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

Copyright statement

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

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

Liability statement

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

Access Agreement

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

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

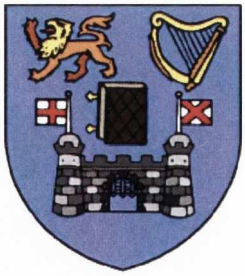
Synthesis, Biophysical and Biochemical studies of New Guanidine-like Derivatives Targeting DNA

by

Padraic Sean Nagle

B.A. Mod. (Chemistry), P. Dip (Statistics)

A thesis presented to the University of Dublin
for the degree of Doctor of Philosophy



Under the supervision of Prof. Isabel Rozas

School of Chemistry
Trinity College
Dublin

April 2010

Declaration

I hereby declare that this thesis has not been submitted as an exercise for a degree at this or any other university. The work contained herein is entirely my own, except where otherwise cited, referenced, acknowledged or accredited. I agree that the library of the University of Dublin may at their discretion lend or copy this thesis upon request.

Acknowledgements

I would like to thank all the people that have contributed to this work. Firstly, I wish to express my gratitude to my supervisor Prof. Isabel Rozas who has been an outstanding support and guidance for the duration of this work. I worked with her for my undergraduate project and learnt much from her, and will always be indebted to her. I would also like to thank her for all her encouragement that I received during my research here at Trinity College Dublin.

It is difficult to overstate my admiration for the postdoctoral fellows Dr. Alessandra Cordiero Machado and Dr. Fernando Rodriguez for their enormous support and help in the laboratory and all their useful advice they shared with me. I would also like to thank them for all the good times I shared with them while we worked together.

I would like to express my sincerest gratitude to Prof. David Wilson from Georgia State University who was responsible for carrying out all the surface plasmon resonance experiments shown in this thesis. I would also like to thank Dr. Amir Kahn for his advice and help with the isothermal titration calorimetry in the biochemistry department. I would also like to thank Prof. John M. Kelly and Dr. Susan Quinn for their most helpful discussions and advice during during the course of this research.

I would like to thank Dr. John O' Brien and Dr. Manuel Ruether for running all my NMR samples and Dr. Martin Feeney for carrying out all my mass spectra even at short notice. Special thanks should go to all the technical staff in the school of chemistry for making my time here most enjoyable.

I would like to thank the members of the Rozas' group for all the enjoyable times in the past three years. Firstly, thanks to the third years Amila Kahvedzic and Aoife Flood for being great friends and for all the support when times were hard for me. I would like to wish them the best for their future. Thanks should go to the second years in the group Caitriona McKeever and Daniel O' Donovan for all their helpful discussions

and great nights out. And to the first years in the group Brendan Kelly and Elena Diez, unfortunately I did not work with them for that long, but thanks for bringing a great atmosphere to the lab and making my last six months far more enjoyable. I would also like to thank Prof. Thorri Gunnlaugson and his group for brightening up the atmosphere in the lab.

I would like to thank all my friends but especially David Barron and Giuseppe La Spina for their laughs especially for the helpful discussions and advice during this research. I should also thank them for all their support and help while this thesis was being written up.

I cannot end without thanking my parents Mr. Patrick J. Nagle and Mrs. Siobhan Nagle who are responsible for guiding me through my educational studies and for whom I dedicate this work to. A special thanks to them also for giving me support while I was at Trinity College. I would also like to thank my late grandparents Thomas and Treasa MacGabhann who never stopped supporting me in my early years. My sincerest gratitude also goes to my sisters Bebhinn, Doireann and Aoife who have helped me a lot and have always supported me especially in the past year. Special thanks should go to Doireann for all the helpful advice and direction to pursue a career in chemistry.

Abstract

The literature describes many DNA minor groove binders, some of them having similar characteristics such as crescent shape, being dicationic and containing hydrogen bond donating or accepting groups.

This project aims at the preparation of new families of DNA minor groove binders and from them, dual action agents combining a known DNA intercalator and these minor groove binders, in the hope that these compounds will be more specific and stronger in their interaction with DNA with the potential to be used for therapeutic purposes in the treatment of cancer.

Thus, the first step in the development of this project is the synthesis of asymmetric diaromatic dicationic molecules, presented in Chapter 3. Two possible synthetic pathways are possible:

i) Starting from a series of aromatic *mono*-guanidines the 2-aminoimidazoline moiety can be added or, oppositely,
ii) guanidylating different aromatic *mono*-2-aminoimidazolines

For methodological and practical reasons, the first approach was pursued even though both series of *mono*-cations were prepared. Once these asymmetric dications were obtained their potential as minor groove binders was assessed by means of different biophysical experiments.

In Chapter 4, their DNA binding affinity was investigated by means of thermal denaturation experiments with natural DNA and the poly(dA-dT)$_2$ homopolymer. UV-vis, CD and LD were then performed to assess their binding strength quantitively and to confirm their mode of binding. In addition, the cytotoxicity of these asymmetric dications was evaluated in the Paterson Institute, Manchester by Dr. Margison. All this information will help in the development of the dual action agents and their biophysical and biochemical evaluation.

Chapter 5 discusses the techniques employed to evaluate the thermodynamic quantities such as the enthalpy and entropy of binding using isothermal titration calorimetry. Surface plasmon resonance experiments were carried out by Prof. David Wilson from Georgia State University providing more specific information (i.e. binding constants) on the binding of these molecules to AT-rich oligonucleotides. In addition, chapter 6 describes biochemical evaluation of the asymmetric dications that was evaluated in the Paterson Institute, Manchester by Dr. Margison. All this information will help in the development of the dual action agents and their biophysical and biochemical evaluation.

The final two chapters focus on the synthesis of novel intercalator dual agents where acridine was attached to a selected few of the promising compounds and their binding affinity is discussed in chapter 8.

Table of Contents

Abbreviations

A	Adenine
Ar	argon
Boc	*tert*-butyloxy-carbonyl
Bp	base pair
br	broad
C	cytosine
d	doublet
DCM	dichloromethane
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EtOAc	ethyl acetate
EtOH	ethanol
Fmoc	9-fluoromethyloxy carbonyl
g	gram
G	guanine
Hz	hertz
IR	infrared
J	coupling constant
m	multiplet
MeOH	methanol
ml	millilitre
mol	mole
MS (ES)	mass spectrometry (electron spray)
NMR	Nuclear Magnetic Resonance
Ph	phenyl
q	quartet
R_f	retention factor
r.t.	room temperature

s	singlet
SPR	surface plasmon resonance
t	triplet
T	thymine
TEA	triethylamine
THF	tetrahydrofuran
TLC	thin layer chromatography
UV	ultraviolet
Z	benzyl carbonate

I would like to dedicate this work to my parents

Patrick J. and Siobhan Nagle

Chapter 1

Introduction

1.1 Cancer: Statistics

Cancer is a leading cause of death worldwide accounting for 7.4 million deaths worldwide in 2004 (~13%) and this figure is expected to rise with an estimated 12 million deaths occurring by 2030.[1] Cancer, by definition, is a set of diseases that share a characteristic medically known as a malignant neoplasm.[2] This can be understood as being the undefined control of cell growth and proliferation that is usually caused by mutations to the cell.[3]

In Ireland today, cancer is the second largest cause of death following heart disease.[4] Even though the incidence of the disease is seen to be increasing due to the larger and increasing age of the population, the survival rate in Ireland is also rising.[4] This is attributed to better understanding of the disease mechanisms and significant advances in surgical techniques and research into improved compounds that combat the disease.

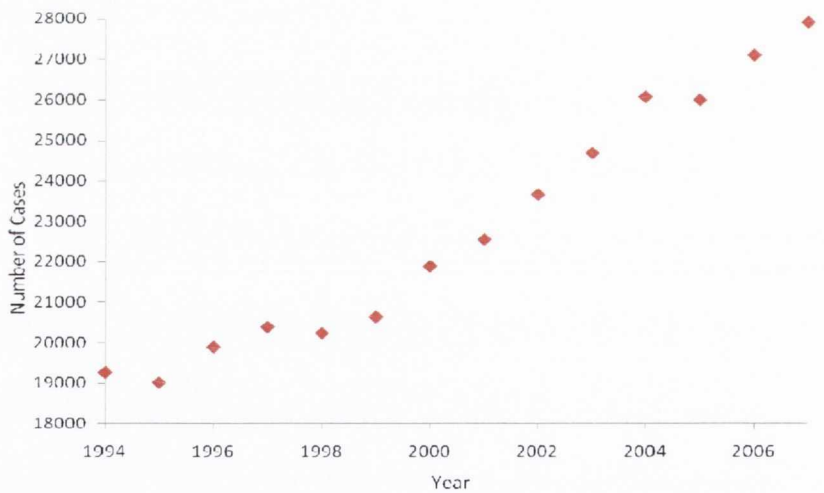


Figure 1.1. Graph showing the differences in the number of cancer cases in Ireland between 1994-2007[4]

Figure 1.1 shows an overall increase that can be attributed to an increment and aging of the population and the national council registry of Ireland predicts that the number of new cases will rise by 108% overall by 2030.[4] This emphasises the crucial need in the development of treatments for this devastating disease. Subsequently, the need for funding in cancer research can never be underestimated.

1.2 Cancer: Introduction

Cancer is a condition often seen as one disease, whereas, in fact, it describes a group of diseases.[5] For example lung cancer, colorectal cancer and leukemia are all different diseases. However, there is one common feature: the uncontrolled growth and proliferation of abnormal cells that can be caused by certain risk factors:

(i) Environmental factors: These include the radiation from the decay of uranium deposits in the ground that forms radon gas. This is a serious problem in Ireland since it accounts for 56% of the total radiation dose received by the Irish population and accounts for around 200 deaths a year from lung cancer.[6] Another environmental factor is the tobacco use among the population which is one of the main causes.[7] Chemical carcinogens may also cause cancer and can be defined as chemicals that may cause cancer such as polyaromatic hydrocarbons (PAHs). Biological carcinogens can be defined as viruses that cause cancer. Human papillomavirus infection is an example of a virus that causes cervical cancer. It is a necessary factor in most cervical cancers which is the second major cause of death among women.[8]

(ii) Hereditary or Genetics: A mutated gene can be passed down from one of the parents (examples include ovarian and breast cancers). They generate from the same genes called BRCA1 and BRCA 2 which are tumour suppressor genes that prevent uncontrolled cell growth.[9]

(iii) Age: is a major factor in the development of cancer. Due to the aging population and the increased life expectancies, the level of cancer incidence is expected to increase dramatically in the next twenty years. When a body ages, the repair mechanisms of the cell decrease their efficiency. Subsequently, the chances of a genetic mutation or combination of them that usually leads to tumour formation increases with age.[10]

(iv) Diet is shown to have an effect on the chances of being diagnosed with cancer. Obesity has been found to increase the overall cancer incidence.[10]

Thus, the main question remains: what differentiates a tumour cell from an ordinary eukaryotic cell. If we contrast these two categories (Table 1), we can split the main differences into five sections.

Table 1.1.-Main differences between healthy and malignant cells[11]

Characteristic	Healthy Cells	Cancer Cells
Function	• Cells divide in an orderly way and are produced only when required. • They avoid uncontrolled proliferation by apoptosis. • Enzymes and hormones are normal	• Cells undergo uncontrolled growth and proliferation. • Apoptosis pathways are altered. • Some are either over active or under active
Metabolism	Derive 70% of their energy from the Krebs cycle	Obtain most of their energy from the glycolytic pathway
Blood Vessels	Cells have a built in system	They require more of certain amino acids to build one
Growth Factors	The amount of growth factors are in balance	Cells require more growth factors and are over active

Table 1.1 shows that the main difference between the two categories lays in the structural differences. For example, a healthy cell divides in an orderly way, with the tumour suppressor and p53 genes fully functional. These genes prevent uncontrolled growth which is a step towards the formation of a tumour. In a cancer cell, these genes are mutated, thereby, allowing the characteristic uncontrolled growth to occur.

The metabolism in tumour cells differs significantly from normal eukaryotic cells. In cancer cells, the uptake of glucose is significantly increased for the uncontrolled growth of

tumour cells, and the rate of glycolysis increases substantially since most of their energy is derived from this pathway. A reason why the increased rate of glucose is required is because the glycolytic pathway yields only two molecules of ATP per glucose molecule, whereas oxidative phosphorylation yields 18 molecules of ATP per molecule of lactate. According to this, cancer cells can be divided into two classes, the hypoxic cells and the aerobic tumour cell that grow in environments of low and high oxygen respectively. The hypoxic tumour cell converts glucose to lactate, whereas the aerobic cell metabolises the lactate to form the mentioned 18 molecules of ATP.[12] A consequence of this metabolism is that the environment in which cancer cells grow is acidic. By consequence, this low pH environment is favourable for tumour growth in hypoxic conditions.

Another characteristic that distinguish tumour cells from normal eukaryotic cells is that tumour cells invade other areas of the body through metastasis. This has a strong correlation between metathesis and angiogenesis which involves the formation of extra blood vessels, important to obtain an oxygen supply for tumour growth. The angiogenic process has been seen as an important process in tumour development, because agents administered to block this process have an adverse effect on the tumour growth.[13] Conversely, tumour growth is enhanced by factors that stimulate angiogenesis.[14]

1.3 Tumorigenesis

One of the most important issues is to know how a cancer cell is derived from the corresponding healthy cell, and which is the sequence of events that a cell undergoes to form a cancer cell.

1.3.1 Normal-Precancer-Cancer Sequence

Fearon and Vogelstein[15] have proposed a model to describe tumour growth showing that more than one mutation is required for tumour formation. For every mutation that occurs, the probability of a tumour increases substantially. By taking the colorectal cancer model, it can be seen that the mutation of the APC gene in a normal colon cell causes hyper-proliferation (Figure 1.2).

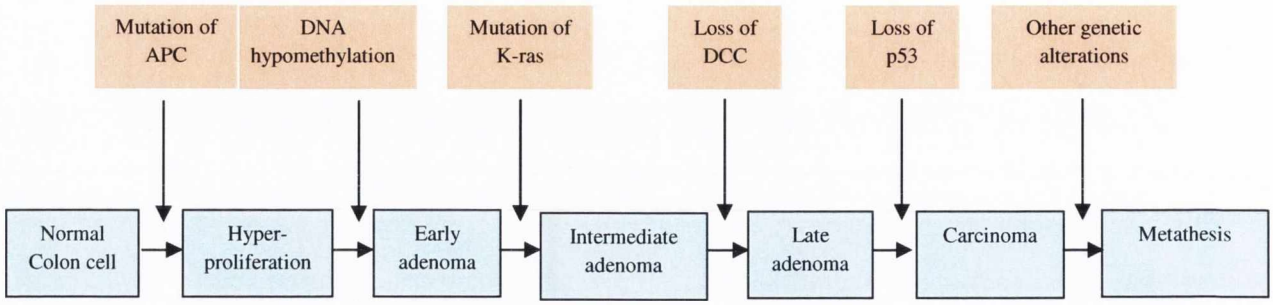


Figure 1.2. Model for the formation of colorectal cancer[15]

From only one mutation, a tumour cannot form; and, consequently more point mutations would be required. DNA can be methylated and three further mutations (in the K-ras, DCC and p53 gene) occur for the formation of a carcinoma, which is known as the pre-cancerous stage. If the cancer was found at this stage, the prognosis would be favourable, even though, surgery followed by chemotherapy would be required. Further mutations, especially in the growth regulation gene would then favour metastasis to occur, at this stage, secondary cancers would form.

1.3.2 Carcinogenesis

Carcinogenesis can be defined as the process that leads to genetic alterations due to chemical or environmental factors.[16,17] It can be divided into three main steps called initiation, promotion and progression. The initiation step consists of the metabolism of genotoxic compounds such as polyaromatic hydrocarbons (PAHs) by converting the procarcinogen into a carcinogenic substance. Alternatively, detoxification can occur by excreting this carcinogenic material (Figure 1.3).

Once the substance is transformed into a carcinogenic compound, it binds to the DNA forming a carcinogen-DNA adduct. From here, three possibilities arise. Firstly, the cell recognises the adduct and eventually apoptosis, a DNA damage response occurs. Secondly, the DNA can be repaired, consequently removing the effects of the carcinogenic compound. Thirdly, the DNA can replicate itself with the carcinogen-DNA adduct present resulting in a mutation. This replication forms an initiated cell and may disturb the functioning of the cell.

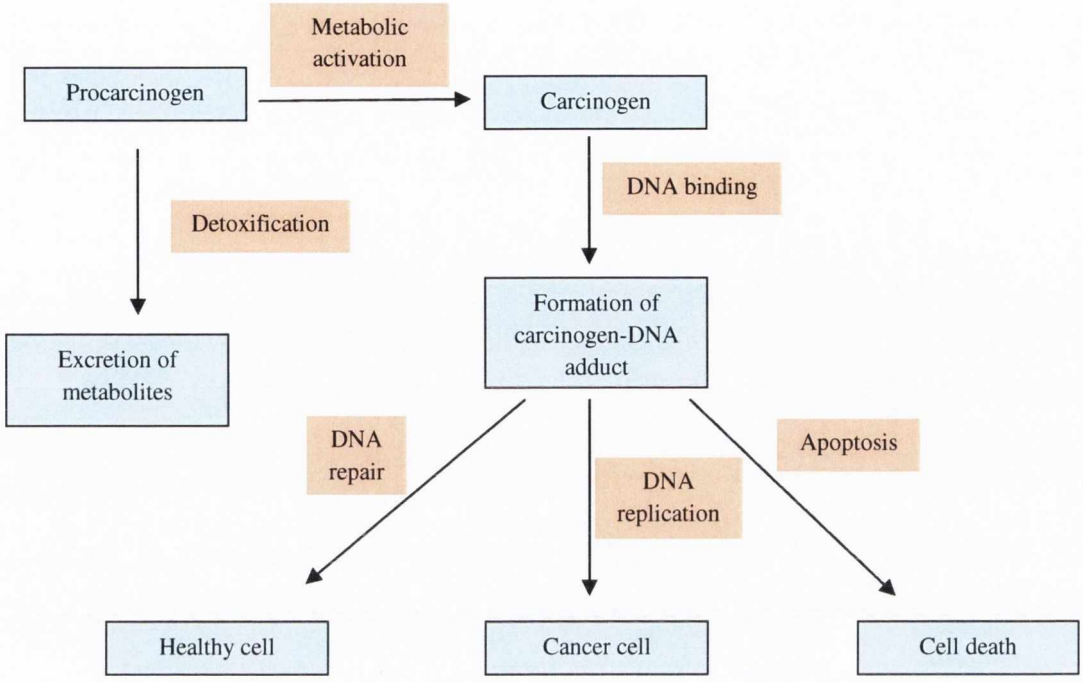


Figure 1.3. Model for carcinogenesis[16]

Further mutations could occur at the growth regulating or tumour suppressor genes, causing the cell to divide uncontrollably, thereby forming a cancer cell that is able to replicate and survive in harsh conditions that would result in the formation of a tumour.

1.3.3 Cancer related genes

Tumour Suppressor Genes

Tumour suppressor genes contrast to oncogenes because they inhibit cell growth or promote stability in the cell.[16] Examples of such genes include the p53 and Adenomatosis Polyposis Coli (APC) genes.

Mutations in the p53 tumour suppressor gene have been seen as a critical factor for tumour growth.[18] This protein works as a transcriptional factor by stimulating several factors that would bring about cell arrest, DNA repair and encourage cell death by apoptosis.[19] Elimination of its function leads to increased mutational rates and resistance to apoptosis.

The structure of the p53 protein can be divided into three main parts, the N-terminus that interacts with many transcriptional factors and proteins that can modify its function, the central domain that is responsible for binding to the DNA where the majority of mutations arise to eliminate the DNA binding activity and the C-terminus that can bind to other p53 proteins.[20]

Mutations in the APC suppressor gene occur in most colon cancers.[21] The importance of this gene has been seen in the fact that it regulates apoptosis and cellular adhesion. Elimination or mutation of this gene would thereby cause uncontrolled cell growth and proliferation.

The APC protein works by binding to glycogensynthasekinase 3 beta (GSK 3β) and Axin. Then this complex binds to β-catenins in the cytoplasm that have dissociated from adherens contacts between cells. It works by phosphorylating the β-catenin a second time after the initial phosphorylation carried out by casein kinase 1 (CK 1). This targets the β-catenin for degradation by cellular proteosomes, thereby preventing it from translocating into the nucleus, where it would act as a transcription factor for proliferation genes.

1.4 Ways of targeting cancer

There are several ways of treating cancer patients that are highly dependent on the stage of cancer. Each of these treatments can be advantageous and disadvantageous. For example, if a patient was treated with a cytotoxic drug, it could have the potential to shrink the tumour size; however, these would also display toxicity against normal epithelial cells. Specific examples of cancer treatments include radiotherapy, surgery, photodynamic therapy, anti-angiogenesis drugs and chemotherapy drugs.

Radiotherapy is usually given externally by a beam of protons or gamma rays. It displays many advantages, for example, it has the ability to shrink the tumour size or even eliminate the tumour if the quantity of radiation provided is efficient. However, there is a limit to the amount of radiation treatment that one can withstand due to its effects on the normal surrounding tissue. Radiopharmaceuticals are drugs that contain radioactive materials,

such as iodine-131. These drugs are taken in by thyroid cancer cells destroying them since the thyroid gland absorbs most of the iodine in the blood.[22]

Photodynamic therapy involves the use of photosensitizers and visible light with the combination of oxygen to produce the cytotoxic singlet oxygen species that induce tumour cell necrosis and the shutdown of microvessels (Fig. 1.4).

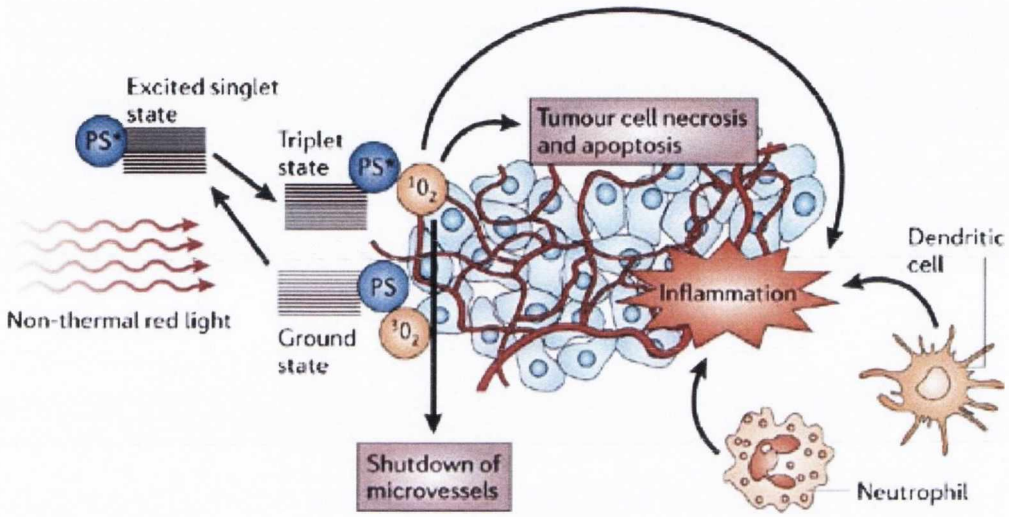


Figure 1.4. Photodynamic Therapy: mechanism of action[23]

The process involves the excitation of the photosensitiser, i.e. the energy is transferred via intersystem crossing to the photosensitisers triplet state. Energy is then transferred to the reactive singlet oxygen species which induces tumour cell apoptosis and necrosis. Examples of photosensitisers include porphyrins.[23]

Tumour cells require an oxygen rich environment, therefore they produce new blood vessels to deliver them the required oxygen. The system responsible for the production of new blood vessels is the vascular endothelial growth factor (VEGF). Therefore, the preparation of drugs that inhibit its activity is an active area of research. An example of a drug that binds to VEGF is bevacizumab.[24]

One of the most common treatments used to treat cancer is cytotoxic therapy. The clinical name for this treatment is *chemotherapy* and involves the administration of cytotoxic compounds with the target to reduce the tumour size or eliminate the remaining cancer cells. However, a limiting factor for *chemotherapy* is the several side effects that accompany the

treatment. For example, these drugs target fast growing cells such as the cancer ones; however, gastrointestinal cells divide rapidly and are also affected by these cytotoxic drugs. Consequently, people that undergo *chemotherapy* generally suffer from nausea or diarrhoea. In these days however, drugs can be administered to alleviate these symptoms. Likewise, the hair follicle cells divide rapidly, and as a consequence people that undergo *chemotherapy* usually suffer from hair lose or thinning; however, these symptoms usually disappear after a few weeks post chemotherapy treatment.[25]

Most commonly, these treatments are usually given in combination, for example, chemotherapy may be used to shrink a tumour size prior to surgery, when these procedures are needed to fully remove the tumour. Following surgery, chemotherapy is usually administered again to kill off the remaining tumour cells.

Consequently, there is a huge interest in the development of drugs that could be selective against tumour cells and have minimal effect on normal cells. Since chemotherapeutic drugs such as *cis*-platin or brostallicin (currently in the clinical trials phase II,[26]) target DNA, there has been much research carried out in compounds that target DNA and their selectivity. In order to understand this, a comprehensive explanation of the structure of DNA must be considered, and that will be discussed in the next section.

1.5 The DNA macromolecule

In this project we focus on molecules that bind to the DNA macromolecule, a common target for many chemotherapy drugs. Compounds can bind to DNA through a variety of modes such as intercalation, groove binding or alkylation. These modes of action are supported through various interactions such as hydrophobic forces (intercalation and minor groove binding), π-π stacking (intercalation) or electrostatic interactions (most intercalators and groove binders).

In order to understand the modes of DNA-ligand interactions, an understanding of the DNA structure is required. DNA is a long macromolecule with a helical structure, composed of bases, sugars and phosphate groups (Figure 1.5). The four heterocyclic bases are divided

into two groups, adenine and guanine bases which are called purines and the pyrimidines which consist of the cytosine and thymine bases.

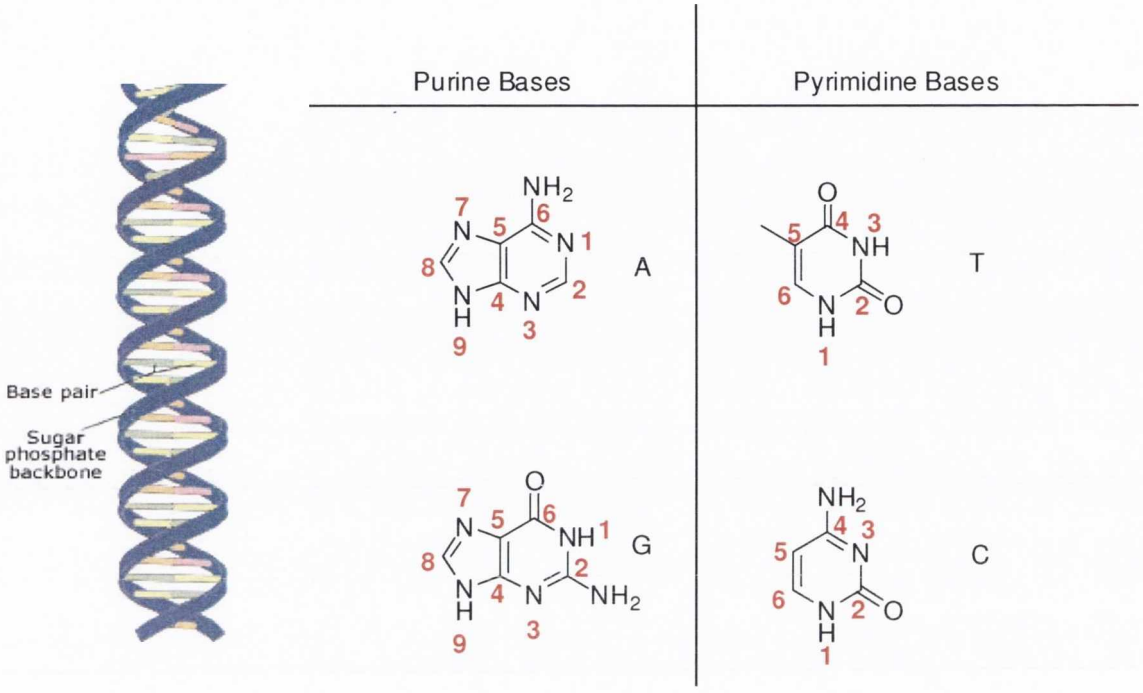


Figure 1.5. Diagram of the DNA structure illustrating the DNA double helix (left), [27] *and DNA bases (right)*

Each of these bases is covalently bound to a five membered sugar ring called deoxyribose. This structure is called the nucleoside structure. When a nucleoside is attached to a phosphate group, the name nucleotide is given. Each nucleotide is hydrogen bonded to a corresponding nucleotide depending on the base that is contained in it. For instance, adenine hydrogen bonds to thymine with two hydrogen bonds, whereas, cytosine hydrogen bonds with guanine with three hydrogen bonds.

When this occurs, two strands of DNA merge resulting in a helical structure, called the DNA double helix. It can be defined as being a long polymer consisting of repeating units called nucleotides, such that the human chromosome 1 is around 220 million base pairs long.[28] The DNA structure is held together by hydrogen bonds between the base pairs with the negatively charged phosphate group situated on the outside of the helix. The charge on these phosphates is balanced by positively charged counterions. Also located in the helix are

water molecules, more commonly known as the 'spine of hydration'. Another important aspect of the DNA double helix is the presence of two grooves. These groove are named according to their size, the major groove and the minor groove. The consequence of having the grooves is that in the major groove, the bases are more exposed, consequently, enabling proteins to have access.

There are two important contrasting structural features for the 5 membered deoxy-ribose ring in nucleotides. The first main feature is that the five membered ring is non-planar. This non-planarity, called puckering, can be defined by a simple qualitative description of the conformation in terms of atoms deviating from ring coplanarity.[29] The naming of the contrasting puckers, are dependent on the side of the major deviation. For example, if the major deviation from the plane is on the side of the base, this is given the term *endo*. The major puckers observed are the C2'-*endo* and the C3'-*endo* (Figure. 1.6).[29] This is where the C2' and C3' atoms deviate from coplanarity on the same side as the base respectively. Subsequently, this can be seen as a really important determinant for the overall structure of the polynucleotide.

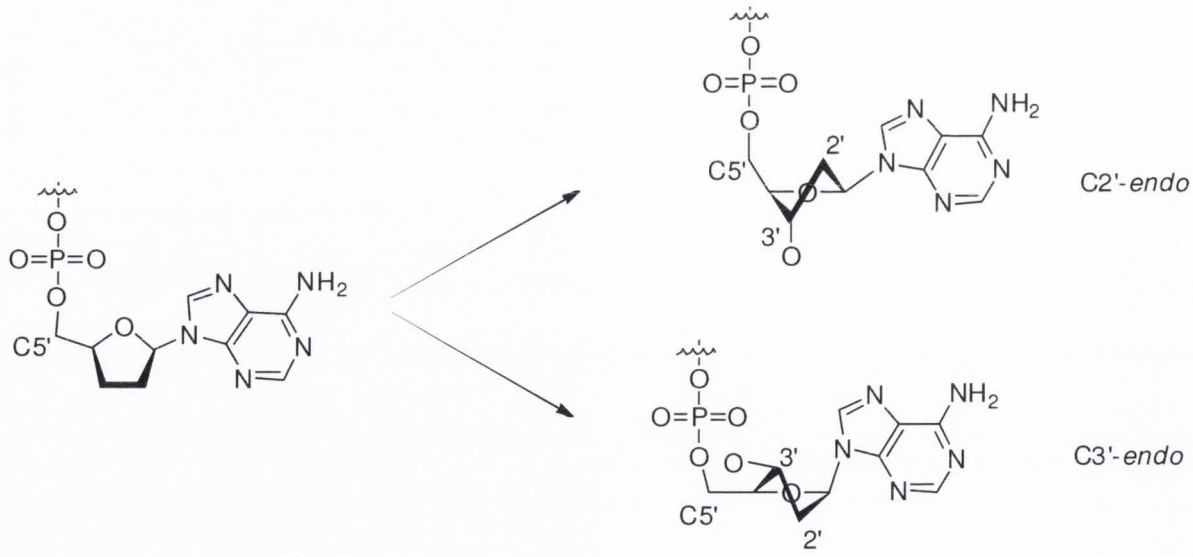


Figure 1.6. DNA sugar puckers, (top right) C2' endo, (bottom right) C3' endo[29]

Another important structural determinant is the glycosidic bond. This covalent bond is formed between the sugar and the base, and can result in two conformations: *syn* and *anti*. Most nucleotides prefer the *anti* conformation as it is lower in energy; however, guanine containing nucleotides are an exception. The reason behind this unusual preference for the *anti* conformation is the favourable electrostatic interaction between the exocyclic amino group of the guanine base and the 5'phosphate group. This *anti* conformation is highly important since it determines the polymorph of DNA that is present. Also, mismatching can result having mutational consequences unless repaired, as for example, the A•G mismatch which is the least efficiently repaired.[30] Webster *et al.* have reported a crystal structure of the d(CGCAAGCTGGCG) sequence,[30] where they observed that the glycosidic angle changed depending on the sequence. In this case it is reported that the A•G mismatch is in the A*(syn)*•G*(anti)* conformation which was determined by x ray crystallography. Interestingly, in this mismatch, two hydrogen bonds between these bases are observed, between the N7 of adenine and the N1 of guanine and also between the N6 of adenine and the O6 of guanine.[30]

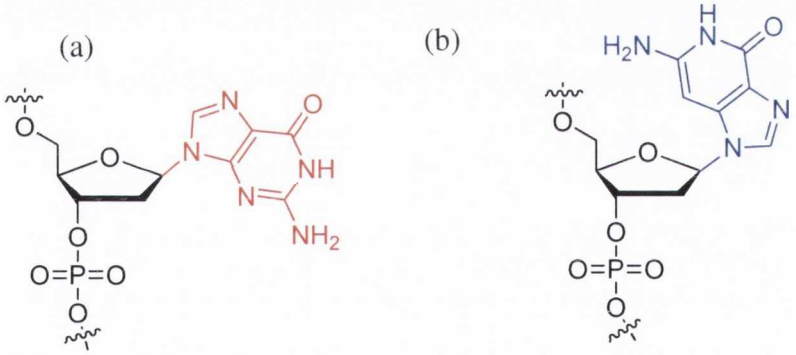


Figure 1.7. Guanine base in the anti-conformation (a), syn-conformation (b)

1.6 Duplex forms: A- B- and Z-DNA

DNA is highly polymorphic meaning that it can exist in many forms. For example, it can exist in the A, B or Z forms (Figure. 1.8) depending on the base pair composition and the conditions. For example, by altering the solution conditions, the B- polymorph can change into A-DNA. Also, if we have polymeric DNA that consists of guanine and cytosine base

pairs, and we increase the salt concentration to approximately 4 M, Z- DNA will be formed from B-DNA.

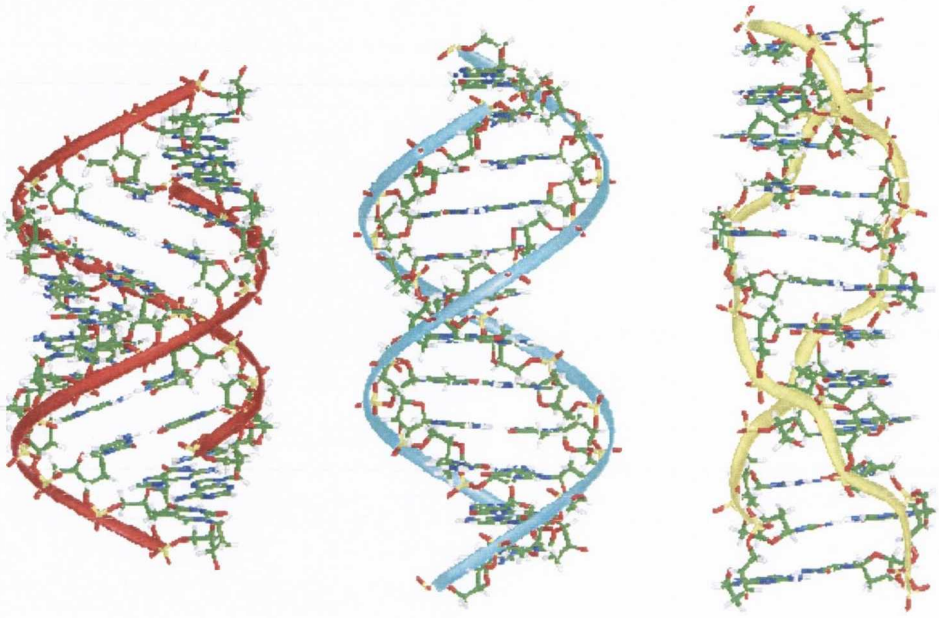


Figure 1.8. A-form DNA(left), B-form DNA (middle) and Z-form DNA (right)[31]

The differences between A-, B- and Z-DNA are tabulated in Table 1.2. The main difference between these three DNA forms is the helix handedness since the Z-form helix is left handed whereas both of the A- and B-forms are right handed helices.

Two other differences between these helices are the sugar puckers and the glycosidic bond. In Table 1.2 can be seen that the conformation of the sugar pucker is C3'-*endo* in the A-form, C2'-*endo* in the B-form and can be C2'-*endo* for the cytosine base or C3'-*endo* for the guanine base in the Z-form DNA. The glycosidic angle is in an *anti*-conformation for all forms of DNA apart from the guanine base of Z-DNA.

Table 1.2.- Main differences between A, B and Z-DNA[29, 32]

Geometrical Attribute	A-DNA	B-DNA	Z-DNA
Helix Sense	right-handed	right-handed	left-handed
Repeating Unit	1 bp	1 bp	2 bp
Rotation/bp*	32.7°	35.9°	60°/2
bp*/turn	11	10.5	12
Inclination of bp* to axis	+19°	-1.2°	-9°
Rise/bp* along axis	2.3 Å	3.32 Å	3.8 Å
Pitch/turn of helix	28.2 Å	33.2 Å	45.6 Å
Mean propeller twist	+18°	+16°	0°
Glycosyl angle	*anti*	*anti*	C: *anti* G: *syn*
Sugar Pucker	C3'-*endo*	C2'-*endo*	C: C2'-*endo* G: C2'-*exo*
Helix Diameter	23 Å	20 Å	18 Å

bp* = base pair

An example of a B-form DNA was crystallised in 1979 and the structure was solved by multiple isomorphous replacement methods.[29,33] The structure obtained was d(CGCGAATTCGCG) and many structural features were observed from its analysis that are summarised below:

- A right handed anti-parallel B-form helix was observed [29]
- The minor groove was exceptionally narrow with a width of 3.2 Å in the 5'-AATT region whereas the major groove was exceptionally wide [29]

- There was a well ordered network of water molecules in the minor groove,[29,34] called the 'spine of hydration'. The first layer of water molecules are bonded to the N3 of adenine and O2 of thymine. These outer water molecule layers are linked by a second layer. This was observed by NMR analysis.[29, 35]

Many other sequences of DNA have been crystallised to date such as the alternating d(CGCATATAGCG).[29,36] It was observed that the minor groove width is equally as narrow as the Dickerson-Drew sequence. In fact, a narrow minor groove is generally preferred by AT sequences. For instance, the d(CGCAAGCTGGCG) oligomer has an average minor groove width of 7 Å in the 5'-AGCT section,[29, 30] indicating that the GC base pairs increase the groove width. However, this is only a preference since in the decamers d(CCAACGTTGG) and d(CCAGCGCTGG) both display narrowing in the central region of the minor groove indicating that GC base pairs do not always cause widening of the groove width.[29, 37]

Double helices of the A-form are formed from random sequence DNA under low humidity and in solution where the water content is reduced by the addition of alcohols.[29] A large number of A-DNA octamers have been crystallised, for example d(GGGGCCCC)[29, 38] and d(GGGTACCC).[29, 39] These sequences differ by their minor groove width, for example d(GGGGCCCC) has a minor groove width of around 15 Å whereas d(GGGTACCC) has a groove width of around 9.7 Å. This illustrates how both A- and B- forms of DNA have a wider minor groove when the proportion of GC base pairs increases.[29]

Z-form DNA double helices are formed from poly(dG-dC)$_2$ nucleotides by increasing the salt concentration beyond 4 M. This stabilises the helix by reducing the electrostatic repulsions between the intrastrand phosphate groups[29] since they are held closer together in Z-form DNA in comparison to the A- and B-forms. Thus, the DNA transforms from the B-form into the Z-form. Crystal structures have been obtained from sequences such as d(CGCGCGCG)[29, 40] and d(CGCGCGCGCG),[29, 41] which are examples of Z-form DNA. Interestingly, it has been shown that by replacing the GC base pair for an AT base pair, the left handed helix will remain intact.[29, 42] However, AT sequences greatly destabilise the Z-form DNA since they are unable to support the groove hydration that is a necessary component to maintain the Z-DNA structure.[43] Another reason why the Z-form prefer the guanine base in comparison to the adenine base is due to the ability of guanosine nucleotides to adopt the *syn* glycosidic conformation in comparison to the *anti*-conformation.[29]

Z-DNA is said to be biologically relevant for two reasons:

- the C8 and N7 atoms of the guanine base are susceptible to electrophilic attack, consequently areas of Z-DNA within the genome could have an increased incidence of mutation.[29]
- another proposed function of Z-DNA is to act as a signal for the induction of transcription.[29, 44]

To summarise, there are three polymorphs of DNA present, and their existence is highly dependent on the sequence and the solution conditions, but from now on the present report will concentrate on the most naturally abundant form, the B-DNA

1.7 DNA binders

There are many different classes of compounds that bind to DNA. For example, cations such as sodium ions bind electrostatically to the DNA backbone, and consequently, stabilise the negatively charged DNA backbone. Other examples include the drugs netropsin[45] and distamycin[46] that interact with DNA by binding to the minor groove, therefore being called minor groove binders. Examples of synthetic drugs include the well known anti-cancer drug called cis-platin. This class of compounds is called an alkylating agent and as the name suggests, it acts by alkylating onto the DNA bases. In the following subsections, our purpose is to discuss different families of DNA targeting agents based on their particular type of interaction with DNA.

1.7.1 Alkylators and Interactions

Molecules that alkylate onto the DNA macromolecule exert their activity by binding covalently to the DNA bases, sugars and backbone. These structures, known as alkylators, have the following characteristics:

- they are sometimes water soluble, consequently, enabling them to be biodistributed and

- they generally have a leaving group present so they can bind covalently to the DNA bases as seen in *cis*-platin.

Different alkylators are used nowadays in the clinic and the most representative will be discussed in the following sections.

1.7.1.1 Nitrogen Mustards

Nitrogen mustards are compounds that attach covalently to the guanine base of DNA (Figure 1.9). They were first used medicinally in 1942, the first one being chlormethine. Unfortunately, a drawback for the use of this reagent, and indeed all mustards, is that tumours can build up resistance by increasing the glutathione concentration[47] and since these compounds are highly reactive, this would result in decreased activity among tumour cells since the mustards can react with glutathione in the cell.

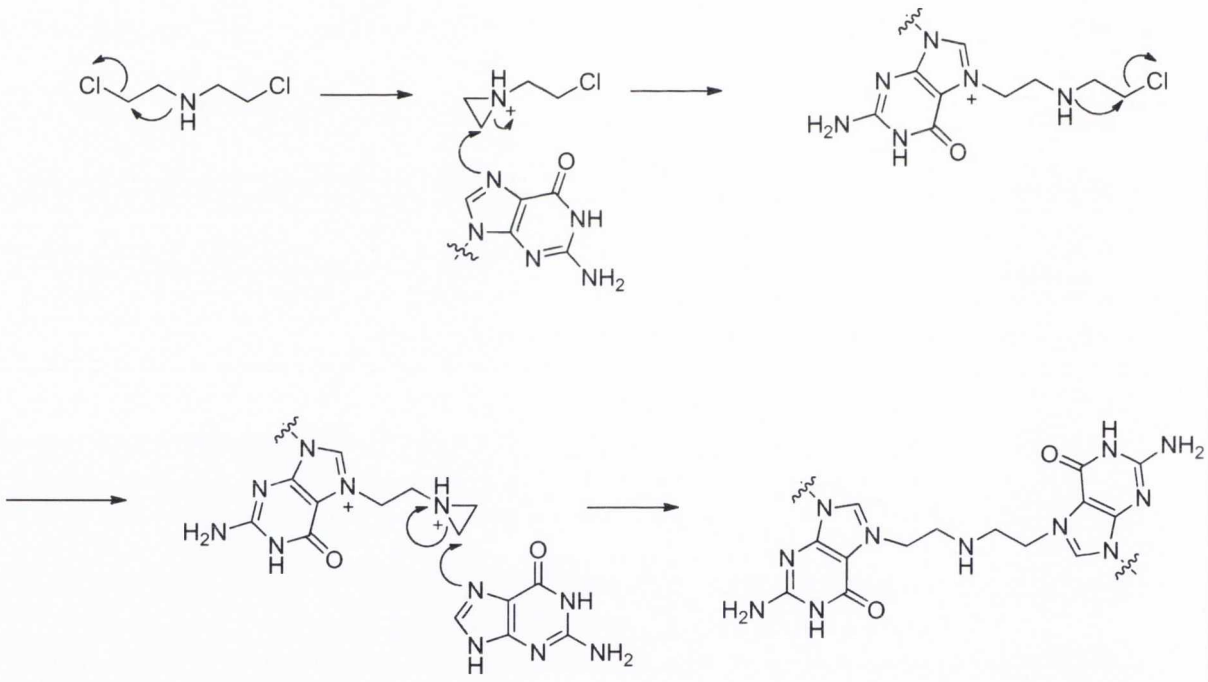


Fig 1.9. Mechanism of alkylation of nitrogen mustards to the N7 of the guanine base

Nitrogen mustards react in a two step mechanism. Initially, the basic nitrogen attacks intramolecularly the electrophilic carbon by the S_N2 mechanism, with the chloride atom leaving and resulting in the formation of an aziridinium cation. This highly reactive

intermediate is readily attacked by the nucleophilic N7 of the guanine base resulting in the neutralisation of the the positive charge on the nitrogen atom forming the *mono* adduct. This step can be repeated to form a *bis* adduct which is the reactive complex since it prevents DNA unwinding upon transcription, thereby resulting in cell death. However, the *mono* adduct is formed in a ratio of 20:1. Unfortunately, it is this product that is responsible for its high genotoxicity.[47]

1.7.1.2 Diaiminodichloroplatinum(II) (cis-platin) and its analogues

Other examples of molecules that alkylate the DNA macromolecule include the family of platinum drugs such as cisplatin and its derivatives: carboplatin and picoplatin. Cisplatin was the first platinum drug that was discovered and was first administered to patients in 1971. It exerts its activity by alkylating the N7 of the guanine base, thus once inside the cell, activation of the cisplatin occurs by aquation (water molecules replace the chlorine groups) in the cytoplasm (Fig 1.10).[48] This is due to the fact that the chloride concentration inside the cell is less than the corresponding concentration outside the cell.

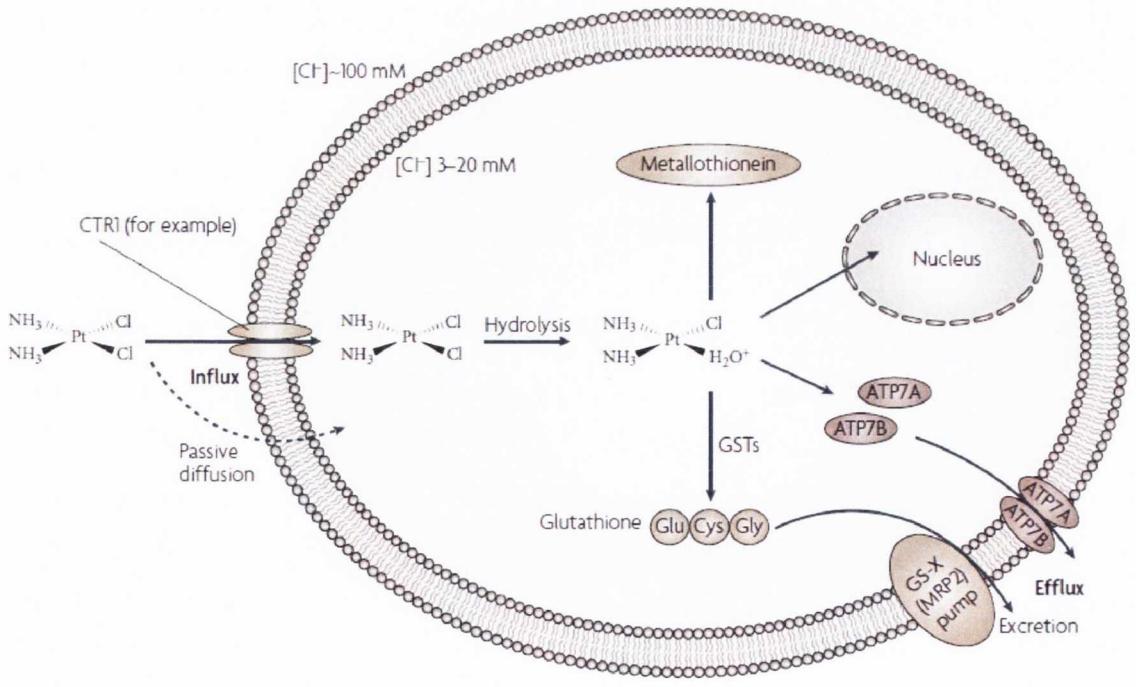


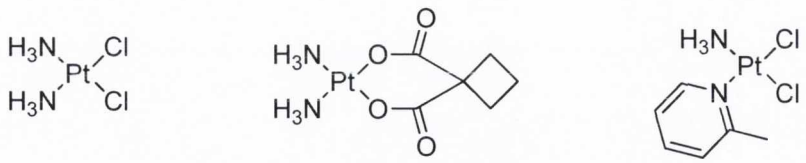
Figure 1.10. Activity of cisplatin in the cell[48]

This activated platinum complex can then enter the nucleus where it binds covalently to the N7 of two guanine bases. It must be noted that the interstrand cisplatin-DNA adduct is formed when the cisplatin drug binds to opposite guanine bases. This is preferred over the intrastrand adduct where the cisplatin binds to guanine bases on the same DNA strand.

Two limitations exist when administering a chemotherapeutic drug such as its cytotoxicity to healthy cells and the ability that the cancer cells have to build up resistance to the drug. Due to the severe side effects that cisplatin displays, research into new derivatives of this compound was carried out. Cisplatin is severely toxic to the kidneys and the gastrointestinal tract. It was therefore hypothesised that by replacing the chloride groups with a more stable group, the toxicity of the compound would be lowered without affecting anti tumour activity. This hypothesis proved to be correct and one of these new compounds carboplatin (Figure 1.11) was developed by substituting both chlorine atoms by a malonate derivative. However, due to the relative inertness of this compounds, the kinetics of adduct formation was around ten-fold slower. Consequently, around 20-40 times more carboplatin has to be administered in comparison to cisplatin.[48]

In addition, tumour cells can build up resistance against platinum containing drugs. This can be achieved mainly by two methods: i) The drugs fail to reach the DNA target or ii) the DNA actively repairs the DNA once the drug is bound. One reason why platinum drugs appear to fail to reach the DNA target can be that there are increased levels of cytoplasmic thiol-containing species in the cell (*i.e.* metallothioneins and glutathione). Once the drug has bound to these sulphur containing proteins, they become more readily exported from the cell by the glutathione S-conjugate export pump (GS-X). If the drug manages to reach the DNA target, it forms the DNA platinum complex; however, many cis-platin resistant cell lines were seen to have an increased DNA repair mechanism that is known as the nucleotide-excision repair mechanism.

Recently, another platinum complex was derived, called pico-platin (Figure 1.11). It was hypothesised that by adding steric bulk around the platinum centre, a reduction in the binding of the sulphur containing residues would be achieved. This was indeed observed and this compound, which displayed activity against a wide range of tumour cell lines, is currently in clinical trials phase III for small cell lung cancer.



cis-platin carboplatin pico platin

Figure 1.11. Some of the platinum drugs used for cancer chemotherapy

1.7.2 Metal Complexes and their interactions

In the last section, the discussion was focused on the DNA targeting drugs consisting of a platinum metal centre bound to different ligands. However, some cancers do not respond to the platinum drugs since deactivation of the metal centre can occur in the cellular cytoplasm as described in the previous section. A way to avoid this is by using different metal centres such as ruthenium since ruthenium is less thiophilic than platinum.

An important property in devising possible ruthenium based drugs is that the metal centre can vary in oxidation state from -2 in $[Ru(CO)_4]^{2-}$ to +8 in RuO_4.[50] The ruthenium metal centre, in many ruthenium based drugs, is in the +3 oxidation state and in an octahedral environment with six ligands. The ruthenium(III) centre is relatively inactive until it reaches the tumour site. Since the tumour environment contains a lower oxygen concentration and is lower in pH in comparison to healthy tissue, reduction to the more active Ru(II) species occurs. This process is called activation by reduction.[51]

Another way by which ruthenium drugs exert their *anti*-tumour activity is by binding strongly to transferrin which is a blood plasma protein used for iron delivery.[52] This provides the possibility of targeting tumour cells with a high concentration of transferrin receptors thereby potentially increasing their *anti*-cancer activity. Once the drugs are administered, the ruthenium(III) complex can bind competitively to transferrin. This is a protein with two iron binding pockets and supplies the cell with iron. Once the two iron pockets are occupied, the transferrin can bind strongly to the receptor of the cell and is taken up. Once inside the cell, the ruthenium(III) complex dissociates form the transferrin protein. However, this can only happen under acidic conditions. There are currently two ruthenium complexes in clinical trials (NAMI-A and KP1019, Fig. 1.12). Once inside the cell, the metal complex must be activated by reduction under the hypoxic conditions of the cell, and considering the increased

reductive environment of the tumour cell compared to healthy cells. Since the Ru(II)-Cl bond is significantly more labile than the Ru(III)-Cl bond, the reduced form may lead to an increased coordination to biomolecules.[53]

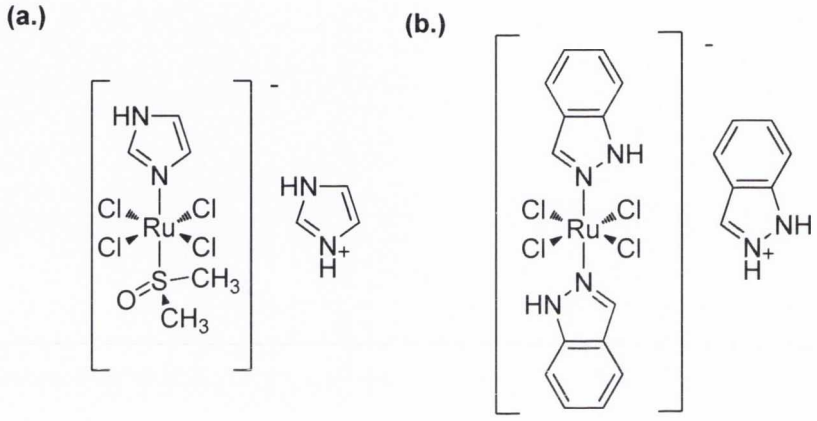


Figure 1.12. Structure of (a) NAMI-A and (b) KP1019

One way of chemoresistance among platinum complexes is the inactivation due to coordination with glutathione (GSH). However, it has been observed that once the glutathione is bound to the ruthenium centre, the bound sulfur atom can be oxidised to the sulfonate. A consequence of this is that the Ru-S bond weakens allowing the guanine base to coordinate and the ruthenium complex to exert its activity.[53]

These ruthenium complexes bind covalently to the DNA bases, sugars and backbone. Interestingly though, by substituting the ligand, the binding mode of the complex can change. For example, in the absence of a good leaving group, complexes of 2,2'-bipyridine will result in electrostatic binding in the grooves.[53] If one of these 2,2'-bipyridine were replaced with a larger polyaromatic ligand, the complex would then bind *via* intercalation (Fig. 1.13).[54]

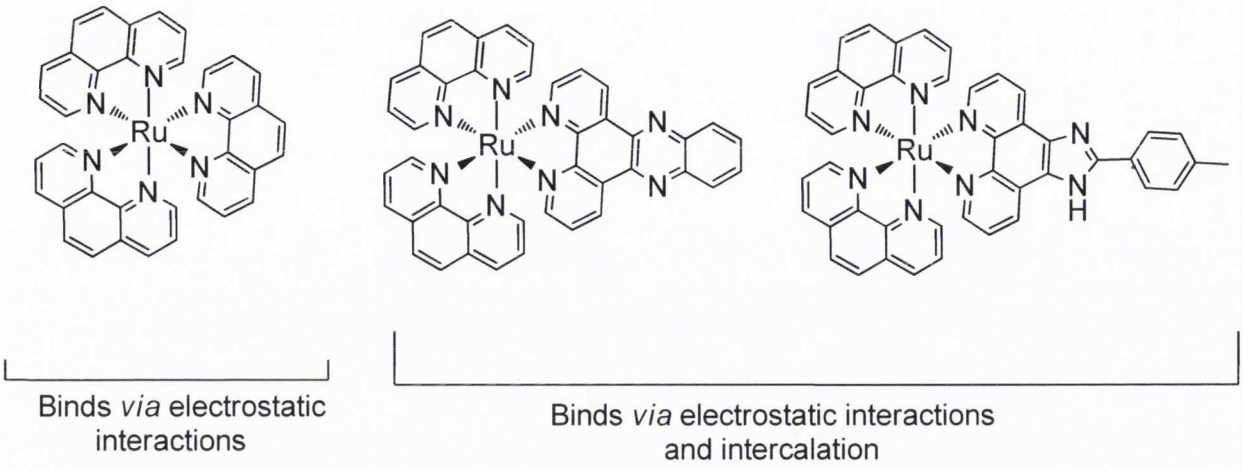


Figure 1.13. Examples of Ruthenium(II) complexes with different DNA binding modes[55]

1.7.3 Intercalators and their interactions

Intercalation was first proposed by L. S. Lerman and was defined as the 'insertion between base pairs'[56]. Intercalators are planar polyaromatic compounds that can insert in between the base pairs of DNA according to the nearest neighbour exclusion principle, which states that a molecule will insert in between every second base pair. This event causes the unwinding and lengthening of the DNA structure.

Intercalators can be either cationic or neutral. If they are cationic, an electrostatic force of attraction will occur between the DNA macromolecule and the molecule. This will be accompanied by the insertion of the planar moiety in between the DNA base pairs due to hydrophobic and π-π interactions. Examples of intercalators include proflavin and ethidium bromide. Proflavin is a nitrogen heterocycle with three planar aromatic rings (acridine core) that allow its insertion in between the base pairs. This, subsequently, forces the lengthening and unwinding of the helical structure causing potential deletions in the coding region (Figure 1.14).

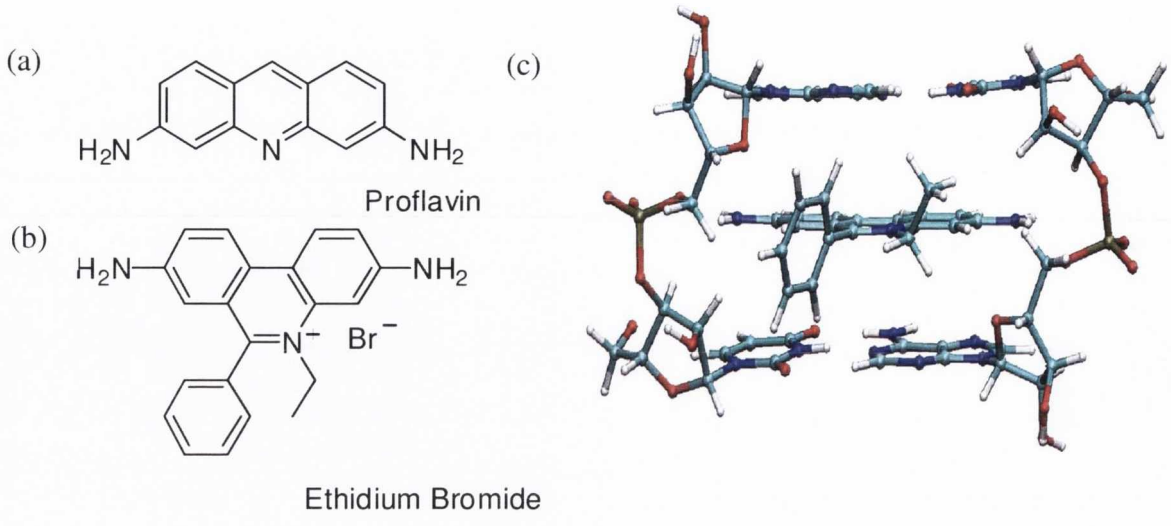


Figure 1.14. Examples of polyaromatic intercalating molecules (left) and ethidium bromide molecule inserted between DNA base pairs[57] (right)

Acridine containing drugs can be considered as the 'model' intercalators, and there is a number of aspects about acridines that can be considered relevant when explaining their anti-cancer activity:

- Firstly, their mode of binding (intercalation) is related to their anti-cancer activity[58]
- Secondly, the strength of DNA binding is known to influence the drug's potency
- Thirdly, the kinetics of binding is highly relevant since the average residence time of the drug in a binding site is related to its biological activity.[60,61]

The main way in which acridine exerts its activity is by poisoning the topoisomerase enzymes after intercalating into DNA. Topoisomerase enzymes are essential for the removal of supercoiling of DNA at the start of replication and can be divided into two main types, I and II. Topoisomerase I enzymes work by breaking single strands of DNA,[52] thereby preventing DNA supercoiling during replication whereas topoisomerase II works by breaking double strands.[62,63] Topoisimerase II enzymes can be divided into two subtypes, the topoisomerase IIα and IIβ. Since topoisomerase enzymes are required for replication, investigation into topoisomerase inhibitors or poisons is an active field of research (figure 1.15).

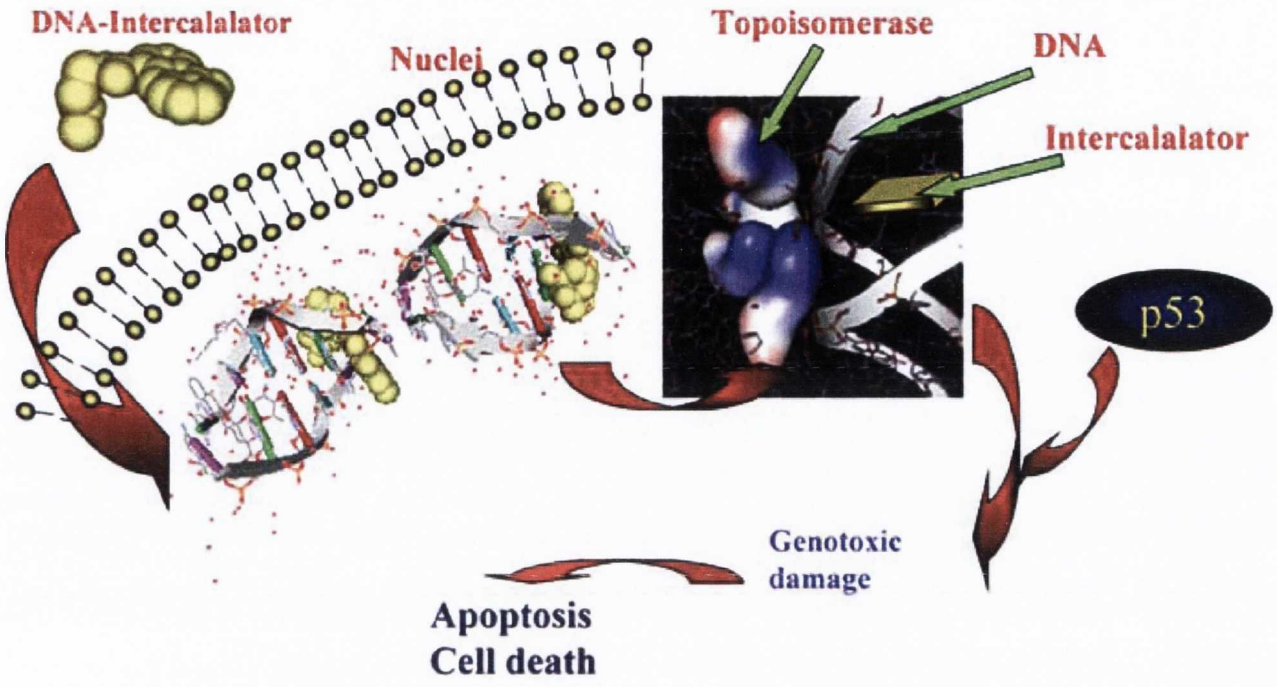


Figure 1.15. Intercalation mechanism and its effect on the topoisomerase activity[64]

An intercalator has cytotoxic ability when poisoning the topoisomerase enzymes, by the stabilisation of the DNA-intercalator-enzyme complex in a way that the replication process cannot be carried on or reversed.[64] This ternary complex is then detected by the cell as a damaged area which initiates a series of events involving the p53 protein that leads to apoptosis.[65] Since tumour cells divide much more rapidly in comparison to the normal cells, it will be the tumour cells that will be mostly affected (Figure 1.15).

There has been a large amount of interest for the preparation of cytotoxic intercalating molecules containing the acridine chromophore. Acridine containing drugs can be divided into three main groups, the 9-anilinoacridines, acridinecarboxamides and the nitroacridines. Each of these families differ by their structural features, metabolism and mechanism of action. Well known examples of a 9-anilinoacridine and an acridinecarboxamide include AMSA and DACA (Figure 1.16).

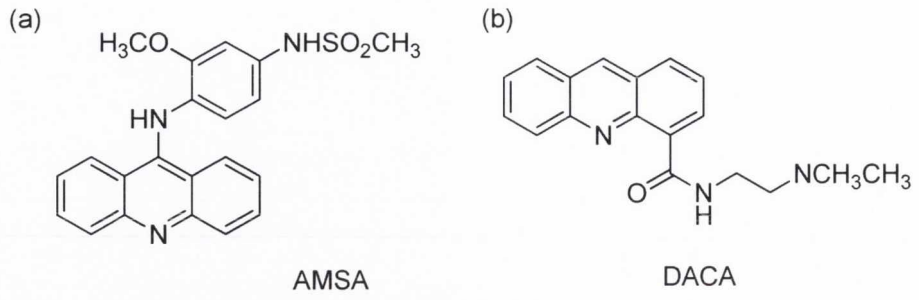


Fig. 1.16. Structures of known 9-anilinoacridines

AMSA was one of the first drugs observed to form a ternary complex with DNA and topoisomerase II.[65] It was observed to intercalate between the DNA base pairs *via* π-π interactions with the acridine chromophore, while the side chain was situated in the minor groove.[66] It is currently used for the treatment against acute myeloid leukemia, however, in contrast, it has not been effective against solid tumours.

DACA is different from AMSA in its binding to DNA. The acridine chromophore intercalates between the base pairs, but the side chain is located in the DNA major groove as verified by crystal structures.[67, 68] When it intercalates, it stimulates the formation of cleavable complexes between the DNA and topoisomerase. Interestingly, it was observed that DACA induced breakages in the presence of topoisomerase II, whereas the chloro analogue displayed converse activity, and induced breakages in the presence of topoisomerase I. This indicates that substituents on the acridine chromophore can influence their interaction with topoisomerase enzymes.[69] In addition, cytotoxicity has been observed to be dependent on the steric bulk of the side chain rather than the electronic properties.[70]

Another class of acridine compounds used as cytotoxic agents are the nitroacridines. These compounds consist of the acridine chromophore that is functionalised with a nitro group on the 1-position (Figure 1.17).

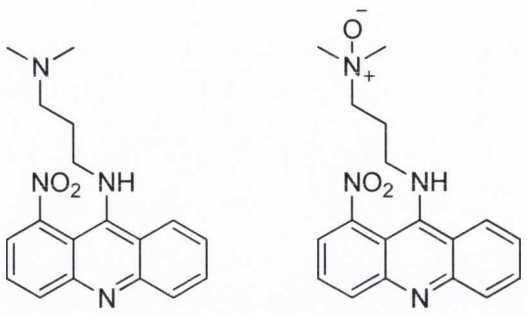


Fig. 1.17. Nitracrine (left) and its N-oxide analogue (right)

Both of these compounds were shown into intercalate to DNA, nitracrine has a 10-fold selectivity towards hypoxic cells due to the prerequisite that it must be first activated by reduction.[71] The corresponding N-oxide compound is a *bis*-bioreductive compound that must be reduced for activation and consequently 1000-fold more selective for hypoxic cells over normal cells.[72]

1.7.4 Sequence specific drugs: compounds that bind to the DNA minor groove

In section 1.6, it was explained that the B-form DNA consists of two grooves, a minor groove and a major groove. The minor groove is the narrower groove where molecules can fit with limited flexibility. Another difference between these two grooves is the larger negative electrostatic potential in the minor groove consisting of runs of AT sequences.[73] Considering that drugs that interact in the minor groove can prevent the replication of DNA or block key enzymes during transcription, the preparation of molecules that interact with that minor groove has become an important area of research.

Minor groove binders generally form hydrogen bonds with the DNA base pairs. Hydrogen bond donation is favoured in the minor groove of AT tracts, due to the ability of the N3 and O2 of adenine and thymine respectively to accept hydrogen bonds (Fig. 1.18).[73] In the case of the GC base pair, there is an exocyclic amino group from the guanine base that protrudes into the minor groove, thereby preventing close contacts between the minor groove and the molecule.[73]

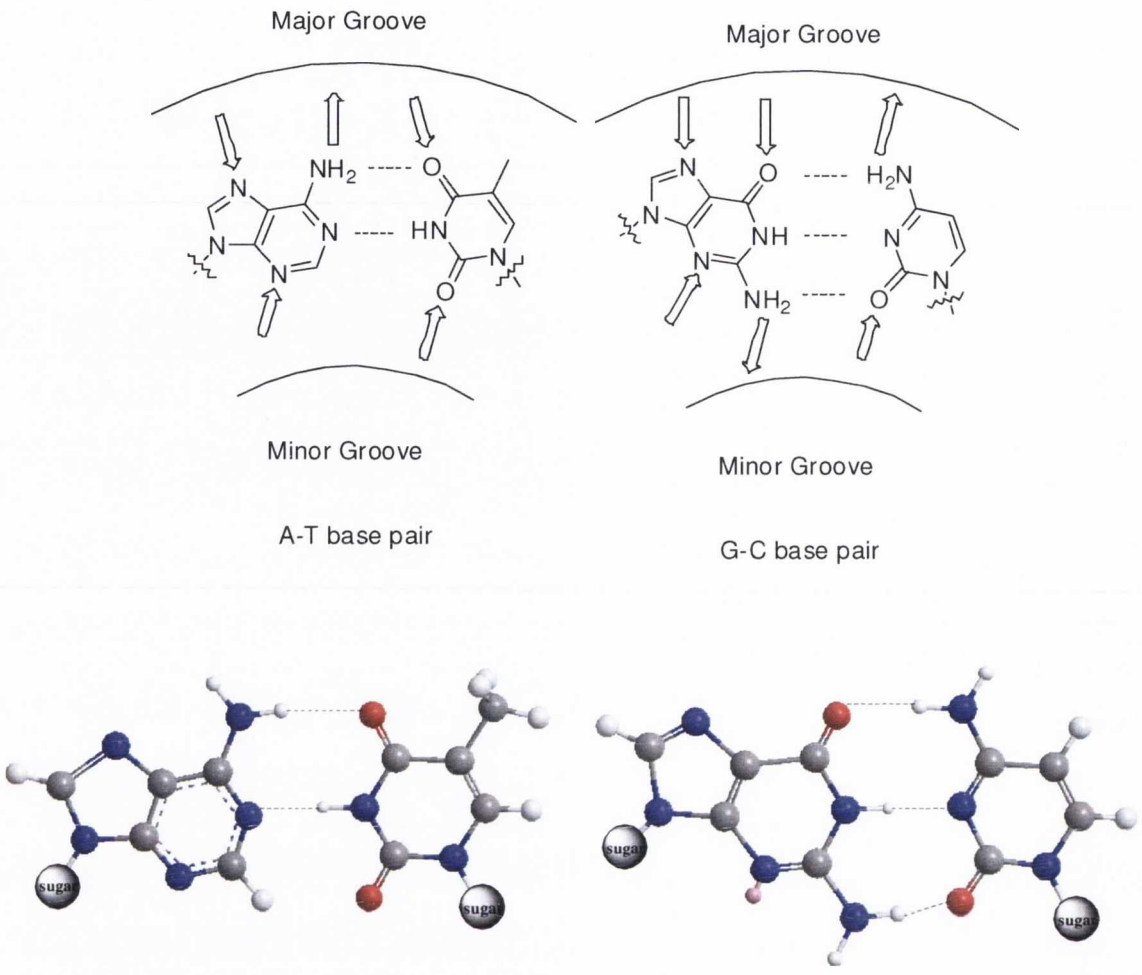


Figure 1.18. Watson-Crick base pairs showing hydrogen-bond donors and acceptors (arrows in and out) in the minor and major grooves (top), model of the AT and GC base pairs (bottom)

Minor groove binders exert their activity by interfering with those proteins that interact in that groove by causing large scale deformations. An example of these binders is the TATA-box binding protein which produces a 90° bending.[74] Thus, minor groove binders can interfere with the normal DNA processes, such as transcription, by inhibiting the DNA from undergoing the necessary structural changes or by binding at an adjacent site, changing the DNA structure and preventing the protein from readily binding.

Minor groove binders have generated huge interest due to their potential to recognise specific DNA sequences. Thereby, the therapeutic use of a molecule that could target a specific sequence in a tumour cell leading to apoptosis is enormous. Many of these

compounds that bind to the minor groove contain aromatic rings but they differ significantly from intercalators since these aromatic rings are linked rather than fused. Thereby these molecules are not necessarily planar in their free form. By definition, minor groove binders are molecules which bind to the minor groove through various interactions such as van der Waals contacts and hydrophobic interactions and to a lesser extend hydrogen bonding and electrostatic interactions. Wilson *et al.* have shown that by varying the length, number of hydrogen bonding groups and amount of cationic functionalities of the benzimidazole scaffold, the van der Waals contacts are the dominant factor for DNA binding.[75] Molecules from this class of compounds must have the following characteristics:

- positive charges
- linked rather than fused aromatic rings
- some are crescent shaped to complement the convex shape of the groove.[73]

Different factors may contribute to favouring minor groove binding over the major groove due to the structural features of the grooves. For example, the minor groove is much deeper and narrower in comparison to the major groove (Fig. 1.19). Subsequently some molecules would fit better into the minor groove.

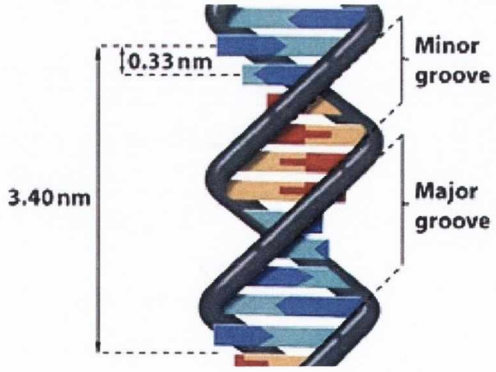


Figure 1.19. DNA minor groove and major groove[76]

Additionally, minor groove binders are more AT sequence specific than GC sequence specific because the DNA minor groove is richer in AT sequences. Therefore, correctly twisted aromatic rings in minor groove binders fit better into the AT 'channel' forming

optimal van der Waals contacts with the helical chains that define the walls of the groove. Additional specificity is derived from van der Waals contacts between the bound ligand and the edges of the base pairs on the floor of the groove.[77] A further reason why AT sequences are favoured is the ability of AT- base pairs to form hydrogen bonds through the carbonyl O-2 of thymine or the nitrogen N-3 of adenine. Even though similar functional groups are available in GC base pairs, the amino group of guanine presents a steric block to hydrogen bond formation at the N-3 nitrogen. Also, the inter-base hydrogen bond between the guanine amino group and the cytosine carbonyl oxygen, lies in the minor groove and as a result of this interaction, the penetration of small molecules into the minor groove at GC sites is inhibited. Finally, another reason because these binders generally prefer the minor groove of AT sequences is that the negative electrostatic potential is greater in AT minor grooves than in GC minor grooves,[73] and minor groove binding ligands are often positively charged.

As mentioned previously, a good example of a minor groove binder is netropsin (Figure 1.20) because it exhibits several key structural features that make it an excellent minor groove binder

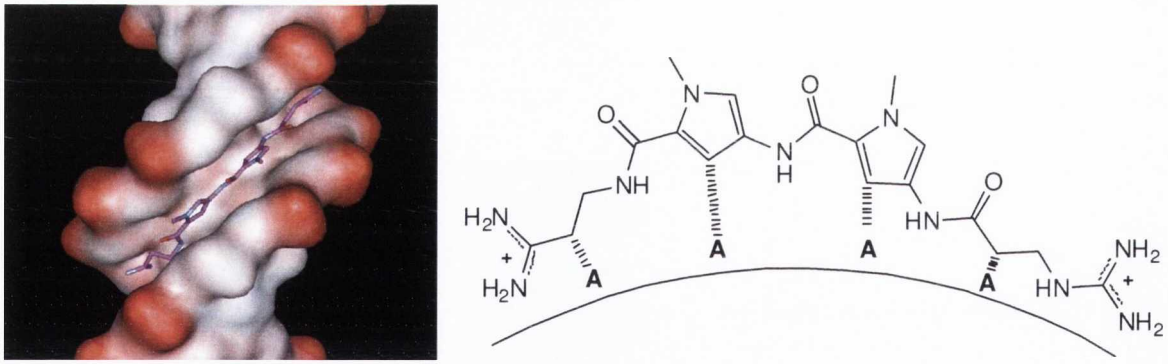


Fig. 1.20. Netropsin bound into the minor groove[73] and van der Waals contacts between the DNA base pairs (A=adenine)[73,78,79]

Firstly, this drug fits in the narrow minor groove and the complex is stabilised by a combination of electrostatic interactions between the negative potential in the minor groove of DNA and the cationic end groups of netropsin. It was reported that netropsin binds over

four consecutive AT base pairs and when binding, the deep penetration is highly enthalpically driven and the positive entropy values result from molecular events such as the expulsion of the spine of hydration.[78] Stabilisation also results from the hydrogen bonds between the ligand and the N-3 of adenine and O-2 of thymine in the minor groove. Thus, the amino group of the three peptide bonds point inwards and form three centred hydrogen bonds as hydrogen bond donors with the O-2 of thymine.[79] Another interaction between netropsin and the DNA minor groove comes from van der Waals contacts (Figure 1.20) contributing to the selectivity for the adenines over the guanine bases.

Another oligopeptide natural product with similar structural features as netropsin is distamycin which was extracted from *Streptomyces distamyces.*[73] This drug inihibits the binding of RNA polymerase by binding to AT tracts in the minor groove.[80] Two principle differences exist between these drugs:

- distamycin has an extra pyrrole-amide linker (Fig. 1.21) making the molecule significantly longer than netropsin. Subsequently, distamycin spans five base pairs on DNA binding,
- distamycin is mono cationic since the guanidinium end on netropsin is replaced by a terminal aldehyde group (Fig 1.21).

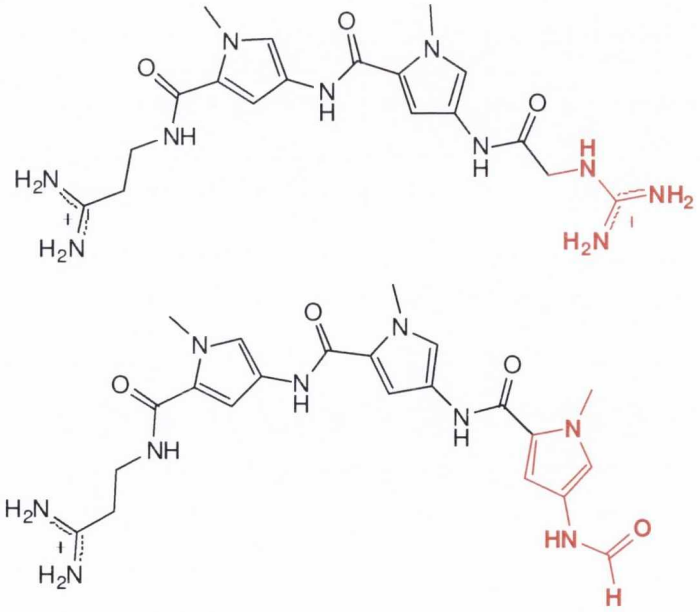


Fig. 1.21. Differences between distamycin and netropsin

These structural features contribute to the different binding of these molecules to DNA. At low concentrations, it was observed by NMR that distamycin displays primarily a 1:1 binding mode to the d(GCGCAAATTTGCGC) hairpin; however, at higher drug concentrations, a 2:1 binding mode was preferred.[81] This 2:1 binding motif resulted in the widening of the narrow minor groove of the 5'-AAATTT tract, from 3.4 Å to 6.8 Å. Therefore, we can deduce that the minor groove is not rigid but is very flexible especially upon the binding of a molecule.[82]

Following the investigation of netropsin and distamycin, it was suggested by Dickerson[79] and Lown[83] that replacing the pyrrole ring on the netropsin or distamycin molecule for an imidazole ring (Figure 1.22) will introduce GC selectivity due to the hydrogen bond accepting ability of the extra nitrogen. It was found that these newly modified molecules, termed lexitropsins, displayed an increased affinity for GC base pairs and decreased affinity for AT sequences.[84] This was due to the width of the minor groove of AT tracts being considerably narrower than the minor groove of natural DNA that contains a random sequence of DNA bases. Therefore, it seemed that a wide groove would not bind a single molecule well since such a molecule would not be able to contact both walls of the minor groove. Although, some GC selectivity was observed, the consequence of the widened groove was to obtain reduced binding constants because the van der Waals distances were too large. Considering that the van der Waals attraction forces are inversely proportional to the radius to the power of six, these forces of attraction would be significantly lower than when binding to a groove that contained only AT base pairs.

Further studies on GC selectivity were successful. Thus, since the GC minor groove is wider than the corresponding AT one, it was proposed that a 2:1 distamycin motif would be necessary for GC selectivity. Two netropsin molecules could not show this motif since the cationic ends would repel each other; consequently, only distamycin molecules were modified[85] yielding a new class of compounds called polyamides. Evidence of this dimeric bridging was obtained analysis of these molecules that exhibited the 2:1 binding while maintaining GC selectivity was resolved.[86] In these molecules, the positioning of a hydrogen bond accepting group proved to be a very important factor since it was required for the interaction with the amino group of guanine.[87]

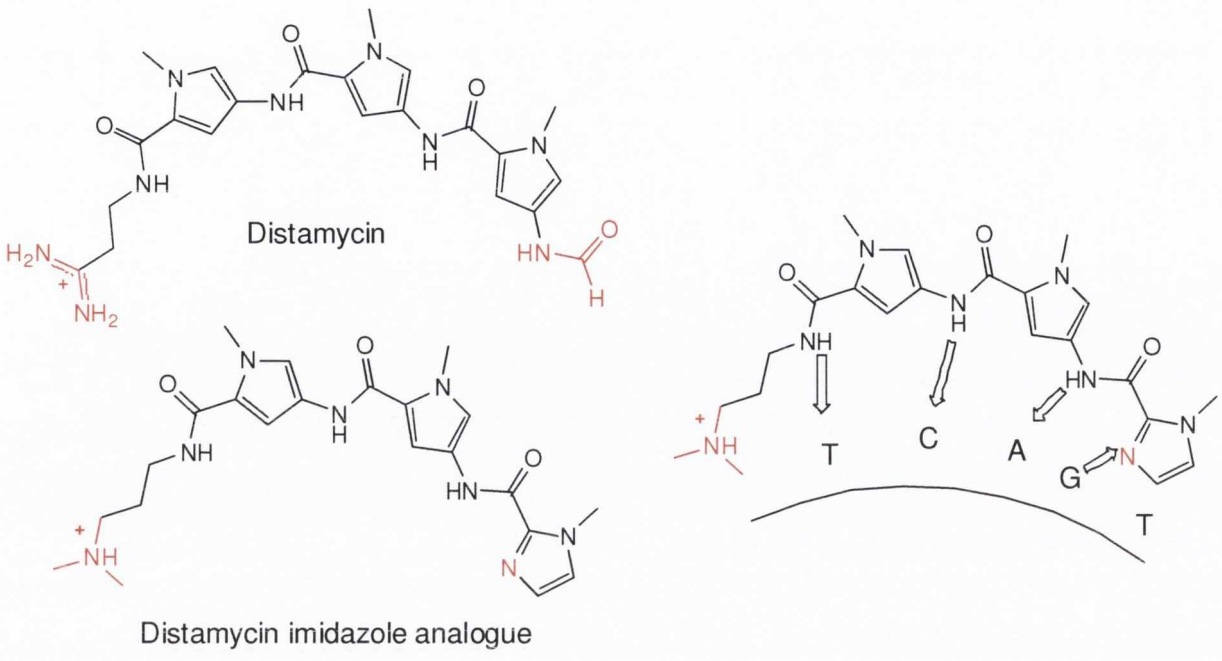


Figure 1.22. Distamycin and an imidazole analogue[84] (left) Hydrogen bonding between the DNA bases and the distamycin analogue (right).[73]

On the basis of this and other observed examples of GC selective molecules, two characteristics seemed to be required to achieve GC selectivity:

- the 2:1 binding motif should be used for increased GC selectivity and
- the tails of the molecules must be interacting with AT base pairs.[85]

Another class of minor groove binders are the *bis*-amidinium molecules. These are molecules that contain a diaromatic scaffold connected with a linker of various length and composition, and are crescent shaped to complement the convex shape of the groove forming optimal hydrogen bonds with the DNA bases. Examples of this series include pentamidine and berenil (Figure 1.23).[73]

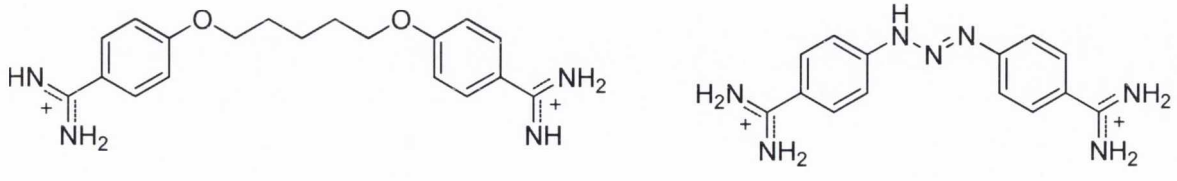


Fig. 1.23. Some bis-amidinium molecules

A crystal structure was obtained of berenil with the oligonucleotide d(CGCGAATTCGCG)$_2$.[73] Interestingly, the main difference between the distamycin and netropsin DNA complexes and the berenil one is a water molecule that was observed to act as a bridge for hydrogen bonding to the N3 of adenine. This molecule was also seen to lie parallel to the groove walls in order to form close van der Waals contacts with the DNA bases. The triazine linker was seen to have no contacts with the DNA bases.[73]

By analysing the structure of the complex of pentamidine with DNA, hydrogen bonds were observed between the amidinium cations and the N3 of adenine.[86] However, the water molecules were seen to have no effect on the binding of this molecule to DNA since it is long enough to reach the base pairs. Also the phenyl rings are twisted by 35° relative to each other so they form strong van der Waals interactions with the DNA bases.[86]

Among these *bis*-amidinium molecules, one of the most important systems is that of furamidine and its analogues.[89] Furamidine has two aromatic moieties, linked by a furan heterocycle, and basic amidine ends which are cationic at physiological pH. The crescent shape of the furamidine molecules complements the convex shape of the minor groove, allowing it to optimise the van der Waals contacts and hydrogen bond interactions with the DNA bases. Furamidine is currently in phase III of clinical trials for African trypanosomes but now abandoned due to toxicity problems.

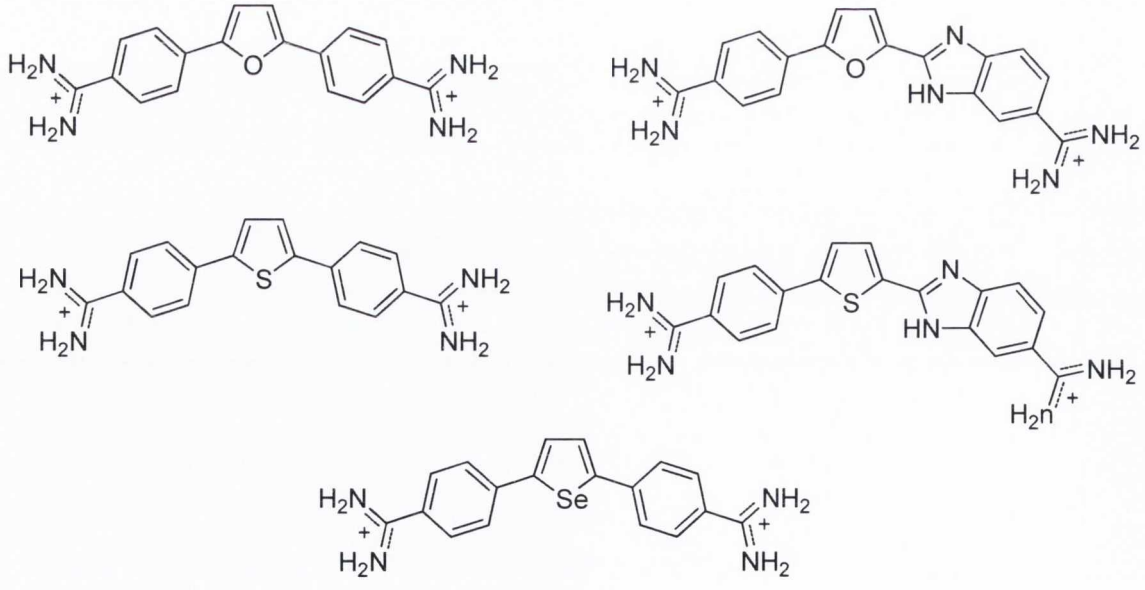


Figure 1.24. Furamidine and some of its analogues

By changing the central furan linker, the aromatic moieties and the nature of the cationic groups, the length of the molecule and the dihedral angle between the two aromatic moieties could be modified,[90,91] thus affecting the binding affinity (Figure 1.24).

When one of the phenyl rings was replaced with a benzimidazole moiety, the distance between the central furan linker and the amidine cation increased; however, a significant increase in the DNA binding was not observed. Likewise, when the central furan ring was replaced by a thiophene ring, a similar result was observed. However, when the furan linker and one of the phenyl rings were replaced with thiophene and benzimidazole respectively, the binding affinity increased by ten-fold.[90] This was due to the slight deviation from the crescent shape requirement for classical minor groove binders, thus allowing for the formation of increased favourable hydrogen bonding and van der Waals contacts.

The analysis of the binding of these compounds was completed by obtaining a full thermodynamic DNA binding profile. It was found that the thiophene-benzimidazole analogue of furamidine displayed a more favourable enthalpy; however this was compensated by a less favourable entropic value. This was due to the sulphur atom being slightly larger than oxygen resulting in a decreased crescent shape, what, affected the position of the amidine groups, allowing optimal hydrogen bonding and van der Waals contacts between the drug and DNA bases. Due to these strong interactions, the compound was twisted and 'locked' into place in the minor groove, thereby, increasing the order within the DNA-drug complex.[90] This would explain the slightly unfavoured entropy observed in the benzimidazole analogue.

Another furamidine derivative studied by Boykin *et al.* was the selenium analogue.[90] It was shown that by substituting the oxygen atom of furamidine for the larger selenium heteroatom, the curvature of the molecule decreased while the helical twist among the five and six membered rings increased. This subsequently, caused a less favourable entropy value on DNA binding which was compensated by a more favourable negative enthalpic value. This is possibly due to the difference in the hydrophobicity of the selenium in comparison to the oxygen atom.[90] However, without a crystal structure this difference in binding affinity cannot be properly proved.

Another revolutionary result in this field was the binding of linear molecules to the minor groove. It was expected that curved minor groove binders would display a favoured binding due to its complementary shape than linear molecules. However, Wilson *et al.* have shown that the linear compounds display a higher binding affinity than its curved counterpart. x ray crystallography showed the involvement of a water molecule in the binding of this molecule. By incorporating water molecules, this compound would achieve the necessary crescent shape required to complement the convex shape of the groove. [91]

Recently, the increased binding affinity by non-classical minor groove binders has been extensively studied.[92] It has been found that some non-classical groove binders such as CGP 40215A bind with increased affinity in comparison to furamidine.[93] It is expected that this molecule would display poor binding affinity for the DNA minor groove due to the absence of the normal crescent shape. However, compared to furamidine, it displayed a 5-fold increase in binding affinity. The crystal structure of the drug-DNA complex showed some interesting features. Firstly, one of the cationic amidine ends form direct hydrogen bonds with the O2 of T8 along with indirect hydrogen bonds through a bridging water molecule, with the N3 of the complementary A17 and O2 of C9. The –NH- group on the linker forms bifurcated hydrogen bonds with the interstrand O2 of T7 and T19,[89] while the amidine on the other end of this molecule protrudes away from the groove. Consequently, two water molecules form a hydrogen bonding bridge between this amidine group and the base pairs. Therefore, this molecule displays optimal binding regardless of the lack of the curvature leading one to question the model of classical minor groove binders (Figure 1.25).

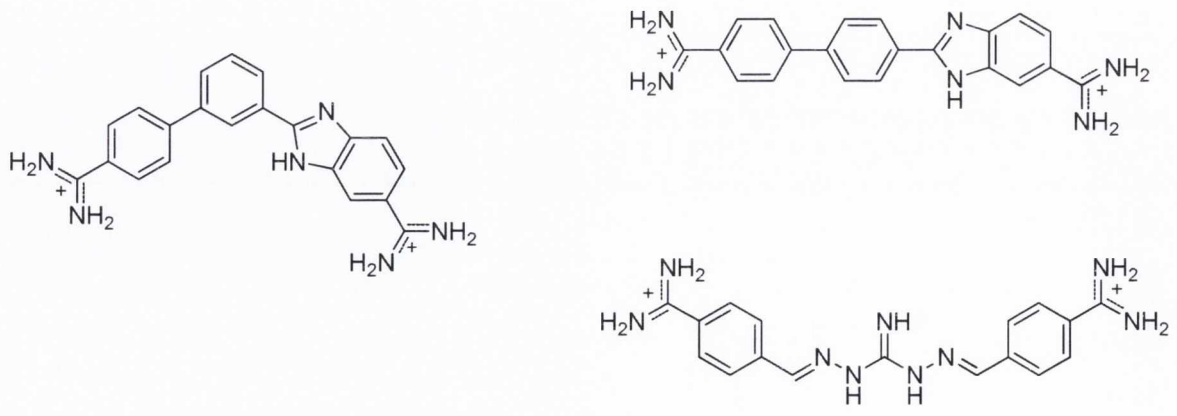


Figure 1.25. Molecules that deviate from the classical minor groove binder

Another class of groove binders are compounds with the ability to bind to GC sequences with an affinity comparable to that for AT sequences. Boykin *et al.* have studied a class of compounds based on the structure shown in Figure 1.26, with the idea that an increase in the GC selectivity should be observed. Given the nature of the guanine base and the way the exocyclic amino protrudes into the minor groove, this would be an unlikely result. However, these molecules were seen to display not only affinity for AATT but also slightly reduced affinity for AAAGTTT sites.[93] Considering this surprising result, it was thought that these molecules could either bind favourably to the GC bases at the ends of the AT sites, or adopt a folded conformation. No definite conclusions could be made due to the lack of a crystal structure. Regardless, this result leads to exciting possibilities of design molecules that could recognise specific sequences over others, therefore leading to potential selective *anti*-tumour drugs.

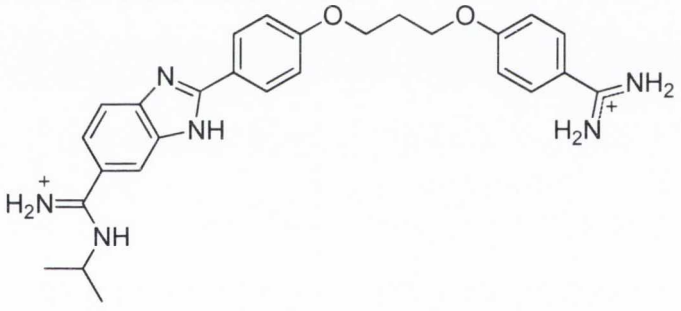


Figure 1.26.Examples of the potential GC selective molecules prepared by Boykin et al.[93]

1.8 *Bis*-guanidinium and *bis*-2-aminoimidazolinium molecules

Another important class of compounds was developed by Rozas *et al.*[94] The main difference of these compounds from furamidine is that they contain different cationic functionalities. By replacing the amidine for the guanidine functional group, an extra point of contact is introduced, thereby potentially increasing the DNA binding affinity. The central linker was also varied to investigate the optimum geometry for DNA binding. Examples of linkers that were studied include the -NH-, -CH_2- and -CO- linkers (Figure 1.27).

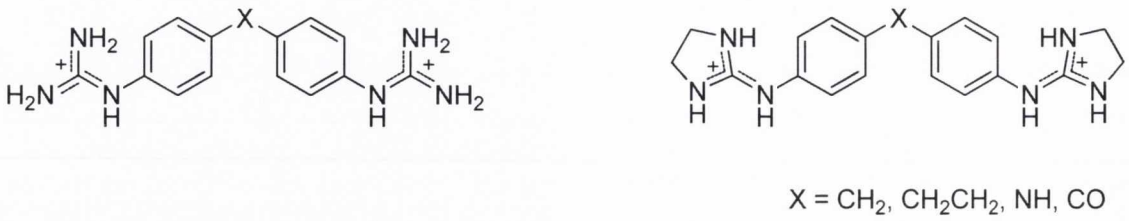


Figure 1.27.Bis-guanidinium and bis-aminoimidazolinium compounds prepared by Rozas et al.

These compounds were tested for *anti*-trypanosomal activity, and initially, their DNA binding affinity to AT sequences was investigated by carrying out thermal denaturation assays (Table 1.3). The results of these experiments indicate that these compounds show interesting binding with evidence for sequence selectivity. [95]

Table 1.3.- Thermal denaturation results obtained for selected bis-guanidines and bis-2-aminoimidazolines[95]

Linker	*Bis*-guanidine ΔT_m/°C [a]	*Bis*-imidazoline ΔT_m/°C [a]	$\Delta\Delta T_m^{diff}$ [b]
CH_2	**15.0**	**18.8**	**3.8**
CH_2CH_2	**26.1**	**31.2**	**5.1**
NH	**29.6**	**38.5**	**8.9**
CO	**27.6**	**-**	**-**

[a] Thermal denaturation experiments were carried out with poly(dA-dT)$_2$ DNA in order to assess the relative binding strength for AT sequences. Experiments were carried out in MES buffer with a Bp/D ratio 3.

[b] $\Delta\Delta T_m^{diff}$ = *Bis*-2-aminoimidazolinium ΔT_m – *Bis*-guanidinium ΔT_m

By analysing the thermal denaturation assays, it can be concluded that, firstly, by comparing the results for the *bis*-guanidine and *bis*-2-aminoimidazoline family, the latter displayed better DNA binging affinity. It was suggested that a reason for this was the potential for the formation of extra van der Waals contacts with the DNA bases through the ethylene bridge on the imidazoline functionality.

Secondly, by comparing the bridges that link the diaromatic scaffold (X group in Figure 1.27), we can see that the order of binding corresponds to NH > CO > CH_2CH_2 > CH_2.

The main difference between these molecules is their different geometries at the linker, and this can affect the formation of hydrogen bonds to the DNA base pairs. For example, the NH linker has a trigonal geometry and the corresponding dication has the characteristics of a classical minor groove binder as described before. It was therefore expected to bind to DNA with high affinity. By comparing the geometrical differences between the molecules with the CO and CH_2 linkers, it can be seen that the tetrahedral geometry of the CH_2 linker corresponds to the molecule with the worst binding. This can be attributed to the molecules non-planarity. However, more definite conclusions can only be made with thermodynamical information and a crystal structure of the DNA complex so that knowledge of the enthalpy and entropy of binding can be correlated to the x-ray structure. Other possible investigations that could be carried out would be on the selectivity of these molecules for different sequences of DNA in order to analyse if these symmetrical molecules can bind selectively to specific sequences.

Fortunately, a crystal structure of the *bis*-2-aminoimidazoline derivative where the diaromatic scaffold is linked with an –NH- group complexed to the 5'AATT site was obtained (Figure 1.28).[96] The crystal structure, shows that this molecule bonds to the N3 and O2 of adenine and thymine respectively via hydrogen bonds with the imidazolinium cations (Fig. 1.28). Further experiments were then carried out to investigate the sequence selectivity of these molecules. From Surface Plasmon Resonance experiments, it was shown that this compound is more selective for this site than the alternating 5'ATAT site. This illustrates that these compounds display a high degree of sequence selectivity, and the need for further investigation into compounds of this nature for potential *anti*-cancer activity.

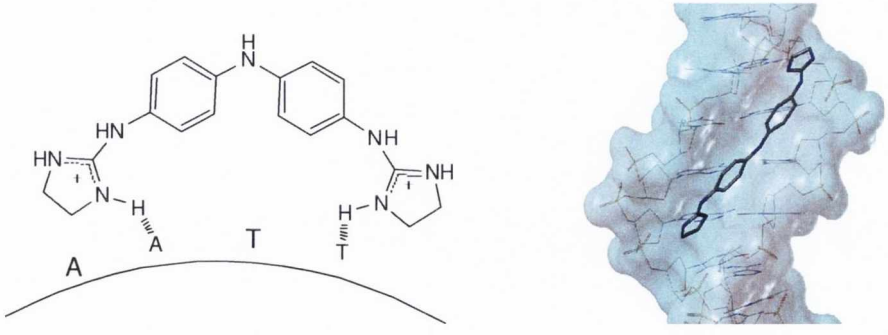


Fig 1.28. Hydrogen bonding between the bis-2-aminoimidazolinium derivative (NH linker) and the DNA bases (left), and the x-ray crystal structure of the molecule bound to the DNA minor groove.

1.9 Combilexins and other dual binding molecules

Recently, an active area of research has started to be developed, which involves combining two different DNA targeting functionalities. An important series of compounds within this group are prepared by attaching sequence specific molecules to an intercalator. By doing so, the sequence specific part of the molecule will serve as a hook with the minor groove binding part serving as a director for specific sequences while the intercalator inserts in between the base pairs, thereby exerting its activity in certain areas of the DNA macromolecule. Examples of such structures are shown below (Figure 1.29).

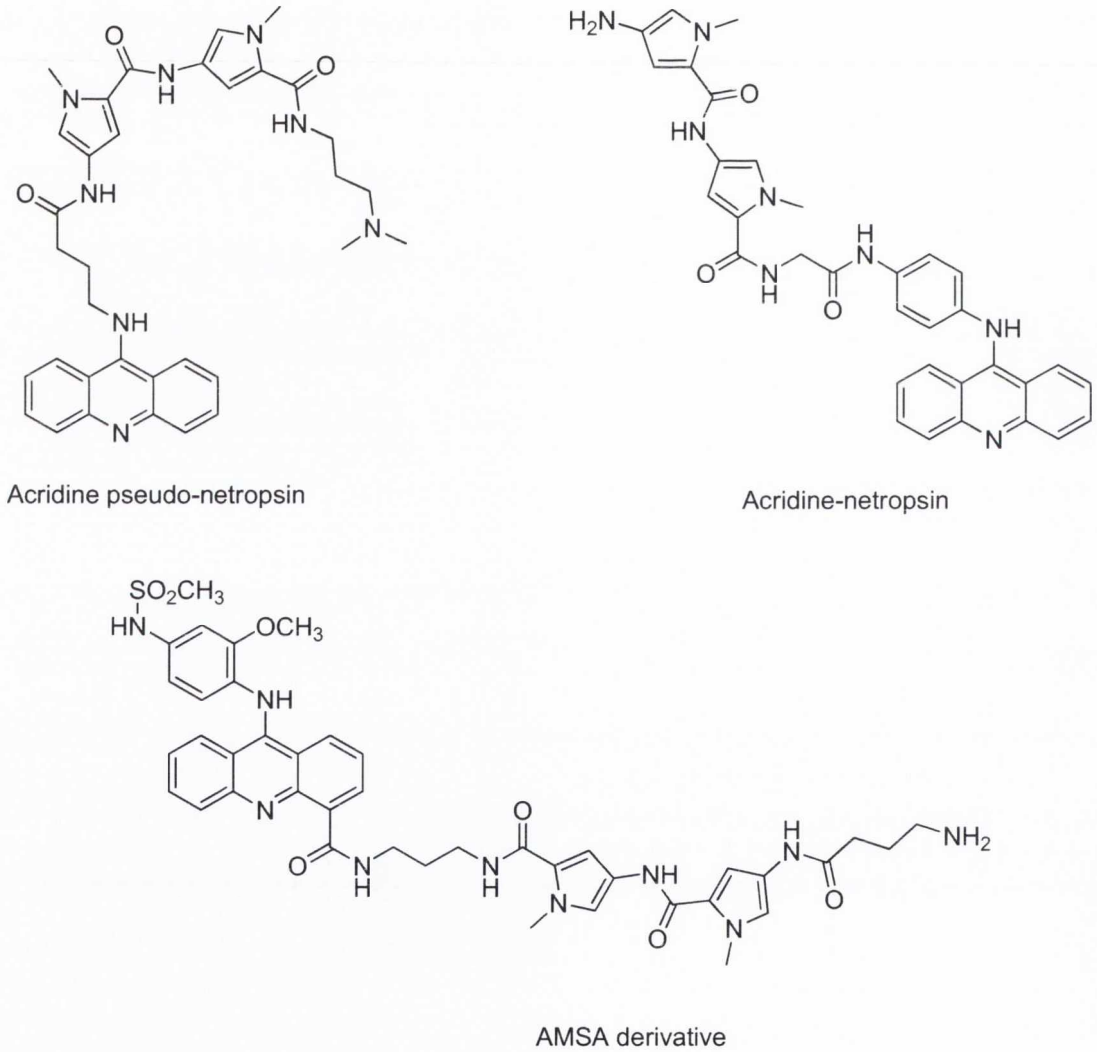


Figure 1.29. Some of the combilexins studied to date

These compounds were shown to display an increased binding affinity in comparison to the minor groove binding analogues due to the increased binding site size that was found to be 5-6 base pairs long. Importantly, the presence of the acridine ring, in the acridine pseudo-netropsin and acridine-netropsin, did not affect the AT sequence selectivity.[97] Another combilexin drug, the AMSA derivative[97,98] displayed an excellent binding affinity for AT base pairs due to the sequence specific netropsin binding unit.

Recently, Pindur *et al.* have prepared a series of combilexin molecules based on the intercalator oligopyrrole carboxyamide. These molecules display dual binding mode that was investigated by circular dichroism. Again the minor groove binding subunit was used to direct the molecule to AT base pairs and a series of intercalators and linker lengths were studied. Interestingly, it was found that there was a limit in the linker length between the two binding subunits and no increase in the binding affinity of these molecules was observed after this limit.[99]

Other dual action molecules have also been developed contain nitrogen mustards attached to intercalators or sequence selective molecules. As previously stated, nitrogen mustards exert their cytotoxicity by covalently attaching into the DNA non specifically and usually to the N7 of the guanine base. However, by attaching an intercalator the selectivity shifted from the N7 of the guanine base to the adenine N3.[100] Most of the work was carried out on alkylator-groove binder dual agents since these offer higher selectivity over intercalators. For example, analogues of distamycin were attached to a nitrogen mustard since the distamycin analogue displays its adenine-thymine selectivity. An example is tallimustine[101] which was selected for clinical trials. Interestingly, this molecule react by mono alkylation at the N3 position of adenine in the sequence $d(TTTTCA)_2$ and contrasting to this, a single base modification completely eliminated alkylation.[102] It was seen that by replacing the chlorides of the nitrogen mustard for bromides, the cytotoxicity increased considerably. The number of pyrroleamide units significantly affected the sequence selectivity, the mono pyrrole analogue mainly alkylated the N7 of guanine whereas the tripyrrole analogue showed increased selectivity for the N3 of adenine. It was also found that when the nitrogen mustard was directly attached to the pyrrole ring, a decrease in selectivity and higher levels of intrastrand alkylation were observed.[103, 104]

1.10 Physicochemical Investigations for evaluating DNA affinity

There are many biophysical techniques that can be employed to characterise the DNA binding affinity of our molecules. These techniques can be divided into two sections, the optical and non-optical techniques. Optical techniques rely on the absorption of light such as thermal denaturation assays, UV and fluorescence spectroscopy. Non-optical techniques involve measurements that do not utilise light such as Isothermal Titration Calorimetry (ITC), Surface Plasmon Resonance (SPR) and Nuclear Magnetic Resonance (NMR).

Optical techniques such as UV spectroscopy are often utilised when studying DNA interactions. UV spectroscopy relies on the principle of measuring the difference in absorption of radiation before and after the interaction with the sample. The absorption of a sample is directly related to the sample concentration, subsequently, a very concentrated sample will absorb more radiation. This principle is applied to study DNA interactions. For example, in a thermal denaturation experiment, a sample of DNA is heated and its corresponding absorbance changes as a result of the DNA double helix unwinding into two single helices, since the bases that are responsible for the light absorption are more exposed in that form than when they were in the double helix. When a molecule binds to DNA, it stabilises the double helix, thus extra energy in the form of heat would be required to unwind the DNA double helix. The difference in this required energy is a good indication of the binding strength of the compound. Another use of UV spectroscopy is to observe absorption changes when a titration is carried out. The principle behind this technique involves monitoring spectral changes upon the addition of a compound to a DNA solution. An advantage of this technique over thermal denaturation assays is that it provides information on binding affinities quantitively without using temperature.

Another optical technique used in the study of DNA targeting is circular dichroism (CD) which relies on the difference in the absorption of right handed and left handed circularly polarised light. Chiral molecules absorb different amounts of circularly polarised light. Since DNA in the double helix form is chiral, it will produce a CD spectrum. This technique is utilised in measuring the binding affinities and the induced structural changes upon the binding of an achiral molecule to DNA. It has many advantages such as the low concentrations of sample that are required and that a CD spectrum could be collected within minutes.[105]

Linear Dichroism (LD) is another optical technique used to study the interactions of compounds with DNA, based on the difference in the absorption of parallel and perpendicularly polarised light. In this experiment, the DNA absorbs different amounts of these lights, resulting in an LD spectrum. Upon the addition of a compound the difference in absorption will change and an induced LD signal would be observed indicating the mode of binding for the compound (intercalation/minor groove binding).[105]

Non-optical techniques that are commonly used to study interactions with DNA and were used in this project include isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR). SPR involves measuring the binding constants to specific hairpins, and can be defined as the excitation of surface plasmons by light, taking advantage of the refractive index of light. In these experiments, DNA hairpin strands that have specific base composition such as $d(GCGCAAATTTGCGC)_2$ or $d(GCGCATATGCGC)_2$ and their complementary strands are attached to a sensor chip which is normally gold. The compound is then injected over the sensor which causes a change in the refractive index upon binding. This is observed in the sensorgram. The molecules with a strong DNA binding affinity will produce large signals whereas weak DNA binders will display weak signals.[106]

ITC is a powerful technique used to measure the thermodynamics of binding. In only one experiment, the binding stoichiometry, binding constant, and the enthalpy of binding can be determined. It is based on the principle of measuring the heat changes when a compound binds to DNA. Using thermodynamic equations, the free energy and entropy of binding can be evaluated, thus, much information regarding the binding event can be determined. Its true potential cannot be realised without structural data since when used in combination, one could correlate the differences in the enthalpy and entropy of binding to structure features such as the number of hydrogen bonds and the geometry of the ligand inside the DNA.[107] One of the techniques that can be used to obtain structural information is called x-ray crystallography which is based on deducing the coordinates of the atoms in a given crystal using x-rays.

Other techniques not used in the present work but yet relevant for studying DNA interactions are DNAase I Footprinting which detects DNA protein interactions because the protein bound to DNA will protect it from enzymatic cleavage; and Nuclear Magnetic

Resonance (NMR) which allows measuring DNA interactions and monitoring changes in chemical shift upon the binding of a molecule.

1.11 References:-

1. World Health Organisation (www.who.int)
2. D. Voet and J. G. Voet, *Biochemistry*, John Wiley and Sons, Inc., California, 2002
3. National Registry of Ireland (www.ncri.ie)
4. National Registry of Ireland (www.ncri.ie)
5. (www.medterms.com)
6. Radiological Protection of Ireland (www.rpii.ie)
7. Boffetta, P.; Pershagen, G.; Jockel, K. H.; Forastiere, F.; Gaborieau, V.; Heinrich, J.; Jahn, I.; Kreuzer, M.; Merletti, F.; Nyberg, F.; Rosch, F.; Simonato, L.; *J. Natl. Can. Ins.*, **1999**, *91*, 697
8. Lowry, D. R., Schiller, J. T., *J. Clin. Invest.* **2006,** *116,* 1167
9. http://cancer.about.com/od/causes/a/geneticcancer_2.htm
10. American Chemical Society (www.cancer.org)
11. www.cancerhelp.org.uk/about_cancer/what_is_cancer/cells/thecancercell
12. Semenza, G. L., *J. Clin. Invest.* **2008,** *118,* 3835
13. Kim, K. J., Li, B., Winer, J., Armanini, M., Gillett, N., Phillips, H. S., Ferrara, N., *Nature,* **1993,** *362,* 841
14. Ueki, N., Nakazato, M., Ohkawa, T., Ikeda, T., Amuro, Y., Hada, T., Higashino, K., *Biochim. Biophys. Acta.,* **1992,** *1137,* 189
15. Fearon, E. R., Vogelstein, B., *Cell,* **1990,** *61,* 759

16. Burger's Medicinal Chemistry and Drug Discovery Sixth Edition Volume 5 by Donald Abraham

17. Hennings, H.; Glick, A. B.; Greenhalgh, D. A.; Morgan, D. L.; Strickland, J. E.; Tennenbaum, T.; Yuspa, S. H.; *Proc. Soc. Exp. Biol.*, **1993**, *202*, 1

18. Hollstein, M., Sidransky, D., Vogelstein, B., Harris, C. C., *Science,* **1991,** *253,* 49

19. Ventura, A., Kirsch, D. G., McLaughlin, M. E., Tuveson, D. A., Grimm, J., Lintault, L., Newman, J., Reczek, E. E., Weissleder, R., Jacks, T., *Nature,* **2007,** *445,* 661

20. Baker, S. J., Fearon, E. R., Nigro, J. M., Hamilton, S. R., Preisinger, A. C., Milburn Jessup, J., VanTuinen, P., Ledbetter, D. H., Barker, D. F., Nakamura, Y., White, R., Vogelstein, B., *Science,* **1989,** *244,* 217

21. Markowitz, S. D., Bertagnolli, M. M., *N. Engl. J. Med.,* **2009,** *361,* 2449

22. Volkert, W. A., Hoffman, T. J., *Chem. Rev.*, **1999**, *99*, 2269

23. Castano, A. P., Mroz, P., Hamblin, M. R., *Nature Reviews,* **2006,** *6,* 535

24. Valentini, A. M., Armentano, R., Pirrelli, M., Caruso, M. L., *Cancer Treatment Rev.*, **2006**, *32*, 607

25. American Chemical Society (www.cancer.org)

26. Geroni, C., Marchini, S., Cozzi, P., Galliera, E., Ragg, E., Colombo, T., Battaglia, R., Howard, M., D'Incalci, M., Broggini, M., *Cancer Research,* **2002,** *62,* 2332

27. http://images2.clinicaltools.com/images/gene/molecular/dna_bases_nhgri.jpg

28. Gregory, S., Barlow, K. F., McLay, K. E., Kaul, R., Swarbreck, D., Dunham, A., Scott, C. E., Howe, K. L., *Nature,* **2006,** *441,* 315-321

29. Stephen Neidle, *Nucleic Acid Structure and Function*, Oxford Press 2002

30. Webster, G. D., Sanderson, M. R., Skelly, J. V., Neidle, S., Swann, P. F., Li, B. F., Tickle, I. J., *Proc. Natl. Acad. Sci. USA.,* **1990,** *87,* 6693

31. http://upload.wikimedia.org/wikipedia/commons/1/13/A-B-Z-DNA_Side_View_Transparent.png

32. Rich, A., Norheim, A., Wang, A. H. J., *Annual Review of Biochemistry,* **1984,** *53,* 791

33. Wing, R. M., Drew, H. R., Takano, T., Broka, C., Takana, S., Itakura, K., Dickerson, R. E., *Nature,* **1980,** *287,* 755

34. Drew, H. R., Dickerson, R. E., *J. Mol. Biol.,* **1981,** *151,* 535

35. Kubinec, M. G., Wemmer, D. E., *JACS,* **1992,** *144,* 8739

36. Yoon, C., Prive, G. G., Goodshell, D. S., Dickerson, R. E., *Proc. Natl. Acad. Sci. USA,* **1988,** *85,* 6332

37. Chiu, T. K., Dickkerson, R. E., *J. Mol. Biol.,* **2000,** *301,* 915

38. McCall, M., Brown, T., Kennard, O., *J.Mol. Biol.,* **1985,** *183,* 385

39. Eisenstein, M., Frolow, F., Shakked, Z., Rabinovich, D., *Nucleic Acids Res.,* **1990,** *18,* 3185

40. Fujii, S., Wang, A. H.-J., Quigley, G. J., Westerink, H., van der Marel, G., van Boom, J. H., Rich, A., *Biopolymers,* **1985,** *24,* 243

41. Ban, C., Ramakrishnan, B., Sundaralingam, M., *Biophys. J.,***1996,** *7,* 1215

42. Wang, A. H.-J., Hakoshima, T., van der Marel, G., van Boom, J. H., Rich, A., *Cell,* **1984,** *37,* 321

43. Egli, M., Williams, L. D., Gao, Q., Rich, A., *Biochemistry,* **1991,** *30,* 11388

44. Rahmouni, A. R., Wells, R. D., *J. Mol. Biol.,* **1992,** *223,* 131

45. Coll, M., Aymami, J., van der Marel, G. A., van Boom, J. H., Rich, A., Wang, A. H-J., *Biochemistry,* **1989,** *28,* 310

46. Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., Dickerson, R. E., *Proc. Natl. Acad. Sci. USA,* **1985,** *82,* 1376

47. Denny, W. A.; *Curr. Med. Chem.,***2001**, *8*, 533

48. Coll, M., Frederick, C. A., Wang, A. H-J., Rich, A., *Proc. Natl. Acad. Sci. USA,* **1987,** *84,* 8385

49. Kelland, L., *Nature Reviews,* **2007,** *7,* 573

50. Davies, M. S., Berners-Price, S. J., Hambley, T. W., *J. Inorg. Biochem,* **2000,** *79,* 167
51. Knox, R. J., Friedlos, F., Lydall, D. A., Roberts, J. J., *Cancer Res.,* **1986,** *46,* 1972
52. Yang, F.; Lum, J. B.; McGill, J. R.; Moore, C. M.; Naylor, S. L.; van Bragt P. H.; Baldwin, W. D.; Bowman, B. H.; *Proc. Natl. Acad. Sci. USA,* **1984**, *81*, 2752
53. Kostova, I., *Current Med. Chem.,* **2006,** *13,* 1085
54. Wang, F., Bella, J., Parkinson, J. A., Sadler, P. J., *J. Biol. Inorg. Chem.,* **2005,** *10,* 147
55. O' Donoghue, K.; Carlos Penedo, J.; Kelly, J. M.; Kruger, P. E.; *Dalton Trans.*, **2005**, 1123
56. Lerman, L. S., *J. Mol. Biol.,* **1961,** *3,* 18
57. http://en.wikipedia.org/wiki/File:DNA_intercalation2.jpg
58. Denny, W. A., *Anti-Cancer Drug Design,* **1989,** *4,* 241
59. Feigon, J., Denny, W. A., Leupin, W., Kearns, D. R., *J. Med. Chem.,* **1984,** *23,* 450
60. Denny, W. A., Atwell, G. J., Baguley, B. C., Wakelin, L. P. G., *J. Med. Chem.,* **1985,** *28,* 1568
61. Stewart, L., Redinbo, M. R., Qiu, X., Hola, W. W. J., Champoux, J. J., *Science,* **1998,** *279,* 1534
62. Berger, J. M., Gamblin, S. J., Harrison, S. G., Wang, J. C., *Nature,* **1996,** *379,* 225
63. Berger, J. M., *Biochim. Biophys. Acta,* **1998,** *1400,* 3
64. Martinez, R., Chacon-Garcia, L., *Curr. Med. Chem.,* **2005,** *12,* 127
65. Liu, Y. –X., *Ann. Rev. Biochem.,* **1989,** *58,* 351
66. Butler, C. A.,Cooney, R. P., Denny, W. A., *Appl. Spect.,* **1994,** *48,* 822
67. Adams, A., Guss, J. M., Collyer, C. A., Denny, W. A., Wakelin, L. P. G., *Biochemistry,* **1999,** *38,* 9221
68. Adams, A., *Curr. Med. Chem.,* **2002,** *9,* 1667
69. Bridgewell, D. J. A., Finlay, G. J., Baguley, B. C., *Cancer Chemother.Pharmacol.,* **1999,** *43,* 302
70. Spicer J. A., Gamage, S. A., Atwell, G. J., Finlay, G. J., Baguley, B. C., Denny, W. A., *J. Med. Chem.,* **1997,** *40,* 1919
71. Wilson, W. R., Denny, W. A., Twigden, S. J., Baguley, B. C., Probert, J. C., *Br. J. Cancer,* **1984,** *49,* 215
72. Wilson, W. R., Van Zijl, P., Denny, W. A., *Int. J. Radiat. Oncol. Biol. Phys.*, **1992,** *22*, 693

73. Neidle, S., *Nat. Prod. Rep.*, **2001,** *18,* 291

74. Nikolov, D. B., Chen, H., Haley, E. D., Hoffman, A., Roeder, R. G., Burley, S. K., *Proc. Natl. Acad. Sci. U.S.A.*, **1996,** *93,* 4862

75. Czarny, A., Boykin, D. W., Wood, A. A., Nunn, C. M., Neidle, S., Zhao, M., Wilson, W. D., *J. Am. Chem. Soc.*, **1995,** *117,* 4716

76. http://3.bp.blogspot.com/_DZH2cmCoois/Rp_K_5IavqI/AAAAAAAACII/9GdTDazv-oQ/s400/figure+19-14.jpg

77. Cotter, T. G., Lennon, S. V.,Glynn, J. G., Martin, S. J., *Anticancer Res.*, **1990,** *10,* 1153

78. Marky, L., Breslauer, K. J., *Proc. Natl. Acad. Sci. USA.*, **1987,** *84,* 4359

79. Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., Dickerson, R. E., *Proc. Natl. Acad. Sci. USA.*, **1985,** *82,* 1376

80. Zimmer, C., Wahnert, U., *Prog. Biophys. Mol. Biol.*, **1986,** *47,* 31

81. Pelton, J. G., Wemmer, D. E., *J. Am. Chem. Soc.*, **1990,** *112,* 1393

82. Bostock-Smith, C. E., Laughton, C. A., Searle, M. S., *Biochemical Journal,* **1999,** *342,* 125

83. Lown, J. W., Krowicki, K., Bhat, U.G., Skorobogaty, A., Ward, B., Dabrowiak, J. C., *Biochemistry,* **1986,** *25,* 7408

84. Mrksich, M., Wade, W. S., Dwyer, T. J., Geierstanger, B. H., Wemmer, D. E., Dervan, P. B., *Proc. Natl. Acad. Sci. USA.*, **1992,** *89,* 7586

85. Wemmer, D. E., *Annu. Rev. Biophys. Biomol. Struct.*, **2000,** *29,* 439

86. Pearl, L. H.; Skelly, J. V.; Hudson, B. D.; Neidle, S.; *Nucleic Acids Res.*, **1987**, *15*, 3469

87. Pilch, D. S.; Kirolos, M. A.; Liu, X.; Plum, G. E.; Breslauer, K. J.; *Biochemistry*, **1995**, *34*, 9962

88. Lansiaux, A.; Tanious F.; Mishal, Z.; Dassonneville, L.; Kumar, A.; Stephens, C.; Hu, Q.; Wilson, W. D.; Boykin D. W.; Bailly, C.; *Cancer Research*, **2002**, *62*, 7219

89. Mallena, S.; Lee, M. P. H.; Bailly, C.; Neidle, S.; Kumar, A.; Boykin, D. W.; Wilson W. D.; *J. Am. Chem. Soc.*, **2004**, 126, 13659

90. Liu, Y.; Collar, C.; Kumar, A.; Stephens, C.; Boykin D. W.; Wilson, W. D.; *J. Phys. Chem. B.*, **2008**, *112*, 11809

91. Miao, Y.; Lee, M. P. H.; Parkinson, G. N.; Batista-Parra, A.; Ismail, M. A.; Neidle, S.; Boykin, D. W.; Wilson, W. D.; *Biochemistry*, **2005**, *44*, 14701

92. Nguyen, B.; Lee, M. P. H.; Hamelberg, D.; Joubert, A.; Bailley, C.; Brun, R.; Neidle, S.; Wilson W. D.; *J. Am. Chem. Soc.*, **2002**, *124*, 13680

93. Rahimian, M.; Kumar, A.; Say, M.; Bakunov, S. A.; Boykin, D. W.; Tidwell, R. R.; Wilson, W. D.; *Biochemistry*, **2009**, *48*, 1573

94. Dardonville, C.; Goya, P.; Rozas, I.; Alsasua, A.; Martin, I.; Borrego, J.; *Bioorg. Med. Chem.*, **2000**, *8*, 1567

95. Rodriguez, F.; Rozas, I.; Kaiser, M.; Brun, R.; Nguyen, B.; Wilson, W. D.; Nelson Garcia, R.; Dardonville, C.; *J. Med. Chem.*, **2008**, *51*, 909

96. Glass, La Teca S.; Nguyen, B.; Goodwin, K. D.; Dardonville, C.; Wilson W. D.; Long, E. C.; Georgiadis, M. M.; *Biochemistry*, **2009**, *48*, 5943

97. Eliadis, A., Phillips, D. R., Reiss, J. A., Skorobogaty, A., *J. Chem. Soc. Chem. Commun.*, **1988,** *15,* 1069

98. Bailly, C., Helbecque, N., Henichart, J. P., Colson, P., Houssier, C., Rao, K. E., Shea, R. G., and Lown, J. W., *J. Mol. Recognit.*, **1990,** *3*, 26

99. David-Cordonnier, M-H., Hildebrand, M-P., Baldeyrou, B., Lansiaux, A., Keuser, C., Benzschawel, K., Lemster, T., Pindur, U., *Eur. J. Med. Chem.*, **2007,** *42*, 752

100. Kohn, K.W., Orr, A., O'Connor, P.M., Guziec, L.J., Guziec, F.S., *J. Med. Chem.*, **1994**, *37*, 67.

101. Arcamone, F.M., Animati, F., Barbieri, B., Configliacchi, E., D'Alessio, R., Geroni, C., Giuliani, F.C., Lazzari, E., Menozzi, M., Mongelli, N., Penco, S., Verini, M.A., *J. Med. Chem.*, **1989**, *32*, 774.

102. Marchini, S., Cozzi, P., Beria, I., Geroni, C., Capolongo, L., D'Incalci, M., Broggini, M., *Anti-Cancer Drug Design*, **1998**, *13*, 193.

103. Wyatt, M.D.; Lee, M.; Garbiras, B.J.; Souhami, R.L.; van der Hartley, J.A. *Biochemistry*, **1995**, *34*, 13034.

104. Xie, G.; Gupta, R.; Lown, J.W. *Anti-Cancer DrugDesign*, **1995**, *10*, 389.

105. Circular and Linear Dichroism, Alison Roger and Bengt Norden, *Oxford University Press,* 2008

106. Nguyen, B.; Tanious, F. A.; Wilson, W. D.; *Methods*, **2007**, *42*, 150

107. Holdgate, G. A.; *Biotechniques*, **2001**, *31*, 164

Chapter 2

Objectives

2.1. Objectives

The main goals pursued in this research can be summarised as follows:

- Firstly, we will prepare three families of asymmetrical dicationic minor groove binders with structural characteristics related to furamidine (Figure 2.1). The rationale behind this synthesis is to optimise the essential features of a molecule that binds to the minor groove. Thus, we will explore the importance of coulombic interactions between the cationic functional groups and the minor groove of DNA.

Moreover, we will evaluate the significance of compound geometry by preparing molecules with different geometry in the central linkers. This is fundamental for the formation of optimal hydrogen bonds and van der Waals contacts between the guanidine or 2-aminoimidazoline functionalities and the DNA minor groove because the central functional group can influence the dihedral angle between the two aromatic moieties. This is highly relevant considering that most minor groove binders that contain a diaromatic scaffold have a dihedral angle close to zero inside the minor groove facilitating the fit into the narrow minor groove.

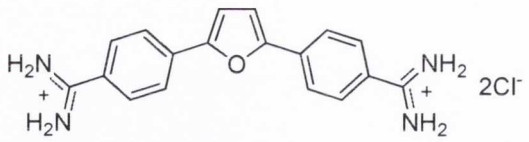


Figure 2.1. Furamidine

Comparisons will be made between the binding affinities of the *mono*-functionalised molecules and the di-functionalised ones to investigate the importance of the extra hydrogen bonding ability of the second guanidinium or 2-aminoimidazolinium groups (Figure 2.2). The consequences of lengthening the central linker will also be investigated in order to observe whether the distance between the cationic ends influences the binding.

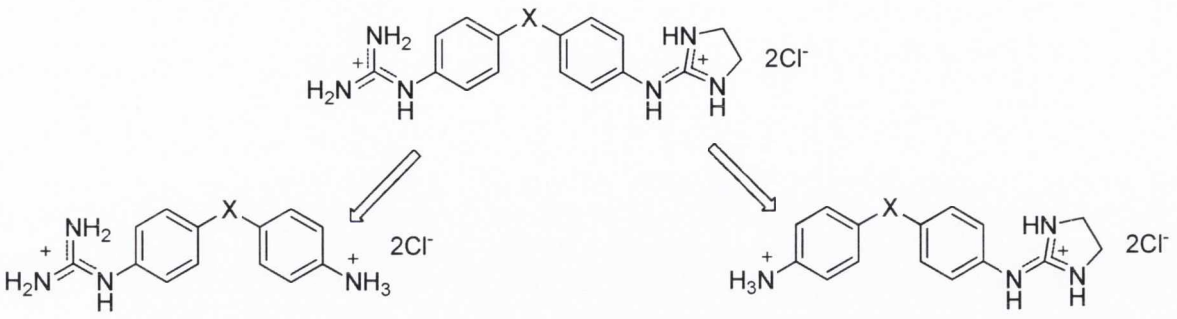


Figure 2.2. Molecules synthesised in this research project

• The second target will consist of carrying out physicochemical measurements to assess the DNA binding affinity of the molecules prepared. Therefore, we will examine whether or not these molecules display any affinity for DNA. Thermal denaturation assays will be carried out with wild type DNA to measure their binding affinity for mixed sequences. Then, we will investigate the sequence selectivity of these molecules. This will be evaluated by repeating the thermal denaturation experiments with poly(dA-dT)$_2$. The reason for this second set of experiments is that the minor groove is rich in AT sequences and thus, binding to these sequences would indicate preference for the DNA minor groove.

Thermal denaturation experiments provide limited information on binding to DNA. Further experiments are required to deepen the knowledge of the nature of the binding of the compounds here proposed. Initially, UV titrations will be performed to determine the equilibrium binding constants of our molecules. From this, the free energy of binding will be evaluated. This allows to numerically compare the binding strength of the molecules. When the absorbance of the molecule is close to the DNA absorbance region, circular dichroism experiments will be carried out to determine these binding constants and consequently the free energy of binding can then be calculated from them.

Once the free energies are evaluated, the binding enthalpy will be measured by using Isothermal Titration Calorimetry. We will obtain a full thermodynamic profile of the DNA interactions of all our molecules by comparing the enthalpy and free energy of binding. This could present us with valuable information on the binding, for example, if the enthalpy of

binding is much higher for one molecule in comparison to another, one could conclude that there is an extra hydrogen bond between the molecule and the DNA base pairs, for example.

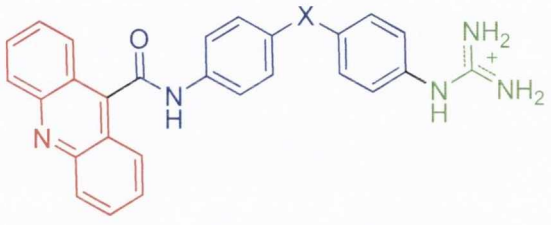
Surface Plasmon Resonance experiments will be then performed on the molecules. This technique that is carried out by Prof. David Wilson from Georgia State University (USA) within an ongoing collaboration, gives very accurate information on the binding affinity to different DNA hairpins. These experiments will allow us to obtain the order of binding preference of our molecules to different runs of bases.

The final physicochemical measurement that will be carried on our molecules is linear dichroism. This technique provides information on the binding mode of molecules to DNA. For example, planar molecules or those with a small dihedral angle between the aromatic moieties could be expected to be intercalators, whereas, molecules that have a large dihedral angle between the aromatic rings can be considered as potential minor groove binders.

- Thirdly, we will assess the cytotoxicity of the molecules, by evaluating their activity against breast cancer cell lines. This study is performed by Dr. Geoff Margisson from the Paterson Institute (UK) within an ongoing collaboration. These biochemical assays will be carried out not only to assess the cytotoxicity of our compounds, but also to compare the DNA binding affinity with the cytotoxicity results and see whether a correlation can be found between both sets of data. If such a correlation is found, we could conclude that the cytotoxicity was related to the DNA binding affinity.

- Fourthly, we will synthesise dual action molecules using acridine to replace one of the cationic groups. In this family, the acridine moiety will be attached to those diaromatic molecules that had displayed the largest binding affinity. It is expected that the acridine chromophore would intercalate between the base pairs, and that the diaromatic guanidine functionality would fit in the minor groove. Thus, we will utilise different linkers between the aromatic moieties of the minor groove binder scaffold to investigate if by increasing the flexibility of the linker, there would be a large entropic cost and this would influence the binding. As well, we will incorporate a hydrogen bond donor such as the NH group between the aromatic rings of the minor groove binder part. This will allow us to evaluate the effect of

increasing the number of hydrogen bond donors on the molecule. In summary, we intend to investigate the importance of hydrogen bond donors and acceptors in these molecules, the effect of increasing the distance between the cationic moiety and the acridine moiety and also to investigate the effect of the geometry of the minor groove binder part of the dual action molecule.



Red = Acridine (intercalator)
Blue = the spacer linking the intercalator and the guanidine group
Green = the cationic guanidine group

Figure 2.3. Diagram of the dual action molecules

- The final aim of this research is to perform physicochemical measurements to assess the DNA binding affinity of the dual action molecules. Initially, thermal denaturation experiments will be conducted with wild type DNA to assess the binding affinity of the molecules relative to each other. If good binding is observed, these experiments will then be repeated in the presence of poly(dA-dT)$_2$ to investigate the sequence selectivity of the molecules. Considering that these tests provide good DNA binding results, UV experiments will be used to evaluate the equilibrium binding constant and linear dichroism to determine the mode of binding.

In summary, we will prepare four families of compounds with different functionalities. Their binding affinity will be then characterised and assessed carefully using physicochemical measurements and finally their cytotoxicity will be evaluated.

Chapter 3

Synthesis of the Asymmetric Derivatives

3.1 Introduction

The preparation of crescent shaped molecules as potential minor groove binders has been actively explored by Boykin *et al.*[1] His research group has reported the synthesis of many of these compounds that have been found to latch onto the convex shape of the DNA minor groove,[2] such as furamidine and its analogues.[3]

Furamidine[4] and furimidazoline (Figure 3.1) both exhibit all the requirements of a minor groove binder:

- crescent shape to complement the convex shape of the groove,
- cationic functionalities for Coulombic interactions with the negative potential of the minor groove,
- hydrogen bond donors for bonding with the N3 of adenine and the O2 of thymine.[5]

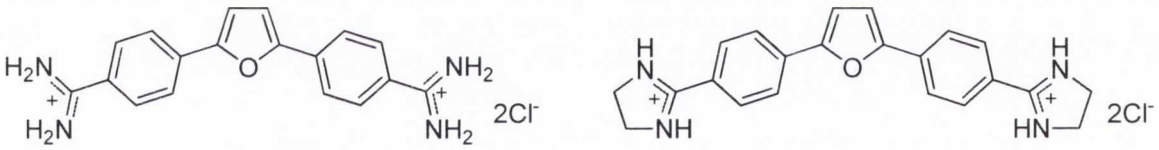


Figure 3.1. Structure of furamidine (left) and furimidazoline (right)

Boykin[6,7] and Wilson have shown that these molecules and their analogues bind to the DNA minor groove with increased AT sequence selectivity due to their induced fit into the narrow groove. On the other hand, Wilson[8] has shown that the previously seen crescent shape of 'classical' minor groove binders was not a necessity for binding because some linear molecules bind to the DNA minor groove with the help of a water molecule. This acts as a bridge for hydrogen bonding between the molecule and base pairs.[9]

Molecules previously prepared in Rozas' group,[10] have shown similar characteristics as furamidine such as crescent shape, cationic ends and hydrogen bond donor groups (Figure 3.2).

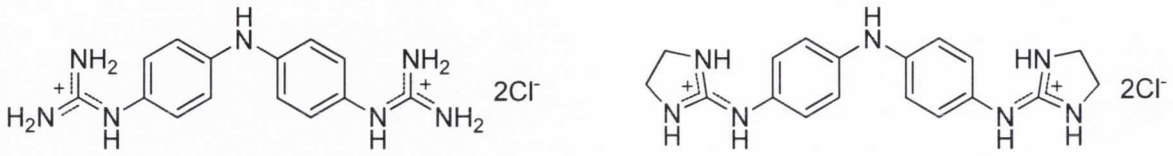
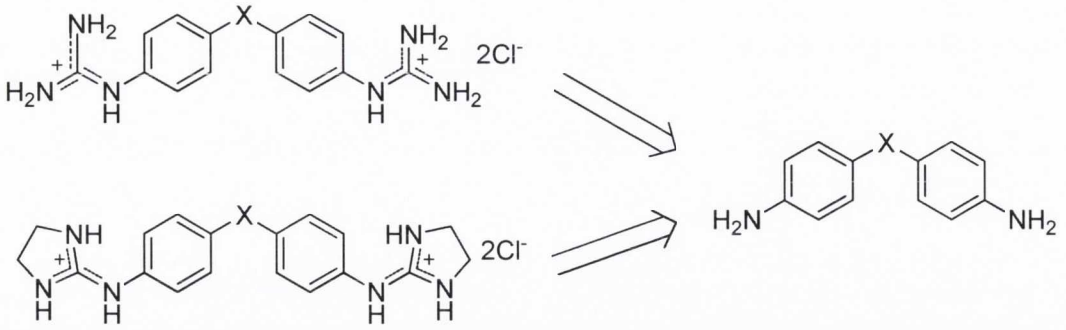


Figure 3.2. Examples of the compounds previously prepared in the Rozas' group

By examining these molecules we observe some key differences. The amidine functional group is replaced by the guanidine group, and thus these molecules will probably display higher DNA affinity due to the extra NH hydrogen bond donor. The guanidine functional group is also longer than the corresponding amidine one, consequently this could be beneficial by providing an optimum distance for hydrogen bonding. Another difference is in the linker between the phenyl rings. The furamidine-like binders show a furan group as a linker whereas in Rozas series, simpler groups were present such as -NH-, -CO- or -CH_2-. This provides the molecules with much greater flexibility and adaptability to bind to the convex minor groove. Since our cationic functional groups were larger than the corresponding amidine group of furamidine, it was decided to investigate the optimum linker length between the two aromatic scaffolds. The importance of the central linker was seen as crucial in DNA binding since the linkers would influence the overall geometry of the molecule as well as the amount of hydrogen bond donors and acceptors in the molecule. The interaction of these molecules with DNA was evaluated[11] thus confirming their suitability as potential minor groove binders.

The synthesis of these symmetric derivatives consisted on the introduction of the guanidine or 2-aminoimidazoline groups to di-aromatic di-amines (Scheme 3.1).



Scheme 3.1.

Considering that compounds already prepared by Rozas' group were all symmetric *bis*-guanidinium or *bis*-2-aminoimidazolinium a question arose if both and similar cationic groups would be required for the binding to the minor groove. Thus, the preparation of *mono*-functionalised and asymmetric di-functionalised molecules should be considered. Preparation of *mono*-guanidinium, *mono*-2-aminoimidazolinium and the corresponding asymmetric guanidinium/2-aminoimidazolinium derivatives will enable improved understanding of the interactions of these molecules with DNA.

Therefore, three families of compounds were prepared (Figure 3.3):

- Family I- *mono*-guanidine molecules.
- Family II - *mono*-2-aminoimidazolines.
- Family III - asymmetric di-functionalised guanidine/2-aminoimidazoline derivatives.

Bearing in mind the preparation of the asymmetric derivatives (family III), different strategies could be considered as for the order of introduction of the different cationic groups on the di-aromatic scaffold. For that reason, the preparation of the corresponding *mono*-guanidinium and *mono*-2-aminoimidazolinium di-aromatic derivatives was first approached.

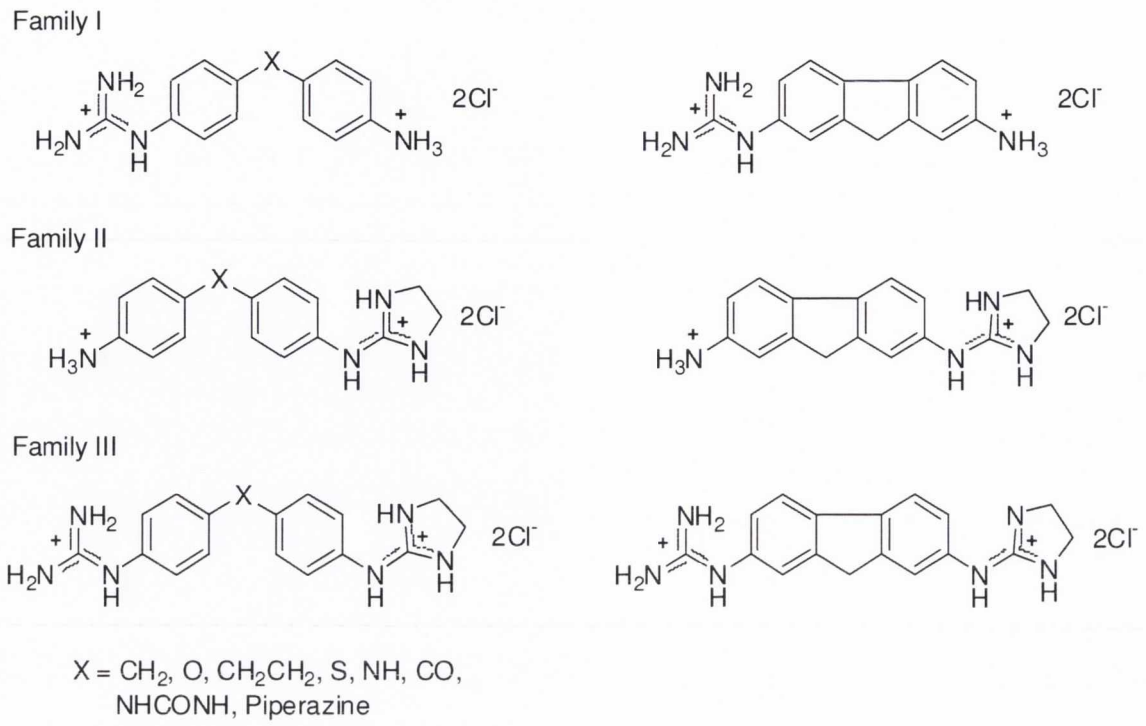


Figure 3.3. General scheme of Families I, II and III (mono-guanidine, mono-2-aminoimidazoline and guanidine/2-aminoimidazolidine derivatives respectively)

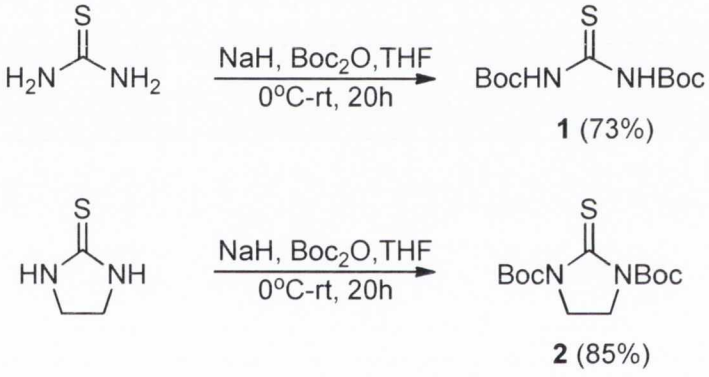
3.2. Synthesis of the "amidylating" agents

Taking into account that the synthesis of all the proposed molecules involves the introduction of the guanidine or the 2-aminoimidazoline ("amidine-like") moieties, the first step was the preparation of such "amidylating" agents.[10]

A good approach to create these amidine-like groups is to use thiourea or 2-thio imidazolidinone as amidylating agents to react with aromatic amines by nucleophilic attack. However, this type of amines are poor nucleophiles and, therefore, in order for the reaction to occur, the electrophilicity of the carbon atom on the C=S atom should be increased. Kim and Quian[12] suggested that the introduction of Boc groups on the amino groups of thiourea increases the electrophilicity of the carbon atom, facilitating the guanidylation reaction of aromatic amines. Dardonville and Rozas[10] extended that idea to 2-thioimidazolidinone and, hence, Boc protection of the amino groups allowed the creation of the 2-aminoimidazoline functionality from an aromatic amine.

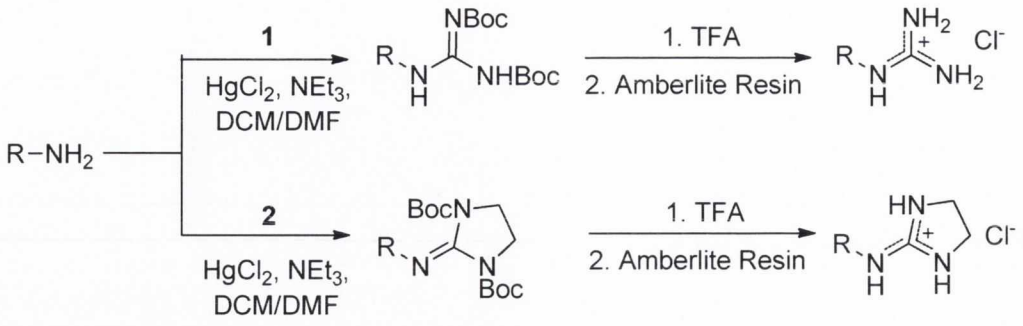
The preparation of the guanidylating and 2-aminoimidazolidylating reactants was carried out by reacting thiourea and imidazolidine-2-thione respectively with Boc_2O in the presence of sodium hydride in anhydrous THF as shown in Scheme 3.2.[10] First, the thiourea or the imidazolidine-2-thione is dissolved in anhydrous THF and stirred at 0 °C and then NaH was added slowly. The reaction was brought to room temperature and stirred for a further ten minutes and was brought back down to 0 °C when Boc_2O was added. It was seen that both solutions turned to a pale yellow colour after one hour.

By comparing the yields in the preparation of these molecules, we can observe that the Boc protected imidazolidine always gave a higher yield in comparison to the corresponding Boc protected thiourea molecule.



Scheme 3.2.

The generic procedure for the preparation of the guanidinium or 2-aminoimidazolinium derivatives is shown in Scheme 3.3. Firstly, the aromatic amine is reacted with the Boc protected thiourea or thioimidazolidinone in the presence of mercury (II) chloride, and triethylamine in the appropriate solvent. The purified Boc protected guanidylated/2-aminoimidazolinium products were then deprotected in a mixture of TFA and DCM. Conversion to the hydrochloride salt was carried out by treating the TFA salt with amberlite resin in the chloride form.

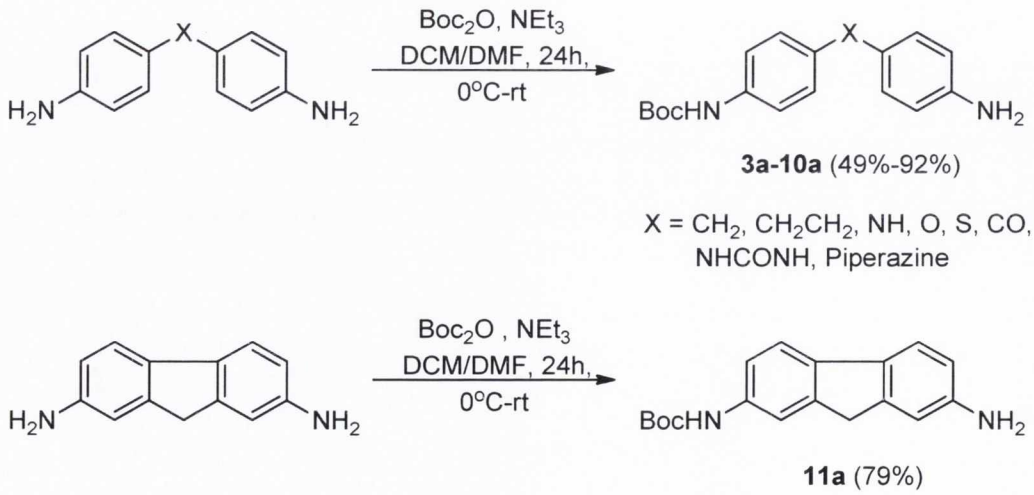


Scheme 3.3.

3.3. Synthesis of the *mono*-Boc protected derivatives

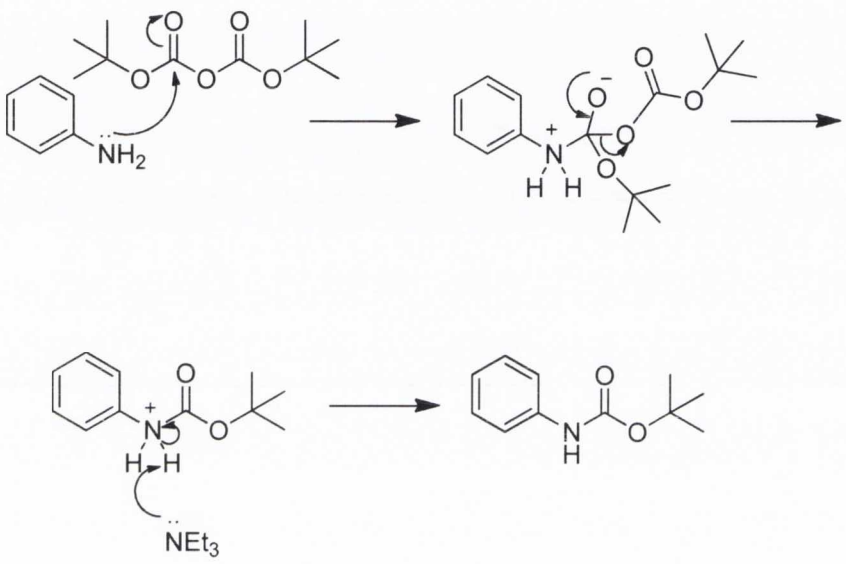
Considering the preparation of the *mono*-guanidinium or the *mono*-2-amino-imidazolinium molecules, we firstly need to protect one of the amino groups of the starting diaromatic diamines to avoid selectivity problems. The Boc protecting group was chosen for several reasons: this group is stable in basic conditions, there was previous experience working with this group in our laboratory and finally, the molecule can be readily deprotected all at once using trifluoroacetic acid.

For the preparation of this set of molecules, Boc_2O was added to a solution of the starting diamine with NEt_3 in the DCM or DMF depending on the polarity of the starting material (Scheme 3.4). Three equivalents of the starting di-amine were reacted with one equivalent of Boc_2O in order to avoid Boc protection occurring on both amino groups. In that way, less than 10% of di-substituted side product was observed.



Scheme 3.4.

The mechanism for Boc protection of an aromatic amine is shown in Scheme 3.5. The nucleophilic amino group attacks the highly electrophilic carbon of Boc_2O causing the movement of electrons towards the oxygen atom. The electrons on the oxygen atom reform the carbonyl group resulting in the loss of the *tert*-butyl carbonate group. Deprotonation with NEt_3 resulted in the required Boc protected molecules.



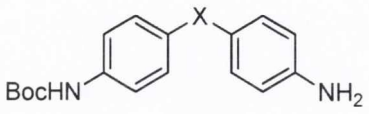
Scheme 3.5.

All the reactions were monitored by thin layered chromatography (TLC). On completion of the reaction as judged by TLC, the reaction solution was washed with water, dried over Na_2SO_4 and concentrated under vacuum. Following purification by silica gel chromatography, the required compounds were obtained in moderate to good yields. In Table 3.1 the yields obtained for all the *mono*-Boc protection reactions performed are presented.

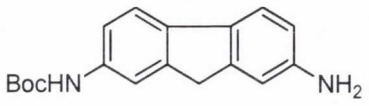
All of the resulting compounds (**3a**-**11a**) were characterised by means of NMR (^1H and ^{13}C), IR and melting point. From the ^1H MNR, we could see a peak in the aliphatic region that is indicative of the nine protons from the Boc group. For example, the NMR of **3a** shows a peak at 1.51 ppm that is associated to Boc protons. Evidence of the free amino group can be observed through the presence of a broad peak at 3.60 ppm. The asymmetry of the

compound can be clearly seen through the four individual aromatic doublets, instead of two aromatic peaks expected for a symmetric compound. From the IR spectrum, peaks at 1691 cm^{-1} indicate the CO stretch from the Boc group.

Table 3.1.- Yields obtained for the mono Boc protection of the diamines



3a-10a



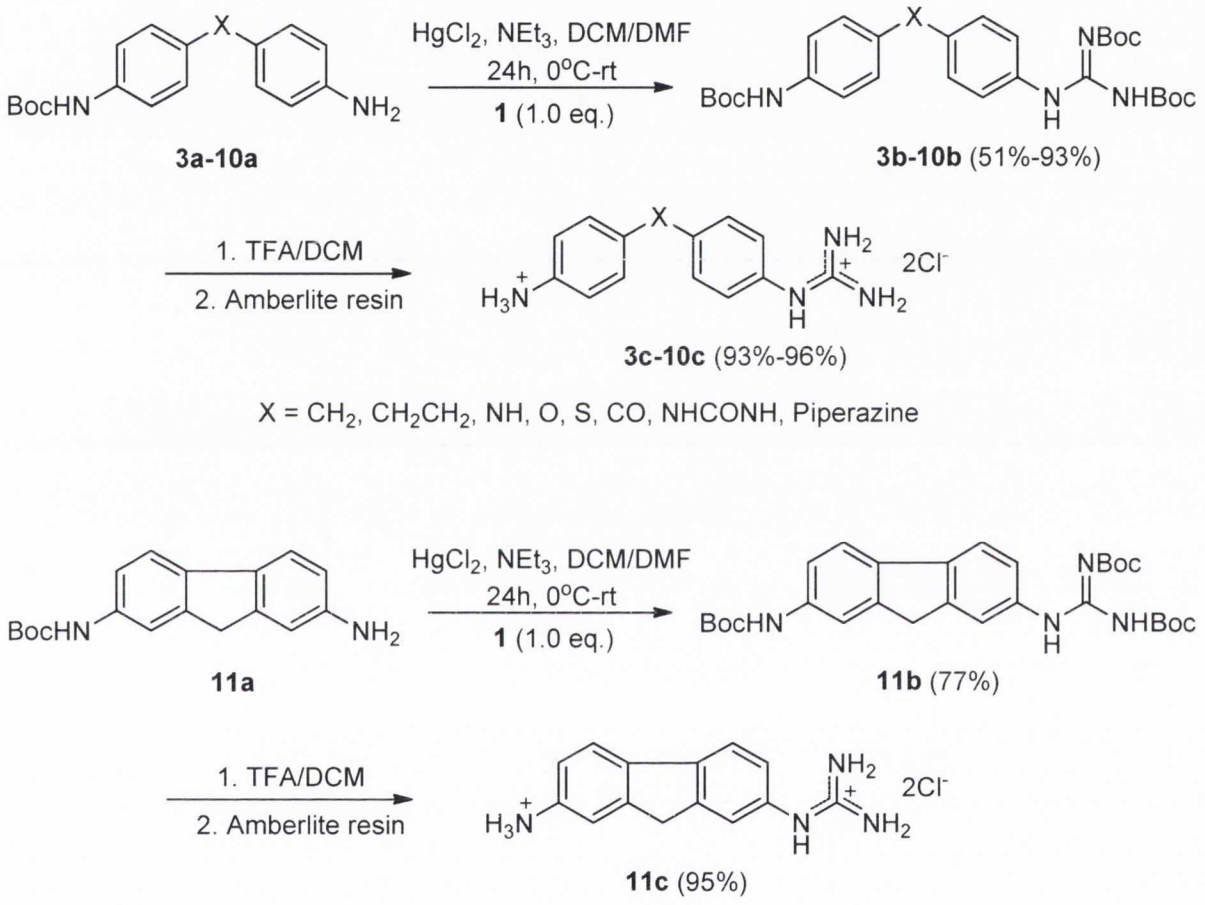
11a

Compound Number	X	% Yield
3a	CH_2	82
4a	CH_2CH_2	74
5a	NH	70
6a	O	76
7a	S	77
8a	CO	46
9a	HNCONH	92
10a	Piperazine	49
11a	-	79

3.4. Preparation of family I: The *mono*-guanidine derivatives

Taking into account the preparation of the *mono*-guanidinium derivatives, the *mono*-Boc protected starting material (**3a-11a**) was reacted with the guanidylating agent **1** in a 1:1 ratio, using mercury (II) chloride (Scheme 3.6). The reaction was stirred for one hour at 0 °C followed by 23 hours at room temperature. When the reaction reached completion, as judged

by TLC, it was diluted with EtOAc, which was followed by the removal of the mercury (II) sulfide by filtering the reaction mixture through a pad of celite. Then, the crude product was washed with water and brine, dried over anhydrous sodium sulphate and concentrated under vacuum to give an oil that was purified by silica gel chromatography, with yields ranging between 51%-93% as shown in Scheme 3.6 and in Table 3.2.

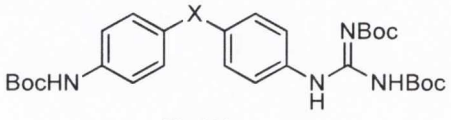


Scheme 3.6.

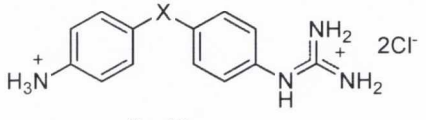
The Boc protected compounds were fully characterised by means of NMR (^{1}H and ^{13}C) and IR. The ^{1}H NMR of these compounds all contained an additional two peaks in the aliphatic region of the spectrum corresponding to the Boc protected guanidine moiety. The disappearance of the peak for the NH_2 group is also indicative of the formation of **3b**. By analysing the IR spectrum, we observe stretches at 3290, 3185 and 3150 cm^{-1} that correspond

to NH stretches. Other characteristic peaks can be seen at 1719, 1626 and 1603 cm^{-1} that correspond to the C-N, C=N and C=O stretchings respectively. By analysing each of the spectra for the compounds in this family, we could conclude that they were successfully prepared.

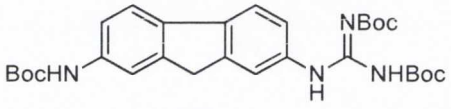
Table 3.2.- Yields for the mono-guanidylation, deprotection and overall yields for the formation of the mono-guanidine family



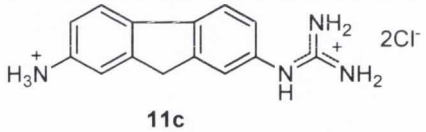
3b-10b



3c-10c



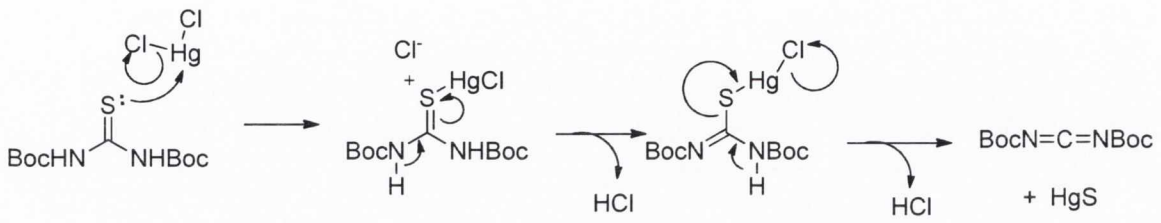
11b



11c

Compound Number	X=	% Yield 3b-11b	% Yield 3c-11c	% Yield Total
3	CH_2	77	94	72
4	CH_2CH_2	51	94	48
5	NH	77	95	73
6	O	53	94	50
7	S	53	96	51
8	CO	-	-	-
9	HNCONH	93	95	88
10	Piperazine	76	93	71
11		77	95	73

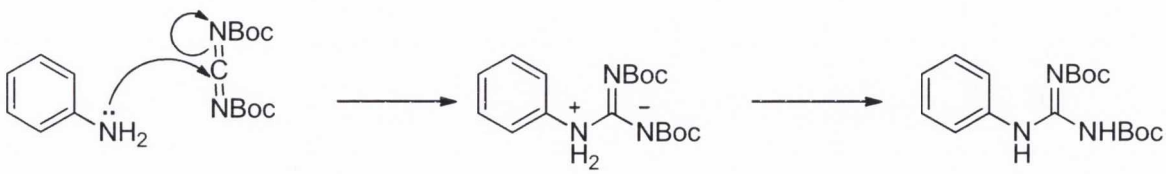
The proposed mechanism for the functionalisation of the amino group with the Boc protected thiourea in the presence of mercury chloride, which involves the generation of the highly reactive carbodiimide intermediate was investigated by Kim and Qian[12] and is shown in Scheme 3.7.



Scheme 3.7.

Initially, the sulfur of the *N,N'*(*tert*-butoxycarbonyl)thiourea forms a strong bond with the thiophilic Hg(II), resulting in the loss of one of the chlorine atoms. This increases the electrophilicity of the carbon atom. A pair of electrons from the C=S double bond neutralise the positive charge on the sulphur atom, while simultaneously the pair of electrons on the N-H bond forms a C=N double bond. This is accompanied by the loss of HCl that is neutralised by NEt_3. Then, the pair of electrons from the N-H bond forms a C=N double bond. Simultaneously, the C-S bond breaks and the pair of electrons forms this bond form a double bond between the mercury and the sulphur. This is accompanied by the loss of HCl that is neutralised by NEt_3.

After the formation of the carbodiimide intermediate, the nucleophilic aromatic NH_2 attacks the highly unstable intermediate, which results in the migration of electrons to the nitrogen atom. Following proton transfer allows formation of the Boc protected guanidine group (Scheme 3.8).



Scheme 3.8.

The Boc protected derivatives (**3b-11b**) were deprotected by treating them in a mixture of trifluoroactic acid and DCM for four hours. The solvent was removed under reduced pressure and the corresponding trifluoroactate salts were treated with basic amberlite resin exchange in Millipore water overnight. The resin was filtered off and the water layer was washed with DCM. The water was then evaporated under vacuum to yields the required di-hydrochloride salts (**3c-11c**).

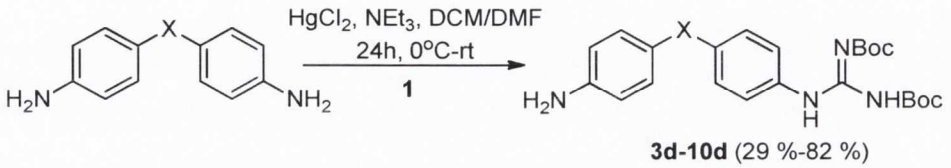
All final compounds (**3c-11c**) were characterised by ^{1}H NMR, ^{13}C NMR, mass spectroscopy, melting point and elemental analysis. By analysing the ^{1}H NMR, we observe the absence of the Boc hydrogen signals that indicates the full deprotection of these compounds. The absence of trifluoroacetate was checked by ^{19}F NMR. The purity of this compound was proved by elemental analysis.

From Table 3.2, it can be seen that the molecule with the carbonyl linker was not obtained. Its synthesis was attempted by reacting **8a** with **1** in the presence of mercury (II) chloride and NEt_3 in DMF at 0 °C for one hour and a further 23 hours at room temperature. No reaction occurred as observed by TLC. The reaction was then heated to 30 °C and was left for a further 24 hours; however, it did not proceed. This was probably due to the fact that the carbonyl group in the *para* position withdraws electron density from the NH_2 group by resonance (Scheme 3.9). Consequently, the amino end is less nucleophilic and less prone to attack the electrophilic carbon of the carbodiimide intermediate.

Scheme 3.9.

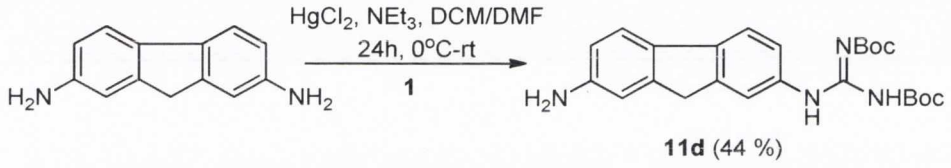
An alternative strategy for the preparation of the *mono*-guanidine family was explored by reacting the starting diamine directly with the Boc protected guanidylating agent **1**

(Scheme 3.10). An excess of the diamine was reacted with **1** in a 3:1 ratio, in the presence of mercury (II) chloride, NEt_3 and the appropriate solvent. An excess of diamine was used in order to minimise the formation of the di-functionalised derivatives. The compounds obtained (**3d-11d**) were purified by silica gel chromatography giving moderate to high yields (see Scheme 3.10).



3d-10d (29 %-82 %)

X = CH_2, CH_2CH_2, NH, O, S, CO, NHCONH, Piperazine



11d (44 %)

Scheme 3.10.

Following this guanidylation reaction, compounds **3d**, **5d**, **9d**, **10d** and **11d** were obtained in high yields. However, compounds **4d**, **6d** and **7d** were obtained in yields lower than 26 %, probably due to partial decomposition in the acidic environment of the silica gel.

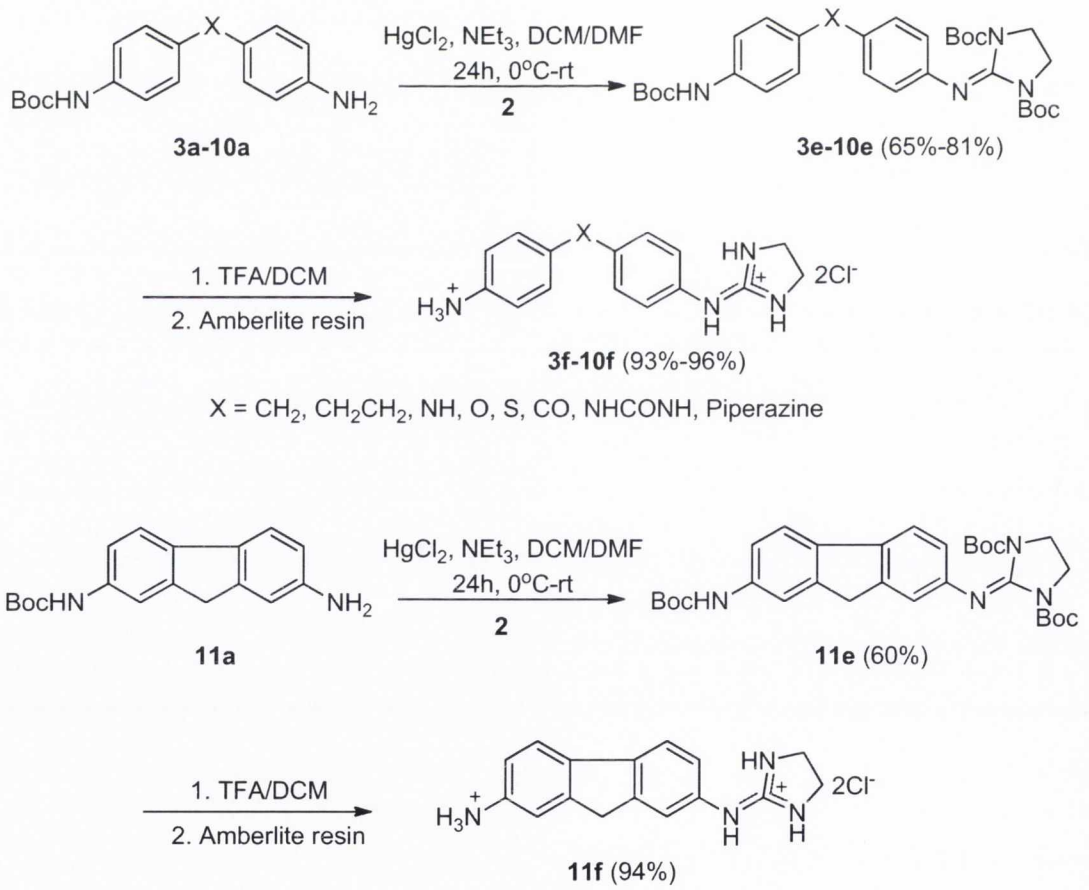
All these *mono*-guanidylated compounds were characterised by ^{1}H NMR, ^{13}C NMR, IR, melting point and mass spectroscopy. For example, the ^{1}H NMR of compound **3d** (X = CH_2) showed peaks at 1.50 and 1.53 ppm that correspond to the Boc protons of the guanidine group. We also observe two extra peaks at 10.26 and 11.66 ppm indicative of the NH groups of the guanidine group. By analysing the IR spectrum, we observed peaks at 1717, 1633 and 1608 cm^{-1} that correspond to different CO and CN stretches. All this spectral analysis was performed for each compound indicating the successful preparation of this family of derivatives.

3.5. Preparation of Family II: The *mono*-2-aminoimidazoline derivatives

The preparation of the *mono*-2-aminoimidazolines was carried out under the same conditions as the *mono*-guanidines except by replacing the Boc protected guanidylating agent with the Boc protected imidazolidylating agent **2** (Scheme 3.11). Therefore, by reacting the

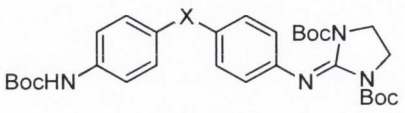
starting *mono*-Boc protected diamine with **2** in a 1:1 ratio in the presence of mercury (II) chloride and NEt_3 in DCM/DMF yielded the corresponding molecule that was purified by alumina chromatography. These compounds were unable to be purified by silica gel chromatography was not possible due to decomposition as seen previously in the Rozas' laboratory. Hence, alumina column chromatography under pressure was used to avoid decomposition with yields ranging between 65%-81%.

The Boc protected derivatives (**3e-11e**) were characterised by means of NMR (^1H and ^{13}C) and IR. For example in the case of the spectrum for **3e**, we observed a peak at 1.26 ppm that is indicative of the Boc protons of the imidazolidine functionality. From the IR spectrum, peaks were observed at 1754 cm^{-1} and 1694 cm^{-1} that is indicative of the CO and CN stretching frequencies.

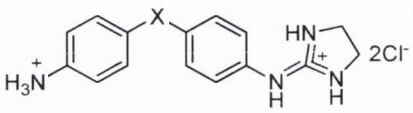


Scheme 3.11.

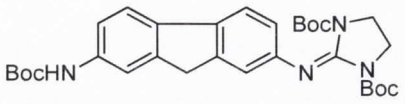
Table 3.3: Yields for the (a) mono-imidazolydation, (b) deprotection and (c) overall yields for the formation of the mono 2-aminoimidazoline family



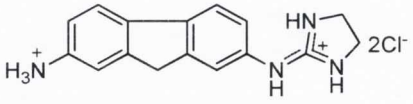
3e-10e



3f-10f



11e



11f

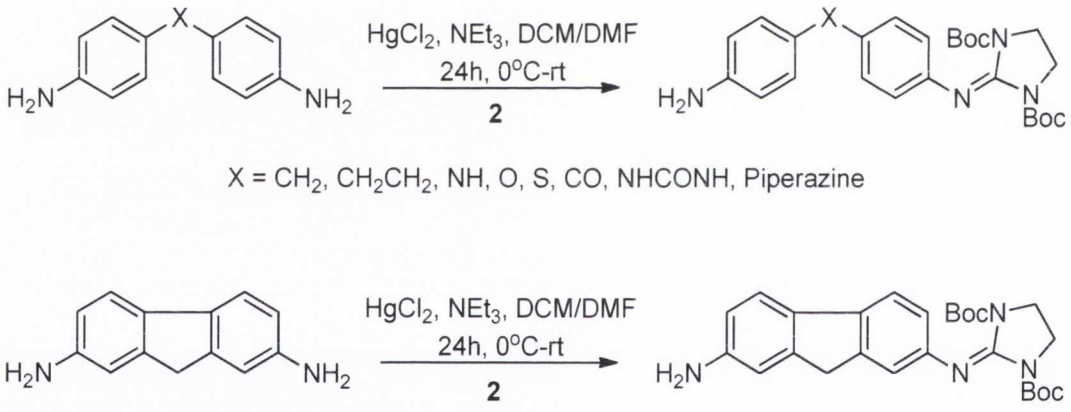
Compound Number	X=	% Yield 3e-11e	% Yield 3f-11f	% Yield Total
3	CH_2	81	93	75
4	CH_2CH_2	72	94	68
5	NH	80	94	75
6	O	65	96	62
7	S	72	95	68
8	CO	-	-	-
9	HNCONH	67	93	62
10	Piperazine	77	94	72
11		60	94	56

Finally, these Boc protected derivatives were reacted in a mixture of trifluoroacetic acid and DCM in order to remove the Boc groups. The trifluoroacetate salts were then treated with basic amberlite resin to give the di-hydrochloride salts (**3f-11f**) which were all characterised by ^{1}H NMR, ^{13}C NMR, mass spectroscopy, melting point and elemental

analysis. In the ^{1}H NMR spectra, we did not observe any peaks corresponding to the Boc protons, indicating that the molecules were fully deprotected. The absence of trifluoroacetate was checked by ^{19}F NMR.

Apart from the molecule **9e** (X = NHCONH), the general trend was that the *mono*-imidazolidines produced higher or equal yields than the corresponding *mono*-guanidines with yields varying from 60%-81% (Table 3.3). Again, synthesis of the corresponding derivatives **8e** and **8f** were not possible to prepare due to the deactivating effect of the CO group.

An alternative strategy was attempted by reacting the starting diamine with **2** in a 3:1 ratio in the presence of mercury (II) chloride and NEt_3 in DCM (Scheme 3.12). The starting diamine was used in such an excess in order to avoid di-functionalisation. However, a mixture of the starting diamine, the *mono*-functionalised derivative and the di-functionalised derivative was obtained.

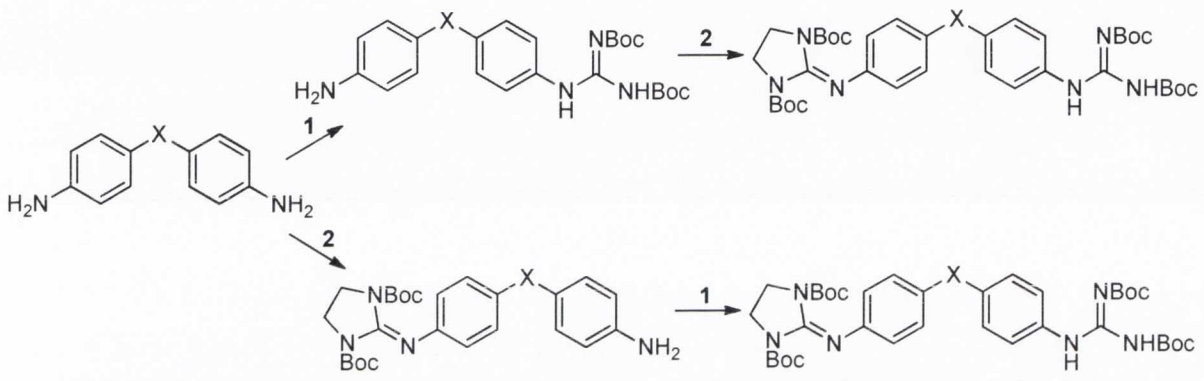


Scheme 3.12.

The required *mono*-functionalised compound proved too difficult to purify for two reasons. Firstly, using different eluents and evaluating by TLC starting diamine and the di-functionalised showed very close Rf values. Secondly, the 2-aminoimidazoline functionalised compounds needed to be purified by alumina gel chromatography under pressure and not by gravity proving difficult to isolate the required compound. Therefore, the only synthetic route that was utilised was the imidazolidylation using compound **2** with the *mono*-Boc protected derivatives (**3a**-**11a**) as shown in Scheme 3.11.

3.6. Preparation of Family III: The guanidine/2-aminoimidazoline derivatives

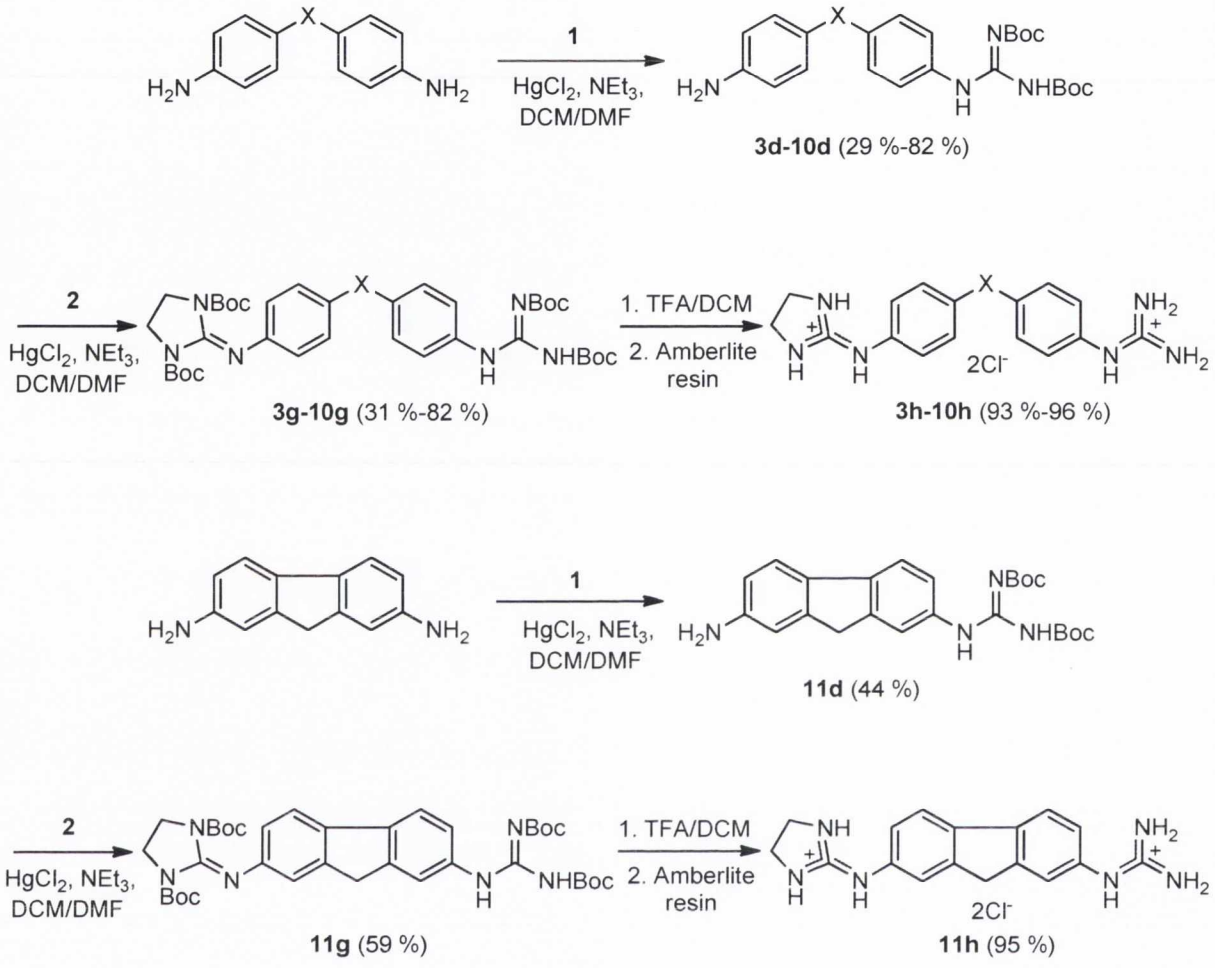
Two synthetic strategies could be considered for the preparation of the molecules in Family III. Firstly, we could introduce the guanidine functional group followed by the 2-aminoimidazoline one. Secondly, we could introduce the 2-aminoimidazoline functional group followed by the guanidine group as shown in Scheme 3.13.



Scheme 3.13.

Attempts were made with the two different synthetic methodologies for the preparation of this third family of molecules. Following the first methodology proved to be the most feasible approach due to the fact that we encountered problems in the purification of the *mono*-imidazolidine compounds as described in the previous section. Thus, it was decided to introduce the guanidine functional group first, followed by the 2-aminoimidazoline functional group according to the Scheme 3.14 below.[12]

Initially, the starting diamine was reacted with **1** in a 3:1 ratio in the presence of mercury (II) chloride and NEt_3 in DCM/DMF for one hour at 0 °C and for a further 23 hours at room temperature. The reaction was monitored by TLC and on completion, it was diluted with EtOAc, washed with distilled water and brine, dried over anhydrous Na_2SO_4 and concentrated under vacuum to give a crude product that was purified by silica gel column chromatography yielded **3d-11d** (Table 3.4). The corresponding compounds were characterised by NMR (^1H and ^{13}C), IR, melting point and mass spectroscopy.



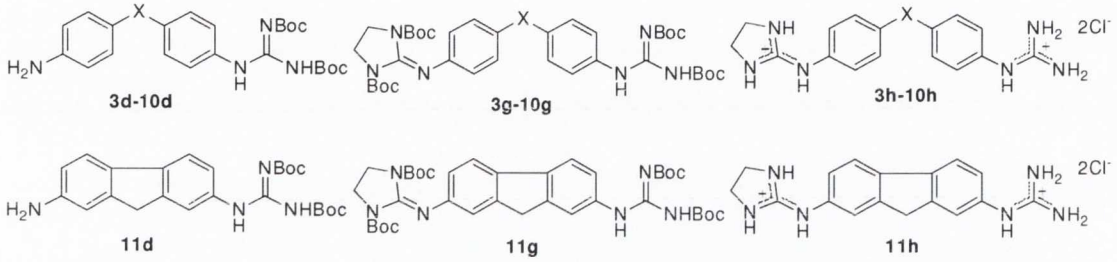
Scheme 3.14.

After careful analysis of these compounds (**3d-11d**), they were reacted with the imidazolidylating agent **2** in a 1:1 ratio under the same conditions as the guanidylation step. The reactions were worked up as before and thus, the crude product was purified using alumina column chromatography under pressure to yield compounds **3g-11g** (Table 3.4) which were characterised by ^{1}H and ^{13}C NMR, IR, melting point and mass spectroscopy.

By analysing the ^{1}H NMR, the disappearance of the free amino peak followed by the appearance of further peaks in the aliphatic region that corresponded to the Boc protons of the Boc protected imidazolidine functionality provided evidence that this product was obtained. In addition, a peak at 3.78 ppm, that is indicative for the protons on the 2-

aminoimidazoline group, was observed. Interestingly, from the IR spectrum of **3g** and compared to that of the starting material **3d** the disappearance of the -NH_2 peaks was observed. The additional two peaks at 1756 and 1668 cm^{-1} corresponds to the extra CO groups of the Boc protected 2-aminoimidazoline functionality. The same spectral analysis was carried out for compounds **4g-11g,** indicating that all the compounds in this family were successfully prepared.

Table 3.4.- Yields for the mono-guanidine, guanidine/2-aminoimidazoline, deprotection and overall yields for the formation of the asymmetric aromatic dicationic family



Compound Number	X=	% Yield 3d-11d	% Yield 3g-11g	% Yield 3h-11h	% Yield Total
3	CH_2	29	67	96	19
4	CH_2CH_2	54	62	94	31
5	NH	63	45	94	27
6	O	82	51	94	39
7	S	39	43	95	16
8	CO	43	31	94	12
9	HNCONH	53	65	94	32
10	Piperazine	65	82	93	50
11		44	59	95	25

As before, all Boc protected compounds **3g-11g** were reacted in a mixture of trifluoroacetic acid and DCM in order to remove the Boc groups. The trifluoroacetate salts were then treated with basic amberlite resin to give the di-hydrochloride salts (**3h-11h**, see Table 3.4). All hydrochloride salts were characterised by ^{1}H and ^{13}C NMR, IR, melting point, mass spectroscopy and elemental analysis.

Comparing the yields obtained for the preparation of the *mono*-guanidine precursors (Table 3.4), we observed that the highest yield that was obtained was for compound **6d** (X= O) followed by **10d** (X= piperazine) and **5d** (X= NH). This could be due to the electronic nature of the linkers since they are electron donating and would thus increase the nucleophilicity of the *para* amino group. However, compound **7d** displayed a moderate yield that could be attributed to the coordination of the sulphur linker with the mercury catalyst.

Looking at the yields for the *mono*-imidazolidylation reactions (Table 3.4), we observed good overall yields; for example, **3g** and **4g** were obtained with yields of 67% and 62%. The poorest yield was obtained for compound **8g** (X= CO) and this could be attributed to the electron withdrawing nature of the linker. On average when comparing all the yields, the imidazolydation reactions appeared to have given the best yields overall which was probably due to the purification process since the *mono*-imidazolidines were purified by a column under pressure whereas the *mono*-guanidines were run though the column for a longer period of time.

Finally, as shown in Table 3.4, the Boc deprotection of **3g-11g** gave yields between 93% and 95%. This showed that the nature of the linker was irrelevant to the percentage yields and that this reaction displayed excellent results.

3.7. Synthetic Strategies for the preparation of the amide linked minor groove binders

Considering that in the symmetric series previously prepared in Rozas' laboratory, those compounds with an amide (CONH) linker had shown very good affinity for DNA,[13] we

decided to prepare the corresponding peptide linked asymmetric derivatives. Since the linker is asymmetric and the cations are different in nature, two structures are possible (Figure. 3.4).

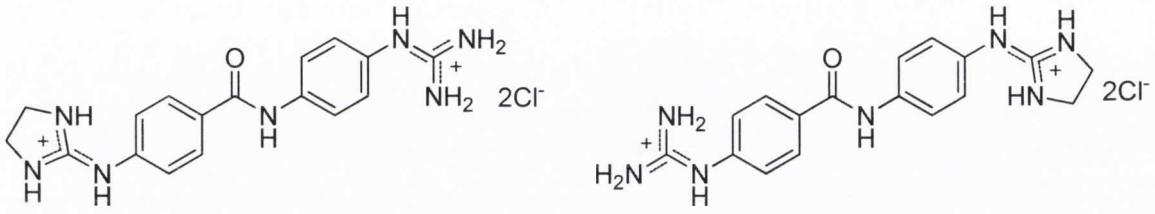
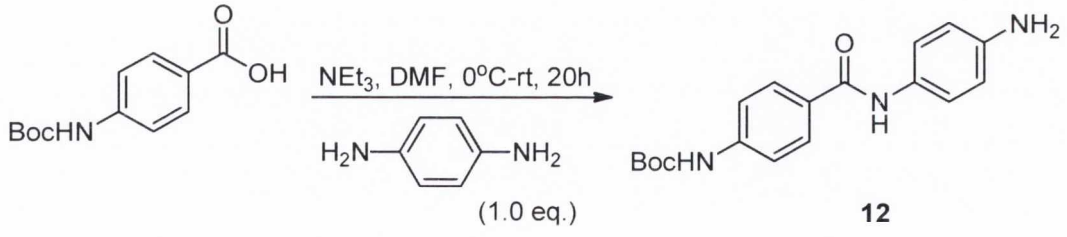


Figure 3.4. Amide linked asymmetric derivatives

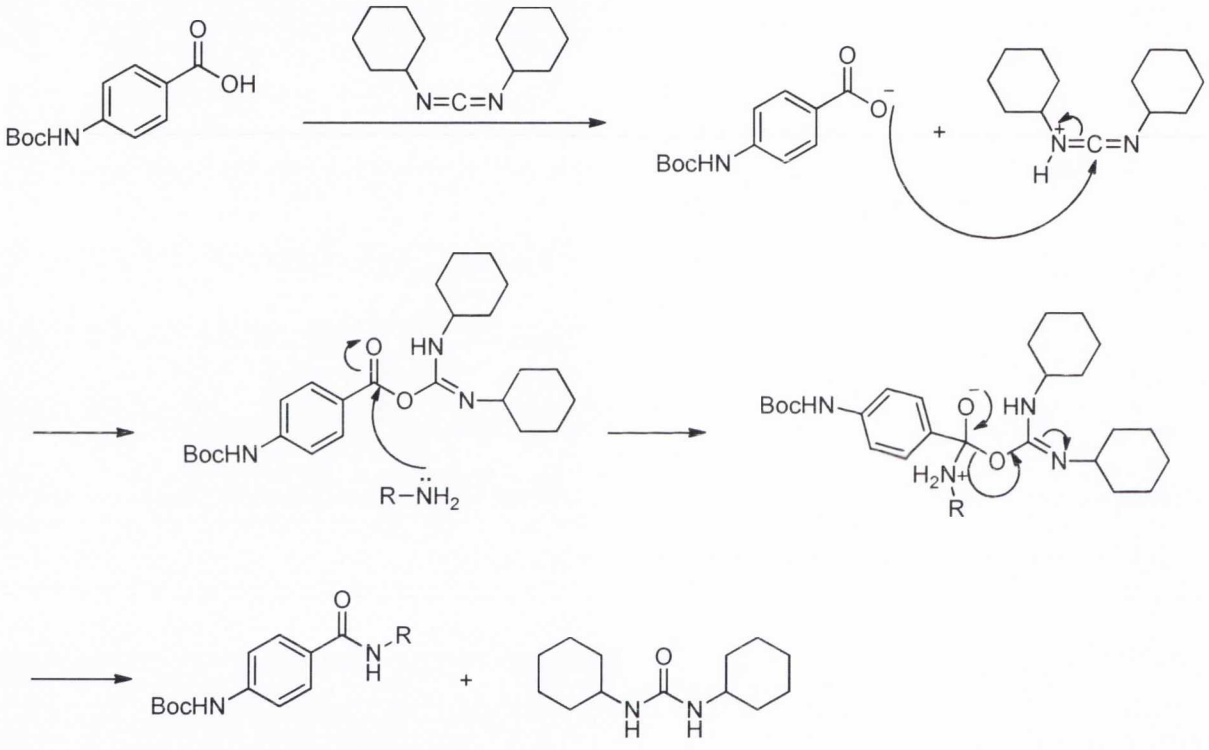
First, we prepared the *mono*-functionalised derivatives. This was carried out by reacting Boc protected 4-aminobenzoic acid (commercially available) with *N*-phenylenediamine (Scheme 3.15) in the presence of DCC and triethylamine in DMF for one hour at 0 °C and a further 19 hours at room temperature. The reaction was monitored by TLC and, on completion, the mixture was diluted with EtOAc. The organic phase was washed with water and brine, dried over anhydrous Na_2SO_4 and concentrated down under vacuum. The crude product was then purified by silica gel chromatography and characterised by NMR (^1H and ^{13}C), IR, melting point and mass spectroscopy.



Scheme 3.15.

The mechanism for the formation of an amide in the presence of DCC is shown in Scheme 3.16. First proton transfer occurs between the benzoic acid and the carbodiimide thereby increasing the electrophilicity of the carbon atom form the carbodiimide. This is due to the difference in acidity between the Boc protected 4-amino benzoic acid and the Boc protected 4-aminobenzoate since the negative charge on the carboxylate anion is resonance

stabilised. Following this, the carboxylate ion attacks the highly electrophilic C atom from the carbodiimide resulting in the movement of a pair of electrons from the C=N double to neutralise the positive charge located on the N atom. This results in the formation of an excellent leaving group. Then, the nucleophilic amine attacks the electrophilic carbon causing the C=O bond to break. A pair of electrons localise on the O atom, and then move back to reform the thermodynamically favourable carbonyl group. This results in the expulsion of the good leaving group and after proton transfer dicyclohexylurea is formed as a side-product.

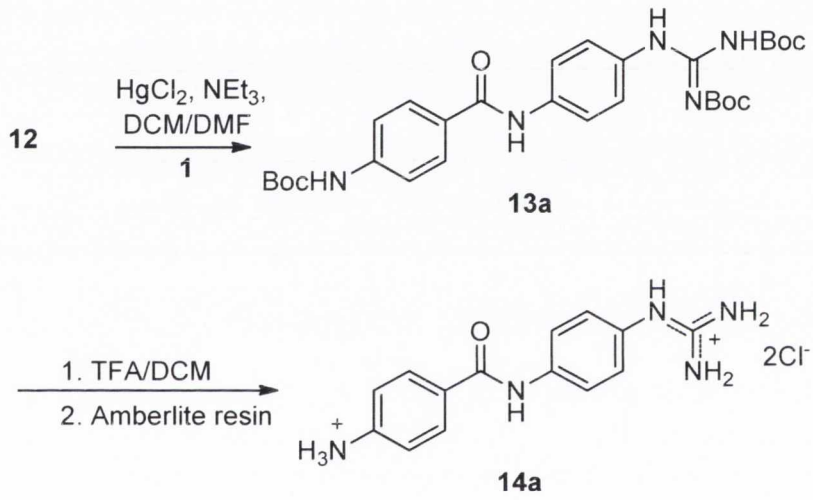


Scheme 3.16.

This compound was then reacted with **1** in the presence of mercury (II) chloride and NEt_3 in DMF for one hour at 0 °C and a further 23 hours at room temperature (Scheme 3.17). Following usual work up, the crude product was purified by silica gel chromatography. The compound was characterised by NMR (^1H and ^{13}C), IR, melting point and mass spectroscopy.

Evidence for the formation of the product was observed by the presence of the Boc hydrogens in the ^{1}H NMR spectra.

Compound **13a** was deprotected in a mixture of trifluoroacetic acid and DCM for four hours. The solvent was then removed under vacuum to yield the trifluoroactate salt. This salt was then treated with basic amberlite resin overnight. The resin was then filtered off and the aqueous phase was washed with DCM. The water was then removed under vacuum to give the dihydrochloride salt that was characterised by NMR (^{1}H and ^{13}C), IR, mass spectroscopy and melting point. The purity of the compound was verified by elemental analysis.

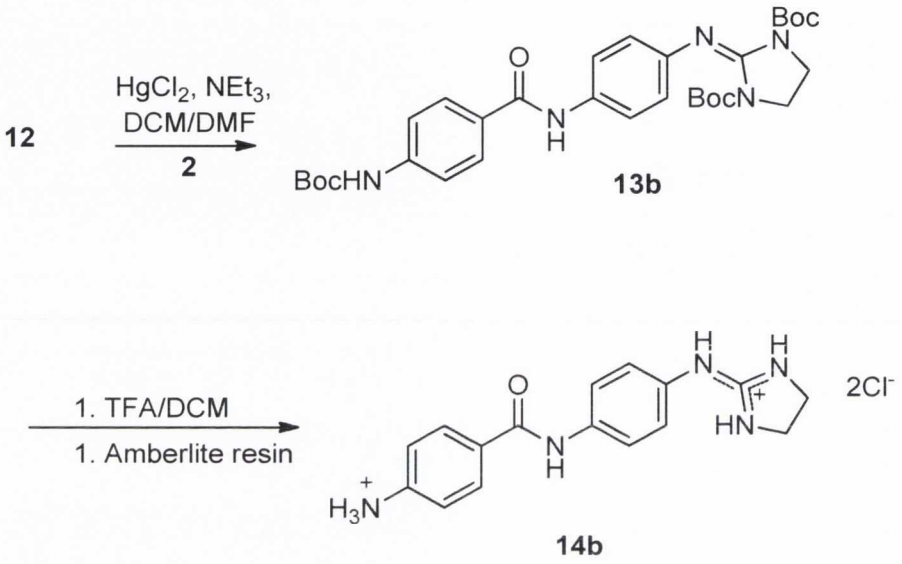


Scheme 3.17.

The corresponding Boc-protected *mono*-imidazoline compound was prepared through a similar method except by reacting the *mono*-Boc protected diamine with **2** as shown in Scheme 3.18. The reaction was monitored by TLC. On completion of the reaction, it was worked up and the crude product was purified by alumina column chromatography to give **13b** which was characterised by NMR (^{1}H and ^{13}C), IR, melting point and mass spectroscopy. The formation of the compound was observed by the presence of the Boc hydrogens in the ^{1}H NMR spectrum.

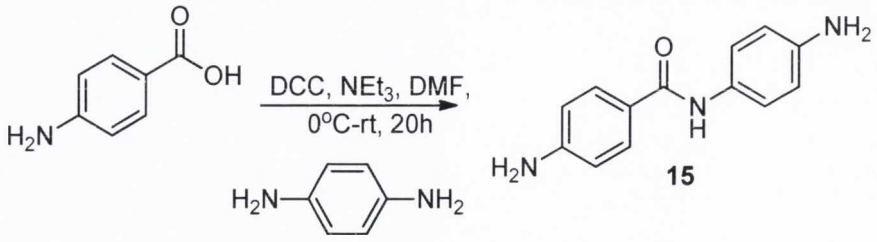
Compound **13b** was deprotected in a mixture of trifluoroacetic acid and DCM for four hours. The solvent was then removed under vacuum to yield the trifluoroactate salt, which

was then treated with basic amberlite resin overnight. The resin was then filtered off and the aqueous phase was washed with DCM. The water was then removed under vacuum to give the dihydrochloride salt that was characterised by NMR (^{1}H and ^{13}C), IR, mass spectroscopy and melting point. The purity of the compound was verified by elemental analysis.



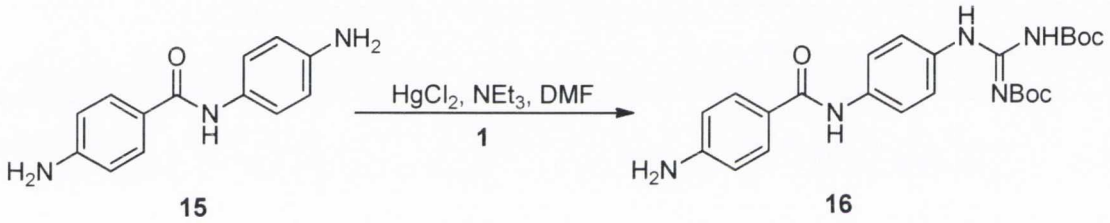
Scheme 3.18.

For the synthesis of the di-functionalised compound with the amide linker, firstly the starting diamine had to be prepared. This was carried out according to the literature[13] by reaction of 4-aminobenzoic acid with 4-aminoaniline in the presence of DCC and NEt_3 in DMF at 0 °C for one hour and a further 19 hours at room temperature. When the reaction was completed according to TLC, it was worked up. The crude product was purified by silica gel chromatography to yield the diamine (**15**) with the amide linker as shown in Scheme 3.19.



Scheme 3.19.

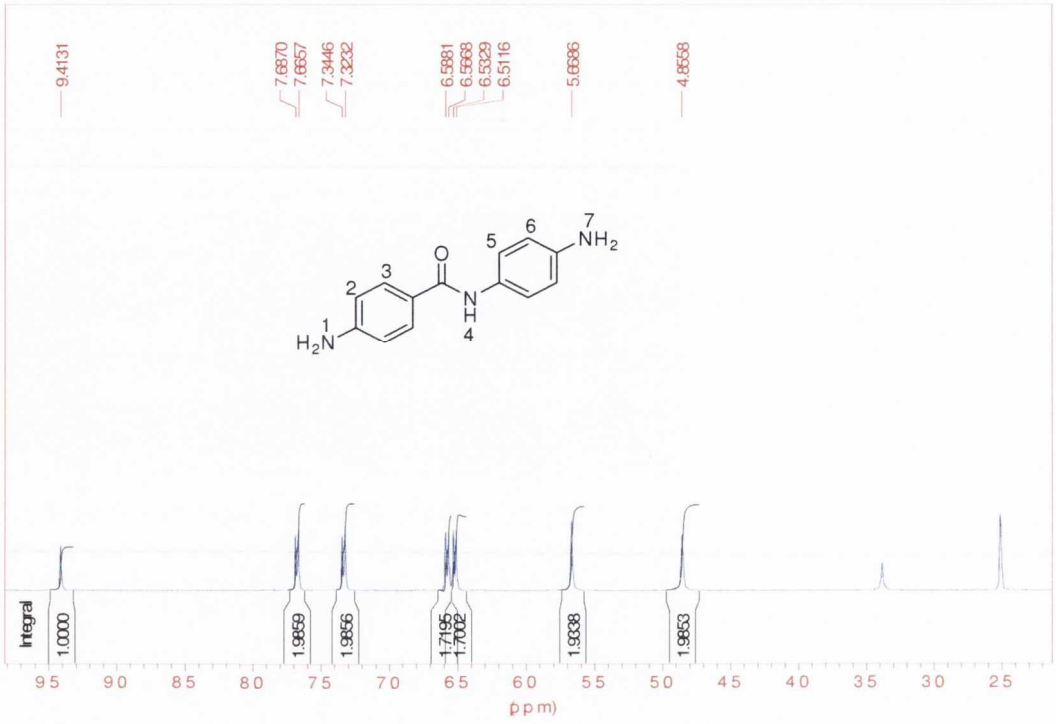
For the preparation of the asymmetrical *bis*-amidine like molecules, it was considered that the amino group *para* to the NH of the linker was more nucleophilic than the amino group *para* to the carbonyl group due to electron donation by resonance. As a result, compound **15** was reacted with the guanidylating agent **1** in a 3:1 ratio in the presence of mercury (II) chloride and NEt_3 in DMF for one hour at 0 °C and a further 23 hours at room temperature (Scheme 3.20). When the reaction reached completion it was worked up and the crude product was purified by silica gel chromatography. Three spots as seen in the TLC were separated by column chromatography. It was anticipated that the middle spot in the TLC was the required product (**16**) and that the top spot was the di-functionalised one so we decided to characterise the possible product.



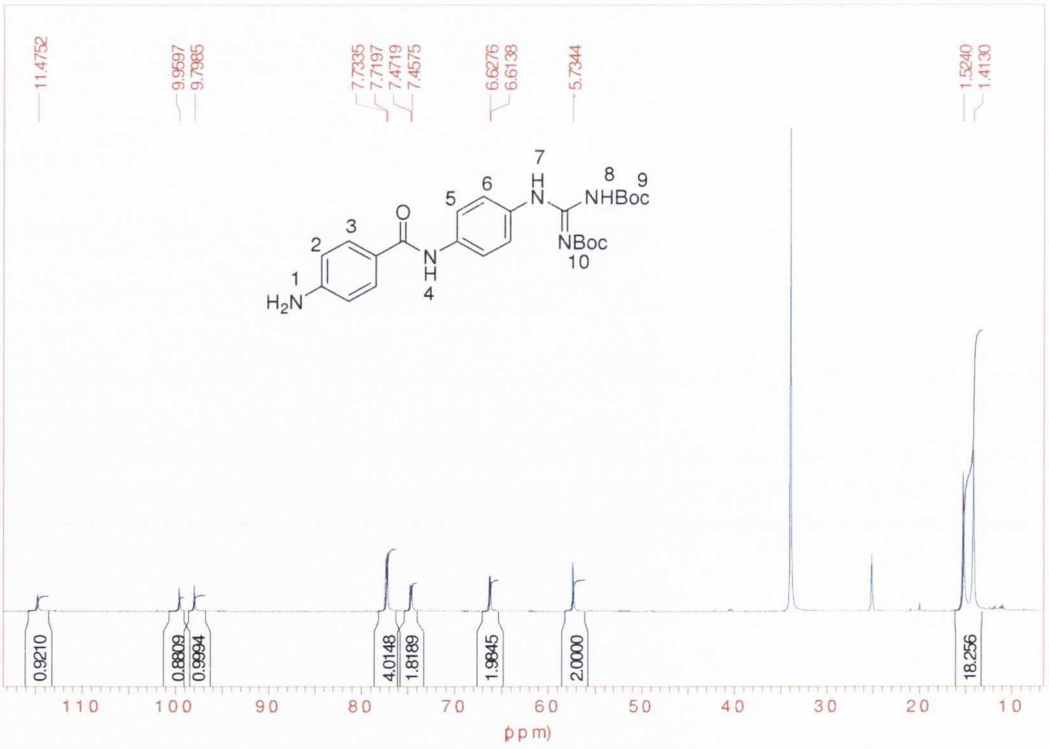
Scheme 3.20.

An extensive NMR analysis was carried out on the molecule in order to determine if the required product was obtained. To help with our analysis we compared the ^1H NMR spectrum of the starting diamine with that of the product as shown in Figure 3.5.

By analysing the spectrum of **15**, we were able to assign the amide proton at 9.41 ppm. All the aromatic peaks were observed between 7.69 and 6.51 ppm. Finally, the two NH_2 signals were seen at 5.67 and 4.86 ppm and the signal at 5.67 ppm was assigned to the amino group *para* to the carbonyl function whereas the peak at 4.86 ppm was assigned to the one *para* to the NH of the amide linker. This is due to the fact that the NH_2 *para* to the carbonyl group is more deshielded due to the electron withdrawal by resonance whereas the opposite occurs for the other NH_2.



H	δ ppm
N-1	5.67
H-2	7.67
H-3	6.58
N-4	9.41
H-5	7.33
H-6	6.52
N-7	4.86



H	δ ppm
N-1	5.73
H-2	6.62
H-3	7.72
N-4	9.79
H-5	7.46
H-6	7.72
N-7	9.96
N-8	11.48
Boc-9	1.41
Boc-10	1.52

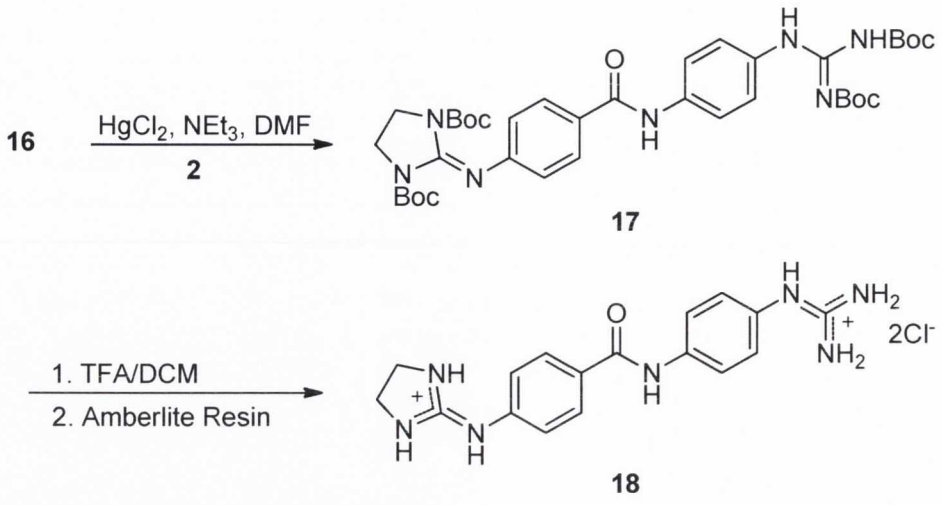
*Figure 3.5. ^{1}H NMR spectra of **15** and **16***

Comparing this ^{1}H NMR spectrum to that for **16**, we observed two extra NH peaks due to the presence of the Boc protected guanidine group. The aromatic signals were observed in the region between 7.73 to 6.61 ppm. The main difference between both spectra was that there was only one amino peak present at 5.73 ppm. This resembles the amino signal at 5.67 ppm in the spectrum of the starting material whereas that at 4.86 ppm has disappeared due to the guanidylation of the starting material.

Finally, the Boc protons were observed at 1.52 and 1.41 ppm. Since the peak at 4.86 ppm had disappeared, we deduced that the NH_2 group *para* to the NH of the amide linker was the one guanidylated. Following this analysis we used C-H and H-H COSYs for extra proof to confirm that the structure was that of the required product.

The next step for the preparation of the asymmetric amide derivative is the reaction between **16** and the Boc protected imidazolidine (**2**) to form the Boc protected asymmetrical molecule. In this reaction the mercury (II) chloride was added to a cooled solution of **16** and **2** with NEt_3 in DMF and was left stirring for one hour and a further 23 hours at room temperature. The resulting black mixture was diluted with EtOAc and the mercury sulphide was filtered off. The reaction mixture was washed with water and brine, dried over anhydrous Na_2SO_4 and concentration under vacuum to give a yellow oil that was purified by silica gel chromatography to give **17** as a white solid.

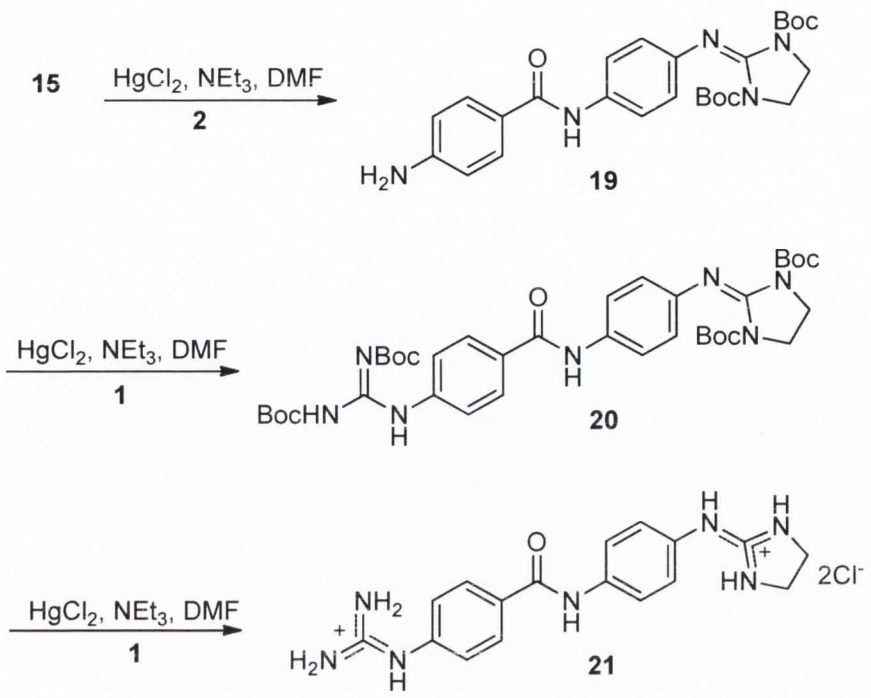
This compound was characterised by ^{1}H and ^{13}C NMR, IR, mass spectroscopy and melting point. From the ^{1}H NMR, we observed the disappearance of the free amino peak at 5.73 ppm, indicating the absence of **16.** The appearance of an extra peak in the aliphatic region, accounts for the hydrogens from the imidazolidine group. Both the IR and mass spectroscopy supported the evidence that **17** was obtained.



Scheme 3.21.

Following the preparation of the Boc protected precursor **17**, the final step was to liberate the Boc groups. This was carried out by reacting **17** in a mixture of TFA and DCM for four hours. The TFA was removed under vacuum to yield the trifluororacetate salt. This was treated with amberlite resin exchange in the chloride form to yield the di-hydrochloride salt **18**. This compound was characterised by ^{1}H NMR, ^{13}C NMR, IR, mass spectroscopy and microanalysis. By analysis of the ^{1}H NMR spectrum, we can see that the Boc groups were removed due to the absence of the corresponding peaks. Mass spectroscopy and elemental analysis supports the purity of **18**.

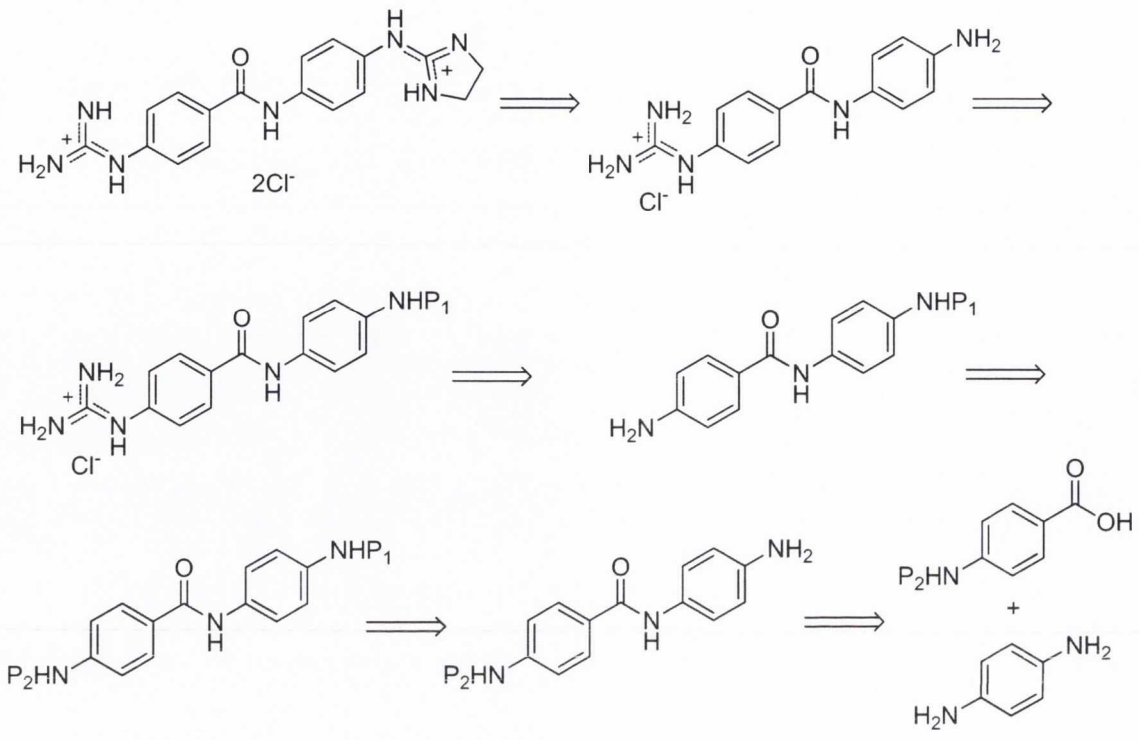
The second asymmetric amide product was attempted by a similar pathway as shown in Scheme 3.22. Here, we intended to introduce the imidazolidine functional group followed by the introduction of the guanidine group.



Scheme 3.22.

In the first step of this reaction, the starting diamine was reacted with **2** in the presence of mercury (II) chloride and NEt_3 in DMF. The reaction was monitored by TLC and when the Boc protected imidazolidine had disappeared, the reaction, diluted with EtOAc, was worked up as previously described. However, TLC indicated that the starting diamine was still present, along with two other spots that were presumed to be the *mono*-functionalised and di-functionalised molecules. Purification of this mixture was attempted by alumina chromatography under pressure to avoid decomposition. However, this method proved unsuccessful as the required product was never isolated as observed by ^1H NMR. Consequently, this synthetic pathway was abandoned.

Then, other synthetic routes were explored trying to introduce the imidazolidine functionality in the last step. A full retrosynthetic analysis (Scheme 3.23) was carried out, and from this we decided to exploit orthogonal protection.



Scheme 3.23

In order to achieve this, we had to protect both sides with different protecting groups as we had to functionalise them separately. The first protecting group that we used was the Boc group due to its stability and ease of removal. The other option was to use the Fmoc group; however, this group is easily liberated with bases such as NEt_3 so this ruled out its use. According to the literature,[14] a Cbz (or Z) protection of aromatic and aliphatic amines works with excellent yields under solvent free conditions. It was decided to exploit this procedure and use Cbz as a second protective group since it can easily be removed by catalytic hydrogenation and is stable to most bases and aqueous acids.

According to the retrosynthetic pathway, the first step of this synthesis involved the coupling of the commercially available Boc protected 4-aminobenzoic acid with 4-amino aniline in the presence of the isouronium-based coupling reagent TBTU (Figure 3.6). This proved to be a successful method however it was low yielding. Hence, we decided to use another coupling reagent called BOP (Figure. 3.6) which is a phosphonium-based coupling reagent.[15]

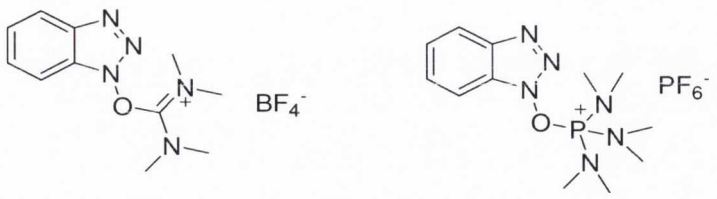
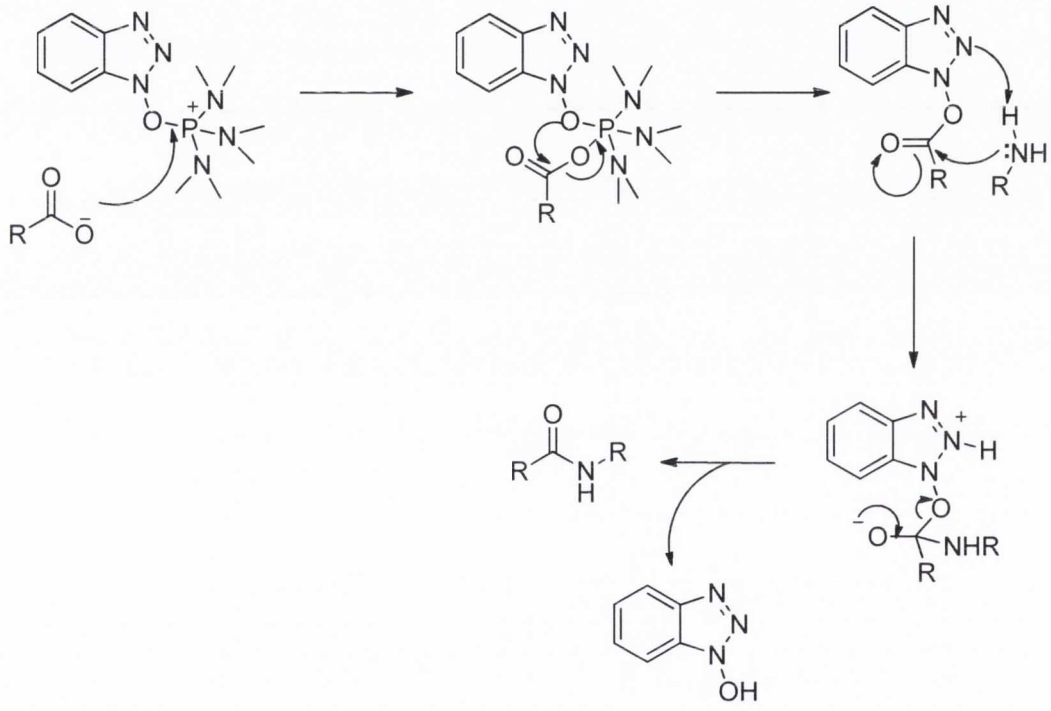


Figure 3.6. Structures of TBTU (left) and BOP (right).

It was thought that this coupling reagent would be more efficient since the phosphorus atom would form a strong bond with the oxygen atom of the carboxylic acid functional group (Scheme 3.24).



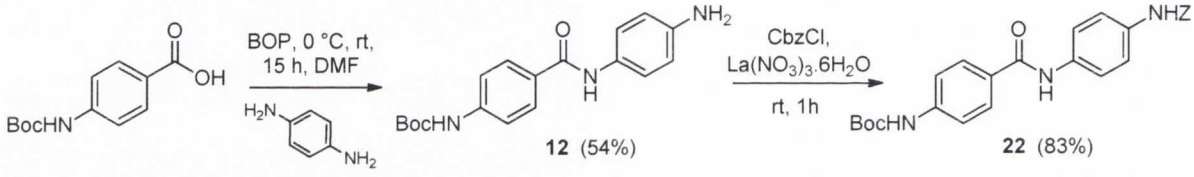
Scheme 3.24.

Following work up the crude product was purified by silica gel column chromatography to yield the required amide linked compound **12** (Scheme 3.25) in good yields and was characterised by ^{1}H and ^{13}C NMR, IR, melting point and mass spectroscopy.

The next step of the reaction was to add the Cbz group to the free amino group in order to protect this end from guanidylation. It was anticipated that the conditions for Cbz-protection should be analogous to that of Boc protection. Therefore, the amide compound

was reacted with CbzCl in the presence of NEt_3 in DMF. This produced low yields so we decided to react the aromatic amine with CbzCl in the presence of a Lewis acid catalyst under solvent free conditions. The Lewis acid used was lanthanum (III) nitrate hexahydrate, and there are many advantages of using lanthanide based catalysts. For example, the ions coordinate preferentially to oxygen or nitrogen based ligands, and they are also stable in air, water and versus many other functionalities.[15]

The reaction was carried out by adding CbzCl to **12** in the presence of a catalytic amount for lanthanum (III) nitrate for one hour as shown in Scheme 3.25. When the reaction was judged complete by TLC, the reaction was worked up. The crude product was purified by silica gel chromatography giving a beige solid. This compound **22** was characterised by ^1H and ^{13}C NMR, IR, HRMS and melting point.

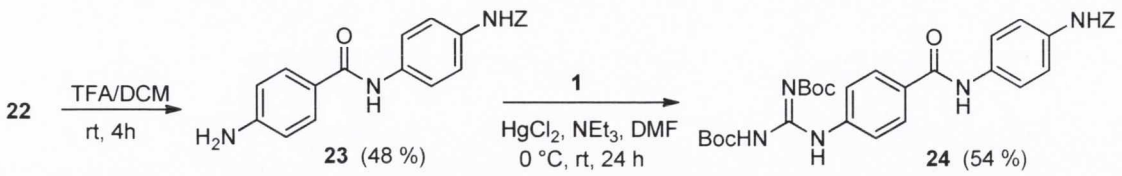


Scheme 3.25.

The next step was to deprotect the Boc group. Initially, we utilised 1.25M HCl in MeOH mixture; however, the compound was insoluble in methanol. Consequently, we dissolved the compound in a mixture of TFA/DCM and left it stirring at room temperature for four hours. The solvent was then removed under reduced pressure to yield the trifluoroacetate salt. This salt was then redissolved in a 2M solution of NaOH and the product was extracted using EtOAc. The solvent was then removed under reduced pressure to yield the free amine (**23**, in Scheme 3.26) as a beige solid. This compound **23** was characterised by ^1H and ^{13}C NMR, IR, mass spectroscopy and melting point.

Once deprotection was complete, the next step was to guanidylate the free amino group (Scheme 3.26). This was carried out using the well known protocol described earlier. The amine **23** was reacted with **1** in the presence of mercury(II) chloride and NEt_3 in DMF for one hour at 0 °C and a further 23 hours at room temperature. When the reaction was complete as judged by TLC, it was worked up and purified by silica gel chromatography to

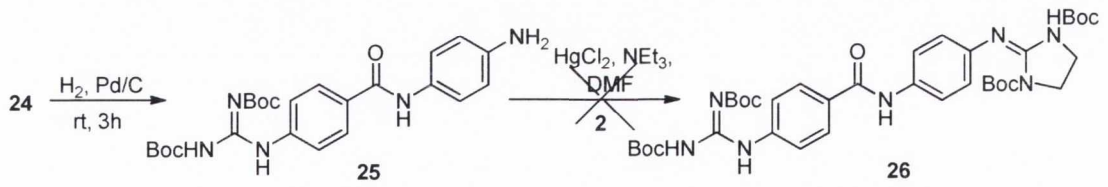
yield a white solid that was characterised as compound **24** by ^{1}H and ^{13}C NMR, IR, HRMS and melting point.



Scheme 3.26.

The next step in this synthetic route was the deprotection of the Cbz group. It is well known that this group is labile with catalytic hydrogenation. Consequently, the compound was hydrogenated in DMF in the presence of a palladium catalyst according to Scheme 3.27. This reaction was monitored by TLC and when it was complete, the solvent was evaporated to yield a crude product that was purified by silica gel chromatography. The purified product **25** was characterised by ^{1}H and ^{13}C NMR, IR, HRMS and melting point.

The final step to carry out in this pathway was the imidazolidylation of the free amino group. This was carried out by reacting the amine with the Boc-protected imidazoline **2** in the presence of mercury (II) chloride and NEt_3 in DMF. However, this reaction was unsuccessful, as the starting material was isolated after column chromatography. Different conditions were attempted such as the variation of temperature and reaction time. Yet, no product was formed. Thus, the synthesis of **26** was then abandoned due to time constraints.



Scheme 3.27.

3.8. Conclusions for the various Synthetic Strategies employed for the preparation of the minor groove binders

In summary, various synthetic strategies have been employed to prepare three families of potential minor groove binders. A very common procedure used in the Rozas laboratory is the functionalisation of the amino groups with the guanidylating and 2-aminoimidazolidylating agents with the use of mercury (II) chloride and triethylamine. This proved to be the most efficient method for the preparation of these compounds since when zinc (II) chloride was used in place of mercury (II) chloride, the yield of the reaction decreased.[16] This is due to the fact that mercury is a thiophilic element.

- ***Families I and II***:

  - I have prepared a family of precursor molecules (the mono Boc protected molecules) used for the preparation of the mono functionalised families. I was able to deduce that the functionalisation with the 2-aminoimidazoline group gave slightly higher yields in most cases than the corresponding functionalisation with the guanidine group. In addition, I observed increased yields with molecules that contain electron donating linkers in comparison to those that contain electron withdrawing ones due to the increased nucleophilicity of the amino groups.

  - Interestingly to note is the way the mono functionalised compounds were prepared. Each of the diamines was mono Boc protected initially, followed by functionalisation of the other free amino group. This was carried out to compare the yields between the guanidylation and 2-aminoimidazolydation steps.

  - Another way of preparing the mono guanidine family is to directly guanidylate the starting diamine which was a more time efficient pathway. However, due to decomposition problems associated with the preparation of the mono 2-aminoimidazoline family using this route, it was decided not to use it for these derivatives.

- ***Family III***: The mono guanidine compounds directly prepared were used as precursors for the synthesis of the asymmetric di-functionalised compounds and, thus, the synthetic strategy was to introduce the Boc protected guanidine group first to avoid decomposition problems. Then, the Boc protected 2-aminoimidazoline system was generated to proceed afterwards to the full Boc deprotection using trifluoroacetic acid and the generation of the hydrochloride salt using basic Amberlyte resin.

  - A different strategy had to be considered in order to prepare the functionalised molecules with asymmetric linkers such as the amide linker. Through the use of a detailed NMR analysis, we were able to prepare one of the asymmetric molecules with the amide linker; however, this was not the case for the other molecule as it would have required the initial addition of the 2-aminoimidazoline functional group. Due to previous decomposition problems associated with this functional group it was decided to add this group on last. A retrosynthetic pathway was designed in order to investigate the ideal synthetic route and it was decided to make use of orthogonal protection of the amino ends with Boc and Cbz groups so that I could introduce the Boc protected guanidine group first followed by the introduction of the imidazolidine functional group. This synthesis was carried out; however, problems with the final 2-aminoimidazolidylation step did not allow the preparation of the desired asymmetric amide derivative.

3.9 References

1. Wilson, W. D.; Nguyen, B.; Tanious, F. A.; Mathis, F.; Hall, J. E.; Stephens, C. E.; Boykin, D. W.; *Current Medicinal Chemistry-Anti-Cancer Agents*, **2005**, *5*, 389
2. Arafa, R. K.; Brun, R.; Wenzler, T.; Tanious, F. A.; Wilson, W. D.; Stephens C. E.; Boykin, D. W.; *J Med Chem*, **2005**, *48*, 5480
3. Neidle, S.; Kelland, L.R.; Trent, J.O.; Simpson, I.J.; Boykin, D.W.; Kumar, A.; Wilson, W.D. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1403
4. Francesconi, I.; Wilson, W.D.; Tanious, F.A.; Hall, J.E.; Bender, B.C.; Tidwell, R.R.; McCurdy, D.; Boykin, D.W. *J. Med. Chem.* **1999**, *42*, 2260
5. Chen, Y-H.; Lown, J. W. *J. Am. Chem. Soc.* **1993,** *116,* 6996
6. Tanious, F.; Hamelberg, D.; Bailly, C.; Czarny, A.; Boykin, D.; Wilson, W.D.; *J. Am. Chem. Soc.* **2004,** *126,* 143
7. Mallena, S.; Lee, M. P. H.; Bailly, C.; Neidle, S.; Kumar, A.; Boykin, D.; Wilson, W. D. *J. Am. Chem. Soc.* **2004,** *126,* 13659
8. Miao, Y.; Lee, M. P. H.; Parknison, G. N.; Batista-Parra, A.; Ismail, M. A.; Neidle, S; Boykin, D. W.; Wilson, W. D.; *Biochemistry*, **2005**, *44*, 14701
9. Ismail, M. A.; Arafa, R. K.; Brun, R.; Wenzler, T.; Miao, Y.; Wilson, W. D.; Generaux, C.; Bridges, A.; Hall, J. E.; Boykin, D. W. *J. Med.Chem.* **2006**, *49*, 5324
10. Dardonville, C.; Goya, P.; Rozas, I.; Alsasua, A.; Martin, I.; Borrego, M.J. *Bioorg. Med. Chem.* **2000**, *8*, 1567
11. Kim, K.; Qian, L.; *Tetrahedron Letters,* **1994,** *34,* 7677
12. Nagle, P.S.; Rodriguez, F.; Kahvedzic, A.; Quinn, S.J.; Rozas, I.; *J. Med. Chem.* **2009,** *52,* 7113
13. Han, S.; Kim, Y.; *Tetrahedron.* **2004,** *60*, 2447
14. Chinni Mahesh, K.; Narasimhulu, M.; Srikanth Reddy, T.; Suryakiran, N.; Venkateswarlu, Y. *Tetrahedron Letters,* **2007,** *48,* 55
15. Sieroslawski, K.; Picur, B.; Lis, T. *Journal of Molecular Structure*. 2003, *657,* 93
16. Suryakiran, N.; Malla Reddy, S.; Srinivasulu, M.; Venkateswarlu, Y.; *Synthetic Communications.* **2008,** *38,* 170

Chapter 4

Asymmetric Dications-DNA Interactions: Optical Physicochemical Techniques

4.1. Introduction

Once the synthesis of the potential minor groove binders was completed, the next objective consisted of characterising their binding to DNA by using physicochemical techniques. Several questions remained to be answered such as whether these compounds bind to DNA, do they display sequence selectivity, can we analyse their binding quantitatively, are we able to obtain any information to the mode of binding and, finally, can we dissect their equilibrium binding constants to relevant thermodynamic quantities such as the entropy and enthalpy of binding.

One of the main objectives of this project was to investigate the criteria for a molecule to bind to the DNA minor groove. In B-DNA the bases are situated inside the double helix shielded by the external phosphate groups. Therefore, the negative potential inside the groove could interact electrostatically with cationic groups. This will be investigated by UV spectroscopy as described in this chapter.

Once this energetically favourable interaction brings the drug closer to the double helix, the molecule would then potentially insert itself inside the minor groove. The minor groove in B-DNA contains AT base pairs, is narrow[1] and has a spine of hydration located inside the groove.[2,3] Consequently, when a molecule binds to the minor groove it would displace a certain number of water molecules from the helix increasing the overall entropy of the system. Binding would also result in a reduction in the enthalpy value due to the formation of hydrogen bonds between the ligand and the DNA base pairs.

Once the molecule is in the minor groove, it may interact with the DNA bases. In the literature,[4] it has been shown that the majority of these molecules interact by non covalent forces such as hydrogen bonding, van der Waals contacts and hydrophobic interactions. For a molecule to form hydrogen bonds with the bases, it must fit optimally in the minor groove, and thus a crescent shape would be favourable to complement the convex shape of the groove. However, it was has been shown that some linear molecules display an increased binding affinity in comparison to their crescent shaped counterparts due to the formation of hydrogen bonds with the DNA bases by linking water molecules from the groove.[5] Wilson *et*

al. have shown that the linear molecule CPG 40215A with an amidine linker serves as a 'see-saw' allowing the molecule adopt various conformations in the minor groove due to the bridging water molecules.[6] In order to compare the binding affinity of our drugs, which exhibit different geometries, techniques such as UV spectroscopy, Surface Plasmon Resonance and circular dichroism can be used. In addition, to analyse the mode of the interaction, linear dichroism is a useful technique and, if a thermodynamic analysis is required to understand the binding, Isothermal Calorimetry is a very good technique to be used.

Furthermore, it is important to investigate the sequence selectivity of these drugs and this can be performed by carrying out DNA binding experiments on oligonucleotides of specific base composition such as poly(dA-dT)$_2$ that consist of a run of alternating adenine-thymine bases. The results obtained from these experiments would then be compared to DNA binding experiments that were carried out in the presence of wild type DNA which has a 32% GC content.

Due to the large number of techniques required for a complete analysis, those used in this study will be discussed separately. First, the results obtained with optical techniques will be presented and, then, the non optical methods will be described. Optical techniques require radiation whereas non optical techniques require other energy sources such as heat and will be discussed fully in the next chapter. Therefore, this chapter focuses in optical techniques such as UV spectroscopy, circular and linear dichroism. The principles behind these techniques will be explained followed by the description and analysis of the experiments carried out to determine DNA binding affinity. Quantities such as the equilibrium binding constants will be evaluated allowing us to calculate the relevant thermodynamic quantities. The overall assessment of the binding of our molecules will enable us to generate guidelines of the characteristics that a drug must have to bind to the minor groove.

The discussion on DNA binding will begin with the determination of the pK_a. This quantity is of upmost importance in medicinal chemistry since it allows an insight into the protonation state at physiological pH. This determination will indicate if our compounds are capable of forming electrostatic interactions with the negatively charged DNA macromolecule.

4.2. Evaluation of pK_a

The pK_a of a compound is clearly a vital quantity to determine in medicinal chemistry since much information can be obtained from the protonation state of a compound and how this affects its properties. For example, if a compound is in its protonated form, it will be probably soluble at physiological pH and, therefore, easily transported into cells. More importantly, with regards to our aims in this work, the pK_a value will allow us to determine whether the molecules are mono- or dicationic when in the minor groove. As the environment of the DNA minor groove is considerably more electron rich than its surroundings,[7] the pK_a value allows us to determine whether the molecules are protonated in the minor groove. Three possibilities arise:

- both protonatable sites of the molecules are protonated in the acidic environment of the minor groove and the molecule is dicationic,
- one of the protonatable sites is protonated and hence the molecule is *mono*-cationic
- none of the protonatable sites are protonated and the molecule is neutral as a consequence.

Before the pK_a values were determined, a comprehensive analysis on the different pK_a values of our functional groups was carried out. Initially, we needed an estimate for the pK_a value of the amino group at the end of our *mono* functionalised molecules. The pK_a of aniline was observed as being 4.58 at 25 °C.[8] This value varied depending on the substituent on the *para* position of the aromatic ring. An increase in the acidity is observed if an extra amino group is present *para* to the amino group yielding a pK_a value of 6.08, whereas if this was a nitro group, the pK_a changes to 1.02.[8] This illustrates the large effect that withdrawing or donating groups have on the pK_a value. Interestingly, as the pK_a value varies, so does the (Ar)C-N bond length, and hence, a smaller pK_a value will result in a decrease in the (Ar)C-N bond length.[8] Therefore, considering the groups present in our *mono*-functionalised molecules, it can be assumed that the $-NH_2$ groups will not be protonated at physiological pH similar to the amino group of aniline.

On the other hand, it has been previously determined in the Rozas group[9] that the guanidine and 2-aminoimidazoline groups are highly basic with pK_a values of more than 9.5, the former showing greater basicity. Subsequently, it was expected that these groups would be protonated at physiological pH in the DNA minor groove. We, therefore, expect the *mono*-functionalised molecules to be *mono*-cationic and the di-functionalised molecules to be dicationic.

A comprehensive pK_a study was carried out on some of the proposed minor groove binders with two questions in mind. Firstly, we intended to confirm if our molecules were dicationic at physiological pH. Secondly, we studied the pK_a of the molecules to see if there was any correlation with the DNA binding affinity. Henceforth, we carried out UV experiments to determine the pK_a of our compounds. Since these compounds are asymmetric, we expect to observe two pK_a values, one for each functionality, as seen in Figure 4.1.

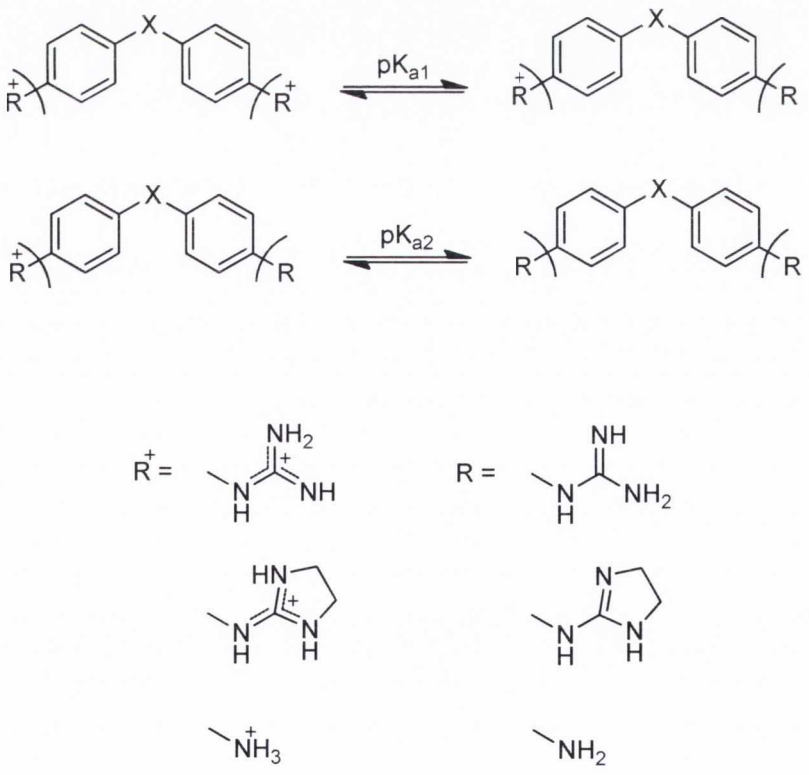


Figure 4.1. Illustration of the protonatable sites on our molecules.

4.2.1. Determination of pK_a using UV spectroscopy

We carried out titrations on some of the *mono*-guanidine compounds to confirm that these compounds were *mono*-cationic. We were expecting two distinct pK_a values since we knew the pK_a value for aniline was 4.63 and the pK_a value of the guanidine group was estimated to be above 9.5.[9]

For these measurements, changes to the UV spectrum were monitored over a wide range of pH values. Initially, a solution of the molecule was prepared in acidic conditions and aliquots of a base were added to increase the pH of the solution, recording the corresponding UV spectra. This was continued until the second pK_a was reached.

First, we carried out the titration of **4c** (linker = CH_2CH_2 from the *mono*-guanidine family). By adding aliquots of NaOH (0.1 M), we observed the formation of a large peak at 236 nm that is indicative of the increased conjugation of the aniline part of the molecule due to the deprotonation of the amino group (Figure 4.2). By plotting pH against normalised absorbance, we observed that the spectral changes occurred in the pH region of 3.5 to 6. Consequently, we were able to evaluate the pK_a for the amino group from these plots.

By increasing the pH of the solution further, we observed another change occurring in the spectrum in the pH range of 10 to 13. This was due to the deprotonation of the guanidine group and the consequent formation of the fully neutral species. Interestingly, the spectral changes were not as pronounced as with the deprotonation of the amino group possibly due to less conjugation with the aromatic ring. We then plotted the normalised absorbance against pH to obtain the second pK_a value. Again, by using the Hasselbach-Henderson equation we were able to evaluate the pK_a of the guanidine functionality as being 11.36. Through the determination of the molecule's two pK_a values, we were able to conclude that compound **4c** is *mono*-cationic at physiological pH and that the positive charge is located at the guanidine functionality.

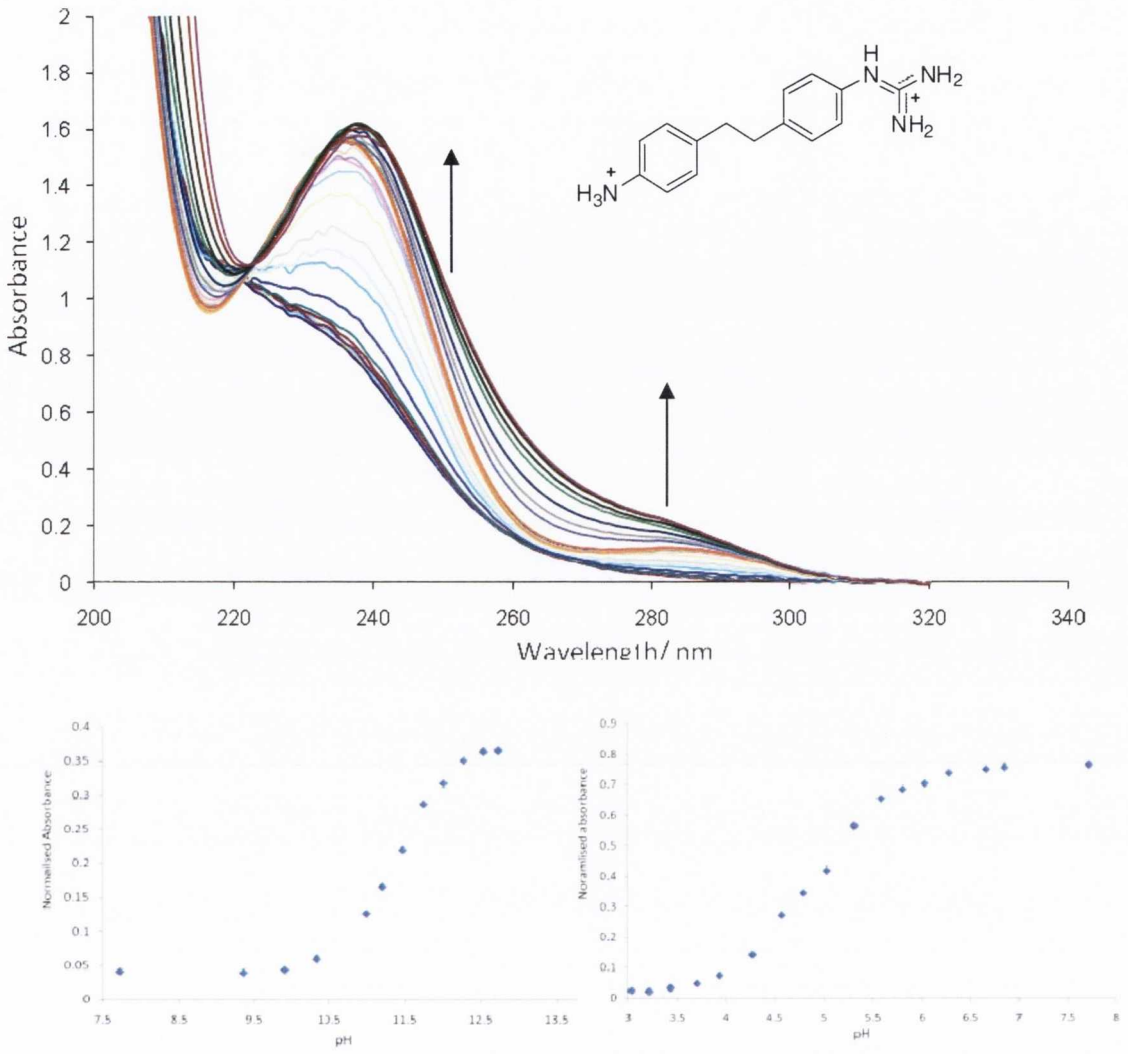


Figure 4.2. (top) spectral changes occurring due to the addition of NaOH (0.1 M), (bottom) the pK_a values were evaluated as being 4.95 (-NH_2) and 11.36 (-NHC[NH_2]NH_2).

A similar experiment was performed for compound **6c** (linker = O, *mono*-guanidine family), and the results obtained are shown in Figure 4.3. The correponding pK_a values obtained were 5.0 for the -NH_3^+ deprotonation and 11.8 for the guanidinium group deprotonation.

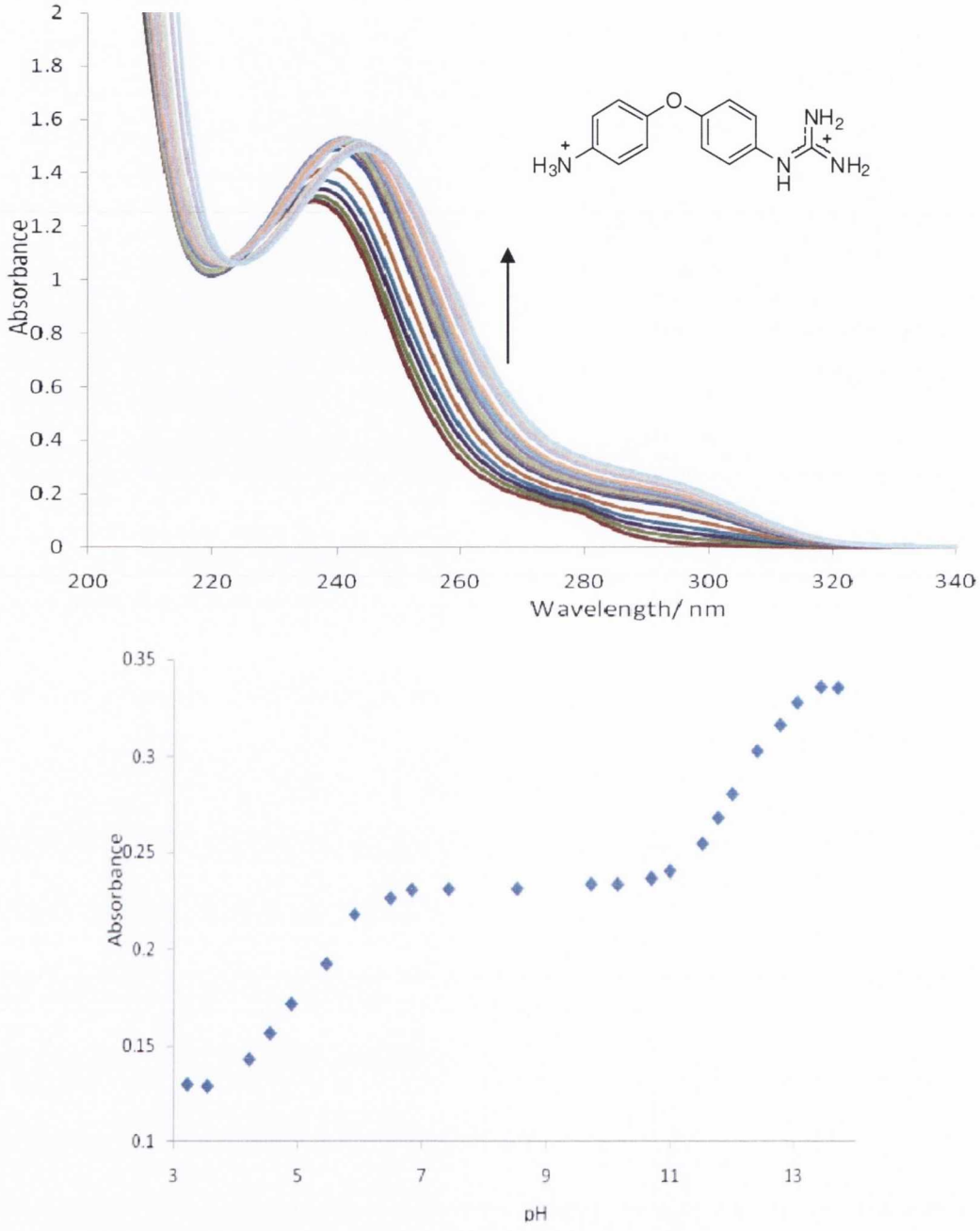


Figure 4.3. (top) spectral changes occurring due to the addition of NaOH (0.1M), (bottom) the pK_a values were evaluated as being 5.00 (-NH_2) and 11.8 (NHC[NH_2]NH_2)

The corresponding titration results for compound **4f** (linker = CH_2CH_2 from the *mono*-2-aminoimidazoline family) are shown in Figure 4.4. In this spectrum, we can see an increase in the absorbance at 231 nm accompanied with a slight bathochromic shift indicating disappearance of the dicationic species and the corresponding formation of the *mono*-protonated species.

By plotting pH against normalised absorbance, we observed that these changes occurred in the pH range of 3.3 to 5.5. As the pH was increased, no change was observed, until the solution reached a pH value of 9.8. Beyond this pH, further changes to the UV spectrum were noticed, that is indicative of the formation of the neutral species at pH 13. The pK_a was calculated to be 4.9 for the amino group and the second pK_a was 10.9 for the 2-aminoimidazoline cation.

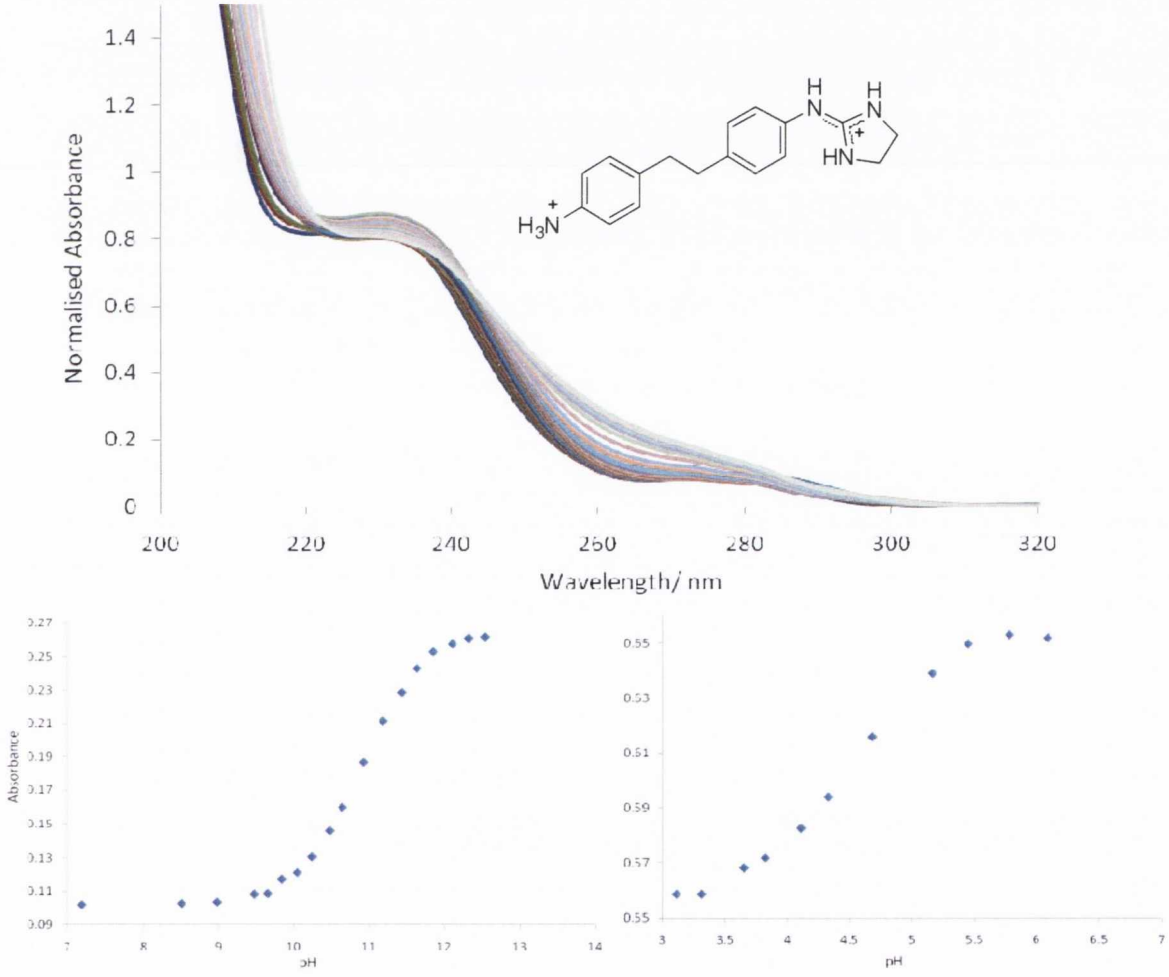


*Figure 4.4. (top) spectral changes occurring for compound **4f** due to the addition of NaOH (0.1M), (bottom) the pK_a values were evaluated as being 4.88 (-NH_2) and 10.93 (2-amino-imidazoline)*

UV titrations were also carried out for some of the hybrid molecules. However, as the pK_a values of the cations are similar, it proved difficult to identify the individual pK_a values. A way that could be used to evaluate these values is by carrying out potentiometric measurements. However, this will not be further discussed since these experiments was not carried out in this project.

4.3. DNA thermal Denaturation Assays

4.3.1. Introduction

Thermal denaturation is an optical technique used for measuring the stability of various molecules on binding to DNA. It relies on the use of absorption spectroscopy and measuring the absorbance of a sample of DNA as a function of temperature.

When DNA is in its double helix form, the bases are stacked away in the helix and are shielded by the external phosphate groups from the light, thus the bases do not absorb much light. However, as a sample of DNA is heated, the hydrogen bonds between the base pairs are given enough heat energy to break. At this point the double helix begins to split apart giving the bases freedom to rotate. Consequently, an increasing amount of light is absorbed by the DNA bases that correlate with the gradual hyperchromic effect seen in the spectrum (Figure 4.5)

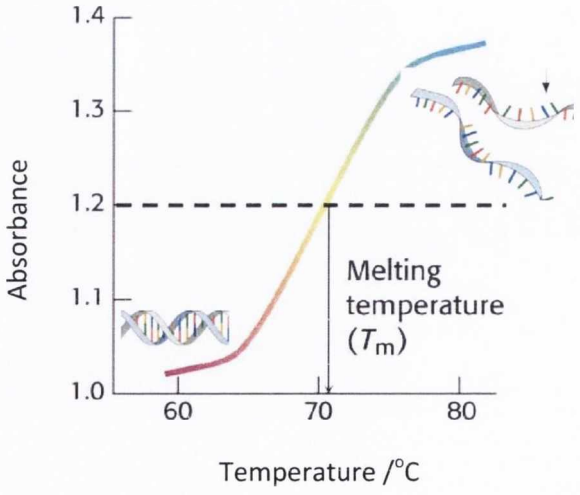


Figure 4.5. Change in DNA absorbance as a result of Thermal Denaturation[10,11]

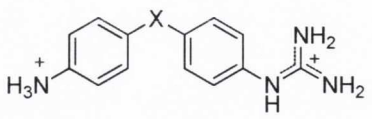
The midpoint of this increase in a population of DNA molecules is called the melting temperature (T_m) correlating to where half of the population of DNA molecules has melted. These experiments are usually run initially with DNA alone followed by experiments with the DNA bound to a ligand and, by comparing the two thermal melting temperatures, we can obtain the change in the thermal melting temperature (ΔT_m). This value can be used as an initial measure of the relative stabilsation to DNA. An increase in the melting temperature would be expected since a compound that binds to the minor groove would stabilise the double helical structure significantly due to its ability to form strong van der Waals contacts and hydrogen bonds with the DNA base pairs.

The melting temperature of DNA is dependent on the sequence type. For example, the observed melting temperature for wild type DNA is 68 ^{o}C whereas for poly(dA-dT)$_2$ DNA, it was found to be 47 ^{o}C. The lower melting temperature of this DNA is due to the lack of GC base pairs with three hydrogen bonds between the bases whereas there are only two hydrogen bonds between the AT bases.

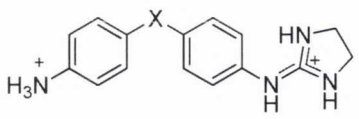
4.3.1. DNA thermal Denaturation Assays on Families I and II

Thermal denaturation experiments were carried out for the mono-functionalised molecules with natural DNA from salmon testes and the results are presented in Table 4.1

Table 4.1. Thermal denaturation assay results in the presence of natural DNA.[13]



3c-10c



3f-10f

Linker	Compound Number	ΔT_m [a]	Compound Number	ΔT_m [a]
CH_2	**3c**	3	**3f**	3
CH_2CH_2	**4c**	0	**4f**	0
NH	**5c**	2	**5f**	2
O	**6c**	2	**6f**	0
S	**7c**	2	**7f**	0
CO	**8c**	0	-	-
NHCONH	**9c**	2	**9f**	4
Piperazine	**10c**	2	**10f**	1
CONH	**14a**	1	**14b**	2

[a] Experiments were carried out with natural DNA (150 μM), phosphate buffer and a Bp/D = 5. Errors for these experiements were not above 5 %.

All *mono*-functionalised molecules displayed poor binding to DNA ($\Delta T_m \leq 4$). These compounds are all *mono*-functionalised, thereby a possible conclusion into the poor binding observed is that the compounds need to be di-functionalised for efficient binding to DNA. An argument could then be presented for the relevance of placing an extra functional group to the free amino end. By adding the extra functional group, this end becomes cationic since the guanidine and 2-aminoimidazoline groups are highly basic and secondly extra hydrogen bonding groups capable of forming hydrogen bonds with the DNA base pairs are introduced.

4.3.3. DNA thermal Denaturation Assays on Family III

The results obtained in the thermal denaturation experiments for the di-functionalised compounds are shown in Table 4.2. These experiments were carried out with natural DNA as before for two reasons:

- to investigate the influence of the extra cation for DNA binding, and
- if significant affinity was observed, to investigate the effects of the various linkers for DNA binding.

There are a few differences between the *mono*-functionalised and di-functionalised molecules that may result in significant increases in the affinity for DNA. Firstly, the length of the di-functionalised molecule is larger than the corresponding *mono*-functionalised compounds which could lead to the formation of optimum hydrogen bonds with the DNA base pairs. Secondly, these molecules have two highly basic functionalities attached and thereby, they would be dicationic in the acidic environment of the minor groove. Thirdly, comparing these molecules with family I (*mono*-guanidine family) they have an extra ethylene moiety from the 2-aminoimidazoline functionality that can form van der Waals contacts with the DNA bases

By studying the results obtained from these experiments (Table 4.2), we noticed that the introduction of a second cationic group significantly increased binding affinity to DNA.

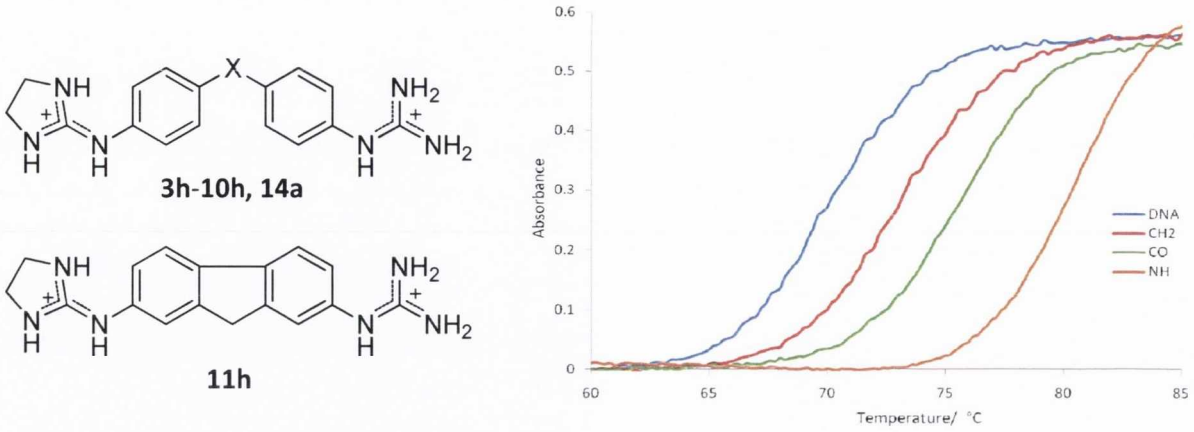


*Figure 4.6. Di-functionalised molecules tested for DNA binding affinity (left) and examples of thermal denaturation experiments that were carried out with [DNA] = 150 μM, phosphate buffer and with a Bp/D = 5, DNA (——), **3h** (——), **5h** (——), **8h** (——)*

Table 4.2. Thermal denaturation assay results for the di-functionalised molecules.

No	*X*	*ΔT_m (°C)[a,b] Nat. DNA*	*ΔT_m (°C)[a,c] poly(dA-dT)$_2$*	*Rel. T_m $\Delta T_{m(AT)}$- $\Delta T_{m(ST)}$*	*Selectivity Diff. T_m $\Delta T_{m(AT)}/\Delta T_{m(ST)}$*	*Sequence Specificity*
3h	CH_2	6	6	0	0	None
4h	CH_2CH_2	8	16	8	2	AT
5h	NH	12	17.5	5.5	1.45	AT
6h	O	7	10	3	1.43	AT
7h	S	7	9	2	1.29	AT
8h	CO	8	15	7	1.88	AT
9h	NHCONH	11.5	14	2.5	1.22	AT
10h	Piperazine	9	15	6	1.67	AT
11h	Fluorene	4	22.2	18.2	5.5	AT
14a	CONH	10.4	26.1	15.7	2.51	AT

[a] Experiments were carried out with natural DNA (150 μM) with a Bp/D ratio = 5, [b] Melting temperature for natural DNA was found to be 68 °C, [c] Melting temperature for poly(dA-dT)$_2$ was found to be 47 °C. Values obtained had 5 % error.

We initially carried out these experiments in the presence of natural DNA (salmon testes) to observe if these molecules displayed good binding affinity relative to the *mono*-functionalised molecules. This natural DNA has 32 % GC content. Due to the presence of the GC base pairs, the minor groove will be average in width, and would possibly accommodate large flexible molecules. Experiments with this DNA are described below.

Analysing the results for molecules **5h, 3h** and **8h** (linkers -NH-, $-CH_2-$ and -CO- respectively) it was seen that **5h** displayed the highest binding affinity (ΔT_m = 12 °C) in comparison to **8h** which in turn was seen to have a more favourable result (ΔT_m = 8 °C) than **3h** (ΔT_m = 6 °C). These molecules differ by their linkers, changing the molecular geometry which affects their curvature and the ability to form optimum hydrogen bonds. The fact that the tetrahedral geometry in the $-CH_2-$ linker was seen to give the worst binding was enforced by the results from molecules **6h** and **7h** (tetrahedral linkers -O- and -S-) which, as anticipated, gave one of the poorest results probably because they did not form optimum hydrogen bonds with the DNA bases. Thereby, we could conclude that when considering the preparation of molecules that bind to the minor groove, a trigonal linker such as -NH- is more favourable than a tetrahedral one. This can be attributed to the ability to form optimum hydrogen bonds or van der Waals contacts with the DNA base pairs within the minor groove

By comparing the thermal denaturation results for molecules **3h** and **4h** (linkers $-CH_2-$ and $-CH_2CH_2-$) it was seen that **4h** displayed higher binding affinity (ΔT_m = 8 °C) than **3h** (ΔT_m = 6 °C). This difference can be attributed to a number of factors such as an extra degree of rotation due to the extra CH_2 group present in the linker, thereby giving the molecule an extra degree of freedom. When comparing the binding affinities we had to consider that the distance between the cations in **4h** is larger than that of **3h**. This could result in the formation of optimum hydrogen bonds with the DNA base pairs.

The effect of lengthening the linker and therefore increasing flexibility could be observed in the increasing binding affinity found for compounds **8h**, **14a** and **9h** (linkers -CO-, -CONH- and -NHCONH-, respectively) in the thermal denaturation experiments. However, this could also be attributed to the extra hydrogen bond donors added on the linkers that would be able to bond to the DNA base pairs (ΔT_m -CO- < ΔT_m -CONH- < ΔT_m -NHCONH-).

It is interesting to compare compounds **3h** (linker -CH_2-) with **11h** (fluorene derivative) since **3h** is more flexible than **11h**. Thus, the aromatic moieties are free to rotate in **3h** whereas in compound **11h** the aromatic rings are planar with respect to each other. In addition, the cations are in the 2 and 7 positions in compound **11h**. It was seen that **11h** displayed worse binding in comparison to **3h** indicating that molecular flexibility is a factor in the DNA binding to natural DNA, and that the position of the cations may also be a critical factor in regulating the binding affinity.

Molecule **10h,** (linker -piperazine-) displayed an average result. By considering the length and flexibility of this molecule, it would have been expected to display a high affinity for DNA. However, given the bulky nature of this group, it would be plausible to suggest that it does not fit well into the minor groove thereby, decreasing its affinity towards DNA.

After carrying out the experiments on salmon sperm DNA, we investigated whether these molecules display sequence selectivity carrying out these experiments in poly(dA-dT)$_2$ DNA. Since this DNA lacks GC base pairs, the minor groove is subsequently narrower than the minor groove of natural DNA. Thereby, we would expect that molecules that have a better fit into the groove would display an increased binding affinity. These molecules would be expected to be relatively planar with a small dihedral angle between the aromatic rings. A detailed discussion is carried out on the results from these experiments.

First, by comparing these results, we observed that all molecules display selectivity for AT sequences except compound **3h** which stabilizes AT DNA to the same extent as natural DNA (see Table 4.2). Secondly, by defining the difference of thermal melting temperature between AT and natural DNA as ***Rel.*** $T_m = \Delta T_{m(AT)} - \Delta T_{m(ST)}$, we could see that these molecules could be divided in two groups. Group one with a ***Rel.*** $T_m > 5$ (linkers -CH_2CH_2-, -NH-, -CO- and -Piperazine-) and group two with a ***Rel.*** $T_m < 5$ (linkers -CH_2-, -O-, -S- and -NHCONH-). All the tetrahedral linker molecules seem not to greatly enhance sequence selectivity thereby, the tetrahedral geometry in the linker does not seem to allow optimum hydrogen bonds to be formed between the cationic ends of the molecules and the DNA bases.

By comparing the ***Rel.*** T_m of molecules **3h** and **4h** (linkers -CH_2- and -CH_2CH_2-) it can be seen that **4h** has enhanced binding affinity for AT sequences in comparison to **3h**. It

seems that by increasing the length of the linker, an increase in the DNA binding affinity and sequence selectivity is observed. The extra methylene in the linker could increase the binding affinity by forming extra van der Waals contacts with the DNA base pairs, or by latching around the DNA minor groove pocket.

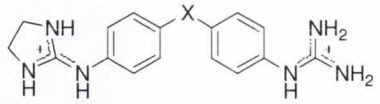
Apart from the tetrahedral linker molecules and that with the urea linker, significant values for the ***Rel. T_m*** were observed. This indicates that molecules that have a trigonal or trigonal pyramidal linker and those with flexible linkers (as in the case of the molecules with the ethylene and piperazine linkers) favour DNA sequence selectivity. This is due to the fact that these molecules can establish close contacts with the DNA bases so hydrogen bonds could be formed.

Interestingly, compound **11h** (fluorene derivative), which displayed poor binding affinity with natural DNA as described before, was the molecule with one of the strongest binding to poly(dA-dT)$_2$ DNA (ΔT_m = 22.2 ^{o}C). This can be attributed to the narrower minor groove of AT sequences, thereby since this molecule has no rotational freedom around the linker, it can fit well into the groove without much structural changes.

Moreover, it was noted that compound **14a** (linker CONH) in the presence of AT sequences displayed the highest binding affinity illustrating that in the design of future potential drugs, the amide linker would be favourable for binding. However, although **14a** showed a strong binding affinity for AT oligonucleotides, its selectivity factor (2.5) was considerably less than that of **11h** (5.5).

In order to verify the AT sequence selectivity of these compounds, thermal denaturation assays were carried out with poly(dA).poly(dT). This polynucleotide differs from poly(dA-dT)$_2$ by the fact that has non-alternating AT bases. The results obtained are shown in Table 4.3.

Table 4.3. Thermal denaturation results for dicationic derivativs with poly(dA).poly(dT).



Compd. No.	X	$\Delta T_{m(A)(T)}$ (°C)[a] Poly(dA)–poly(dT) DNA
3h	CH_2	17.2
4h	CH_2CH_2	29
5h	NH	35.1
6h	O	25.3
7h	S	23
8h	CO	32
9h	NHCONH	30
10h	Piperazine	31.1

[a] Experiments were carried out in poly(dA).poly(dT) in MES buffer with a Bp/D = 3

These experiments were carried out in Georgia State University by Prof. David Wilson's group using a different Bp/D ratio (3 instead of 5) and buffer (MES instead of phosphate), what explains the difference in magnitude of the values obtained compared to those resulting from our thermal denaturation experiments with natural DNA and poly(dA-dT)$_2$. However, a strong correlation was seen between the results obtained in both laboratories as seen in the Figure 4.7.

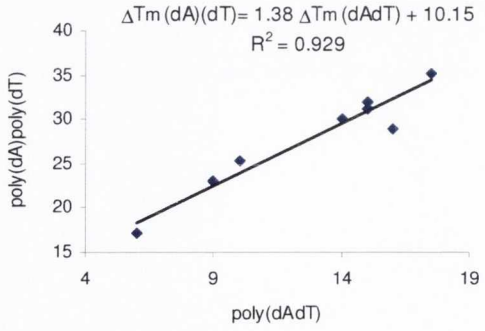


Figure 4.7. Correlation found between the ΔT_m poly(dA–dT)$_2$ vs. ΔT_m poly(dA).poly(dT)

Overall from these studies we can conclude that the *mono*-functionalised molecules display poor binding to DNA, however, by replacing the free amino group by another cationic group, we increase the binding affinity significantly. The results for the asymmetric di-functionalised molecules show that the affinity is dependent on the geometry and the length of the as well as their ability to form hydrogen bonds.

4.4. UV Titration studies

4.4.1. Introduction: Studies of the changes in the UV-Vis spectra of ligands upon the addition of DNA

UV spectroscopy is a powerful optical technique that can be used to measure the binding affinity of molecules to DNA. It has a wide range of applications:

i.) Sequence selectivity can be investigated

ii.) The polyelectrolyte theory[13] can be studied and the dependence of DNA binding on the electrostatic interactions analysed

In this technique the free molecule binds to the DNA macromolecule to form a DNA-ligand complex with a certain binding constant (Figure 4.8). This binding constant is a quantitative measurement of the binding strength of a molecule to DNA allowing for the comparison with other molecules.

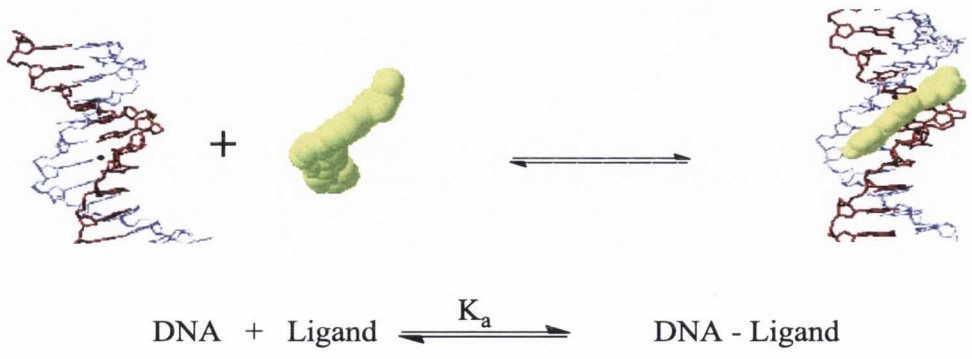


Figure 4.8. DNA-ligand binding equilibrium[14]

The principle behind the technique involves the addition of aliquots of DNA to a sample of the molecule in the appropriate buffer while recording its UV spectra. When the molecule binds to DNA, spectral changes would be observed. Generally, a hypochromic shift will occur which represents a decrease in the concentration of the free molecule accompanied by the formation of a new band at a different wavelength representing the DNA-ligand complex.

Due to the fact that most of our molecules absorb too close to the DNA region, UV spectroscopy experiments could not be carried out for most of our molecules. This was due to the fact that the DNA band was too intense and broad and, consequently, would overtake the spectrum of these molecules. However, some of our molecules absorbed near the visible region of the spectrum far from the DNA one. Consequently, UV titrations were attempted on compounds **5h**, **8h** and **11h** (with the -NH-, -CO- linkers and fluorene derivative respectively).

4.4.2. UV Spectroscopic Studies on some Asymmetric Dications (Family III)

4.4.2.1. Results for compound ***5h*** *(NH linker)*

The UV spectrum of **5h** was recorded (Figure 4.9), and we observed the π-π absorption band at 296 nm with ε = 25,640 $M^{-1}cm^{-1}$. This indicated that a UV titration experiment could be attempted and if the molecule would bind to DNA, we should observe a shift in this absorption band.

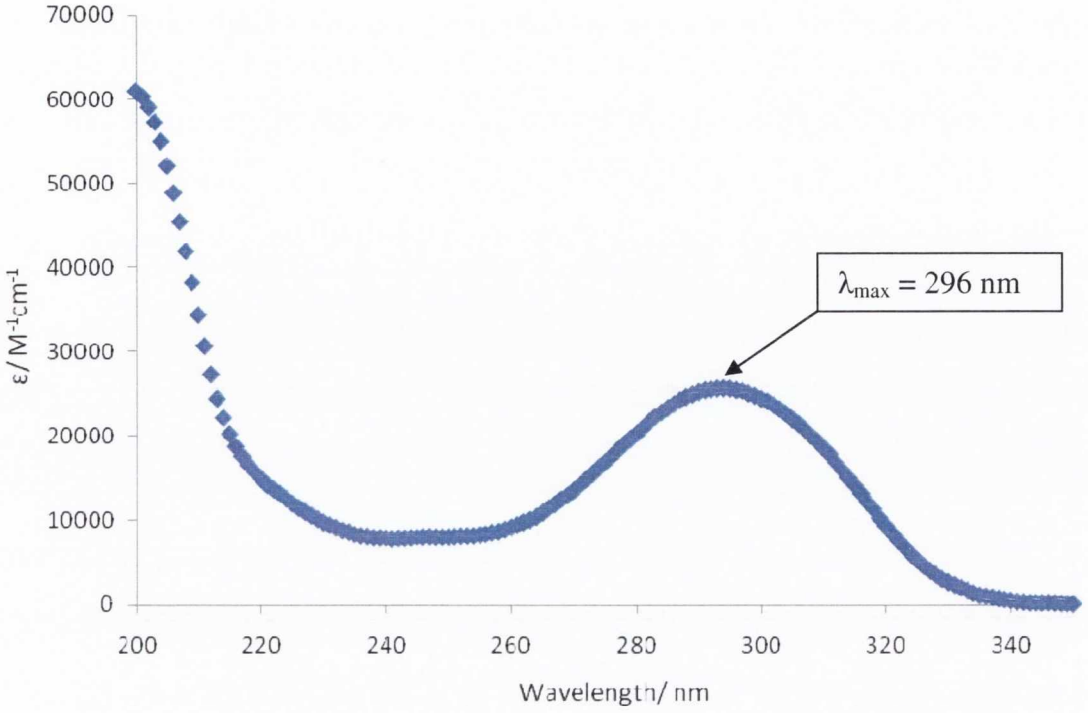


Figure 4.9. UV absorption spectrum for ***5h***

Aliquots of a 2 mM natural DNA solution were added to the compound solution (1.6 μM), and a small hypochromic shift was observed indicating the disappearance of the free molecule. This was followed by a bathochromic shift showing the formation of a new DNA-ligand species (Figure 4.10). To assess the sequence selectivity of this compound, the experiment was repeated with the poly(dA-dT)$_2$ polynucleotide. A large bathochromic shift was observed along with the formation of a new band at 313 nm indicating the formation of the new DNA-ligand species. This new band had not been observed in the titration with natural DNA indicating that **5h** displays a higher affinity for the poly(dA-dT)$_2$ oligonucleotide.

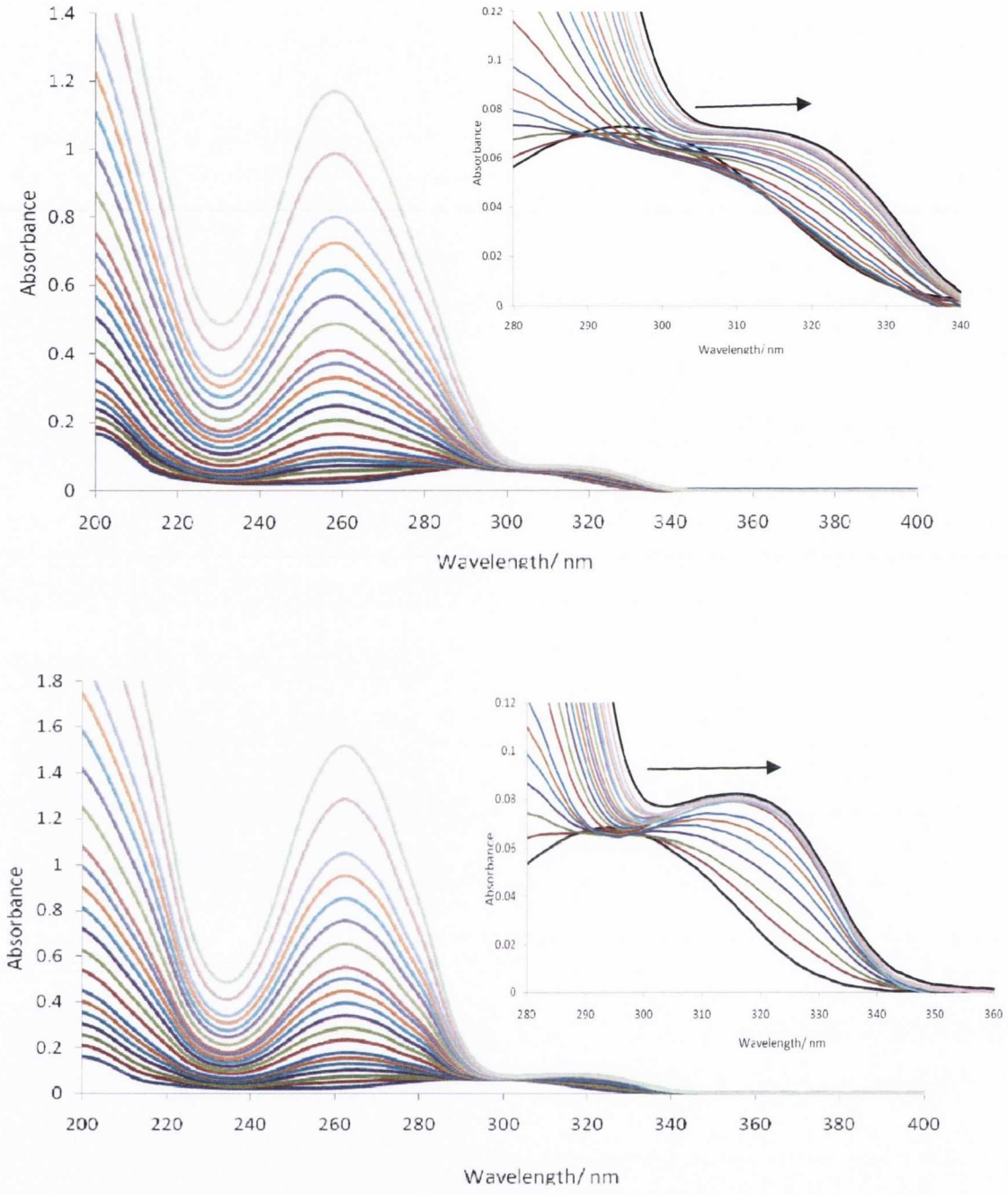


Figure 4.10. UV titration results with compound ***5h*** *(concentration = 1.6 μM) using natural DNA (above), and poly(dA-dT)$_2$ (below) working from a Bp/D 0-15*

Since the changes in the spectra were not clear due to the intense DNA band at 260 nm, it was decided to subtract the spectrum of the molecule from all other spectra (Figure 4.11). Thus, we could observe a DNA-molecule band forming at λ_{max} = 323 nm, confirming

the increasing concentrations of the DNA–ligand complex due to the strong binding of this molecule to DNA.

We then plotted saturation curves for the two titrations on the same graph for comparison purposes. Saturation curves are plots of absorbance against the Bp/D ratio and are used to indicate when the molecule ceases to bind. If we analyse the saturation curve for the titration with natural DNA, we observe large changes in the absorbance until a Bp/D ratio of 4, with slight changes occurring from a Bp/D ratio of 4 to 10 until saturation occurs indicating that binding has ended. If we compared the saturation plot for the titration with the AT polynucleotide we observe large increases in the curve until a Bp/D ratio of 2 is reached, at which point saturation occurs.

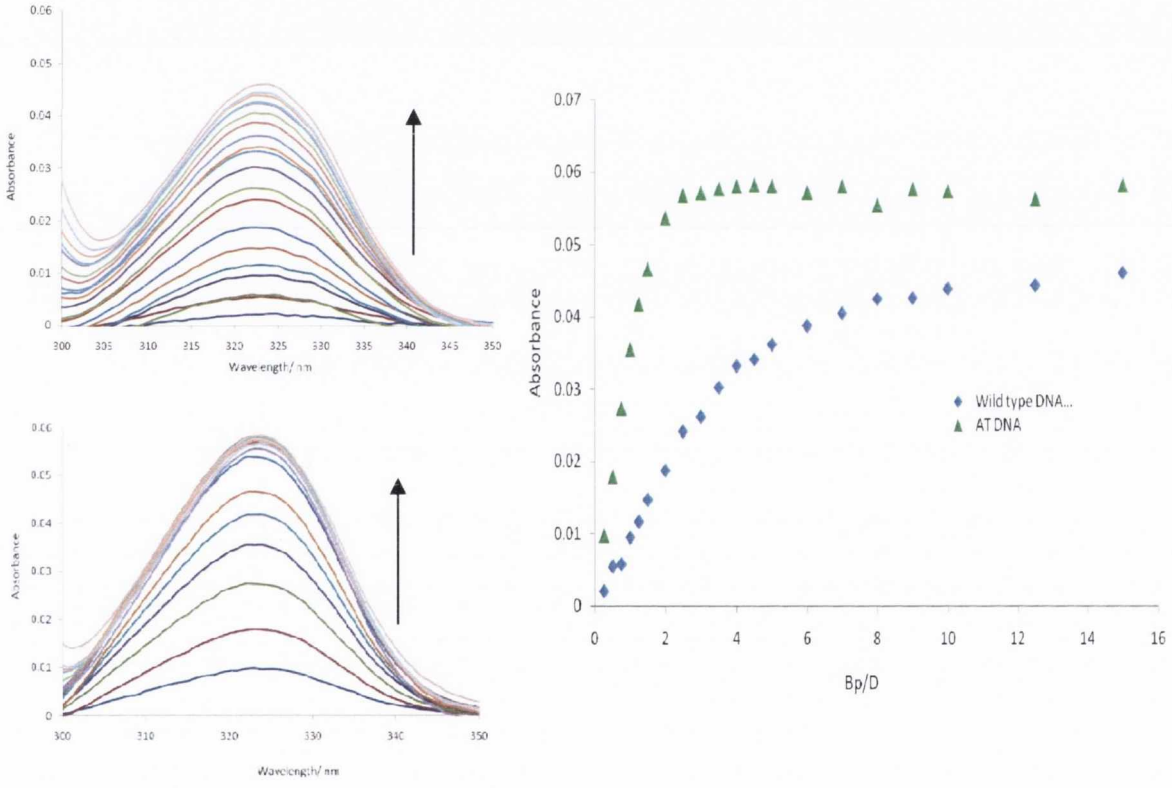


Figure 4.11. Subtraction plots for the titrations carried out for ***5h*** *with natural DNA (top) and poly(dA-dT)$_2$ and the corresponding saturation plots (right)*

From these experiments it was possible to calculate the corresponding binding constants of **5h** both to natural DNA and to the poly(dA-dT)$_2$ oligonucleotide and the results are presented in Table 4.4 confirming that **5h** binds more strongly to poly(dA-dT)$_2$ than to natural DNA.

*4.4.2.2. Results for compound **8h** (CO linker)*

The UV spectrum for **8h** (Figure 4.12) was recorded showing that the molecule absorbs strongly at 294 nm (π-π transition) with an ε = 17,170 $M^{-1}cm^{-1}$. Like with the previous molecule we carried out a UV titration experiment to see if a shift of the absorption band of the same magnitude than for **5h** was observed.

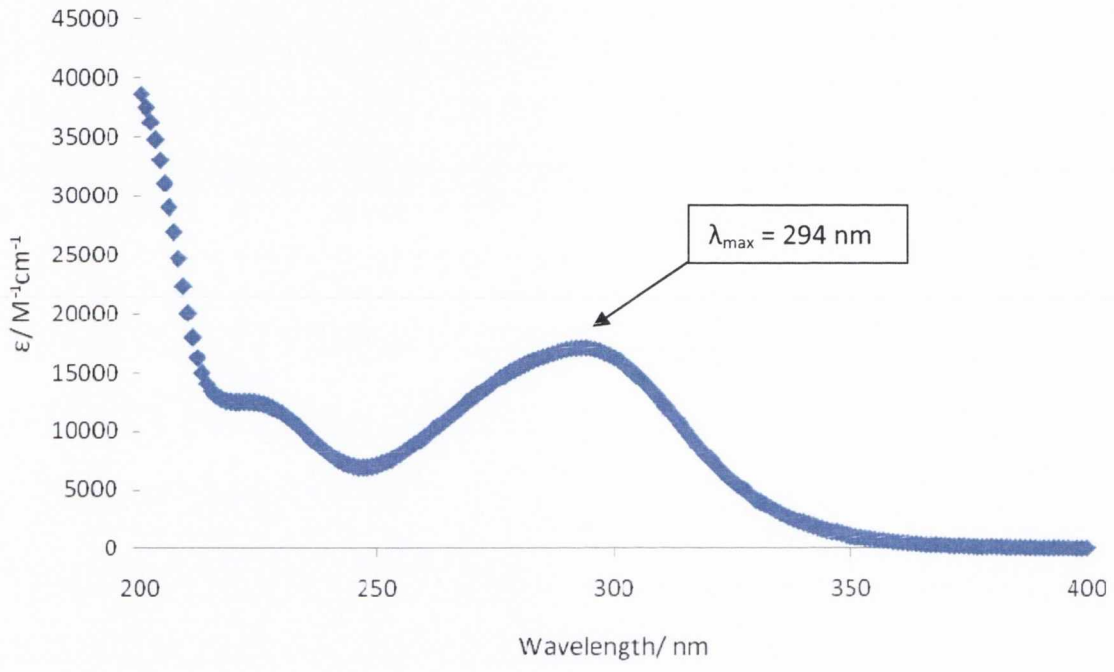


*Figure 4.12. UV spectrum of **8h***

Thus, a solution of the compound of specific concentration (1.6 μM) was prepared and the corresponding absorption spectrum was recorded. Aliquots of natural DNA solution (2 mM) were added and the corresponding absorption spectra recorded until saturation was observed. Results are presented in Figure 4.13 showing a small hypochromic shift that indicates the binding of this molecule to DNA. This was followed by a slight bathochromic shift that is also indicative of a new species in solution which in this case is the DNA-ligand complex. However, in comparison to the titration with **5h**, only small spectral changes were observed due to the weaker binding of this molecule, which correlates well with the thermal denaturation assay results. Consequently, a binding constant could not be evaluated.

To verify further that this molecule displays selectivity for AT sequences, the UV titration experiment was repeated with poly(dA-dT)$_2$ DNA. We observed a larger

hypochromic shift accompanied with a bathochromic shift indicating the formation of the DNA-molecule complex. The fact that we found larger shifts in the presence of AT sequences is proof that this molecule also displays AT sequence selectivity.

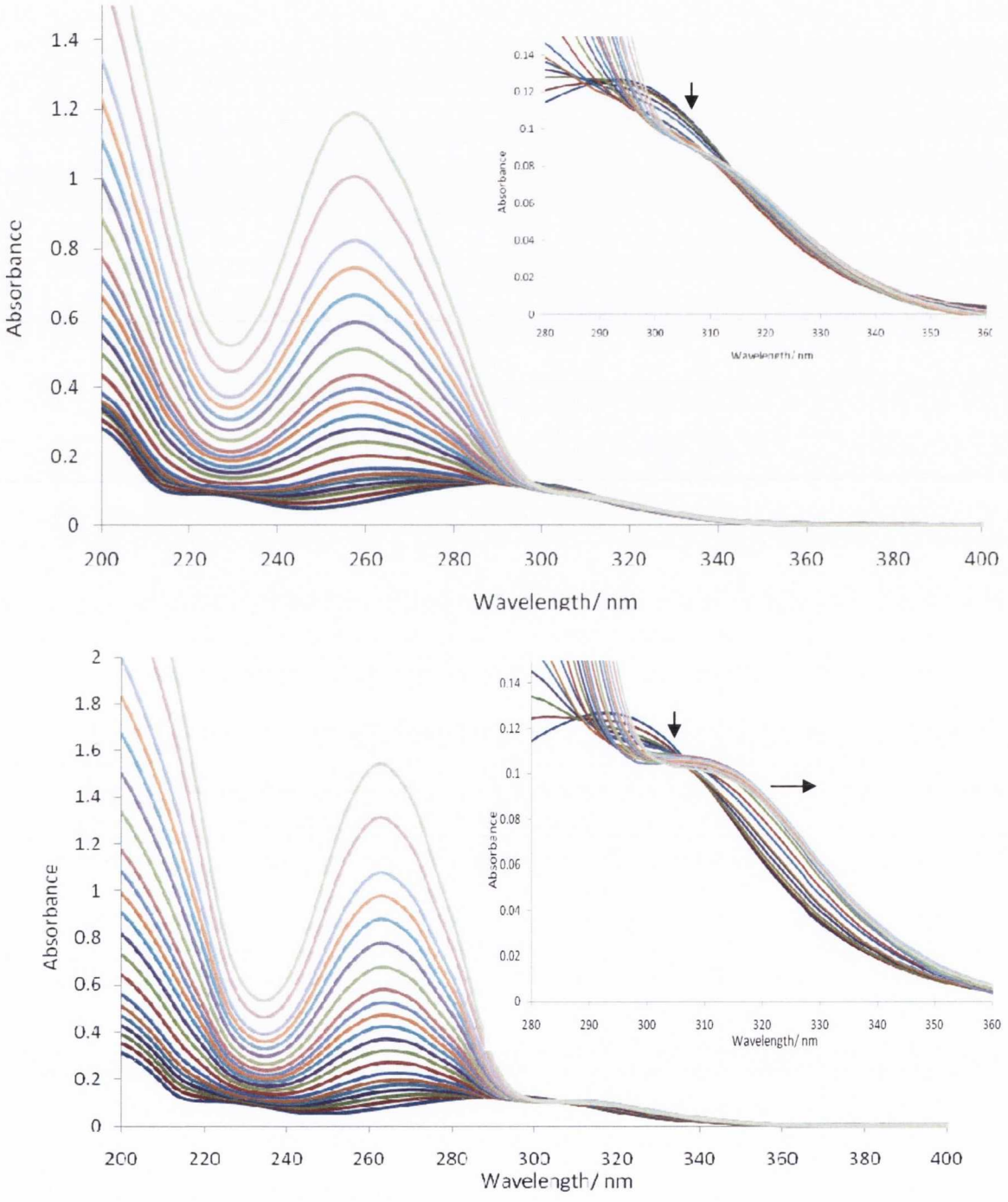


Figure 4.13. UV titrations for ***8h****. Spectral changes observed upon addition of natural DNA (upper) and poly(dA-dT)$_2$ (lower)*

As in the previous example, we subtracted the spectrum of the molecule from all the other spectra (Figure 4.14) allowing to observe a AT-DNA-molecule band forming at λ_{max} = 323 nm. This shows that by increasing the AT DNA concentration, the formation of the AT-DNA-ligand complex results. The corresponding graph plotting absorbance against Bp/D ratio, shows that the concentration of the AT DNA-ligand complex increased as more AT DNA solution is added until a Bp/D ratio of 3.5 is reached. At this point, saturation occurs and no more binding takes place.

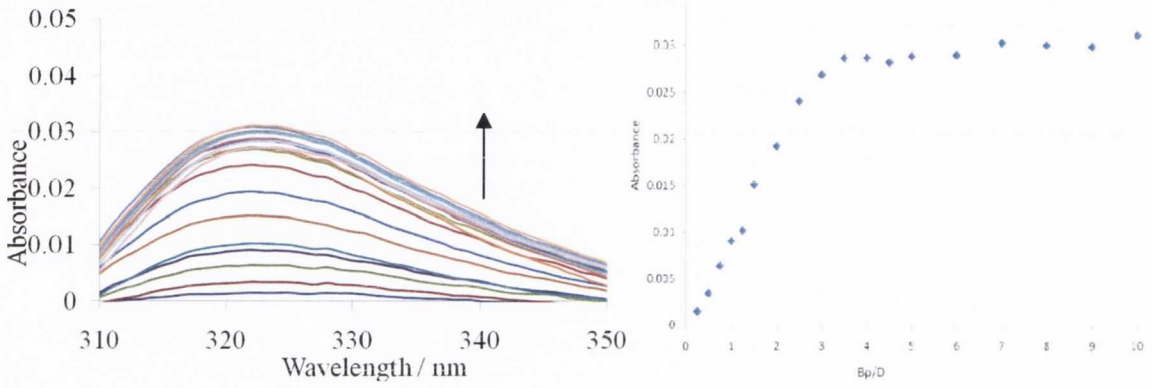


*Figure 4.14.Subtraction plot for the tiration of **8h** with AT DNA (left) and the saturation plot (right)*

In this case, it was possible to calculate the corresponding binding constants of **8h** to the poly(dA-dT)$_2$ oligonucleotide and the result is presented in Table 4.4.

*4.4.2.3. Results for compound **11h** (fluorene derivative)*

Like previously, the UV spectrum for **11h** was recorded (Figure 4.15), showing that the molecule absorbs at 211 nm and 273 nm, which corresponds to the π-π transitions. Interestingly, the peak at 273 nm is accompanied by a shoulder, which was far enough from the DNA absorption region, and a UV titration experiment was attempted.

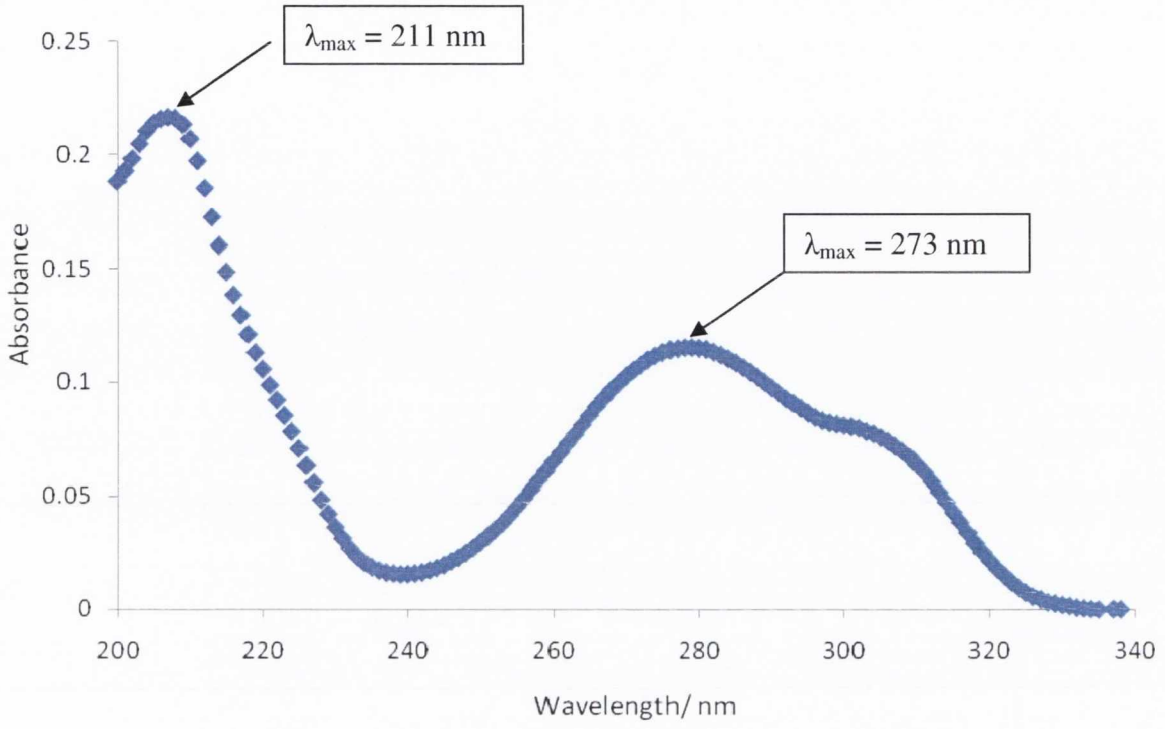


Figure 4.15. UV absorption of compound ***11h***

Subsequently, aliquots of natural DNA (2 mM) were added to the compound solution (1.67 μM) and the corresponding UV spectra were recorded to give the overall titration seen in Figure 4.16. As DNA was added, small hypochromic and bathochromic shifts were observed indicating that binding was occurring. The same experiment was then carried out with poly(dA-dT)$_2$ and larger and more pronounced spectral changes than with natural DNA were observed. A very large bathochromic shift was noticed with the formation of a new band at 320 nm.

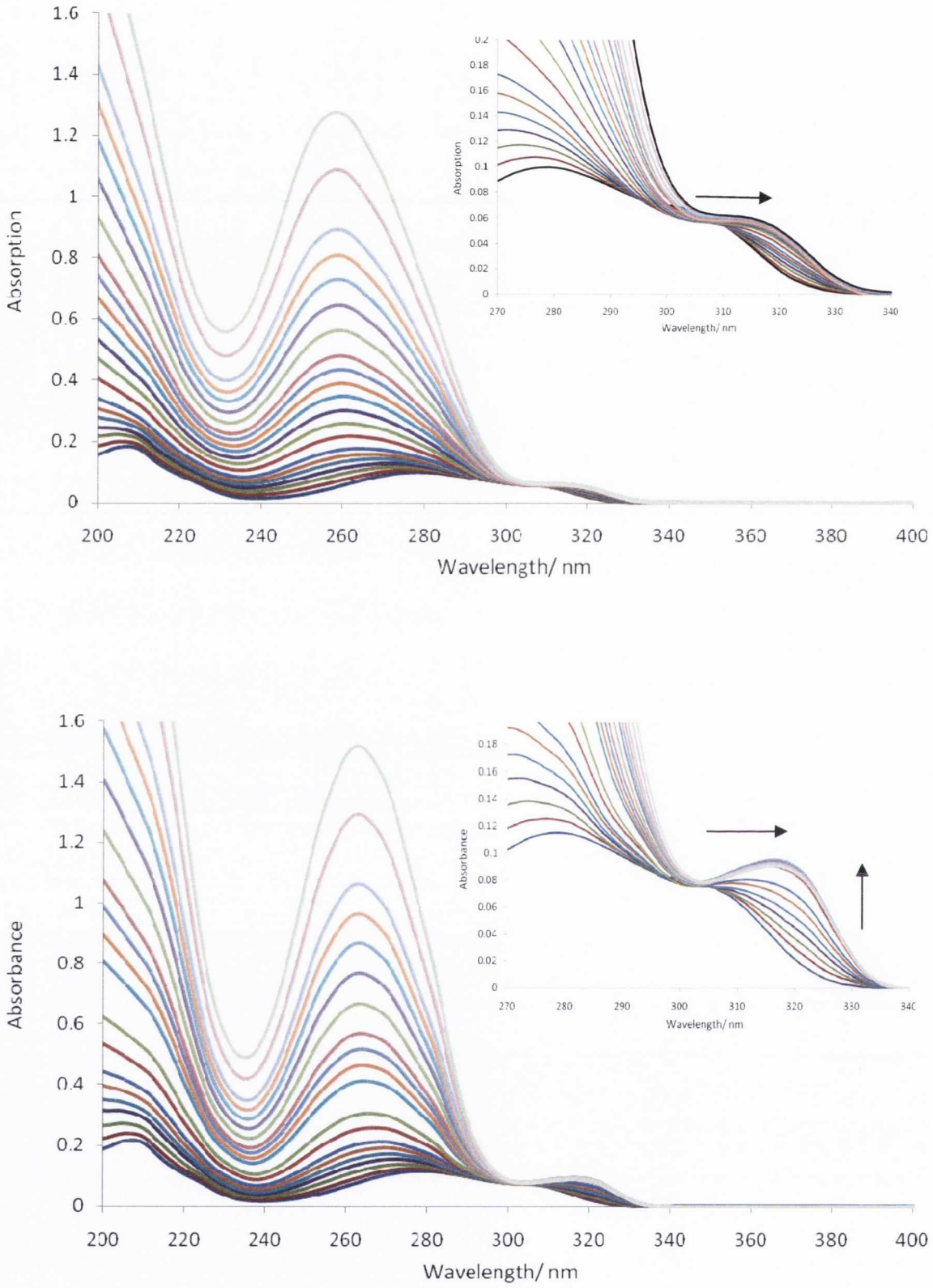


Figure 4.16. UV titration experiments for ***11h*** *in the presence of natural DNA (top) and poly(dA-dT)$_2$ (bottom)*

Since the changes in the spectra were not clear in the case of the experiment with natural DNA, due to the intense DNA band, it was decided to subtract the spectrum of the molecule from all the other spectra and then we could observe the formation of a DNA-ligand band at λ_{max} = 323 nm in both cases. This shows that through the addition of DNA, not only the DNA concentration was being increased, but also the concentration of the DNA–ligand complex. From the saturation curves shown in Figure 4.17, we concluded that when DNA was added to the compound solution, this bound to DNA. This equilibrium was faster in the case of AT sequences than with wild type DNA, showing that **11h** binds preferably to AT sequences rather than to wild type DNA. In the case of the titration with AT DNA, equilibrium was reached at a Bp/D ratio of 2. As more DNA was added beyond this point, no more binding occurred since all the ligand molecules were already bound to the DNA minor groove. Comparing the experiments with AT sequences and wild type DNA, it can be seen that in the latter, equilibrium was not completely reached at a Bp/D ratio of 2 since beyond this point, there was some evidence of DNA binding. Also, the intensity of the absorbance for the titration with AT sequences was larger than for the titration with wild type DNA.

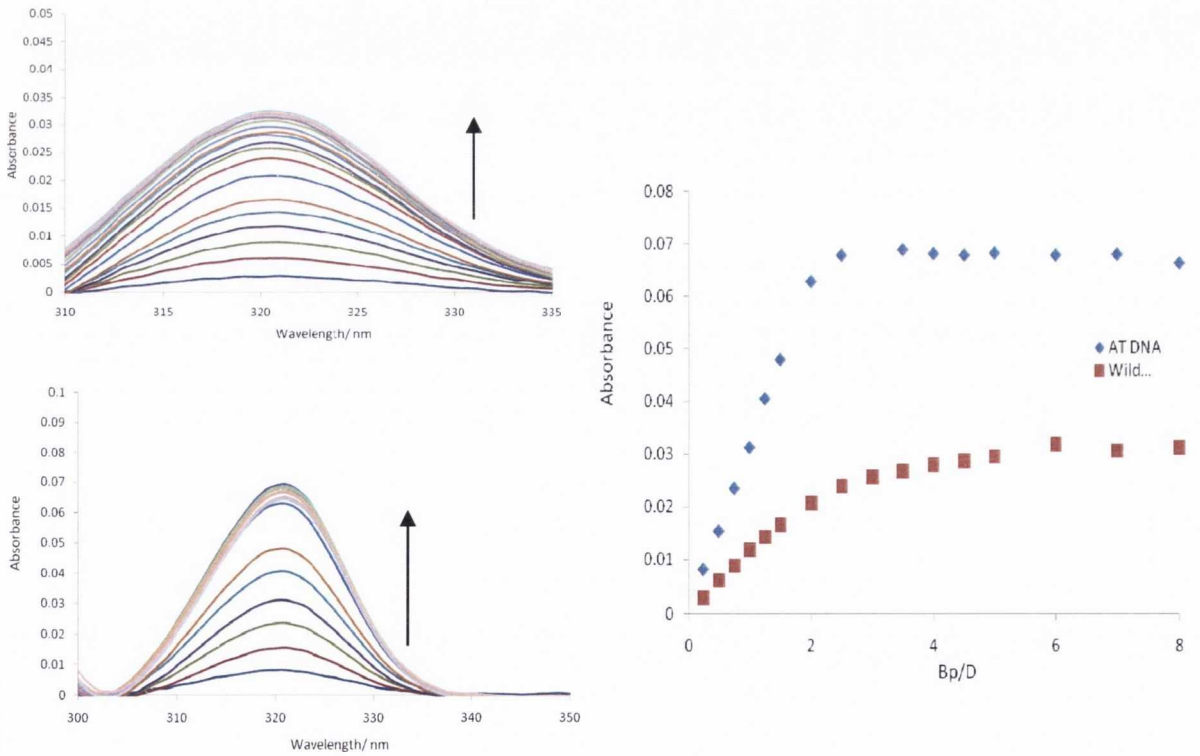


Figure 4.17. Subtraction plots for the UV titration of ***11h*** *with natural DNA (top left) and poly(dA-dT)$_2$ (bottom left) and the corresponding saturation plots (right)*

Both experiments with the natural DNA and the poly(dA-dT)$_2$ oligonucleotide allowed us to calculate the corresponding binding constants of **11h** and the results are presented in Table 4.4. By comparing the results obtained for the three compounds in poly(dA-dT)$_2$ it can be concluded that compound **11h** is the one that binds stronger to this sequence specific DNA and the trend observed is in agreement with the thermal denaturation results obtained.

*Table 4.4.- Binding constants calculated from scatchard plot analysis for compounds **5h**, **8h** and **11h** with natural and AT DNA*

	Binding Constant (K) ($\times 10^5$ M^{-1})		
	5h	**8h**	**11h**
Natural DNA	0.32 ± 0.01	-	0.12±0.01
poly(dA-dT)$_2$	2.5 ± 0.2	0.64 ±0.04	14.6±0.3

4.4.3 Studying the dependence of increasing the ionic strength on the binding affinity

The reverse salt titration has found many applications to this day and consists of increasing the ionic strength of the DNA-ligand solution through the addition of NaCl. Subsequently, the competition for the binding sites between the ligand and the cation from the salt increases and as a consequence the bound ligand is displaced upon rising salt concentration. This is detected in the UV spectra when a decrease in the DNA-ligand band is observed (Figure 4. 18). One application for this technique is to distinguish between binding modes (intercalation *vs.* minor groove binding) since the minor groove binder can be easily displaced from the groove upon increasing ionic strength whereas the intercalator is not easily displaced since it would be inserted in between the base pairs and would not be affected by the increasing salt concentration as much. Figure 4.18 shows that upon increasing the ionic strength by increasing the salt concentration, large spectral alterations were observed for the UV spectrum of compound **5h** with poly(dA-dT)$_2$. This indicates that significant quantities of the ligand were being displaced from the helix demonstrating that this molecule binds to the minor groove.

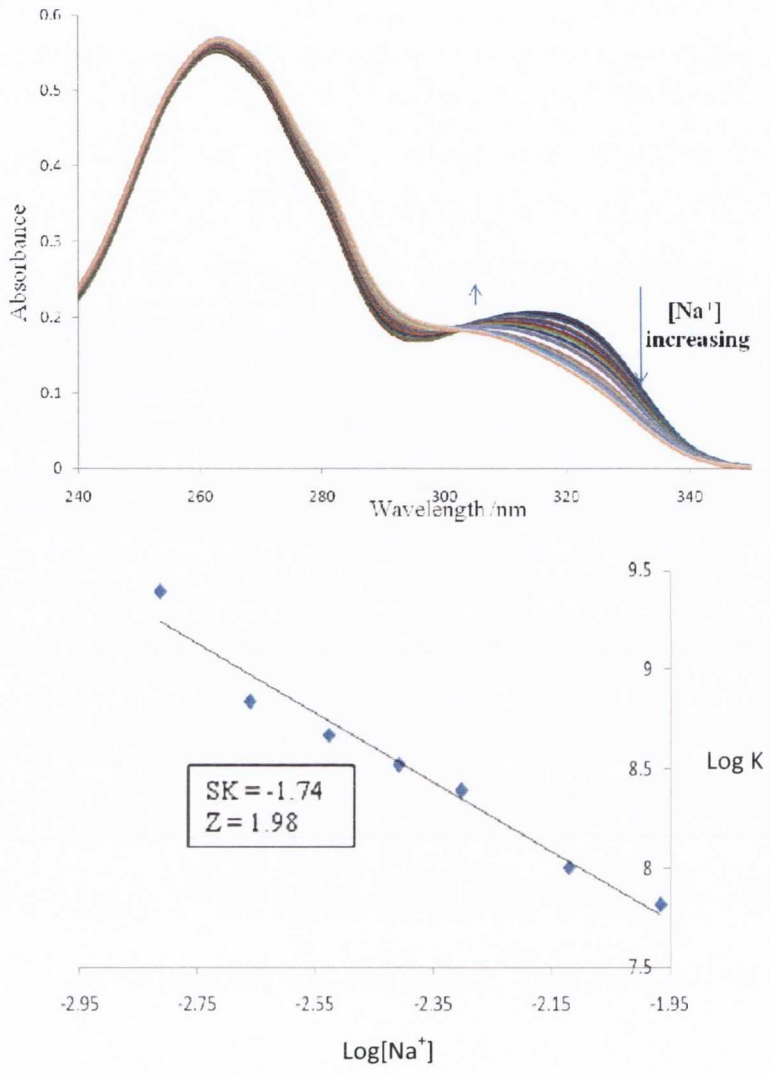


*Figure 4.18. UV-Vis absorption spectrum for the reverse salt titration of compound **5h** (up), the dependence of the binding constant on the salt concentrations is shown on the graph down*

Since the binding constants were evaluated from the UV titration experiments, it was possible to calculate the free energy term from equation (1). According to the polyelectrolyte theory developed by Record *et al.*,[13] the observed binding constant is a function of the charge on the cation and the fraction of the counterion associated to each DNA phosphate. The dependence for the binding of **5h** was investigated by plotting logK against log[Na^+] and the slope of the graph (SK) is calculated according to equation (2) where Z is the charge on the ligand and ψ is the charge on the phosphate:[15, 16, 17]

$$\Delta G^0 = -RT\ln K \qquad (1)$$

$$SK = \delta \log K / \delta \log[Na^+] = -Z\psi \qquad (2)$$

The slope of the graph (SK) was found to be -1.74 and this was used to partition the free energy term that was obtained from the UV experiment described in section 4.4.2.1. This observed free energy (ΔG_{OBS}) was calculated to be -28.9 kJmol^{-1}and then it was split into the electrostatic (ΔG_{PE}) and the non-electrostatic ($\Delta G_{NON\text{-}PE}$) terms by using equations (3) and (4). The corresponding results are shown in Table 4.7.

$$\Delta G_{PE} = SKRTIn[Na^+] \qquad (3)$$

$$\Delta G_{OBS} = \Delta G_{NON\text{-}PE} - \Delta G_{PE} \qquad (4)$$

Table 4.5. Free energies from the reverse salt titration of compound ***5h***

	K_{AT} $\times 10^5 M^{-1}$	ΔG_T (kJmol^{-1})	ΔG_{PE} (kJmol^{-1})	$\Delta G_{NON\text{-}PE}$ (kJmol^{-1})
poly(dA-dT)$_2$	2.5	*-28.9*	*-19.9*	*-9.0*

Table 4.5 shows that the electrostatic term is over two fold larger than the non electrostatic one indicating that binding is more dependent on the electrostatic interactions than on the non electrostatic ones. This shows that the cationic functionalities are essential for DNA binding which is in direct correlation with the thermal denaturation experiment results.

4.5. Circular Dichroism Studies

4.5.1. Introduction

Circular dichroism (CD) is a spectroscopic technique based on the difference in absorption of left and right handed circularly polarised light by a chiral molecule.[19] CD has many applications, such as the study of chiral molecules, as each enantiomer will interact with different amounts of right handed and left handed light thus resulting in a CD spectrum. If the molecule is achiral, no spectrum will be observed since it would interact with the same amount of left handed and right handed photons.

Additionally, CD can be used to measure the interaction between small molecules and macromolecules such as proteins or DNA. It can also be used to measure the interaction between DNA/RNA and DNA/protein interactions which cannot be evaluated with UV spectroscopy.

Single stranded DNA is in fact achiral and produces little or no spectrum; however, when two strands come together to form duplex DNA, this is chiral due to the rearrangement of the DNA bases inside the helix. When DNA is in the B-form, the CD spectrum has a positive signal at 275 nm and a negative one at 245 nm, passing zero at 258 nm.

The main uses of CD involving DNA, are to probe DNA conformational changes and its interaction with small molecules.[19,20] By adding small quatities of our achiral drugs, we can monitor spectral changes which are indicative of structural modifications occurring to the double helix. The higher the binding affinity of the molecule, the more structural changes will occur to the DNA. Subsequently, this will be accompanied by large spectral changes. Through this, we can deduce the binding mode and the binding affinity of the molecules, by titrating them into a DNA solution and observing the spectral shifts.

4.5.2. Circular Dichroism studies on the asymmetric di-functionalised molecules

A CD titration was carried on **5h** (NH linker) with poly(dA-dT)$_2$. First, the CD spectrum of the AT DNA was recorded. Following this, aliquots of the small molecule in a 0.5 mM solution were added and the corresponding spectra recorded as shown in Figure 4.19. The spectra show that structural changes occur in the DNA due to the binding of **5h**. We can observe the formation of an induced circular dichroism (ICD) peak at 323 nm. The magnitude of the changes is indicative of the molecule binding to the minor groove instead of intercalating between the DNA base pairs. Thus, by plotting the concentration of the ligand against the spectral rises in the ICD peak, we obtained a saturation curve, and this ICD peak illustrates that the dication was binding to the chiral environment of the DNA helix. Interestingly, by comparing the changes in the ICD peak with the changes in the DNA environment (represented from the changes in the DNA spectrum at 265 nm), we can see that the binding of the drug changes the DNA structure, but it does not induce a large structural change to the DNA backbone. Other information that could be obtained from these plots is

the fact that that binding was occurring until 5 μM was added. Beyond this point no spectral changes were observed upon addition of the drug, indicating that saturation had occurred. Moreover, there is no evidence of a bathochromic shift which together with the presence of an isosbestic point, indicates that there is only one form of binding. Also the magnitude of the ICD peak is consistant with minor groove binding since a large peak is indicative of minor groove binding whereas a small peak is indicative of intercalation since the former has more interaction with the chiral sugars.

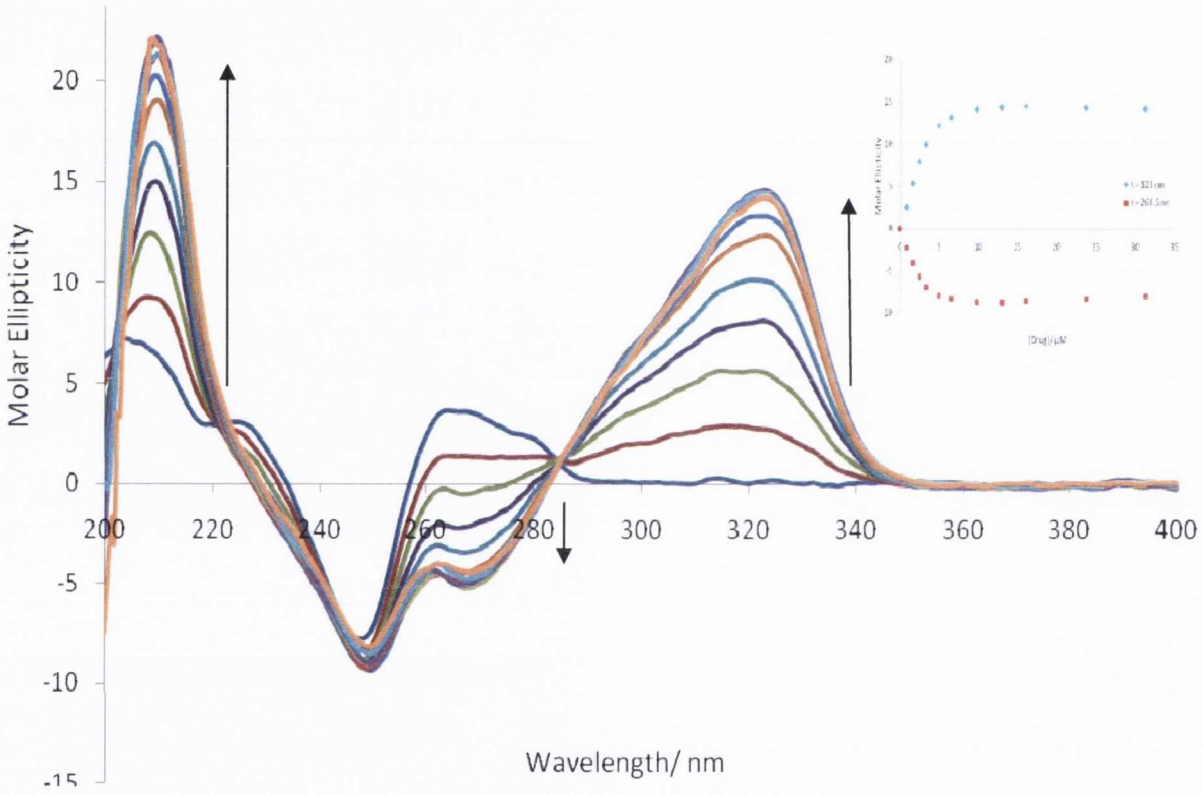


Figure 4.19. CD titration spectra obtained for ***5h*** *produced upon addition of DNA in phosphate buffer. Arrows correspond to increasing drug concentration*

The same experiment was carried out for compound **14a** (CONH linker) and the results are shown in Figure 4.20. In these spectra, we observe the formation of a large ICD peak at 300 nm which is consistent with minor groove binding. Similarly, we can see the decrease in the peak at 265 nm. Both of these spectral changes indicate the structural changes occurring to DNA due to the binding of this dication. Also, the fact that there is no

observation of any bathochromic shift and the presence of an isosbestic point tells us that only one species is formed in solution and that there is only one binding mode associated with this ligand.

By plotting the concentration of compound **14a** against the spectral rises in the ICD peak, a saturation curve is obtained. This ICD peak illustrate that the ligand was binding to the chiral environment of the DNA helix. Again, comparing the changes in the ICD peak with those in the DNA environment (changes in the DNA spectrum at 265 nm), we observe that the dication changes the DNA structure, but does not induce a large structural change to the DNA backbone. Binding occurs until 7.5 μM was added and beyond this point no spectral changes were seen.

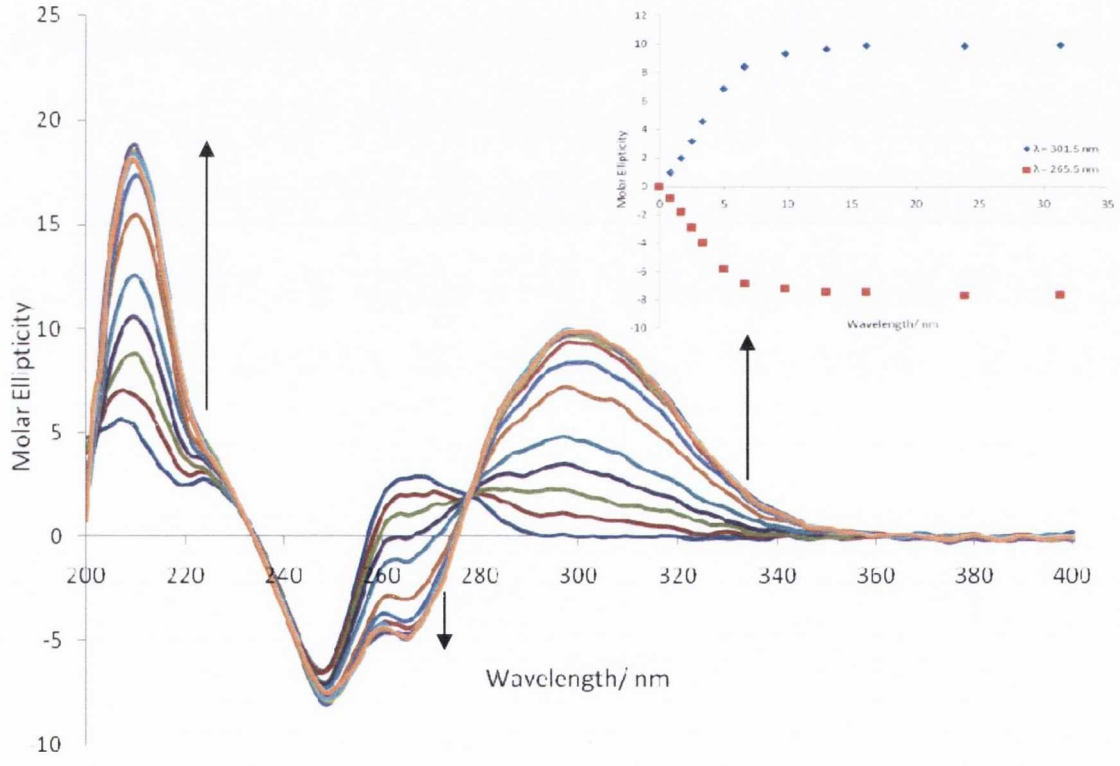


*Figure 4.20. CD titration spectra obtained for **14a** produced upon addition of DNA*

CD titrations were also performed in the presence of **8h** (CO linker) and similar trends were observed in comparison to the previous two experiments (Figure 4.21). We found a strong ICD signal which is indicative of minor groove binding. As aliquots of the ligand were added to the DNA solution, spectral changes indicating structural changes occurring to DNA were observed.

As aliquots of a 0.5 mM solution of **8h** were added, we could see a decrease in the peak at 270 nm indicating DNA binding, and an isosbestic point at 288 nm indicating the formation of a single species and, therefore, only one mode of binding.

By plotting the concentration of the dication against the spectral rises in the ICD peak, we obtained the corresponding saturation curve. Once again, the changes in the ICD peak compared to those in the DNA environment (changes in the DNA spectrum at 265 nm), show that the binding of the drug changes the DNA structure but not the DNA backbone in the same magnitude. Binding occurs until only 10 μM was added.

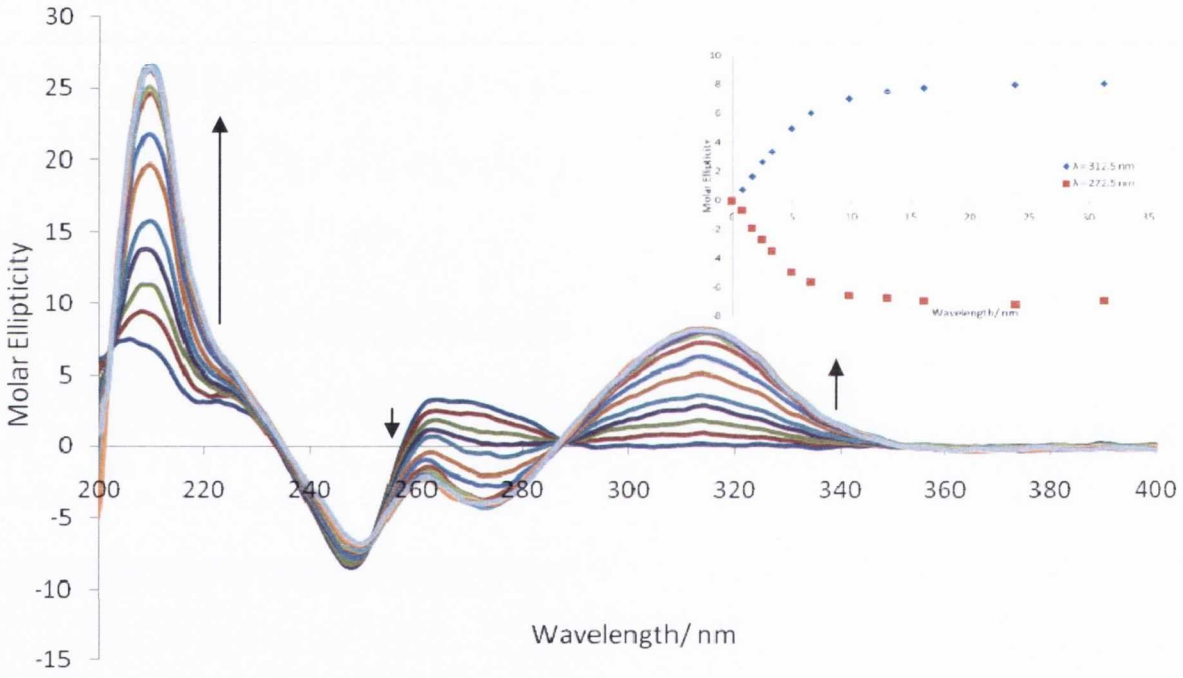


*Figure 4.21. CD titration spectra obtained for **8h** produced upon addition of AT DNA*

These CD experiments in AT DNA were also performed for compounds **3h**, **4h**, **9h** and **10h** and the results are presented in Figure 4.22, showing very similar pattern to those already discussed, that is minor groove binding was observed in all cases.

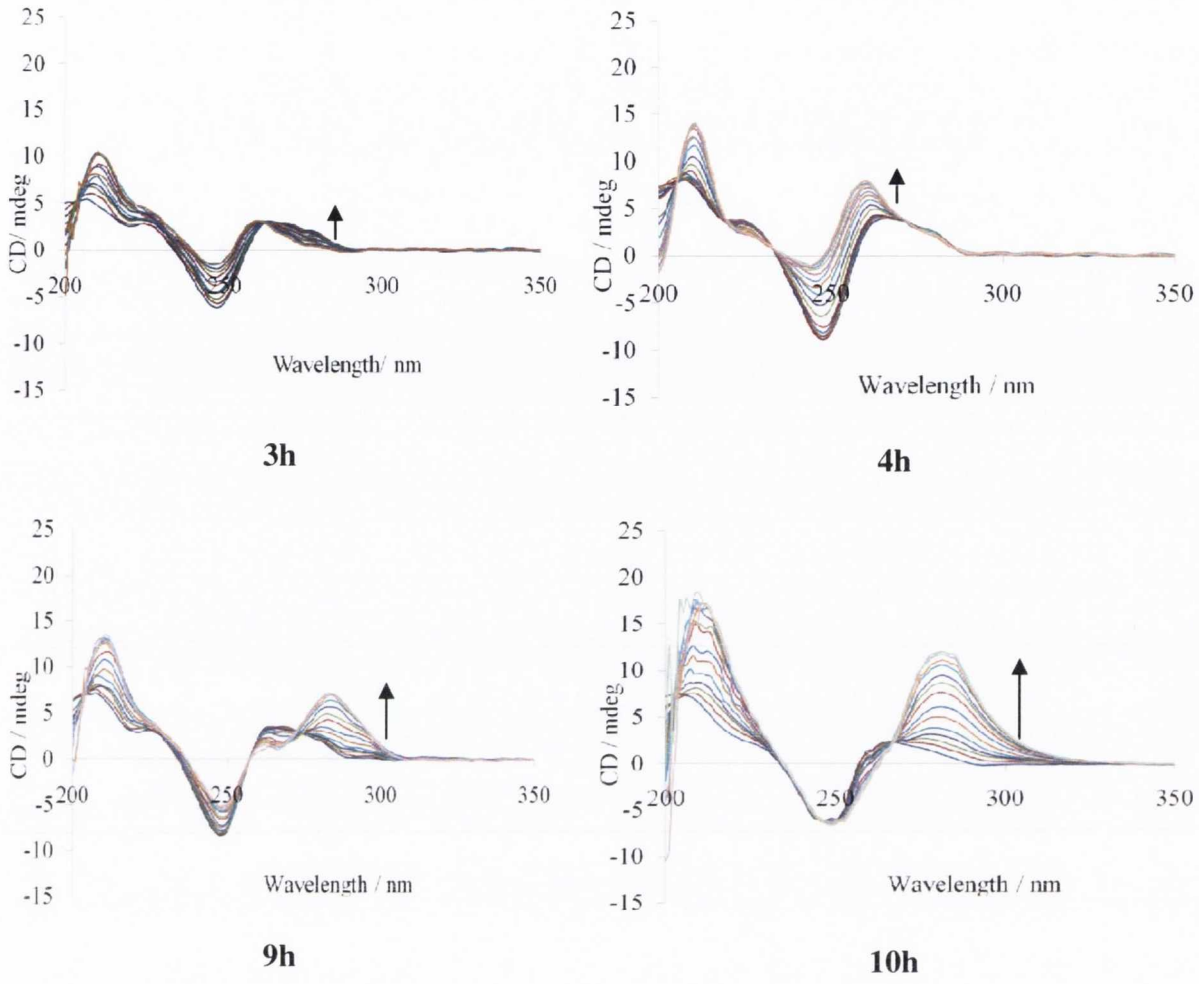


Figure 4.22. CD titration spectra for ***3h***, ***4h***, ***9h*** *and* ***10h*** *with poly(dA-dT)*$_2$

CD was also used to yield quantitative information on the binding of these molecules. By using the intrinsic method and Scatchard plot analysis[18], we were able to evaluate the binding constants of our family of compounds and the binding site size. The results for all the compounds studied are shown in Table 4.6.

Table 4.6.- Binding constants determined by circular dichroism using poly(dA-dT)$_2$

Compound No.	Linker	K $\times 10^5 M^{-1}$
3h	CH_2	0.1
4h	CH_2CH_2	0.2
5h	NH	1.2
8h	CO	0.3
9h	NHCONH	0.4
10h	Piperazine	1.0
14a	Amide	16.0

The binding constants obtained from the CD experiments illustrate two important aspects of these molecules:

- The molecules with CONH and NH linkers have the highest AT DNA binding affinity while the molecule with the CH_2 linker displays the worst binding verifying the thermal denauration results
- Geometry of the linker and the number of hydrogen bonding groups may be an important factor in AT DNA binding.

4.6 Electric Flow Linear Dichroism

4.6.1. Introduction

Linear Dichroism (LD) is a powerful method for the determination of the binding mode of small molecules to DNA. It is defined as the difference in the absorption of linearly polarised light ($A_{parallel}$) and perpendicularly polarised light ($A_{perpendicular}$) according to the following equation

$$LD = A_{parallel} - A_{perpendicular} = 3/2S(3Cos^2\alpha - 1) \quad (5)$$

where S is the orientation factor (S = 1 if the DNA is aligned whereas S = 0 if the DNA is not aligned), and α is the angle that represents the orientation of the of the transition moment that is responsible for the absorption of light.[19]

In this experiment, the DNA is spun around so that the bases are perpendicular with respect to the incoming light. Consequently, at 260 nm the bases absorb perpendicular light and subsequently, we observe a large negative peak at 260 nm. If a molecule intercalates between the base pairs the intercalating molecule would be parallel with respect to the bases therefore the angle α would be equal to 90°. If we solve equation (5) when α = 90°, we get a negative LD signal. If a molecule binds to the DNA minor groove, then α would be equal to 45°. By calculating the LD signal, we would expect a positive induced signal. Therefore, LD experiments are very informative of the binding mode of the ligand.

An advantage of this technique is that we can evaluate the binding mode of a molecule irrespective of where it absorbs. For example, if the molecule absorbed strongly in the DNA region, and is a known minor groove binder, we would observe an increase in the LD signal at that range.

4.6.2. Linear Dichroism studies on some asymmetric dications

Thus, LD spectra were recorded for natural DNA (salmon testes) titrated with compounds **5h** and **8h** in phosphate buffer at 25 °C. Titrations were carried out with a DNA (blue line) concentration of 378.8 μM working with a P/D ratio of 0, 2 and 10, varying the P/D ratio from 10 to 2 over 2 additions (Figure 4.23).

In these experiments, the DNA and compound concentration was kept constant to avoid dilution effects. The reason for this is that we are dealing with very small quantities (70 μL). Consequently, each solution was prepared individually and the corresponding spectra were recorded.

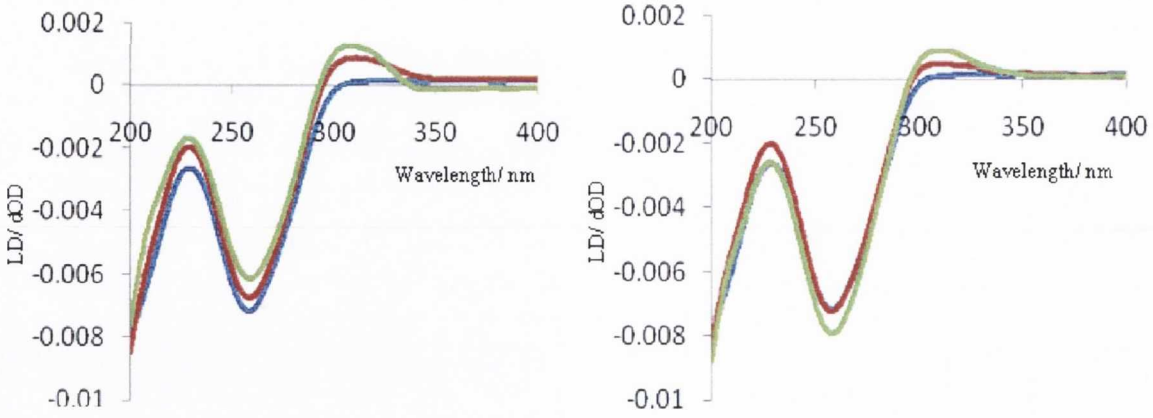


*Figure 4.23. LD spectra for natural DNA with **5h** (left) and **8h** (right)*

Both compounds showed a notable increase in the absorbance at 310 nm where the DNA does not absorb. Positive induced signals are indicative of minor groove binding and, hence, these LD results with unspecific DNA indicate that the asymmetric compounds **5h** and **8h** (NH and CO linkers respectively), were found to bind in the minor groove.

4.8 Conclusions

This chapter presents many of the physicochemical studies carried out for the binding to DNA (both random sequence and poly(dA-dT)$_2$) of the molecules whose synthesis has been previously described in chapter 3. The experiments presented here are based on optical techniques that depend on the use of the UV-Vis spectrometer, CD and LD.

The first experiment described was the determination of pK_a. This was carried out to assess the protonation state of the molecule at physiological pH. It was found that in the *mono*-functionalised molecules, the amino groups have pK_a values less than that of physiological pH, whereas the highly basic guanidine and 2-aminoimidazoline functionalities were found to be protonated at physiological pH. Thus, for the mono functionalised molecules whose pK_a was measured, we found them to be *mono*-cationic whereas the di-functionalised molecules were dicationic.

Following from these experiments, the DNA binding affinity of these molecules to natural DNA was investigated by using thermal denaturation experiments. The *mono*-functionalised molecules displayed poor affinity to natural DNA. In contrast, the di-functionalised molecules displayed high affinity for natural DNA. It was observed from these experiments that the binding was dependent on the linker, which governed many of their properties such as geometry, length and the number of hydrogen bonding groups on the molecule.

Using, also thermal denaturation techniques, the next target was to investigate the sequence selectivity of the dicationic molecules. It was observed that all of them, apart from **3h** (linker = CH_2), displayed selectivity for AT sequences, which is common for minor groove binders. An interesting case was that of compound **11h** since it displayed the highest selectivity for AT sequences over natural DNA.

In order to obtain more information about the binding of these dications such as the equilibrium binding constants, UV titration experiments were carried out to evaluate these quantities by using Scatchard plot analysis. These results were found to be in agreement with the thermal denaturation assay results. Then, a reverse salt titration was performed to confirm that the compound was dicationic on DNA binding. From these experiments the free energy could be divided into an electrostatic and a non electrostatic term. From these results we observed that the binding was dominated by the electrostatic interactions between the molecule and the DNA backbone.

Next, we carried out circular dichroism experiments to obtain numerical values to characterise the binding to AT sequences. Large induced signals were observed upon binding of these molecules confirming that they are minor groove binders.

Lastly, linear dichroism was carried out. The purpose of this experiment was to confirm the mode of binding to DNA. In this technique, a positive induced signal is indicative of minor groove binding whereas a negative induced signal would be indicative of intercalation. From these experiments, we observed positive induced signals confirming that both **5h** and **8h** molecules bind to the minor groove of DNA.

Overall, much information has been obtained from these optical techniques. The following chapter will focus on the non optical techniques concerning the thermodynamics of the binding and surface plasmon resonance results.

4.9 References

1. Yoon, C.; Privt, G. G.; Goodsell, D. S.; Dickerson, R. E.; *Proc. Nati. Acad. Sci. USA,* **1988,** *85,*6332
2. Soler-Lopez, M.; Malinina, L.; Liu, J.; Huynh-Dinh, T.; Subirana, J.; *J. Bio. Chem.* **1999,** *274,* 23683
3. Drew, H. R.; Dickerson, R. E.; *J. Mol. Biol.* **1981,** *151,* 535
4. Neidle, S.; *Nat. Prod. Rep.* **2001**, *18*, 291
5. Ismail, M.A.; Arafa, RK; Brun, R; Wenzler, T; Miao, Y; Wilson, WD; Generaux, C; Bridges, A; Hall, JE; Boykin, D. W.; *J. Med. Chem.* **2006**, *49*, 5324
6. Nguyen, B.; Lee, M. P. H.; Hamelberg, D.; Joubert, A.; Bailly, C.; Brun, R.; Neidle, S.; Wilson, W. D.; *J. Am. Chem. Soc.,* **2002**, *124*, 13680
7. Hanlon, S., Wong, L., Pack, G. R., *Biophys. Journal,* **1996**, *72*, 291
8. http://ifs.massey.ac.nz/outreach/resources/chem/orgbases.php
9. Kinsella, G.; Rodriguez, F.; Watson, G.; Rozas, I.; *Bioorg. Med. Chem.,* **2007,** *15,* 2850
10. *http://www.biochem.arizona.edu/classes/bioc461/GRAPHICS/Chapter5-1/Slide38.JPG*
11. *http://www.siumed.edu/~bbartholomew/images/chapter5/F05-14.jpg*
12. Nagle, P. S., Rodriguez, F., Kahvedzic, A., Quinn, S. J., Rozas, I., *J. Med. Chem.,* **2009,** *52,* 7113
13. Record, M. T., Anderson, C. F., Lohman, T. M., *Q. Rev. Biophys.,* **1978**, *2*, 103
14. Han, F.; Taulier, N.; Chalikian, T. V.; *Biochemistry*, **2005**, *44*, 9785
15. Misra, V. A.; Sharp, K. A.; Friedman, R. A.; Honig, B.; **1994**, *238*, 245
16. Haq, I.; Ladbury, J. E.; Chowdhry, B. Z.; Jenkins, T. C.; Chaires, J. B.; *J. Mol. Biol.,* **1997**, *271*, 244
17. Jenkins, T. C.; *Methods in Molecular Biology,*Vol.90, Humana Press Inc. Totowa, NJ
18. Circular and Linear Dichroism, Alison Roger and Bengt Norden, *Oxford University Press,* 2008
19. Norden, B., Kubista, M., and Kurucsev, T., *Q. Rev. Biophys.* **1992**, *25*, 51

Chapter 5

Asymmetric Dications-DNA Interactions: Non-Optical Physicochemical Techniques

5.1. Introduction

In the previous chapter, optical techniques were discussed in measuring the DNA binding affinity to natural DNA and to AT polynucleotides. The next target was to investigate the DNA binding affinity to specific sequences of DNA using non optical techniques such as Surface Plasmon Resonance (SPR)[1,2] or Isothermal Titration Calorimetry (ITC).[3]

SPR is a technique where biomolecules such as DNA are immobilised onto one side of a metallic chip and then light is focused on the other side to excite the surface plasmons. Upon binding of a ligand there is a change in the refractive index that can be measured.[3] This is a technique used to measure DNA binding affinities to specific sequences of DNA such as d(CGAATTCGTCTCCGAATTCG) or d(CGATATCGTCTCCGATATCG). The purpose of these investigations is to observe the selectivity factor among our compounds which can be defined as the difference in the binding affinity for a molecule between two different sequences. It is generally desirable to prepare compounds that can recognize specific sequences to enable us to target a specific part of the genome.

Another non-optical technique that was used to characterise the DNA binding of our molecules is Isothermal Titration Calorimetry (ITC).[4] This method is based on the difference in heat supplied between the control and the experiment and allows the determination of the thermodynamics of binding such as the changes in the enthalpy, entropy and Gibbs free energy. From these parameters we can deduce thermodynamic aspects about the binding of our molecules to the DNA macromolecule, for example, a positive entropy change means that more disorder was introduced to the system as a result of the binding, and a more negative enthalpic change would signal the formation of more hydrogen bonds between the molecule and the DNA bases.

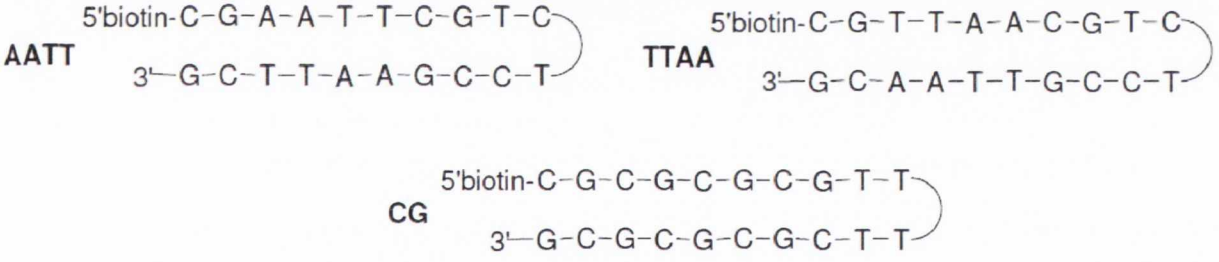
Finally, a crystal structure was obtained for one of our molecules. This structure was the dihydrochloride salt from the mono guanidine family (linker = CH_2CH_2). The importance of this compound is that it tells us about the potential hydrogen bonding of our molecules to the DNA base pairs. Unfortunately, this compound displayed poor binding to DNA; however, the potential hydrogen bonds that it could form upon binding will be discussed, and compared to the x-ray crystal structure of the DNA complex with a related *bis*-2-amino-imidazoline compound recently published by *Glass et al.*[5]

5.2. Surface Plasmon Resonance

5.2.1. Introduction

Surface Plasmon Resonance experiments were carried out in Georgia State University by Prof. David Wilson's group on some of the difunctionalised molecules that have shown significant results for the optical biophysical measurements. Surface Plasmon Resonance is used to measure binding affinities to hairpin DNA. An advantage of SPR over other optical techniques is that it can measure binding affinities and kinetics,[1] hence it allows the investigation of the binding of these molecules to specific sequences of DNA and the evaluation of the corresponding binding constants.

Unlike the previously described experiments, hairpin DNA sequences were used with this technique. Three different sequences were utilised in order to evaluate the differences in binding affinity:



In a SPR experiment, the hairpin is immobilized onto a sensor chip, which is normally made of gold. The technique is based on the fact that light will reflect with a different angle from this gold chip when it has DNA immobilized onto it than when it has DNA bound to a compound (Figure 5.1.a). Steady-state binding analyses are performed with multiple injections of different concentrations of the compound over the immobilized DNA surface. These solutions of the ligand are injected through the flow cells until a constant steady-state response is obtained. Compound solution flow is then replaced by buffer flow, resulting in dissociation of the ligand from the DNA. The reference response from the blank cell is subtracted from the response in each cell containing DNA to give a signal (RU, response units) that is directly proportional to the amount of bound compound.[1,2]

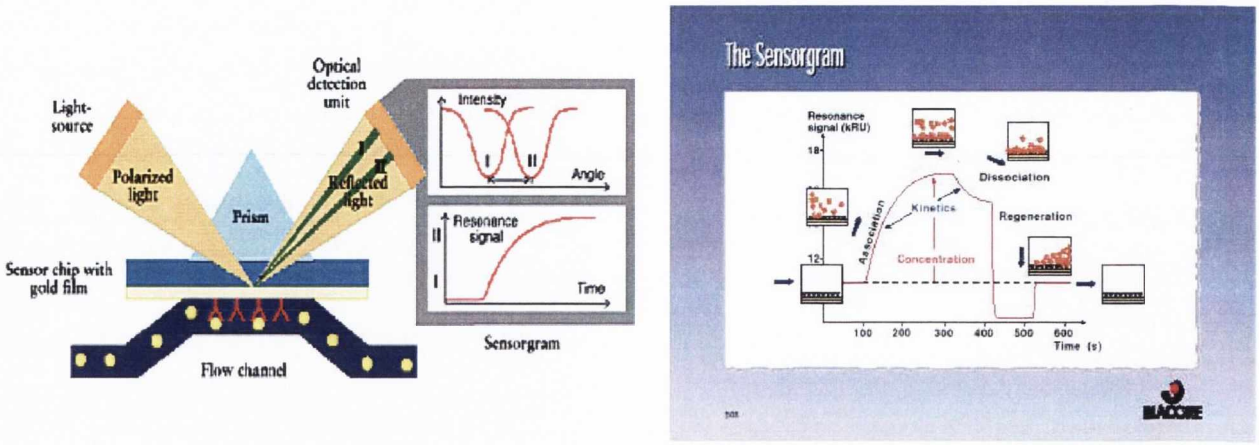


Figure 5.1. (a) Reflection of light on the sensor chip with the hairpin DNA and the influx of the ligand solution.[6] *(b) Sensorgram resulting from the association-dissociation of the ligand to the oligonucleotide.*[7]

The predicted maximum response per bound compound in the steady-state region (RUmax) is determined from the DNA molecular weight, the amount of DNA on the flow cell, the compounds molecular weight, and the refractive index gradient ratio of the compound and DNA. A plot known as a sensorgram can be produced by registering the changes in the RU with time (Fig. 5.1b).

5.2.2. Surface Plasmon Resonance (SPR) Results

SPR experiments were carried out in all asymmetric molecules (**3h-10h**). All the SPR experiments were performed using three types of oligonucleotides, AT-rich (AATT or TTAA) and CG-rich (CGGC) hairpins, to investigate the specificity of the different molecules for each sequence assessing their potential as selective minor groove binders.[1,2]

The results by representative compounds such as **5h** and **3h** in particular will be discussed next. Compound **5h** (NH linker), binds strongly to the AATT hairpin producing strong changes in the sensorgram as can be observed in Figure 5.2a. However, if the molecule

binds not so strongly to the AATT-hairpin, such is the case of compound **3h** (CH_2 linker), only a small change will be observed in the sensorgram as seen in Figure 5.2b.

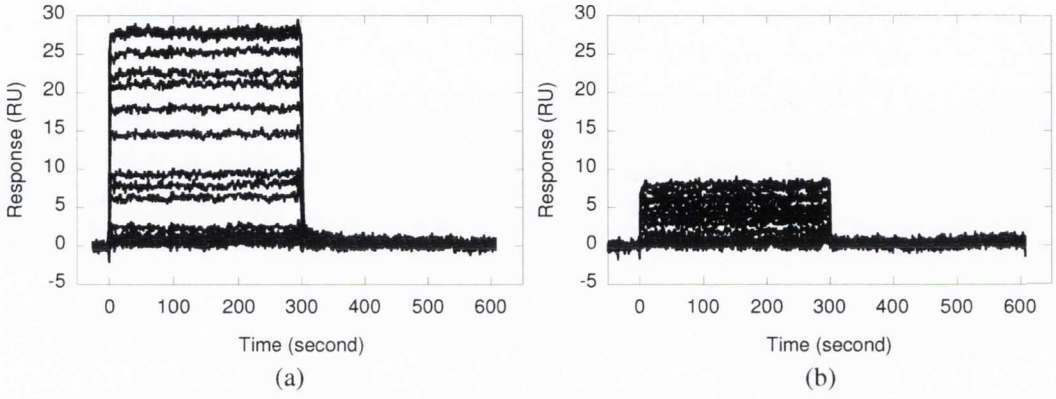


*Figure 5.2. SPR sensorgrams obtained using the AATT hairpin for molecule **5h** (a), and for molecule **3h** (b).*

The number of binding sites and the equilibrium constant can be obtained from fitting plots of RU *versus* the concentration of free compound. A plot of RU in the steady-state plateau regions from the sensorgrams shown in Figure 5.3 *versus* the compound concentration in the flow solution (free compound concentrations) can be obtained for **5h** (circles) and **3h** (squares) with the AATT hairpin, indicating strong bind (exponential fitting) and weak bind (linear fitting).

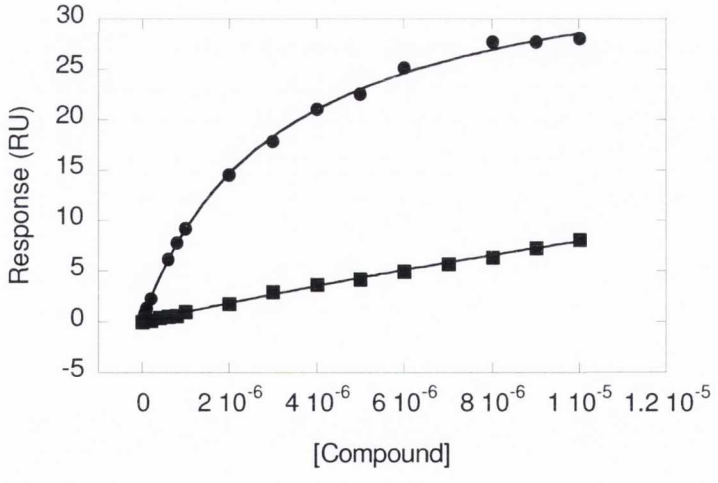


*Figure 5.3. A plot of compound concentration vs. response units for the experiments carried out with compounds **5h** (circles) and **3h** (squares).*

From the SPR experiments it is possible to extrapolate the corresponding binding constants. Thus, in Table 5.1 the results obtained for all the compounds studied with the three hairpins are presented.

Table 5.1.- Binding constants determined from the SPR experiments performed with the difunctionalised molecules and three different hairpins (AATT, TTAA and CG).

Compound	X	K_{AATT} $\times 10^5\ M^{-1}$	K_{TTAA} $\times 10^5\ M^{-1}$	K_{CG} $\times 10^5\ M^{-1}$	Selectivity index[a]
3h	CH_2	0.5	0.5	<<1	1.0
4h	CH_2CH_2	2.9	0.8	<<1	3.9
5h	NH	6.4	2.4	<<1	2.7
6h	O	4.7	0.7	<<1	7.0
7h	S	1.2	2.4	1.8	0.5
8h	CO	2.7	4.6	3.6	0.6
9h	NHCONH	1.3	0.7	<<1	2
10h	Piperazine	2.0	0.8	<<1	2.5

[a] Selectivity index = the ratio between the binding constants evaluated for the AATT and TTAA sequences (K_{AATT}/K_{TTAA}). Maximum errors were 10 %.

In general, larger binding constants are obtained for the interactions with AATT than with TTAA. Exceptions are compound **3h,** with similar and very small values indicating poor binding in both cases, and compounds **4h** and **7h** with TTAA binding constants double than the corresponding AATT ones (still small values indicating not too strong interactions). It has been found in several crystal structures that the AATT sequences minor groove is narrower than the TTAA one and the latter does not contain an ordered array of water molecules. Therefore, our molecules seem to have a preference for narrower and hydrated minor grooves.

As shown with other techniques, the compound with stronger binding to AATT was **5h** (with the NH linker) with almost 3–fold AATT/TTAA selectivity. The compound with the

largest selectivity for the AATT oligonucleotide was compound **3h** (X= O) with a 7–fold ratio and a good binding strength. Compounds **2** (X= CH_2CH_2) and **6** (X= piperazine) even though showing a medium binding strength exhibit very good 4– and 3–fold selectivity for AATT.

Interestingly, most of the compounds show very poor binding constants with the CG oligonucleotides indicating preference for AT sequences and, therefore, for the minor groove. However, compounds **7h** (X= CO) and **4h** (X= S) exhibited unexpected affinity for the CG oligonucleotide (see Table 5.1). In particular, compound **7h** shows a larger K value for CG than for AATT and very similar to that for TTAA. This is a surprising result since all the experiments performed so far indicated that compound **7h** is a minor groove binder and it is generally assumed that preference for CG sequences is characteristic of intercalators. The binding mode of these compounds clearly requires more detailed investigation.

5.3 Isothermal Titration Calorimetry

5.3.1. Introduction

Up to now the discussion has been focussed on the binding affinities of our molecules to different sequences of DNA. Another important topic that can help with our understanding of the DNA binding of these molecules is the thermodynamics of the process. In Chapter four we have been able to dissect the binding free energy for **5h** into the free energy change due to the electrostatic interactions and due to the non-electrostatic interactions. Now this discussion is based on dividing the free energy into the enthalpic and entropic contributions. This was investigated by using Isothermal Titration Calorimetry (ITC).

ITC is a powerful analytical tool to measure the change in enthalpy in a bimolecular process and is the only method known to measure this quantity directly.[12] Its usefulness can never be underestimated since in a single experiment the association constant, binding stoichiometry, enthalpy, free energy and the entropy can be determined. The combination of structural data and thermodynamic information can improve our overall understanding of the binding interactions to DNA. The principle behind this technique involves measuring the difference of temperature between a reference cell and the sample cell. The reference cell contains a buffer solution whereas the sample cell contains the DNA solution. A constant

power is supplied to the reference cell whereas a variable power is supplied to the sample cell. This is to maintain a very small temperature difference between the two cells.[13,14]

In the ITC experiments carried out for this project, the DNA macromolecule was placed in the sample cell and the ligand under investigation was taken up into the syringe (Figure 5.4.). After thermal calibration of the instrument, aliquots of the ligand were injected into the DNA solution. The reaction between the ligand and DNA resulted in large release of heat causing the power supplied to the sample cell to be adjusted accordingly to maintain thermal equilibrium. This variation in the power is registered in the ITC graph (Figure 5.4).

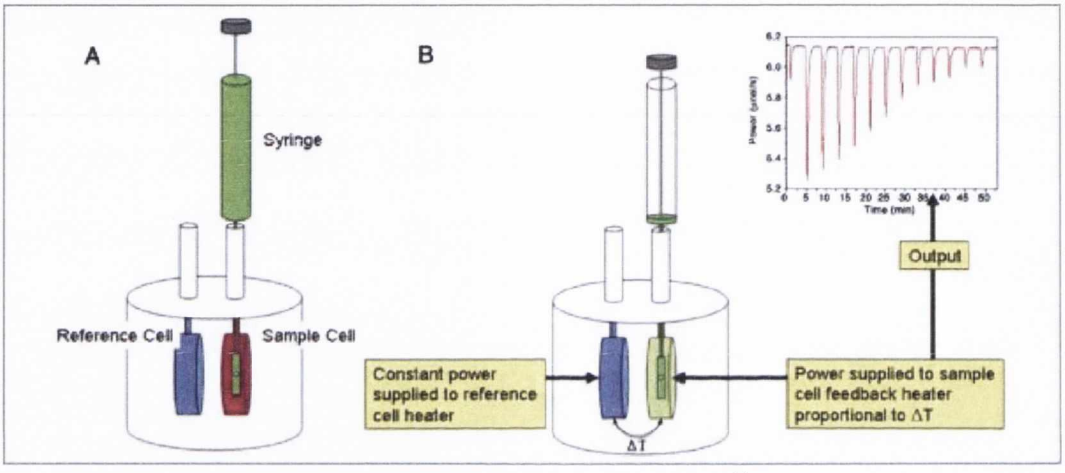


Figure 5.4. Illustration of the calorimeter[14]

ITC data were fit according to a standard model that assumes a single set of equivalent and independent binding sites. The midpoint (inflection) of the sigmoidal binding isotherm curve corresponds to the stoichiometry (n) of the interaction, given the input concentrations.

5.3.2. Isothermal Titration Calorimetry (ITC) Results

ITC experiments were performed, with both natural and poly(dA-dT)$_2$ DNA, for compounds **8h**, **5h** and **14a**. Results are shown in Figure 5.5 and the corresponding thermodynamic parameters obtained are presented in Table 5.2. From these experiments we observed that binding between our molecules and DNA produced exothermic reactions (Figure 5.5).

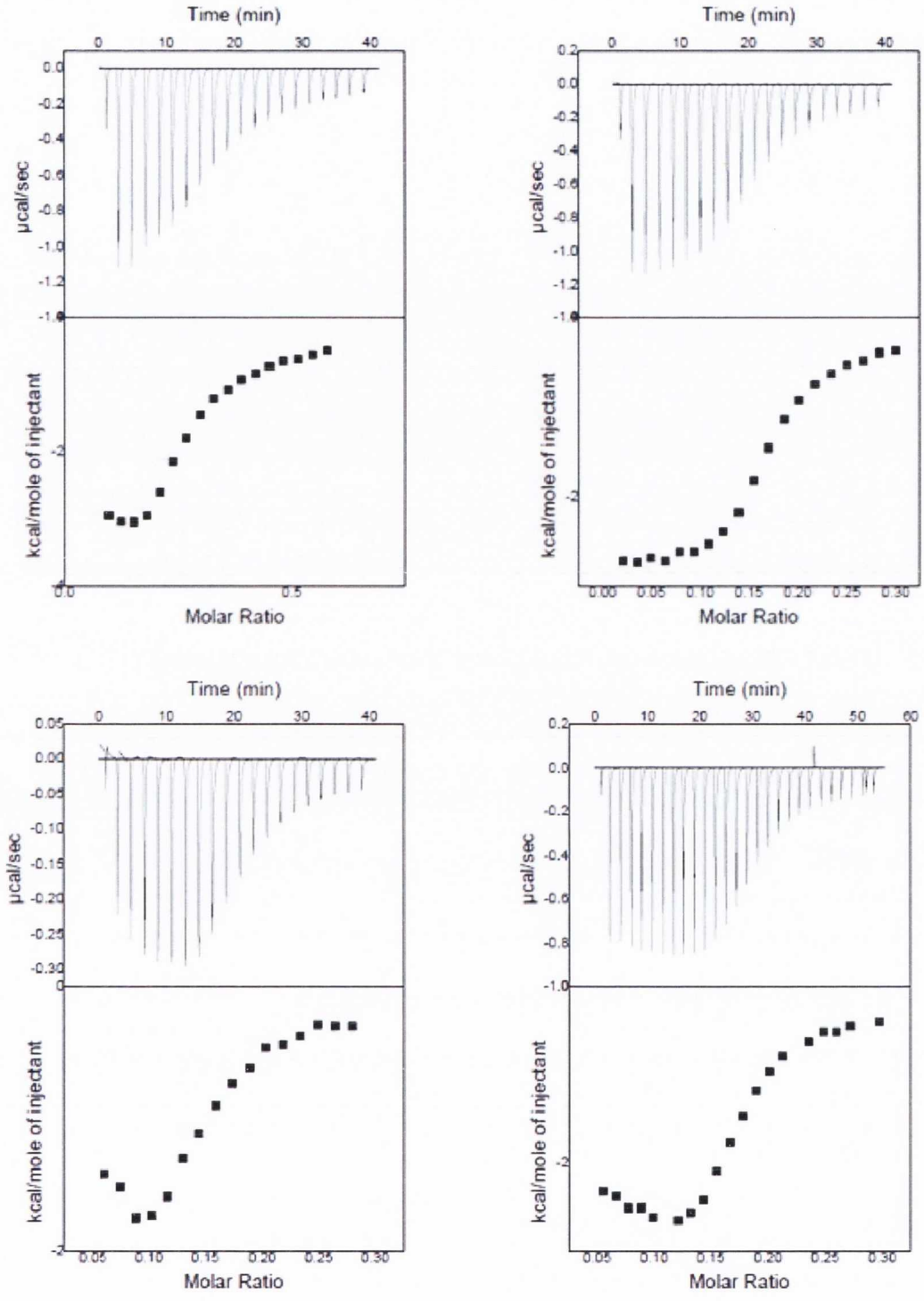
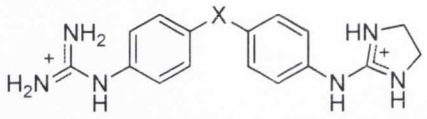


Figure 5.5. ITC graphs for compounds ***5h***, ***8h*** *and* ***14a:*** *top left = NH with natural DNA; top right = NH with poly(dA-dT)$_2$ DNA; bottom left = CO with natural DNA, bottom right = CONH with poly(dA-dT)$_2$ DNA*

As mentioned before, ITC experiments indicate that the binding of compounds **8h**, **5h** and **14a** results in an exothermic reaction since each of these compounds display favourable thermodynamic quantities (negative enthalpy and positive entropy values, Table 5.2). Interestingly, by comparing the binding thermodynamics of **5h** with natural DNA and poly(dA-dT)$_2$ DNA we can see that when bound to natural DNA, compound **5h** displays a slightly more favourable enthalpy. However, this is compensated in AT DNA by a larger positive entropy value which determines the overall favourable binding for AT sites in comparison to natural DNA. Natural DNA exhibits a random sequence, whereas poly(dA-dT)$_2$ contains only adenine-thymine base pairs, which facilitate the incorporation of water molecules into the groove. This could explain the slightly less favourable enthalpic value and the more favourable entropic value observed.

*Table 5.2.- ITC results for the compounds **8h**, **5h** and **14a** both with natural and AT DNA.*



Compd. (X)	DNA	K $\times 10^4$ /M^{-1}	ΔH kcal/mol	ΔS kcal/mol °C	n[a]
8h (CO)	natural	5.00 ± 1.73	-1.12 ± 0.1	17.7	0.298 ± 0.014
5h (NH)	natural	3.44 ± 0.863	-3.74 ± 0.13	8.21	0.299 ± 0.010
5h (NH)	AT	6.12 ± 0.942	-2.98 ± 0.07	11.9	0.355 ± 0.006
14a (CONH)	AT	15.4 ± 6.3	-2.55 ± 0.09	15.9	0.193 ±0.005

[a] Binding stoichiometry is in Bp/D ratio.

5.4. Structural Studies through X-Ray Crystallography

5.4.1. X-Ray Diffraction (XRD): Crystal Structure Determination of 4c

Compound **4c** was crystallised from slow evaporation in a mixture methanol and dichloromethane (Figure 5.6). The crystal structure was measured by Dr. Tom McCabe from the School of Chemistry in Trinity College Dublin.

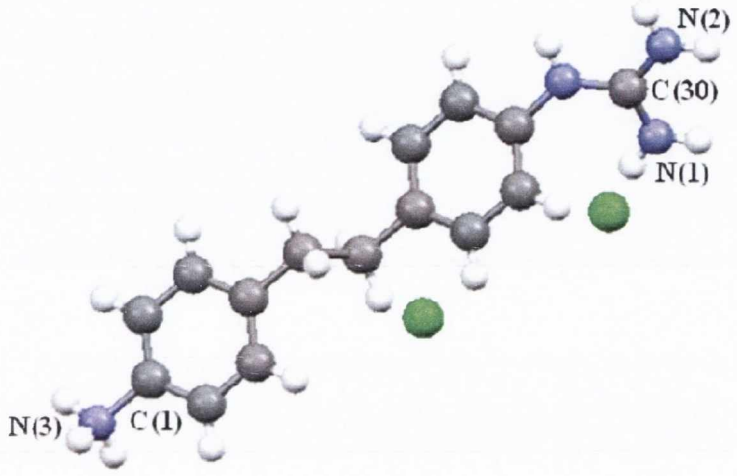


Figure 5.6. View of the asymmetric unit of 4c along the c-axis.Colour code: C grey, H white, N blue, Cl green.

Several key features can be observed: the amino group and the highly basic guanidine functionality are protonated and this is in agreement with the presence of the chloride anions in the structure. It can also be rationalised by comparing the C–NH_2 bond distance of the aniline group, 1.398(1) Å, with the C(1)–N(1) bond length of compound **4c**, being 1.466(1) Å. The bond lengths of the two NH_2 groups in the guanidine moiety, C(30)-N(1), and C(30)-N(2) are 1.321(2) Å and 1.326(2) Å respectively. These analysed bond distances involving the carbon and nitrogen atoms suggest an intermediate single/double bond character in structure **4c**. In addition, it can be deduced that the positive charge is delocalised among the nitrogen atoms and the central carbon atom in the guanidine functionality. Furthermore, the two aromatic rings of the molecule are approximately coplanar with respect to the central linker which forms an angle of around 112°. This central linker provides the molecule with a curvature that fits well inside the minor groove convex structure.

5.5. Conclusions

The non-optical techniques to study the DNA binding or our asymmetric dications were described in this chapter. First, Surface Plasmon Resonance was carried out to investigate the selectivity of our molecules for three different sequences. Interestingly, it was found that the dication **6h** (linker = O) displays the highest selectivity for AATT over TTAA sequences. A reason for this lies in the possible difference in groove width in these sequences; however, further investigation would be required by preparing other molecules and investigating their selectivity. Some correlations were found for these compounds; thus, **5h** displayed the highest binding affinity in the SPR experiments which is in agreement with the results obtained using optical techniques excluding **11h** and **14a** (fluorene derivative and linker = CONH respectively). An unexpected result was observed for **8h** (linker = CO) where these experiments detected good binding affinity for GC sequences.

The thermodynamics of the binding to DNA (both natural and AT) of several dications was investigated by ITC obtaining, in all cases, a favourable negative enthalpy and positive entropy values upon binding of the molecules. We were able to obtain the binding stoichiometry which correlated well with one molecule binding to every four base pairs (the normal value observed for minor groove binders). However, ITC is most useful when it is correlated with structural data which we have not been able to obtain yet.

The final aspect described in this chapter is the crystal structure obtained of **4c**. This structure illustrated many important aspects such as that the positive charge is distributed equally over the guanidine functionality, and that the amino group was protonated under these conditions. Moreover, it was found that the two aromatic rings only slightly deviate colinearity. However, when this molecule is in solution is only *mono*-cationic which does not favour minor groove binding. Interestingly, stacking interactions between the aromatic rings were observed in the crystal unit to have a stabilising effect on the structure.

Overall, the past two chapters illustrate that these molecules exhibit all the necessary characteristics of a minor groove binder. These include cationic ends for ionic interactions, hydrogen bonding groups and aromatic rings for potential van der Waals contacts with the DNA base pairs. The asymmetric dicationic compounds have shown good binding affinity to the minor groove and lay a foundation for future research in this area.

5.6. References

1. Nguyen, B., Tanious, F. A., Wilson, W. D., *Methods*, **2007**, *42*, 150
2. Freyer, M. W., Buscaglia, R., Nguyen, B., Wilson, W. D., Lewis, E. A., *Anal. Biochem.*, **2006**, *355*, 259
3. www.reference.md/files/D020/mD020349.html
4. Velazquez-Campoy, A., Ohtaka, H., Nezami, A., Muzammil, S., Freire, E., *Current Protocols in Cell Biology*, **2004**, *17.8*, 1
5. Glass, L. S., Nguyen, B., Goodwin, K. D., Dardonville, C., Wilson, W. D., Long, E. C., Georgiadis, M. M., *Biochemistry*, **2009**, *48*, 5943
6. www.lerner.ccf.org/services/molecbiotech/images/biacore1.gif
7. www.astbury.leeds.ac.uk
8. Dickerson, R. E.; and Drew, H. R.; *J. Mol. Biol.*, **1981**, *149*, 761
9. Fratini, A.V.; Kopka, M.L.; Drew, H.R.; and Dickerson, R. E.; *J.Biol. Chem.*, **1982**, *257*, 686
10. Quintana, J.R., Grzeskowiak, K. Yanagi, K. and Dickerson, R.E. *J.Mol. Biol.*, **1992**, *225*, 379
11. Liepinsh, E., Leupin, W. and Otting, G. *Nucleic Acids Res.*, **1994**, *22*, 2249
12. Bjelic, S., Jelesarov, I., *J. Mol. Recognit.*, **2008**, *21*, 289
13. Holdgate, G. A., *Biotechniques*, **2001**, *31*, 164
14. Ladbury, J. E., *Biotechniques*, **2004**, *37*, 885
15. Fukuyo, M., Hirotsu, K., Higuchi, T., *Acta Cryst.*, **1982**, *B38*, 640

Chapter 6

Cytotoxicity evaluation of the Asymmetric Dications

6.1. Introduction

Following the experiments to characterise the DNA binding of our molecules, the next step was to evaluate the cytotoxicity of these compounds. Investigating the cytotoxicity of the dications against a particular cancer line will help our research in two ways. Firstly, the necessary characteristics for a compound to display significant anti-cancer activity could be determined and this will help us to design new potential anti-cancer compounds. Secondly, we would be able to understand if the anti-cancer activity is solely due to DNA binding or to some other process. Only those compounds that displayed significant DNA binding affinity were subjected to biochemical testing. Consequently, we only investigated the asymmetric dicationic molecules belonging to the Family III.

The cytotoxicity of these compounds was evaluated by cellular growth inhibition assays and was determined by our collaborators at Dr. Geoff Margison's laboratory from the Paterson Institute for Cancer Research, UK. The cancer cell line that was used is called MCF-7 which is derived from a breast tumour from an elderly Caucasian woman, which was isolated in 1970. MCF-7 stands for Michigan Cancer Foundation-7 referring to the institute in Detroit where the cell line was established in 1973.[1]

In these assays a monolayer of MCF-7 cells were grown in RPMI 1640 medium supplemented with 10 % foetal calf serum and 1 % L-glutamine. These cells were plated in a 96-well flat bottomed cell culture plate at a density of 2000 cells per well. Doses ranging from 0 μM- 100 μM were added to the plates and these were incubated at 37 °C over 6 days.

6.2. Cytotoxicity experiments

The effect of increasing the concentration of the compounds **3h** to **10h** (from 0 to 100 μg) upon the growth of MCF-7 cell lines was recorded and the results obtained for each compound are displayed in Figure 6.1. From these assays it was possible to calculate the corresponding IC_{50} values (concentration required to inhibit 50% of the cellular growth) for the asymmetrical derivatives and the results are presented in Table 6.1.

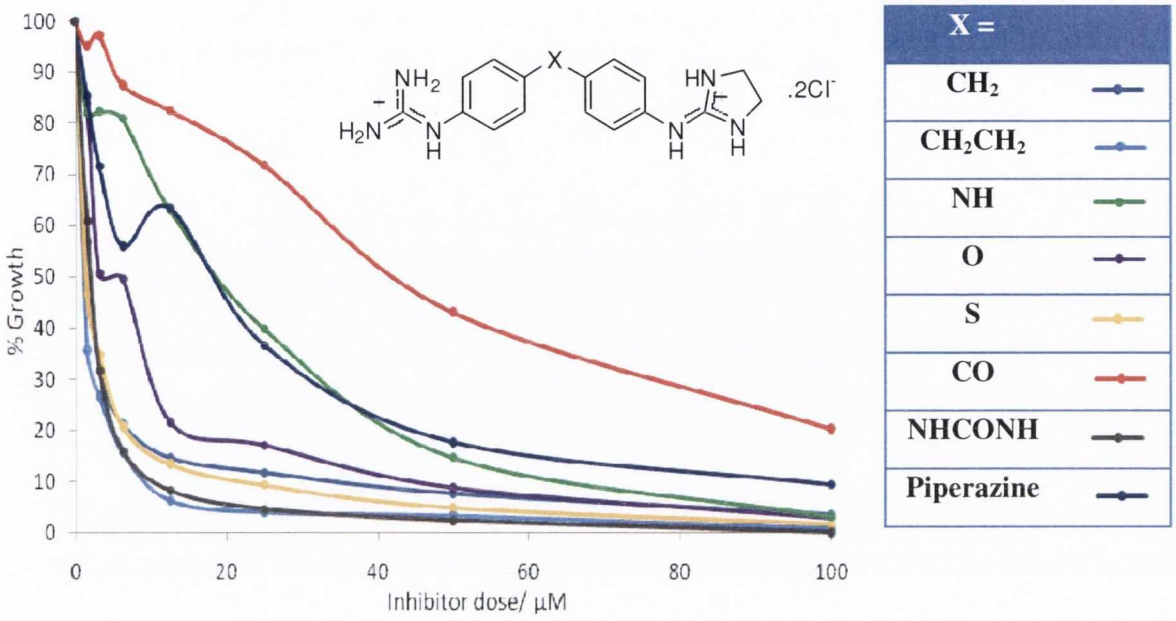


*Figure 6.1. Growth inhibition curves for **3h-10h***

Interestingly, we observed that the IC_{50} value of **4h** (X = CH_2CH_2) was two-fold higher than that of **3h** (X = CH_2). This compound **4h** is more lipophilic than **3h** so its higher cytotoxicity could be due to its ease of crossing the cell membrane. In contrast to this result, compound **5h** (X = NH) displayed a nine-fold decrease in activity in comparison to **3h**. In this case, the differences between these two compounds are not only of lipophilicity (**3h** more lipophillic than **5h** due to the linker) but also the DNA binding affinity of compound **5h** is much higher than that of compound **3h**, as described in the predeceasing chapters. Therefore, even though **5h** has better DNA affinity than **3h**, the difference in the IC_{50} values between these two compounds may be attributed to the lipophilicity of compound **3h** and its rate of diffusion into the cancer cell, as diffusion would be favoured for the more lipophilic compounds.

Table 6.1.- IC_{50} values obtained for the di-functionalised molecules with MCF-7 cells line

Compound Number	X =	IC_{50}/ µM
3h	CH_2	3
4h	CH_2CH_2	1.5
5h	NH	28
6h	O	9
7h	S	3
8h	CO	63.5
9h	NHCONH	3
10h	Piperazine	25

Further evidence could be obtained from Table 6.1 since compound **8h** (X= CO) displays the worst results in this assay (IC_{50} = 63.5 µM). This corresponds to over 21-fold decrease in the cytotoxicity activity against the cancer cell line in comparison to **3h**. Again, despite the good DNA affinity of **8h** and since the CO linker is less lipophilic than the CH_2 one.

Comparing the results obtained for **6h** and **7h** (X = O and S respectively), the less polar sulfur bridge shows a three-fold increase in the cytotoxicity in comparison to the oxygen linker. This again correlates to the lipophilicity of the compounds. However, different molecular factors cannot be dismissed.

6.3. Correlations: DNA binding against Cytotoxicity

One of the purposes of the cytotoxicity experiments was to examine if there was any correlations with the DNA binding affinities discussed in the previous chapters. Thus the potential correlation between both properties was explored, but no direct relation was found between the DNA affinity of these compounds as measured by the ΔT_m values obtained in the thermal denaturation experiments both with natural DNA and with the poly(dA-dT)$_2$ oligonucleotide. Thus, the compound with one of the strongest binding affinities (**5h**) has

been shown to display one of the worst results from the cytoxicity assays. Nevertheless, relevant compounds such as those with an ethylene (**4h**) or a urea (**9h**) linker showed very good IC_{50} values and had shown to strongly bind to DNA (see Table 6.1).

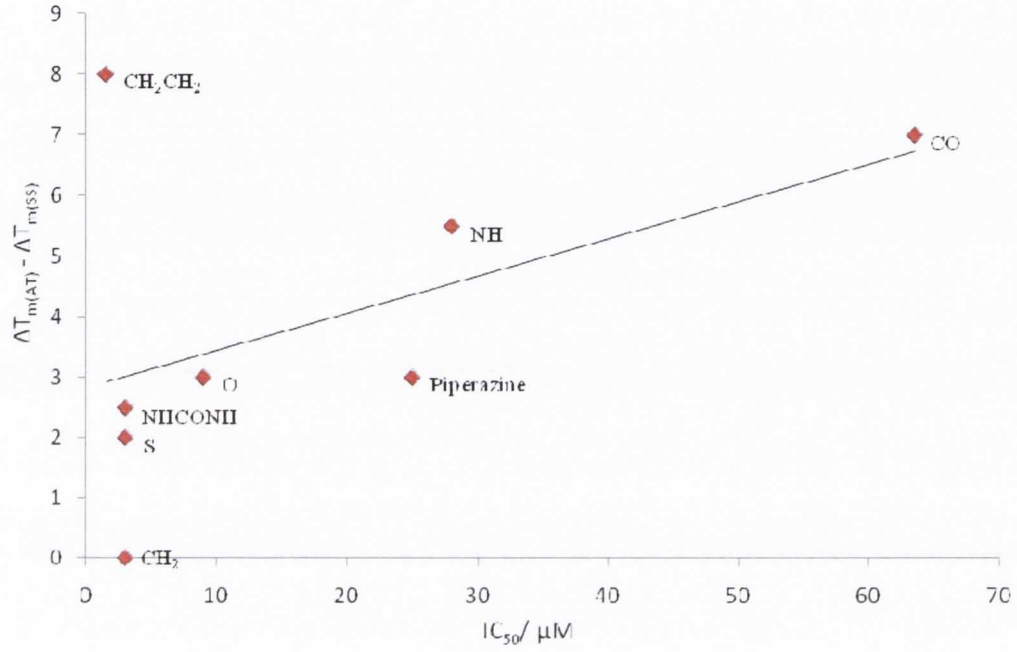


Figure 6.2. Correlation plot between the DNA binding affinity and the IC_{50} values from the cytotoxicity experiments

Next, the possible correlation between cytotoxicity values and AT selectivity (as measured by the difference in ΔT_m values in salmon sperm and AT DNA) was analysed (Figure 6.2) but, even though some similar trends were observed for compounds **8h** (X= CO), **5h** (X= NH), **6h** (X= O) and **9h** (X= NHCONH), again no clear correlation was found.

When considering the best approach for constructing a molecular scaffold to facilitate DNA binding interactions, the chemical nature of the bridging units must also be considered. It is interesting to note that, with the exception of the urea linker, the more chemically inert functional groups (CH_2, CH_2CH_2 and S) convey the greatest cytotoxicity (see Table 6.1).

6.4 Correlations: Log(IC_{50}) against Log P

Interestingly, we decided to investigate the correlation between the lipophilicity of the asymmetric dications **3h** to **10h** and their corresponding cytotoxicity. For this we plotted the theoretical value for the logP (as calculated using the program ChemDraw Version 12.0) against the log(IC_{50}) values and the graph obtained is shown in Figure 6.3.

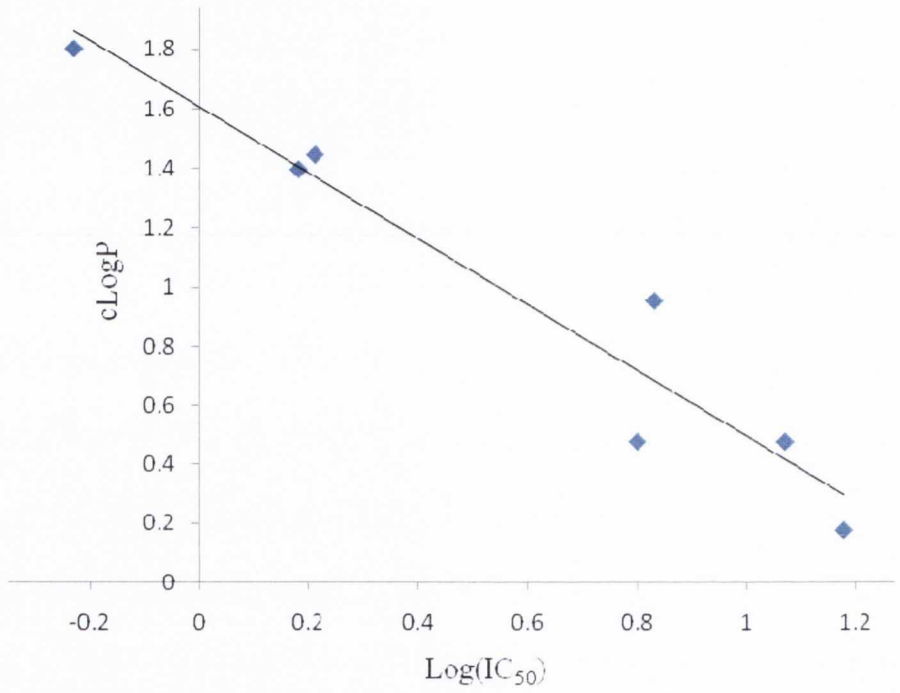


*Figure 6.3.Correlation between the Log(IC_{50}) and the cLogP values for compounds **3h** to **10h**.*

The linear relationship found between the cytotoxicity and the calculated LogP value is shown in the following equation:

$$cLogP = -1.114(Log(IC_{50}) + 1.607 \qquad (R^2 = 0.93)$$

A good correlation was found indicating that when designing new DNA targeting anticancer derivatives the lipophilicity of the compounds should be taken into account. However, the result for compound **9h** (X = NHCONH) was excluded for this correlation as the result obtained decreased the corresponding R^2 coefficient. It should be noticed that these

cLogP values are only theoretical but they provide a good basis for the design of future anticancer compounds in our group.

6.5 Conclusions

Asymmetric dicationic compounds **3h** to **10h** displayed IC_{50} values ranging from 1.5 μM to 63.5 μM. This provides us with information of the most favourable linkers when preparing new compounds. It was observed that the compounds with the less polar linkers displayed the most favourable results, indicating that for future synthesis, molecules with a non polar linker should be considered. Also, it was noted that there was some correlation between the IC_{50} values and lipophilicity of the linker, indicating that the cytotoxicity could be related to the diffusion of the compounds into the cell prior to DNA binding.

It was originally thought that cytotoxicity could be related to the DNA binding affinity of the molecules. However, plots of the IC_{50} values against the thermal denaturations results and AT sequence selectivity indicate poor correlation. This could be due to the cytotoxicity being related to other cellular mechanisms.

6.6 References

1. Soule, H. D.; Vazquez, J.; Long, A.; Albert, S.; Brennan, M.; *J. Nat. Cancer Institute*, **1973**, *51*, 1409

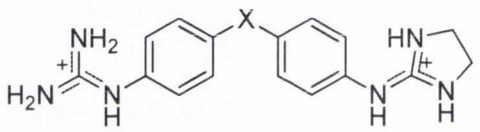
2. Levenson, A. S.; Jordan, V. C.; *Cancer Res.*, **1997**, *57*, 307

Chapter 7

Synthesis of the Asymmetric dications-Acridine Conjugates

7.1. Introduction

In the previous chapters, the interactions and cytotoxicity of minor groove binders have been studied. These compounds bind to DNA with high affinity and we have been able to evaluate aspects of their DNA interactions such as their binding constants, thermodynamics and mode of binding by means of different biophysical techniques. Moreover, we have studied the impact of the linker in minor groove binding and the order of the linker influence on the binding affinity of the diaromatic scaffold.[1] We observed that compound **5h** (X = NH, Figure 7.1) displayed the highest binding affinity to natural DNA that is thermodynamically favourable with a strongly negative enthalpy and positive entropy. We have also observed that most of these molecules displayed selective binding for adenine-thymine (AT) sequences.



X = CH_2, CH_2CH_2, NH, O, S, CO, NHCONH, Piperazine

Figure 7.1. Asymmetric compounds whose DNA affinity and AT selectivity was evaluated in this project

It is known from the literature, that many intercalators are good cytotoxic drugs.[2] In most cases they also work as topoisomerase poisons[3] meaning that they stabilise the ternary DNA-topoisomerase complex thereby inhibiting transcription that is a known cause of apoptosis.[4,5] There are several chemotherapeutic drugs that are DNA intercalators and act as topoisomerase poisons such as DACA[6] and Amsacrine (AMSA)[7] which are used in the clinic (Figure 7.2). Both of these drugs have in common the acridine chromophore used for interacting into the DNA double helix by intercalation. This causes the unwinding and lengthening of the double helix and is stabilised by hydrophobic, π-π and charge transfer interactions.[8]

As mentioned before, the common feature of these molecules is the acridine chromophore with a side chain attached. Considering that our aim was to prepare DNA dual acting agents it was decided to attach our molecules to the acridine chromophore in order to obtain an enhancement in the DNA binding affinity and cytotoxicity of the molecules would hopefully be observed. Thus, we intended to prepare dual action molecules that would

intercalate between the DNA base pairs with the acridine moiety and bind the minor groove by means of the diaromatic guanidine, simultaneously.

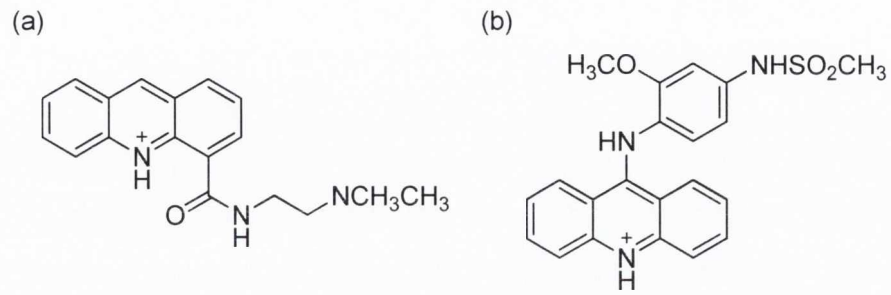


Figure 7.2. (a) DACA, (b) AMSA

In doing so, we selected a few minor groove binding molecules to attach on the acridine chromophore, on the basis of their DNA binding affinity and cytotoxicity. Since **5h** (X= NH) and **4h** (X= CH_2CH_2) display the highest DNA binding affinity and the most favourable results from the cytotoxicity experiments respectively, we decided to attach the corresponding mono guanidine analogues to the acridine chromophore. Additionally, the acridine conjugates of the mono guanidine **3h** (X= CH_2) since it displayed poor DNA binding affinity and no sequence selectivity, that of **8h** (X= CO) and that of *para*-toluidine were also considered for the sake of comparison. Hence, all the molecules proposed for synthesis are shown in Figure 7.2.

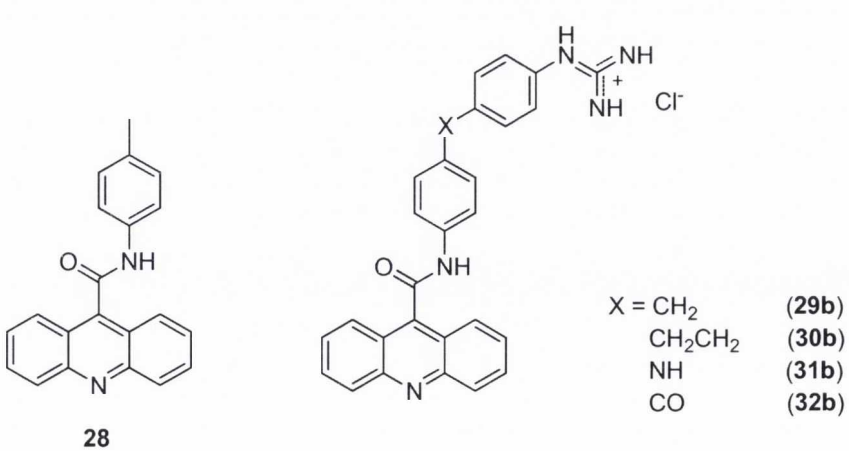
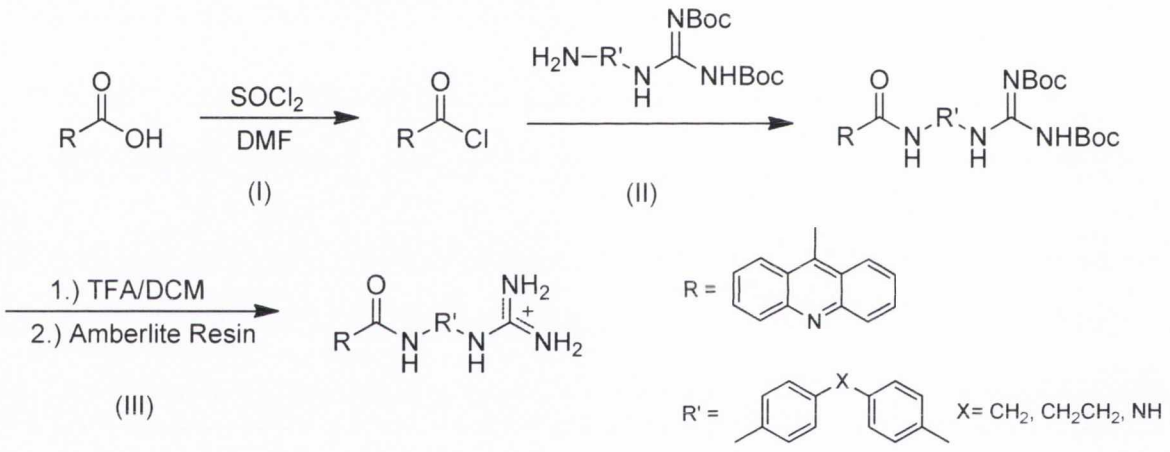


Figure 7.2. Acridine dual agents proposed to be synthesised

Our hypothesis is that by having a dual action molecule, the binding affinity of the intercalator and the groove binding part would be summed, thereby having an increased effect on the binding. From this, it is hoped that the cytotoxicity of these acridine conjugated potential drugs will be increased.

The generic procedure for the preparation of these molecules is shown in Scheme 7.1. Thus, the 9-acridine carboxylic acid will be reacted with thionyl chloride in the presence of a catalytic amount of DMF to form the corresponding acid chloride (I in Scheme 7.1).[9] This will be then reacted with the corresponding Boc protected *mono*-guanidine derivative to form an amide bond (II in Scheme 7.1) which will be then deprotected in the presence of a mixture of TFA and DCM following by treatment with basic amberlite resin exchange to yield the hydrochloride salts (III in Scheme 7.1).



Scheme 7.1.

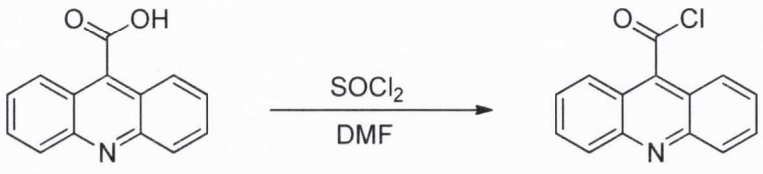
To summarise, the target of this research is the preparation of a series of DNA dual action molecules so they display intercalation and minor groove binding simultaneously and hopefully increased cytotoxicity against tumour cell lines.

7.2. Synthesis of the chloride of 9-acridine carboxylic acid

Considering the various strategies that are employed for peptide synthesis, as previously discussed (Chapter 3), we decided to synthesise the acridine conjugates by coupling the di-aromatic amine with the acid chloride of the 9-acridine carboxylic acid. There

are various reagents that could be used for the formation of the acid chloride, for example $SOCl_2$, $SOCl_2$ catalysed by DMF, oxalyl chloride or $POCl_3$.[10, 11]

We decided to react 9-acridine carboxylic acid with $SOCl_2$ catalysed with DMF as shown in Scheme 7.2.

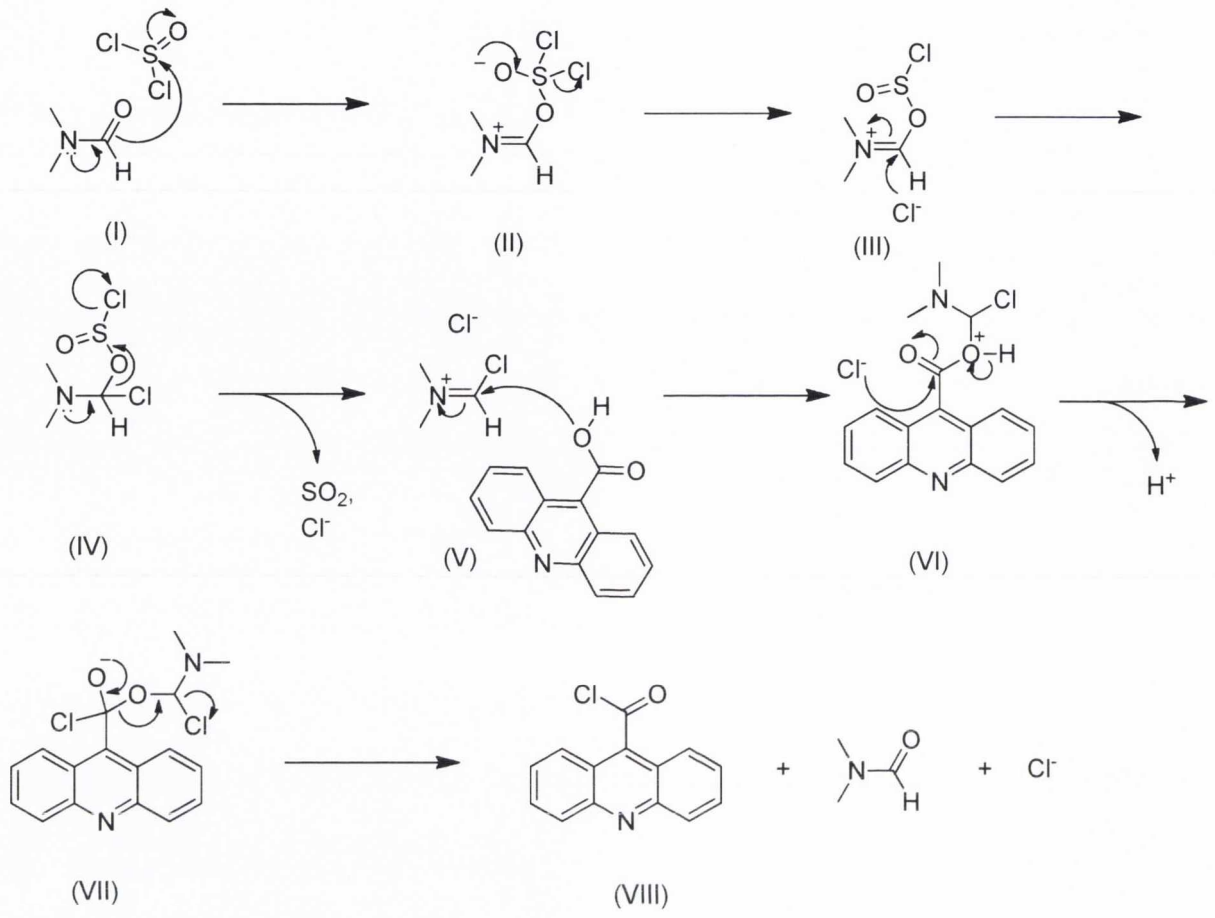


27

Scheme 7.2

First, a solution of the acridine carboxylic acid with an excess of thionyl chloride was heated to 50 °C. A catalytic amount of DMF was added slowly. The solution immediately dissolved and was left reacting for four hours. The reaction was concentrated under vacuum to yield a yellow solid that was used without further purification.

The mechanism for the formation of the acid chloride catalysed by DMF[12] is shown in Scheme 7.3. Initially, the lone pair of electrons from the nitrogen of the DMF molecule forms a C=N double bond (I). Simultaneously, a pair of electrons from the C=O bond attacks the electrophilic sulphur atom from the thionyl chloride, causing a pair of electrons from the S=O to go on the electronegative oxygen. This electrons go back to reform the S=O double bond, causing one of the chloride ions to leave (II). This chloride ion attacks the carbon of the iminium cation, at which point, a pair of electrons from the C=N neutralise the positive charge on the nitrogen atom (III). Then, a pair of electrons from the nitrogen atom reforms the iminium cation (IV), causing the pair of electrons form the C-O bond to form an S=O double bond and overall resulting in the loss of a chloride anion and SO_2 gas which is the driving force for this reaction.

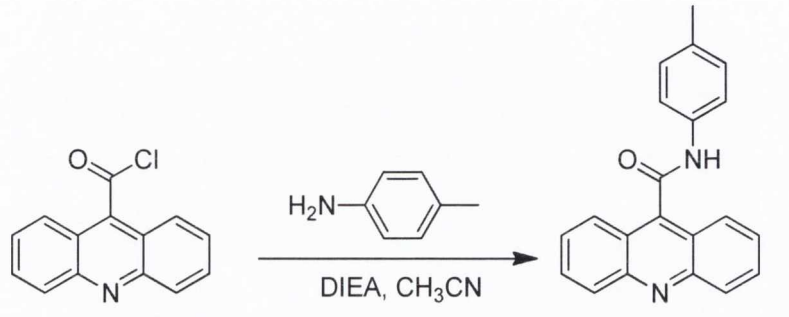


Scheme 7.3

The oxygen of the -OH from the carboxylic acid attacks the carbon atom of the iminium cation (V), resulting in the migration of a pair of electrons from the C=N double bond to neutralise the cation on the nitrogen atom, accompanied by the loss of a H^+ thereby increasing the acidity of the solution. The free chloride anion attacks the electrophilic carbon (VI) causing the pair of electrons from the carbonyl group to migrate to the oxygen atom. These electrons then return to reform the carbonyl group (VII), expelling dimethylformamide as a leaving group, thus recovering the catalyst resulting in the formation of the acid chloride (VIII).

7.3. Synthesis of N-(p-tolyl)acridine-9-carboxamide

Following the synthesis of the highly reactive acid chloride, the next target was the preparation of the acridine conjugates. Initially, and as a test reaction, the chloride of the acridine-9-carboxylic acid was reacted with *p*-toluidine according to Scheme 7.4.

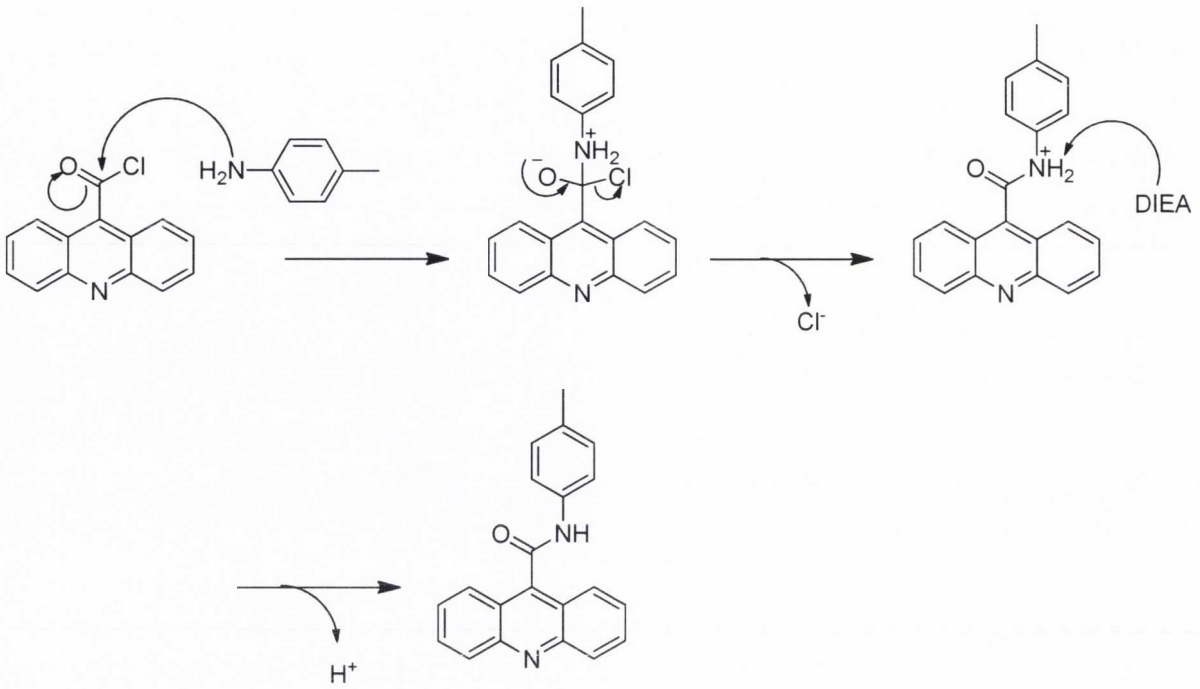


27 **28** (45%)

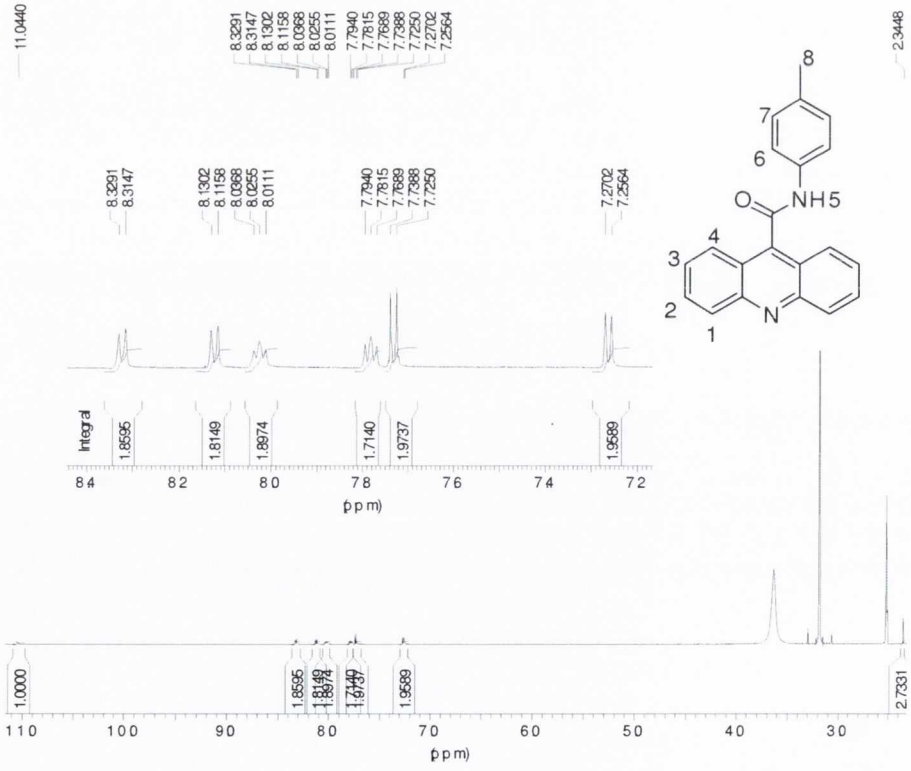
Scheme 7.4

This reaction required the preparation of two solutions. The first solution was prepared by dissolving the chloride of the 9-acridine carboxylic acid in anhydrous acetonitrile (CH_3CN) and the second solution was prepared by dissolving *p*-toluidine in anhydrous CH_3CN. The solution of the acid chloride was cooled to 0 °C at which point DIEA was added followed by the dropwise addition of the amine solution. The resulting mixture was allowed to stir for one hour at 0 °C and a further 20 hours at room temperature. The formation of a precipitate was observed which was filtered, washed with CH_3CN and purified by silica chromatography giving a yellow solid.

The mechanism of the CONH bond formation is shown in Scheme 7.5 and the ^1H NMR spectrum of the resulting compound **28** is presented in Figure 7.3.



Scheme 7.5

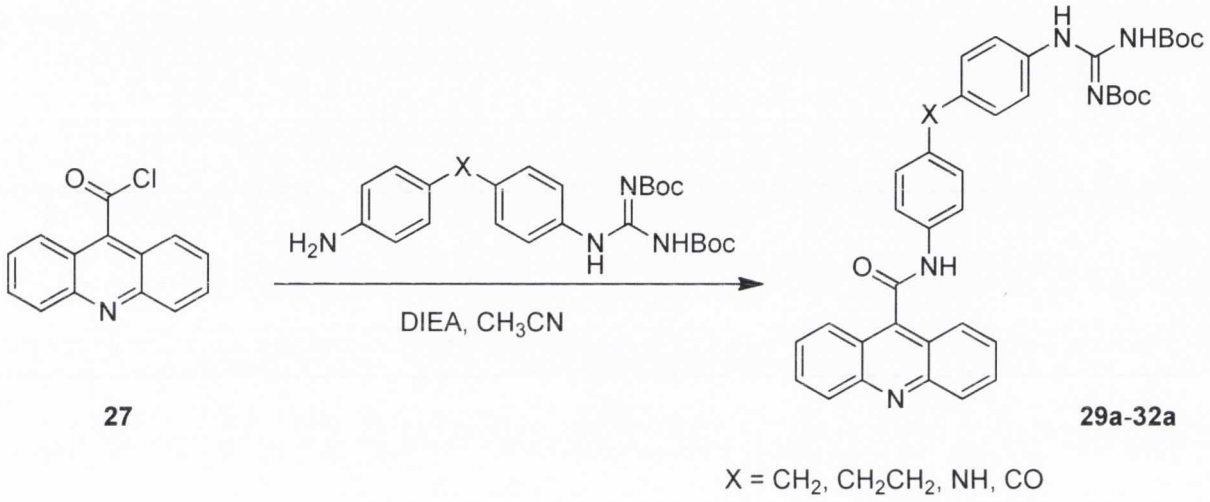


H	δ_H /ppm
C-1	8.32
C-2	7.78
C-3	8.03
C-4	7.73
N-5	11.04
C-6	8.12
C-7	7.26
C-8	2.34

*Figure 7.3. ^{1}H NMR analysis on compound **28***

7.4. Synthesis of guanidine diphenyl conjugates of acridine-9-carboxamide

Once the conditions for the acridine conjugation reactions were established, we proceeded to attach the minor groove binding scaffolds. First, we prepared the corresponding Boc protected *mono*-guanidine compounds as previously described in Chapter 3. Following the preparation of these compounds, they were reacted with the chloride of the 9-acridine carboxylic acid as shown in Scheme 7.6.



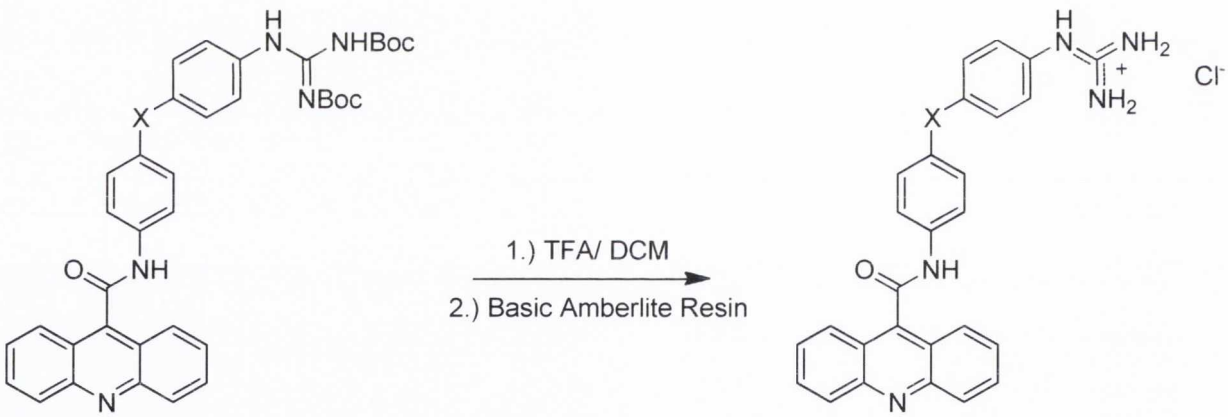
Scheme 7.6

As before, two solutions were prepared. The first solution consisted of acid chloride dissolved in anhydrous acetonitrile (CH_3CN) and the second of the corresponding amine dissolved in anhydrous CH_3CN. The acid chloride solution was cooled to 0 °C at which point DIEA was added followed by the dropwise addition of the amine solution. The resulting solution was allowed to stir for one hour at 0 °C and a further 20 hours at room temperature. The formation of a precipitate was observed which was filtered, washed with CH_3CN and purified by silica chromatography giving in each case a solid that was characterised by ^1H and ^{13}C NMR, C-H COSY, HRMS and melting point. Evidence for the formation of the products was seen in the ^1H NMR spectra from the presence of the amide NH peak, the eight signals in the aromatic region and the peaks for the Boc protons.

Unfortunately, the molecule with the carbonyl linker (**32a**) was not obtained. Even though the same procedure was followed, no product was observed by TLC and careful NMR analysis. The fact that the derivative with X= CO was not obtained could be explained because the deactivating nature of the carbonyl group, as it reduces the nucleophilicity of the *para* amino group. Consequently, the amino group is less prone to attack the electrophilic carbon of the acid chloride.

The Boc protected derivatives (**29a-31a**) were reacted in a mixture of trifluoroacetic acid and DCM to remove the Boc groups (Scheme 7.7). The trifluoroacetate salts were then treated with basic amberlite resin to give the corresponding hydrochloride salts. All final hydrochloride salts were characterised by ^{1}H and ^{13}C NMR, HRMS and melting point.

In the ^{1}H NMR spectra, we did not observe any peak corresponding to the Boc protons indicating that the molecules were fully deprotected. The absence of the trifluoroacetic acid was checked by ^{19}F NMR.



29a-31a

29b-31b

X = CH_2, CH_2CH_2, NH

Scheme 7.7

Comparing the yield obtained for compound **31b** with those of compounds **29b** or **30b** (Table 7.1) it can be seen that the electron donating effect of the *para* NH group on the NH_2 enhances its nucleophilicity producing better overall yields.

*Table 7.1.- Partial and overall yields obtained in the preparation of compounds **29b** to **31b***

Compound No.	X	% Yield	% Yield	Total % Yield
29b	CH_2	57	92	52
30b	CH_2CH_2	64	93	60
31b	NH	73	93	68

7.5. Conclusions

In summary, a series of intercalator-minor groove binding molecules were prepared. These were synthesised by a two step mechanism. First, the carboxylic acid was reacted with $SOCl_2$ in the presence of a catalytic amount of DMF to form the acid chloride. This was then used without purification and reacted with a series of Boc protected *mono*-guanidine compounds to yield the expected conjugated derivatives upon deprotection

The nature of the linker affected the nucleophilicity of the *para* amino group what in turn affected the yield. Where the linker was an electron donating group (X = NH), we obtained the highest yields, whereas when the linker was electron withdrawing (X = CO), no product was observed. Average yields were seen for those linkers that were neither strong electron donating nor withdrawing (X = CH_2, CH_2CH_2).

7.6 References

1. Nagle, P. S.; Rodriguez, F.; Kahvedzic, A.; Quinn, S. J.; Rozas, I.; *J. Med. Chem.*, **2009**, *52*, 7113
2. Pauwels, O.; Kiss, R.; Pasteels, J-L.; Atassi, G.; *Pharmaceutical Research*, **1995**, *12*, 1011
3. Liu, L. F.; *Annual Review of Biochemistry*, **1989**, *58*, 351
4. Zwelling, L. A.; Michaels, S.; Erickson, L. C.; Ungerleider, R. S.; Nichols, M.; Kohn, K. W.; *Biochemistry*, **1981**, *20*, 6553
5. Nelson, E. M.; Tewey, K. M.; Liu, L. F.; *Proc. Natl. Acad. Sci. USA.*, **1984**, *181*, 1361
6. Atwell, G.; Rewcastle, G. W.; Baguley, B. C.; Denny, W. A.; *J. Med. Chem.*, **1987**, *30,* 664
7. Cain, B. F.; Atwell, G. J.; Denny, W. A.; *J. Med. Chem.*, **1975**, *18*, 1110
8. Xu, Z.; Bai, G.; Dong, C.; *Bioorg. Med. Chem.*, **2005**, *13*, 5694
9. Montalbetti, C.; Falque, V.; *Tetrahedron*, **2005**, *61*, 10827

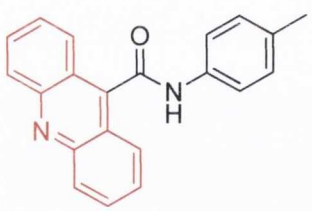
Chapter 8

Physicochemical Studies of the Acridine Conjugates Interactions with DNA

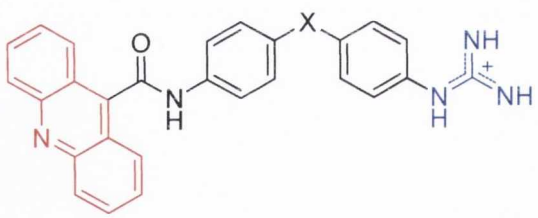
8.1. Introduction

Once the synthesis of the acridine-"guandine-binders" conjugates was complete, the next target was to characterise their DNA binding affinity. Several questions remain to be answered such as whether or not they bind to DNA and if they display higher binding affinity than the corresponding minor groove binders described in Chapter 4.

These molecules contain an intercalator moiety (acridine) and a minor groove binder moiety (-Ph-X-Ph-Guanidine), which are connected by an amide bridge. Our hypothesis is that these molecules should display a higher binding affinity than the minor groove binder due to the presence of the intercalator since we expect the intercalator moiety to insert in between the bases of DNA directing the minor groove binder moiety to fall into the DNA minor groove where the guanidinium cation could form hydrogen bonds and ionic interactions. This would be due to the highly favourable π-π interactions and hydrophobic effects[1] that can be established by both types of moieties. The model compound prepare to explore the synthesis of the acridine conjugates, the *p*-toluidine acridine conjugate, was also included in this study.



Acridine-p-toluidine (**28**)



X = CH_2 (**29**), CH_2CH_2 (**30**), NH (**31**)

Figure 8.1. Acridine containing molecules that were tested for DNA affinity

Our aim was to investigate the DNA binding affinity of these molecules by using thermal denaturation experiments with natural DNA. This technique is easy to use and can serve as a preliminary experiment to investigate DNA binding affinity. Further experiments would then be carried out if binding is observed.

8.2. Results: Thermal Denaturation Assays

Thermal denaturation assays were carried out using natural DNA for all acridine-guanidine binder conjugates to determine if binding was observed. Furthermore, for comparison purposes, we measured the binding affinity of the starting material (9-acridine carboxylic acid), the known intercalator (proflavin) and the acridine *p*-toluidine (**28**). The results are presented in Table 8.1.

Table 8.1.- Results for the Thermal Denaturation Assays for the acridine conjugates. Proflavin and 9-acridine carboxylic acid are shown for comparison

Compound Number	X	ΔT_m /°C
Proflavin	-	3.4
9-Acridine carboxylic acid	-	4
28	-	2
29	CH_2	2
30	CH_2CH_2	2
31	NH	2

[a] Salmon testes DNA melting temperature was found to be 68 °C

These acridine/binder conjugates, could act upon DNA in three possible ways; (i) only by intercalation, (ii) only by binding into the minor groove or (iii) by both intercalating and minor groove binding simultaneously. From the results shown in Table 8.1, we can see that none of these possible three scenarios has occurred. Regarding the possibility of only DNA intercalation, we observe that by comparing the binding affinity of proflavin and 9-acridine carboxylic acid with that of compounds **28-31** our compounds display a decreased binding affinity which is an unexpected result. However, this could be due to the way that the

rigid amide bond attached to the acridine 9 position does not allow the acridine moiety to insert fully in between the DNA base pairs and would, therefore, result in a decreased binding affinity.

Regarding the second possible scenario, binding to the minor groove, and comparing the acridine conjugates **29-31** with the corresponding asymmetric dicationic molecules we observe that, for example compound **5h**, which only differs from compound **31** in the substitution of a 2-aminoimidazolinium group by a 9-amidoacridine one (Figure 8.2), binds much more strongly to DNA (ΔT_m = 12 °C). However, compound **5c**, which is a *mono*-guanidinium derivative (Family I) and has an NH_2 group in the place of the 9-amidoacridine moiety (Figure 8.2), shows the same DNA affinity to **31** (ΔT_m = 2 °C). Thus, the presence of a second "guanidine-like" cation seems to be an essential requirement for the minor groove binding of these compounds.

Obviously, if no intercalation and no minor groove binding are observed individually, no possible simultaneous interactions could take place, discarding the third possible scenario. Therefore, these results seem to indicate that for good binding to DNA we need to (a) use a more flexible link to the acridine molecule than CONH to allow the chromophore to intercalate more freely and (b) to use a dicationic minor groove binder moiety to allow for optimum electrostatic and hydrogen bond interactions to occur within the groove. In that way, at least one of the possible DNA dual interactions would take place and hopefully both could occur.

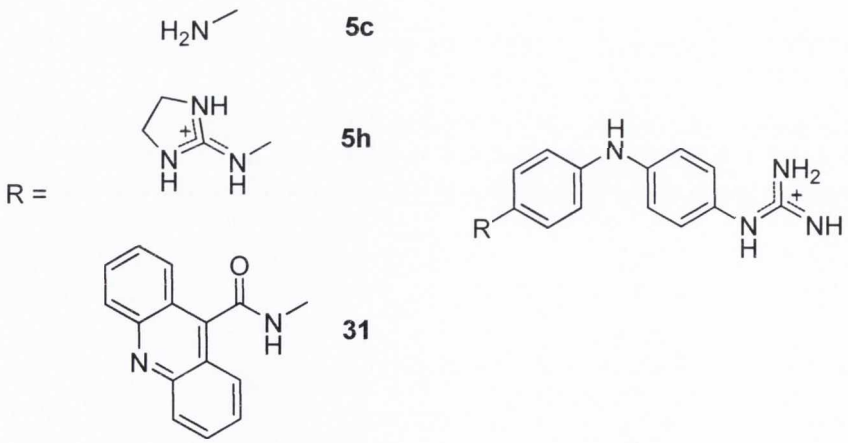


Fig. 8.2. Comparison between the structures of ***5c***, ***5h***, *and* ***31***.

8.3 Conclusions

This chapter describes the experiments that were carried out to determine the DNA binding affinity of the acridine containing molecules using thermal denaturation experiments. It was observed that these compounds display poor affinity for salmon sperm DNA, and that they show decreased DNA affinity than 9-acridine carboxylic acid or proflavin. As well, their DNA binding results were lower than those obtained for the asymmetric dicationic structures **3h**-**10h** but similar to those obtained for the *mono*-functionalised molecules **3c**-**10c**. However cytotoxicity experiments still need to be carried out to determine if they display an increased cytotoxicity relative to the compounds discussed in chapter 6.

For the design of future compounds, several features could be considered. For example,

(i) we could extend the linker between the guanidine functionality and the acridine core,

(ii) connect these moieties with a linker different than CONH or

(iii) keep both "guanidine-like" cations in the minor groove moiety.

Chapter 9

Conclusions and Future Work

9.1. Conclusions

- We have successfully prepared two family of precursor molecules (the *mono*-Boc protected guanidines and 2-aminoimidazolines) used for the preparation of the *mono*-functionalised Families I and II. We were able to deduce that the functionalisation with the 2-aminoimidazoline group gave slightly higher yields in most cases than the corresponding guanidinylation. In addition, we observed increased yields with molecules that contain electron donating linkers in comparison to those that contain electron withdrawing ones due to the increased nucleophilicity of the amino groups.

- The *mono*-guanidine compounds (Family I) were directly prepared and used as precursors for the synthesis of the asymmetric *di*-functionalised derivatives (Family III) and, thus, the synthetic strategy was to introduce the Boc protected guanidine group first to avoid decomposition problems. Then, the Boc protected 2-aminoimidazoline system was generated to proceed afterwards to the full Boc deprotection using trifluoroacetic acid and the generation of the hydrochloride salts using basic Amberlyte resin.

- With the help of detailed NMR analysis, we were able to prepare one of the asymmetric molecules with the amide linker; however, this was not the case for the other molecule as it would have required the initial addition of the 2-aminoimidazoline functional group. Due to previous decomposition problems associated with this functional group it was decided to add this group on last. A retrosynthetic pathway was designed in order to investigate the ideal synthetic route and it was decided to make use of orthogonal protection of the amino ends with Boc and Cbz groups so that we could introduce the Boc protected guanidine group first followed by the introduction of the imidazolidine functional group. This synthesis was carried out; however, problems with the final 2-aminoimidazolidylation step did not enable the preparation of the desired asymmetric amide derivative to work.

- The pK_a of some of the compounds from three Familes (I, II and III) was determined to assess the protonation state of the molecule at physiological pH. It was found that in the *mono*-functionalised molecules, the NH_2 groups have pK_a values less than that of physiological pH, whereas the highly basic guanidine and 2-aminoimidazoline

functionalities were found to be protonated at physiological pH. Thus, for the mono functionalised molecules (Families I and II) whose pK_a was measured, we found them to be *mono*-cationic whereas the di-functionalised molecules (Family III) were dicationic.

- Following from these experiments, the DNA binding affinity of these molecules to natural DNA was investigated by using thermal denaturation experiments. The *mono*-functionalised molecules displayed poor affinity to natural DNA. In contrast, the di-functionalised molecules displayed a high affinity for natural DNA. It was observed from these experiments that the binding was dependent on the linker which governed many of their properties such as geometry, length and the number of hydrogen bonding groups on the molecule. From this technique we were also able to investigate their sequence selectivity. It was observed that all of them, apart from **3h** (linker = CH_2), displayed selectivity for AT sequences which is common for minor groove binders.

- In order to obtain more information about the binding of these dications such as the equilibrium binding constants, UV and CD titration experiments were carried out to evaluate these quantities by using scatchard plot analysis. These results were found to be in agreement with the thermal denaturation assay results. Then, a reverse salt titration was performed to confirm that the compound was dicationic on DNA binding. From these experiments the free energy could be split into an electrostatic and a non electrostatic term. From these results we observed that the binding was dominated by the electrostatic interactions between the molecule and the DNA backbone.

- Circular dichroism experiments were carried out to obtain numerical values to characterise the binding to AT sequences. Further information was obtained from these experiments such as the structural changes that occurred to the DNA backbone upon binding, i.e. the widening of the minor groove. Also, large induced signals were observed upon binding of these molecules confirming that they are minor groove binders.

- Following on from this, linear dichroism was carried out to determine the mode of binding to DNA. From these experiments, we observed positive induced signals confirming that both **5h** and **8h** molecules (linkers = NH and CO respectively) bind to the minor groove of DNA.

- Surface Plasmon Resonance experiments were performed to investigate the selectivity of our molecules for three different sequences. Interestingly, it was found that the dication **6h** (linker = O) displays the highest selectivity for AATT over TTAA sequences. An unexpected result was observed for **8h** (linker = CO) which in these experiments showed a good binding affinity for GC sequences.

- The thermodynamics of the binding to DNA (both natural and AT) of several di-cations was investigated by ITC, showing, in all cases, a favourable negative enthalpy and positive entropy values upon binding of the molecules. The final aspect described in this chapter is the crystal structure obtained for the *mono*-guanidinium derivative **4c**.

- The cytotoxicity of these compounds was meassured against the MCF-7 cell line and IC_{50} values ranging from 1.5 to 63.5 μM were obtained providing us with information of the most favourable linkers when preparing new potential anticancer compounds. It was observed that the derivatives with the less polar linkers displayed the most favourable results, indicating that for future synthesis, molecules with a non polar linker should be considered. Also, it was noted that there was some correlation between the IC_{50} values and the lipophilicity of the compounds, indicating that the cytotoxicity could be related to the diffusion of the compounds into the cell prior to DNA binding. However, other cellular mechanisms of aciton cannot been discarded.

- A series of acridine containing dual agents were prepared by a two step mechanism involving firstly, the reaction between the carboxylic acid was reacted with $SOCl_2$ in the presence of a catalytic amount of DMF to form the acid chloride and, secondly, the reaction of this acid chloride with a series of Boc protected *mono*-guanidine compounds to yield the expected conjugated derivatives upon deprotection.

- Finally, thermal denaturation experiments were carried out to determine the DNA binding affinity of these acridine conjugates. It was observed that these compounds display poor affinity for salmon sperm DNA, and that they show decreased DNA affinity than 9-acridine carboxylic acid or proflavin. As well, their DNA binding results were lower than those obtained for the asymmetric dicationic structures **3h-10h** but similar to those obtained for the *mono*-functionalised molecules **3c-10c**.

9.2 Future Work

- Further research could be carried out by preparing a series of minor groove binders by changing the cation to introduce an extra point of contact with the minor groove. The aromatic moiety could also be changed from a phenyl ring to a pyrrole ring to imitate the structure of netropsin and distamycin. An interesting idea could be the preparation of a molecule with the same scaffold but having the guanidine group as a linker for extra hydrogen bonds donors. Further molecules could be prepared by adding an electron withdrawer onto the phenyl ring so as to increase the pK_a of the cation.

- A possible route that could be followed up is the synthesis of longer minor groove binders by having a tri-aromatic scaffold rather than a di-aromatic one. This would increase the points of contacts between the compound and the DNA minor groove.

- Another idea could be the complexation of our minor groove binders to different metals such as iron or ruthenium. This would probably form a major groove binder that could exhibit its cytotoxicity by binding to the major grove, thus inhibiting some DNA processes.

- Further acridine conjugates could be prepared by changing the amide linker for a more flexible one to enable the acridine intercalate efficiently and exhibit its cytotoxicity. Further compounds could be synthesised by varying the linker length between the acridine chromophore and the minor groove binding scaffold. This would hopefully increase the binding affinity and direct the conjugate to AT sequences. Work in this direction is ongoing in Rozas' laboratory.

- The acridine conjugates could be complexed to a metal such as iron(II). This could potentially increase their cytotoxicity against tumour cell lines.

- Further physicohemical experiments could be carried out on the aymmetric minor groove binders. For example, Isothermal Titration Calorimetry could be carried out for all the compounds (**3h-11h** and **14a**) and DNAase I Footprinting could be attempted for further studies on the selectivity of these compounds.

- Further cytotoxicity experiments could be carried out on the asymmetric minor groove on different tumour cell lines to investigate if they exhibit any selectivity for a particular one.
- Cytotoxicity experiments need to be carried out on the acridine conjugates to see if they exhibit increased cytotoxicity relative to the asymmetric minor groove binders. These should be carried out on the MCF-7 cell line and possibly other tumour cell lines.

Chapter 10

Experimental Section

10.1. Chemistry

All the commercial chemicals were obtained from Sigma-Aldrich or Fluka and were used without further purification. Deuterated solvents for NMR use were purchased from Apollo. Dry solvents were prepared using standard procedures, according to Vogel,[1] with distillation prior to use. Chromatographic columns were run using Silica gel 60 (230-400 mesh ASTM) or Aluminium Oxide (activated, Neutral Brockman I STD grade 150 mesh). Solvents for synthesis purposes were used at GPR grade. Analytical TLC was performed using Merck Kieselgel 60 F_{254} silica gel plates or Polygram Alox N/UV_{254} aluminium oxide plates. Visualisation was by UV light (254 nm). NMR spectra were recorded on a Bruker DPX-400 Advance spectrometer, operating at 400.13 MHz and 600.1 MHz for ^{1}H-NMR and 100.6 MHz and 150.9 MHz for ^{13}C-NMR. Shifts are referenced to the internal solvent signals. NMR data were processed using Bruker Win-NMR 5.0 software. Electrospray mass spectra were recorded on a Mass Lynx NT V 3.4 on a Waters 600 controller connected to a 996 photodiode array detector with methanol, water or ethanol as carrier solvents. Melting points were determined using an Electrothermal IA9000 digital melting point apparatus and are uncorrected. Infrared spectra were recorded on a Mattson Genesis II FTIR spectrometer equipped with a Gateway 2000 4DX2-66 workstation and on a Perkin Elmer Spectrum One FT-IR Spectrometer equipped with Universal ATR sampling accessory. Elemental analysis was carried out at the Microanalysis Laboratory, School of Chemistry and Chemical Biology, University College Dublin. LogP and cLogP values have been calculated using the ChemDraw Ultra package, version 12.

10.1.1. Preparation of the compounds: General procedures

Method A: General method for the preparation of the mono Boc-protected amines

10.0 mmol of TEA and 10.0 mmol of di-*tert*-butyl dicarbonate (Boc_2O) were added at 0 °C over a solution containing an excess (30.0 mmol) of the corresponding diamine in DCM (50 mL). The resulting mixture was stirred at 0 °C for 1 h and 16 h more at room temperature. Then, the reaction was concentrated under vacuum to give a residue that was purified by silica gel column chromatography, eluting with the appropriate hexane:EtOAc mixture. The *mono*-Boc protected compound were obtained as a solid. The excess of unreacted starting material diamine was recovered from the column.

Method B: General method for the preparation of the di-Boc-protected amino guanidines

The corresponding amine (3.0 mmol) was treated in DCM or DMF (5 ml) at 0 °C with 3.3 mmol of mercury (II) chloride, 3.0 mmol of *N,N'*-di(*tert*-butoxycarbonyl)thiourea and 9.3 mmol of TEA. The resulting mixture was stirred at 0 °C for 1 hour and for the appropriate duration at room temperature. Then, the reaction mixture was diluted with EtOAc and filtered through a pad of Celite to eliminate the mercury sulfide formed. The filter cake was rinsed with EtOAc. The organic phase was extracted with water (2×30 mL), washed with brine (1×30 mL), dried over anhydrous Na_2SO_4 and concentrated under vacuum to give a residue that was purified by neutral alumina column flash chromatography, eluting with the appropriate hexane:EtOAc mixture.

Method C: General method for the preparation of the di-Boc protected amino imidazolidines

The corresponding amine (3.0 mmol) was treated in DCM or DMF (5 ml) at 0 °C with 3.3 mmol of mercury (II) chloride, 3.0 mmol of *N,N'*-di(*tert*-butoxycarbonyl)imidazolidine-2-thione and 9.3 mmol of TEA. The resulting mixture was stirred at 0 °C for 1 hour and for the appropriate duration at room temperature. Then, the reaction mixture was diluted with EtOAc and filtered through a pad of Celite to eliminate the mercury sulfide formed. The filter cake was rinsed with EtOAc. The organic phase was extracted with water (2×30 mL), washed with brine (1×30 mL), dried over anhydrous Na_2SO_4 and concentrated under vacuum to give a residue that was purified by neutral alumina column flash chromatography, eluting with the appropriate hexane:EtOAc mixture.

Method D: General method for the preparation of the mono Boc protected guanidines

The starting diamine (9 mmol) was treated in DCM or DMF (5 ml) at 0 °C with 3.3 mmol of mercury (II) chloride, 3.0 mmol of *N,N'*-di(*tert*-butoxycarbonyl) thiourea 9.3 mmol of TEA. This mixture was left stirring for 1 hour at 0°C and a further 23 hours at room temperature. The resulting mixture was diluted with EtOAc and filtered through a pad of Celite. The filter cake was rinsed with EtOAc. The organic phase was washed with water (2×30 ml), dried over Na_2SO_4 and concentrated under vacuum. Purification by flash chromatography with silica gel yielded the required product.

Method E: General method for the preparation of the di-Boc protected 2-aminoimidazoline-guanidine family

The corresponding *mono*-guanidine derivative (3 mmol), was treated in DCM or DMF (5 ml) at 0 °C with 3.3 mmol of mercury (II) chloride, 3.0 mmol of *N,N'*-di(*tert*-butoxycarbonyl) imidazolidine-2-thione and 9.3 mmol of TEA. This mixture was left stirring for 1 hour at 0°C and a further 23 hours at room temperature. The resulting mixture was diluted with EtOAc and filtered through a pad of Celite. The filter cake was rinsed with EtOAc. The organic phase was washed with water (2×30 ml) dried over Na_2SO_4 and concentrated under vacuum. Purification by flash chromatography with alumina gel afforded the required product.

Method F: General method for the Boc deprotection and preparation of the hydrochloride salts

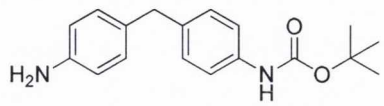
A solution of the corresponding Boc protected derivative (0.5 mmol) in DCM:TFA (1:1) (20 ml) was stirred at room temperature for 3 h; then, the solvent was eliminated under vacuum to generate the trifluoroacetate salt. This salt was dissolved in 20 ml of water and treated for 24 h with IRA400 Amberlyte resin in its Cl^- form. Then, the resin was removed by filtration and the aqueous solution washed with DCM (2×10 ml). Evaporation of the water afforded the corresponding pure hydrochloride salt.

Method G: General method for the preparation for the Boc protected acridine conjugates:

A catalytic amount of DMF was added to a solution of 9-acridine carboxylic acid (1 mmol) in $SOCl_2$ (10 ml), and was refluxed for 5 hours. Removal of the $SOCl_2$ yielded the required compound as a yellow solid. The yellow solid was used without further purification and was dissolved in CH_3CN (20 ml) with NEt_3 under argon. A solution of the corresponding amine (1 mmol) in CH_3CN (10 ml) was added to the acridine solution at 0°C for one hour and at room temperature for 3 hours. The crude product was filtered and washed with EtOAc. Recrystallisation from MeOH yielded the required compound. Purification by silica gel chromatography gave the required molecules.

10.1.2. Preparation of the compounds

4-(4-*tert*-Butoxycarbonylaminobenzyl)aniline (3a[2])



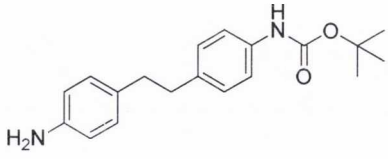
Following *Method* A and after purification with flash chromatography (Hexane:Ethyl Acetate 3:2), 2437 mg (82%) of **3a**[1] as a white solid were obtained. Mp: 129-131 °C.

δ_H (400 MHz, $CDCl_3$): 1.51 (s, 9H, $(CH_3)_3$), 3.60 (broad s, 2H, NH_2), 3.82 (s, 2H, CH_2), 6.40 (broad s, 1H, NH), 6.62 (d, 2H, J = 8.5 Hz, Ar.), 6.96 (d, 2H, J = 8.5 Hz, Ar.), 7.10 (d, 2H, J = 8.5 Hz, Ar.), 7.26 (d, 2H, J = 8.5 Hz, Ar.)

δ_C (100 MHz, $CDCl_3$): 28.2 ($(CH_3)_3$), 40.1 (CH_2), 80.0 ($\underline{C}(CH_3)_3$), 115.1, 118.6, 129.0, 129.4, 131.1, 136.1, 136.3, 144.2 (Ar.), 152.8 (CO)

ν_{max} (film)/cm^{-1}: 3433, 3363, 3359 (NH, NH_2), 1691 (CO)

4-(4-*tert*-Butoxycarbonylaminophenylethyl)aniline (4a)



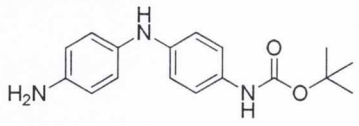
Following *Method* A and after purification with flash chromatography (Hexane:Ethyl Acetate 3:2), 2327 mg (74%) of **4a** as a white solid were obtained. Mp: 124-126 °C.

δ_H (400 MHz, $CDCl_3$): 1.61 (s, 9H, $(CH_3)_3$), 2.80-2.93 (m, 4H, $2CH_2$), 3.62 (broad s, 2H, NH_2), 6.67 (d, 2H, J = 8.0 Hz, Ar.), 6.86 (broad s, 1H, NH), 7.28 (d, 2H, J = 8.0 Hz, Ar.), 7.15 (d, 2H, J = 8.5 Hz, Ar.), 7.41 (d, 2H, J = 8.5 Hz, Ar.)

δ_C (100 MHz, $CDCl_3$): 28.2 ($(CH_3)_3$), 36.9, 37.3 (CH_2), 80.0 ($\underline{C}(CH_3)_3$), 115.0, 118.4, 128.7, 129.0, 131.5, 136.0, 136.4, 144.1 (Ar.), 152.8 (CO)

ν_{max} (film)/cm^{-1}: 3416, 3346, 3211 (NH, NH_2), 1705 (CO)

4-(4-*tert*-Butoxycarbonylaminophenylamino)aniline (5a)



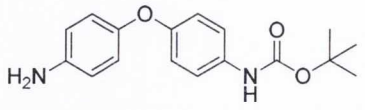
Following *Method* A and after purification with flash chromatography (Hexane:Ethyl Acetate 3:2), 2112 mg (70%) of **5a** as a brown solid were obtained. Mp: 168-170 °C.

δ_H (400 MHz, $CDCl_3$): 1.47 (s, 9H, $(CH_3)_3$), 4.72 (broad s, 2H, NH_2), 6.54 (d, 2H, J = 8.0 Hz, Ar.), 6.74 (d, 2H, J = 8.5 Hz, Ar.), 6.80 (d, 2H, J = 8.5 Hz, Ar.), 7.20 (d, 2H, J = 8.0 Hz, Ar.), 7.28 (broad s, 1H, PhNHPh), 8.96 (broad s, 1H, NHCO)

δ_C (100 MHz, $CDCl_3$): 26.7 ($(CH_3)_3$), 76.8 ($\underline{C}(CH_3)_3$), 113.0, 113.4, 118.4, 119.6, 128.5, 131.3, 140.0, 141.6 (Ar.), 151.5 (CO)

ν_{max} (film)/cm^{-1}: 3403, 3353, 3332, 3232 (NH, NH_2), 1693 (CO)

4-(4-*tert*-Butoxycarbonylaminophenoxy)aniline (6a[3])



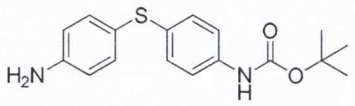
Following *Method* A and after purification with flash chromatography (Hexane:Ethyl Acetate 3:2), 2283 mg (76%) of **6a** as a white solid were obtained. Mp: 129-131 °C.

δ_H (400 MHz, $CDCl_3$): 1.54 (s, 9H, $(CH_3)_3$), 3.63 (broad s, 2H, NH_2), 6.64 (d, 2H, J = 8.6 Hz, Ar.), 6.84 (d, 2H, J = 8.6 Hz, Ar.), 6.88 (d, 2H, J = 9.0 Hz, Ar.), 7.10 (broad s, 1H, NHCO), 7.28 (d, 2H, J = 9.0 Hz, Ar.)

δ_C (100 MHz, $CDCl_3$): 28.1 ($(CH_3)_3$), 79.9 ($\underline{C}(CH_3)_3$), 116.0, 117.7, 120.1, 120.2, 132.8, 142.2, 148.8, 153.1 (Ar.), 153.9 (CO)

ν_{max} (film)/cm^{-1}: 3413, 3343, 3231 (NH, NH_2), 1718 (CO)

4-(4-*tert*-Butoxycarbonylaminophenylsulfanyl)aniline (7a)



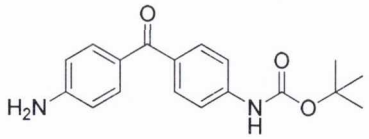
Following *Method* A and after purification with flash chromatography (Hexane:Ethyl Acetate 3:2), 2420 mg (77%) of **7a** as a yellowish solid were obtained. Mp: 108-110 °C.

δ_H (400 MHz, $CDCl_3$): 1.51 (s, 9H, $(CH_3)_3$), 3.79 (broad s, 2H, NH_2), 6.43 (broad s, 1H, NH), 6.65 (d, 2H, J = 8.5 Hz, Ar.), 7.15 (d, 2H, J = 8.5 Hz, Ar.), 7.18-7.27 (m, 4H, Ar.)

δ_C (100 MHz, $CDCl_3$): 28.2 ($(CH_3)_3$), 80.4 ($\underline{C}(CH_3)_3$), 115.7, 119.1, 121.8, 129.5, 132.1, 134.6, 136.4, 146.4 (Ar.), 152.7 (CO)

ν_{max} (film)/cm^{-1}: 3471, 3431, 3311 (NH, NH_2), 1698 (CO)

4-[4-(4-*tert*-Butoxycarbonylaminophenyl)piperazin-1-yl]aniline (8a)



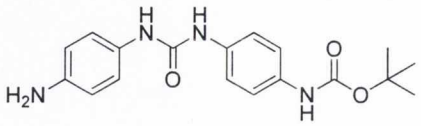
Following *Method A* and after purification with flash chromatography (Ethyl Acetate), 1435 mg (46 %) of **8a** as a yellowish solid were obtained. Mp: decomposes over 220 °C.

δ_H (400 MHz, $CDCl_3$): 1.50 (s, 9H, $(CH_3)_3$), 6.08 (broad s, 2H, NH_2), 6.60 (d, 2H, J = 9.0 Hz, Ar.), 7.50 (d, 2H, J = 9.0 Hz, Ar.), 7.58 (m, 4H, Ar.), 9.72 (broad s, 1H, NH).

δ_C (100 MHz, $CDCl_3$): 28.5 ($(CH_3)_3$), 80.0 ($\underline{C}(CH_3)_3$), 112.9, 117.5, 122.4, 123.8, 132.7, 132.8, 143.0, 153.1 (Ar.), 153.8 (CO)

ν_{max} (film)/cm^{-1}: 3483, 3380, 3311 (NH, NH_2), 1716, 1624 (CO)

{4-[3-(4-Aminophenyl)ureido]phenyl}carbamic acid *tert*-butyl ester (9a)



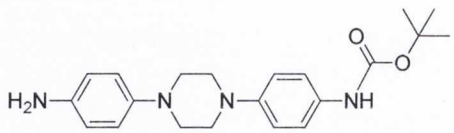
Following *Method* A and after purification with flash chromatography (Ethyl Acetate), 1806 mg (49%) of **9a** as a white solid were obtained. Mp: decomposes over 205 °C.

δ_H (400 MHz, DMSO-d_6): 1.48 (s, 9H, $(CH_3)_3$), 4.78 (broad s, 2H, NH_2), 6.54 (d, 2H, J = 8.5 Hz, Ar.), 7.10 (d, 2H, J = 8.5 Hz, Ar.), 7.30-7.41 (m, 4H, Ar.), 8.09 (broad s, 1H, NH), 8.36 (broad s, 1H, NH), 9.18 (broad s, 1H, NH)

δ_C (100 MHz, DMSO-d_6): 26.7 ($(CH_3)_3$), 77.2 ($\underline{C}(CH_3)_3$), 112.7, 117.0, 117.4, 119.2, 127.3, 132.0, 133.2, 142.5 (Ar.), 151.4, 151.6 (CO)

ν_{max} (film)/cm^{-1}: 3478, 3413, 3370, 3289, 3028 (NH, NH_2), 1708, 1628 (CO)

4-[4-(4-*tert*-Butoxycarbonylaminophenyl)piperazin-1-yl]aniline (10a)



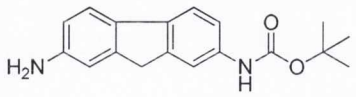
Following Method A and after purification with flash chromatography (Ethyl Acetate), 1806 mg (49%) of **10a** as a yellowish solid were obtained. Mp: 220-222 °C.

δ_H (400 MHz, $CDCl_3$): 1.52 (s, 9H, $(CH_3)_3$), 3.16-3.23 (m, 4H, CH_2 Pip), 3.24-3.31 (m, 4H, CH_2), 3.47 (broad s, 2H, NH_2), 6.37 (broad s, 1H, NH), 6.69 (d, 2H, J = 8.5 Hz, Ar.), 6.87 (d, 2H, J = 9.0 Hz, Ar.), 6.93 (d, 2H, J = 9.0 Hz, Ar.), 7.22-7.33 (m, 2H, Ar.)

δ_C (400 MHz, $CDCl_3$): 28.4 ($(CH_3)_3$), 50.2, 51.1 (CH_2), 80.2 ($\underline{C}(CH_3)_3$), 116.2, 117.2, 118.8, 120.1, 131.1, 140.4, 144.4, 147.5 (Ar.), 153.1 (CO)

ν_{max} (film)/cm^{-1}: 3360, 3292, 3196 (NH, NH_2), 1701 (CO)

(7-Amino-9*H*-fluoren-2-yl)-carbamic acid *tert*-butyl ester (11a)



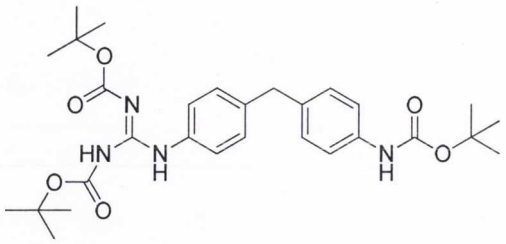
Following *Method* A and after purification with flash chromatography (Hexane:Ethyl Acetate, 1:4), 2345 mg (79%) of **11a** as a orange solid were obtained. Mp: 152-154 °C.

δ_H (400 MHz, $CDCl_3$): 1.54 (s, 9H, $(CH_3)_3$), 3.72 (broad s, 2H, NH_2), 3.78 (s, 2H, CH_2), 6.51 (broad s, 1H, NH), 6.69 (d, 1H, J = 7.8 Hz, Ar.), 6.86 (s, 1H, Ar.), 7.14 (d, 1H, J = 7.8 Hz, Ar.), 7.49 (d, 1H, J = 7.8 Hz, Ar.), 7.52 (d, 1H, J = 7.8 Hz, Ar.), 7.67 (s, 1H, Ar.)

δ_C (100 MHz, $CDCl_3$): 28.3 ($(CH_3)_3$), 36.8 (CH_2), 80.2 ($\underline{C}(CH_3)_3$), 111.8, 113.8, 115.5, 117.3, 118.6, 120.0, 132.6, 135.7, 137.4, 143.3, 144.8, 145.1 (Ar.), 152.9 (CO)

ν_{max} (film)/cm^{-1}: 3391, 3360, 3311 (NH, NH_2), 1697 (CO)

_N_,_N_'-di(*tert*-Butoxycarbonyl)-_N_''-[4-(4-*tert*-butoxycarbonylaminobenzyl)phenyl] guanidine (3b)



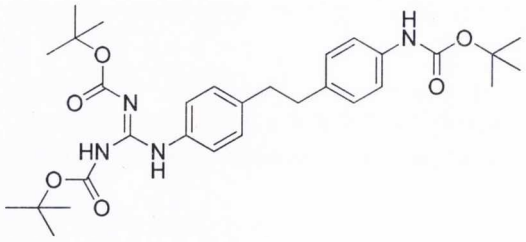
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 2:1) 1249 mg (77 %) of **3b** as a white solid were obtained. Mp: 101-103 °C.

δ_H (400 MHz, $CDCl_3$): 1.52 (s, 9H, $(CH_3)_3$), 1.53 (s, 9H, $(CH_3)_3$), 1.55 (s, 9H, $(CH_3)_3$), 3.89 (s, 2H, CH_2), 6.64 (broad s, 1H, NHCO), 7.03-7.14 (m, 4H, Ar.), 7.29 (d, 2H, J = 7.5 Hz, Ar.), 7.50 (d, 2H, J = 8.0 Hz, Ar.), 10.28 (broad s, 1H, NH), 11.67 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.1, 28.2 ($(CH_3)_3$), 40.5 (CH_2), 79.4, 80.1, 83.5 ($\underline{C}(CH_3)_3$), 118.6, 122.3, 129.1, 129.3, 134.6, 135.5, 136.4, 137.8 (Ar.), 152.8, 153.2, 153.5 (CO), 163.5 (CN)

ν_{max} (film)/cm^{-1}: 3290, 3185, 3150 (NH), 1719, 1626, 1603 (CO, CN)

_N_,_N_'-di(*tert*-Butoxycarbonyl)-_N_''-{4-[2-(4-*tert*-butoxycarbonylaminophenyl)ethyl] phenyl}guanidine (4b)



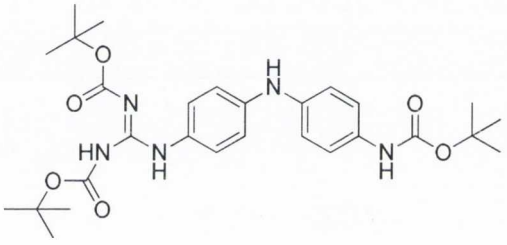
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 5:2) 850 mg (51 %) of **4b** as a white solid were obtained. Mp: 143-145 °C.

δ_H (400 MHz, $CDCl_3$): 1.51 (s, 9H, $(CH_3)_3$), 1.52 (s, 9H, $(CH_3)_3$), 1.54 (s, 9H, $(CH_3)_3$), 2.83 (s, 4H, $2CH_2$), 6.48 (broad s, 1H, NHCO), 7.05-7.12 (m, 4H, Ar.), 7.26 (d, 2H, J = 8.2 Hz, Ar.), 7.49 (d, 2H, J = 8.2 Hz, Ar.), 10.28 (broad s, 1H, NH), 11.65 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.2, 28.3 ($(CH_3)_3$), 37.2, 37.4 (CH_2), 79.5, 80.2, 83.6 ($\underline{C}(CH_3)_3$), 118.5, 122.1, 128.8, 128.9, 134.6, 136.2, 136.3, 138.2 (Ar.), 153.3, 153.5 (CO), 163.6 (CN)

ν_{max} (film)/cm^{-1}: 3295, 3250, 3118 (NH), 1723, 1642, 1621, 1605 (CO, CN)

_N_,_N'_-di(*tert*-butoxycarbonyl)-_N''_-[4-(4-*tert*-butoxycarbonylaminophenylamino)phenyl] guanidine (5b)



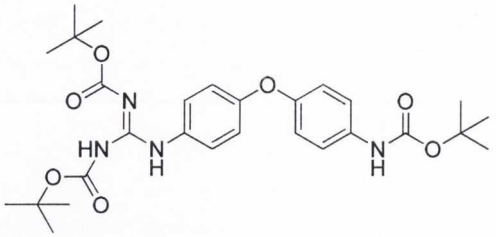
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 3:2) 1254 mg (77 %) of **5b** as a white solid were obtained. Mp: decomposes over 190 °C.

δ_H (400 MHz, $CDCl_3$): 1.45 (s, 9H, $(CH_3)_3$), 1.50 (s, 9H, $(CH_3)_3$), 1.53 (s, 9H, $(CH_3)_3$), 5.86 (broad s, 1H, PhNHPh), 6.71 (broad s, 1H, NHCO), 6.82 (d, 2H, J = 8.6 Hz, Ar.), 6.90 (d, 2H, J = 8.0 Hz, Ar.), 7.21 (d, 2H, J = 8.0 Hz, Ar.), 7.31 (d, 2H, J = 8.6 Hz, Ar.), 10.11 (broad s, 1H, NH), 11.67 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.1, 28.3 ($(CH_3)_3$), 79.4, 79.8, 83.4 ($\underline{C}(CH_3)_3$), 116.8, 118.9, 120.1, 124.1, 128.6, 131.8, 138.4, 141.3 (Ar.), 153.1, 153.2, 153.9 (CO), 163.5 (CN)

ν_{max} (film)/cm^{-1}: 3304, 3124, 3039 (NH), 1720, 1647, 1621, 1595 (CO, CN)

_N_,_N_'-di(_tert_-butoxycarbonyl)-_N_''-[4-(4-_tert_-butoxycarbonylaminophenoxy)phenyl] guanidine (6b)



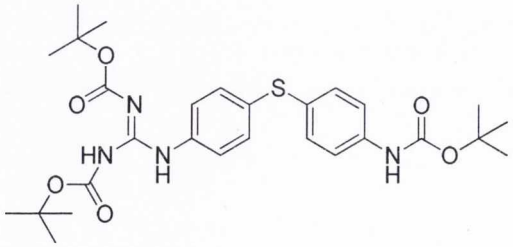
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 5:2) 864 mg, (53 %) of **6b** as a white solid were obtained. Mp: 149-151 °C.

δ_H (400 MHz, $CDCl_3$): 1.50 (s, 9H, $(CH_3)_3$), 1.52 (s, 9H, $(CH_3)_3$), 1.54 (s, 9H, $(CH_3)_3$), 6.44 (broad s, 1H, NHCO), 6.91-6.98 (m, 4H, Ar.), 7.32 (d, 2H, J = 9.0 Hz, Ar.), 7.53 (d, 2H, J = 8.5 Hz, Ar.), 10.27 (broad s, 1H, NH), 11.64 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.1, 28.2, 28.3 ($(CH_3)_3$), 79.6, 80.5, 83.7 ($\underline{C}(CH_3)_3$), 118.6, 119.6, 120.2, 123.8, 131.8, 133.8, 152.6, 152.9 (Ar.), 153.3, 153.6, 154.6 (CO), 163.5 (CN)

ν_{max} (film)/cm^{-1}: 3342, 3217, 3145 (NH), 1719, 1654, 1623, 1602 (CO, CN)

_N_,_N_'-di(_tert_-butoxycarbonyl)-_N_''-[4-(4-_tert_-butoxycarbonylaminophenylsulfanyl)phenyl] guanidine (7b)



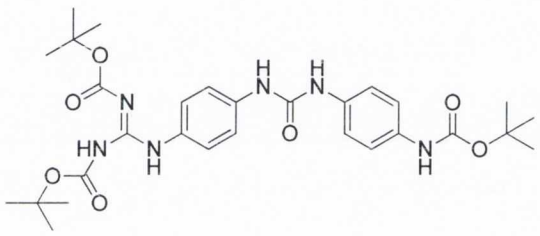
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 5:2) 890 mg, (53 %) of **7b** as a white solid were obtained. Mp: 122-124 °C.

δ_H (400 MHz, $CDCl_3$): 1.43-1.52 (m, 27H, $(CH_3)_3$), 6.91 (broad s, 1H, NHCO), 7.18 (d, 2H, J = 8.5 Hz, Ar.), 7.25-7.38 (m, 4H, Ar.), 7.50 (d, 2H, J = 8.8 Hz, Ar.), 10.32 (broad s, 1H, NH), 11.65 (broad s, 1H, NH)

δ_C (400 MHz, $CDCl_3$): 27.9, 28.0, 28.2 ($(CH_3)_3$), 79.6, 80.4, 83.7 ($\underline{C}(CH_3)_3$), 119.0, 122.7, 128.0, 130.2, 132.9, 135.1, 138.1, 152.5 (Ar.), 153.1, 153.4 (CO), 163.3 (CN)

ν_{max} (film)/cm^{-1}: 3261, 3255, 3171 (NH), 1719, 1623, 1588 (CO, CN)

4-[2,3-di(*tert*-butoxycarbonyl)guanidine]-4'-(*tert*-butoxycarbonylamino)-diphenyl urea (9b)



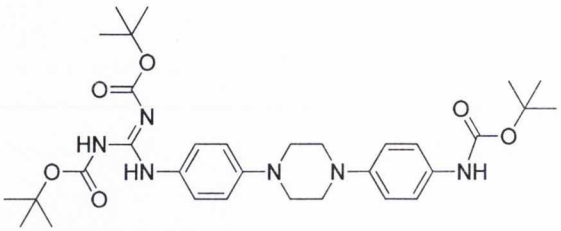
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 1:1) 1632 mg, (93 %) of **9b** as a white solid were obtained. Mp: decomposes over 235 °C.

δ_H (400 MHz, $CDCl_3$): 1.41 (s, 9H, $(CH_3)_3$), 1.46 (s, 9H, $(CH_3)_3$), 1.54 (s, 9H, $(CH_3)_3$), 6.96-7.20 (m, 9H, NH & Ar.), 7.60-7.70 (m, 2H, NHCONH), 10.03 (broad s, 1H, NH), 11.63 (broad s, 1H, NH);

δ_C (100 MHz, $CDCl_3$): 27.9, 28.0, 28.3 ($(CH_3)_3$), 79.6, 79.9, 83.8 ($\underline{C}(CH_3)_3$), 119.6, 119.8, 120.9, 125.0, 129.8, 133.6, 133.8, 137.2 (Ar.), 153.0, 153.2, 153.6, 155.3 (CO, CN), 166.1 (CO)

ν_{max} (film)/cm^{-1}: 3299, 3169 (NH), 1720, 1692, 1644, 1613 (CO, CN)

_N,N'_-di(_tert_-butoxycarbonyl)-_N''_-{4-[4-(4-_tert_-butoxycarbonylaminophenyl)piperazin-1-yl]phenyl}guanidine (10b)



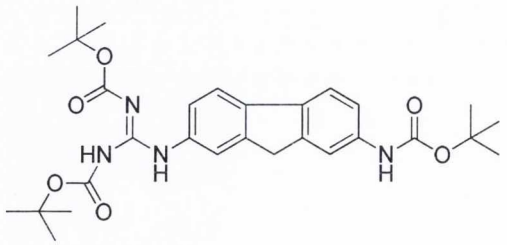
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 7:3) 1394 mg, (76 %) of **10b** as a white solid were obtained. Mp: 108-110 °C.

δ_H (400 MHz, $CDCl_3$): 1.50 (s, 9H, $(CH_3)_3$), 1.51 (s, 9H, $(CH_3)_3$), 1.54 (s, 9H, $(CH_3)_3$), 3.26 (s, 4H, CH_2), 3.27 (s, 4H, CH_2 Pip), 6.44 (broad s, 1H, NHCO), 6.90-6.96 (m, 4H, Ar.), 7.28 (d, 2H, J = 8.0 Hz, Ar.), 7.48 (d, 2H, J = 8.5 Hz, Ar.), 10.18 (broad s, 1H, NH), 11.66 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.1, 28.3 ($(CH_3)_3$), 49.6, 49.9 (CH_2), 79.3, 79.8, 83.4 ($\underline{C}(CH_3)_3$), 116.6, 117.1, 119.9, 123.6, 129.0, 131.4, 147.1, 148.5 (Ar.), 153.0, 153.2, 153.6 (CO), 163.5 (CN);

ν_{max} (film)/cm^{-1}: 3363, 3297, 3261 (NH), 1723, 1694, 1636, 1623, 1607 (CO, CN)

_N,N'_-di(_tert_-butoxycarbonyl)-_N''_-(7-_tert_-butoxycarbonylamino-9_H_-fluoren-2-yl) guanidine (11b)



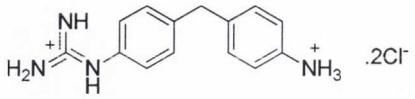
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 3:1) 1249 mg, (77 %) of **11b** as a white solid were obtained. Mp: decomposes over 206 °C.

δ_H (400 MHz, $CDCl_3$): 1.50 (s, 9H, $(CH_3)_3$), 1.53 (s, 9H, $(CH_3)_3$), 1.55 (s, 9H, $(CH_3)_3$), 3.79 (s, 2H, CH_2), 6.87 (broad s, 1H, NHCO), 7.11 (d, 1H, J = 8.0 Hz, Ar.), 7.38 (d, 1H, J = 8.0 Hz, Ar.), 7.52 (d, 1H, J = 8.0 Hz, Ar.), 7.57 (d, 1H, J = 8.0 Hz, Ar.), 7.69 (s, 1H, Ar.), 7.77 (s, 1H, Ar.), 10.40 (broad s, 1H, NH), 11.73 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.1, 28.4 ($(CH_3)_3$), 36.9 (CH_2), 79.6, 79.8, 83.6 ($\underline{C}(CH_3)_3$), 114.9, 117.3, 119.2, 119.4, 120.1, 121.8, 134.2, 135.9, 137.2, 138.9, 143.7, 144.2 (Ar.), 152.9, 153.3, 154.1 (CO), 163.3 (CN)

ν_{max} (film)/cm^{-1}: 3305, 3267 (NH), 1717, 1646, 1606 (CO, CN)

Dihydrochloride salt of *N*-[4-(4-aminobenzyl)phenyl]guanidine (3c)



Following *Method* F, 113 mg of the pure dihydrochloride salt **3c** were obtained as a white solid (94%). Mp: 84-86 °C; clogP: 0.66

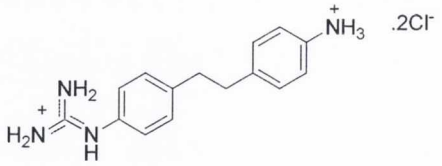
δ_H (400 MHz, D_2O): 4.00 (s, 2H, CH_2), 7.18 (d, 2H, J = 8.5 Hz, Ar.), 7.26-7.34 (m, 4H, Ar.), 7.37 (d, 2H, J = 8.5 Hz, Ar.)

δ_C (100 MHz, D_2O): 39.5 (CH_2), 122.6, 125.1, 127.2, 129.7, 129.8, 131.4, 140.0, 141.7 (Ar.), 155.4 (CN)

HRMS (*m/z* -ES): Found: 241.1452 (M^+ + H). $C_{14}H_{17}N_4$ Requires: 241.1453

Anal. ($C_{14}H_{18}Cl_2N_4 \cdot 0.5H_2O$) Calcd: C, 52.18; H, 5.94; N, 17.39. Found: C, 52.31; H, 6.08; N, 17.01

Dihydrochloride salt of *N*-{4-[2-(4-aminophenyl)ethyl]phenyl}guanidine (4c)



Following *Method* F, 120 mg of the pure dihydrochloride salt **4c** were obtained as a white solid (94%). Mp: decomposes over 190 °C; clogP: 1.039

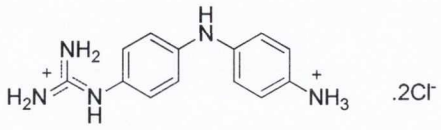
δ_H (400 MHz, D_2O): 2.99 (s, 4H, $2CH_2$), 7.18 (d, 2H, J = 7.5 Hz, Ar.), 7.23-7.31 (m, 4H, Ar.), 7.34 (d, 2H, J = 7.5 Hz, Ar.)

δ_C (100 MHz, D_2O): 35.4, 35.5 ($2CH_2$), 122.2, 125.4, 127.0, 129.6, 129.7, 131.4, 141.0, 142.3 (Ar.), 155.9 (CN)

HRMS (*m/z* -ES): Found: 255.1603 ($M^+ + H$). $C_{15}H_{19}N_4$ Requires: 255.1610

Anal. ($C_{15}H_{20}Cl_2N_4 \cdot 0.4H_2O$) Calcd: C, 53.87; H, 6.27; N, 16.76. Found: C, 54.13; H, 5.96; N, 16.85.

Dihydrochloride salt of *N*-[4-(4-aminophenylamino)phenyl]guanidine (5c)



Following *Method* F, 115 mg of the pure dihydrochloride salt **5c** were obtained as a dark grey solid (95%). Mp: 162-164 °C; clogP: 0.071

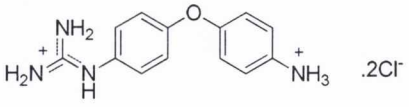
δ_H (400 MHz, D_2O): 6.97-7.11 (m, 6H, Ar.), 7.20 (d, 2H, J = 8.0 Hz, Ar.)

δ_C (100 MHz, D_2O): 117.7, 118.3, 121.6, 123.4, 126.2, 127.0, 141.8, 143.1 (Ar.), 155.9 (CN)

HRMS (*m/z* -ES): Found: 242.1330 ($M^+ + H$). $C_{13}H_{16}N_5$ Requires: 242.1329

Anal. ($C_{13}H_{17}Cl_2N_5 \cdot 1.8H_2O$) Calcd: C, 45.04; H, 5.99; N, 20.20. Found: C, 45.36; H, 5.65; N, 20.18.

Dihydrochloride salt of N-[4-(4-aminophenoxy)phenyl]guanidine (6c)



Following *Method* F, 114 mg of the pure dihydrochloride salt **6c** were obtained as a white solid (94%). Mp: decomposes over 175 °C; clogP: 0.691

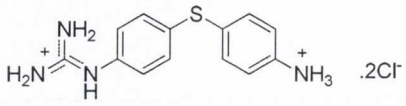
δ_H (400 MHz, D_2O): 7.13 (d, 2H, J = 8.5 Hz, Ar.), 7.17 (d, 2H, J = 8.5 Hz, Ar.), 7.32 (d, 2H, J = 8.5 Hz, Ar.), 7.40 (d, 2H, J = 8.5 Hz, Ar.)

δ_C (100 MHz, D_2O): 119.5, 119.8, 124.2, 124.7, 127.7, 129.3, 155.2 (Ar.), 156.0 (CN), 156.5 (Ar.)

HRMS (*m/z* -ES): Found: 243.1246 ($M^+ + H$). $C_{13}H_{15}N_4O$ Requires: 243.1246

Anal. ($C_{13}H_{16}Cl_2N_4O \cdot 1.0H_2O$) Calcd: C, 46.86; H, 5.44; N, 16.81. Found: C, 46.91; H, 5.12; N, 16.46.

Dihydrochloride salt of *N*-[4-(4-aminophenylsulfanyl)phenyl]guanidine (7c)



Following *Method* F, 124 mg of the pure dihydrochloride salt **7c** were obtained as a white solid (96%). Mp: 82-84 °C; clogP: 0.931

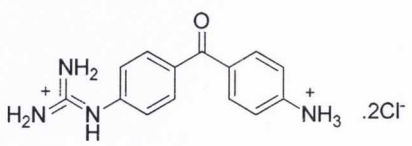
δ_H (400 MHz, D_2O): 7.20 (d, 2H, *J* = 8.5 Hz, Ar.), 7.32 (d, 2H, *J* = 8.0 Hz, Ar.), 7.35-7.41 (m, 4H, Ar.)

δ_C (100 MHz, D_2O): 123.5, 125.8, 128.3, 131.1, 132.6, 132.7, 133.6, 136.4 (Ar.), 155.5 (CN)

HRMS (*m/z* -ES): Found: 259.1017 ($M^+ + H$). $C_{13}H_{15}N_4S$ Requires: 259.1017

Anal. ($C_{13}H_{16}Cl_2N_4S \cdot 1.2H_2O$) Calcd: C, 44.70; H, 5.19; N, 16.04. Found: C, 44.35; H, 4.83; N, 15.67.

Dihydrochloride salt of *N*-[4-(4-aminobenzoyl)phenyl]guanidine (8c)



Following *Method* F, 119 mg of the pure dihydrochloride salt **8c** were obtained as a white solid (93%). Mp: decomposes over 175 °C; clogP: -0.369

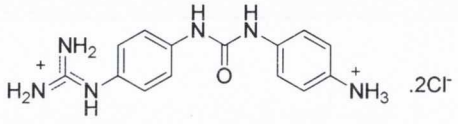
δ_H (400 MHz, D_2O): 7.42 (d, 4H, J = 8.0 Hz, Ar.), 7.79-7.87 (m, 4H, Ar.)

δ_C (100 MHz, D_2O): 121.0, 123.5, 131.5, 131.7, 134.2, 134.4, 137.5, 138.9 (Ar.), 155.3 (CN), 197.4 (CO)

HRMS (*m/z* -ES): Found: 255.1244 (M^+ + H). $C_{14}H_{21}N_2O_2$ Requires: 255.1246

Anal. ($C_{14}H_{16}Cl_2N_4O \cdot 0.4H_2O$) Calcd: C, 50.28; H, 5.06; N, 16.75. Found: C, 50.44; H, 4.84; N, 16.99.

Dihydrochloride salt of 1-(4-aminophenyl)-3-(4-guanidinophenyl)urea (9c)



Following *Method* F, 134 mg of the pure dihydrochloride salt **9c** were obtained as a white solid (94%). Mp: decomposes over 215 °C; clogP: -0.539

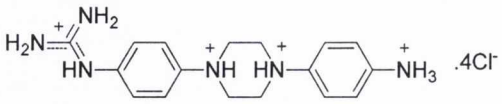
δ_H (400 MHz, D_2O): 7.07 (d, 2H, J = 8.3 Hz, Ar.), 7.23-7.29 (m, 4H, Ar.), 7.35 (d, 2H, J = 8.8 Hz, Ar.)

δ_C (100 MHz, D_2O): 121.0, 121.1, 123.5, 124.5, 126.2, 129.0, 137.2, 138.6 (Ar.), 154.4, 156.0 (CO, CN)

HRMS (*m/z* -ES): Found: 285.1456 (M^+ + H). $C_{14}H_{18}N_6O$ Requires: 285.1464

Anal. ($C_{14}H_{18}Cl_2N_6O \cdot 1.5H_2O$) Calcd: C, 43.76; H, 5.51; N, 21.87. Found: C, 44.09; H, 5.55; N, 21.62.

Tetrahydrochloride salt of *N*-{4-[4-(4-aminophenyl)piperazin-1-yl]phenyl}guanidine (10c)



Following *Method* F, 145 mg of the pure dihydrochloride salt **10c** were obtained as a brown solid (93%). Mp: 168-170 °C; clogP: 0.041

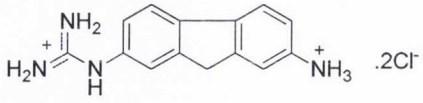
δ_H (400 MHz, D_2O): 3.68-3.77 (m, 4H, CH_2), 3.78-3.86 (m, 4H, CH_2 Pip), 7.35 (d, 2H, J = 9.0 Hz, Ar.), 7.39-7.46 (m, 4H, Ar.), 7.58 (d, 2H, J = 8.5 Hz, Ar.)

δ_C (100 MHz, D_2O): 48.4, 52.3 (CH_2), 119.0, 121.1, 124.0, 124.9, 126.7, 133.7, 141.0, 146.7 (Ar.), 155.6 (CN)

HRMS (*m/z* -ES): Found: 311.1993 ($M^+ + H$). $C_{17}H_{25}N_6O_2$ Requires: 311.1984

Anal. ($C_{17}H_{26}Cl_4N_6 \cdot 2.2H_2O$) Calcd: C, 41.18; H, 6.18; N, 16.95. Found: C, 41.32; H, 5.86; N, 17.05.

Dihydrochloride salt of *N*-(7-Amino-9*H*-fluoren-2-yl)guanidine (11c)



Following *Method* F, 114 mg of the pure dihydrochloride salt **11c** were obtained as a white solid (95%). Mp: decomposes over 176 °C; clogP: 0.526

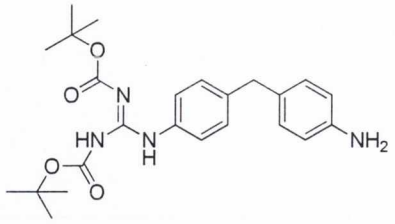
δ_H (400 MHz, D_2O): δ 3.53 (s, 2H, CH_2), 7.06 (d, 1H, J = 8.3 Hz, Ar.), 7.20 (s, 1H, Ar.), 7.24 (d, 1H, J = 7.9 Hz, Ar.), 7.41 (s, 1H, Ar.), 7.52 (d, 1H, J = 7.9 Hz, Ar.), 7.58 (d, 1H, J = 8.3 Hz, Ar.)

δ_C (100 MHz, D_2O): δ 36.2 (CH_2), 119.4, 121.0, 121.2, 121.3, 121.6, 123.8, 128.2, 133.0, 138.9, 140.8, 145.0, 145.3 (Ar.), 156.0 (CN)

HRMS (*m/z* -ES): Found: 239.1297 ($M^+ + H$). $C_{14}H_{15}N_4$ Requires: 239.1297

Anal. ($C_{14}H_{16}Cl_2N_4 \cdot 0.7H_2O$) Calcd: C, 51.93; H, 5.42; N, 17.30. Found: C, 52.14; H, 5.30; N, 16.61.

***N*,*N*'-di(*tert*-butoxycarbonyl)-*N*''-[4-(4-aminobenzyl)phenyl]guanidine (3d)**



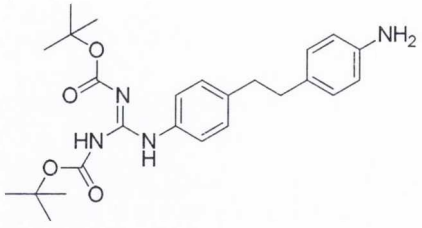
Following *Method* D and after purification with flash chromatography (Hexane:Ethyl Acetate, 2:1) 384 mg, (29 %) of **3d** as a pale yellow solid were obtained. Mp: 106-108 °C.

δ_H (400 MHz, $CDCl_3$): 1.50 (s, 9H, $(CH_3)_3$), 1.53 (s, 9H, $(CH_3)_3$), 3.59 (broad s, 2H, NH_2), 3.82 (s, 2H, CH_2), 6.60 (d, 2H, J = 7.5 Hz, Ar.), 6.94 (d, 2H, J = 7.5 Hz, Ar.), 7.11 (d, 2H, J = 8.0 Hz, Ar.), 7.48 (d, 2H, J = 8.0 Hz, Ar.), 10.26 (broad s, 1H, NH), 11.66 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.9, 28.0 ($(CH_3)_3$), 40.3 (CH_2), 79.3, 83.4 ($\underline{C}(CH_3)_3$), 115.0, 122.1, 129.0, 129.6, 130.8, 134.4, 138.3, 144.4 (Ar.), 153.1, 153.4 (CO), 163.4 (CN)

ν_{max} (film)/cm^{-1}: 3456, 3391, 3259, 3152 (NH, NH_2), 1717, 1633, 1608 (CO, CN)

***N*,*N*'-di(*tert*-butoxycarbonyl)-*N*''-{4-[2-(4-aminophenyl)ethyl]phenyl}guanidine (4d)**



Following *Method* D and after purification with flash chromatography (Hexane:Ethyl Acetate, 2:1) 738 mg, (54 %) of **4d** as a yellow solid were obtained. Mp: decomposes over 245 °C.

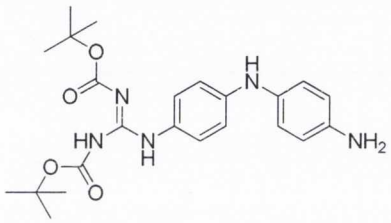
δ_H (400 MHz, $CDCl_3$): 1.51 (s, 9H, $(CH_3)_3$), 1.54 (s, 9H, $(CH_3)_3$), 2.72-2.88 (m, 4H, $2CH_2$), 3.57 (broad s, 2H, NH_2), 6.62 (d, 2H, J = 8.5 Hz, Ar.), 6.97 (d, 2H, J

= 8.5 Hz, Ar.), 7.12 (d, 2H, *J* = 8.5 Hz, Ar.), 7.50 (d, 2H, *J* = 8.5 Hz, Ar.), 10.27 (broad s, 1H, NH), 11.65 (broad s, 1H, NH);

δ_C (100 MHz, $CDCl_3$): 28.0, 28.2 (($CH_3)_3$), 37.0, 37.7 (CH_2), 79.5, 83.6 ($\underline{C}(CH_3)_3$), 115.2, 122.1, 128.8, 129.2, 131.8, 134.5, 138.6, 144.2 (Ar.), 153.3, 153.5 (CO), 163.6 (CN)

ν_{max} (film)/cm^{-1}: 3366, 3261, 3155 (NH, NH_2), 1718, 1623, 1606 (CO, CN)

_N_,_N'_-di(*tert*-butoxycarbonyl)-_N''_-[4-(4-aminophenylamino)phenyl]guanidine (5d)



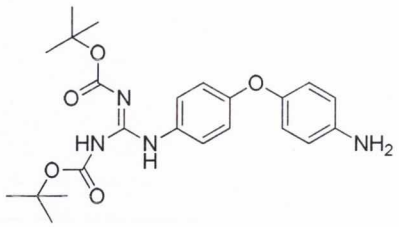
Following *Method* D and after purification with flash chromatography (Hexane:Ethyl Acetate, 2:1) 840 mg, (63 %) of **5d** as a pale pink solid were obtained. Mp: decomposes over 230 °C.

δ_H (400 MHz, $CDCl_3$): 1.49 (s, 9H, ($CH_3)_3$), 1.54 (s, 9H, ($CH_3)_3$), 3.54 (broad s, 2H, NH_2), 5.49 (broad s, 1H, PhNHPh), 6.64 (d, 2H, *J* = 7.5 Hz, Ar.), 6.79 (d, 2H, *J* = 8.0 Hz, Ar.), 6.91 (d, 2H, *J* = 7.5 Hz, Ar.), 7.35 (d, 2H, *J* = 8.0 Hz, Ar.), 10.11 (broad s, 1H, NH), 11.66 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.1 (($CH_3)_3$), 79.3, 83.3 ($\underline{C}(CH_3)_3$), 115.4, 116.0, 122.5, 123.9, 128.0, 134.0, 141.7, 143.0 (Ar.), 153.2, 153.6 (CO), 163.6 (CN)

ν_{max} (film)/cm^{-1}: 3399, 3304, 3299, 3270, 3104 (NH, NH_2), 1719, 1635, 1602 (CO, CN)

_N,N'_-di(_tert_-butoxycarbonyl)-_N''_- [4-(4-aminophenoxy)phenyl]guanidine (6d)



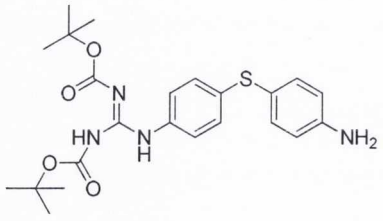
Following *Method* D and after purification with flash chromatography (Hexane:Ethyl Acetate, 3:2) 1093 mg, (82 %) of **6d** as a brown solid were obtained. Mp: decomposes over 260 °C.

δ_H (400 MHz, $CDCl_3$): 1.49 (s, 9H, $(CH_3)_3$), 1.53 (s, 9H, $(CH_3)_3$), 3.60 (broad s, 2H, NH_2), 6.65 (d, 2H, J = 9.0 Hz, Ar.), 6.84 (d, 2H, J = 9.0 Hz, Ar.), 6.88 (d, 2H, J = 9.0 Hz, Ar.), 7.48 (d, 2H, J = 9.0 Hz, Ar.), 10.22 (broad s, 1H, NH), 11.65 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.1 ($(CH_3)_3$), 79.4, 83.5 ($\underline{C}(CH_3)_3$), 116.1, 117.5, 120.8, 123.7, 130.9, 142.6, 148.6 (Ar.), 153.2, 153.5 (CO), 155.7 (Ar.), 163.5 (CN)

ν_{max} (film)/cm^{-1}: 3431, 3345, 3257, 3159 (NH, NH_2), 1722, 1650, 1626 (CO, CN)

_N,N'_-di(_tert_-butoxycarbonyl)-_N''_-[4-(4-aminophenylsulfanyl)phenyl]guanidine (7d)



Following *Method* D and after purification with flash chromatography (Hexane:Ethyl Acetate, 2:1) 541 mg, (39 %) of **7d** as a yellow solid were obtained. Mp: 103-105 °C.

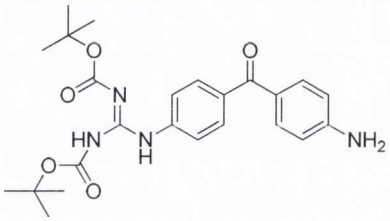
δ_H (400 MHz, $CDCl_3$): 1.48 (s, 9H, $(CH_3)_3$), 1.52 (s, 9H, $(CH_3)_3$), 3.84 (broad s, 2H, NH_2), 6.60 (d, 2H, J = 8.5 Hz, Ar.), 7.09 (d, 2H, J = 8.5 Hz, Ar.), 7.24 (d,

2H, *J* = 8.5 Hz, Ar.), 7.44 (d, 2H, *J* = 8.5 Hz, Ar.), 10.26 (broad s, 1H, NH), 11.64 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.8, 28.0 ($(CH_3)_3$), 79.4, 83.5 ($\underline{C}(CH_3)_3$), 115.6, 120.4, 122.6, 128.2, 134.2, 135.2, 135.4, 146.9 (Ar.), 153.0, 153.3 (CO), 163.3 (CN)

ν_{max} (film)/cm^{-1}: 3472, 3376, 3251, 3140 (NH, NH_2), 1717, 1621, 1593 (CO, CN)

_N_,_N'_-di(*tert*-butoxycarbonyl)-_N''_-[4-(4-aminobenzoyl)phenyl]guanidine (8d)



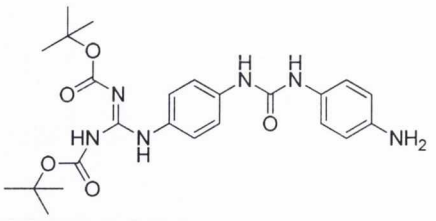
Following *Method* D and after purification with flash chromatography (Hexane:Ethyl Acetate, 3:2) 591 mg, (43 %) of **8d** as a pale yellow solid were obtained. Mp: decomposes over 205 °C.

δ_H (400 MHz, $CDCl_3$): 1.53 (s, 9H, $(CH_3)_3$), 1.55 (s, 9H, $(CH_3)_3$), 4.17 (broad s, 2H, NH_2), 6.67 (d, 2H, *J* = 8.5 Hz, Ar.), 7.70 (d, 2H, *J* = 8.5 Hz, Ar.), 7.72-7.76 (m, 4H, Ar.), 10.57 (broad s, 1H, NH), 11.65 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.1 ($(CH_3)_3$), 80.0, 84.0 ($\underline{C}(CH_3)_3$), 113.6, 121.0, 127.5, 130.8, 132.8, 134.6, 139.9, 150.8 (Ar.), 153.2, 153.3 (CO), 163.3 (CN), 194.3 (PhCOPh)

ν_{max} (film)/cm^{-1}: 3460, 3351, 3240, 3150 (NH, NH_2), 1790, 1721, 1631, 1582 (CO, CN)

1-{4-[*N'*,*N''*-di(*tert*-butoxycarbonyl)guanidino]phenyl}-3-(4-aminophenyl)urea (9d)



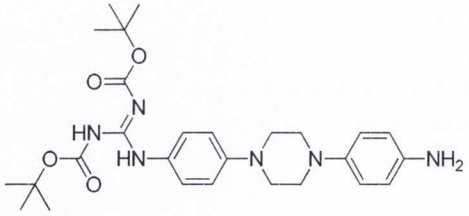
Following *Method* D and after purification with flash chromatography (Ethyl Acetate) 775 mg, (53 %) of **9d** as a yellow solid were obtained. Mp: decomposes over 260 °C.

δ_H (400 MHz, $CDCl_3$): 1.44 (s, 9H, $(CH_3)_3$), 1.55 (s, 9H, $(CH_3)_3$), 3.56 (broad s, 2H, NH_2), 6.56 (d, 2H, J = 8.5 Hz, Ar.), 7.00 (d, 2H, J = 8.0 Hz, Ar.), 7.09 (d, 2H, J = 8.5 Hz, Ar.), 7.15-7.22 (m, 3H, Ar. + NHCOPh), 7.45 (broad s, 1H, NHCONH), 10.08 (broad s, 1H, NH), 11.63 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.1 ($(CH_3)_3$), 80.1, 83.9 ($\underline{C}(CH_3)_3$), 115.6, 119.5, 123.1, 124.8, 129.4, 129.9, 137.4, 143.0 (Ar.), 153.1, 153.9, 155.2 (CO), 163.2 (CN)

ν_{max} (film)/cm^{-1}: 3373, 3305, 3270, 3108 (NH, NH_2), 1717, 1646, 1620, 1601 (CO, CN)

***N*,*N'*-di(*tert*-butoxycarbonyl)-*N''*-{4-[4-(4-aminophenyl)piperazin-1-yl]phenyl}guanidine (10d)**



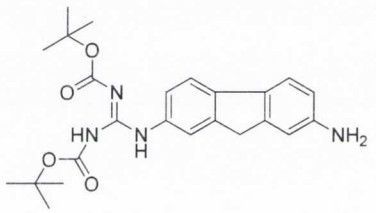
Following *Method* D and after purification with flash chromatography (Hexane:Ethyl Acetate, 1:1) 997 mg, (65 %) of **10d** as a brown solid were obtained. Mp: decomposes over 220 °C.

δ_H (400 MHz, $CDCl_3$): 1.50 (s, 9H, $(CH_3)_3$), 1.54 (s, 9H, $(CH_3)_3$), 3.12-3.23 (m, 4H, CH_2 Pip), 3.25-3.36 (m, 4H, CH_2 Pip), 3.48 (broad s, 2H, NH_2), 6.68 (d, 2H, J = 8.5 Hz, Ar.), 6.87 (d, 2H, J = 8.5 Hz, Ar.), 6.94 (d, 2H, J = 8.5 Hz, Ar.), 7.48 (d, 2H, J = 8.5 Hz, Ar.), 10.18 (broad s, 1H, NH), 11.66 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.2 ($(CH_3)_3$), 49.8, 51.1 ($4CH_2$), 79.3, 83.4 ($\underline{C}(CH_3)_3$), 116.1, 116.7, 118.8, 123.4, 129.2, 140.4, 144.3, 148.6 (Ar.), 153.3, 153.5 (CO), 163.6 (CN);

ν_{max} (film)/cm^{-1}: 3445, 3371, 3288, 3266 (NH, NH_2), 1718, 1620, 1601 (CO, CN)

_N_,_N'_-di(*tert*-butoxycarbonyl)-_N''_-(6-amino-9H-fluoren-3-yl)guanidine (11d)



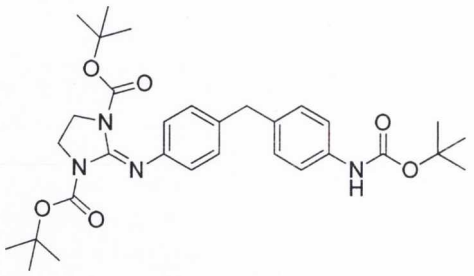
Following *Method* D and after purification with flash chromatography (Hexane:EtOAc, 2:1) 554 mg, (44 %) of **11d** as a orange solid were obtained. Mp: decomposes over 220 °C.

δ_H (400 MHz, $CDCl_3$): 1.53 (s, 9H, $(CH_3)_3$), 1.55 (s, 9H, $(CH_3)_3$), 3.62, (broad s, 2H, NH_2), 3.78 (s, 2H, CH_2), 6.68 (d, 1H, J = 7.8 Hz, Ar.), 6.82 (s, 1H, Ar.), 7.44 (d, 1H, J = 7.8 Hz, Ar.), 7.49 (d, 1H, J = 7.8 Hz, Ar.), 7.55 (d, 1H, J = 7.8 Hz, Ar.), 7.82 (s, 1H, Ar.), 10.40 (broad s, 1H, NH), 11.71 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 28.0, 28.2 ($(CH_3)_3$), 36.8 (CH_2), 79.5, 83.5 ($\underline{C}(CH_3)_3$), 111.7, 113.9, 118.5, 119.0, 120.3, 120.9, 132.5, 134.0, 139.0, 143.0, 145.2, 145.5 (Ar.), 153.3, 153.4 (CO), 163.6 (CN)

ν_{max} (film)/cm^{-1}: 3458, 3374, 3261, 3161 (NH, NH_2), 1716, 1633, 1611 (CO, CN)

2-[4-(4-*tert*-Butoxycarbonylaminobenzyl)phenylimino]imidazolidine-1,3-dicarboxylic acid di-*tert*-butyl ester (3e)



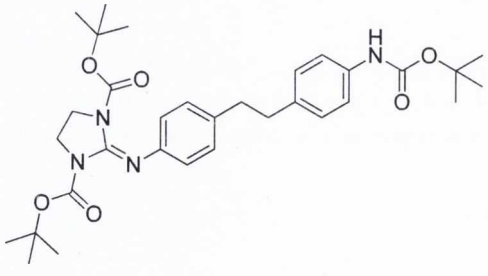
Following *Method* C and after purification with flash chromatography (Hexane:Ethyl Acetate, 5:2) 1378 mg, (81 %) of **3e** as a white solid were obtained. Mp: 118-120 °C.

δ_H (400 MHz, $CDCl_3$): 1.26 (s, 18H, $(CH_3)_3$), 1.47 (s, 9H, $(CH_3)_3$), 3.77 (s, 4H, $2CH_2$ Im), 3.81 (s, 2H, $PhCH_2Ph$), 6.62 (broad s, 1H, NH), 6.88 (d, 2H, J = 8.0 Hz, Ar.), 6.99 (d, 2H, J = 8.0 Hz, Ar.), 7.03 (d, 2H, J = 8.0 Hz, Ar.), 7.21 (d, 2H, J = 8.0 Hz, Ar.)

δ_C (100 MHz, $CDCl_3$): 27.7 ($2(CH_3)_3$), 28.2 ($(CH_3)_3$), 40.5 ($PhCH_2Ph$), 42.9 ($2CH_2$ Im), 80.0 ($\underline{C}(CH_3)_3$), 82.5 ($2\underline{C}(CH_3)_3$), 118.4, 121.3, 128.9, 129.1, 135.3, 136.0, 136.1, 138.8 (Ar.), 146.1, 150.1 (CO), 152.7 (CN)

ν_{max} (film)/cm^{-1}: 3333 (NH), 1754, 1695 (CO, CN)

2-{4-[2-(4-*tert*-Butoxycarbonylaminophenyl)ethyl]phenylimino}imidazolidine-1,3-dicarboxylic acid di-*tert*-butyl ester (4e)



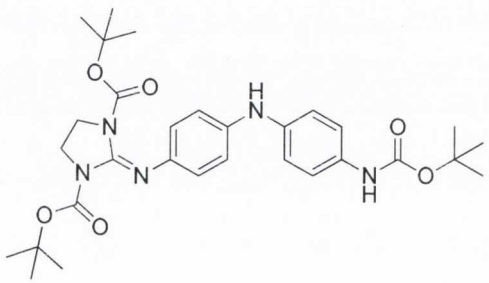
Following *Method* C and after purification with flash chromatography (Hexane:Ethyl Acetate, 5:2) 1262 mg, (72 %) of **4e** as a white solid were obtained. Mp: 181-183 °C.

δ_H (400 MHz, $CDCl_3$): 1.31 (s, 18H, $(CH_3)_3$), 1.51 (s, 9H, $(CH_3)_3$), 2.79 (s, 4H, $2CH_2$), 3.81 (s, 4H, $2CH_2$ Im), 6.49 (broad s, 1H, NHCO), 6.91 (d, 2H, J = 8.2 Hz, Ar.), 7.03 (d, 2H, J = 8.2 Hz, Ar.), 7.09 (d, 2H, J = 8.2 Hz, Ar.), 7.24 (d, 2H, J = 8.2 Hz, Ar.)

δ_C (100 MHz, $CDCl_3$): 27.8 ($2(CH_3)_3$), 28.3 ($(CH_3)_3$ NH), 37.4, 37.5 (CH_2), 43.0 ($2CH_2$ Im), 80.2 ($\underline{C}(CH_3)_3$), 80.6 ($2\underline{C}(CH_3)_3$), 118.5, 121.4, 128.5, 128.7, 135.9, 136.0, 136.6, 138.8 (Ar.), 146.1, 150.3 (CO), 152.8 (CN)

ν_{max} (film)/cm^{-1}: 3283 (NH), 1751, 1714, 1669 (CO, CN)

2-[4-(4-*tert*-Butoxycarbonylaminophenylamino)phenylimino]imidazolidine-1,3-dicarboxylic acid di-*tert*-butyl ester (5e)



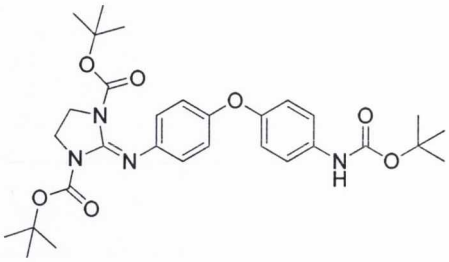
Following *Method* C and after purification with flash chromatography (Hexane:Ethyl Acetate, 3:1) 1364 mg, (80 %) of **5e** as a pink solid were obtained. Mp: 209-211 °C.

δ_H (400 MHz, $CDCl_3$): 1.37 (s, 18H, $2(CH_3)_3$), 1.52 (s, 9H, $(CH_3)_3$), 3.83 (s, 4H, $2CH_2$ Im), 5.47 (broad s, 1H, PhNHPh), 6.33 (broad s, 1H, NHCO), 6.89-6.96 (m, 6H, Ar.), 7.20 (d, 2H, J = 7.5 Hz, Ar.)

δ_C (100 MHz, $CDCl_3$): 27.9 ($2(CH_3)_3$), 28.3 ($(CH_3)_3$), 43.1 ($2CH_2$ Im), 80.1 ($\underline{C}(CH_3)_3$), 82.6 ($2\underline{C}(CH_3)_3$), 117.5, 119.0, 120.4, 122.4, 131.0, 138.3, 138.6, 140.0 (Ar.), 142.2, 150.3 (CO), 153.1 (CN)

ν_{max} (film)/cm^{-1}: 3389, 3329 (NH), 1749, 1709, 1653 (CO, CN)

2-[4-(4-*tert*-Butoxycarbonylaminophenoxy)phenylimino]imidazolidine-1,3-dicarboxylic acid di-*tert*-butyl ester (6e)



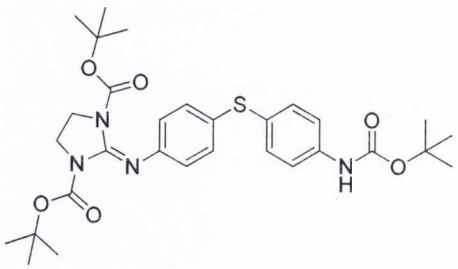
Following *Method* C and after purification with flash chromatography (Hexane:Ethyl Acetate, 5:2) 1117 mg, (65 %) of **6e** as a white solid were obtained. Mp: 149-151 °C.

δ_H (400 MHz, $CDCl_3$): 1.35 (s, 18H, $2(CH_3)_3$), 1.49 (s, 9H, $(CH_3)_3$), 3.81 (s, 4H, $2CH_2$ Im), 6.66 (broad s, 1H, NH), 6.81-6.91 (m, 4H, Ar.), 6.94 (d, 2H, J = 8.1 Hz, Ar.), 7.28 (d, 2H, J = 7.6 Hz, Ar.);

δ_C (100 MHz, $CDCl_3$): 27.8 ($2(CH_3)_3$), 28.2 ($(CH_3)_3$), 43.0 ($2CH_2$ Im), 80.1 ($\underline{C}(CH_3)_3$), 82.6 ($2\underline{C}(CH_3)_3$), 118.5, 119.0, 120.1, 122.4, 133.2, 139.1, 143.8, 150.0 (Ar.), 152.4, 152.9 (CO), 153.4 (CN)

ν_{max} (film)/cm^{-1}: 3286 (NH), 1750, 1715, 1673 (CO, CN)

2-[4-(4-*tert*-Butoxycarbonylaminophenylsulfanyl)phenylimino]imidazolidine-1,3-dicarboxylic acid di-*tert*-butyl ester: (7e)



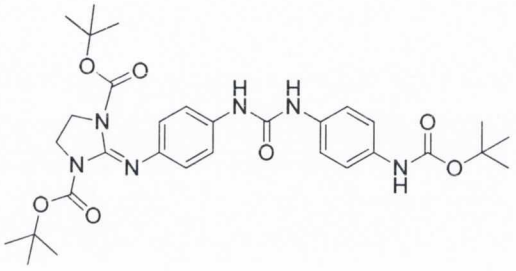
Following *Method* C and after purification with flash chromatography (Hexane:Ethyl Acetate, 5:2) 1260 mg, (72 %) of **7e** as a white solid were obtained. Mp: 174-176 °C.

δ_H (400 MHz, $CDCl_3$): 1.33 (s, 18H, $2(CH_3)_3$), 1.49 (s, 9H, $(CH_3)_3$), 3.81 (s, 4H, $2CH_2$ Im), 6.67 (broad s, 1H, NHCO), 6.91 (d, 2H, J = 8.2 Hz, Ar.), 7.19-7.25 (m, 4H, Ar.), 7.26 (d, 2H, J = 8.2 Hz, Ar.)

δ_C (100 MHz, $CDCl_3$): 27.8 ($2(CH_3)_3$), 28.2 ($(CH_3)_3$), 43.0 ($2CH_2$ Im), 80.4 ($\underline{C}(CH_3)_3$), 82.8 ($2\underline{C}(CH_3)_3$), 118.9, 122.1, 128.3, 130.4, 131.3, 132.1, 137.2, 139.5 (Ar.), 147.7, 150.0 (CO), 152.5 (CN)

ν_{max} (film)/cm^{-1}: 3301 (NH), 1742, 1713, 1659 (CO, CN)

2-{4-[3-(4-*tert*-Butoxycarbonylaminophenyl)ureido]phenylimino}imidazolidine-1,3-dicarboxylic acid di-*tert*-butyl ester (9e)



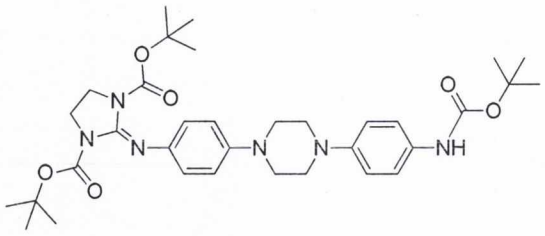
Following *Method* C and after purification with flash chromatography (Hexane:Ethyl Acetate, 1:4) 1231 mg, (67 %) of **9e** as a white solid were obtained. Mp: 150-152 °C.

δ_H (400 MHz, $CDCl_3$): 1.26 (s, 18H, $2(CH_3)_3$), 1.45 (s, 9H, $(CH_3)_3$), 3.79 (s, 4H, $2CH_2$ Im), 6.77 (d, 2H, J = 7.5 Hz, Ar.), 6.84 (broad s, 1H, NH), 7.03-7.16 (m, 6H, Ar.), 7.73-7.78 (m, 2H, NHCONH)

δ_C (100 MHz, $CDCl_3$): 27.7 ($2(CH_3)_3$), 28.2 ($(CH_3)_3$), 43.1 ($2CH_2$ Im), 79.8 ($\underline{C}(CH_3)_3$), 82.9 ($2\underline{C}(CH_3)_3$), 119.7, 120.0, 120.4, 121.5, 133.1, 134.2, 134.3, 139.4 (Ar.), 142.8, 150.2, 153.2 (CO), 153.7 (CN)

ν_{max} (film)/cm^{-1}: 3341 (NH), 1756, 1704, 1683, 1600 (CO, CN)

2-{4-[4-(4-*tert*-Butoxycarbonylaminophenyl)piperazin-1-yl]-phenylimino}imidazolidine-1,3-dicarboxylic acid di-*tert*-butyl ester (10e)



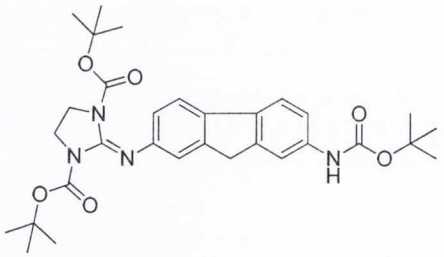
Following *Method* C and after purification with flash chromatography (Hexane:Ethyl Acetate, 1:1) 1473 mg, (77 %) of **10e** as a white solid were obtained. Mp: 111-113 °C.

δ_H (400 MHz, $CDCl_3$): 1.33 (s, 18H, $2(CH_3)_3$), 1.51 (s, 9H, $(CH_3)_3$), 3.19-3.29 (m, 8H, $4CH_2$ Pip), 3.82 (s, 4H, $2CH_2$ Im), 6.37 (broad s, 1H, NHCO), 6.86-7.00 (m, 6H, Ar.), 7.23-7.32 (m, 2H, Ar.)

δ_C (100 MHz, $CDCl_3$): 27.8 ($2(CH_3)_3$), 28.3 ($(CH_3)_3$), 43.0 ($2CH_2$ Im), 50.0, 50.5 ($4CH_2$ Pip), 80.0 ($\underline{C}(CH_3)_3$), 82.5 ($2\underline{C}(CH_3)_3$), 117.2, 117.4, 121.1, 122.2, 131.2, 138.6, 141.6, 147.2 (Ar.), 147.4, 150.3 (CO), 153.1 (CN)

ν_{max} (film)/cm^{-1}: 3360 (NH), 1754, 1699, 1679 (CO, CN)

2-(7-*tert*-Butoxycarbonylamino-9*H*-fluoren-2-ylimino)imidazolidine-1,3-dicarboxylic acid di-*tert*-butyl ester (11e)



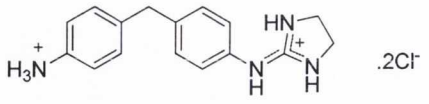
Following *Method* C and after purification with flash chromatography (Hexane:Ethyl Acetate, 3:1) 1018 mg, (60 %) of **11e** as a white solid were obtained. Mp: 110-112 °C.

δ_H (400 MHz, $CDCl_3$): 1.27 (s, 18H, $2(CH_3)_3$), 1.51 (s, 9H, $(CH_3)_3$), 3.75 (s, 2H, $PhCH_2Ph$), 3.83 (s, 4H, $2CH_2$ Im), 6.73 (broad s, 1H, NH), 6.98 (d, 1H, J = 8.0 Hz, Ar.), 7.12 (s, 1H, Ar.), 7.16 (d, 1H, J = 8.0 Hz, Ar.), 7.53 (d, 1H, J = 8.0 Hz, Ar.), 7.54 (d, 1H, J = 8.0 Hz, Ar.), 7.67 (s, 1H, Ar.)

δ_C (100 MHz, $CDCl_3$): 27.7 ($2(CH_3)_3$), 28.2 ($(CH_3)_3$), 36.8 ($PhCH_2Ph$), 43.0 ($2CH_2$ Im), 80.1 ($\underline{C}(CH_3)_3$), 82.7 ($2\underline{C}(CH_3)_3$), 115.5, 117.4, 117.9, 119.2, 119.3, 120.4, 136.1, 136.2, 137.1, 138.7, 143.5, 144.0 (Ar.), 146.6, 150.2 (CO), 152.9 (CN)

ν_{max} (film)/cm^{-1}: 3330 (NH), 1754, 1698 (CO, CN)

Dihydrochloride salt of 2-[4-(4-aminobenzyl)phenylimino]imidazolidine (3f)



Following *Method F*, 124 mg of the pure dihydrochloride salt **3f** were obtained as a white solid (93%). Mp: decomposes over 165 °C; clogP: 1.688

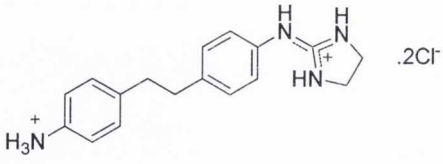
δ_H (400 MHz, D_2O): 3.64 (s, 4H, $2CH_2$ Im), 3.93 (s, 2H, $PhCH_2Ph$), 7.12 (d, 2H, J = 8.0 Hz, Ar.), 7.22-7.30 (m, 4H, Ar.), 7.31 (d, 2H, J = 8.0 Hz, Ar.)

δ_C (100 MHz, D_2O): 39.5 ($PhCH_2Ph$), 42.2 ($2CH_2$ Im), 122.6, 123.9, 127.3, 129.6, 129.8, 132.6, 139.7, 142.0 (Ar.), 158.1 (CN)

HRMS (*m/z* -ES): Found: 267.1712 ($M^+ + H$). $C_{16}H_{19}N_4$ Requires: 267.1708

Anal. ($C_{16}H_{20}Cl_2N_4 \cdot 0.8H_2O$) Calcd: C, 54.34; H, 6.16; N, 15.84. Found: C, 54.65; H, 6.05; N, 15.82.

Dihydrochloride salt of 2-{4-[2-(4-aminophenyl)ethyl]phenylimino}imidazolidine (4f)



Following *Method F*, 132 mg of the pure dihydrochloride salt **4f** were obtained as a white solid (94%). Mp: decomposes over 210 °C; clogP: 2.067

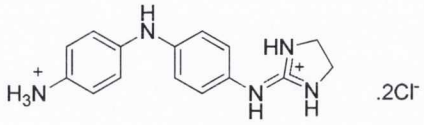
δ_H (400 MHz, D_2O): 2.89 (s, 4H, $2CH_2$), 3.66 (s, 4H, $2CH_2$ Im), 7.09 (d, 2H, J = 8.0 Hz, Ar.), 7.20 (d, 2H, J = 8.0 Hz, Ar.), 7.26 (d, 2H, J = 8.0 Hz, Ar.), 7.28 (d, 2H, J = 8.0 Hz, Ar.)

δ_C (100 MHz, D_2O): 35.4, 35.5 ($2CH_2$), 42.2 ($2CH_2$ Im), 122.2, 123.6, 127.0, 129.4, 129.7, 132.3, 140.3, 142.3 (Ar.), 158.1 (CN)

HRMS (*m/z* -ES): Found: 281.1939 ($M^+ + H$). $C_{17}H_{21}N_4$ Requires: 281.1933

Anal. ($C_{17}H_{22}Cl_2N_4 \cdot 1.0H_2O$) Calcd: C, 54.99; H, 6.52; N, 15.09. Found: C, 54.80; H, 6.17; N, 14.71.

Dihydrochloride salt of 2-[4-(4-aminophenylamino)phenylimino]imidazolidine (5f)



Following *Method F*, 126 mg of the pure dihydrochloride salt **5f** were obtained as a pink solid (94%). Mp: decomposes over 195 °C; clogP: 1.099

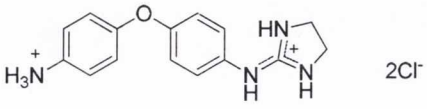
δ_H (400 MHz, D_2O): 3.67 (s, 4H, $2CH_2$, Im), 7.05-7.20 (m, 6H, Ar.), 7.28 (d, 2H, J = 7.5 Hz, Ar.)

δ_C (100 MHz, D_2O): 42.2 ($2CH_2$, Im), 117.5, 118.3, 121.5, 123.4, 125.4, 127.4, 141.3, 143.3 (Ar.), 158.4 (CN)

HRMS (*m/z* -ES): Found: 268.1562 ($M^+ + H$). $C_{15}H_{18}N_5$ Requires: 268.1562

Anal. ($C_{15}H_{19}Cl_2N_5 \cdot 0.5H_2O$) Calcd: C, 51.58; H, 5.77; N, 20.05. Found: C, 51.60; H, 5.40; N, 19.86.

Dihydrochloride salt of 2-[4-(4-aminophenoxy)phenylimino]imidazolidine (6f)



Following *Method F*, 129 mg of the pure dihydrochloride salt **6f** were obtained as a pink solid (96%). Mp: decomposes over 180 °C; clogP: 1.719

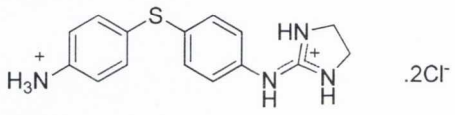
δ_H (400 MHz, D_2O): 3.72 (s, 4H, $2CH_2$, Im), 7.11 (d, 2H, J = 8.5 Hz, Ar.), 7.15 (d, 2H, J = 8.5 Hz, Ar.), 7.29 (d, 2H, J = 8.5 Hz, Ar.), 7.41 (d, 2H, J = 8.5 Hz, Ar.)

δ_C (100 MHz, D_2O): 42.2 ($2CH_2$, Im), 119.4, 119.8, 124.2, 124.6, 126.2, 130.3, 154.7, 156.5 (Ar.), 158.4 (CN)

HRMS (*m/z* -ES): Found: 269.1400 ($M^+ + H$). $C_{15}H_{17}N_4O$ Requires: 269.1402

Anal. ($C_{15}H_{18}Cl_2N_4O{\cdot}0.5H_2O$) Calcd: C, 51.44; H, 5.47; N, 16.00. Found: C, 51.76; H, 5.10; N, 15.83.

Dihydrochloride salt of 2-[4-(4-aminophenylsulfanyl)phenylimino]imidazolidine (7f)



Following *Method F*, 135 mg of the pure dihydrochloride salt **7f** were obtained as a white solid (95%). Mp: decomposes over 200 °C; clogP: 1.959

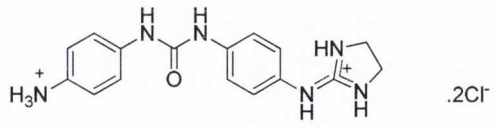
δ_H (400 MHz, D_2O): 3.73 (s, 4H, $2CH_2$, Im), 7.24 (d, 2H, J = 7.5 Hz, Ar.), 7.33 (d, 2H, J = 7.5 Hz, Ar.), 7.39-7.47 (m, 4H, Ar.)

δ_C (100 MHz, D_2O): 42.2 ($2CH_2$, Im), 123.4, 124.4, 128.4, 130.9, 131.9, 132.8, 134.5, 136.5 (Ar.), 157.9 (CN)

HRMS (*m/z* -ES): Found: 285.1173 ($M^+ + H$). $C_{15}H_{18}N_4S$ Requires: 285.1174

Anal. ($C_{15}H_{18}Cl_2N_4S{\cdot}0.6H_2O$) Calcd: C, 48.94; H, 5.26; N, 15.22. Found: C, 49.08; H, 5.05; N, 15.02.

Dihydrochloride salt of 1-(4-aminophenyl)-3-[4-(imidazolidin-2-ylideneamino)phenyl] urea (9f)



Following *Method F*, 145 mg of the pure dihydrochloride salt **9f** were obtained as a white solid (93%). Mp: decomposes over 230 °C; clogP: 0.489

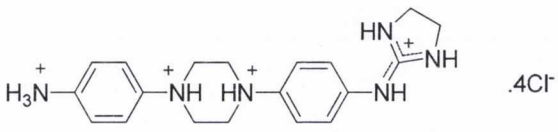
δ_H (400 MHz, D_2O): 3.69 (s, 4H, $2CH_2$, Im), 7.20 (d, 2H, J = 8.3 Hz, Ar.), 7.31-7.39 (m, 4H, Ar.), 7.45 (d, 2H, J = 8.8 Hz, Ar.)

δ_C (100 MHz, D_2O): 42.6 ($2CH_2$, Im), 121.6, 121.8, 123.6, 124.9, 125.1, 130.5, 136.8, 138.6 (Ar.), 155.2 (CN), 158.7 (CO)

HRMS (*m/z* -ES): Found: 311.1611 ($M^+ + H$). $C_{16}H_{19}N_6O$ Requires: 311.1620

Anal. ($C_{16}H_{20}Cl_2N_6O \cdot 1.4H_2O$) Calcd: C, 47.04; H, 5.63; N, 20.57. Found: C, 47.12; H, 5.51; N, 20.24.

Tetrahydrochloride salt of 2-{4-[4-(4-aminophenyl)piperazin-1-yl]phenylimino} imidazolidine (10f)



Following *Method F*, 158 mg of the pure dihydrochloride salt **10f** were obtained as a brown solid (94%). Mp: 196-198 °C; clogP: 1.069

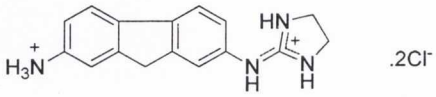
δ_H (400 MHz, D_2O): δ 3.69-3.74 (m, 4H, $2CH_2$ Pip), 3.75 (s, 4H, $2CH_2$ Im), 3.77-3.83 (m, 4H, $2CH_2$ Pip), 7.35 (d, 2H, J = 9.0 Hz, Ar.), 7.42 (d, 2H, J = 9.0 Hz, Ar.), 7.43 (d, 2H, J = 9.0 Hz, Ar.), 7.56 (d, 2H, J = 9.0 Hz, Ar.)

δ_C (100 MHz, D_2O): δ 42.3 ($2CH_2$ Im), 48.2, 52.6 ($2CH_2$ Pip), 118.8, 121.2, 123.9, 124.7, 125.2, 134.8, 140.3, 147.0 (Ar.), 158.0 (CN)

HRMS (*m/z* -ES): Found: 337.2153 ($M^+ + H$. $C_{19}H_{27}N_6$ Requires: 337.2141)

Anal. ($C_{19}H_{28}Cl_4N_6 \cdot 0.5H_2O$) Calcd: C, 46.45; H, 5.95; N, 17.11. Found: C, 46.41; H, 6.10; N, 16.80.

Dihydrochloride salt of N^2-imidazolidin-2-ylidene-9*H*-fluorene-2,7-diamine (11f)



Following *Method F*, 125 mg of the pure dihydrochloride salt **11f** were obtained as a white solid (94%). Mp: decomposes over 225 °C; clogP: 1.554

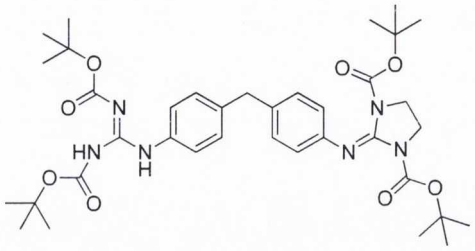
δ_H (400 MHz, D_2O): 3.57 (s, 4H, CH_2 Im), 3.60 (s, 2H, $PhCH_2Ph$), 6.98 (d, 1H, *J* = 8.0 Hz, Ar.), 7.13 (s, 1H, Ar.), 7.25 (d, 1H, *J* = 8.0 Hz, Ar.), 7.43 (s, 1H, Ar.), 7.53 (d, 1H, *J* = 8.0 Hz, Ar.), 7.61 (d, 1H, *J* = 8.0 Hz, Ar.)

δ_C (100 MHz, D_2O): 35.8 ($PhCH_2Ph$), 42.1 (CH_2 Im), 119.0, 119.1, 120.4, 120.6, 120.9, 121.3, 127.7, 133.5, 137.7, 140.4, 144.5, 144.7 (Ar.), 157.3 (CN)

HRMS (*m/z* -ES): Found: 265.1449 ($M^+ + H$). $C_{16}H_{17}N_4$ Requires: 265.1453

Anal. ($C_{16}H_{18}Cl_2N_4 \cdot 0.3H_2O$) Calcd: C, 56.08; H, 5.47; N, 16.35. Found: C, 56.23; H, 5.37; N, 16.10.

4-[2,3-di(*tert*-butoxycarbonyl)guanidine]-4'-[1,3-di(*tert*-butoxycarbonyl)-2-imidazo lidinylimino]diphenyl methane (3g)



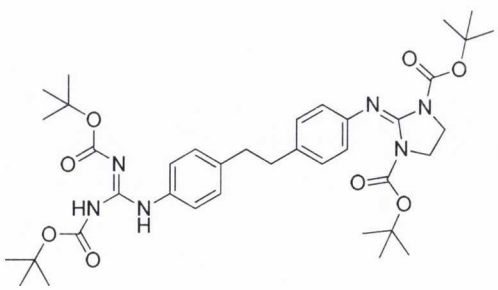
Following *Method E* after purification with flash chromatography (Hexane:Ethyl Acetate, 5:2) 1425 mg, (67 %) of **3g** as a white solid were obtained. Mp: decomposes over 226 °C

δ_H (400 MHz, $CDCl_3$): 1.28 (s, 18H, $2(CH_3)_3$ Im), 1.47 (s, 9H, $(CH_3)_3$ Gu), 1.51 (s, 9H, $(CH_3)_3$ Gu), 3.78 (s, 4H, $2CH_2$ Im), 3.84 (s, 2H, $PhCH_2Ph$), 6.89 (d, 2H, J = 8.2 Hz, Ar.), 7.00 (d, 2H, J = 8.2 Hz, Ar.), 7.08 (d, 2H, J = 8.2 Hz, Ar.), 7.44 (d, 2H, J = 8.2 Hz, Ar.), 10.22 (broad s, 1H, NH); 11.62 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.7 ($(CH_3)_3$ Im), 27.9, 28.0 ($(CH_3)_3$ Gu), 40.7 ($PhCH_2Ph$), 42.9 (CH_2 Im), 79.3 ($\underline{C}(CH_3)_3$ Gu), 82.5 ($\underline{C}(CH_3)_3$ Im), 83.4 ($\underline{C}(CH_3)_3$ Gu), 121.4, 122.0, 129.0, 134.4, 135.0, 138.1, 138.8, 146.2 (Ar.), 150.2, 153.1, 153.3, 163.4 (CO, CN)

ν_{max} (film)/cm^{-1}: 3261, 3156 (NH), 1756, 1720, 1668, 1635, 1609 (CO, CN)

{4-[2,3-di(*tert*-butoxycarbonyl)guanidine]-4'-[1,3-di(*tert*-butoxycarbonyl)-2-imidazo lidinylimino]}-1,2-diphenyl ethane (4g)



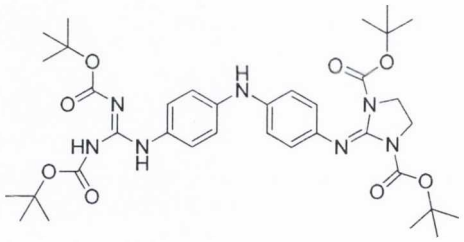
Following *Method* E after purification with flash chromatography (Hexane:Ethyl Acetate, 1:1) 1345 mg, (62 %) of **4g** as a white solid were obtained. Mp: decomposes over 220 °C

δ_H (400 MHz, $CDCl_3$): 1.32 (s, 18H, $2(CH_3)_3$ Im), 1.50 (s, 9H, $(CH_3)_3$ Gu), 1.53 (s, 9H, $(CH_3)_3$ Gu), 2.80 (s, 4H, $2CH_2$), 3.82 (s, 4H, CH_2 Im), 6.92 (d, 2H, *J* = 8.5 Hz, Ar.), 7.05 (d, 2H, *J* = 8.5 Hz, Ar.), 7.14 (d, 2H, *J* = 8.5 Hz, Ar.), 7.49 (d, 2H, *J* = 8.5 Hz, Ar.), 10.26 (broad s, 1H, NH), 11.64 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.8 ($(CH_3)_3$ Im), 28.0, 28.2 ($(CH_3)_3$ Gu), 37.4, 37.7 ($2CH_2$), 43.0 (CH_2 Im), 79.4 ($\underline{C}(CH_3)_3$ Gu), 82.7 ($\underline{C}(CH_3)_3$ Im), 83.5 ($\underline{C}(CH_3)_3$ Gu), 121.4, 122.1, 128.6, 128.8, 134.5, 136.0, 138.5, 138.9 (Ar.), 146.1, 150.3, 153.3, 153.4, 163.6 (CO, CN)

ν_{max} (film)/cm^{-1}: 3312, 3283 (NH), 1714, 1703, 1683, 1633, 1603 (CO, CN)

4-[2,3-di(*tert*-butoxycarbonyl)guanidine]-4'-[1,3-di(*tert*-butoxycarbonyl)-2-imidazo lidinylimino]diphenyl amine (5g)



Following *Method E* after purification with flash chromatography (Hexane:Ethyl Acetate, 1:1) 963 mg, (45 %) of **5g** as a pink solid were obtained. Mp : decomposes over 245 °C

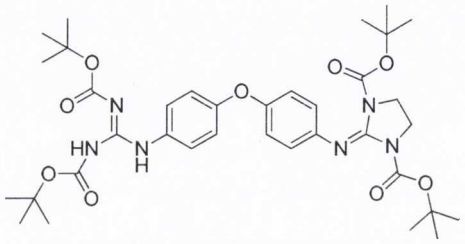
δ_H (400 MHz, $CDCl_3$): 1.34 (s, 18H, $2(CH_3)_3$ Im), 1.47 (s, 9H, $(CH_3)_3$ Gu), 1.51 (s, 9H, $(CH_3)_3$ Gu), 3.80 (s, 4H, CH_2), 5.68 (broad s, 1H, PhNHPh), 6.87 (d, 2H, *J* = 8.0 Hz, Ar.), 6.88-7.02 (m, 4H, Ar.), 7.36 (d, 2H, *J* = 7.5 Hz, Ar.), 10.13 (broad s, 1H, NH), 11.64 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.8 ($(CH_3)_3$ Im), 28.0, 28.1 ($(CH_3)_3$ Gu), 43.0 (CH_2), 79.3 ($\underline{C}(CH_3)_3$ Gu), 82.6 ($\underline{C}(CH_3)_3$ Im), 83.3 ($\underline{C}(CH_3)_3$ Gu), 116.4, 119.4, 122.4,

123.7, 128.7, 137.7, 138.6, 141.5 (Ar.), 142.4, 150.3, 153.2, 153.5, 163.6 (CO, CN)

ν_{max} (film)/cm^{-1}: 3310, 3263, 3114 (NH), 1746, 1723, 1698, 1634, 1606 (CO, CN)

4-[2,3-di(*tert*-butoxycarbonyl)guanidine]-4'-[1,3-di(*tert*-butoxycarbonyl)-2-imidazo lidinylimino]diphenyl ether (6g)



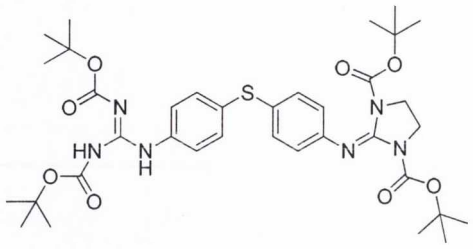
Following *Method* E and after purification with flash chromatography (Hexane:Ethyl Acetate, 3:2) 1090 mg, (51 %) of **6g** as a white solid were obtained. Mp : decomposes over 240 °C

δ_H (400 MHz, $CDCl_3$): 1.36 (s, 18H, 2$(CH_3)_3$ Im), 1.49 (s, 9H, $(CH_3)_3$ Gu), 1.53 (s, 9H, $(CH_3)_3$ Gu), 3.82 (s, 4H, CH_2), 6.82-6.94 (m, 4H, Ar.), 6.97 (d, 2H, *J* = 8.5 Hz, Ar.), 7.48 (d, 2H, *J* = 8.5 Hz, Ar.), 10.23 (broad s, 1H, NH), 11.63 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.9 ($(CH_3)_3$ Im), 28.0, 28.1 ($(CH_3)_3$ Gu), 43.1 (CH_2), 79.5 ($\underline{C}(CH_3)_3$ Gu), 82.7 ($\underline{C}(CH_3)_3$ Im), 83.6 ($\underline{C}(CH_3)_3$ Gu), 118.1, 119.6, 122.5, 123.6, 131.4, 139.2, 144.2, 150.1 (Ar.), 152.1, 153.3, 153.5, 155.1, 163.5 (CO, CN)

ν_{max} (film)/cm^{-1}: 3261, 3164 (NH), 1750, 1722, 1662, 1636, 1621 (CO, CN)

4-[2,3-di(*tert*-butoxycarbonyl)guanidine]-4'-[1,3-di(*tert*-butoxycarbonyl)-2-imidazoli dinylimino]diphenyl thioether (7g)



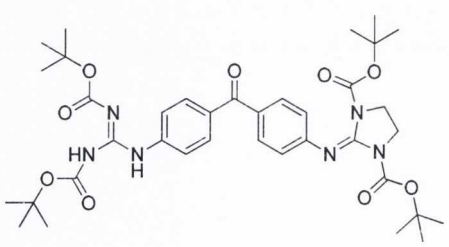
Following *Method* E and after purification with flash chromatography (Hexane:Ethyl Acetate, 5:2) 943 mg, (43 %) of **7g** as a yellow solid were obtained. Mp: decomposes over 232 °C

δ_H (400 MHz, $CDCl_3$): 1.36 (s, 18H, $2(CH_3)_3$ Im), 1.50 (s, 9H, $(CH_3)_3$ Gu), 1.54 (s, 9H, $(CH_3)_3$ Gu), 3.84 (s, 4H, CH_2), 6.96 (d, 2H, J = 8.8 Hz, Ar.), 7.19 (d, 2H, J = 8.8 Hz, Ar.), 7.29 (d, 2H, J = 8.2 Hz, Ar.), 7.49 (d, 2H, J = 8.2 Hz, Ar.), 10.30 (broad s, 1H, NH), 11.62 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.8 ($(CH_3)_3$ Im), 27.9, 28.0 ($(CH_3)_3$ Gu), 43.1 (CH_2), 79.5 ($\underline{C}(CH_3)_3$ Gu), 82.8 ($\underline{C}(CH_3)_3$ Im), 83.6 ($\underline{C}(CH_3)_3$ Gu), 122.2, 122.5, 127.1, 129.8, 133.2, 133.5, 135.0, 139.5 (Ar.), 148.2, 150.0, 153.1, 153.2, 163.3 (CO, CN)

ν_{max} (film)/cm^{-1}: 3276, 3139 (NH), 1747, 1717, 1702, 1674, 1639 (CO, CN)

4-[2,3-di(*tert*-butoxycarbonyl)guanidine]-4'-[1,3-di(*tert*-butoxycarbonyl)-2-imidazoli dinylimino]diphenyl methanone (8g)



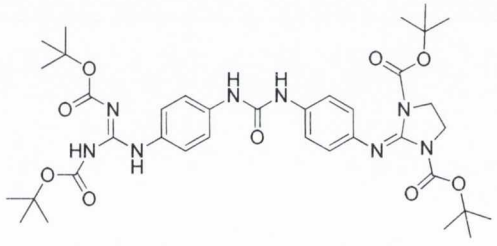
Following *Method* E after purification with flash chromatography (Hexane:Ethyl Acetate, 5:2) 680 mg, (31 %) of **8g** as a white solid were obtained. Mp: decomposes over 180 °C

δ_H (400 MHz, $CDCl_3$): 1.35 (s, 18H, $2(CH_3)_3$ Im), 1.51 (s, 9H, $(CH_3)_3$ Gu), 1.53 (s, 9H, $(CH_3)_3$ Gu), 3.85 (s, 4H, CH_2), 7.02 (d, 2H, J = 8.5 Hz, Ar.), 7.68-7.76 (m, 6H, Ar.), 10.58 (broad s, 1H, NH), 11.64 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.8 ($(CH_3)_3$ Im), 28.0, 28.1 ($(CH_3)_3$ Gu), 43.2 (CH_2), 79.9 ($\underline{C}(CH_3)_3$ Gu), 83.1 ($\underline{C}(CH_3)_3$ Im), 84.0 ($\underline{C}(CH_3)_3$ Gu), 120.8, 120.9, 130.9, 131.4, 134.1, 140.1, 140.2, 149.9 (Ar.), 152.7, 152.8, 153.1, 153.2, 163.2 (CO, CN), 194.7 (PhCOPh)

ν_{max} (film)/cm^{-1}: 3279, 3174 (NH), 1752, 1726, 1638, 1594 (CO, CN)

4-[2,3-di(*tert*-butoxycarbonyl)guanidine]-4'-[1,3-di(*tert*-butoxycarbonyl)-2-imidazoli dinylimino]diphenyl urea (9g)



Following *Method* E and after purification with flash chromatography Hexane:Ethyl Acetate, 1:3) 1478 mg, (65 %) of **9g** as a white solid were obtained. Mp : decomposes over 220 °C

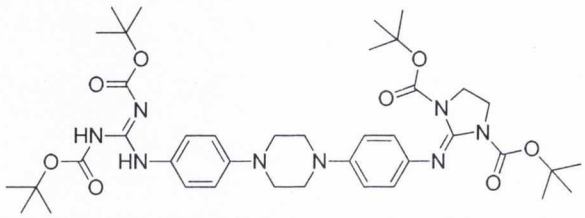
δ_H (400 MHz, $CDCl_3$): 1.28 (s, 18H, $2(CH_3)_3$ Im), 1.37 (s, 9H, $(CH_3)_3$ Gu), 1.50 (s, 9H, $(CH_3)_3$ Gu), 3.78 (s, 4H, $2CH_2$), 6.85 (d, 2H, J = 9.0 Hz, Ar.), 7.08 (d, 2H, J = 9.0 Hz, Ar.), 7.12 (d, 2H, J = 9.0 Hz, Ar.), 7.20 (d, 2H, J = 9.0 Hz, Ar.), 7.56 (broad s, 1H, NHCONH), 7.63 (broad s, 1H, NHCONH), 10.02 (broad s, 1H, NH), 11.57 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.7 ($(CH_3)_3$ Im), 27.8, 27.9 ($(CH_3)_3$ Gu), 43.0 (CH_2), 79.7 ($\underline{C}(CH_3)_3$ Gu), 82.7 ($\underline{C}(CH_3)_3$ Im), 83.7 ($\underline{C}(CH_3)_3$ Gu), 118.8, 119.1, 121.6,

124.6, 129.3, 134.4, 137.7, 139.0 (Ar.), 142.5, 150.2, 152.6, 152.9, 155.2, 162.9 (CO, CN)

ν_{max} (film)/cm^{-1}: 3372, 3304, 3277 (NH), 1756, 1712, 1636, 1620, 1604 (CO, CN)

{4-[2,3-di(*tert*-butoxycarbonyl)guanidine]-4'-[1,3-di(*tert*-butoxycarbonyl)-2-imidazolidinylimino]}-1,4-diphenyl piperazine (10g)



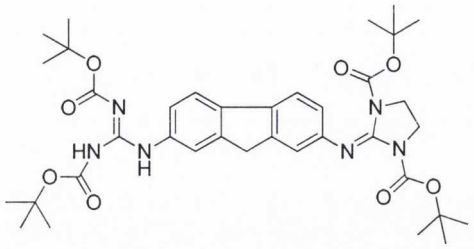
Following *Method* E and after purification with flash chromatography (Hexane:Ethyl Acetate, 1:1) 1917 mg, (82 %) of **10g** as a white solid were obtained. Mp: 88-90 °C

δ_H (400 MHz, $CDCl_3$): 1.32 (s, 18H, 2$(CH_3)_3$ Im), 1.48 (s, 9H, $(CH_3)_3$ Gu), 1.52 (s, 9H, $(CH_3)_3$ Gu), 3.15-3.24 (m, 4H, 2CH_2 Pip), 3.24-3.35 (m, 4H, 2CH_2 Pip), 3.80 (s, 4H, CH_2 Im), 6.82-7.05 (m, 6H, J = 8.0 Hz, Ar.), 7.47 (d, 2H, J = 8.0 Hz, Ar.), 10.17 (broad s, 1H, NH), 11.64 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.8 ($(CH_3)_3$ Im), 28.0, 28.1 ($(CH_3)_3$ Gu), 43.0 (CH_2 Im), 49.7, 50.4 (CH_2 Pip), 79.3 ($\underline{C}(CH_3)_3$ Gu), 82.5 ($\underline{C}(CH_3)_3$ Im), 83.4 ($\underline{C}(CH_3)_3$ Gu), 116.7, 117.4, 122.2, 123.4, 129.2, 138.6, 141.6, 147.2 (Ar.), 148.6, 150.3, 153.2, 153.5, 163.6 (CO, CN)

ν_{max} (film)/cm^{-1}: 3291, 3174 (NH), 1755, 1703, 1635, 1620 (CO, CN)

2-[2,3-di(*tert*-butoxycarbonyl)guanidine]-7-[1,3-di(*tert*-butoxycarbonyl)-2-imidazolidinyl imino]-9*H*-fluorene (11g)



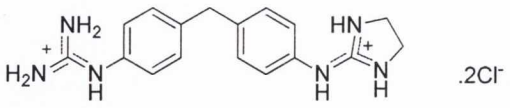
Following *Method* E and after purification with flash chromatography (Hexane:Ethyl Acetate, 3:1) 1260 mg, (59 %) of **11g** as a yellow solid were obtained. Mp: decomposes over 192 °C

δ_H (400 MHz, $CDCl_3$): 1.26 (s, 18H, 2$(CH_3)_3$ Im), 1.49 (s, 9H, $(CH_3)_3$ Gu), 1.51 (s, 9H, $(CH_3)_3$ Gu), 3.80 (s, 2H, $PhCH_2Ph$), 3.82 (s, 4H, 2CH_2 Im), 6.98 (d, 1H, J = 8.0 Hz, Ar.), 7.13 (s, 1H, Ar.), 7.38 (d, 1H, J = 8.0 Hz, Ar.), 7.55 (d, 1H, J = 8.0 Hz, Ar.), 7.58 (d, 1H, J = 8.0 Hz, Ar.), 7.85 (s, 1H, Ar.), 10.37 (broad s, 1H, NH), 11.69 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.7 ($(CH_3)_3$ Im), 27.9, 28.0 ($(CH_3)_3$ Gu), 36.8 ($PhCH_2Ph$), 43.0 (CH_2 Im), 79.4 ($\underline{C}(CH_3)_3$ Gu), 82.6 ($\underline{C}(CH_3)_3$ Im), 83.4 ($\underline{C}(CH_3)_3$ Gu), 117.9, 119.0, 119.1, 119.6, 120.3, 120.8, 134.3, 135.8, 138.7, 143.7, 143.8, 147.0 (Ar.), 150.2, 153.2, 153.4, 163.4 (CO, CN)

ν_{max} (film)/cm^{-1}: 3241, 3160 (NH), 1761, 1723, 1713, 1639, 1614 (CO, CN)

Dihydrochloride salt of *N*-{4-[4-(imidazolidin-2-ylideneamino)benzyl]phenyl}guanidine (3h)



Following *Method* F, 148 mg of the pure dihydrochloride salt **3h** were obtained as a white solid (96%). Mp: 78-80 °C; clogP: 0.801

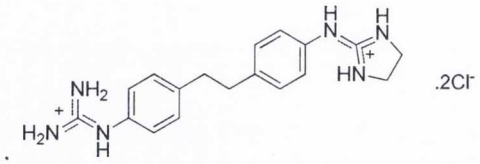
δ_H (400 MHz, D_2O): 3.58 (s, 4H, $2CH_2$ Im), 3.76 (s, 2H, $PhCH_2Ph$), 6.98 (d, 2H, J = 8.0 Hz, Ar.), 7.01 (d, 2H, J = 8.0 Hz, Ar.), 7.15 (d, 2H, J = 8.0 Hz, Ar.), 7.16 (d, 2H, J = 8.0 Hz, Ar.)

δ_C (100 MHz, D_2O): 39.6 ($PhCH_2Ph$), 42.2 ($2CH_2$ Im), 123.3, 125.1, 129.6, 129.7, 131.5, 132.6, 139.7, 140.4 (Ar.), 155.5, 157.6 (CN)

HRMS (*m/z* -ES): Found: 309.1732 ($M^+ + H$). $C_{17}H_{21}N_6$ Requires: 309.1728

Anal. ($C_{17}H_{22}Cl_2N_6 \cdot 1.5H_2O$) Calcd: C, 50.00; H, 6.17; N, 20.58. Found: C, 50.00; H, 6.13; N, 20.34.

Dihydrochloride salt of *N*-(4-{2-[4-(imidazolidin-2-ylideneamino)phenyl]ethyl}phenyl) guanidine (4h)



Following *Method* F, 152 mg of the pure dihydrochloride salt **4h** were obtained as a white solid (94%). Mp: decomposes over 215 °C; clogP: 1.18

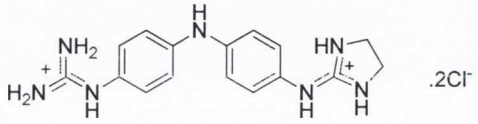
δ_H (400 MHz, D_2O): 2.84 (s, 4H, $2CH_2$), 3.67 (s, 4H, $2CH_2$ Im), 7.03-7.13 (m, 4H, Ar.), 7.15-7.25 (m, 4H, Ar.)

δ_C (100 MHz, D_2O): 35.4, 35.5 ($2CH_2$), 42.2 ($2CH_2$ Im), 123.3, 125.1, 129.4, 129.5, 131.3, 132.3, 140.3, 140.9 (Ar.), 155.6, 157.8 (CN)

HRMS (*m/z* -ES): Found: 323.1988 ($M^+ + H$). $C_{18}H_{23}N_6$ Requires: 323.1984

Anal. ($C_{18}H_{24}Cl_2N_6 \cdot 0.8H_2O$) Calcd: C, 52.76; H, 6.30; N, 20.51. Found: C, 53.14; H, 6.03; N, 20.15.

Dihydrochloride salt of *N*-{4-[4-(imidazolidin-2-ylideneamino)phenylamino]phenyl} guanidine (5h)



Following *Method* F, 146 mg of the pure dihydrochloride salt **5h** were obtained as a green solid (94%). Mp: 66-68 °C; clogP: 0.212

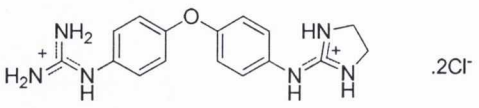
δ_H (400 MHz, D_2O): 3.60 (s, 4H, $2CH_2$, Im), 6.95-7.09 (m, 8H, Ar.)

δ_C (100 MHz, D_2O): 42.2 ($2CH_2$, Im), 117.8, 118.0, 125.2, 126.0, 126.9, 127.3, 141.3, 142.0 (Ar.), 155.9, 158.1 (CN)

HRMS (*m/z* -ES): Found: 310.1724 ($M^+ + H$). $C_{16}H_{20}N_7$ Requires: 310.1722

Anal. ($C_{16}H_{21}Cl_2N_7 \cdot 1.7H_2O$) Calcd: C, 46.54; H, 5.96; N, 23.74. Found: C, 46.87; H, 5.66; N, 23.87.

Dihydrochloride salt of *N*-{4-[4-(imidazolidin-2-ylideneamino)phenoxy]phenyl} guanidine (6h)



Following *Method* F, 146 mg of the pure dihydrochloride salt **6h** were obtained as a yellow solid (94%). Mp: 48-50 °C; clogP: 0.832

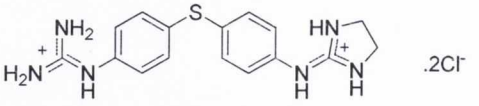
δ_H (400 MHz, D_2O): 3.70 (s, 4H, $2CH_2$, Im), 6.97-7.06 (m, 4H, Ar.), 7.17-7.26 (m, 4H, Ar.)

δ_C (100 MHz, D_2O): 42.3 ($2CH_2$, Im), 119.5, 119.6, 125.8, 127.4, 129.0, 130.1, 154.8, 155.4 (Ar.), 155.8, 158.1 (CN)

HRMS (*m/z* -ES): Found: 311.1595 ($M^+ + H$). $C_{16}H_{19}N_6$ Requires: 311.1594

Anal. ($C_{16}H_{20}Cl_2N_6O \cdot 2.8H_2O$) Calcd: C, 44.31; H, 5.95; N, 19.38. Found: C, 44.62; H, 5.64; N, 19.54.

Dihydrochloride salt of *N*-{4-[4-(imidazolidin-2-ylideneamino)phenylsulfanyl]phenyl} guanidine (7h)



Following *Method* F, 155 mg of the pure dihydrochloride salt **7h** were obtained as a yellow solid (95%). Mp: 94-96 °C; clogP: 1.072

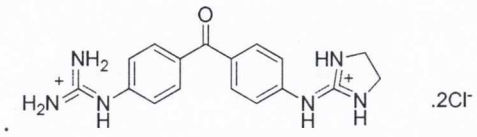
δ_H (400 MHz, D_2O): 3.67 (s, 4H, $2CH_2$, Im), 7.05-7.12 (m, 4H, Ar.), 7.19-7.27 (m, 4H, Ar.)

δ_C (100 MHz, D_2O): 42.2 ($2CH_2$, Im), 123.7, 125.5, 131.5, 132.0, 132.4, 132.9, 133.6, 134.1 (Ar.), 155.3, 157.4 (CN)

HRMS (*m/z* -ES): Found: 327.1295 ($M^+ + H$). $C_{16}H_{19}N_6S$ Requires: 327.1293

Anal. ($C_{16}H_{20}Cl_2N_6S \cdot 1.8H_2O$) Calcd: C, 44.51; H, 5.51; N, 19.46. Found: C, 44.75; H, 5.38; N, 19.33.

Dihydrochloride salt of *N*-{4-[4-(Imidazolidin-2-ylideneamino)benzoyl]phenyl} guanidine (8h)



Following *Method* F, 152 mg of the pure dihydrochloride salt **8h** were obtained as a white solid (94%). Mp: 99-101 °C; clogP: -0.228

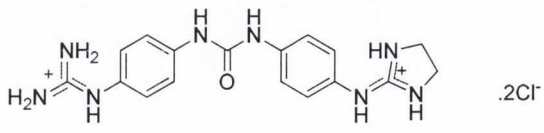
δ_H (400 MHz, D_2O): 3.80 (s, 4H, $2CH_2$, Im), 7.38 (d, 2H, J = 8.5 Hz, Ar.), 7.42 (d, 2H, J = 8.5 Hz, Ar.), 7.78-7.84 (m, 4H, Ar.)

δ_C (100 MHz, D_2O): 42.3 ($2CH_2$, Im), 121.8, 123.5, 131.5, 131.6, 133.7, 134.2, 138.9, 139.5 (Ar.); 155.3, 157.5 (CN), 197.5 (CO)

HRMS (*m/z* -ES): Found: 323.1619 ($M^+ + H$). $C_{17}H_{19}N_6O$ Requires: 323.1620

Anal. ($C_{17}H_{20}Cl_2N_6O \cdot 2.0H_2O$) Calcd: C, 47.34; H, 5.61; N, 19.48. Found: C, 47.14; H, 5.38; N, 19.87.

Dihydrochloride salt of 1-(4-guanidinophenyl)-3-[4-(imidazolidin-2-ylideneamino) phenyl]urea (9h)



Following *Method* F, 166 mg of the pure dihydrochloride salt **9h** were obtained as a white solid (94%). Mp: decomposes over 200 °C; clogP: -0.398

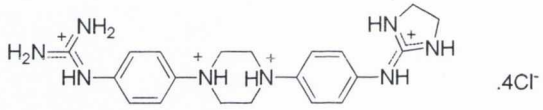
δ_H (400 MHz, D_2O): 3.57 (s, 4H, $2CH_2$, Im), 6.98 (d, 2H, J = 8.5 Hz, Ar.), 7.03 (d, 2H, J = 8.5 Hz, Ar.), 7.16 (d, 2H, J = 8.5 Hz, Ar.), 7.20 (d, 2H, J = 8.5 Hz, Ar.)

δ_C (100 MHz, D_2O): 42.1 ($2CH_2$, Im), 120.0, 120.1, 123.4, 125.4, 128.3, 129.3, 136.1, 136.8 (Ar.), 153.4, 155.4, 157.3 (CO, CN)

HRMS (*m/z* -ES): Found: 353.1832 ($M^+ + H$). $C_{17}H_{21}N_8O$ Requires: 353.1838

Anal. ($C_{17}H_{22}Cl_2N_8O \cdot 2.0H_2O$) Calcd: C, 44.26; H, 5.68; N, 24.29. Found: C, 44.12; H, 5.36; N, 24.07

Tetrahydrochloride salt of *N*-(4-{4-[4-(imidazolidin-2-ylideneamino)phenyl]-piperazin-1-yl}phenyl)guanidine (10h)



Following *Method* F, 176 mg of the pure dihydrochloride salt **10h** were obtained as a brown solid (93%). Mp: 168-170 °C; clogP: 1.438

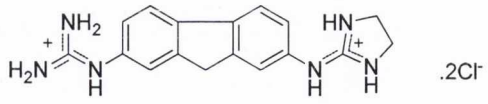
δ_H (400 MHz, D_2O): 3.62 (s, 4H, $2CH_2$ Im), 3.71 (s, 8H, CH_2 Pip), 7.23-7.31 (m, 4H, Ar.), 7.37 (d, 2H, J = 9.0 Hz, Ar.), 7.41 (d, 2H, J = 8.5 Hz, Ar.)

δ_C (100 MHz, D_2O): 42.3 ($2CH_2$, Im), 50.2, 50.9 (CH_2 Pip), 120.0, 120.3, 125.0, 126.6, 131.4, 133.0, 141.9, 143.1 (Ar.), 155.6, 157.8 (CN)

HRMS (*m/z* -ES): Found: 379.2362 ($M^+ + H$). $C_{20}H_{29}N_8$ Requires: 379.2359

Anal. ($C_{20}H_{30}Cl_4N_8 \cdot 1.0H_2O$) Calcd: C, 44.29; H, 5.95; N, 20.66. Found: C, 44.67; H, 5.85; N, 20.80.

Dihydrochloride salt of *N*-[7-(imidazolidin-2-ylideneamino)-9*H*-fluoren-2-yl]guanidine (11h)



Following *Method* F, 146 mg of the pure dihydrochloride salt **11h** were obtained as a white solid (95%). Mp: 168-170 °C; clogP: 0.667

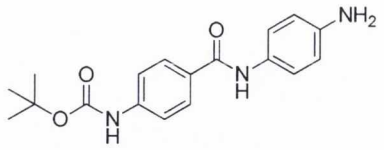
δ_H (400 MHz, D_2O): 3.71 (s, 4H, $2CH_2$ Im), 3.78 (s, 2H, $PhCH_2Ph$), 7.16 (d, 1H, J = 6.0 Hz, Ar.), 7.22 (d, 1H, J = 6.0 Hz, Ar.), 7.32 (s, 1H, Ar.), 7.38 (s, 1H, Ar.), 7.67-7.80 (m, 2H, Ar.)

δ_C (400 MHz, D_2O): 35.9 ($PhCH_2Ph$), 42.2 ($2CH_2$ Im), 119.8, 120.5, 121.6, 121.9, 123.7, 130.2, 132.3, 133.4, 138.5, 139.2, 144.7, 144.8 (Ar.), 155.7, 157.8 (CN)

HRMS (*m/z* -ES): Found: 307.1662 ($M^+ + H$). $C_{17}H_{19}N_6$ Requires: 307.1671

Anal. ($C_{17}H_{20}Cl_2N_6 \cdot 2.6H_2O$) Calcd: C, 47.92; H, 5.96; N, 19.72. Found: C, 48.20; H, 5.57; N, 19.23.

4-(*tert*-butoxycarbonylamino]-*N*-(4-aminophenyl)benzamide (12)



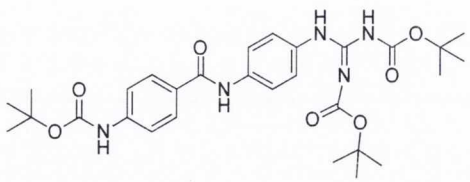
BOP (13967 mg, 31.58 mmol) was added to a cooled solution of Boc-4-Abz-OH (7492 mg, 31.58 mmol), 4-amino aniline (6829 mg, 63.15 mmol) and triethylamine (17.5 ml, 126 mmol) in anhydrous DMF (20 ml). The solution was stirred at 0°C for one hour and a further 20 hours at room temperature. The solution was diluted with EtOAc (30 mL) and washed with deionised water (3×30ml) and brine (1×30ml), dried over anhydrous Na_2SO_4 and concentrated under vacuum to give a residue that was purified with flash chromatography (Hexane:Ethyl Acetate, 1:2) 5218 mg, (54 %) of **12** as a yellow solid. Mp: 88-90 °C.

δ_H (400 MHz, DMSO-d_6): 1.50 (s, 9H, $(CH_3)_3$), 4.92 (broad s, 2H, NH_2), 6.55 (d, 2H, J = 8.8 Hz, Ar.), 7.36 (d, 2H, J = 8.8 Hz, Ar.), 7.56 (d, 2H, J = 8.8 Hz, Ar.), 7.86 (d, 2H, J = 8.8 Hz, Ar.), 9.65 (broad s, 1H, NH), 9.71 (broad s, 1H, NH)

δ_C (100 MHz, DMSO-d_6): 28.6 ($(CH_3)_3$), 80.0 ($\underline{C}(CH_3)_3$), 114.1, 117.6, 122.8, 128.7, 128.8, 129.9, 142.7, 145.5, (Ar.), 153.1, 164.6 (CO)

ν_{max} (film)/cm^{-1}: 3370, 3288, 2985 (NH), 1699, 1637, 1614, 1589, (CO, CN)

4-(*tert*-butoxycarbonylamino]-*N*-{4-[*N'*,*N''*-di(*tert*-butoxycarbonyl)guanidino]phenyl} benzamide (13a)



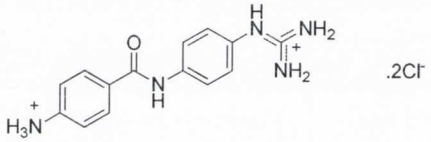
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 1:2) 618 mg, (36 %) of **13a** as a white solid were obtained. Mp: decomposes over 195 °C.

δ_H (400 MHz, DMSO-d_6): 1.42 (s, 9H, $(CH_3)_3$), 1.51 (s, 9H, $(CH_3)_3$), 1.54 (s, 9H, $(CH_3)_3$), 7.50 (d, 2H, J = 9.0 Hz, Ar.), 7.60 (d, 2H, J = 8.5 Hz, Ar.), 7.75 (d, 2H, J = 9.0 Hz, Ar.), 7.90 (d, 2H, J = 8.5 Hz, Ar.), 9.72 (broad s, 1H, NH), 9.98 (broad s, 1H, NH), 10.14 (broad s, 1H, NH), 11.48 (broad s, 1H, NH)

δ_C (100 MHz, DMSO-d_6): 29.0, 29.2, 29.4, $((CH_3)_3)$, 80.0, 80.9, 84.7, $(\underline{C}(CH_3)_3)$, 118.4, 121.8, 124.6, 129.2, 129.9, 133.2, 137.6, 144.0 (Ar.), 153.5, 153.9, 154.3 (CO), 164.0 (CN), 166.1 (CO)

ν_{max} (film)/cm^{-1}: 3300, 3268, 3174 (NH), 1726, 1649, 1634, 1604, 1596 (CO, CN)

Dihydrochloride salt of 4-Amino-*N*-(4-guanidino-phenyl)-benzamide (13b)



Following *Method* F, 171 mg of the pure dihydrochloride salt **13b** were obtained as a white solid (95 %). Mp decomposes over 220 °C; clogP : -0.889

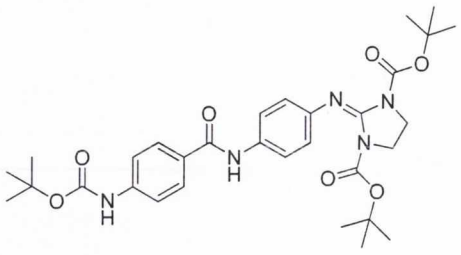
δ_H (400 MHz, D_2O): 7.27 (d, 2H, J = 7.5 Hz, Ar.), 7.50 (d, 2H, J = 7.5 Hz, Ar.), 7.56 (d, 2H, J = 7.5 Hz, Ar.), 7.92 (d, 2H, J = 7.5 Hz, Ar.)

δ_C (100 MHz, D_2O): 122.7, 122.9, 126.0, 128.9, 130.7, 133.5, 133.6, 135.8 (Ar.), 155.7 (CN), 167.5 (CO)

HRMS (*m/z* -ES): Found: 270.1342 (M^+ + H). $C_{14}H_{18}N_5O$ Requires: 270.1355

Anal. ($C_{14}H_{17}Cl_2N_5O{\cdot}1.0H_2O$) Calcd: C, 46.68; H, 5.32; N, 19.44. Found: C, 46.32; H, 5.17; N, 19.08.

2-{4-[3-(4-*tert*-Butoxycarbonylamino-phenyl)amide]phenylimino}imidazolidine-1,3-dicarboxylic acid di-*tert*-butyl ester (14a)



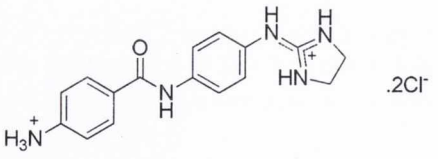
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 2:5) 1164 mg, (65 %) of **14a** as a white solid were obtained. Mp: 216-218 °C.

δ_H (400 MHz, DMSO-d_6): 1.30 (s, 18H, $2(CH_3)_3$), 1.51 (s, 9H, $(CH_3)_3$), 3.77 (s, 4H, $2CH_2$ Im), 6.83 (d, 2H, J = 8.5 Hz, Ar.), 7.59 (d, 2H, J = 8.5 Hz, Ar.), 7.63 (d, 2H, J = 8.5 Hz, Ar.), 7.91 (d, 2H, J = 8.5 Hz, Ar.), 9.70 (broad s, 1H, NH), 9.96 (broad s, 1H, NH), 10.14 (broad s, 1H, NH)

δ_C (100 MHz, DMSO-d_6): 28.8, ($2(CH_3)_3$), 29.4 ($(CH_3)_3$), 44.2 ($2CH_2$ Im), 80.8 ($\underline{C}(CH_3)$), 82.7 ($2\underline{C}(CH_3)$), 118.4, 122.0, 122.1, 129.4, 129.8, 135.1, 140.2, 143.8, (Ar.), 145.6, 151.0, (CO), 153.9 (CN), 165.7 (CO)

ν_{max} (film)/cm^{-1}: 3364, 3276 (NH), 1731, 1710, 1655, 1594 (CO, CN)

Dihydrochloride salt of 4-Amino-*N*-(4-((4,5-dihydro-1*H*-imidazol-2-yl)amino)phenyl) (14b)



Following *Method* F, 180 mg of the pure dihydrochloride salt **14b** were obtained as a white solid (92 %). Mp decomposes over 210 °C; clogP: 0.129

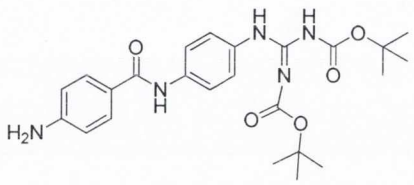
δ_H (400 MHz, D_2O): 3.67 (s, 4H, CH_2), 7.19 (d, 2H, J = 8.5 Hz, Ar.), 7.45-7.56 (m, 4H, Ar.), 7.89 (d, 2H, J = 8.0 Hz, Ar.)

δ_C (100 MHz, D_2O): 42.2 (CH_3) 122.7, 123.0, 124.1, 128.8, 131.6, 132.9, 134.0, 135.3 (Ar.), 157.9 (CN), 167.2 (CO)

HRMS (*m/z* -ES): Found: 296.1510 (M^+ + H). $C_{16}H_{18}N_5O$ Requires: 296.1511

Anal. ($C_{16}H_{19}Cl_2N_5O \cdot 1.3H_2O$) Calcd: C, 49.06; H, 5.56; N, 17.88. Found: C, 48.82; H, 5.17; N, 17.52.

1-{4-[*N'*,*N''*-di(*tert*-butoxycarbonyl)guanidino]phenyl}-3-(4-aminophenyl)amide (16)



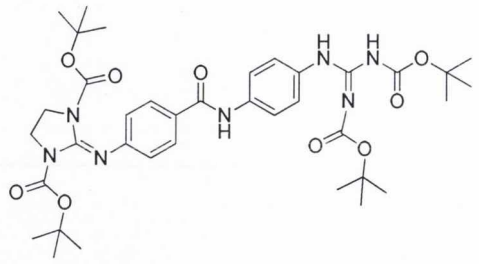
Following *Method* B and after purification with flash chromatography (Hexane:Ethyl Acetate, 2:1) 738 mg, (58 %) of **16** as a white solid were obtained. Mp: 136-138 °C.

δ_H (400 MHz, $CDCl_3$): 1.41 (s, 9H, $(CH_3)_3$), 1.53 (s, 9H, $(CH_3)_3$), 5.74 (broad s, 2H, NH_2), 6.61 (d, 2H, J = 8.4 Hz, Ar.), 7.46 (d, 2H, J = 8.4 Hz, Ar.), 7.72 ('d', 4H, Ar.), 9.79 (broad s, 1H, NH), 9.95 (broad s, 1H, NH), 11.47 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.6, 27.9 ($(CH_3)_3$), 78.7, 83.3 ($\underline{C}(CH_3)_3$), 112.5, 120.0, 120.9, 123.1, 123.2, 129.3, 136.8, 144.4 (Ar.), 152.1, 153.0 (CO), 162.7 (CN), 165.1 (CO)

ν_{max} (film)/cm^{-1}: 3474, 3373, 2984 (NH), 1718, 1647, 1624, 1608 (CO, CN)

4-[2,3-di(*tert*-butoxycarbonyl)guanidine]-4'-[1,3-di(*tert*-butoxycarbonyl)-2-imidazolidinylimino]diphenyl amide (17)



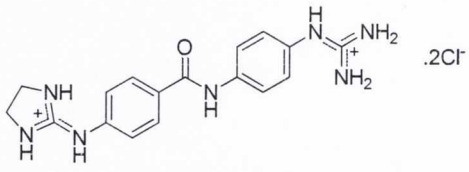
Following *Method* E and after purification with flash chromatography (Hexane:Ethyl Acetate, 1:1) 1437 mg, (65 %) of **17** as a white solid were obtained. Mp: 112-114 °C

δ_H (400 MHz, $CDCl_3$): 1.28 (s, 18H, $2(CH_3)_3$ Im), 1.47 (s, 9H, $(CH_3)_3$), 1.52 (s, 9H, $(CH_3)_3$), 3.79 (s, 2H, $2CH_2$ Im), 6.93 (d, 2H, J = 8.5 Hz, Ar.), 7.49 (d, 2H, J = 9.0 Hz, Ar.), 7.76 (d, 2H, J = 9.0 Hz, Ar.), 7.88 (d, 2H, J = 8.5 Hz, Ar.), 9.97 (broad s, 1H, NH), 10.08 (broad s, 1H, NH), 11.47 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.2 ($(CH_3)_3$ Im), 27.5, 27.6 ($(CH_3)_3$ Gu), 43.0 (CH_2), 59.8 ($\underline{C}(CH_3)_3$ Gu), 64.9 ($\underline{C}(CH_3)_3$ Im), 81.7 ($\underline{C}(CH_3)_3$ Gu), 120.3, 120.4, 123.3, 127.4, 128.4, 131.9, 136.5, 140.2 (Ar.), 152.1, 152.2, 153.0, 162.8, 164.9, 170.4 (CO, CN)

ν_{max} (film)/cm^{-1}: 3386, 3252, 3155 (NH), 1764, 1731, 1717, 1668, 1645, 1634 (CO, CN)

Dihydrochloride salt of *N*-(4-guanidinophenyl)-4-(imidazolidin-2-ylideneamino) benzamide (18)



Following *Method* F, 159 mg of the pure dihydrochloride salt **18** were obtained as a white solid (94 %). Mp decomposes over 220 °C; clogP: -0.758

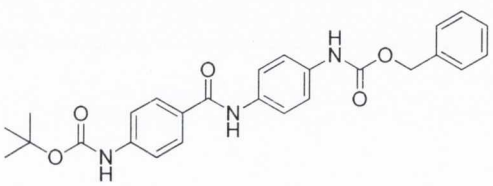
δ_H (400 MHz, D_2O): δ 3.68 (s, 4H, CH_2), 7.24 (d, 2H, *J* = 7.0 Hz, Ar.), 7.41 (d, 2H, *J* = 8.5 Hz, Ar.), 7.51 (d, 2H, *J* = 8.5 Hz, Ar.), 7.89 (d, 2H, *J* = 8.5 Hz, Ar.)

δ_C (100 MHz, D_2O): δ 42.7 (CH_2) 122.6, 123.6, 123.8, 125.1, 132.2, 133.4, 135.2, 135.9 (Ar.), 158.7 (CN), 168.4 (CO)

HRMS (*m/z* -ES): Found: 338.1734 ($M^+ + H$). $C_{17}H_{20}N_7O$ Requires: 338.1729

Anal. ($C_{17}H_{21}Cl_2N_7O{\cdot}2.7H_2O$) Calcd: C, 44.49; H, 5.80; N, 21.36. Found: C, 44.11; H, 5.38; N, 20.97.

4-(*tert*-butoxycarbonylamino)-*N*-(*N*-methyl-3-phenylethylamide)benzamide (22)



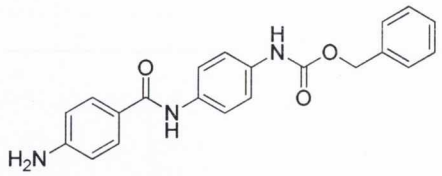
A catalytic amount of $La(NO_3)_3.6H_2O$ was added to a cooled solution of **12** (3 mmol) and Cbz (3 mmol). The reaction was monitored be thin layered chromatography. When the reaction reached completion, the reaction mixture was diluted with EtOAc and washed with water,dried over anhydrous Na_2SO_4 and concentrated under vacuum to give a residue that was purified by flash chromatography (Hexane:EtOAc, 2:1) to give **22** as a brown solid (1148 mg, 83 %). Mp: 128-130 °C.

δ_H (400 MHz, $CDCl_3$): δ 1.50, (s, 9H, $(CH_3)_3$), 5.15 (s, 2H, CH_2), 7.36 (m, 7H, Ar.), 7.67 (d, 2H, *J* = 8.5 Hz, Ar.), 7.65 (d, 2H, *J* = 9.0 Hz, Ar.), 7.88 (d, 2H, *J* = 8.5, Ar.), 9.70 (s, 1H, NH), 9.74 (s, 1H, NH), 10.01 (s, 1H, NH),

δ_C (100 MHz, $CDCl_3$): 28.0, (CH_3), 65.6 (CH_2), 79.5 ($\underline{C}(CH_3)$), 117.1, 118.3, 121.0, 128.0, 128.4, 128.6, 131.5, 131.7, 134.1, 134.7, 136.7, 142.5, (Ar.), 153.4, 164.5, 166.9 (CO)

ν_{max} (film)/cm^{-1}: 3335, 3143, 2923 (NH), 1699, 1640, 1607 (CO)

4-Amino-*N*-(*N*-methyl-3-phenylethylamide)benzamide (23)



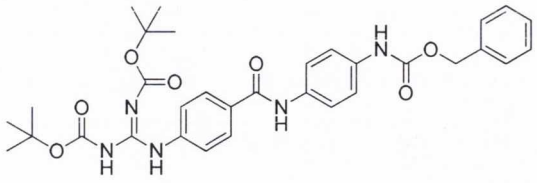
A solution of **22** (0.5 mmol) in DCM:TFA (1:1) (20 ml) was stirred at room temperature for 3 h. After that time, the solvent was eliminated under vacuum to generate the trifluoroacetate salt. This salt was redissolved in 20 ml of an aqueous solution of NaOH (2M) and washed with EtOAc (3 × 15 ml). The organic layer was washed with water (2 × 10 ml), dried over anhydrous Na_2SO_4, filtered and the solvent was removed under vacuum to yield the free amine as a yellow solid (170 mg, 94 %) . Mp : 152-154 °C.

δ_H (400 MHz, $CDCl_3$): 5.15 (s, 2H, CH_2), 5.71 (s, 2H, NH_2), 6.60 (d, 2H, J = 5.8 Hz, Ar.), 7.36 (m, 7H, Ar.), 7.65 (d, 2H, J = 4.5 Hz, Ar.), 7.71 (d, 2H, J = 5.8 Hz, Ar.), 9.67 (s, 1H, NH), 9.98 (s, 1H, NH), 10.14

δ_C (100 MHz, $CDCl_3$): 65.6 (CH_2), 112.5, 118.4, 120.7, 121.2, 128.0, 128.1, 128.4, 129.2, 134.2, 134.6, 136.7, 152.0 (Ar.), 153.4, 165.0 (CO)

ν_{max} (film)/cm^{-1}: 3389, 3310, 3088, 2923 (NH), 1699, 1640 (CO)

4-[N'N''-di(*tert*-butoxycarbonyl)guanidine]phenyl-*N*-(*N*-methyl-3-phenylethylamide) benzamide (24)



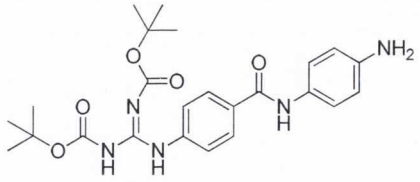
Following *Method* C and after purification with flash chromatography (Hexane:Ethyl Acetate, 2:1) 524 mg, (29 %) of **24** as a brown solid were obtained. Mp : 86-88 °C.

δ_H (400 MHz, $CDCl_3$): 1.44 (s, 18H, $2(CH_3)_3$), 5.16 (s, 2H, CH_2), 7.39 (m, 7H, Ar.), 7.68 (d, 2H, J = 9.0 Hz, Ar.), 7.74 (d, 2H, J = 8.0 Hz, Ar.), 7.74 (d, 2H, J = 8.5 Hz, Ar.), 9.75 (s, 1H, NH), 10.13 (s, 1H, NH), 10.18 (s, 1H, NH), 11.35 (s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 65.6 (CH_2), 112.5, 118.4, 120.7, 121.2, 128.0, 128.1, 128.4, 129.2, 134.2, 134.6, 136.7, 152.0 (Ar.), 153.4, 165.0 (CO)

ν_{max} (film)/cm^{-1}: 3330, 3294, 3114, 2941 (NH), 1715, 1699, 1640, 1756, 1618, 1603 (CO, CN)

4-[N'N''-di(*tert*-butoxycarbonyl)guanidine]phenyl-*N*-(4-aminophenyl)benzamide (25)



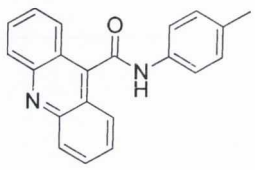
Compound **24** (0.5 mmol) was dissolved in methanol and hydrogenated at 3mbar for 4 hours. The mixture was then filtered over ceilite and washed with water (2×10 ml) and brine (10 ml). The solvent was then removed under reduced pressure to give a solid that was purified by column chromatography (Hexane:Ethyl Acetate, 2:1) to give a yellow solid (235 mg, 89 %). Mp : decomposes over 220 °C.

δ_H (400 MHz, $CDCl_3$): 1.52 (s, 9H, $(CH_3)_3$), 6.60 (d, 2H, *J* = 8.5 Hz, Ar.), 7.48 (d, 2H, *J* = 8.3 Hz, Ar.), 7.59 (d, 2H, *J* = 8.5 Hz, Ar.), 7.18-7.27 (d, 2H, *J* = 8.3 Hz, Ar.)

δ_C (100 MHz, $CDCl_3$): 27.8, 28.0 ($(CH_3)_3$), 83.2, 83.9 ($\underline{C}(CH_3)_3$), 113.6, 122.0, 122.1, 127.8, 128.5, 131.3, 139.1, 144.3 (Ar.), 153.1, 153.5, 163.1, 165.0 (CO)

ν_{max} (film)/cm^{-1}: 3281, 3140, 2976, 2930, 2871 (NH, NH_2), 1791, 1718, 1631, 1597, (CO, CN)

***N*-*p*-tolylacridine-9-carboxamide (28)**



Following *Method* G and after recrystallisation from MeOH, 140 mg (45%) of **28** as a yellow solid was obtained. Mp : decomposes over 240 °C; logP: 4.83

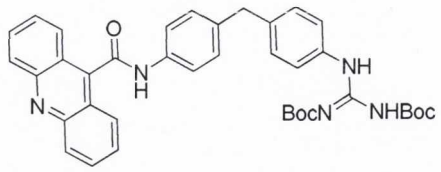
δ_H (400 MHz, $CDCl_3$): 2.34 (s, 3H, CH_3), 7.27 (d, 2H, *J* = 8.3 Hz, Ar.), 7.73 (d, 2H, *J* = 8.3 Hz, Ar.), 7.78 (t, 2H, Ar.), 8.03 (t, 2H, Ar.), 8.13 (d, 2H, *J*=8.6 Hz, Ar.), 8.32 (d, 2H, *J* = 8.6 Hz, Ar.)

δ_C (100 MHz, $CDCl_3$): 30.8 (CH_3), 120.3, 122.1, 126.0, 127.9, 129.7, 132.5, 134.0, 136.3, 146.9, 164.4 (Ar.), 206.8 (CO)

HRMS (*m/z* -ES): Found: 312.1254 ($M^+ + H$). $C_{21}H_{16}N_2O$ Requires: 312.1263

Anal. ($C_{21}H_{16}N_2O$) Calcd: C, 80.75; H, 5.16; N, 8.97. Found: C, 79.36; H, 5.43; N, 9.43.

1-(4-(4-(acridine-9-carboxamido)benzyl)phenyl)-[*N'N''*-di(*tert*-butoxycarbonyl) guanidine] (29a)



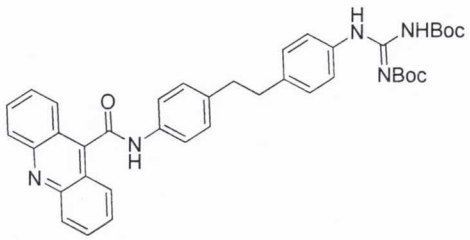
Following *Method* G and after recrystallisation from MeOH, 367 mg (57 %) of **29a** as a yellow solid was obtained. Mp : decomposes over 220 °C.

δ_H (400 MHz, $CDCl_3$): 1.41 (s, 9H, $(CH_3)_3$), 1.52 (s, 9H, $(CH_3)_3$), 3.96 (s, 2H, CH_2), 7.25 (d, 2H, *J* = 8.3 Hz, Ar.), 7.31 (d, 2H, *J* = 8.3 Hz, Ar.), 7.48 (d, 2H, *J* = 8.3 Hz, Ar.), 7.77 (d, 2H, *J* = 8.3 Hz, Ar.), 8.06 (d, 2H, *J* = 8.7 Hz, Ar.), 8.25 (d, 2H, *J* = 8.7 Hz, Ar.), 9.96 (broad s, 1H, NH), 11.00 (broad s, 1H, NH), 11.44 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.6, 27.9 ($C(CH_3)$), 48.5 (CH_2), 56.6 ($\underline{C}(CH_3)$, 120.0, 123.0, 125.4, 127.1, 128.8, 129.2, 129.3, 130.7, 146.9, 164.4 (Ar.), 186.8 (CN), 204.3 (CO)

HRMS (*m/z* -ES): Found: 645.2943 ($M^+ + H$). $C_{38}H_{39}N_5O_5$ Requires: 645.2951

1-(4-(4-(acridine-9-carboxamido)benzyl)phenethyl)-[*N'N''*-di(*tert*-butoxycarbonyl) guanidine] (30a)



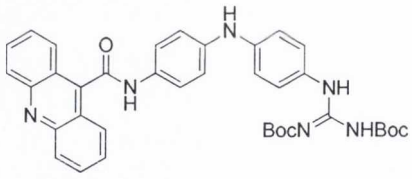
Following *Method* G and after recrystallisation from MeOH, 422 mg (64 %) of **30a** as a yellow solid was obtained. Mp : decomposes over 220 °C.

δ_H (400 MHz, $CDCl_3$): 1.40 (s, 9H, $(CH_3)_3$), 1.51 (s, 9H, $(CH_3)_3$), 2.75 (s, 4H, $2CH_2$), 7.26 (m, J = 8.0 Hz, Ar.), 7.46 (d, 2H, J = 8.0 Hz, Ar.), 7.71 (m, 4H, Ar.), 7.93 (t, 2H, Ar.), 8.05 (d, 2H, J=8.5 Hz, Ar.), 8.24 (d, 2H, J = 8.0 Hz, Ar.), 9.97 (broad s, 1H, NH), 11.01 (broad s, 1H, NH), 11.46 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.7, 27.9 ($C(CH_3)$), 36.2 ($2CH_2$), 79.1, 83.5, ($\underline{C}(CH_3)$), 118.1, 119.9, 121.7, 122.8, 125.4, 127.1, 128.6, 128.9, 129.3, 130.8, 134.5, 136.5, 137.6, 138.2, 141.6, 148.3 (Ar.), 152.2, 152.9, 162.7 (CO), 164.6 (CN)

HRMS (*m/z* -ES): Found: 660.3186 ($M^+ + H$). $C_{39}H_{41}N_5O_5$ Requires: 660.3141

1-(4-((4-(acridine-9-carboxamido)phenyl)amino)phenyl)-[*N'N''*-di(*tert*-butoxycarbonyl) guanidine] (31a)



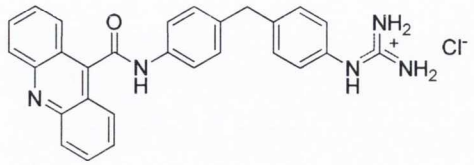
Following Method G and after recrystallisation from MeOH, 471 mg (73 %) of **31a** as a green solid was obtained. Mp : decomposes over 250 °C.

δ_H (400 MHz, $CDCl_3$): 1.41 (s, 9H, $(CH_3)_3$), 1.53 (s, 9H, $(CH_3)_3$), 7.07 (d, 2H, J = 8.6 Hz, Ar.), 7.16 (d, 2H, J = 9.0 Hz, Ar.), 7.38 (d, 2H, J = 8.6 Hz, Ar.), 7.73 (m, 4H, Ar.), 7.94 (t, 2H, Ar.), 8.25 (d, 2H, J = 7.9 Hz, Ar.), 8.31 (d, 2H, J = 7.9 Hz, Ar.), 9.87 (broad s, 1H, NH), 10.89 (broad s, 1H, NH), 11.50 (broad s, 1H, NH)

δ_C (100 MHz, $CDCl_3$): 27.6, 27.9 ($C(CH_3)$), 45.5 (CH_2), 55.8 ($\underline{C}(CH_3)$), 116.1, 117.5, 121.2, 121.7, 124.5, 125.5, 127.0, 128.2, 129.3, 130.7, 131.3, 139.9, 141.2, 141.7, 148.2 (Ar.), 153.1, 164.1 (CO)

HRMS (*m/z* -ES): Found: 646.2916 ($M^+ + H$). $C_{37}H_{38}N_6O_5$ Requires: 646.2904

Hydrochloride salt of 1-(4-(4-(acridine-9-carboxamido)benzyl)phenyl)guanidinium (29b)



Following *Method* F, 205 mg of the pure hydrochloride salt **29b** was obtained as a yellow solid (92 %). Mp : decomposes over 210 °C; logP: 4.51

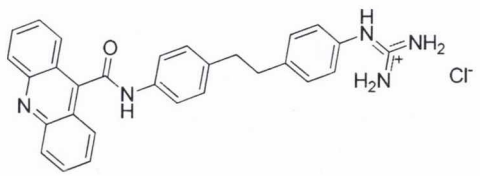
δ_H (400 MHz, D_2O): 4.01 (s, 2H, CH_2), 7.18 (d, 2H, *J* = 8.0 Hz, Ar.), 7.35 (m, 4H, Ar.), 7.61 (d, 2H, *J* = 7.5 Hz Ar.), 7.87 (t, 2H, Ar.), 8.20 (d, 2H, *J* = 7.5 Hz, Ar.), 8.31 (m, 4H, Ar.)

δ_C (100 MHz, D_2O): 30.3 (CH_3), 118.9, 119.9, 120.5, 121.5, 122.2, 125.7, 125.9, 126.7, 127.7, 128.7, 129.1, 130.0, 130.4, 132.4, 137.4 (Ar.), 141.2, 149.3 (CO, CN)

HRMS (*m/z* -ES): Found: 446.1991 ($M^+ + H$). $C_{28}H_{23}N_5O$ Requires: 446.1981

Anal. ($C_{28}H_{24}ClN_5O.1.3H_2O$) Calcd: C, 66.54; H, 5.30; N, 13.86. Found: C, 67.06; H, 5.68; N, 13.57.

Hydrochloride salt of 1-(4-(4-(acridine-9-carboxamido)phenethyl)phenyl)guanidine (30b)



Following *Method* F, 145 mg of the pure hydrochloride salt **30b** was obtained as a yellow solid (93 %). Mp : decomposes over 220 °C; log P: 5.48

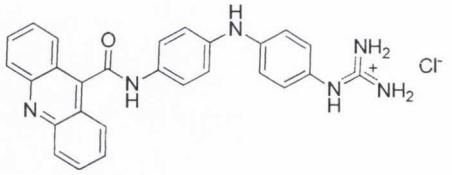
δ_H (400 MHz, D_2O): 2.19 (s, 4H, $2CH_2$), 7.26-7.36 (m, 6H, Ar.), 7.62 (d, 2H, J = 8.3 Hz, Ar.), 7.97 (t, 2H, Ar.), 8.31 (t, 2H, Ar.), 8.37-8.41 (m, 4H, Ar.)

δ_C (100 MHz, D_2O): 30.8 ($2CH_3$), 118.6, 119.9, 121.2, 121.9, 122.7, 123.5, 126.3, 127.4 128.3, 129.4, 129.7, 130.3, 137.4, 137.9, 138.2 , (Ar.), 141.5, 148.6 (CO)

HRMS (*m/z* -ES): Found: 460.2085 ($M^+ + H$). $C_{29}H_{25}N_5O$ Requires: 460.2093

Anal. ($C_{29}H_{26}ClN_5O$) Calcd: C, 66.14; H, 5.63; N, 13.30. Found: C, 65.83; H, 5.92; N, 13.54.

1-(4((4-(acridine-9-carboxamido)phenyl)amino)phenyl)guanidine (31b)



Following *Method* F, 208 mg of the pure dihydrochloride salt **31b** was obtained as a green solid (93 %). Mp : decomposes over 250 °C; log P: 4.93

δ_H (400 MHz, D_2O): 7.04 (m, 6H, Ar.), 7.48 (d, 2H, J = 8.5, Ar.), 7.79 (t, 2H, Ar.), 7.87 (t, 2H, Ar.), 8.17 (m, 6H, Ar.).

δ_C (100 MHz, D_2O): 118.4, 118.6, 118.9, 119.9, 121.1, 121.9, 123.0, 123.5, 126.2, 127.8, 128.9, 129.4, 137.9, 137.4, 139.6 (Ar.), 141.1, 149.9 (CO, CN).

HRMS (*m/z* -ES): Found: 447.1937 ($M^+ + H$). $C_{27}H_{22}N_6O$ Requires: 447.1933

Anal. ($C_{27}H_{23}N_6ClO.0.7H_2O$) Calcd: C, 65.57; H, 4.77; N, 16.99. Found: C, 64.98; H, 5.13; N, 16.87.

10.2 Biophysical experiments

10.2.1. DNA Binding Assays

Thermal melting experiments were conducted with a Varian Cary 300 Bio spectrophotometer equipped with a 6x6 multicell temperature-controlled block. Temperatures were monitored with a thermistor inserted into a 1-mL quartz curvet containing the same volume of water as in the sample cells. Absorbance changes at 260 nm were monitored from a range of 20 °C to 90 °C with a heating rate of 1 °C/min and a data collection rate of five points per °C. The salmon sperm DNA was purchased from Sigma Aldrich (extinction coefficient ε_{260}=6600 $cm^{-1}M^{-1}$ base). A quartz cell with a 1-cm path length was filled with a 1-mL solution of DNA polymer or DNA-compound complex. The DNA polymer (150 μM base) and the compound solution (15 μM) were prepared in a phosphate buffer [0.01 M Na_2HPO_4/NaH_2PO_4], adjusted to pH 7) so that a compound to DNA base ratio of 0.1 was obtained. The thermal melting temperatures of the duplex or duplex-compound complex obtained from the first derivative of the melting curves are reported. There were 5 % errors involved in the measurements

10.2.2. UV Spectroscopy: Determination of the pK_a

All UV absorbance experiments were conducted on a Cary 300 UV spectrophotometer. A quartz cell with a 1 cm path length was used for all absorbance studies. Compound stock solutions were 30 μM and contained 10 mM sodium phosphate buffer (pH 2). The pH of the solution was increased from the addition of aliquots of a NaOH solution (0.1 M) and the corresponding absorption spectra were recorded.

10.2.3. UV Spectroscopy: Determination of DNA binding affinity

For the evaluation of the DNA binding affinity compound stock solutions were 6.67 μM that contained 10 mM sodium phosphate buffer (pH 7) were used. The DNA at increasing ratios was then titrated into the compound solution and the corresponding absorption spectra were recorded under the same conditions. All concentrations were determined using the appropriate extinction coefficients.

10.2.4. UV Spectroscopy: The effect of increasing the ionic strength

A solution DNA-complex where 90 % of the compound was initially bound was prepared in phosphate buffer (10 mM) at pH 7. Aliquots of a salt solution (5 M NaCl) were titrated until

the salt concentration reached 0.2 M. From a plot of $\log(Na^+)$ against log(K), the slope SK (=$Z\psi$) was determined.

10.2.5. Circular Dichroism Spectroscopy

CD spectra were collected with a JASCO J–800 spectrometer at different ratios of compound to DNA [poly(dA-dT)$_2$ was used] at 25 °C in phosphate buffer. Titrations were carried out by addition of aliquots of the relevant compound (0.5 mM) to a DNA solution (18.75 μM in base pair) in a 1 cm quartz cuvette and scanned over a desired wavelength range.

10.2.6. Linear Dichroism Spectroscopy

LD spectra were collected with a JASCO J–810 spectrometer at different ratios of compound to DNA at 25 °C in phosphate buffer. Each flow LD spectrum was acquired from 200 nm to 400 nm and reflects the average of two scans. The couette flow orientation was used for sample orientation for all LD studies. The DNA solutions were 189 μM with a compound to DNA base pair ratio of 1 to 5.

10.2.7. Surface Plasmon Resonance

Biosensor–SPR experiments were conducted as previously described with a BIAcore 2000 instrument (Biacore AB) using degassed MES buffer (10 mM 2–(N–morpholino)ethanesulfonic acid, 1 mM EDTA, 92mM NaCl, 0.0005%v/v of surfactant P20, pH 6.25) at 25 °C. The 5'–biotinlabeled DNA hairpins were purchased from Midland Certified Reagent Co., Inc. (Midland, TX), with HPLC purification. The DNA hairpin sequences examined included 5'biotin–CGAATTCGTCTCCGAATTCG–3', 5'biotin-CGTTAACGTCTCCGTTAACG–3', and 5'biotin–CGCGCGCGTTTTCGCGCGCG–3', referred to in the text as AATT, TTAA, and CG respectively. The DNA hairpins were immobilized on a streptavidin–derivatized gold chip (SA chip from BIAcore) by manual injection of a 25 nM hairpin DNA solution with a flow rate of 1 μL/min until the response units (RUs) reach about 375–415. Flow cell 1 was left blank for reference subtraction while flow cells 2, 3, and 4 were immobilized with three different DNA hairpins. Typically, a series of different concentrations of ligand was injected onto the chip at 25 °C with a flow rate of 20 μL/min for a period of 5-minute followed by 5-minute dissociation. After the dissociation process, the chip surface was regenerated with a 20-μL injection of 200 mM NaCl and 10mM NaOH solution, injection tube rinsing, and multiple 1-minute buffer injections. The observed steady–state responses, RU_{obs}, are proportional to the amount of ligand bound, and the

maximum response per ligand bound (RU_{max}) was calculated as previously described.[4,5] The binding constants were obtained from fitting RU_{obs} vs free ligand concentration using $RU_{obs} = RU_{max} (K_1L+2K_1K_2L^2)/(1+K_1L+K_1K_2L^2)$; (L= ligand concentrations in the flow solution).

10.2.8. Isothermal Titration Calorimetry

ITC experiments were performed using a MicroCal VP–ITC instrument (MicroCal Inc., Northampton, MA) interfaced with a computer equipped with VP–2000 viewer instrument control software. ITC data were analysed with Origin 7.0 software. In ITC experiments, 10 μL of 2.48 mM compound solution in 10 mM phosphate buffer were injected every 300 seconds for a total of 29 injections into a solution of DNA in the calorimeter cell at 1 mM (in base pair). The observed heat for each injection (peak) was measured by area integration of the power peak with respect to time.

10.3. References

1. A. I. Vogel, B. S. Furniss, A. J. Hannaford, P.W.G. Smith, A.R. Tatchell, *Vogel's Textbook of Practical Organic Chemistry* (5th Edition), 1989, Pearson Education Limited

2. Dumas, J.; Khire, U.; Lowinger, T. B.; Paulson, H.; Riedl, B; Scott, W. J.; Smith, R. A.; Wood, J. E.; Hatoum-Mokdad, H.; Johnson, J.; Lee, W.; Redman, A.; Sibley, R.; Renick, J.; Patent US 244120, **2007**

3. Botez, I.; David-Basei, C.; Gourlaoueen, N.; Nicolaie, E.; Balavoine, F.; Valette, G.; Serradeil-Le Gal, C.; Patent WO 108965, **2006**

4. Nguyen, B.; Tanious, F. A.; Wilson, W. D. *Methods*, **2007,** *42*, 150

5. Tanious, F. A., Nguyen, B.; Wilson W.D. *Methods Cell Biol.* **2008**, *84*, 53

Appendix

Table 1. Crystal data and structure refinement for **4c**.

Identification code	shelxl	
Empirical formula	C15 H20 C2 N4	
Formula weight	1309.00	
Temperature	150(2) K	
Wavelength	0.71075 Å	
Crystal system	Monoclinic	
Space group	C2/c	
Unit cell dimensions	a = 18.096(4) Å	□ = 90°.
	b = 19.551(4) Å	□ = 90°.
	c = 4.7445(10) Å	□ = 90°.
Volume	1678.6(6) $Å^3$	
Z	8	
Density (calculated)	1.295 Mg/m^3	
Absorption coefficient	0.386 mm^{-1}	
F(000)	688	
Crystal size	0.5 x 0.3 x 0.2 mm^3	
Theta range for data collection	2.37 to 24.99°.	
Index ranges	-21<=h<=21, -18<=k<=23, -5<=l<=3	
Reflections collected	6895	
Independent reflections	2420 [R(int) = 0.0232]	
Completeness to theta = 24.99°	99.0 %	
Absorption correction	None	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	2420 / 1 / 192	
Goodness-of-fit on F^2	1.080	
Final R indices [I>2sigma(I)]	R1 = 0.0232, wR2 = 0.0611	
R indices (all data)	R1 = 0.0240, wR2 = 0.0620	
Absolute structure parameter	0.03(5)	
Largest diff. peak and hole	0.155 and -0.173 $e.Å^{-3}$	