-diaminoethane (0.5 g, 0.0079 mol) in dichloromethane (10 mL). All spectra were calibrated to DMSO (2.49 or 39.5 ppm) or H9 of C2-DACA dimer (9.07 ppm). The numbering system employed for the assignment of the C2-DACA dimer is shown in **Figure 2.31**

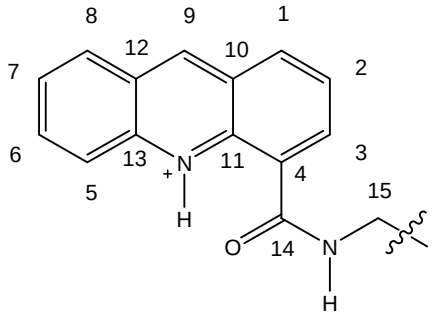



**Figure 2.31** Numbering system for C2-DACA dimer. The wavy line indicates the mirror plane of the dimer.

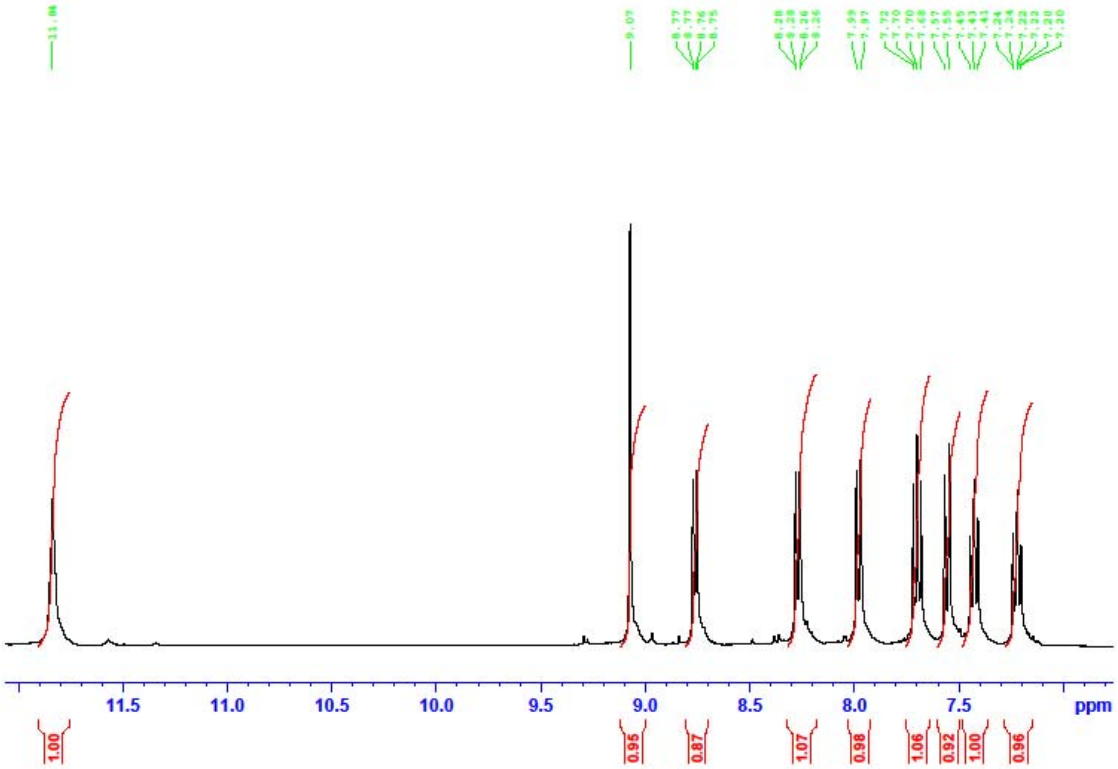



**Figure 2.32** Expanded low field section of a 400 MHz $^{1}$H NMR spectrum of the C2-DACA dimer in d6-DMSO showing the splitting pattern of the acridine protons and the NH at 11.84 ppm.

**Figure 2.32** shows the aromatic section with the expected splitting pattern of a 400 MHz $^{1}$H spectrum of the C2-DACA dimer. The main difference between the aromatic section of the dimer and monomer of C2-DACA is that none of the aromatic protons overlap. **Figure 2.33** is of the same spectrum but shows the H15 "singlet" at 3.97 ppm which was coupled to the NH at 11.84 ppm (not shown), and a large water peak at 3.3 ppm.

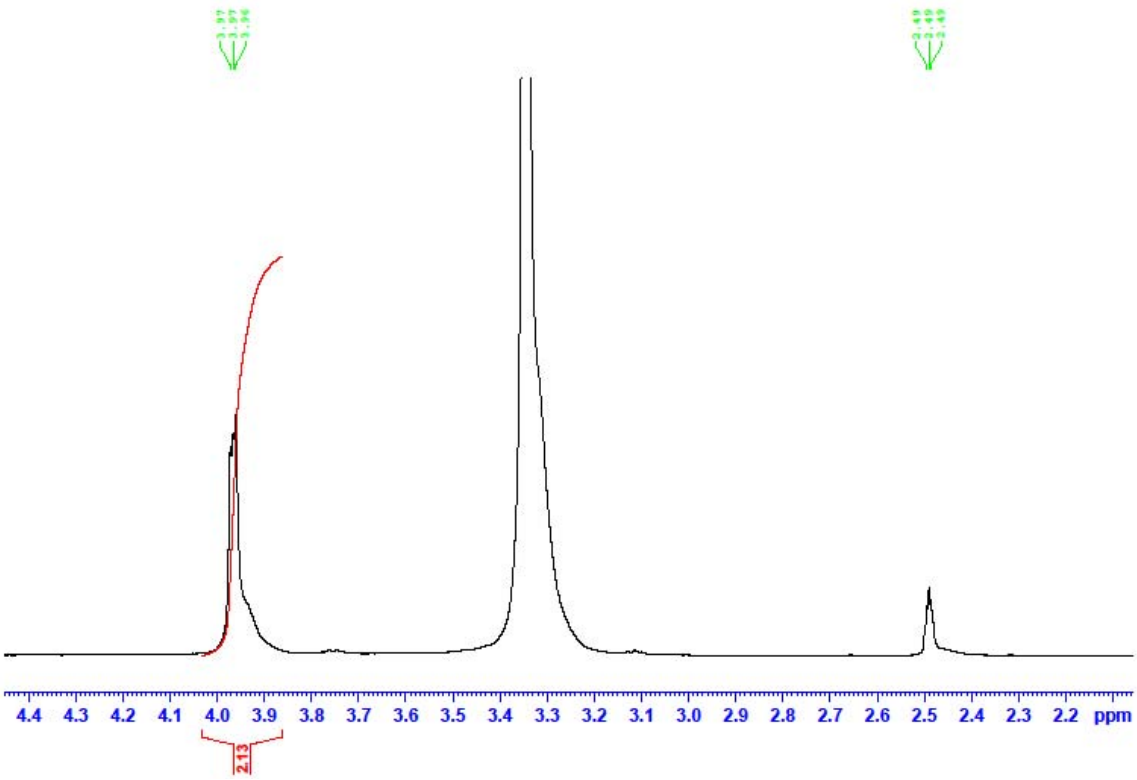



**Figure 2.33** Expanded high field section of a 400 MHz $^{1}$H NMR spectrum of the C2-DACA dimer in d6-DMSO showing the H15 protons. There is a large water peak at approximately 3.3 ppm and the DMSO peak at 2.49 ppm.

The 400 MHz $^{13}$C spectrum shown in **Figure 2.34** reveals that the C15 carbon peak at 39.17 ppm is hidden underneath the DMSO multiplet, indicated by the red arrow. As expected there are 15 carbon peaks in total between **Figures 2.34** and **2.35** due to the symmetry of the dimer.

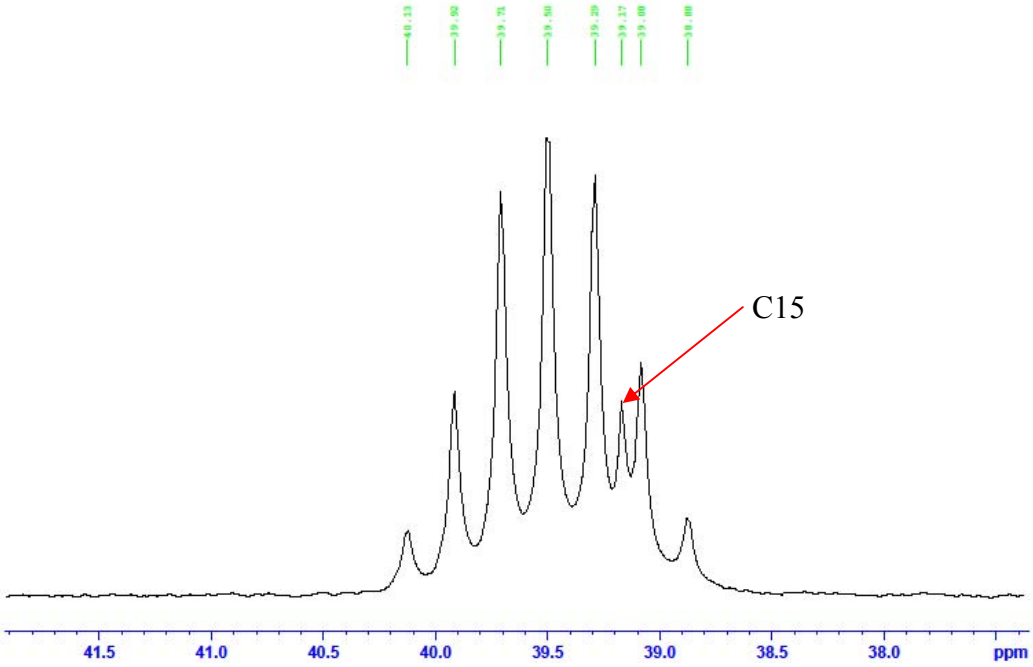



**Figure 2.34** Expanded section of a 400 MHz $^{13}$C NMR spectrum of the C2-DACA dimer in d6-DMSO showing the C15 carbon at 39.2 ppm under the DMSO multiplet.

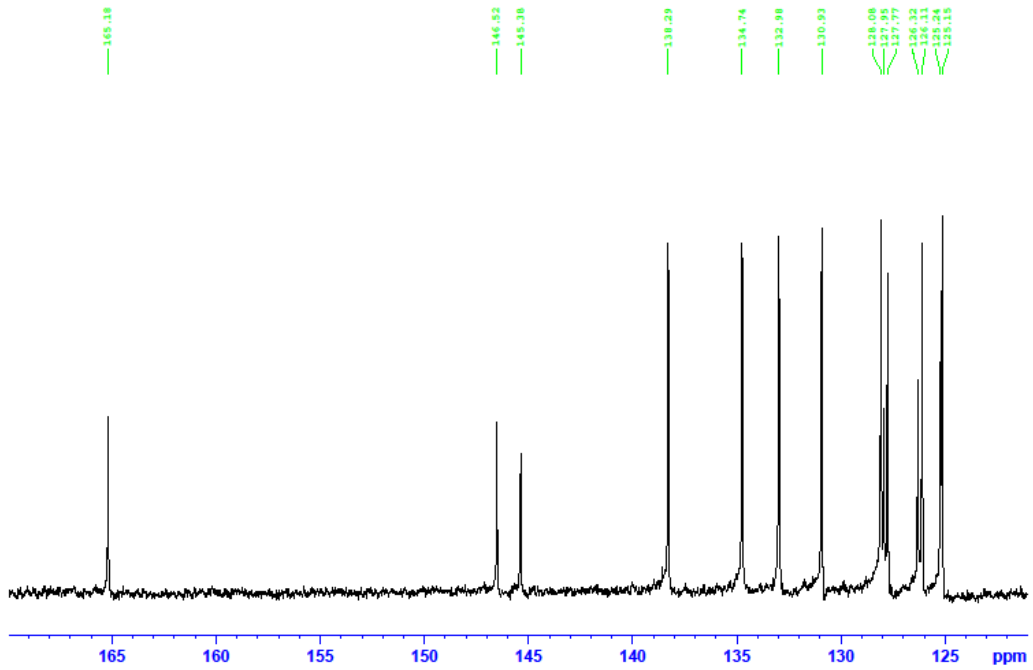



**Figure 2.35** Expanded aromatic section of a 400 MHz $^{13}$C NMR spectrum of the C2-DACA dimer in d6-DMSO showing the aromatic carbons with expected integration.

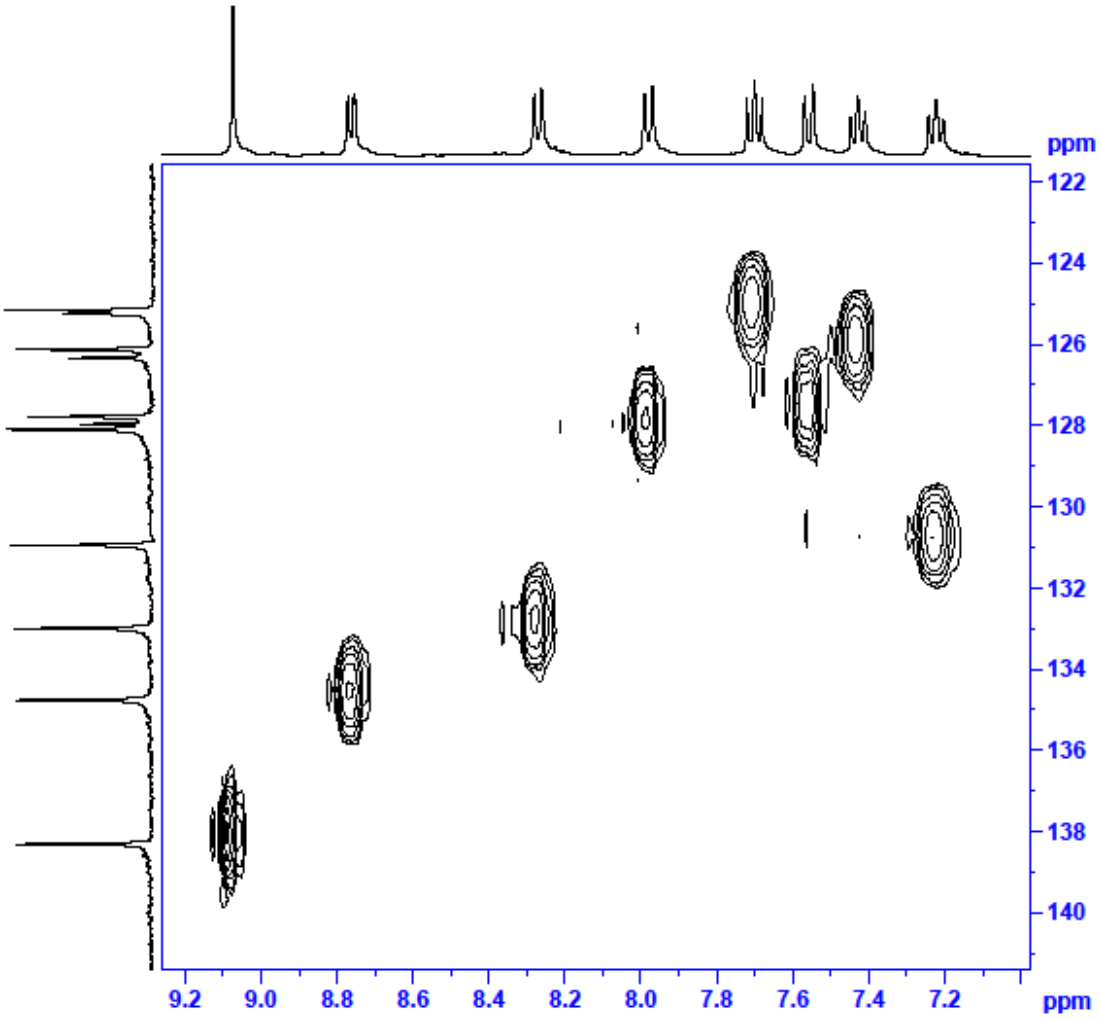



**Figure 2.36** Expanded aromatic section of a 400 MHz HSQC NMR spectrum of the C2-DACA dimer in d6-DMSO.

The 400 MHz HSQC NMR spectrum shown in **Figure 2.36** allowed the assignment of the acridine protons to their corresponding carbon peaks. This spectrum also allowed the identification of the C15 carbon at 39.17 ppm via **Figure 2.37**, indicated by the red arrow. These carbon shifts are also consistent with model $^{13}C$ prediction using the ChemDraw NMR software.

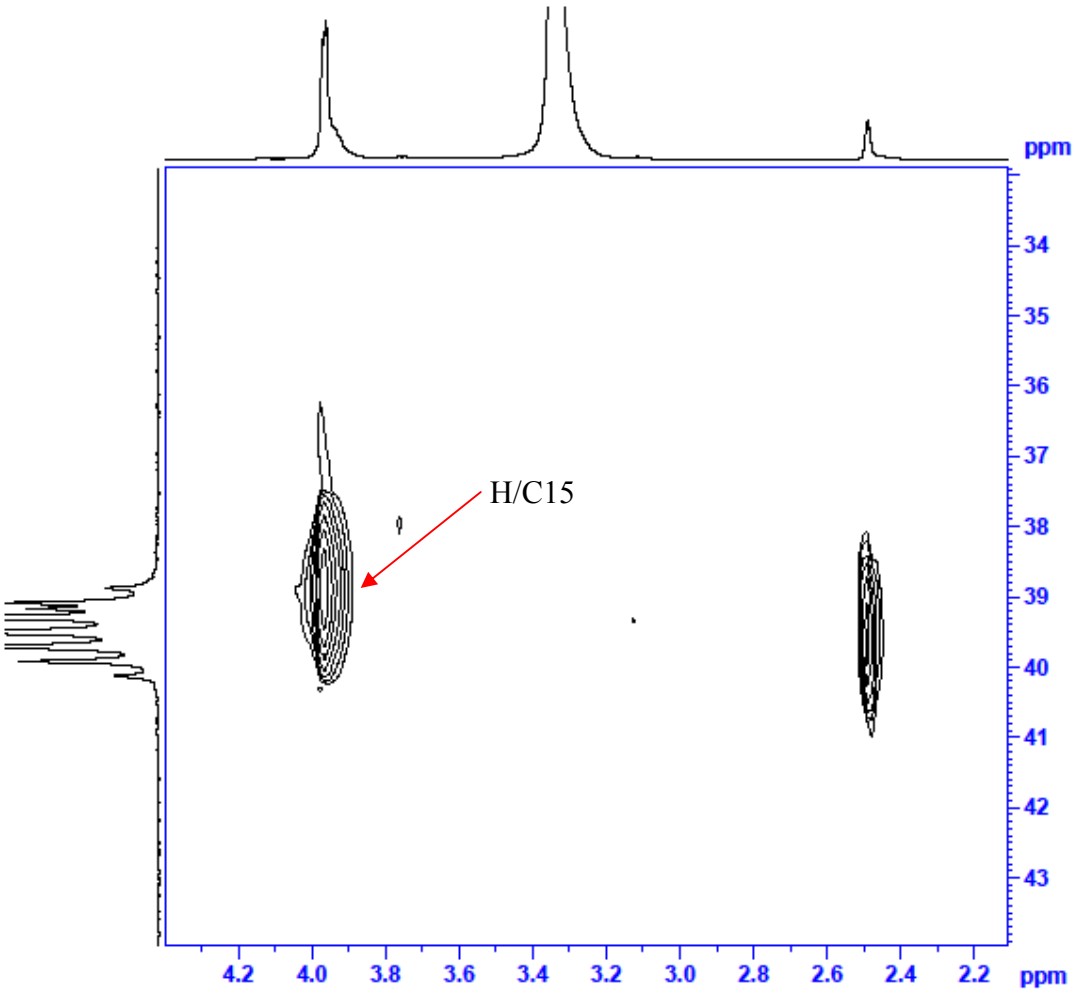



**Figure 2.37** Expanded high field section of a 400 MHz HSQC NMR spectrum of the C2-DACA dimer in d6-DMSO showing the H15/C15 crosspeak.

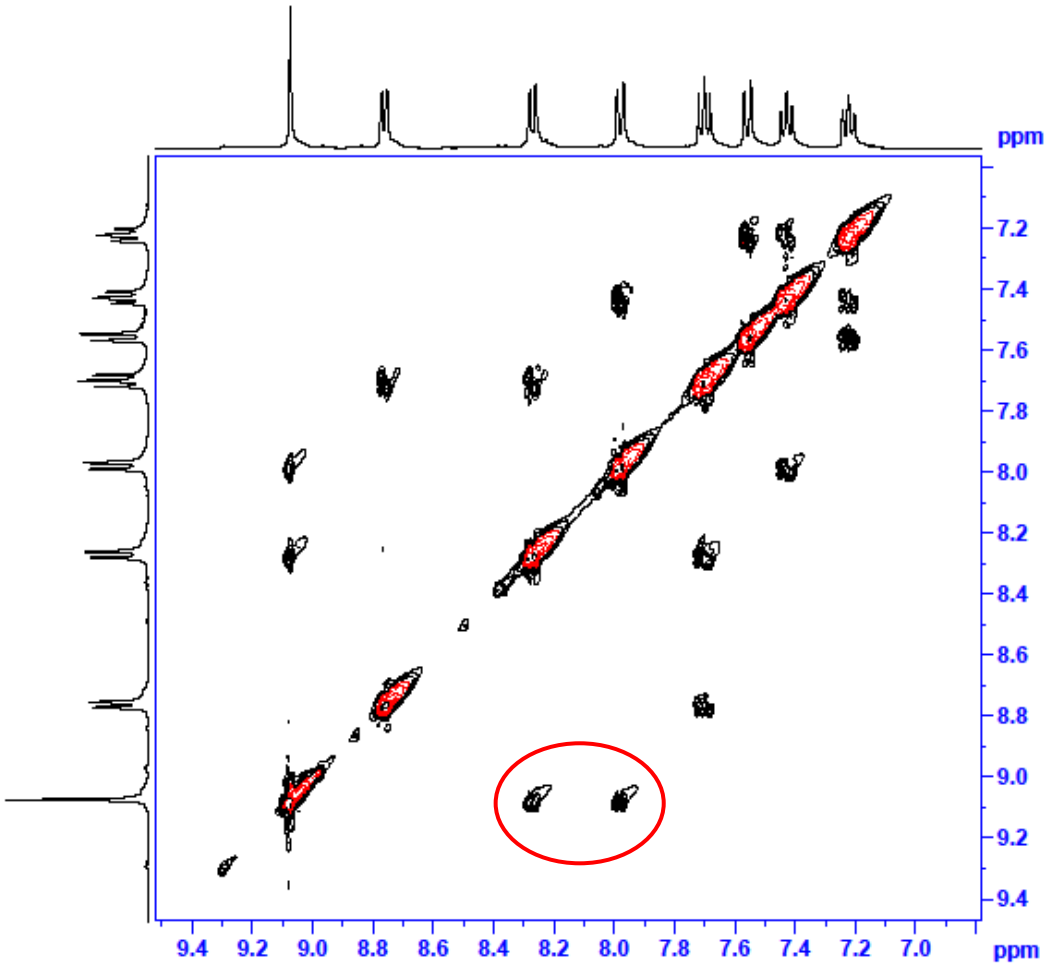



**Figure 2.38** Expanded aromatic section of a 400 MHz NOESY NMR spectrum of the C2-DACA dimer in d6-DMSO showing the NOE peaks between the H9 proton at 9.07 ppm and the H1 proton at 8.27 ppm and to the H8 proton at 7.98 ppm, indicated by the red oval.

In much the same way as for C2-DACA, the NOESY spectrum in **Figure 2.38** showed the cross peak of the H9 "singlet" at 9.07 ppm to the H1 "doublet" at 8.27 ppm and to the H8 "doublet" at 7.98 ppm as a result of their proximity (indicated by the red oval). This information was used as a starting point for assigning the protons of the acridine ring from the magnitude COSY spectrum in **Figure 2.39**. The coupled protons of the acridine ring are in a group of three protons (highlighted in red) and a group of four protons (highlighted in green) which together with the coupling allowed the complete assignment. Once the group of four and three have been identified, the "doublet" at 8.27 ppm must be the H1 proton and the "doublet" at 7.98 ppm must be the H8 proton. This allowed unequivocal identification of all other acridine protons.

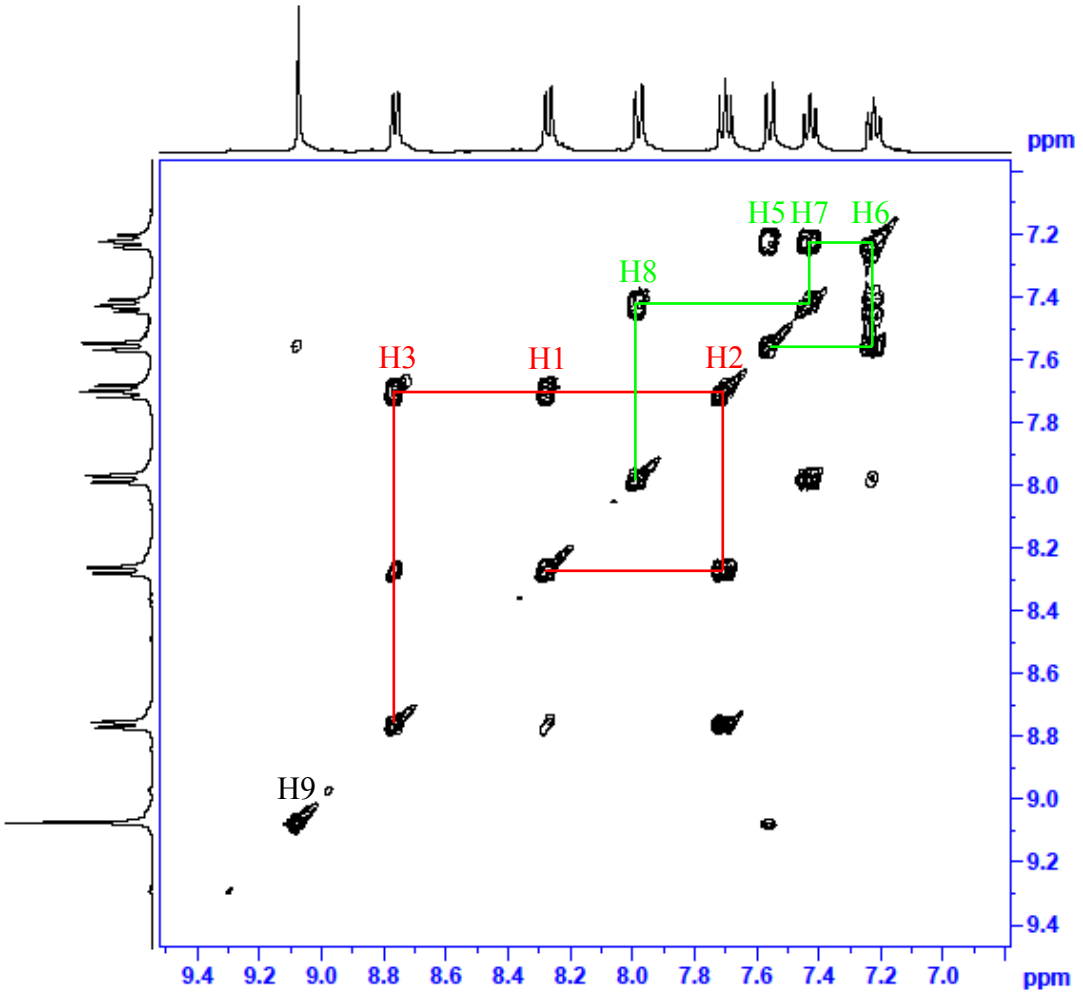



**Figure 2.39** Expanded aromatic section of a 400 MHz magnitude COSY NMR spectrum of the C2-DACA dimer in d6-DMSO. This shows the groups of three protons (H1 to H3) and four protons (H5 to H8) and defines the H6 and H7 protons via the crosspeaks from the H5 to the H6 protons, and also the H8 to H7 protons.

### 2.5.8. NMR Assignments of C2-DACA Dimer

The dimer was produced as an unexpected product of C2-DACA synthesis and thus spectroscopic methods were used to unequivocally identify and assign this structure. Infrared spectroscopy on the drug was inconclusive towards defining the absolute structure of the molecule. A 400 MHz $^{1}H$ NMR spectrum showed only one high field peak. In an attempt to separate possible overlapping peaks by increasing the temperature, increasing to 390 K did not resolve the peak at 3.94 ppm, indicating the dimer structure was present. A DEPT135 (not shown) and HSQC (**Figure 2.37**) confirmed the presence of only one $CH_2$ carbon peak below 120 ppm (39.2 ppm). These peaks are identified in **Table 2.4**, where the numbering system from **Figure 2.31** is employed.

**Table 2.4** $^{1}H$ and $^{13}C$ chemical-shift values in ppm of C2-DACA dimer in d6-DMSO at 300 K and 400 MHz. Carbon peaks assigned using ChemDraw and compared to literature compounds are indicated by an asterisk (*).

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **H-1** | 8.27 | 7.70 | 8.77 | N/A | 7.56 | 7.22 | 7.43 | 7.98 | 9.07 | N/A | N/A | N/A | N/A | N/A | 3.97 |
| **C-13** | 133.0 | 125.2 | 134.8 | 127.9* | 127.8 | 130.9 | 126.1 | 128.1 | 138.3 | 126.3* | 146.5* | 125.3* | 145.4* | 165.2* | 39.2 |

### 2.5.9. Electrospray Ionisation (ESI) Mass Spectrum of C2-DACA Dimer

The ESI mass spectrum of C2-DACA dimer, **Figure 2.41**, shows the molecular ion at 471.2 m/z as expected which corresponds to the structure shown in **Figure 2.40**.

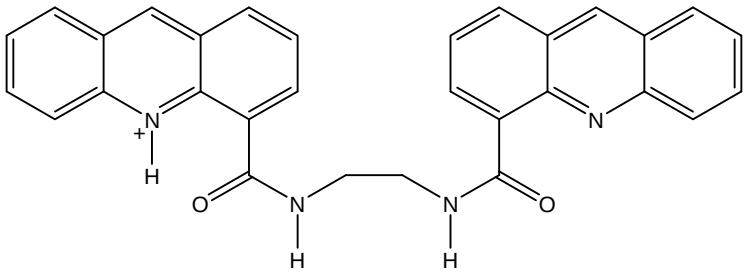



**Figure 2.40** Structure of C2-DACA dimer molecular ion (MW 471.18 g/mol).

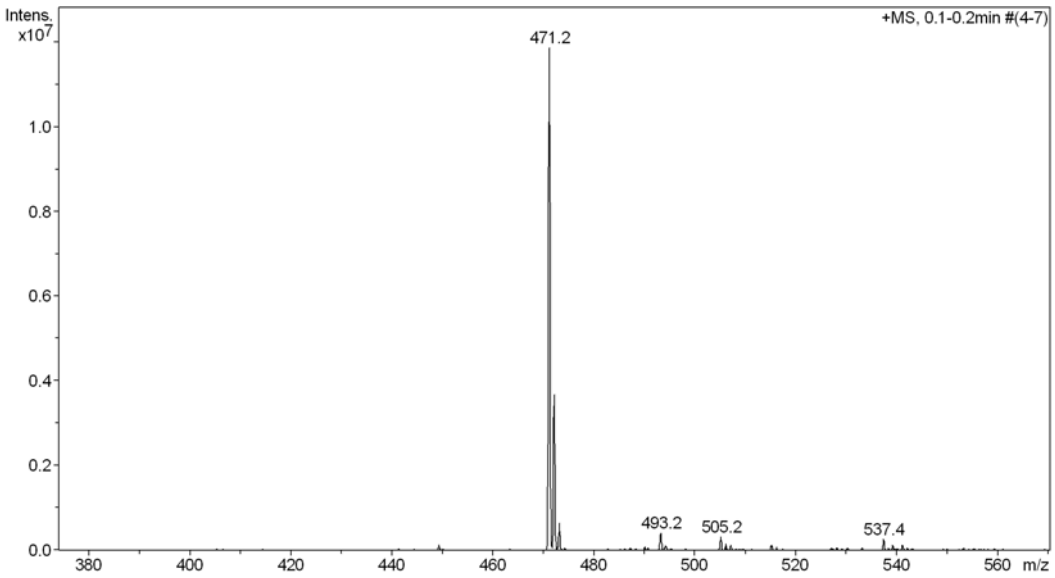



**Figure 2.41** ESI mass spectrum of C2-DACA dimer.

### 2.5.10. C4-DACA Absorbance Data

Very small quantities of the drugs were used for the preparation of samples for the DNA binding and transcription experiments described in **Chapters 4** and **5**. These concentrations were estimated from the UV/VIS absorbance data. The maximum absorbance and extinction coefficient (ε) were found spectrophotometrically. Assume 1:1 ratio of DACA:oxalate.

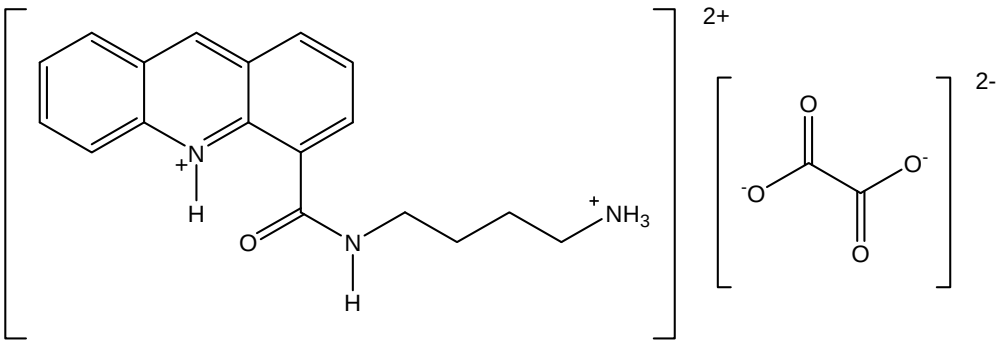



**Figure 2.42** C4-DACA.oxalate salt (MW = 383.403).

The absorbance of a solution of approximately 0.1 mM C4-DACA was run from 549.9 nm to 300.0 nm. The maximum absorbance occurred at 356.0 nm. The

absorbance of various standard concentrations, shown in **Figure 2.43**, revealed the extinction coefficient of C4-DACA at 356.0 nm.

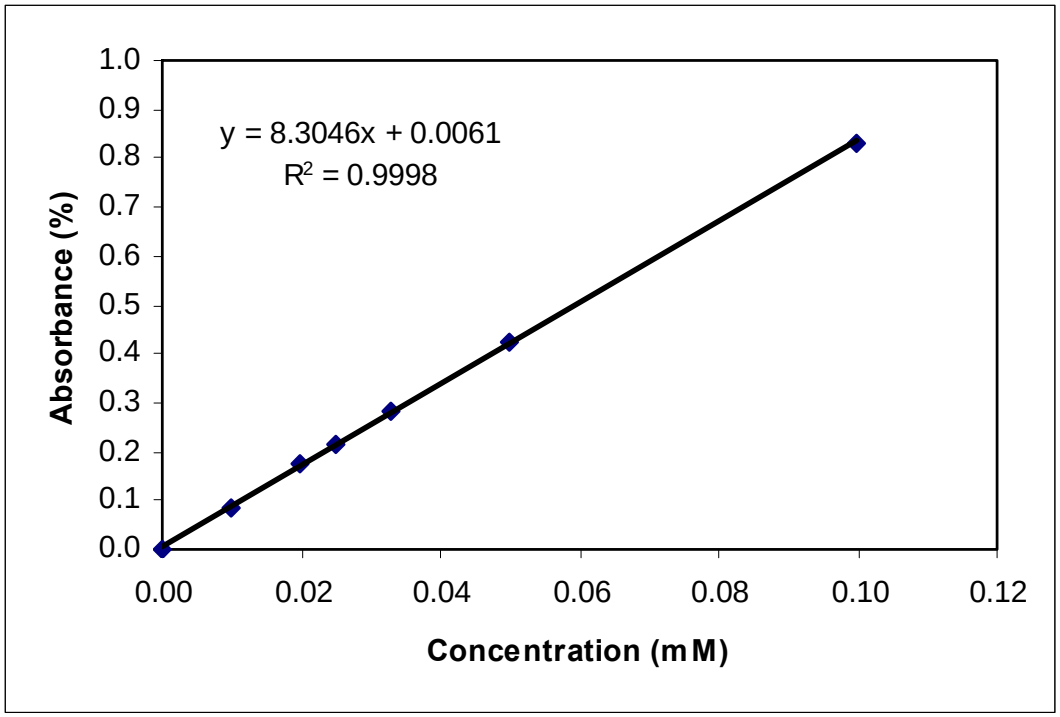



**Figure 2.43** Absorbance vs. concentration of C4-DACA in $H_2O$ at 356.0 nm.

A standard absorbance versus concentration graph revealed an extinction coefficient of ε = 8305 $M^{-1}cm^{-1}$. This ε value was used to calculate the concentration of all C4-DACA solutions.

## 2.6. Conclusions

The synthesis of *N*-(4-aminobutyl)acridine-4-carboxamide.oxalate salt afforded a pale yellow solid which was identified by NMR spectroscopic methods and ESI mass spectrometry to be C4-DACA. The protons of C4-DACA were assigned using magnitude COSY and NOESY spectra, whereas the carbons were assigned using HSQC data and carbon shift prediction software. The extinction coefficient was found to be 8305 $M^{-1}cm^{-1}$ which was used to calculate the concentration of all C4-DACA solutions. Hereafter, any reference to C4-DACA is to the pale yellow solid reported in this chapter.

Initially *N*-(2-aminoethyl)acridine-4-carboxamide, C2-DACA, and it's dimer were synthesised for experimentation using the same synthetic method as described for C4-DACA. The only alteration to the synthetic method was that the 1,4-diaminobutane was replaced with 1,2-diaminoethane. Their complete assignment was achieved by use of HSQC, NOESY and magnitude COSY NMR spectra, carbon shift prediction software, and was confirmed by ESI mass spectrometry. However, binding studies reported in **Chapter 3** and molecular modelling in **Chapter 8** led us to re-evaluate the compounds to be studied. This is discussed in more detail in **Chapter 3**.

# CHAPTER 3. DNA BINDING STUDIES OF C2-DACA AND THE C2-DACA DIMER

## 3.1. Introduction

This chapter reports on the DNA binding of the first two drugs used in this study. Two drugs were available because slight variations in the synthetic protocol for the synthesis of *N*-(2-aminoethyl)acridine-4-carboxamide, C2-DACA, enabled the production of both the monomer and dimer. Even though the dimer does not have a primary amine group, there was a possibility that it might act as a bis-intercalator, thus binding studies were undertaken on this compound. These compounds were both analysed using methods described in this chapter. Experimentation with these drugs showed that the structure of C2-DACA needed to be altered in order to obtain formaldehyde-mediated drug-DNA adducts. This led to molecular modelling (**Chapter 8**) to find a more suitable compound that was capable of forming drug-DNA adducts, followed by synthesis of the new drug C4-DACA (presented earlier in **Chapter 2**), and biochemical experimentation (**Chapters 4** and **5**) to elucidate if and where the drug-DNA adducts were formed.

Binding of the drugs to DNA was studied using a crosslinking assay. A fully detailed description of this method is presented in **Chapter 4**. Briefly, the *in vitro* crosslinking assay involved preparing radiolabelled DNA and measuring DNA migration properties by electrophoresis in the presence or absence of drug and formaldehyde. Binding to DNA is indicated by stabilisation of double stranded DNA during denaturisation which could be observed in the assay. This method was based on that employed by Hartley *et al*. [95], and it has been employed to show that Adriamycin and mitoxantrone form crosslinks with DNA in the presence of formaldehyde [59, 96].

It is possible to get some insight into how a drug is interacting with DNA by use of absorbance changes. Finding the maximum absorbance of the desired drug in solution allows us to see quantitatively if the drug remains associated with DNA after a phenol/chloroform extraction. If the absorbance after extraction is less than that of the initial sample, it indicates that some of the drug has been removed. This change in absorbance was employed to indicate whether each drug

covalently bound to DNA. The phenol/chloroform extractions ensured the removal of unincorporated drug and intercalated drug from DNA after drug reaction. This is an essential feature exploited in the crosslinking assay, as it is possible for the DNA to be held together by intercalated drugs under mild denaturing conditions.

## 3.2. Results and Discussion

### 3.2.1. Studies of the C2-DACA Dimer

The following results were obtained without repetition. The first compound to be studied was the dimer of C2-DACA (MW = 470.17 g/mol, **Figure 3.1**). This compound is a possible potential bis-intercalator, and thus the interaction of this compound with DNA was therefore of interest.

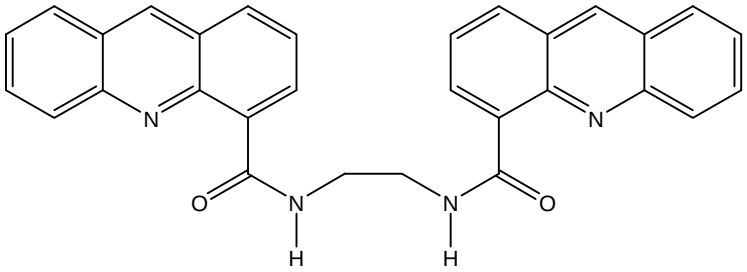



**Figure 3.1** Structure of the C2-DACA dimer.

### 3.2.2. *In Vitro* Crosslinking Assay of the C2-DACA Dimer

A crosslinking assay was performed on the drug, with a similar procedure to that described above with experimental detail in **Chapter 4**. In this case Adriamycin was used as a positive control with concentrations of 1 μM and 10 μM (**Figure 3.2**, lanes 3 and 4). The double stranded (ds) and single stranded (ss) controls are in lanes 1 and 2 respectively. Lanes 5-10 were samples for another student in the laboratory and hence not relevant to this thesis. The concentrations of the C2-DACA dimer were 1, 10, 50 and 100 μM (lanes 11-14, respectively), with the formaldehyde concentration set at 2 mM for all lanes.

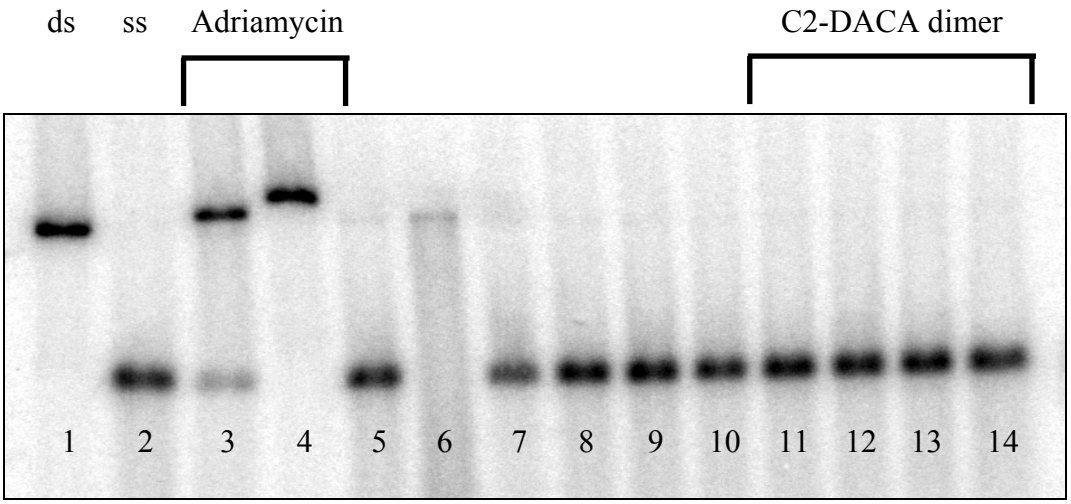



**Figure 3.2** Crosslinking assay of C2-DACA dimer with Adriamycin-DNA adduct as a control. Adduct formation following overnight incubation at 37°C. Samples containing the C2-DACA dimer are in lanes 11-14.

It can be seen in **Figure 3.2** that Adriamycin stabilises the ds DNA, however there was no evidence of ds DNA and hence crosslinking for the C2-DACA dimer. The lack of interaction with DNA is consistent with the expectation that this compound was not designed as a bis-intercalator and that the free amine is required for activation by formaldehyde and consequent drug-DNA adduct formation [66].

### 3.2.3. DNA-Binding Studies of C2-DACA

Analysis of the routine spectra was sufficient to confirm the identity of the C2-DACA oxalate salt (MW = 355.12 g/mol, **Figure 3.5**). C2-DACA was synthesised using conditions to minimise the formation of the dimer as described previously in **Chapter 2**.

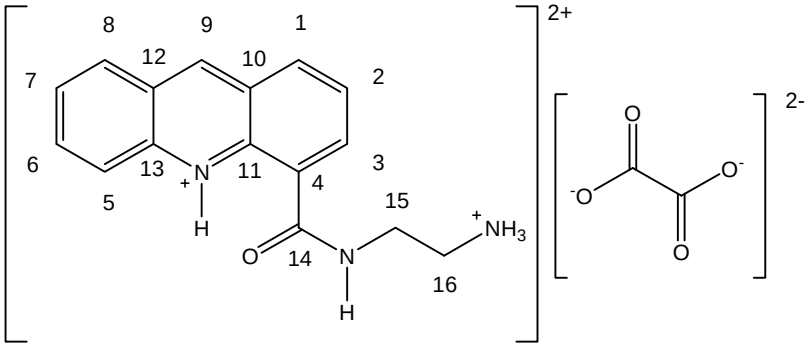


**Figure 3.5** Structure of C2-DACA.oxalate.

### 3.2.4. *In Vitro* Crosslinking Assays of C2-DACA

One way to discover if a drug forms covalent adducts with DNA is to exploit the way it interacts with DNA. A drug-DNA solution has a change in absorbance after phenol/chloroform extractions as any unbound/intercalated drug is removed from the DNA solution. The phenol/chloroform extractions ensured the removal of unincorporated drug and intercalated drug from DNA after drug reaction in the crosslinking assays.

The maximum absorbance of C2-DACA occurred at 354.9 nm, and the extinction coefficient (ε) was determined to be approximately 9300 $M^{-1}cm^{-1}$. Only one extraction was required to totally remove C2-DACA from DNA in solution, which implied the drug had only a weak, reversible intercalation affinity. This also confirmed that C2-DACA does not form covalent bonds with DNA in the presence of up to 100 μM C2-DACA and 20 mM formaldehyde.

### 3.2.5. Varied Concentration of C2-DACA

The ability of C2-DACA to form crosslinks with DNA was determined by performing a crosslinking assay with a similar procedure to that described in **Chapter 4**. **Figure 3.6** shows the ability of C2-DACA to form drug-DNA adducts in the presence of formaldehyde in a crosslinking assay. The concentration of C2-DACA was varied from 1 to 100 μM and the formaldehyde concentration was set at 2 mM for all samples. All samples contained 50 $\mu M_{bp}$ DNA and 1 x Tc buffer. The ds and ss controls contained no C2-DACA or formaldehyde. Samples were incubated for 2 h at 37°C.

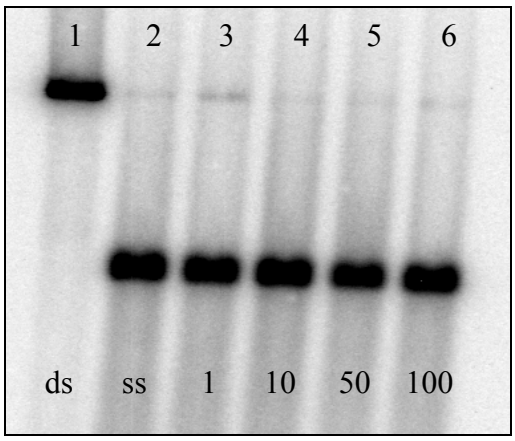



**Figure 3.6** Varied concentration of C2-DACA. Lanes 3 to 6 contain C2-DACA at concentrations from 1 to 100 μM. Samples were incubated for 2 h at 37°C.

The first lane contained no drug, was not denatured and served as the ds control. The second lane also lacked drug, but was denatured, and represented the ss control. The other four lanes contained C2-DACA at concentrations of 1, 10, 50, and 100 μM. Although there was a minor band for ds DNA for all of the samples, there was no real evidence of crosslinks when using C2-DACA as the crosslinking drug. The lack of crosslinking was thought to be due to insufficient formaldehyde concentration, or to insufficient reaction time. Thus, a crosslinking assay was undertaken with varied concentrations of formaldehyde and reaction times (see **Figure 3.7**).

### 3.2.6. Dependence on Formaldehyde for C2-DACA-DNA Adduct Formation

**Figure 3.7** shows the ability of C2-DACA to form drug-DNA adducts in the presence of formaldehyde with two different reaction times. The concentration of formaldehyde was varied from 2 to 20 mM. The C2-DACA concentration was set at 100 μM to ensure an excess of drug was available for adduct formation. All samples contained 100 $\mu M_{bp}$ DNA and 1 x Tc buffer. The ds and ss controls contained no formaldehyde or C2-DACA. Samples were incubated for 2 or 8 h.

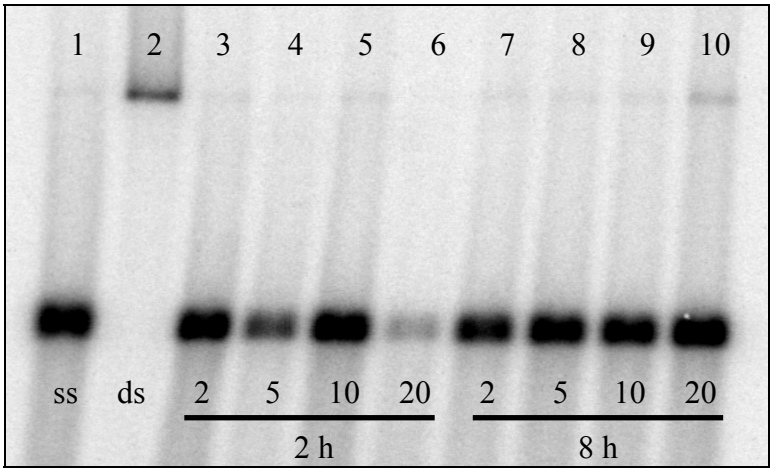



**Figure 3.7** Varied concentration of formaldehyde. Lanes 3 to 6 were incubated for 2 h, and lanes 7 to 10 were incubated for 8 h. Lanes 3 to 10 contained formaldehyde at concentrations from 2 to 20 mM and C2-DACA at 100 μM.

The ss and ds lanes represented the ss and ds control lanes (as for **Figure 3.6**). The first well contained no drug, but was denatured, to act as the ss control. The second, also lacking drug, was not denatured, to act as the ds control. The other eight wells contained drug at 100 μM, with formaldehyde at varying concentrations of 2, 5, 10, and 20 mM as indicated in **Figure 3.7**. Lanes 3 to 6 were incubated at 37°C for 2 h, and 7 to 10 incubated at 37°C for 8 h. Although there was a slight band at the double stranded level, it was present throughout the samples, indicating background levels of ds DNA. As seen in lane 4 of **Figure 3.2**, Adriamycin shows almost 100% ds DNA at 10 μM and 2 mM formaldehyde. Even though Lane 10 from **Figure 3.7** appears to have a slightly increased level of ds DNA the result could not be replicated. The samples for 5 and 20 mM formaldehyde incubated for 2 h were not as dark as the rest of the samples as some of the DNA pellet was lost during washing. This assay therefore failed to provide evidence of crosslinks when using C2-DACA as the crosslinking drug.

## 3.3. Discussion

The initial drug used was proven by various methods in **Chapter 2** to be the dimer of C2-DACA. The synthesis was altered to produce C2-DACA, which was unequivocally confirmed in **Chapter 2** by NMR analysis and ESI mass spectrometry.

The crosslinking assay of C2-DACA showed little evidence of drug-DNA adducts. It was expected that the acridine would intercalate as seen in the literature, but this alone was not sufficient to stabilise the DNA double helix. These results led to the belief that C2-DACA was binding to the DNA, but not stabilising the double helix to any significant extent. Thus, when heated and denatured, the drug did not have sufficient stability to maintain the two strands of DNA together. To avoid this, the DNA was not denatured in the binding studies, and was extracted with phenol and its absorbance taken. It was believed that this would yield a change in absorbance indicating that a percentage of unbound drug was removed. The binding study showed that the drug only has a weak intercalation with DNA as the majority of the drug was removed following a single phenol extraction.

It was proposed that there were four possible reasons for the drugs limited interaction with DNA. One was that the acridine section was intercalating, but the side chain was not being activated by formaldehyde. This may imply that the amount of reacted drug would increase in proportion to an increase in formaldehyde. This was tested experimentally and was shown not to account for the lack of adducts detected.

The second proposal was that the drug did not have access to the reaction site, as it was intercalating from the major groove (instead of the minor groove). However, the likelihood is that the drug is able to intercalate to some degree from both sides [97, 98].

The third hypothesis was that the Schiff base was not being formed due to the reaction taking place at a non-optimal pH. Evidence from more complex drugs such as Adriamycin, mitoxantrone and pixantrone shows that drugs containing primary amines can undergo Schiff base formation in the presence of formaldehyde [4, 59, 61]. This was investigated further using amines similar to C2-DACA by increasing the basicity of the solution, then adding the strongly

acidic formaldehyde solution. However this line of investigation was inconclusive as to whether the Schiff base was actually forming or not due to the complexity of the resulting mixture.

The fourth possibility was that the side chain of the drug was not long enough to enable the intercalated drug to form a methylene bridge with the DNA. C2-DACA may well be able to intercalate in a manner similar to Adriamycin, but adduct formation would not proceed if the formaldehyde activated side chain was not long enough to reach a free amine of guanine. This possibility prompted further investigation into the three dimensional shape of C2-DACA-DNA adducts, resulting in the molecular modelling shown in **Chapter 8**.

## 3.4. Conclusions

The crosslinking assay showed that the C2-DACA dimer does not induce stable crosslinks between the DNA strands, and implies that the free amine is essential for the formation of formaldehyde mediated drug-DNA adducts. The C2-DACA compound showed a weak intercalation affinity with DNA when subjected to phenol/chloroform extraction, and also showed no evidence of stable crosslinks between the DNA strands in crosslinking assays. These results proved that C2-DACA needed to be redesigned to allow the side chain to form stable formaldehyde-activated drug-DNA adducts and consequently for the acridine section to be able to intercalate simultaneously.

# CHAPTER 4. C4-DACA *IN VITRO* CROSSLINKING ASSAY

## 4.1. Introduction

Modelling studies described in **Chapter 8** found that the optimum length for the DACA side chain to bind to DNA was 4 carbon atoms long. The structure with the optimal length side chain, C4-DACA, is shown in **Figure 4.1**.

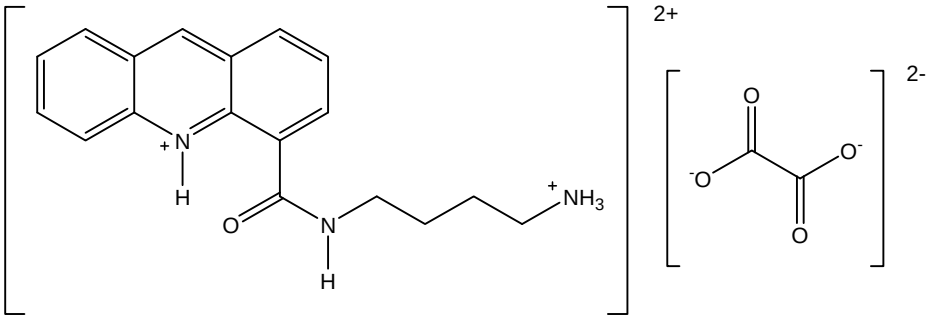



**Figure 4.1** Structure of C4-DACA.

This chapter describes the methods used to study the binding of the drug C4-DACA to DNA to form drug-DNA adducts. The aim was to see if the drug would bind covalently to DNA and stabilise ds DNA. A crosslinking assay is a routine biochemical method for determining whether the chromophore intercalates and binds covalently. This method has been used in our laboratory for drugs such as Adriamycin, daunomycin, mitoxantrone and pixantrone. The crosslinking assay gives a positive result if the drug in question stabilises ds DNA (such that it does not denature when heated) and is therefore a marker of the extent of covalent DNA binding. There are many possible conditions that need to be optimised for this method, including concentration of drug and formaldehyde, time allowed for activation (pre-reaction), temperature dependence, pH, and phenol/chloroform extraction of non-covalently bound drug.

### 4.1.1. Description of the Crosslinking Assay Method

The method used to study the crosslinking ability of C4-DACA was based on that employed by Hartley *et al.* [95], and is used routinely in our laboratory.

The circular bacterial pCC1 plasmid (approximately 3500 base pairs) was isolated from *E. coli* and purified by conventional procedures (see **Appendix 1** for transcription fragment and vector sequence). This was then linearised with the restriction endonuclease *Hind*III which leaves a 5' overhang called a 'sticky end'. The linearised plasmid was purified by phenol/chloroform extractions (equal volume of aqueous/organic wash), then precipitated with ethanol, sodium acetate and glycogen to form a DNA pellet. The DNA was then end-labelled at the 'sticky ends' with [α-$^{32}$P]dATP using the Klenow fragment of *E. coli* DNA polymerase I. This mixture was cleaned up by micro-column and resuspended in calf thymus DNA and TE buffer (Tris-EDTA buffer protects DNA from degradation). The end label was incorporated purely to enable detection of the ds and ss DNA using autoradiography. Phosphor-imaging allows the $^{32}$P end labels to identify whether the DNA is ss or ds DNA following separation by electrophoresis.

The drug C4-DACA was dissolved in Milli-Q water to an approximate concentration of 2 mM. The end labelled DNA was reacted with C4-DACA for a given period of time (typically overnight) in transcription buffer in the presence or absence of formaldehyde at a given temperature to enable drug-DNA adduct formation. Samples were placed on ice to reduce any further reaction, and any unreacted drug was extracted using phenol/chloroform, followed by ethanol precipitation. Phenol/chloroform extractions ensure the removal of unincorporated drug and intercalated drug from DNA. Samples were resuspended in TE buffer and denatured at 65°C for a given time, quenched on ice, mixed with loading dye and loaded onto a 0.8 % agarose gel. The denaturation step is designed to denature the DNA however drug-DNA adducts hold the DNA together at the adduct sites. This enables the DNA to 'zipper' back into the double stranded helix in the subsequent renaturation step. Any DNA which was not stabilised will remain as ss DNA. Thus, ds DNA indicates that drug-DNA adducts are present. The DNA in the agarose gel was then subjected to electrophoretic separation overnight, then dried and exposed to a phosphorscreen. The time of exposure was 4-16 h and was dependent on the age of the radiolabel. The exposed plate was subject to autoradiography and the radioactive bands quantitated by ImageQuant [99].

### 4.1.2. Why Choose the Crosslinking Assay?

The crosslinking assay method was chosen to mimic results involving Adriamycin-DNA adducts reported by Cullinane *et al.* [7]. Their results showed irreversible transcription blockages detected *in vitro* which displayed the characteristics of covalent adducts. These adducts were formed in the presence of formaldehyde. Adriamycin, in the presence of formaldehyde, forms covalent adducts with DNA almost exclusively at GpC sequences [2, 7, 66, 77]. It has also been shown that adduct formation has an absolute requirement for the exocyclic amine of guanine [2, 66]. High resolution X-ray crystallography (1.5 Å) revealed that the anthracycline chromophore intercalates between the DNA base pairs, with the amino sugar in the minor groove [79]. One of the common features of the aforementioned anthracycline based drugs is that once intercalated in the DNA there are many side groups on the flat aromatic structure which are attracted in varying degrees to the base pairs. This key feature would assist with their ability to stabilise the double helix, give longer adduct life-times, and hence a lower concentration of drug would be required to yield a positive result in a crosslinking assay. As described in **Chapter 1**, Adriamycin stabilises ds DNA by forming a covalent bond between the N3'of Adriamycin (primary amine of the sugar) and the N2 of guanine [66, 79]. C4-DACA is a simple derivative modelled on this drug, and as it does not contain any side groups (other than the side arm) it is expected that it would have a shorter life-time as an adduct.

It is obvious that C4-DACA is a simplified version of Adriamycin, however retaining the ability of intercalation and binding to DNA. The binding affinity will not be the same as that of Adriamycin so it proved necessary to vary experimental conditions of the experimental crosslinking assay considerably in order to detect C4-DACA-DNA adducts.

### 4.1.3. Cell Death

DNA replication is obstructed by both interstrand and intrastrand crosslinks. Crosslinks within DNA block both DNA replication and transcription and ultimately result in cell death. This feature resulting from crosslinks in DNA can be exploited when treating cancer. Cancer cells generally exhibit uncontrolled cell growth, so creating crosslinks within the cancer cells should target them for

cell death. This method of treatment is known as chemotherapy. The drugs used in chemotherapy are referred to as cytotoxic drugs and target any rapidly growing cells. However, some level of specificity occurs as, unlike normal cells, some cancer cells are unable to repair DNA damage.

## 4.2. Materials and Methods

### 4.2.1. Materials

The acridine C4-DACA was as described in **Chapter 2**. The Plasmid Isolation Kit was obtained from Qiagen while restriction enzymes *Hind*III and *Pvu*II were from Promega. The DNA Polymerase I, Large (Klenow) Fragment was from New England Biolabs. The [$\alpha^{32}$P] dATP (3000 Ci/mmol) was from Perkin Elmer and Probe-Quant G-50 micro columns were purchased from GE Healthcare. Molecular biology grade urea was from MP Biomedicals and acrylamide:bisacrylamide solution was purchased from Amresco. The plasmid pCC1 was constructed by directional ligation of the 479 bp *Sal*I/*Pvu*II restriction fragment of pRW1 [100] into pSP64 linearised with the same two enzymes using standard techniques [101]. The calf thymus DNA was purchased from Sigma-Aldrich. The remaining chemicals and reagents were of analytical grade. Distilled water passed through a four stage Milli-Q purification system was used to prepare all solutions.

### 4.2.2. Drugs

The identity of the C4-DACA supplied was confirmed using 1D and 2D NMR methods. Analysis of the routine spectra shown in **Chapter 2** was sufficient to confirm that the sample was C4-DACA.oxalate (**Figure 4.1**). A stock solution of C4-DACA (stored at 4°C) was prepared by dissolving the oxalate salt in Milli-Q water to an approximate concentration of 2 mM. An accurate concentration of the drug was determined spectrophotometrically using $\varepsilon$ = 8,305 $M^{-1}cm^{-1}$ at 356 nm as noted in **Chapter 2**. Formaldehyde solutions were prepared fresh on the day of each experiment.

### 4.2.3. Preparation of Calf Thymus DNA

Calf thymus DNA (30 mg) was dissolved in 20 ml of 1 x TAE (Tris-acetate-EDTA buffer). Two drops of chloroform were added (to inhibit bacterial growth), the flask stoppered, and the DNA allowed to dissolve overnight at 4°C on a stirring platform. The solution was sonicated for 1 min and filtered through a 0.44 μm Millipore filter. DNA was characterised by electrophoresis to be between 1650-2000 base pairs (presented in **Figure 4.2**), on a 0.8% agarose gel with 2 μL ethidium bromide, and 1 Kb Plus DNA Ladder. The DNA concentration was determined from the absorbance at 260 nm using an extinction coefficient of 13,200 $M_{bp}^{-1}cm^{-1}$. Aliquots were stored at 4°C.

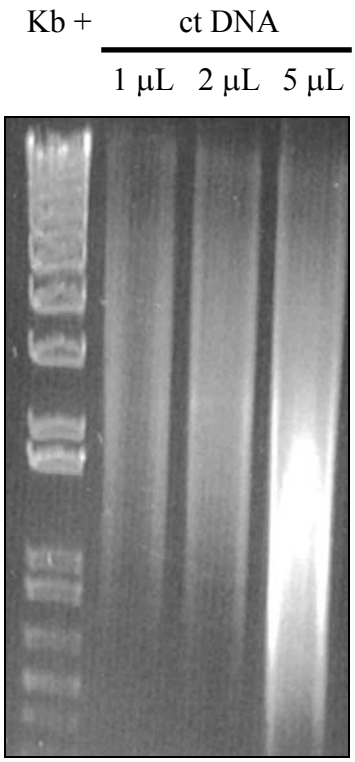



**Figure 4.2** Characterisation of calf thymus DNA. The first lane contains 1 Kb Plus DNA ladder to indicate approximate length of DNA.

### 4.2.4. Preparation of End-labelled DNA

A glycerol stock of *E. coli* XL1 Blue strain (containing the plasmid pCC1) was streaked on agar and allowed to grow overnight at 37°C. A single colony was selected and grown in LB broth over a 16 h period at 37°C. The plasmid was then

extracted from *E. coli* using a Qiagen Maxi plasmid isolation kit and linearised with *Hind*III. The linearised plasmid was then purified using a standard phenol/chloroform extraction procedure and the DNA precipitated with ethanol using 20 μg of glycogen as an inert carrier [78].

The *Hind*III linearised plasmid was end-labelled with [α-$^{32}$P]dATP in the presence of the Klenow fragment of *E. coli* DNA polymerase I. Unincorporated label and protein was removed by passing the labelled DNA solution through a Probe-Quant G-50 micro column. The labelled DNA was resuspended in calf thymus DNA to a final concentration of 400 μM base pair in TE buffer.

### 4.2.5. Efficiency of Phenol/Chloroform Extractions

Phenol/chloroform extraction was used to ensure removal of unincorporated drug, intercalated drug and proteins from DNA after drug reactions. This required that the DNA be extracted with phenol (twice) and chloroform (once). The DNA was precipitated with two volumes of ethanol, sodium acetate (with a final concentration of 0.3 M as a source of monovalent cations to aid in DNA precipitation) and glycogen. The efficiency of this method was tested using spectral analysis of drugs between the wavelengths of 340 and 450 nm. Solutions of drug, formaldehyde and drug-formaldehyde devoid of DNA were prepared for C4-DACA and spectra of these solutions were obtained. The solutions were then extracted with one volume of phenol and the resulting solutions scanned again between 340 and 450 nm.

### 4.2.6. General Method for *In Vitro* Crosslinking Assays

End labelled DNA (25 μM base pair) was reacted with C4-DACA for a given period of time in transcription buffer (40 mM Tris, 100 mM KCl, 3 mM $MgCl_2$, 0.1 mM EDTA, pH 8.0) in the presence or absence of formaldehyde at 4°C, room temperature, or 37°C. Samples were then placed on ice and unreacted drug was extracted first with phenol and then chloroform after which the DNA was precipitated in ethanol (using glycogen as an inert carrier). The samples were then resuspended in TE and denatured at 65°C for 1, 5 or 10 min in a final concentration of 66% loading dye (60% formamide, 6.6 mM EDTA, pH 8.0), and

consequently loaded onto a 0.8 % agarose gel (1 x TAE buffer). The DNA was then separated electrophoretically in 1 x TAE buffer at 45 V for 16 h (overnight). The agarose gel was then dried and subjected to autoradiography using a Molecular Dynamics Model 400B PhosphorImager, and the resultant radioactive bands quantitated by ImageQuant [99].

## 4.3. Results and Discussion of the Optimisation of Experimental Conditions

The following results were obtained without repetition with the exception of the experiments corresponding to **Figures 4.14** and **4.20**.

### 4.3.1. Optimisation of Reaction pH for C4-DACA-DNA Adduct Formation

The optimum pH required in order to maximise the amount of drug available for adduct formation with formaldehyde and DNA was investigated. If the sample solution was too acidic, the free amine on the sidearm of C4-DACA would most likely be protonated, and therefore would not be available to undergo Schiff base formation with formaldehyde. The optimum pH for Schiff base formation was found by reacting formaldehyde and drug in the presence of calf thymus DNA at different pH. It was shown in **Chapter 2** that C4-DACA had a maximum absorbance at 365 nm. A higher residual absorbance at 365 nm (following the extraction of unreacted drug) therefore indicates a greater amount of drug-DNA adducts. The pH optimisation was performed using 100 mM formaldehyde, 100 μM C4-DACA, and 100 μM calf thymus DNA, and the pH adjusted in the range of 5 – 8 using phosphate buffer. Samples were incubated overnight at 37°C, un-reacted drug was removed via one phenol extraction, and absorbance taken from 340 nm to 450 nm using a Cary 50 UV-Vis Spectrophotometer [102], and the pH profile is shown below in **Figure 4.3**.

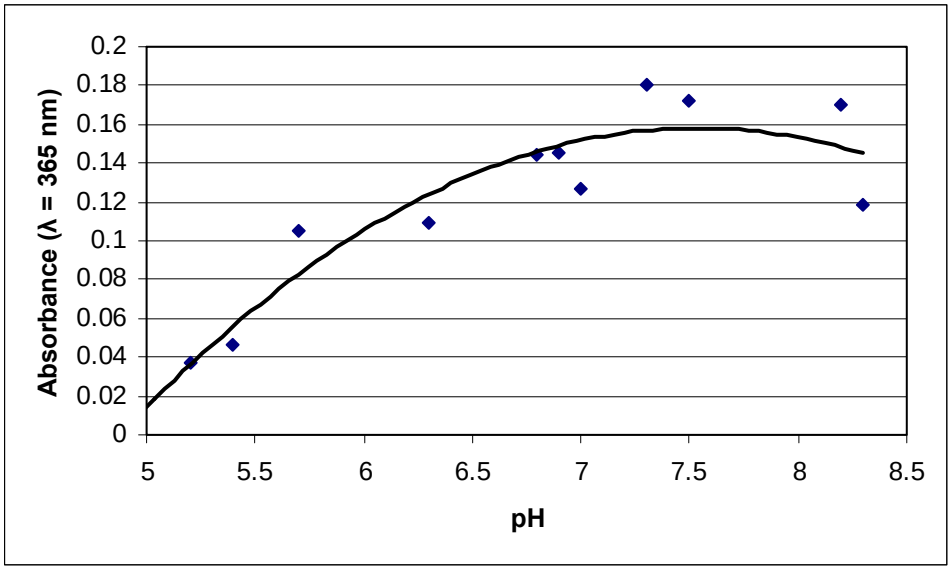



**Figure 4.3** Absorbance at various pH with 100 μM C4-DACA and 100 mM formaldehyde.

The pH of the formaldehyde stock was approximately 4 so the pH of the samples had to be checked after the addition of formaldehyde. Phosphate buffer of less than 40 mM was not enough to prevent a decrease in pH when 100 mM formaldehyde was added and 50 mM phosphate buffer was required to minimise this pH decrease. The results imply that the optimum pH for drug-DNA adduct formation was around 7 to 7.5, and therefore that the use of buffer and controlled pH was imperative for these studies.

### 4.3.2. Does Formaldehyde Alone Stabilise DNA?

The ability for formaldehyde to form crosslinks with DNA was investigated by varying the concentration of formaldehyde (**Figure 4.4**). All samples contained 50 $\mu M_{bp}$ DNA and 50 mM phosphate buffer. The ds and ss DNA controls contained no formaldehyde. The first two lanes were not denatured and served as the ds DNA marker. The second lane contained 80 mM formaldehyde. The four samples had increasing concentration of formaldehyde ranging from 10 – 80 mM. Samples were incubated at 37°C overnight and denatured at 65°C for 5 min. The percentage of ds and ss DNA was determined by ImageQuant [99] and is shown in **Figures 4.4** and **4.5** and reveals a linear relationship between the concentration of formaldehyde and the percentage of

double stranded DNA formed. This confirms that interstrand crosslinks form in direct proportion to the concentration of formaldehyde.

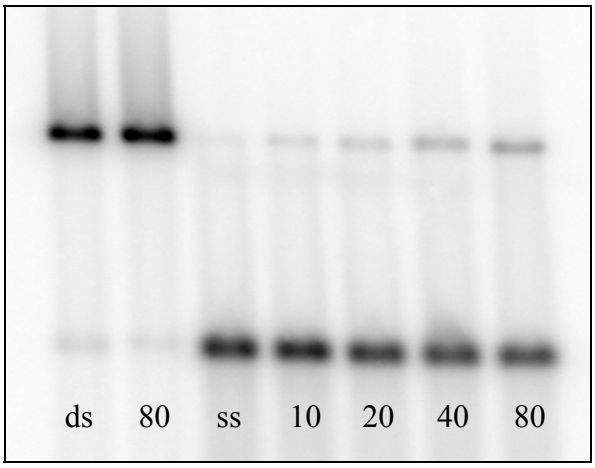



**Figure 4.4** Dependence of adduct formation on formaldehyde concentration (10 to 80 mM).

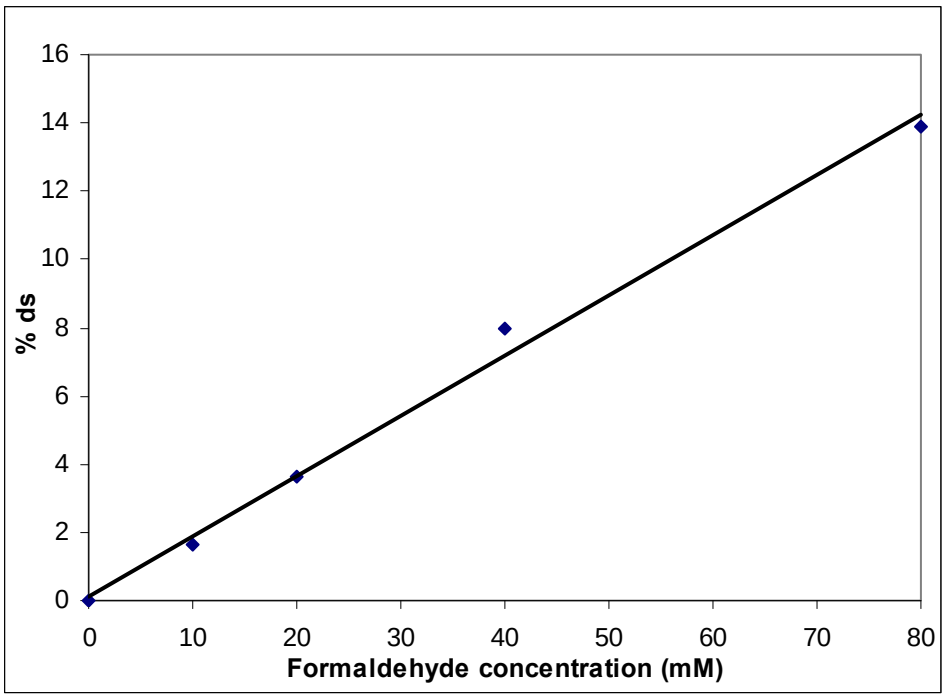



**Figure 4.5** Percentage of ds DNA vs. increasing formaldehyde concentration.

### 4.3.3. Optimisation of Denaturation Temperature (65°C to 75°C)

The minimum temperature for denaturation was investigated in order to ensure that the samples were being fully denatured. All samples contained 50 $\mu M_{bp}$ DNA and 20 μM of C4-DACA but formaldehyde was omitted. Formaldehyde was omitted to ensure that there was only intercalated drug present. The phenol/chloroform wash was also omitted for the same reason. This was to make sure that intercalated drug alone would not stabilise the double helix during denaturation, and guarantee that any ds DNA formed was only from drug-DNA adducts. Samples were denatured for 1 minute at 65°C, 70°C, and 75°C. The ds and ss controls contained no formaldehyde or C4-DACA. The percentage of ds and ss DNA was determined by ImageQuant[99] and is shown in **Figure 4.6** and **Table 4.1**.

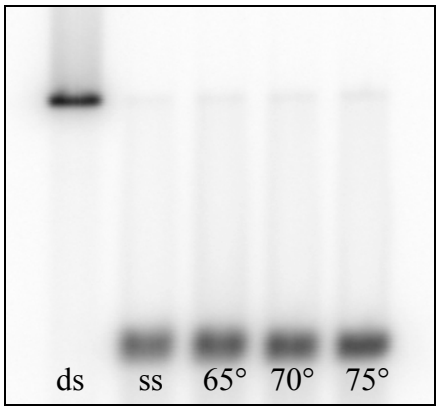



**Figure 4.6** Denaturation temperatures were 65°C, 70°C, and 75°C.

**Table 4.1** Percentage of remaining ss DNA with increasing denaturation temperature. Samples contain 20 μM C4-DACA without formaldehyde.

| Temperature | % of ss DNA |
|---|---|
| ds control | 0.4 |
| ss control | 99.1 |
| 65°C | 99.3 |
| 70°C | 99.2 |
| 75°C | 99.0 |

These results show that with a denaturation temperature of 65°C and above, there was less than 1% ds DNA remaining. This result confirms that 65°C is satisfactory to adequately denature DNA under these conditions.

### 4.3.4. Optimisation of Denaturation Temperature (50°C to 66°C)

A more thorough investigation into the lowest possible denaturation temperature was required. This experiment utilised half the concentration of DNA of the previous procedure and a low concentration of formaldehyde. All samples contained 25 $\mu M_{bp}$ DNA and 50 mM phosphate buffer. The ds and ss controls contained no formaldehyde or C4-DACA. The master mix contained 20 µM C4-DACA and 2 mM formaldehyde. Samples were denatured for 1 minute from 50°C to 66°C. The percentage of ds and ss DNA was determined by ImageQuant[99] and is shown in **Figures 4.7** and **4.8.**

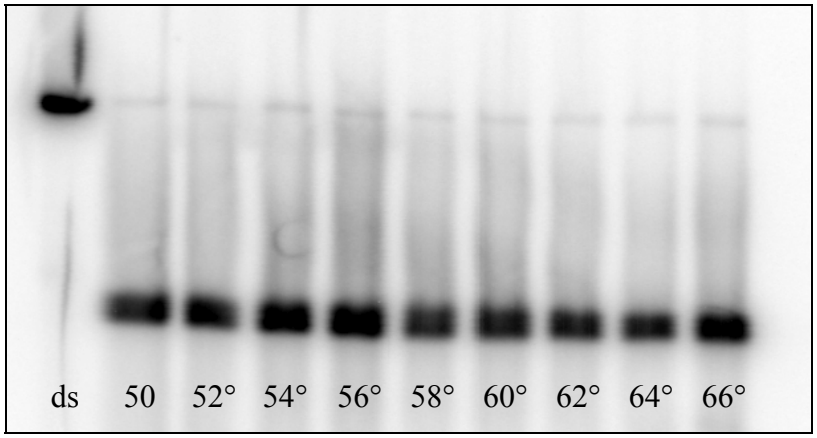



**Figure 4.7** Samples were denatured from 50°C to 66°C and contain 20 µM C4-DACA and 2 mM formaldehyde.

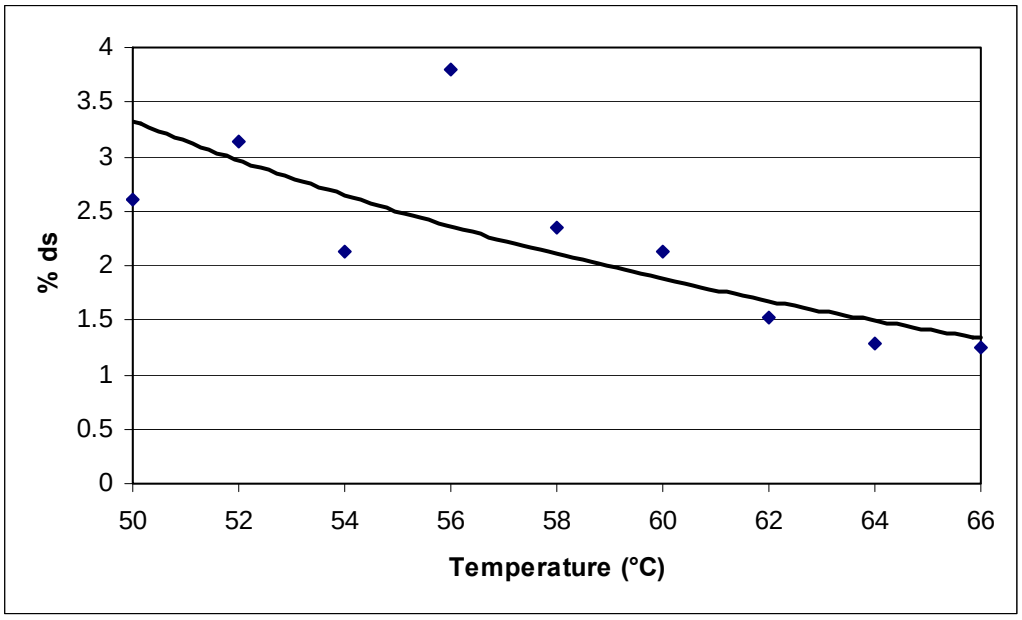



**Figure 4.8** Percentage of remaining ds DNA at increasing denaturation temperatures.

This result shows that even at 50°C at least 97% of DNA was denatured. This implies that all following experimental temperatures were sufficient to achieve full denaturation (65°C ensures that approx 99% of the DNA was denatured).

### 4.3.5. Efficiency of Phenol/Chloroform Extractions

The efficiency of phenol extractions in removing unbound drug at 4°C was investigated. This was examined by reacting the drug and DNA at 37°C over 24 h, followed by work-up in a 4°C cold room. Two separate master mixes (MM) were made up consisting of the ds/ss control (MMc) and the master mix which contained drug (MMd). The concentration of C4-DACA was 40 μM for MMd, whereas formaldehyde was omitted to ensure any drug present was only of the intercalated form. Master mixes were incubated at 37°C for 24 h, and 20 μL samples were removed from the master mixes to create a ds and ss control and four samples. Samples were worked-up in the 4°C cold room involving 0, 1, 2 or 3 phenol extractions followed by one chloroform extraction, and subsequently denatured at 65°C for one min. All samples contain 50 $\mu M_{bp}$ DNA and 50 mM phosphate buffer. The ds and ss controls contained no formaldehyde or C4-DACA. The percentage of ds and ss DNA was determined by ImageQuant[99] and is shown in **Figures 4.9** and **4.10.**

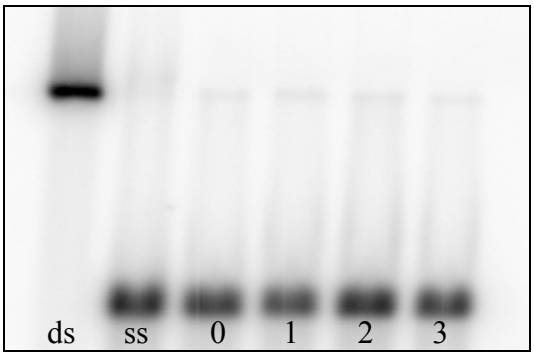



**Figure 4.9** Optimum number of phenol extractions (0-3) containing 40 μM C4-DACA without formaldehyde.

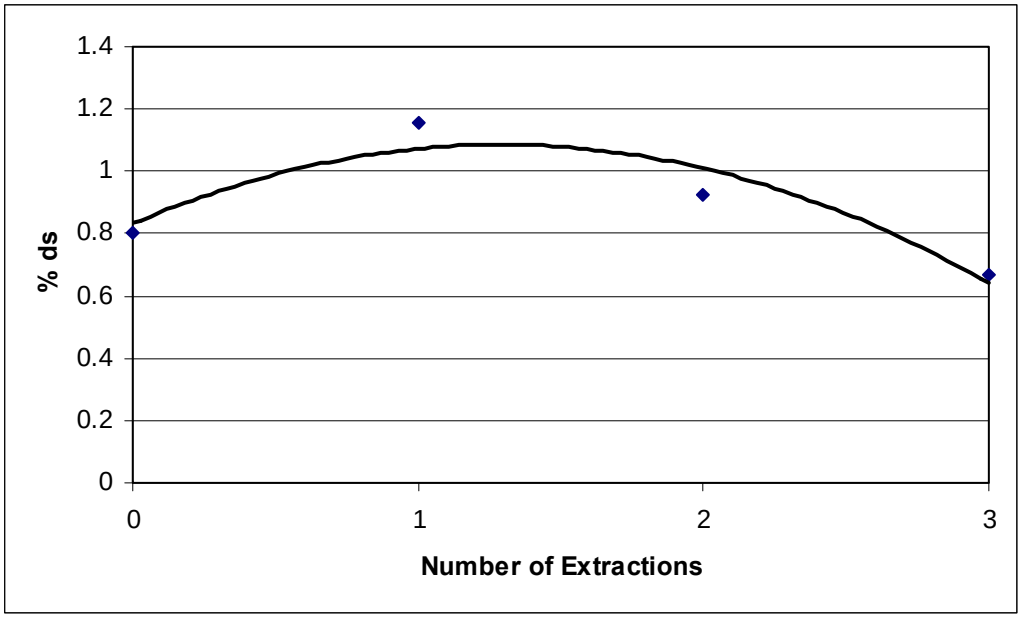



**Figure 4.10** Percentage of remaining ds DNA with increasing phenol extractions.

These results show that one chloroform extraction was required to remove 99% of intercalated drug. These results also demonstrate the absolute requirement for formaldehyde in order for C4-DACA to form covalent drug-DNA adducts.

### 4.3.6. Is a Phenol/Chloroform Wash Necessary?

It was possible that the phenol/chloroform wash may have removed the bound drug as well as the intercalated drug from the DNA. Thus, a crosslinking assay was performed without the subsequent phenol/chloroform clean-up (**Figure 4.11**). The possibility of slight changes in concentration between reaction samples was excluded by the use of master mixes. Three separate master mixes (MM) were employed. A ds/ss control, a master mix containing drug (MMd) and a master mix control without drug (MMc). The concentration of C4-DACA was again kept at 20 μM for MMd (indicated by an asterisk) and formaldehyde at 80 mM for MMd and MMc. Master mixes were incubated at 37°C and 20 μL of both MMd and MMc were removed at each time interval and frozen. The time intervals were 0, 3, 7, 16, 20 and 24 h. Once all time points had been frozen the samples were thawed, 20 μL of formamide loading buffer added and subsequently denatured at 65°C for 1 min. All samples contained 50 $\mu M_{bp}$ DNA and 50 mM

phosphate buffer. The ds and ss controls contained no formaldehyde or C4-DACA and were frozen at the 24 h time period.

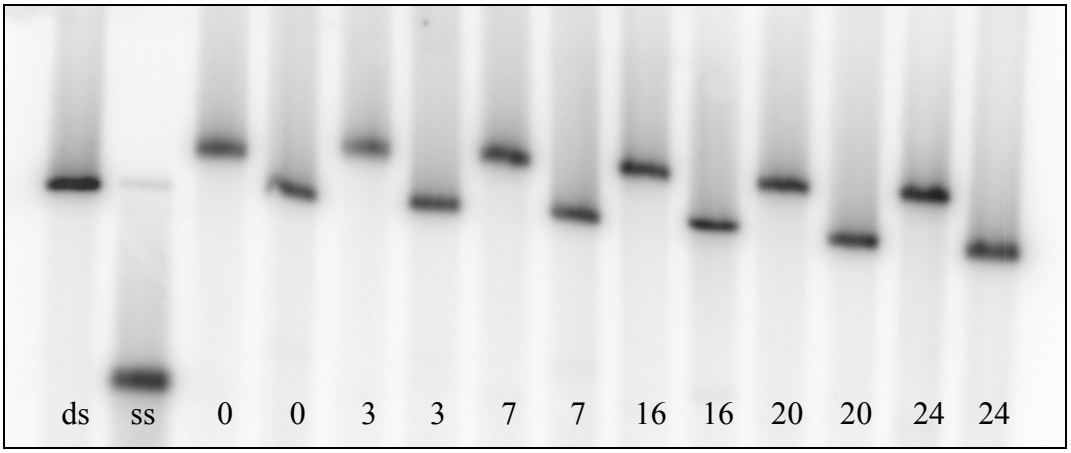



**Figure 4.11** The effect of omitting phenol/chloroform washes. An asterisk indicates the presence of 20 μM C4-DACA.

There was a significant electromobility shift between the samples containing drug and those without. This indicates that the drug was interacting with the DNA but does not reveal if it was forming covalent links or merely intercalating into the DNA. It is clear from **Figure 4.11** that a phenol/chloroform wash was required for the crosslinking assay between formaldehyde, C4-DACA and DNA.

### 4.3.7. Varying Concentration of C4-DACA (20-100 μM)

**Figure 4.12** shows the result of a crosslinking assay with C4-DACA as the ligand. All samples contained 50 $\mu M_{bp}$ DNA, 50 mM phosphate buffer and 75 mM formaldehyde. Samples were incubated at 37°C overnight and denatured at 65°C for 5 min. The ds and ss controls contained no drug. The five samples had increasing concentration of C4-DACA ranging from 20 μM to 100 μM. There were significant $^{32}P$ DNA counts still in the phenol layer during the phenol wash on the 80 and 100 μM samples. This implied that the amount of drug attached may have been causing the DNA to partition into the phenol layer by making the DNA less polar, and suggests that this could be an upper limit for drug concentrations for this assay. The percentage of ds and ss DNA was determined by ImageQuant[99] and is shown graphically in **Figure 4.13.**

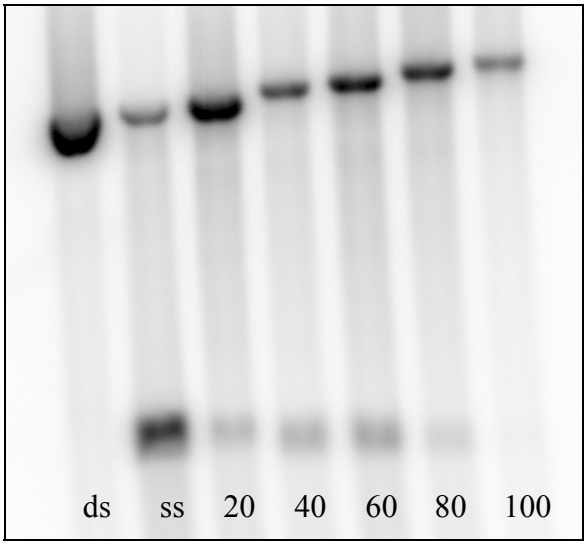



**Figure 4.12** Crosslinking assay of varied concentration of C4-DACA (20 to 100 μM) with 75 mM formaldehyde.

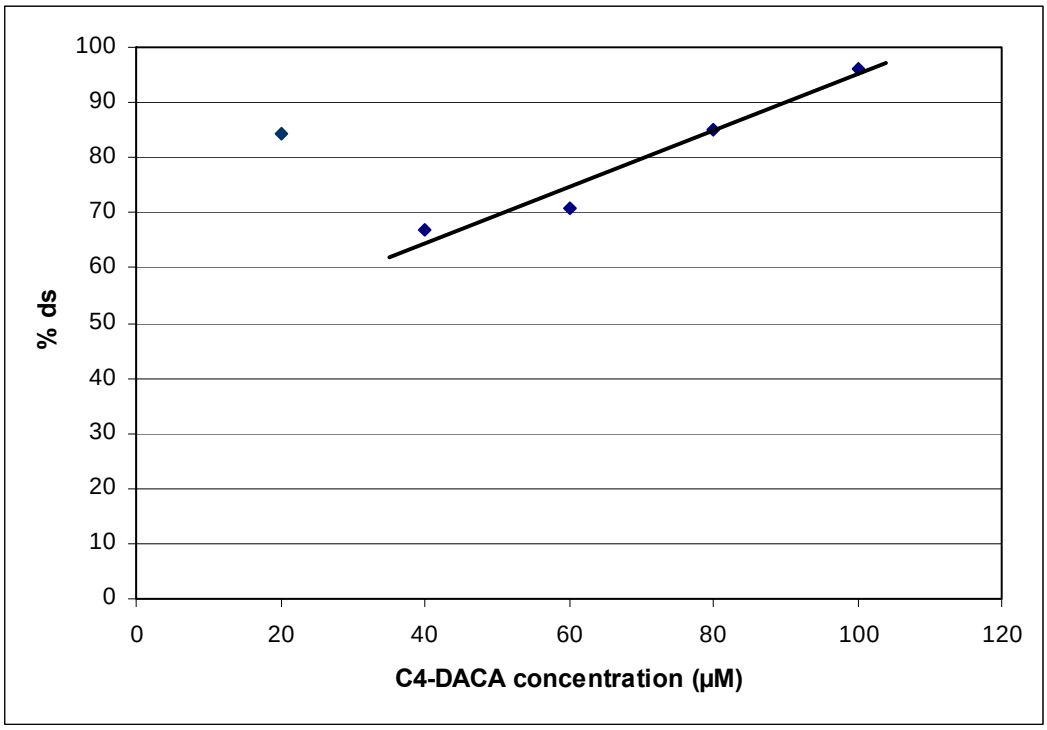



**Figure 4.13** Percentage of ds DNA with increasing C4-DACA concentrations. The point for 20 μM drug appears to be an outlier and was excluded from the line of best fit.

The ss control was not totally denatured, implying either that the formamide buffer was too old, poor resuspension technique was employed, or the concentration of formaldehyde was too high. However, even if only three quarters

of the ss control was denatured, the lanes containing drug show only small amounts of ss DNA, implying that C4-DACA was stabilising the DNA. There is a definite retardation or electromobility shift of the ds DNA evident with increasing drug concentration, implying that C4-DACA had formed a covalent link to the DNA in a drug dependent manner.

### 4.3.8. Concentration Dependence of C4-DACA (5-40 μM)

Another crosslinking assay was completed in order to elucidate the minimum concentration of C4-DACA required to form detectable levels of drug-DNA adducts, and this is shown in **Figure 4.14**. Samples were incubated at 37°C overnight and denatured at 65°C for 5 min. All samples contain 42 $\mu M_{bp}$ DNA, 53 mM phosphate buffer and 79 mM formaldehyde. The ds and ss controls contained no C4-DACA. The six samples had increasing concentrations of C4-DACA ranging from 5 μM to 40 μM. It was noted that the total radioactive DNA present after the phenol/chloroform wash decreased with an increase in C4-DACA concentration. This was attributed to partitioning of the DNA into the phenol layer as noted previously at high adduct levels. The percentage of ds and ss DNA was determined by ImageQuant [99] and is shown graphically in **Figure 4.15**.

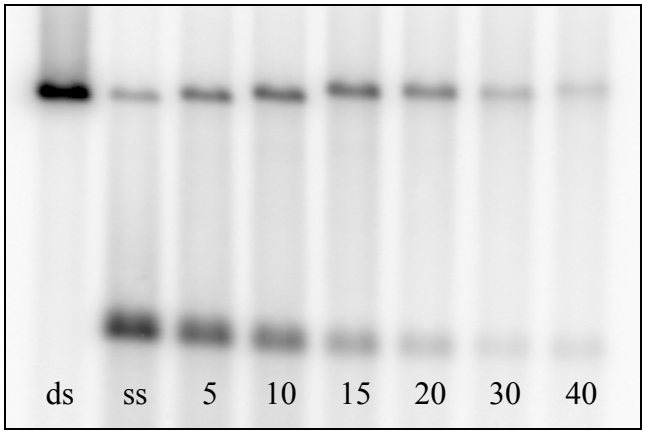



**Figure 4.14** Crosslinking assay of various concentrations of C4-DACA (5 to 40 μM) with 79 mM formaldehyde.

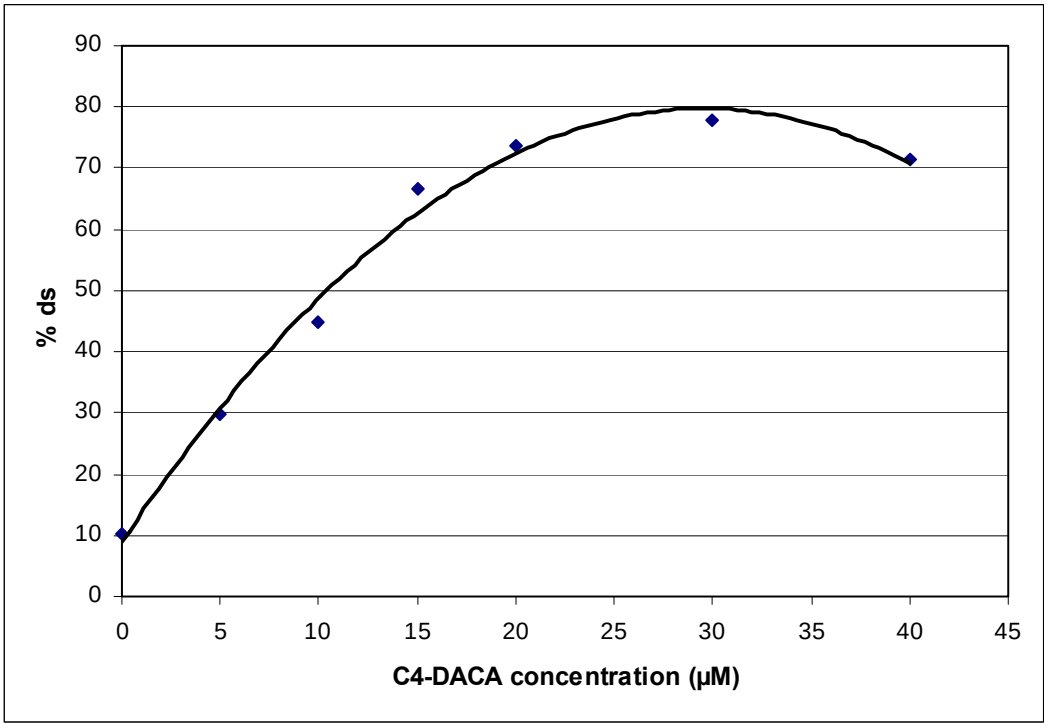



**Figure 4.15** Percentage of ds DNA vs. increasing C4-DACA concentration.

These figures confirm that C4-DACA stabilises ds DNA at 79 mM formaldehyde in a drug-dependent manner and that significant levels of adducts were detectable at drug levels as low as 5 µM. There is also evidence of an electromobility shift indicating the copious amount of drug attaching covalently to the DNA at high drug levels.

### 4.3.9. Time Dependence on Adduct Formation (2 to 24 h)

It was necessary to establish the optimum time of reaction for maximum drug-DNA adduct formation, and this is shown in **Figure 4.16**. The concentration of C4-DACA was kept at 20 µM and formaldehyde at 80 mM, as established from previous experimentation. Samples were incubated at 37°C for 24, 8, 6, 4, 2 and 0 h and denatured at 65°C for 5 min. All samples contained 50 $\mu M_{bp}$ DNA and 50 mM phosphate buffer. The ds and ss controls contained no formaldehyde or C4-DACA. All twelve samples had 80 mM formaldehyde but only those marked with an asterisk contained 20 µM of C4-DACA. The percentage of ds and ss DNA was determined by ImageQuant[99] and is shown graphically in **Figure 4.17**.

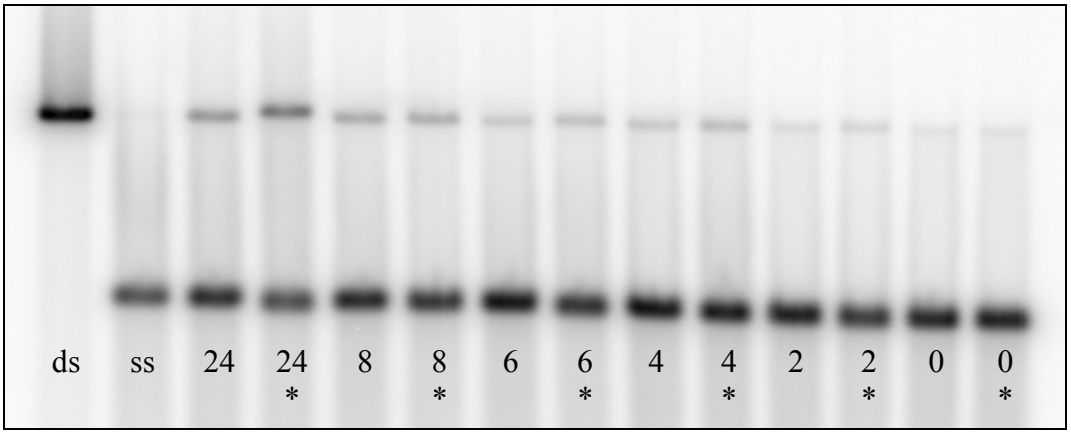



**Figure 4.16** Adduct formation following incubation at 37°C for 0-24 h. Samples containing 20 µM C4-DACA are indicated with an asterisk.

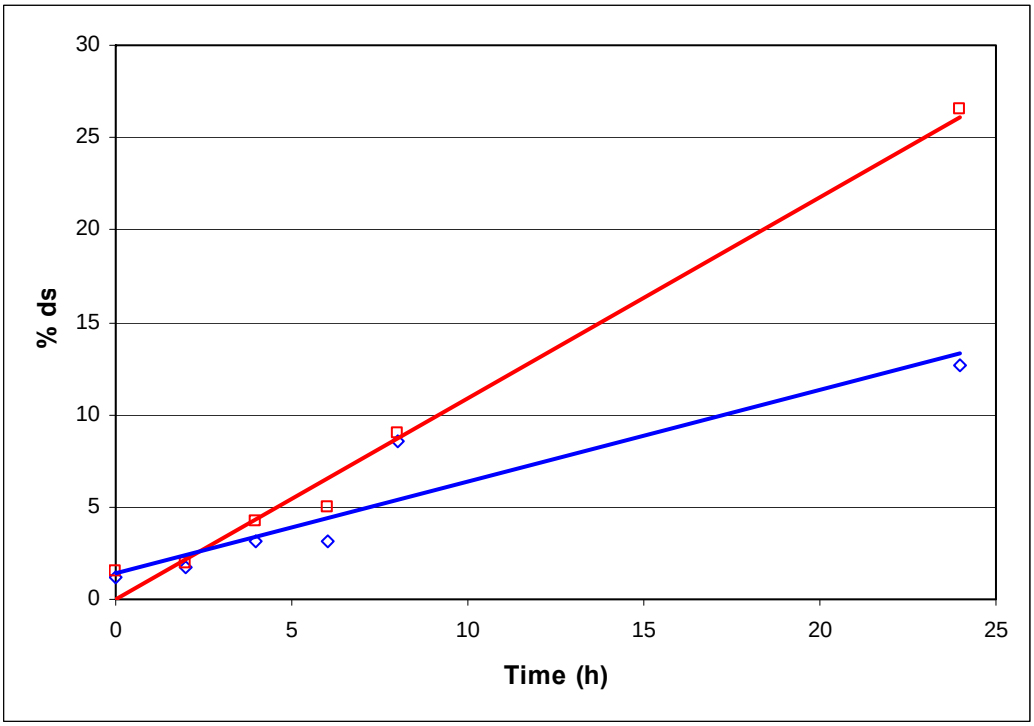



**Figure 4.17** Percentage of ds DNA with increasing reaction time, where the red □ designates a sample containing 20 µM C4-DACA and the blue ◇ indicates the background ds DNA level for that particular time point.

**Figure 4.17** shows that there was no significant difference between the control and drug sample for a reaction time of less than 8 h. The sample at 24 h shows double the amount of ds DNA as the control. An incubation time of less than 24 h was therefore taken as a convenient time in order to obtain a significant

amount of drug-DNA complex. These results prompted clarification of adduct levels formed between 10 and 24 h.

### 4.3.10. Time for Adduct Formation (10 to 24 h)

The reaction time for the period of 10 to 24 h was investigated as shown in **Figure 4.18**. The concentration of C4-DACA was again kept at 20 μM and formaldehyde at 80 mM. Samples were incubated at 37°C for 22, 20, 18, 16, 14, 12 and 10 h and denatured at 65°C for 5 min. All samples contained 50 $\mu M_{bp}$ DNA and 50 mM phosphate buffer. The ds and ss controls contained no formaldehyde or C4-DACA. All twelve samples had 80 mM formaldehyde but only those marked with an asterisk contained 20 μM C4-DACA. The percentage of ds and ss DNA was determined by ImageQuant [99] and is shown graphically in **Figure 4.19**.

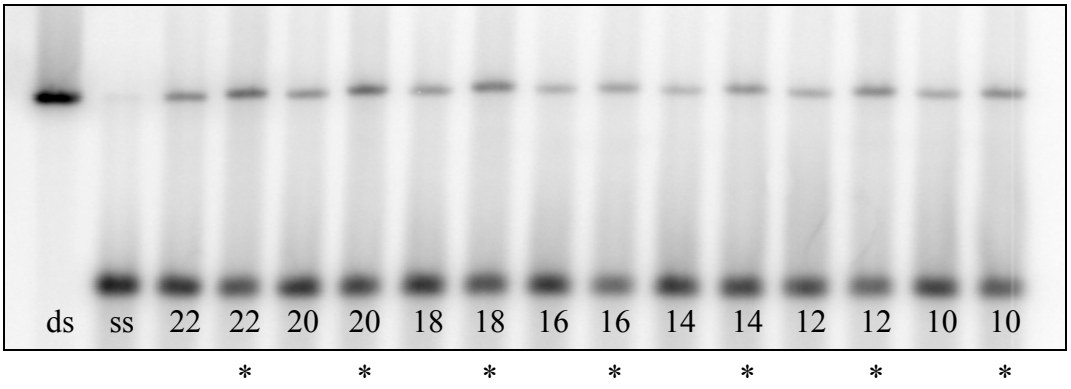



**Figure 4.18** Adduct formation following incubation at 37°C for 10-22 h. Samples containing 20 μM C4-DACA are indicated with an asterisk.

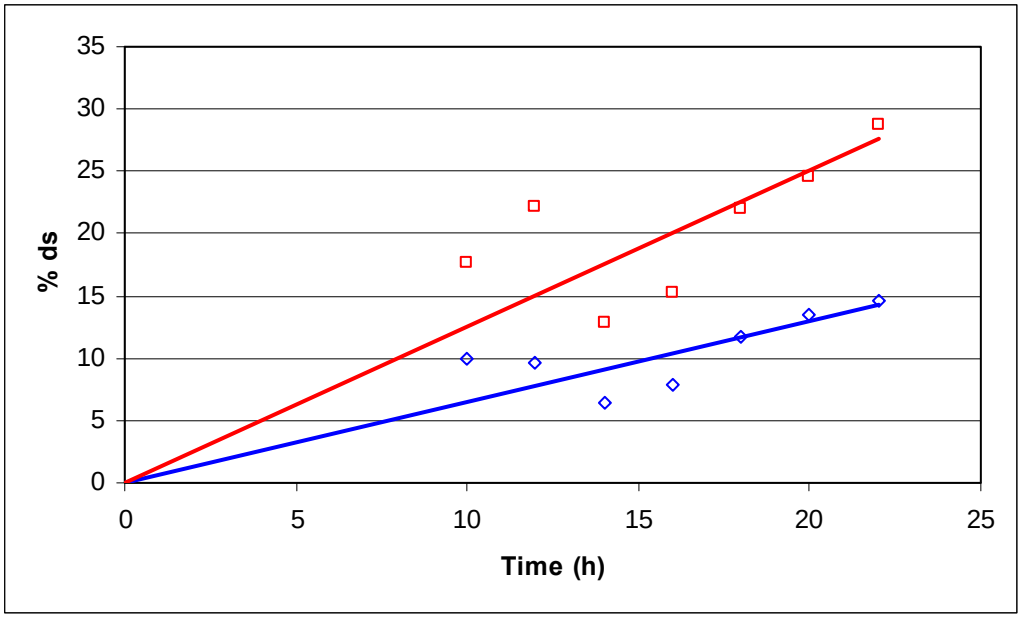



**Figure 4.19** Percentage of ds DNA vs. increasing reaction time, where the red □ designates a sample containing 20 µM C4-DACA and the blue ◇ indicates the background ds DNA level for that particular time point.

Although there is a large degree of error with these results, it is clear that there is approximately double the amount of ds DNA for the samples containing drug as the control for all reaction times. The time required (24 h) for the drug-DNA complex to form with 20 µM C4-DACA and 80 mM formaldehyde is comparable to the previous experiments (ss and background inclusive). These results indicated that the ds DNA occurred in an essentially linear manner up to 22 h. That said, an incubation time of greater than 10 h is sufficient to obtain a significant amount of complex (>13% ds), and thus an overnight reaction between drug/DNA/formaldehyde is adequate to obtain detectable levels of adducts.

### 4.3.11. Varying Concentration of C4-DACA (5-40 µM)

It was decided that the experiment corresponding to **Figure 4.14** be repeated due to the high background levels of adducts detected in these series of crosslinking assays. For **Figure 4.20** samples were incubated at 37°C overnight and denatured at 65°C for 5 minutes. All samples contain 42 $\mu M_{bp}$ DNA, 53 mM phosphate buffer and 79 mM formaldehyde. The ds and ss controls contained no C4-DACA. The six samples had increasing concentration of C4-DACA ranging

from 5 μM to 40 μM. It was noted that the lane intensity decreased with an increase in C4-DACA concentration, as seen previously. The percentage of ds and ss DNA was determined by ImageQuant[99] and is shown graphically in **Figure 4.21**.

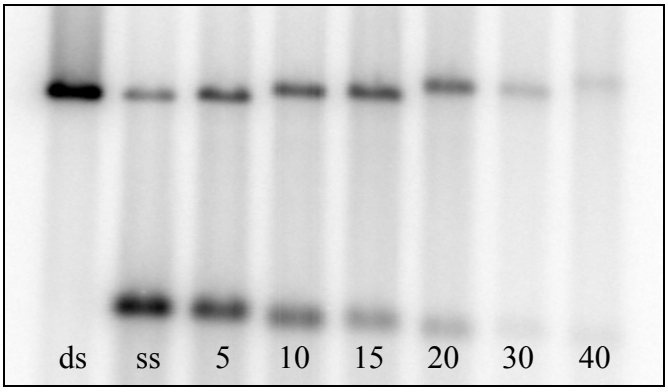



**Figure 4.20** Crosslinking assay of various concentrations of C4-DACA (5 to 40 μM) with 79 mM formaldehyde.

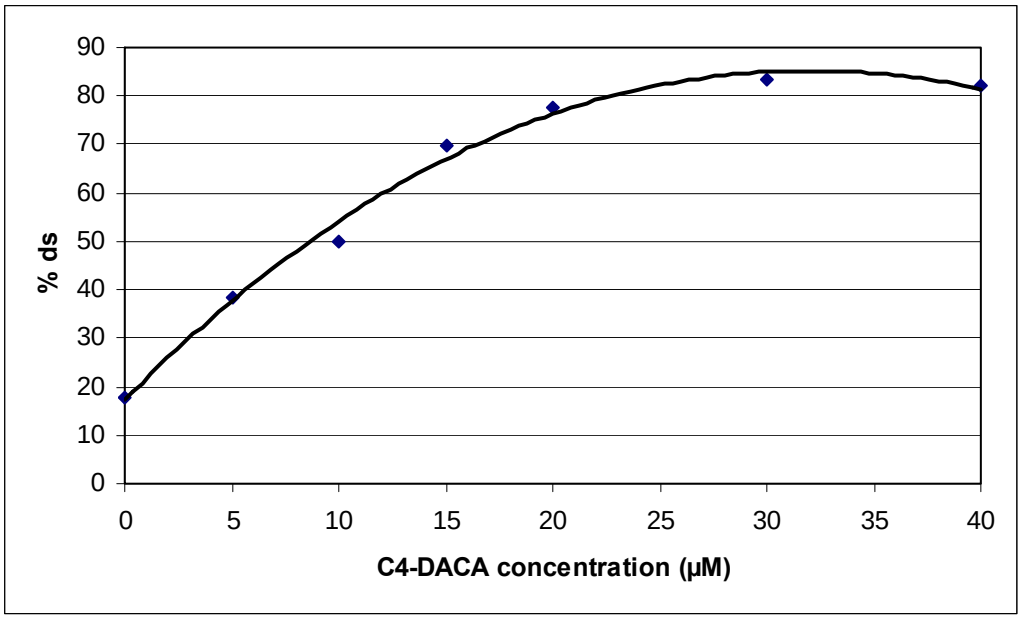



**Figure 4.21** Percentage of ds DNA with increasing C4-DACA concentration.

Compared with the previous experiment, all of the samples showed slightly higher amounts of ds DNA. There was sufficient drug attached to partition the DNA into the phenol layer after approximately 70% of the DNA was in the ds DNA form. This again occurred at approximately 80 mM formaldehyde and concentrations greater than 15 μM C4-DACA. This agrees with previous

experiments that C4-DACA stabilises >65% ds DNA at 80 mM formaldehyde and >15 μM drug (presented in **Figure 4.21**). There is also evidence of an electromobility shift, again indicating the profuse amount of drug attached to the DNA.

### 4.3.12. Temperature Dependence (<4°C) and Varied Time (>Days) for Adduct Formation

The results from reduced temperature experiments and/or varied time (over days) indicated that drug-DNA adducts were not stable if any part of the experiment was performed at 4°C or if samples were subjected to a freeze/thaw process (data not shown).

## 4.4. Conclusions

The ability of the drug C4-DACA to form drug-DNA adducts was investigated. Using crosslinking assays, the drug was found to bind covalently to DNA in the presence of formaldehyde and therefore stabilize ds DNA. There were many possible conditions that needed optimisation using this method, including concentration of drug and formaldehyde, time allowed for activation (pre-reaction), temperature dependence, pH, and phenol/chloroform extraction of unbound drug. The assays revealed that the complex partitioned into the phenol layer at high adduct levels (at approximately 70% ds DNA in the crosslinking assay). This occurred at concentrations greater than 15 μM C4-DACA and 80 mM formaldehyde. There was a distinct difference in migration at the ds DNA level between samples with and without drug. The electromobility shift was attributed to the substantial amount of drug attached to the numerous binding sites on the DNA. C4-DACA was proven to stabilise ds DNA in the presence of formaldehyde. A reaction time of 10 to 24 h yielded a substantial amount of the drug-DNA complex, therefore an overnight reaction was sufficient for these studies. The drug-DNA complex did not form significantly at low temperatures (4°C or -80°C) with a 5 fold decrease detected in the expected amount of adduct formation compared to 37°C. Drug-DNA adducts were not stable after being subjected to freeze/thawing treatments.

The crosslinking assays prompted the requirement for a more definitive answer to how and where C4-DACA binds to DNA. The crosslinking assay merely shows whether the drug stabilises ds DNA, not how and where it binds. The more precise method of a transcription assay would elucidate the finer details of drug-DNA complex formation. It uses a known specific DNA sequence as the template for the controlled process of transcription. If there is something blocking the path of transcription (i.e. covalently bound drug) the progression of RNA polymerase is halted and shortened transcription factors are detected, and this can be utilised to reveal adduct binding sites as well as the adduct stability at these sites.

# CHAPTER 5. C4-DACA *IN VITRO* TRANSCRIPTION ASSAY

## 5.1. Introduction

An *in vitro* transcription assay by Phillips *et al.* [103] showed that Adriamycin blocked transcription at specific sites of DNA in the presence of a reducing environment containing $Fe^{2+}/Fe^{3+}$ and DTT. Adriamycin-DNA adducts were found to have a high sequence selectivity, forming exclusively at GpC sites [103]. It was later found that this adduct formation was mediated by trace amounts of formaldehyde in solution [79]. Furthermore, Adriamycin was found to be covalently linked to only one strand of DNA with the ability to stabilise ds DNA through non-covalent interactions with the opposite strand [66].

After the tantalizingly positive results from the crosslinking assay described in **Chapter 4**, it was decided that a transcription assay would elucidate whether the drug C4-DACA was in fact forming a covalent bond to DNA or if it was only intercalating. The *E. coli lac* UV5 transcription assay has been used extensively in our laboratory to identify the sequence of a number of reversibly binding drugs (actinomycin D, Adriamycin, daunomycin, echinomycin, mithramycin, mitoxantrone, and nogalamycin) as well as drug-induced DNA adducts (Adriamycin, barminomycin, cyanomorpholino-Adriamycin, nitrogen mustards, cisplatin, mitomycin C) [82, 104]. Due to the complexity of set-up of the transcription assay, it is more efficient to examine multiple drugs or drug treatments in one experiment. The transcription assay allows direct visualisation of where a drug binds specifically to DNA. The progression of RNA polymerase is disrupted by blockages on the template strand of DNA (such as drug-DNA adducts) and these resultant truncated transcripts can be used to identify the specific base sequence of the blockage sites.

## 5.2. Overview of Transcription Assay

Transcription is the synthesis of RNA from a template strand of DNA. There are three main phases involved in transcription – initiation, elongation and termination. The initiation step is further separated into the two discrete phases of DNA binding (assembly) and initiation of RNA synthesis. Regulation of transcription is a complex process and can occur at any step in transcription, however most is directed at the polymerase binding and initiation steps. There are proteins which activate or increase transcription of specific genes by assisting binding of RNA polymerase, and proteins which repress transcription by blocking promoter sequences and hence repress synthesis at particular genes.

The assembly or initiation stage involves RNA polymerase binding to a core promoter sequence within the DNA. *E. coli* promoters are characterised by two sets of sequences of six nucleotides each located 10 and 35 base pairs upstream of the transcription start site (+1) [105]. RNA polymerase binds to these promoter regions, and requires all four ribonucleoside 5'-triphosphates (ATP, GTP, UTP and CTP), $Mg^{2+}$, and $Zn^{2+}$ to attach to DNA, separate the two strands, and allow access to the single template strand of DNA. The DNA strand unwinds over about 17 base pairs to create a transcription 'bubble', referred to as the open complex, shown in **Figure 5.1** [106]. Promoter escape occurs after the polymerase has synthesised the first ten nucleotides [107]. During this time there is a proclivity for the RNA transcript to be released and consequently generate truncated transcripts, referred to as abortive initiation. Early elongation complexes are unstable until the transcript reaches about ten nucleotides long, after which it no longer slips and the elongation step can proceed [107]. This growing end of the new RNA strand temporarily forms base-pairs with the template DNA strand, seen in **Figure 5.1**, forming a short hybrid RNA-DNA double helix, measured by X-ray diffraction studies to be 9 base-pairs long [107, 108].

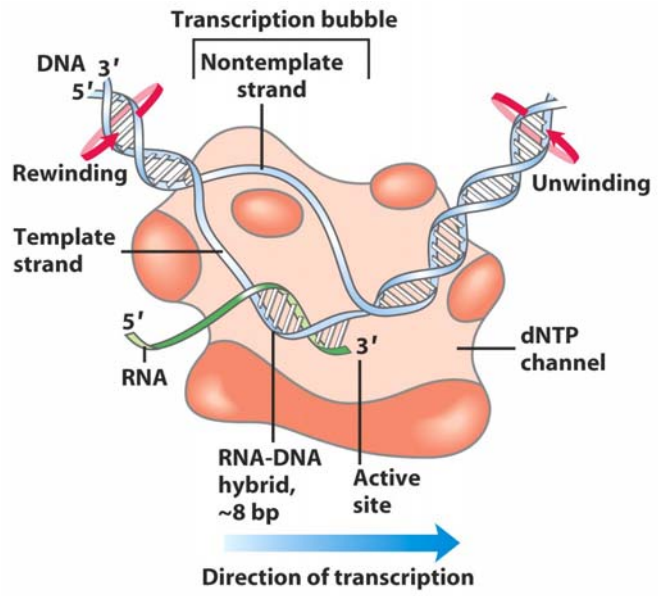



**Figure 5.1** Transcription by RNA polymerase in E. coli [106].

Elongation is essentially the addition of nucleotides in the specific sequence provided by the template DNA strand. The nucleophilic attack of the 3'-hydroxyl group at the end of the growing RNA strand is directed at the α phosphate of the incoming NTP (NTP denotes all four ribonuclease 5'-triphosphates – ATP, GTP, UTP and CTP). The reaction is catalysed by $Mg^{2+}$ ions associated with the phosphate groups of the incoming NTP, and three Asp residues within the RNA polymerase [106]. Movement of the RNA polymerase along the DNA strand induces positive supercoils (overwound) ahead of the transcription bubble and negative supercoils (underwound) behind it. In the cell, it is the purpose of topoisomerases to rapidly remove the positive supercoils and regulate negative supercoiling.

Termination in eukaryotes is not well understood, but for bacteria there are two classes of termination signals. One class relies upon a protein factor called ρ (rho) whereas the other is ρ-independent. In ρ-independent termination, the newly synthesised RNA molecule forms a G-C rich hairpin loop followed by a run of U residues which disrupts the RNA-DNA hybrid. When a polymerase arrives at a termination site with this structure, it pauses. As this section of the new transcript is relatively unstable, the RNA dissociates from the DNA template. In ρ-dependent termination, the protein factor ρ associates with RNA at the termination site and destabilises the RNA-DNA interaction, resulting in RNA dissociation.

The RNA polymerase inhibition assay (transcription assay) results in the accumulation of truncated transcripts arising from the blockage of movement of RNA polymerase along the DNA at specific drug-binding sites [104]. An overview of the procedure is shown in **Figure 5.2** [104]. The length of the truncated RNAs therefore reveals the location of these drug-binding sites.

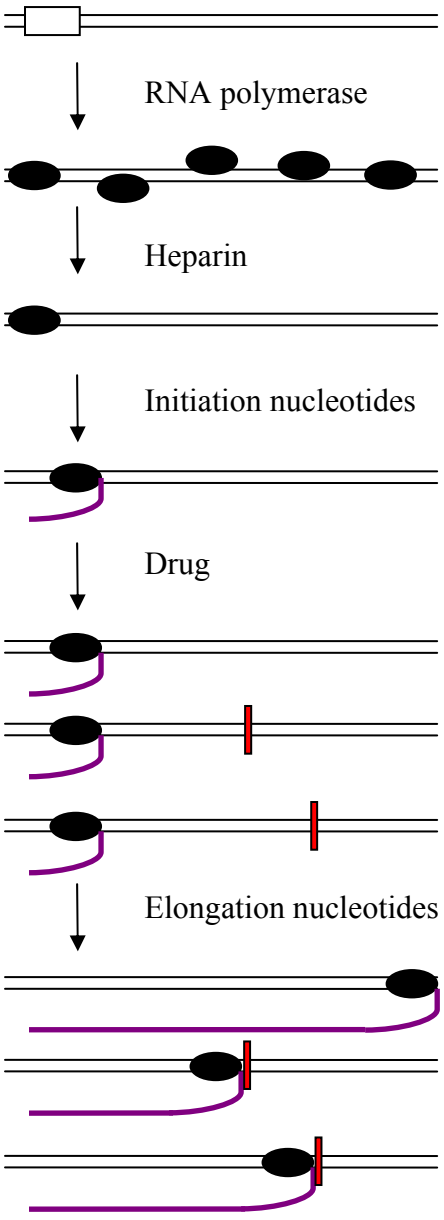



**Figure 5.2** Overview of the transcription assay. The major steps are binding of *E. coli* RNA polymerase selectively to the *lac* UV5 promoter, formation of a synchronised initiated transcription complex, reaction of the initiated transcription complex with the drug of interest, and elongation of the transcription complex to yield drug-induced blocked transcripts [104]. The black oval denotes RNA

polymerase, the growing RNA strand is shown in purple, and the red rectangles represent drug induced blockages.

## 5.3. Overview of Procedure

### 5.3.1. Plasmid Preparation

The 512 base pair *lac* UV5 DNA fragment was isolated from the plasmid pCC1 following restriction by *Pvu*II and *Hind*III. The resulting DNA fragments were separated electrophoretically in an agarose gel, recovered, purified and redissolved in TE buffer. The DNA template contains the *lac* UV5 promoter as this promoter allows for robust production of synchronised transcripts using *E. coli* RNA polymerase [109].

### 5.3.2. Formation of Initiated Transcription Complexes

Initiation of synchronised complexes requires *E. coli* RNA polymerase to bind to the *lac* UV5 promoter. Heparin was added to displace bacterial RNA polymerase from any non-specific binding sites on the DNA. "This procedure ensures that only single-copy transcripts result from the subsequent elongation step because the RNA polymerase will be unable to rebind to the promoter due to competition with heparin" [104]. Nucleotide mix containing a radiolabelled nucleotide (usually UTP) but lacking CTP was added for initiation. The omission of CTP forces the initiation complex to pause at the first dGTP of the template strand (10 nucleotides long). Inclusion of a GpA dinucleotide ensures that the first nucleotide of the newly synthesised transcript is always at the -1 position, and therefore initiated transcripts are synchronised from a common starting point [109]. All nucleotides were maintained at <5 μM to ensure that transcription begins from one site only. "It is important to use high-purity, sterile water to prepare all solutions for this procedure because the presence of trace amounts of metal ions, bacteria or nucleases can completely destroy transcription complexes" [104]. Fresh aliquots of dithiothreitol (DTT) were used, and RNase inhibitor added due to the long reaction time. An optimal $MgCl_2$ concentration ensures efficient transcription and minimises natural pausing. Fresh $^{32}P$ nucleotides ensure minimal amount of radiolytic degradation products that inhibit transcription [110].

### 5.3.3. Preparation of RNA-Sequencing Reactions

Sequencing of the DNA template was achieved by adding chain terminating 3'-*O*-methylnucleotides to the initiated complex, with subsequent incorporation during the elongation phase of transcription [104]. The sequence of transcripts that are terminated by DNA sequence-specific drug blockages to RNA polymerase can then be assigned according to the length of the truncated transcripts. The methoxynucleotides yield C and G sequencing lanes analogous to dideoxy-terminated DNA-sequencing lanes [104]. The 3'-methyoxynucleotide:nucleotide ratio must be reduced for RNA longer than 150 nucleotides to allow RNA polymerase to transcribe further along the DNA before termination.

### 5.3.4. Reaction of Initiated Transcripts with Drug

The drug of interest may be added directly to the initiated transcription complex, or pre-reacted with the DNA so that unreacted drug may be removed. The latter procedure is the preferred method when covalently binding drugs react with other components of the transcription assay (e.g. alkylation of DNA by nitrogen mustards is inhibited by the transcription buffer itself [111]; cisplatin reacts with RNA polymerase) or when drugs interact with the unprotected promoter region, leading to a decrease in the RNA polymerase-promoter complex [96, 104]. Minimum drug concentrations are employed in order to obtain a range of different drug sites for detection. If the DNA is saturated with drug the first drug site would be completely occupied and most of the RNA polymerase would not be able to proceed to elongate and elucidate any additional downstream drug sites [104]. High levels of all four nucleotides and high ionic strength (0.4 M KCl) during the elongation process minimises natural pausing of RNA polymerase. The high level of nucleotides also ensures negligible addition of radiolabel into the growing RNA chain [104].

### 5.3.5. Separation of Blocked Transcripts

Transcripts were separated on high resolution sequencing gels. If control lanes contain a significant amount of background the most likely causes are insufficient purity of the promoter-containing DNA fragment, or degradation of the DNA by the presence of single-strand or double-strand nicks [104]. An additional phenol-chloroform extraction provides high-quality template DNA. If, however, the quality of the DNA is not the problem, fresh stock solutions of all reagents should be made.

### 5.3.6. Quantitation of Blocked Transcripts

There are two methods of determining the relative amounts of each length of RNA. Those are autoradiography and the same phosphorimaging process as described for the crosslinking assays. The phosphorimaging process is preferred because it is faster, more sensitive (by at least 250-fold for $^{32}P$), has a bigger dynamic range (at least 400 times that of film), and is fully computerised [104].

### 5.3.7. Relative Occupancy and Drug Dissociation Kinetics

A true dissociation rate constant can only be determined from the first possible drug site [104]. Any other dissociation rate constants are distorted to some degree as a result of RNA polymerase being released from previously drug-occupied blockages at upstream sites. Using subsaturating concentrations of the desired drug ensures an almost full-length transcript is reached for the majority of initiated transcription complexes before termination occurs. “Decay of transcriptional blockages can be quantitated to reveal the time-dependent loss of drug-DNA adducts at individual sites” [104]. The *in vitro* transcription assay can provide an assessment of adduct lability in a physiological environment. “Loss of drug-DNA adducts during transcription enables the progression of RNA polymerase past the lesion, yielding a clear decay in discrete transcriptional blockages with time.” [112]

## 5.4. Materials

Heparin was purchased from Sigma. RNAguard RNase Inhibitor (Human Placenta), RNase/DNase free BSA, 3′-*O*-methyl NTPs, and GpA dinucleotide were purchased from GE Healthcare. *E. coli* RNA polymerase and glycogen were obtained from Roche. TEMED and the restriction enzyme *Pvu*II were from Promega. DTT and ammonium persulfate were purchased from Bio-Rad. Molecular biology grade urea was from MP Biomedicals and acrylamide:bisacrylamide solution was purchased from Amresco. [$\alpha^{32}$P] UTP (3000 Ci/mmol) was purchased from Perkin Elmer. The remaining chemicals and reagents were of analytical grade. Distilled water passed through a four stage Milli-Q purification system was used to prepare all solutions.

### 5.4.1. Drugs

The acridine C4-DACA was synthesised as described in **Chapter 2** by Dr Les Deady (La Trobe University). C4-DACA stock solution (stored at 4°C) was prepared by dissolving the oxalate salt in Milli-Q water to an approximate concentration of 2 mM. An accurate concentration of the drug was determined spectrophotometrically using $\varepsilon$ = 8,305 $M^{-1}cm^{-1}$ at 356 nm. Formaldehyde solutions were freshly prepared on the day of each experiment.

### 5.4.2. DNA Source

"*E. coli* HB101 cells containing the plasmid pCC1 were grown overnight in selective LB broth containing ampicillin. The cells were harvested and the amplified plasmid isolated using a Qiagen Maxi Plasmid Purification Kit. A restriction digest using the enzymes *Pvu*II and *Hin*dIII was used to liberate a 512 bp fragment containing the *lac* UV5 promoter from the pCC1 plasmid [3]. The digest was subjected to electrophoresis through an agarose gel and the 512 bp fragment excised in a gel slice. The fragment was subsequently electroeluted from

the gel slice using an Elutrap Electroelution apparatus (Whatman) and further purified using a phenol/chloroform extraction. The isolated fragment was ethanol precipitated and resuspended in 1 × TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0)." [112]

## 5.5. *In vitro* Transcription Assay

Covalent drug-DNA adducts were formed in a reaction mixture in which the 512 bp DNA fragment (25 $\mu M_{bp}$) was reacted with C4-DACA and formaldehyde in phosphate-buffered saline (pH 7.0) at 37°C over 20 h. "Drug-reacted DNA was subsequently ethanol precipitated to remove unreacted formaldehyde and then resuspended in 1 × transcription buffer consisting of 40 mM Tris pH 8.0, 100 mM KCl, 3 mM $MgCl_2$ and 0.1 mM EDTA. The *in vitro* transcription assay employing RNA polymerase transcription from the *lac* UV5 promoter has been extensively reviewed elsewhere [104]. Briefly, a transcription complex was formed by incubating 12.5 $\mu M_{bp}$ drug-reacted DNA fragment, 0.025 U/μL *E. coli* RNA polymerase, 10 mM DTT, 125 μg/mL BSA and 2 U/μL RNase inhibitor in 1 × transcription buffer. Non-specifically bound RNA polymerase was removed from the DNA template by the addition of heparin (400 μg/mL). A labelled initiation complex was subsequently formed by adding 200 μM GpA, 5 μM ATP, 5 μM GTP and [$\alpha^{32}$P] UTP. Transcripts were then allowed to elongate for a defined time period by adjusting the concentration of all four nucleotides to 2 mM and KCl to 400 mM. Each reaction was terminated by the addition of an equal volume of loading/termination buffer (9 M urea, 10% sucrose, 40 mM EDTA, 0.1% xylene cyanol and 0.1% bromophenol blue in 2 × TBE). All transcription samples were denatured at 90ºC for 5 min and immediately quenched on ice." [112]

### 5.5.1. Sequencing of RNA Transcripts

"The sequences of RNA transcripts were determined by incubating the non-drug-treated initiated transcription complex with chain-terminating 3′-*O*-

methyl nucleotides in the elongation phase of transcription. This was achieved by incubating the initiated transcription complex with 90 μM 3′-*O*-methoxy-ATP (or 3′-*O*-methoxy-CTP), 10 μM ATP (or CTP), 2 mM CTP, GTP, UTP (or ATP) and 400 mM KCl at 37°C for 5 min. Transcriptional elongation was subsequently terminated by the addition of an equal volume of loading/termination buffer and the samples denatured as described above." [112]

### 5.5.2. Electrophoresis, Phosphorimaging and Quantitation

Denatured transcripts were loaded onto a pre-electrophoresed 12% denaturing polyacrylamide gel (19:1 acrylamide:bisacrylamide, 7.5 M urea). Transcripts were subsequently resolved by electrophoresis at 2000 V for 1.5–2 h in 1 × TBE buffer. Following electrophoresis, the gel was fixed in 10% glacial acetic acid/10% methanol and vacuum dried on a Bio-Rad Model 583 gel drier. The dried gel was exposed to a phosphor screen overnight and the screen scanned using a model 400B PhosphorImager. The image of the gel was analysed and quantitated using ImageQuant software (Molecular Dynamics, CA)." [112]

## 5.6. Results and Discussion

It can be seen from **Figure 5.3** that the *in vitro* transcription assay results in an accumulation of truncated transcripts that arise from the impairment of RNA polymerase movement along ds DNA by the presence of drug-DNA lesions. "The length of each drug-induced truncated transcript provides a direct measure of the location of the lesion on the DNA template. In the absence of drug-induced blockages, RNA polymerase can progress through the DNA template to yield a full length transcript (FLT) that is 379 nucleotides in length." [112] The following *in vitro* transcription assay was kindly performed by Dr Ben Evison (La Trobe University).

### 5.6.1. Transcriptional Detection of C4-DACA-DNA Adducts

After the mixed results from the crosslinking assays, it was necessary to establish if C4-DACA was forming covalent drug-DNA adducts, and a useful and widely employed method was deemed to be the *in vitro* transcription assay. This technique does not recognise drugs that merely intercalate as RNA polymerase has the ability to pause during transcription for up to sixty seconds [113]. A drug that is only intercalating in one groove will have a much shorter half life than that of a covalently bound drug-DNA adduct (e.g. intercalated daunomycin $t_{1/2} \leq 1$ s [114], formaldehyde mediated daunomycin-DNA adduct $t_{1/2} \geq 7.5$ h [26], formaldehyde mediated mitoxantrone-DNA adduct $t_{1/2} \geq 3$ h [115]). For this reason, the transcription method was appropriate for the detection of formaldehyde-activated drug-DNA lesions, assuming that the half-life of these adducts would be at least several minutes. A simple drug concentration-dependence assay was used to confirm this idea. Prior to transcription, the 512 bp fragment containing the *lac* UV5 promoter was reacted with 10 mM formaldehyde and concentrations of C4-DACA of 0, 20 and 80 μM overnight. In the absence of C4-DACA, RNA polymerase transcribed through the DNA template efficiently to yield the 379 FLT, indicating that formaldehyde alone was insufficient to induce any transcriptional blockages (**Figure 5.3**). As seen in **Figure 5.3**, in the absence of formaldehyde (lane 0 mM), RNA polymerase effectively progressed through the 512 bp DNA template to produce FLT. The absence of drug-induced transcriptional blockages in this control indicates that formaldehyde is crucial for the generation of C4-DACA adducts with the ability to inhibit transcription. The progression of RNA polymerase through each drug-reacted DNA template was increasingly impaired at specific sites with increasing concentrations of C4-DACA (**Figure 5.3**).

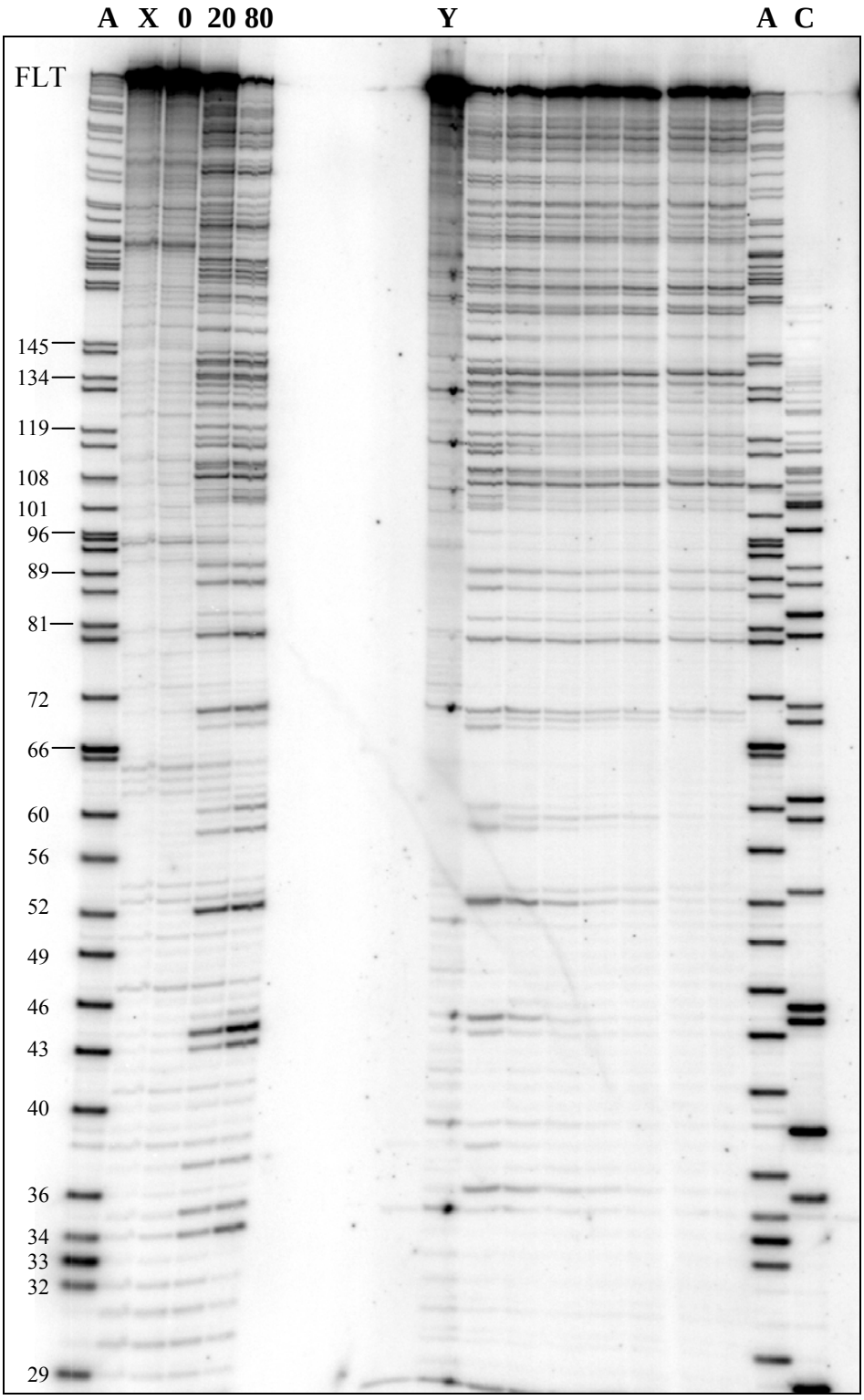



**Figure 5.3** Lane **X** is the control (no formaldehyde or drug). Those designated **0**, **20**, and **80** contain 10 mM formaldehyde with 0, 20, and 80 μM C4-DACA respectively. Lanes **A** and **C** are sequencing lanes using the chain terminators 3′-*O*-methoxy-ATP or 3′-*O*-methoxy-CTP respectively. Lane **Y** denotes the elongation of this complex to FLT. Unlabelled lanes were part of another study on pixantrone of varying concentrations.

The data from **Figure 5.3** allows the individual binding/blockage sites for the drug to be elucidated. An important feature of this experiment is that another DNA intercalator (pixantrone) was run in the unmarked lanes (at varying concentrations, lanes between **Y** and **A**), which shows blockage sites similar to C4-DACA. The sequencing lanes **A** and **C** use the chain terminators 3′-*O*-methoxy-ATP or 3′-*O*-methoxy-CTP respectively, and indicate termination of the transcript by adenine or cytosine respectively. The bold lines in lanes **20** and **80** show the blockage sites resulting from formaldehyde-activated C4-DACA. Due to the 4 fold increase in drug concentration, the intensity of blockages (lane **80**) is noticeably greater than in lane **20**, as expected. The negligible amount of background transcriptional blockages can be seen in lane **X** and is comparable to that seen from 10 mM formaldehyde treatment in lane **0**, indicating that essentially no blockages result without the presence of drug. The two drugs (C4-DACA and pixantrone) have almost identical blockage sites, as can be seen in **Figure 5.3**.

The increase in transcriptional blockages with increasing concentrations of either C4-DACA or formaldehyde was also associated with a decrease in the fraction of full length transcript. This is a direct reflection of transcriptional inhibition by formaldehyde-mediated C4-DACA-DNA adducts. The presence of discrete transcriptional blockages induced by the C4-DACA formaldehyde combination indicates that the drug-DNA adduct demonstrates a degree of sequence specificity.

These results show that the binding of the formaldehyde-activated C4-DACA to DNA was sufficient to interrupt the process of transcription. The disruption of the fundamental process of transcription would undoubtedly have a significant effect on cell viability. The combination of C4-DACA and formaldehyde generated a DNA lesion that blocked the movement of RNA polymerase during the process of transcription. The generation of sequence-specific truncated transcripts and the overall inhibition of transcription indicate impairment of RNA polymerase. The DNA damage caused by formaldehyde-activated C4-DACA-DNA adducts may therefore ultimately contribute towards the drugs cytotoxicity.

Although the activated drug may intercalate reversibly between any base pair combination, the drug-DNA adducts exhibited a distinctively sequence-dependent occurrence, with the sequence-selective blockage of RNA polymerase at and immediately prior to CpG and CpA dinucleotide sequences. The formation of C4-DACA-DNA adducts at these sequence specific sites indicates that there is a favourable 3-dimensional access point to a receptive nucleophilic site on the DNA, suggesting the formaldehyde-activated C4-DACA reacts reversibly with guanine and/or cytosine. The N7 position of guanine has been established as the most reactive site in nucleic acids for numerous alkylating agents in neutral aqueous solutions [116, 117], thus it is most likely the site of adduct formation. It is evident from **Figure 5.3** that C4-DACA has a remarkably similar GpC sequence selectivity to that reported for formaldehyde-activated pixantrone [115, 118] and formaldehyde-activated mitoxantrone [13, 119]. This sequence selectivity differs significantly from the 5'-GpC sequence preferred for formaldehyde-activated doxorubicin adducts [66] and formaldehyde-activated daunorubicin [6, 25]. Despite the differences in sequence selectivity, the four aforementioned formaldehyde-activated drugs have an absolute requirement for the exocyclic N2 amino group of guanine [2, 4, 59, 115]. The structural similarities between C4-DACA and these drugs indicate the likelihood is that C4-DACA forms formaldehyde mediated adducts with DNA via the exocyclic N2 amino group of guanine.

### 5.6.2. Sequence Specificity of Transcriptional Blockages

The sites of highest transcriptional blockages (shown in bold) are summarised in **Table 5.1**, with likely adduct sites indicated by the rectangular boxes. This table highlights the sequence selectivity of formaldehyde-mediated C4-DACA-DNA adduct formation.

**Table 5.1** C4-DACA-DNA adduct blockage sites. Those shown in bold indicate intense blockage sites, with the blockage site shown numerically to the left (as the length of the truncated transcript, beginning from the -1 position). The respective likely adduct attachment sites are shown within the rectangular boxes.

| Blockage site | Blockage Sequence |
|---|---|
| 44 | A C **C** A T G |
| 52 | T **A** C G G |
| 70 | C **T** C A T |
| 80 | G A **C** A G C |
| 109 | A **C** G C C G |
| 112 | G C **C** G G |
| 115 | A **C** G C |

## 5.7. Conclusions

The ability of the drug C4-DACA to form drug-DNA adducts was investigated. Using a transcription assay, the drug was found to bind covalently to DNA in the presence of formaldehyde and therefore produce truncated transcripts. The assay revealed that formaldehyde-mediated C4-DACA-DNA adducts disrupted the progression of RNA polymerase during transcription. The drug-DNA complex was achieved using 10 mM formaldehyde with 20 and 80 μM C4-DACA reacted with DNA overnight. The amount of truncated transcripts increased with an increase in drug concentration in a sequence-dependent manner, with the sequence-selective blockage of RNA polymerase at and immediately prior to CpG and CpA dinucleotide sequences. Formaldehyde-activated C4-DACA was also found to form adducts at strikingly similar positions to that of formaldehyde-activated pixantrone and formaldehyde-activated mitoxantrone, indicating that C4-DACA most likely forms formaldehyde-mediated adducts with DNA via the exocyclic N2 amino group of guanine.

# CHAPTER 6. MODEL ENERGY CALCULATIONS FOR THE REACTIONS OF AMINE DRUGS WITH ALDEHYDES AND DNA

## 6.1. Introduction

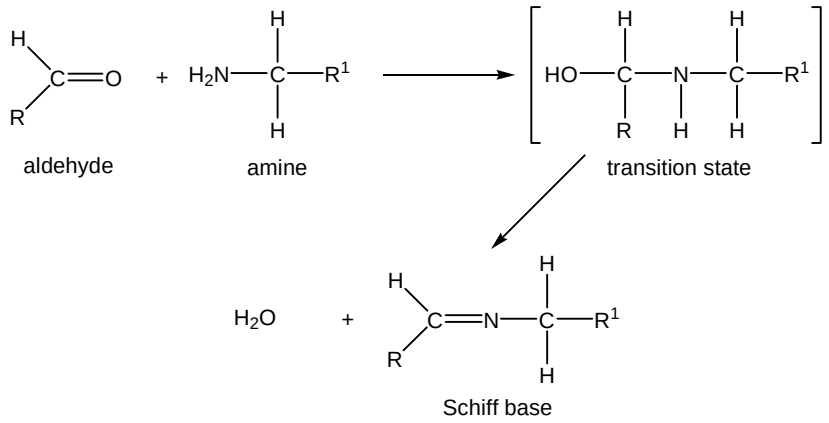
Compounds that contain a primary amine group can react with aldehydes to form a Schiff base through the elimination of water. Hugo Schiff first reported the nucleophilic addition of a primary amine to an aldehyde through an intermediate to create an imine in 1864 [120]. This Schiff base can further react with a second primary amine to form an "aminal" adduct. Adriamycin contains such a primary amine, which forms a Schiff base in the presence of formaldehyde [4]. This activated or imine form of Adriamycin has been shown to react with the free amine of guanine via the minor groove to produce a stable covalent adduct with high biological activity [6, 79]. A schematic model for these two reactions is shown in **Scheme 6.1a** and **6.1b**. The compounds studied in this work contain a primary amine and are expected to react in a similar way to Adriamycin.

The objectives of the work presented in this chapter are to examine the energetics of both the Schiff base and aminal formation, and to determine which aldehyde is more energetically favourable to form the Schiff base. The energy of a reaction can convey information about the stability of products and the conditions at which the reaction may proceed. In particular, the energy involved in the reaction with primary amines and formaldehyde could aid in ascertaining binding information, and so help discover how a drug interacts with DNA. It was anticipated that formaldehyde may be the more energetically favourable aldehyde, since it is found at higher concentrations in cancer cells [80].

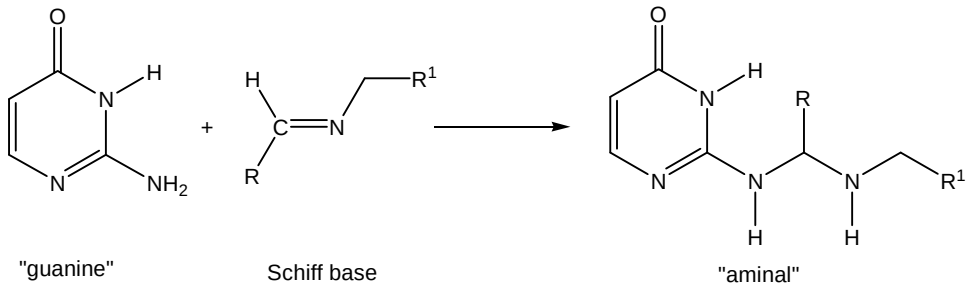
The energetics of these reactions were studied in detail in order to better understand the mechanism behind adduct formation. In particular, it is important to know if there is a significant reduction in energy from the initial reactant to the drug-formaldehyde-DNA adduct, indicative of a favourable forward reaction. Theoretical studies presented in this chapter involve simplified model compounds that replicate the drug and DNA to reduce the complexity and time of calculations. The sheer size of the molecules required to accurately estimate the energies

involved in the adduct formation was not feasible, thus only the model components depicted in **Scheme 6.1a** were used.

The general mechanism for formaldehyde-activated drug-DNA adduct formation has two distinct steps and is shown in **Schemes 6.1a** and **6.1b**. The first step involves nucleophilic addition of the amine to an aldehyde. This produces an intermediate involving proton transfer, which then decomposes with the elimination of water to yield the corresponding Schiff base (**Scheme 6.1a**). The second step is the addition of guanine (in the DNA) via nucleophilic attack of the imine affording the drug-DNA adduct (**Scheme 6.1b**). The product from **Scheme 6.1a** is known to react with the amine of guanine in DNA, which is represented by a simpler model compound in a one-step reaction (**Scheme 6.1b**).




**Scheme 6.1a** Schiff base formation where R = H, $R^1 = CH_3$.




**Scheme 6.1b** Schiff base and pseudo-guanine reaction where R = H, $R^1 = CH_3$.

## 6.2. Quantum Chemical Calculations

It is possible to use quantum mechanical methods to calculate the energy for a single molecule, a group of molecules or a solid. Quantum mechanical methods use elaborate wave equations based on solving the Schrödinger equation to approximate properties such as energy, molecular structure, vibrational frequency and molecular properties of molecules. Methods that are calculated directly from first principles (non-empirical) are referred to as *ab initio* methods. Those that use empirical parameters employ experimental results of accepted models and are referred to as empirical or semi-empirical methods. *Ab initio* calculations generally have lower errors with models involving few electrons (less than 40) so are best suited to small molecules. Semi-empirical methods give good approximations on large molecules, whereas very large molecules require classical mechanics methods due to the computational cost of the more extensive *ab initio* methods.

The energetics of the Schiff base formation reaction was investigated using two *ab initio* methods as well as the semi-empirical AMPAC method. In the case of the *ab initio* methods, early calculations of the electronic energies (enthalpies) used the MacGAMESS program (Version October 26 2000 R4). However, during the course of this work, more sophisticated methods have become available for the calculation of molecular thermodynamic properties. The later calculations were performed using the Gaussian 09 [121] software suite. The semi-empirical calculations employed the AMPAC 7.0 software [122].

Structures for each of the molecules from **Schemes 6.1a** and **6.1b** were created using a molecular builder: in the case of MacGAMESS, MacMolPlot was employed using Z-matrix co-ordinates, and in the case of Gaussian 09, WebMO or GaussView were used. In all cases the structures were geometry optimised at the appropriate calculation level before calculating the thermodynamic properties.

Energy calculations using MacGAMESS and Gaussian 09 give the electronic reaction energy ($E_e$ at 0 K), and when combined with the zero-point energy, the $E_0$ of reaction. In Gaussian 09, a standard temperature of 298.15 K and 1 atm pressure were used in the calculation of thermodynamic properties (enthalpy and Gibbs free energy). This requires the addition of a full thermal correction ($\Delta H_T$) rather than just the zero-point energy correction to the electronic energy.

The AMPAC program additionally gives the energies of the reaction along a reaction pathway at regular intervals, allowing a visualisation of the reaction energetics along with the corresponding total energy. These calculations were performed firstly in the gas phase, followed by simulations in the presence of water and then methanol solvents. It should be noted that in this chapter the words exothermic and endothermic are used generically to describe changes in electronic energy and enthalpy. Analogous changes in the Gibbs free energy are denoted exergonic and endergonic.

The energy involved in the reaction of a simple primary amine with formaldehyde and then with a second amine to form an "aminal" was first studied using MacGAMESS. The Schiff base formation reaction is shown in **Scheme 6.1a** and the aminal reaction in **Scheme 6.1b**. The total energy of **Schemes 6.1a** and **6.1b** is quoted in atomic units (a.u.), where 1 a.u. = 2625.5 kJ/mol.

## 6.3. Reaction Energy Calculations using MacGAMESS

Each component structure was built and optimised at the Hartree-Fock level of theory using a range of progressively more sophisticated basis sets; MIDI, 6-31G and 6-311G. MIDI basis sets are not scaled by MacGAMESS and are derived from the MINI sets. The MINI basis sets are constructed by employing an expansion of three Gaussian-type functions for each atomic orbital (AO) in a similar manner to the STO-3G basis set, however the MINI set gives much lower energies than does STO-3G [123].

Generally one starts calculations with the simpler basis sets (e.g. MIDI, 6-31G), and upon refinement the more complex set is used (e.g. 6-311G). Calculations with more sophisticated basis sets will generally give more accurate results. Increasing the number and range of basis functions will produce energies that converge towards the exact theoretical energy. Once the Hartree-Fock/6-311G optimised geometry was obtained MP2/6-311G single point energy calculations were carried out. The MP2 method affords a more accurate total energy value. The electronic energies of the geometry optimised molecules are given in **Tables 6.1** and **6.2**. Since the MP2 method was the most extensive method used, it may be expected that the MP2/6-311G calculated results are the more accurate data, and are therefore the values used for comparison.

**Table 6.1** Reaction energies for Schiff base formation between ethylamine, formaldehyde and guanine.[1] Energy change ($\Delta E_{rxn}$) is in units of kJ/mol.

| **Energies (Hartree)** | **HF/MIDI** | **HF/6-31G** | **HF/6-311G** | **MP2/6-311G[2]** |
|---|---|---|---|---|
| Ethylamine | -133.43510 | -134.19279 | -134.22206 | -134.53675 |
| Formaldehyde | -113.17700 | -113.80837 | -113.83952 | -114.07290 |
| Ethylimine | -171.05163 | -172.01475 | -172.05047 | -172.45980 |
| Water | -75.55378 | -75.98536 | -76.01095 | -76.14960 |
| Guanine | -390.22319 | -392.43094 | -392.52785 | -393.36365 |
| Ethyl-aminal | -561.31101 | -564.47262 | -564.59838 | -565.83962 |
| **Reaction 1a** | | | | |
| $\Delta E_{rxn}$ | 17.57 | 2.77 | 0.42 | 0.67 |
| **Reaction 1b** | | | | |
| $\Delta E_{rxn}$ | -95.03 | -70.71 | -52.68 | -42.45 |
| **Overall reaction (1a + 1b)** | | | | |
| $\Delta E_{rxn}$ | -77.46 | -67.94 | -52.26 | -41.78 |

[1] Unless noted, geometry optimised at same level of theory as tabled.
[2] Single-point energy at HF/6-311G optimised geometry.

The data in **Table 6.1** shows that the first step of the reaction is slightly endothermic (+0.7 kJ/mol at the MP2/6-311G level of theory) while the second step is exothermic (-42.5 kJ/mol at the MP2/6-311G level of theory). There is reasonable consistency within the calculated results for the first step, with the exception of results with the minimal MIDI basis set, as may be expected. This result implies **Scheme 6.1a** is a finely balanced reaction, and suggests that it may not be feasible to pre-react the drug with formaldehyde to isolate the imine and subsequently react it with DNA. This also indicates that the drug may have to be activated *in situ* with an excess of formaldehyde in order to mediate drug-DNA adducts.

Since the second part of the reaction (**Scheme 6.1b**) involves a condensation reaction, it may be expected to be entropically unfavourable. This aspect of the reaction is studied in more detail in later calculations. **Scheme 6.1b** is known to undergo an exchange reaction [124], and as this addition uses a primary amine (exocyclic $NH_2$ of guanine), this explains the half-life of drug-DNA adducts.

The overall energy of reaction for **Schemes 6.1a** and **6.1b** is -41.8 kJ/mol (MP2/6-311G level of theory). The fact that this is an overall exothermic process indicates that once the drug-DNA adduct has been formed, it is likely to persist

even though the energy change is quite small. As a comparison, an example of a well known electrophilic addition is the reaction of ethylene with hydrobromic acid to give bromoethane. At 298 K, the enthalpy change (ΔH°) for the bromination reaction is -84.1 kJ/mol [125].

The reaction in **Scheme 6.1a** was attempted using acetaldehyde in place of formaldehyde, but at the time these calculations were performed, it was not feasible to perform calculations on larger molecules. For this reason, semi-empirical calculations were undertaken using AMPAC 7.0 software (see section 6.5) to get a better understanding of the energies involved in larger aldehydes. The validity of the MacGAMESS calculation results was subsequently tested by performing more advanced calculations within the Gaussian 09 software suite.

## 6.4. Gaussian 09 Thermochemistry Calculations

Compound thermochemical methods have been shown to yield accurate reaction thermochemistry [126]. In this work the complete basis-set (CBS-Q) and Gaussian-*n* methods (G3/G4) have been utilised. Calculations for **Schemes 6.1a** and **6.1b** were carried out using the Gaussian 09 software suite [121].

The change in Gibbs free energy (ΔG) of a system for a process at constant temperature is represented by the following equation,

$$\Delta G = \Delta H - T\Delta S \qquad (6.1)$$

where ΔG is the change in Gibbs free energy, ΔH is the change in enthalpy, T is the temperature, and ΔS is the change in entropy. Thus, if $\Delta G < 0$ the forward reaction will proceed spontaneously, if $\Delta G > 0$ the forward reaction is not spontaneous, and if $\Delta G = 0$ the system is at equilibrium. In the case of this study, calculations with two aldehydes and two primary amines were possible and reflected the vast increase in computing power since the first calculations were performed. For each reaction the results from the three compound methods are grouped together. The results from the Gaussian calculations are presented in **Tables 6.2-6.5** for reactions involving ethylamine, propylamine, formaldehyde

and acetaldehyde. For comparison with MacGAMESS results in **Table 6.1**, the model reaction of ethylamine, formaldehyde and guanine is given in **Table 6.2**.

**Table 6.2** Reaction thermochemistry for Schiff base formation between ethylamine, formaldehyde and guanine.

| **Enthalpies (Hartree)** | **CBS-QB3** | **G3** | **G4** |
|---|---|---|---|
| Ethylamine | -134.891326 | -135.028177 | -135.053518 |
| Formaldehyde | -114.340362 | -114.427244 | -114.449380 |
| Ethylimine | -172.908485 | -173.088034 | -173.120325 |
| Water | -76.333710 | -76.378265 | -76.393464 |
| Guanine | -394.326935 | -394.675317 | -394.746345 |
| Ethyl-aminal | -567.254111 | -567.781682 | -567.884553 |
| **Thermochemistry of Reaction 1a** | | | |
| ΔH (kJ/mol) | -27.59 | -28.56 | -28.59 |
| ΔG (kJ/mol) | -26.57 | -24.30 | -24.21 |
| ΔS (J/mol/K) | 3.41 | 14.29 | 14.71 |
| **Thermochemistry of Reaction 1b** | | | |
| ΔH (kJ/mol) | -49.07 | -48.13 | -46.95 |
| ΔG (kJ/mol) | -0.28 | 0.85 | 1.91 |
| ΔS (J/mol/K) | 163.67 | 164.26 | 163.89 |
| **Thermochemistry of Overall Reaction (1a + 1b)** | | | |
| ΔH (kJ/mol) | -76.66 | -76.69 | -75.55 |
| ΔG (kJ/mol) | -26.85 | -23.45 | -22.30 |
| ΔS (J/mol/K) | 167.08 | 178.55 | 178.60 |

The results from the various methods show remarkable consistency and indicate that the overall reaction to form the Schiff base is reasonably exothermic and exergonic, with ΔG of approximately -24 kJ/mol. In each case ΔG for the reaction of the Schiff base with guanine (**1b**, a model for the DNA base) is almost zero. This second reaction has a greater entropy penalty as shown in the large ΔS term, which nearly cancels out the enthalpy gain. These results are shown graphically in **Figure 6.1**, which shows that the energy of the overall reaction (**1a** plus **1b**) is exergonic.

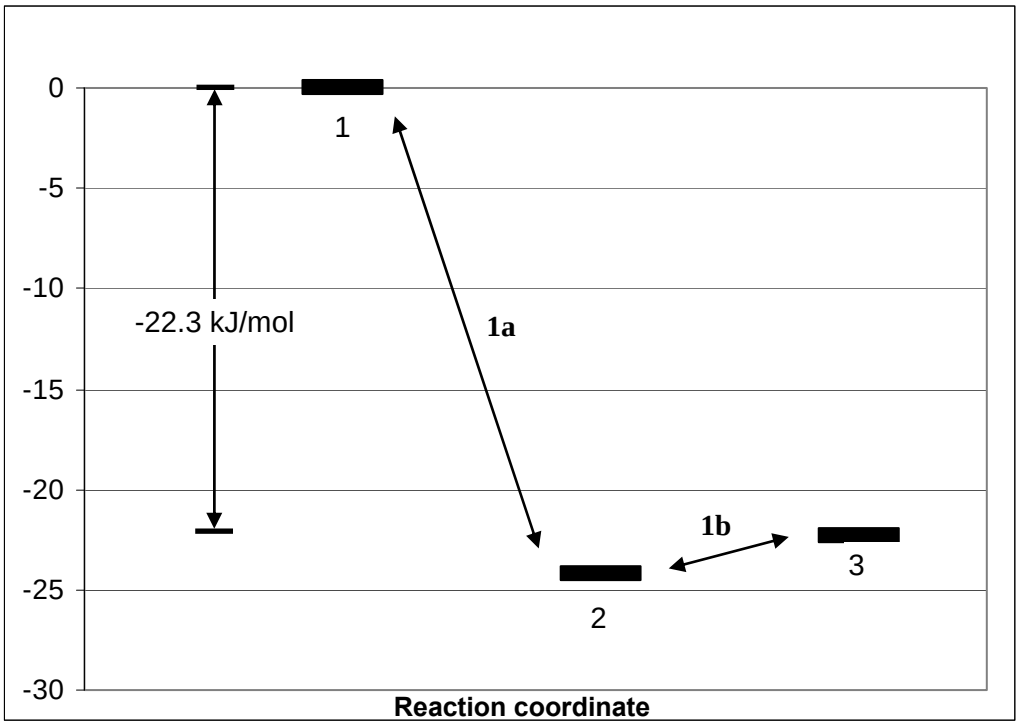



**Figure 6.1** G4 calculated relative Gibbs free energy of Schiff base (**1a**) and consequent aminal formation (**1b**) between ethylamine, formaldehyde and guanine. The y-axis indicates the energy (kJ/mol) relative to the reactants. The x-axis represents the reaction coordinate. Point **1** represents the reactants (formaldehyde, ethylamine and guanine). Point **2** represents the Schiff base, water and guanine. Energy change of **Scheme 6.1a** is the difference between **1** and **2**. Point **3** represents the energy of the aminal and water products. The free energy change of **Scheme 6.1b** is the difference between **2** and **3**. The overall change in reaction free energy (**1** to **3**) is -22.3 kJ/mol.

To explore the part played by the aldehyde in this reaction, formaldehyde was replaced with acetaldehyde, with the results presented in **Table 6.3**. Again, reasonable consistency of results is observed across all methods. Importantly, replacement of formaldehyde with acetaldehyde produces a change in sign for $\Delta G$ of the Schiff base formation (**Scheme 6.1a**). The results indicate that the forward reaction of **Scheme 6.1a** is not favoured. The reaction of the Schiff base with guanine (**Scheme 6.1b**) is predicted to be only moderately exergonic. Thus there is a marked difference between formaldehyde and acetaldehyde in the imine formation reaction, which can be observed by comparing **Figures 6.1** and **6.2**.

This second reaction has a high entropy penalty as evidenced by the large ΔS contribution (**Table 6.3**). It is important to note that ΔG for the overall reaction is negative in sign, implying that if formed, the production of aminal is favoured.

**Table 6.3** Reaction thermochemistry for Schiff base formation between ethylamine, acetaldehyde and guanine.

| **Enthalpies (Hartree)** | **CBS-QB3** | **G3** | **G4** |
|---|---|---|---|
| Ethylamine | -134.891326 | -135.028177 | -135.053518 |
| Acetaldehyde | -153.577613 | -153.709891 | -153.738938 |
| Ethylimine-acet | -212.135718 | -212.360676 | -212.399762 |
| Water | -76.333710 | -76.378265 | -76.393464 |
| Guanine | -394.326935 | -394.675317 | -394.746345 |
| Ethyl-aminal-acet | -606.485574 | -607.058858 | -607.168579 |
| **Thermochemistry of Reaction 1a** | | | |
| ΔH (kJ/mol) | -1.28 | -2.29 | -2.02 |
| ΔG (kJ/mol) | 0.09 | -1.13 | 1.19 |
| ΔS (J/mol/K) | 4.60 | 3.89 | 10.79 |
| **Thermochemistry of Reaction 1** | | | |
| ΔH (kJ/mol) | -60.18 | -60.03 | -59.00 |
| ΔG (kJ/mol) | -5.37 | -3.29 | -4.43 |
| ΔS (J/mol/K) | 183.84 | 190.32 | 183.01 |
| **Thermochemistry of Overall Reaction (1a + 1b)** | | | |
| ΔH (kJ/mol) | -61.46 | -62.32 | -61.02 |
| ΔG (kJ/mol) | -5.28 | -4.42 | -3.24 |
| ΔS (J/mol/K) | 188.44 | 194.21 | 193.80 |

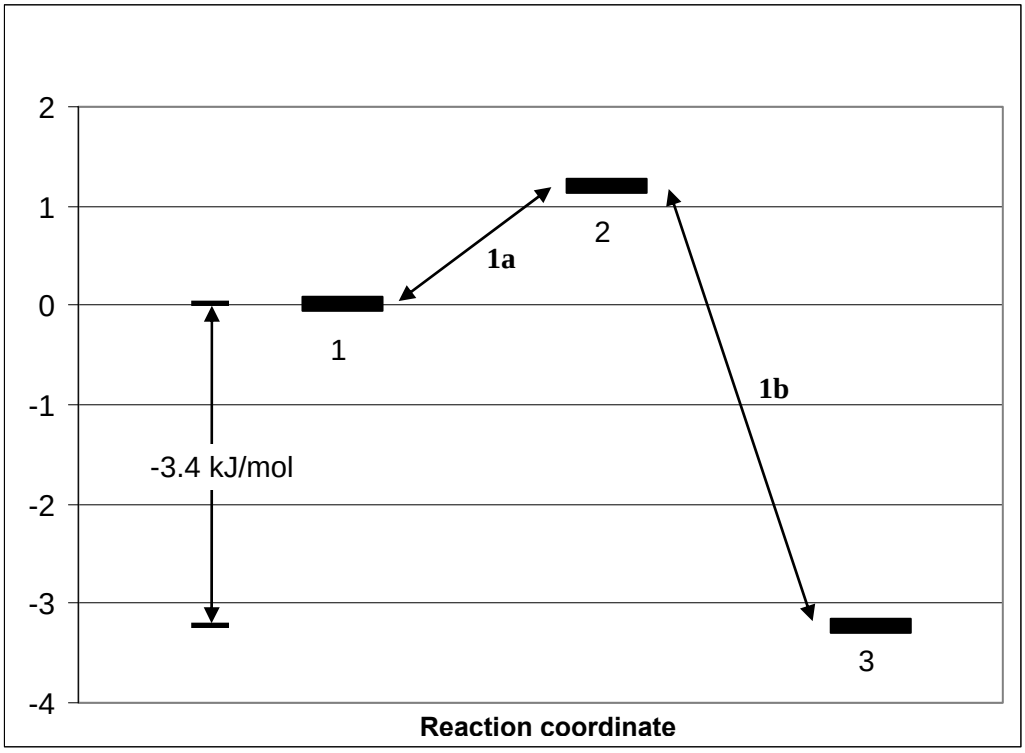



**Figure 6.2** G4 calculated relative Gibbs free energy of Schiff base (**1a**) and consequent aminal formation (**1b**) between ethylamine, acetaldehyde and guanine. See caption of **Figure 6.1** for notation and description. The overall free energy of reaction is -3.4 kJ/mol.

To further explore the importance of the length of the side chain attached to the amine of this reaction, ethylamine was replaced with propylamine. The results are presented in **Table 6.4**. Again, reasonable consistency of results was achieved independent of thermochemical method. The reaction to form the Schiff base is similar to that seen from ethylamine in that the forward reaction is reasonably favoured (ΔG is approximately -15 kJ/mol compared with -24 kJ/mol for ethylamine). The reaction of the Schiff base with guanine (**Scheme 6.1b**) is predicted to be slightly product favoured, as is illustrated in **Figure 6.3**. This second reaction (**Scheme 6.1b**) is predicted to exhibit a large entropy penalty (similar to ethylamine), to the extent that it nearly cancels out the enthalpy gain. However, the reaction does remain product favoured, as illustrated in **Figure 6.3**, with an overall ΔG of -23.0 kJ/mol which is strikingly similar to that for ethylamine (-22.3 kJ/mol).

**Table 6.4** Reaction thermochemistry for Schiff base formation between propylamine, formaldehyde and guanine.

| Enthalpies (Hartree) | CBS-QB3 | G3 | G4 |
|---|---|---|---|
| Propylamine | -174.115457 | -174.297578 | -174.330035 |
| Formaldehyde | -114.340362 | -114.427244 | -114.449380 |
| Propylimine | -212.129735 | -212.354513 | -212.393793 |
| Water | -76.333710 | -76.378265 | -76.393464 |
| Guanine | -394.326935 | -394.675317 | -394.746345 |
| Propyl-aminal | -606.478461 | -607.051268 | -607.161298 |
| **Thermochemistry of Reaction 1a** | | | |
| ΔH (kJ/mol) | -20.02 | -20.89 | -20.59 |
| ΔG (kJ/mol) | -18.35 | -15.76 | -15.23 |
| ΔS (J/mol/K) | 5.59 | 17.19 | 17.96 |
| **Thermochemistry of Reaction 1b** | | | |
| ΔH (kJ/mol) | -57.21 | -56.29 | -55.56 |
| ΔG (kJ/mol) | -9.43 | -8.18 | -7.76 |
| ΔS (J/mol/K) | 160.26 | 161.34 | 160.30 |
| **Thermochemistry of Overall Reaction (1a + 1b)** | | | |
| ΔH (kJ/mol) | -77.23 | -77.17 | -76.14 |
| ΔG (kJ/mol) | -27.78 | -23.94 | -22.99 |
| ΔS (J/mol/K) | 165.85 | 178.53 | 178.26 |

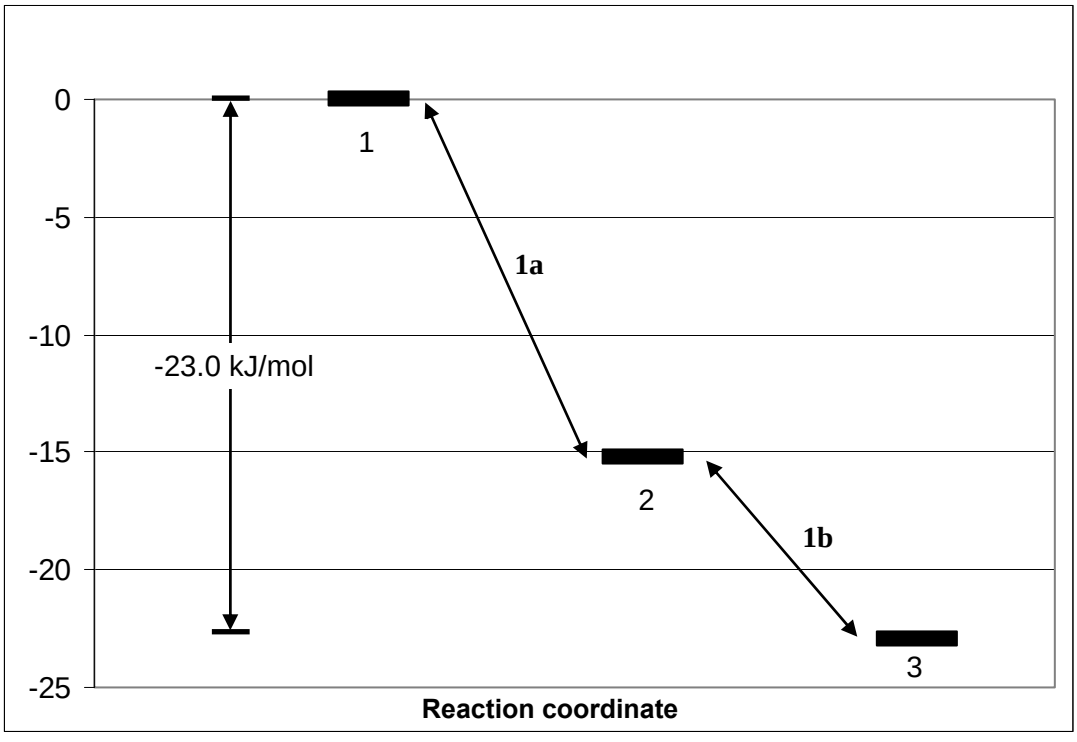



**Figure 6.3** G4 calculated relative Gibbs free energy of Schiff base (**1a**) and consequent aminal formation (**1b**) between propylamine, formaldehyde and guanine. See caption of **Figure 6.1** for notation and description. Overall free energy of reaction is -23.0 kJ/mol.

Finally, the reaction involving propylamine and acetaldehyde was considered, and the results are presented in **Table 6.5**. Again, there is reasonable consistency of results across all methods considered. The reaction to form the Schiff base is similar in energy to the two previous reactions involving formaldehyde in that the forward reaction is reasonably exothermic. Here ΔG is predicted to be approximately -15 kJ/mol, which compares with -24 and -15 kJ/mol for ethylamine and propylamine, respectively. Again, this implies that the forward reaction of **Scheme 6.1a** is favoured. However, the reaction of the Schiff base with guanine is quite different to all other models with a moderately positive value of ΔG. The difference in reaction energetics can be clearly seen by comparing **Figures 6.1** and **6.4**. Again, the entropy penalty for the second reaction (**1b**) is significant. The overall reaction remains exothermic with ΔG being negative (-4.7 kJ/mol compared with -3.3 kJ/mol for the other acetaldehyde reaction), implying that in this case production of the Schiff base is energetically favoured.

**Table 6.5** Reaction thermochemistry for Schiff base formation between propylamine, acetaldehyde and guanine.

| **Enthalpies (Hartree)** | **CBS-QB3** | **G3** | **G4** |
|---|---|---|---|
| Propylamine | -174.115457 | -174.297578 | -174.330035 |
| Acetaldehyde | -153.577613 | -153.709891 | -153.738938 |
| Propylimine-acet | -251.366445 | -251.636440 | -251.683148 |
| Water | -76.333710 | -76.378265 | -76.393464 |
| Guanine | -394.326935 | -394.675317 | -394.746345 |
| Propyl-aminal-acet | -645.710006 | -646.328485 | -646.445426 |
| **Thermochemistry of Reaction 1a** | | | |
| ΔH (kJ/mol) | -18.60 | -19.00 | -20.06 |
| ΔG (kJ/mol) | -14.79 | -13.53 | -14.63 |
| ΔS (J/mol/K) | 12.80 | 18.33 | 18.21 |
| **Thermochemistry of Reaction 1b** | | | |
| ΔH (kJ/mol) | -43.65 | -43.92 | -41.83 |
| ΔG (kJ/mol) | 8.12 | 8.48 | 9.94 |
| ΔS (J/mol/K) | 173.63 | 175.76 | 173.65 |
| **Thermochemistry of Overall Reaction (1a + 1b)** | | | |
| ΔH (kJ/mol) | -62.25 | -62.92 | -61.89 |
| ΔG (kJ/mol) | -6.67 | -5.05 | -4.69 |
| ΔS (J/mol/K) | 186.43 | 194.09 | 191.86 |

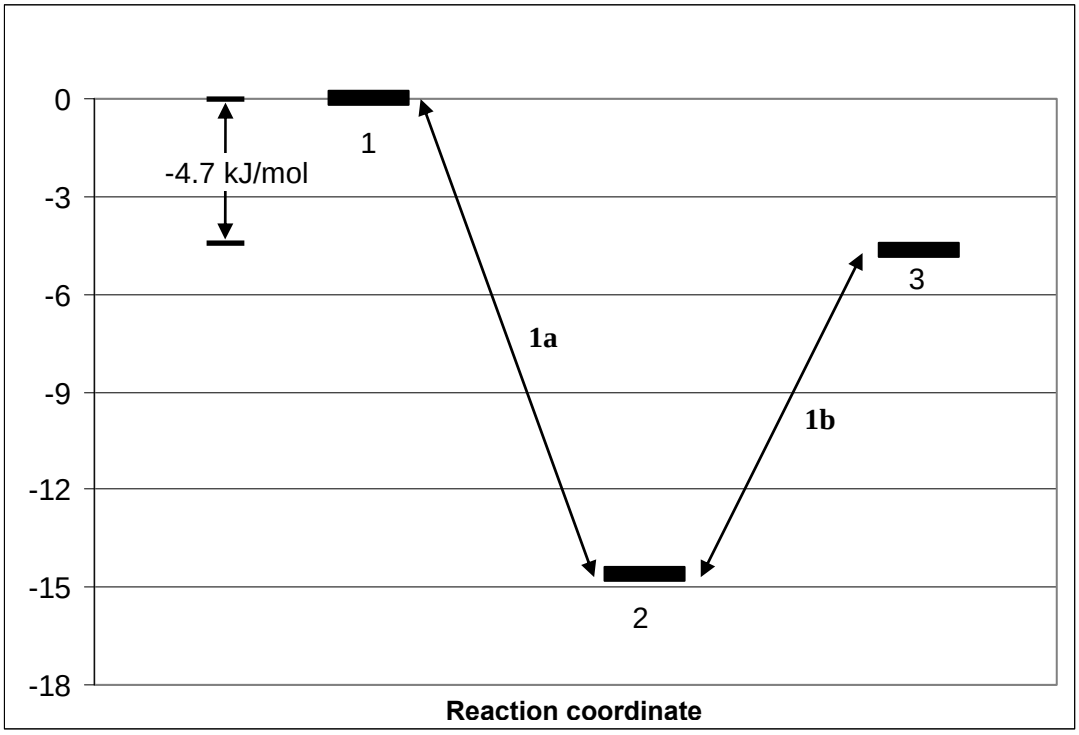


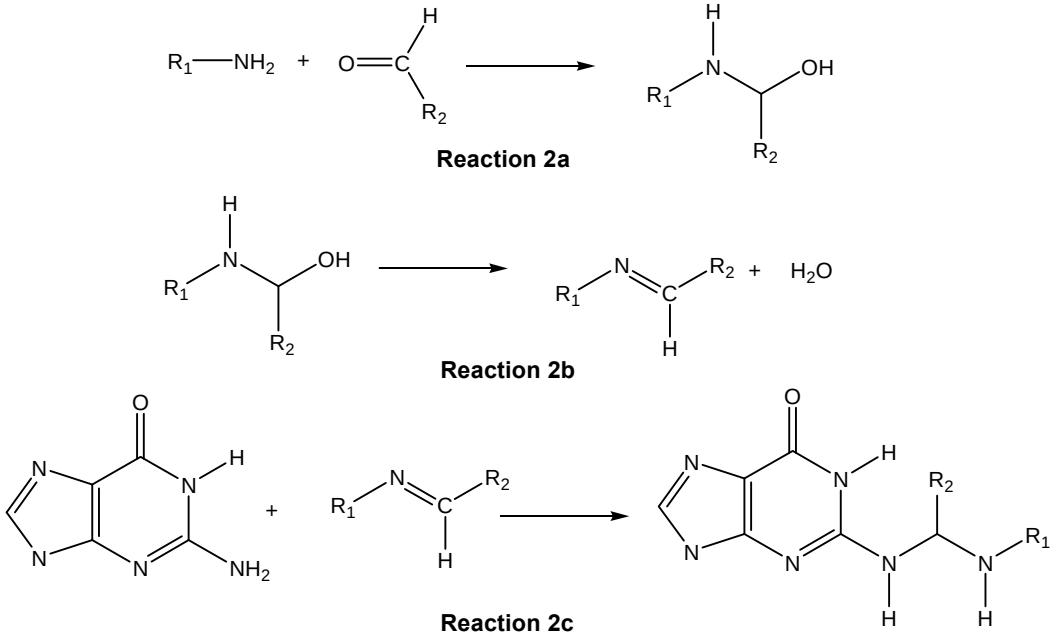

**Figure 6.4** G4 calculated relative Gibbs free energy of Schiff base (**1a**) and consequent aminal formation (**1b**) between propylamine, acetaldehyde and guanine. See caption of **Figure 6.1** for notation and description. Overall free energy of reaction is -4.7 kJ/mol.

Overall the results of these theoretical calculations indicate that the forward reaction is more exothermic and exergonic (product favoured) when using formaldehyde to form the Schiff base compared with acetaldehyde, and that extending the length of the carbon chain of the amine did not affect the energies greatly.

## 6.5. Semi-Empirical AM1 Calculations

Since the energy calculations made using the MacGAMESS software were limited to small molecular systems, it was hoped that additional insight could be obtained using semi-empirical calculations with larger and more realistic molecules. For these calculations, it was necessary to split the reaction **Scheme 6.1a** into two separate reactions, while in **1b** the pseudo-guanine could be replaced by guanine. This resulted in the three main reactions shown in **Scheme 6.2**. The

product from **Reaction 2a** and the reactant from **Reaction 2b** is the intermediate from **Scheme 6.1a**, which was not investigated in the *ab initio* calculation.




**Scheme 6.2** Three reactions considered for AM1 calculations. $R_1$ = methyl – pentyl. $R_2$ = hydrogen (formaldehyde) or methyl (acetaldehyde).

### 6.5.1. Methodology

Semi-empirical AM1 calculations were performed using the AMPAC 7.0 software [122]. Molecules for the reactions were built from scratch (no templates were used). Reactants and products were built and optimised separately. It is a requirement of the program that the numbering of each atom is consistent throughout all three reactions. No torsional bond angles close to 180° are allowed as they give a poorly defined dihedral angle, as shown in **Figure 6.5**.

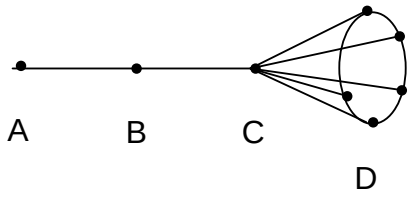
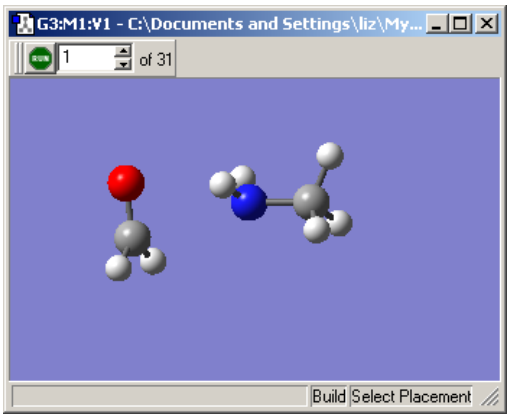


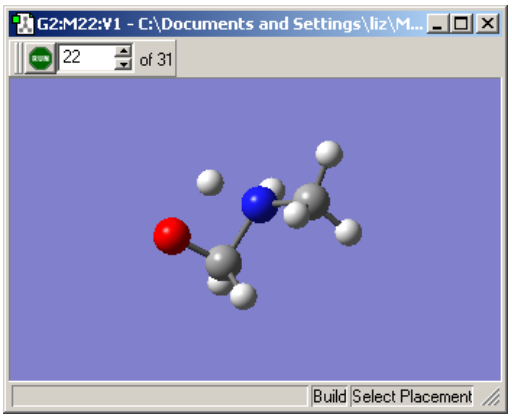

**Figure 6.5** Example of a poor dihedral angle. **D**, which cannot be defined by an **A B C** model since **D** could be anywhere around the ellipse.

Once the structure was optimised, the output file (.out file) was used to create a chain input file (.chn file). The chain input file was used to obtain the overall reaction energy and hence the energy change. After the chain reaction was completed a chain coordinate plot was generated showing the energy at each point. The .arc file allowed visualisation at each step in the reaction. Enthalpies of the reactant, product and transition state were recorded (all energies are quoted in kcal/mol (1 kcal/mol = 4.184 kJ/mol) [127]). If the .arc file showed a problem such as an atom spontaneously breaking off a molecule, more transition states were inserted into the chain file, and the file was re-run. It was possible to have up to five slides (or transition states) per chain file to essentially push the forward reaction.

Initially a test model calculation was performed on a simple model system, $NH_3 + H_2C{=}O \rightarrow NH{=}CH_2 + H_2O$, to ensure the reaction path was being correctly modelled. The amine side chain was successively increased from methylamine ($CH_3NH_2$) to pentylamine ($CH_3(CH_2)_4NH_2$) in order to elucidate any trends. **Figure 6.6** shows screen captures from AMPAC of a chain file for **Reaction 2a**. Each point along the chain coordinate plot (**Figure 6.7**) corresponds to a different step along the reaction pathway (**Figures 6.6** (**a-c**)), which allows visualisation of the reaction mechanism. The enthalpies (kcal/mol) for the first, transition state, and final points of the chain coordinate plot are given in **Figure 6.7** for **Reaction 2a**.

**(a)** **(b)**




**(c)**

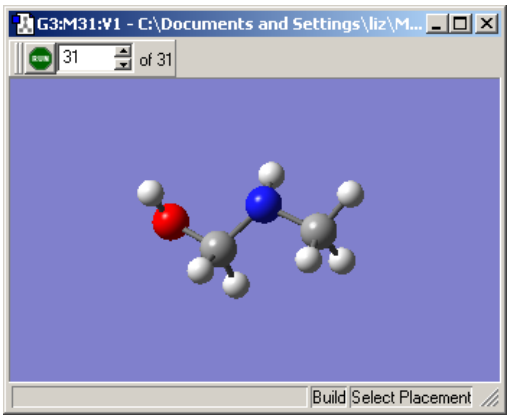



**Figure 6.6** Representation of **Reaction 2a** (**Scheme 6.2**) where **(a)** is the methylamine and formaldehyde reactants, **(b)** is the transition state, and **(c)** is the product.

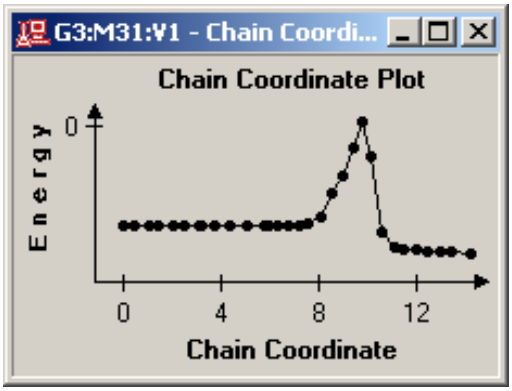



**Figure 6.7** Chain coordinate plot for the reaction of methylamine and formaldehyde to afford the product of **Reaction 2a**. Energy in units of kcal/mol.

**Figure 6.6** (**a-c**) illustrates the pathway of **Reaction 2a**, (**Scheme 6.2**), and represent the initial energy, highest energy and final energy from the chain coordinate plot in **Figure 6.7**, which corresponds to the reactants, transition state and products of the reaction. The chain coordinate plot provides the enthalpy of the reactants (-173.7 kJ/mol), the transition state (11.3 kJ/mol) and the aminal product (-222.1 kJ/mol), indicating that the reaction is exothermic. Once calculations of the initial reaction succeeded, they were re-run separately using the two model solvents of methanol and water.

### 6.5.2. Reaction of Alkylamines with Formaldehyde

Results for **Reaction 2a** and **2b** (formation of the Schiff base) are given in **Tables 6.6** and **6.7** respectively, and indicate that the enthalpy of reaction for the Schiff base formation is approximately zero. This is in agreement with results from MacGAMESS and Gaussian 09, in that the reaction is very finely balanced.

The results from **Reaction 2c** (**Scheme 6.2**, imine reacting with DNA model) are given in **Table 6.8**. The results are not directly comparable to those from the MacGAMESS calculations, although both sets of results suggest an overall exothermic and exergonic process (for the ethylamine model, $\Delta H$ of **1a** +0.7 kJ/mol, **1b** -42.5 kJ/mol, **1b**+**1a** -41.8 kJ/mol compared with **2a**+**2b** -4.2 kJ/mol, **2c** -38.6 kJ/mol, **2c**-(**2a**+**2b**) -34.4 kJ/mol). Some problems occurred with the program when the molecules were large (the guanine from **Reaction 2c**). For example, protons were removed or moved randomly, or the whole molecule collapsed onto itself. Some of these problems were overcome by using more transition states in the chain file (up to five total reaction steps to a chain). However, in **Reaction 2c** the upper limit proved to be for $R_1$ = ethyl.

**Table 6.6** Enthalpy of reaction (kJ/mol) for **Reaction 2a** (formation of adduct where $R_2$ = H, formaldehyde).[1]

| $R_1$ Group | Solvent | Reactants | TS | Products | $E_A$ (fwd) | $E_A$ (rev) | $\Delta H°_{rxn}$ |
|---|---|---|---|---|---|---|---|
| methyl | Gas phase | -173.72 | 11.34 | -222.04 | 185.06 | 233.38 | **-48.33** |
| | Methanol | -198.99 | -42.34 | -255.85 | 156.65 | 213.51 | **-56.86** |
| | Water | -204.35 | -48.33 | -267.69 | 156.02 | 219.37 | **-63.35** |
| ethyl | Gas phase | -206.02 | -17.78 | -247.53 | 188.24 | 229.74 | **-41.51** |
| | Methanol | -233.68 | -71.80 | -283.68 | 161.88 | 211.88 | **-50.00** |
| | Water | -234.39 | -74.85 | -292.21 | 159.54 | 217.36 | **-57.82** |
| propyl | Gas phase | -234.81 | -45.90 | -274.47 | 188.91 | 228.57 | **-39.66** |
| | Methanol | -233.68 | -71.80 | -283.68 | 161.88 | 211.88 | **-50.00** |
| | Water | -264.85 | -101.59 | -318.57 | 163.26 | 216.98 | **-53.72** |
| butyl | Gas phase | -263.30 | -74.52 | -303.17 | 188.78 | 228.66 | **-39.87** |
| | Methanol | -295.64 | -131.75 | -342.80 | 163.89 | 211.04 | **-47.15** |
| | Water | -291.50 | -129.29 | -346.23 | 162.21 | 216.94 | **-54.73** |
| pentyl | Gas phase | -263.42 | -100.29 | -301.46 | 163.13 | 201.17 | **-38.03** |
| | Methanol | -296.31 | -159.08 | -342.21 | 137.24 | 183.13 | **-45.90** |
| | Water | -289.11 | -153.97 | -342.75 | 135.14 | 188.78 | **-53.64** |

[1] $R_1$ indicates the length of the alkyl group of the alkylamine. $E_A$ is the activation enthalpy.

**Table 6.7** Enthalpy of reaction (kJ/mol) for **Reaction 2b** (elimination of water to form imine where $R_2$ = H, formaldehyde).[1]

| $R_1$ Group | Solvent | Reactants | TS | Products | $E_A$ (fwd) | $E_A$ (rev) | $\Delta H°_{rxn}$ |
|---|---|---|---|---|---|---|---|
| methyl | Gas phase | -222.04 | 75.48 | -187.61 | 297.52 | 263.09 | **34.43** |
| | Methanol | -264.26 | 34.89 | -220.25 | 299.16 | 255.14 | **44.02** |
| | Water | -275.89 | 26.99 | -221.12 | 302.88 | 248.11 | **54.77** |
| ethyl | Gas phase | -247.53 | 48.07 | -210.20 | 295.60 | 258.28 | **37.32** |
| | Methanol | -283.68 | 6.99 | -244.81 | 290.66 | 251.79 | **38.87** |
| | Water | -292.84 | 2.05 | -243.34 | 294.89 | 245.39 | **49.50** |
| propyl | Gas phase | -274.47 | 74.89 | -237.02 | 349.36 | 311.92 | **37.45** |
| | Methanol | -312.38 | -21.09 | -273.88 | 291.29 | 252.80 | **38.49** |
| | Water | -318.57 | -23.22 | -269.07 | 295.35 | 245.85 | **49.50** |
| butyl | Gas phase | -303.17 | -2.13 | -265.73 | 301.04 | 263.59 | **37.45** |
| | Methanol | -342.50 | -48.41 | -303.55 | 294.09 | 255.14 | **38.95** |
| | Water | -345.43 | -47.91 | -295.31 | 297.52 | 247.40 | **50.12** |
| pentyl | Gas phase | -328.86 | -27.74 | -291.37 | 301.12 | 263.63 | **37.49** |
| | Methanol | -370.12 | -75.40 | -291.37 | 294.72 | 215.98 | **78.74** |
| | Water | -370.91 | -72.26 | -321.21 | 298.65 | 248.95 | **49.71** |

[1] $R_1$ indicates the length of the alkyl group of the alkylamine. $E_A$ is the activation enthalpy.

**Table 6.8** Enthalpy of reaction (kJ/mol) for **Reaction 2c** (reaction of imine with guanine to form aminal where $R_2$ = H, formaldehyde).[1]

| $R_1$ Group | Solvent | Reactants | TS | Products | $E_A$ (fwd) | $E_A$ (rev) | $\Delta H°_{rxn}$ |
|---|---|---|---|---|---|---|---|
| methyl | Gas phase | 259.66 | 542.04 | 225.10 | 282.38 | 316.94 | **-34.56** |
| | Methanol | 194.26 | 453.09 | 149.79 | 258.82 | 303.30 | **-44.48** |
| | Water | 189.07 | 445.09 | 143.01 | 256.02 | 302.08 | **-46.07** |
| ethyl | Gas phase | 236.06 | 513.13 | 197.48 | 277.06 | 315.64 | **-38.58** |
| | Methanol | - | - | - | - | - | - |
| | Water | 144.64 | 423.50 | 106.82 | 278.86 | 316.69 | **-37.82** |

[1] $R_1$ indicates the length of the alkyl group of the alkylamine. $E_A$ is the activation enthalpy.

Enthalpies of reaction for the overall reaction between amine and formaldehyde to form the DNA adduct are listed in **Table 6.9**. These results suggest that the overall reaction is an exothermic process. This implies that once the Schiff base has formed and has reacted with the guanine, the product is reasonably stable and that the reverse reaction will be slow. Hence, if the drug attaches covalently to DNA, it will most likely remain attached, or at least have a lifetime long enough to show NOE peaks to the DNA in NMR experiments satisfying the slow exchange conditions.

**Table 6.9** Total enthalpy for each reaction with formaldehyde (**2a** to **2c**).[1]

| $R_1$ Group | Solvent | 2a[2] | 2b[2] | 2c[2] | 2a+2b[2] | Overall[2] |
|---|---|---|---|---|---|---|
| methyl | Gas phase | -48.33 | 34.43 | -34.56 | -13.89 | **-20.67** |
| | Methanol | -56.86 | 44.02 | -44.48 | -12.84 | **-31.63** |
| | Water | -63.35 | 54.77 | -46.07 | -8.58 | **-37.49** |
| ethyl | Gas phase | -41.51 | 37.32 | -38.58 | -4.18 | **-34.39** |
| | Methanol | | | | | |
| | Water | -57.82 | 49.50 | -37.82 | -8.33 | **-29.50** |

[1] $R_1$ indicates the length of the alkyl group of the alkylamine.
[2] $\Delta H°$ in kJ/mol for each reaction.

### 6.5.3. Reaction of Alkylamines with Acetaldehyde

Once the process had been completed for formaldehyde, it was repeated using acetaldehyde in place of formaldehyde (i.e. $R_2 = CH_3$). The results for **Reaction 2a**, (**Scheme 6.2**) are listed in **Table 6.10**, with those of **Reaction 2b** given in **Table 6.11**, and of **Reaction 2c** in **Table 6.12**. The results from **Tables 6.10** and **6.11** are similar to the energetics of the reaction with formaldehyde, in that the reaction is very finely balanced. The Schiff base formation (**Reaction 2a** + **2b**) is very slightly exothermic, with a $\Delta H$ of 2.8 kJ/mol for ethylamine. These calculations were too large for the MacGAMESS software package available at the time, so the AMPAC calculations were employed to allow some insight into using acetaldehyde.

**Table 6.10** Enthalpy of reaction (kJ/mol) for **Reaction 2a** (formation of adduct where $R_2 = CH_3$, acetaldehyde).[1]

| $R_1$ Group | Solvent | Reactants | TS | Products | $E_A$ (fwd) | $E_A$ (rev) | $\Delta H°_{rxn}$ |
|---|---|---|---|---|---|---|---|
| methyl | Gas phase | -216.56 | -12.89 | -237.61 | 203.68 | 224.72 | **-21.05** |
| | Methanol | -247.07 | -66.73 | -272.59 | 180.33 | 205.85 | **-25.52** |
| | Water | -248.70 | -70.12 | -281.37 | 178.57 | 211.25 | **-32.68** |
| ethyl | Gas phase | -248.74 | -41.80 | -262.80 | 206.94 | 221.00 | **-14.06** |
| | Methanol | -282.17 | -96.19 | -299.83 | 185.98 | 203.64 | **-17.66** |
| | Water | -281.04 | -96.73 | -305.56 | 184.31 | 208.82 | **-24.52** |
| propyl | Gas phase | -276.19 | -70.46 | -290.54 | 205.73 | 220.08 | **-14.35** |
| | Methanol | -313.34 | -126.40 | -329.95 | 186.94 | 203.55 | **-16.61** |
| | Water | -309.49 | -124.18 | -333.13 | 185.31 | 208.95 | **-23.64** |
| butyl | Gas phase | -301.58 | -99.04 | -319.82 | 202.55 | 220.79 | **-18.24** |
| | Methanol | -341.37 | -156.52 | -360.45 | 184.85 | 203.93 | **-19.08** |
| | Water | -333.80 | -151.71 | -360.66 | 182.09 | 208.95 | **-26.86** |
| pentyl | Gas phase | -335.10 | -127.70 | -319.82 | 207.40 | 192.13 | **15.27** |
| | Methanol | -374.22 | -187.32 | -390.87 | 186.90 | 203.55 | **-16.65** |
| | Water | -364.80 | -179.54 | -388.36 | 185.27 | 208.82 | **-23.56** |

[1] $R_1$ indicates the length of the alkyl group of the alkylamine. $E_A$ is the activation enthalpy.

**Table 6.11** Enthalpy of reaction (kJ/mol) for **Reaction 2b** (elimination of water to form imine where $R_2 = CH_3$, acetaldehyde).[1]

| $R_1$ Group | Solvent | Reactants | TS | Products | $E_A$ (fwd) | $E_A$ (rev) | $\Delta H°_{rxn}$ |
|---|---|---|---|---|---|---|---|
| methyl | Gas phase | -237.61 | -99.58 | -223.43 | 138.03 | 123.85 | **14.18** |
| | Methanol | -272.59 | 7.49 | -261.12 | 280.08 | 268.61 | **11.46** |
| | Water | -281.37 | 2.68 | -259.20 | 284.05 | 261.88 | **22.18** |
| ethyl | Gas phase | -262.80 | -86.15 | -245.94 | 176.65 | 159.79 | **16.86** |
| | Methanol | -299.83 | -15.73 | -285.89 | 284.09 | 270.16 | **13.93** |
| | Water | -305.56 | -17.70 | -281.08 | 287.86 | 263.38 | **24.48** |
| propyl | Gas phase | -291.33 | 57.70 | -274.72 | 349.03 | 332.42 | **16.61** |
| | Methanol | -330.12 | -46.19 | -316.35 | 283.93 | 270.16 | **13.77** |
| | Water | -333.09 | -45.40 | -308.78 | 287.69 | 263.38 | **24.31** |
| butyl | Gas phase | -319.82 | -27.24 | -303.21 | 292.59 | 275.98 | **16.61** |
| | Methanol | -360.45 | -76.27 | -346.56 | 284.18 | 270.29 | **13.89** |
| | Water | -360.66 | -72.30 | -336.23 | 288.36 | 263.93 | **24.43** |
| pentyl | Gas phase | -348.44 | -55.90 | -331.83 | 292.55 | 275.93 | **16.61** |
| | Methanol | -388.32 | -100.25 | -363.88 | 288.07 | 263.63 | **24.43** |
| | Water | -388.32 | -100.25 | -363.88 | 288.07 | 263.63 | **24.43** |

[1] $R_1$ indicates the length of the alkyl group of the alkylamine. $E_A$ is the activation enthalpy.

**Table 6.12** Enthalpy of reaction (kJ/mol) for **Reaction 2c** (reaction of imine with guanine to form aminal where $R_2 = CH_3$, acetaldehyde).[1]

| $R_1$ Group | Solvent | Reactants | TS | Products | $E_A$ (fwd) | $E_A$ (rev) | $\Delta H°_{rxn}$ |
|---|---|---|---|---|---|---|---|
| methyl | Gas phase | 224.68 | 528.23 | 218.61 | 303.55 | 309.62 | **-6.07** |
| | Methanol | 133.68 | 436.73 | 132.80 | 303.05 | 303.93 | **-0.88** |
| | Water | 125.85 | 427.98 | 127.99 | 302.13 | 299.99 | **2.13** |
| ethyl | Gas phase | 201.92 | 503.38 | 195.77 | 301.46 | 307.61 | **-6.15** |
| | Methanol | 102.05 | 403.92 | 103.97 | 301.88 | 299.95 | **1.92** |
| | Water | 104.60 | 403.92 | 103.97 | 299.32 | 299.95 | **-0.63** |
| propyl | Gas phase | 171.88 | 474.21 | 167.15 | 302.34 | 307.06 | **-4.73** |
| | Methanol | 79.16 | 379.11 | 75.48 | 299.95 | 303.63 | **-3.68** |
| | Water | 77.07 | 375.81 | 76.23 | 298.74 | 299.57 | **-0.84** |
| butyl | Gas phase | 143.43 | 445.68 | 138.70 | 302.25 | 306.98 | **-4.73** |
| | Methanol | 48.79 | 348.86 | 45.19 | 300.08 | 303.67 | **-3.60** |
| | Water | 49.83 | 348.36 | 48.74 | 298.53 | 299.62 | **-1.09** |
| pentyl | Gas phase | 110.04 | 417.06 | 110.04 | 307.02 | 307.02 | **0.00** |
| | Methanol | 18.37 | 318.53 | 14.69 | 300.16 | 303.84 | **-3.68** |
| | Water | 22.09 | 320.79 | 21.00 | 298.70 | 299.78 | **-1.09** |

[1] $R_1$ indicates the length of the alkyl group of the alkylamine. $E_A$ is the activation enthalpy.

Enthalpies of reaction for the overall reaction between amine and acetaldehyde ($R_2 = CH_3$) to form the DNA adduct are collected in **Table 6.13**. The last column gives the total energy change for the overall reaction (**2a** to **2c**) using acetaldehyde. This indicates that the formation of the Schiff base and consequent reaction with guanine is relatively enthalpically neutral. This implies that neither the forward nor the backward reaction is strongly favoured, indicating that acetaldehyde would not be an energetically favourable replacement for formaldehyde in these reactions. The total enthalpy is more negative as the alkyl group is extended, although this is not a consistent trend. Again, though not directly comparable, these results also support the *ab initio* methods outlined above, which suggest that the overall reaction is less favoured with acetaldehyde compared to formaldehyde. The results also indicate that solvent effects on this reaction are not large.

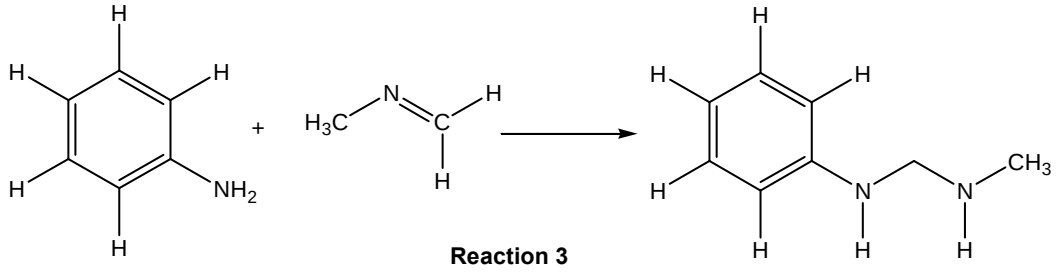
**Table 6.13** Total enthalpy for each reaction **2a** to **2c** with acetaldehyde ($R_2$ = $CH_3$).[1]

| $R_1$ Group | Solvent | 2a[2] | 2b[2] | 2c[2] | 2a+2b[2] | Overall[2] |
|---|---|---|---|---|---|---|
| methyl | Gas phase | -21.05 | 14.18 | -6.07 | -6.86 | **0.79** |
| | Methanol | -25.52 | 11.46 | -0.88 | -14.06 | **13.18** |
| | Water | -32.68 | 22.18 | 2.13 | -10.50 | **12.64** |
| ethyl | Gas phase | -14.06 | 16.86 | -6.15 | 2.80 | **-8.95** |
| | Methanol | -17.66 | 13.93 | 1.92 | -3.72 | **5.65** |
| | Water | -24.52 | 24.48 | -0.63 | -0.04 | **-0.59** |
| propyl | Gas phase | -14.35 | 16.61 | -4.73 | 2.26 | **-6.99** |
| | Methanol | -16.61 | 13.77 | -3.68 | -2.85 | **-0.84** |
| | Water | -23.64 | 24.31 | -0.84 | 0.67 | **-1.51** |
| butyl | Gas phase | -18.24 | 16.61 | -4.73 | -1.63 | **-3.10** |
| | Methanol | -19.08 | 13.89 | -3.60 | -5.19 | **1.59** |
| | Water | -26.86 | 24.43 | -1.09 | -2.43 | **1.34** |
| pentyl | Gas phase | 15.27 | 16.61 | 0.00 | 31.88 | **-31.88** |
| | Methanol | -16.65 | 24.43 | -3.68 | 7.78 | **-11.46** |
| | Water | -23.56 | 24.43 | -1.09 | 0.88 | **-1.97** |

[1] $R_1$ indicates the length of the alkyl group of the alkylamine.
[2] $\Delta H°$ in kJ/mol for each reaction.

### 6.5.4. Reaction of Aniline and Formaldehyde

A modification of **Reaction 2c** (**Scheme 6.3**) was tested due to some problems associated with the calculations of larger molecules (guanine). The reaction using aniline in place of guanine was only partially finished, with results shown in **Table 6.14**. Due to limitations of the program (i.e. spontaneous proton removal and molecule collapse), purchasing Version 7.0 of AMPAC to complete these reactions was not justified.




**Scheme 6.3 Reaction 3**, a modification of **2c** using aniline in place of guanine.

**Table 6.14** Enthalpy of reaction (kJ/mol) for **Reaction 3**, **Scheme 6.3** (formaldehyde and aniline).[1]

| $R_1$ Group | Solvent | Reactants | TS | Products | $E_A$ (fwd) | $E_A$ (rev) | $\Delta H°_{rxn}$ |
|---|---|---|---|---|---|---|---|
| methyl | Gas phase | 149.58 | 415.89 | 112.17 | 266.31 | 303.72 | **-37.40** |
| | Methanol | 109.41 | 267.82 | 69.58 | 158.41 | 198.24 | **-39.83** |
| | Water | 116.23 | 376.43 | 78.74 | 260.20 | 297.69 | **-37.49** |
| ethyl | Gas phase | 121.88 | 387.02 | 84.56 | 265.14 | 302.46 | **-37.32** |
| | Methanol | 87.91 | 256.69 | 41.59 | 168.78 | 215.10 | **-46.32** |
| | Water | 94.35 | 255.06 | 53.51 | 160.71 | 201.54 | **-40.84** |
| propyl | Gas phase | 93.18 | 357.86 | 55.69 | 264.68 | 302.17 | **-37.49** |
| | Methanol | 53.72 | 311.96 | 11.21 | 258.24 | 300.75 | **-42.51** |
| | Water | 66.65 | 327.90 | 25.94 | 261.25 | 301.96 | **-40.71** |

[1] $R_1$ indicates the length of the alkyl group of the alkylamine. $E_A$ is the activation enthalpy.

The results in **Table 6.14** suggest that there were not any significant changes in energy within the first two increases in chain length, hence the methyl, ethyl, and propyl results are significant in determining the range of reaction energy. The formation of the Schiff base is relatively entropically neutral, which may suggest an equilibrium between amine and imine formation. The effect of using a solvent compared to gas phase calculations was that the enthalpy was slightly less. For example, methylamine in the gas phase was 149.6 kJ/mol, compared with 109.4 kJ/mol for methanol solvent (difference of 40.2 kJ/mol). It had little effect on the $\Delta H°_{rxn}$ (difference of 2.4 kJ/mol).

## 6.6. Conclusions

Although all of the calculations used reactions between model compounds out of necessity, and the majority of the calculations were done in the gas phase, significant and useful results were obtained. One issue that could not be addressed related to the possibility that an intercalating drug containing a free amine may be able to organise its structure before aminal formation *in vivo*, and thus may overcome some entropy disadvantage.

In comparing all of the calculations, it was evident that the more advanced thermodynamic calculation methods using the Gaussian 09 software had an internal consistency, even though they were very computer time intensive. For comparison, heats of fusion and vaporization are usually of the order of 10 kJ/mol, bond energies are of the order of 100 kJ/mol, and ionization energies of the order of 1000 kJ/mol [128]. The reactions with acetaldehyde in place of formaldehyde are not as energetically favourable showing that future studies with acetaldehyde may not be productive.

The MacGAMESS calculations showed that the change in electronic energy for Schiff base formation is very slightly endothermic (+0.7 kJ/mol for ethylamine at the MP2/6-311G level of theory), and that formation of the aminal is considerably exothermic (-42.5 kJ/mol for ethylamine at the MP2/6-311G level of theory). However, it was expected that **Scheme 6.1b** would be unfavourable due to the fact that the aminal formation involves a condensation reaction. It is clear from these results that the change in electronic energy alone is not a good prediction of the enthalpy of a reaction. As the energies from the MacGAMESS calculations were purely electronic energy, they did not provide the enthalpy, entropy or Gibbs free energy. Therefore the use of more sophisticated methods was required in order to determine the reaction enthalpy and entropy.

Results from the G3 and G4 methods are considered to be the most reliable of those considered, and showed that the overall reactions were exothermic and exergonic. The majority of this energy comes from the Schiff base formation, with a smaller amount from the aminal formation. In the case of ethylamine, formaldehyde and guanine (G4 level of theory), the $\Delta H$ for Schiff base formation was -28.6 kJ/mol, and the $\Delta H$ of aminal formation was -47.0 kJ/mol. The overall reaction resulted in a $\Delta H$ of -75.6 kJ/mol, $\Delta G$ of -22.3 kJ/mol and $\Delta S$ of 178.6

J/mol/K. The overall reaction involving formaldehyde was more favourable in both enthalpy and entropy when compared to that of acetaldehyde (ΔH of -61.0 kJ/mol, ΔG of -3.3 kJ/mol and ΔS of 193.8 J/mol/K).

The results from the AMPAC AM1 calculations for forming a Schiff base between ethylamine and formaldehyde also suggested that the overall reaction is an exothermic process. Schiff base formation between methylamine and formaldehyde, and consequent reaction with guanine in the gas phase (**Reaction 2a – 2c**, **Scheme 6.2**) had an enthalpy of reaction of -20.7 kJ/mol, and in water solvent it was -37.5 kJ/mol. For ethylamine and formaldehyde it was -34.4 and -29.5 kJ/mol, respectively. An increase in the length of the side chain in water did not significantly change the difference in overall energy (only 8 kJ/mol). This implies that for all side chain lengths that once the Schiff base has formed and reacted with DNA *in vitro*, the reverse reaction is possible, but not likely on energetic arguments.

When using aniline in place of guanine (**Scheme 6.3**), the subsequent reaction with the $NH_2$ of pseudo-guanine to form the methylene mediated bridge proved to be relatively energetically neutral. As the size of the side chain of the Schiff base was increased, the difference in energy between the reactants and products was reduced to the point of being energetically neutral. For the gas phase, there was only a difference of 0.04 kJ/mol when the side chain was between one and three carbons long. In comparison, the difference was 0.80 kJ/mol when using water as solvent. The trend was that the products released slightly more energy with a longer side chain, again implying that energies would be more favourable for the forward reaction *in vitro*. With longer chain amines the balance between the two reactions is altered, but the overall reaction is still exothermic. The solvent effect calculations using the AMPAC 7.0 method showed that the solvent did not have a large effect on the energies of the reaction.

The important result is that to varying degrees, all three of these methods showed that the Schiff base formation is exothermic and exergonic, and that the aminal formation is almost neutral. Thus the overall reaction is exothermic and should proceed in the forward direction. These energetic calculations suggest that if the drug has the opportunity to attach covalently to DNA, it may have a limited bound lifetime due to the minimal energies involved.

# CHAPTER 7. NMR STUDIES

A significant aim of the NMR studies discussed in this chapter was to obtain NOE data on a formaldehyde mediated drug-DNA covalently bound complex and to use that data to compute a 3D molecular structure of that complex. Such a structure would have confirmed the covalent bond and intercalating ability of the drug, and identified the binding site to DNA. This aim was ultimately unsuccessful, but the NMR data has been included for completion. Part of the reason for the unsuccessful result is that formaldehyde proved to be a much more complex molecule than initially imagined, consisting of a very complex mixture of hexamers, dimers, trimers, cyclic and non-cyclic species. This complexity, and the use of a significant molar excess of formaldehyde in preparing the final complex, proved to be a significant issue with the final NMR sample as discussed below.

## 7.1. Methods

### 7.1.1. NMR

The 1D and 2D spectra discussed in this chapter were recorded on a 400 MHz Bruker Avance DRX Nuclear Magnetic Resonance Spectrometer ($^{1}H$ at 400.100 MHz, $^{13}C$ at 100.5877 MHz and $^{31}P$ at 161.9270 MHz) using XWIN-NMR software Version 3.5 at 300 K. The Bruker standard pulse sequences were employed unless otherwise stated. An overview on each method is given in the introduction of **Chapter 2**. All samples were dissolved in $H_2O/D_2O$ and hence water suppression techniques were employed throughout this chapter. The deuterated solvents were purchased from Cambridge Isotopes, and the 25 x PBS was purchased from Amresco and used undiluted.

### 7.1.2. Selection of the DNA sequence: Working with d(GACGTC)

Analysis of the information from the transcription assay data (**Table 1** in **Chapter 5**) showed that the formaldehyde activated C4-DACA attaches at CG and CA rich sites. There were three sites with ACG as high occupancy sites. Thus

for NMR studies where assignments are required of the proton signals of the DNA a short sequence was selected meeting the following requirements. The sequence had to be palindromic so that the apparent concentration of DNA would be essentially doubled under NMR conditions. However, once the formaldehyde-activated drug was attached, the palindromic nature would be lost, and the apparent DNA concentration would be less than the true concentration. The remaining base pairs were chosen to be G or C to increase the proclivity for double stranded DNA in solution. Thus there were two possible oligonucleotides chosen for NMR studies:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 6-mer | 5’ | G | A | C | G | T | C | 3’ | |
| 8-mer | 5’ | G | G | A | C | G | T | C | C | 3’ |

In this case the simpler hexamer, 5’ GACGTC 3’, was selected for NMR studies. This was purchased from GeneWorks Pty Ltd (10 μmole x 2, HPLC grade). PBS was purchased from Astral as 25 x buffer at pH 7.4.

### 7.1.3. DNA NMR Sample Preparation

Although GeneWorks was a preferred supplier of DNA in the School, it became clear that they did not initially have experience in supplying NMR samples nor experience in purifying DNA. The first 10 μmole synthesis ordered was supplied in 31 tubes. There was approximately 3.7 mg in total. $D_2O$ (20 μL of 100%) was added to one vial, vortexed, and pulsed in the centrifuge. 1 μL of this solution was diluted to 100 μl using Milli-Q water. The absorbance of the diluted solution at 260 nm was determined using a Nanodrop. Actual absorbance was 8.06 μg/μL, expected was 6.6 μg/μL.

$D_2O$ (20 μL of 100%) was added to each of the 31 tubes. Each 20 μL was collected in one tube, and all tubes were rinsed with 1 x 100 μL. This 1 x 100 μL was added to the tube of 30 x 20 μL to make a total of 700 μL. This constituted the DNA sample.

A standard $^1H$ spectra was performed on the 400 MHz Bruker NMR. A large peak at ~4.65 ppm indicated HDO and the requirement to freeze dry the

sample to remove any $H_2O$ absorbed during the extraction process. There was also a peak at ~1.83 ppm.

PBS buffer was added to the DNA solution to a final concentration of 1 x PBS. The entire buffered DNA sample was transferred to a 5 ml round bottom flask and freeze dried between ~0.01 and 0.1 mmHg. Freeze drying was stopped when the pressure returned to 0.01 mmHg. The dried sample was transferred to a clean and dry NMR tube (Wilmad #528) by 1 x 200 µl wash of 100% $D_2O$ followed by 4 x 100 µl wash of 100% $D_2O$.

A standard $^1H$ spectrum was performed on the 400 MHz Bruker NMR. This showed that the impurity peak at ~1.83 ppm was still present. This was thought to be the tetraethylammonium ion (and other related ethylamines) – a by product of $Et_3N$ – used in the HPLC purification step. The buffered sample was deemed not suitable for research due to the impurity. The sample was returned to GeneWorks for further purification. GeneWorks were eventually able to provide a purified DNA hexamer sample which was dissolved in 100% $D_2O$ and a check for purity using a proton NMR showed that the impurity had been removed. This purified DNA sample was used hereafter.

### 7.1.4. NMR Assignments of the DNA Hexamer

The first process was to assign the DNA itself (in PBS) for ease of assignment in the final sample. A 30 and 60 ms TOCSY and a 600 ms NOESY were run on the 400 MHz Bruker NMR Spectrometer. It is appreciated that a mixing time of 600 ms for the NOESY is longer than may be required, however the spectrum was of good quality and care was taken to exclude effects of spin diffusion. The numbering system employed for DNA assignment is presented in **Figure 7.1**.

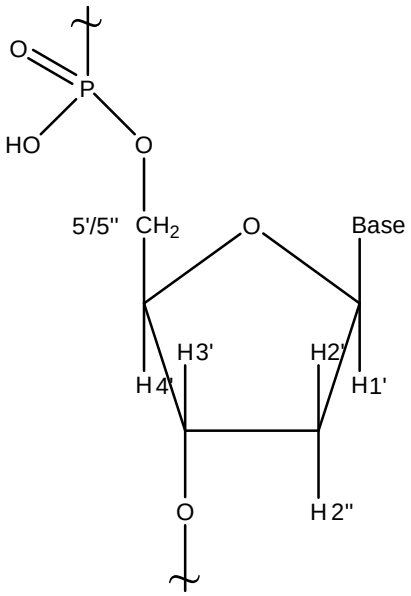



**Figure 7.1** Numbering system for the protons on the DNA sugar backbone.

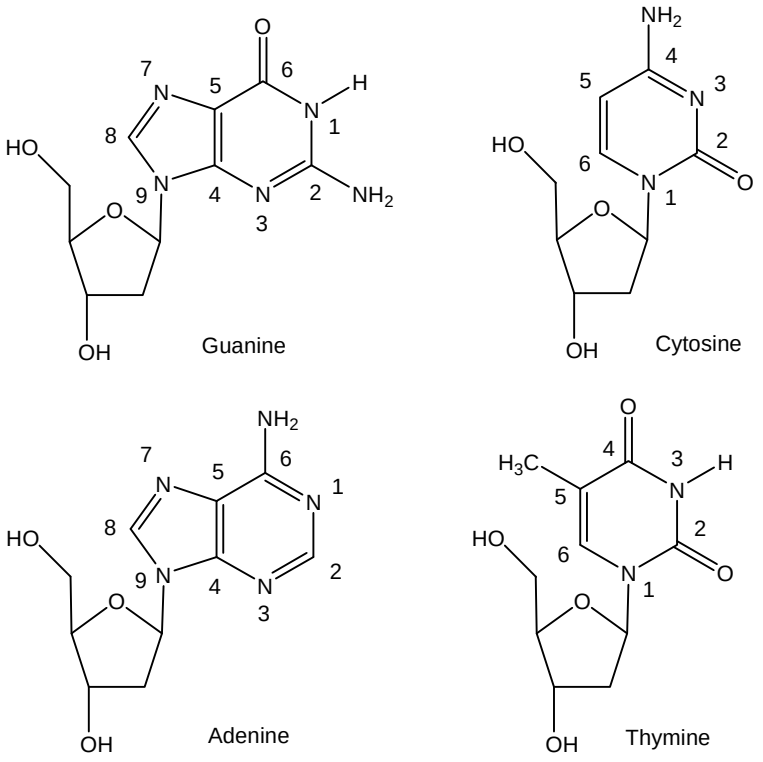



**Figure 7.2** Numbering system employed for the assignment of the aromatic protons on the nucleotides. Guanine has H8, cytosine H5 and H6, adenine H2 and H8, and thymine $CH_3$ and H6. The NOE peaks from the aromatic protons to the sugar backbone allow assignment from one end of the DNA to the other.

The method used to assign the NMR data from the NOESY spectrum of the DNA is given hereafter with reference to **Figures 7.1 - 7.3**;

A = aromatics to H1’ B = H1’ to H2’ and H2”

C = H2’ and H2” to H3’ D = H3’ to H4’, H5’ and H5”

E = aromatics to H2’ and H2” F = aromatic $CH_3$ (thymine) to H2’ and H2”

G = H1’ to H3’, H4’, H5’ and H5”

Beginning in section **A** of **Figure 7.3**, the 5’ end (1GH8) had no other cross peaks in a vertical line (within **A**). A horizontal line from this gave 1GH1’. A vertical cross peak indicates 2AH8, horizontally from there is 2AH1’, and so on along the DNA backbone of sugar rings. Once all of the aromatic peaks and their corresponding H1’ are known, the process moves onto section **B**.

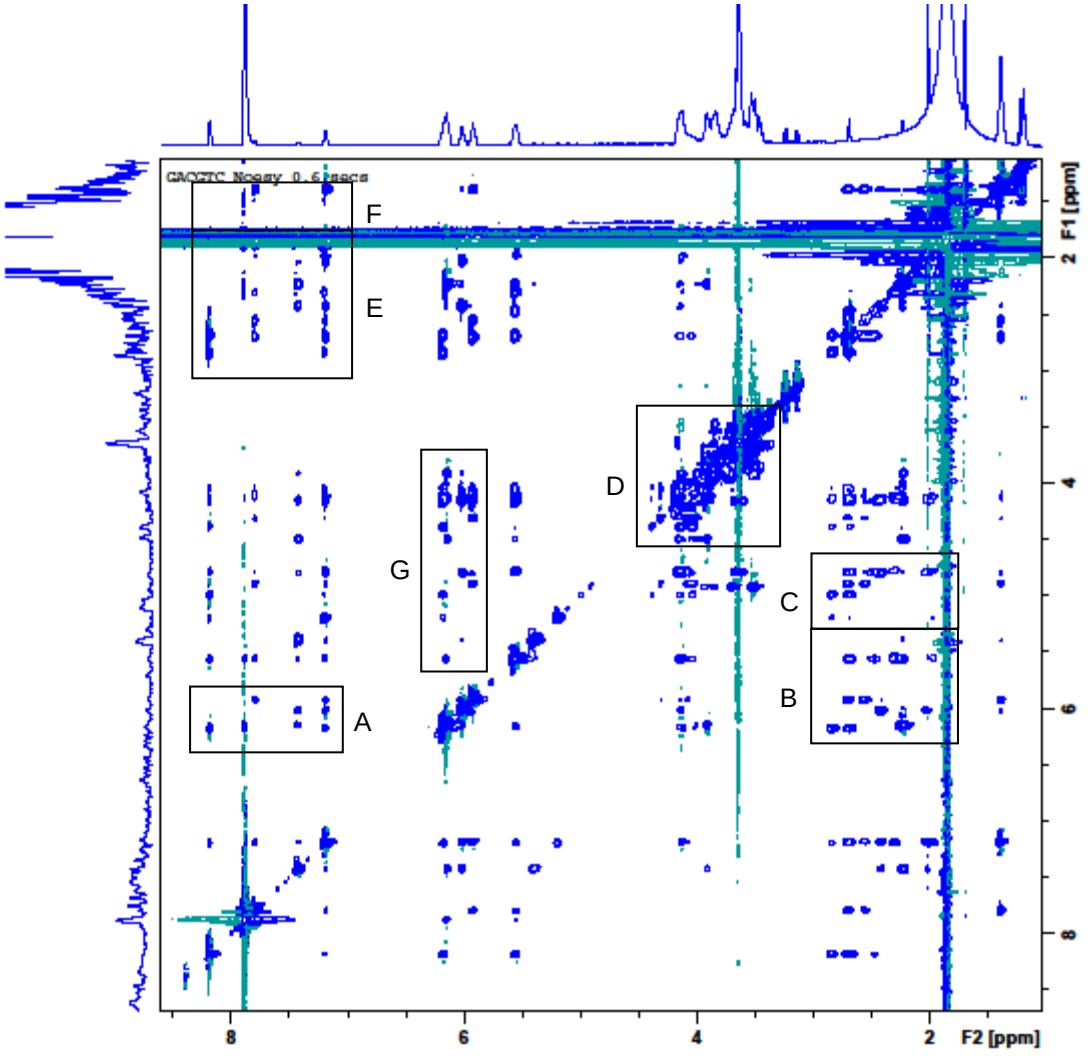



**Figure 7.3** 400 MHz NMR data from the 600 ms NOESY of d(GACGTC)$_2$ DNA.

Using the information from section **A**, a horizontal line from each H1' had one or two cross peaks corresponding to the H2' and H2" protons. The stronger peak was the H2". A vertical line from the H2' and H2" protons in section **B** revealed cross peaks to each H3' proton.

A COSY spectrum afforded a cross peak from the H3' to the H4' protons, and the remaining cross peaks on the H3' line were the H5' and H5" protons. Though these are two individual protons, they are not expected to separate, and thus there was only one peak for the H5' and H5". Sections **E**, **F** and **G** are used as proof of assignment.

These spectra were assigned using Sparky software running on a PC [129]. As seen in **Figure 7.3**, all of the H5' and H5" protons from 4.5 to 3 ppm were unable to be assigned due to overlapping peaks. The nucleotide and sugar backbone proton chemical shifts of $d(GACGTC)_2$ are presented in **Table 7.1**.

**Table 7.1** 400 MHz $^1H$ chemical-shift values in ppm from 600 ms NOESY spectrum of $d(GACGTC)_2$ in $H_2O/D_2O$ at 300 K.

| Base | H1' | H2'/H2" | H3' | H4' | H2/H5/$CH_3$ | H6/H8 |
|---|---|---|---|---|---|---|
| G1 | 5.560 | 2.471/2.689 | 4.784 | 4.137 | - | 7.883 |
| A2 | 6.176 | 2.682/2.839 | 4.989 | 4.386 | 7.882 | 8.185 |
| C3 | 5.552 | 1.971/2.302 | 4.766 | 4.039 | 5.194 | 7.196 |
| G4 | 5.922 | 2.552/2.699 | 4.890 | 4.311 | - | 7.794 |
| T5 | 6.015 | 2.029/2.421 | 4.794 | 4.139 | 1.390 | 7.188 |
| C6 | 6.146 | 2.227/2.433 | 4.492 | 3.911 | 5.393 | 7.420 |

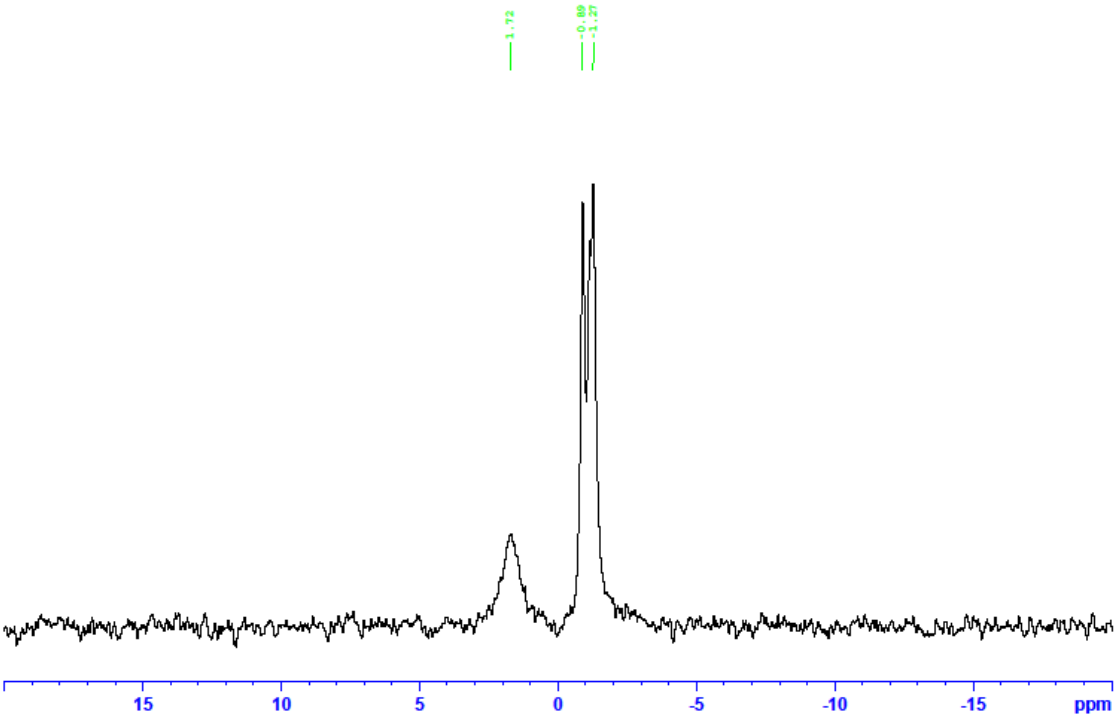



**Figure 7.4** The 162 MHz $^{31}$P NMR spectrum of d(GACGTC)$_2$ in $D_2O$ at 300 K measured on the 400 MHz Bruker DRX spectrometer.

The 162 MHz $^{31}$P NMR spectrum of d(GACGTC)$_2$ presented in **Figure 7.4** was performed in an attempt to observe the drug binding to the DNA. The phosphate in the DNA backbone appears as the sharp multiplet between -0.87 and -1.27 ppm, and the phosphate buffer has a broad peak at 1.72 ppm. This spectrum is consistent with ds DNA.

## 7.2. Individual Component Assignment

Each individual component of the final NMR sample was assigned from multiple spectra. These mainly consisted of a standard $^{1}$H, watergate $^{1}$H, $^{13}$C, DEPT135, TOCSY, NOESY and $^{31}$P (depending on the component). It is essential to assign the drug and the DNA before they are mixed together to assist with the assignment of the complex. A NOESY spectrum should greatly assist in determining the 3D structure.

### 7.2.1. C4-DACA Assignment

A more detailed account of the assignment of C4-DACA is given in **Chapter 2**. There was a slight complexity in the C4-DACA spectrum as the $CH_2$'s (H16/17) in the side chain were quite broad due to the small molecule undergoing dynamics. They had a different proton but same carbon shift hence it was not possible to unequivocally assign the $CH_2$'s in the side chain, although luckily the carbon assignment is not necessary for this particular set of experiments. **Table 7.2** was taken from **Chapter 2**. The H1 to H9 protons were assigned from the 900 ms mixing time NOESY spectra presented in **Figure 7.5**.

**Table 7.2** $^1$H Chemical-shift values in ppm of C4-DACA in $H_2O/D_2O$ at 300 K and 400 MHz.

| H1 | H2 | H3 | H5 | H6 | H7 | H8 | H9 | H15 | H16/17 | H18 | NH |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 8.52 | 8.04 | 8.81 | 8.29 | 8.13 | 8.41 | 8.45 | 9.32 | 3.97 | 2.28 | 3.60 | 11.42 |

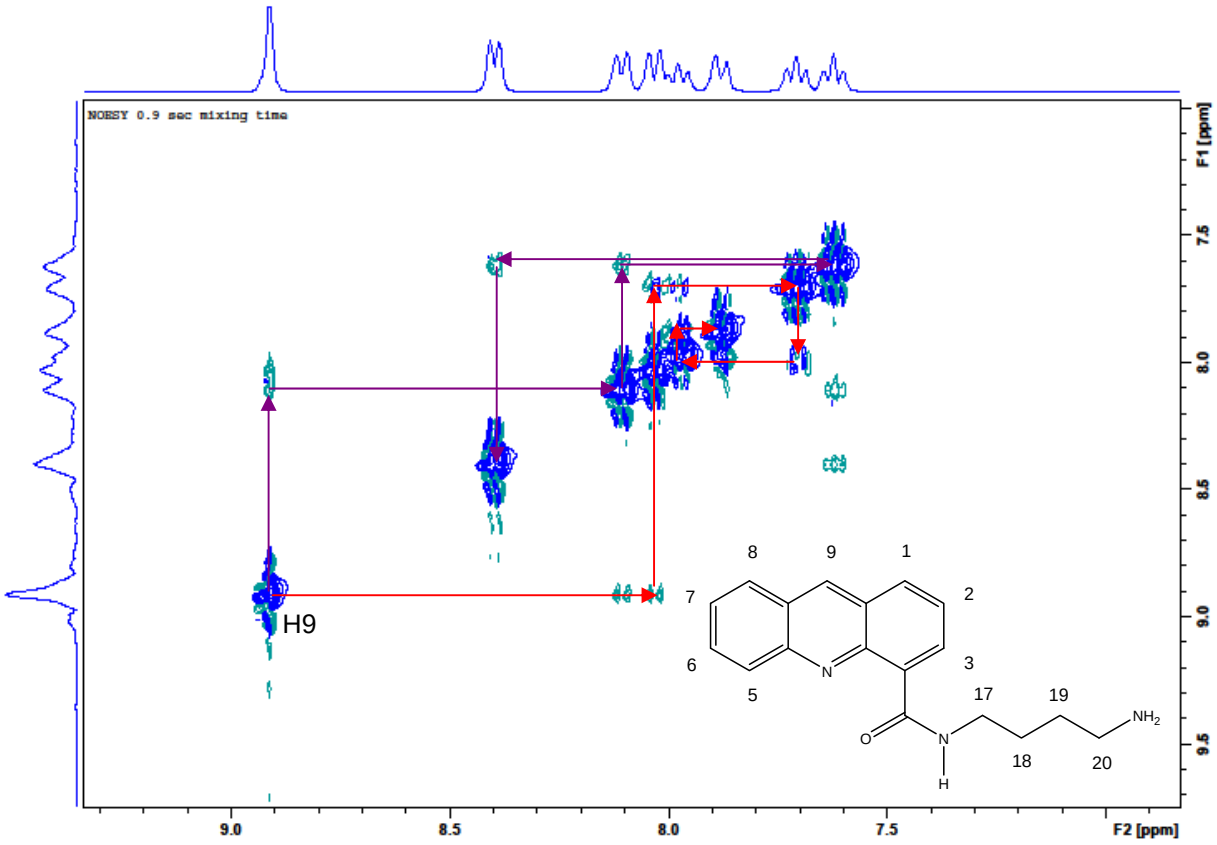



**Figure 7.5** An expanded view of the aromatic section of a 900 ms NOESY (1.8 mM C4-DACA in $D_2O$). The coloured arrows show how the NOESY can see protons through space (NOE peaks). The H9 is not directly next to another proton, but the NOESY spectrum gives NOE peaks to indicate either the H1 to H3 side (purple arrows) or H5 to H8 side (red arrows) of the acridine ring.

The NOESY did not contain any cross peaks at the mixing times relevant to the drug-DNA complex.

### 7.2.2. Formaldehyde Assignment

The proton NMR spectrum of formaldehyde (not shown) proved to be very complex with many peaks representing a full range of cyclic and linear polymers. It was hoped that this complexity could be circumvented using deuterated formaldehyde. However, the spectrum shown in **Figure 7.6** of a stock solution of formaldehyde-$d_2$ in $D_2O$ (undiluted) from Sigma is still very complex, showing none of the residual peaks expected from the formaldehyde monomer.

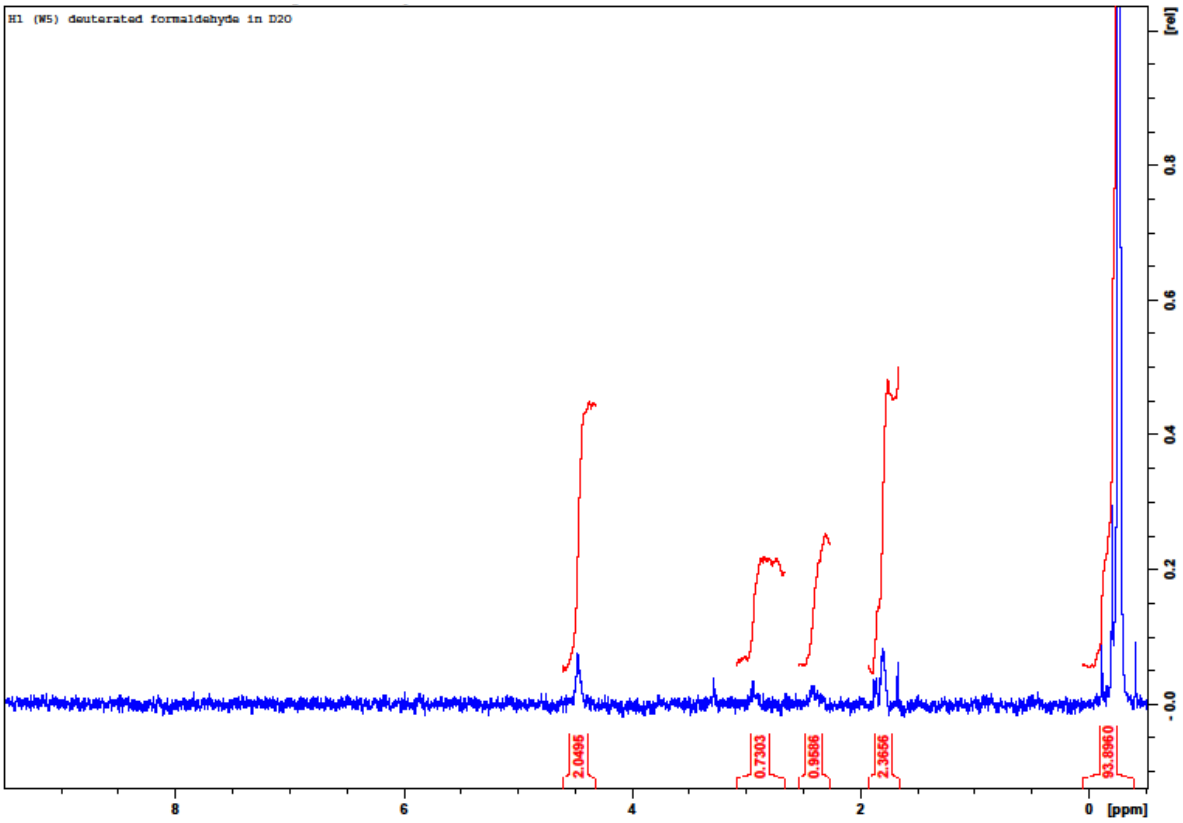



**Figure 7.6** 400 MHz watergate $^{1}$H NMR spectrum of stock formaldehyde-$d_2$ solution.

## 7.3. Drug/DNA Sample Preparation

All concentrations were determined using a NanoDrop. The stock solution of PBS contained sodium azide (12 mg $NaN_3$ in 5 mL PBS) to inhibit bacterial growth. The drug-DNA sample was prepared at a 1:1 ratio (slight excess of drug, calculated for ds DNA) by adding 24 µL of 38 mM C4-DACA (in $D_2O$) to the previous DNA sample from **7.2.1** (~3 mg $d(GACGTC)_2$ in 0.6 mL $D_2O$ and 24 mL of 25 x PBS). Note that formaldehyde was not added at this stage.

A standard $^{1}$H spectrum was performed on the 400MHz Bruker NMR. This was followed by a watergate $^{1}$H (centred on the water peak), a watergate NOESY (**Figure 7.7**), TOCSY and $^{31}$P spectra (**Figure 7.9**) at 300 K.

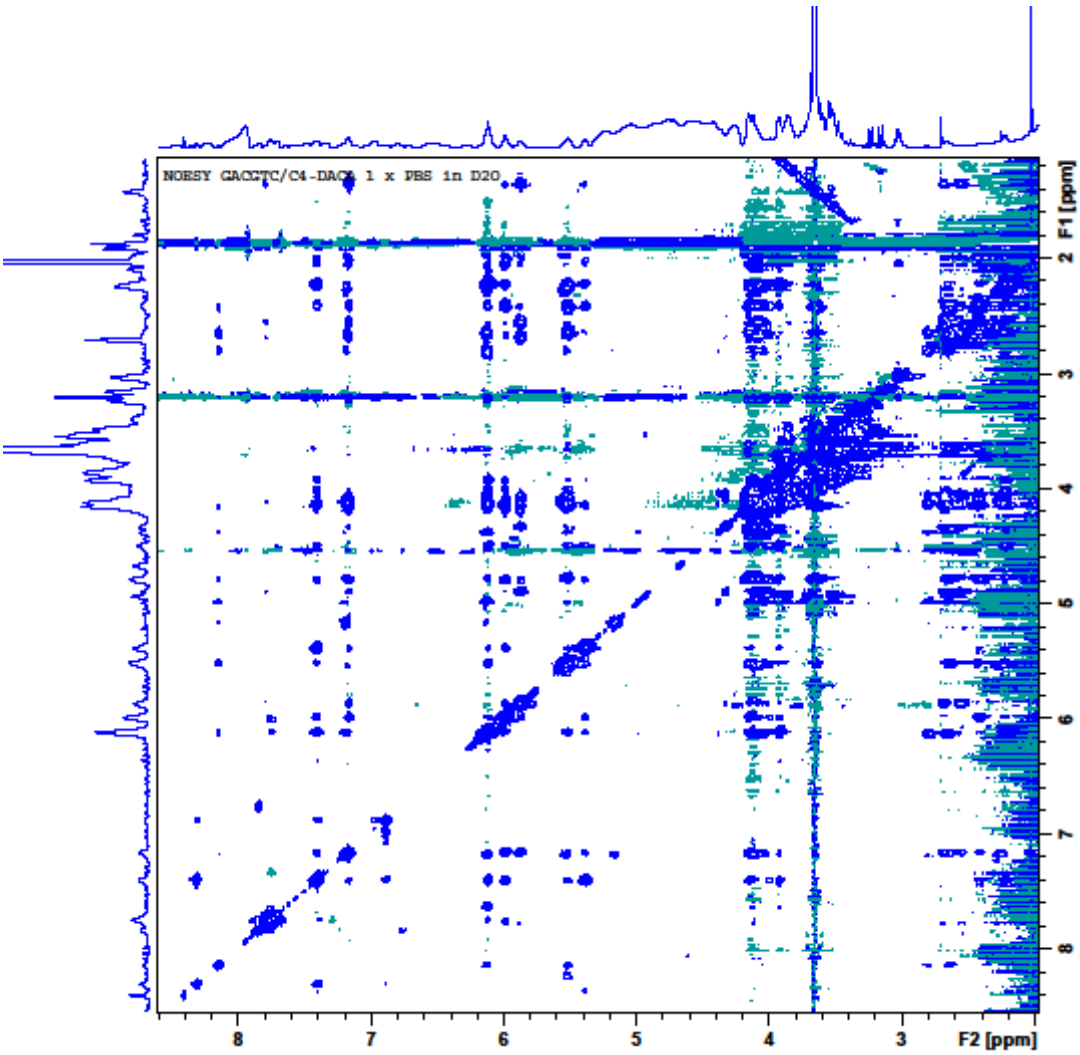



**Figure 7.7** 400 MHz 900 ms mixing time NOESY spectrum of C4-DACA/d(GACGTC)$_2$ in $D_2O$ at 300 K.

The 900 ms NOESY spectra shown in **Figure 7.7** had broad peaks, particularly in the aromatic region that were thought to arise from a very short C4-DACA intercalation time and the spectrum being in "fast exchange". Thus, these spectra were repeated at 310 K and are presented in **Figure 7.8**. A range of temperatures, mixing times, and the full range of window functions were employed, however a lower temperature experiment (280 K) was not recorded.

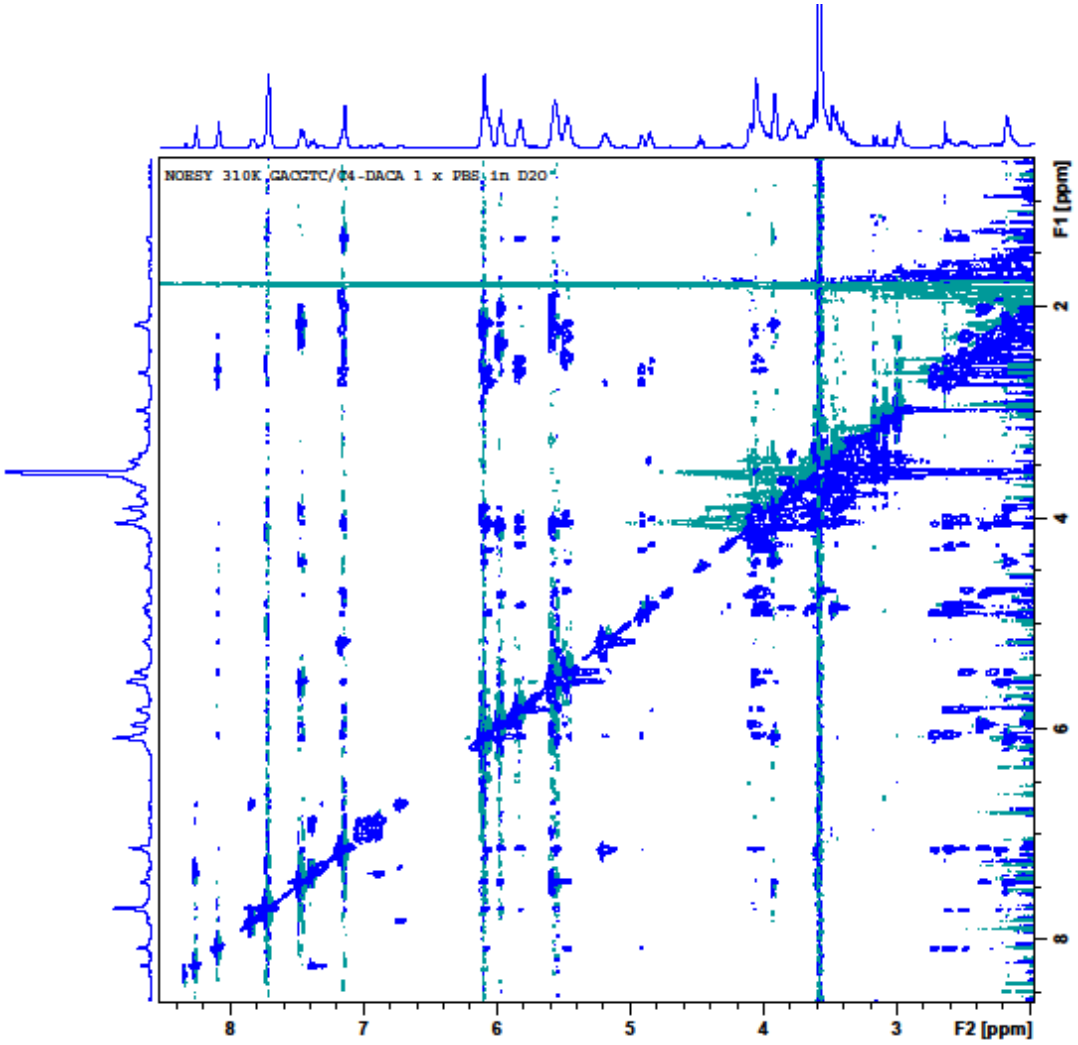



**Figure 7.8** A 400 MHz 900 ms mixing time NOESY spectrum of C4-DACA/d(GACGTC)$_2$ in $D_2O$ at 310 K.

As expected, the increase in temperature improved the line widths of the protons in the aromatic region and moved the chemical shift of the peaks. This is consistent with the drug in fast exchange with the DNA. It was also apparent that there was considerable disruption of the DNA in the rest of the spectrum when compared to **Figure 7.7**. There were two aromatic peaks without any correlating NOE peaks (G1 and G4). It was thought that the drug was interfering with the DNA's ability to form double stranded DNA, or that the fast drug exchange was broadening the G1 and G4 correlations to the point that they were no longer visible. There were no NOE cross peaks observed from the drug aromatic peaks to the DNA.

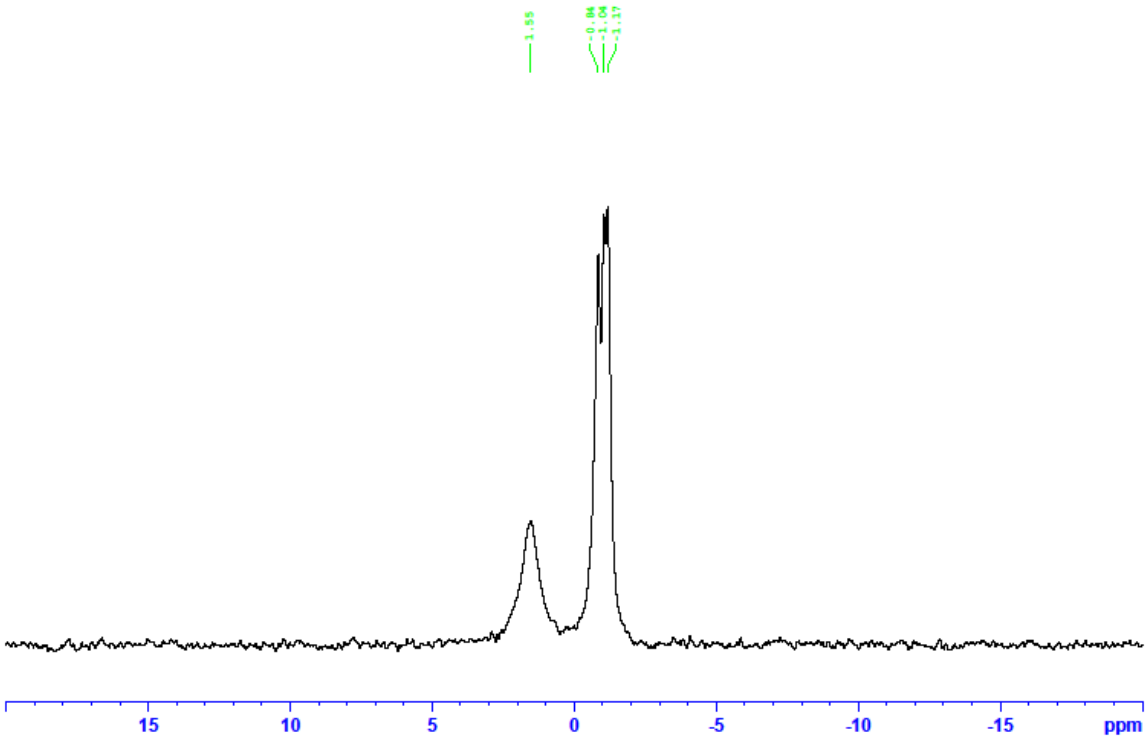



**Figure 7.9** The 162 MHz $^{31}$P NMR spectrum of C4-DACA/d(GACGTC)$_2$ in $D_2O$ at 300 K measured on the 400 MHz Bruker DRX spectrometer.

The $^{31}$P NMR spectrum of C4-DACA/d(GACGTC)$_2$ presented in **Figure 7.9** was performed in an attempt to observe the drug intercalating into the DNA. The phosphate in the DNA backbone appears as the sharp multiplet between -0.84 and -1.17 ppm, and the phosphate buffer has a broad peak at 1.55 ppm. It was of note that the DNA peak in this spectrum is still very similar to that shown in **Figure 7.4** in that it is still quite sharp. The fact that it is not exactly the same supports the assumption that the drug may be in the fast site exchange regime and may be intercalating as observed by the 900 ms mixing time NOESY spectrum in **Figure 7.8**.

## 7.4. Drug/DNA/Formaldehyde Sample Preparation

Formaldehyde was added to the previous drug-DNA sample (**7.3**) in portions. A decision was made to add quite a large excess of formaldehyde to the solution with the purpose of driving the Schiff base reaction to completion, thus it was decided to use an excess of 100 times the concentration of the drug. The

concentration of the stock formaldehyde was 20% wt (~6.24 mol/L), so the sample needed 4 x 3.75 µL portions (some would be lost with the multiple pipetting). Once each portion was added, a standard $^{1}H$, watergate $^{1}H$, and $^{31}P$ spectra was taken. The first portion was added and the three specified spectra taken immediately followed by 1, 2, 3, 18, 24 and 48 hour time points. The second portion was added at 50 hours (shown in **Figure 7.10**) and spectra taken at 50, 66 and 72 hour time points. After 96 hours, only the $^{31}P$ spectra were taken, at 4K and at 8K resolution. After 7 days from the initial formaldehyde addition, only one $^{31}P$ spectra at 10K resolution was taken. This $^{31}P$ (at 10K resolution) spectra was repeated at 8, 9 and 10 days. After this time, it was decided not to add the last two formaldehyde portions as the spectrum was already very convoluted and broad, and was missing some DNA peaks.

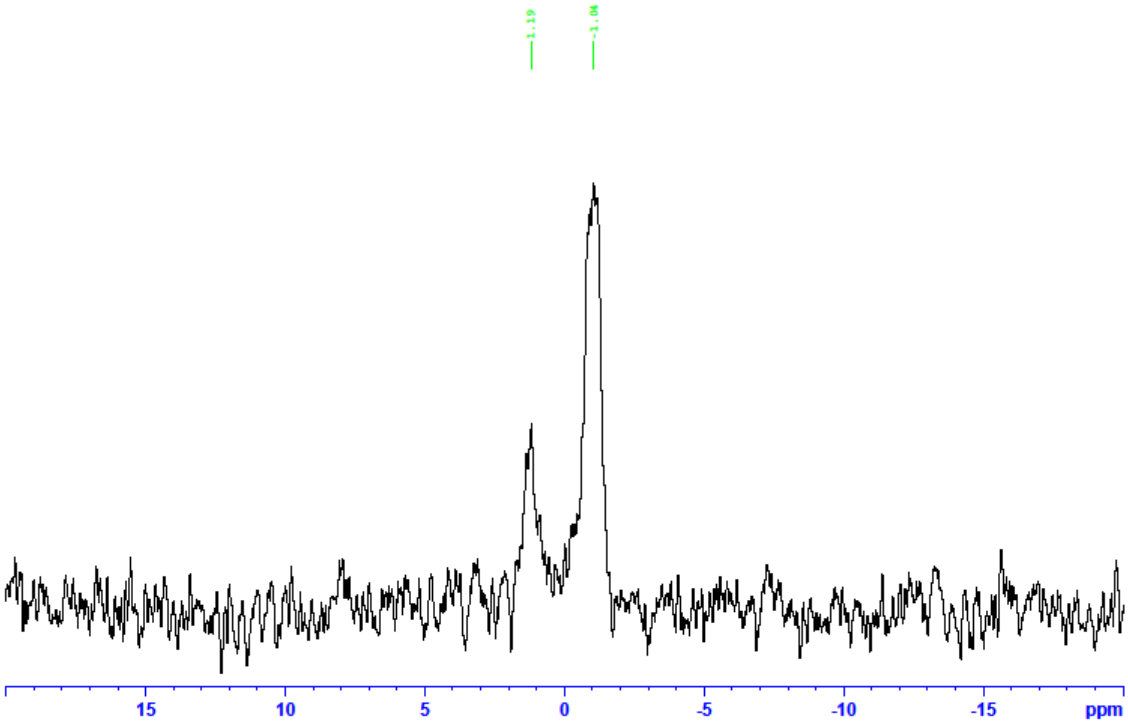



**Figure 7.10** The 162 MHz $^{31}P$ NMR spectrum of formaldehyde/C4-DACA/d(GACGTC)$_2$ in $D_2O$ at 300 K at 50 hours after formaldehyde addition measured on the 400 MHz Bruker DRX spectrometer.

The 162 MHz $^{31}P$ NMR spectrum of formaldehyde/C4-DACA/d(GACGTC)$_2$ presented in **Figure 7.10** was performed in an attempt to observe the drug binding to the DNA. The phosphate in the DNA backbone appears as the broadened multiplet at -1.04 ppm, and the phosphate buffer has a broad peak at 1.19 ppm. It was of note that the DNA peak in this spectrum is no longer similar to those shown in **Figure 7.4** or **7.9** as the peak has broadened. The fact that it is not the same supports the assumption that the drug has formed the formaldehyde-mediated drug-DNA adduct, however it could also be attributed to the effect of the excess formaldehyde. Unfortunately this could not be observed by the 900 ms mixing time NOESY spectrum in **Figure 7.8**.

The 120 ms mixing time TOCSY spectrum of the formaldehyde/C4-DACA/d(GACGTC)$_2$ sample presented in **Figure 7.11** was performed 15 days after the initial formaldehyde addition. It is clear from **Figure 7.11** that there are virtually no correlation peaks apparent. The DNA contained in this sample is effectively unrecognisable in the spectrum arising from the formaldehyde addition. It was suspected that the formaldehyde physically broke up the DNA, essentially destroying the sample.

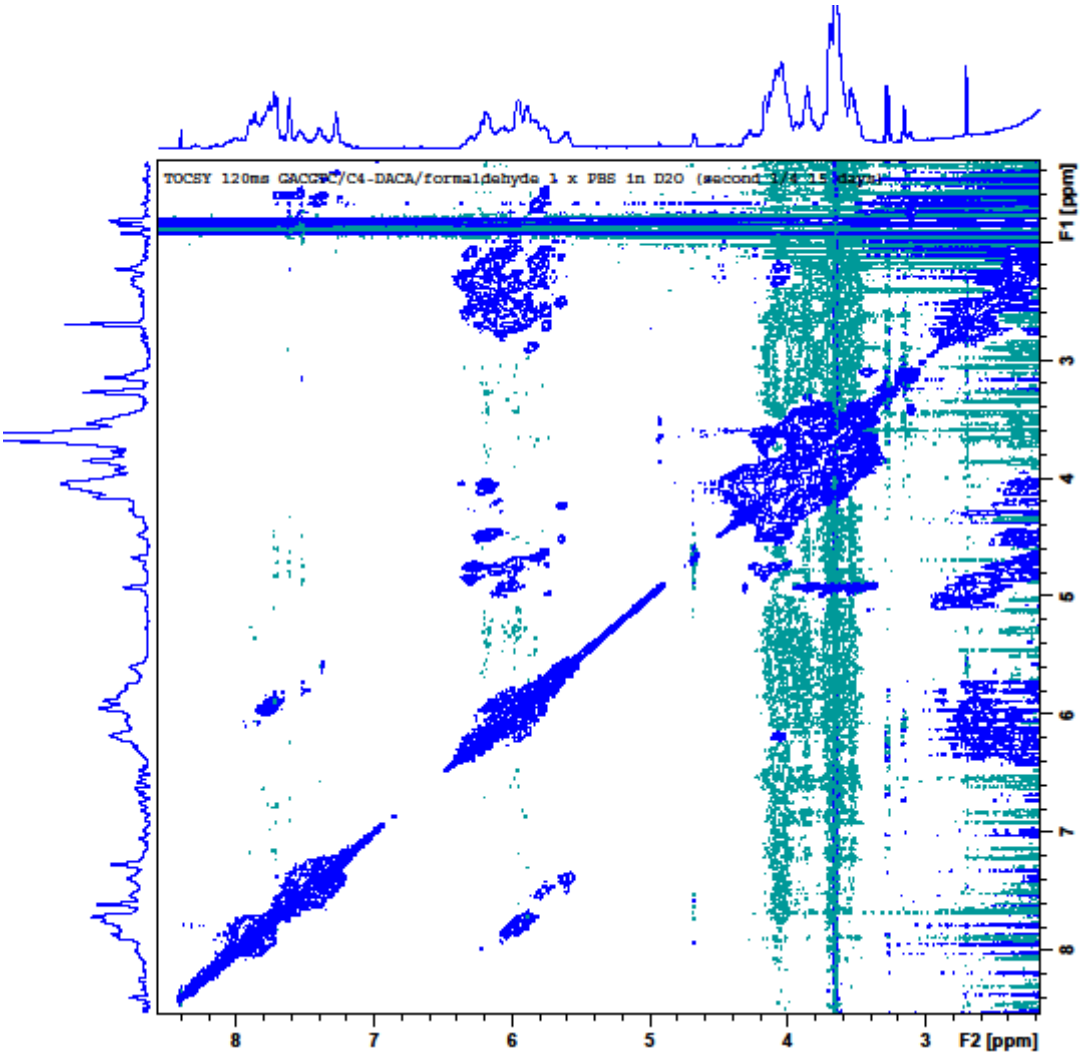



**Figure 7.11** 400 MHz 120 ms mixing time TOCSY spectrum of formaldehyde/C4-DACA/d(GACGTC)$_2$ in $D_2O$ at 300 K.

## 7.5. Conclusions

The full assignment of d(GACGTC)$_2$ was completed (excluding H5' and H5" protons, **Table 7.1**) in $H_2O/D_2O$ at 300 K on a 400 MHz Bruker NMR spectrometer. Results from a standard $^1$H NMR of stock formaldehyde-$d_2$ solution showed a very complicated spectrum consisting of monomers, dimers, polymers and cyclic derivatives (paraformaldehyde) and hydrated species of formaldehyde.

There was evidence of drug-DNA interactions from the spectrum of C4-DACA/d(GACGTC)$_2$ in $H_2O/D_2O$ observed as peak broadening in $^1$H, $^{31}$P and NOESY spectra. It was apparent that there was a considerable disruption of the DNA structure when analysing the multiple spectra of C4-DACA/d(GACGTC)$_2$ in $H_2O/D_2O$. Compared to the spectrum of DNA alone, there were two aromatic

peaks without any NOE cross peaks (G1 and G4), and also no NOE peaks from the drug aromatic protons to the side chain or to the DNA. It is reasonable to assume that the drug was in dynamic exchange with the DNA, and disrupting the ability to form double stranded DNA. Unfortunately this made the assignment of the drug-DNA complex unachievable. It was hoped that the addition of formaldehyde would reduce the exchange time of the drug by forming the formaldehyde-mediated drug-DNA adduct, and therefore aid the assignment.

Addition of formaldehyde to the solution of C4-DACA/d(GACGTC)$_2$ in $H_2O/D_2O$ proved to be inconclusive. The $^{1}H$, $^{31}P$, and TOCSY spectra of the drug/DNA/formaldehyde solution were very convoluted and broad, and missing a number of DNA correlation peaks. There were virtually no correlation peaks apparent in the 120 ms mixing time TOCSY spectrum, making the DNA effectively unrecognisable after the formaldehyde addition. It was suspected that in addition to Schiff base formation that the excess of formaldehyde physically reacted with the DNA thus destroying the integrity the sample.

Unfortunately limitations on the DNA sample did not allow additional experiments.

# CHAPTER 8. MODELLING STUDIES

## 8.1. Overview

The results from the initial experimental work involving C2-DACA which had a short side chain motivated the modelling of various alterations to the drug. Specifically, this involved a modified version of C2-DACA intercalated and/or bound to DNA with or without formaldehyde activation. This was initially to discover which groove (major or minor) the drug was intercalating in, and in particular to find the optimum chain length for the side arm to enable DNA binding. The starting point for this work was a literature [8] X-ray crystal structure of Adriamycin covalently bound to a short section of DNA via a formaldehyde-mediated bridge. Geometry optimisation of the intercalated drug and the drug-DNA adduct were required for both C2-DACA-DNA and the Adriamycin-DNA adduct. A comparison of these two models aided in discovering the most favourable chain length. Modelling was also used to visualise the final structure when using different methods to extend the chain.

The daunosamine side chain of Adriamycin can be activated *in vitro* by formaldehyde and consequently form a covalent drug-DNA adduct [66] as shown in **Figure 8.1**. This mode of action was the inspiration for this body of work.

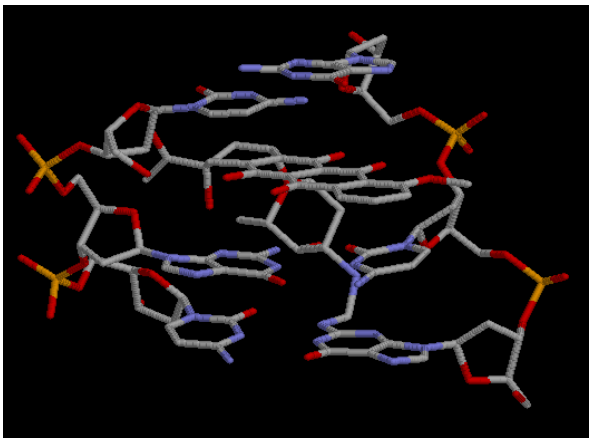


**Figure 8.1** Part of the crystal structure of Adriamycin bound to guanine via a methylene bridge in the minor groove [8].

The literature X-ray structure shows that Adriamycin intercalates into the DNA lengthwise with the daunosamine side chain positioned in the minor groove. The 3'-$NH_2$ of the Adriamycin sugar forms a Schiff base with formaldehyde and consequently reacts with the free amine of guanine via the minor groove to form the drug-DNA adduct [66]. The aim of the modelling was to elucidate a similarly structured compound that mimics the nature in which Adriamycin interacts with DNA. The requirements are to have a flat planar aromatic ring system with a side chain of sufficient length (approximately 8 Å) and flexibility capable of interacting with the primary amine of guanine via the minor groove.

The work described in this thesis is based on the acridine intercalators which can intercalate via both the major and the minor groove depending upon the nature of the side chain. The tricycline acridine part of the molecule slots into the DNA parallel to the base pairs with the side chain determining which groove it resides in by any interactions it may have with the DNA. As proposed in this work, a modified version of *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide (DACA) was chosen as a drug with much potential as DACA was a Topo II poison and had been selected for clinical trials [47]. The literature X-ray structure of the 9-amino derivative (452D depicted on the left hand side of **Figure 8.5**) bound to DNA shows the expected intercalation [38]. In this case the side chain is in the major groove probably due to the interaction of the positive charge on the protonated dimethylamino group with the DNA. Its basic components were the essential flat planar aromatic compound and the side chain. The only part that required manipulation and modification was the side chain which was unable to form a Schiff base with formaldehyde and react with guanine in the minor groove due to the di-methylated amino group. Thus a compound with a free amine at the end of the side chain was required. Modelling was used to determine the optimum length for the side chain in order to minimise base pair puckering and twisting, and to give the biochemical work the greatest potential to succeed.

DNA has the ability to wind and unwind as the base pairs are held together by flexible hydrogen bonds. This supple nature of DNA allows the flat planar aromatic moiety of the drug to intercalate and still maintain the double helical form of DNA. In the case of nogalamycin, the dumbbell shaped molecule adopts a 'threaded' binding orientation, in which the drug chromophore passes through the DNA helix positioning the nogalose sugar in the minor groove, and the positively

charged bicycloamino sugar in the major groove [17]. This suggests that the hydrogen bonding between the base pairs near the binding site must be disrupted for the large sugar to thread through, and subsequently reform once the chromophore is in place. This also implies that DNA has the ability to accommodate chromophores of varying size while still maintaining the stability of base pair hydrogen bonding.

## 8.2. Molecular Modelling

### 8.2.1. Drug-DNA Structures

A series of DNA intercalators and representative pieces of DNA were downloaded in .pdb format from the websites www.pdblite.org and www.rcsb.org (Research Collaboratory for Structural Bioinformatics). These structures were used as a starting point for the modelling studies. The ligand or ligands from these X-ray structures were given shorthand names derived from their original .pdb file (e.g. the ligand from 1g3x.pdb was *N*[α]-[9-acridinoyl]-tetraarginine-amide, but is referred to as 1g3x).

### 8.2.2. X-Ray structures and Modelling Software: Insight and RasMol

The structures were reviewed and geometric information derived using RasMol [130]. Manipulations and calculations on the structures were performed using a Silicon Graphics Iris Computer using the MSI software Insight and Discover. In general the structures needed to be "cleaned up" with the water molecules in the crystal structure deleted, the correct number of hydrogen atoms added, and the atom potentials and charges set according to the CVFF forcefield using Insight II [131].

Of the series of drug-DNA crystal structures downloaded, five were selected since they contained acridine-like ligands with structures shown in **Figure 8.2**.

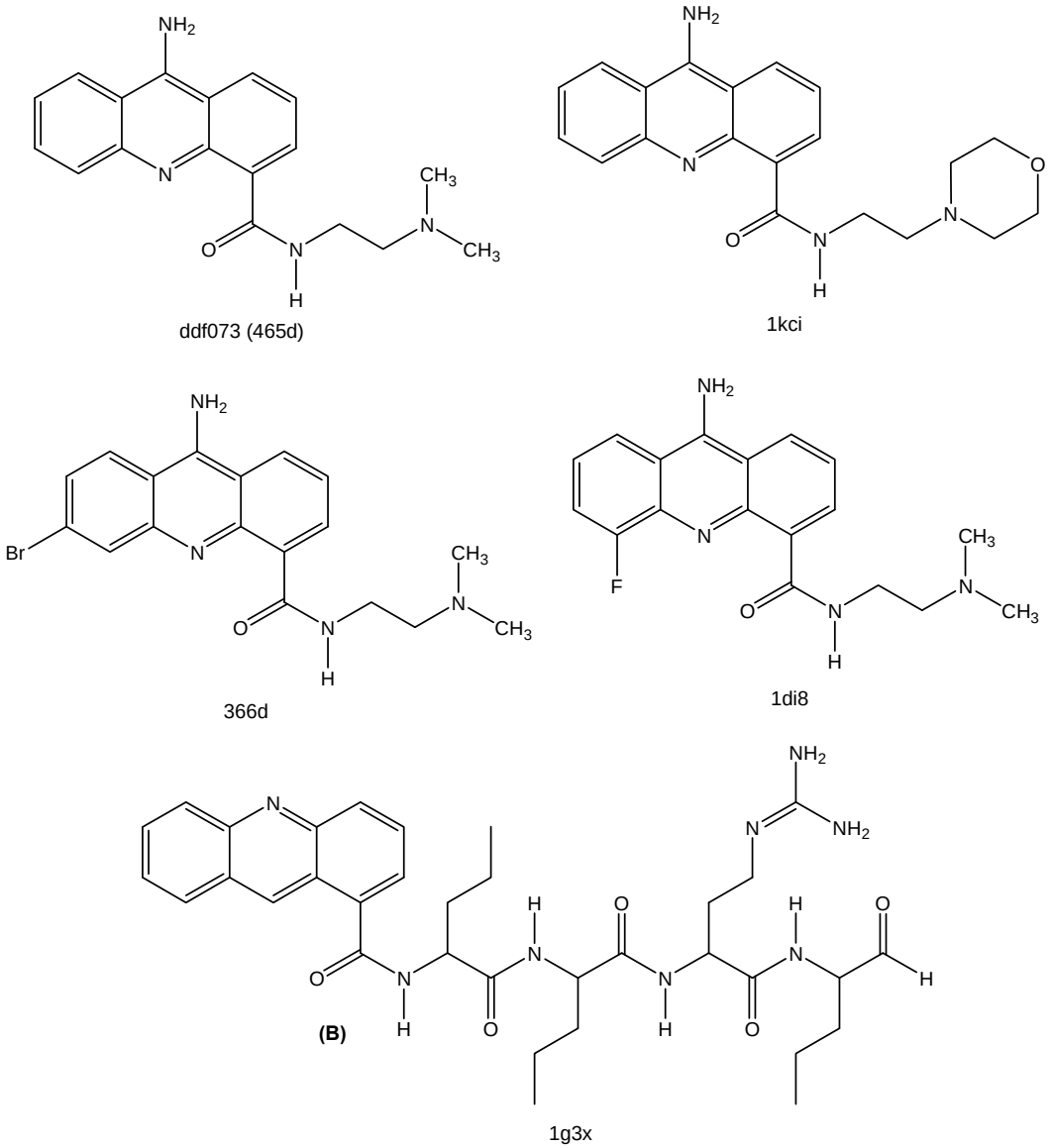



**Figure 8.2** Structures of some of the drugs that were intercalated in the PDB files.

All of these drugs were intercalated with the DNA sequence d(CGTACG)$_2$ **A** as shown in **Figure 8.3**, except *N*[α]-[9-acridinoyl]-tetraarginine-amide (the structure from 1g3x.pdb), which was intercalated with the sequence d(CGCGAATTCGCG)$_2$ **B** also shown in **Figure 8.3**.

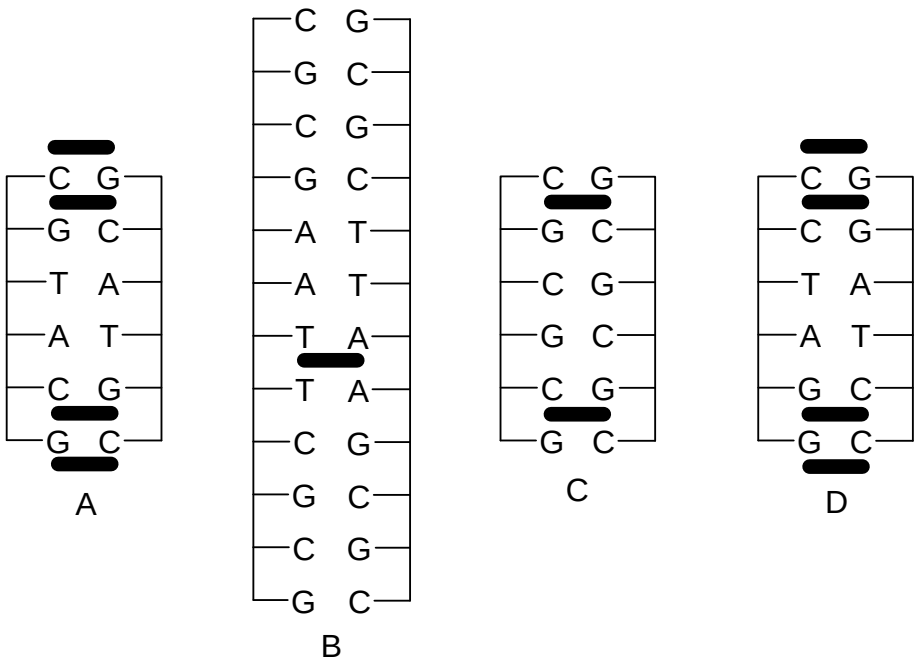



**Figure 8.3** DNA sequences from the various crystal structure PDB files used. The black bar indicates the intercalation sites of the drugs.

The black bar indicates the intercalation site of the drug (i.e. only one molecule of 1g3x intercalates in DNA sequence **B**, whereas the rest of the drugs from **Figure 8.2** have four molecules intercalating into DNA sequence **A**). The ligand from 1g3x proved to be interesting in that the acridine has the side chain attached to the opposite side to DACA and its analogues (at the 9-acridine position), and that it intercalates between a 5'-TT-3' in the minor groove. The long acridine chains were close to the phosphate-sugar backbone, with the $-N=C(NH_2)_2$ relatively close to one of the phosphorus linkers.

A number of other structures were obtained using the keyword "acridine" on the PDB lite website, and are shown in **Figure 8.4**. Structures 1k2l and 1fd5 were intercalated to DNA sequence **A**, while 1k9g was intercalated to DNA sequence **D**. The literature crystal structure of Adriamycin was intercalated and bound to DNA sequence $d(CGCGCG)_2$ **C** [8], while 1k9g was intercalated to DNA sequence $d(CCTAGG)_2$ **D** (intercalation sites are shown in **Figure 8.3**).

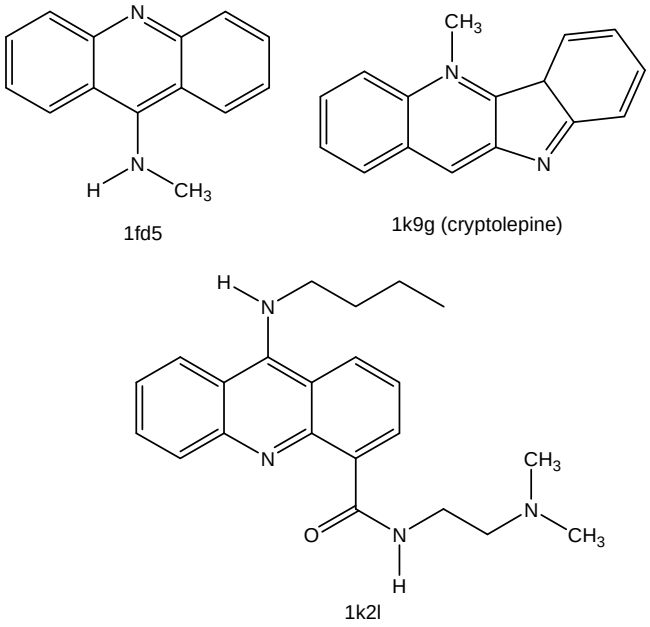



**Figure 8.4** Structures of some of the drugs that were intercalated in the PDB files.

On examining these crystal structures it was apparent that part of the DNA sequences for 1k2l and 1fd5 were not properly paired up (i.e. the last one or two base pairs of the DNA displayed no hydrogen bonding), or one of the bases was missing completely (**Figure 8.5**). Due to this defect, the exact intercalation sites of the molecules are uncertain.

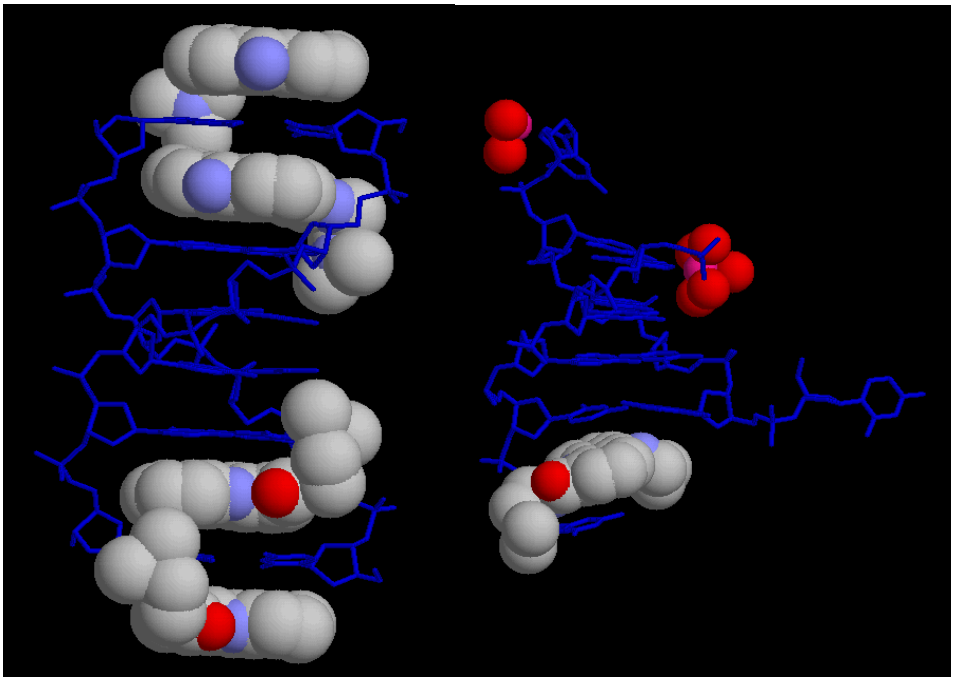


**Figure 8.5** Normal (465d) and abnormal (1k2l) base pair structure with intercalated drug.

It is clear that in general the literature X-ray structures have the acridine ring system parallel to the DNA base pair axis, and although the DNA structures with the protonated side chains are in the major groove, the side chains of acridine intercalators occupy both DNA grooves.

## 8.3. Modelling of C2-DACA and Derivatives

The first compound studied with the DNA binding assay was C2-DACA. This compound did not show any DNA binding presumably because the side chain was too short to allow a covalent bond to the primary amine group of the guanine base. This compound was therefore used as the starting point for the modelling studies to elucidate a better compound for further work.

### 8.3.1. Method

The literature crystal structure from 1qda contained Adriamycin bound to the DNA sequence **C**. The initial model was generated by manually overlapping the intercalated drug from 465d (DACA) onto the Adriamycin chromophore. The acridine ring was placed in the same orientation as the 465d PDB structure, with the acridine ring parallel to the DNA bases. This was to examine the binding sites in more detail, and to investigate the base pair interaction with the intercalating drug. The bottom three base pairs of the DNA structure were removed to simplify the computational study and any later minimisations. This resulted in the $d(GCG)_2$ sequence with the drug in the appropriate position (**Figure 8.1**). As the calculations are generally performed without solvent, distance constraints must be applied to simulate hydrogen bonding between the base pairs (**Figure 8.6**). This is normal practice to maintain the correct DNA structure, but to allow reasonable freedom of the complex to equilibrate. Distance constraints were applied to the DNA and the terminal nitrogen of the DACA side chain. The structures were energy minimised using steepest decent, then conjugate gradient methods (100,000 iterations). The application of distance constraints was an attempt to

create a stable platform for the drug to be optimised in, as the drug itself was not constrained. This was to find out how long the side arm should be to reach the guanine without too much strain or puckering of the DNA, and to discover if the nitrogen of the acridine has a preferred conformation for intercalation.

Distance constraints were applied between the base pairs (where hydrogen bonds would exist in solution), and between each end of the backbone to minimise distortion (shown in **Figure 8.6**). The values for the DNA constraints listed in **Table 8.1** were obtained from the 1qda crystal structure containing Adriamycin, and then given ±0.1 Å to allow minimal movement. Values for the terminal nitrogen of the side chain were obtained from the $N(Me)_2$ group of DACA in the 465d crystal structure.

**Table 8.1** Distance constraints applied to the $d(GCG)_2$ sequence with C2-DACA intercalated.[1]

| **Constrained atom** | | **Distance in angstroms (Å)** |
|---|---|---|
| Base pairs: | A1N:B6H; A3O:B4H, A3H:B4O | 2.10 – 1.90 |
| | A2H:B5O | 1.85 – 1.65 |
| | All other base pair H-bonds | 2.00 – 1.80 |
| 5’ to 3’ end: | A1C1’:A3C1’ | 10.45 – 10.25 |
| 5’ to 3’ end: | B4C1’:B6C1’ | 11.40 – 11.20 |
| Drug to base pair: | N:B5O | 3.60 – 3.40 |
| Drug to base pair: | N:A2H | 3.60 – 3.40 |

[1] The values for the constraints were taken from the 1qda crystal structure containing Adriamycin, and those for the terminal nitrogen of the side chain were obtained from the 465d crystal structure (distance ±0.1 Å).

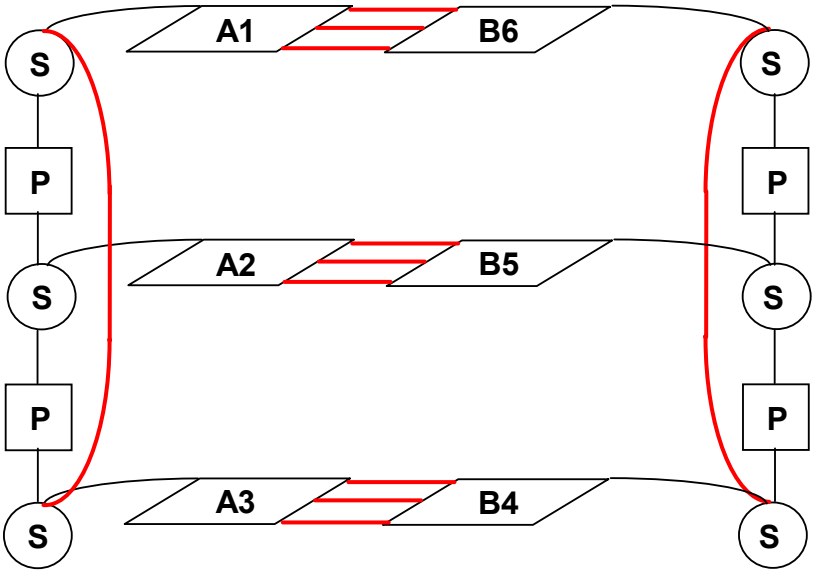



**Figure 8.6** The red lines indicate where distance constraints were applied to the DNA. Though there are three red lines between all base pairs, the A-T base pairs only had two hydrogen bond sites and hence only two points for constraint.

### 8.3.2. Results and Discussion

The structure in **Figure 8.7** consists of C2-DACA intercalated in $d(GCG)_2$. The first drug manually intercalated in $d(GCG)_2$ was DACA, which was minimised initially to verify the constraints used would sufficiently hold the DNA together. After minimisations were finished, the DACA side chain was altered to C2-DACA, and the complete minimisation using the correct DNA restraints was performed. The resulting structure is shown in **Figure 8.7**, and shows a reasonable drug-DNA interaction with minimal steric issues. The DNA bases around the intercalating section are planar with minimal distortion. The terminal base pair shows minor distortion, as expected.

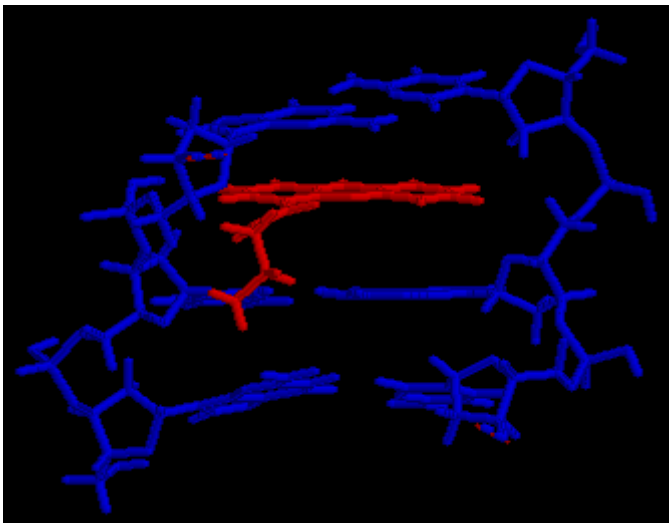


**Figure 8.7** Minimised unbound C2-DACA in d(GCG)$_2$ (view from minor groove).

In order to generate the covalent bond with the amino group of guanine, the C2-DACA side chain was artificially reacted with formaldehyde, and bound to the same free amine on the guanine as the Adriamycin molecule from 1qda. It was desired to evaluate the side arm length required to bind C2-DACA to the guanine without requiring too much strain or puckering of the DNA bases. The energy optimised structure of the formaldehyde-mediated C2-DACA-d(GCG)$_2$ adduct is shown in **Figure 8.8**.

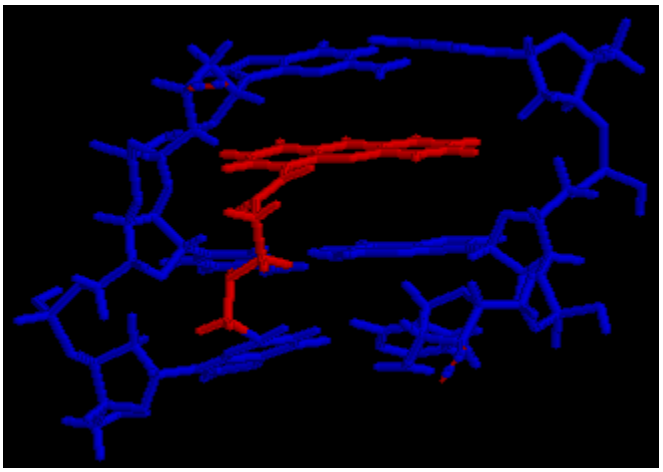


**Figure 8.8** The energy minimised structure of C2-DACA covalently bound to d(GCG)$_2$ (view from minor groove).

The high degree of puckering and twisting of the base pairs can be seen from this view, especially with the bottom base pairs of the DNA. When the side

chain is attached via the methylene bridge it pulls on the attached guanine which in turn twists the corresponding cytosine as a result of the base pair hydrogen bonding. If C2-DACA were to bind in this manner *in vitro*, hydrogen bonding which stabilises the DNA would be seriously disrupted. This may result in the drug bound to only a single strand of DNA, not intercalated and bound to duplex DNA. It is clear from the result in **Figure 8.8** that the side chain of C2-DACA is too short to first intercalate then form a drug-DNA adduct via a methylene bridge.

## 8.4. Side Chain Optimisation

In the following section, computer modelling was employed to determine the optimum length of the side chain in order for the terminal amine to form the formaldehyde-mediated adduct with minimal distortion of the DNA structure. It was also investigated whether the side chain could access the free amine of guanine via the major or minor groove. Thus N-C2-DACA (normal) and R-C2-DACA (reverse) as shown in **Figure 8.9** were used to see which fit better into the model from **Figure 8.8**. The side chain was attached at the 4-position for N-C2-DACA, and at the 8-position for R-C2-DACA.

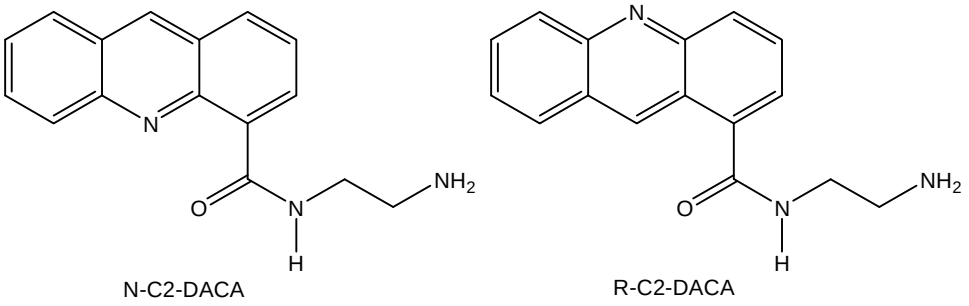



**Figure 8.9** Structures of N-C2-DACA and R-C2-DACA.

Both structures were separately placed into the three base pair model $d(GCG)_2$ in both the major and minor grooves. Models from this experiment illustrated that the length of the side chain was insufficient to reach the preferred binding site regardless of which groove the side chain was in. Although the position of the nitrogen did not appear to be crucial, these models allowed the discovery of the optimum chain length for the drug (e.g. 4 to 5 carbons long).

### 8.4.1. Results and Discussion

In order to investigate the optimal chain length, it is clear that the side chain of C2-DACA needed to be extended by one to three atoms and bound to the DNA $d(GCG)_2$ via the minor groove only (**Figures 8.11** and **8.13**). This was performed for the Schiff base of the drug-formaldehyde complex where a covalent bond was formed from the terminal drug-imine to the guanine amino group. Constraints were again applied only to the DNA, and the drug-DNA adduct was geometry optimised. The geometry optimised drug-DNA adduct of the compound from **Figure 8.10** is bound to $d(GCG)_2$ and shown in **Figure 8.11**. The $d(GCG)_2$ is coloured blue and the drug is highlighted in red.

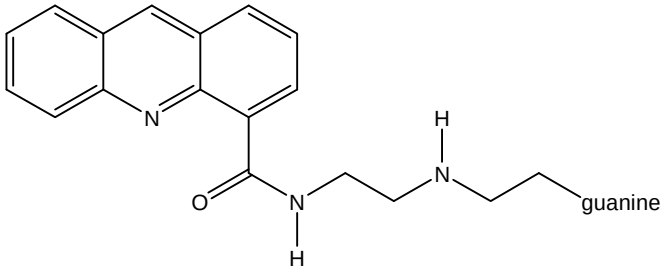



**Figure 8.10** Structure of the drug intercalated and bound to guanine in **Figure 8.11**. This structure retains the N-H in the middle of the side chain.

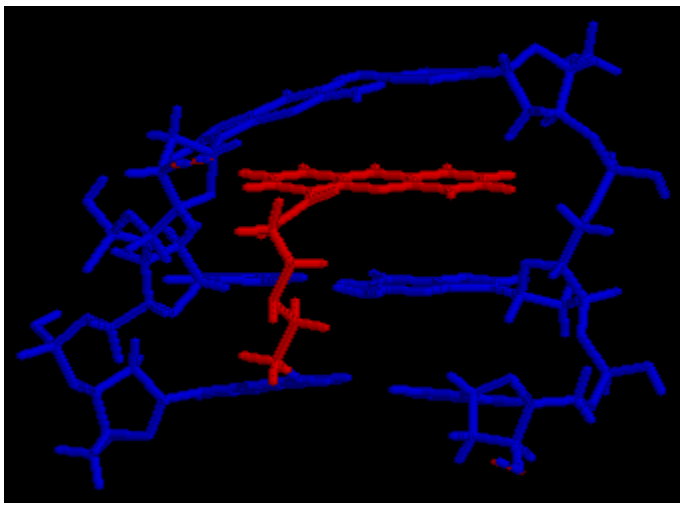


**Figure 8.11** Resultant structure from a geometry optimisation of C2-DACA with two extra carbons added to the chain and bound to $d(GCG)_2$. The drug is highlighted in red, and has the structure as shown in **Figure 8.10**.

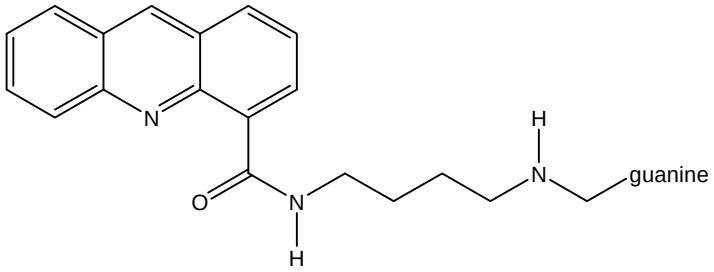



**Figure 8.12** Structure of the drug intercalated and bound to guanine in **Figure 8.13**.

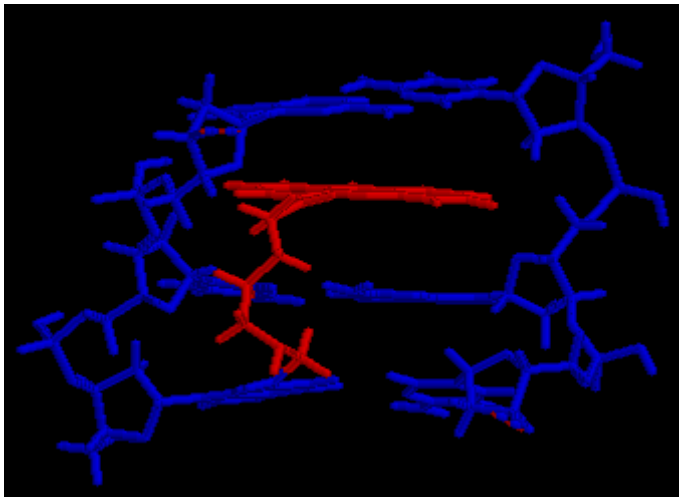
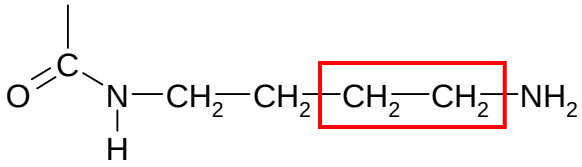


**Figure 8.13** Resultant structure from a geometry optimisation of C4-DACA bound to the 3' guanine of $d(GCG)_2$ via a formaldehyde bridge. The drug is highlighted in red, and has the structure as shown in **Figure 8.12**.

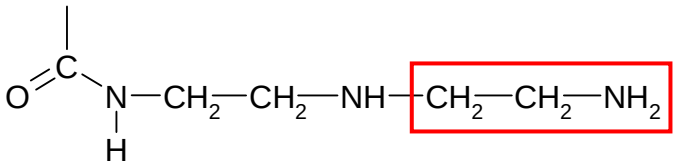
Results from the molecular modelling based on C2-DACA showed that an increase in chain length of approximately two to three atoms (C4-DACA) would produce a more favourable binding interaction. The structure from **Figure 8.12** resulted in minimal disruption of the base pairs to yield the drug-DNA adduct, including reasonable angles and torsion angles in the side chain. This structure was the formaldehyde-mediated C4-DACA-DNA adduct.

Methods for extending the side chain of C2-DACA were required to yield an adduct forming drug. It may be possible to lengthen the side chain of the C2-DACA that had already been synthesised by simply adding to the terminal primary amine. Three main theoretical possibilities for the side chain extension were considered:

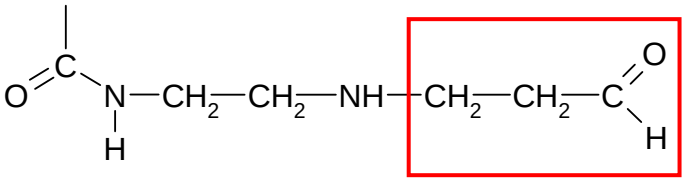
1). Synthesise C2-DACA using 1,4-diaminobutane instead of 1,2-diaminoethane to produce a longer four carbon side chain;



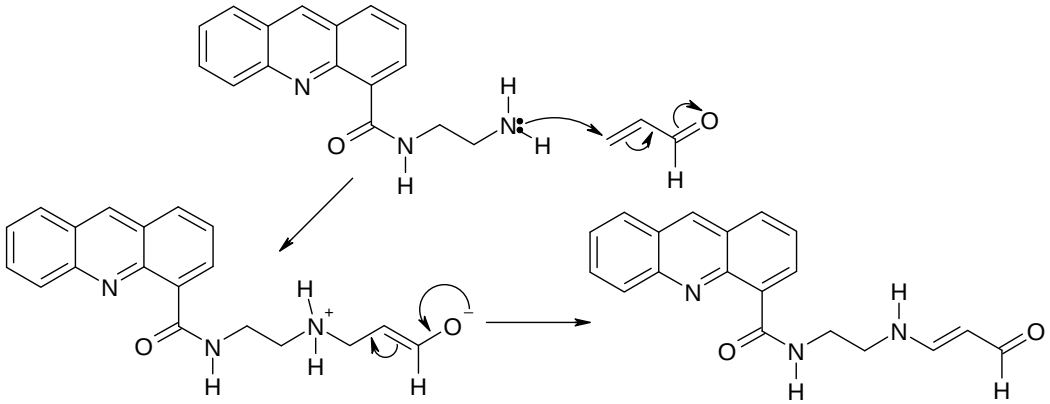
2). Use the original C2-DACA drug and extend the chain synthetically to keep the N-H in the middle of the chain;



3). Use the original drug and one acrolein unit (instead of formaldehyde) in a Michael addition *in vitro*.

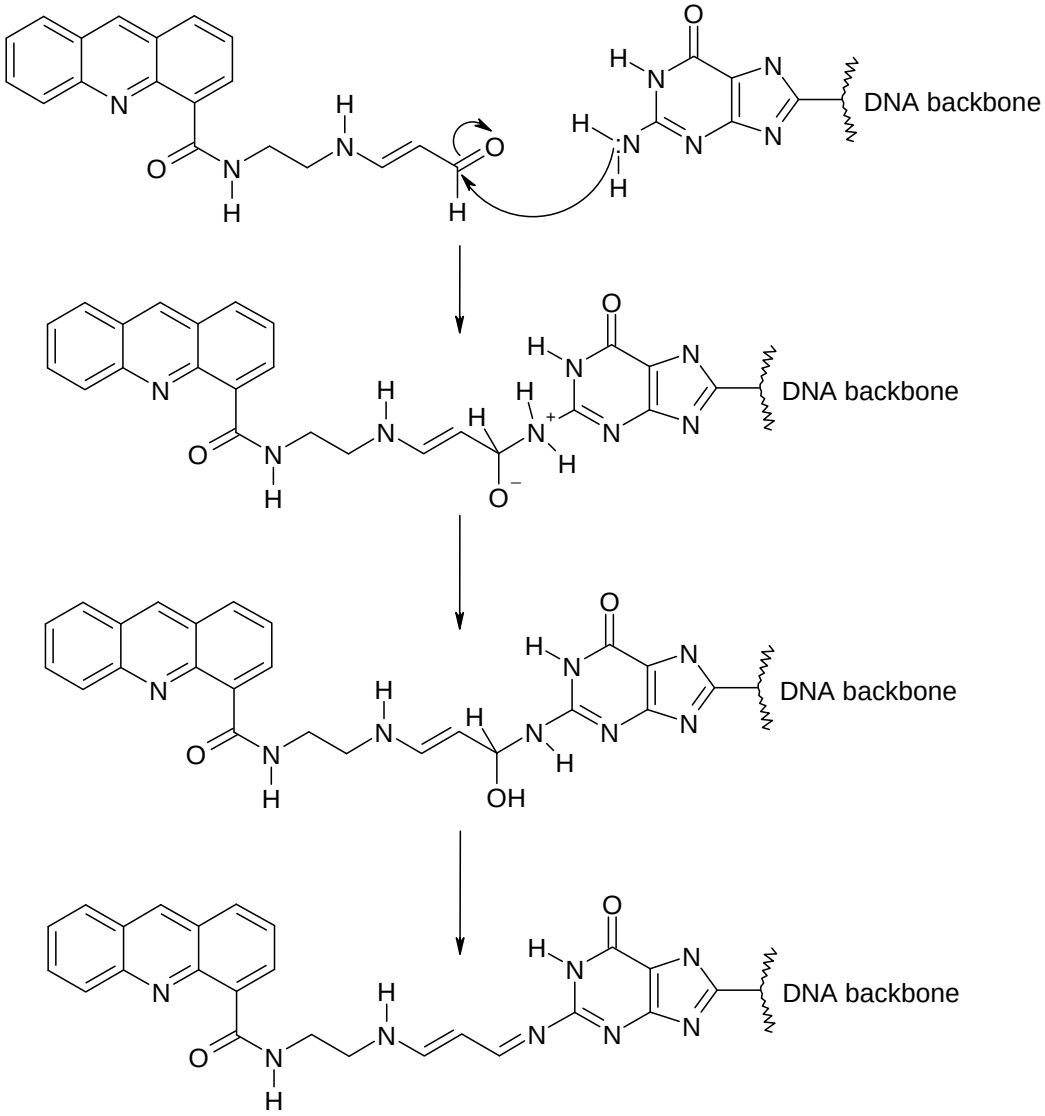


Acrolein is known to react with guanine in animal and human tissue (under physiological conditions) [94, 132]. Thus, for the Michael addition of acrolein to the drug to occur, DNA would have to be absent. The reaction scheme for C2-DACA reacting with acrolein is shown in **Scheme 8.1**. If this reaction proceeds, the use of substituted acroleins or possibly esters could be investigated.



**Scheme 8.1** Michael reaction between C2-DACA and acrolein.

Once C2-DACA has undergone the Michael reaction with acrolein, its desired reaction with DNA is shown in **Scheme 8.2**.




**Scheme 8.2** Reaction of activated drug with guanine.

Molecular modelling on the reaction using acrolein from **Schemes 8.1** and **8.2** gave the structure shown in **Figure 8.14**.

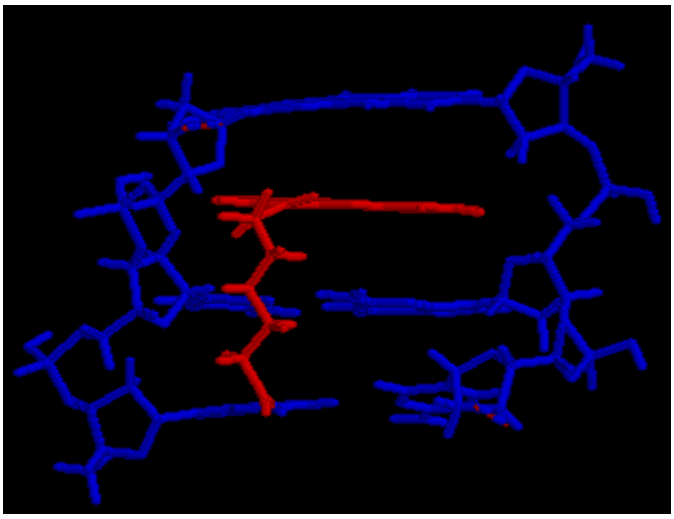


**Figure 8.14** Minimised acrolein-activated C2-DACA bound to the 3' guanine of d(GCG)$_2$ (product from **Scheme 8.2**).

Of all the different length side chains that were modelled, the acrolein-activated C2-DACA-DNA adduct shown in **Figure 8.14** gave the least distortion of the base pair interactions. This structure also shows acceptable bond lengths, angles and torsion angles. This shows that a side chain of five atoms after the carboxamide is sufficiently long to bind to DNA with minimal disruption to the hydrogen bonding. Thus the use of formaldehyde-activated C4-DACA or acrolein-activated C2-DACA would satisfy these requirements.

It was decided that the most practical option was to re-synthesise the drug using 1,4-diaminobutane instead of 1,2-diaminoethane in order to model formaldehyde-activated C4-DACA. This would give complete confidence that the drug was the only compound available to react with and stabilise the double stranded DNA. Also, results from the energetics calculations in **Chapter 6** showed that formaldehyde was the most energetically favourable activator to use. In addition, the reaction with amine-containing drugs such as daunomycin is well documented. Even though the distortion of angles and torsional angles from the modelling of the second option were reasonable, the procedure was dismissed. This was due to the fact that the amino group in the middle of the chain could complicate the interactions with the DNA side chain. The third option of using acrolein *in vitro* was unfeasible as acrolein is a known mutagen that also reacts with guanine to form exocyclic adducts [133] which could compete with the desired reaction. There would also be a problem when attempting to use acrolein

*in vivo* as it would have to somehow be administered along with the drug. Using formaldehyde instead of acrolein bypasses this problem as there are already trace amounts of formaldehyde in cells, with higher concentrations in cancer cells. Thus, the synthesis of DACA with a four-carbon long side chain as reported in **Chapter 2** was undertaken using 1,4-diaminobutane to create the longer side chain. This compound is referred to as C4-DACA (**Figure 8.12**).

## 8.5. Conclusions

Biochemical studies determined unequivocally that neither C2-DACA nor its dimer form covalent drug-DNA adducts in the presence of formaldehyde. Molecular modelling of C2-DACA and related molecules intercalated and/or bound to DNA allowed the binding sites to be studied in greater detail. The length of the C2-DACA side chain proved to be insufficient to reach the preferred binding site. It was confirmed that the side chain of C2-DACA was under considerable steric hindrance once intercalated and covalently bound to DNA.

Extending the chain by two to three carbons demonstrated more favourable interactions between the ligand and DNA under simulated conditions. A greater understanding of the interactions between the ligand and DNA enabled the construction of a new derivative, C4-DACA, which has the potential to form a covalent drug-DNA adduct in the presence of formaldehyde. This derivative of DACA involves a four carbon side chain with a terminal free amino group. As seen in **Chapter 5**, the new compound that was created from this process was observed to interrupt the transcription process by forming formaldehyde-mediated adducts with DNA. Formaldehyde-activated C4-DACA was also found to form adducts at strikingly similar positions to that of formaldehyde-activated pixantrone and formaldehyde-activated mitoxantrone. There is also potential, which was not explored here, to work on compounds that contain secondary amines. It could be advantageous to investigate the optimisation of formaldehyde mediated drug-DNA adducts for derivatives of other clinically important drugs such as mitoxantrone.

# CHAPTER 9. CONCLUSIONS AND FUTURE DIRECTIONS

## 9.1. Overview

The aim of this work was to develop a relatively simple anti-cancer drug. This model drug was based on an acridine framework with an appropriate side chain that would enable covalent binding to DNA in the presence of formaldehyde, similar to Adriamycin and mitoxantrone. This objective has been substantially achieved.

## 9.2. Conclusions

### 9.2.1. Synthesis

The synthesis of *N*-(2-aminoethyl)acridine-4-carboxamide.oxalate salt, C2-DACA, the C2-DACA dimer, and *N*-(4-aminobutyl)acridine-4-carboxamide.oxalate salt, C4-DACA, were characterised and these compounds identified using a range of nuclear magnetic resonance spectroscopic methods and electrospray mass spectrometry. The full assignment was completed for all three synthesised drugs in d6-DMSO and a full proton assignment of C4-DACA in $H_2O/D_2O$.

### 9.2.2. Crosslinking Assays of C2-DACA and its Dimer

The crosslinking assays showed that C2-DACA showed at best only weak binding to DNA in the presence of formaldehyde, and did not stabilise ds DNA. The C2-DACA dimer showed no ability to bind to or stabilise ds DNA in the presence of formaldehyde. Importantly, this led to the re-design of the covalent DNA binding compounds using the molecular modelling methods featured in the Insight and Discover software package.

### 9.2.3. Molecular Modelling

Investigation into a range of published crystal structures with acridine-like intercalators identified the use of a 3 base pair DNA sequence with the ability to insert different compounds into the active site for modelling purposes. Investigations into the effects of binding via the major or the minor groove of the DNA resulted in the requirement for the side chain of the drug to be in the minor groove in order to form a covalent bond with guanine. The range of possible compounds considered eventually led to C4-DACA being deemed the most suitable candidate for synthesis and for subsequent biochemical and NMR studies.

### 9.2.4. *In Vitro* Crosslinking Assay of C4-DACA

The re-designed drug (C4-DACA) was found to form drug-DNA adducts when in the presence of formaldehyde and to weakly stabilise the double helix. There was a distinct difference in migration of the ds DNA between samples with and without drug. An electromobility shift was attributed to the substantial amount of drug attached to the numerous binding sites on the DNA. A range of experimental conditions were investigated to enhance adduct formation, including concentration of drug and formaldehyde, time allowed for activation (pre-reaction), temperature dependence, pH, and phenol/chloroform extraction of unbound drug. The actual adduct sites were investigated using an *in vitro* transcription assay.

### 9.2.5. *In Vitro* Transcription Assay of C4-DACA

The ability of C4-DACA to form drug-DNA adducts, and the site of these adducts, was investigated using an *in vitro* transcription assay. Formaldehyde-mediated drug-DNA adducts produced truncated transcripts, confirming that formaldehyde-mediated C4-DACA-DNA adducts disrupted the progression of RNA polymerase during transcription. The amount of truncated transcripts increased with an increase in drug concentration in a sequence-dependent manner, with the sequence-selective blockage of RNA polymerase at and immediately prior to CpG and CpA dinucleotide sequences. Formaldehyde-activated C4-DACA was also found to form adducts at strikingly similar positions to that of

formaldehyde-activated pixantrone and formaldehyde-activated mitoxantrone, indicating that C4-DACA probably forms formaldehyde mediated adducts with DNA via the exocyclic N2 amino group of guanine. The intricacies of the energy involved in this formaldehyde activation and consequent adduct formation were performed using a simplified model by *ab initio* calculations.

### 9.2.6. Energy Calculations

Advanced *ab initio* calculations were able to identify the energetics of the Schiff base and consequent aminal formation reactions. Among other results, these calculations were able to determine that formaldehyde rather than acetaldehyde was a better candidate for Schiff base formation. Specifically, the calculations gave great insight into the likelihood of which reactions would proceed, the results from which were considered when deciding which reactants to use for biochemical and NMR studies.

Energy calculations of model reactions using the advanced Gaussian 09 method showed that the overall reaction was exothermic. Depending on the starting materials, most of this energy comes from the formation of the Schiff base, with an additional smaller amount from aminal formation. An increase in energy was found to be negligible for side chains longer than ethylamine. The solvent effect calculations using the AMPAC 7.0 method showed that the solvent did not have a large effect on the energies of the reaction. These energy calculations suggested that if the drug has the opportunity to attach covalently to DNA, it may have a limited half-life which would in turn broaden the peaks identified by NMR.

### 9.2.7. NMR Investigation of the Formaldehyde-mediated C4-DACA-d(GACGTC)$_2$ Complex

An assignment of d(GACGTC)$_2$ was completed (excluding H5' and H5" protons) in $H_2O/D_2O$ at 300 K. There was evidence of drug-DNA interactions from the spectrum of C4-DACA/d(GACGTC)$_2$ in $H_2O/D_2O$ as peak broadening was observed. It was apparent that there was a considerable disruption of the DNA structure from analysis of the multiple spectra of C4-DACA/d(GACGTC)$_2$ in

$H_2O/D_2O$. Compared to the spectrum of DNA alone, there were two aromatic peaks without any NOE cross peaks (G1 and G4) and also no NOE peaks from the drug aromatic protons to the side chain. It is therefore reasonable to assume that the drug was in dynamic exchange with the DNA, thereby disrupting the ability to form double stranded DNA.

Addition of formaldehyde to the solution of C4-DACA/d$(GACGTC)_2$ in $H_2O/D_2O$ proved to be inconclusive. NMR studies of the formaldehyde-mediated C4-DACA-DNA adduct using a hexamer of DNA was not successful. It was suspected that in addition to Schiff base formation that the excess of formaldehyde reacted with the DNA. Spectra of the drug/DNA/formaldehyde solution were very convoluted and broad, and missing a number of DNA peaks, indicating that the integrity of the sample was destroyed. Results from a standard $^1H$ NMR of stock formaldehyde-$d_2$ solution showed an exceedingly complicated spectrum consisting of monomers, dimers, polymers and cyclic derivatives (paraformaldehyde) and hydrated species of formaldehyde, ultimately precluding the use of NMR for the analysis of oligonucleotide-formaldehyde reactions.

## 9.3. Significant Outcomes

The work presented in this thesis has proven to be particularly useful to other projects within our laboratory by giving insight into drug constituents and binding site requirements. Our laboratory has been investigating a number of anti-cancer drugs with the ability to form formaldehyde-mediated drug-DNA adducts. The realisation of this ability by Adriamycin [4] encouraged the investigation into similarly equipped compounds. Mitoxantrone has been studied extensively within our laboratory [12, 13, 59, 134] however, formation of the formaldehyde generated drug-DNA adducts occurs at biologically irrelevant drug concentrations [61]. The results from this thesis showed that the incorporation of a free primary amine into a compound similar to mitoxantrone (mitoxantrone contains a secondary amine in each side chain) would be more energetically favourable. This resulted in the investigation of pixantrone as a transcription inhibitor as it contains a terminal primary amine on each side chain and this has yielded exceedingly positive results [61, 112]. When compared to mitoxantrone, pixantrone

demonstrates a 10- to 100-fold greater proclivity to generate adducts at equimolar formaldehyde and drug concentrations [61].

## 9.4. Future Directions

The work presented in this thesis showed that simple drugs such as C4-DACA can bind to DNA in the presence of formaldehyde, although weakly and reversibly. It is therefore theoretically possible to increase the binding affinity of such drugs by modifying the structure by adding substituents at the acridine-9 position, and perhaps including a short side chain with the ability to bind via the minor groove of DNA. The addition of a second four-carbon long side chain could mimic the potential crosslinking abilities of pixantrone and mitoxantrone. Another angle for possible investigation would be adding a heteronuclear side chain with the specific ability to hydrogen bond to sites within the minor groove in a bid to aid stabilisation of the intercalated and covalently bound complex.

Throughout the time this project was undertaken, there have been significant advances in computational abilities. Therefore, theoretical calculations on secondary amino drugs such as mitoxantrone should now also be undertaken to determine how this clinically useful anti-cancer drug could be activated to bind more effectively with DNA.

The concentration of the formaldehyde-drug-DNA NMR sample was based on the results from MACGAMESS calculations. These initial calculations indicated that Schiff base formation was only slightly endothermic, and that the subsequent aminal formation was exothermic, inferring that the use of an excess of formaldehyde could encourage the forward reaction. However, following detailed analysis using Gaussian 09 calculations (which were performed after the NMR experimentation), it was discovered that the use of such a large excess of formaldehyde was not necessary as the formation of the Schiff base is actually substantially exothermic, and the aminal formation slightly endothermic. This excess of formaldehyde employed may have compromised the drug-DNA sample, probably by creating both inter-strand and intra-strand crosslinks in the DNA, thus losing symmetry within the DNA. In hindsight, results from the crosslinking assays showed evidence of formaldehyde-induced ds DNA in a linear fashion, with almost 10% ds DNA at 40 mM formaldehyde. The drug-DNA NMR sample

was made up to 70 mM formaldehyde which would imply that approximately 15-20% of the DNA would be crosslinked by formaldehyde alone. A 1:1:1 ratio of formaldehyde/drug/DNA could be employed in the future in order to model the binding site of the formaldehyde-mediated drug-DNA adduct. This knowledge alone is a significant start towards creating successful NMR conditions and for modelling the formaldehyde-mediated drug-DNA adduct via molecular modelling software such as Amber.

Finally, the work presented in this thesis and the ideas generated have proven to be useful to other existing projects within our laboratory and potential future projects through the insight on drug constituents and binding site requirements. In particular, it has contributed to the examination of anti-cancer drugs that contain a free primary amine such as pixantrone, and the reaction of that compound to form long-lived covalent DNA complexes with enhanced biological response [61].

## CHAPTER 10. REFERENCES[126]


1. Arcamone, F., *Doxorubicin: anticancer antibiotics*. **1981**. New York: Academic Press.
2. Cutts, S. M.; Phillips, D. R. Use of oligonucleotides to define the site of interstrand cross-links induced by Adriamycin. *Nucleic Acids Res.* **1995**, 23(13), 2450-6.
3. Cullinane, C.; Phillips, D. R. Thermal stability of DNA adducts induced by cyanomorpholinoadriamycin *in vitro*. *Nucleic Acids Res.* **1993**, 21(8), 1857-62.
4. Taatjes, D. J.; Gaudiano, G.; Koch, T. H.; Fenick, D. J. Production of formaldehyde and DNA-adriamycin or DNA-daunomycin adducts, initiated through redox chemistry of dithiothreitol/iron, xanthine oxidase/NADH/iron, or glutathione/iron. *Chem. Res. Toxicol.* **1997**, 10(9), 953-61.
5. Taatjes, D. J.; Fenick, D. J.; Koch, T. H. Epidoxoform: a hydrolytically more stable anthracycline-formaldehyde conjugate toxic to resistant tumor cells. *J. Med. Chem.* **1998**, 41(8), 1306-14.
6. Wang, A. H.; Gao, Y. G.; Liaw, Y. C.; Li, Y. K. Formaldehyde cross-links daunorubicin and DNA efficiently: HPLC and X-ray diffraction studies. *Biochemistry* **1991**, 30(16), 3812-5.
7. Cullinane, C.; van Rosmalen, A.; Phillips, D. R. Does Adriamycin induce interstrand cross-links in DNA? *Biochemistry* **1994**, 33(15), 4632-8.
8. Podell, E. R.; Harrington, D. J.; Taatjes, D. J.; Koch, T. H. Crystal structure of epidoxorubicin-formaldehyde virtual crosslink of DNA and evidence for its formation in human breast-cancer cells. *Acta Crystallogr. D Biol. Crystallogr.* **1999**, 55(Pt 9), 1516-23.
9. Denny, W. A. Acridine derivatives as chemotherapeutic agents. *Curr. Med. Chem.* **2002**, 9(18), 1655-65.
10. Atwell, G. J.; Cain, B. F.; Baguley, B. C.; Finlay, G. J.; Denny, W. A. Potential antitumor agents. 43. Synthesis and biological activity of dibasic 9-aminoacridine-4-carboxamides, a new class of antitumor agent. *J. Med. Chem.* **1984**, 27(11), 1481-5.

11. Alison, M. R., *The cancer handbook*, ed. M. Alison. Vol. 2. **2002**. London: Nature.
12. Panousis, C.; Phillips, D. R. DNA sequence specificity of mitoxantrone. *Nucleic Acids Res.* **1994**, 22(8), 1342-5.
13. Parker, B. S.; Cutts, S. M.; Cullinane, C.; Phillips, D. R. Formaldehyde activation of mitoxantrone yields CpG and CpA specific DNA adducts. *Nucleic Acids Res.* **2000**, 28(4), 982-90.
14. Huang, H.; Zhu, L.; Reid, B. R.; Drobny, G. P.; Hopkins, P. B. Solution structure of a cisplatin-induced DNA interstrand cross-link. *Science* **1995**, 270(5243), 1842-5.
15. Williams, H. E.; Searle, M. S. Structure, dynamics and hydration of the nogalamycin-d(ATGCAT)2Complex determined by NMR and molecular dynamics simulations in solution. *J. Mol. Biol.* **1999**, 290(3), 699-716.
16. Searle, M. S.; Maynard, A. J.; Williams, H. E. DNA recognition by the anthracycline antibiotic respinomycin D: NMR structure of the intercalation complex with d(AGACGTCT)2. *Org. Biomol. Chem.* **2003**, 1(1), 60-6.
17. Searle, M. S.; Bicknell, W. Interaction of the anthracycline antibiotic nogalamycin with the hexamer duplex d(5'-GACGTC)2. An NMR and molecular modelling study. *Eur. J. Biochem.* **1992**, 205(1), 45-58.
18. Searle, M. S. Binding of quinomycin antibiotic UK-65,662 to DNA: 1H-NMR studies of drug-induced changes in DNA conformation in complexes with $d(ACGT)_2$ and $d(GACGTC)_2$. *Biochem. J.* **1994**, 304(Pt 3), 967-79.
19. Williams, L. D.; Egli, M.; Qi, G.; Bash, P.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Frederick, C. A. Structure of nogalamycin bound to a DNA hexamer. *Proc. Natl. Acad. Sci. USA* **1990**, 87(6), 2225-9.
20. Fox, K. R.; Brassett, C.; Waring, M. J. Kinetics of dissociation of nogalamycin from DNA: comparison with other anthracycline antibiotics. *Biochim. Biophys. Acta* **1985**, 840(3), 383-92.
21. Arora, S. K. Molecular structure, absolute stereochemistry, and interactions of nogalamycin, a DNA-binding anthracycline antitumor antibiotic. *J. Am. Chem. Soc.* **1983**, 105(5), 1328-32.
22. Ubukata, M.; Osada, H.; Kudo, T.; Isono, K. Respinomycins A1, A2, B, C and D, a novel group of anthracycline antibiotics. I. Taxonomy,

fermentation, isolation and biological activities. *J. Antibiot. (Tokyo)* **1993**, 46(6), 936-41.

23. Arcamone, F.; Cassinelli, G.; Fantini, G.; Grein, A.; Orezzi, P.; Pol, C.; Spalla, C. Adriamycin, 14-hydroxydaunomycin, a new antitumor antibiotic from *S. Peucetius var. caesius*. *Biotechnol. Bioeng.* **1969**, 11(6), 1101-10.
24. Arcamone, P.; Canssinelli, G.; Franceschi, G.; Orezzi, P.; Mondelli, R. The total absolute configuration of daunomycin. *Tetrahedron Lett.* **1968**, 9(30), 3353-6.
25. Leng, F.; Savkur, R.; Fokt, I.; Przewloka, T.; Priebe, W.; Chaires, J. B. Base specific and regioselective chemical cross-linking of Daunorubicin to DNA. *J. Am. Chem. Soc.* **1996**, 118(20), 4731-8.
26. van Rosmalen, A.; Cullinane, C.; Cutts, S. M.; Phillips, D. R. Stability of Adriamycin-induced DNA adducts and interstrand crosslinks. *Nucleic Acids Res.* **1995**, 23(1), 42-50.
27. Minotti, G.; Menna, P.; Salvatorelli, E.; Cairo, G.; Gianni, L. Anthracyclines: Molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. *Pharmacol. Rev.* **2004**, 56(2), 185-229.
28. Zunino, F.; Capranico, G. DNA topoisomerase II as the primary target of anti-tumor anthracyclines. *Anti Canc. Drug Des.* **1990**, 5(4), 307-17.
29. Cortes-Funes, H.; Coronado, C. Role of anthracyclines in the era of targeted therapy. *Cardiovasc. Toxicol.* **2007**, 7(2), 56-60.
30. Kennedy, S.; DiCesare, J. C.; Sheaff, R. J. Topoisomerase I/II inhibition by a novel naphthoquinone containing a modified anthracycline ring system. *Biochem. Biophys. Res. Commun.* **2011**, 408(1), 94-8.
31. Bigioni, M.; Zunino, F.; Capranico, G. Base mutation analysis of topoisomerase II-idarubicin-DNA ternary complex formation. Evidence for enzyme subunit cooperativity in DNA cleavage. *Nucleic Acids Res.* **1994**, 22(12), 2274-81.
32. Zunino, F.; Pratesi, G.; Perego, P. Role of the sugar moiety in the pharmacological activity of anthracyclines: development of a novel series of disaccharide analogs. *Biochem. Pharmacol.* **2001**, 61(8), 933-8.
33. Gao, Y. G.; Liaw, Y. C.; Li, Y. K.; van der Marel, G. A.; van Boom, J. H.; Wang, A. H. Facile formation of a crosslinked adduct between DNA and

the daunorubicin derivative MAR70 mediated by formaldehyde: molecular structure of the MAR70-d(CGTnACG) covalent adduct. *Proc. Natl. Acad. Sci. USA* **1991**, 88(11), 4845-9.

34. Atwell, G. J.; Rewcastle, G. W.; Baguley, B. C.; Denny, W. A. Potential antitumor agents. 50. *In vivo* solid-tumor activity of derivatives of *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide. *J. Med. Chem.* **1987**, 30(4), 664-9.

35. 2003 [cited; Available from: http://www.nlm.nih.gov/medlineplus/druginfo/uspdi/500109.html.

36. Haldane, A.; Holdaway, K. M.; Finlay, G. J.; Baguley, B. C. Cytokinetic differences in the action of *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide as compared with that of amsacrine and doxorubicin. *Cancer Chemother. Pharmacol.* **1993**, 32(6), 463-70.

37. Wakelin, L. P.; Adams, A.; Denny, W. A. Kinetic studies of the binding of acridinecarboxamide topoisomerase poisons to DNA: implications for mode of binding of ligands with uncharged chromophores. *J. Med. Chem.* **2002**, 45(4), 894-901.

38. Todd, A. K.; Adams, A.; Thorpe, J. H.; Denny, W. A.; Wakelin, L. P.; Cardin, C. J. Major groove binding and 'DNA-induced' fit in the intercalation of a derivative of the mixed topoisomerase I/II poison *N*-(2-(dimethylamino)ethyl)acridine-4-carboxamide (DACA) into DNA: X-ray structure complexed to d(CG(5-BrU)ACG)$_2$ at 1.3-Å resolution. *J. Med. Chem.* **1999**, 42(4), 536-40.

39. Fortune, J. M.; Velea, L.; Graves, D. E.; Utsugi, T.; Yamada, Y.; Osheroff, N. DNA topoisomerases as targets for the anticancer drug TAS-103: DNA interactions and topoisomerase catalytic inhibition. *Biochemistry* **1999**, 38(47), 15580-6.

40. Adams, A.; Guss, J. M.; Collyer, C. A.; Denny, W. A.; Wakelin, L. P. Crystal structure of the topoisomerase II poison 9-amino-[*N*-(2-dimethylamino)ethyl]acridine-4-carboxamide bound to the DNA hexanucleotide d(CGTACG)$_2$. *Biochemistry* **1999**, 38(29), 9221-33.

41. Caponigro, F.; Dittrich, C.; Sorensen, J. B.; Schellens, J. H.; Duffaud, F.; Paz Ares, L.; Lacombe, D.; de Balincourt, C.; Fumoleau, P. Phase II study

of XR 5000, an inhibitor of topoisomerases I and II, in advanced colorectal cancer. *Eur. J. Cancer* **2002**, 38(1), 70-4.

42. Adams, A.; Guss, J. M.; Collyer, C. A.; Denny, W. A.; Wakelin, L. P. A novel form of intercalation involving four DNA duplexes in an acridine-4-carboxamide complex of d(CGTACG)$_2$. *Nucleic Acids Res.* **2000**, 28(21), 4244-53.
43. Haldane, A.; Finlay, G. J.; Hay, M. P.; Denny, W. A.; Baguley, B. C. Cellular uptake of *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide (DACA). *Anti Canc. Drug Des.* **1999**, 14(3), 275-80.
44. Adams, A.; Guss, J. M.; Denny, W. A.; Wakelin, L. P. Crystal structure of 9-amino-*N*-[2-(4-morpholinyl)ethyl]-4-acridinecarboxamide bound to d(CGTACG)$_2$: implications for structure-activity relationships of acridinecarboxamide topoisomerase poisons. *Nucleic Acids Res.* **2002**, 30(3), 719-25.
45. Spicer, J. A.; Finlay, G. J.; Baguley, B. C.; Velea, L.; Graves, D. E.; Denny, W. A. 5,7-Disubstituted analogues of the mixed topoisomerase I/II poison *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide (DACA): DNA binding and patterns of cytotoxicity. *Anti Canc. Drug Des.* **1999**, 14(1), 37-45.
46. Twelves, C.; Campone, M.; Coudert, B.; van den Bent, M.; de Jonge, M.; Dittrich, C.; Rampling, R.; Sorio, R.; Lacombe, D.; de Balincourt, C.; Fumoleau, P. Phase II study of XR5000 (DACA) administered as a 120-h infusion in patients with recurrent glioblastoma multiforme. *Ann. Oncol.* **2002**, 13(5), 777-80.
47. Dittrich, C.; Coudert, B.; Paz-Ares, L.; Caponigro, F.; Salzberg, M.; Gamucci, T.; Paoletti, X.; Hermans, C.; Lacombe, D.; Fumoleau, P. Phase II study of XR 5000 (DACA), an inhibitor of topoisomerase I and II, administered as a 120-h infusion in patients with non-small cell lung cancer. *Eur. J. Cancer* **2003**, 39(3), 330-4.
48. Propper, D. J.; de Bono, J.; Saleem, A.; Ellard, S.; Flanagan, E.; Paul, J.; Ganesan, T. S.; Talbot, D. C.; Aboagye, E. O.; Price, P.; Harris, A. L.; Twelves, C. Use of positron emission tomography in pharmacokinetic studies to investigate therapeutic advantage in a phase I study of 120-hour intravenous infusion XR5000. *J. Clin. Oncol.* **2003**, 21(2), 203-10.

49. Saleem, A.; Harte, R. J.; Matthews, J. C.; Osman, S.; Brady, F.; Luthra, S. K.; Brown, G. D.; Bleehen, N.; Connors, T.; Jones, T.; Price, P. M.; Aboagye, E. O. Pharmacokinetic evaluation of *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide in patients by positron emission tomography. *J. Clin. Oncol.* **2001**, 19(5), 1421-9.

50. Neale, M. H.; Charlton, P. A.; Cree, I. A. *Ex vivo* activity of XR5000 against solid tumors. *Anti-Cancer Drugs* **2000**, 11(6), 471-8.

51. Bridewell, D. J.; Finlay, G. J.; Baguley, B. C. Topoisomerase I/II selectivity among derivatives of *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide (DACA). *Anti Canc. Drug Des.* **2001**, 16(6), 317-24.

52. Spicer, J. A.; Gamage, S. A.; Atwell, G. J.; Finlay, G. J.; Baguley, B. C.; Denny, W. A. Dimeric analogues of non-cationic tricyclic aromatic carboxamides are a new class of cytotoxic agents. *Anti Canc. Drug Des.* **1999**, 14(3), 281-9.

53. Padget, K.; Stewart, A.; Charlton, P.; Tilby, M. J.; Austin, C. A. An investigation into the formation of *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide (DACA) and 6-[2-(dimethylamino)ethylamino]-3-hydroxy-7H-indeno[2, 1-C]quinolin-7-one dihydrochloride (TAS-103) stabilised DNA topoisomerase I and II cleavable complexes in human leukaemia cells. *Biochem. Pharmacol.* **2000**, 60(6), 817-21.

54. Denny, W. A.; Baguley, B. C. Dual topoisomerase I/II inhibitors in cancer therapy. *Curr. Top. Med. Chem.* **2003**, 3(3), 339-53.

55. Twelves, C. J.; Gardner, C.; Flavin, A.; Sludden, J.; Dennis, I.; de Bono, J.; Beale, P.; Vasey, P.; Hutchison, C.; Macham, M. A.; Rodriguez, A.; Judson, I.; Bleehen, N. M. Phase I and pharmacokinetic study of DACA (XR5000): a novel inhibitor of topoisomerase I and II. CRC Phase I/II Committee. *Br. J. Cancer* **1999**, 80(11), 1786-91.

56. Finlay, G. J.; Marshall, E.; Matthews, J. H.; Paull, K. D.; Baguley, B. C. *In vitro* assessment of *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide, a DNA-intercalating antitumour drug with reduced sensitivity to multidrug resistance. *Cancer Chemother. Pharmacol.* **1993**, 31(5), 401-6.

57. Adams, A.; Guss, J. M.; Collyer, C. A.; Denny, W. A.; Prakash, A. S.; Wakelin, L. P. Acridinecarboxamide topoisomerase poisons: structural and kinetic studies of the DNA complexes of 5-substituted 9-amino-(*N*-(2-

dimethylamino)ethyl)acridine-4-carboxamides. *Mol. Pharmacol.* **2000**, 58(3), 649-58.

58. Cornbleet, M. A.; Stuart-Harris, R. C.; Smith, I. E.; Coleman, R. E.; Rubens, R. D.; McDonald, M.; Mouridsen, H. T.; Rainer, H.; van Oosterom, A. T.; Smyth, J. F. Mitoxantrone for the treatment of advanced breast cancer: single-agent therapy in previously untreated patients. *Eur. J. Cancer Clin. Oncol.* **1984**, 20(9), 1141-6.
59. Parker, B. S.; Cullinane, C.; Phillips, D. R. Formation of DNA adducts by formaldehyde-activated mitoxantrone. *Nucleic Acids Res.* **1999**, 27(14), 2918-23.
60. Feofanov, A.; Sharonov, S.; Fleury, F.; Kudelina, I.; Nabiev, I. Quantitative confocal spectral imaging analysis of mitoxantrone within living K562 cells: intracellular accumulation and distribution of monomers, aggregates, naphtoquinoxaline metabolite, and drug-target complexes. *Biophys. J.* **1997**, 73(6), 3328-36.
61. Evison, B. J.; Mansour, O. C.; Menta, E.; Phillips, D. R.; Cutts, S. M. Pixantrone can be activated by formaldehyde to generate a potent DNA adduct forming agent. *Nucleic Acids Res.* **2007**, 35(11), 3581-9.
62. Routier, S.; Bernier, J. L.; Catteau, J. P.; Riou, J. F.; Bailly, C. Synthesis, DNA binding, topoisomerase II inhibition and cytotoxicity of two guanidine-containing anthracene-9,10-diones. *Anti Canc. Drug Des.* **1998**, 13(5), 407-15.
63. Hande, K. R. Clinical applications of anticancer drugs targeted to topoisomerase II. *Biochim. Biophys. Acta* **1998**, 1400(1-3), 173-84.
64. Engert, A.; Herbrecht, R.; Santoro, A.; Zinzani, P. L.; Gorbatchevsky, I. EXTEND PIX301: a phase III randomized trial of pixantrone versus other chemotherapeutic agents as third-line monotherapy in patients with relapsed, aggressive non-Hodgkin's lymphoma. *Clin. Lymphoma Myeloma* **2006**, 7(2), 152-4.
65. Krapcho, A. P.; Petry, M. E.; Getahun, Z.; Landi, J. J., Jr.; Stallman, J.; Polsenberg, J. F.; Gallagher, C. E.; Maresch, M. J.; Hacker, M. P.; Giuliani, F. C.; et al. 6,9-Bis[(aminoalkyl)amino]benzo[g]isoquinoline-5,10-diones. A novel class of chromophore-modified antitumor

anthracene-9,10-diones: synthesis and antitumor evaluations. *J. Med. Chem.* **1994**, 37(6), 828-37.

66. Zeman, S. M.; Phillips, D. R.; Crothers, D. M. Characterization of covalent Adriamycin-DNA adducts. *Proc. Natl. Acad. Sci. USA* **1998**, 95(20), 11561-5.

67. Cordes, E. H.; Jencks, W. P. On the mechanism of schiff base formation and hydrolysis. *J. Am. Chem. Soc.* **1962**, 84(5), 832-7.

68. Zamble, D. B.; Lippard, S. J. Cisplatin and DNA repair in cancer chemotherapy. *Trends Biochem. Sci.* **1995**, 20(10), 435-9.

69. Sherman, S. E.; Lippard, S. J. Structural aspects of platinum anticancer drug interactions with DNA. *Chem. Rev.* **1987**, 87(5), 1153-81.

70. Malinge, J. M.; Sip, M.; Blacker, A. J.; Lehn, J. M.; Leng, M. Formation of a DNA monofunctional cis-platinum adduct cross-linking the intercalating drug *N*-methyl-2,7-diazapyrenium. *Nucleic Acids Res.* **1990**, 18(13), 3887-91.

71. Sip, M.; Leng, M., *DNA, cis-platinum and intercalators: catalytic activity of the DNA double helix*. Nucleic Acids Mol. Biol., ed. F. Eckstein; D. Lilley. Vol. 7. **1993**. Berlin: Springer-Verlag. 1-15.

72. Milkevitch, M.; Shirley, B. W.; Brewer, K. J. Mixed-metal polymetallic platinum complexes designed to interact with DNA. *Inorg. Chim. Acta* **1997**, 264(1), 249-56.

73. Takahara, P. M.; Rosenzweig, A. C.; Frederick, C. A.; Lippard, S. J. Crystal structure of double-stranded DNA containing the major adduct of the anticancer drug cisplatin. *Nature* **1995**, 377(6550), 649-52.

74. Rosenberg, B. Fundamental studies with cisplatin. *Cancer* **1985**, 55(10), 2303-l6.

75. Bradford, C. W., *Handbook on toxicity of inorganic compounds*, ed. H. G. Seiler; H. Sigel; A. Sigel. **1988**. New York: Dekker.

76. Reedijk, J. The mechanism of action of platinum anti-tumor drugs. *Pure AppI. Chem.* **1987**, 59(2), 81-192.

77. Cullinane, C.; Phillips, D. R. Induction of stable transcriptional blockage sites by Adriamycin: GpC specificity of apparent Adriamycin-DNA adducts and dependence on iron(III) ions. *Biochemistry* **1990**, 29(23), 5638-46.

78. Cullinane, C. M., *Detection and characterisation of Adriamycin-DNA adducts*, in *Biochemistry*. 1994, La Trobe University: Bundoora.

79. Zhang, H.; Gao, Y. G.; van der Marel, G. A.; van Boom, J. H.; Wang, A. H. Simultaneous incorporations of two anticancer drugs into DNA. The structures of formaldehyde-cross-linked adducts of daunorubicin-d(CG(araC)GCG) and doxorubicin-d(CA(araC)GTG) complexes at high resolution. *J. Biol. Chem.* **1993**, 268(14), 10095-101.

80. Thorndike, J.; Beck, W. S. Production of formaldehyde from N5-methyltetrahydrofolate by normal and leukemic leukocytes. *Cancer Res.* **1977**, 37(4), 1125-32.

81. Edwards, S. W.; Swan, T. F. Regulation of superoxide generation by myeloperoxidase during the respiratory burst of human neutrophils. *Biochem. J.* **1986**, 237(2), 601-4.

82. Perrin, L. C.; Cullinane, C.; Kimura, K.; Phillips, D. R. Barminomycin forms GC-specific adducts and virtual interstrand crosslinks with DNA. *Nucleic Acids Res.* **1999**, 27(8), 1781-7.

83. Kikuchi, Y.; Niwano, M.; Yajima, N.; Nakamura, G.; Miyata, N. A novel macromolecular antibiotic, SN-07. Taxonomy of producing organism, isolation, characterization and biological activities. *J. Antibiot. (Tokyo)* **1985**, 38(12), 1670-6.

84. Kimura, K.; Takahashi, H.; Nakayama, S.; Miyata, N.; Kawanishi, G. Reconstruction of the macromolecular antitumor antibiotic SN-07 from SN-07 chromophore and DNA (Biological Chemistry). *Agri. Biol. Chem.* **1989**, 53(7), 1797-1803.

85. Kimura, K.; Spencer, D. M.; Bilardi, R.; Swift, L. P.; Box, A. J.; Brownlee, R. T.; Cutts, S. M.; Phillips, D. R. Barminomycin, a model for the development of new anthracyclines. *Anti Cancer Agents Med. Chem.* **2010**, 10(1), 70-7.

86. Watson, J. D.; Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, 171(4356), 737-8.

87. *Molecular Biology Web Book*. 2008 [cited; Available from: http://www.web-books.com/MoBio/Free/Ch3B3.htm.

88. Saenger, W.; Hunter, W. N.; Kennard, O. DNA conformation is determined by economics in the hydration of phosphate groups. *Nature* **1986**, 324(6095), 385-8.

89. Arnott, S.; Chandrasekaran, R.; Hall, I. H.; Puigjaner, L. C.; Walker, J. K.; Wang, M. DNA secondary structures: helices, wrinkles, and junctions. *Cold Spring Harb. Symp. Quant. Biol.* **1983**, 47 (Pt 1), 53-65.

90. Mahendrasingam, A.; Forsyth, V. T.; Hussain, R.; Greenall, R. J.; Pigram, W. J.; Fuller, W. Time-resolved X-ray diffraction studies of the B in equilibrium D structural transition in the DNA double helix. *Science* **1986**, 233(4760), 195-7.

91. *An Introduction to Nucleic Acids*. 2003 [cited 2003; Available from: http://ndbserver.rutgers.edu/NDB/archives/NAintro/index.html.

92. Stryer, L., *Biochemistry*. 4th ed. **1995**. New York: W.H. Freeman.

93. Dickerson, R. E.; Drew, H. R. Structure of a B-DNA dodecamer. II. Influence of base sequence on helix structure. *J. Mol. Biol.* **1981**, 149(4), 761-86.

94. Kim, H. Y.; Voehler, M.; Harris, T. M.; Stone, M. P. Detection of an interchain carbinolamine cross-link formed in a CpG sequence by the acrolein DNA adduct gamma-OH-1,*N*( 2)-propano-2'-deoxyguanosine. *J. Am. Chem. Soc.* **2002**, 124(32), 9324-5.

95. Hartley, J. A.; Berardini, M. D.; Souhami, R. L. An agarose gel method for the determination of DNA interstrand crosslinking applicable to the measurement of the rate of total and "second-arm" crosslink reactions. *Anal. Biochem.* **1991**, 193(1), 131-4.

96. Cutts, S. M.; Parsons, P. G.; Sturm, R. A.; Phillips, D. R. Adriamycin-induced DNA adducts inhibit the DNA interactions of transcription factors and RNA polymerase. *J. Biol. Chem.* **1996**, 271(10), 5422-9.

97. Baruah, H.; Bierbach, U. Unusual intercalation of acridin-9-ylthiourea into the 5'-GA/TC DNA base step from the minor groove: implications for the covalent DNA adduct profile of a novel platinum-intercalator conjugate. *Nucleic Acids Res.* **2003**, 31(14), 4138-46.

98. Fan, J. Y.; Ohms, S. J.; Boyd, M.; Denny, W. A. DNA adducts of 9-anilinoacridine mustards: characterization by NMR. *Chem. Res. Toxicol.* **1999**, 12(12), 1166-72.

99. *ImageQuant*. 1993-1998, Molecular Dynamics, Inc.

100. Skorobogaty, A.; White, R. J.; Phillips, D. R.; Reiss, J. A. Elucidation of the DNA sequence preferences of daunomycin. *Drug Des. Deliv.* **1988**, 3(2), 125-51.

101. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **1983**, 166(4), 557-80.

102. *Simple Reads, and Scan*. 1999, Varian Australia Pty. Ltd. UV-Vis operating program.

103. Phillips, D. R.; White, R. J.; Cullinane, C. DNA sequence-specific adducts of Adriamycin and mitomycin C. *FEBS Lett.* **1989**, 246(1-2), 233-40.

104. Phillips, D. R.; Cutts, S. M.; Cullinane, C. M.; Crothers, D. M. High-resolution transcription assay for probing drug-DNA interactions at individual drug sites. *Meth. Enzymol.* **2001**, 340, 466-85.

105. Cooper, G. M., *The Cell - A Molecular Approach*. 2nd ed. **2000**. Sunderland (MA): Sinauer Associates, Inc.

106. Nelson, D. L.; Cox, M. M., *Lehninger Principles of Biochemistry*. 4th ed. **2005**. New York: W. H. Freeman and Company. 1119.

107. Dvir, A. Promoter escape by RNA polymerase II. *Biochim. Biophys. Acta* **2002**, 1577(2), 208-23.

108. Kent, T.; Kashkina, E.; Anikin, M.; Temiakov, D. Maintenance of RNA-DNA hybrid length in bacterial RNA polymerases. *J. Biol. Chem.* **2009**, 284(20), 13497-504.

109. Evison, B. J.; Phillips, D. R.; Cutts, S. M. *In vitro* transcription assay for resolution of drug-DNA interactions at defined DNA sequences. *Meth. Mol. Biol.* **2010**, 613, 207-22.

110. Phillips, D. R.; Crothers, D. M. Kinetics and sequence specificity of drug-DNA interactions: an *in vitro* transcription assay. *Biochemistry* **1986**, 25(23), 7355-62.

111. Gray, P. J.; Cullinane, C.; Phillips, D. R. *In vitro* transcription analysis of DNA alkylation by nitrogen mustard. *Biochemistry* **1991**, 30(32), 8036-40.

112. Evison, B., *The Structural and Functional Characterisation of Covalent Pixantrone-DNA Adducts*, in *Biochemistry*. 2007, La Trobe University: Melbourne.

113. Lee, D. N.; Phung, L.; Stewart, J.; Landick, R. Transcription pausing by *Escherichia coli* RNA polymerase is modulated by downstream DNA sequences. *J. Biol. Chem.* **1990**, 265(25), 15145-53.

114. Chaires, J. B.; Dattagupta, N.; Crothers, D. M. Kinetics of the daunomycin--DNA interaction. *Biochemistry* **1985**, 24(2), 260-7.

115. Evison, B. J.; Chiu, F.; Pezzoni, G.; Phillips, D. R.; Cutts, S. M. Formaldehyde-activated Pixantrone is a monofunctional DNA alkylator that binds selectively to CpG and CpA doublets. *Mol. Pharmacol.* **2008**, 74(1), 184-94.

116. Lawley, P. D. Effects of some chemical mutagens and carcinogens on nucleic acids. *Prog. Nucleic Acid Res. Mol. Biol.* **1966**, 5, 89-131.

117. Hemminki, K.; Ludlum, D. B. Covalent modification of DNA by antineoplastic agents. *J. Natl. Cancer Inst.* **1984**, 73(5), 1021-8.

118. Evison, B. J.; Bilardi, R. A.; Chiu, F. C.; Pezzoni, G.; Phillips, D. R.; Cutts, S. M. CpG methylation potentiates pixantrone and doxorubicin-induced DNA damage and is a marker of drug sensitivity. *Nucleic Acids Res.* **2009**, 37(19), 6355-70.

119. Parker, B. S.; Buley, T.; Evison, B. J.; Cutts, S. M.; Neumann, G. M.; Iskander, M. N.; Phillips, D. R. A molecular understanding of mitoxantrone-DNA adduct formation: effect of cytosine methylation and flanking sequences. *J. Biol. Chem.* **2004**, 279(18), 18814-23.

120. Sprung, M. M. A summary of the reactions of aldehydes with amines. *Chem. Rev.* **1939**, 26, 297-338.

121. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; J. A. Montgomery, J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.;

Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox., D. J., *Gaussian 09*. 2009, Gaussian, Inc.: Wallingford CT.

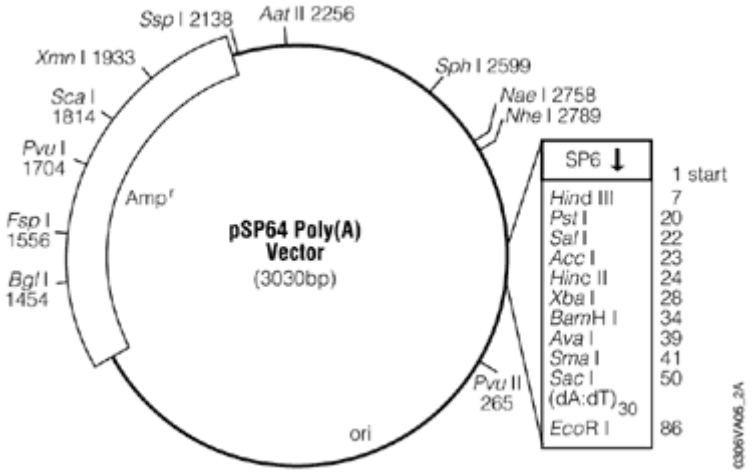
122. *Ampac*. 2001, Semichem: Kansas, USA. AM1 Semiempirical Methods.
123. McC, B. J. Gaussian basis sets for molecular calculations : edited by S. Huzinaga, with as co-authors, J. Andzelm, M. Klobukowski, E. Radzio-Andzelm, Y. Sakai and H. Tatewaki, Elsevier, Amsterdam, Oxford, New York, Tokyo, 1984, pp. viii + 426, price Dfl 265. *Journal of Molecular Structure: THEOCHEM* **1986**, 136(1-2), 201-201.
124. Layer, R. W. The Chemistry of Imines. *Chem. Rev.* **1962**, 489-94.
125. McMurry, J., *Organic chemistry*. 4th ed. **1996**. Pacific Grove: Brooks/Cole Pub. Co.
126. Foresman, J. B.; Frisch, A., *Exploring Chemistry With Electronic Structure Methods*. 2nd ed. **August 1996**. Pittsburgh, PA: Gaussian, Inc.
127. Alberty, R. A.; Silbey, R. J., *Physical chemistry*. 1st ed, ed. N. Rose. **1992**. New York: Wiley.
128. Chang, R., *Chemistry*. 7th ed. **2002**. Boston: McGraw-Hill.
129. Goddard, T. D.; Kneller, D. G., *SPARKY 3*. 2007, University of California: San Fransisco. 2D NMR Assignment.
130. Sayle, R.; Bernstein, H., *RasMol*. 2001. Molecular Renderer.
131. *Insight II*. 1997, Molecular Simulations Inc., A.I.: San Diego, CA. Molecular Modeling System.
132. Dooley, P. A.; Zhang, M.; Korbel, G. A.; Nechev, L. V.; Harris, C. M.; Stone, M. P.; Harris, T. M. NMR determination of the conformation of a trimethylene interstrand cross-link in an oligodeoxynucleotide duplex containing a 5'-d(GpC) motif. *J. Am. Chem. Soc.* **2003**, 125(1), 62-72.
133. VanderVeen, L. A.; Hashim, M. F.; Nechev, L. V.; Harris, T. M.; Harris, C. M.; Marnett, L. J. Evaluation of the mutagenic potential of the principal DNA adduct of acrolein. *J. Biol. Chem.* **2001**, 276(12), 9066-70.
134. Parker, B. S.; Cutts, S. M.; Phillips, D. R. Cytosine methylation enhances mitoxantrone-DNA adduct formation at CpG dinucleotides. *J. Biol. Chem.* **2001**, 276(19), 15953-60.

APPENDIX 1

http://www.promega.com/vectors/psp64pa.txt

**pSP64 Poly(A) Vector**

pSP64 poly(A) Vector Sequence




pCC1 is constructed from pSP64 496bp fragment from pRW1 digested with Pvu II/Sal I 497bp fragment inserted into Pvu II/Sal I digested pSP64.

PvuII/SalI transcription fragment from pRWI 497 bp (+ 4ss base over-hang)


⟼ = Transcription +1 site

NB 1. Full length transcription = 397 bases
2. Transcription lengths correlate to using GpA as a primer or dGGA