UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Centre for Biological Sciences

Volume 1 of 1

Interaction of Pyrrolobenzodiazepine conjugates with DNA

by

Mohammad Anwarul Basher

Thesis for the degree of Doctor of Philosophy

September 2015

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Centre for Biological Sciences

Thesis for the degree of Doctor of Philosophy

INTERACTION OF PYRROLOBENZODIAZEPINE CONJUGATES WITH DNA

Mohammad Anwarul Basher

Pyrrolobenzodiazepines (PBD) are a group of naturally occurring antitumor antibiotics that consist of a tricyclic system containing anthranilate, diazepine and pyrrolidine ring. The compounds bind preferentially to 5'-Pu-G-Pu-3' sequences in the minor groove of double stranded DNA by covalently linking the diazepine ring with the 2-amino group of guanine. The PBD binding sites can be extended by conjugating the C8 position of the anthranilate ring with heterocyclic moieties such as pyrrole, imidazole and thiazole. Alternatively, benzofused rings such as benzothiophene, benzofuran and indole can be used for conjugate formation; some natural compounds containing these benzofused heterocycles have been found to show anticancer activities.

We have examined the sequence selectivity of PBD-conjugates using DNase I footprinting with substrates that contain every tetranucleotide (MS1 and MS2) and every symmetrical hexanucleotide (HexA and HexB). Conjugates with pyrrole, imidazole and thiazole were found to bind adjacent to AT-rich sequences; their binding preferences towards AT-rich sequences were confirmed with synthetic DNA fragments in which the compounds bound at ATATAT with a guanine at either the 5'- or 3'-end. Conjugation with benzothiophene and benzofuran benzofused rings produced similar binding profiles, though indole conjugates did not appear to bind. The benzofused-conjugates bound at lower concentrations than the polyamide conjugates. These binding sites indicate that the conjugated component plays a vital role in determining the sequence selectivity, while the PBD unit forms the covalent linkage with guanine. PBD-conjugates were also studied by fluorescence melting to evaluate their effect on melting profiles of different oligonucleotide duplexes; the compounds produced biphasic melting curves with large increases in melting temperature.

We also describe a simple ligation assay for determining the sequence selectivity of PBD-dimers. The assay involves chemically joining two DNA duplexes by a PBD-dimer when a binding site is generated by overhanging single-stranded 'sticky ends'. Oligonucleotides containing GATC or GTAC or GAATC overhangs were used to demonstrate that SJG-136 has strongest affinity for GATC; it also produces cross-links with GTAC, but not with GAATC.

The final chapter uses DNase I footprinting to examine the sequence specificity of two pyrrole/imidazole hairpin polyamides that were expected to bind to the sequences (A/T)GG(A/T)G(A/T) and (A/T)GG(A/T)C(A/T). We confirm the expected selectivity, but demonstrate that they can also form strong interactions with some secondary sites.

ACKNOWLEDGEMENTS

In the name of Allah, the Beneficent, the Merciful.

I would like to express my deepest thanks to Allah for giving me the efforts to finish this thesis successfully. Without His permission, I would never be able to reach at this stage.

I would like to extend my deepest gratitude to my parents for their love, care and enduring all the sufferings caused by me.

Next, I would like to express my deepest gratitude to my supervisor, Professor Keith Fox for his invaluable suggestions, guidance and encouragement that have substantially helped in my work through the last four years; it is truly a great experience to work under his supervision. Thanks to all my colleagues- Mohammed, Tara, Doreen, Ibrahim, David, Scott, Helen, Michael and James for their support.

Last but not the least, I would like to thank, my wife- Loba for her unimaginable support and encouragement in the last four years.

Table of Contents

List of Figures

List of Tables

Abbreviations

Ac	Acetate
ATP	Adenosine Triphosphate
DNA	Deoxyribonucleic acid
Hp	Hairpin polyamide
Im	Imidazole
IPTG	isopropylthio-β-galactoside
PBD	Pyrrolobenzodiazepine
Py	Pyrrole
Thz	Thiazole
X-gal	5-bromo-4-chloro-3-indoyl-β-*D*-galactoside

1 Introduction

Cancer is one of the leading causes of morbidity and mortality worldwide and caused 8.2 million deaths in 2012 (1). Many research efforts have been made to combat cancer that includes early detection, prevention of causative agents and invention with new chemotherapeutics. As cancer is characterized by uncontrolled cell growth, chemotherapeutic drugs primarily work by impairing cell division, targeting mainly the fast-dividing cancer cells. This cytotoxicity broadly results from the DNA damage of target cells and inhibition of the cellular machinery involved in cell division. The chemotherapeutic agents can be classified in various ways based on their chemical structure, mechanism of action, cell cycle specificity, etc. Alkylating agents, one of the oldest classes of anticancer agents, act by binding covalently with DNA. These agents, which are either from natural sources or of synthetic origin, can target specific DNA sequences. The binding of these ligands to specific sequences can affect gene transcription, perturb the cell cycle or involve DNA repair mechanisms. Other chemotherapeutic classes include antimetabolites, topoisomerase inhibitors, cytotoxic antibiotics, etc. Antimetabolites are structural analogues of naturally occurring compounds; antimetabolites especially antifolates are effective against a variety of malignancies where they block the production of nucleic acids by competing for binding sites on enzymes and incorporation of nucleic acids (2). Topoisomerase inhibitors affect the enzyme topoisomerase, which is involved in DNA replication by relieving the superhelical stress generated during replication process; the inhibitors stabilize the complex between the enzyme and DNA, thus preventing DNA replication (3,4). Cytotoxic antibiotics such as anthracyclines (5), actinomycin (6), bleomycin (7), mitomycin (8), etc. affect the function and synthesis of nucleic acids in different ways.

1.1 Alkylating agents

Alkylating agents are compounds which alkylate nucleophilic moieties of DNA or protein by covalent attachment of an alkyl group (9,10). In general, alkylation is a nucleophilic substitution reaction in which an alkyl group replaces a hydrogen atom. There are a number of preferential sites for DNA alkylation depending on the nature of nucleophile and alkylating agent (11-13). Mono-functional agents form covalent adducts while bi-functional compounds produce inter-strand or intra-strand DNA cross-links. Alkylation usually occurs at guanine N7 and O6, adenine N1 and N3 and cytosine N3 (12,14,15). Mitomycin and new classes of minor groove binders also revealed guanine N2 as a target of DNA alkylation. Alkylation is also observed at adenine N6, N7, thymine O2, N2, O4 and cytosine O2 to a lesser extent (13,16) (Figure 1.1). The

stability of a DNA alkylation adduct depends on the site of attachment – the most stable adducts are observed with O6G, N1G, N2G and O4T, while alkylation at other positions produces chemically less stable DNA adducts (17). O-alkylation and N-alkylation also differ in their cellular responses. N-alkylation usually produces toxic effects as it blocks DNA replication or forms DNA lesions. On the other hand, O-alkylation is mutagenic and cytotoxic. For example, O6MeG is mutagenic as it forms an O6MeG:T mispair which leads to an A:T transition mutation (12,13,18). Besides alkylation of the DNA bases, the phosphodiester bond of the DNA backbone can also be alkylated as it offers numerous negative charges on its oxygens (11).

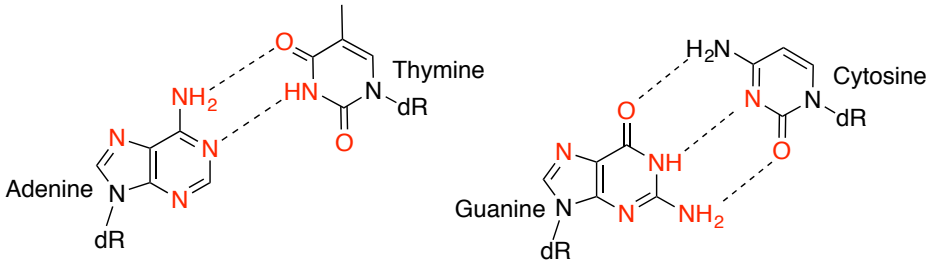


Figure 1.1: Different sites of base alkylation (red coloured nitrogen and oxygen)(19).

1.1.1 Nitrogen mustards

Nitrogen mustards are derived from sulphur mustards and share a common bis-(2-chloroethyl)amino motif that causes formation of an aziridinium ion; this ion forms a covalent link with N7 of guanine (Figure 1.2) forming a mono-adduct. Intra- or inter-strand crosslinks are also observed when a second chloro group reacts with the N7 of an adjacent guanine (20). Important nitrogen mustard alkylating agents such as melphalan and chlorambucil are mainly effective against retinoblastoma (21) and chronic lymphatic leukaemia (22) respectively.

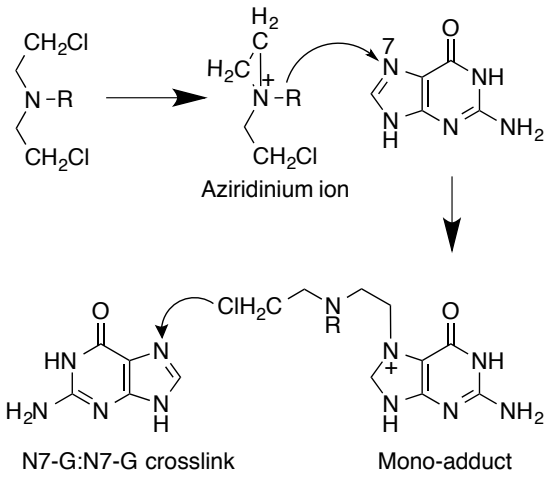


Figure 1.2: Formation of aziridinium ion responsible for the alkylation of guanine N7 by nitrogen mustards (19)

1.1.2 Oxazaphosphorine

Oxazophosphorines, such as cyclophosphamide, are an important class of alkylating agents with a wide spectrum of anticancer activity. The Oxazaphosphorine structure is composed of an oxazaphosphorine-2-amine 2-oxide heterocycle containing one atom of phosphorous, one atom of nitrogen and one of oxygen (23). Important derivatives of this class include cyclophosphamide and ifosfamide, These compounds are metabolized by liver cytochrome P450 enzymes, that result in ring opening of the heterocycle and subsequent production of the nucleophile - chloroethylaziridine (Figure 1.3) (19); this nucleophile then crosslinks between N7G:N7G. Cyclophosphamide is widely used in combination with cisplatin and etoposide for the treatment of small lung cancer, in combination with methotrexate, fluorouracil and doxorubicin for breast cancer. Ifosfamide is frequently administered to treat adult and paediatric tumours in both haematological and non-haematological disease (23).

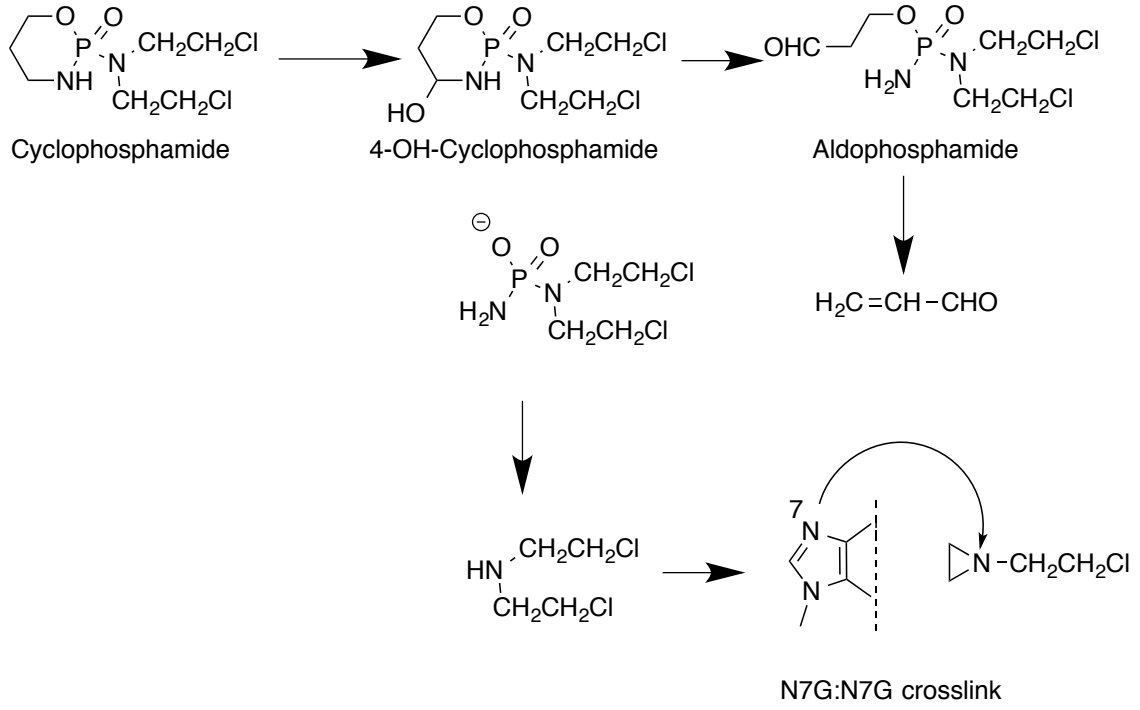


Figure 1.3: Mechanism of DNA alkylation by cyclophosphamide (24)

1.1.3 Ethylene amines

This class of alkylating agents is characterized by the presence of one or more aziridine rings. Thiotepa (triethylenethiophosphomide) and mitomycin C are the important members of this group. Thiotepa is a trifunctional alkylating agent; *in vitro* reaction of thiotepa with DNA shows multiple alkylations (25,26) at the N1 position of

thymidine, O2 position of cytidine, N1, N6 or N7 of adenosine and N1, O6 or N7 position of guanine. When incubated with carcinoma cells (27,28), thiotepa acts as a prodrug for aziridine performing as a cell-penetrating carrier for aziridine, which is released intracellularly after hydrolysis. The released aziridine then alkylates N7 of guanine and N3 or N7 of adenine.

In case of mitomycin C, reactive electrophilic centres are generated at C-1 and C-10, leading to cross-linking of DNA, following one- or two-electron reduction of the quinone moiety (29-31). Two-electron reduction generates the hydroquinone followed by loss of tertiary methoxy group forming mitosene, presumably assisted by the indole nitrogen and opening of the aziridine to form an electrophile at C-1 capable of alkylating DNA at N2 of guanine. After elimination of carbamate group, again initiated by the nitrogen lone pair of the indole ring, a second electrophile is formed at C-10 that binds DNA once more, thereby completing the cross-linking (Figure 1.4).

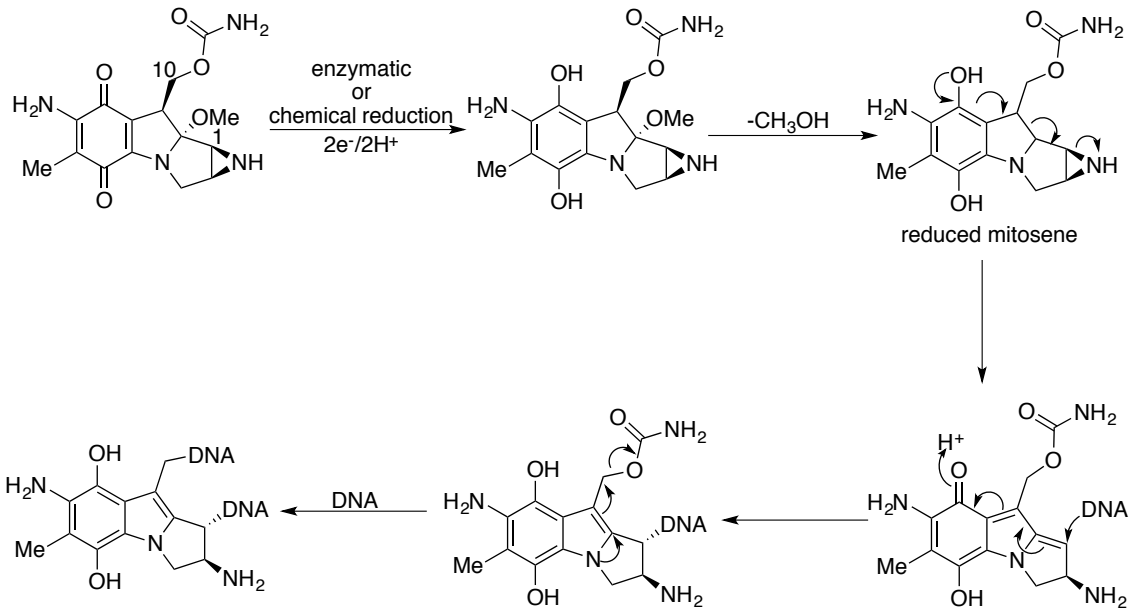


Figure 1.4: Mechanism of action of mitomycin C, activated by enzymatic or chemical reduction (32)

1.1.4 Nitrosoureas

These agents are used for the treatment of brain tumours, melanomas and pancreatic tumours. Nitrosoureas show antitumor activity by alkylating N1 position of guanine and N3 of cytosine. Their mechanism of action depends on the formation of diazohydroxide, which in turn generates the reactive chloroethyl cation. The chloroethyl cation first forms mono-adduct at O6 of guanine, which ultimately leads to formation of N1G:N3C inter-strand cross-link (Figure 1.5).

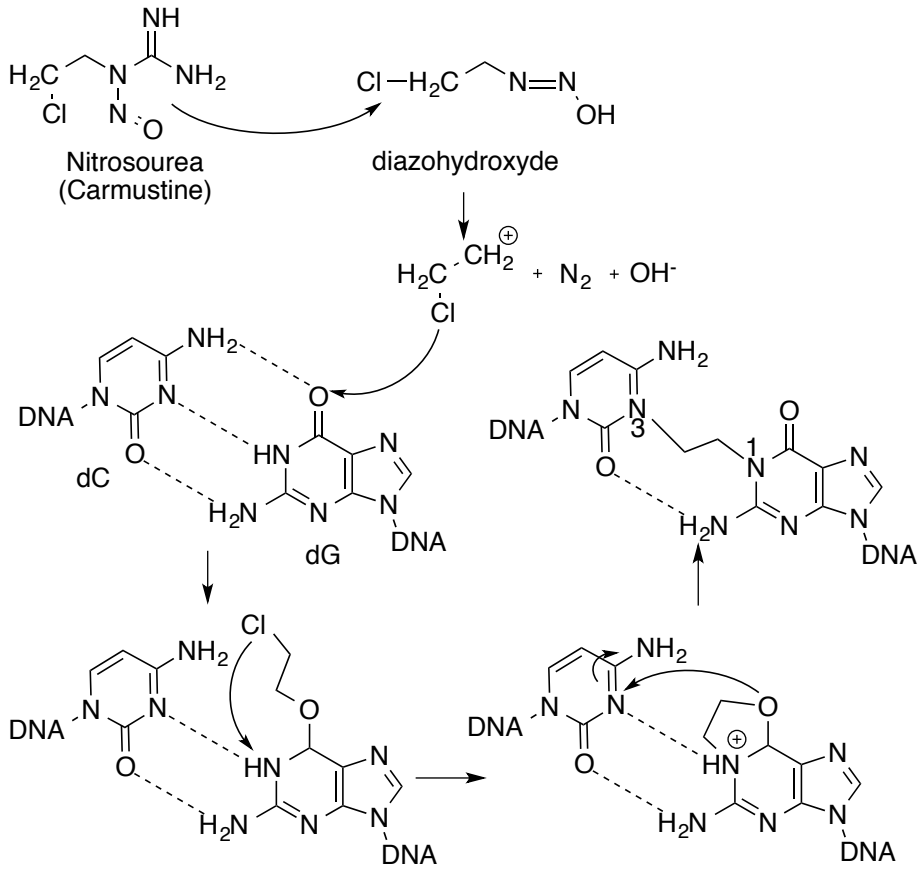


Figure 1.5: Mechanism of action of nitrosoureas by generating diazohydroxide (19)

1.1.5 Alkyl alcane sulfonates

Alkyl alkane sulfonates represent a homologous series of alkanediol dimethanesulfonates of general formula $H_3C.SO_2O.(CH_2)_n.OSO_2.CH_3$. The best-known agent of this series is busulfan (n=4). Busulfan produces cross-linking in DNA through a GG bridge (34); a diguanyl derivative (35) was isolated after incubation of DNA with busulfan, which upon alkaline elution showed the presence of an interstrand crosslinking product (Figure 1.6). The compound also forms an intrastrand crosslinking adduct between N7 of guanine and N3 of adenine. Busulfan is used for the treatment of refractory chronic myeloid leukaemia and prior to hematopoietic stem cell transplant in grafts (10,36).

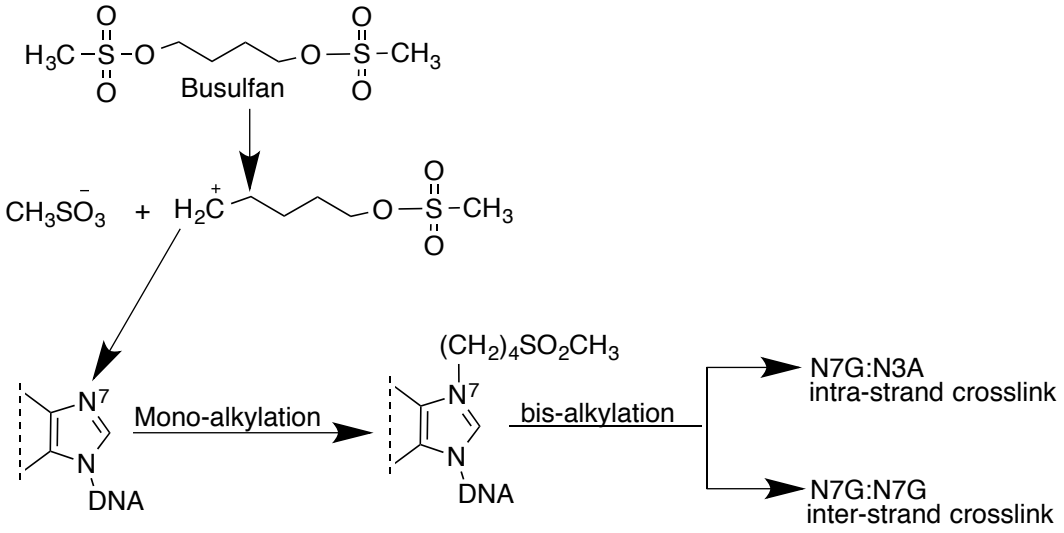


Figure 1.6: Mechanism of DNA alkylation by alkyl alcane sulfonates (19)

1.1.6 Triazenes and hydrazines

The triazenes are a group of alkylating agents containing a triazenyl group as the active moiety. The group includes dacarbazine, temozolomide, mitozolomide (Figure 1.7); the active moiety- triazenyl- is composed of three adjacent nitrogen atoms, which are responsible for their chemical, physical and antitumor activity. Dacarbazine, which is structurally related to purines, was designed to act as an antimetabolite by interfering with purine biosynthesis; however, its mechanism of action matches with the antitumor activity of other alkylating agents. The compound is first activated by liver microsomes, spontaneously producing molecular nitrogen and a methyldiazonium cation (37). This cation reacts with DNA purine bases, forming methyl adducts. The most favourable site for alkylation is at the N7 position of guanine but O^6-methyguanine (O^6-meG) is thought to be the main lesion responsible for its cytotoxicity (38). Temozolomide also causes DNA lesions by an identical pathway, but it does not require enzyme activation, rather it is spontaneously activated in aqueous solution at physiological pH. Dacarbazine is used for the treatment of melanomas, lymphomas and sarcomas and temozolomide are for brain tumours (39,40); mitozolomide is not used clinically due to its severe myelosuppressive effects.

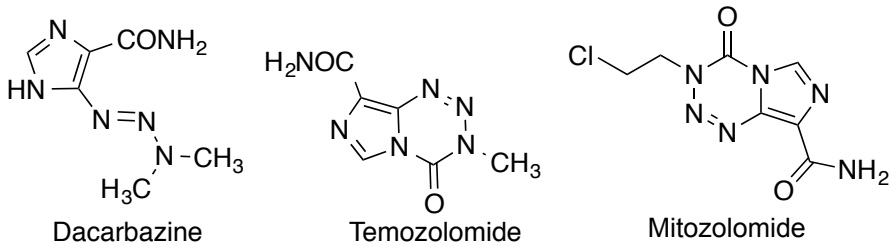
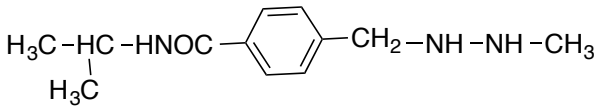


Figure 1.7: Chemical structure of dacarbazine, temozolomide and mitozolomide

Procarbazine (Figure 1.8) belongs to the hydrazine family and is used for the treatment of leukaemias. The compound is first oxidized to azoprocarbazine by cytochrome P450 and monoamine oxidase (41-43). Azoprocarbazine then undergoes further oxidation to a mixture of methylazoxyprocarbazine and benzylazoxyprocarbazine, finally producing a diazonium ion- a highly reactive group responsible for cytotoxicity via mono-adduct formation at the O6 and N7 positions of guanine.



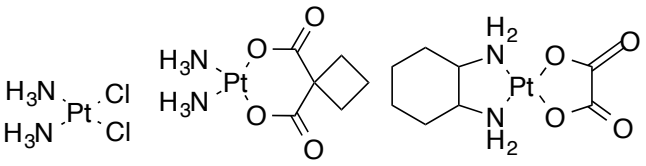
Procarbazine

Figure 1.8: Chemical structure of procarbazine

1.1.7 Platinum derivatives

Platinum derivatives emerged as potential anticancer therapeutics in 1970s with the use of cisplatin (36,44,45). Since then, twenty-three other platinum based compounds have entered into clinical trials, though only two- carboplatin and oxaliplatin (Figure 1.9) have received global approval. For anticancer activity, cisplatin is first aquated inside the cell by loosing one or both of the chloride ligands (46). This aquated cisplatin forms a covalent coordinate bond with the N7 position of guanine and subsequent coordination with second guanine produces a 1,2-GpG intrastrand adduct (Figure 1.10). The adduct causes DNA to bend and unwinds the helix (47). This distortion prevents DNA replication, which ultimately leads to cellular apoptosis. Cisplatin in combination with other anticancer drugs is used for the treatment of testicular cancer, ovarian, bladder, melanoma, non-small cell lung cancer, small cell lung cancer, lymphomas and myelomas. However, the use of cisplatin is reduced because of some severe side effects such as nephrotoxicity, neurotoxicity, ototoxicity and myelosuppression as well as the development of drug resistance. Resistance to cisplatin is occurred mainly by two broad mechanisms- first, failure of a sufficient amount of platinum to reach the target DNA. Many tumour cells show acquired resistance to cisplatin because of reduced platinum accumulation in comparison to normal cells (48). Due to high polarity of cisplatin molecule, the compound enters cells relatively slowly that other classes of small molecular cancer drugs. Copper transporter (CTR1), a major plasma membrane transporter involved in copper homeostasis, has been found to play a substantial role in cisplatin intake (49,50). Mouse embryonic fibroblasts that were exposed to clinically relevant concentration of 2 μM cisplatin showed only 35% accumulation of drug in comparison to wild type cells; loss of CTR1 led to a 2-3 fold increase in cisplatin resistance (51). In addition to this, resistance to

cisplatin is also correlated with increased levels of cytoplasmic thiol-containing species such as tripeptide glutathione and metalothiones. These thiol-containing species are rich in cysteine and methionine that lead to detoxification of cisplatin as platinum binds avidly with sulphur; a significant correlation between sensitivity to cisplatin and levels of glutathione has been reported (52). The conjugation of cisplatin with glutathione is catalysed by glutathione-S-transferase that makes the compound more anionic and more readily exported from cells by the ATP-dependent glutathione-S-conjugate export pump, thus further increasing resistance towards cisplatin (53). A study of two ovarian cancer cell lines that are derived from the same patient before and after the onset of drug resistance showed 2.9 fold higher levels of GST in the cells derived from the drug resistance tumour (54). After platinum-DNA adducts have been formed, resistance to cisplatin is developed by the initiation of nucleotide excision repair (NER) pathway (55). NER pathway first recognises DNA damage by cisplatin followed by the formation of a complex to unwind the damaged area to excise it and finally ligating the DNA strand to maintain normal helical structure; a protein known as excision repair cross-complementation group 1 (ERCC1) protein plays a vital role in this repair process. Deficiency in NER system and reduced levels of ERCC1 in testicular cell lines results cisplatin therapy effective in treating metastatic testicular cancer (56,57). The DNA repair mechanism has been discussed further in section 1.3. The toxicity of platinum drugs is directly related to the ease with which the leaving groups are aquated. Cisplatin, having two highly liable chloride ligands, is easily aquated while carboplatin is aquated at a much slower rate than cisplatin due to the replacement of chloride by cyclobutanedicarboxylate (58). Once aquated, carboplatin produces same the active component and thus forms the same DNA adduct. In oxaliplatin, the two ammine ligands are replaced by 1,2-diaminecyclohexane and two chlorides are replaced by oxalate. Oxaliplatin also forms GpG intrastrand adducts but the presence of the oxalate ion reduces the severity of side effects from cisplatin (59). Additionally, after forming an intrastrand adduct, the bulky cyclohexane moiety points into DNA major groove which prevents binding of DNA repair proteins (60) thus oxaliplatin overcomes cisplatin resistance.



Cisplatin Carboplatin Oxaliplatin

Figure 1.9: Chemical structures of the main platinum derivatives used in the clinic

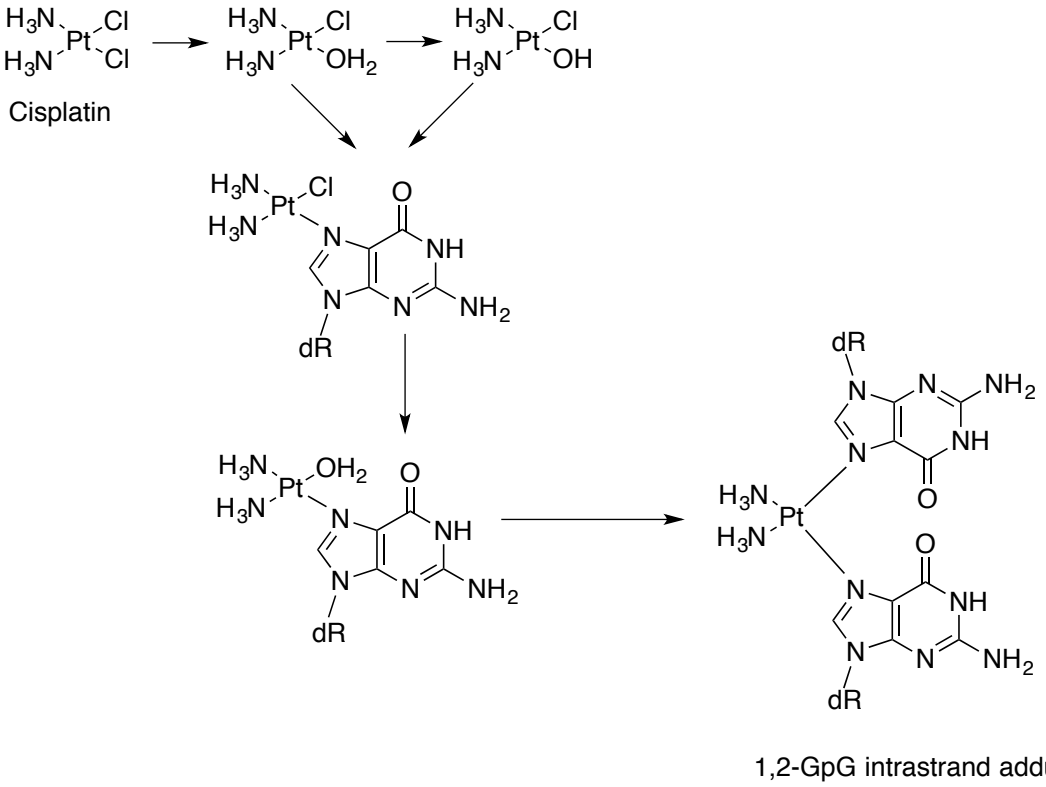


Figure 1.10: Intrastrand formation induced by cisplatin (19)

1.2 Minor groove binding alkylating agents

The alkylating agents discussed above (except mitomycin C) primarily interact with the major groove of DNA. The major groove of DNA is deeper and wider, thus provides easy access for bulky molecules (61). Also major groove has multiple sites of interaction offering comparatively strong binding to ligands (62,63). On the other hand, the minor groove is narrow and shallow. The minor groove becomes a site of great focus for developing new drugs as it offers non-covalent highly sequence selective interactions for a large number of small molecules such as distamycin, benzimidazoles, bis-phenylamidiniums etc. (64). Many minor groove binders exhibit antiprotozoal, antiviral and antibacterial activities. Some of these compounds show antitumor properties and are considered for therapeutic applications. Most of these simple agents bind at AT-rich sequences within the minor groove. These minor groove binders can be categorized as either i) compounds that induce permanent DNA damage through direct or indirect irreversible interaction with DNA nucleotides; for example, mitomycin C, anthramycin derivatives (pyrrolobenzodiazepines), CC-1065, ecteinascidine derivatives, *etc.* and ii) compounds that cause reversible inhibition of DNA-dependent functions by physical interaction with DNA, such as distamycin A, diarylamidines, bis-benzimidazoles, *etc*. Here, I will discuss a few classes of minor groove binders with main emphasis on pyrrolobenzodiazepines.

1.2.1 Distamycin derivatives

The tripyrrole peptide distamycin A is a naturally occurring antibiotic isolated from *Streptomyces distallicus* (65). Distamycin (Figure 1.11) is characterized by an oligopeptide pyrrolcarbamoyl frame ending with an amidino moiety. The compound binds reversibly to DNA with a preference for regions containing four or more AT base pairs (66,67). Distamycin A shows no antitumor properties because of its non-covalent binding with DNA through van der Waals, hydrogen bonds and electrostatic forces (68,69) (the details of the sequence specificity and mode of binding of distamycin has been described in section 1.2.4.6.3). The addition of a pyrrole ring to distamycin to form a tetrapyrrole derivative exerts very low cytotoxicity and the pentapyrrole (*N*-methyl pyrrole carboxamide) analogue is able to inhibit telomerase (70). Though devoid of anticancer activity, distamycin A is widely used as a backbone for the development of new minor groove binders- either by adding alkylating agents or by substituting pyrrole rings by pyrrazole or benzofuran rings, producing new alkylating agents with increased cytotoxicity (71-73). Tallimustine (Figure 1.11) is a benzoyl nitrogen mustard derivative of distamycin that produces monoalkylation at 3'-adenine-N3 atom in the sequence 5'TTTTGA-3' in the minor groove of DNA and inhibits the binding of ubiquitous transcription factors such as OTF-1 and NFE1 (74,75) to specific AT-rich promoter sequences and TATA-box binding proteins, leading to a decrease in the basal level of transcription (76). Tallimustine showed some activities against acute myeloid leukaemia in Phase I trials but it produced severe myelotoxicity. Another distamycin derivative- brostallicin showed potent *in vitro* cytotoxic activity against tumour cells with IC_{50} values in the low nanomolar range (77) and also increased activity was observed against melphalan resistant L1210 cells (78), which are characterized by high levels of glutathione compared with wild-type L1210 cells. The presence of glutathione is crucial for brostallicin action as it can only alkylate DNA in the presence of cellular thiols and was found to bind at the sequence 5'-AAAG (79).

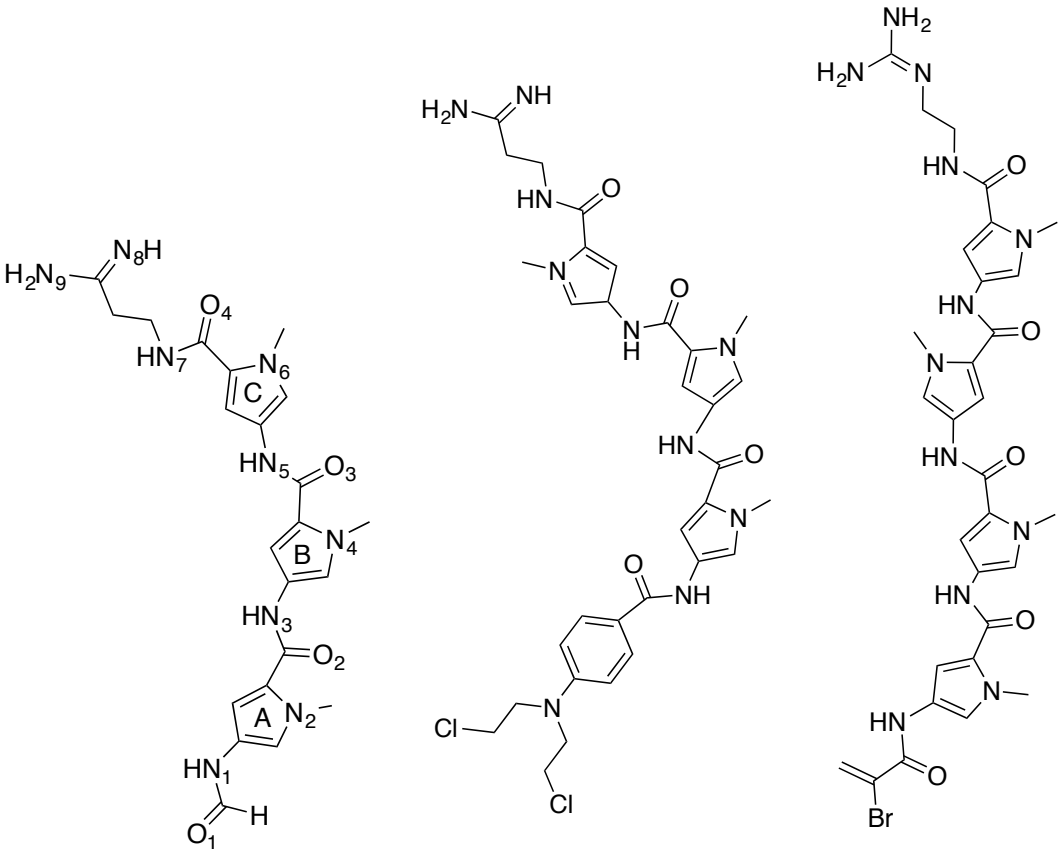
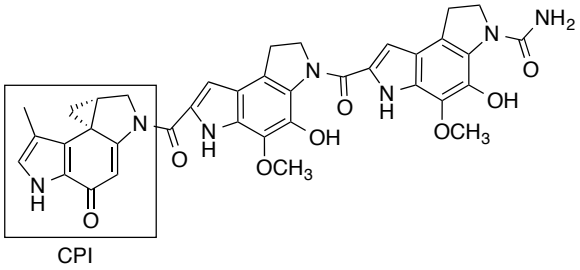


Figure 1.11:Chemical structures distamycin A, tallimustine and brostallicin

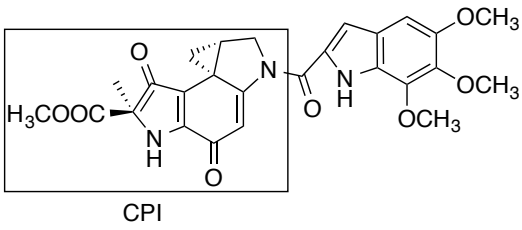
1.2.2 (+)-CC-1065 and duocarmycins

(+)-CC-1065 (80,81) and duocarmycins (82,83) are natural antibiotics isolated from *Streptomyces* bacteria, which show potent antitumor activities by selective alkylation in the minor groove of DNA. These compounds possess a cyclopropylpyrroloindolone (CPI) ring, known as the alkylation unit (Figure 1.12) and involve covalent bond formation with N3 of adenines in AT regions. (+)-CC-1065 was found to bind best at 5'-PuNTTA (Pu=Purine, N=any base) and 5'-AAAAA (with the reactive adenine underlined) (84). On the other hand, duocarmycins usually target 3 bp AT-rich sequences with the following order: 5'-AAA > 5'-TTA > 5'-TAA > 5'-ATA (84). These compounds display cytotoxicity against leukaemia L1210 cell lines in the picomolar range; C(+)-C-1065 also showed good antitumor activity *in vivo,* while duocarmycins had weak activity. However both compounds produced bone marrow toxicity, which did not permit their use in the clinic. Several derivatives of (+)-CC-1065 have entered clinical trials, among them, adozelesin, carzelesin and bizelesin are more notable (72). Adozelesin fits snugly into the minor groove of B-form DNA with a sequence preference of 5'-Pu,Py/Pu,TTA* (asterisk indicates site of binding); the compound forms covalent

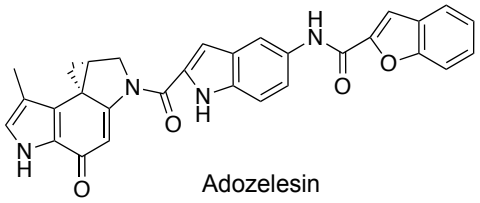
bond with N3 position of adenine spanning 5 base pairs region in the minor groove. NMR study of adozelesin with 5'-d(CGATTAATCG)$_2$ showed two types of adducts; both are formed from overlapping of aromatic benzofuran C rings of two adozelesin molecules but differ in hydrogen bonding pattern (85). In the first adduct, binding of adozelesin molecules with DNA retains Watson-Crick base pairing with minor distortion of the duplex; both molecules sit in the minor groove with the adozelesin benzofuran C rings stacked edge-on to the groove and stabilized by π-stacking between the rings. In case of second adduct, the two central A.T base pairs are in Hoogsteen base-paired position. Both types of duplex adducts were present in equivalent concentrations indicating approximately equal stability. Bizelesin, consists of two open-ring homologs of the (+)-CC-1065 CPI subunit connected by a rigid linker moiety, forms interstrand cross-links through N3 of adenine spaced six base pairs apart (inclusive of adenines) while occupying the intervening minor groove (86). NMR study of bizelesin with d(CGTAATTACG)$_2$ has been reported to show similar phenomenon like adozelesin. Bizelesin also showed two adducts where the first adduct contained an open base-pair arrangement and the second a Hoogsteen base-paired adduct (87). In the duocarmycin family only one derivative, KW-2189 has been advanced for clinical trials.



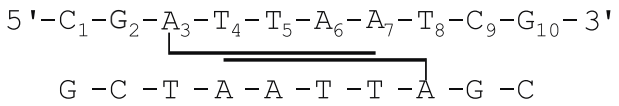
(+)-CC-1065

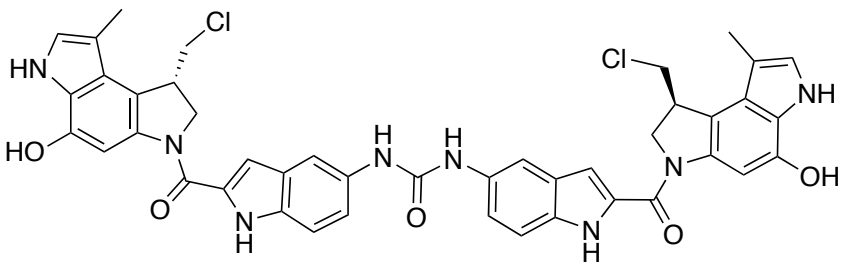


Duocarmycin A



Adozelesin





Bizelesin

Figure 1.12: Structure of (+)-CC-1065, duocarmycin A, adozelesin and its adduct (85) and bizelesin

1.2.3 Ecteinascidines

Ecteinascidins, which are tetrahydroisoquinoline alkaloids, were isolated from the Caribbean marine tunicate *Ecteinascidia turbinata* and also from the sea slug *Jorunna funebris*. One compound of this group, Ecteinascidin 743 showed promising antitumor effects in *in vitro* and in xenograft models. DNA alkylation by ecteinascidin 743 results from the generation of an iminium intermediate from N21 of the compound (Figure 1.13) that subsequently binds with N2 of guanine (88). The compound inhibited the binding of specific transcription factors such as SRF/TCF and NF-Y (89) that caused selective alteration of the transcription of genes with CCAAT-containing promoters (90,91). The compound entered in Phase I and II clinical trials and showed activities against a variety of soft tissue sarcomas (92,93).

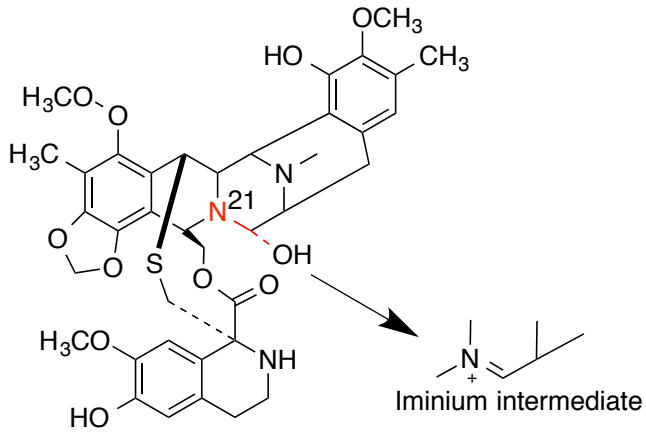


Figure 1.13: Structure of ecteinascidin 743 and generation of iminium intermediate for alkylation

1.2.4 Pyrrolobenzodiazepines (PBD)

Pyrrolobenzodiazepines (the main topic of this thesis) are sequence selective DNA minor groove binders with potent anticancer activity. The compounds exert their effects by binding covalently to the 2-amino group of guanine. Naturally produced, PBD was discovered in 1960s, and its chemotherapeutic potential encouraged researchers to use it as a lead compound, from which numerous PBD derivatives have been synthesized and screened as anticancer agents - such as PBD dimers, conjugates and dilactams. One of the PBD dimers- SJG136 has passed Phase I clinical trial and is now in Phase II clinical trial against ovarian cancer (94,95).

1.2.4.1 *Discovery*

The pyrrolobenzodiazepine group of antibiotics was discovered in the 1960s from cultures of *Streptomyces* (96) and more recently has been isolated from *Micrococcus* species (97). The first PBD to be discovered was anthramycin and since then eighteen other PBDs have been isolated. The PBDs are tricyclic compounds having different groups attached to their extremities. More recently, ten new pentacyclic PBDs named circumdatins have been isolated from *Aspergillus* fungi (98-102). The tricyclic PBDs have received considerable attention as a result of their potent activity against cancer cells. This is caused by their covalent binding to DNA. The PBDs have therefore been used as a scaffold to synthesis new anticancer molecules with enhanced DNA binding and greater sequence selectivity.

1.2.4.2 *Composition*

PBDs are composed of tricyclic systems (Figure 1.14) constituted by an anthranilate A-ring, a 1-4-diazepine-5-one B-ring and a pyrrolidine C-ring.

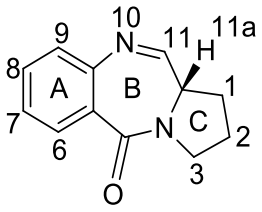


Figure 1.14: Basic structure of pyrrolobenzodiazepine (PBD)

1.2.4.3 *Epimerization*

The PBDs isolated from natural sources or produced synthetically or semi-synthetically exist as a mixture of three interconvertible forms- imine, carbinolamine and carbinolamine ether at N10-C11 position. The predominance of the different forms depends on various factors, including the substituents in the A and C rings, the method of synthesis, and the solvents and conditions used for final purification step. In aqueous solution, the imine form is in equilibrium with the carbinolamine form and is stabilized by the presence of electron-withdrawing groups at C8 (96,103,104); whereas in protic solvents such as in alcohol, the carbinolamine ether and carbinolamine forms are in equilibrium (105,106) (Figure 1.15)

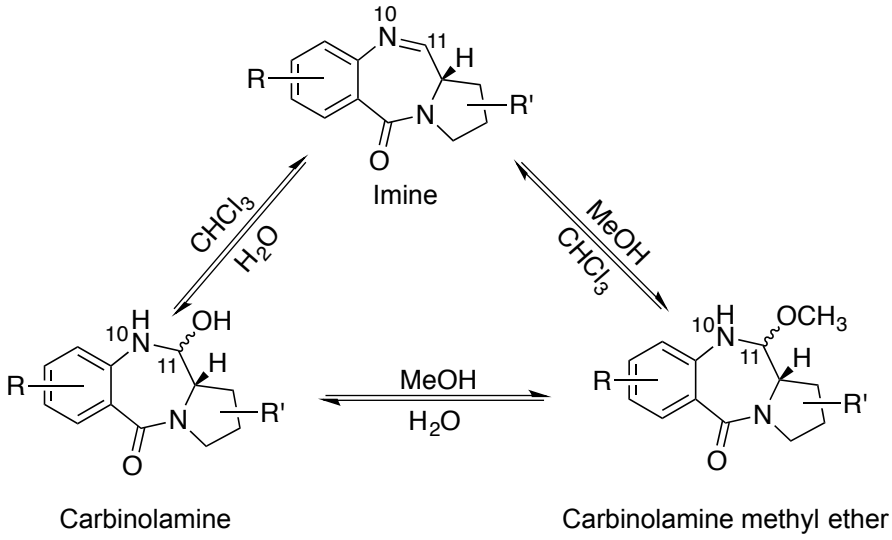


Figure 1.15: Epimerization of PBDs (107)

1.2.4.4 *Mechanism of action of PBDs*

All natural and synthetic pyrrolobenzodiazepines exert their effects by covalent linking of their N10-C11 position with the 2-amino group of guanine, which is located in the DNA minor groove. PBDs are highly selective for covalent bond formation in minor groove in either duplex or hairpin DNA and do not bind to single strand DNA regardless of its GC content (108); also not with RNA because of lack of minor groove environment (109). It has also been reported that PBDs do not interact with G-quadruplexes as these structures do not possess a B-form DNA structure or a minor groove environment (110). The covalent binding of PBDs to DNA is presumed to occur in two steps (111). In step one, molecular recognition of a favoured low-energy binding site occurs by reversible non-covalent association of the drug in the minor groove (112); the recognition is driven by hydrogen-bonding, van der Waals and electrostatic contacts. If this association is weak, the drug dissociates and then reassociates at another site until the PBD finds a suitable low energy triplet with the C2-NH_2 of a central guanine available to bind with PBD C11 position. The second step then occurs with covalent bond formation between guanine and PBD, ultimately locking the molecule into position.

Step 1: Formation of noncovalent complex formation

$$\text{DNA} + \text{PBD} \rightleftharpoons (\text{DNA}\bullet\text{PBD})_{\text{noncov}}$$

Step 2: Formation of covalent linkage

$$(\text{DNA}\bullet\text{PBD})_{\text{noncov}} \longrightarrow (\text{DNA - PBD})_{\text{cov}}$$

For step 2, three mechanisms have been proposed for covalent bond formation-(Figure 1.16):

a) Nucleophilic attack of the protonated carbinolamine or methyl ether form of PBD by the biological nucleophile *i.e*. NH_2 group of guanine (113).
b) Formation of a Schiff's base between the acyclic amino aldehyde of PBD with N2 of guanine, followed by intramolecular cyclization via attack of the aromatic amino moiety (114).
c) Direct attack of a biological nucleophile on the imine moiety of PBD (115).

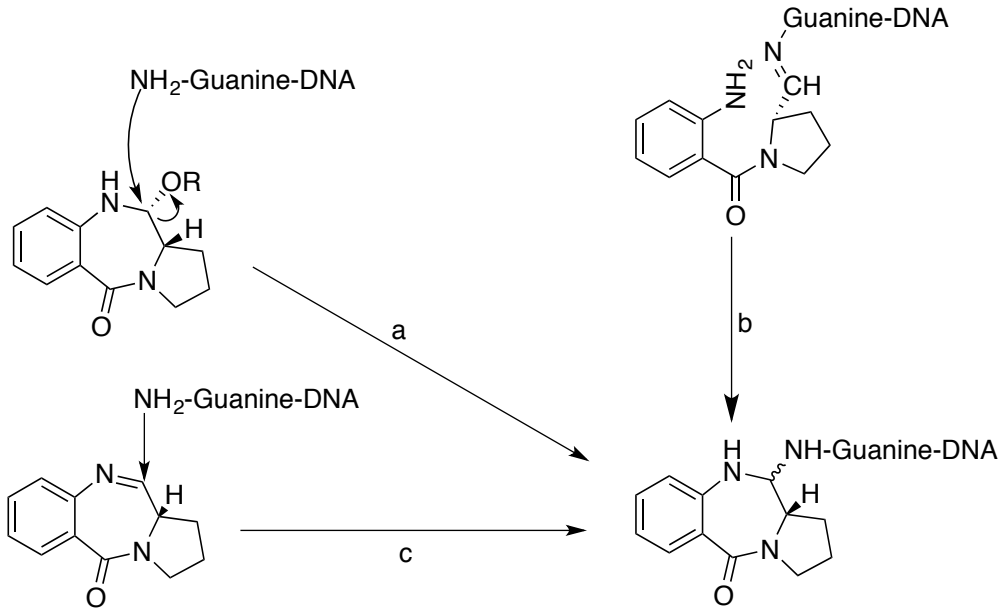


Figure 1.16: Proposed reaction mechanism of action of PBDs (116)

1.2.4.5 *Structure activity relationships (SAR)*

Knowledge of the structure activity relationships of the PBDs should enable the design of molecules with improved or altered DNA binding characteristics. Among the three rings of the PBD system, the anthranilate ring and the pyrrolidine ring have the greatest importance on overall activity of the molecules as antitumor agents.

Anthranilate ring

The nature and position of substituents in the A ring affect the potency of these compounds. Substituting the anthranilate ring with a heterocyclic system such as pyridine, pyrazine, pyrimidine or pyrrazole reduces the binding efficiency (117). The *S*-isomer at the C11a position is active whereas the *R*-isomer not (104). The presence of electron-donating groups, such as a methoxy group at C7 and C8, increases the activity of PBDs by shifting the equilibrium from the carbinolamine to the imine form (117).

Pyrrolidine ring

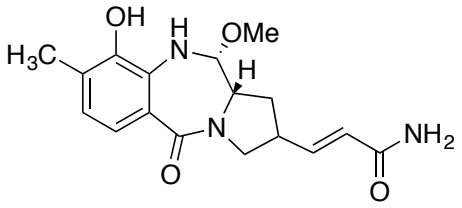
The binding of PBDs within the minor groove of DNA requires flattening of the ring, which is usually achieved by C2 *endo* or *exo* unsaturation (117). Full saturation of the ring greatly reduces DNA binding, whereas complete unsaturation impedes any DNA binding (118).

1.2.4.6 *Classification of PBDs*

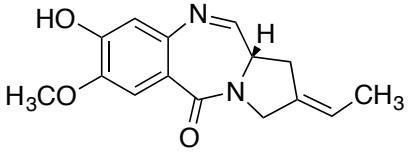
Pyrrolobenzodiazepines can be classified according to their synthesis and chemical structure. Here we describe natural PBDs and synthetic PBDs that are again categorized based on their chemical structure.

1.2.4.6.1 Natural PBD monomers

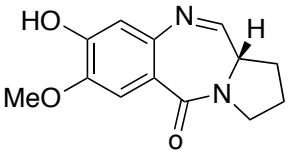
Natural PBD monomers exhibit their biological effects by binding selectively with the 5´-purine-G-purine-3' DNA sequence by forming a covalent bond with the amino group at 2-position of the guanine (103,119). It was observed from footprinting studies with four natural PBDs- sibiromycin, tomaymycin, anthramycin and neothramycin and one synthetic PBD named DC81 that the adduct spanned 3 base pairs with the preferred order of 5'-Pu-G-Pu-3' > 5'-Pu-G-Py-3' or 5'-Py-G-Pu-3' > 5'-Py-G-Py-3' (103,120). These five PBDs were also assessed by a quantitative *in vitro* transcription assay, examining the sequence specific inhibition of bacteriophage T7 RNA polymerase demonstrating 5'AGA as the most preferred binding site. This binding site was then designed with six different flanking sequences. It was observed that PBDs exerted their greatest effect with the sequence AC<u>AGA</u>AA-5' (121). This flanking sequence is believed to present the most favoured sequence energetically for adduct formation with PBDs where the adducts formed by other flanking sequences can easily be read through by polymerase (121). The (*S*)-configuration at C11a results in a right-handed helical conformation and enables the PBD molecules to fit perfectly in the minor groove of DNA. The potency of natural PBDs increases with an *endo-exo* unsaturation at the C2 position. Sibiromycin and sibanomicin are glycosylated at the C7 position of the A-ring. On the other hand, the ring C is fully saturated in neothramycin and DC-81, unsaturated at C2-C3 in anthramycin, or exocyclically unsaturated at C2 as found in tomaymycin (Figure 1.17).



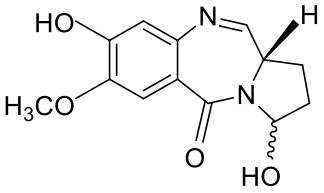
Anthramycin



Tomaymycin



DC81



Neothramycin

Figure 1.17: Structure of some natural PBD monomers

Numerous experiments have been conducted to define the nature and characteristics of the PBD monomer-DNA adduct. Carbon-13 and proton NMR experiments have demonstrated that the anthramycin C11 position is the reactive site (122). ^{1}H NMR studies were performed on an anthramycin modified deoxyribonucleotide, d(AT^{am}GCAT).d(ATGCAT) in which anthramycin was covalently attached to the exocyclic amino group of one (denoted as ^{am}G) of the two guanines in the duplex; the other guanine was found to be sterically blocked from reaction with anthramycin demonstrating only 1:1 anthramycin-duplex adduct (123). The experiment demonstrated that the conformation of B-DNA had not been significantly perturbed by adduct formation. Two key factors determine the basic structure of the anthramycin adduct - first, stereochemistry of the anthramycin C11 and secondly, the orientation of anthramycin side chain at the C2 position. The C11 position of anthramycin can be attacked by the exocyclic amino group of guanine from either face of the molecule, leading to potential formation of *R*- and *S*-isomers of the adduct. NOESY and TOCSY 2D NMR spectral data showed the existence of the 11(*S*) stereoisomer of anthramycin (124). The single bond between anthramycin C2 and C12 allows for possible free rotation of the side chain. Molecular mechanics calculations indicated that the orientation of the anthramycin C12-C13 side chain toward the 5'-end of the modified duplex was favored by 5 kcal/mol (125) to 9.3 kcal/mol (126), as compared to opposite orientation in which the side chain pointed toward the 3'-end of the modified duplex, revealing that the 3'-orientation for anthramycin binding to a dG.dC duplex is extremely disfavored (127). Crystal structure analysis had been conducted on anthramycin, covalently bound to a ten base pair DNA double helix of sequence CCAACGTTGC

(128). From this study, it was found that one anthramycin molecule bound within the minor groove at each end of the double helix, with its six membered ring directed towards the end of the helix and its acrylamide tail pointing backward toward the middle of the helix (Figure 1.18). The two drug molecules and the two ends of the helix are related by a crystallographic symmetry axis. Here, the primary attachment of the drug to DNA occurs by a covalent bond between the C11 atom of the anthramycin seven-membered ring and the N2 of guanine 9. The complex is stabilized further by hydrogen bonds. The O2 atom of cytosine 2 accepts a bond from both anthramycin atoms N10 (3.4 Å) and C9-OH (3.1 Å). Longer interactions extend from anthramycin N10-H to the N3 of guanine 20 (3.7 Å) and from anthramycin C9-OH to the N2 of guanine (3.9 Å). At the other end of the drug molecule, the amide nitrogen of the acrylamide tail donates an H-bond to the O2 of thymine 18 (2.9 Å) and the O4' of sugar (3.3 Å) (128). This crystallographic experiment was further extended to investigate the interaction between anthramycin and a dodecanucleotide d[GCCAACGTTG*GC]$_2$ (reactive guanine denoted as G*) by molecular simulation. Molecular dynamics (MD) on both the hydroxy and anhydro (C11 position) putative reactive forms of anthramycin were carried out to obtain an insight into the formation of the initial noncovalent anthramycin-DNA complex (129). The simulations showed that both forms of anthramycin, docked in front of the reactive site, induce structural deformations on the DNA frame at its central region, while significant distortions were observed at the binding region for covalent adduct formation. It was also found that the DNA frame greatly influenced the electronic distribution of ligand and revealed the anhydro- anthramycin more favourable for the alkylation reaction.

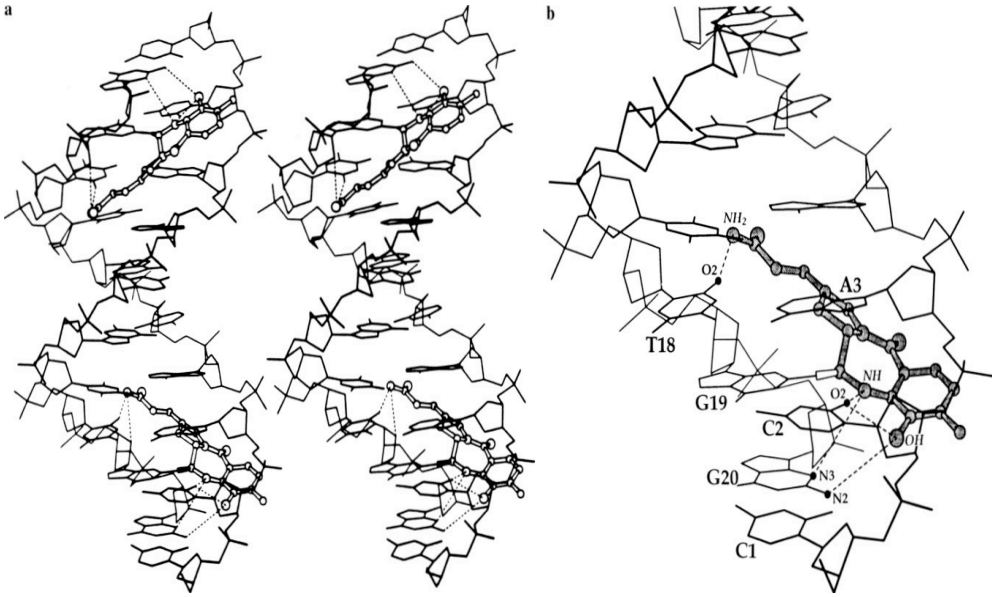


Figure 1.18: Structure of the B-DNA decamer CCAACGTTGG with two anthramycin molecules covalently bound within the minor groove to the N2 of guanine on the second base pair from each end of the helix. DNA to drug H-bonds are dotted. (a) View

into the minor groove at top and major at bottom, displaying anthramycin binding. (b) Close-up of the bottom half of (a), showing details of interactions involving base pairs C1.G20, C2.G19 and G3-C18. Three of the six distances drawn are between 2.8 and 3.5 Å; all six are below 4.0 Å. The open bond linkage is the covalent bond to guanine 19 (128).

Sibiromycin, the most cytotoxic PBD monomer, differs from other naturally occurring PBDs by the presence of a glycosyl residue (sibirosamine sugar) at the C7 position. The enhanced cytotoxicity of sibiromycin is assumed to be related with the C7 sugar moiety. Advanced molecular dynamics simulations showed that sibiromycin accommodated snugly in the minor groove with either *R* or *S* stereochemistry at C11 and the bulky C7- sibirosamine sugar protruded orthogonally from the minor groove (130); this protrusion of the sugar moiety enables the compound to inhibit the interaction of transcription factors with DNA and ultimately produces enhanced cytotoxicity (131). To gain an insight into optimized binding of sibiromycin with DNA sequences, this simulation experiment was conducted on a set of eight duplex sequences each containing a triplet binding site (GGG, GGA, AGG, AGA, AGT, AGC, TGG and TGA) (130). Free energy of binding calculations indicated that a guanine base was favoured on the 3'-side of the reacting triplet (5'-Pu-G-G-3') due to hydrogen bonding contacts. Furthermore, molecular simulation on the duplex containing AGG showed that the C9-OH of the PBD formed an H-bond with the exocyclic C2-NH_2 group of the adjacent G7 for both C11(*R*) and C11(*S*) adducts. A second H-bond was also observed between C9-OH of the PBD and the ring N3 nitrogen of G7 (Figure 1.19). Though both *R* and *S* stereomers were observed, H-bonding of C11(*R*) was found to be more transient than C11(*S*) suggesting the preference for the S configuration, as with other PBDs.

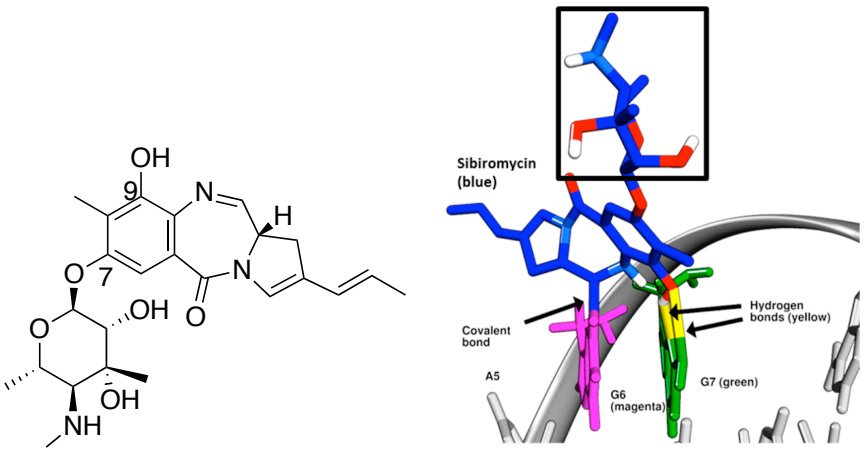


Figure 1.19: Structure of sibiromycin (left). On the right panel shows a low energy snapshot from a molecular dynamics simulation of sibiromycin covalently bound in the C11(*S*) A-ring-3' configuration to G6 (magenta) of 5'TATAAGGTATA-3' showing the ligand perfectly accommodated in the DNA minor groove. Two H-bonds (yellow) were

observed between C9-OH and N3 of G6, and the C9-O and exocyclic C2-NH_2 of guanine G7 (green) adjacent to the covalently modified guanine (130).

1.2.4.6.2 Synthetic PBDs - Monomers

PBD monomers have been synthesized chemically by introducing different moieties at C2 with an *endo*-*exo* unsaturation as in the natural PBDs. The substituents varying from single aryl rings to fused ring systems and also styryl moieties have involved more than eighty different synthetic monomers. Among the synthetic monomers, C2-aryl C2/C3–unsaturated PBDs (132) (Figure 1.20) seem very important due to their significant cytotoxicity against various cancer cell lines at the sub-nanomolar level.

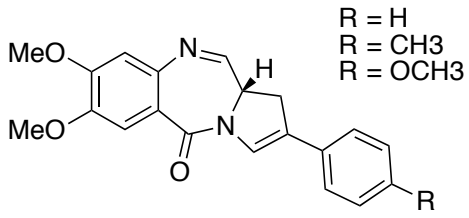


Figure 1.20: C2-aryl/C2-C3 unsaturated PBDs

The compounds were found to be active at nanomolar concentrations (LC_{50}) against six melanoma, two non-small lung cell, one CNS and two colon cancer lines in the standard NCI 60 cell line screening. The compounds were then tested against six more specific cell lines comprising two melanoma and one each of ovarian, colon and breast cell lines, where they exhibited nanomolar potency at the GI_{50} level against melanoma and picomolar potency against an ovarian cell line (132).

1.2.4.6.3 PBD monomers with heterocyclic moieties

The prominent activity of PBDs as antitumor agents encouraged researchers to synthesize analogues of PBDs with conjugates to increase their sequence selectivity and cytotoxic potency. Here, the sequence selectivity of distamycin will be covered first and then PBD-conjugates will be discussed.

Determination of sequence specificity and mode of binding of distamycin

Different footprinting techniques were used to identify the preferred binding sites for distamycin. DNase I footprinting showed distamycin to protect four regions, each centred around an AT rich site (133). However, it is not easy to identify accurately the extent of binding by distamycin because these bound regions occur at sites where DNase I cuts the native DNA (without ligand) with low efficiency. Using the same DNA fragment, Portugal and Waring (134) performed hydroxyl radical footprinting and found the binding at AT-rich sites by distamycin. Binding of distamycin at AT-rich sites has

also been reported by another footprinting technique. With methidiumpropyl (MPE)- Fe (II) footprinting, it was also found that distamycin bound preferentially to AT-rich regions with minimal protected sites of four base pairs. With different arrangement of adenines and thymines, distamycin showed highest affinity for AATT and then ATTA, TATA with least affinity for ATAT (135).

The binding specificity of distamycin for AT-rich sites in DNA can be explained in two ways- first, the narrow minor groove of AT base pairs promotes tight fitting of the distamycin molecule. The three pyrrole rings (A, B and C rings) in the distamycin have a slight twist, so that the molecule is not planar. There is an angle of 15° between the plane of pyrrole rings A and B and 11° between B and C (136). The planar amides connect the pyrrole rings such a way that the molecule conforms to a crescent shape that is complementary to the natural curvature of minor groove of B-DNA. Secondly, the presence of the amino group at C2 of guanine sterically hinders the pyrroloamide chain from docking fully to the floor of minor groove of GC rich regions. To elucidate the mode of binding of distamycin, the compound was subjected to interact with a wide number of duplexes; these duplexes differ in number of adenines and thymines in the middle of the sequences such as AATT, AAATT, AAATTT flanked by GC rich sequences at both 5' and 3'-end. From NMR and crystallographic studies, it has been demonstrated that distamycin can interact with such oligonucleotide duplexes in more than one manner forming either 1:1 complex or 2:1 (stoichiometry of distamycin to DNA) complex (66,136-141).

High resolution ^{1}H NMR was used to study the complex between distamycin and the self complementary DNA dodecamer d(CGCGAATTCGCG)$_2$ (141). Each of distamycin amides gave an NOEs to an adenine C2H resonance, confirming that the compound bound in the minor groove of the B-form DNA duplex. Pelton and Wemmer (66) used the same dodecamer duplex and subjected to 2D NMR; the experiment also demonstrated positioning of distamycin in the AATT minor groove, with each of three pyrrole rings approximately parallel to the walls of the groove; but the assignment of hydrogen bonding differed from ^{1}NMR (Figure 1.21).

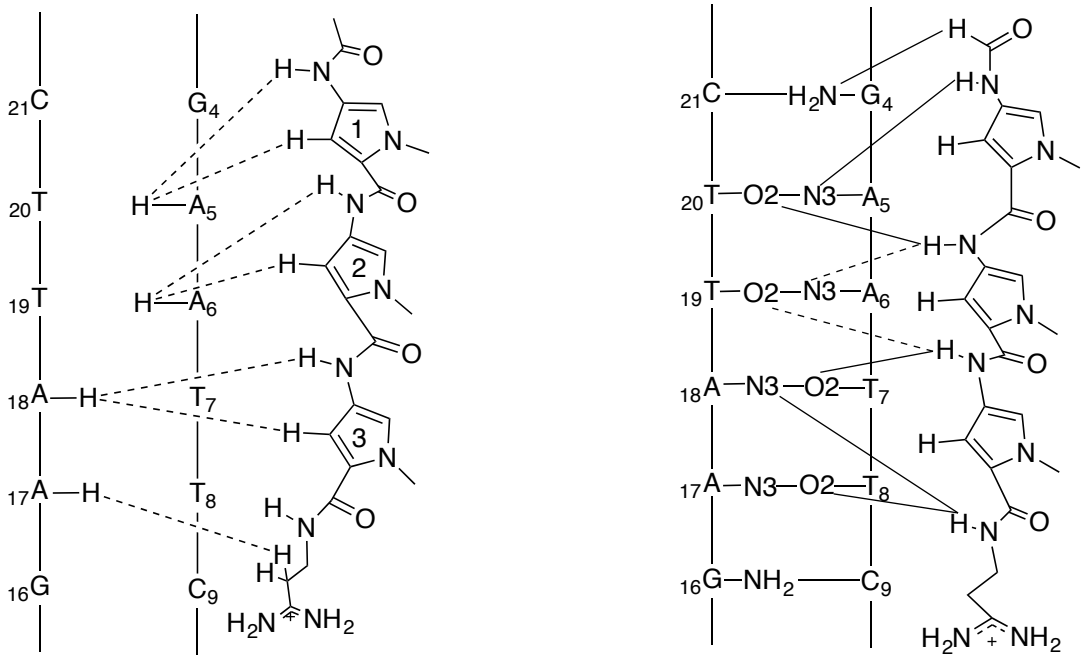
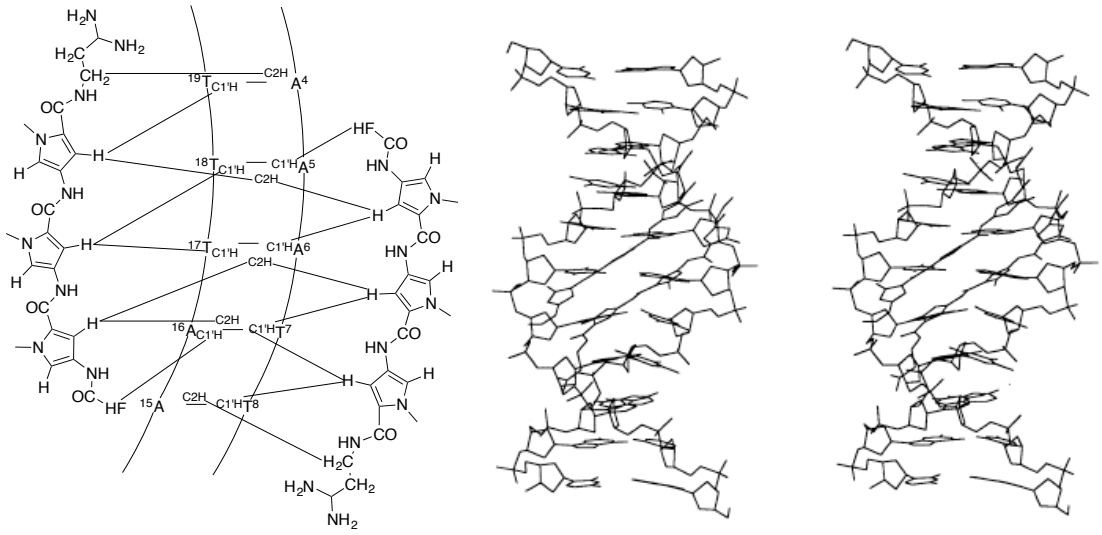


Figure 1.21: (A) Diagrammatic representation of the distamycin binding site in the central AATT region of d(CGCGAATTCGCG)$_2$ as determined by ^{1}H NMR. The observed NOE contacts from assigned C2H resonances to distamycin amide protons (dashed lines) and to nonexchangeable distamycin protons (dotted lines) are indicated (141). (B) Hydrogen bonding scheme from 2D NMR. Solid lines denote hydrogen bonds in the highly constrained structure (H_D) and dashed lines denote the additional hydrogen bonds that formed when NMR constraints were relaxed (W_D) (66).

Pelton and Wemmer (137) further reported the structure of distamycin complexed with d(CGCAAATTGGC).d(GCCAATTTGCG) using 2D NMR and molecular modelling, in an attempt to examine which binding sites were occupied in solution when a sequence with more than the minimal four AT base pairs was available. Here, two types of drug-NOEs were observed. In the first type, pyrrole rings span the A5-A6-T7 sequence. In second type, one formyl proton near the 5'-side of the A-rich strand while another formyl proton near the 5'-side of the T-rich strand were reported. These data indicated that two distamycin molecules were positioned in the minor-groove in the central portion of this dodecamer in an antiparallel orientation, with one distamycin molecule in close contact with the sequence 5'-AATT-3' of the A-rich strand and the second distamycin with 5'-ATTT-3' of the T-rich strand, demonstrating formation of 2:1 complex. Orientation in an antiparallel order is favoured because it reduces charge-charge repulsion between the positively charged propylamidinium groups (Figure 1.22). To identify the exact positioning of pyrrole and amide linkages, molecular modelling was done and found that the two distamycin molecules were staggered with respect to one another so that the pyrrole rings of one molecule stack with the amide linkages of the opposing drug on one side, and with DNA sugar O1' atoms on the other side.



(A) (B) (C)

Figure 1.22: (A) Schematic of distamycin binding sites on d(CGCAAATTGGC). Intermolecular contacts are presented as lines. (B) and (C) Stereo drawing of the distamycin-DNA complex obtained from NOESY of 2D NMR (137).

X-ray crystallography (136) with d(CGCAAATTTGCG) showed that distamycin lay in the narrow minor groove covering 5 of the 6 AT base pairs producing a 1:1 complex. van der Waals interactions between the walls and floor of the minor groove and distamycin are evident, with the oxygen atom of the amide linkages pointing away from the minor groove; solvent water molecules are also found within hydrogen bonding distances. In contrast analysis of this dodecamer by 2D NMR demonstrated the predominance of an antiparallel 2:1 complex along with formation of 1:1 complex (142) in the central AT rich regions. NMR spectra obtained at intermediate points in the titration of this AAATTT duplex with distamycin indicated the presence of both one-drug and two-drug binding modes. To explain these results, Pelton and Wemmer (142) concluded that in the one-drug binding mode, distamycin either bound to only one site (as in the crystal structure) or the drug was in fast exchange between several binding sites. On the other hand, in the two-drug binding mode, rapid exchange occurs between the two binding sites. At low drug concentrations the one-drug mode was significantly populated whereas at high drug concentrations, 2:1 complex predominated that facilitated assignment and structural characterization.

The antiparallel 2:1 complex formation of distamycin was further examined in the octamer duplex d(ICICICIC) (140). The minor groove of ICICICIC is significantly wider (average 7.7(±0.2 Å) than B-DNA (5.7 Å) (143) while the major groove widths are similar (average 11.3(±0.1 Å) and 11.7 Å respectively) (140). More importantly inosine is devoid of amino group at the C2 position of guanine. In this wider groove, two

distamycin molecules were sandwiched in an antiparallel orientation with the pyrrole rings of one molecule stacked on the amide groups of the other. By forming a 2:1 complex, individual distamycin molecules interacted almost exclusively with the DNA strand in closest proximity while the second molecule shielded contact with complementary strand. Each distamycin bound to the lower 5 residues of each strand; the amide N1 atom of the distamycin made bifurcated hydrogen bonds to the O2 of cytosine-4 and to the O4' sugar oxygen of the same residue while the amide N3 formed only one hydrogen bond to the N3 base atom of inosine-5. The amide of N5 and N7 also made bifurcated hydrogen bonds to the base atoms O2 of cytosine-6 and N2 of inosines-7 respectively and to the O4' sugar oxygen atoms of the following residues (Figure 1.23).

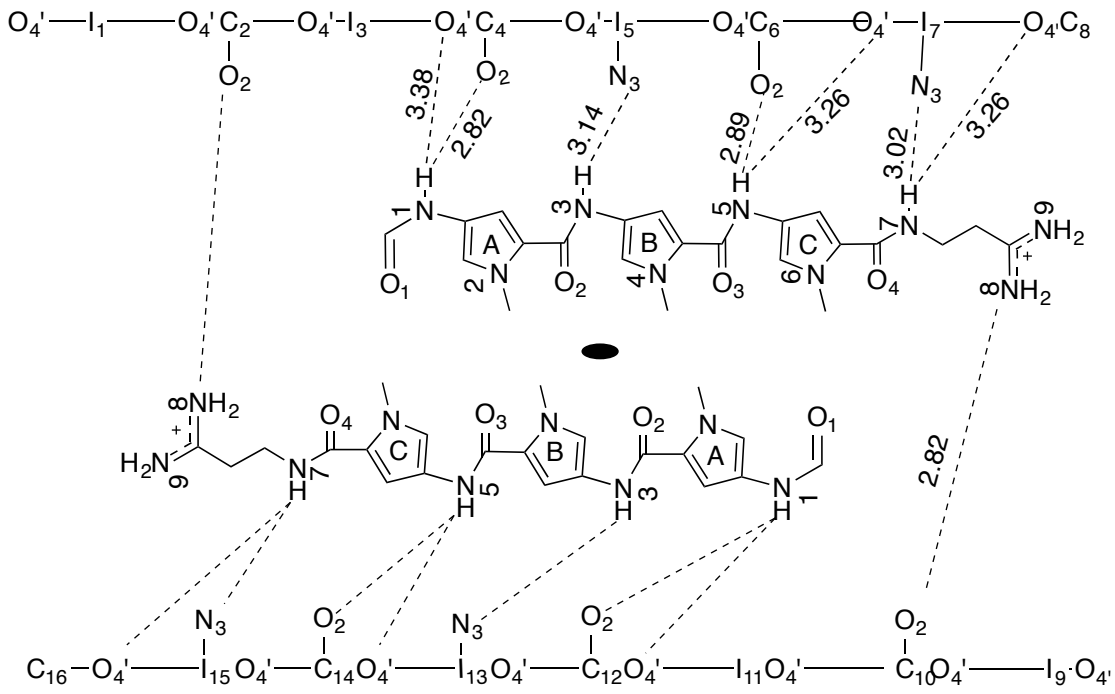
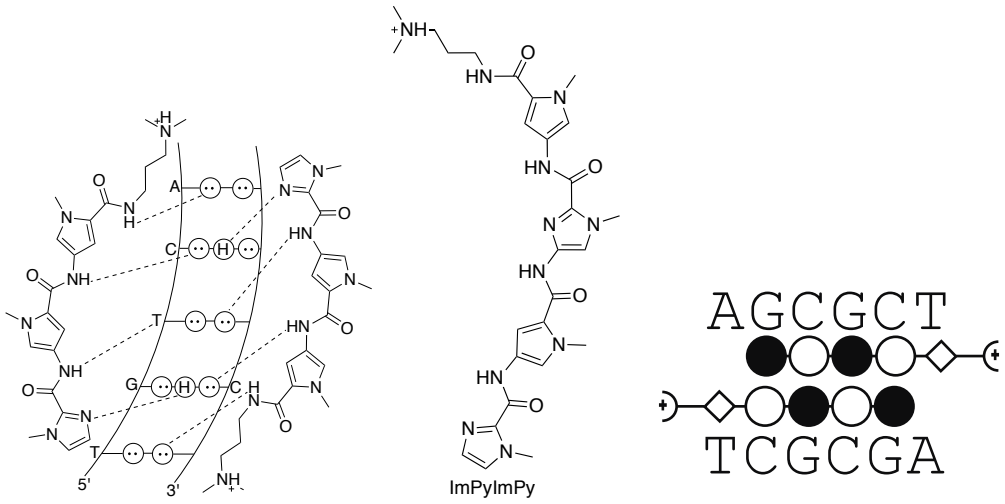


Figure 1.23: Schematic diagram showing hydrogen bonds between distamycin and DNA (140).

Polyamides formed by pyrrole, imidazole and thiazole

The mode of binding distamycin to AT rich sequences has been utilized by researchers to design and synthesize derivatives with increasing numbers of *N*-methylpyrrolecarboxamides that can bind longer sequences of AT-rich double helical DNA. To achieve this, numerous compounds have been synthesized containing 3 to 10 carboxamide NH's, mostly by Peter Dervan and his group from Caltech (144-150). For recognition of mixed sequences *i.e.* for G.C selectivity, Dervan's group synthesized Imidazole-Pyrrole-Pyrrole complex and anticipated to obtain 1:1 complex with altered

specificity due to the presence of *N*-methylimidazole (151). Surprisingly, the compound bound at 5'TGTCA, which they demonstrated was due to the formation of a 2:1 complex in an antiparallel orientation. Another four-ring tripeptide containing alternating *N*-methylimidazole and *N*-methylpyrrole carboxamides (ImPyImPy) showed binding specifically at 5'(A,T)GCGC(A,T)-3' in the minor groove where two molecules bound side-by-side in an antiparallel orientation in the minor groove (152). Each ligand molecule bound with one of the DNA strands in the minor groove. The N3 of imidazole of each ligand forms hydrogen bonds with NH_2 group at C2 position of guanine. Thus the side-by-side combination of an imidazole ring on one ligand and a pyrrole ring on the second ligand recognizes G.C, where pyrrole-imidazole pair targets C.G base pairs; the pyrrole-pyrrole combination binds either at A.T or T.A base pairs (151,153) (Figure 1.24). Differentiation between T.A from A.T was finally achieved by substitution of C3-H by C3-OH (hydroxypyrrole) within a Py/Py pair creating Hp/Py that could identify T.A whereas Py/Hp for A.T (154).



(A) (B)

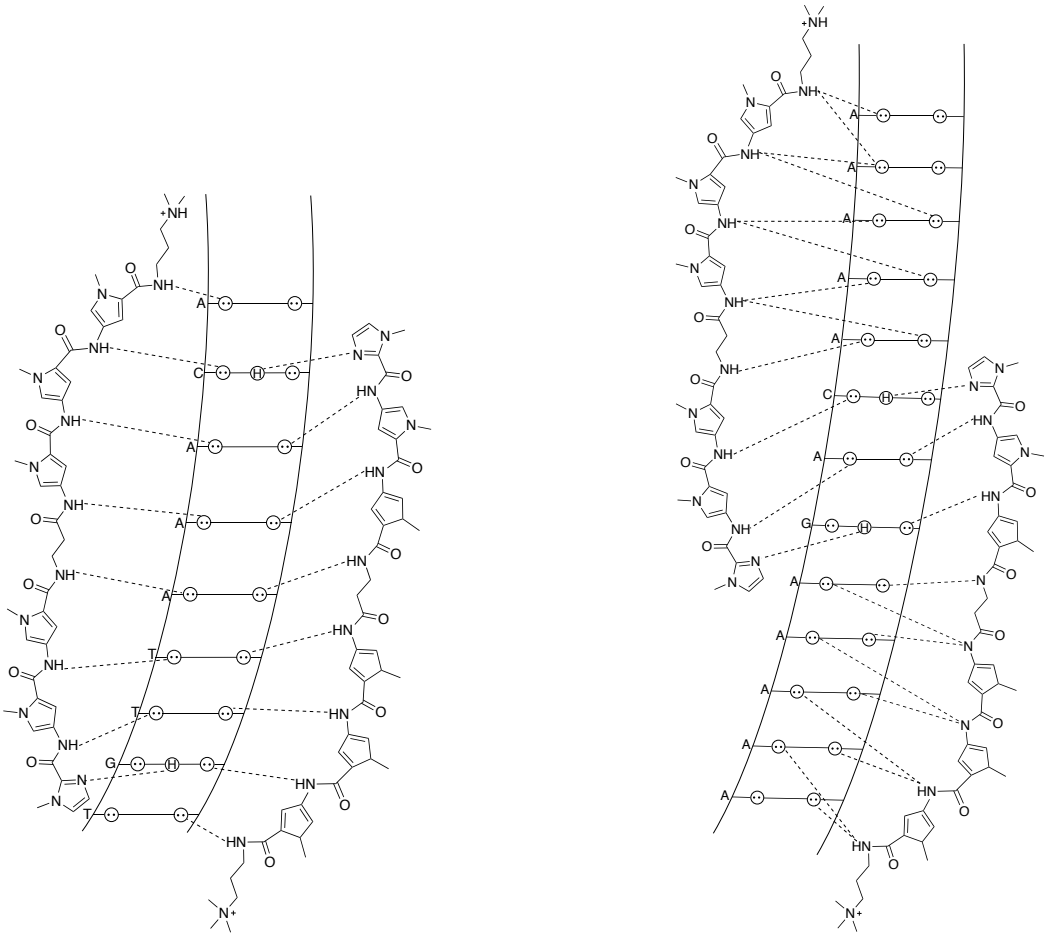
Pair	G.C	C.G	T.A	A.T
Im/Py	+	-	-	-
Py/Im	-	+	-	-
Hp/Py	-	-	+	-
Py/Hp	-	-	-	+

Favoured (+), disfavoured (-)

Figure 1.24: (A) Recognition of base pairs by Im/Py polyamides with binding models of ImPyPy at 5'TGTCA (151,155). (B) Sequence recognition by ImPyImPy carboxamides (152). The table shows the pairing code for minor groove recognition.

The formation of dimeric complexes by polyamides has become very useful for targeting specific sequences, since Py/Py binds at A.T or T.A and Py/Im and Im/Py target C.G and G.C respectively. For recognition of longer sequences, the length of pyrrole-imidazole polyamide can be extended, but the question arises as to what will be the maximum length of polyamide that can identify target site accurately. In an experiment with a polyamide series of Im-$(Py)_n$-Dp (n = 2 to 7), DNase I footprinting showed that polyamides with three to seven rings could bind to 5 to 9 base pairs respectively with good specificity. The binding affinity (equilibrium association constant) to target sites increased with increasing number of rings up to five rings, showed no increase for six and seven rings and dramatically decreased for a polyamide with eight rings (156). Polyamides with only three rings showed least tendency to bind at mismatch site (one base pair mismatch in the middle of the target site); four and five rings had a higher tendency to bind to mismatches, while longer polyamides (specially eight rings) had very high affinity for mismatches within their target sites. Longer polyamides lack isohelicity with the DNA minor groove resulting in poor correspondence between the binding surfaces of the polyamides and the DNA sites, whereas the curvature of polyamides with three and four rings match very closely with that of DNA. Thus, for specific recognition of 10 base pairs or more, it is necessary to replace one of the central pyrroles or imidazoles with a flexible linker, which allows the antiparallel dimer to register at target site with good affinity and specificity. *β*-alanine has been proved a good linker achieving this, enabling recognition of longer sequences (157); *β*-alanine has a preference for binding opposite an A.T or T.A base pair relative to G.C or C.G.(158). Polyamides composed of six-rings *i.e.* two molecules of ImPyPy-*β*-PyPyPy bound opposite one another at the anticipated site 5'TGTTAAACA3' (157) (Figure 1.25A). However, this polyamide can also bind in a 'slipped' mode, in which the two ligand molecular are not fully overlapped, at the sequence 5'AAAAAGACAAAAA where the 5'AGACA3' in the middle of the sequence is bound in a 2:1 manner, as with the ImPyPy homodimer, and the PyPyPy moiety binds to the A,T flanking sequences in a 1:1 manner (Figure 1.25B). This slipped mode of binding was also observed when a eight ring polyamide ImPy-*β*-ImPy-*β*-ImPy-*β*-PyPy-*β*-Dp was found to recognise a 16 base pair sequence of 5'ATAAGCAGCTGCTTTT3' (159); in this polyamide-DNA complex, eight pairs of amide residues form a fully overlapped core which properly positions the six Im-Py pairs for recognition of six G,C base pairs and two *β/β* pairs for recognition of two A,T base pairs. To overcome the formation of slipped mode binding, a 'hairpin' motif of polyamide has been developed where amino- and carboxyl termini of the antiparallel dimers are connected by an aliphatic amino acid, thus creating a U-

shaped motif that also binds in the minor groove of DNA with very high affinity and specificity. This hairpin polyamide has been discussed in details in Chapter 6.



(A) (B)

Figure 1.25: Complexes of ImPyPy-*β*-PyPyPy with target sites (A) 5'TGTTAAACA3' (9 bp overlapped mode) and (B) 5'AAAAAGACAAAAA3' (13 bp slipped mode). Circles with dots represent lone pairs on N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dashed lines (157).

Thiazole

The thiazole unit has been introduced into these polyamides in order to control the overlap between the two molecules in the 2:1 binding mode, so as to achieve recognition of longer DNA sequences. The first example of this was with thiazotropsin A (160), which contains an isopropyl-substituted thiazole in place of the *N*-methylpyrrole. The two molecules of this compound bind side-by-side in the minor groove, but due to the presence of bulky thiazole ring, the molecules are staggered from each other and so recognise the 6 base pair sequence ACTAGT. Similarly, when *N*-methylpyrrole is replaced with *N*-methylimidazole, producing thiazotropsin B (161)

this recognises the sequence (A/T)CGCG(A/T), as illustrated in Figure 1.26. In these compounds, the thiazole of one molecule and the formyl or acetyl end of the second molecule can recognise a G.C base pair.

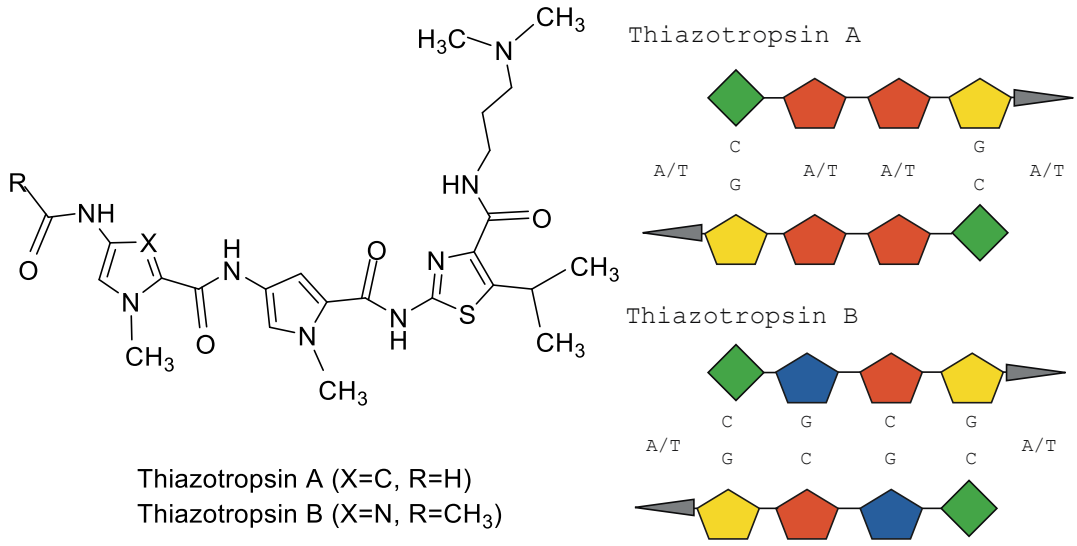


Figure 1.26: Structure of Thiazotropsin A and B and schematic representation of the binding of thiazotropsins A and B in the DNA minor groove. Green, formyl (thiazotropsin A) or acetyl (thiazotropsin B) 'head'; red, *N*-methylpyrrole; yellow, thiazole; blue, *N*-methylimidazole; grey, DMAP 'tail' (161).

Rao *et al* (162) also reported the synthesis of thiazole-containing ligands related to distamycin and netropsin, in which the location of the CH, N and S were varied around position 3 of the ring. The orientation of the sulphur atom was modified by placing it facing either towards or away from the DNA minor groove, comparing the sequence recognition properties of the different ligands and identifying the factors that affected binding, such as steric hindrance, electrostatic interactions and ligand chirality. The ligands bound to calf thymus DNA, which comparable affinities to distamycin and netropsin and two of these compounds (Figure 1.27) bound stronger than distamycin.

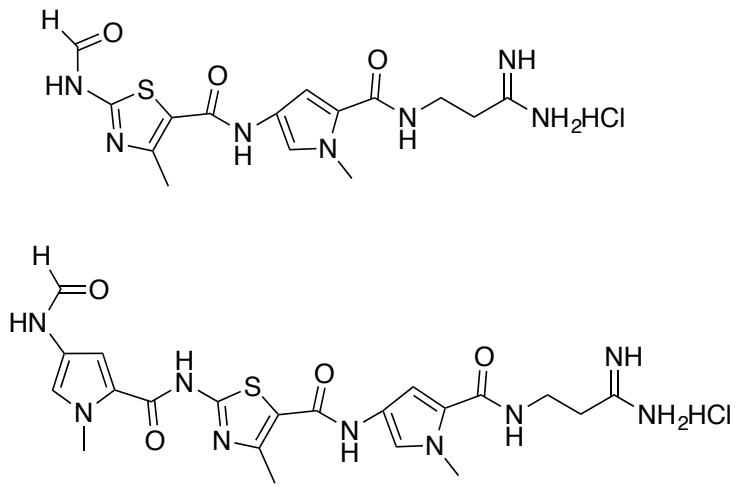
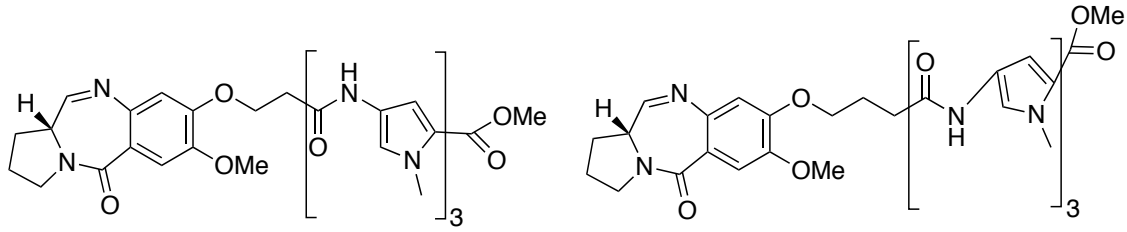


Figure 1.27: Thiazole containing oligopeptides related to netropsin and distamycin (162)

PBD conjugates with pyrrole, imidazole and thiazole

The alkylating ability of PBD at guanine in the minor groove and the recognition of A/T regions by distamycin have been combined to achieve a synergistic effect from both groups. Molecular modelling studies of PBD-dimers (two PBD molecules joined by a linker; PBD dimers will be discussed in Section 1.2.4.6.4) suggested that C8-linked PBD dimers have greater isohelicity with the minor groove of DNA compared to the C7-linked dimer (163); the same position is also found appropriate to conjugate PBD with polyamides (164-167). Another important aspect of these conjugates is the length of the linker that is required to maintain appropriate isohelicity of the conjugate; for covalent binding of PBD unit it is essential to position the imine functionality (alkylation site of PBD) suitably against C2-NH_2 of guanine while maintaining other favourable and potential sequence-selective interactions by the remainder of the molecule. In case of a dimethylene linker, the polyamide (*N*-methylpyrroles) fits snugly and isohelically within the minor groove and the PBD unit rotates partially out of the groove. However, optimal interaction of the PBD moiety with the minor groove results in the first *N*-methylpyrrole ring being 'bulged' out of the groove (167). When the linker is increased to trimethylene, a superior isohelical ligand conformation can be achieved (Figure 1.28). Binding energy calculations showed that trimethylene (-26 ± 6 kcal/mol) is a better choice than dimethylene linker (-14 ± 3 kcal/mol) (167). After fixing the trimethylene chain as a suitable linker, a series of PBD-polypyrrole (as *N*-methylpyrrole) conjugates was synthesized, varying the number of pyrroles (from 1 to 6) and examined by DNase I footprinting, thermal denaturation studies and *in vitro* cytotoxicity assays (167). All six conjugates showed binding to similar sites and followed a consensus motif- 5'XGXW_Z,

Z = 3 ± 1, W = A or T, X = any base but preferable purine. However, these varied in the length of binding site as well as the binding affinity. As expected the longer conjugates appeared to have larger binding sites. The PBD conjugate with only one pyrrole showed very low binding affinity, while conjugates with two and six had moderate affinity and three, four and five pyrroles demonstrated similar, higher affinities compared with the individual components – the PBD and polypyrrole chain. In addition, thermal denaturation studies with calf thymus DNA demonstrated significant increases in melting temperature (8.1° C to 21.2° C). The highest ΔT_m (21.2° C) was produced by PBD-Py-Py-Py; this change in melting temperature was 53-fold higher than for the PBD linker unit alone and 5-fold greater than the tripyrrole unit alone. All these PBD-conjugates produced significant cytotoxicity across the cell lines when tested against 60 cells lines of the NCI 60 panels. Moreover, Kotecha and co-workers reported that PBD conjugated with two pyrroles (PBD-Py-Py), coded as GWL-78, could inhibit the binding of NF-Y to the promoter of DNA Topo II*α* and displace NF-Y bound to several CCAAT motifs, known as inverted CCAAT boxes (ICBs), within the promoters of genes involved in cell cycle progression (166).



PBD-tripyrrole with dimethylene linker

PBD-tripyrrole with trimethylene linker

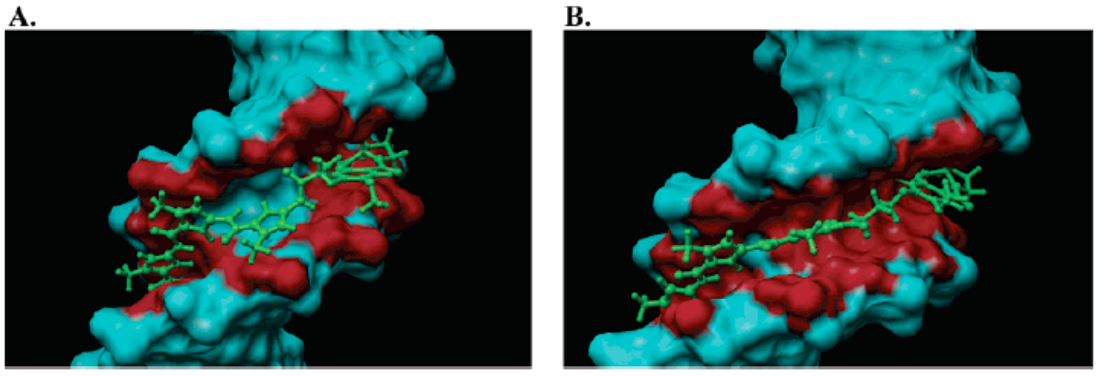
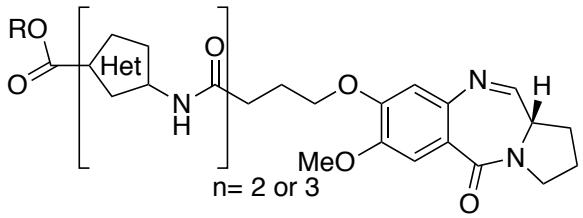


Figure 1.28: Time-averaged structure of dimethylene (A) and trimethylene (B) linked PBD-tripyrrole conjugates bound to the 17-mer DNA duplex 5'TTTTGCAAAAAGCTTCA3'. In each case, sections of the molecular surface of the DNA less than 5 Å from the ligand are coloured red to emphasize the superior fit of trimethylene-linked polyamides in the minor groove (167).

The ability of GWL-78 to recognise ICBs was capitalised to enhance the potency and differentiate between two ICB sites- ICB1 (5'CAGGG*ATTGG*CTGGT3') and ICB2 (5'CTACG*ATTGG*TTCTT3') by replacing pyrrole(s) by imidazole (Im) or thiazole (Thz)

rings. From a series of six analogues (Figure 1.29), the thiazole containing conjugate PBD-Thz-Py was found to be superior to PBD-Py-Py in terms of overall DNA binding affinity, but lacked differentiation between ICBs; DNA footprinting potencies of PBD-Py-Py and PBD-Thz-Py were 1 μM and 0.2 μM respectively (165). To achieve selectivity, PBD-trimeric conjugates were synthesized and examined against 170 bp of the human Topo II*α* promoter containing the ICB1 and ICB2 sequences; among these, PBD-Im-Py-Py exhibited 10-fold selectivity for ICB2 compared to the ICB1 site (164). These results suggest that it is possible to design PBD-conjugates with pyrrole, imidazole and thiazole that can discriminate between binding sites.



Het = Pyrrole (Py), Imidazole (Im) or Thiazole (Thz)
R = Me or Et

Figure 1.29: Structures of the C8-linked PBD conjugates with different combinations of pyrrole (Py), imidazole (Im) and thiazole (Thz) rings (164).

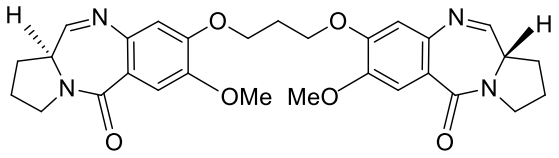
1.2.4.6.4 PBD-Dimers

Several PBD dimers have been designed in order to enhance their binding and to achieve sequence selectivity over a greater length of DNA base pairs. These dimeric molecules are designed to conserve the inherent chemical properties of monomeric PBDs for binding, with increased potential to form covalent adducts to guanine bases of specific DNA sequences. The first molecule of this class, DSB120, was synthesized in the Thurston laboratory (168) by joining two DC-81 molecules with a propyldioxy linker at the C8/C8´ position. The compound was found to be more reactive towards DNA than many monomeric PBDs. A number of other PBD dimers have been synthesized, which show potent DNA binding (169); here I will focus on a few examples of this category such as DSB120, SJG-136, ELB21 and SG2285.

DSB120

The design of DSB120 was inspired by comparison with the antitumor activity of mitomycin, cis-platinum and nitrogen mustard, which all produce intra or interstrand cross linking of DNA (170,171). Gel electrophoresis and thermal denaturation studies of DSB120 with DNA *in vitro* revealed high affinity for DNA, forming irreversible interstrand cross links (168,172). Cross-linking of DSB120 is thought to occur by recognition of the 5´-GATC motif, with additional hydrogen bonding to the spanned

adenine bases (Figure 1.30). This was confirmed with the self-complementary decanucleotide d(CICGATCICG), by using 2D-NMR NOE distance restrained simulated annealing procedures (173). Here two guanines were replaced by inosines to offer one unique bonding site for the ligand.



DSB120

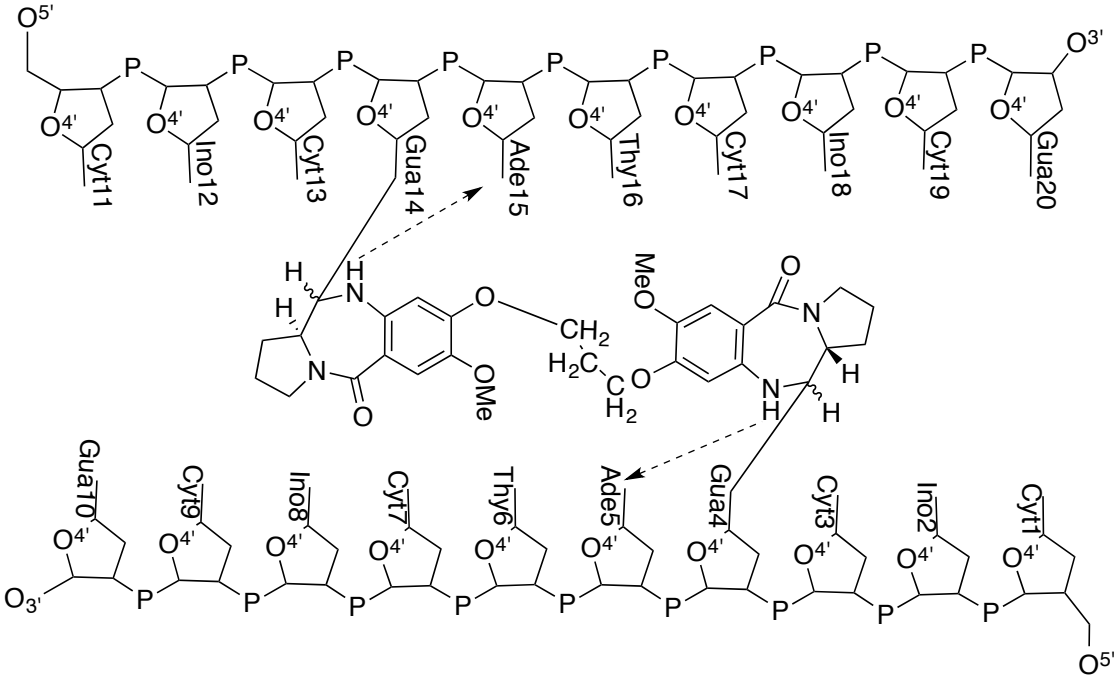
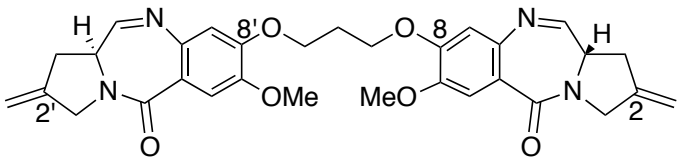


Figure 1.30: Binding of DSB120 with the sequence d(CICGATCICG) (173)

SJG-136

SJG-136 is the most important of the PBD dimers (Figure 1.31), which is now under Phase II clinical trial. The compound differs from DSB120 with two additional methylene groups attached at the C2 position. Though DSB120 has potential to cross link with specific guanine sequences its reactivity with thiol-containing molecules (174) reduces its usefulness as an antitumor agent. The presence of the C2-exo-methylene significantly increases the cytotoxicity of SJG-136 as well as decreasing its affinity for thiol groups.



SJG-136

Figure 1.31: Structure of SJG-136

The cross linking property of SJG-136 arises from its recognition of 5′-Pu-GATC-Py-3′. Like other PBDs, the compound also forms covalent bonds with guanines, along with hydrogen bonding with the adjacent adenines by N10/N10′ protons. The compound is also capable of intrastrand adduct formation with two guanines separated by 3 bases, along with hydrogen bonding with adenines on alternate strands (175). Interstrand and intrastrand crosslinking adducts have also been confirmed by high-field NMR study (176). SJG-136 formed interstrand crosslinking adduct with d(CICGATCICG)$_2$ showed deep penetration of the ligand in the minor groove and the two PBD residues tilted such that the unsaturated methylenes were raised up from the base of the minor groove. In case of intrastrand adduct with d(CTCATCAC).(GTGATGAG), covalent linkages were formed with the exocyclic NH_2 groups of DNA G11 and DNA G14 (Figure 1.32); the stereochemical configurations with C11 of PBD molecule and NH_2 of guanine have been confirmed as *S* at both ends of the duplex that further confirmed that the *S*-configuration at C11 of SJG-136 is energetically favourable. The strong binding affinity of the dimer was shown by a thermal denaturation assay using calf thymus DNA. With a DNA: ligand ratio of 5:1, after incubation at 37 °C for 18 hours, SJG-136 increased the melting temperature by 33.6 °C, which is highest among all PBDs to date (177). In addition, at a concentration > 0.3 μM this compound produced 100% cross-linking using linear duplex DNA, as demonstrated using agarose gel electrophoresis (177).

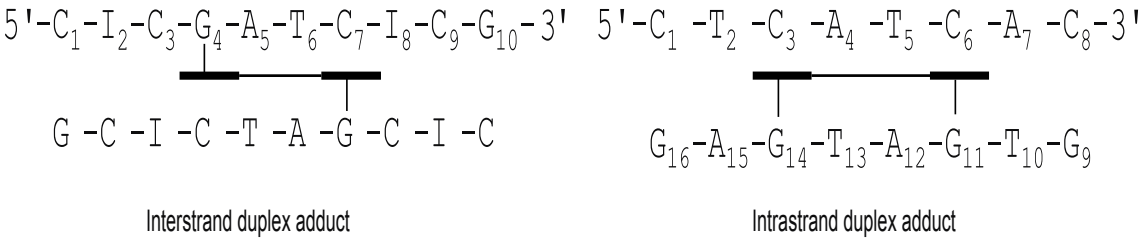
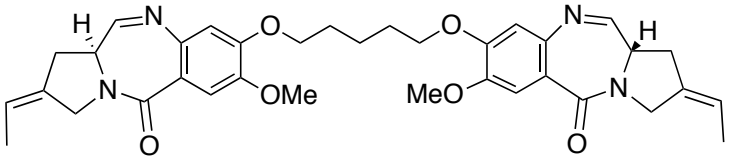


Figure 1.32: A. Interstrand adduct with d(CICGATCICG)$_2$ B. Intrastrand adduct with d(CTCATCAC).(GTGATGAG) (176).

SJG-136 was extensively investigated to evaluate its *in vitro* and *in vivo* antitumor activity using different biochemical techniques (178). SJG-136 was tested in the standard NCI hollow fibre assay against 12 cell lines, with doses of 0.5 and 0.4 mg/kg given once daily for 4 days. It was found that more than 50% cell growth was inhibited and 10% to 47% cells were killed in 5 of these 12 cell lines. The highest toxicity was observed in ovarian adenocarcinoma cells and subsequent testing in ovarian cancer xenograft models revealed marked *in vivo* antitumor activity. The ligand was also examined against a cisplatin resistant tumour where it showed significant growth delay at a dose of 0.2 mg/kg (118).

ELB21

ELB21 is composed of two PBD units tethered with a five-carbon diether linker (Figure 1.33). The preferred binding site for this ligand is thought to be 5′-Pu-**G**ATTC-Py and 5′-Pu-**G**AATC-Py (Pu = purine; Py = pyrimidine, reactive guanines shown in bold), producing interstrand cross-linking between the guanines (179). Recent footprinting studies using a fragment containing symmetrical hexanucleotide sequences revealed high affinity binding sites including 5′-**G**AATTC and 5′-**G**TATAC flanked by a purine at the 5′-end and a pyrimidine at the 3′-end (179). The compound is also capable of forming adducts with guanines separated by 2 to 6 bases in the same strand.



ELB21

Figure 1.33: Structure of ELB21

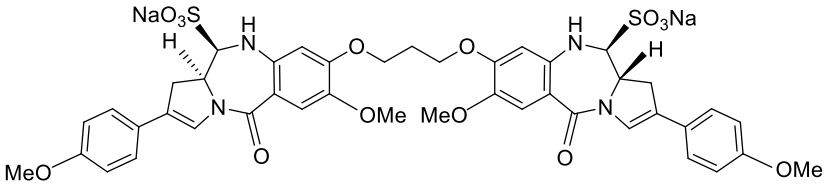
ELB21 is the most potent antibacterial agent in this group. It was tested against methicillin resistant *Staphylococcus aureus* (MRSA) showing significant antibacterial activity with a MIC50 value of 0.03 mg/L. There are hundreds of 5′-Pu-**G**A(T/A)TC-Py in close proximity to transcription sites in the *S. aureus* EMRSA-16 genome. Binding of ELB21 with these sites prevents the separation of two DNA strands, resulting in blocking of transcription (179).

SG2285

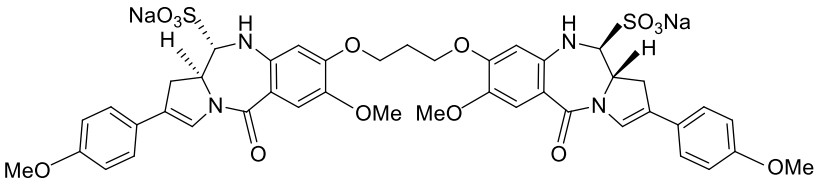
SG2285 is the prodrug form of SG2202 (Figure 1.34). The active molecule SG2202 has been synthesized from DSB120 with an additional aryl substitution at the C2/C2´ positions (180). The compound was observed to be more cytotoxic *in vitro* than DSB120, but *in vivo* analysis of the molecule was difficult due to its poor water solubility. The problem was overcome by introducing bisulphite groups at C11/C11´ positions. The presence of bisulphite gave rise to two water-soluble diastereomeric forms; 11S, 11´R and 11S, 11´S, which retained the *in vitro* antitumor activity of SG2202.

The antitumor activity of SG2285 was studied both *in vitro* and *in vivo* using different cell lines. The molecule recognizes the 8-base pair long 5´-XAGATCTX-3´ sequence (X = any base) in duplex DNA, producing interstrand crosslink between guanine N2 groups spanning 4 bases (181) (Figure 1.35). The cross linking ability *in vitro* was determined in the human prostate cancer cell line DU145 which showed significant levels of cross linking after 8 hours at 1 nmol/L. The antitumor activity *in vivo* was

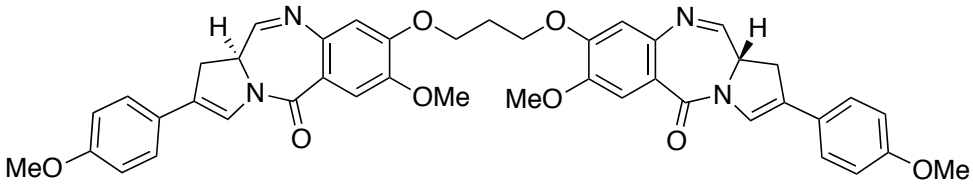
measured in a variety of human tumour xenograft models, including early stage melanoma and ovarian models using intravenous dosing schedules where the molecule was found to be highly potent (181).



A. SG2285 (11S,11R´)



B. SG2285 (11S, 11S´)



C. SG2202

Figure 1.34: Structures of diastereomers of SG2285 (A and B) and The active form of SG2202 (C) (180,181)

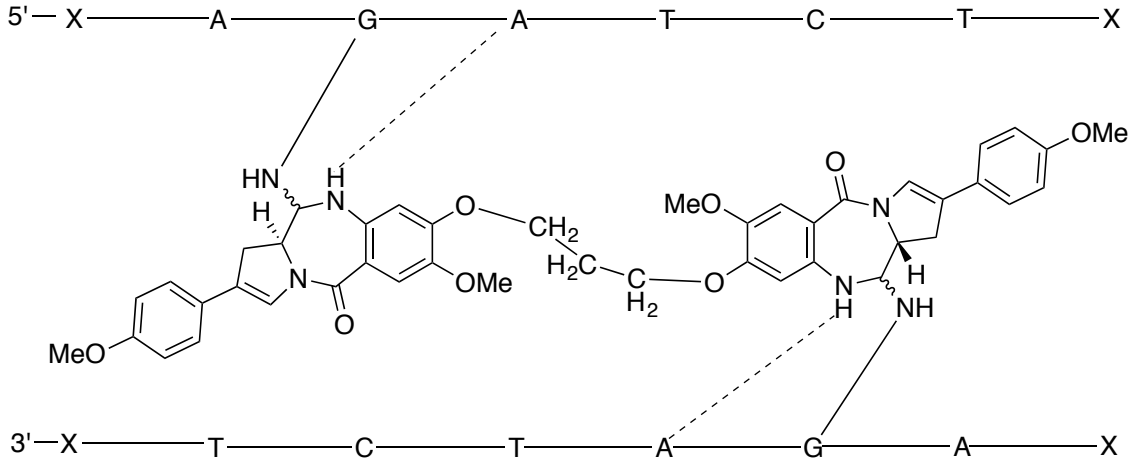
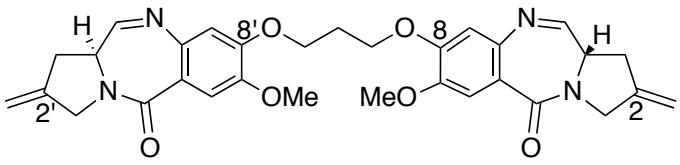


Figure 1.35: Interstrand cross-linking of SG2202 with GATC (181)

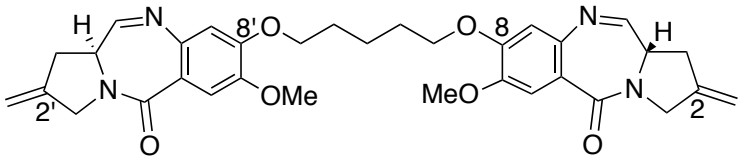
Interstrand, intrastrand or mono-adduct by PBD-dimer?

A bifunctional PBD-dimer, having PBD units at each end, is able to alkylate two guanines in the duplex DNA. Covalent interaction between the DNA and PBD-dimer is

approximately twice as favourable as that for the corresponding monomers, which reflects the bifunctional characteristics of PBD-dimer (182). Both halves of the dimer behave in identical fashion without disrupting the host DNA integrity. The linker between two PBD units provides a flexible connection and makes favourable non-bonding contacts with the floor and walls of the minor groove that promotes overall binding. The alkylation can produce an interstrand cross-linked adduct by reacting with two guanines on complementary strands or intrastrand adducts by joining two guanines of the same strand. At the same time, one of the two PBD units could react with one guanine producing a monoalkylated adduct. It has been reported that the PBD dimers (with C8/C8' linking) can form all three types of adducts: inter- and intrastrand cross-linked adducts and also mono-alkylated product (183). To examine the type and distribution of these adducts produced, two PBD dimers SJG1-136 and DRG-16 (C8-O-$(CH_2)_n$-O-C8') (n = 3 and 5 respectively) (Figure 1.36) were tested against a set of oligonucleotide duplexes of varying length and sequences. The duplexes were designed to contain potential binding sites for either two guanines on the same strand or one guanine in each strand, separated by 1, 2, 3 or 4 A/T base pairs. An HPLC/MS assay of these DNA-ligand complexes revealed that the type and relative distribution of the adducts formed depend on i) the positioning of two reactive guanines on the same or opposite strands, ii) number of bases (A/Ts) between the guanines and iii) the length of the linker (183). SJG-136 (trimethylene linker) prefers to form interstrand cross-linking adducts with Pu-GAC-Py and Pu-GATC-Py over intrastrand product with Pu-GAG and Pu-GATG-Py respectively *i.e.* two reactive guanines are separated by one or two A/T base pair(s), whereas binding sites containing three A/T base pairs are more likely to produce intrastrand cross-linking adducts over interstrand (Pu-GAATG-Py over Pu-GAATC-Py). The same phenomenon was observed for DRG-16 (pentamethylene linker) but in relation to DNA sequences with two guanines separated by 2, 3 and 4 rather than 1, 2 and 3 A/T base pairs. Thus two extra methylenes in the linker of DRG-16 cover one additional A/T base pair (183).



SJG-136



DRG-16

Figure 1.36: Structure of SJG-136 and DRG-16

The preference for forming interstrand over intrastrand adducts and *vice versa* have been explained by molecular modelling studies (184). Molecular modelling simulations with SJG-136 on Pu-GATC-Py, Pu-GATG-Py, Pu-GAATC-Py and Pu-GAATG-Py revealed that the ligand bound with all four duplexes with relatively well-tolerated distortions of the DNA helices (Figure 1.37). With GATC and GATG, one PBD unit covalently links with one guanine of either duplex and the electrophilic C11 atom of the second PBD unit appears well positioned to react with the C2-NH_2 of the second guanine, whether on the same or opposite strand; thus the free PBD unit is able to form either inter- or intrastrand cross-links with equal ability. On the other hand, binding of one PBD unit to guanine in GAATC and GAATG causes helical distortion in both duplexes; in this case, GAATC is more distorted than GAATG, which promotes intrastrand cross-link than interstrand adduct formation. From a thermodynamic point of view, GATC and GAATG have slightly lower energy compared to GATG and GAATC, which again facilitates adduct formation.

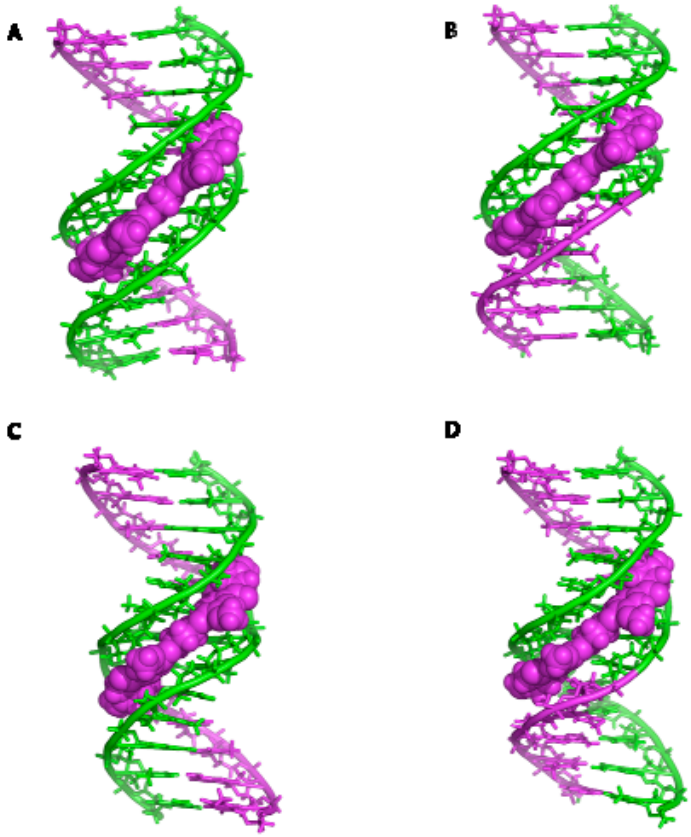


Figure 1.37: Molecular models of the adducts of SJG-136 with different DNA binding sequences (A) Pu-GATC-Py (B) Pu-GATG-Py (C) Pu-GAATC-Py (D) Pu-GAATG-Py. All configurations appear to be relatively well-tolerated with negligible distortion of the helices (184).

The formation of inter- and intrastrand cross-linked adducts is also accompanied by presence of minor amount of monoalkylated adduct (184). This monoalkylated adduct is presumably an intermediate entity formed on-route to the duplex adduct. Here, one PBD unit binds covalently producing monoalkylated adduct; if a second guanine is available at an appropriate distance from the first guanine, either on the same or opposite strand, the other PBD unit will bind to form an intra or interstrand adduct. With two adjacent guanines (Pu-GG-Py or Pu-GC-Py), only one PBD unit can react with one guanine; reaction with the second guanine is impossible due to steric clash. A monoalkylated adduct was also observed with Pu-GAG-Py, which was supposed to produce intrastrand cross-linking; binding of one PBD with guanine in GAG causes significant distortion in double helix that makes the C2-NH_2 of second guanine unavailable for the second PBD moiety. On the other hand, if no neighbouring guanine is available, then only a monoalkylated adduct will be observed. Adduct formation with the inosine-containing duplex (Pu-GATI-Py) showed only monoalkylated adduct (175) as this duplex offers only one guanine 2-amino group. Also, no cross-linked adducts were observed with Pu-GAAATG-Py or Pu-GAAATC-Py with SJG-136, as the ligand

cannot stretch this far between two guanines on the same or opposite strand; while for DRG-16, the monoalkylated adduct-forming sequences are Pu-GAAAATG-Py or Pu-GAAAATC-Py *i.e.* one base pair longer than for SJG-136 (183).

Besides the effect of base sequence on the type of adduct formation, C2-exo-unsaturation facilitates cross-linking; C2-exo-unsaturation causes a more isohelical fit of the molecule in the minor groove, due to the flattening of the C-rings by C2/C2'-methylene groups. On the other hand, steric clash between the sugar wall and the C2-H atoms occurs with a saturated C-ring that prevents full isohelical fit of the PBD units and the linker (Figure 1.38).

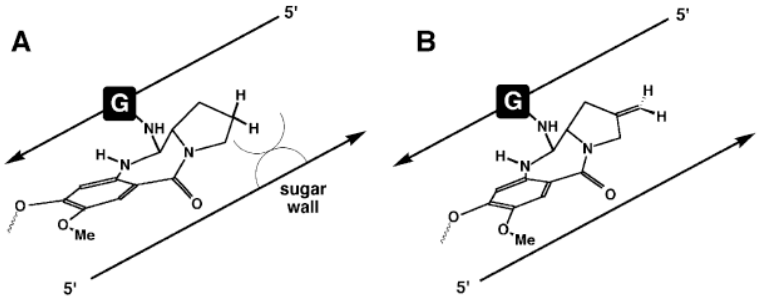


Figure 1.38: Schematic representation showing the fit of a PBD subunit within the minor groove of a host DNA duplex for PBD dimers (A) with C2-saturated C-rings (B) exo-unsaturated rings (185)

1.2.4.6.5 PBD-antibody conjugates

Kung Sutherland *et al* (186) have reported the conjugation of a novel synthetic pyrrolobenzodiazepine dimer (SGD-1882) (Figure 1.39) with the anti-CD33 monoclonal antibody for the treatment of acute myeloid leukaemia (AML). CD33 is a transmembrane receptor that is expressed in malignant cells of majority of AML patients. Treatment with unconjugated anti-CD33 antibodies was unsuccessful (187,188) while gemtuzumab ozogamicin (GO), composed of anti-CD33 antibody complexed with calicheamicin, showed some activity in patients with relapsed AML and also improved survival in some patients with newly diagnosed AML (24,33). Based on this outcome, a PBD dimer (SGD-1882) has been conjugated with an anti-CD32 antibody and tested against human AML cell lines and primary patient samples. The conjugate was found to be highly active against 12 tested AML cell lines with a mean IC_{50} of 22 ng/ml. In samples from AML patients, this conjugate also showed significant potency having an IC_{50} of 8 ng/ml whereas gemtuzumab ozogamicin had 27 ng/ml. Another drug-antibody conjugate (189) was synthesized from this PBD-dimer with an anti-CD70 antibody. Treatment against CD-70 positive renal cell carcinoma and non-Hodgkin carcinoma displayed pronounced antitumor activities.

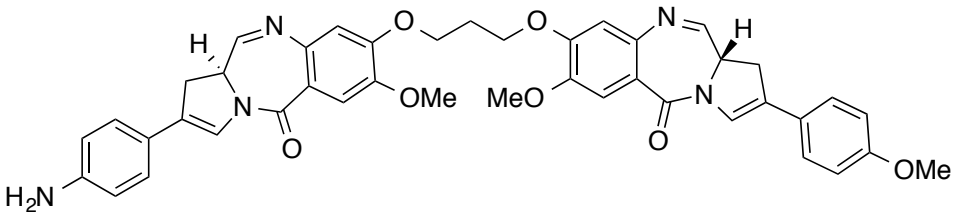


Figure 1.39: Structure of SGD-1882

1.3 DNA repair mechanisms for damages induced by alkylating agents

1.3.1 MGMT repair

O^6-methylguanine-DNA methyltransferase (MGMT) plays a vital role in defence against the DNA lesion of O^6-alkylguanine and to a lesser extent O^4-alkylthymine, restoring the base to its original form. The repair mechanism consists of a one-step alkyl transferase reaction that transfers the damaging alkyl group to a cysteine residue in the catalytic site of MGMT. Although MGMT has preference for O^6-MeG, it also removes longer alkyl groups such as ethyl, *n*-propyl, *n*-butyl and 2-chloroethyl from the O^6-position. If unrepaired, O^6-MeG forms an O^6-MeG:T mismatch during replication, that will give rise to A:T transition mutation in the next round of replication. Cells deficient in MGMT are unable to repair O^6G-methylation damage and are therefore more sensitive to the effects of methylating agents than normal cells (190,191). Xenograft model in mice with 0% to 10% MGMT-positive cells were highly sensitive to BCNU, whereas increasing percentages of MGMT-positive cells showed enhanced resistance to the drug (192). Inhibition of MGMT with O^6-benzylguanine and subsequent administration of temozolomide showed increased antitumor activity both *in vitro* and *in vivo* (193-195). In addition, the expression of MGMT can be inhibited by methylation of its gene promoter region (196,197). Tumours with methylated MGMT promoters are more sensitive to alkylating agents while tumours with unmethylated MGMT promoters express high levels of the enzyme and are more resistant to alkylating agents (198).

DNA base adducts at N1MeA, N3MeC, N3MeT and N1MeG are repaired by human AlkB homologues 2 and 3 (ABH2 and ABH3) (199,200). ABH2 and ABH3 belong to the alpha ketoglutarate and Fe(II)-dependent dioxygenase super-family. The repair mechanism involves oxidation of the alkyl group, which results in demethylation *i.e.* direct damage reversal and restoration of the undamaged base.

1.3.2 Mismatch repair (MMR)

The cytotoxicity of monofunctional alkylating agents through O^6-methylguanine requires a functional MMR in the targets cells. MMR deals with base mismatches that are caused by spontaneous or induced base deamination, oxidation, methylation and by

replication errors and consists of several proteins (MutSα, MutLα, etc.). Guanine/thymine mismatches could naturally arise from the deamination of 5'-methylcytosine; this mismatch is recognised by MutSα. MMR removes the nucleotides along with a long stretch of DNA strand specifically from the newly synthesized strand. During DNA replication, DNA polymerase mispairs O^6-MeG with thymine that ought to be repaired by MGMT. If the mismatch (O^6-MeG:T) is not repaired by MGMT, the base pair is also recognized by MutSα initiating MMR pathway (201). However, MMR is not able to process the lesion correctly as it inserts a new thymine opposite to O^6-MeG leading to repetitive cycle of MMR (202-204). This futile MMR cycle produces gapped DNA that causes double strand breaks (DSBs) and leads to p53 dependent cell cycle arrest and apoptosis (205). Therefore, cells with deficiency in MMR become tolerant with O^6-MeG at the expense of mutations and tumours lack of MMR are found relatively resistant to the cytotoxic effects alkylating agents. In fact, mammalian cells proficient in MMR are found 100-fold more sensitive to alkylating agents than MMR-deficient cells (206,207).

1.3.3 Base excision repair (BER)

The alkylation adducts N7MeG, N3MeA and N3MeG are repaired by base excision repair (BER) system. BER is initiated by recognition of the modified bases by a DNA glycosylase, that excises the base by hydrolysing the N-glycosidic bond producing an apurinic (AP) site (208). The resulting AP site is recognised by an apurinic endonuclease (APE1) that cleaves the damaged strand leaving 3'-OH and 5'-deoxyribosephosphate (5'-dRP). DNA polymerase β (pol β) then inserts a single nucleotide in the repair site while removing the cytotoxic 5'-dRP group. For long patch repair, further DNA synthesis is performed by pol ε and pol δ together with PCNA, resulting in repair patches up to 10 nucleotides (209,210). Finally, the ligation step is conducted by DNA ligase I, or a complex of DNA ligase III and XRCC1. In the absence of pol β, cells cannot repair the BER intermediate 5'-dRP and are hypersensitive to the alkylating agents (211-213). Mouse fibroblasts defective in pol β were found to have a higher frequency of chromosomal aberrations and apoptosis following alkylating damage (214). XRCC1 mutant hamster cells (215) and ligase I deficient human fibroblasts (216) are also sensitive to alkylating agents indicating N-methylation lesions can be cytotoxic if not repaired.

1.3.4 Nucleotide excision repair (NER) for crosslink adducts

Interstrand crosslinks (ICLs) are highly toxic DNA lesions that prevent transcription and replication by creating an absolute block for strand separation. The covalent linkages

between bases on opposite strands are usually irreversible. If left unrepaired, interstrand crosslinks can lead to cell death and it has been reported that as few as 20 interstrand crosslinks in the bacteria or mammalian genome can be lethal to cells that are deficient to repair mechanism of such lesion (217,218). Bonding between bases in the same strand forming intrastrand crosslink can be bypassed by some DNA polymerases making this type of lesions less toxic during replication than crosslinks between two DNA strands. Repair of interstrand crosslinks requires involvement of multiple repair pathways- nucleotide excision repair (NER), homologous recombination (HR) and translesion synthesis (TLS). In brief, NER consists of four steps- damage recognition and formation of a preincision complex, asymmetric incisions on the 5' and 3' sides of the DNA damage to produce a short DNA fragment that contains the adduct, DNA synthesis and finally ligation to complete the repair of the DNA. In *E.coli*, DNA damage is recognised by an ATP-dependent process mediated by the $(UvrA)_2(UvrB)_1$ complex and results in the loading of UvrB with the DNA. After dissociation of UvrA, the remaining UvrB recruits UvrC to the damaged site followed by DNA incisions. The incision complex is then removed and DNA pol I synthesizes the incised gap. The process is then completed by the action of DNA ligase. In mammalian cells, NER is initiated by the recognition of the crosslink by XPC protein. The first incision step on either side of the lesion on one strand of the duplex is performed by XPF-ERCC1 complex and generates a gapped structure while flipping out a monoadduct-like structure of the opposite strand. DNA synthesis or gap filling is carried out by either DNA pol δ or pol ε followed by ligation by human DNA ligase I. Again, the NER pathway will initiate the second cycle of repair and remove the remaining adduct on the opposite strand. Cellular sensitivity of PBD dimer (SJG-136) was found to depend on XPF-ERCC1 complex and homologous recombination repair factors XRCC2 and XRCC3 (219) .

1.3.5 Role of Glutathione-S-transferase

Glutathione-S-transferases (GST) are a family of phase II detoxification enzymes that protect cellular macromolecules from attack of reactive electrophiles by catalyzing the conjugation of glutathione (GSH) to a wide number of endogenous and exogenous electrophilic compounds. GSTs have been implicated in the development of resistance towards alkylating agents, insecticides, herbicides and microbial antibiotics (220-223). The enzymes produce drug resistance in two distinct ways - firstly, by direct detoxification of the drug and secondly, acting as an inhibitor of the MAP kinase pathway. Many alkylating agents *e.g*. chlorambucil, melphalan, nitrogen mustard, etc. are directly inactivated by GST through catalytic conjugation with glutathione by

thioether bond formation; many cancer drugs which are converted into electrophilic species can also be detoxified by glutathione metabolism (224). The mitogen-activated protein (MAP) kinase pathway is involved in cellular survival and death signaling; GST interacts with one MAP kinase - c-Jun N-terminal kinase 1 (JNK1). JNK1 is involved in stress response and apoptosis (225). One polymorph of GST family (GSTπ) plays a vital role in the regulation of MAP kinase pathway; this GST sequesters JNK1 forming GSTπ:JNK1 complex which is obvious in nonstressed cells by the presence of very low JNK1 activity (225). In many drug-resistant tumors, high levels of GSTs are observed that indicate the correlation between GST expression and sensitivity toward alkylating agents (226). In addition to this, some anticancer agents such as cisplatin require JNK1 activation for cytotoxicity; inhibition of JNK1 signaling pathway by high level of GST leads to a decrease in cisplatin induced apoptosis while overexpression of JNK increases the sensitivity of cells towards cisplatin (227). In some cases, increased levels of GST are also associated with other proteins such as p28 (228), fructose-1,6-bisphosphatase (229) that also cause enhanced resistance to apoptosis.

1.4 Biochemical techniques

The binding efficiency, formation and stability of cross-linked products of PBDs with DNA can be evaluated by a several biochemical techniques. The techniques used in this thesis include footprinting, ligation assays and fluorescence melting studies to determine the preferred ligand binding sequences, and each of these is described below.

1.4.1 Footprinting

Footprinting is a method for identifying the binding sites for ligands on double stranded DNA. The method was first used in 1978 to examine the interaction between proteins and DNA (230). It is now widely used to explore the binding sites of many small molecules to DNA and has been successfully used to determine the sequence specificity of ligands such as actinomycin (231,232), mithramycin (233,234), minor groove binding ligands (135,231,232) and triplex forming oligonucleotides (235-237). The method involves radiolabelling of one strand of a template duplex DNA, usually at the 3´-end; though 5´-end labelled DNA can also be used. The radiolabelled DNA is then digested with an appropriate cleavage agent that cleaves DNA non-specifically at all sites, except at the sites to which the ligand is bound. When the cleavage products are run on a denaturing polyacrylamide gel, they produce a series of radiolabelled bands depending on their size; the ligand-protected sites are then apparent as a gap in the ladder of bands (Figure 1.40).

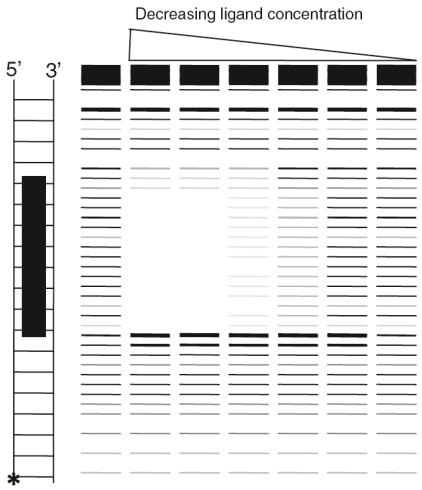


Figure 1.40: The radiolabelled template DNA on the left (asterisk indicates the position of the radiolabel). The footprinting gel is on the right. The first lane shows cleavage pattern of the DNA without any ligand; the gap in second and third lanes (the footprint) indicates the site to which the ligand is bound (238).

1.4.1.1 *Cleavage agents*

The commonly used cleavage agents in footprinting experiments are DNase I, hydroxyl radicals and methidiumpropyl-EDTA (MPE) (239); among these, DNase I is the most widely used. DNase I is an endonuclease that cleaves one strand of double stranded DNA at the O3'-P bond, yielding 5´-phosphoterminated nucleotides with a free hydroxyl group at 3´-position. It also cuts single stranded DNA at a lower rate. The enzyme is a glycoprotein with molecular weight 30,400. For enzymatic action, it requires divalent cations such as calcium or magnesium (240-242). Magnesium is about 10-fold more efficient than calcium and is so routinely used in footprinting buffers. The reaction is further enhanced by inclusion of manganese (241). DNase I performs its cleavage action by inserting an exposed loop into the minor groove of the substrate (241,243,244). This is accompanied by bending of the DNA towards the major groove (243,244). That's why, the width and flexibility of minor groove are very important for its action. For these reasons, DNase I cuts poorly at runs of As and Ts, which typically possess very narrow minor grooves, and also at GC-rich sequences, which have much lower flexibility. DNase I binds to about 10 base pairs (one turn of the DNA helix); at least 4 base pairs to the 5'-side and 6 base pairs to the 3'-side of the cutting site are covered by DNase I, thus limiting the accuracy for estimating preferred ligand binding sites (245). As DNase I binds across the minor groove and the phosphodiester backbone of DNA is inclined relative to the helical axis, the region protected from cleavage is staggered across the two strands. Thus the footprint is staggered in the 3'-end and the enzyme overestimates the ligand site size by about 3 base pairs (239).

Although the enzyme produces an uneven ladder of cleavage product, it can work over a wide pH range (5.0 to 9.0), temperatures (4-70°C) and ionic concentrations (up to 1M NaCl) (239). DNase I is cheap, easy to use and its enzymatic action can be stopped by adding EDTA to chelate the essential divalent cation.

Footprinting can also be performed with hydroxyl radicals, which, on account of their very small size, are able to cleave DNA close to the bound ligands, and thus enable the position of the binding site to be determined with greater accuracy. The process produces even ladder of cleavage products with very little sequence dependence (246-248) and is quite useful for probing the interaction of minor groove binders (249-251). The experiment is performed by producing hydroxyl radicals via the Fenton reaction (252) between Fe^{2+} and H_2O_2 ($Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + {}^{\cdot}OH + OH^-$). Hydroxyl radicals attack at the C4' or C1' positions via the minor groove, which ultimately leads to base removal and strand scission. Interaction of Fe^{2+} with negatively charged DNA substrates is prevented by the presence of EDTA that chelates the iron as $[Fe(EDTA)]^2$. Although hydroxyl radical footprints are more accurate that those of DNase I, the method has some limitations. The method is time consuming as the reaction is much slower and it requires precipitating the reaction products before loading onto a polyacrylamide gel. Also it is sensitive to the presence of solvent such as DMSO or ethanol that scavenge free radicals. However, hydroxyl radical footprinting is useful for probing the interaction of minor groove binding ligands with DNA (134,253,254). Methidiumpropyl-EDTA (MPE) footprinting is another technique which produces single strand breaks in double helical DNA in the presence of Fe(II) and O_2 (255). MPE.Fe(II) is a relatively non-sequence specific DNA cleaving agent with lower sequence specificity than DNase I.

1.4.1.2 *DNA substrates*

DNA substrates for footprinting are usually 50-200 base pairs long (239). For identifying the binding site of a ligand, it is essential to have the preferred sequence in the substrate. Ligands with short sequence selectivity can be easily studied by footprinting with common restriction fragments as their short binding sites will be present; for example the preferred binding sequences of echinomycin (CpG) and actinomycin (GpC) have been examined with great ease by footprinting; there are only 10 possible dinucleotide sequences, and each of these will be present in all DNA substrates. On the other hand, ligands selective for four or more base pairs require particular DNA fragments with special combinations of base pairs, as the preferred sites may not be present within any 200 base pair restriction fragment. For such ligands, synthetic DNA fragments such as MS1/MS2 (256) and HexA/HexB (257) have

been prepared. MS1 and MS2 contain all 134 possible tetranucleotide sequences whereas HexA and HexB together contain all 64 symmetrical hexanucleotide sequences. Footprinting with these fragments enables the identification primary binding preferences for most small molecules.

1.4.2 Fluorescence melting studies

The stability of nucleic acid-ligand complexes is often measured by thermal denaturation experiments. Heating a sample containing duplex DNA leads to the melting or denaturation of the duplex *i.e.* breaking the hydrogen bonds between base pairs with subsequent separation of double strands. Upon cooling the same sample, reestablishment of hydrogen bonding occurs thus two strands become annealed. In the melting step, the temperature at the mid-point of the transition indicates the stability of the structure (T_m) and changes in this value (ΔT_m) are used to compare the effect of experimental conditions, base substitution and drug binding. Ligands that bind to double stranded DNA increase the melting temperatures and these changes can be used to compare the binding of different ligands or to assess the relative stabilization of different sequences. Melting transitions can be detected by various techniques such as fluorescence (258), NMR (259,260), circular dichroism (261-263), Raman (264) and UV spectroscopy (265). Among these, UV spectroscopy is the most commonly used technique that involves measuring changes in absorbance at 260 nm. UV spectroscopy is simple to use and it is easy to interpret the data qualitatively. However, most spectrophotometers measure no more than four samples at a time, thus the technique has low throughput, which is a big limitation. The technique also requires relatively large sample volumes (1-3 ml) and the changes in absorbance are not large (~25%). In addition to these, higher order nucleic acid structures such as triplexes and quadruplexes show multiple components in their melting profiles; these components often overlap and UV spectroscopy cannot resolve different transitions. To overcome these limitations, a novel technique (266) has been devised, based on molecular beacon methods (267); the technique measures changes in the fluorescence yield of oligonucleotides containing suitably placed fluorophores and quenchers. This is based on the use of molecular beacons as a method for detecting specific nucleic acids in homogenous solutions by constructing single stranded nucleic acid probe molecules. The probe possesses a stem-and-loop structure (267) in which the loop is complementary to a predetermined sequence of target nucleic acid while the stem is formed by the annealing of the two complementary arms forming a stem structure (Figure 1.41). A fluorophore is attached to the end of one arm and a quencher to the end of other arm. The stem keeps these two moieties in close proximity, thus the

fluorescence of the fluorophore becomes quenched. When the probe detects a target DNA sequence, it forms a hybrid that is longer and more stable than the hybrid formed by the arm sequences; thus hybrid of the arm sequences can no longer exists. As a result, the fluorophore and quencher of two arms move away and gives a high fluorescence. Tyagi *et al* (267) named the method as 'molecular beacons' as they emit a fluorescent signal only when hybridized to target DNA sequences.

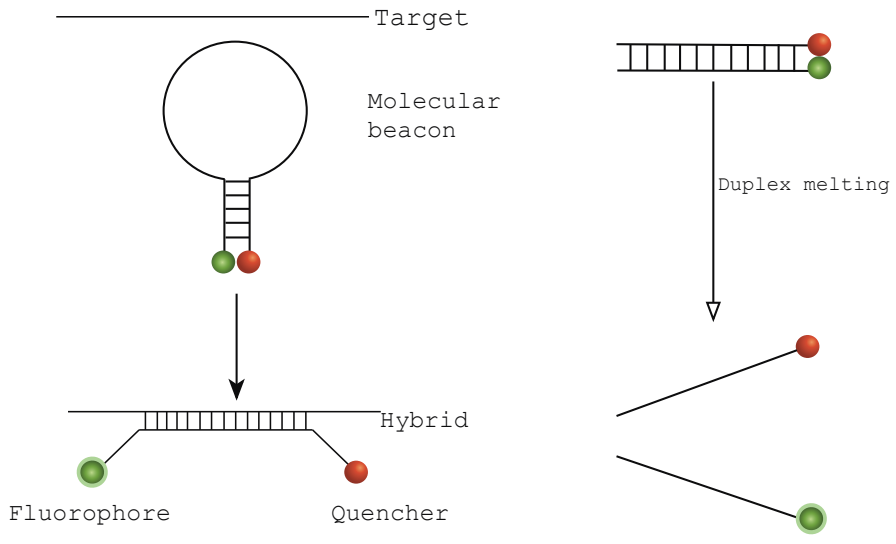


Figure 1.41: (A) Principle of operation of molecular beacons (267). (B) Schematic representation of the melting of oligonucleotides designed to show duplex melting (266).

Based on molecular beacon method, Darby *et al* (266) developed a method to determine the stability of duplex, triplex and quadruplex DNA structures. In case of duplex DNA, an oligonucleotide pair or hairpin is used where one oligonucleotide contains a fluorophore and the other has a quencher moiety. In the duplex state, the fluorophore and quencher are in close proximity giving very low fluorescence. As the duplex DNA melts, these two groups are separated and there is a large increase in fluorescence. In presence of ligand that has a suitable binding site in the duplex DNA, melting of DNA occurs at higher temperature to show the fluorescence; thus the stronger the ligand-DNA binding, the higher the melting temperature it requires to show fluorescence.

1.5 Aim of the study:

Conjugation of pyrrolobenzodiazepines with different heterocycles provides a very promising method for targeting specific DNA sequences. The aim of this research is to

determine the sequence specificity of two classes of pyrrolobenzodiazepine conjugates- i) conjugates with pyrrole, imidazole and thiazole and ii) conjugates with benzofused heterocycles- benzofuran, benzothiophene and indole. The sequence selectivity will be assessed by DNase I footprinting experiments with different DNA fragments. Fluorescence melting profiles are also used with selected synthetic oligonucleotides. Furthermore, a novel ligation assay is tested to explore the sequence selectivity of a bis-PBD.

2 Materials and methods

2.1 Materials

2.1.1 Chemicals, enzymes and reagents

All chemicals for use in standard buffers were purchased from Sigma Aldrich (Poole, UK). Restriction enzymes were purchased from Promega (Southampton, UK) and stored at -20 °C. T4 DNA ligase, polyncleotide kinase (PNK) and Klenow fragment were purchased from New England Biolabs, UK and reverse transcriptase from Sigma Life Sciences, UK. Plasmid purification kit (Mini prep) was purchased from Qiagen. Radioactive [α-^{32}P] dATP and [γ-^{32}P] ATP were purchased from Perkin Elmer with an initial activity of 3000 Ci/mmol. Polyacrylamide concentrates for gel electrophoresis as Accugel (40% w/v 19:1 acrylamide: bis-acrylamide solution) and Sequagel (25% w/v 19:1 acrylamide: bis-acrylamide solution containing 8M urea) were purchased from National Diagnostics (Hull, UK).

2.1.2 Ligands

PBD dimers and PBD-conjugates with benzofused rings were synthesized in Professor David Thurston's Laboratory (King's College, London). PBD-conjugates with pyrrole, imidazole and thiazole were synthesized in Dr Federico Brucoli's Laboratory (University of Reading). The hairpin polyamides were synthesized in Dr Glen Burley's laboratory (University of Leicester). The ligands were initially in powder form and were dissolved in DMSO (Sigma Aldrich, UK) to prepare stock solution varying from 1 mM to 10 mM, which were stored at -20 °C; these were diluted to working concentrations in aqueous buffers immediately before use.

2.1.2.1 PBD conjugates with pyrrole, imidazole and thiazole ring (detail structures are given at section 3.3).

Compound	Identity
RMH-35	Thz-Py-PBD
RMH-71	Thz-Thz-PBD
RMH-41	Py-Py-Im-PBD
RMH-53	Im-Im-Py-PBD
RMH-43	Py-Py-Thz-PBD
RMH-47	Py-Thz-Py-PBD

Table 2.1: Chemical composition of PBD conjugates with pyrrole, imidazole and thiazole. The name of the compounds in the left column is as given by Dr Federico Brucoli.

2.1.2.2 PBD conjugates with benzofused rings (detail structures are given at section 4.3)

Compound	Identity
KMR30	Indole-Py-PBD
KMR31	Benzofuran-Py-PBD
KMR32	Benzothiophene-Py-PBD
KMR33	Benzothiophene (acetate)-Py-PBD
KMR164	Benzothiophene (acetate)-Py-Ph-PBD
KMR171	Indole-Im-PBD
KMR173	Benzothiophene (acetate)-Im-PBD
KMR175	Benzothiophene-PBD

Table 2.2: Chemical composition of PBD conjugates with benzofused rings. The name of the compounds in the left column is as given by Dr Khondaker Rahman.

2.1.2.3 Hairpin polyamides (detailed structures are given at section 6.3)

Compound	Identity
PA1	ImImPyPy-γ-ImPyPyPy-β-Dp
PA2	ImImPyIm-γ-PyPyPyPy-β-Dp

Table 2.3: Chemical composition of hairpin polyamides

2.1.3 Oligonucleotides

Oligonucleotides for ligation studies and for fluorescence melting experiments were synthesized in Professor Tom Brown's Laboratory (Department of Chemistry, University of Southampton) and stored at -20°C. Oligonucleotide sequences are provided in the relevant chapters.

2.1.4 Buffers and solutions

5X TBE	108 g Tris, 55 g Boric acid, 9.4 g EDTA (2L)
Diluent	50% (w/v) urea
Transformation buffer	50 mM $CaCl_2$, 10 mM Tris-HCl, pH 7.4
DNA elution buffer	10 mM Tris-HCl, 0.1 mM EDTA, pH 7.5
Buffer for ligands	10 mM Tris-HCl, 10 mM NaCl, pH 7.5
Buffer for ligands (cross-linking assay)	10 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, pH 7.5
DNase I buffer	20 mM NaCl, 2 mM $MgCl_2$, 2 mM $MnCl_2$
Formamide stop solution (DI stop)	10 mM EDTA, 1 mM NaOH, 0.1% bromophenol blue, 80% Formamide

Loading dye	20% Ficoll, 10 mM EDTA, 0.01% w/v bromophenol
Lithium phosphate buffer	40 mM lithium hydroxide, adjusted to pH 7.4 with phosphoric acid
Sodium Phosphate buffer	10 mM Sodium phosphate containing 100 mM NaCl, pH 7.4

2.2 Footprinting

Three duplex sequences named HexA, HexB and MS1 (Figure 2.1) were used in the footprinting experiments. Three other duplex sequences named GXC3, GXC2 and CXG2 (Figure 2.2) were used only for PBD-conjugates with pyrrole, imidazole and thiazole; DNA fragments LE1 and LE2 (Figure 2.3) were used for hairpin polyamides. The HexA, HexB and MS1 sequences were originally cloned (257) into the *BamH*I site of the pUC18 polylinker. HexA and HexB fragments together contain all 64 symmetrical hexanucleotide sequences whereas MS1 contains all 136 tetranucleotide sequences. All the sequences were also cloned in the opposite orientation named as HexArev, HexBrev and MS2 respectively, for ease of visualising the complementary strand in the footprinting gels.

HexA

```
5′GGATCCCGGGATATCGATATATGGCGCCAAATTTAGCTATAGATCTAGAATTCCGGACCGCG
3′CCTAGGGCCCTATAGCTATATACCGCGGTTTAAATCGATATCTAGATCTTAAGGCCTGGCGC

GTTTAAACGTTAACCGGTACCTAGGCCTGCAGCTGCGCATGCTAGCGCTTAAGTACTAGTGCAC
CAAATTTGCAATTGGCCATGGATCCGGACGTCGACGCGTACGATCGCGAATTCATGATCACGTG

GTGGCCATGGATCC3′
CACCGGTACCTAGG5′
```

HexB

```
5′GGATCCGGCCGATCGCGAGCTCGAGGGCCCTAATTAGCCGGCAATTGCAAGCTTATAAGCGC
3′CCTAGGCCGGCTAGCGCTCGAGCTCCCGGGATTAATCGGCCGTTAACGTTCGAATATTCGCG

GCTACGTATACGCGTACGCGCGTATATACATATGTACATGTCGACGTCATGATCAATATTCGAA
CGATGCATATGCGCATGCGCGCATATATGTATACATGTACAGCTGCAGTACTAGTTATAAGCTT

TTAATGCATGGATCC-3′
AATTACGTACCTAGG-5′
```

MS1

```
5′GGATCCGCATTCGAGGCTGAGATGACAAAACAGACCCCACCGGACGTACTTTACATAACTCT
3′CCTAGGCGTAAGCTCCGACTCTACTGTTTTGTCTGGGGTGGCCTGCATGAAATGTATTGAGA

TCACGCCCTAATTGCTATACCAGGATAGAACGGGAGCTTAACCTTGATCGCGCTACGACTAGTG
AGTGCGGGATTAACGATATGGTCCTATCTTGCCCTCGAATTGGAACTAGCGCGATGCTGATCAC

CAGTTGGAAATCGGCCATGTGTATTGCCGCATATGGATCC-3′
GTCAACCTTTAGCCGGTACACATAACGGCGTATACCTAGG-5′
```

Figure 2.1: Sequences of the inserts for HexA, HexB and MS1 (239)

DNA substrates for PBD-conjugates

GXC3

```
5′GATCCATATATGCTTGCATATATCGTTCGATATATGAATGCATATATCGTAAGATATATGGT
3′    GTATATACGAACGTATATAGCAAGCTATATACTTACGTATATAGCATTCTATATACCA

ACCATATATG3′
TGGTATATACCTAG5′
```

GXC2

```
5′GATCCATATATGCTAGCATATATGTATACATATATGGCGCCATATATGTTAACATATATGCT
3′    GTATATACGATCGTATATACATATGTATATACCGCGGTATATACAATTGTATATACGA

AGCATATATG3′
TCGTATATACCTAG5′
```

CXG2

```
5′GATCCATATATCGATCGATATATCATATGATATATCCGCGGATATATCAATTGATATATCGA
3′    GTATATAGCTAGCTATATAGTATACTATATAGGCGCCTATATAGTTAACTATATAGCT

TCGATATATG3′
AGCTATATACCTAG5′
```

Figure 2.2: Sequences of DNA substrates for PBD-conjugates

DNA substrates for hairpin polyamides

LE1

```
5′GATCACGTTATATGGTGATATATGGTCATATGATCACGTTATATGGTGATATATGGTC
3′CTAGTGCAATATACCACTATATACCAGTATACTAGTGCAATATACCACTATATACCAT

ATATGATC3′
TATACTAG5′
```

LE2

```
5′GATCACGTTATATGGTGATATATGGTCATATGATCATATGACCATATATCACCATATAA
3′CTAGTGCAATATACCACTATATACCAGTATACTAGTATACTGGTATATAGTGGTATATT

CGTGATCATATGACCATATATCACCATATAACGTGATC3′
GCACTAGTATACTGGTATATAGTGGTATATTGCACTAG5′
```

Figure 2.3: Sequences of DNA substrates for hairpin polyamides

2.2.1 Preparation of competent cells

E. coli TG2 cells were grown in 5 ml of 2YT media (16 g tryptone, 10 g yeast extract, 5 g NaCl/L) without carbenicillin at 37 °C in a shaker for overnight. 1 ml of the overnight culture was inoculated into 100 ml 2YT media and allowed to grow to an optical density between 0.5 and 0.8 at 600 nm. Media were then transferred to a centrifuge tube and spun at 3000 rpm for 10 minutes at 4 °C. The supernatant was removed; the pellet was suspended in 20 ml transformation buffer and left on ice for 30 minutes. Cells were again centrifuged at 3000 rpm for 10 minutes; the supernatant removed and the pellet was resuspended in 4 ml transformation buffer. Competent cells were stored at 4 °C for a maximum of 2 weeks.

2.2.2 Transformation

5 µl (~100 ng/ml) of plasmid DNA was added to 95 µl of competent cells in a sterile 1.5 ml microcentrifuge tube. The mixture was kept on ice for 30 minutes, placed in a water bath at 42 °C for about 1 minute, promoting the entrance of plasmid into the cells, and kept again in ice for 10 minutes. The transformed cells were plated on agar plates containing 100 µg/ml carbenicillin. The plates were incubated overnight at 37 °C.

2.2.3 Plasmid purification

A colony was picked from the plate using a sterile metal loop and grown overnight at 37 °C in 5 ml 2YT media containing 100 µg/ml carbenicillin. The cultures were then divided between three microcentrifuge tubes and centrifuged at 5000 rpm for 5 minutes to pellet the cells. The supernatant was discarded and the plasmids were prepared using the Qiagen QIAprep kit following the manufacturer's protocol. First the pellets were re-suspended in a total of 250 µl of Buffer P1, which contains RNaseA and then lysed under alkaline condition by addition of 250 µl of Buffer P2. The mixture was then neutralised by addition of 350 µl of Buffer N3. The mixture was then centrifuged at 13000 rpm for 10 minutes. The supernatant was transferred to a Qiagen spin column and centrifuged at 13000 rpm for 1 minute and the supernatant was discarded. Then 500 µl of Buffer PB was added to remove any remaining cellular debris and centrifuged again at 13000 rpm for 1 minute. The supernatant was again discarded and 750 µl of Buffer PE was added and centrifuged twice at 13000 rpm for 1 minute discarding the supernatant in between each centrifuge. This was to remove salt within the solution and any remaining contaminants. Finally, the column was placed in a 1.5 ml microcentrifuge tube and 55 µl of either Buffer EB or water was added; the solution was spun at 13000 rpm for 1 minute to elute the plasmid. The concentration of the plasmid was determined using a Nanodrop 2000c (Thermo Scientific) and was stored at -20 °C for further experiments.

2.2.4 Cloning

The sequences (Figure 2.2 and 2.3) were inserted into pUC18 cloning vector at the BamHI restriction site as described below. The samples were then transformed into *E. coli* and grown overnight on X-gal blue-white selection plates and successful clones were identified as white colonies.

2.2.4.1 *Annealing of oligonucleotides (1:1 mixture)*

Equimolar amounts of the complementary oligonucleotides were annealed by mixing in dH_2O at a final concentration of 10 µM; the samples were incubated for 10 minutes at

95 °C and then left to cool slowly to room temperature.

2.2.4.2 *Preparation of cloning vector*

10 µl of plasmid pUC18 (140 ng/µl) was added to 2 µl of 10x Buffer E (Promega), water (7 µl) and BamHI (1 µl, 10 units). This was then incubated at 37 °C for 2 hours. To precipitate the DNA, 80 µl of absolute ethanol was added and the mixture was chilled on dry ice for 15 minutes. The mixture was then centrifuged at 13000 rpm for 10 minutes; the supernatant was removed and the pellets were dried under speed-vacuum for 10 minutes to remove any residual supernatant. Finally the pellets were dissolved in 20 µl dH_2O and the concentration was determined by Nanodrop spectrophotometer.

2.2.4.3 *Ligation of oligonucleotide to cloning vector*

4 µl of ethanol-precipitated BamHI-digested pUC18 (51 ng/µl) was mixed with 6 µl DNA fragment (~ 1000 ng), 2 µl of 10x ligase buffer, 1 µl of T4 DNA ligase (10 units) and 7 µl dH_2O. The mixture was then left overnight to ligate at room temperature. 200 µl of TG2 competent cells were added to the ligated mixture and the plasmid was transformed as described above (section 2.2.2). 20 µl of the transformation mixture was plated onto agar plates containing X-gal (0.02%, w/v), IPTG (0.1 mM) and carbenicillin (100 µg/ml) and grown overnight at 37 °C.

2.2.4.4 *Blue-White selection*

From the agar plates grown overnight, white colonies were picked and grown in 2YT sterile media (5 ml) containing carbenicillin (100 µg/µl) overnight at 37 °C with constant agitation. The plasmid was extracted and purified as described in section 2.2.3.

2.2.4.5 *Sequencing*

The sequencing of cloned inserts was determined by commercial sequencing by Eurofins MWG Operon (The sequences are given in Appendix 1, Figure A-E).

2.2.5 Radiolabelling of DNA fragments

Radiolabelling of DNA fragments was done in two different approaches based on the source of the DNA (synthetic oligonucleotides or DNA restriction fragments). For footprinting experiments, the double stranded DNA restriction fragments were labelled at the 3´-end by attaching a radioactive phosphorous from [α-^{32}P]-dATP using reverse transcriptase (footprinting experiments of Chapter 3) or Klenow fragment (Chapter 4 and 6). Single stranded synthetic oligonucleotides were labelled at the 5´-end by using [γ-^{32}P]-ATP and polynucleotide kinase. Here, alpha and gamma indicate the position of radioactive phosphorous in the ATP (Figure 2.4).

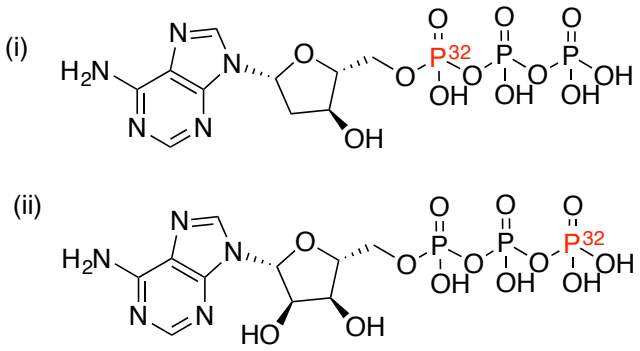


Figure 2.4: Alpha and gamma labelling positions; radioactive phosphorous is shown in red colour. i) [α-^{32}P]-dATP ii) [γ-^{32}P]-ATP

2.2.5.1 *3´ labelling*

Radiolabelling at the 3´-end involves the phosphorylation of 3´-OH of the nucleotide by dATP, containing radioactive phosphorous at the alpha position. The adenosine is attached by filling in the sticky ends from restriction enzyme cleavage (A/AGCTT for *Hind*III or G/AATTC for *EcoR*I) opposite the thymines in the complementary strand.

For labelling of the HexAfor, HexArev, MS1 and MS2 sequences, 5 µl of Promega buffer E, 10 units of *Hind*III and 10 units of *Sac*I were added to 45 µl of the plasmid preparation (as described above) containing the DNA fragment. Since HexBfor and HexBrev contain internal HindIII sites they cannot be labelled in this way and so were obtained by cutting with *EcoR*I and *Pst*I. For these fragments, 45 µl of plasmid were mixed with 5 µl Promega Buffer H, 10 units *Pst*I and 10 units *EcoR*I. Both the mixtures were incubated at 37 °C for 6 hours. After digestion with these enzymes the DNA was precipitated by addition of 180 µl ethanol and the mixture was left on dry ice overnight. The DNA was recovered by centrifuging at 13000 rpm for 15 minutes; the supernatant was removed, again washed with 70% ethanol. Finally, the pellet was dried in a speed vac. This was dissolved in 45 µl of water and 5 µl of reverse transcriptase buffer, 1 µl of [α-^{32}P]-dATP and 10 units of AMV reverse transcriptase or Klenow fragment were added; incubated in 37°C for 2-3 hours.

2.2.5.2 *5´ labelling*

Radiolabelling at 5´-end of the oligonucleotides involves the attachment of radioactive phosphate at the 5´-position (Figure 2.5). Here 1 µl of 10 µM each oligonucleotide was incubated with 2 µl PNK buffer (Promega), 10 units polynucleotide kinase, 1 µl of [γ-^{32}P]-ATP and 13 µl of dH_2O at 37°C for 2 hours.

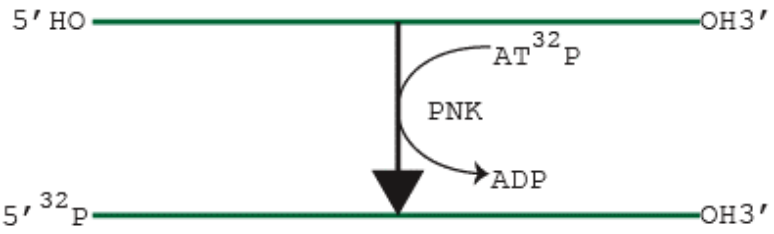


Figure 2.5: Attachment of phosphate at the 5´-end of an oligonucleotide. Solid green line indicates oligonucleotide sequence

2.2.5.3 *Purification of radiolabelled fragments*

After labelling, the desired radiolabelled DNA sequence was purified from the enzymes and other components of the mixture by polyacrylamide gel electrophoresis. For 3´ labelling, a 8% non-denaturing polyacrylamide gel (20 cm long and 0.3 mm thick) was used. 30 µl of Ficoll dye (20% Ficoll, 10 mM EDTA and 0.01% w/v bromophenol blue) was added to the labelling mixture before loading onto the gel. The gel was run at 400 V for 1 hour. For 5´ labelling of synthetic oligonucleotides, 10 µl of DNase I stop solution (80% formamide, 10 mM EDTA, 1 mM NaOH and 0.01% w/v bromophenol blue) was added to the labelled DNA mixture. The mixture was then heated at 95°C for 3 minutes and immediately crash cooled into the ice. The sample was then loaded onto a 14% denaturing polyacrylamide gel (containing 8M urea) and run at 800 V for about 50 minutes.

After completion of electrophoresis, the gel was exposed to X-ray film for about 5 minutes to identify the location of the radiolabelled DNA, followed by excision of the band. The excised band was immersed in 400 µl of elution buffer containing 10 mM Tris-HCl pH 7.5 and 0.1 mM EDTA and slowly agitated overnight. After elution the DNA was precipitated by adding 1000 µl of ethanol and kept on dry ice for 2 hours. The sample was then centrifuged at 13000 rpm for 15 minutes. The supernatant was removed and checked with a hand held Geiger-Muller counter to ensure that it contained no radioactivity. The pelleted DNA was washed with 70% ethanol, centrifuged again at 13000 rpm for 2 minutes and dried in a speed vac for maximum 5 minutes. Finally the dried pellets were dissolved in 10 mM Tris-HCl pH 7.5 containing 0.1 mM EDTA so as to give about 10 cps per microlitre of the radiolabelled DNA. For the 5´-labelled single-stranded oligonucleotides, 120 µl of 3M NaOAc was added to the precipitation mixture and the DNA pellet was washed with 80% ethanol instead of 70% ethanol.

2.2.6 DNA footprinting

2.2.6.1 *Incubation of DNA and ligand*

Ligands were diluted from 10 mM stocks in DMSO into 10 mM Tris-HCl pH 7.5 containing 10 mM NaCl, immediately before use. 1.5 µl ligand solution was mixed with 1.5 µl radiolabelled DNA and incubated at room temperature for various times to allow binding. In addition to the ligand samples, a control sample was also prepared by combining 1.5 µl radiolabelled DNA with 1.5 µl Tris-HCl pH 7.5 containing 10 mM NaCl.

2.2.6.2 *DNase I digestion*

After incubation, the template DNA was cleaved with DNase I. The final enzyme concentration was typically 0.01 units/ml and was obtained by diluting 2 µl of stock DNase I (7200 units/ml) into 1 ml of 20 mM NaCl, 2 mM $MgCl_2$ and 2 mM $MnCl_2$; 4 µl of this first dilution were then diluted into 1 ml of the same buffer. For cleavage with DNase I, 2 µl of the diluted enzyme was added to each of the samples and allowed 1 minute for digestion. The digestion was stopped by adding 3.5 µl of DNase I stop solution (10 mM EDTA, 1 mM NaOH, 0.01% bromophenol blue, 80% Formamide). The samples were heated at 100°C for 3 minutes followed by rapid cooling in ice prior to electrophoresis.

2.2.6.3 *GA track marker*

The GA marker indicates the positions of guanines and adenines within the target DNA sequence by cleaving only at the purine sites, thus the sequences in the footprinting sites can be identified. The marker was prepared by mixing 1.5 µl of radiolabelled DNA, 5 µl of DNase I stop and 20 µl of dH_2O. The mixture was heated at 100°C for 20 minutes with the cap open to allow evaporation followed by cooling on ice.

2.2.6.4 *Polyacrylamide gel electrophoresis*

The products of DNase I digestion of the HexAfor, HexArev, HexBfor, HexBrev, MS1 and MS2 DNA fragments were run on 40 cm long, 0.3 mm thick 10% denaturing polyacrylamide gels containing 8M urea. The gel was run at 1500 V for about 2 hours until the dye reached the bottom of the gel. The plates were separated and the gel was soaked in 10% acetic acid (v/v) for 10 minutes. The gel was transferred to 3 MM Whatman paper and dried at 80 °C for 1 hour. The dried gel was then exposed to a phosphoimager screen overnight for scanning.

2.3 Fluorescence melting studies

2.3.1 Oligonucleotide sequences

Synthetic oligonucleotides designed for fluorescence melting studies were obtained with a fluorophore (F) and a quencher (Q) attached at the 5'-end and 3'-end respectively; here, FAM was used as fluorophore whereas dabcyl was used as quencher.

1. Oligonucleotides containing two consecutive guanines

5′-F-AAAAGGTACCAAAA-3′
3′-Q-TTTTCCATGGTTTT-5′

5′-F-AAAAGGATCCAAAA-3′
3′-Q-TTTTCCTAGGTTTT-5′

2. Oligonucleotides containing two guanines separated by three other bases

5′-F-AAAAGCTTGCAAAA-3′
3′-Q-TTTTCGAACGTTTT-5′

5′-F-AAAAGAATGCAAAA-3′
3′-Q-TTTTCTTACGTTTT-5′

5′-F-AAAAGCTAGCAAAA-3′
3′-Q-TTTTCGATCGTTTT-5′

3. Oligonucleotides containing one guanine

5′-F-AAAAGTATACAAAA-3′
3′-Q-TTTTCATATGTTTT-5′

5′-F-AAAAGTTAACAAAA-3′
3′-Q-TTTTCAATTGTTTT-5′

2.3.2 Preparation of samples

For melting studies, samples (20 µl) were prepared containing 0.25 µM duplex DNA, dissolved in 10 mM sodium phosphate pH 7.4 containing 100 mM NaCl, with varying concentrations of the ligand. Six different ligand concentrations (10 µM, 5 µM, 2 µM, 1 µM, 0.5 µM and 0.1 µM) were used and incubated for 24 hours. Samples were then heated from 30 °C to 95 °C at 0.1 °C/sec, left at 95 °C for 5 minutes and then re-annealed by cooling to 30 °C, held at this temperature for 5 minutes before heating again to 95 °C (Figure 2.6).

Fluorescence melting curves were performed in Roche LightCycler, using a total reaction volume of 20 µl. The LightCycler (version 1.5) is an instrument to perform real-time PCR combined with ultra rapid cycling. The machine has one excitation source (488 nm) and for our studies, the fluorescence emission *i.e.* the changes in fluorescence were measured at 520 nm. Recordings were taken during both the heating and annealing phases and the T_m values were determined from the first

derivatives of the melting profiles. Before examining the effect of different ligand concentrations, control sample (without any ligand) and a sample with highest ligand concentration (10 μM) were run in triplicate and the T_m values usually differed by <0.5 °C that indicate the accuracy of succeeding experiments.

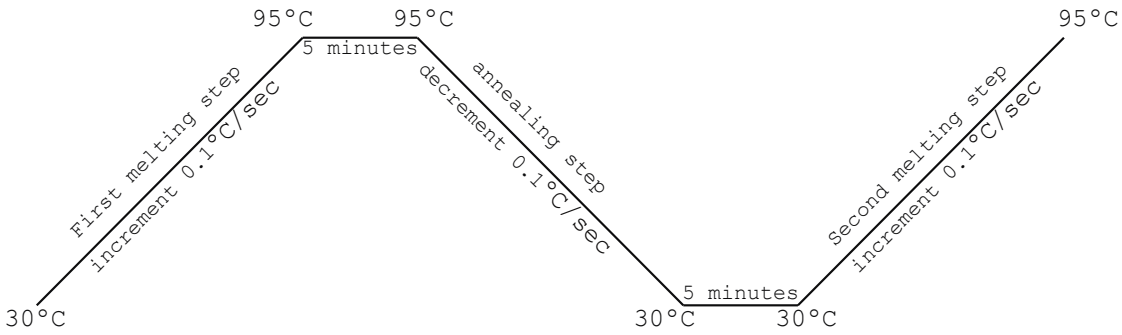


Figure 2.6: Steps of fluorescence melting studies

2.4 Ligation assay

2.4.1 Oligonucleotide sequences

Oligonucleotides were radiolabelled at 5´-end with γ-^{32}P-ATP as described in section 2.2.5.2. The radiolabelled oligonucleotides were stored at -20 °C in 10 mM Tris-HCl pH 7.5 containing 0.1 mM EDTA.

The oligonucleotides used in these experiments can be categorized into three groups as follows-

1. Self-complementary strands having two (GXXC) binding sites:

```
Oligo-1: 5’GATCAATGTGTACACATT3’
             3’TTACACATGTGTAACTAG3’

Oligo-2: 5’GTACAATGTGTACACATT3’
             3’TTACACATGTGTAACATG5’

Oligo-3: 5’AGATATATCTATGATCAT3’
                    3’TACTAGTATCTATATAGA5’
```

2. Binding sites (GXXC) in the middle of strands with complementary strands in pair:

```
Oligo-4: 5’ATCATATATGATCTATACTA3’
Oligo-5:       3’ATACTAGTATAGATTAGTAT5’

Oligo-6: 5’TAATTAATATGTACTA3’
Oligo-7:       3’TATACATGATATTAAT5’

Oligo-8: 5’ATACTAATGAATCATTAA3’
```

```
Oligo-9:       3'TACTTAGTAATTTATGAT5'
```

3. Binding sites at overhang with complementary strands in pair-

```
Oligo-10: 5'GATCATATAATAAATA3'
Oligo-11:     3'TATATTATTTATCTAG5'

Oligo-12: 5'GTACATATAATAAATA3'
Oligo-13:     3'TATATTATTTATCATG5'
```

2.4.2 Preparation of samples

A 50 μM concentration of ligand (SJG-136) was used for these experiments. The ligand was dissolved in 10 mM Tris-HCl pH 7.5 containing various concentrations of NaCl and $MgCl_2$. Then 1.5 μl of each complementary oligonucleotide (radiolabelled) was mixed with 3.0 μl of the ligand. 4 μl of Ficoll dye was added to each sample before electrophoresis. The samples were incubated for definite period. In some experiments, the total DNA concentration was increased by adding unlabelled oligonucleotides as well. The ligand-oligonucleotide complexes were incubated for 24 hours at room temperature.

2.4.3 Polyacrylamide gel electrophoresis

The samples were run on a 40 cm long, 0.3 mm thick 16% non-denaturing gel. The gel was run at 800V until the dye run about three quarters of the gel. The plates were separated and the gel was soaked with 10% acetic acid (v/v) for 10 minutes. The gel was transferred to 3 MM Whatman paper and dried at 80 °C for 1.5 hours. The dried gel was then exposed to a phosphoimager screen overnight for scanning.

3 PBD conjugates with pyrrole, imidazole and thiazole

3.1 Introduction

The pyrrolobenzodiazepines (PBD) are a group of naturally occurring antitumor antibiotics and show chemotherapeutic activity, which is thought to arise from the binding of N10 of the benzodiazepine ring with C2 of guanine. The basic structure of PBDs is a tricyclic system consisting of anthranilate, diazepine and a pyrrolidine ring. Due to their potential as anticancer agents, their basic structure has been modified in several different ways. Numerous PBD compounds have been synthesized such as PBD dimers produced by joining two PBD units with a suitable linker (168,172,175,177,179,180). Another novel approach is the formation of PBD conjugates by attaching heterocyclic rings, such as pyrrole, imidazole or thiazole linked together and to the PBD nucleus by amide bonds. The rationale for appending these rings is due to their ability to recognise specific DNA sequences. As explained in the general introduction, these polyamides can bind as antiparallel dimers in the DNA minor groove, such that pyrrole-pyrrole pairs recognise AT or TA whereas pyrrole-imidazole and imidazole-pyrrole bind with CG and GC respectively. Benzofused rings such as benzothiophene, benzofuran and indole have also been used for conjugate formation with PBDs; though much less is known about the binding specificity of these rings. Benzofused rings are present in many natural and synthetic compounds and have shown antitumor activity, thus their conjugation with PBD can be an effective way to enhance the sequence selectivity of PBD (Chapter 4).

3.2 Experimental design

3.2.1 Footprinting studies

The DNA sequence selectivity of PBD conjugates with pyrrole, imidazole and thiazole was first examined using a relatively high concentration (10 μM) in footprinting experiments, to obtain an initial impression of which compounds bind to DNA. Compounds that produced footprints at 10 μM were then examined individually over a range of concentrations from 10 μM to 0.5 μM. Samples were prepared by mixing 1.5 μl of each ligand (diluted in 10 mM Tris-HCl pH 7.5 containing 10 mM NaCl) with 1.5 μl of radiolabelled DNA and were incubated at room temperature for 24 hours. Beside footprinting studies over a range of ligand concentrations, the binding kinetics of each ligand were examined by performing footprinting with a fixed concentration at different time intervals. For these kinetic studies, two different concentrations were chosen (10

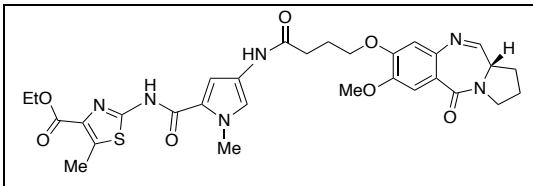
μM and 5 μM) with incubation periods from 1 minute to 24 hours. For these experiments a stock solution was prepared by mixing the ligand (7.5 μl) with radiolabelled DNA (7.5 μl), removing 3 μl at each time interval for digestion with DNase I.

3.2.2 Melting studies

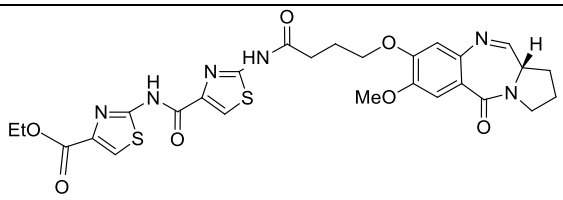
For melting studies, samples (20 μl) were prepared containing 0.25 μM duplex DNA, dissolved in 10 mM sodium phosphate pH 7.4 containing 100 mM NaCl, with varying concentrations of the ligand. Six different concentrations (10 μM, 5 μM, 2 μM, 1 μM, 0.5 μM and 0.1 μM) were used and incubated for 12 hours. Samples were then heated from 30 °C to 95 °C at 0.1 °C/sec, left at 95 °C for 5 minutes and then re-annealed by cooling to 30 °C, held at this temperature for 5 minutes before heating again to 95 °C. Recordings were taken during both the heating and annealing phases.

3.3 Chemical structures of PBD-conjugates with pyrrole, imidazole and thiazole ring

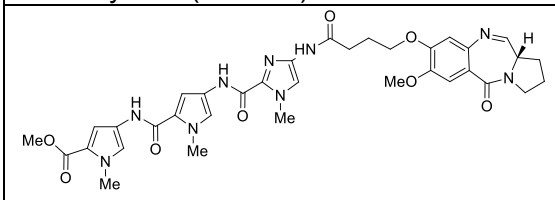
The compounds were synthesized in Dr Federico Brucoli's laboratory in University of Reading. Here, pyrrole, imidazole and thiazole are abbreviated as 'Py', 'Im' and 'Thz' respectively.



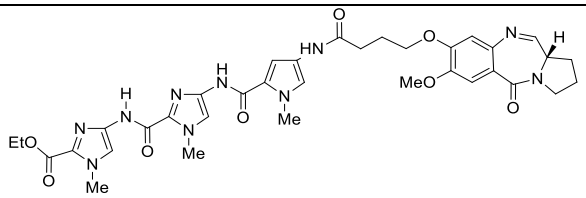
PBD-Py-Thz (RMH35)



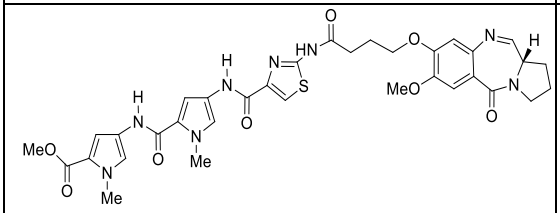
PBD-Thz-Thz (RMH71)



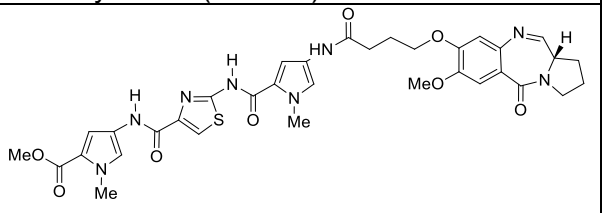
PBD-Im-Py-Py (RMH41)



PBD-Py-Im-Im (RMH53)



PBD-Thz-Py-Py (RMH43)



PBD-Py-Thz-Py (RMH47)

3.4 Results

3.4.1 Footprinting studies

The sequence selectivity of these PBD conjugates was studied using DNA fragments HexA and HexB (which together contain all 64 symmetrical hexanucleotide sequences) and MS1 (which contains all 136 tetranucleotide sequences).

Figure 3.1 (Left panel) shows footprinting gels for the interaction of PBD conjugates (10 μM) with DNA fragments of HexAfor and HexBrev. From the gels, it is evident that compounds RMH41, RMH43 and RMH47 produce large footprints in both HexAfor and HexBrev. Compound RMH53 shows several footprints in HexBrev but less in HexAfor. All these compounds contain three rings conjugated with PBD. Weaker footprints are seen in HexAfor and HexBrev for RMH35 (thiazole and pyrrole) whereas RMH71 (two thiazole rings) protects only in HexBrev. The cleavage protection by these compounds at different concentrations is described in the following paragraphs.

3.4.1.1 *PBD-Py-Thz (RMH35) and PBD-Thz-Thz (RMH71)*

Figure 3.1 (Right panel) shows DNase I footprints for the interaction of compounds PBD-Py-Thz (RMH35) and PBD-Thz-Thz (RMH71) with HexAfor and HexBrev. These compounds, which only contain two conjugated rings, do not greatly affect the DNase I cleavage pattern. However, both ligands produce two sites of protection at the highest ligand concentration (10 μM). These are most clearly seen with HexBrev, for which the sites are closest to the middle of the gel and are located around the sequence AATTAGGGC (site 1) and GCAATTGCCGG (site 2). Protection is observed with HexAfor only by RMH35, though this is higher up the gel and therefore less well resolved, but appears to be around the sequence GAATTCCGG only. These sites are very similar, and contain AT rich sequences at the 5'-end and GC rich regions at the 3'-end.

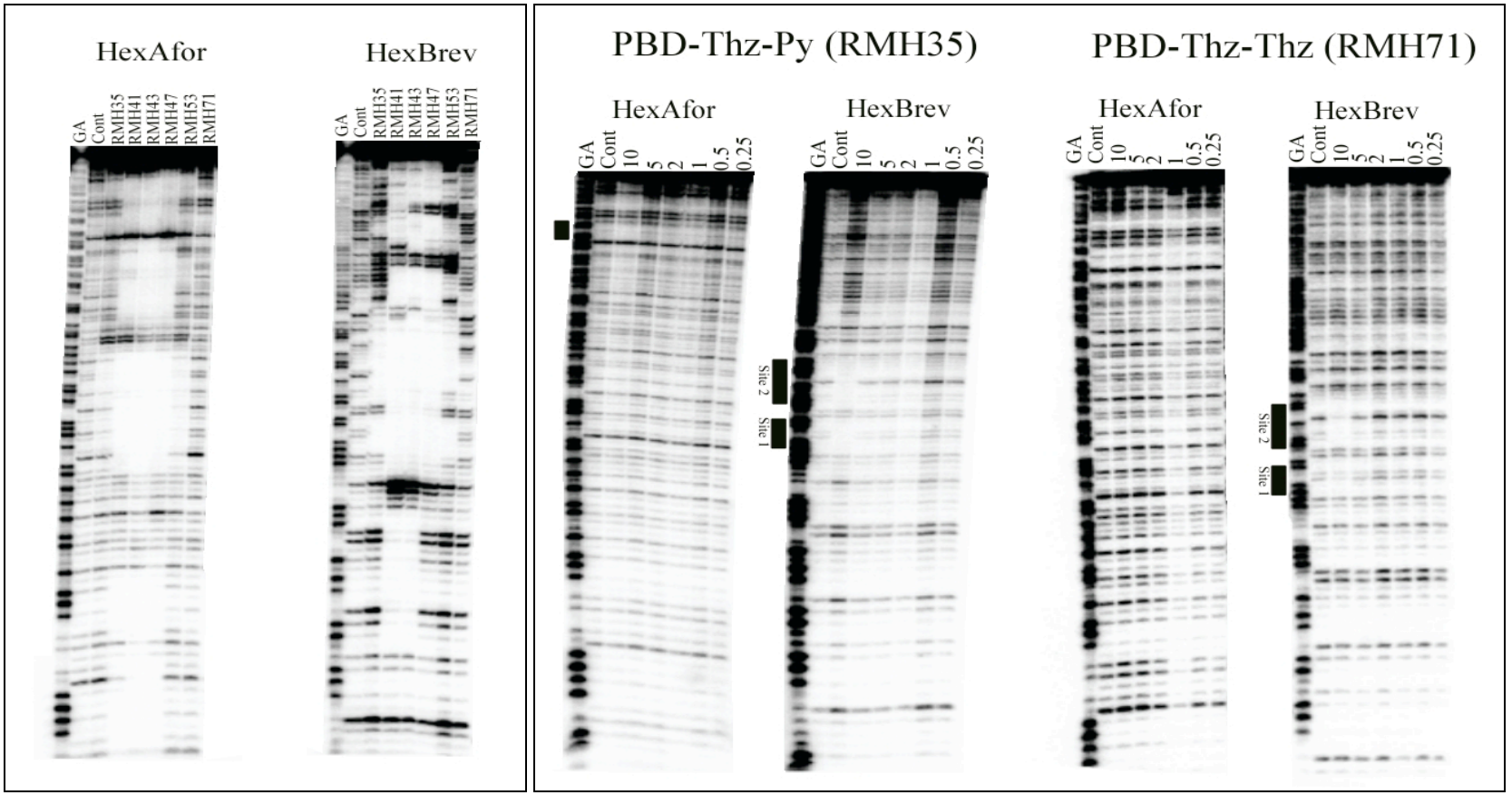


Figure 3.1: DNase I footprinting gels showing the interaction of PBD-conjugates (10 μM) (Left panel) and various concentrations of RMH35 and RMH71 (Right panel) with HexAfor and HexBrev. The samples were incubated with the ligands at room temperature for 24 hours before digesting with DNase I. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands. Filled bars indicate the location of the footprints.

3.4.1.2 *PBD-Im-Py-Py (RMH41)*

Figure 3.2 shows DNase I footprints for the interaction of compound PBD-Im-Py-Py (RMH41) with HexAfor, HexArev, HexBfor, HexBrev, MS1 and MS2. The sequences protected from cleavage in these DNA fragments are summarised in Figure 3.3. In DNA fragment HexA, several footprints are observed with 10 μM with only one site with 5 μM. Protections are found at 5'-GGATCC (site 1), 5'-GGTACCTTGGCC (site 3) and 5'-TAAACGTT (site 4) in HexAfor; these sites are also protected in HexArev (sites 1, 5 and 4 respectively) showing an additional protection at 5'-CATATATC (site 2). One large footprint spanning 20 bp is found in both HexAfor (site 5) and HexArev (site 3). Another large footprint covering 21 bp is observed in HexAfor (site 2) with 10 μM that becomes shortened to 5'TTAAGTACTAG with 5 μM; in HexArev, that region (site 6) is only protected with 10 μM.

The footprints with HexB are more pronounced than with HexA and some protections are still observed 1 μM ligand. Very large footprints spanning as much as 30 bp are found with 10 μM, though these are resolved into smaller sites with lower ligand concentrations. In HexBfor, the footprint at 5'-TACGTATACGCGTACGCGC (site 4) becomes shortened to 5'-TACGTATACGC when the concentration is decreased from 10 μM to 5 μM. Protections are also observed with HexBfor at 5'-TACATATGTA (site 3) and 5'-TGTCGAC (site 2). One large footprint 5'-GATCAATATTCGAATTAATGCATGGA (site 1) observed with 10 μM is divided into two smaller sites with 5 μM at 5'-GATCAATAT and 5'-GAATTAATGCATGGA; on lowering the ligand concentration to 2 μM only the lower (3') site is protected and this is shorter by a few bases at the 3'-end. With HexBrev, footprints at 5'-ACATATGTA (site 4), 5'-GTATACG (site 3) and 5'-AATTAG (site 2) are observed with 2 μM; protection is also seen at 5'GAGCTCGCGATCGGCCGGA (site 1) with 10 μM that is truncated to 5'-GCTCGCGTACGGC with 5 μM.

For MS1, protections are seen with 5 μM with two sites still protected with 2 μM. Some of these sites are found in MS2 but only with 10 μM ligand. Protections with 2 μM are observed at 5'GTACTTTACATAACTC (site 5) and 5'AATTGCTA (site 4) in MS2; for MS1, footprints are found at 5'-AGCAATTAGGGG (site 3), 5'-AAGAGTTATGTAAAGTACG (site 2) and 5'-TTGTCATCTCAGCCTCGAATGCG (site 1).

Footprints such as 5'-GAATTC (site 5), 5'-GGTACC (site 3) and 5'-GCTAGC (site 2) in HexAfor, 5'-GCAAGC (site 5) in HexBfor and 5'-GGATCC in HexAfor (site 1) and MS2 (site 1) can be grouped as 5'-GX(A/T)(A/T)XC and 5'-CAATTG (site 5/HexBfor) and 5'-

CATATG (site 3/HexBfor and site 1/MS2) as 5'-CX(A/T)(A/T)XG. Protections are also visible at AT-rich locations at 5'(G/C)-$(A/T)_{4,6}$-(G/C) such as 5'-CTTTAC (site 5), 5'-CTTAAC (site 3), 5'-CTATAC (site 4) and 5'-GTATTG (site 1) in MS2, 5'-CAATATTC (site 1/HexBfor) and 5'-GTTTAAAC (site 4) and 5'CATATATC (site 2) in HexArev.

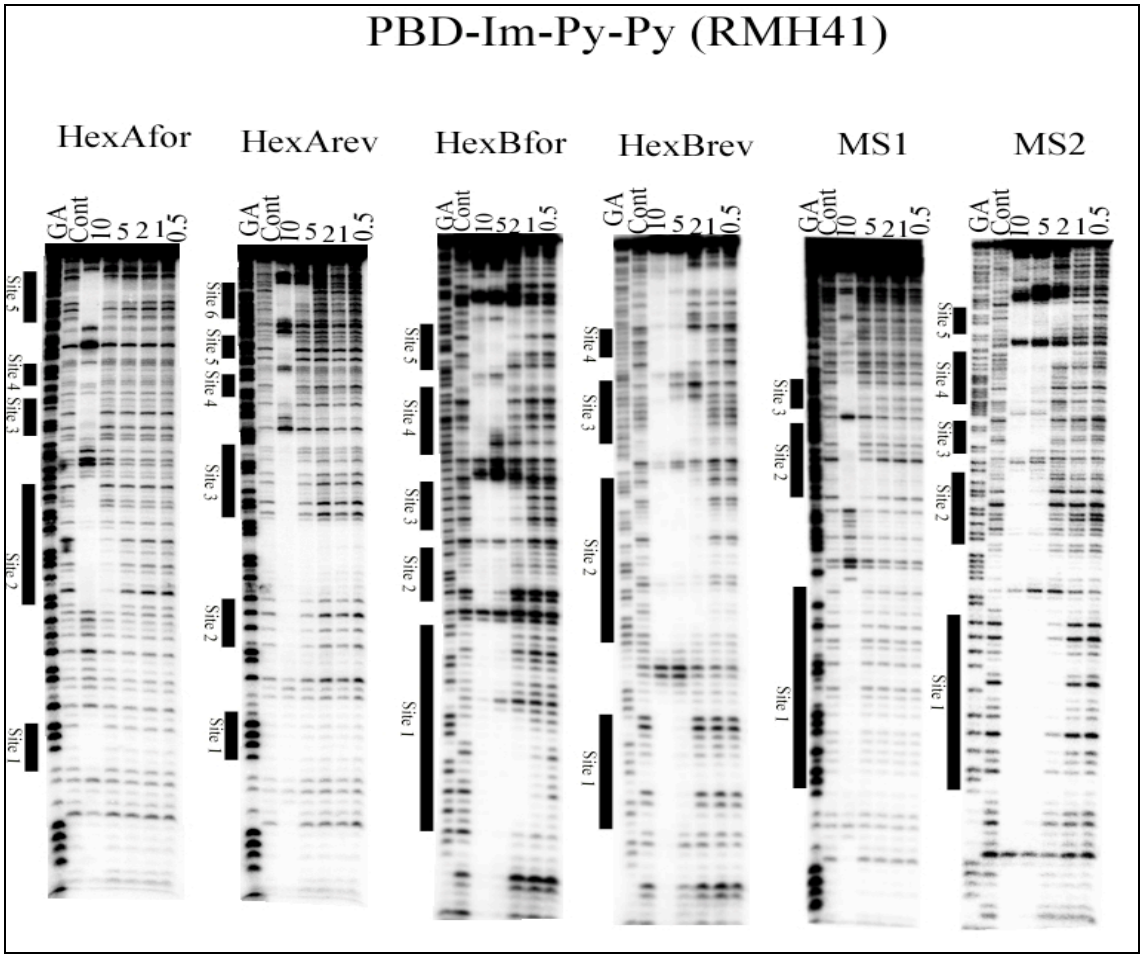


Figure 3.2: DNase I footprinting gels showing the interaction of various concentrations of PBD-Im-Py-Py (RMH41) with HexAfor, HexArev, HexBfor, HexBrev, MS1 and MS2. The samples were incubated with the ligand at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled "GA" correspond to markers specific for purines, while "Cont" indicates DNase I cleavage in the absence of added ligands. Filled bars indicate the location of the footprints.

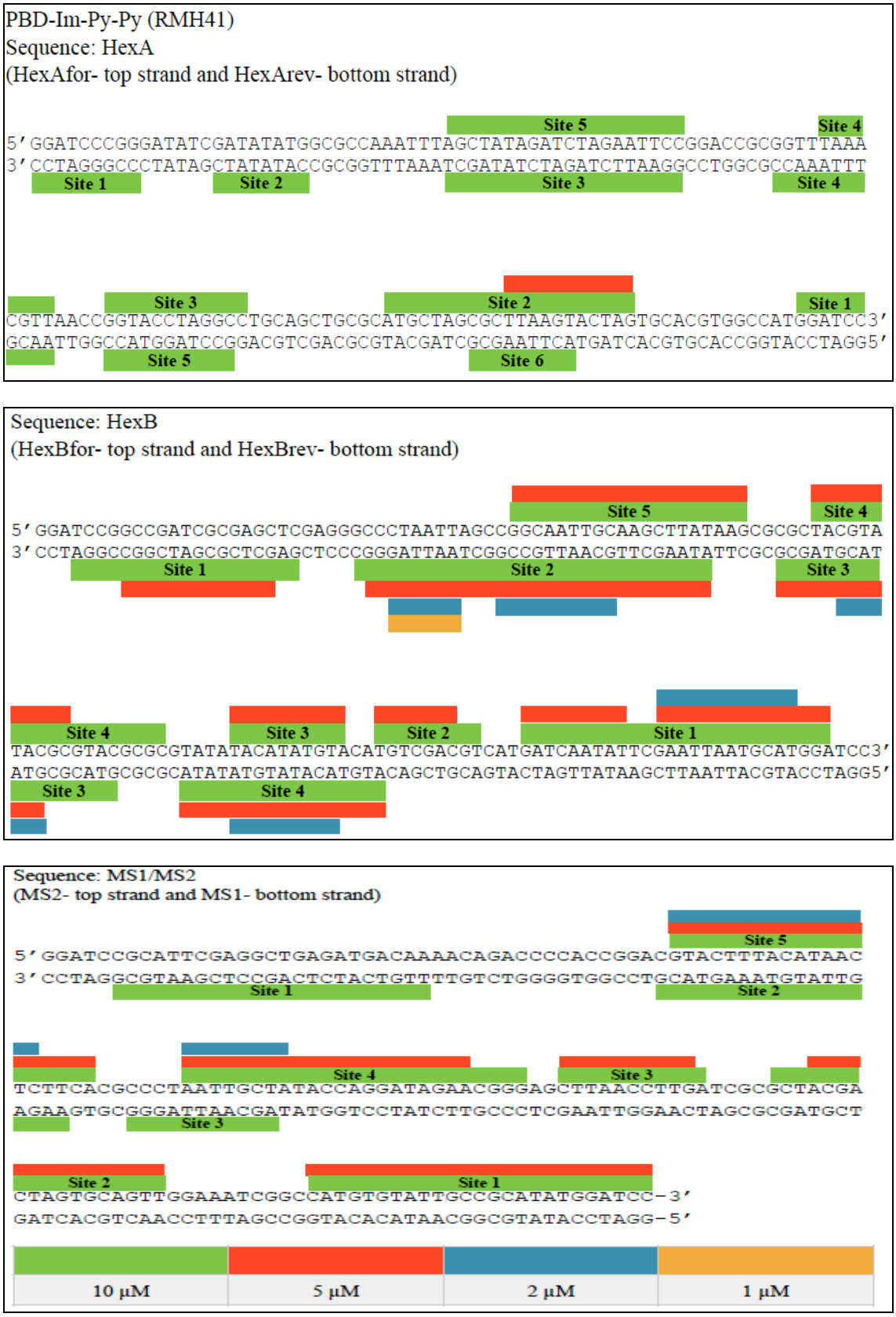


Figure 3.3: Sequences of the footprinting substrates indicating the regions protected by PBD-Im-Py-Py (RMH41) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand.

3.4.1.3 *PBD-Py-Im-Im (RMH53)*

DNase I footprinting gels showing the interaction of PBD-Py-Im-Im (RMH53) with HexAfor, HexArve, HexBfor, HexBrev, MS1 and MS2 are shown in Figure 3.4 and the protected sequences are summarised in Figure 3.5. With HexAfor, two large footprints are observed with 10 μM; on decreasing the ligand concentration to 5 μM, one of these (site 2) splits into two protections at 5'-AGAATTCCGGA and 5-'GTTAACCG, while the other one is shortened to 5'-ATGCTAGCGCTTAAGTAC (site 1). In HexArev, protections are observed at 5'-AGGCCTAG (site 5), 5'-GGTTAACGTTT (site 4), 5'-CCGGAA (site 3) and 5'-CTAGA (site 2) and 5'ATATCGATATCCCG (site 1) only with 10 μM; another footprint at the top of the gel protecting 5'-ACTTAAGCGCTAGCATG (site 6) is only evident with highest concentration. With HexBfor and HexBrev, the binding sites are again smaller than those seen with HexA.

In HexBfor, 10 μM ligand protects two large footprints at 5'-AATTGCAAGCTTATAAGCG (site 5) and 5'-TACATATGTACATGT (site 3), which resolve into two footprints around 5'-AATTGCAAG and 5'-TACATATGTA with 5 μM. These two sites are also protected in HexBrev (site 2 and 4 respectively). Footprints are also found at 5'-TTAG (site 6) and 5'-ATATTC (site 2) for HexBfor and 5'-AATTAG (site 1) with HexBrev; and the protection at 5'-AATTAG (site 1) is evident even with 2 μM.

Several footprints that persist to 2 μM ligand are observed with MS1/MS2. In MS2, these are located around 5'-ACATAACTC (site 5), 5'-AATTGCT (site 4), 5'-AACCTT (site 3) with 5 μM. A large footprint is observed at 5'-CCGCATATGGA (site 1) with 10 μM ligand, which divides into two regions of protection- 5'-CCG and 5'-ATGGA at 5 μM. Adjacent to this large footprint, another small footprint (5'-GTAT/site 2) is seen at 5 μM. For MS1, a few larger footprints are seen at 5'-AAGGTTAAGCT (site 5), 5'-ATAGCAATTAGGG (site 4), 5'-TTGTCATCTCAGC (site 2) and 5'-CGAATGCG (site 1) with 10 μM. On lowering the ligand concentration to 2 μM, footprints are evident at 5'-GGTTAAGCT, 5'-AATTAGGG, 5'-TTGTCAT and 5'CGAATGC. A 20 bp footprint (site 3) is also observed with 10 μM, which is shortened to 14 bp with 2 μM ligand protecting at 5'-TTATGTAAAGTACG.

On examining the footprints with these DNA fragments, some common sequences are observed. For example, the larger footprints contain 5'-GTTAAC (site 2), 5'-GCTAGC (site 1) and 5'-GCATGC (site 1) of HexAfor, 5'-GCTTGC (site 2) of HexBrev. 5'GAATGC (site 1) and 5'GTCATC (site 2) of MS1 (*i.e.* with the common sequence 5'-GXXXXC), while footprints are also seen at 5'-CTTAAG (site 1/HexAfor), 5'-CAATTG (site 2/HexBrev), 5'-GTTATG (site 3) and 5'GTAAAG (site 3) of MS1 (*i.e.* 5'-(G/C)-$(A/T)_4$-(G/C)).

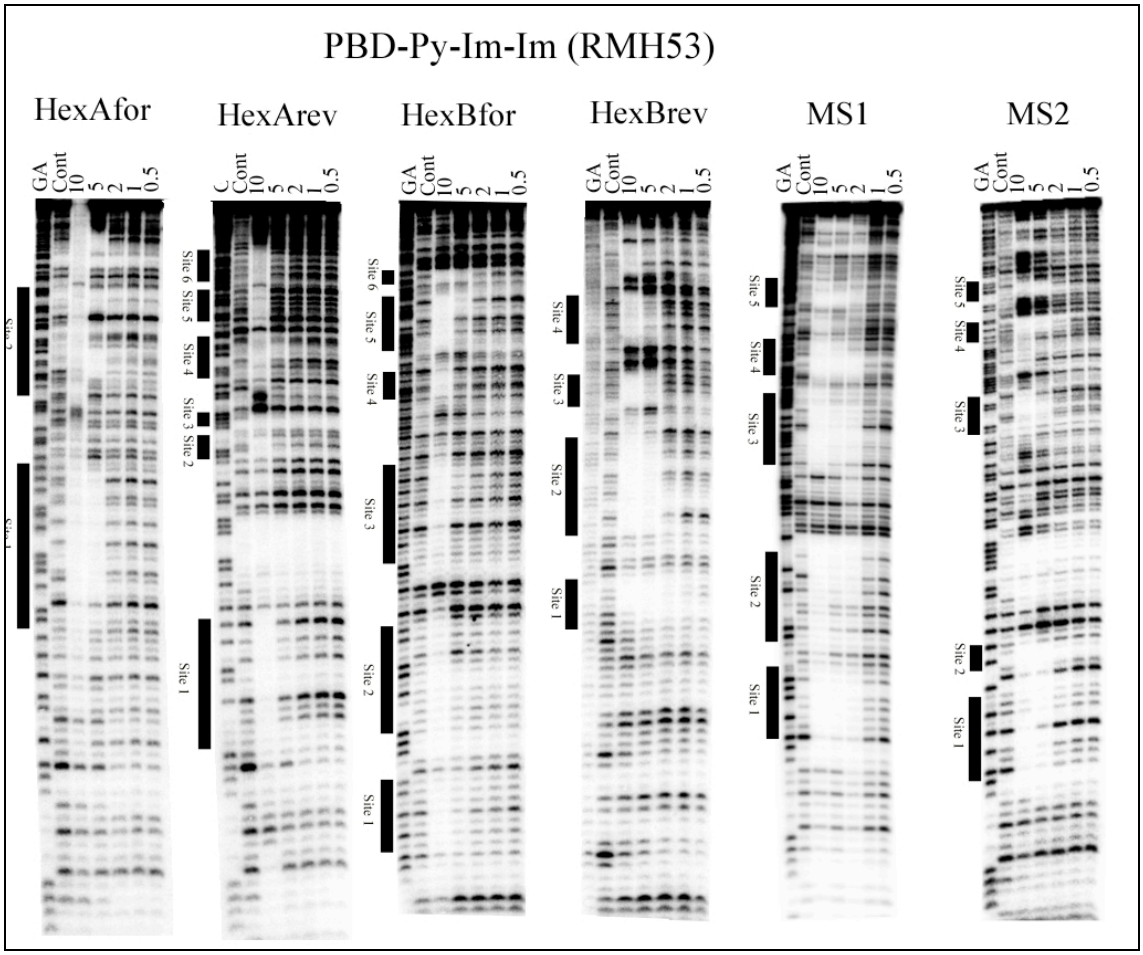


Figure 3.4: DNase I footprinting gels showing the interaction of various concentrations of PBD-Py-Im-Im (RMH53) with HexAfor, HexArev, HexBfor, HexBrev, MS1 and MS2. The samples were incubated with the ligand at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands. Filled bars indicate the location of the footprints.

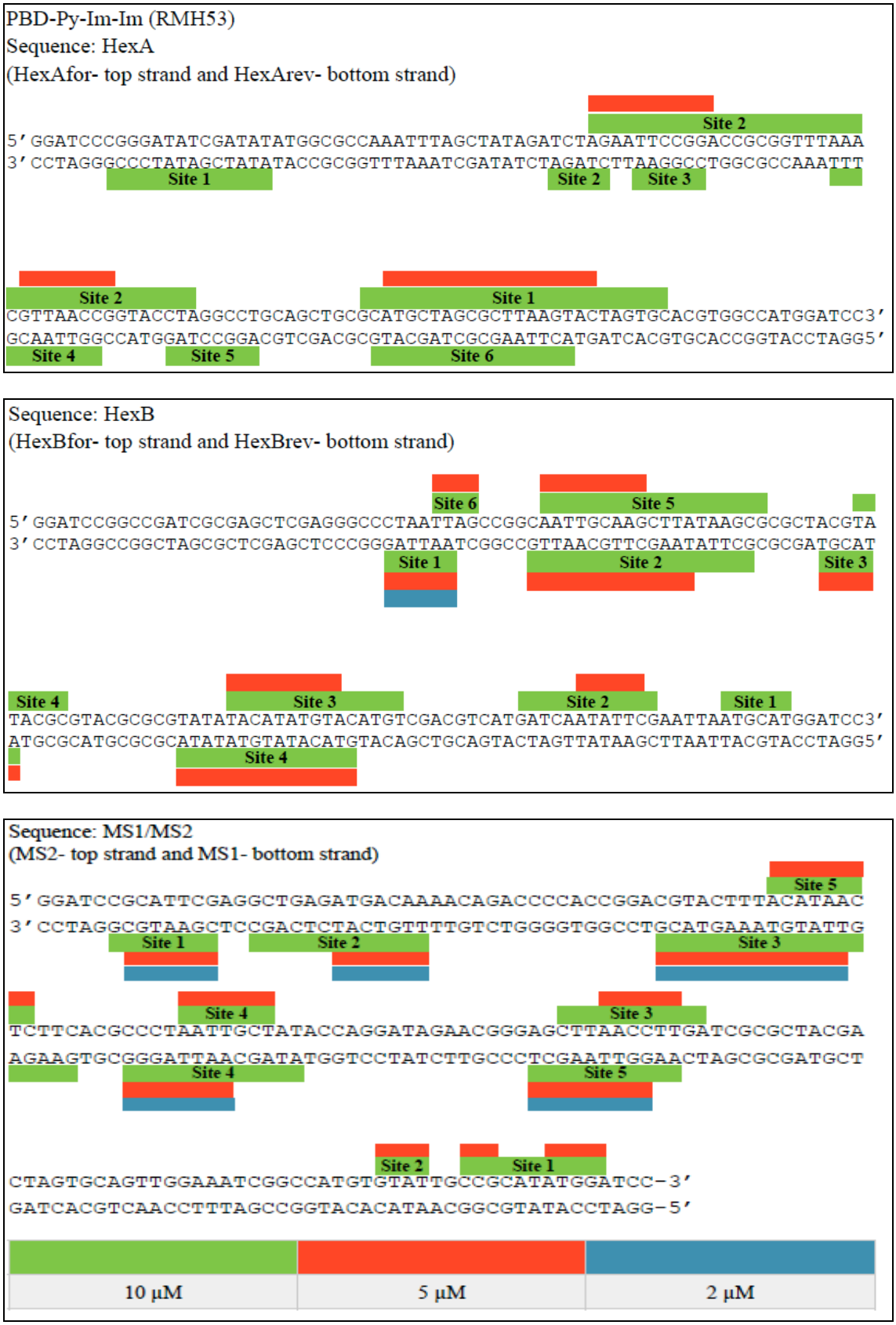


Figure 3.5: Sequences of the footprinting substrates indicating the regions protected by PBD-Py-Im-Im (RMH53) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand.

3.4.1.4 *PBD-Thz-Py-Py (RMH43)*

Figure 3.6 shows DNase I footprints for the interaction of compound PBD-Thz-Py-Py (RMH43) with HexAfor, HexBfor, MS1 and MS2, and the sequences protected from cleavage in these DNA fragments are summarised in Figure 3.7. In HexAfor, three large footprints are observed at 5'-TATAGATCTAGAATTCC (site 4), 5'-GGTTTAAACGTTAACCGGTACCTAGG (site 3) and 5'-ATGCTAGCGCTTAAGTACT (site 2) with 10 μM; the first of these binding sites is still protected with 2 μM, while the other two sites become smaller at 5'-AAACGTT and 5'-TAAGTAC with footprints that persist to 1 μM and 0.5 μM respectively. Another small protection at 5'-GGATCC (site 1) is seen with 5 μM ligand.

With HexBfor, footprints are produced at 5'-AATTGCAAGCTTATAA (site 5), 5'-GTATACGCG (site 4) and 5'-ACATATGTACATG (site 3) with 10 μM and 5 μM ligand. Another footprint with 10 μM at 5'-TGATCAATAT (site 2) is only protected at the AT rich regions at the 3'end when the concentration is lowered to 5 μM. The footprint adjacent to this site at 5'-CGAATTAATG (site 1) is truncated at 2 μM and only protects the 5'-end at 5'-TTAATG.

With MS1/MS2, footprints are observed at the 3'-end of the DNA fragments (*i.e*. bottom of the gel) at 5'-GAATGCG (site 1), 5'-TTGTCATCTCA (site 2) for MS1 and 5'-ATGTGTATTGCCGCATATG (site 1) and 5'GCTACGACTAGTGC (site 3) for MS2 with 10 μM ligand. The footprint at site 2 (MS1) becomes shorter at the 3'-end with 5 μM showing protection at 5'-TTGTCA, whereas site 3 (MS2) protects the smaller regions at the 5'-end at 5'-AGTGC. In the middle of the gel, both fragments show protections at same sites with footprints at 5'-CTTTACATAACTCTT (site 6 for MS2 and site 3 for MS1), 5'-AACCTTG (site 4/MS2 and site 5/MS1) and 5'-ATAGCAATT (site 5/MS2 and site 4/MS1).

Sequences such as 5'-GCTAGC (site 2), 5'-CTTAAG (site 2), 5'-GTTAAC (site 3), 5'-GGTACC (site 3), 5'-GAATTC (site 4) of HexAfor; 5'-CATATG (site 2), 5'-GTATAC (site 4), and 5'-GCAAGC (site 5) of HexBfor; 5'GAATGC (site 1) and 5'GTATAC (site 4) of MS1 and 5'GTATTG (site 1/MS2) are often found in these footprints with the common feature of 5'-(G/C)X(A/T)(A/T)X(G/C). A footprint is also seen at GGTACC (site 3) of HexAfor, which is of the type 5'-GXXXC overlooking the first guanine. Other binding locations are seen at 5'-GAATTCC (site 4) of HexAfor and 5'GTATAGC (site 4) of MS1, which can be grouped as 5'-GXX(A/T)XXC.

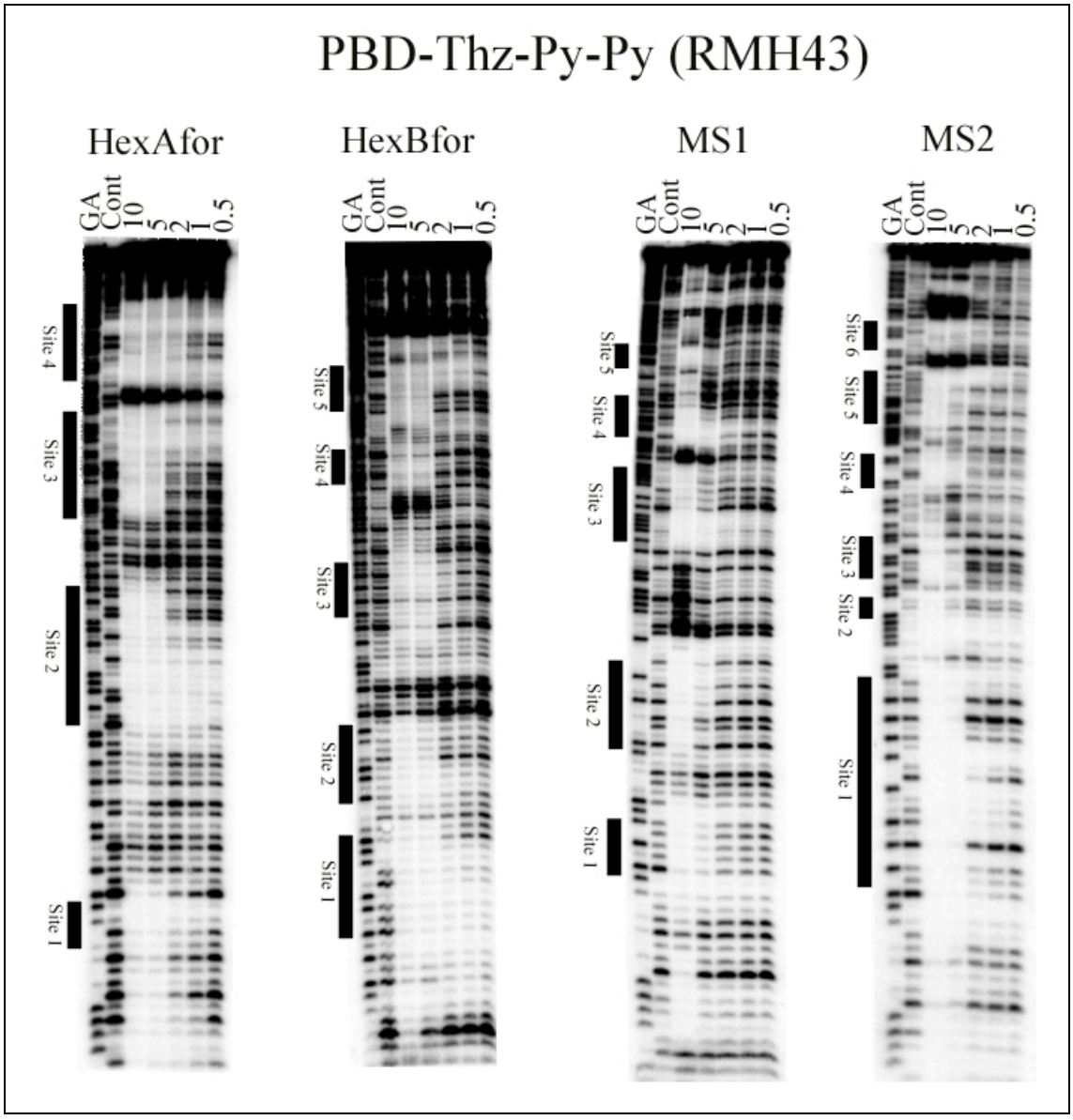


Figure 3.6: DNase I footprinting gels showing the interaction of various concentrations of PBD-Thz-Py-Py (RMH43) with HexAfor, HexBfor, MS1 and MS2. The samples were incubated with the ligand at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands. Filled bars indicate the location of the footprints.

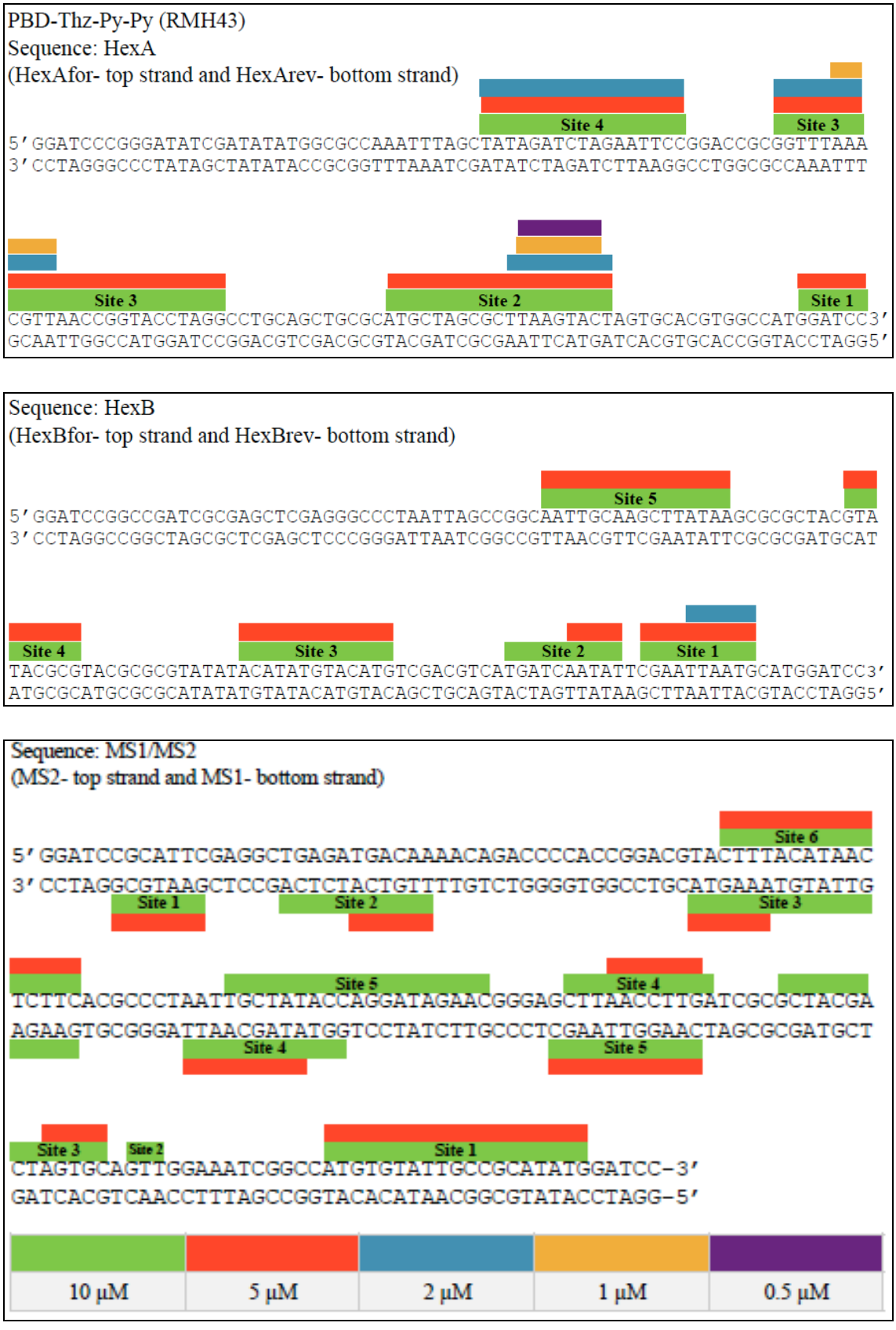


Figure 3.7: Sequences of the footprinting substrates indicating the regions protected by PBD-Thz-Py-Py (RMH43) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand.

3.4.1.5 *PBD-Py-Thz-Py (RMH47)*

The interaction of RMH47 (PBD-Py-Thz-Py) with HexAfor, HexBrev, MS1 and MS2 is shown in Figure 3.8 and the protected sequences are summarised in Figure 3.9. With HexAfor, the compounds show two large footprints covering 21 bp (site 2) and 20 bp (site 3) with 10 μM. Protection at site 2 is also evident with 5 μM. Another large footprint with 5'-CCGGACCGCGGT (site 4) is found with 10 μM and 5 μM, whereas a small footprint at 5'-TGGATCC at the bottom of the gel (site 1) is seen only with 10 μM ligand. No footprint is observed with concentrations of 2 μM and below.

With HexBrev, one very large footprint protecting 42 bp (site 1) is observed with 10 μM. With 5 μM this large footprint is broken into two regions; one footprint is at 5'-TTAGG, while the other still covers about 30 bp. On decreasing the ligand concentration further, the protection at 5'-TTAGG is retained while the longer footprint is shortened to only 7 bp, protecting 5'-AATTGCC. The ligand produces two more footprints at the top of the gel at 5'-ACATATGTATA (site 2) and 5'-CATGACGTCGA (site 1) with 10 μM; the first one is still protected at 2 μM, while the second one is only evident to 5 μM.

With MS1, the compound produces few footprints, with 10 μM protecting 5'-GAATGCG (site 1), 5'-TTGTCA (site 2), 5'-TATGTAAAGTAC (site 3), 5'-ATAGCAATTAG (site 4) and 5'AAGGTTAAG (site 5); the footprints at sites 2, 3 and 4 become shorter when the ligand concentration is decreased to 5 μM, With MS2, a large footprint (site 1) at the bottom of the gel covering 23 bp is found with 10 μM; this footprint is split into two sites of 5'-GGCCATGTGTAT and 5'-CATATG with 5 μM. Another footprint is observed at 5'-TAGTGC (site 2) both with 10 μM and 5 μM ligand. Other footprints in the centre of the gel with MS2 correspond to those seen with MS1.

These footprints seem to contain the sequences GGATCC (site 1), GCTAGC (site 2), GTTAAC (site 3) and GGTACC (site 3) of HexAfor; GCTTGC (site 1) of HexBrev and GAATGC (site 1) of MS1, which share the common feature of 5'-GX(A/T)(A/T)XC; a few sites have similar sequences except the position of guanine and cytosine are interchanged, such as 5'-CTTAAG (site 2) of HexAfor, 5'-CAATTG (site 1) and 5'-CATATG (site 2) of HexBrev.

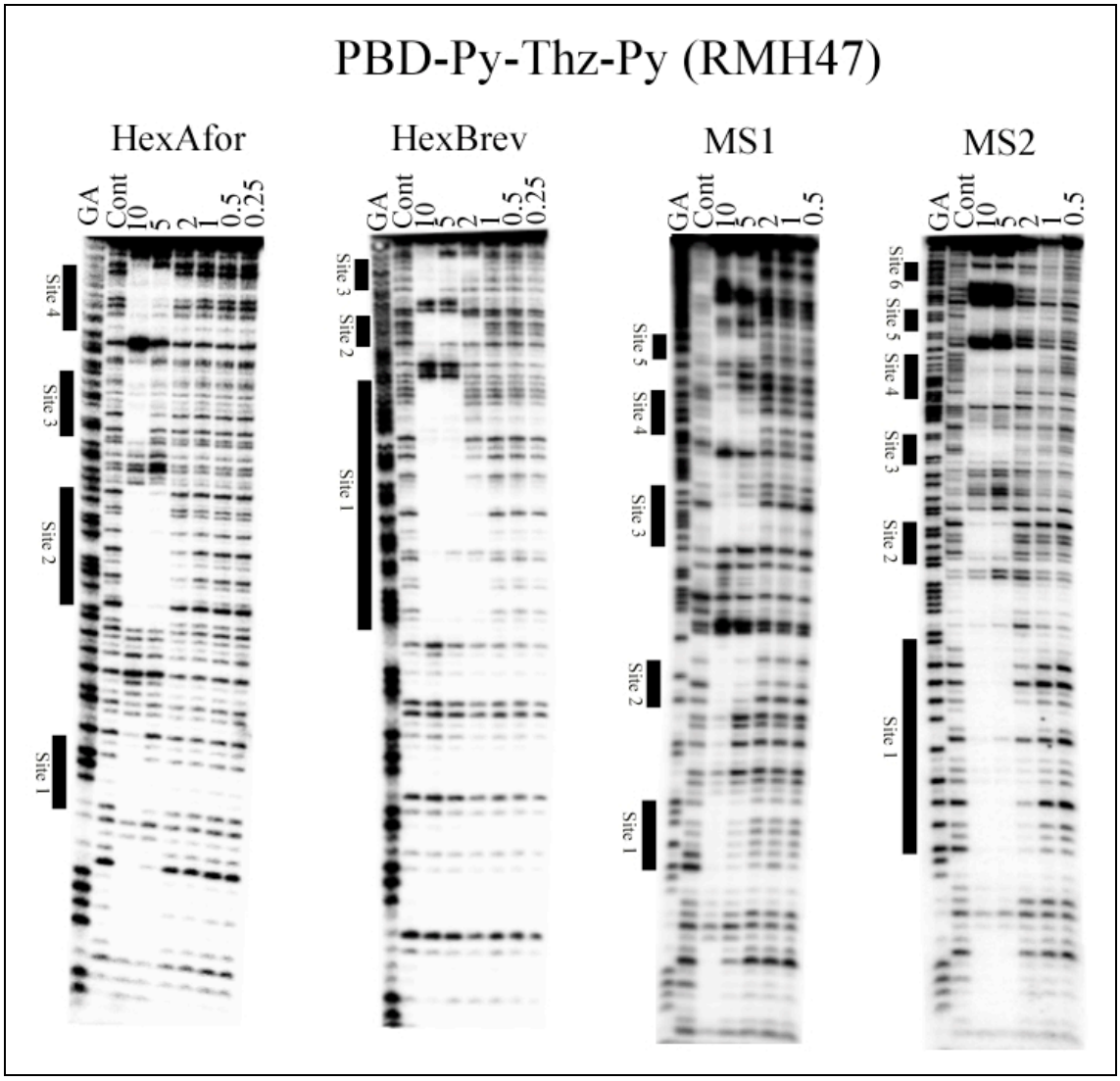


Figure 3.8: DNase I footprinting gels showing the interaction of various concentrations of PBD-Py-Thz-Py (RMH47) with HexAfor, HexBrev, MS1 and MS2. The samples were incubated with the ligand at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled "GA" correspond to markers specific for purines, while "Cont" indicates DNase I cleavage in the absence of added ligands. Filled bars indicate the location of the footprints.

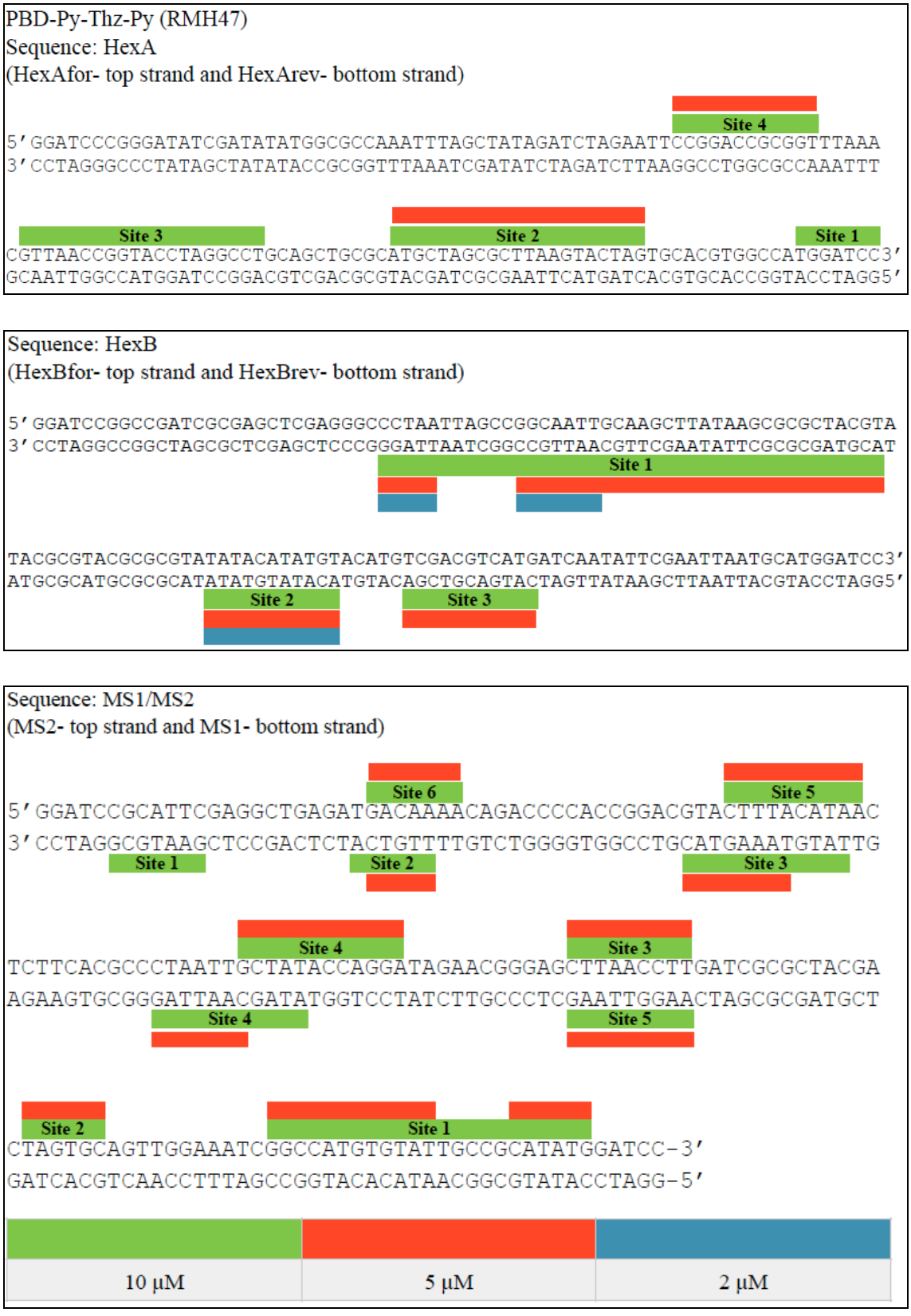


Figure 3.9: Sequences of the footprinting substrates indicating the regions protected by PBD-Py-Thz-Py (RMH47) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand.

3.4.2 Footprinting with DNA fragments containing potential binding sites

On the basis of the results obtained above we prepared three designed DNA fragments (Figure 3.10) that contain different variations of the potential ligand binding sites. Each of these fragments contains a set of potential six base pair binding sites, which are each separated by the sequence ATATAT, to which we assume the ligands will not bind. DNA fragment GXC2 contains four different 6 bp sites with guanine at the 5'-end and cytosine at the 3'-end (*e.g.* GCTAGC). CXG2 contain the inverted binding sites of GXC2 (*e.g.* CGATCG). GXC3 contains a number of other sequence variations such as 5'-GCTTGC, 5'-CGTTCG, 5'-GAATGC and 5'-CGTAAG. All of these inserts were cloned into the *Bam*HI site of pUC18 (GGATCC) and this potential ligand binding site is therefore present at both ends of the DNA fragments. The interaction of PBD conjugates (RMH41, RMH53, RMH43 and RMH47) with these DNA fragments is shown in Figures 3.11 to 3.14.

GXC3

5'GATCCATATATGGTACCATATATCTTACGATATATGCATTCATATATCGAACGATATATGCAAGCATATAGA3'
3' GTATATACCATGGTATATAGAATGCTATATACGTAAGTATATAGCTTGCTATATACGTTCGTATATCTAG5'

GXC2

5'GATCCATATATGCTAGCATATATGTATACATATATGGCGCCATATATGTTAACATATATGCTAGCATATATGGA3'
3' GTATATACGATCGTATATACATATGTATATACCGCGGTATATACAATTGTATATACGATCGTATATACCTAG5'

CXG2

5'GATCCATATATCGATCGATATATCRYRYGATATATCCGCGGATATATCATATGATATATCGATCGATATATGGA3'
3' GTATATAGCTACGTATATAGYRYRCTATATAGGCGCCTATATAGTATACTATATACGTAGCTATATACCTAG5'

Figure 3.10: Sequences of the designed DNA fragments- GXC3, GXC2 and CXG2 with hexanucleotide potential binding sites (red).

3.4.2.1 *PBD-Im-Py-Py (RMH41)*

Footprinting gels of this compound with the GXC3, GXC2 and CXG2 along with the protected sequences are shown in Figure 3.11. With insert GXC3, the compound produces six footprints. Strongly protected regions are found at 5'AGCATATAG (site 1), 5'ATATG (site 2), 5'CGATATATGCAT (site 4) and 5'TATATGGT (site 6) whereas other two footprints with attenuated bands are found at 5'ATATATC (site 3) and 5'ATATA (site 5). All these sites are protected with 5 μM and 2 μM, and the stronger ones at 1 μM. All the footprints indicate binding within the linker AT sequences rather than protecting 6 bp sequences in between the AT tracts. With insert GXC2, two large footprints can be seen at the top (site 4) and bottom (site 1) of the gel, as well as two further protections in the middle of the gel at 5'TATGTTA (site 2) and 5'TATATGGCGCCA (site 3). The large footprints at sites 4 and 1 contain the same 6

bp predicted site (GCTAGC) with footprints continuing into adjacent AT regions. The size of the footprint suggests that it contains two binding sites. With insert CXG2, although several regions (site 2, 3, 4, 5 and 6) are attenuated throughout the fragment, two clear footprints can be seen at 1 μM, one, near the bottom of the gel, covers 5'ATATGGA (site 1) while another, close to the top of the gel (site 7), protects 5'TATATCGA.

3.4.2.2 *PBD-Py-Im-Im (RMH53)*

The interaction of RMH53 with the designed sequences is shown in Figure 3.12. With insert GXC3, this ligand protects same regions as RMH41 producing six footprints with little variation in the length of footprints. Clear footprints are found at 5'AGCATATAGA (site 1), 5'TAGC (site 2), 5'ATATAT (site 3), 5'ATATGCA (site 4) and 5'TATATGG (site 6). [For this fragment data are not available for 2 and 1 μM]. Another footprint with attenuated bands is observed at 5'ATATCT (site 5). With GXC2, footprints are seen only with highest concentration (5 μM). Here, six strong protections are observed at 5'GCATATATGGA (site 1), 5'TATATGCT (site 2), 5'TATGTTA (site 3), 5'TATATGG (site 4), 5'TATATGT (site 5) and 5'ATATGCT (site 6). Footprints at site 2, 3, 4, 5 and 6 contain guanine at 3'-end with preceding AT regions of the linker. The large footprint at site 1 contain several guanines indicating multiple binding sites for this PBD. With insert CXG2, only two strong footprints are found with 5 μM ligand; one at the top of gel (site 2) in the sequence 5'-ATATATCGATC and another close to the bottom of the gel (site 1) protecting 5'-ATATGGA.

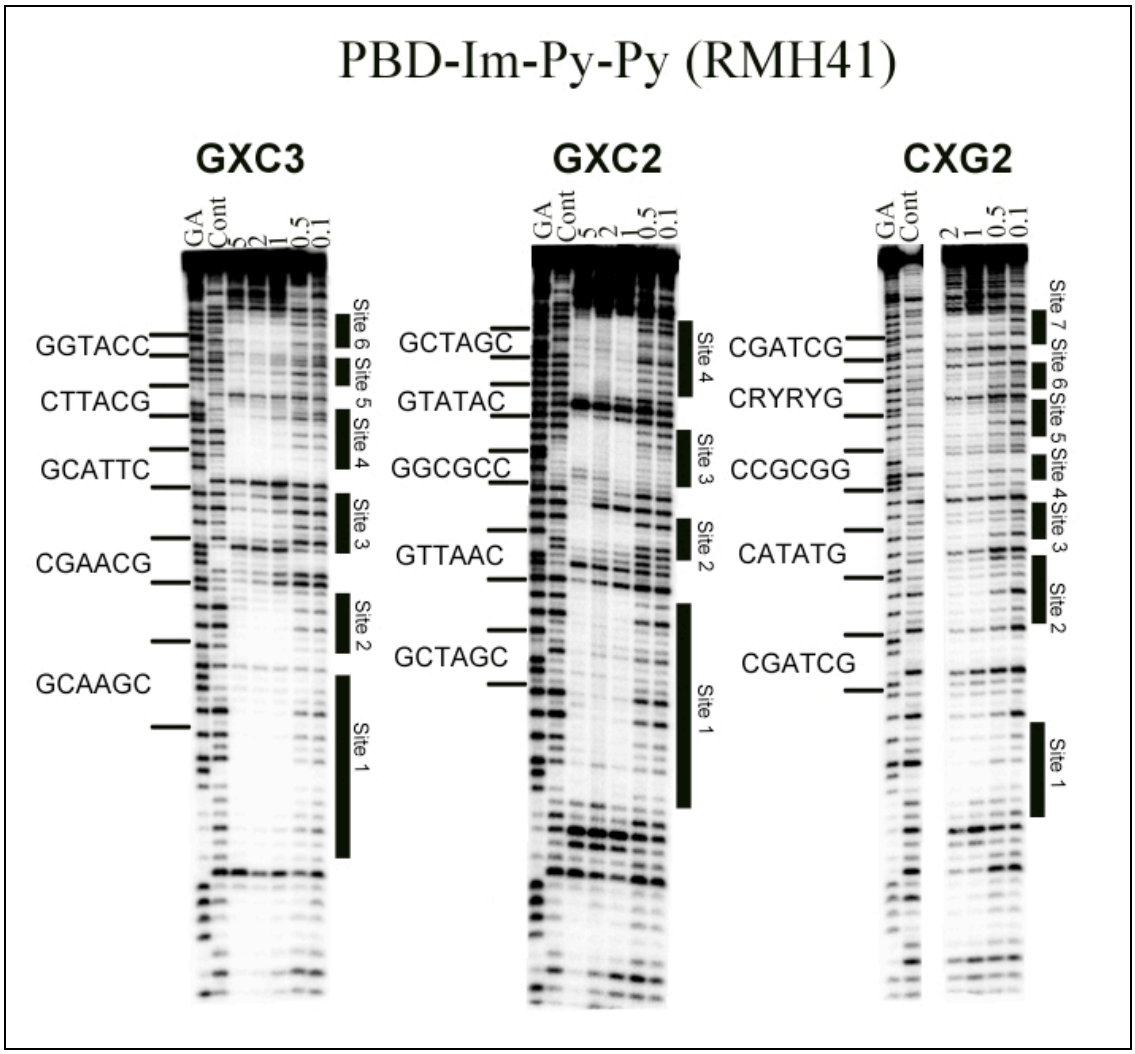
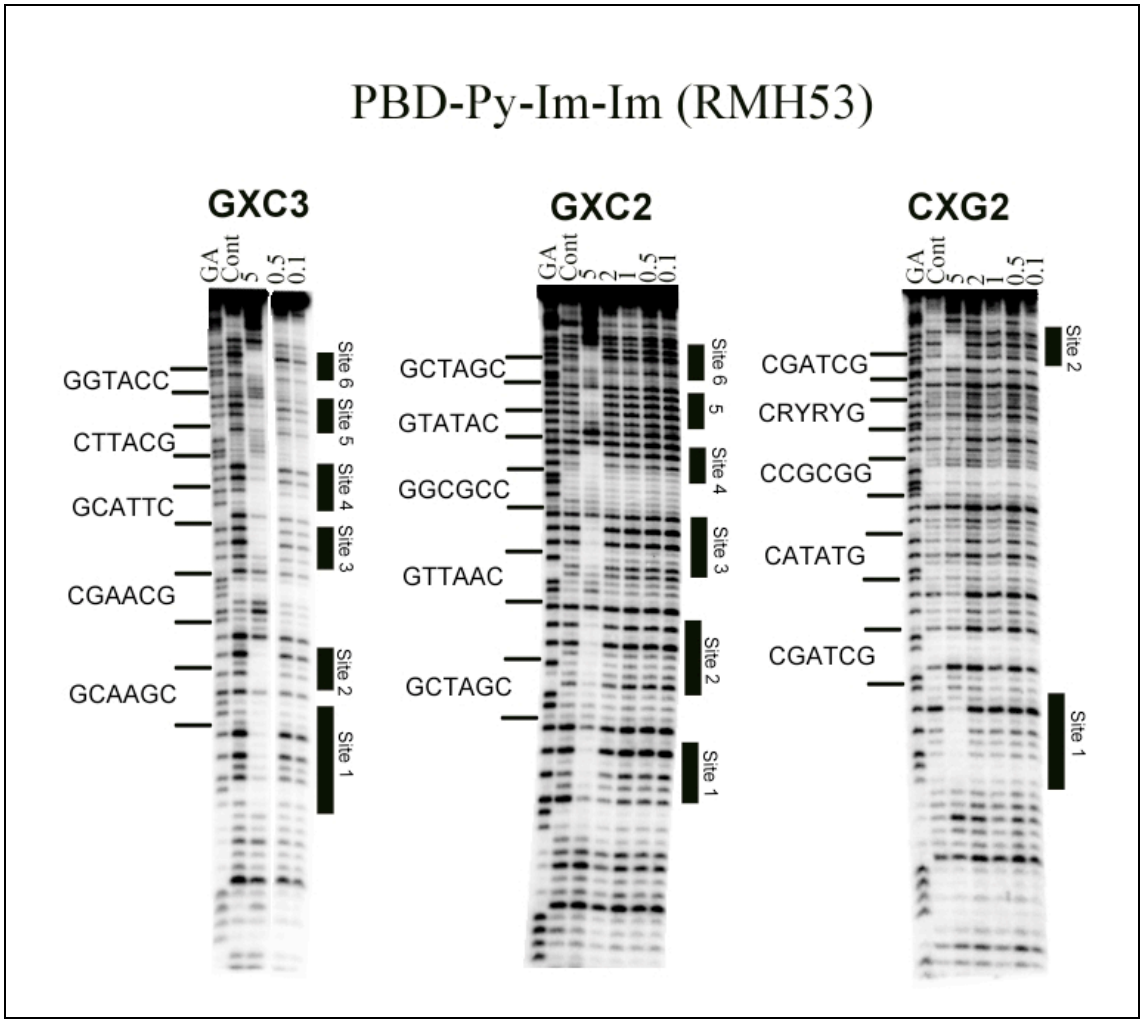


Figure 3.11: DNase I footprinting gels (top panel) showing the interaction of various concentrations of PBD-Im-Py-Py (RMH41) with insert GXC3, GXC2 and CXG2. The samples were incubated with the ligand at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel line. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligand. Filled bars indicate the positions of the footprints. In the bottom part, sequences of the footprinting substrates indicating the regions (coloured boxes) protected from DNase I cleavage. Only the sequence of radiolabelled strand is shown.



PBD-Py-Im-Im (RMH53)

GXC3

Site 6 Site 5 Site 4 Site 3 Site 2 Site 1

5′GATCCATATATGGTACCATATATCTTACGATATATGCATTCATATATCGAACGATATATGCAAGCATATAGA3′

GXC2

Site 6 Site 5 Site 4 Site 3 Site 2 Site 1

5′GATCCATATATGCTAGCATATATGTATACATATATGGCGCCATATATGTTAACATATATGCTAGCATATATGGA3′

CXG2

Site 2 Site 1

5′GATCCATATATCGATCGATATATCRYRYGATATATCCGCGGATATATCATATGATATATCGATCGATATATGGA3′

Figure 3.12: DNase I footprinting gels (top panel) showing the interaction of various concentrations of PBD-Py-Im-Im (RMH53) with inserts GXC3, GXC2 and CXG2. The samples were incubated with the ligand at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel line. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligand. Filled bars indicate the position and sequence of the footprints. In the bottom part, sequences of the footprinting substrates indicating the regions (coloured boxes) protected from DNase I cleavage. Only the sequence of radiolabelled strand is shown.

3.4.2.3 *PBD-Thz-Py-Py (RMH43)*

Footprinting gels of RMH43 with the GXC3, GXC2 and CXG2 along with the protected sequences are shown in Figure 3.13. With DNA fragment GXC3, the ligand protects six regions. Clear footprints are found at 5'AGCATATAGA (site 1), 5'ATATG (site 2), 5'GATATATGCAT (site 4) and 5'TATATGG (site 6). Two additional footprints with attenuated bands are observed at 5'ATATATC (site 3) and 5'ATATCT (site 5). The footprints at sites 1, 2 and 3 resemble those seen with compound RMH41. The footprints are sites 4 and 5 are one base shorter at their 5'-end and 3'-end respectively when compared with RMH41. With insert GXC2, protections are seen at six locations with clear footprints evident at 5 μM, 2 μM and 1 μM. Footprints are observed at the sequences 5'TATATGGA (site 1), 5'ATATGCT (site 2), 5'ATATGTT (site 3), 5'ATATGG (site 4), 5'ATATGT (site 5) and 5'ATATGC (site 6). With 1 μM, clear bands are observed between site 1 and 2 and also between site 3 and 4. With the highest concentrations, 5 μM and 2 μM, these middle bands become attenuated. With insert CXG2, four footprints with attenuated bands are produced, along with two more strong footprints. Clear footprints are seen at 5'TGGA (site 1) and 5'GATATA (site 4) while protections with attenuated bands are observed at site 2, 3, 5 and 6.

3.4.2.4 *PBD-Py-Thz-Py (RMH47)*

Figure 3.14 shows the interaction of RMH47 with GXC3, GXC2 and CXG2. Insert GXC3 shows six clear footprints and protections are observed at 5'AGCATATAGA (site 1), 5'ATATG (site 2), 5'ATATA (site 3), 5'ATATGCAT (site 4), 5'ATATCT (site 5) and 5'TATATG (site 6). Sites 2 and 6 are observed up to 1 μM while the footprints at sites 1, 3, 4 and 5 are only produced with 5 μM and 2 μM. With insert GXC2, one large footprint at the bottom of the gel (site 1) is seen with 5 μM, 2 μM and 1 μM, though few attenuated bands are observed in this footprint. Clear footprints are found at 5'ATATATGT (site 2), 5'TATATGGC (site 3), 5'ATATGTA (site 4) and 5'ATATGCT (site 5); with 5 μM and 2 μM, these footprints are clear while with 1 μM, attenuated bands are seen in site 2 and 4. With CXG2, two clear binding sites are observed at the top (site 6) and bottom (site 1) of the gel, while other footprints are produced, displaying weaker cleavage, at sites 2, 3, 4 and 5.

To compare the interaction of these four compounds with GXC3, GXC2 and CXG2, the footprints produced are combined and shown in Figure 3.15.

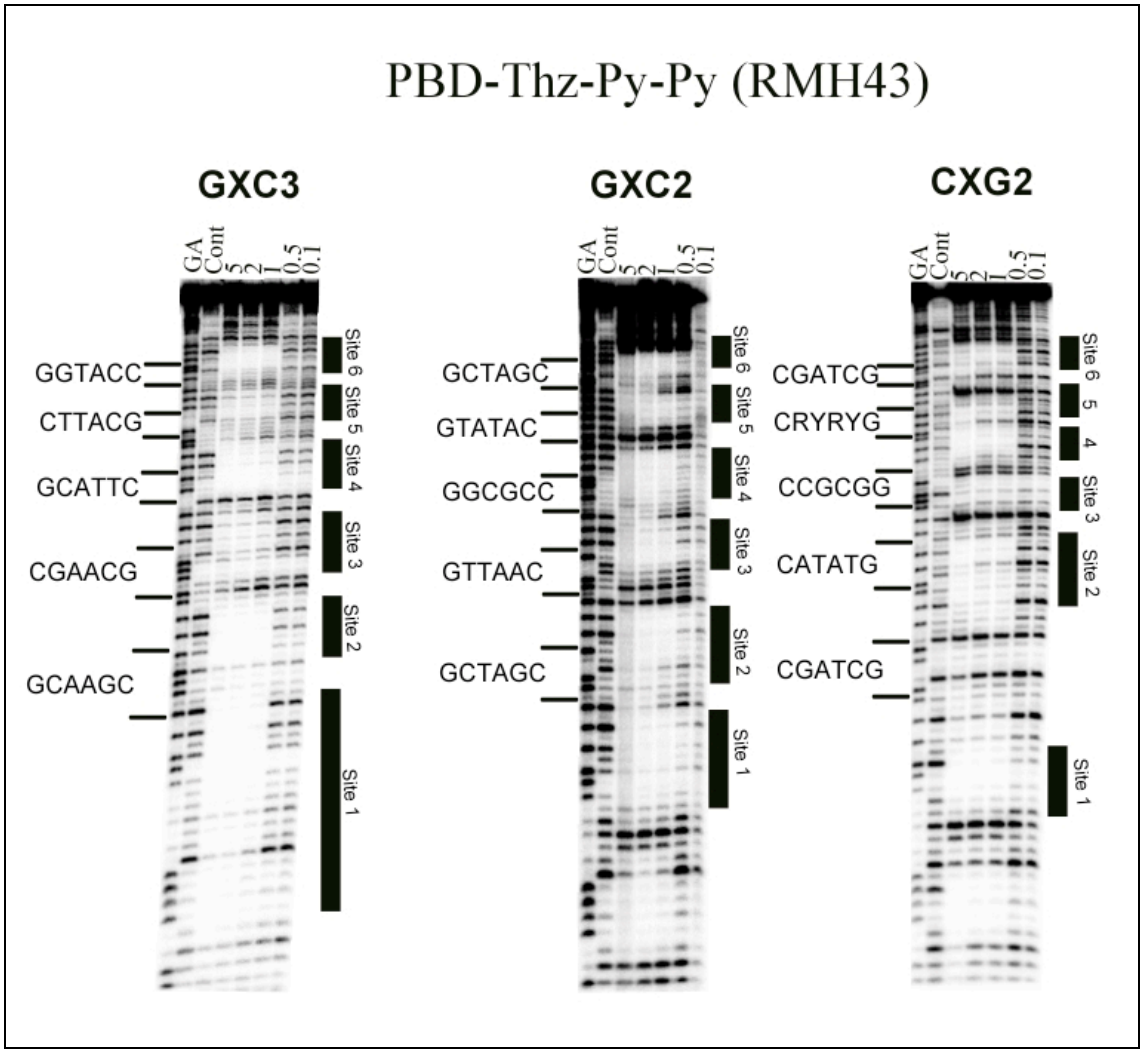


Figure 3.13: DNase I footprinting gels (top panel) showing the interaction of various concentrations of PBD-Thz-Py-Py (RMH43) with inserts GXC3, GXC2 and CXG2. The samples were incubated with the ligand at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel line. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligand. Filled bars indicate the position and sequence of the footprints. In the bottom part, sequences of the footprinting substrates indicating the regions (coloured boxes) protected from DNase I cleavage. Only the sequence of radiolabelled strand is shown.

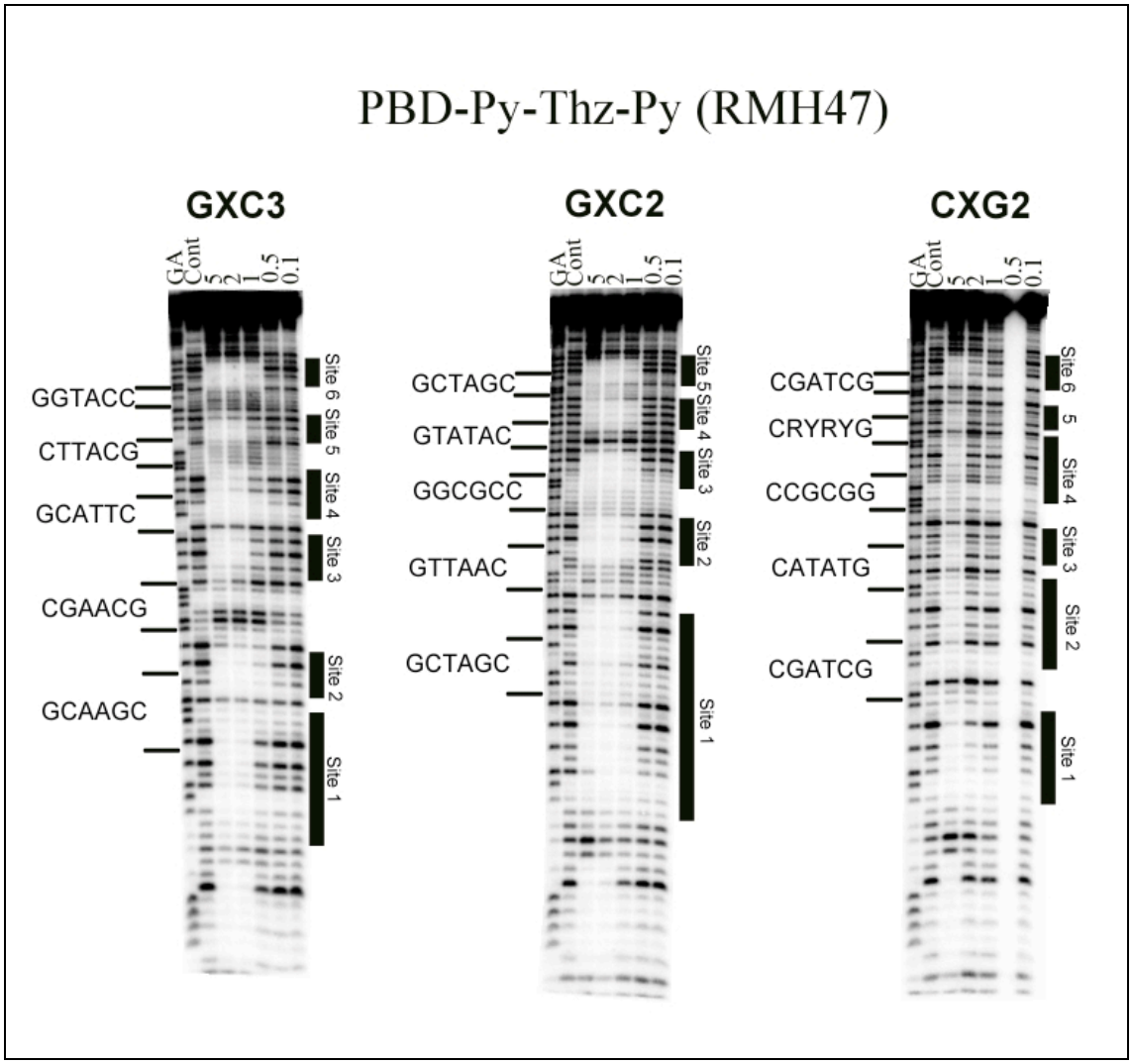


Figure 3.14: DNase I footprinting gels (top panel) showing the interaction of various concentrations of PBD-Py-Thz-Py (RMH47) with Insert GXC3, GXC2 and CXG2. The samples were incubated with the ligand at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel line. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligand. Filled bars indicate the position and sequence of the footprints. In the bottom part, sequences of the footprinting substrates indicating the regions (coloured boxes) protected from DNase I cleavage. Only the sequence of radiolabelled strand is shown.

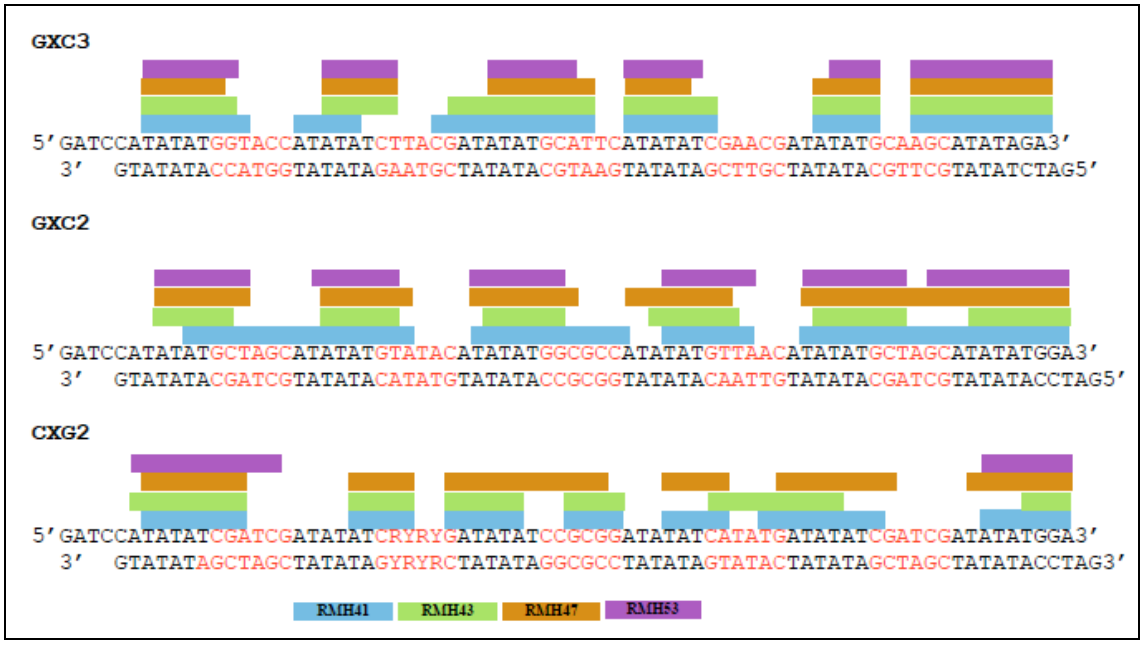


Figure 3.15: Sequences of the footprinting substrates indicating the regions (coloured boxes) protected from DNase I cleavage by all PBD-conjugates.

3.4.3 Effect of incubation period

Pyrrolobenzodiazepine is selective for guanine, with which it forms a covalent bond with the 2-amino group. The added polyamide chain consists of pyrrole, imidazole and thiazole and is intended to increase the selectivity of PBD-conjugates over longer sequences. Some footprints with the highest concentrations cover very long sequences (even 42 bp in HexB with RMH47), consisting of multiple overlapping ligand binding sites, and become shorter at lower concentrations. Since these compounds form covalent adducts with DNA it is possible that they could become 'trapped' on less preferred sites, with which they interact more quickly. The observed binding sites might therefore represent sequences with which the ligands interact more quickly, rather than those with which they bind more tightly. The best binding sites might therefore only become apparent after long incubation times. All the experiments in the preceding section were performed after incubating the ligand and DNA for about 24 hours. We therefore performed some experiments with MS1, examining whether the footprinting patterns change with different incubation times. Figure 3.16 shows the interaction of RMH41, RMH43, RMH47 and RMH53 with MS1 at different time intervals and the sequences protected from cleavage are summarised in Figure 3.17.

3.4.3.1 *PBD-Im-Py-Py (RMH41)*

With 10 μM RMH41 the footprint at site 2 appears within 1 minute of mixing the ligand with DNA. With 5 μM ligand this footprint is only observed after 5 minutes and is complete by one hour. In contrast the footprint at the bottom of the gel (site 1) appears

more slowly and is only evident after 3 hours and 24 hours with 10 μM and 5 μM ligand respectively. The protection at site 3 only becomes visible after 1 hour with 10 μM ligand; on incubating to longer times this footprint is extended in the 3'- (lower) direction end; this extended protection is seen after 3 hours with 5 μM ligand. Two other smaller footprints at the top of the gel (sites 4 and 5) appear after 1 hour with 10 μM ligand and 3 hours at the lower concentration (5 μM).

3.4.3.2 *PBD-Py-Im-Im (RMH53)*

The protections at sites 1, 2 and 3 appear after 1 and 3 hours incubation with 10 and 5 μM ligand respectively. The other two footprints at sites 4 and 5 appear at similar times but extend in the 5'-(upper) direction at longer incubation times.

3.4.3.3 *PBD-Thz-Py-Py (RMH43)*

The bands in site 2 are attenuated after 1 minute incubation with 10 μM ligand, though the footprint is not complete until 1 hour. Enhanced DNase I cleavage is evident at the top of this site at the shorter time points, but this disappears after longer incubation times as the footprint extends in the 3'- (lower) direction. This extended protection is also seen after 1 hour with 5 μM ligand. The protections at sites 3, 4 and 5 also appear after 5 minutes with 10 μM ligand and 1 hour with 5 μM. At site 4 the footprint extends further in the 3'-direction at longer incubation times. Another two footprints at sites 1 and 6 appear more slowly and are observed after 1 hour with 10 μM and 3 hours with 5 μM ligand.

3.4.3.4 *PBD-Py-Thz-Py (RMH47)*

The footprint at site 3 appears within 1 minute of mixing 10 μM ligand with the DNA and extends further in the 5'-(upper) direction at longer incubation times. With 5 μM ligand this footprint is slower to develop and does not show the time-dependent extension. Site 4 shows a similar profile, though another footprint towards the 5'-end of this site is observed after 1 hour, with only one unprotected band between them. These footprints coalesce after 3 hours producing a single large footprint; 5 μM shows same binding pattern at this site. The small footprint at site 5 is observed after 5 minutes and 1 hour with 10 μM and 5 μM respectively. Another footprint at the bottom of the gel (site 1) is evident after 1 hour and 3 hours respectively with 10 μM and 5μM. The small footprint at site 2 is observed after 1 and 3 hours with 10 μM and 5 μM ligand respectively, and with 10 μM ligand the protection extends in the 3' (lower)-direction end at longer incubation times. The proection at site 6 is also appears between 1 and 3 hours with 10 μM and 5 μM ligand.

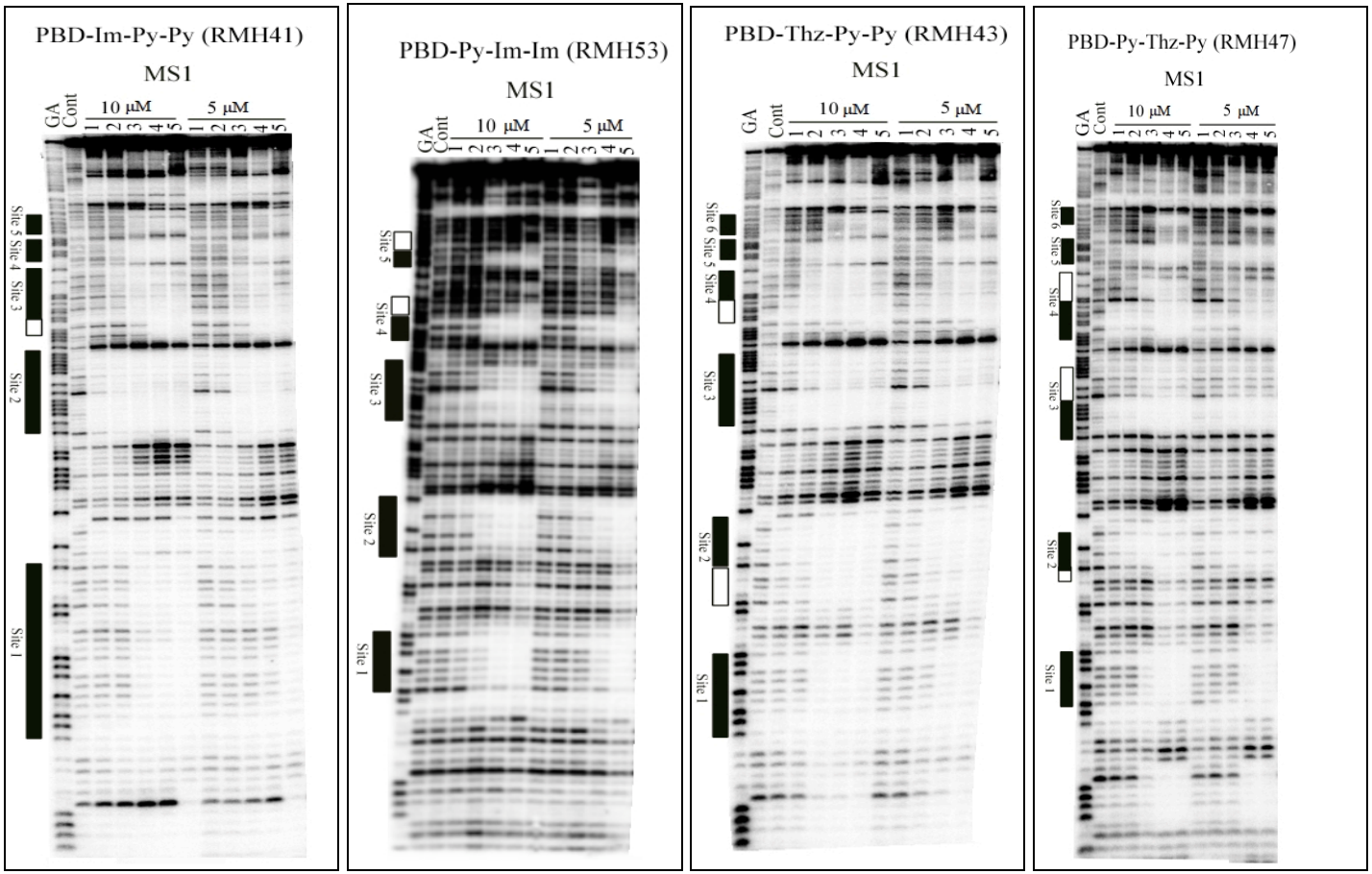
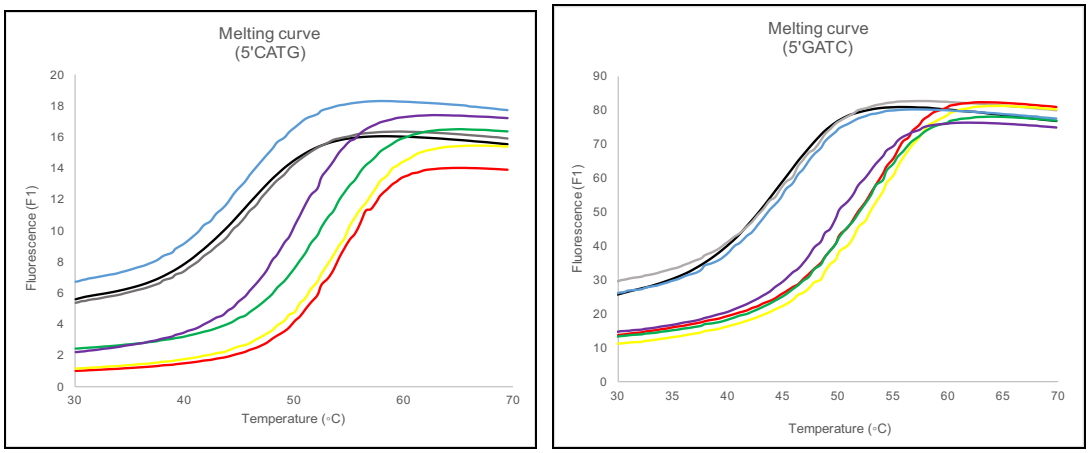


Figure 3.16: DNase I footprinting gels showing the time course of interaction of various concentrations of PBD-conjugates with MS1. The samples were incubated with the ligands (10 μM and 5 μM) at room temperature before digesting with DNase I. Incubation periods are shown at the top of each gel lane; 1 = 1 minute, 2 = 5 minutes, 3 = 1 hour, 4 = 3 hours and 5 = 24 hours. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands. The boxes indicate to the position of the footprints and correspond to those identified in Figure 3.17; the open boxes show regions where the size of the footprint increases with time.

Figure 3.17: Sequences of the footprinting substrates indicating the regions protected by PBD-conjugates from DNase I cleavage. Filled boxes (coloured) correspond to initial footprints at different locations; open boxes correspond to any extension/annexation of footprints over time. Only the sequence of radiolabelled strand is shown.

3.4.4 Melting studies

Melting studies were first performed with two different fluorescently labelled oligonucleotides with a single ligand concentration (10 μM). These oligonucleotides were used as they were already available in the laboratory, and the preliminary experiments were designed to test the feasibility of performing fluorescence melting experiment with these compounds. However, it should be noted that these self-complementary sequences are not well designed for this work as each F-labelled strand can bind to another F-strand (as too can two Q-labelled strands). The complexes were equilibrated for 12 hours and then heated from 30 °C to 70 °C while recording the fluorescence. Figure 3.18 shows the melting curves of these PBD-conjugates with these oligonucleotide duplexes and their corresponding melting temperatures are shown in the Table below. Both sequences show monophasic transitions in both the absence and presence of the ligands, with clear ligand-induced increase in T_m for RMH41, 43, 47 and 53, but very little change with RMH35 and 71.



	5′-F-AAAAACATGTTTTT 3′-Q-TTTTTGTACAAAAA (Control T_m = 44.9 °C)	5′-F-AAAAAGATCTTTTT 3′-Q-TTTTTCTAGAAAAA (Control T_m = 44.2 °C)
Ligand	Melting temperature (T_m) (ΔT_m) (°C)	
RMH35	45.5 ° (0.6 °)	44.2 ° (0 °)
RMH71	44.4 ° (-0.5 °)	44.7 ° (0.5 °)
RMH41	54.5 ° (9.6 °)	51.9 ° (7.7 °)
RMH43	54.1 ° (9.2 °)	52.9 ° (8.7 °)
RMH47	52.1 ° (7.2 °)	51.9 ° (7.7 °)
RMH53	50.0 ° (5.1 °)	50.1 ° (5.9 °)

Figure 3.18: Melting curves of oligonucleotide duplexes containing the sequences CATG and GATC in the presence of PBD-conjugates (10 μM)- control (black), RMH35 (grey), RMH71 (light blue), RMH41 (red), RMH43 (yellow), RMH47 (green), RMH53 (purple). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to melting and then annealing. 'Control' indicates samples without any added ligand. The table shows melting temperatures (T_m) of PBD-conjugates with each oligonucleotide duplexes along with the change in melting temperatures (ΔT_m).

These preliminary experiments confirm that fluorescence melting curves can be used to investigate the binding of these ligands and so other, more appropriate (non-self-complementary), sequences were designed based on the some of the best sites from the footprinting experiment. Each of these seven synthetic duplexes contained multiple guanines (Figure 3.19) and were tested with ligands RMH41, RMH43, RMH47 and RMH53 at concentrations between 10 μM and 0.1 μM. Since the concentration of the duplex is only 0.25 μM there is an excess DNA over ligand at the lowest concentration.

Oligonucleotides containing two consecutive guanines in each strand

5′-F-AAAAGGTACCAAAA 5′-F-AAAAGGATCCAAAA

3′-Q-TTTTCCATGGTTTT 3′-Q-TTTTCCTAGGTTTT

Oligonucleotides containing one guanine in each strand

5′-F-AAAAGTATACAAAA 5′-F-AAAAGTTAACAAAA

3′-Q-TTTTCATATGTTTT 3′-Q-TTTTCAATTGTTTT

Oligonucleotides containing two guanines separated by three other bases

5′-F-AAAAGCTTGCAAAA 5′-F-AAAAGCTAGCAAAA 5′-F-AAAAGAATGCAAAA

3′-Q-TTTTCGAACGTTTT 3′-Q-TTTTCGATCGTTTT 3′-Q-TTTTCTTACGTTTT

Figure 3.19: Sequences of the oligonucleotide duplexes designed for fluorescence melting studies with hexanucleotide binding sites shown in the middle (red).

In these experiments the complexes were equilibrated for about 12 hours before first heating from 30 °C to 95 °C at 0.1 °C/sec, while recording the fluorescence. Although the PBDs form covalent links to guanine, these bonds are broken at high temperatures and the ligand dissociates from the DNA. These melting curves were followed by cooling the samples to 30 °C at 0.1 °C/sec. The ligands only slowly associate with the DNA during this annealing reaction and in most cases these annealing profiles were similar to those of the free DNA. This was then followed by a second heating step recording the new melting transitions. Some of the ligands had reattached to the DNA during the annealing phase and these melting curves typically showed ligand-induced stabilisation relative to the drug-free controls, though this was lower than in the first melting transition. In the following description the melting curves are named as first melting curve and second melting curve. The compounds produce large changes in the melting profiles, showing multiphase transitions with all the oligonucleotides. In each melting curve, up to three transitions are observed: the first transition corresponds to that of the unmodified duplex having no ligand attached; this is followed by two further transitions, which we assume correspond to covalent attachment of one or two drug molecules, producing two further T_m values, which are denoted as 1st adduct and 2nd adduct. The proportions of unmodified duplex, 1st adduct and 2nd adduct were estimated as fractions of the total melting transition as shown in Figure 3.20. For these, and all the melting profile observed described below, the positions (T_m) of the different

melting transitions do not change with ligand concentration but their relative proportions are altered, so that a greater fraction of the higher T_m is evident with higher ligand concentration.

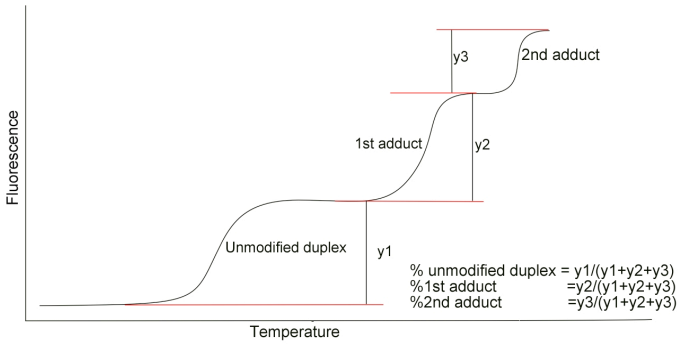


Figure 3.20: Schematic of a melting curve displaying three transitions

3.4.4.1 *PBD-Im-Py-Py (RMH41)*

Figures 3.21 and 3.22 show the first and second melting curves of PBD-Im-Py-Py (RMH41) with seven different oligonucleotide duplexes. The percentages of the unmodified duplex, 1st adduct and 2nd adduct are given in Tables 3.1 to 3.4 while the T_m values of unmodified duplex, 1st adduct and 2nd adduct are tabulated and presented in the Appendix.

GGTACC and GGATCC

The oligonucleotide containing GGTACC displays four melting temperatures at 41.4° C, 65.9° C, 76.8° C and 87.6° C in the first melting curve. The transition at 41.4° C corresponds to melting of unmodified duplex while the other three belong to different adducts formed by binding of ligand with duplex. In Table 3.1 the transitions at 65.9° C and 76.8° C have been included in same column as the 1st adduct, while the transition at T_m 87.6° C is termed as 2nd adduct. The transition at T_m 76.8° C is observed at ligand concentrations from 0.5 μM to 10 μM while the other two are only seen with 2 - 10 μM. With 0.5 μM ligand, about 47% of the melting profile corresponds to the 1st transition, with the rest as the unmodified duplex. Increasing the ligand concentration to 1 μM increases the proportion of this transition to 78%. With 2 μM, a new T_m of 65.9 °C is seen, constituting 36% of total melting transition with about 14% in a second transition at 87.6 °C. With 5 μM, the percentages of the transitions change to 47% (T_m of 65.9 °C), 24% (T_m of 76.8 °C) and 29% (T_m of 87.6 °C), which become 27%, 18% and 45% with 10 μM. When these sample were re-annealed the T_ms were all identical to the control,as the ligand had dissociated from the duplex at the higher temperatures, with insufficient time to reattach. In the second melting curve, no 2nd adduct is observed

at any concentration and more than 50% of the melting curve corresponds to the unmodified duplex. Two transitions of T_m 65.7 °C and 76.9 °C are observed with 0.5 μM, each constituting about 10% of the melting profile. The fraction in the transition at 65.7 °C increases with ligand concentration, finally reaching 80% with 10 μM. The percentage of the transition at 76.9 °C also increases a little, showing only 20% with 10 μM. The oligonucleotide containing GGATCC also shows three transitions at 61.9 °C, 75.8 °C and 90.0 °C in the presence of the ligand while the T_m of the unmodified duplex is 42.3 °C. The transition at 61.9 °C is first observed with 0.5 μM occupying 36% of the melt, which becomes more than 90% with 2 μM; at this point, no unmodified duplex is seen with very little (~7%) 2nd adduct. With 5 μM and 10 μM ligand, a new transition (T_m of 61.9 °C) is found accounting for 20% of total duplex. The proportion of the 2nd transition (T_m ~ 90.0 °C) also increases to 40%. In the second melting curve, transitions at 64.4 °C and 76.4 °C are observed, each constituting about 25% with 1 μM while the remainder appears as the unmodified duplex. These proportions increase to 41% and 34% with 2 μM. On increasing the ligand concentration to 5 μM and 10 μM, the amount of the transition at 64.4 °C increases to about 65%, while the transition at 76.4 °C remains same accounting for about 35% of the total.

GTATAC and GTTAAC

The oligonucleotides containing binding sites GTATAC and GTTAAC show similar melting profiles in the presence of RMH41. With GTATAC transitions are observed at 36.5 °C, 64.3 °C and 83.3 °C corresponding to the unmodified duplex, 1st adduct and 2nd adduct respectively. The transition corresponding to the 1st adduct accounts for only ~14% of the melt with 0.5 μM ligand, while the remainder appears as unmodified duplex. With 1 μM, the proportion of the unmodified duplex decreases with increasing amount of the 1st adduct (~43%) as well as the appearance of a transition corresponding to the 2nd adduct (~14%). Upon increasing the ligand concentration to 2 μM, the amount of the 1st adduct increases to around 70% with 30% as the 2nd adduct. In the second melting curve the transition corresponding to the 1st adduct is first seen with 2 μM (~22%) and this increases at higher ligand concentrations, accounting for about two-thirds of the melting profile with 10 μM; no transition equivalent to the 2nd adduct is observed at any concentration in these second melts. For the oligonucleotide containing GTTAAC, transitions are observed at 37.8 °C, 66.2 °C and 83.6 °C. The transition for the 1st adduct appears (~18%) with 0.5 μM and reaches 50% with 1 μM, with the remainder as the unmodified duplex. The proportion of the 1st adduct increases as the ligand concentration is raised, along with the appearance of a transition for the 2nd adduct; with 10 μM, about 10% of the melt corresponds to the 2nd

adduct with remainder as the 1st adduct. In the second melting curve, about 75% of the melting curve corresponds to the unmodified duplex at concentrations up to 2 μM, with only 27% as the 1st adduct; the proportion is reversed with 5 μM and with 10 μM the amount of 1st adduct increases to about ~84% with no evidence for the formation of a 2nd adduct.

GCTTGC, GCTAGC and GAATGC

For oligonucleotide GCTTGC, the transition at T_m 76.4 °C, which corresponds to the 1st adduct, starts to form with 0.5 μM ligand, at which it constitutes about 42% of the total melting curve; the proportion increases to 62% with 1 μM with the remainder as the unmodified duplex. With 2 μM ligand the majority of the melt appears as the 1st adduct, with only 23% as 2nd adduct. On increasing the ligand concentration to 5 μM, the amount of the 2nd adduct increases while that of 1st adduct decreases; both accounting for about 50% of the melting curve. In the second melting curve, the majority of the transition appears as the unmodified duplex at ligand concentrations up to 1 μM; with 2 μM, the 1st adduct accounts for about two-thirds the melt, rising to 100% with 5 μM. On increasing the concentration to 10 μM the amount of the 1st adduct decreases a little with the appearance of a transition corresponding to the 2nd adduct as about 12% of the total. The related oligonucleotide GCTAGC also shows three transitions at 43.9 °C, 76.4 °C and 92.5 °C. The transition at 76.4 °C accounts for about 35% of the melting curve at 0.5 μM ligand with remainder as unmodified duplex; this increases to 83% with 1 μM and 93% with 2 μM ligand. With 5 μM, an additional transition appears at 64.8 °C along with the formation of 2nd adduct (~92.5 °C); thus at this concentration, three transitions having T_m of 64.8 °C, 76.4 °C and 92.5 °C are seen sharing 17%, 21% and 62% of the melting transition respectively; the proportions remain same with 10 μM ligand. In the second melting curve, the 1st adduct is observed with 0.5 μM and this gradually increases to 85% of the total transition with 5 μM ligand along with 15% from the 2nd adduct. With 10 μM ligand, the transition from the 1st adduct reduces to 77% with 23% as the 2nd adduct. Unlike first melting curve, no transition around 64.8 °C is observed at any ligand concentration. Oligonucleotide GAATGC shows similar transitions demonstrating the formation of two adducts. With 0.5 μM ligand about, 25 % of the melt corresponds to the 1st adduct, this increases to over 50% with 1 μM, with the remainder as the unmodified duplex. Further increases in ligand concentration cause the appearance of the 2nd adduct. With 5 μM, the proportions of the 1st and 2nd adducts are 57% and 43% respectively; each of which account for 50% with 10 μM ligand. In the second melting curve, the unmodified duplex predominates, even with highest concentration (10 μM); only 20% and 32% of 1st adduct is observed with 5 μM and 10 μM ligand respectively with no appearance of 2nd adduct.

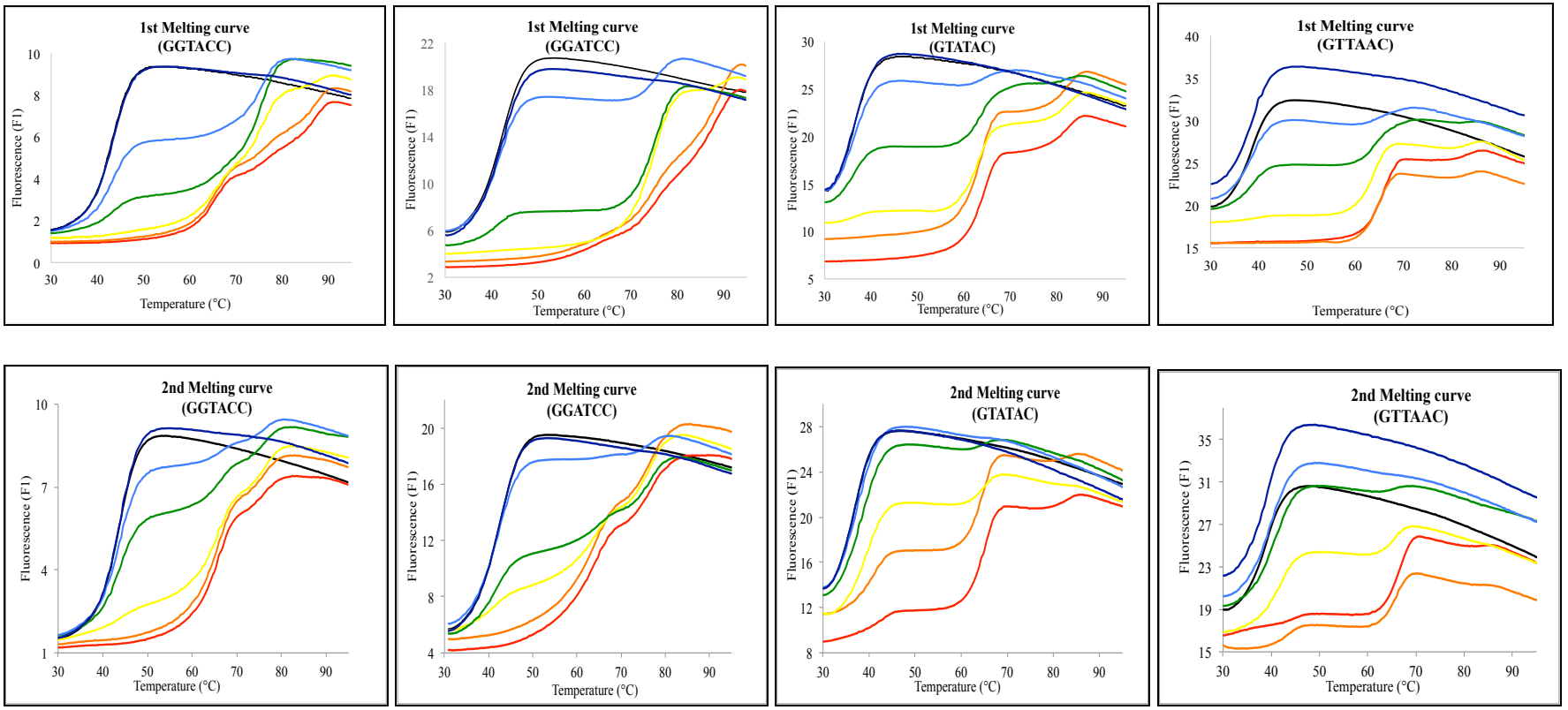


Figure 3.21: First and second melting curves of oligonucleotide duplexes containing the sequences GGTACC, GGATCC, GTATAC and GTTAAC in the presence of different concentrations of PBD-Im-Py-Py (RMH41)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green); 0.5 μM (light blue) and 0.1 μM (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting.

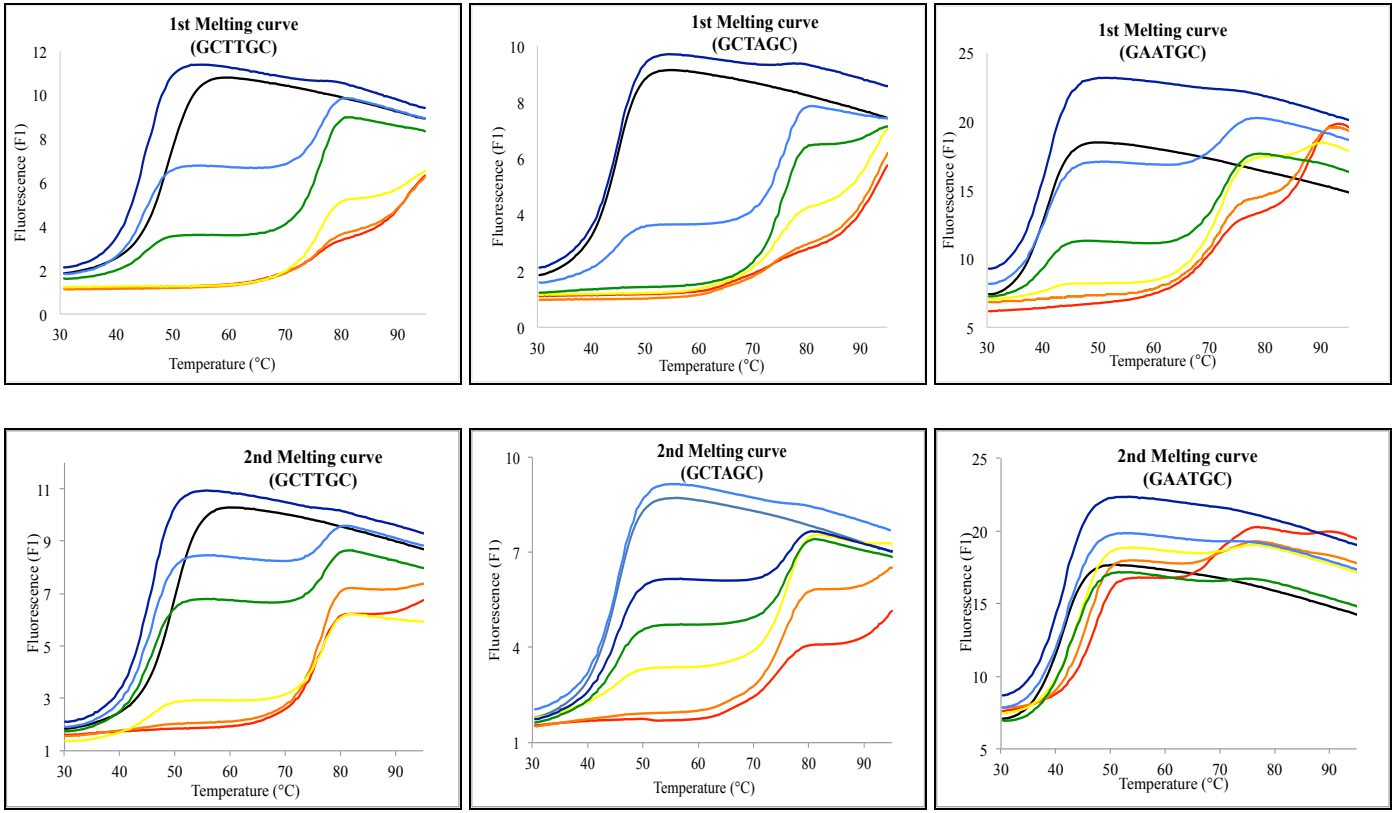


Figure 3.22: First and second melting curves of oligonucleotide duplexes containing the sequences GCTTGC, GCTAGC and GAATGC in the presence of different concentrations of PBD-Im-Py-Py (RMH41)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green); 0.5 μM (light blue) and 0.1 μM (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting.

Table 3.1: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Im-Py-Py (RMH41) with oligonucleotide duplexes containing the sequences GGTACC and GGATCC

	1st melting curve			2nd melting curve		
5'GGTACC	Unmodified duplex T_m 41.4 °C	Duplex+Ligand (1st adduct) T_m 65.9 -76.8 °C	Duplex+Ligand (2nd adduct) T_m 87.0 -87.6 °C	Unmodified duplex T_m 41.7 °C	Duplex+Ligand (1st adduct) T_m 65.7 -76.9 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	53%	-, 47%	0%	77%	12%*, 11%*	0%
1 μM	22%	-, 78%	0%	56%	27%*, 17%*	0%
2 μM	0%	36%*, 50%*	14%	15%	58%*, 27%*	0%
5 μM	0%	47%*, 24%*	29%	0%	78%*, 22%*	0%
10 μM	0%	36%*, 19%*	45%	0%	80%*, 20%*	0%

*Two transitions are observed in close proximity and are included in the same column

	1st melting curve			2nd melting curve		
5'GGATCC	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 61.9 -75.8 °C	Duplex+Ligand (2nd adduct) T_m 88.9 -90.0 °C	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 64.4 -76.4 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	65%	-, 36%	0%	91%	-, 9%*	0%
1 μM	17%	-, 83%	0%	46%	25%*, 29%*	0%
2 μM	0%	-, 93%	7%	25%	41%*, 34%*	0%
5 μM	0%	20%*, 40%*	40%	0%	62%*, 38%	0%
10 μM	0%	20%*, 40%*	40%	0%	68%, 32%*	0%

Table 3.2: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Im-Py-Py (RMH41) with oligonucleotide duplexes containing the sequences GTATAC and GTTAAC

	1st melting curve			2nd melting curve		
5'GTATAC	Unmodified duplex T_m 36.5 °C	Duplex+Ligand (1st adduct) T_m 63.6 -64.3 °C	Duplex+Ligand (2nd adduct) T_m 82.3 -83.3 °C	Unmodified duplex T_m 36.6 °C	Duplex+Ligand (1st adduct) T_m 64.7 -64.8 °C	Duplex+Ligand (2nd adduct) T_m 82.8 -83.0 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	86%	14%	0%	100%	0%	0%
1 μM	43%	43%	14%	95%	0%	0%
2 μM	7%	65%	28%	78%	22%	0%
5 μM	0%	75%	25%	37%	63%	0%
10 μM	0%	71%	29%	25%	75%	0%

	1st melting curve			2nd melting curve		
5'GTTAAC	Unmodified duplex T_m 37.8 °C	Duplex+Ligand (1st adduct) T_m 64.3 -66.2 °C	Duplex+Ligand (2nd adduct) T_m 83.6 °C	Unmodified duplex T_m 38.1 °C	Duplex+Ligand (1st adduct) T_m 65.4 -66.5 °C	Duplex+Ligand (2nd adduct) T_m 83.8 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	82%	18%	0%	100%	0%	0%
1 μM	50%	50%	0%	96%	0%	0%
2 μM	11%	84%	5%	73%	27%	0%
5 μM	0%	94%	6%	27%	73%	0%
10 μM	0%	90%	10%	16%	84%	0%

Table 3.3: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Im-Py-Py (RMH41) with oligonucleotide duplexes containing the sequences GCTTGC and GCTAGC

	1st melting curve			2nd melting curve		
5'GCTTGC	Unmodified duplex T_m 44.6 °C	Duplex+Ligand (1st adduct) T_m 75.9 -76.4 °C	Duplex+Ligand (2nd adduct) T_m 91.3 -91.8 °C	Unmodified duplex T_m 44.8 °C	Duplex+Ligand (1st adduct) T_m 76.1 -77.0 °C	Duplex+Ligand (2nd adduct) T_m 92.0 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	58%	42%	0%	82%	18%	0%
1 μM	38%	62%	0%	71%	29%	0%
2 μM	0%	77%	23%	33%	67%	0%
5 μM	0%	46%	54%	0%	100%	0%
10 μM	0%	46%	54%	0%	88%	12%

	1st melting curve			2nd melting curve		
5'GCTAGC	Unmodified duplex T_m 43.9 °C	Duplex+Ligand (1st adduct) T_m 64.8 -76.4 °C	Duplex+Ligand (2nd adduct) T_m 92.0 -92.5 °C	Unmodified duplex T_m 44.3 °C	Duplex+Ligand (1st adduct) T_m 74.6 -76.6 °C	Duplex+Ligand (2nd adduct) T_m 89.9 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	65%	35%	0%	72%	28%	0%
1 μM	17%	83%	0%	56%	44%	0%
2 μM	0%	93%	7%	28%	72%	0%
5 μM	0%	17%, 21%*	62%	0%	-, 85%	15%
10 μM	0%	17%, 21%*	62%	0%	-, 77%	23%

*Two transitions are observed in close proximity and are included in the same column

Table 3.4: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Im-Py-Py (RMH41) with oligonucleotide duplex containing the sequence GAATGC

	1st melting curve			2nd melting curve		
5'GAATGC	Unmodified duplex T_m 40.7 °C	Duplex+Ligand (1st adduct) T_m 71.6 -72.7 °C	Duplex+Ligand (2nd adduct) T_m 87.7 -88.2 °C	Unmodified duplex T_m 40.9 °C	Duplex+Ligand (1st adduct) T_m 68.9 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	75%	25%	0%	100%	0%	0%
1 μM	42%	58%	0%	100%	0%	0%
2 μM	0%	85%	14%	100%	0%	0%
5 μM	0%	57%	43%	80%	20%	0%
10 μM	0%	50%	50%	68%	32%	0%

3.4.4.2 *PBD-Thz-Py-Py (RMH43)*

GGTACC and GGATCC

Figures 3.23 and 3.24 show the effect of PBD-Thz-Py-Py (RMH43) on the melting profiles of the different oligonucleotides and the percentages of different adducts are given in Table 3.5 to 3.8. Oligonucleotides containing GGTACC and GGATCC display similar melting characteristics. For GGTACC, transitions are observed at 41.4 °C, 73.2 °C and 90.3 °C corresponding to the unmodified duplex, 1st adduct and 2nd adduct respectively, while for GGATCC, these are 42.3 °C, 77.6 °C and 92.4 °C. For both oligonucleotides nearly all the transition appears as the 1st adduct with 0.5 μM ligand. The transition for the 2nd adduct appears at 1 μM, and constitutes about 25% of the melting profile. On increasing the concentration of ligand, the proportion of the 1st adduct decreases whereas the 2nd adduct increases; finally showing 50% of each at 10 μM ligand for GGTACC and 36% and 64% for the 1st and 2nd adduct with GGATCC. In the second melting curve of GGTACC, a new transition of T_m 64.9 °C is found with 5 μM and 10 μM, though there is no evidence for a 2nd adduct at any ligand concentration. As seen with the first melting curve, a transition at 73.4 °C is observed from 0.5 μM which accounts for 91% of the change with 2 μM ligand. With 5 μM, the transition at 64.9 °C appears, sharing 45% of the total melt, rising to 64% with 10 μM ligand. The same phenomenon is observed in the second melting curve of GGATCC; the transition at 63.9 °C is first observed with 0.5 μM constituting 54% of total melt; On increasing the ligand concentration to 2 μM, this rises to 93% and another transition appears with a T_m of 63.9 °C, which accounts for about 7% of the total. With 10 μM ligand these transitions account for 22% and 78% of the total respectively.

GTATAC and GTTAAC

For oligonucleotide GTATAC, transitions are found at 36.5 °C, 65.0 °C and 87.0 °C. With 0.5 μM, contributions from all three transitions are observed in the ratio 5:85:10. The amount of 1st adduct then decreases to 62% while the 2nd adduct rises to 38% with 1 μM. With higher concentrations, the 1st and 2nd adduct make equal contributions to the melting curve. In the second melting curve, the transition of the unmodified duplex predominates up to 2 μM; after that the 1st adduct occupies about 70% of melting curve with less than 10% as 2nd adduct. With 10 μM, the 2nd adduct increases to 14% of the total. With oligonucleotide GTTAAC the 1st and 2nd adduct are observed at 69.6 °C and 77.6 °C. With 0.5 μM, two-thirds of duplex is in the 1st adduct with the remainder as the unmodified duplex. On increasing the ligand concentration to 1 μM the 2nd adduct starts to form accounting for 35%; which then rises to 44% with 10 μM ligand. In the second melting curve, 75% of the melting transition appears as the unmodified duplex with 2

μM ligand. With 5 μM and 10 μM, more than 60% of the profile is found as the 1st adduct though there is no formation of the 2nd adduct.

GCTTGC, GCTAGC and GAATGC

In the first melting curves of GCTTGC and GCTAGC, 0.5 μM ligand produces less than 10% of the melting transition as the 1st adduct, though this increases to more than 80% along with formation of 2nd adduct at higher concentrations. The proportion of the 1st adduct then decreases with a concomitant increase in the 2nd adduct; both oligonucleotides show about 40% as the 1st adduct and 60% as the 2nd adduct with 5 μM and 10 μM ligand. In the second melting curves more than 80% corresponds to the 1st adduct form with 1 μM ligand. With GCTTGC, the entire transition appears as the 1st adduct from 2 μM to 10 μM, while GCTAGC shows 100% as the 1st adduct with 2 μM, which then decreases to 71% at 10 μM with the remainder appearing as the 2nd adduct. For oligonucleotide GAATGC, transitions are observed at 40.7 °C, 73.0 °C and 91.3 °C. The 1st adduct is evident at 0.1 μM ligand while the 2nd adduct begins to appear with 0.5 μM at which concentration 87% is present as the 1st adduct with the remainder as 2nd adduct With 1 μM the proportion changes to 55% of 1st adduct with the remainder as 2nd adduct. With 2 μM, 5 μM and 10 μM, the 1st and 2nd adduct account for about 45% and 55% of the transition respectively. In the second melting curve, at concentrations up to 2 μM more than 80% of the melt appears as the unmodified duplex. The 2nd adduct is only seen with 5 μM and 10 μM accounting for only 17% with 44% and 39% as the unmodified duplex and 1st adduct respectively.

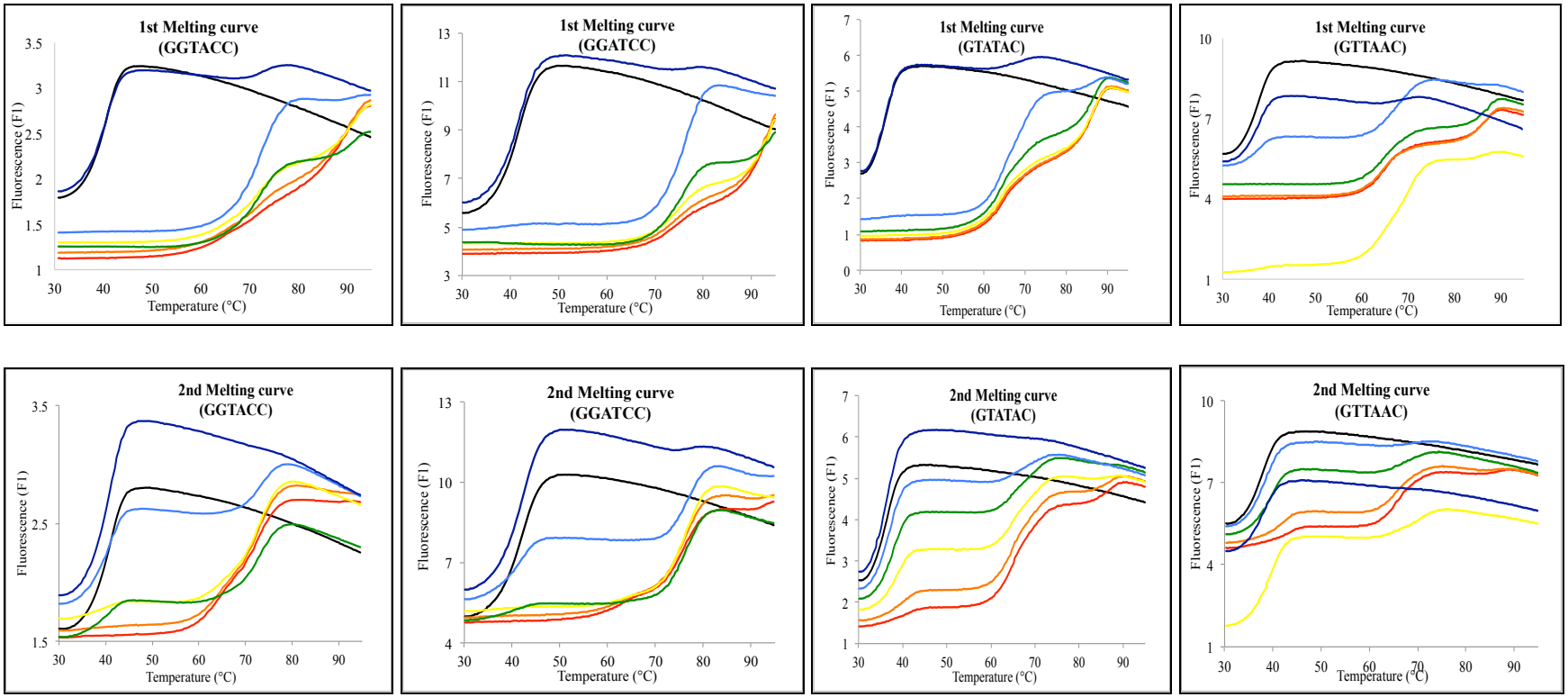


Figure 3.23: First and second melting curves of oligonucleotide duplexes containing the sequences GGTACC, GGATCC, GTATAC and GTTAAC in the presence of different concentrations of PBD-Thz-Py-Py (RMH43)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green); 0.5 μM (light blue) and 0.1 μM (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting.

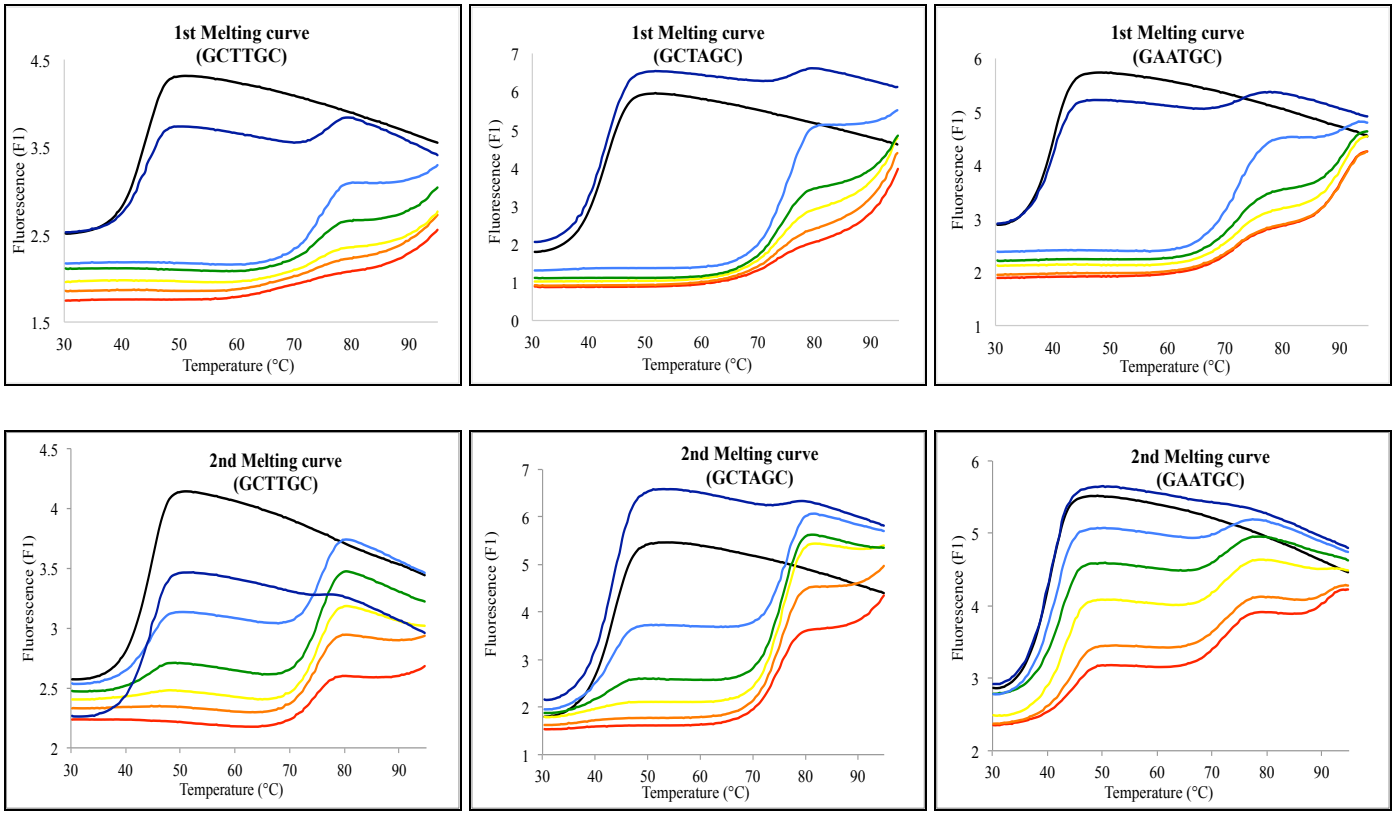


Figure 3.24: First and second melting curves of oligonucleotide duplexes containing the sequences GCTTGC, GCTAGC and GAATGC in the presence of different concentrations of PBD-Thz-Py-Py (RMH43)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green); 0.5 μM (light blue) and 0.1 μM (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting.

Table 3.5: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Thz-Py-Py (RMH43) with oligonucleotide duplexes containing the sequences GGTACC and GGATCC

	1st melting curve			2nd melting curve		
5'GGTACC	Unmodified duplex T_m 41.4 °C	Duplex+Ligand (1st adduct) T_m 72.0 -73.2 °C	Duplex+Ligand (2nd adduct) T_m 89.7 -90.3 °C	Unmodified duplex T_m 41.7 °C	Duplex+Ligand (1st adduct) T_m 64.9 -73.4 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0. 1μM	87%	13%	0%	100%	0%	0%
0.5 μM	0%	100%	0%	67%	-, 33%*	0%
1 μM	0%	75%	25%	26%	-, 74%*	0%
2 μM	0%	60%	40%	9%	-, 91%*	0%
5 μM	0%	53%	47%	0%	45%, 55%*	0%
10 μM	0%	50%	50%	0%	36%, 64%*	0%

*Two transitions are observed in close proximity and are included in the same column

	1st melting curve			2nd melting curve		
5'GGATCC	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 74.8 -77.6 °C	Duplex+Ligand (2nd adduct) T_m 92.3 -92.4 °C	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 63.9 -76.9 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	7%	93%	0%	46%	-, 54%*	0%
1 μM	0%	73%	27%	14%	-, 86%*	0%
2 μM	0%	47%	53%	0%	7%*, 93%*	0%
5 μM	0%	36%	64%	0%	18%*, 82%*	0%
10 μM	0%	36%	64%	0%	22%*, 78%*	0%

Table 3.6: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Thz-Py-Py (RMH43) with oligonucleotide duplexes containing the sequences GTATAC and GTTAAC

	1st melting curve			2nd melting curve		
5'GTATAC	Unmodified duplex T_m 36.5°C	Duplex+Ligand (1st adduct) T_m 64.3 -65.0 °C	Duplex+Ligand (2nd adduct) T_m 86.0 -87.0 °C	Unmodified duplex T_m 36.6 °C	Duplex+Ligand (1st adduct) T_m 64.8 -65.2 °C	Duplex+Ligand (2nd adduct) T_m 86.8 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	88%	12%	0%	100%	0%	0%
0.5 μM	5%	85%	10%	79%	21%	0%
1 μM	0%	62%	38%	62%	38%	0%
2 μM	0%	53%	47%	60%	40%	0%
5 μM	0%	51%	49%	21%	70%	9%
10 μM	0%	51%	49%	14%	72%	14%

	1st melting curve			2nd melting curve		
5'GTTAAC	Unmodified duplex T_m 37.8 °C	Duplex+Ligand (1st adduct) T_m 64.4 -69.6 °C	Duplex+Ligand (2nd adduct) T_m 74.8 -77.6 °C	Unmodified duplex T_m 38.1 °C	Duplex+Ligand (1st adduct) T_m 65.5-68.8 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	92%	8%	0%	100%	0%	0%
0.5 μM	32%	68%	0%	100%	0%	0%
1 μM	0%	65%	35%	76%	24%	0%
2 μM	0%	91%	9%	74%	26%	0%
5 μM	0%	56%	44%	37%	63%	0%
10 μM	0%	56%	44%	29%	71%	0%

Table 3.7: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Thz-Py-Py (RMH43) with oligonucleotide duplexes containing the sequences GCTTGC and GCTAGC

	1st melting curve			2nd melting curve		
5'GCTTGC	Unmodified duplex T_m 44.6 °C	Duplex+Ligand (1st adduct) T_m 73.6-76.5 °C	Duplex+Ligand (2nd adduct) T_m 92.3 °C	Unmodified duplex T_m 44.8 °C	Duplex+Ligand (1st adduct) T_m 75.0-76.0 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	88%	12%	0%	100%	0%	0%
0.5 μM	0%	82%	18%	80%	20%	0%
1 μM	0%	61%	39%	18%	82%	0%
2 μM	0%	50%	50%	0%	100%	0%
5 μM	0%	35%	65%	0%	100%	0%
10 μM	0%	40%	60%	0%	100%	0%

	1st melting curve			2nd melting curve		
5'GCTAGC	Unmodified duplex T_m 43.9 °C	Duplex+Ligand (1st adduct) T_m 73.1 -76.5 °C	Duplex+Ligand (2nd adduct) T_m 92.2 -92.5 °C	Unmodified duplex T_m 44.3 °C	Duplex+Ligand (1st adduct) T_m 75.3 -76.8 °C	Duplex+Ligand (2nd adduct) T_m 89.9 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	94%	6%	0%	96%	4%	0%
0.5 μM	0%	91%	9%	43%	57%	0%
1 μM	0%	66%	34%	19%	81%	0%
2 μM	0%	50%	50%	0%	100%	0%
5 μM	0%	41%	59%	0%	88%	12%
10 μM	0%	43%	57%	0%	71%	29%

Table 3.8: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Thz-Py-Py (RMH43) with oligonucleotide duplex containing the sequence GAATGC

	1st melting curve			2nd melting curve		
5'GAATGC	Unmodified duplex T_m 40.7 °C	Duplex+Ligand (1st adduct) T_m 71.8 -73.0 °C	Duplex+Ligand (2nd adduct) T_m 90.6 -91.3 °C	Unmodified duplex T_m 40.9 °C	Duplex+Ligand (1st adduct) T_m 72.5 -73.0 °C	Duplex+Ligand (2nd adduct) T_m 89.9 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	89%	11%	0%	100%	0%	0%
0.5 μM	0%	87%	13%	88%	12%	0%
1 μM	0%	55%	45%	82%	18%	0%
2 μM	0%	46%	54%	84%	16%	0%
5 μM	0%	43%	57%	53%	37%	10%
10 μM	0%	43%	57%	44%	39%	17%

3.4.4.3 *PBD-Py-Thz-Py (RMH47)*

GGTACC and GGATCC

Figures 3.25 and 3.26 show the effect of PBD-Py-Thz-Py (RMH47) on the melting profiles of the different oligonucleotides; the parameters derived from these are presented in Tables 3.9 to 3.12. Oligonucleotides containing GGTACC and GGATCC show similar profiles for the first melting curves, but differ in the 2nd melting curves. Three transitions are observed for both oligonucleotides; for GGTACC, these are 41.4 °C, 68.9 °C and 89.3 °C corresponding to the unmodified duplex, 1st adduct and 2nd adduct respectively while for GGATCC, these are 42.3 °C, 72.7 °C and 92.0 °C. In the presence of 0.1 μM ligand about 46% of GGTACC, appears as the 1st adduct; on increasing the concentration to 0.5 μM the unmodified duplex disappears and the 2nd adduct starts to appear. The 1st and 2nd adduct account for 57% and 43% respectively at 1 μM and these proportions remain the same at higher concentrations. In the presence of 0.1 μM ligand about 75% of the melting profile for GGATCC corresponds to the unmodified oligonucleotide with the remainder as the 1st adduct; with 0.5 μM ligand, three different transitions are evident with relative proportions of 27%, 50% and 23% the for unmodified duplex, 1st adduct and 2nd adduct respectively. No unmodified duplex is seen with ligand concentrations of 1 μM and above and the proportions of the 1st and 2nd adduct remain at about 40% and 60%. No 2nd adducts are observed in the second melting curves of GGTACC and GGATCC at all ligand concentrations. However, with GGTACC, the 1st adduct predominates at 0.5 μM, accounting for 58% of the melting profile and reaches 100% with 5 μM ligand. Interestingly, GGATCC shows two transitions one at 62 °C and another at 73.6 °C, as well as a transition for unmodified duplex at 42.3 °C. With 0.5 μM, about 62% melting profile is from the unmodified oligonucleotide with the remainder almost equally distributed between the transitions at 62 °C and 73.6 °C. No unmodified duplex is observed at concentrations of 1 μM and above; the adduct at 73.6 °C remains unchanged at about 20% while the other adduct at 62 °C increases to 80%.

GTATAC and GTTAAC

With oligonucleotides GTATAC and GTTAAC, 0.1 μM ligand produced melting profiles for which about 50% of the transition corresponds to the unmodified state with the remainder as the 1st adduct. The 2nd adduct starts to appear with 1 μM but the majority is still present as the 1st adduct, with about 70% as the 2nd adduct. With GTATC the proportion of the 2nd adduct then increases to about 50% with the remainder as the 1st adduct; for GTTAAC the amount of the 2nd adduct stays fairly constant from 0.5 μM to 10 μM. In the second melting curves, GTATAC shows formation of 2nd adduct with 2 μM ligand and above, while no 2nd adduct is observed for GTTAAC at any ligand

concentration. In the second melting curves 1 μM ligand and GTATAC show transitions corresponding to the unmodified duplex and the 1st adduct with proportions of 43% and 57% respectively; the amount of unmodified duplex then decreases with simultaneous formation of the 2nd adduct, with ratios of 27%, 53% and 20% at 10 μM ligand. For GTTAAC the 1st adduct starts to appear with 0.5 μM ligand at which concentration it corresponds to about half of the transition with the remainder as the unmodified duplex; this amount of the 1st transition then increases up around 65% at 10 μM ligand with the remaining 35% duplex as the unmodified duplex.

GCTTGC, GCTAGC and GAATGC

Oligonucleotide GCTTGC displays three transitions at 44.6 °C, 75.0 °C and 92.0 °C for unmodified duplex, 1st adduct and 2nd adduct respectively. The 1st adduct starts to appear with 0.1 μM ligand with about 75% of the melting profile in this form. The proportion of the 1st adduct then decreases to 30%, with an increase in the fraction of the transition corresponding to the 2nd adduct. The proportion of the 1st and 2nd adducts is then fairly constant (30%: 70%) at concentrations of 0.5 μM and above. In the second melting curve, about 30% of the melting profile corresponds to the 1st adduct with 0.5 μM ligand, increasing to 100% with 2 μM. With 5 μM and 10 μM, 2nd adduct is observed and accounts for 33% and 40% of the profile with the remainder as the 1st adduct. For oligonucleotide GCTAGC, about one-third of profile appears as the 1st adduct with a concentration of 0.1 μM, with other two-thirds as the unmodified duplex. The amount of 1st adduct increases with 0.5 μM ligand along with the appearance of a transition corresponding to the 2nd adduct (~37%). With 10 μM ligand the 1st and 2nd adduct correspond to 56% and 44% of the melting profile. For the second melting curve, about of half of the transition is observed as the 1st adduct, rising to 67% at higher concentrations, with the remainder as the unmodified duplex. The amount of the 1st adduct remains unchanged at higher concentrations while the unmodified duplex decreases and is replaced by the transition corresponding to the 2nd adduct. With 5 μM and 10 μM ligand about two-thirds of the melting profile is present as the 1st adduct while other one-third corresponds to the 2nd adduct. For the oligonucleotide containing GAATGC, three transitions are observed at 40.7 °C, 76.2 °C and 92.0 °C corresponding to the unmodified duplex, 1st adduct and 2nd adduct respectively. With 0.1 μM ligand about 58% of the melting profile corresponds to the 1st adduct, increasing to 76% at 0.5 μM, with the remaining 24% as the 2nd adduct. The proportion of 1st adduct then gradually decreases while 2nd adduct formation increases. The proportions of the 1st and 2nd adduct stay around 65% and 35% respectively with ligand concentrations between 1-5 μM, with a slight increase in the 2nd adduct with 10 μM. In the second melting curve, 52% of the melting profile appears as the 1st adduct in the

presence of 0.5 μM ligand, with the remainder as unmodified duplex. The amount of the 1st adduct then increases with ligand concentration, rising to 88% with 10 μM; here, no 2nd adduct is observed at any concentration in the second melt.

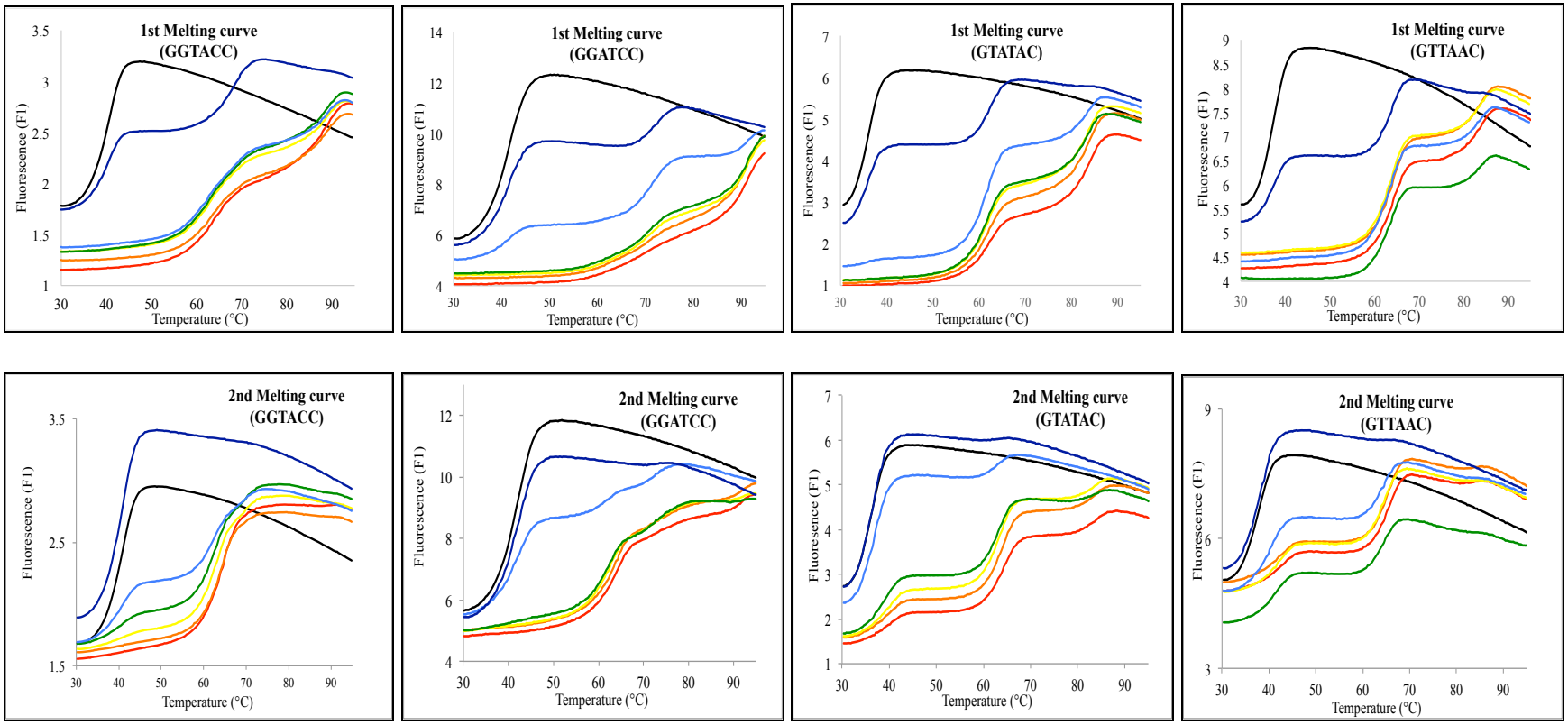


Figure 3.25: First and second melting curves of oligonucleotide duplexes containing the sequences GGTACC, GGATCC, GTATAC and GTTAAC in the presence of different concentrations of PBD-Py-Thz-Py (RMH47)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green); 0.5 μM (light blue) and 0.1 μM (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting.

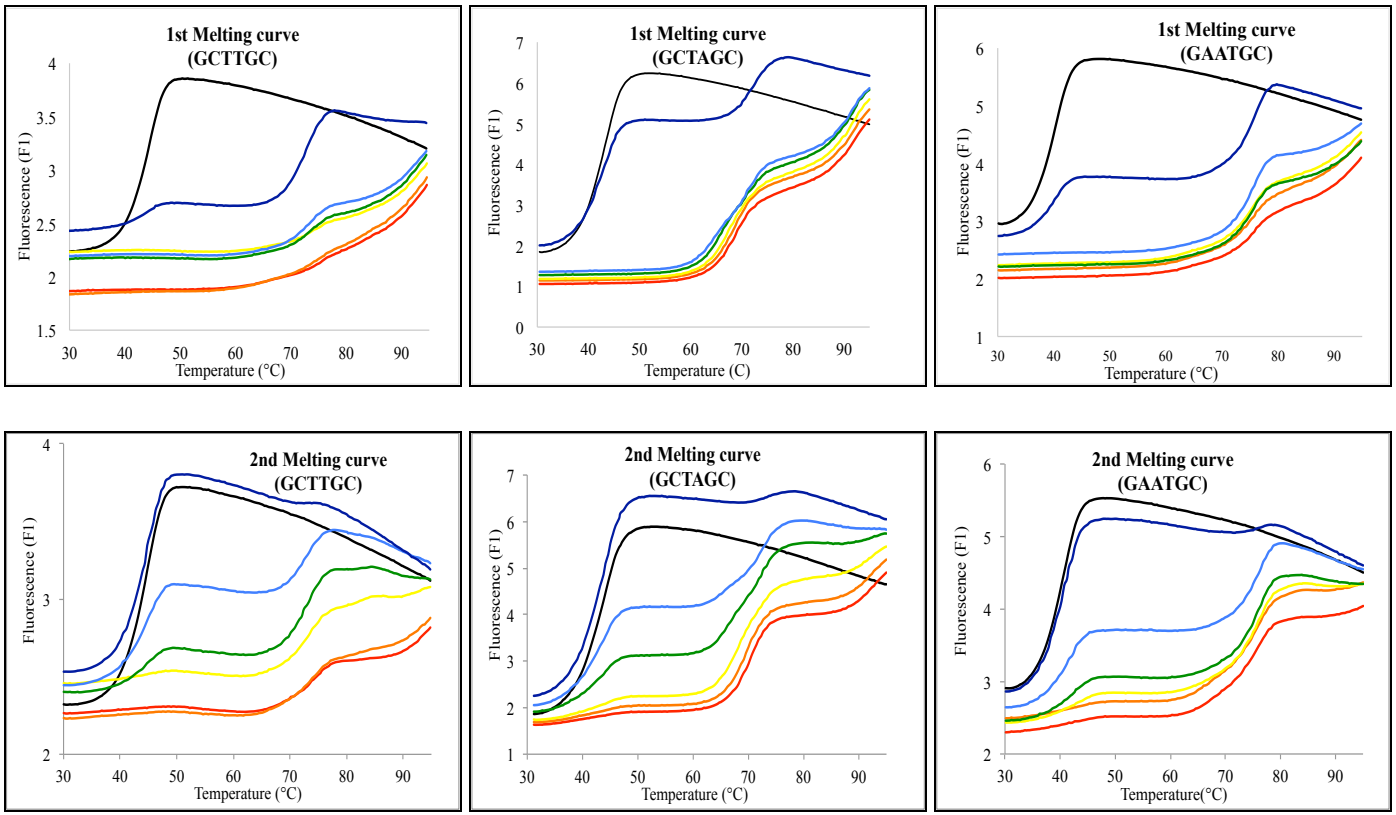


Figure 3.26: First and second melting curves of oligonucleotide duplexes containing the sequences GCTTGC, GCTAGC and GAATGC in the presence of different concentrations of PBD-Py-Thz-Py (RMH47)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green); 0.5 μM (light blue) and 0.1 μM (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting.

Table 3.9: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Thz-Py (RMH47) with oligonucleotide duplexes containing the sequences GGTACC and GGATCC

	1st melting curve			2nd melting curve		
5'GGTACC	Unmodified Duplex T_m 41.4 °C	Duplex+Ligand (1st adduct) T_m 63.4 -68.9 °C	Duplex+Ligand (2nd adduct) T_m 89.0 -89.3 °C	Unmodified duplex T_m 41.7 °C	Duplex+Ligand (1st adduct) T_m 63.2 -64.5 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	54%	46%	0%	100%	0%	0%
0.5 μM	0%	62%	38%	42%	58%	0%
1 μM	0%	57%	43%	17%	83%	0%
2 μM	0%	64%	35%	17%	83%	0%
5 μM	0%	54%	46%	0%	100%	0%
10 μM	0%	53%	47%	0%	100%	0%

	1st melting curve			2nd melting curve		
5'GGATCC	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 70.7 -72.7 °C	Duplex+Ligand (2nd adduct) T_m 91.6 -92.0 °C	Duplex (No ligand binds) T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 62.0 -73.6 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	74%	26%	0%	100%	0%	0%
0.5 μM	27%	50%	23%	62%	21%*, 17%*	0%
1 μM	0%	46%	54%	0%	80%*, 20%*	0%
2 μM	0%	46%	54%	0%	80%*, 20%*	0%
5 μM	0%	40%	60%	0%	80%*, 20%*	0%
10 μM	0%	40%	60%	0%	83%*, 17%*	0%

*Two transitions are observed in close proximity and are included in the same column

Table 3.10: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Thz-Py (RMH47) with oligonucleotide duplexes containing the sequences GTATAC and GTTAAC

	1st melting curve			2nd melting curve		
5'GTATAC	Unmodified duplex T_m 36.5 °C	Duplex+Ligand (1st adduct) T_m 61.7 -62.4 °C	Duplex+Ligand (2nd adduct) T_m 84.0 °C	Unmodified duplex T_m 36.6 °C	Duplex+Ligand (1st adduct) T_m 63.2 -64.5 °C	Duplex+Ligand (2nd adduct) T_m 83.7 -84.0 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	54%	46%	0%	100%	0%	0%
0.5 μM	0%	70%	30%	85%	15%	0%
1 μM	0%	57%	43%	43%	57%	0%
2 μM	0%	54%	46%	30%	58%	12%
5 μM	0%	49%	51%	24%	58%	18%
10 μM	0%	44%	56%	27%	53%	20%

	1st melting curve			2nd melting curve		
5'GTTAAC	Unmodified duplex T_m 37.8 °C	Duplex+Ligand (1st adduct) T_m 63.5 -64.3 °C	Duplex+Ligand (2nd adduct) T_m 84.0 °C	Unmodified duplex T_m 38.1 °C	Duplex+Ligand (1st adduct) T_m 64.5 -65.6 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	47%	53%	0%	100%	0%	0%
0.5 μM	0%	73%	27%	53%	47%	0%
1 μM	0%	75%	25%	50%	50%	0%
2 μM	0%	69%	31%	36%	64%	0%
5 μM	0%	69%	31%	29%	71%	0%
10 μM	0%	69%	31%	35%	65%	0%

Table 3.11: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Thz-Py (RMH47) with oligonucleotide duplexes containing the sequences GCTTGC and GCTAGC

	1st melting curve			2nd melting curve		
5'GCTTGC	Unmodified duplex T_m 44.6 °C	Duplex+Ligand (1st adduct) T_m 72.6 -75.0 °C	Duplex+Ligand (2nd adduct) T_m 92.0 °C	Unmodified duplex T_m 44.8 °C	Duplex+Ligand (1st adduct) T_m 72.7 -74.3 °C	Duplex+Ligand (2nd adduct) T_m 92.0 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	23%	77%	0%	100%	0%	0%
0.5 μM	0%	31%	69%	67%	33%	0%
1 μM	0%	37%	63%	38%	62%	0%
2 μM	0%	31%	69%	0%	100%	0%
5 μM	0%	33%	67%	0%	67%	33%
10 μM	0%	33%	67%	0%	60%	40%

	1st melting curve			2nd melting curve		
5'GCTAGC	Unmodified duplex T_m 43.9 °C	Duplex+Ligand (1st adduct) T_m 64.5 -69.0 °C	Duplex+Ligand (2nd adduct) T_m 91.5 -91.8 °C	Unmodified duplex T_m 44.3 °C	Duplex+Ligand (1st adduct) T_m 69.5 -73.4 °C	Duplex+Ligand (2nd adduct) T_m 90.0 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	67%	33%	0%	100%	0%	0%
0.5 μM	0%	63%	37%	55%	45%	0%
1 μM	0%	59%	41%	33%	67%	0%
2 μM	0%	58%	42%	14%	69%	17%
5 μM	0%	56%	44%	0%	68%	32%
10 μM	0%	56%	44%	0%	69%	31%

Table 3.12: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Thz-Py (RMH47) with oligonucleotide duplex containing the sequence GAATGC

	1st melting curve			2nd melting curve		
5'GAATGC	Unmodified duplex T_m 40.7 °C	Duplex+Ligand (1st adduct) T_m 75.6 -76.2 °C	Duplex+Ligand (2nd adduct) T_m 92.0 °C	Unmodified duplex T_m 40.9 °C	Duplex+Ligand (1st adduct) T_m 75.7 -76.2 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	42%	58%	0%	100%	0%	0%
0.5 μM	0%	76%	24%	48%	52%	0%
1 μM	0%	67%	33%	32%	68%	0%
2 μM	0%	68%	32%	17%	83%	0%
5 μM	0%	64%	36%	13%	87%	0%
10 μM	0%	57%	43%	12%	88%	0%

3.4.4.4 *PBD-Py-Im-Im (RMH53)*

GGTACC and GGATCC

The effect of PBD-Py-Im-Im (RMH53) on the melting profiles of different oligonucleotides is shown in Figures 3.27 and 3.28 and the percentages of the various transitions are presented in Tables 3.13 to 3.16. The oligonucleotide containing the binding site GGTACC shows three transitions at 41.4 °C, 69.5 °C and 89.6 °C corresponding to the unmodified duplex, 1st adduct and 2nd adduct respectively. With 0.5 μM ligand the transition corresponding to the 1st adduct accounts for ~48% of the melting profile with the remainder as the unmodified duplex. With 1 μM ligand the 1st adduct corresponds to 62% of the profile, with 38% as the 2nd adduct. The proportion of the 1st adduct then decreases with a corresponding increase in the transition for the 2nd adduct, which accounts for about two-thirds of the melting profile with 10 μM ligand. In the second melting curve, the 1st adduct is only observed with 5 μM and 10 μM with around 70% as the unmodified duplex. The oligonucleotide containing GGATCC displays similar melting characteristics. Very little of the 1st adduct is evident with 0.5 μM ligand, which increases to 56% with 1 μM with the remainder of the melting profile corresponding to the unmodified duplex. The transition from the 2nd adduct is observed with 2 μM, 5 μM and 10 μM ligand, accounting for about 65% of the total melting profile. In the second melting curve the unmodified duplex predominates up to 2 μM with little 1st adduct; with 5 μM and 10 μM ligand the amount of 1st adduct increases to about 50% of the melting profile.

GTATAC and GTTAAC

Ligand RMH53 has similar effects on oligonucleotides GTATAC and GTTAAC. Three transitions are observed for GTATAC at 36.5 °C, 64.1 °C and 92.6 °C for unmodified duplex, 1st adduct and 2nd adduct respectively while for GTTAAC, these are 37.8 °C, 65.7 °C and 84.5 °C. The 1st adduct is observed with 0.5 μM ligand for both oligonucleotides; all three transitions are seen with 1 μM, with more than 70% as the 1st adduct and only 10% as the 2nd adduct. The proportion of the 1st adduct then decreases while that of 2nd adduct increases and with 10 μM ligand 59% of the transition with GTTAAC corresponds to the 1st adduct with 41% as the 2nd adduct. With GATAC each adduct accounts for about half of the melting profile. In the second melting curves, the unmodified duplex predominates at all ligand concentrations with no evidence of the 2nd adduct. For GTATAC, the unmodified duplex and 1st adduct each account for about 50% of the melting profile, while these ratios are 65% and 35% respectively for GTTAAC.

GCTTGC, GCTAGC and GAATGC

The oligonucleotides containing the binding sites GCTTGC and GCTAGC also have similar melting profiles. Both oligonucleotides show melting temperatures around 72 °C and 92 °C for the 1st and 2nd adduct respectively in their first melting curves. With GCTTGC about two-thirds of total duplex is found as the 1st adduct with 0.5 μM ligand, with the remainder as the unmodified duplex. For GCTAGC, these proportions are 31% and 69%. Increasing the ligand concentration from 1 μM to 10 μM increases the proportion of the 2nd adduct with subsequent decrease of 1st adduct, showing about 70% as 2nd adduct and 30% as 1st adduct in the presence of 10 μM ligand. In the second melting curves, no transition for the 2nd adduct is observed with GCTTGC, though this accounts for about 13% of the melting profile for GCTAGC in the presence of 10 μM ligand. Most of the melting profile corresponds to the unmodified duplex at ligand concentrations up to 5 μM, though about 50% appears as the first adduct with 10 μM ligand; For GAATGC, three transitions are observed at 40.7 °C, 76.2 °C and 89.7 °C. Transitions for the 1st adduct are observed with 0.5 μM, accounting for about 72% of the melting profile, rising to above 90% with 1 μM, with less than 10% as the 2nd adduct. The proportion of the 1st adduct decreases at higher concentrations as it is replaced by that of the 2nd adduct, and these two transitions each account for about 50% of the total with 10 μM ligand. In the second melting curve, the unmodified duplex predominates up to 2 μM with little presence of the 1st adduct; with 10 μM ligand the proportion of the1st adduct is just over 50% with the remainder as the unmodified duplex.

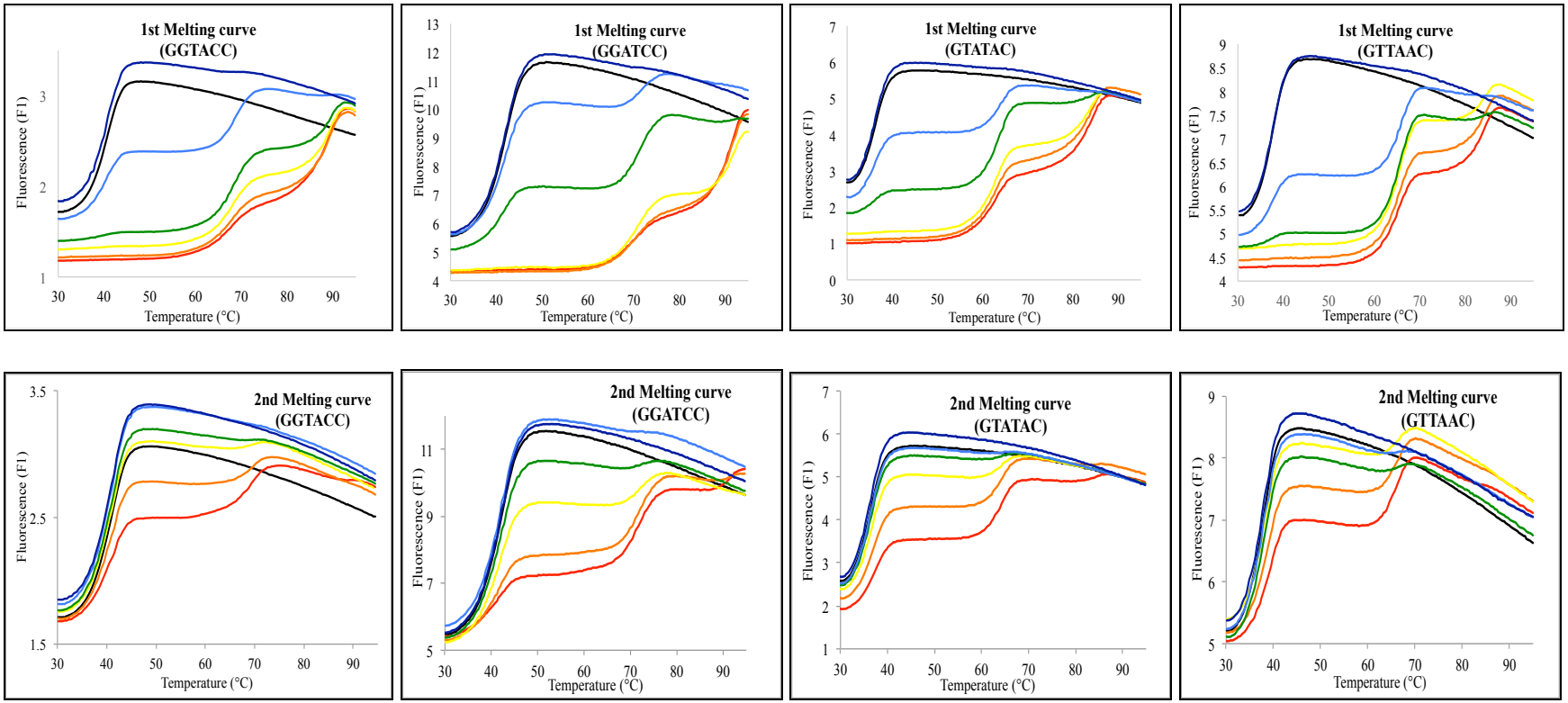


Figure 3.27: First and second melting curves of oligonucleotide duplexes containing the sequences GGTACC, GGATCC, GTATAC and GTTAAC in the presence of different concentrations of PBD-Py-Im-Im (RMH53)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green); 0.5 μM (light blue) and 0.1 μM (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting.

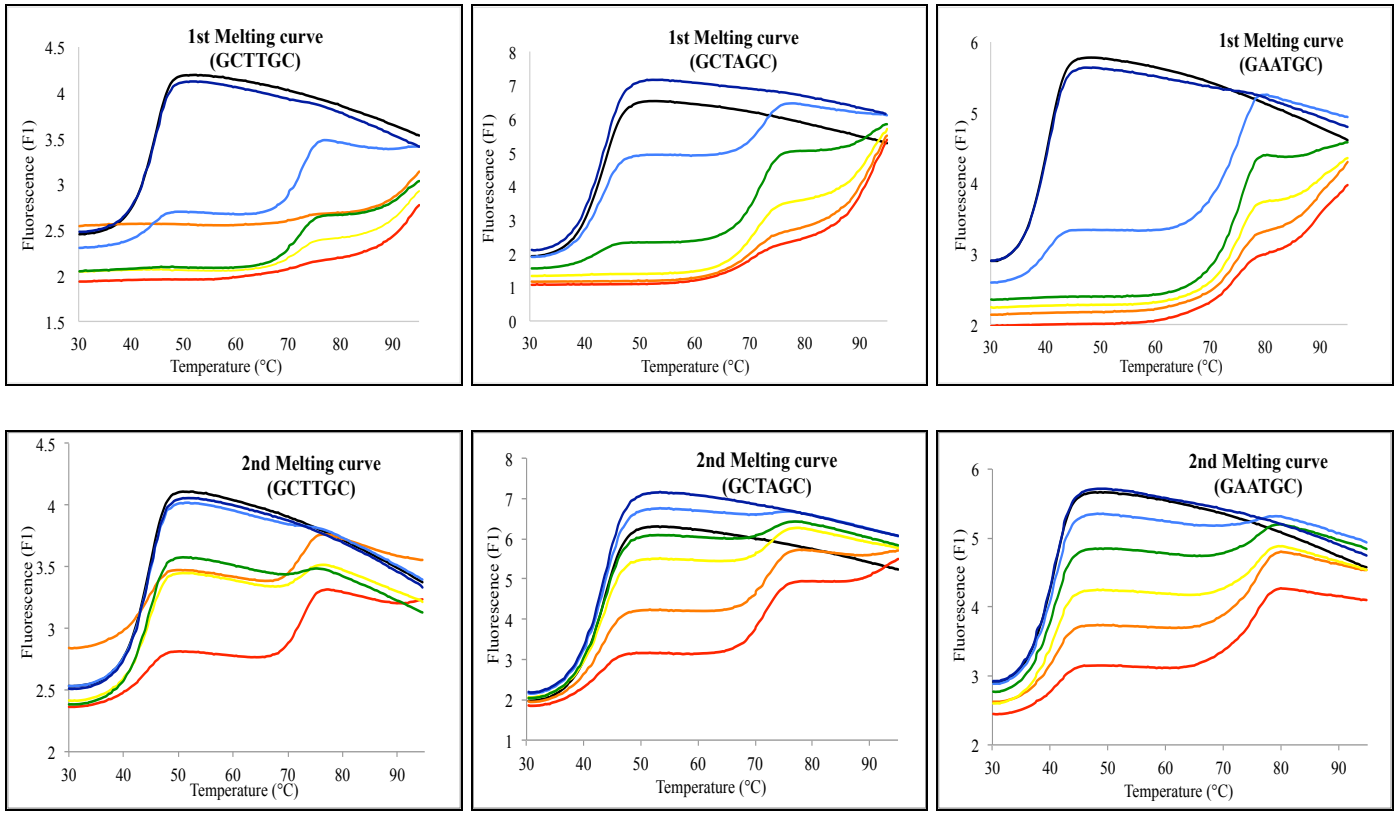


Figure 3.28: First and second melting curves of oligonucleotide duplexes containing the sequences GCTTGC, GCTAGC and GAATGC in the presence of different concentrations of PBD-Py-Im-Im (RMH53)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green); 0.5 μM (light blue) and 0.1 μM (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting.

Table 3.13: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Im-Im (RMH53) with oligonucleotide duplexes containing the sequences GGTACC and GGATCC

	1st melting curve			2nd melting curve		
5'GGTACC	Unmodified duplex T_m 41.4 °C	Duplex+Ligand (1st adduct) T_m 67.4 -69.5 °C	Duplex+Ligand (2nd adduct) T_m 88.8 -89.6 °C	Unmodified duplex T_m 41.7 °C	Duplex+Ligand (1st adduct) T_m 69.5 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	52%	48%	0%	100%	0%	0%
1 μM	0%	62%	38%	100%	0%	0%
2 μM	0%	48%	52%	100%	0%	0%
5 μM	0%	40%	60%	86%	15%	0%
10 μM	0%	34%	66%	67%	33%	0%

	1st melting curve			2nd melting curve		
5'GGATCC	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 70.5 -71.8 °C	Duplex+Ligand (2nd adduct) T_m 91.4 -92.6 °C	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 72.1 -72.4 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	82%	18%	0%	100%	0%	0%
1 μM	44%	56%	0%	96%	4%	0%
2 μM	0%	32%	68%	83%	17%	0%
5 μM	0%	39%	64%	52%	48%	0%
10 μM	0%	31%	69%	43%	57%	0%

Table 3.14: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Im-Im (RMH53) with oligonucleotide duplexes containing the sequences GTATAC and GTTAAC

	1st melting curve			2nd melting curve		
5'GTATAC	Unmodified duplex T_m 36.5 °C	Duplex+Ligand (1st adduct) T_m 63.3 -64.1 °C	Duplex+Ligand (2nd adduct) T_m 91.8 -92.6 °C	Unmodified duplex T_m 36.6 °C	Duplex+Ligand (1st adduct) T_m 64.2°C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	55%	45%	0%	100%	0%	0%
1 μM	19%	71%	10%	100%	0%	0%
2 μM	0%	63%	37%	84%	16%	0%
5 μM	0%	54%	46%	50%	50%	0%
10 μM	0%	49%	51%	52%	48%	0%

	1st melting curve			2nd melting curve		
5'GTTAAC	Unmodified duplex T_m 37.8 °C	Duplex+Ligand (1st adduct) T_m 65.1 -65.7 °C	Duplex+Ligand (2nd adduct) T_m 83.9 -84.5 °C	Unmodified duplex T_m 38.1 °C	Duplex+Ligand (1st adduct) T_m 66.2 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	39%	61%	0%	100%	0%	0%
1 μM	10%	83%	7%	97%	3%	0%
2 μM	0%	74%	23%	88%	12%	0%
5 μM	0%	65%	35%	77%	23%	0%
10 μM	0%	59%	41%	65%	35%	0%

Table 3.15: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Im-Im (RMH53) with oligonucleotide duplexes containing the sequences GCTTGC and GCTAGC

	1st melting curve			2nd melting curve		
5'GCTTGC	Unmodified duplex T_m 44.6 °C	Duplex+Ligand (1st adduct) T_m 71.9 -73.2 °C	Duplex+Ligand (2nd adduct) T_m 92.0 °C	Unmodified duplex T_m 44.8 °C	Duplex+Ligand (1st adduct) T_m 72.8 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	33%	67%	0%	100%	0%	0%
1 μM	0%	61%	39%	96%	4%	0%
2 μM	0%	38%	62%	87%	13%	0%
5 μM	0%	18%	82%	67%	33%	0%
10 μM	0%	25%	75%	44%	56%	0%

	1st melting curve			2nd melting curve		
5'GCTAGC	Unmodified duplex T_m 43.9 °C	Duplex+Ligand (1st adduct) T_m 71.1 -72.6 °C	Duplex+Ligand (2nd adduct) T_m 92.2 -92.6 °C	Unmodified duplex T_m 44.3 °C	Duplex+Ligand (1st adduct) T_m 71.6 -73.0 °C	Duplex+Ligand (2nd adduct) T_m 90.0 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	69%	31%	0%	100%	0%	0%
1 μM	18%	65%	17%	93%	7%	0%
2 μM	0%	49%	51%	83%	17%	0%
5 μM	0%	37%	63%	59%	41%	0%
10 μM	0%	29%	71%	41%	46%	13%

Table 3.16: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Im-Im (RMH53) with oligonucleotide duplex containing the sequence GAATGC

	1st melting curve			2nd melting curve		
5'GAATGC	Unmodified duplex T_m 40.7 °C	Duplex+Ligand (1st adduct) T_m 74.7 -76.2 °C	Duplex+Ligand (2nd adduct) T_m 88.8 -89.7 °C	Unmodified duplex T_m 40.9 °C	Duplex+Ligand (1st adduct) T_m 75.8 -76.2 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	28%	72%	0%	83%	17%	0%
1 μM	0%	91%	9%	83%	17%	0%
2 μM	0%	71%	29%	70%	30%	0%
5 μM	0%	57%	43%	50%	50%	0%
10 μM	0%	50%	50%	42%	58%	0%

3.5 Discussion

3.5.1 Footprinting studies

The PBD conjugates examined in this chapter contain a common structural motif: two compounds RMH35 (PBD-Py-Thz) and RMH71 (PBD-Thz-Thz) contain a polyamide chain with two heterocycles, while the other four compounds – RMH41, RMH43, RMH47 and RMH53 have conjugation of PBD with polyamides of three heterocycles. The connection between the PBD moiety and polyamide chain is made up of trimethylene linker. Thus three different components are present in these compounds- i) the PBD moiety, ii) the trimethylene linker and iii) the polyamide containing different combinations of pyrrole and imidazole or thiazole. A trimethylene linker has previously been found to provide suitable isohelicity for positioning of the polyamide chain the DNA minor groove (167). However, manual docking experiments showed that the PBD moiety made poor contacts with the minor groove if the polyamide chain (Py-Py-Py) fitted snugly, on the other hand, if the PBD was arranged to have optimal interaction with minor groove then the polyamide chain bulged out of the groove (167). Also molecular modelling of PBD-Im-Py-Py with two 15 base pair oligonucleotide duplexes (5'CAGG**G**G*ATTGG*CTGGT3' and 5'CTA**C**G*ATTGG*TTCTT3', italic underlined indicates target site) showed that the molecule could span six to seven base pairs forming a 1:1 complex (164) in which the binding of PBD with guanine also protected the neighbouring base adjacent to guanine (guanine and cytosine in bold). The footprinting experiments with these PBD conjugates produced similar protections in the AT-rich sequences for all the derivatives. The presence of protections at the same sites indicates the predominance of a common binding motif *i.e.* PBD. However, these sites differ in their concentration dependence, which suggests that the polyamide has some role in affecting the ligand's position in the DNA minor groove. The DNA fragments used in the footprinting studies - HexA, HexB and MS1 contain numerous AT-rich regions; most of these regions are protected with higher concentrations (10 μM and 5 μM) of the PBD conjugates. However, if we consider lower concentrations (2 μM, 1 μM and 0.5 μM), the compounds still retain the tendency to bind at AT-sites, but have a few variations in their binding preferences.

RMH41 (PBD-Im-Py-Py) binds at AT-rich regions adjacent to a guanine(s) and showed highest affinity at 5'AATTAG in HexB with a footprint at 1 μM. Other footprints included 5'GCAATTGCCG, 5'GTATACGTAGCG, 5'ACATATGTA and 5'GAATTAATGCAT in HexB and 5'GTACTTTACATAACTC and 5'AATTGCTA in MS2. The presence of AT-rich sequences seems a prerequisite for positioning the polyamide chain that is followed by covalent bond formation by the PBD moiety; all of the binding sites contain

G or GC with four or more base pairs of adenines and/or thymines. Larger footprints contain multiple guanines, which obviously offer multiple binding sites for the PBD. 5'GCAATTGCCG contains several guanines and so overlapping binding sites can occur with the PBD moiety directed towards either the 5' or 3'-end (164). The footprint at 5'GAATTAATGCAT is flanked by AATATTC at its 5'-end, which could also have been a potential binding site for the conjugate, but which did not show any binding. Covalent binding of the PBD moiety to one guanine will causes steric hindrance preventing another PBD from binding to an adjacent guanine (underlined) in either the same or the complementary strand (5'AATATTCGAATTAATGCAT); the same phenomenon was observed for footprint at 5'AATTGC (5'AATTGCTATA) that has TATA at the 3'-end but which is not included in the footprint.

The presence of thiazole in place of imidazole in RMH43 (PBD-Thz-Py-Py) does not significantly affect the binding. This compound produced large footprints at high concentrations, like RMH41, but showed greater affinity to some sites at lower concentrations. The compound bound with highest affinity at 5'TAAGTAC in HexA with 0.5 μM; other footprints include 5'AAACGTT in HexA and 5'TATAGATCTAGAATTCC and 5'TTAATG in HexBfor. In the case of 5'TAAGTAC, the ligand could have bound to either of the guanines (in the middle or complementary to cytosine), but most probably is attached to the guanine opposite the cytosine (underlined), spanning seven base pairs in the 5'-direction protecting the rest of the sequence. Binding site 5'AAACGTT can offer one guanine for covalent bonding, as with 5'TTAATG. The large footprint at 5'TATAGATCTAGAATTCC contains multiple AT-rich sequences interspersed with guanines, thus presenting several overlapping protections.

The polyamide chain composed of thiazole in between two pyrroles reduced the binding affinity, as observed with RMH47 (PBD-Py-Thz-Py). This compound also bounds to AT-rich sequences like the other ligands but with decreased affinity. Binding sites are observed at 5'TTAGG, 5'AATTGC and 5'ACATATGTATA in HexBrev with equal affinity at 2 μM. The footprint at 5TTAGG is flanked by TAA at the 5'-end that indicates presence of four or more adenines and/or thymines for binding; the other two footprints also comply with this requirement. The arrangement of the adenines and thymines seems to affect the binding as more favourable binding is observed with 5'ACATATGTATA, containing alternating purines and pyrimidines that with stretches of continuous adenines or thymines such as 5'GACAAAA or 5'CTTTACATAA, which only show footprints at higher concentrations.

The compound RMH53 (PBD-Py-Im-Im) shows its highest affinity for 5'AATTAG in HexBrev and 5'CGAATGC, 5'TTGTCAT, 5'TTATGTAAAGTACG, 5'AATTAGGG and

5'GGTTAAGCT in MS1, which produce footprints with 2 μM ligand. All these sites are also rich in adenines and thymines, but the sequences that flank the footprints are different to the other ligands. The footprint at 5'AATTAG is flanked by GGC and GGCC at the 5'- and 3'-end respectively; also the footprints at 5'CGAATGC, AATTAGGG and GGTTAAGCT have GC rich flanking sequences. Replacement of *N*-methylpyrrole by *N*-methylimidazole provides space for NH_2 of guanine and introduces a hydrogen bond acceptor that may favour interaction with GC base pairs (68,268); thus the presence of two imidazoles in RMH53 might produce more favourable interactions with GC-rich sequences.

The rationale for synthesising the conjugates between PBD and polyamide chains was to exploit the sequence selectivity of polyamides. The polyamides are well known for their antiparallel homodimer formation in the DNA minor groove and can recognise A/T or T/A by pairing of pyrrole-pyrrole and G/C and C/G by pairing of imidazole-pyrrole and pyrrole-imidazole respectively. Attachment of an additional minor groove binder like PBD might also permit this dimeric orientation with increased binding affinity due to covalent bond formation by PBD. It was therefore envisioned that the PBD conjugates might also form 2:1 complexes, with the PBDs at either end of a polyamide dimer. PBD conjugates RMH35 and RMH71, both contain a thiazole moiety, which would be expected to have the same effect on their sequence selectivity. Thiazole is too large to stack side-by-side with another thiazole or pyrrole or imidazole and so results in staggering of the two molecules, with imperfect overlap, ultimately forcing binding in the 1:1 mode in the minor groove (161). These compounds were synthesized (165) to recognise the target sites of a number of general transcription factors that bind to CCAAT boxes, which can be present in either direct or reverse orientation. Both these compounds bind to the HexA and HexB sequences at the same locations. These DNA fragments do not contain any copies of the sequence CCAAT, but footprints are found in closely similar sequences- 5'GCAATTGCCGG and 5'GAATTCCGG. The sequence GAATTC matches with proposed sequence GXXXXC (Figure 3.29a and b) where two PBD units bind with guanine in complementary strands.

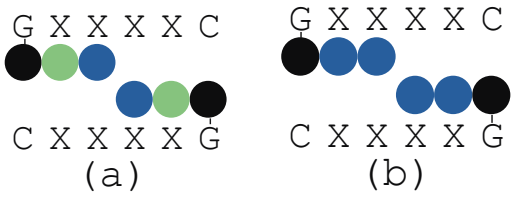


Figure 3.29: Proposed binding sites of (a) PBD-Py-Thz (RMH35) and (b) PBD-Thz-Thz (RMH71) (PBD unit, pyrrole and thiazole are presented as black, green and blue coloured circles respectively).

RMH41 possesses an imidazole ring next to the PBD, followed by two pyrrole rings and in principle might be able to form a 2:1 complex with DNA recognising the sequence GG(A/T)CC (Figure 3.30a) if there is full overlap between the two parts of the dimer; one sequence of this kind (5'GGACC) is found only in HexA, but this is not protected by the ligand. In the large footprints, sequences like 5'XG(A/T)CX or 3'XC(A/T)GX (Figure 3.30b) are observed and can be considered as possible binding sites where the arrangement of imidazole-pyrrole, pyrrole-pyrrole and pyrrole-imidazole identify GAC or CAG without any contribution from the PBD; in contrast to these large footprints, several other 5'XG(A/T)CX or 5'XC(A/T)XC sites are present in the DNA fragments that are not protected. In addition, overlapping of only one pair of pyrrole-pyrrole could generate footprints at sites having common sequences of 5'GXX(A/T)XXC or 3'CXX(A/T)XXG (Figure 3.30c); these results suggest that the covalent binding of the PBD units to Gs predominates in the interaction and that this occurs in regions in which any allowable combination of heterocycles is possible. Footprints are observed at 5'GGATCCC, 5'GAATTCC, 5'GCTATAC and 5'GTATTGC. These sites can be considered as GXX(A/T)XXC; though GGATCC and GAATTC can also be grouped as GX(A/T)(A/T)XC, ignoring cytosine at the end of these sequences. On the other hand, 5'GCTATAC and 5'GTATTGC are best classed as AT-rich sequences, having either guanine or cytosine at both the 5'-end and the 3'-end. Sequences of the type 5'CXX(A/T)XXG such as 5'CGGACCG are not protected by this ligand. In the footprints, 5'GX(A/T)(A/T)XC sequences such as 5'GAATTC, 5'GGTACC, 5'GCTAGC and 5'GCAAGC or 5'CX(A/T)(A/T)XG such as 5'CAATTG and 5'CATATG appears to be among the best binding sites present in the large footprints suggesting that the preferred combination may be the one shown in Figure 3.30 (d, e), with the two pyrrole-pyrrole pairs.

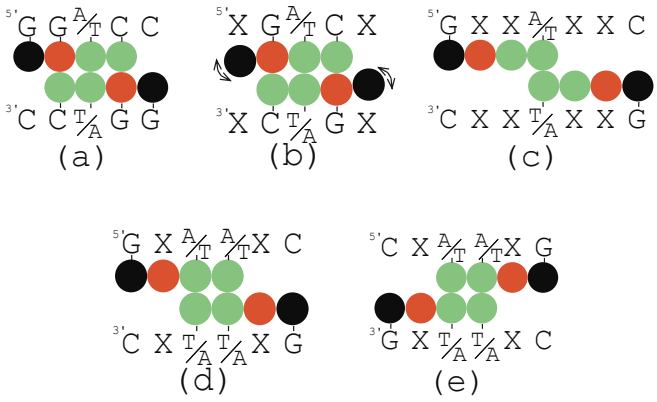


Figure 3.30: Proposed binding sites of PBD-Im-Py-Py (RMH41). The symbol (↷) indicates free rotation of adjacent group (s) without any specificity of binding sequence

(PBD unit, pyrrole and imidazole are presented as black, green and red coloured circles respectively)

In PBD-Thz-Py-Py (RMH43), two types of 2:1 complex are possible. Unlike thiazotropsin A (160) and thiazotropsin B (161), the ligands have no formyl or acetyl amide head where the pairing of thiazole with formyl or acetyl has been reported to recognise GC base pairs. Very few footprints were evident with 5 base pairs between two guanines (Figure 3.31a). In Figure 3.31b,c there are pyrrole-pyrrole pairs, recognising the 6 bp sequence GX(A/T)(A/T)XC. Examining the footprints at higher concentrations of RMH43, most of the footprints conform with the sequence 5'GX(A/T)(A/T)XC or 5'CX(A/T)(A/T)XG, suggesting recognition might involve two pyrrole-pyrrole pairs. In the case of RMH47, the presence of a thiazole in between the two pyrroles can only support 2:1 complex formation by the terminal pyrrole-pyrrole pair recognising GXX(A/T)XXC. Sequences like 5'GTTAACC, 5'GCGATGC, 5'GCTATAC, 5'GTATAGC in the large footprints can be considered as examples of 5'GXX(A/T)XXC whereas only one footprint contains 5'CCATGTG of the type CXX(A/T)XXG (Figure 3.31d,e).

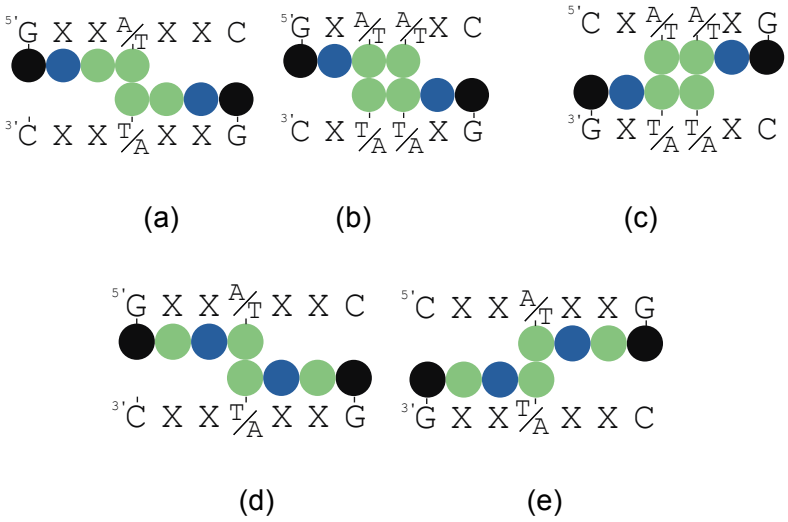


Figure 3.31: Proposed binding sites of PBD-Thz-Py-Py (RMH43) (a, b and c) and PBD-Py-Thz-Py (RMH47) (d and e) (PBD unit, pyrrole and thiazole are presented as black, green and blue coloured circles respectively)

With RMH53, three different 2:1 binding orientations can be envisaged, with different arrangements of the rings (as shown in Figure 3.32). Among the observed footprints, 5'GXXXXC is the most abundant one, which would be formed when both the imidazole rings are juxtaposed against an imidazole in the second molecule and where the two PBD units bind with guanines. Yang *et al* (269) reported a tri-imidazole polyamide DNA minor groove binder which bound to a C/G-containing four bp core, preferably CCGG; thus sequences like GGCC, GCGC CCGG may be favoured by imidazole-imidazole

without a suitable flanking sequence for the PBD. No footprints were observed at the sequence 5'GCXGC, though this is not present in HexB and MS1 and there is only one (5'GCAGC) in HexA, which does not produce a footprint with this compound.

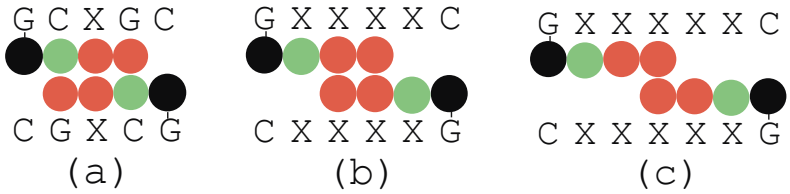


Figure 3.32: Proposed binding sites of PBD-Py-Im-Im (RMH53) (PBD unit, pyrrole and imidazole are presented as black, green and red coloured circles respectively)

The footprints of PBD conjugates at lower concentrations indicate their preferences for AT-rich sequences adjacent to a guanine for covalent bond formation with the PBD, which conforms to the 1:1 binding mode. With the footprints at higher concentrations a few 6 base pairs sequences with the common motif 5'GXXXXC appear to be the preferred binding sites; these sites can only be occupied if binding occurs in the 2:1 mode. To determine the preferred binding sites of these PBD conjugates, footprinting experiments were conducted with DNA fragments having various potential 6 base pairs binding sites separated with the sequence ATATAT. The experiments showed that all the compounds bound at similar sites demonstrating preferences for ATATAT sequences with a guanine at the 5'-end or 3'-end. Six base pairs sites such as GTATAC, GTTAAC, and CATATG having ATATAT at both the 5' -end and 3' -end showed two protections *i.e*. one PBD moiety binds with guanine (underlined) and the polyamide chain is positioned with ATATAT sequences. Six base pair sequences having two adjacent guanines in the same strand or complementary strand such as GGTACC, GGATCC, GCATTC, GCAAGC are also not protected, rather the conjugates covered the AT-rich regions in the linker between these sites and protected one or more guanine. The PBD moiety can bind with any guanine but binding of PBD with one guanine precludes attachment of a second PBD to an adjacent guanine. The sequence GCTAGC was protected fully by RMH41, though a large footprint was also produced covering the ATATAT sequences on both ends of the site; this could be either due to binding of PBD with any of the available guanines in a 1:1 binding mode as well as binding in a 2:1 mode protecting GCTAGC. However, 2:1 binding seems unlikely at this site as the ATATAT sites would remain unprotected if only the ligand only prefers GCTAGC. Thus AT-rich sequences are more favourable for PBD conjugates to bind in the 1:1 mode.

The footprints taken at different time intervals indicate that the PBD conjugates interact with their preferred binding sites at different rates. All the compounds bound at the AT-rich sequences but differed in their rate of complex formation. At 5 μM, RMH41 (PBD-Im-Py-Py) bound within 5 minutes to AT-rich sequences containing one guanine between the AT-sites such as 5'GAAGAGTTATGTAAAGTACG; similar sequences containing two guanines in the same or complementary strands, such as 5'GGTATAGCAATTAGG are also bound but less quickly as protection is observed after 3 hours with the same concentration; the same phenomenon is seen with 5'GGTTAAGC. On the other hand, RMH43 (PBD-Thz-Py-Py) does not differ in the rate of reaction with these sites at 5 μM and all the reactions are completed within 1 hour. For compound RMH53 (PBD-Py-Im-Im), the rate of complex formation is much slower than that of RMH41 and RMH43 and all the AT-rich sites are protected by RMH53 after 3 hours. RMH47 (PBD-Py-Thz-Py) differed in rate of reaction between different AT-rich sites. The fastest binding occurred at 5'GTAAAGTAC within 1 minute at 5 μM; another four AT base pair sequence 5'GGTTAAGC was protected but at slower rate after 1 hour. Complex formation with longer sequences containing multiple AT-rich regions was much slower than at the other two sites and 5'GGTATAGCAATTAG is bound after 3 hours. For the kinetic studies, it can be suggested that these PBD-conjugates show preferences for AT-rich sequences having four base pairs of adenines and/or thymines. Two pyrroles in the polyamide chain of RMH41, RMH43 and RMH47 play an important role in placement of these ligands in the minor groove, thus show similar rate of reactions. On the other hand, presence of two imidazoles causes delayed binding of RMH53 in these sites.

3.5.2 Melting studies

The oligonucleotide duplexes that were used in the melting experiments were designed to examine the preferences of PBD-conjugates towards different six base pair binding sites. Preliminary studies with oligonucleotide duplexes containing one guanine and with a single ligand concentration, produced monophasic melting curves, which clearly demonstrated that only one PBD conjugate molecule was accommodated within each duplex. The polyamide must play a vital role in the binding as RMH35 (PBD-Py-Thz) and RMH71 (PBD-Thz-Thz) showed little or no shift in the melting temperature despite the presence of guanine for covalent binding, in contrast to the other polyamide conjugates, with three heterocycles, which produced significant changes. The oligonucleotide duplexes containing potential six base pair ligand binding sites each have at least two guanines, which can in theory be occupied with two PBD molecules. All the oligonucleotides showed biphasic or triphasic melting curves *i.e.* at least two

further melting transitions at higher temperatures than the free oligonucleotide duplex. We assume that these correspond to the addition of one and two ligand molecules, generating the1st and 2nd adducts for which their relative proportions indicate the fraction of singly and doubly bound duplexes. From these melting data three different phenomena can be addressed: i) the formation of mono-adduct and bis-adduct, ii) the variation of the T_m values for different ligands iii) presence of an additional transition in oligonucleotide duplexes of GGTACC, GGATCC and GCTAGC.

The appearance of the 1st adduct starts, which is produced by the lower concentrations, indicates binding of one molecule of PBD, which is assumed to be at one of the guanines. As the binding is determined by covalent attachment of the PBD the rest of the molecule could in theory be oriented in either direction. The binding of a second PBD molecule to another guanine in the complementary strand or the same strand produces a 2nd adduct and, proportion of the 1st adduct inevitably decreases at higher ligand concentrations as the amount of the 2nd adduct increases; thus, the sample contains a mixture of 1st and 2nd adducts and their relative percentages indicates the fraction of 1st and 2nd adducts. An important consideration for the production of the 2nd adduct is the rate of attachment of second PBD unit. The samples were incubated for 12 hours before performing first melting phase, which allows adequate time for binding of the PBD. On heating the samples to 95 °C, and maintaining them at this temperature for 5 minutes, the ligand detaches from the DNA. The ligands then do not have sufficient time to reattach during the annealing reaction and so the annealing reaction is independent of the added ligand concentration. In the second melting phase, on the other hand, the samples were left for only 5 minutes after reducing the temperature to 30 °C, which is not sufficient time for complete reaction and reattachment to the DNA. In most cases, no transition for the 2nd adduct is seen in the second melting curve having mostly in unmodified duplex state with a minor proportion of the 1st adduct. The variations in the relative abundance of the unmodified duplex, 1st adduct and 2nd adduct in the first and second melting curves demonstrate that the ligands show different rates of reaction from unmodified duplex into 1st adduct and 2nd adduct.

The oligonucleotide duplexes contained different numbers of guanines in various arrangements. Thus the attachment of one PBD to the oligonucleotide duplexes GGTACC, GGATCC, GCTAGC and GCTTGC, could occur in any one of four different positions. Similarly there are three potential sites of attachment for GAATGC, but only two for GTTAAC and GTATAC. For example, in Figure 3.33, four different mono adducts can be formed with GCXXGC, though in each one the polyamide side chain

could be oriented in either direction. The mono adducts (❶ and ❸) are very similar as the central region is symmetrical, though they have different flanking sequences as one strand has 3'- and 5'- As while the other contains Ts. Similarly ❷ and ❹ also differ in their flanking base pairs. The first transition could therefore contain a mixture of mono adducts, which might have different effects on the T_m. The observed single smooth melting transitions suggest that either these sites have similar effects on the T_m or that one of these complexes predominates. Addition of a second PBD molecule to this complex could also give rise to a mixture of bis–adducts, each of which could have different effects on the T_m values.

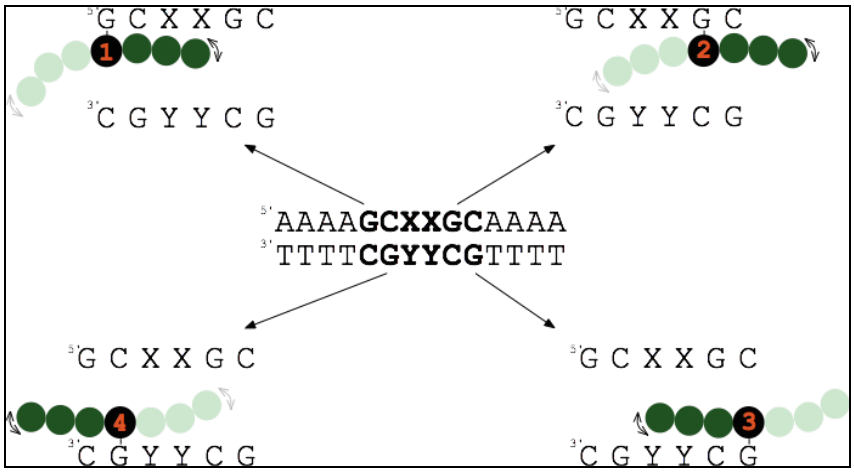


Figure 3.33: The diagram shows the possible formation of 1st adducts of PBD conjugates with (5'GCXXGC)(3'CGYYCG5')). Duplex is shown in the middle. Molecules shown as dark green indicates orientation towards 5' to 3' whereas light green refers binding towards other direction.

In the duplex GGATCC, compounds RMH41, RMH43 and RMH47 show an additional transition at high ligand concentrations (in Result section, table 3.1, 3.5 and 3.9 respectively); the same effect is observed for GGTACC with RMH41 (Table 3.1) and RMH43 (Table 3.5), and GCTAGC with RMH41 (Table 3.3). In all these cases, the additional T_m is lower than that of 1st adduct. All these oligonucleotide duplexes contain AT or TA in the middle. In the ligands, positioning of terminal two heterocycles might occur first at AT and TA without any covalent bond formation between the PBD and guanine; next, one PBD molecule can covalently bind with guanine producing the 1st adduct; subsequent binding of the second PBD molecule will results in the 2nd adduct. The rate of conversion between these species will determine the relative abundance of each adduct in the final mixture.

The adducts formed by each ligand with the different duplexes produces significant changes in their melting temperatures, which vary slightly between the different ligands (Figure 3.17). The compound RMH41 (PBD-Im-Py-Py) causes largest change (35.4

°C) in melting temperature with GGTACC forming 1st adduct while GTATAC produces least change in temperature (27.8 °C). On the other hand, greatest stability (48.6 °C) is produced with GCTAGC in 2nd adduct while GGTACC causes least change. Likewise, GGATCC (33 °C) and GTATAC (50.5 °C) are found to have greatest stability as 1st adduct and 2nd adduct formation with compound RMH43 (PBD-Thz-Py-Py). Compound RMH47 (PBD-Py-Thz-Py) demonstrates highest preferences with GAATGC for both 1st adduct (34.9 °C) and 2nd adduct (51.3 °C). On the other hand, RMH53 (PBD-Py-Im-Im) also shows greatest stability in 1st adduct (34 °C) with GAATGC while for 2nd adduct formation, GGATCC (49.1 °C) produces largest change in melting temperature. Thus it can be suggested that, though, the covalent bond formation between PBD and guanine results in large increase in melting temperatures of these duplexes, however, the polyamide chain have preferences for some specific sites.

Table 3.17: Changes in melting temperatures (ΔT_m) in the first melting curves at highest (10 μM) ligand concentration of PBD conjugates (Melting temperatures in each concentration is given in Appendix 2, Table I to VIII)

	RMH41		**RMH43**		**RMH47**		**RMH53**	
	Change in melting temperature (ΔT_m) (°C)							
	1st ΔT_m	2nd ΔT_m	1st ΔT_m	2nd ΔT_m	1st ΔT_m	2nd ΔT_m	1st ΔT_m	2nd ΔT_m
GGTACC	24.5, 35.4	45.6	30.6	48.3	22.3	47.6	26.0	47.4
GGATCC	19.6, 32.7	46.6	33.0	50.0	28.4	49.3	28.7	49.1
GTATAC	27.8	46.1	28.0	50.5	25.9	47.5	26.8	48.0
GTTAAC	28.4	45.8	27.0	48.5	26.5	46.2	27.5	46.1
GCTAGC	21, 29.2	48.6	29.6	48.6	25.2	47.9	27.4	48.7
GAATGC	31.3	47.5	31.3	49.9	34.9	51.3	34.0	48.1
GCTTGC	31.3	46.7	29.3	47.7	30.4	47.6	27.6	47.5

The melting temperatures of the oligonucleotides also depend on their sequences; stability of the DNA double helix with respect to separation of complementary strands is known to depend on the base composition and base stacking of the duplex. The stability of DNA is directly related with the G-C content (mole percentage of guanine plus cytosine) in the duplex (270,271); thus GGTACC, GGATCC, GCTTGC, GCTAGC and GAATGC have higher melting temperatures than GTATAC and GTTAAC (Table 3.18). In addition to this, base stacking between nearest-neighbour bases performs as the main stabilizing factor in the DNA double helix (272,273); nearest-neighbour base interaction between CpG provides highest stability and the order of the stability is CpG,

GpC, GpG, ApA, ApT, CpT, GpT, TpA, CpA and GpA (272). The melting temperatures of the duplexes used in this study are consistent with respect to base composition and base stacking; the order of stability is GCTTGC, GCTAGC, GGATCC, GGTACC, GAATGC, GTTAAC and GTATAC.

Table 3.18: Melting temperatures of different oligonucleotide duplexes

Melting temperatures (T_m) of the oligonucleotide duplexes						
GGTACC	GGATCC	GTATAC	GTTAAC	GCTTGC	GCTAGC	GAATGC
41.4 °C	42.3 °C	36.5 °C	37.8 °C	44.6 °C	43.9 °C	40.7 °C

In addition to variations in melting temperatures, percentages of adducts in first and second melting curves for each oligonucleotide also varied that again indicate their varying affinities for these sequences. From the footprinting studies, RMH41 (PBD-Im-Py-Py) protects GGATCC (site 1/MS2) at a concentration of 5 μM while GGTACC (site 3/HexAfor) is protected with 10 μM. In the melting studies with highest ligand concentration (10 μM), GGATCC produces a greater amount of 1st adduct than GGTACC. The rate of formation of the adducts is also higher for GGATCC than GGTACC as is evident from the greater amount of the 1st adduct in the second melting curve. GTATAC (site 3/HexBrev) is protected with 2 μM ligand while GTTAAC (site 4/HexAfor) is not fully protected, though the first half of this site (GTT) is part of a footprint with 10 μM ligand. In the first melting curves the amount of the 2nd adduct is greater for GTATAC than for GTTAAC. However, in the second melting curves only the first adduct is formed with both these duplexes, and GTTAAC produced a greater amount of the 1st adduct than GTATAC. These percentages suggest that the ligand has greater affinity for GTATAC after long incubation times, though the rate of reaction may be faster with GTTAAC. RMH41 protects GCTAGC (site 2/HexAfor) at a concentration of 10 μM, while GCTTGC (site 2/HexBrev) is protected with 5 μM and GAATGC (site 1/MS1) is protected with 10 μM. These preferences are also evident in the melting studies for which GCTAGC and GCTTGC shows a higher percentages of the 2nd adduct in the first melting curve, as well as the formation of some 2nd adduct in the second melting curve, whereas GAATGC does not form any 2nd adduct in the second melting curve.

RMH43 (PBD-Thz-Py-Py) shows equal affinity for all its sites in the footprinting studies, which are all protected with 5 μM. In the melting studies at 10 μM, GGATCC shows a

greater amount of adduct formation than GGTACC in the first melting curve, but the rate of adduct formation seems faster for GGTACC as observed in the second melting curve. GTATAC is a better binding site than GTTAAC, in terms of both adduct formation in the first melting curve, and rate of reaction as observed in the second melting curves. Oligonucleotides GCTAGC, GCTTGC and GAATGC form 1st and 2nd adducts of equal concentration dependence in the first melting curves. However in the second melting curves, GCTAGC is seemed to be preferred over the other two.

In the footprinting experiments RMH47 (PBD-Py-Thz-Py) protects both GGATCC (site 1) and GGTACC (site 3) in HexAfor at a concentration of 10 μM. In the first melting curve, GGTACC shows a greater percentage of the 2nd adduct than GGATCC. In the second melting curve GGATCC shows only 1st adduct, while a small amount of an adduct with higher T_m is also observed with GGTACC. GGTACC converts all the unmodified duplex into the 1st adduct at a concentration of 5 μM, while for GGATCC formation of the 1st adduct is completed with 1 μM ligand. GTTAAC shows a greater amount of the 2nd adduct in both melting curves, indicating a faster rate of reaction. Formation of the 1st and 2nd adduct is completed with 2 μM ligand with GTTAAC, while with GTATAC the maximum percentages is only achieved with 10 μM. In the footprinting experiment GCTAGC (site 2/HexAfor) and GCTTGC (site 1/HexBrev) are both protected with 5 μM ligand while GAATGC (site1/MS1) is only protected at 10 μM. In the melting studies, GCTTGC shows a greater percentage of the 2nd adduct in the first melting curve, while GCTAGC and GAATGC show equal amounts of 1st and 2nd adduct. In the second melting curves, both GCTTGC and GCTAGC show 1st and 2nd adducts while GAATGC does not show any 2nd adduct. Thus it appears that GCTTGC is the most preferred of these sites.

In the footprinting experiments RMH53 (PBD-Py-Im-Im) protects GGTACC (site 2/HexAfor) at a concentration of 10 μM, while GGATCC is not protected that an adjacent footprint covers only GGA (site1/MS2) with 5 μM. In the melting studies, GGATCC shows a slightly greater amount of the 2nd adduct than GGTACC suggesting a greater rate of reaction to convert the 1st adduct to the 2nd adduct. In the second melting curves, GGATCC also shows greater formation of the 1st adduct than GGTACC, which again suggests faster formation of the 1st adduct with GGATCC than GGTACC. Comparing the results with GTATAC and GTTAAC suggests that GTATAC is better a binding site as it forms a greater percentage of the 2nd adduct, as well as having greater rate of reaction as observed in the second melting curve. Comparing GCTAGC, GCTTGC and GAATGC, in the first melting curves GCTTGC shows the highest percentage of the 2nd adduct, and GAATGC shows the least. In terms of rate of

formation of the adducts, GCTAGC shows fastest rate as it forms both 1st and 2nd adducts in the second melting curve while GCTTGC and GAATGC can produce only 1st adduct.

3.6 Conclusion

From the footprinting studies, it is evident that these PBD conjugates bind at AT-rich sequences. Presence of two consecutive pyrroles in RMH41 (PBD-Im-Py-Py) and RMH43 (PBD-Thz-Py-Py) appear to provide stronger binding; presence of thiazole between two pyrroles in RMH47 (PBD-Py-Thz-Py) slightly reduces the affinity. Compound with two imidazoles (RMH53) shows least affinity among these conjugates; moreover, the compound acquires some affinity for GC-rich sequence as well. Binding of these conjugates is dominated by PBD unit that forms covalent linkage with guanine. Footprinting with sequences containing 6-bp sequences reveals that the conjugates prefers to bind in 1:1 mode. All the compounds produce large increases in melting temperatures indicating strong binding with six base-pair sites.

4 PBD conjugates with benzofused rings

4.1 Introduction

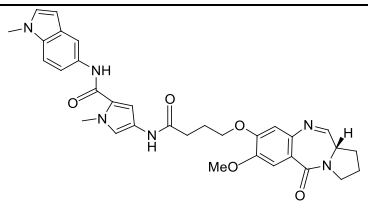
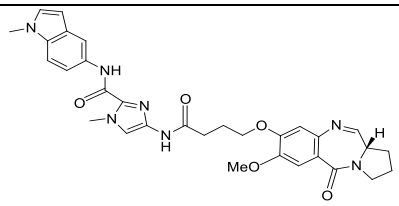
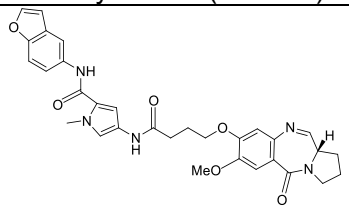
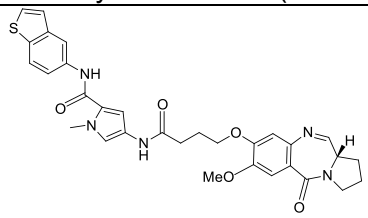
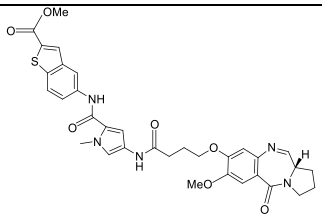
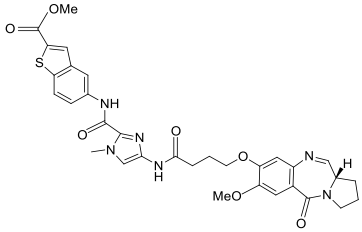
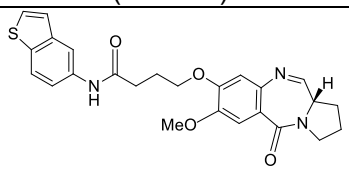
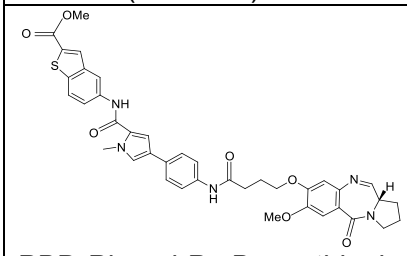
Pyrrolobenzodiazepines (PBD) are sequence selective minor groove binders with chemotherapeutic potential. These compounds bind covalently with the 2-amino group of guanine, as described in the Introduction (section 1.2.4). Conjugation of pyrrolobenzodiazepine with pyrrole and imidazole groups can potentially enhance their binding specificity over longer sequences. Pyrrole and imidazole polyamides have been well studied and ligands containing these subunits are known to bind at specific A.T and G.C base pairs respectively (151,153). The specificity of benzofused ring systems is yet to be established, though compounds containing indole, benzofuran and benzothiophene have been reported to show anticancer activity. Several synthetic and natural indole-containing compounds have been found to bind to DNA with sequence selectivity. Among them, ellipticine (274,275) inhibits DNA topoisomerase-II; cryptolepine (276) has antimalarial and cytotoxic properties by intercalating between cytosine-cytosine sites of DNA duplex and Hoechst-33258 (277,278) is a well-known DNA minor groove binder. From their common ability to bind to DNA, it is concluded that the indole ring in these molecules has an important role in their sequence selective binding. Benzofurans are ubiquitous in nature and some benzofuran-containing compounds such as cyclopenta[*b*]benzofurans are pharmacologically effective as antitumor and selective cytotoxic agents against tumorigenic cell lines (279,280). Taking benzofuran as lead compound, aroylbenzofuran-3-ols and their pyrazolyl derivatives have been evaluated against carcinoma cell lines; all the compounds have shown moderate to weak antiproliferative activities (281). Amongst benzothiophene-containing compounds thienothiophene carboxanilides and quinolones have been reported to exhibit strong inhibitory activities in several human tumour cell lines (282). The compounds displayed diverse but significant antiproliferative activity. Among them, compounds having either mono-imidazolinyl or bis-imidazolinyl substitutions at position-2 were found to increase DNA melting temperature (T_m) by up to 23.9 °C. The sequence selectivity of these mono- and bis-imidazolinyl compounds was determined by DNase I footprinting where bis-substituted carboxanilides revealed sequence selectivity for AT-rich sites (282).

4.2 Experimental design

The footprinting and melting experiments of PBD-conjugates with benzofused rings were performed as described in Chapter 3.2.

4.3 Chemical structures of PBD conjugates with benzofused rings

The compounds were synthesized in Professor David Thurston's laboratory in King's College, UK.

<table>
<tr><td>PBD-Indole derivative</td><td> 
PBD-Py-Indole (KMR30)</td><td> 
PBD-Im-Indole (KMR171)</td></tr>
<tr><td>PBD-Benzofuran derivatives</td><td colspan="2"> 
PBD-Py-Benzofuran (KMR31)</td></tr>
<tr><td rowspan="3">PBD-Benzothiophene derivatives</td><td> 
PBD-Py-Benzothiophene (KMR32)</td><td> 
PBD-Py-Benzothiophene-Acetate (KMR33)</td></tr>
<tr><td> 
PBD-Im-Benzothiophene-Acetate (KMR173)</td><td> 
PBD-Benzothiophene (KMR175)</td></tr>
<tr><td colspan="2"> 
PBD-Phenyl-Py-Benzothiophene-Acetate (KMR164)</td></tr>
</table>

4.4 Result

4.4.1 Footprinting studies with concentration gradient

Figure 4.1 shows DNase I footprints for the interaction of PBD conjugates with benzofused rings with the DNA fragments HexAfor, HexArev and MS1. The compounds were first tested using a high concentration (10 μM) in order to exclude non-binding derivatives from further detailed studies. From these gels it is evident that compounds KMR32, KMR33, KMR173 and KMR31 produce large footprints. Compound KMR175 shows a few footprints whereas KMR30, KMR164 and KMR171 do not show any footprints. The latter three compounds were not studied further.

With DNA fragment HexAfor, compounds KMR31, KMR32, KMR33, KMR173 and KMR175 show footprints at the sequence 5'-CCATGGAT (site 1), located near the bottom of the gel. In the middle of the gel KMR31, KMR32 and KMR33 produce large footprints at the sequence 5'-ATGCTAGCGCTTAAGTACTAGTG (site 2). KMR173 protects the same region but this is split into two separate footprints at ATGCTAGCG and TAAGTACTAGTG, whereas KMR175 binds only at the beginning of this sequence, producing a footprint at ATGCTAGC. With HexArev, which contains the same sequence in the opposite orientation, KMR31, KMR32, KMR33 and KMR173 produce similar footprints at the bottom of the gel in the sequence 5'-CATATATC (site 1); with KMR175 an extra band (C) appears in the first base of this sequence. Another large footprint is produced by KMR31 in the sequence 5'-TTCTAGATCTATAGCT (site 2); KMR32, KMR33 and KMR173 show two separate footprints at this site, corresponding to TTCTAG and CTATAGCT. In addition KMR33 produces a footprint at 5'-TTAACGTTTAAAC (site 3) while KMR31, KMR32 and KMR173 show footprints at this site, which do not include the 5'-TT. Compounds KMR31, KMR32, KMR33 and KMR173 also produce footprints at CTAGCATGCGCAG (site 4) and CAGGCCTAGGTAC (site 5).

With DNA fragment MS1, KMR31, KMR32, KMR33 and KMR173 show footprints near the bottom of the gel at 5'-GAATGCGGA (site 1). Just above this site, KMR33 and KMR173 produce footprints covering a long sequence of 5'-GTTTTGTCATCTCAG (site 2); within this sequence KMR175 binds at GTTTTG whereas KMR31 and KMR32 protect GTTTTGTCAT. All these compounds bind at 5'-GTTATG (site 3). KMR33 and KMR173 also produce a large footprint higher up the gel at the sequence 5'-ATCCTGGTATAGCAATTAGGG (site 4). KMR31 and KMR32 also bind at this site but an extra band of in the middle splits this into two separate footprints at ATCCTG and TATAGCAATTAGGG, whereas KMR175 binds only at GTATAGCAATTAG. All the

footprints for KMR31, KMR32, KMR33 and KMR173 are at AT-rich sequences that contain one or more guanines within them, while KMR175 shows shorter protections.

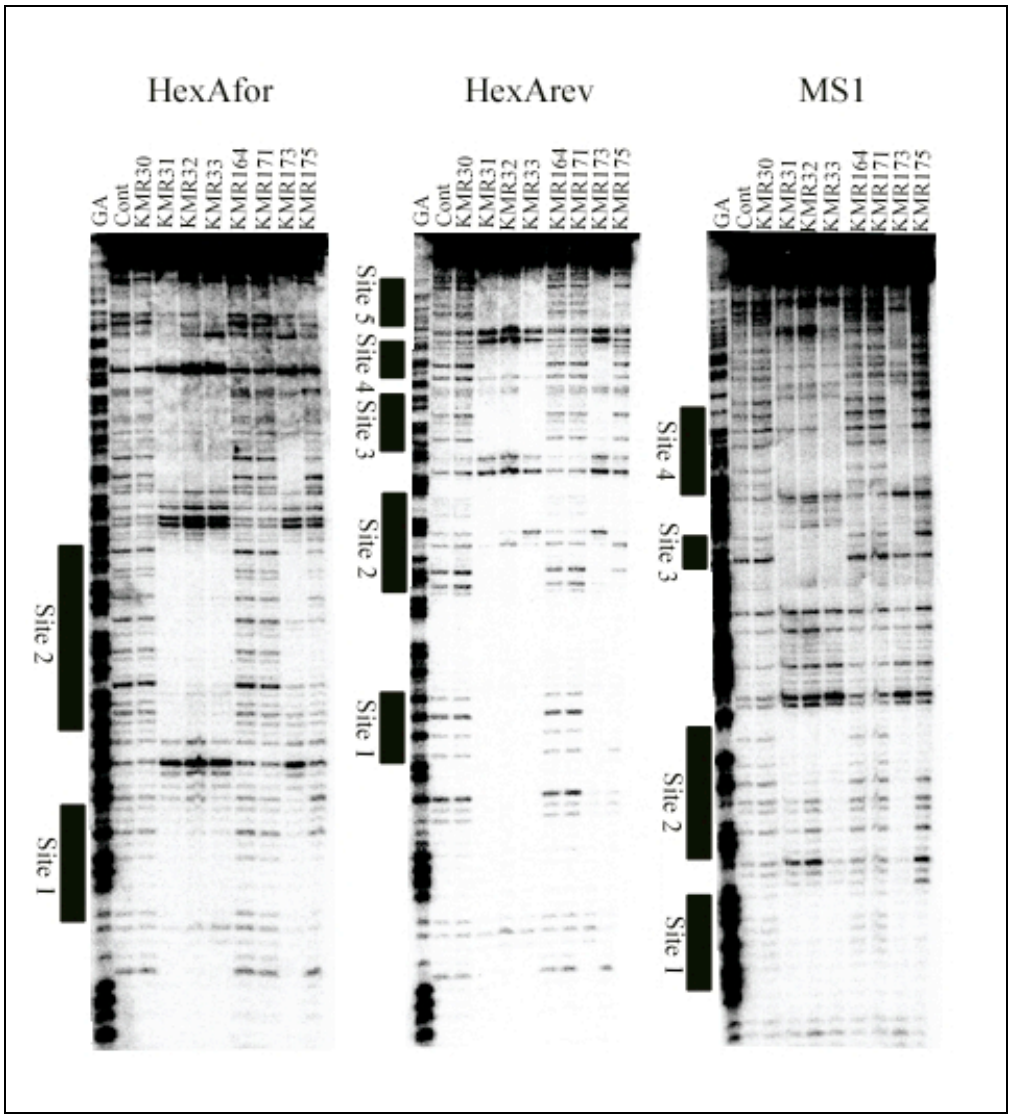


Figure 4.1: DNase I footprinting gels showing the interaction of PBD-conjugates with benzofused rings with HexAfor, HexArev and MS1. The samples were incubated with the ligands (10 μM) at room temperature for 24 hours before digesting with DNase I. Tracks labelled "GA" correspond to markers specific for purines, while "Cont" indicates DNase I cleavage in the absence of added ligands. The numbers (*e.g*. KMR30) indicated the compounds shown in section 4.3.

4.4.1.1 *PBD-Py-Benzofuran (KMR31)*

Figure 4.2 shows DNase I footprints for the interaction of compound PBD-Py-benzofuran (KMR31) with HexA, HexB and MS1 and the sequences protected from cleavage in these three DNA fragments are summarised in Figure 4.3. For HexAfor, footprints are observed even at the lowest concentration of 0.5 μM. In HexAfor, a large footprint covering about 24 bp (site 1) is observed at 10 μM and 5 μM, which becomes smaller at 2 μM and is further reduced to show a footprint at 5'-TTAAGTA at 1 μM and

0.5 μM. Another large footprint of about 23 bp (site 2) can be seen at 10 μM which reduces to a smaller protected region around 5'-TTTAAACGTT at 5 μM and 2 μM and is further truncated to 5'AACGTT at 1 μM. A site of protection of 5'-ATAGATCT (site 3) is observed at 10 μM and 5 μM. A large footprint around the sequence 5'-GGGATATCGATA (site 4) is observed at 10 μM, which is shortened to 5'-TCGATA at concentrations from 5 μM to 0.5 μM. For the sequence HexArev (the same as HexAfor, but in the opposite orientation), footprints are observed at concentrations of 2 μM and above; no footprint is found at 1 μM and 0.5 μM. At 2 μM, footprints are found at 5'-ATATATCGA (site 1) and 5'-TTCT (site 3). Other protected sites are observed at 5'-AACGTTTAAACC (site 4), 5'-ATAGCT (site 2), 5'-GGCCTAGG (site 5) and 5'-ACTTAAGCG (site 6) at 10 μM.

Several footprints are also observed with HexBfor and two sites are still protected at the lowest concentration of 0.5 μM. A large footprint covering the sequence 5'-TAGCCGGCAATTGCAA (site 5) is observed from 10 μM to 0.5 μM, another protection is found at 5'-CTTATA (site 4); these two sites are separated by one guanine. The same site is protected in HexBrev, showing one large footprint (site 1). Footprints at the same location are again found at lower concentrations in the sequences 5'-CATATGTA (site 3) and 5'-TACATATGTAT (site 2) in HexBfor and HexBrev respectively.

The interaction of this compound with MS1/MS2 shows a few large footprints. One footprint in MS2 in the sequence 5'-AATTGCTATACCAGGATA (site 4) is also evident in MS1 in which it appears as two footprints at 5'-ATCCTGG (site 5) and 5'-ATAGCAATTAGG (site 4) that are separated by one thymine. Smaller footprints such as 5'-GCATTCG (site 7), 5'-GACAAAACA (site 6), 5'-CTTTACATAAC (site 5) and 5'-CTTAACCTTGATC (site 3) are observed in MS2, and are also evident in MS1.

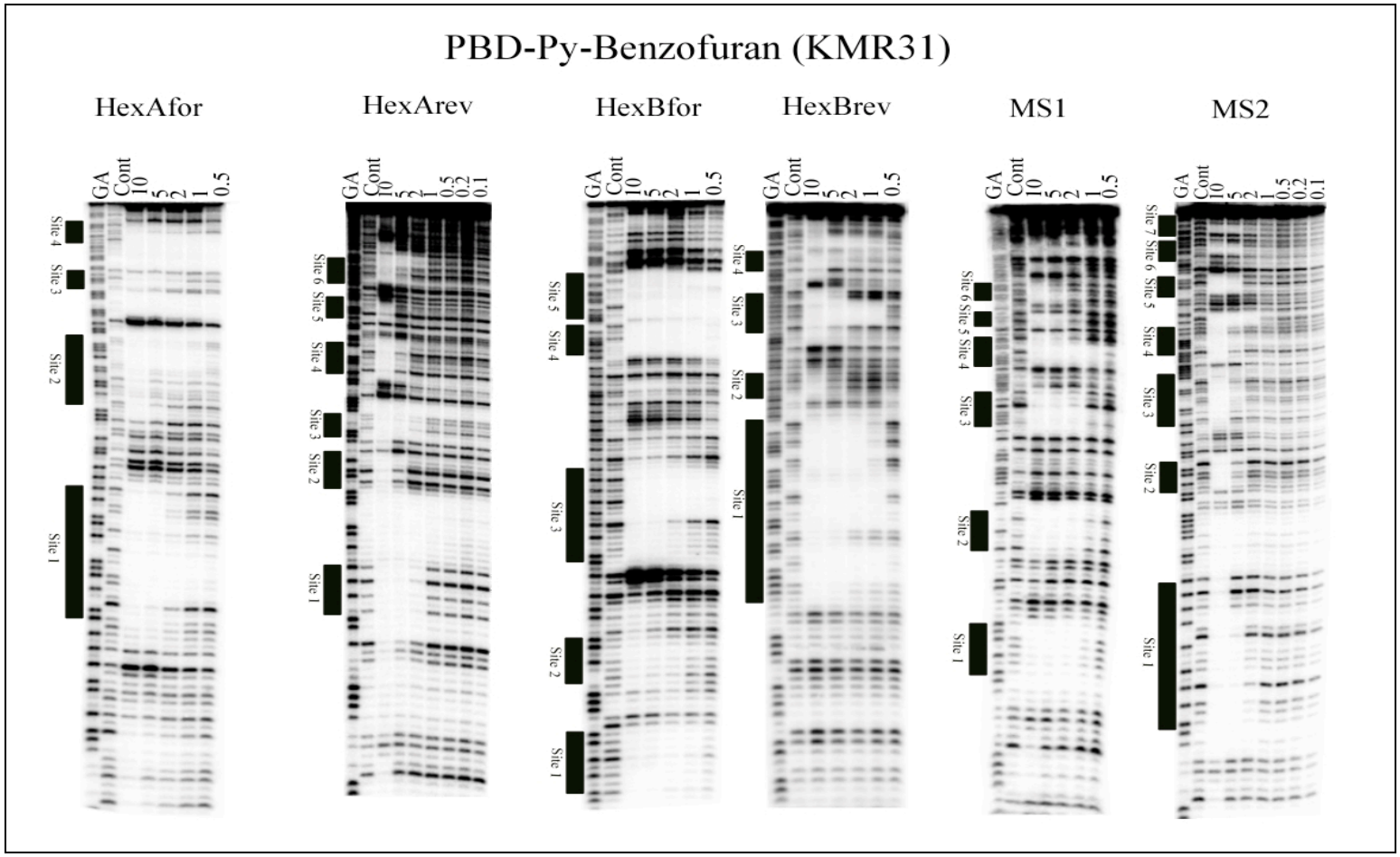


Figure 4.2: DNase I footprinting gels showing the interaction of various concentrations of PBD-Py-Benzofuran (KMR31) with HexAfor, HexArev, HexBfor, HexBrev, MS1 and MS2. The samples were incubated with the ligands at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands.

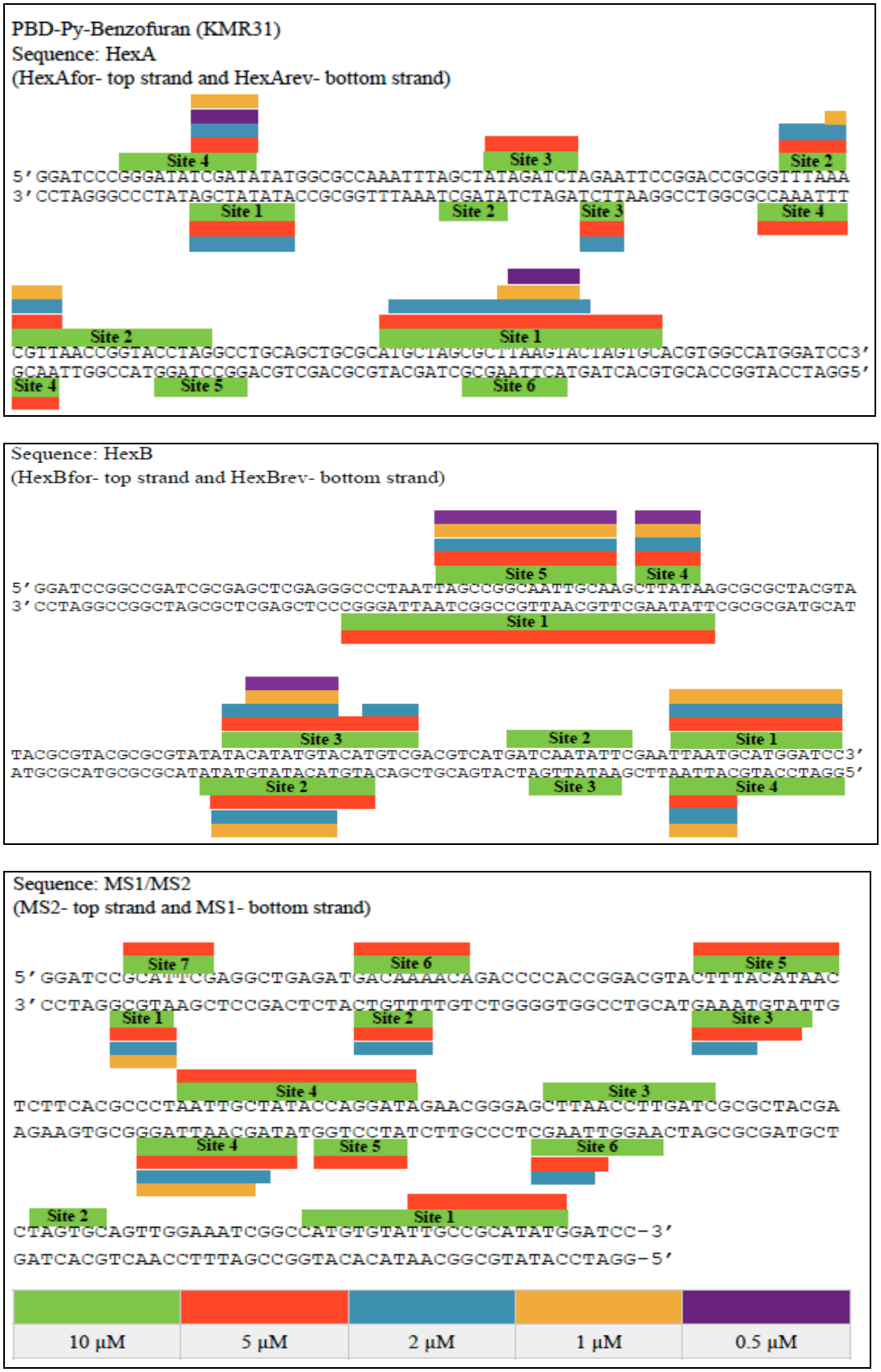


Figure 4.3: Sequences of the footprinting substrates indicating the regions protected by PBD-Py-Benzofuran (KMR31) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand.

4.4.1.2 *PBD-Py-Benzothiophene (KMR32) and PBD-Py-Benzothiophene-Acetate (KMR33)*

Figures 4.4 and 4.6 show the interaction of PBD-Py-Benzothiophene (KMR32) and PBD-Py-Benzothiophene-Acetate (KMR33) with HexA, HexB and MS1 respectively and the sequences protected from cleavage in these three DNA fragments are summarised in Figures 4.5 and 4.7 respectively. KMR32 protects several sites in HexA and HexB at the lowest concentration (0.5 μM); while for MS1 and MS2, it only produces footprints at higher concentrations with no protection at 0.5 μM. KMR33 has a similar chemical composition to KMR32, except that it possesses an additional acetate moiety attached to the benzothiophene. KMR33 produces footprints at similar locations to KMR32, with only a few variations.

With DNA fragment HexA, KMR32 and KMR33 both protect at 5'-GGGATATCGATA (site 4/KMR32 and site 5/KMR33) as well as producing two other large footprints (site 1 and site 2 for KMR32; site 2 and site 3 for KMR33), each of about 24 bp, that become shorter at lower concentrations (1 μM and 0.5 μM). For KMR32, protections are observed at 5'-TCGATA (site 4), 5'-TTTAAACGTT (site 2) and 5'-CTTAAGTA (site 1) at 1 μM along with another footprint of 5'-ATAGATCT (site 3) at 2 μM. Similar protections are observed for KMR33 with an additional footprint at the bottom of the gel at 5'CATGGATCC (site 1). With HexArev, footprints are evident for both KMR32 and KMR33 at concentrations 10 μM, 5 μM and 2 μM. KMR32 shows footprints at 5'-ATATATCGA (site 2), 5'-TAACGTTTAAACC (site 5), 5'-GGCCTAGG (site 6) and 5'-ACTTAAGCG (site 7), which correspond to the sites that were protected in the forward sequence. KMR33 also binds to these sequences, except 5'-GGCCTAGG.

Interaction of these two compounds with HexB depicts a similar binding profile. KMR32 produces a large footprint at 5'-CTAATTAGCCGGCAATT (site 5) that is protected from 10 μM to 0.5 μM whereas the same site is split into two sites (5'-AATTAGGGC and 5'-TTGCCG) in HexBrev (site 1) when the concentration is lowered. For KMR33, footprints are observed at the same sites. Other sites protected by both KMR32 and KMR33 include 5'-CATATGTA (site 3), 5'-TGTCG (site 3), 5'ATATTC (site 2) and 5'AATGCAT (site 1) in HexBfor. The same sites are also protected in HexBrev.

Comparing the binding of KMR32 and KMR33 with MS1/MS2, it is evident that these compounds protect the same sites; but they differ in terms of concentrations or length of binding sites. 5'-TTTACATAA of MS2 (site 6) is protected by both compounds but this is only evident with KMR32 at the highest concentration (10 μM) whereas KMR33 still binds at 5 μM. 10 μM KMR32 produces a large footprint at 5'-

TTGCTATACCAGGATAG (site 5), which is truncated to 5'-TTGCT at 5 μM and 2 μM; in contrast with KMR33, the binding sites are the same at 10 μM and 5 μM, but there is no footprint at 2 μM. In MS1, KMR32 binds at 5'-AATTAGG (site 4) at 1 μM while KMR33 protects a shorter region (site 4) (5'-TTAGG) at 2 μM. KMR32 and KMR33 both protect 5'-GGAAATCGGC at 10 μM and 5 μM respectively.

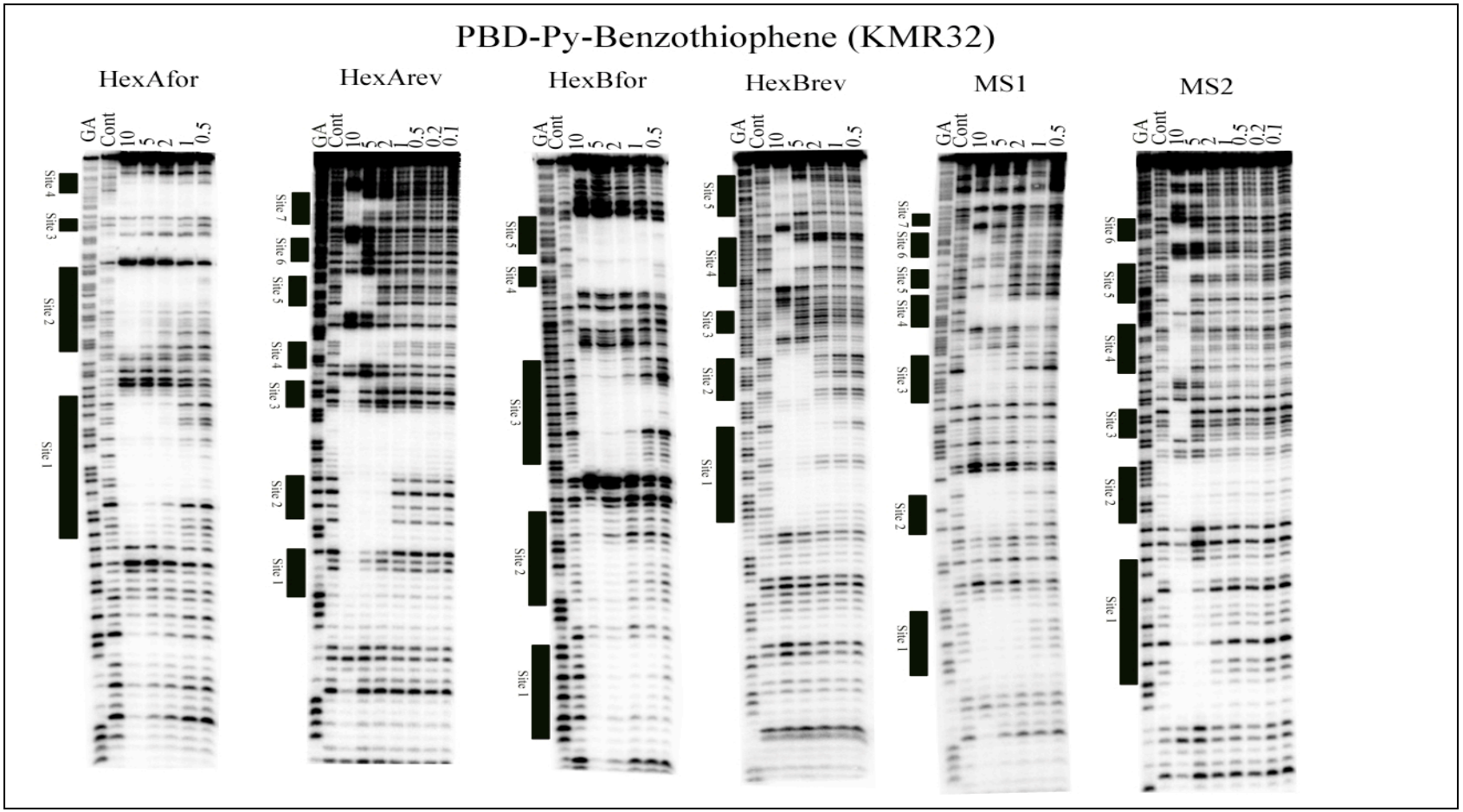


Figure 4.4: DNase I footprinting gels showing the interaction of various concentrations of PBD-Py-Benzothiophene (KMR32) with HexAfor, HexArev, HexBfor, HexBrev, MS1 and MS2. The samples were incubated with the ligands at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands.

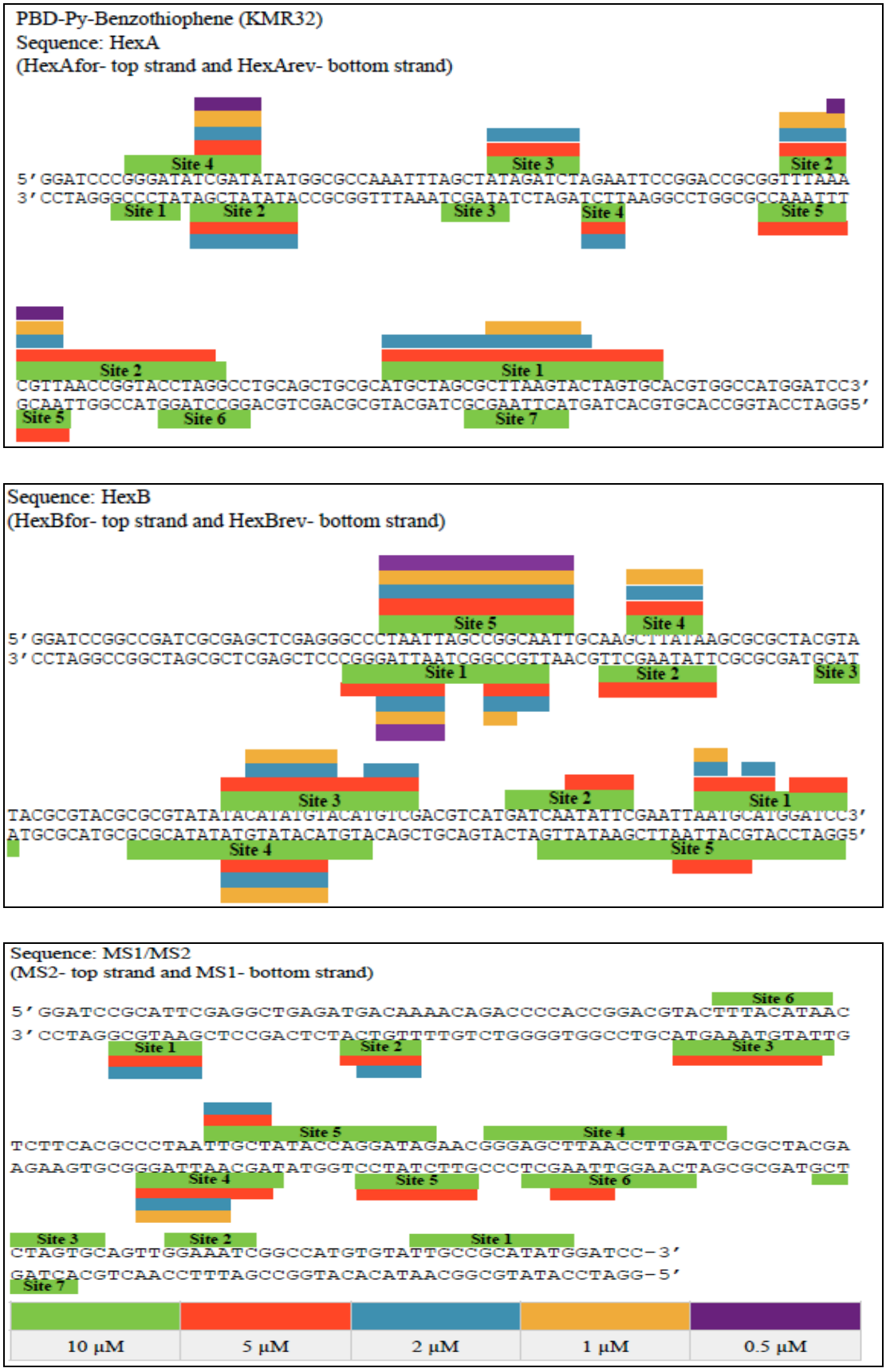


Figure 4.5: Sequences of the footprinting substrates indicating the regions protected by PBD-Py-Benzothiophene (KMR32) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand.

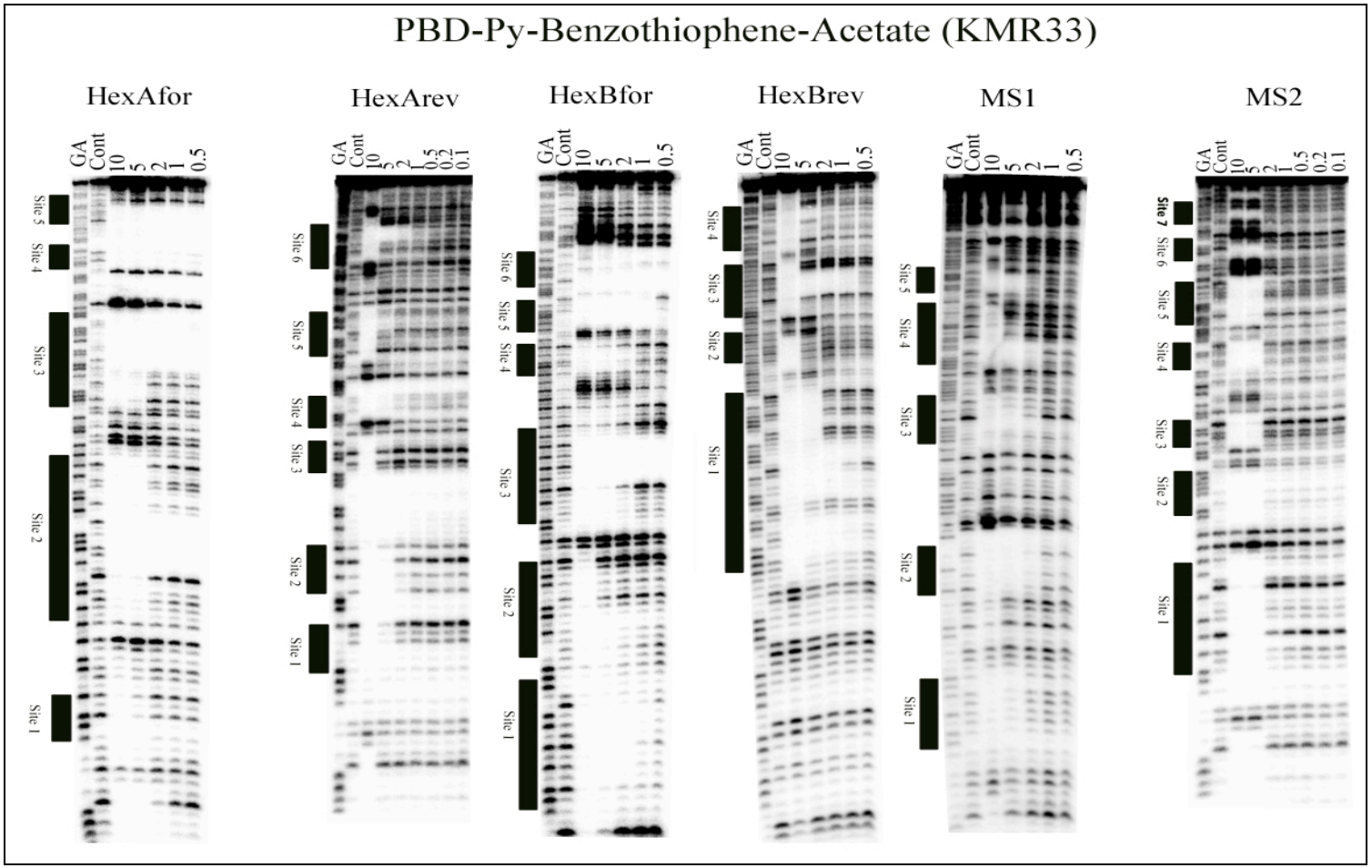


Figure 4.6: DNase I footprinting gels showing the interaction of various concentrations of PBD-Py-Benzothiophene-Acetate (KMR33) with HexAfor, HexArev, HexBfor, HexBrev, MS1 and MS2. The samples were incubated with the ligands at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands.

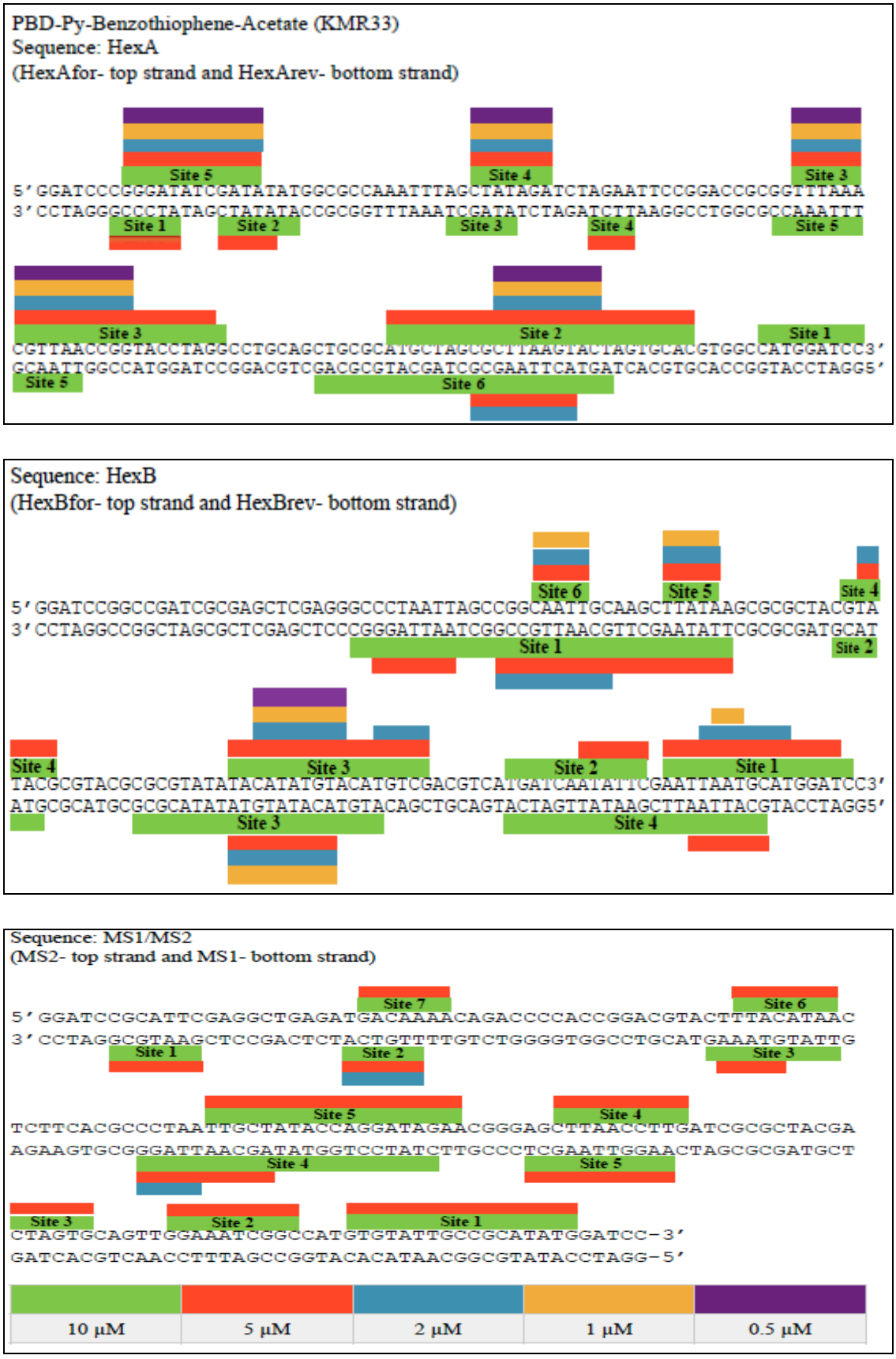


Figure 4.7: Sequences of the footprinting substrates indicating the regions protected by PBD-Py-Benzothiophene-Acetate (KMR33) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand.

4.4.1.3 *PBD-Im-Benzothiophene-Acetate (KMR173)*

The interaction of PBD-Im-Benzothiophene-Acetate (KMR173) with HexA, HexB and MS1 is shown in Figure 4.8 and the sequences protected from cleavage in these DNA fragments are summarised in Figure 4.9. This compound is similar to KMR33, except for replacement of the pyrrole of KMR33 with an imidazole ring. The compound shows footprints with HexA only at the highest concentrations (10 μM and 5 μM). Two large footprints (~20 bp) (site 1 and site 2) are observed in HexAfor at 10 μM and one of these footprints (site 2) is truncated to show binding at 5'-AAGTAC at 5 μM. Protection at the same site is found in HexArev. A few shorter binding sites such as 5'-ATAGATCT (site 5), 5'-AACGTTAA (site 4) and 5'GTACCTAG (site 3) are observed at 10 μM with HexAfor.

The binding of KMR173 is more pronounced with HexB and a few sites are protected at 1 μM. Two large footprints, 5'-GGCAATTGCAA (site 7) and 5'-CTACGTATAC (site 6) are observed at 2 μM and 5 μM respectively in HexBfor. The same sites are protected at 10 μM in HexBrev; with 5 μM of this ligand the sequence 5'-TTGCAATTGCCG (site 2) is split into two sites of 5'-TTGCA and 5'-TGCCG leaving cleavage at the middle AT. When the concentration is further lowered, protection is restricted to 5'-GCCG. A few short sites such as 5'-GCTCGA (site 9), 5'GCCCTAA (site 8), 5'-CGCGTA (site 5), 5'-CATAT (site 4), 5'-TACAT (site 3) and 5'-GATCAA (site 2) are also protected in HexBfor at various concentrations; 5'-GCCCTAA and 5'-CGCGTA of HexBfor are also protected in HexBrev shown as site 1 and site 4 respectively.

For MS1/MS2 KMR173 footprints are mainly observed at the highest concentration (10 μM) with only two footprints at 5 μM. No protections are found at 2 μM or less. For MS1, shorter binding sites are observed compared with its reverse sequence, MS2. In MS2, sites rich in As and Ts having one or more guanine in middle, such 5'-CTTTAC (site 6), 5'-TTGCTA (site 5), 5'-GATAGAA (site 4), 5'-CAGTTGGAAA (site 3) and 5'-ATGGA (site 1) are protected. In MS1, four large footprints (~20 bp) (site 1, 2, 3 and 6) are found at 10 μM; two footprints rich in As and Ts such as 5'-AAAGTAC (site 2) and 5'-TTAGG (site 3) are observed at 5 μM.

4.4.1.4 *PBD-Benzothiophene (KMR175)*

Figure 4.10 shows DNase I footprints for the interaction of PBD-Benzothiophene (KMR175) with HexA, HexB and MS1 and the protection sites are showed in Figure 4.11. With HexAfor the compound protects 5'-TCCGG (site 2) at concentrations down to 0.5 μM. Other sites such as 5'-TAAG (site 1 of HexAfor) and 5'-ATAT (site 1), 5'-TTC (site 2) and 5'-TTAAG (site 3) of HexArev are protected only at highest

concentration. KMR175 shows better protection with HexB, binding to more sequences than HexA. In this DNA fragment, most of the binding sites are rich in As and Ts, showing protection at 5'-AATT (site 6), 5'-TTG (site 5), 5'-AAGCG (site 4), 5'-TATGT (site 3), and 5'-ATGCAT (site 1) in HexBfor. These sites are also protected in HexBrev. Binding at short sequences like 5'-TTG, 5'-ATG, 5'-ATTA and 5'-TATAA can be seen at lower concentration up to 1 μM. In MS1/MS2, few footprints are found with 10 μM and 5 μM of this ligand and among these, one site is protected at 2 μM. Footprints are observed at 5'-CAATT (site 4), 5'-AAAGTA (site 3), 5'-TTGTC (site 3) and 5'ATGCG (site 1) in MS1. With MS2, only two footprints are found at 5'-TTG and 5'-ATG.

Comparing the footprints of all these PBD-conjugates, it is evident that they share some common binding sites, but differ in the concentrations dependence and the length of footprints. In HexA, KMR31, KMR32, KMR33 and KMR173 all bind at 5'-TCGATA, 5'-ATAGATCT, 5'-AACGTT, 5'-CTTAAGTA and 5'-AGGCCTAGGTA. Protection is also found at a comparatively small site of 5'-TCTT. In HexB, these four compounds bind at 5'-TAGCCGGCAATTGCAA, 5'-CTTATA, 5'-CATATGTA and 5'-TTAATGCATGGATCC. With MS1, footprints are observed at 5'-CTTTACATAACTC, 5'-AATTGCTATACCAGGATAGAA and 5'-CTTAACCTTGATC.

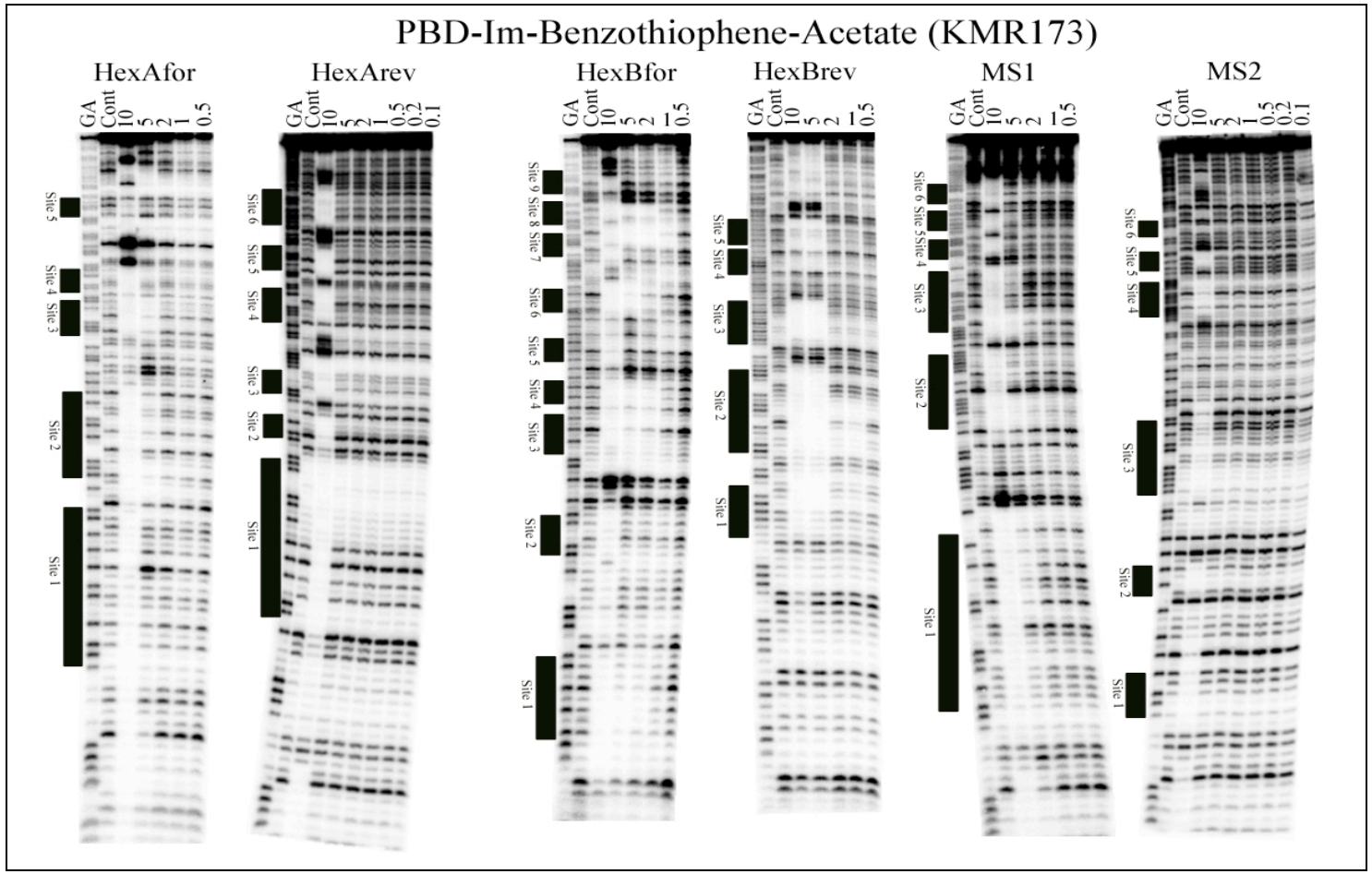


Figure 4.8: DNase I footprinting gels showing the interaction of various concentrations of PBD-Im-Benzothiophene-Acetate (KMR173) with HexAfor, HexArev, HexBfor, HexBrev, MS1 and MS2. The samples were incubated with the ligands at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands.

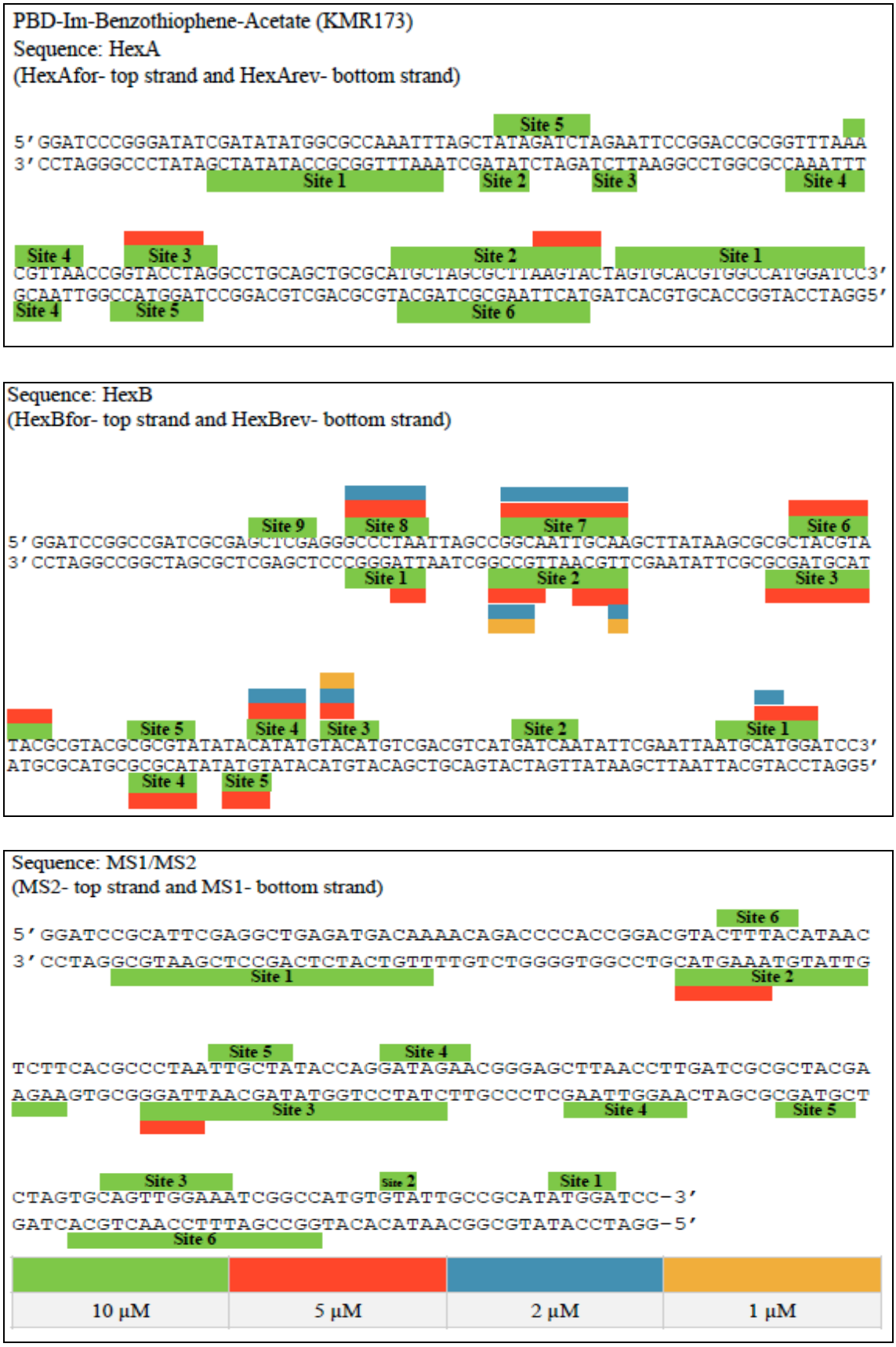


Figure 4.9: Sequences of the footprinting substrates indicating the regions protected by PBD-Im-Benzothiophene-Acetate (KMR173) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand.

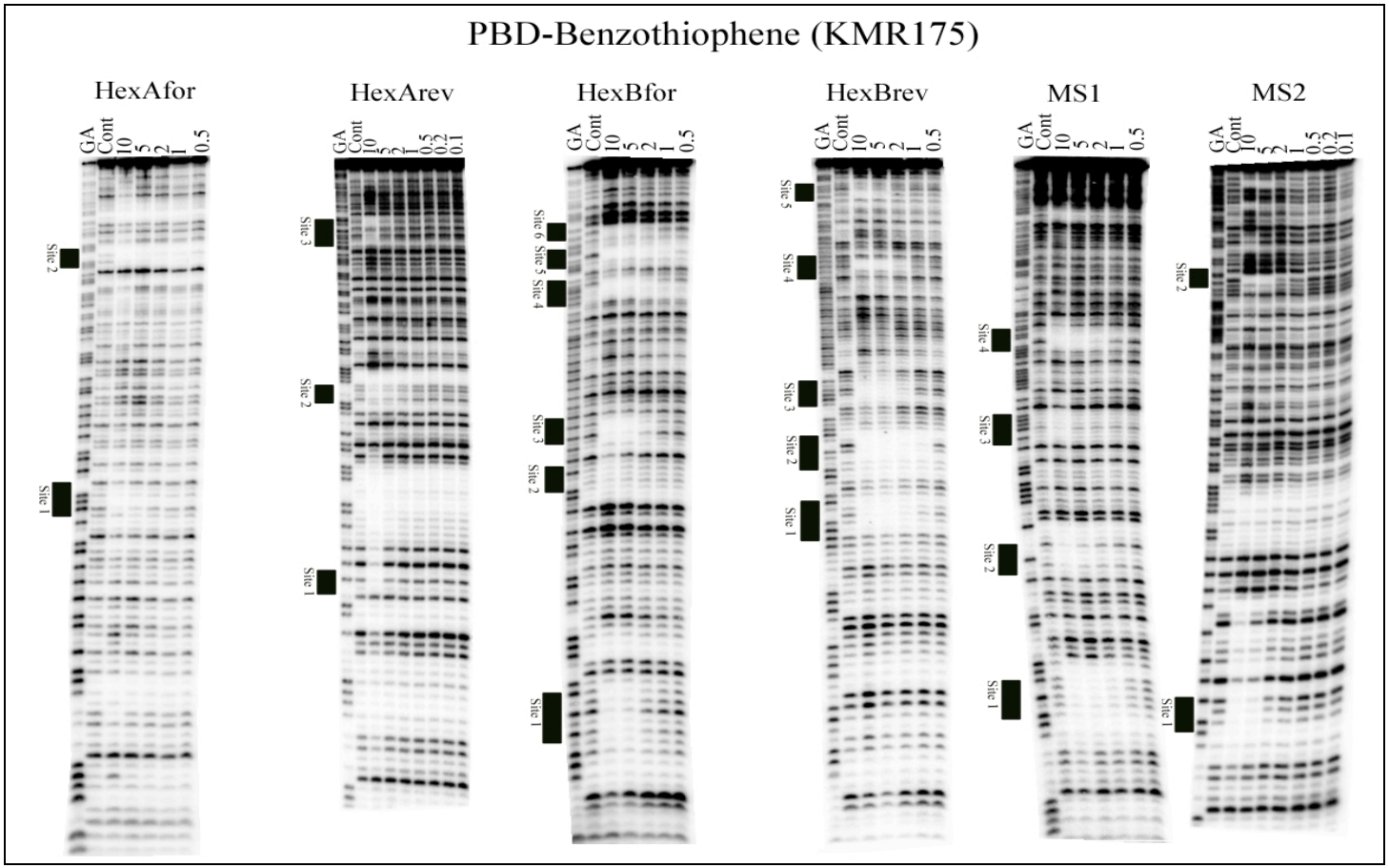


Figure 4.10: DNase I footprinting gels showing the interaction of various concentrations of PBD-Benzothiophene (KMR175) with HexAfor, HexArev, HexBfor, HexBrev, MS1 and MS2. The samples were incubated with the ligands at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands.

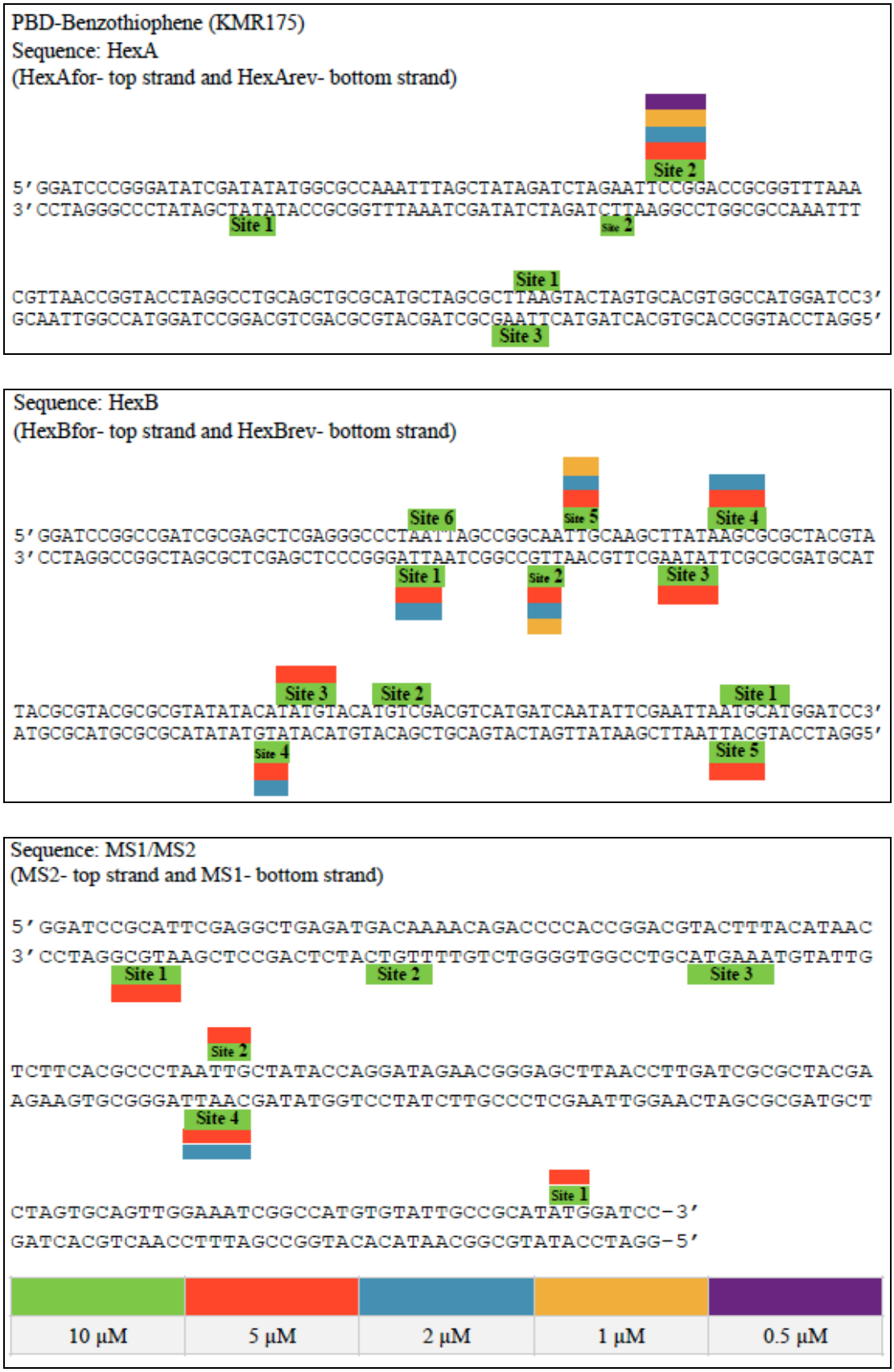


Figure 4.11: Sequences of the footprinting substrates indicating the regions protected by PBD-Benzothiophene (KMR175) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand.

4.4.2 Effect of incubation period

Since these compounds form covalent adducts with DNA it is possible that they could become 'trapped' on less preferred sites, with which they interact more quickly. The observed binding sites might therefore represent sequences with which the ligands interact more quickly, rather than those with which they bind more tightly. The best binding sites might therefore only become apparent after long incubation times. All the experiments in the preceding section were performed after incubating the ligand and DNA for about 24 hours. We therefore performed some experiments examining whether the footprinting patterns change with different incubation times. Figure 4.12 shows the interaction of KMR31, KMR32 and KMR33 with HexB and KMR173 with HexA at different time intervals and the protected sequences are shown in Figure 4.13.

4.4.2.1 *PBD-Py-Benzofuran (KMR31)*

With 5 μM KMR31 the footprint at 5'-ATATTCG (site 2) appears within 1 minute of mixing the ligand and DNA. With 2 μM ligand this footprint is only observed after 24 hours, whereas no protection is observed with 1 μM. The footprint at 5'-TACATATGTAC (site 3) is evident after 1 minute with both 5 μM and 2 μM ligand; the protection is extended towards 3'-end and shows binding with ATGTC. With 1 μM ligand, the footprint is observed after 1 hour with no changes at longer incubation times. Another common site of protection is 5'-TTAGCCGGCAATTGCAAG (site 6), which is observed after 1 minute for 5 μM and 2 μM but only after 5 minutes for 1 μM; the binding is then extended over time towards 3'-end and shows binding with CTTATA. Protection at 5'-GTAT (site 4) is observed after 5 minutes with 5 μM and 2 μM ligand whereas this footprint is only evident after 3 hours with 1 μM ligand. The footprint at the sequence 5'-TATA (site 5) is evident after 1 hour with 5 μM but after 3 hours with 2 μM and 1 μM. The binding at 5'-ATGCA (site 1) is seen after 3 hours with 5 μM KMR31 while this site is only evident after 24 hours with 2 μM.

4.4.2.2 *PBD-Py-Benzothiophene (KMR32) and PBD-Py-Benzothiophene-acetate (KMR33)*

Compounds KMR32 and KMR33 show similar binding time-dependent profiles with HexB. KMR32 shows protection at 5'ATGATGGA (site 1) after 24 hours with 5 μM while with 2 μM and 1 μM, the ligand protects only first five bases (underlined) after 24 hours. For KMR33, this site 5'ATGCA (site 1) portion is protected with 5 μM after 1 hour and extends to cover TGGA after 24 hours; with 2 μM and 1 μM, protection of 5'ATGCA is only seen after 24 hours. KMR32 produces footprint at 5'TGATCAATATTC (site 2) after 24 hours with 5 μM; no footprint is found with 2 μM and 1 μM. For KMR33, the ligand binds at 5'ATATTC (site 2) (underlined) after 3 hours and 24 hours with 5 μM

and 2 μM respectively; the ligand binds with the remaining TGATCA after 24 hours only with 5 μM. Both compounds show protection at 5'CATATGTA (site 3). Protection of this sequence is observed after 5 minutes with 5 μM of KMR32; with 2 μM and 1 μM, the protection is observed after 1 hour and 3 hours respectively. The footprint is extended for all three concentrations; but extension is seen after 3 hours with 5 μM while with 2 μM and 1 μM, it is observed after 24 hours. In case of KMR33, this site 3 is seen after 1 minute of incubation with all concentrations. Extension of binding is observed after 3 hours and 24 hours with 5 μM and 2 μM respectively while no extension is found with 1 μM. KMR32 shows binding at 5'GTAT (site 4) with 5 μM and 2 μM after 1 hour and 3 hours respectively; with 1 μM, the footprint is seen after 24 hours. KMR33 does not show any binding with this site. KMR32 binds with 5'TATA (site 5) after 1 hour with 5 μM while with lower concentrations, the protection is observed after 24 hours. KMR33 also binds at the same site but protects a slight longer sequence of 5'TATACGC (site 4) after 5 minutes and 3 hours with 5 μM and 2 μM respectively; no binding is observed with 1 μM. KMR32 shows a large footprint at 5'TTAGCCGGCAATTGCAAG (site 6) after 1 minute of incubation with 5 μM; with 2 μM and 1 μM, the footprint is observed after 5 minutes. The protection is then extended towards 3'-end covering CTTATAAGCG. Compound KMR33 protects same location (site 5) but produces two footprints at 5'CCGGCAATTGCAAG and 5'TATAAG after 1 minutes and 5 minutes with 5 μM and 2 μM respectively; with 1 μM KMR33, protection is only seen at the former one. These two sequences become coalesced with 5 μM and 2 μM after 5 minutes and 1 hour respectively while with 1 μM, the footprint is extended to cover entire sequence after 3 hours.

4.4.2.3 *PBD-Im-Benzothiophene-Acetate (KMR173)*

The interaction of KMR173 with HexA also shows different binding sites protected at various time intervals. Protections at 5'-GCCATGGATCC (site 1) and 5'-GTGCA (site 2) are observed only after 24 hours with 5 μM ligand while no binding is observed with 2 μM and 1 μM. The compound binds to 5'-AGTAC after 1 hour and 3 hours with 5 μM and 2 μM respectively; with 5 μM another footprint adjacent to this site is observed after 3 hours, these two sites then merge after 24 hours giving a large footprint at 5'-ATGCTAGCGCTTAAGTAC (site 3). However, no time-dependent extension at over time is observed at this site with 2 μM. The compound protects 5'-GGTACCTAGGCCT (site 4) after 3 hours with 5 μM whereas the same footprint is only observed after 24 hours with 2 μM; with 5 μM this binding site extends at the 5'-end to GTTAACC after 24 hours. A small footprint at 5'TCCGGA (site 5) is observed after 1 minute with 5 μM while with 2 μM, the site is evident after 1 hour.

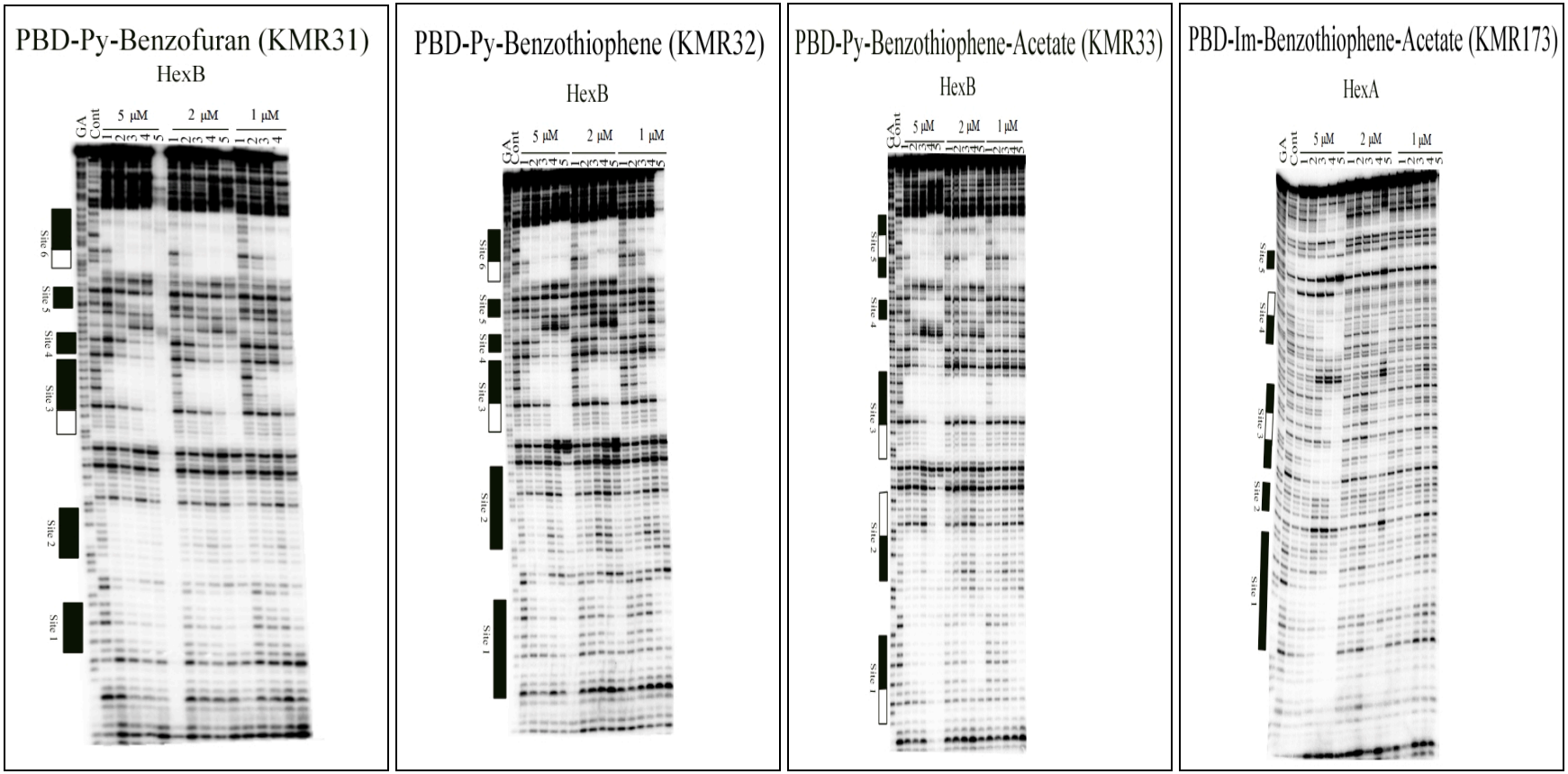


Figure 4.12: DNase I footprinting gels showing the time course of interaction of various concentrations of PBD-conjugates with benzofused rings with HexB (KMR31, KMR32 and KMR33) and HexA (KMR173). The samples were incubated with the ligands (5 μM, 2 μM and 1 μM) at room temperatures before digesting with DNase I. Incubation periods are shown at the top of each gel lane; 1 = 1 minute, 2 = 5 minutes, 3 = 1 hour, 4 = 3 hours and 5 = 24 hours. Tracks labelled "GA" correspond to markers specific for purines, while "Cont" indicates DNase I cleavage in the absence of added ligand. The boxes indicate to the position of footprints and correspond to those identified in Figure 4.13; the open boxes show regions where the size of the footprint increases with time.

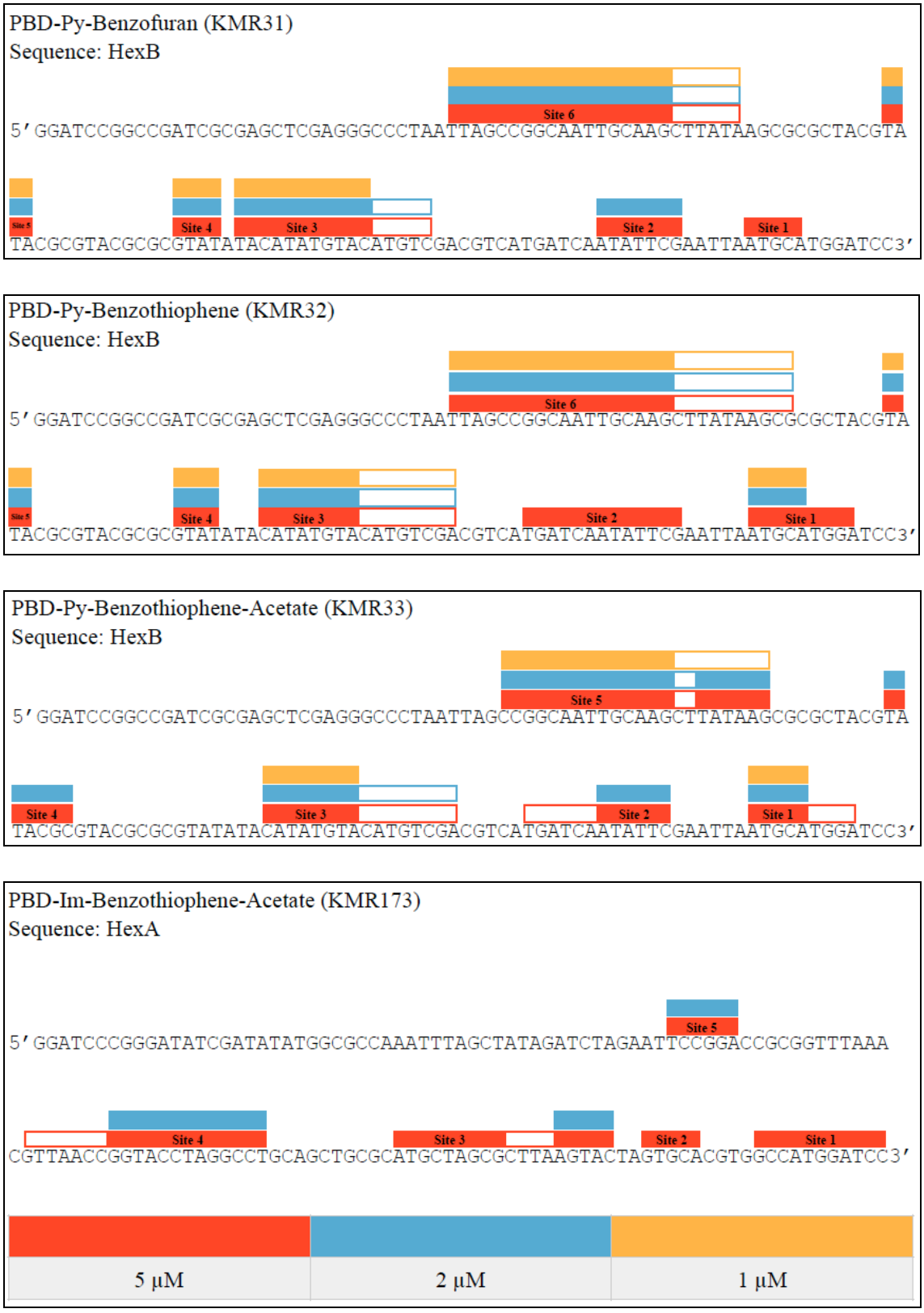


Figure 4.13: Sequences of the footprinting substrates indicating the regions protected by PBD-conjugates from DNase I cleavage. Filled boxes (coloured) correspond to initial footprints at different locations; open boxes correspond to any extension/annexation of footprints over time. Only the sequence of radiolabelled strand is shown.

4.4.3 Melting studies

Melting studies were performed with seven different fluorescently labelled oligonucleotides with ligand concentrations ranging from 10 μM to 0.1 μM. Since the concentration of the duplex is only 0.25 μM there is an excess DNA over ligand at the lowest concentration. In these experiments the complexes, which had been equilibrated 12 hours were first heated from 30 °C to 95 °C, while recording the fluorescence. Although the PBDs form covalent links to their target guanines, these bonds are broken at high temperatures and the ligand completely dissociates from the DNA. These melting curves were followed by cooling the samples to 30 °C. The ligands only slowly reassociate with the DNA during this annealing reaction and in most cases these annealing profiles were similar to those of the free DNA. This was then followed by a second heating step recording the new melting transitions. The ligands had reattached to the DNA during the annealing phase and these melting curves typically showed ligand-induced stabilisation relative to the drug-free controls, though this was lower than in the first melting transition. In the following description the melting curves are named as first melting curve and second melting curve. In each melting curve, up to three transitions are observed: the first transition corresponds to that of the unmodified duplex having no ligand attached; this is followed by up to two further transitions, which correspond to covalent attachment of one or two drug molecules, producing two further T_m values, which are denoted as 1st adduct and 2nd adduct. The proportions of unmodified duplex, 1st adduct and 2nd adduct were calculated as ratios of ordinates of each corresponding transitions as shown in Figure 4.14).

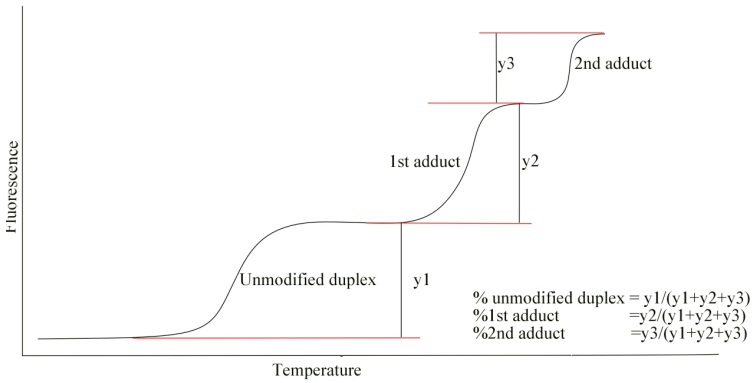


Figure 4.14: Schematic of a melting curve displaying three transitions

4.4.3.1 *PBD-Py-Benzofuran (KMR31)*

GGTACC and GGATCC

Figures 4.15 and 4.16 show the first and second melting curves of PBD-Py-Benzofuran (KMR31) with seven different oligonucleotide duplexes. The compound produces large changes in the melting profiles, showing multiphase transitions with all the oligonucleotides. For these, and all the melting profiles described below, the positions (T_m) of the different melting transitions do not change with ligand concentration (in contrast to studies with reversible ligand-DNA interactions), but their relative proportions are altered, so that a greater fraction of the higher T_m is evident with higher ligand concentrations. The percentages of unmodified duplex, 1st adduct and 2nd adduct are given in Tables 4.1 to 4.4, while the T_m values of unmodified duplex, 1st adduct and 2nd adduct are tabulated and presented in the Appendix. In the presence of KMR31 the oligonucleotide containing GGTACC shows three T_ms of 41.4 °C, 70.2 °C and 88.8 °C in the first melting curve corresponding unmodified duplex, 1st adduct and 2nd adduct respectively and two T_m of 41.7 °C and 70.1 °C in the second melting curve, for which there is no third transition. Oligonucleotide GGATCC shows similar changes in melting profiles with T_ms of 42.3 °C, 72.5 °C and 90.6 °C for first melting curve and 42.3 °C and 71.8 °C for second melting curve. With 2 μM ligand almost the entire transition of GGTACC corresponds to that of the 1st adduct, while for GGATCC 62% corresponds the 1st adduct with 38% represented by the 2nd adduct. Similarly, at the highest concentration, GGTACC shows melting transitions for the 1st and 2nd adducts of equal amplitude, while with GGATCC 39% corresponds to the 1st adduct and 61% to the 2nd adduct. For the second melting curve, both oligonucleotides do not show any transition for the 2nd adduct. For GGTACC, ligand concentrations up to 1 μM produce melting profiles that correspond to the unmodified duplex, while with GGATCC 17% of the transition corresponds to that of the 1st adduct. These results suggest that KMR31 binds better to GGATCC than GGTACC.

GTATAC and GTTAC

Oligonucleotides containing only one GC base pair at each end of the duplex, such as GTATAC and GTTAAC also show multiple transitions in the presence of the ligand. GTATAC shows three transitions with 0.5 μM and 1 μM ligand with T_ms of 36.5 °C, 65.2 °C and 86.1 °C. At higher concentrations, only the two transitions of the modified duplexes are observed. At 2 μM, about 32% of the transition corresponds to the 1st adduct and 68% for the 2nd adduct; this does not change much at higher concentrations, though the melting profiles show a slight shift to the more stable

complexes. For the second melting curve, the profile for GTATAC correspond to the unmodified duplex at ligand concentrations up to 1 μM; the other transitions reappear at higher ligand concentrations so that with 10 μM three transitions are observed with relative amplitudes of 55%, 23% and 22%. Note that this contrasts with the results for GGTACC and GGATCC for which no transition for the 2nd adduct was observed in the second melting curves. GTTAAC also shows three melting transitions at ligand concentrations of 0.5 μM, 1 μM and 2 μM with T_m values of 37.6 °C, 64 °C and 86.5 °C. No transition for the unmodified duplex is seen with 5 μM and 10 μM for which 39% corresponds to the 1st adduct and 61% to the 2nd adduct with 10 μM. In the second melting curve, most of the transition corresponds to the unmodified duplex with only 11% of the 1st adduct observed at 5 μM which then increases to 21% at 10 μM with no evidence for the second adduct. These results suggest that GTATAC is a better binding site than GTTAAC.

GCTTGC, GCTAGC and GGATGC

Oligonucleotide GCTTGC shows three transitions at different ligand concentrations. With 0.1 μM ligand about 20% of the duplex is modified into the 1st adduct and this rises to about 80% at 0.5 μM. At higher concentrations no unmodified duplex is observed. With 1 μM, 78% corresponds to the 1st adduct with a T_m 73.5 °C and 22% as the 2nd adduct with a T_m of 91.8 °C. The percentage of 1st adduct then gradually decreases, while that of 2nd adduct increases showing 40% and 60% of these at 10 μM. In the second melting curve, the unmodified duplex predominates up to 2 μM where only 28% corresponds to the 1st adduct. With 5 μM and 10 μM, three transitions are observed; the percentage of the unmodified duplex remains same for both concentrations, while proportion of the 1st adduct changes from 34% to 64% and the 2nd adduct from 41% to 12%. Oligonucleotide GCTAGC shows three transitions at different ligand concentrations. With 0.1 μM ligand about 13% of the duplex is modified into the 1st adduct, and by 0.5 μM three transitions are evident with T_m of 43.9 °C, 73.3 °C and 91.7 °C. At 1 μM, the transition corresponds to two modified duplexes with 74% as the 1st adduct and 26% in the 2nd adduct. The percentage of the 1st adduct then decreases while the 2nd adduct increases to 66% with 10 μM. In the second melting curve, the transitions of the unmodified duplex and 1st adduct are observed with 0.5 μM, 1 μM and 2 μM, while three transitions are seen with 5 μM with relative amplitudes of 39%, 54% and 7%; the percentage of the 2nd adduct then increases to 17% at 10 μM. For oligonucleotide GAATGC about 40% of the duplex is modified into the 1st adduct with 0.1 μM ligand and with 0.5 μM the melt appears as a single transition corresponding to the 1st adduct with a T_m of 73.9 °C. A higher melting transition,

corresponding to the 2nd adduct, appears at higher concentrations so that at 10 μM 43% corresponds to the 1st adduct and 57% to the 2nd adduct. In the second melting curve of this oligonucleotide there was no evidence of the 2nd adduct, even at highest ligand concentration. 20% of the transition corresponds to the 1st adduct at 0.5 μM and this increased to over 90% with 10 μM.

By comparing these changes in T_m, it is clear that KMR31 binds to all these oligonucleotides. The melting transitions also suggest that each of the oligonucleotides can bind at least 2 drug molecules at higher concentrations. However the proportions of each species show different concentration dependence. The second melts, before which the ligand will have only had a relatively short time to reattach to the duplex, are also very different. GGTACC, GGATCC and GAATGC do not show any evidence of the 2nd adduct in the second melting curves, while this is present with GTATAC, GTTAAC, GCTTGC and GCTAGC, though its proportion is greatest for GCTAGC and GTATAC. These results will be considered further in the Discussion, but it appears that the rank order of these binding sites is GCTAGC=GTATAC > GCTTGC=GTTAAC > GGTACC=GGATCC=GAATGC.

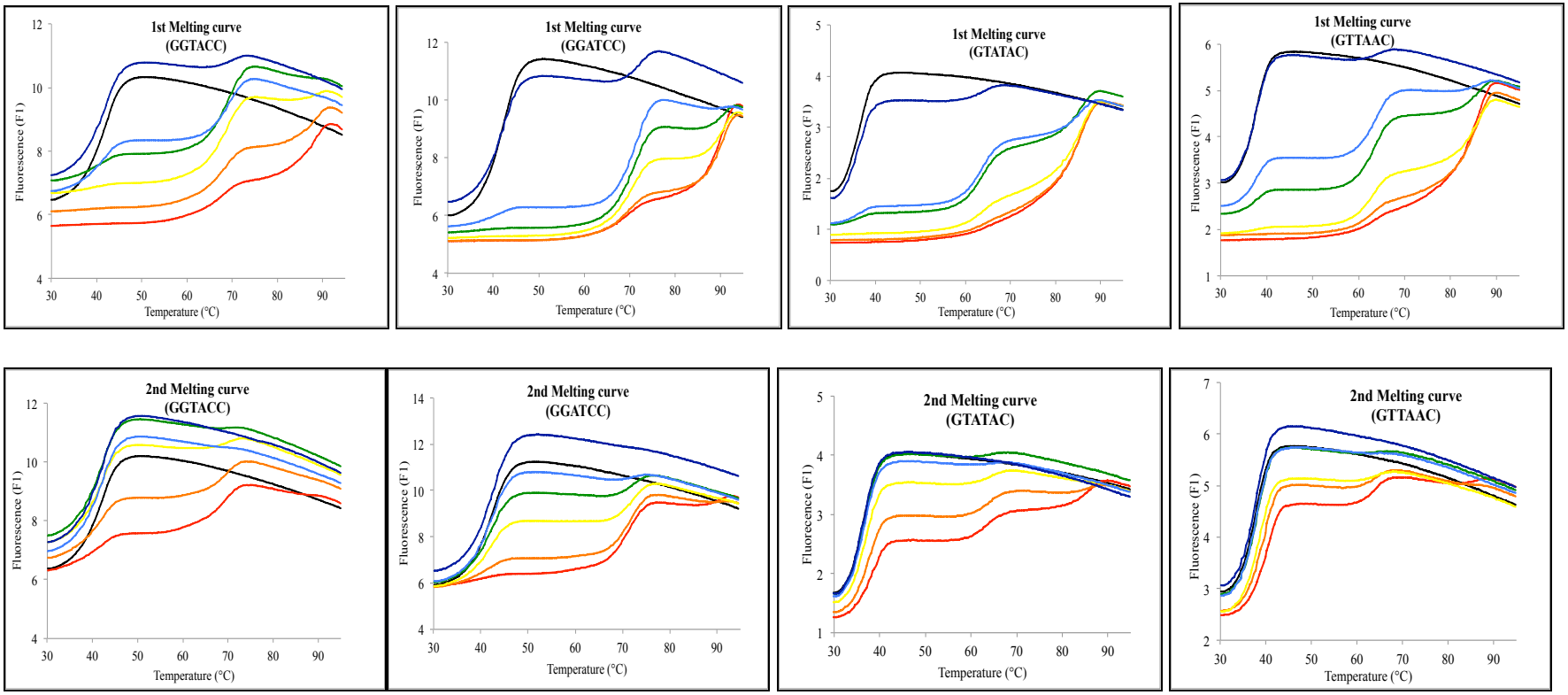


Figure 4.15: First and second melting curves of oligonucleotide duplexes containing the sequences GGTACC, GGATCC, GTATAC and GTTAAC in the presence of different concentrations of PBD-Py-Benzofuran (KMR31)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green), 0.5 μM (light blue) and 0.1 (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting. 'Control' indicates sample without any added ligand.

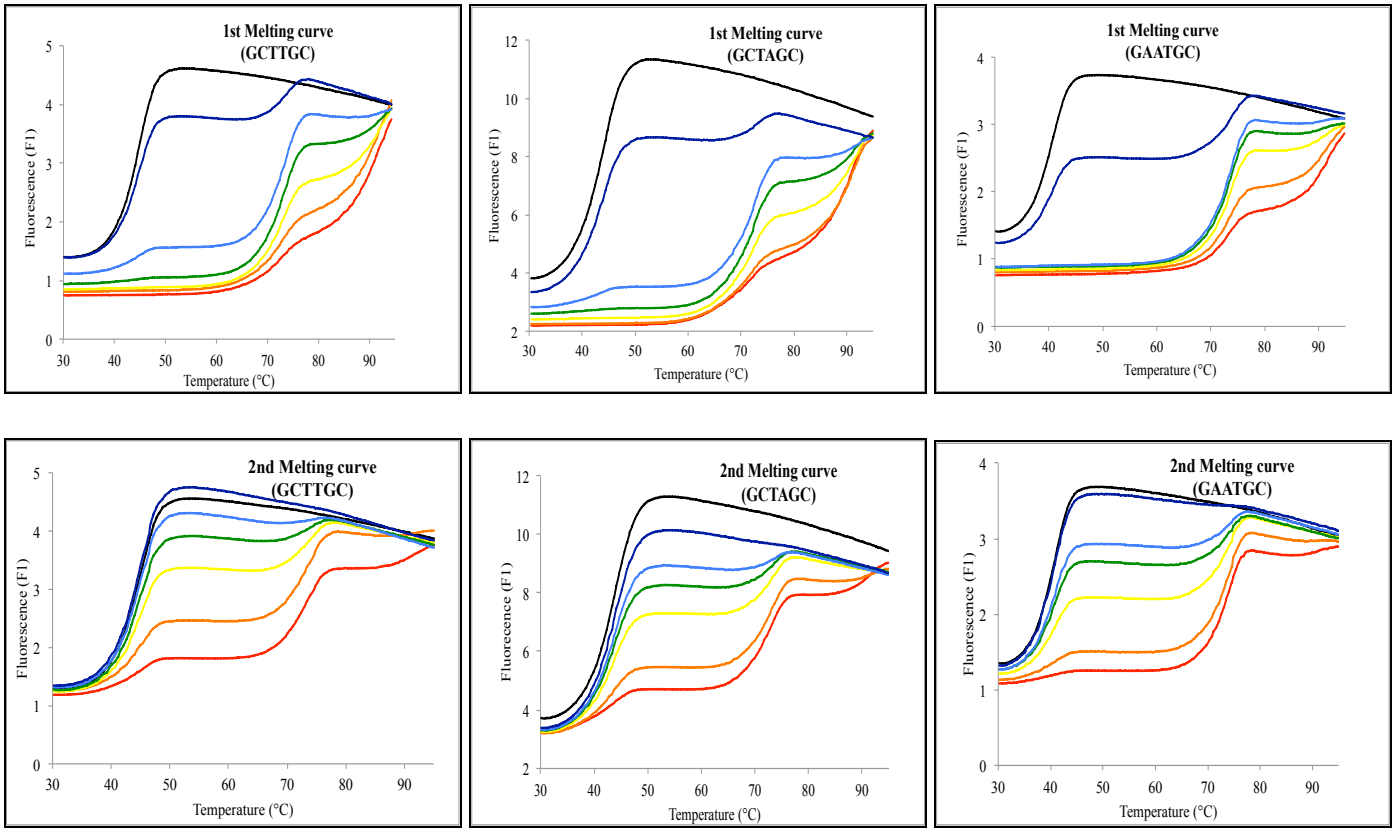


Figure 4.16: First and second melting curves of oligonucleotide duplexes containing the sequences GCTTGC, GCTAGC and GAATGC in the presence of different concentrations of PBD-Py-Benzofuran (KMR31)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green), 0.5 μM (light blue) and 0.1 (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting. 'Control' indicates sample without any added ligand.

Table 4.1: Percentages of the different melting transitions, corresponding to 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzofuran (KMR31) with oligonucleotide duplexes containing the sequences GGTACC and GGATCC

	1st melting curve			2nd melting curve		
5'-GGTACC	Unmodified duplex T_m 41.4 °C	Duplex+Ligand (1st adduct) T_m 67.6 -70.2 °C	Duplex+Ligand (2nd adduct) T_m 88.2 - 88.8 °C	Unmodified duplex T_m 41.7 °C	Duplex+Ligand (1st adduct) T_m 69.3 -70.1 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	90%	10%	0%	100%	0%	0%
0.5 μM	50%	50%	0%	100%	0%	0%
1 μM	20%	80%	0%	100%	0%	0%
2 μM	0%	100%	0%	89%	11%	0%
5 μM	0%	53%	47%	63%	37%	0%
10 μM	0%	50%	50%	43%	57%	0%

	1st melting curve			2nd melting curve		
5'-GGATCC	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 69.5 - 72.5 °C	Duplex+Ligand (2nd adduct) T_m 90.4 - 90.6 °C	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 71.4 °C -71.8 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	81%	19%	0%	100%	0%	0%
0.5 μM	16%	84%	0%	96%	4%	0%
1 μM	0%	85%	15%	83%	17%	0%
2 μM	0%	62%	38%	64%	37%	0%
5 μM	0%	50%	50%	33%	67%	0%
10 μM	0%	39%	61%	25%	75%	0%

Table 4.2: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzofuran (KMR31) with oligonucleotide duplexes containing the sequences GTATAC and GTTAAC

5'-GTATAC	1st melting curve			2nd melting curve		
	Unmodified duplex T_m 36.5 °C	Duplex+Ligand (1st adduct) T_m 63.9 -65.2 °C	Duplex+Ligand (2nd adduct) T_m 85.4 - 86.1	Unmodified duplex T_m 36.6 °C	Duplex+Ligand (1st adduct) T_m 63.6 -64.2°C	Duplex+Ligand (2nd adduct) T_m 85.8
Control	100%	0%	0%	100%	0%	0%
0.1 μM	86%	14%	0%	100%	0%	0%
0.5 μM	17%	54%	29%	100%	0%	0%
1 μM	12%	46%	42%	100%	0%	0%
2 μM	0%	32%	68%	91%	9%	0%
5 μM	0%	37%	63%	73%	18%	9%
10 μM	0%	37%	63%	55%	23%	22%

5'-GTTAAC	1st melting curve			2nd melting curve		
	Unmodified duplex T_m 37.8 °C	Duplex+Ligand (1st adduct) T_m 62.8 - 64.0 °C	Duplex+Ligand (2nd adduct) T_m 85.6 - 86.5 °C	Unmodified duplex T_m 38.1 °C	Duplex+Ligand (1st adduct) T_m 63.9 - 64.2 °C	Duplex+Ligand (2nd adduct) T_m 86.1°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	93%	7%	0%	100%	0%	0%
0.5 μM	41%	52%	7%	100%	0%	0%
1 μM	16%	60%	24%	100%	0%	0%
2 μM	31%	28%	41%	93%	7%	0%
5 μM	0%	44%	56%	89%	11%	0%
10 μM	0%	39%	61%	73%	21%	6%

Table 4.3: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzofuran (KMR31) with oligonucleotide duplexes containing the sequences GCTTGC and GCTAGC

5'-GCTTGC	1st melting curve			2nd melting curve		
	Unmodified duplex T_m 44.6 °C	Duplex+Ligand (1st adduct) T_m 72.3 -73.5 °C	Duplex+Ligand (2nd adduct) T_m 91.4 - 91.8 °C	Unmodified duplex T_m 44.8 °C	Duplex+Ligand (1st adduct) T_m 73.4 -73.9 °C	Duplex+Ligand (2nd adduct) T_m 89.8°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	80%	20%	0%	100%	0%	0%
0.5 μM	21%	79%	0%	97%	3%	0%
1 μM	0%	78%	22%	89%	11%	0%
2 μM	0%	63%	37%	72%	28%	0%
5 μM	0%	50%	50%	25%	34%	41%
10 μM	0%	40%	60%	24%	64%	12%

5'-GCTAGC	1st melting curve			2nd melting curve		
	Unmodified duplex T_m 43.9 °C	Duplex+Ligand (1st adduct) T_m 72.1 - 73.3 °C	Duplex+Ligand (2nd adduct) T_m 90.7 - 91.7 °C	Unmodified duplex T_m 44.3 °C	Duplex+Ligand (1st adduct) T_m 73.1 -73.6°C	Duplex+Ligand (2nd adduct) T_m 90.0°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	87%	13%	0%	100%	0%	0%
0.5 μM	13%	73%	14%	90%	10%	0%
1 μM	0%	74%	26%	80%	20%	0%
2 μM	0%	56%	44%	67%	33%	0%
5 μM	0%	42%	58%	39%	54%	7%
10 μM	0%	34%	66%	28%	55%	17%

Table 4.4: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzofuran (KMR31) with oligonucleotide duplexes containing the sequence GAATGC

	1st melting curve			2nd melting curve		
5'-GAATGC	Unmodified duplex T_m 40.7 °C	Duplex+Ligand (1st adduct) T_m 73.0 -73.9 °C	Duplex+Ligand (2nd adduct) T_m 90.9 - 91.4 °C	Unmodified duplex T_m 40.9 °C	Duplex+Ligand (1st adduct) T_m 74.0 -74.2 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	59%	41%	0%	100%	0%	0%
0.5 μM	0%	100%	0%	80%	20%	0%
1 μM	0%	95%	5%	70%	30%	0%
2 μM	0%	79%	21%	48%	52%	0%
5 μM	0%	55%	45%	20%	80%	0%
10 μM	0%	43%	57%	6%	94%	0%

4.4.3.2 *PBD-Py-Benzothiophene (KMR32)*

Similar melting studies with PBD-Py-Benzothiophene (KMR32) are presented in Figures 4.17 and 4.18 and the percentages of the different species are shown in Tables 4.5 to 4.8.

GGTACC and GGATCC

For GGTACC, the T_m values for the unmodified duplex, 1st adduct and 2nd adduct are 41.4 °C, 70.4 °C and 89.3 °C respectively. More than 80% of the melting transition appears as the 1st adduct at concentrations 0.5 μM and 1 μM with no 2nd adduct. With the addition of 2 μM, no unmodified duplex is seen and 81% appears as the 1st adduct; the proportion the 2nd adduct then rises to 47% with 10 μM. In the second melting curve, no 2nd adduct is observed, even at the highest concentration. The percentage of the 1st adduct rises to 59% with 10 μM ligand. For oligonucleotide GGATCC, three transitions with T_ms of 42.3 °C, 72.6 °C and 90.5 °C are observed. About 80% appears as the 1st adduct at a concentration of 0.5 μM and no 2nd adduct is observed; the 2nd adduct appears with 1 μM and increases to about 80% with 10 μM. Considering the second melting curve, more than 50% of the unmodified duplex is observed at ligand concentrations up to 2 μM; with 5 μM ligand 84% appear as the 1st adduct with only 16% unmodified duplex, while about 10% appears as the 2nd adduct with 10 μM.

GTATAC and GTTAAC

For oligonucleotides GTATAC and GTTAAC, the changes in T_m and the proportions of the different species are very similar. The T_ms for the 1st adduct and 2nd adduct are 65.3 °C and 85.9 °C for GTATAC and 64.1 °C and 85.6 °C for GTTAAC. For both oligonucleotides the 1st and 2nd adduct account for around 50% of the transition with 1 μM. The percentage of the 2nd adduct increases to about 80% with 10 μM. In the second melting curves, the unmodified duplex predominates with little or no transition for the 2nd adduct, even at highest ligand concentration.

GCTTGC, GCTAGC and GAATGC

Oligonucleotide GCTTGC shows T_m values of 44.6 °C, 74.4 °C and 91.9 °C for the unmodified duplex, 1st adduct and 2nd adduct respectively. Oligonucleotide GCTAGC shows melting of the 1st and 2nd adducts at similar temperatures of 73.6 °C and 91.7 °C. Both oligonucleotides produce about 77% of 1st adduct with 1 μM with the remainder as 2nd adduct. The proportion of the 1st adduct then decreases and the 2nd adduct increases to about 65% with 10 μM. In the second melting curve, three transitions are seen with 10 μM with both oligonucleotides and with 5 μM with

GCTAGC. For oligonucleotide GAATGC, almost all the melting transition corresponds to the 1st adduct with 0.5 μM and 1 μM ligand. The 2nd adduct appears at higher concentrations and accounts for 57% with 10 μM. In the second melting curve, no transition is seen for the 2nd adduct, but the duplex is fully converted into the 1st adduct with 10 μM.

Comparing the melting profiles of all oligonucleotides with KMR32, it can be seen that GGTACC and GGATCC show significant differences in adduct formation, even though their sequences are similar. In the first melting curve 10 μM produces 80% in the 2nd adduct with GGATCC but only 47% with GGTACC. In the second melting curve, GGTACC shows no transition for the 2nd adduct with 59% in the 1st adduct, while for GGATCC, around 90% is in 1st adduct form with the remainder in 2nd adduct. Oligonucleotides GTATAC and GTTAAC show almost same transitions in the first melting curves, but in the second melting curves, no 2nd adduct is seen with GTTAAC but this accounts for 17% of the transition with GTATAC. Oligonucleotides GCTTGC and GCTAGC also produce similar melting profiles with 10 μM for both the first and second melting curves. GAATGC produces a similar profile in first melting curve but produces less stabilisation in the second melting curve.

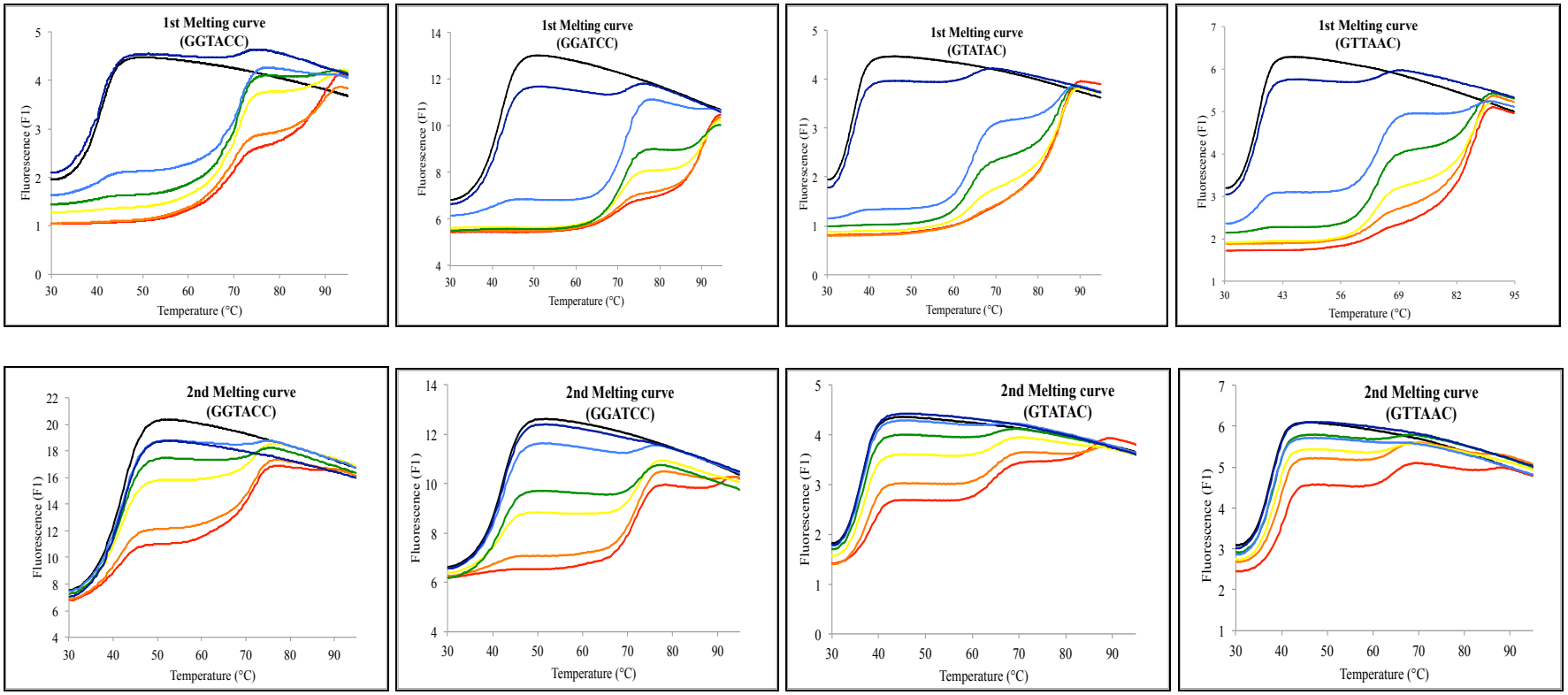


Figure 4.17: First and second melting curves of oligonucleotide duplexes containing the sequences GGTACC, GGATCC, GTATAC and GTTAAC in the presence of different concentrations of PBD-Py-Benzothiophene (KMR32)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green), 0.5 μM (light blue) and 0.1 (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting. 'Control' indicates sample without any added ligand.

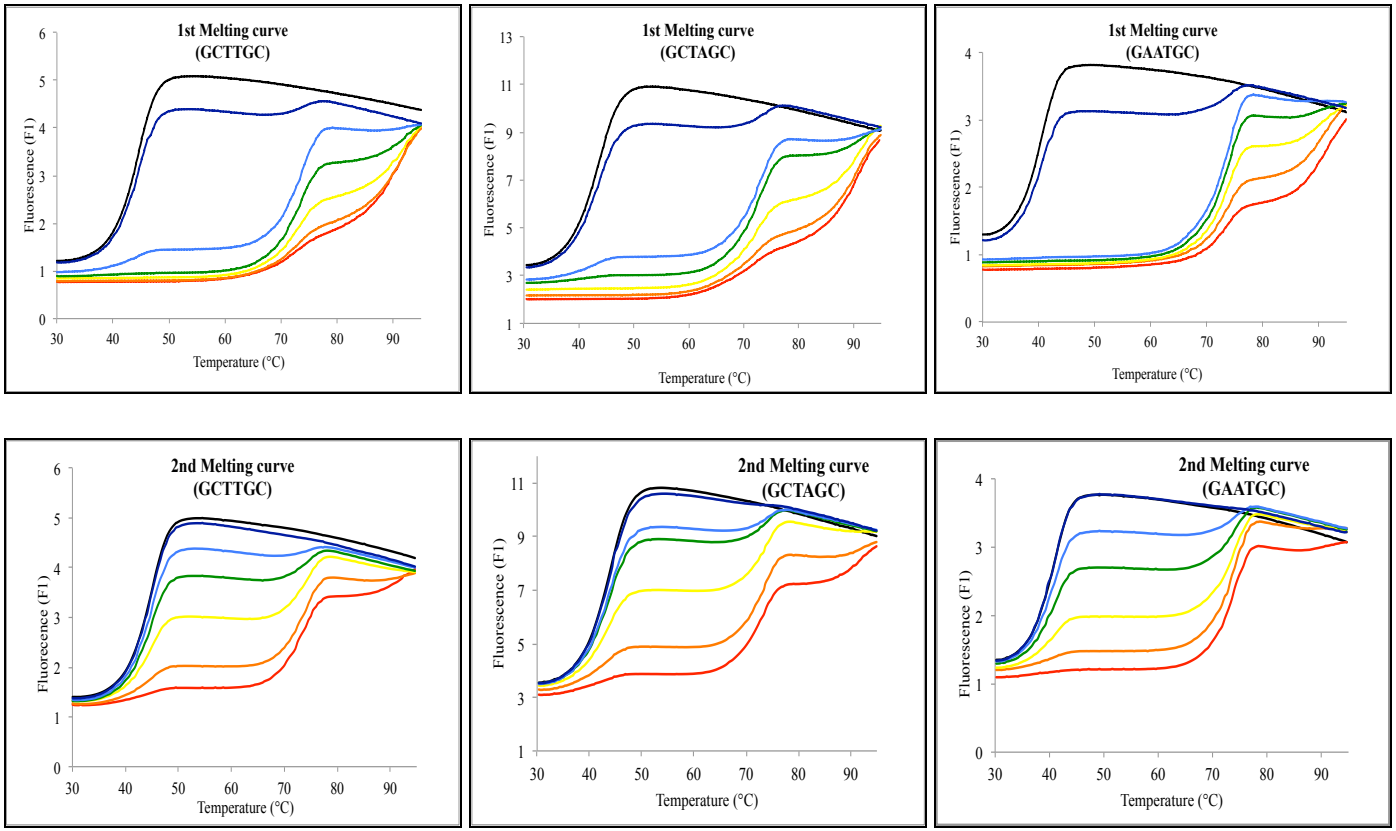


Figure 4.18: First and second melting curves of oligonucleotide duplexes containing the sequences GCTTGC, GCTAGC and GAATGC in the presence of different concentrations of PBD-Py-Benzothiophene (KMR32)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green), 0.5 μM (light blue) and 0.1 (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting. 'Control' indicates sample without any added ligand.

Table 4.5**:** Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzothiophene (KMR32) with oligonucleotide duplexes containing the sequences GGTACC and GGATCC

5'-GGTACC	1st melting curve			2nd melting curve		
	Unmodified duplex T_m 41.4 °C	Duplex+Ligand (1st adduct) T_m 68.2 - 70.4 °C	Duplex+Ligand (2nd adduct) T_m 88.7 - 89.3 °C	Unmodified duplex T_m 41.7 °C	Duplex+Ligand (1st adduct) T_m 70.1 -70.6 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	96%	4%	0%	100%	0%	0%
0.5 μM	19%	81%	0%	97%	3%	0%
1 μM	8%	92%	0%	94%	6%	0%
2 μM	0%	81%	19%	78%	22%	0%
5 μM	0%	65%	35%	51%	49%	0%
10 μM	0%	53%	47%	41%	59%	0%

5'GGATCC	1st melting curve			2nd melting curve		
	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 68.5 -72.6 °C	Duplex+Ligand (2nd adduct) T_m 90.1 - 90.5°C	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 71.5 -71.8 °C	Duplex+Ligand (2nd adduct) T_m 89.7 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	93%	7%	0%	100%	0%	0%
0.5 μM	12%	88%	0%	93%	7%	0%
1 μM	0%	82%	18%	73%	27%	0%
2 μM	0%	62%	38%	55%	45%	0%
5 μM	0%	33%	67%	16%	84%	0%
10μM	0%	20%	80%	0%	89%	11%

Table 4.6: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzothiophene (KMR32) with oligonucleotide duplexes containing the sequences GTATAC and GTTAAC

5'-GTATAC	1st melting curve			2nd melting curve		
	Unmodified duplex) T_m 36.5 °C	Duplex+Ligand (1st adduct) T_m 64.8 - 65.3 °C	Duplex+Ligand (2nd adduct) T_m 85.6 - 85.9 °C	Unmodified duplex T_m 36.6 °C	Duplex+Ligand (1st adduct) T_m 63.7 - 64.6 °C	Duplex+Ligand (2nd adduct) T_m 86.1°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	92%	8%	0%	100%	0%	0%
0.5 μM	4%	71%	25%	100%	0%	0%
1 μM	0%	48%	52%	92%	8%	0%
2 μM	0%	31%	69%	87%	13%	0%
5 μM	0%	21%	79%	73%	27%	0%
10 μM	0%	21%	79%	54%	29%	17%

5'-GTTAAC	1st melting curve			2nd melting curve		
	Unmodified duplex T_m 37.8 °C	Duplex+Ligand (1st adduct) T_m 63.8 - 64.1 °C	Duplex+Ligand (2nd adduct) T_m 85.6 °C	Unmodified duplex T_m 38.1 °C	Duplex+Ligand (1st adduct) T_m 63.9 - 64.2 °C	Duplex+Ligand (2nd adduct) T_m 89.2°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	90%	10%	0%	100%	0%	0%
0.5 μM	30%	63%	7%	100%	0%	0%
1 μM	0%	53%	47%	97%	3%	0%
2 μM	0%	40%	60%	93%	7%	0%
5 μM	0%	16%	84%	86%	14%	0%
10 μM	0%	18%	82%	79%	18%	0%

Table 4.7: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzothiophene (KMR32) with oligonucleotide duplexes containing the sequences GCTTGC and GCTAGC

5'-GCTTGC	1st melting curve			2nd melting curve		
	Unmodified duplex T_m 44.6 °C	Duplex+Ligand (1st adduct) T_m 73.3 - 74.4 °C	Duplex+Ligand (2nd adduct) T_m 91.1 -91.9 °C	Unmodified duplex T_m 44.8 °C	Duplex+Ligand (1st adduct) T_m 73.4 -74.4 °C	Duplex+Ligand (2nd adduct) T_m 89.2°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	91%	9%	0%	100%	0%	0%
0.5 μM	14%	86%	0%	94%	6%	0%
1 μM	0%	77%	23%	83%	17%	0%
2 μM	0%	52%	48%	59%	41%	0%
5 μM	0%	44%	56%	27%	73%	0%
10 μM	0%	38%	62%	14%	64%	22%

5'-GCTAGC	1st melting curve			2nd melting curve		
	Unmodified duplex T_m 43.9 °C	Duplex+Ligand (1st adduct) T_m 71.9 -73.6 °C	Duplex+Ligand (2nd adduct) T_m 91.5 - 91.7 °C	Unmodified duplex T_m 44.3 °C	Duplex+Ligand (1st adduct) T_m 73.0 -74.1 °C	Duplex+Ligand (2nd adduct) T_m 89.4°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	88%	12%	0%	100%	0%	0%
0.5 μM	16%	78%	6%	88%	12%	0%
1 μM	5%	77%	18%	81%	19%	0%
2 μM	0%	52%	48%	60%	40%	0%
5 μM	0%	38%	62%	30%	59%	11%
10 μM	0%	33%	67%	13%	60%	27%

Table 4.8: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzothiophene (KMR32) with oligonucleotide duplexes containing the sequence GAATGC

	1st melting curve			2nd melting curve		
5'-GAATGC	Unmodified duplex T_m 40.7 °C	Duplex+Ligand (1st adduct) T_m 73.5 - 74.3 °C	Duplex+Ligand (2nd adduct) T_m 90.0 -91.1 °C	Unmodified duplex T_m 40.9 °C	Duplex+Ligand (1st adduct) T_m 74.0 -74.3 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	83%	17%	0%	100%	0%	0%
0.5 μM	0%	100%	0%	83%	17%	0%
1 μM	0%	95%	5%	61%	39%	0%
2 μM	0%	71%	29%	32%	68%	0%
5 μM	0%	57%	43%	14%	86%	0%
10 μM	0%	43%	57%	0%	100%	0%

4.4.3.3 *PBD-Py-Benzothiophene-Acetate (KMR33)*

The effect of PBD-Py-Benzothiophene-Acetate (KMR33) on the melting profiles of these oligonucleotides is shown in Figures 4.19 and 4.20 and the percentages of each species are shown in Tables 4.9 to 4.12.

GGTACC and GGATCC

GGTACC shows T_m values of 41.4 °C, 71.5 °C and 89.5 °C for the unmodified duplex, 1st adduct and 2nd adduct respectively, whereas for GGATCC these are 42.3 °C, 74.7 °C and 92.1 °C. Both oligonucleotides show a single melting transition corresponding to the 1st adduct with 0.5 μM ligand. The T_m of the 2nd adduct appears at higher concentrations, rising to about 75% with 10 μM ligand for both oligonucleotides. In the second melting curve, GGTACC shows transitions for only the unmodified duplex and the 1st adduct at concentrations up to 5 μM, though the 2nd adduct appears with 10 μM. In contrast, GGATCC shows transitions for the 1st adduct and 2nd adduct with 5 μM at which concentration no unmodified duplex remains. The percentage of the 2nd adduct then reaches 28% with 10 μM.

GTATAC and GTTAAC

For oligonucleotide GTATAC, 0.5 μM KMR33 produces melting transitions for the 1st and 2nd adducts at 65.3 °C and 87.2 °C**.** Interestingly, with 5 μM and 10 μM, a single transition is observed, which mainly corresponds to the 2nd adduct, and it not possible to assess the small percentage of the 1st adduct that remains in the leading edge of the transition. In contrast with GTTAAC, 0.5 μM ligand produces a transition which is predominantly that of the 1st adduct. The transition of the 2nd adduct appears at higher ligand concentrations, reaching 72% with 10 μM. In the second melting curve, no 2nd adduction formation is observed in contrast to the result with GTATAC, and with 10 μM ligand, for which 60% of the unmodified duplex is still evident.

GCTTGC, GCTAGC and GAATGC

KMR33 produces similar effects on the melting profiles of oligonucleotides GCTTGC and GCTAGC. The three transitions are observed at 44.6 °C, 75.7 °C and 92.1 °C for GCTTGC and 43.9 °C, 74.8 °C and 92.3 °C for GCTAGC. In the first melting curves both oligonucleotides start to form adducts with 0.5 μM ligand, with similar proportions of each adduct at 2 μM; the percentage of 2nd adduct then increases to about 60% for both sequences with 10 μM ligand. In the second melting curves, three transitions are seen for both oligonucleotides with 5 μM and 10 μM ligand, with 25% and 42% of the 2nd adduct with 10 μM for GCTTGC and GCTAGC respectively. For GAATGC, T_m values are observed at 40.7 °C, 74.1 °C and 91.2 °C in the first melting curve. About

87% of the transition corresponds to the 1st adduct with 0.5 μM ligand; the proportion of the 2nd adduct increases at higher concentrations, reaching about 67% with 10 μM. In the second melting curve, the transitions of unmodified duplex and 1st adduct are seen up to 2 μM. With 5 μM, only 5% is in 2nd adduct form rising to 16% with 10 μM.

Comparing the melting profiles of all oligonucleotide duplexes with KMR33, it is observed that GGTACC and GGATCC differ in adduct formation in the second melting curve. Both oligonucleotides duplexes show almost equal amplitude of 1st and 2nd adduct with 10 μM in the first melting curve. But, in the second melting curve, no unmodified duplex is seen in GGATCC with 28% of 2nd adduct whereas in GGTACC, only 8% of 2nd adduct is formed with 31% of unmodified duplex. In case of GTATAC and GTTAAC, both oligonucleotides show complete formation of adducts in their first melting curves with 10 μM. However, in the second melting curve, GTATAC shows preference over GTTAAC as 28% of 2nd adduct is observed in GTATAC while no 2nd adduct is evident in GTTAAC. For oligonucleotide duplexes GCTTGC and GCTAGC, about 40% and 60% of 1st and 2nd adduct are seen respectively with 10 μM in their first melting curves. However, GCTAGC forms a greater percentage (42%) of the 2nd adduct than GCTTGC (25%) in the second melting curve, while GAATGC shows the lowest percentage of the 2nd adduct (16%).

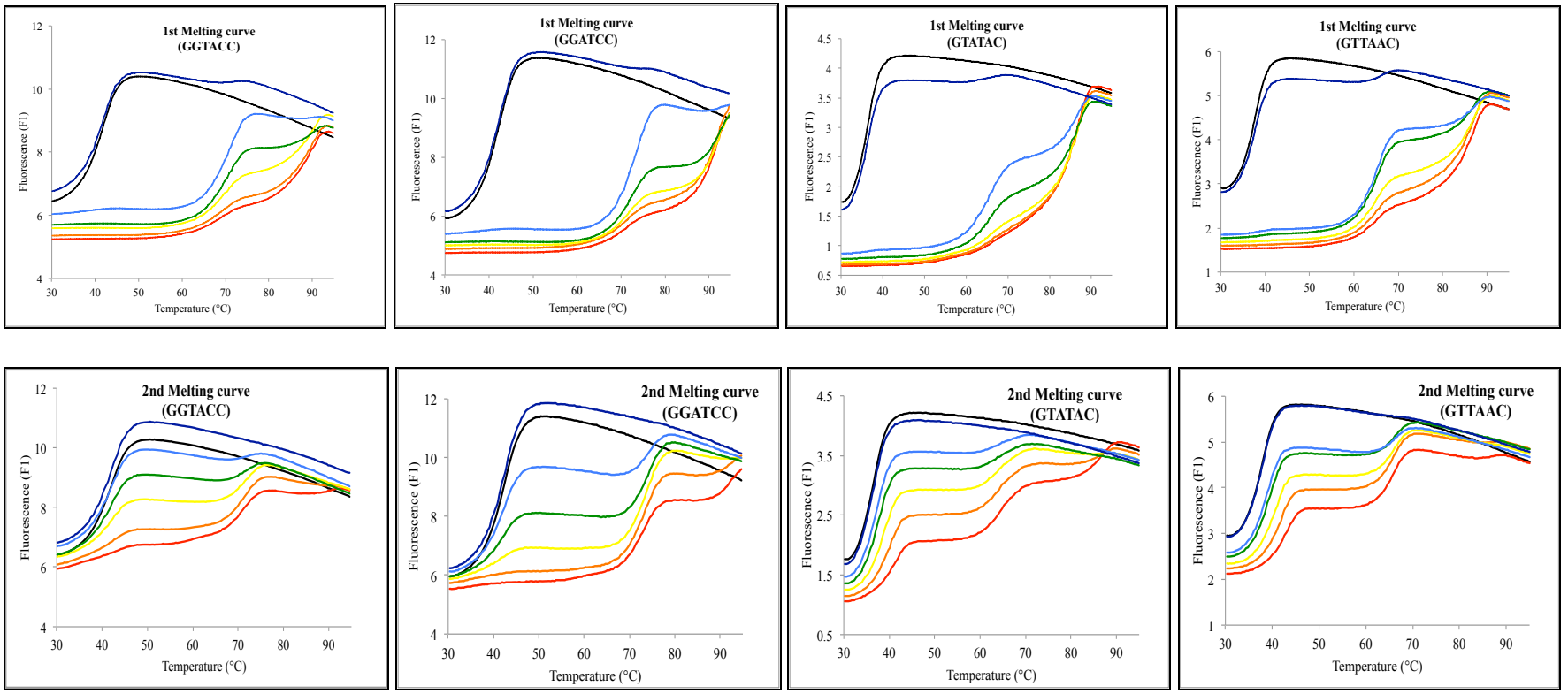


Figure 4.19: First and second melting curves of oligonucleotide duplexes containing the sequences GGTACC, GGATCC, GTATAC and GTTAAC in the presence of different concentrations of PBD-Py-Benzothiophene-Acetate (KMR33)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green), 0.5 μM (light blue) and 0.1 (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting. 'Control' indicates sample without any added ligand.

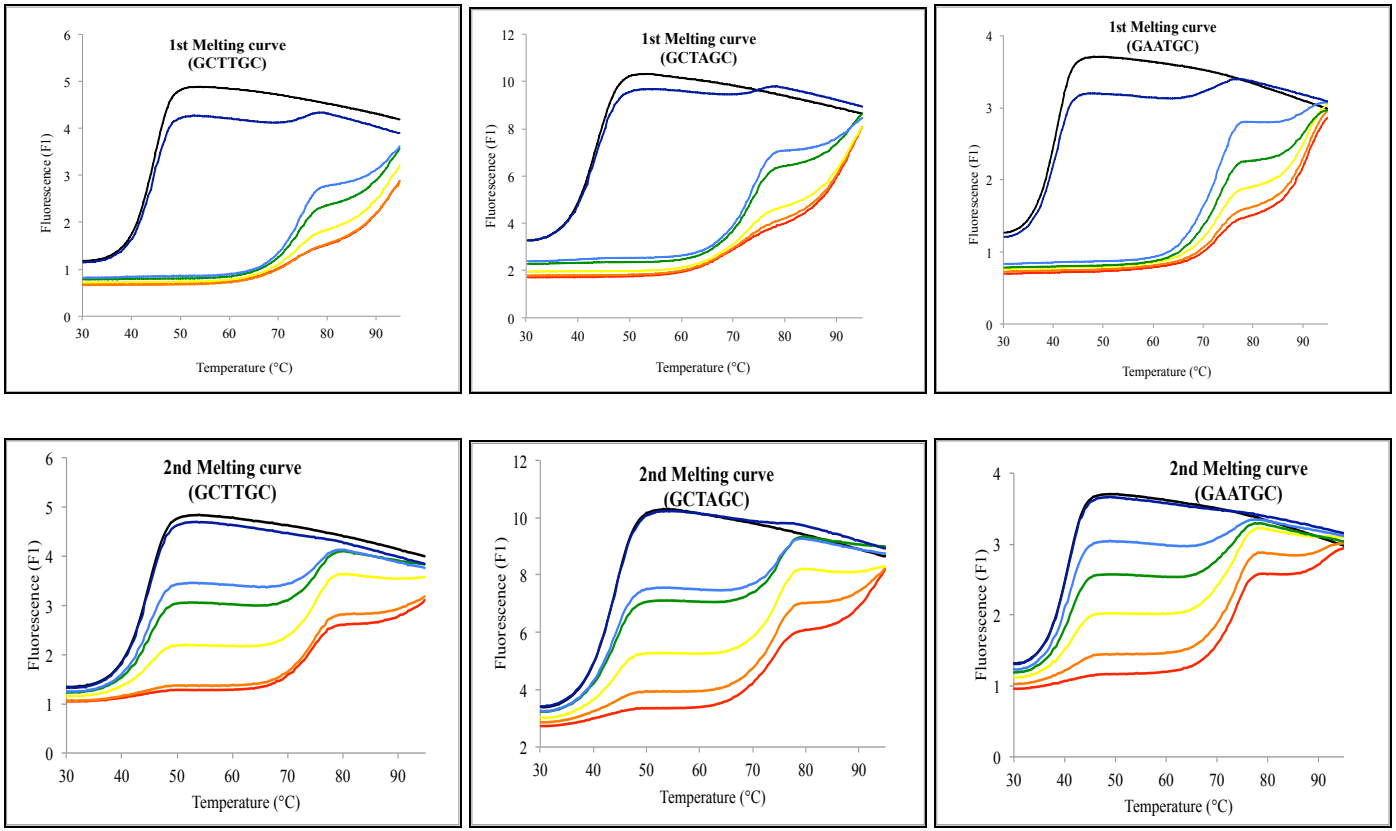


Figure 4.20: First and second melting curves of oligonucleotide duplexes containing the sequences GCTTGC, GCTAGC and GAATGC in the presence of different concentrations of PBD-Py-Benzothiophene-Acetate (KMR33)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green), 0.5 μM (light blue) and 0.1 (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting. 'Control' indicates sample without any added ligand.

Table 4.9: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzothiophene-Acetate (KMR33) with oligonucleotide duplexes containing the sequences GGTACC and GGATCC.

	1st melting curve			2nd melting curve		
5'-GGTACC	Unmodified duplex T_m 41.4 °C	Duplex+Ligand (1st adduct) T_m 69.5 -71.5 °C	Duplex+Ligand (2nd adduct) T_m 89.5 °C	Unmodified duplex T_m 41.7 °C	Duplex+Ligand (1st adduct) T_m 71.1 -72.1°C	Duplex+Ligand (2nd adduct) T_m 89.6°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	0%	100%	0%	94%	6%	0%
1 μM	0%	84%	16%	83%	17%	0%
2 μM	0%	44%	56%	63%	37%	0%
5 μM	0%	35%	65%	40%	60%	0%
10 μM	0%	23%	77%	31%	61%	8%

	1st melting curve			2nd melting curve		
5'-GGATCC	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 71.5 -74.7 °C	Duplex+Ligand (2nd adduct) T_m 92.0 -92.1 °C	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 73.0 -74.2 °C	Duplex+Ligand (2nd adduct) T_m 89.7°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	100%	0%	0%	100%	0%	0%
0.5 μM	0%	100%	0%	72%	28%	0%
1 μM	0%	57%	43%	46%	54%	0%
2 μM	0%	41%	59%	19%	81%	0%
5 μM	0%	29%	71%	0%	84%	16%
10 μM	0%	26%	74%	0%	72%	28%

Table 4.10: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzothiophene-Acetate (KMR33) with oligonucleotide duplexes containing the sequences GTATAC and GTTAAC.

	1st melting curve			2nd melting curve		
5'-GTATAC	Unmodified duplex T_m 36.5 °C	Duplex+Ligand (1st adduct) T_m 65.0 -65.3 °C	Duplex+Ligand (2nd adduct) T_m 86.9 - 87.2 °C	Unmodified duplex T_m 36.6 °C	Duplex+Ligand (1st adduct) T_m 63.7 - 64.5 °C	Duplex+Ligand (2nd adduct) T_m 86.7°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	96%	4%	0%	100%	0%	0%
0.5 μM	0%	60%	40%	85%	15%	0%
1 μM	0%	44%	56%	83%	17%	0%
2 μM	0%	35%	65%	70%	30%	0%
5 μM	0%	100%*		55%	33%	12%
10μM	0%	100%*		37%	35%	28%

* Difficult to distinguish which entity it corresponds

	1st melting curve			2nd melting curve		
5'-GTTAAC	Unmodified duplex T_m 37.8 °C	Duplex+Ligand (1st adduct) T_m 64.9 - 65.7 °C	Duplex+Ligand (2nd adduct) T_m 87.0 °C	Unmodified duplex T_m 38.1 °C	Duplex+Ligand (1st adduct) T_m 65.7 - 66.0 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	90%	10%	0%	100%	0%	0%
0.5 μM	0%	73%	27%	82%	18%	0%
1 μM	0%	66%	34%	79%	21%	0%
2 μM	0%	39%	61%	69%	31%	0%
5 μM	0%	33%	67%	60%	40%	0%
10 μM	0%	28%	72%	56%	44%	0%

Table 4.11: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzothiophene-Acetate (KMR33) with oligonucleotide duplexes containing the sequences GCTTGC and GCTAGC

	1st melting curve			2nd melting curve		
5'-GCTTGC	Unmodified duplex T_m 44.6 °C	Duplex+Ligand (1st adduct) T_m 72.8 -75.7 °C	Duplex+Ligand (2nd adduct) T_m 92.1 °C	Unmodified duplex T_m 44.8 °C	Duplex+Ligand (1st adduct) T_m 74.6 -75.8 °C	Duplex+Ligand (2nd adduct) T_m 89.6°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	94%	6%	0%	100%	0%	0%
0.5 μM	0%	68%	32%	61%	39%	0%
1 μM	0%	56%	44%	62%	38%	0%
2 μM	0%	52%	48%	42%	58%	0%
5 μM	0%	43%	57%	19%	62%	19%
10 μM	0%	43%	57%	15%	60%	25%

	1st melting curve			2nd melting curve		
5'-GCTAGC	Unmodified duplex T_m 43.9 °C	Duplex+Ligand (1st adduct) T_m 73.3 -74.8°C	Duplex+Ligand (2nd adduct) T_m 92.3 °C	Unmodified duplex T_m 44.3 °C	Duplex+Ligand (1st adduct) T_m 74.5 -75.3 °C	Duplex+Ligand (2nd adduct) T_m 89.7°C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	94%	6%	0%	100%	0%	0%
0.5 μM	0%	76%	24%	68%	32%	0%
1 μM	0%	65%	35%	67%	33%	0%
2 μM	0%	47%	53%	42%	58%	0%
5 μM	0%	42%	58%	22%	56%	22%
10 μM	0%	39%	61%	8%	50%	42%

Table 4.12: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Py-Benzothiophene-Acetate (KMR33) with oligonucleotide duplexes containing the sequence GAATGC

5'-GAATGC	1st melting curve			2nd melting curve		
	Unmodified duplex T_m 40.7 °C	Duplex+Ligand (1st adduct) T_m 73.1 -74.1 °C	Duplex+Ligand (2nd adduct) T_m 90.2 - 91.2 °C	Unmodified duplex T_m 40.9 °C	Duplex+Ligand (1st adduct) T_m 73.9 °C	Duplex+Ligand (2nd adduct) T_m 89.9 °C
Control	100%	0%	0%	100%	0%	0%
0.1 μM	91%	9%	0%	100%	0%	0%
0.5 μM	0%	87%	13%	86%	14%	0%
1 μM	0%	68%	32%	67%	33%	0%
2 μM	0%	48%	52%	50%	50%	0%
5 μM	0%	36%	64%	21%	74%	5%
10 μM	0%	33%	67%	11%	73%	16%

4.4.3.4 *PBD-Im-Benzothiophene-Acetate (KMR173)*

Figures 4.21 and 4.22 show the effect of PBD-Im-Benzothiophene-Acetate (KMR173) on the different oligonucleotides; the percentages of unmodified duplex, 1st adduct and 2nd adduct are presented in Tables 4.13 to 4.16.

GGTACC and GGATCC

Oligonucleotide GGTACC shows no 2nd adduct transition with KMR173 in both the first and second melting curves. The first melting curve shows a single transition with a T_m of 67.5 °C at concentrations of 2 μM and above. Oligonucleotide GGATCC shows transitions for unmodified duplex and 1st adduct, with T_ms of 42.3 °C and 72 °C respectively at concentrations up to 2 μM. The 2nd adduct appears at higher concentrations and corresponds to 43% of the melting profile with 10 μM. In the second melting curve, no transition for the 2nd adduct is observed. Some unmodified duplex is still evident at concentrations up to 2 μM, though a single melting transition for the 1st adduct is observed at higher concentrations.

GTATAC and GTTAAC

Oligonucleotide GTATAC shows three transitions with 0.5 μM at 36.5 °C, 63.7 °C and 82.9 °C for unmodified duplex, 1st adduct and 2nd adduct respectively forming 69% of 1st adduct with only 7% of 2nd adduct. Percentages of 1st and 2nd adduct reach about 50% with 5 μM and 10 μM with no unmodified duplex. On the other hand, GTTAAC shows three transitions even at highest concentration forming only 59% and 28% of 1st and 2nd adduct respectively with the remainder as unmodified duplex with 10 μM. The second melting curve of both oligonucleotides corresponds to the unmodified duplex even at highest ligand concentration 10 μM.

GCTTGC, GCTAGC and GAATGC

For GCTTGC, the three transitions are observed at 44.6 °C, 76.4 °C and 90.5 °C. The 1st adduct represent 17% of the profile with 0.1 μM ligand and this is the only melting transition seen with 1 μM and 2 μM. The 2nd adduct starts to appear with 5 μM and increases to 28% with 10 μM. In the second melting curve, the unmodified duplex predominates up to 1 μM with no 2nd adduct evident even at the highest concentration. With 2 μM, 75% of the melting profile corresponds to the 1st adduct, and this is the only transition with 5 μM and 10 μM. For GCTAGC, transitions are seen at 43.9 °C, 72.1 °C and 89 °C for unmodified duplex, 1st adduct and 2nd adduct respectively. Very little (~4%) of 1st adduct is found with 0.1 μM, but this increases up to 93% with 2 μM; at this point the 2nd adduct begins to appears (~7%) increasing to 46% with 10 μM. In the second melting curve, no transition for the 2nd adduct is observed. The transition for the

1st adduct does not appear until 1 μM ligand; the proportion of this transition increases at higher concentrations and corresponds to the entire melting curve at 5 μM. For GAATGC the 1st adduct, with a Tm of 73.4 °C, corresponds to 73% of the profile at 0.5 μM with the remainder as the unmodified duplex and this is the only transition with 1 μM and 2 μM. The 2nd adduct corresponds to 11% of the profile with 5 μM, rising to 35% with 10 μM. In the second melting curve, the proportion of 1st adduct increases to about 80% with 10 μM, with no 2nd adduct formation.

These results will be considered further in the Discussion, but some indication of the selectivity of this ligand can be estimated by comparing the effects of the highest ligand concentration (10 μM) on the first melting curves. Only the 1st adduct is observed with GGTACC, while the closely related sequence GGATCC shows 43% as the 2nd adduct. Similarly, for GTATAC 10 μM ligand only produces the two transitions for the 1st and 2nd adduct, whereas with GTTAAC some unmodified duplex is still present. For GCTTGC, 28% is found as the 2nd adduct while GCTAGC shows about 46%. The sequence GAATGC, containing only one GC step behaves in an intermediate fashion with 35% of the transition as the 2nd adduct.

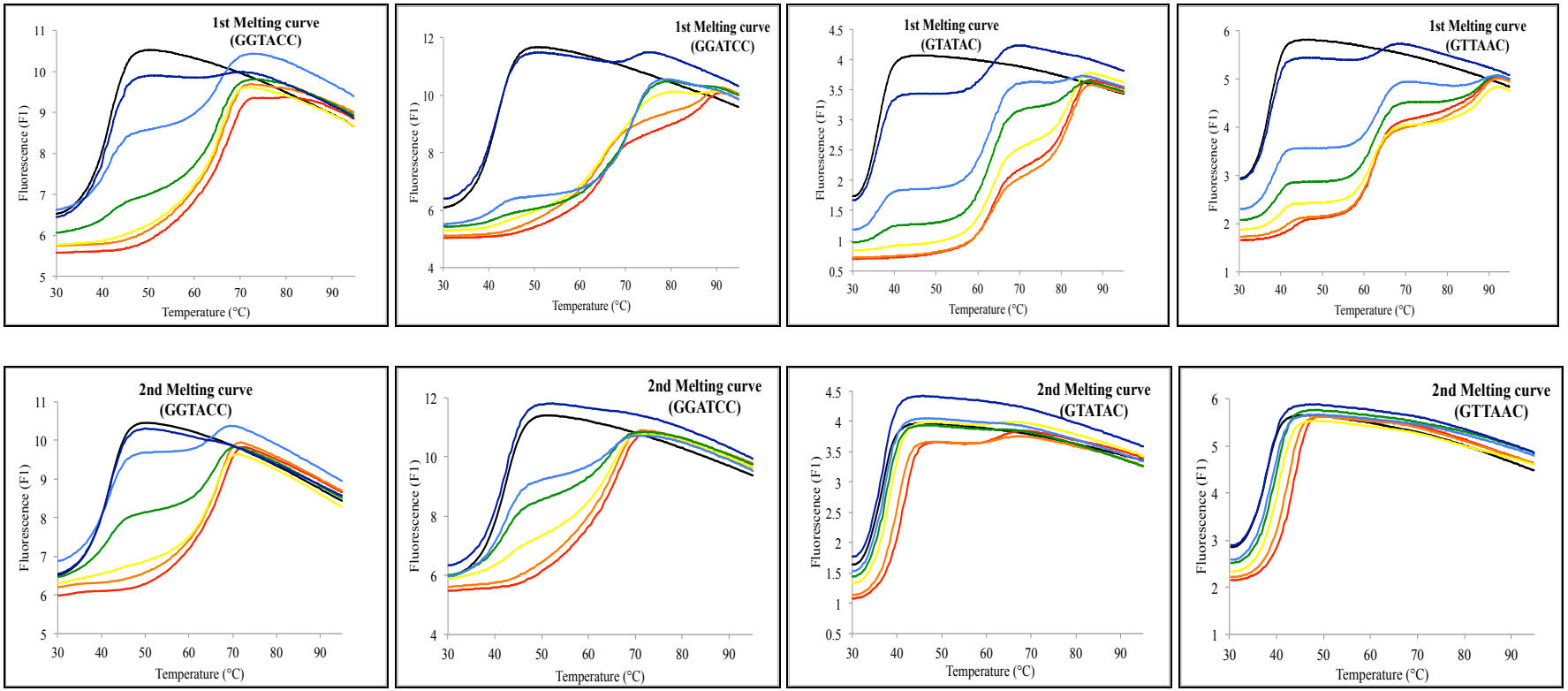


Figure 4.21: First and second melting curves of oligonucleotide duplexes containing the sequences GGTACC, GGATCC, GTATAC and GTTAAC in the presence of different concentrations of PBD-Im-Benzothiophene-Acetate (KMR173)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green), 0.5 μM (light blue) and 0.1 (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting. 'Control' indicates sample without any added ligand.

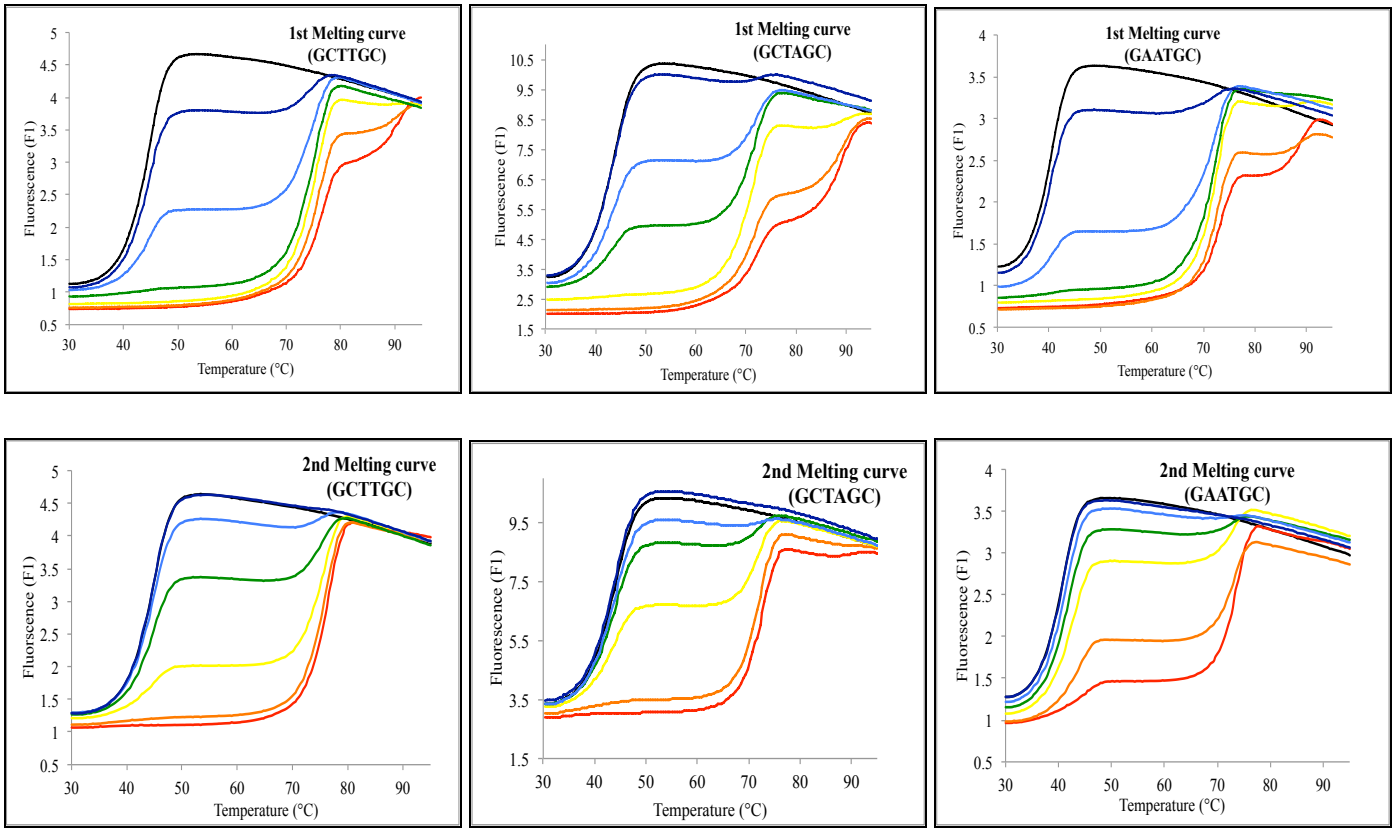


Figure 4.22: First and second melting curves of oligonucleotide duplexes containing the sequences GCTTGC, GCTAGC and GAATGC in the presence of different concentrations of PBD-Im-Benzothiophene-Acetate (KMR173)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green), 0.5 μM (light blue) and 0.1 (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting. 'Control' indicates sample without any added ligand.

Table 4.13: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Im-Benzothiophene-Acetate (KMR173) with oligonucleotide duplexes containing the sequences GGTACC and GGATCC

	1st melting curve			2nd melting curve		
5'-GGTACC	Unmodified duplex T_m 41.4 °C	Duplex+Ligand (1st adduct) T_m 65.6 - 67.5°C	Duplex+Ligand (2nd adduct)	Unmodified duplex T_m 41.7 °C	Duplex+Ligand (1st adduct) T_m 65.9 -67.8 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 µM	94%	6%	0%	100%	0%	0%
0.5 µM	49%	51%	0%	74%	26%	0%
1 µM	22%	78%	0%	45%	55%	0%
2 µM	0%	100%	0%	0%	100%	0%
5 µM	0%	100%	0%	0%	100%	0%
10 µM	0%	100%	0%	0%	100%	0%

	1st melting curve			2nd melting curve		
5'-GGATCC	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 64.6 -72.0 °C	Duplex+Ligand (2nd adduct) T_m 87.4 -87.8 °C	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 64.8 - 66.7 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 µM	94%	6%	0%	100%	0%	0%
0.5 µM	17%	83%	0%	60%	40%	0%
1 µM	10%	90%	0%	51%	49%	0%
2 µM	13%	87%	0%	27%	73%	0%
5 µM	0%	78%	22%	0%	100%	0%
10 µM	0%	57%	43%	0%	100%	0%

Table 4.14: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Im-Benzothiophene-Acetate (KMR173) with oligonucleotide duplexes containing the sequences GTATAC and GTTAAC

	1st melting curve			2nd melting curve		
5'-GTATAC	Unmodified duplex T_m 36.5 °C	Duplex+Ligand (1st adduct) T_m 62.6 - 63.7 °C	Duplex+Ligand (2nd adduct) T_m 82.0 - 82.9 °C	Unmodified duplex T_m 36.6 °C	Duplex+Ligand (1st adduct) T_m 62.1 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	69%	31%	0%	100%	0%	0%
0.5 μM	24%	69%	7%	100%	0%	0%
1 μM	10%	75%	15%	100%	0%	0%
2 μM	0%	58%	42%	100%	0%	0%
5 μM	0%	46%	54%	96%	4%	0%
10 μM	0%	48%	52%	94%	6%	0%

	1st melting curve			2nd melting curve		
5'-GTTAAC	Unmodified duplex T_m 37.8 °C	Duplex+Ligand (1st adduct) T_m 61.4 -63.4 °C	Duplex+Ligand (2nd adduct) T_m 87.3 -88.1 °C	Unmodified duplex T_m 38.1 °C	Duplex+Ligand (1st adduct)	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	96%	4%	0%	100%	0%	0%
0.5 μM	48%	48%	4%	100%	0%	0%
1 μM	26%	59%	15%	100%	0%	0%
2 μM	17%	55%	28%	100%	0%	0%
5 μM	13%	56%	31%	100%	0%	0%
10 μM	13%	59%	28%	100%	0%	0%

Table 4.15: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Im-Benzothiophene-Acetate (KMR173) with oligonucleotide duplexes containing the sequences GCTTGC and GCTAGC

	1st melting curve			2nd melting curve		
5'-GCTTGC	Unmodified duplex T_m 44.6 °C	Duplex+Ligand (1st adduct) T_m 74.5 -76.4 °C	Duplex+Ligand (2nd adduct) T_m 90.5 °C	Unmodified duplex T_m 44.8 °C	Duplex+Ligand (1st adduct) T_m 75.9 -76.8 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	83%	17%	0%	100%	0%	0%
0.5 μM	37%	63%	0%	91%	9%	0%
1 μM	0%	100%	0%	68%	32%	0%
2 μM	0%	100%	0%	26%	74%	0%
5 μM	0%	87%	13%	0%	100%	0%
10 μM	0%	72%	28%	0%	100%	0%

	1st melting curve			2nd melting curve		
5'-GCTAGC	Unmodified duplex T_m 43.9 °C	Duplex+Ligand (1st adduct) T_m 71.8 -72.1 °C	Duplex+Ligand (2nd adduct) T_m 89.0 °C	Unmodified duplex T_m 44.3 °C	Duplex+Ligand (1st adduct) T_m 72.1 -73.0 C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	96%	4%	0%	100%	0%	0%
0.5 μM	61%	39%	0%	97%	3%	0%
1 μM	32%	68%	0%	85%	15%	0%
2 μM	0%	93%	7%	54%	46%	0%
5 μM	0%	58%	42%	0%	100%	0%
10 μM	0%	54%	46%	0%	100%	0%

Table 4.16: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Im-Benzothiophene-Acetate (KMR173) with oligonucleotide duplexes containing the sequences GAATGC

	1st melting curve			2nd melting curve		
5'-GAATGC	Unmodified duplex T_m 40.7 °C	Duplex+Ligand (1st adduct) T_m 72.1 -73.4 °C	Duplex+Ligand (2nd adduct) T_m 88.9 °C	Unmodified duplex T_m 40.9 °C	Duplex+Ligand (1st adduct) T_m 72.9 -73.8 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	88%	12%	0%	100%	0%	0%
0.5 μM	27%	73%	0%	98%	2%	0%
1 μM	4%	96%	0%	92%	8%	0%
2 μM	0%	100%	0%	75%	25%	0%
5 μM	0%	89%	11%	47%	53%	0%
10 μM	0%	65%	35%	20%	80%	0%

4.4.3.5 *PBD-Benzothiophene (KMR175)*

The effect of PBD-benzothiophene (KMR175) on the melting profiles of the oligonucleotides is shown in Figures 4.23 and 4.24 with the percentages of each species in Tables 4.17 to 4.20.

GGTACC and GGATCC

GGTACC shows all three transitions with 2 μM ligand producing T_ms of 41.4 °C, 63.7 °C and 79 °C with 73% as the 1st adduct and only about 9% as the 2nd adduct. The amount of the 2nd adduct increases with 5 μM and 10 μM, representing more than 50% at highest concentration. No adduct formation is observed in the second melting curve. GGATCC also shows three transitions with 1 μM and the percentage of the 2nd adduct increases up to 67% with 10 μM. The second melting curve corresponds to the unmodified duplex at ligand concentrations up to 2 μM. At concentrations of 5 μM and 10 μM, the second melting transition, with a T_m close to 1st adduct, is also observed, though this corresponds to less than 20% of the transition even with 10 μM.

GTATAC and GTTAAC

GTATAC displays three transitions of T_m 36.5 °C, 58.1 °C and 74.1 °C. Some unmodified duplex is still seen at ligand concentrations up to 2 μM at which about 45% and 41% of the profile corresponds to the 1st and 2nd adduct respectively. The proportion of the 2nd adduct then increases up to 82% with 10 μM. GTTAAC shows similar effects. For both oligonucleotides, no change in the melting profiles is observed at any concentration and the curves resemble that of the unmodified duplex.

GCTTGC, GCTAGC and GAATGC

The melting profiles for GCTTGC and GCTAGC are also similar and all three species are seen in the first melting curves at a ligand concentration of 1 μM. The proportion of the 2nd adduct rises at higher concentrations to about 60% with 10 μM. In the 2nd melting curve, both oligonucleotides only show a small amount of the 1st adduct at 5 μM and 10 μM with no formation of 2nd adduct. For GAATGC, three transitions with T_ms of 40.7 °C, 67.4 °C and 85 °C are observed with 1 μM ligand. The amount of the 2nd adduct increases to 86% at 10 μM. In the second melting curve there is no evidence for the 2nd adduct, and only 24% corresponds to the 1st adduct at the highest concentration.

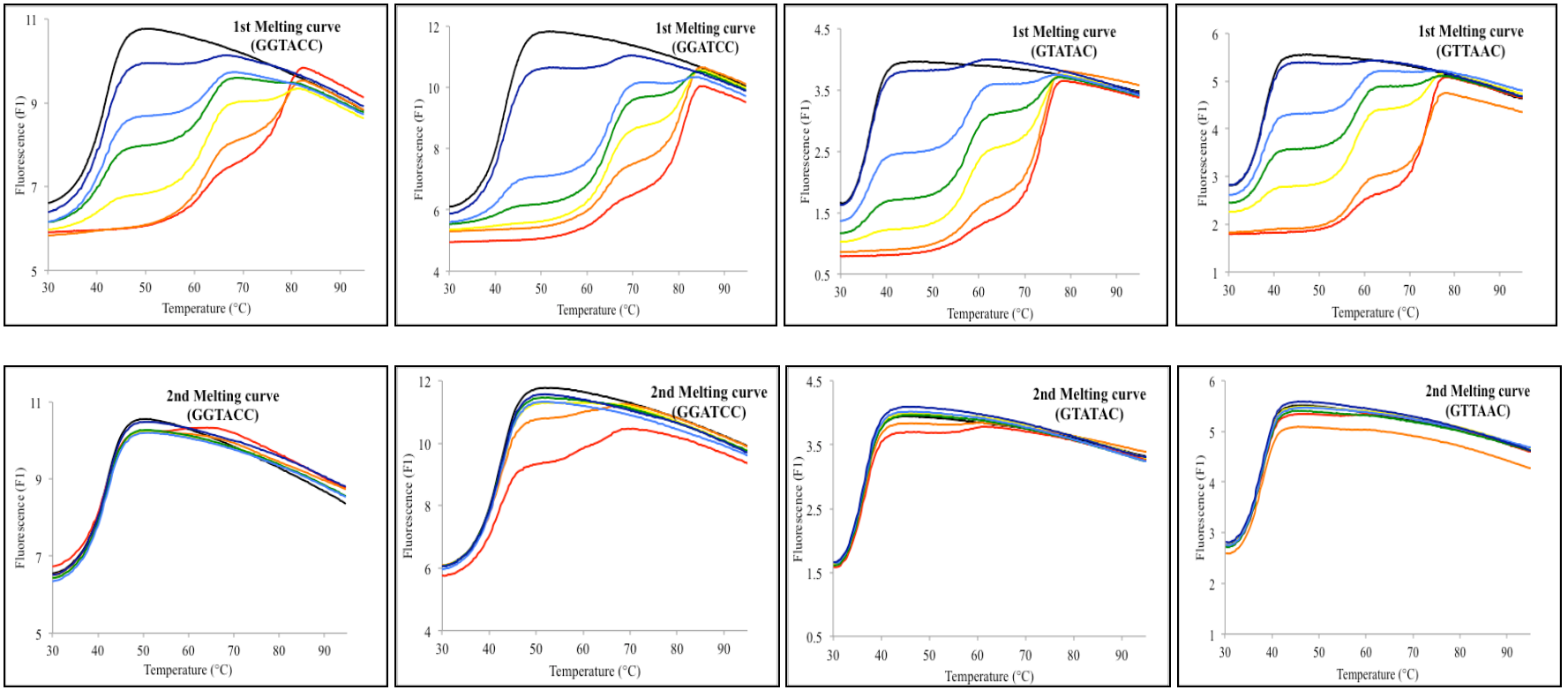


Figure 4.23: First and second melting curves of oligonucleotide duplexes containing the sequences GGTACC, GGATCC, GTATAC and GTTAAC in the presence of different concentrations of PBD-Benzothiophene (KMR175)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green), 0.5 μM (light blue) and 0.1 (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting. 'Control' indicates sample without any added ligand.

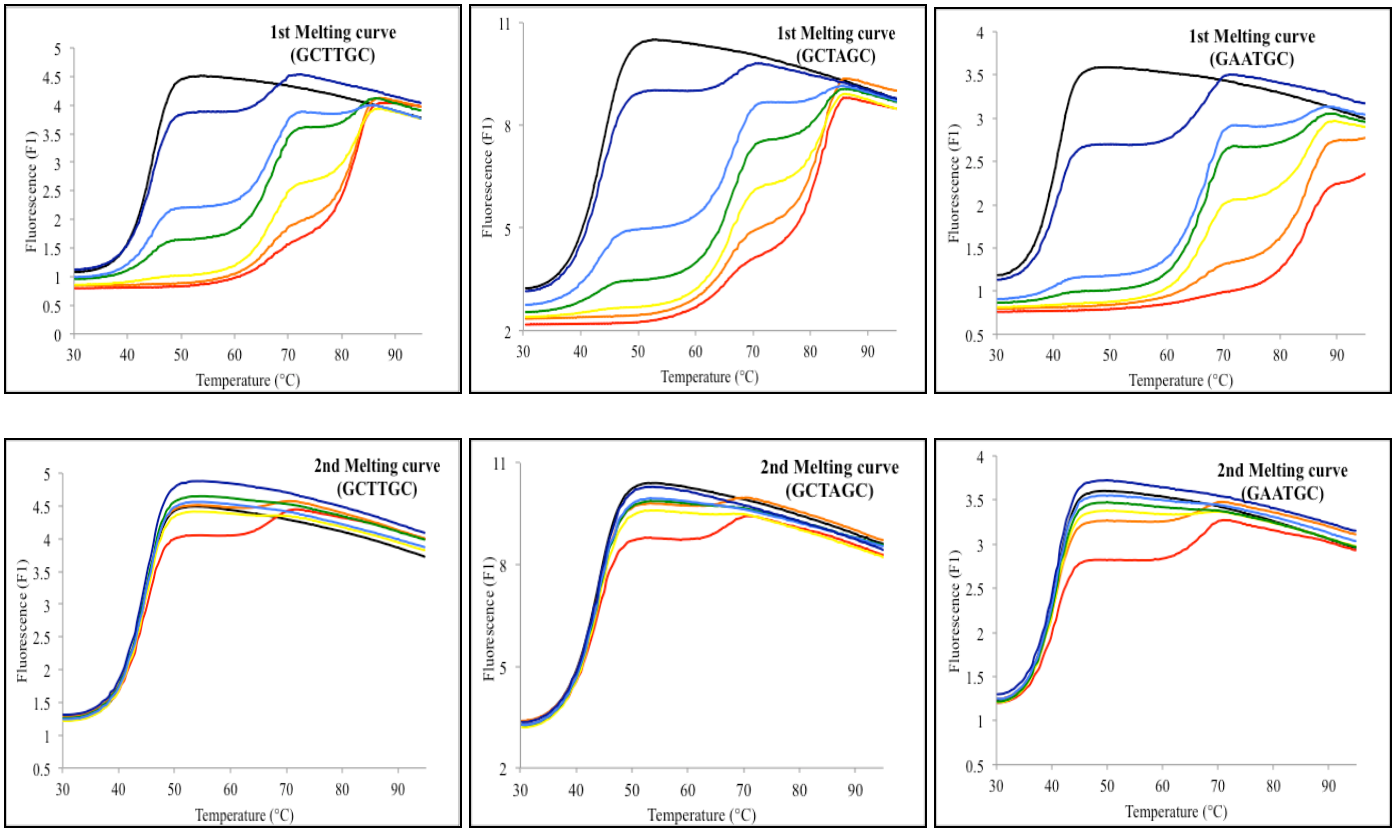


Figure 4.24: First and second melting curves of oligonucleotide duplexes containing the sequences GCTTGC, GCTAGC and GAATGC in the presence of different concentrations of PBD-Benzothiophene (KMR175)- Control (black), 10 μM (red), 5 μM (orange), 2 μM (yellow), 1 μM (green), 0.5 μM (light blue) and 0.1 (dark blue). The y-axis shows the fluorescence, expressed in arbitrary units. Each oligonucleotide duplex was incubated with the ligand at room temperature for 12 hours before subjecting to first melting, annealing and second melting. 'Control' indicates sample without any added ligand.

Table 4.17: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Benzothiophene (KMR175) with oligonucleotide duplexes containing the sequences GGTACC and GGATCC

	1st melting curve			2nd melting curve		
5'-GGTACC	Unmodified duplex T_m 41.4 °C	Duplex+Ligand (1st adduct) T_m 62.0 -63.7 °C	Duplex+Ligand (2nd adduct) T_m 79.0 °C	Unmodified duplex T_m 41.7 °C	Duplex+Ligand (1st adduct) T_m 57.8 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	95%	5%	0%	100%	0%	0%
0.5 μM	74%	26%	0%	100%	0%	0%
1 μM	52%	48%	0%	100%	0%	0%
2 μM	18%	73%	9%	100%	0%	0%
5 μM	0%	59%	41%	100%	0%	0%
10 μM	0%	44%	56%	100%	0%	0%

	1st melting curve			2nd melting curve		
5'-GGATCC	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 64.0 - 66.0 °C	Duplex+Ligand (2nd adduct) T_m 73.1 -74.1 °C	Unmodified duplex T_m 42.3 °C	Duplex+Ligand (1st adduct) T_m 57.9 -65.3 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	92%	8%	0%	100%	0%	0%
0.5 μM	32%	68%	0%	100%	0%	0%
1 μM	12%	67%	21%	100%	0%	0%
2 μM	0%	63%	37%	100%	0%	0%
5 μM	0%	44%	56%	92%	4%, 4%	0%
10 μM	0%	33%	67%	79%	12%, 9%	0%

Table 4.18: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Benzothiophene (KMR175) with oligonucleotide duplexes containing the sequences GTATAC and GTTAAC

	1st melting curve			2nd melting curve		
5'-GTATAC	Unmodified duplex T_m 36.5 °C	Duplex+Ligand (1st adduct) T_m 57.2 -58.1 °C	Duplex+Ligand (2nd adduct) T_m 73.1 -74.1 °C	Unmodified duplex T_m 36.6 °C	Duplex+Ligand (1st adduct)	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 µM	91%	9%	0%	100%	0%	0%
0.5 µM	46%	46%	8%	100%	0%	0%
1 µM	26%	52%	22%	100%	0%	0%
2 µM	14%	45%	41%	100%	0%	0%
5 µM	0%	28%	72%	100%	0%	0%
10 µM	0%	18%	82%	100%	0%	0%

	1st melting curve			2nd melting curve		
5'-GTTAAC	Unmodified duplex T_m 37.8 °C	Duplex+Ligand (1st adduct) T_m 57.8 -59.4 °C	Duplex+Ligand (2nd adduct) T_m 74.6 °C	Unmodified duplex T_m 38.1 °C	Duplex+Ligand (1st adduct)	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 µM	100%	0%	0%	100%	0%	0%
0.5 µM	65%	35%	0%	100%	0%	0%
1 µM	44%	48%	8%	100%	0%	0%
2 µM	19%	54%	27%	100%	0%	0%
5 µM	0%	36%	64%	100%	0%	0%
10 µM	0%	22%	78%	100%	0%	0%

Table 4.19: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Benzothiophene (KMR175) with oligonucleotide duplexes containing the sequences GCTTGC and GCTAGC

	1st melting curve			2nd melting curve		
5'-GCTTGC	Unmodified duplex T_m 44.6 °C	Duplex+Ligand (1st adduct) T_m 66.5 - 68.1°C	Duplex+Ligand (2nd adduct) T_m 82.9 -83.3 °C	Unmodified duplex T_m 44.8 °C	Duplex+Ligand (1st adduct) T_m 67.4 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	82%	18%	0%	100%	0%	0%
0.5 μM	39%	56%	5%	100%	0%	0%
1 μM	20%	66%	14%	100%	0%	0%
2 μM	0%	52%	48%	100%	0%	0%
5 μM	0%	40%	60%	97%	3%	0%
10 μM	0%	35%	65%	89%	11%	0%

	1st melting curve			2nd melting curve		
5'-GCTAGC	Unmodified duplex T_m 43.9 °C	Duplex+Ligand (1st adduct) T_m 65.8 -67.0 °C	Duplex+Ligand (2nd adduct) T_m 82.5 -82.8 °C	Unmodified duplex T_m 44.3 °C	Duplex+Ligand (1st adduct) T_m 66.7 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	86%	14%	0%	100%	0%	0%
0.5 μM	33%	57%	10%	100%	0%	0%
1 μM	13%	65%	22%	100%	0%	0%
2 μM	0%	55%	45%	100%	0%	0%
5 μM	0%	35%	65%	97%	3%	0%
10 μM	0%	40%	60%	89%	11%	0%

Table 4.20: Percentages of the different melting transitions, corresponding to the 1st and 2nd adducts, in the presence of different concentrations of PBD-Benzothiophene (KMR175) with oligonucleotide duplexes containing the sequence GAATGC

	1st melting curve			2nd melting curve		
5'-GAATGC	Unmodified duplex T_m 40.7 °C	Duplex+Ligand (1st adduct) T_m 66.5 -67.4 °C	Duplex+Ligand (2nd adduct) T_m 84.9 -85.0 °C	Unmodified duplex T_m 40.9 °C	Duplex+Ligand (1st adduct) T_m 67.7 °C	Duplex+Ligand (2nd adduct)
Control	100%	0%	0%	100%	0%	0%
0.1 μM	73%	27%	0%	100%	0%	0%
0.5 μM	11%	78%	11%	100%	0%	0%
1 μM	9%	75%	16%	100%	0%	0%
2 μM	0%	40%	60%	100%	0%	0%
5 μM	0%	26%	74%	89%	11%	0%
10 μM	0%	14%	86%	76%	24%	0%

4.5 Discussion

4.5.1 Footprinting studies

All these compounds contain a common structural motif composed of either PBD-pyrrole or PBD-imidazole, and the ligands have similar sequence selectivity. Compounds KMR32 and KMR33 are identical except that there is an additional acetyl group on the benzothiophene in KMR33. These compounds are composed of a PBD with an amide linkage having a pyrrole between them. In compound KMR31, the benzothiophene is replaced by benzofuran. Compound KMR173 has similar structure to KMR33 and differs by having an imidazole ring in place of pyrrole, while compound KMR175 has no pyrrole or imidazole in the linker between PBD and benzothiophene. Here, four different components are present in these compounds- i) PBD, ii) pyrrole or imidazole, iii) benzofuran, benzothiophene or indole and iv) amide linkage. Each component has its own particular effect on DNA binding and their binding affinity and selectivity should not be dependent on a single component. Thus, the sequence selectivity of PBD-benzofused conjugates can be postulated as for the PBD conjugates in the previous section with pyrrole, imidazole and thiazole. Binding will either be controlled by the PBD (*i.e.* PBD will covalently bind to guanine with subsequent recognition of pyrrole or imidazole in the linker and the benzofused rings) or the PBD may have no role in determining the sequence specificity, and the binding will be dictated only by the heterocyclic moieties, though covalent attachment will still require the presence of guanine.

The footprinting results suggest that the benzofused ring must have some effect on the binding of these ligands as compounds KMR30 (PBD-Py-Indole) and KMR171 (PBD-Im-Indole), which only contain an added indole did not produce any footprints, whereas KMR31 (PBD-Py-Benzofuran) and KMR32 (PBD-Py-Benzothiophene) showed large footprints. The sequence selectivity of KMR31 and KMR32 seemed very similar; both compounds produced footprints at the same locations but differ in concentration dependence. If we consider the protections of sequences at lower concentrations, the binding sites found in HexAfor at 5'TCGATA, 5'ATAGATCT, 5'AACGTT and 5'TAAGTA clearly demonstrating these predominance of AT-rich sequences having one guanine (underlined). Protections at AT-rich sequences are also evident in HexBfor at 5'TTAATGCAT, 5'ATATTC, 5'CATATGTA and 5'CTTATA and in MS1 at 5'GAATGCG, 5'TTTGTC, 5'TATGTAAAGTA, 5'GCAATTAGG and 5'TTAAG. Examining the flanking regions of most of these binding sites, it is observed that these sites contain 'G' or 'GC' with any combination of four base pairs of adenine and/or thymine such as TATA, ATAT, AATT and TTAA at 5'-end or 3'-end. Binding at these AT-rich sequences

suggests that the ligand is first positioned with AT rich regions with subsequent attachment of the PBD moiety to guanine. Large footprints at higher concentrations such as at 5'-ATGCTAGCGCTTAAGTACTAGTGC (in HexA) and 5'-TAGCCGGCAATTGCAA (in HexB) are also rich in adenine and thymine; these large protections clearly indicate the presence of overlapping binding sites. Though the large footprints that are seen at high concentration contain some GC steps these are not protected at lower concentrations. Moreover, a number of GC-rich sequences such as GGCGCC, CCGCGG, GCGC, GGCC, CCGGCC, GCGCGC, CGCGCG, GGGCG, CCCG, GGCCG, etc. are present in each of the DNA fragments and none of these were protected by KMR31 and KMR32, which again emphasizes the preference of KMR31 and KMR32 for AT-rich regions.

Addition of an acetyl moiety to KMR32 clearly has some effects on the binding as is evident from the footprints of KMR33 (PBD-Py-Benzothiophene-Acetate). KMR33 bound to the same regions AT-rich sequences as KMR32, but some of these footprints are found protecting longer sequences at lower concentrations. The compound showed same preferences for AT-rich regions in HexB and MS1. Two large footprints in HexAfor at 5'-TTTAAACGTTAACCGGTACCTAGG and 5'-GGGATATCGATA were produced at the highest concentrations; but, considerable differences in size of binding sites were seen at lowers concentrations. For the former footprint, KMR33 showed binding at 5'-TTTAAACGTTAACCGG with lowest concentration while protection at 5'AACGTT was found with KMR32; the later footprint was shortened to 5'TCGATA with KMR32 while KMR33 did not show any change in size of binding sites at lower concentrations.

Replacement of pyrrole with imidazole reduces the affinity of the ligand as seen with KMR173 (PBD-Im-Benzothiophene-Acetate), which requires higher concentrations to produce footprints as seen with HexA and MS1. The binding sites of KMR173 in HexA and MS1 are similar to those of KMR33 (PBD-Py-Benzothiophene-Acetate) and are again at AT-rich regions that contain one guanine (underlined) such as 5'ATAGATCT, 5'AACGTTAA, 5'AAGTAC, 5'TTGCTA, 5'GATAGAA, etc. The imidazole may impart some preference towards GC-rich regions and binding at 5'GCCGTTAACGTT was observed for both KMR33 and KMR173 at the highest concentrations, though KMR173 still bound to 5'GCCG at lower concentrations. Binding was also observed at 5'CGCGTA, though only at highest concentration; it should be noted there are plenty of other GC-rich sequences (as mentioned above), which are not protected by KMR173.

The presence of pyrrole or imidazole in these PBD-benzofused conjugates is clearly important as KMR175 (PBD-benzothiophene) displays different binding selectivity. This

ligand bound to fewer sequences than the ligands with pyrrole- or imidazole-containing linkers. Although this ligand still bound to AT-rich sequences containing one guanine most of the footprints with KMR175, were shorter sequences even at highest concentration. In addition, this ligand exclusively bound at 5'TCCGG and 5'AAGCG; both of these binding sites have similar flanking sequences (italic) at their 5'-ends: 5'-*GAAT*TCCGG and 5'-*TTAT*AAGCG. Thus it can be assumed that KMR175 prefers to bind to a guanine with AT, AA, TA or TT on either the 5'-end or 3'-end, as similar binding sites were observed at 5'-GAA, 5'-TAAG, 5'-TTG, 5'-ATG and 5'-TAG.

The similar sequence specificity of KMR31 (PBD-Py-Benzofuran), KMR32 (PBD-Py-Benzothiophene), KMR33 (PBD-Py-Benzothiophene-Acetate) and KMR173 (PBD-Im-Benzothiophene-Acetate) might be attributed to possible intercalation of the benzofuran and benzothiophene rings. Shi *et al* (283,284) reported that glycosylated 2-phenyl-benzofuran, 2-phenyl-benzothiophene and 2-phenyl-indole, which have potential antitumor activities, showed intercalation of the benzofuran and benzothiophene rings. Intercalation producing cytotoxicity has also been reported for benzofuran-3-ols (281) and benzothiophene carboxanilides (282). Thus, in PBD conjugates with benzofuran and benzothiophene, the pyrrole or imidazole area possibly positioned in the DNA minor groove with simultaneous intercalation by benzofuran or benzothiophene and finally covalent bond formation by PBD with guanine. On the other hand, indole-containing compounds (283,284) were found to produce weak intercalation. The indole moiety of duocarmycin SA (285) was reported to be positioned in the DNA minor groove. Duocarmycin SA is a potent antitumor antibiotic, which is composed of a spirocyclopropyl-indole linked to trimethyoxy-indole by a carbonyl moiety; the compound binds at AATTA in the minor groove. The cyclopropane ring, which is fused to indole, is responsible for alkylation. Replacing the attached trimethyoxy-indole with a simple indole significantly slowed the rate of alkylation indicating the importance of substituents in indole ring interaction with the target sequences. Several small indole derivatives (286) have also been described as minor groove binders, though the interaction was again greatly influenced by nature and position of the side chains. In these PBD-indole conjugates, the indole ring might not facilitate positioning the PBD moiety along with pyrrole (in KMR30) or imidazole (in KMR171) in DNA minor groove thus the ligands could not bind at the sites like other conjugates. With KMR164, the presence of the phenyl ring may also hinder its binding to DNA.

Footprints at different time intervals also demonstrate the preferences of KMR31, KMR32 and KMR33 towards AT-rich regions at different rates. All three compounds

binds promptly within 1 minute with a long sequence of 5'-TTAGCCGGCAATTGCAAG-3' at a concentration of 5 μM; the site is composed to AT-rich regions separated by stretches of guanines and cytosines. The site is protected even protected at 1 μM with 5 minutes of incubation. Another AT-rich site- 5'-ATATTCG is protected within 1 minute by KMR31 (5 μM) while KMR32 and KMR33 bind after 24 hours and 3 hours respectively. On the other hand, 5'-CATATGTA is protected within 5 minutes with 1 μM of KMR33 while the same site is observed after 1 hour and 3 hours by KMR31 and KMR32 respectively. Comparing all these sequences, it can be suggested that KMR31, KMR32 and KMR33 do not differ in preferences for AT-rich sites but they do prefer different combinations of adenines and thymines. The binding sites of KMR173 displays presence of multiple guanines but these sites are also adjacent to AT rich regions. Binding at 5'-TCCGGA is observed after one minute- this site has AAT at 5'-end. Another binding at 5'-AGTAC is also flanked by TTA. All these compounds initially bind at their most preferred sites, which is then followed by extension or annexation of sites over a longer incubation period, producing multiple sites in larger footprints.

4.5.2 Melting studies

Melting studies were performed with oligonucleotides containing different potential 6-bp binding sites. These oligonucleotides differ in the number and position of the guanines; two of these duplexes contain two guanines at either end of an A/T tract, while the other five duplexes each contain four guanines, with two at either end of the central region as GpC or GpG steps. These multiple guanines present several potential attachment sites for the PBD moiety, several of which will be mutually exclusive. For example, with GGATCC, four different PBD mono-adducts can be formed, two at the inner guanines and two at the outer guanines in the asymmetric duplex, though they may have similar effects on the T_m. Examining the T_m values of the 1st adduct, show that these are very similar for all the ligand concentrations. The melting curves display multi-transitions at different ligand concentrations clearly indicating the formation of adducts containing two PBD units with each duplex. Since the attachment is covalent, the T_m shifts are all-or-none *i.e.* the T_m of each transition does not change with increasing ligand concentration, but the proportion does. This contrasts with melting curves obtained with ligands that bind reversibly, in which the T_m usually increases with increasing ligand concentration. This observation provides further confirmation that these PBD conjugates do indeed for covalent adducts with their targets. The changes in the melting curves are always less pronounced in the second melting curve, and in several instances these show no evidence for the 2nd adduct. The covalent attachment of the ligand is broken at high temperatures during the first melting reaction and,

although it can re-attach at lower temperatures, this is a slow process, which presumably requires the prior reannealing of the DNA duplex, which accounts for the observation that there are no ligand-induced changes in the annealing temperature of any of these duplexes. Once the duplex has reannealed the ligand can slowly rebind to its target site during the remainder of the annealing phase and the 5 minutes incubation at 30 °C before beginning the second melting curve. Thus the samples do not have sufficient time to re-equilibrate and these second melting curves will represent sites to which the ligand has bound more quickly. This is consistent with the footprinting studies at different time intervals in which only a few footprints are observed at short time intervals after incubation, whereas new protections appear at prolonged incubation.

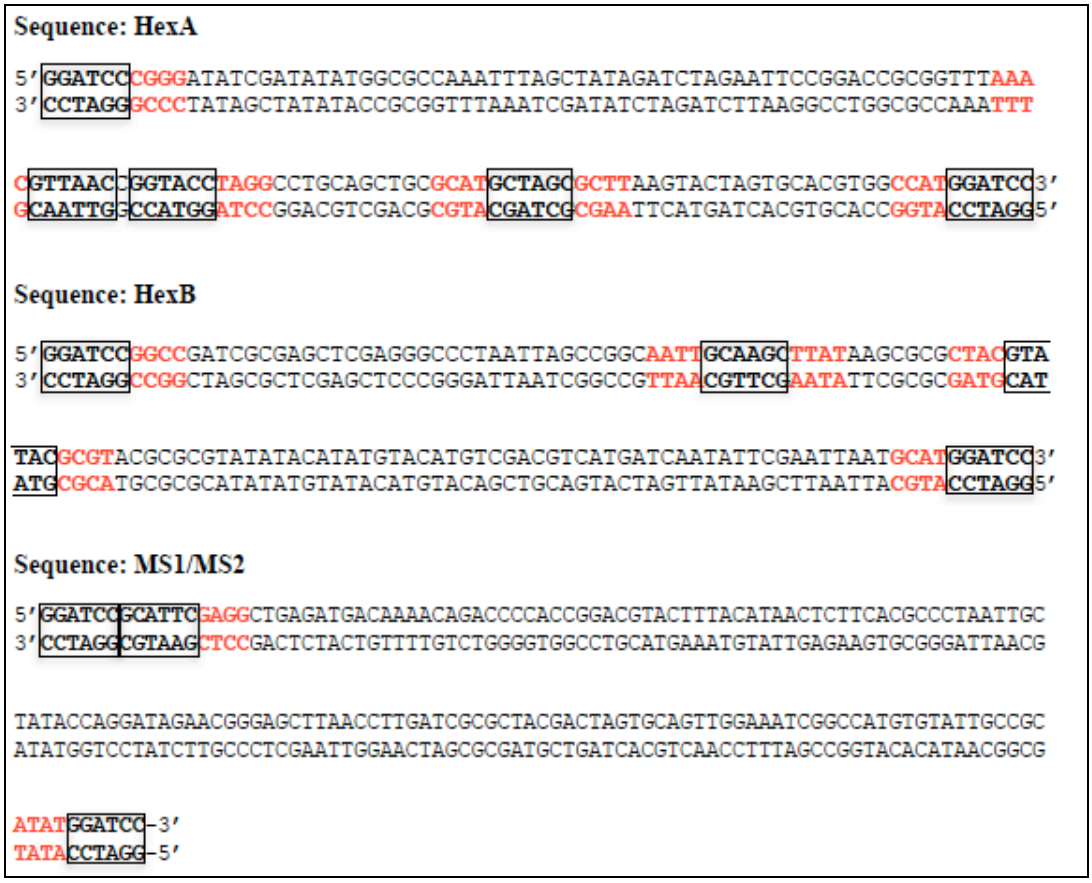


Figure 4.25: Sequences of HexA, HexB and MS1/MS2; the hexanucleotide duplexes used for melting study are shown in boxes with their flanking sequences (red).

The oligonucleotide duplexes designed for the melting study have identical flanking sequences; four AT base pairs ($A_4.T_4$) on both the 5'-3'-sides, whereas in the DNA footprinting fragments, these sites are flanked by several different sequences (Figure 4.25). Moreover, in the footprinting fragments these sites will be present along with a large number of other potential binding sites, which will compete for interaction with the

ligands. However, The affinity of the ligands for each oligonucleotide duplex is clearly different as they produced varying percentages of the various adducts in the first melting curve. Furthermore, differences in the presence of 1st and 2nd adducts in the second melting transition indicates differences in the rates of formation of these adducts. Furthermore, these hexanucleotide sequences are present in some of the footprints but are protected at different concentrations, which again emphasizes their varying affinities for these sequences. Changes in melting temperatures of different ligands also vary with these duplexes where each ligand produces 1st and 2nd adducts having different melting temperatures at highest ligand concentration (Table 4.21).

Table 4.21: Changes in melting temperatures (ΔT_m) in the first melting curves at highest (10 μM) ligand concentration (melting temperatures in each concentrations are given in Appendix 2, Table IX-XVIII).

	KMR31		**KMR32**		**KMR33**		**KMR173**		**KMR175**	
	Change in melting temperature (ΔT_m) (°C)									
	1st ΔT_m	2nd ΔT_m	1st ΔT_m	2nd ΔT_m	1st ΔT_m	2nd ΔT_m	1st ΔT_m	2nd ΔT_m	1st ΔT_m	2nd ΔT_m
GGTACC T_m = 41.4 °C	26.2	46.8	26.8	47.6	28.1	48.1	26.1	-	20.6	37.7
GGATCC T_m = 42.3 °C	27.2	48.1	26.2	47.8	29.2	49.7	23.1	45.1	21.7	38.4
GTATAC T_m = 36.5 °C	28.7	49.0	28.8	49.4	-	50.4	27.2	46.4	20.7	37.6
GTTAAC T_m = 37.8 °C	25.6	47.8	26	47.8	27.1	49.2	25.2	49.5	20.3	36.8
GCTAGC T_m = 43.9 °C	28.1	46.8	28	47.6	29.4	48.4	28.2	45.1	21.9	38.9
GAATGC T_m = 40.7 °C	32.3	50.7	32.8	50.4	32.4	50.5	32.7	48.2	26.7	44.2
GCTTGC T_m = 44.6 °C	27.7	46.8	28.7	46.5	28.2	47.5	31.8	45.9	21.9	38.3

From the footprint studies, it is seems that KMR31 (PBD-Py-Benzofuran) has highest affinity for GCTAGC (site 1/HexAfor) and GCTTGC (site 1/HexBrev), which are protected with 2 μM; GGATCC (site 4/HexBrev) and GAATGC (site 7/MS2) have intermediate affinity producing footprints with 5 μM while GTTAAC (HexAfor site 2) and GGTACC (site 2/HexAfor) are only bound at the highest concentration whereas GTATAC is not protected. In the melting studies, GGATCC has higher affinity than GGTACC as evident from the amount of the 1st and 2nd adducts and in terms of rate of reaction in the second melting transition, and GGATCC forms more 1st adduct than GGTACC. Oligonucleotide duplexes GTATAC and GTTAAC show similar affinities for the ligand, though in the second melting curve GTATAC forms a greater amount of 2nd adduct than GTTAAC, indicating that GTATAC has a higher rate of adduct formation than GTTAAC. As in the footprinting studies, GCTTGC and GCTAGC have greater

affinity for KMR31 than GAATGC; the rate of formation of adducts is also higher for GCTTGC and GCTAGC as evident from the transitions in the second melting curve.
KMR32 (PBD-Py-Benzothiophene) shows greater affinity for GGATCC than GGTACC, producing higher percentages of the adducts in the first melting transition. The rate of reaction is also higher at GGATCC, as seen from the second melting transition in which GGTACC shows both 1st adduct and 2nd adducts, while GGTACC only forms the 1st adduct. In the footprint experiments, both sequences- GGTACC (site 2/HexAfor) and GGATCC (site 1/HexBfor) are found in the footprints with KMR32. The ligand binds to GTTAAC and GTATAC with similar affinities though GTATAC is bound more rapidly than GTTAAC, as evidenced from the formation of both adducts in the second melting curve while GTTAAC (HexAfor site 2) is better protected than GTATAC (HexBrev site 3) in footprints. KMR32 binds well to both GCTAGC (site 1/HexAfor) and GAATGC (MS1 site 1) as both sequences produce footprints with 2 μM ligand, while GCTTGC (HexBrev site 2) is partly protected at this concentration. In the melting studies, KMR32 has highest affinity for GCTAGC at which is also shows a greater rate of reaction, while GCTTGC is of intermediate affinity and rate of adduct formation; GAATGC shows the lowest affinity and doesn't show any formation of the 2nd adduct.

In the footprinting experiments KMR33 (PBD-Py-Benzothiophene-Acetate) shows higher affinity for GGTACC (site 3/HexAfor) (protected with 5 μM) than GGATCC (site 1/HexAfor) (protected with 10 μM); this result is consistent with the melting studies, in which GGTACC shows slightly greater percentages of adducts in the first melting transition. GGATCC has higher rate of reaction as it forms greater amount of 1st and 2nd adducts in second melting transition than GGTACC. Both GTATAC and GTTAAC, are both fully modified in the first melting curve but they differ in the rate of reaction; GTATAC forms both 1st and 2nd adducts in the second melting curve while 2nd adduct is formed with GTTAAC and a significant amount remains in the unmodified state. Footprints are produced at GTTAAC (site 3/HexAfor) with 0.5 μM ligand, while GTATAC-containing sites (site 4/HexBfor) only produce footprints with the highest concentration. GAATGC forms greatest amount of 2nd adduct in the first melting curve among GC containing oligonucleotide; in terms of rate of reaction, GCTAGC shows fastest rate while GCTTGC has highest amount of unmodified duplex in the second melting curve.

The footprinting studies with KMR173 (PBD-Im-Benzothiophene-Acetate) show only a few binding sites at the highest concentrations. The melting studies also indicate that this ligand has a lower affinity for the hexanucleotide sequences. Both GGTACC (site 5/HexArev) and GGATCC (site 1/HexAfor) are protected only at highest concentration.

In the melting studies, GGTACC can only form the 1st adduct while GGATCC forms both 1st and 2nd adduct. However the second melting transitions are similar for these sequences, in which only the 1st adduct is evident. KMR173 shows higher affinity for GTATAC than GTTAAC, though both oligonucleotides are unable to produce any adducts in the second melting curve. Oligonucleotides GCTTGC and GCTAGC show equal rates of reaction with KMR173, though it shows greater affinity for GCTAGC, then GCTTGC and finally GAATGC. In the footprinting experiments, GCTAGC (site 2/HexAfor) and GAATGC (site 1/MS1) are protected with only highest concentration, while GCTTGC (site 2/HexBrev) is partially protected. KMR175 (PBD-Benzothiophene) produces 1st and 2nd adducts in the melting experiments though these sequences are not protected in the footprinting studies. The ligand shows the formation of 1st and 2nd adducts in the first melting transition with all the oligonucleotide sequences, but had very little or no adduct formation in the second melting transition.

4.6 Conclusion

From the footprinting studies, it can be suggested that these PBD-conjugates bind with AT-rich sites. Presence of pyrrole seems to provide higher affinity in binding as evident from KMR31, KMR32 and KMR33. Presence of imidazole (KMR173) lacks affinity compared to other pyrrole containing conjugates; moreover, it binds with some GC-rich squences. Either pyrrole or imidazole is essential in the polyamide chain as KMR173 (PBD-Benzothiophene) demonstrates least footprints. Benzofused rings also greatly affect the binding; benzothiophene and benzofuran shows binding while no binding is observed in PBD-conjugates containing indole. From the melting studies, it is also evident that these conjugates bind very strongly with 6-bp sequences producing large increase in melting temperatures; variations in melting temperatures and amount of adducts formation indicate that these PBD-conjugates differ slightly in terms of preferences for these sequences.

5 Development of a ligation assay for assessing the sequence selectivity of bis-PBDs

5.1 Introduction

Pyrrolobenzodiazepines (PBDs) are a family of DNA minor-groove binding antitumor antibiotics derived from various *Streptomyces* species (287). They exert their biological activity by alkylating the C2-amino group of guanine in the minor groove of DNA. The naturally occurring PBD monomers can form singly alkylated DNA adducts; the PBD monomer-DNA adduct spans 3 base pairs with the preferred order of 5'-Pu-G-Pu > 5'-Pu-G-Py or 5'-Py-G-Pu > 5'-Py-G-Py (Pu = purine and Py = pyrimidine) (103,120). Tethering two PBD units together would be expected to extend the size of recognition site, to twice that of a PBD monomer subunit. The first PBD-dimer (DSB120) reported in 1992 was composed of two PBD units tethered by a propyldioxy linker at the C8/C8' positions of their aromatic A-rings (168); the compound recognized 6 base pairs in a DNA duplex with a cross-linking preference for the sequence 5'-Pu-GATC-Py-3' (288,289). DSB120 showed potent *in vitro* cytotoxicity toward a panel of human cancer cell lines (168) but lacked significant *in vivo* antitumor activity (290). Dimers with longer linkers revealed another PBD-dimer with a pentyldioxy linker that favoured the extended sequence 5'-Pu-GAATC-Py-3', while linkers with four and six methylenes did not produce cross-links with any of the $GAT_{1\text{-}4}C$ combinations; molecular modelling of these homologous series demonstrated that PBD-dimers with odd-numbers of methylene linkers are energetically favourable for forming interstrand cross-linking adducts (291). Subtle modification of compound DSB120 with C2/C2' *exo*-unsaturation showed dramatic results having both *in vitro* (177,291-293) and *in vivo* (294) activity, forming sequence-selective interstrand cross-links at 5'-Pu-GATC-Py-3' (178); this modified PBD-dimer (SJG-136) is presently being evaluated in clinical trials (94,295). Details of different PBD-dimers have been discussed in the Introduction (Chapter 1, Section 1.2.4.6.4).

The aim of the work described in this chapter is to develop a simple assay for determining the sequence selectivity of PBD-dimers. As described in the Introduction, these compounds are comprised of two PBD units that are tethered by linkers of varying length. Since each PBD reacts with a different guanine the PBD-dimers should cross-link two guanines together in an intra- or inter-strand reaction and should therefore have a greater and extended selectivity than the parent mono-PBDs. SJG-136 is one of the best studied of these compounds and is thought to bind to the symmetrical sequence GATC, forming intermolecular cross-links between the guanines

on opposite strands that are separated by two base pairs (175,178) However, there is little rigorous experimental evidence for this selectivity or details of how the ligand might cross-link other G-containing sequences.

The basis of the assay being developed in this chapter is that the bis-PBDs should be able to chemically join (ligate) two DNA duplexes together if a good ligand binding site is generated when their sticky-ends anneal (non-covalently). For example, if we assume that a ligand binds to GATC, then duplexes that contain single stranded GATC overhangs will be ligated together between the two guanines on different molecules as illustrated in Figure 5.1. In theory this could generate a long ladder of chemical ligation products, in much the same way as would be achieved by using DNA ligase. If the ligand does not bind to GATC then clearly no cross-linking will occur. By comparing the efficiency of cross-linking of different overhangs (*i.e.* GATC, GTAC, CATG, CTAG, GAATC *etc.*) it should be possible to determine which sequence(s) the ligand binds to best.

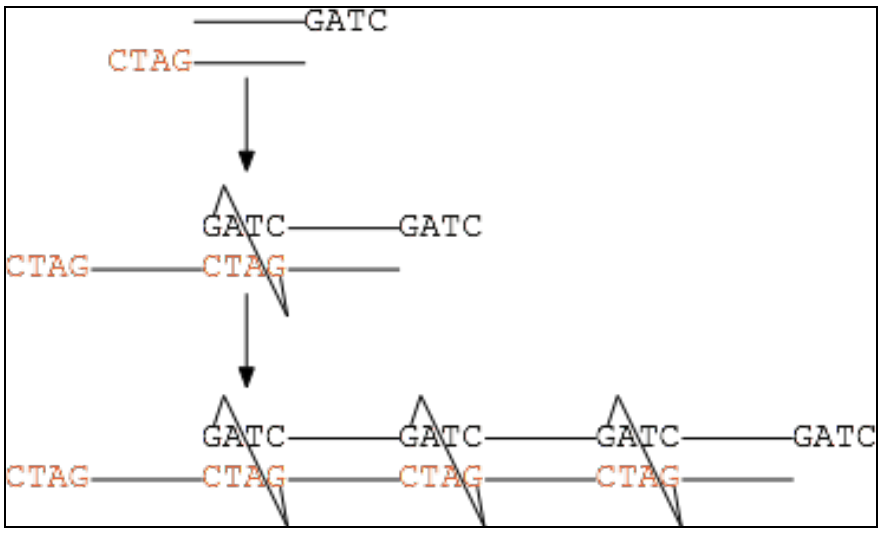


Figure 5.1: Illustration of cross-linking of DNA duplexes that possess a GATC overhang (top) to generate a dimer (middle) and tetramer (bottom). The ligand is illustrated by the staple, between the guanines on opposite strands.

5.2 Experimental design

The oligonucleotide sequences and preparation of samples were described in section 2.4.1 and 2.4.2.

5.3 Results

Preferred binding sites in oligonucleotides containing two different binding sequences

In the first ligation assay we examined the ability of SJG-136 to cross-link short self-complementary oligonucleotides containing two binding sites with different overhangs. Oligo-1 and oligo-2 contain two potential (GTAC) binding sites in the middle of the sequence as well as having GATC and GTAC in the overhangs respectively, while oligo-3 only contains a central GATC binding site. Each of these is a self-complementary sequence. Figure 5.2 shows the effect of varying SJG-136 concentration (25-500 μM) on the gel mobility of these radiolabelled self-complementary oligonucleotides. Oligo-1 and oligo-2 produce a single species with retarded mobility (a and b), which appears to be stronger in oligo-1 (GATC overhang). The lower dark band corresponds to the labelled oligonucleotide alone; although it is not clear whether this runs as a duplex or single strand, we would expect this 14-mer duplex to run as a duplex under these conditions. The retarded band could be either a stable duplex (if the other band is single stranded) or an end-to-end dimer. The absence of any higher concatamers suggests that it is unlikely to be a simple dimer. However, these self-complementary sequences can also form hairpin structures and the retarded band might then correspond to a dimeric dumbbell structure, as shown below. This would not be able to form any higher order structures, as it does not possess any overhangs. Whatever the nature of the bands, the retarded species appears to be stronger with the GATC overhang (oligo-1).

In contrast oligo-3 does not produce a highly retarded species. This oligo would only produce an 8-compementary duplex, which is unlikely to be stable under these conditions. Moreover this oligonucleotide will not be able to form any hairpin structures. Reaction with SJG-136 across the central GATC could stabilise this duplex. It is therefore possible to propose that the lower band, which appears in a concentration-dependent fashion, corresponds to the duplex.

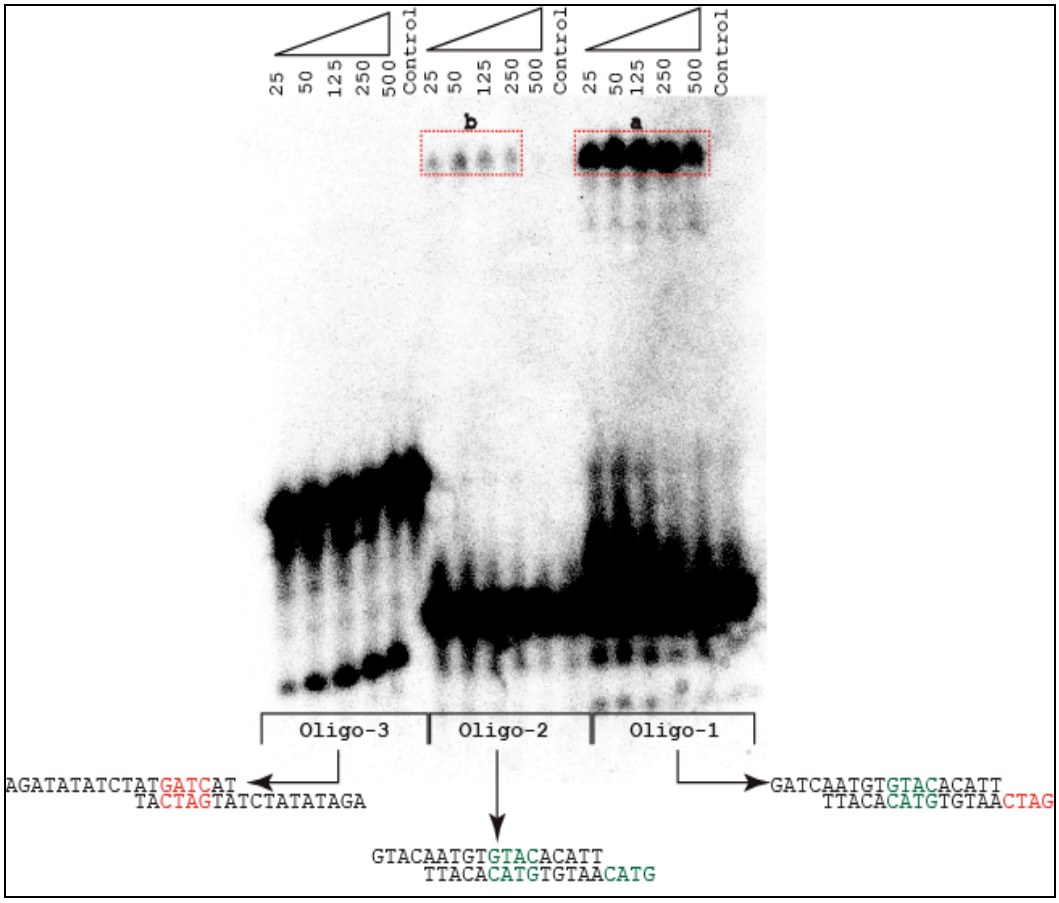


Figure 5.2: Autoradiograph of a non-denaturing polyacrylamide gel showing the interaction of 5'-^{32}P labelled self-complementary oligonucleotides with SJG-136. The samples were incubated with various concentrations of SJG-136 at room temperature for 24 hours before running on the gel. Ligand concentrations (μM) are shown at the top of each gel lane. "Control" indicated samples without added ligand. "a" and "b" with dashed red box correspond to ligated products.

Since there is little concentration-dependent change in the intensity of the retarded species for oligos-1 and -2, a comparatively low concentration (50 μM) was chosen for further experiments. In later experiments $MgCl_2$ (10 mM) was included, along with NaCl (100 mM) in order to increase the duplex stability, as reported by Owczarzy *et al* (296); at the same time, EDTA is omitted from the buffer due to its ability to chelate magnesium.

Cross-link product formation in oligonucleotides with binding sites in the middle of the sequence

Figure 5.3 shows the effect of SJG-136 (50 μM) on the mobility of oligonucleotides that have different binding sites in the middle of the sequences. In these experiments the oligos are not self-complementary and so are used in pairs (4+5; 6+7; 8+9). In each case the "control" lane has the same mobility as the single stranded oligos, suggesting that, in the absence of the ligand these mini-duplexes are not stable under these

conditions. With oligos-4/5 (with a central GATC) a retarded band is evident in the presence of SJG-136. Surprisingly a band of faster mobility is also observed below the control lane. With oligos-6/7 (with a central GTAC) and oligos-8/9 (with GAATC), no retarded products are observed. Since each of the oligonucleotide pairs has a different mobility the cross-linking was also examined by mixing the oligonucleotides. Two pairs of complementary oligonucleotides (oligos-4/5 and 6/7, oligos-6/7 and 8/9, oligo-4/5 and 8/9) and another mixture with all the oligonucleotides. In these mixtures the only ones producing the retarded species are those containing oligos-4/5 with a central GATC, and these also show the faster moving species. The mixture of GTAC and GAATC does not show any retarded products whereas the GAATC/GATC mixture shows the slow and fast moving species. When all the oligonucleotides are mixed together, SJG-136 produces bands with altered mobility, which are the same as for GATC alone.

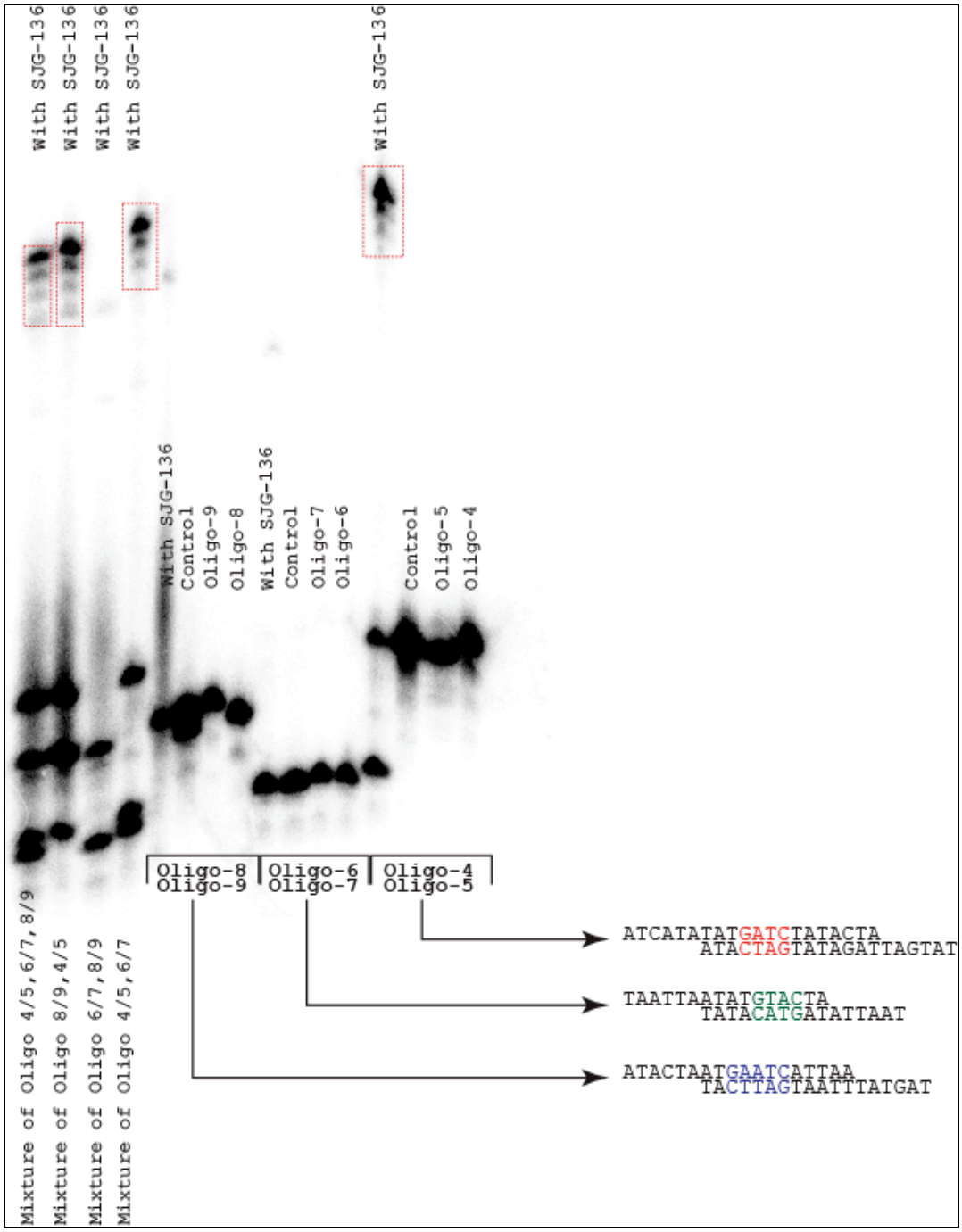


Figure 5.3: Autoradiograph of a non-denaturing polyacrylamide gel showing reaction of 5'-end ^{32}P labelled complementary oligonucleotides with SJG-136 (50 μM). The oligonucleotides contain binding sites in the middle (shown in colours) of the sequences. The samples were incubated with SJG-136 at room temperature for 24 hours before running into gel. Oligo-4, oligo-5, oligo-6, oligo-7, oligo-8 and oligo-9 indicate lanes running only radiolabelled single-stranded oligonucleotide. "Control" indicates mixtures of the two strands without added ligand. Dashed red boxes correspond to ligated products.

Ligation products stabilized by SJG-136 in both Figure 5.2 and Figure 5.3 show only one retarded band along with few faint bands. Here, most of the oligonucleotides are found in monomeric form as they travel the same distance as the control. In these

experiments the concentration of radiolabelled oligonucleotides is low (probably in nanomolar range). Thus the band of ligated products observed in the gel occurs when a fraction of these radiolabelled oligonucleotides are annealed, while most of the oligonucleotides remain in a monomeric form. The probability of ligation might therefore be increased by increasing the overall oligonucleotide concentration. As a result in the following experiments unlabelled oligonucleotides (2.5 μM) were added to the reaction mixtures.

Increased cross-linking in presence of unlabelled oligonucleotides

Figure 5.4 shows the effect of unlabelled oligonucleotides (2.5 μM) on the reaction. In this experiment, two controls were run for each pair of oligonucleotides by annealing one radiolabelled oligonucleotide with its complementary unlabelled oligonucleotide; for example, radiolabelled oligo-4 with unlabelled oligo-5 and *vice versa*. The samples with ligands were prepared in the same way, by adding ligand to a mixture of radiolabelled and unlabelled oligonucleotides. Radiolabelled oligonucleotides were also run individually without mixing with their complementary strands. It can be seen that oligo-4 and oligo-5 (with GATC) produce a retarded band with a few weaker bands beneath this (“a”). Interestingly, in the presence of the ligand no band is observed in the same position as the control, suggesting that all radiolabelled oligonucleotides have reacted with the ligand. The faster species observed in Figure 5.3 with this pair of oligos is also not evident. Oligo-6 and oligo-7 (with a central GTAC) also show a retarded product (“b”), and none of the unreacted species are evident. With GAATC (oligo-8 and oligo-9), no retarded species is observed, though the ligand-treated bands appear to have a slightly lower mobility, which is more smeared.

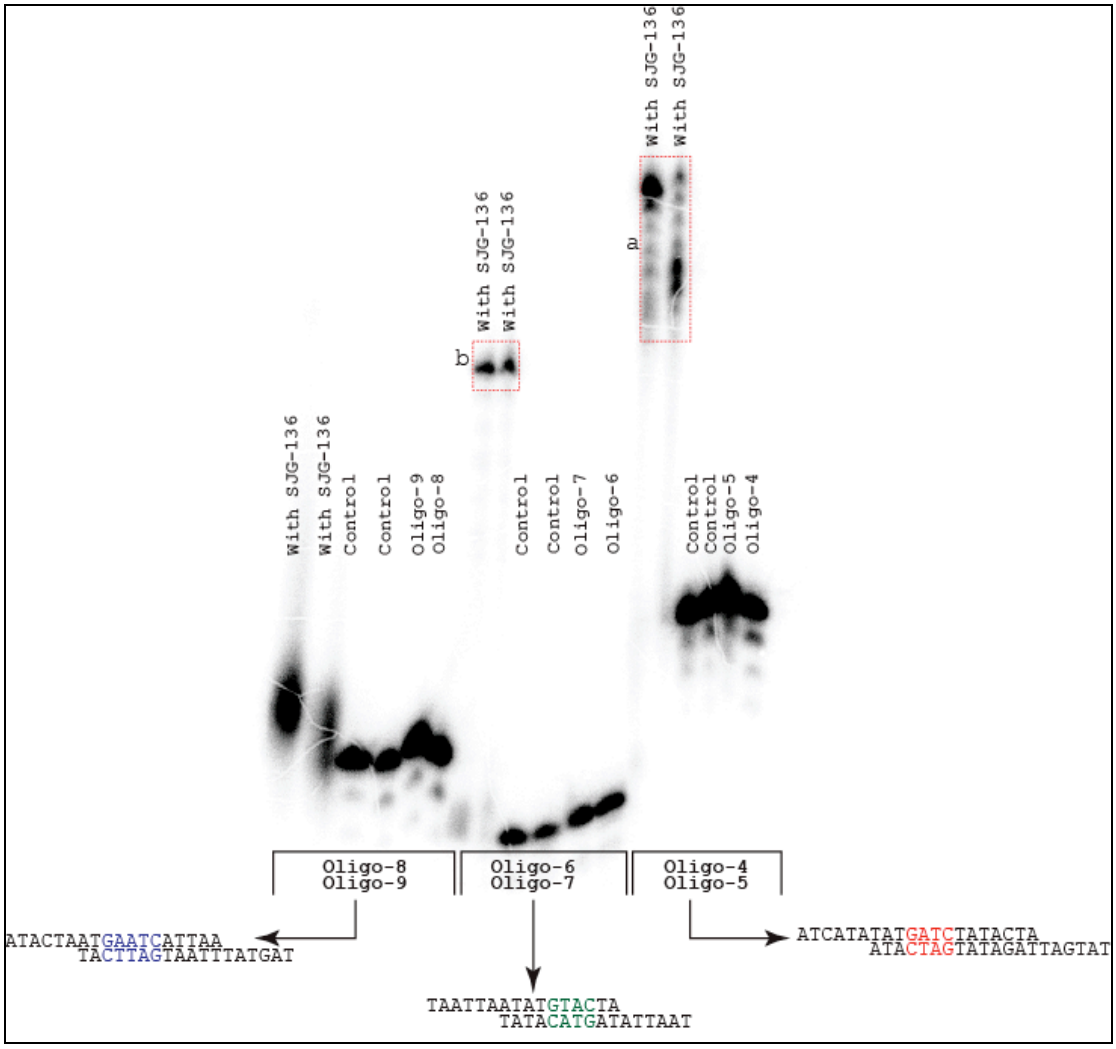


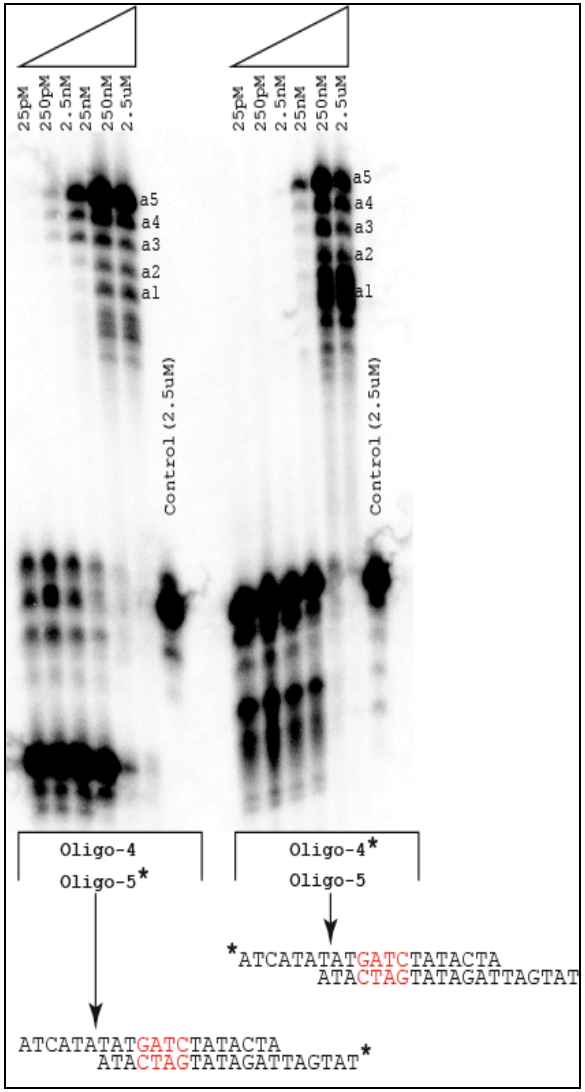
Figure 5.4: Autoradiograph of a non-denaturing polyacrylamide gel showing cross-linking of 5'-end ^{32}P labelled complementary oligonucleotides by SJG-136 (50 μM). The oligonucleotides contain binding sites in the middle (shown in colours) of the sequences. The samples were incubated with SJG-136 at room temperature for 24 hours before running into gel. Oligo-4, oligo-5, oligo-6, oligo-7, oligo-8 and oligo-9 indicate lanes running only radiolabelled single-stranded oligonucleotide. "Control" indicates samples without added ligand. "a" and "b" with dashed red boxes correspond to products with altered mobility. For each pair of oligonucleotides, upper strands correspond to oligo-4, oligo-6 and oligo-8 whereas lower strands correspond to their complementary sequence.

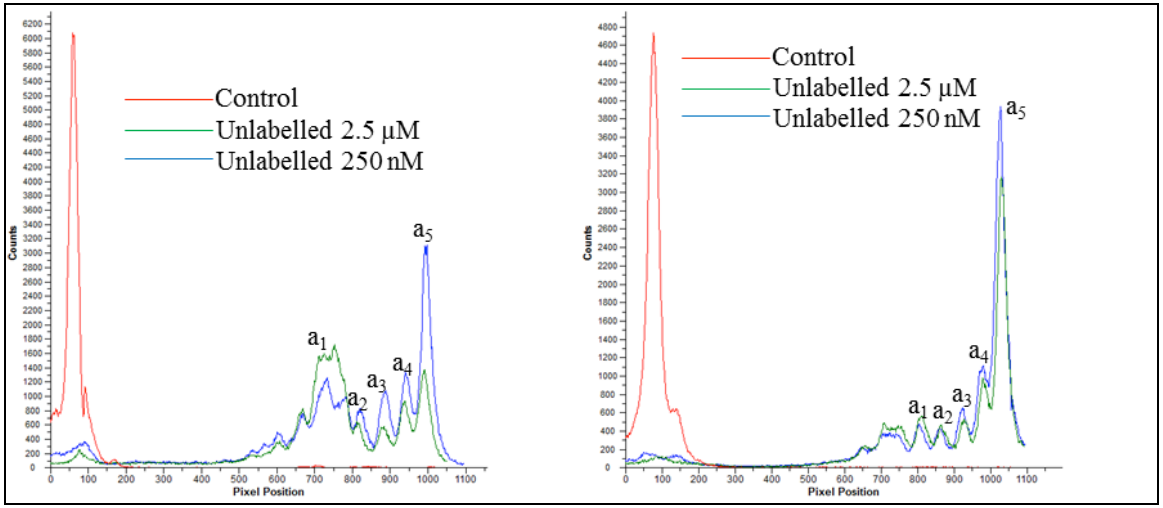
Effect of concentration of unlabelled oligonucleotides on cross-linking

Since increasing the total oligonucleotide concentration by the addition of unlabelled oligonucleotides (2.5 μM) altered the cross-linking efficiency, further experiments were performed with varying concentrations of unlabelled oligonucleotides. These experiments, used SJG-136 (50 μM) with 100 mM NaCl and 10 mM $MgCl_2$. Figure 5.5 shows the effect of varying concentrations of unlabelled oligonucleotides (2.5 μM to 25 pM) on the reaction with oligos-4/5 having GATC in the middle. The experiment was done in duplicate by mixing unlabelled oligo-5 (25 pM to 2.5 μM) with radiolabelled

oligo-4 and *vice versa*. Control samples were also prepared in the same way. With unlabelled oligo-5, several bands of reduced mobility are observed in the presence of higher concentrations of the oligonucleotide (2.5 μM and 250 nM), and in these lanes there is no unreacted oligonucleotide with the same mobility as the control sample. The patterns were analysed by ImageQuant TL software from which peaks of different height are identified; the peaks (a1, a2, a3, a4 and a5 in Figure 5.5) indicate ligated products and the height of the peak corresponds to the intensity of the band of ligated products. Samples with 25 nM and lower oligonucleotide concentrations show a much less of this retarded species, but bands with a faster mobility than the control are again evident. Repeating the experiment by mixing unlabelled oligo-4 (2.5 μM t0 25 pM) with radiolabelled oligo-5 produces a similar pattern of bands. It therefore appears that reaction of SJG-136 with high oligonucleotide concentrations produces a pattern of retarded bands, while low oligonucleotide concentrations generate a species with faster mobility.

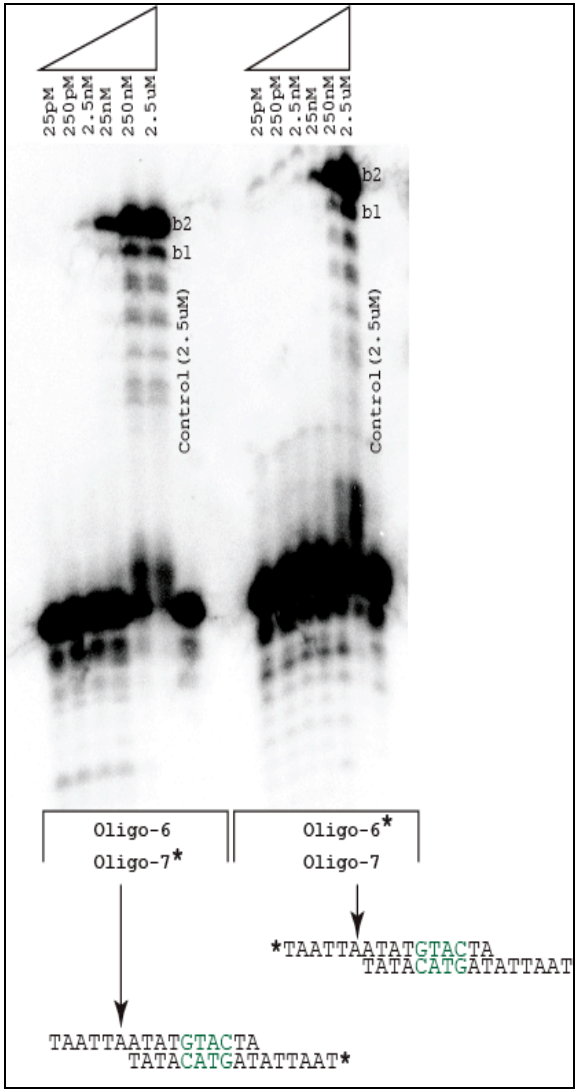
Figure 5.6 shows similar experiments on the effect of varying concentrations of unlabelled oligonucleotides (2.5 μM to 25 pM) on the reaction with oligos-6/7, which possesses GTAC in the middle of the sequences (oligo-6 and oligo-7). This experiment was also done in duplicate and the samples were prepared by adding unlabelled oligo-7 (2.5 μM t0 25 pM) with radiolabelled oligo-6 and *vice versa*. A strong retarded band is evident with both combinations with a series of much weaker bands beneath this. These bands are identified as peaks (b1 and b2). In each case some unreacted material is evident with the same mobility as the control and this becomes stronger at lower oligonucleotide concentrations, so that no additional bands are produced with 2.5 nM oligo and below. In contrast to the results with oligos-4/5 there are no species with faster mobility than the control.

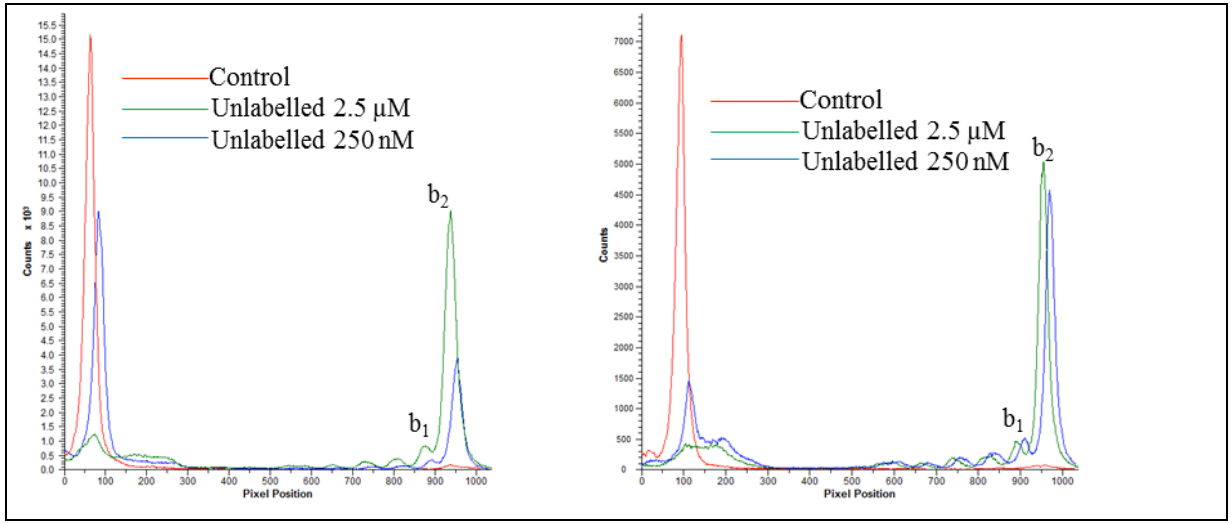




(I) (II)

Figure 5.5: ***In the left panel***, autoradiograph of a non-denaturing polyacrylamide gel showing reaction of 5'-end ^{32}P labelled complementary oligonucleotides with SJG-136 (50 µM). The oligonucleotides contain binding sites in the middle (shown in colours) of the sequences. The concentration of unlabelled oligonucleotides is shown at the top of each gel lane. The samples were incubated with SJG-136 at room temperature for 24 hours before running into gel. "Control" indicated samples without added ligand. "a1", "a2", "a3" "a4" and "a5" indicate ligated products. Oligo-4 and oligo-5 correspond to upper and lower strand respectively. Radiolabelled strand is shown as asterisk (*). ***In the top panel***, Densitometric analysis of GATC containing oligonucleotides (oligo-4 and oligo-5) - (I) Radiolabelled oligo-4 with two different concentrations (2.5 µM and 250 nM) of unlabelled oligo-5. (II) Radiolabelled oligo-5 with two different concentrations (2.5 µM and 250 nM) of unlabelled oligo-4. The peaks are obtained by analysing the image in left panel with ImageQuant TL. "a1", "a2", "a3", "a4" and "a5" indicate peaks for ligated products.





(I) (II)

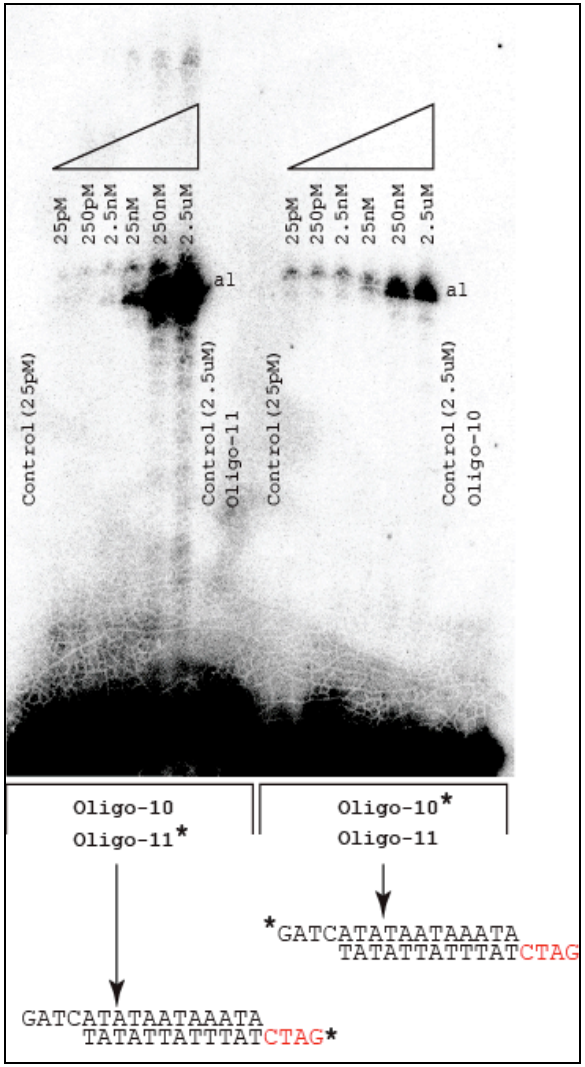
Figure 5.6: ***In the left panel***, autoradiograph of a non-denaturing polyacrylamide gel showing reaction of 5'-end ^{32}P labelled complementary oligonucleotides with SJG-136. The oligonucleotides contain binding sites in the middle (shown in colours) of the sequences. The concentration of unlabelled oligonucleotides is shown at the top of each gel lane. The samples were incubated with SJG-136 at room temperature for 24 hours before running into the gel. Ligand concentrations (µM) are shown at the top of each gel lane. "Control" indicates samples without added ligand. "b1" and "b2" indicate ligated products. Oligo-6 and oligo-7 correspond to upper and lower strand respectively. The radiolabelled strand is indicated by an asterisk (*). ***In the top panel***, Densitometric analysis of GTAC containing oligonucleotides (oligo-6 and oligo-7) - (I) Radiolabelled oligo-6 with two different concentrations (2.5 µM and 250 nM) of unlabelled oligo-7 (II) Radiolabelled oligo-7 with two different concentrations (2.5 µM and 250 nM) of unlabelled oligo-6. The peaks are obtained by analysing the image in left panel with ImageQuant TL. "b1" and "b2" indicate peaks for ligated products.

Ligation of oligonucleotides containing binding sites in the overhangs in presence of unlabelled oligonucleotides

The success of detecting reaction products with oligonucleotides having GATC or GTAC in the middle of the sequences with addition of unlabelled complementary oligonucleotides was extended by examining oligonucleotides only having potential binding sites in the self-complementary overhangs. Oligo-10 is complementary to oligo-11 and these produce a GATC overhang whereas oligo-12 and oligo-13 produce a GTAC site. These oligonucleotides were examined with varying concentrations of unlabelled oligonucleotides (2.5 μM to 25 pM) and 50 μM SJG-136 in the presence of NaCl (100 mM) and $MgCl_2$ (10 mM). In the previous two gels, samples with lower concentrations of unlabelled oligonucleotides do not produce any ligated products; for confirmation, two control samples were prepared for the following experiments - radiolabelled oligonucleotides with i) 2.5 μM unlabelled oligonucleotides and ii) 25 pM unlabelled oligonucleotides assuming that both control samples as well as ligand-treated samples with lower concentrations of unlabelled oligonucleotides would travel same distance.

Figure 5.7 shows formation of ligated products with oligonucleotides having GATC (oligo-10 and oligo-11) in the overhangs with different concentrations of unlabelled oligonucleotides (2.5 μM to 25 pM). The samples were prepared by mixing radiolabelled oligo-10 with different concentrations of unlabelled oligo-11 and *vice versa*. In these experiments a single retarded species is evident at oligonucleotide concentrations of 25 nM and above; though this is very faint at low concentrations, though in all cases there is a large amount of unretarded control DNA. Analysing the gel with ImageQuant software, only one peak of ligated product is observed (a1 in figure 5.7). No peak is observed at lower concentrations of unlabelled oligonucleotides (not shown in figure 5.7).

Figure 5.8 shows the results of similar experiments with oligonucleotides having GTAC (oligo-12 and oligo-13) overhangs with 50 μM SJG136 and different concentrations of the unlabelled oligonucleotides (2.5 μM to 25 pM). The samples were prepared by mixing radiolabelled oligo-12 with different concentrations of unlabelled oligo-13 and *vice versa*. Both mixtures produce a single retarded band, shown as “b1” peak in Figures 5.8, in the presence of 250 nM oligonucleotide and above, though again a very large proportion of the sample is unaffected. Lower concentrations of unlabelled oligonucleotides (25 nM to 25 pM) show a very faint band, which is not quantifiable (not shown in Figure 5.8).



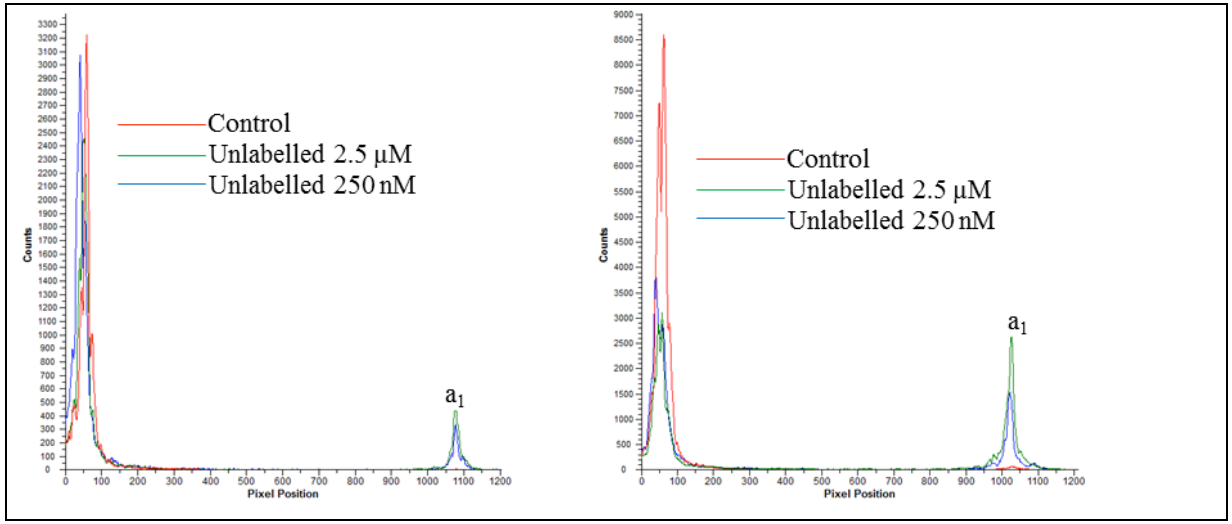
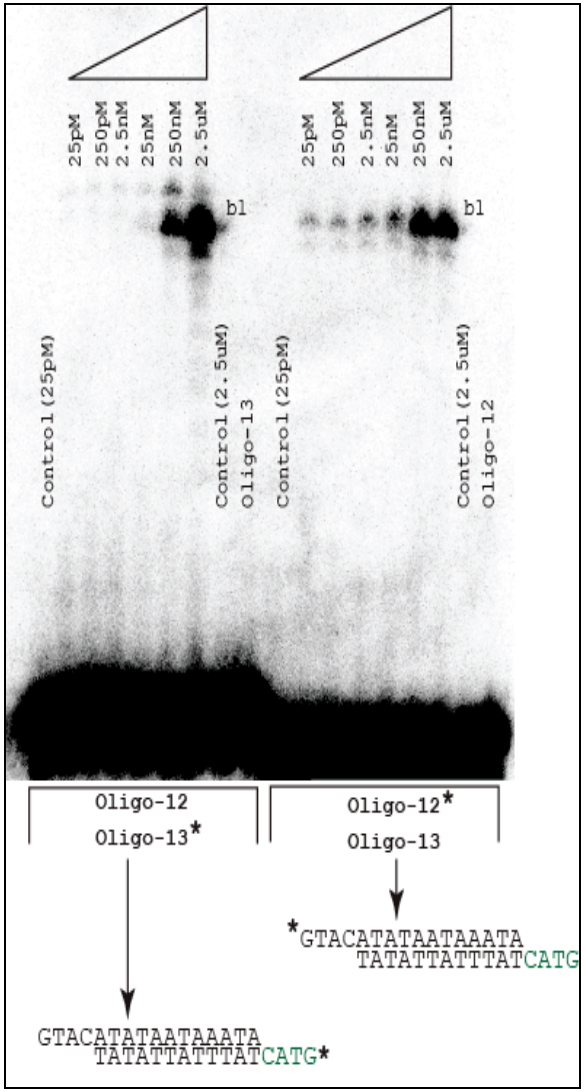
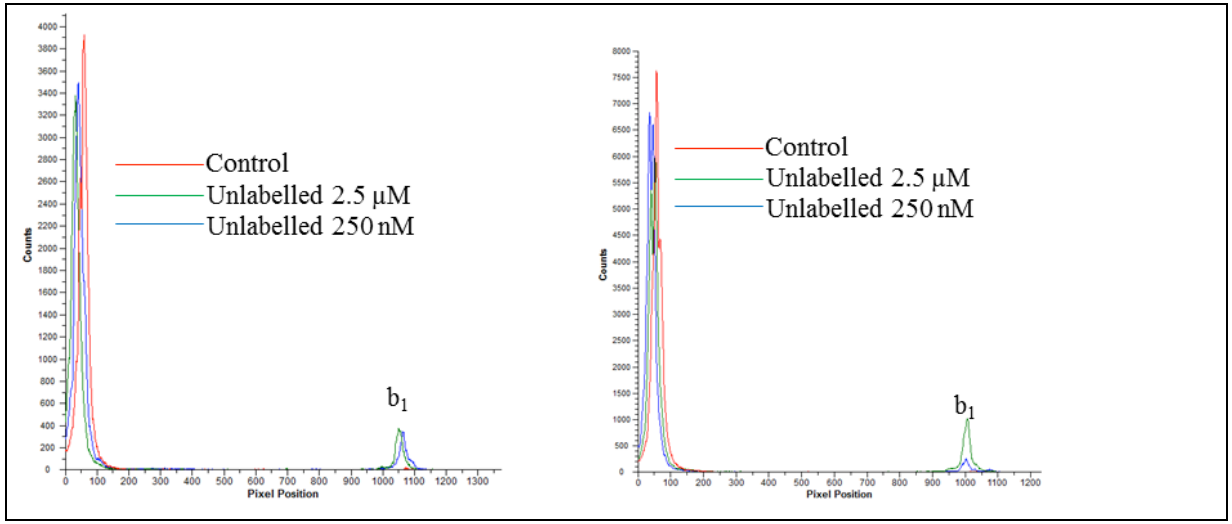


Figure 5.7: ***In the left panel,*** autoradiograph of a non-denaturing polyacrylamide gel showing reaction of 5'-end ^{32}P labelled complementary oligonucleotides with SJG-136 (50 µM). The oligonucleotides contain potential GATC ligand binding sites in the 5'-overhangs (shown in colours). The concentration of unlabelled oligonucleotides is shown at the top of each gel lane. The samples were incubated with SJG-136 at room temperature for 24 hours before running into gel. "Control" indicates samples without added ligand with concentration of unlabelled oligonucleotide. "a1" indicates ligated products. Oligo-10 and oligo-11 correspond to upper and lower strand respectively. The radiolabelled strand is indicated by an asterisk (*). ***In the top panel,*** densitometric analysis of GATC-containing oligonucleotides (oligo-10 and oligo-11) - (I) radiolabelled oligo-10 with two different concentrations (2.5 µM and 250 nM) of unlabelled oligo-11 (II) radiolabelled oligo-11 with two different concentrations (2.5 µM and 250 nM) of unlabelled oligo-10. The peaks are obtained by analysing the image in left panel with ImageQuant TL. "a1" indicates peak for ligated product.





(I) (II)

Figure 5.8: ***In the left panel,*** autoradiograph of a non-denaturing polyacrylamide gel showing reaction of 5'-end ^{32}P labelled complementary oligonucleotides with SJG-136 (50 µM).The oligonucleotides contain potential GTAC binding sites in the 5'-overhang (shown in colours). The concentration of unlabelled oligonucleotides is shown at the top of each gel lane. The samples were incubated with SJG-136 at room temperature for 24 hours before running into the gel. "Control" indicates samples without added ligand with concentration of unlabelled oligonucleotide. "b1" indicates ligated products. Oligo-12 and oligo-13 correspond to upper and lower strand respectively. The radiolabelled strand is shown by an asterisk (*). ***In the top panel,*** densitometric analysis of GTAC containing oligonucleotides (oligo-12 and oligo-13) - (I) radiolabelled oligo-12 with two different concentrations (2.5 µM and 250 nM) of unlabelled oligo-13 (II) radiolabelled oligo-11 with two different concentrations (2.5 µM and 250 nM) of unlabelled oligo-10. The peaks are obtained by analysing the image in left panel with ImageQuant TL. "b1" indicates peak for ligated product.

5.4 Discussion

The cross-linking of SJG-136 is thought to result from its recognition of 5'GXXC3' (178) and the ligand forms covalent bonds with guanines generating interstrand cross-linked complexes. In our ligation assay, complementary oligonucleotides anneal to form duplexes, which are then stabilized by interstrand cross-linking by SJG-136 at its binding sites. Formation of intense bands of ligated product with oligo-1, which possess both GATC and GTAC, compared with the less intense bands in oligo-2 having two GTAC sites, suggests that the ligand has a higher affinity for GATC than GTAC. Oligo-3 is also expected to generate ligated products like oligo-1 and oligo-2, but due to its long 10 bp overhang, and the presence of only eight base pairs in the complementary region, the probability of annealing in exact complementary bases is reduced.

With the oligonucleotides having potential binding sites in the middle, SJG-136 shows a strong effect on the one containing GATC. Ligated products of different mobility are achieved upon addition of unlabelled oligonucleotides. At higher concentrations of unlabelled oligonucleotides, GATC-containing sequences produced multiple bands indicating that the duplex DNA is firmly stabilized when cross-linked at this site; on the other hand, GTAC-containing sequences produce fewer peaks confirming that SJG-136 has greater affinity for GATC than GTAC. The presence of high concentration of oligonucleotides will drive duplex formation, providing enough duplex target sites which are then stabilized by cross-linking with the ligand; the annealed DNA would join at its overhang to form dimer. The process is expected to continue, generating a ladder of ligated DNA molecules. However this is not observed; the probability of joining the overhangs will gradually decrease with progression of long ligated products. No product was observed at lower concentration of unlabelled oligonucleotides, suggesting that a pool of oligonucleotides must be present to allow efficient ligation in order to produce a visible band in the gel.

Oligonucleotides offering binding sites in their overhangs (oligo-10/11 and oligo-12/13) produce a single ligated product in the presence of the ligand, which indicates binding of SJG-136 to both GATC and GTAC sites. These oligonucleotides were also expected to produce multiple bands like previous ones, but due to short length of the overhang, only 4 bp, and low melting temperatures (27.8 °C), a ligation ladder is not observed.

This preference for interstrand cross-linking of GATC is consistent with molecular modelling studies reported on compound DSB120 as well as SJG-136; the two compounds have an identical chemical structure except for exocyclic unsaturation (at the C2 position) of SJG-136. Molecular modelling of (5'-CIC^{4}**G**ATCCICG-3')/(5'-

CGC[14]**G**ATCGCG-3') with DSB120 demonstrated that the distance between the two covalent linkage sites (guanines shown in bold) in the cross-linked adduct was 12.5 Å, that corresponds favourably with an interstrand distance (11.5 Å) between the NH_2 groups of the two guanines; thus the linker (propyldioxy) provides optimum separation for interstrand cross-linking of a 4-bp bonding site (173). Modelling of the self-complementary oligonucleotide 5'-TATAGATCTATA-3' with SJG-136 also showed perfect positioning of the ligand in the central GATC sequence without any distortion of the DNA structure (175,183). GTAC might be expected to offer same environment; however, the adenine-N3 on the 3'-side of guanine forms a hydrogen bond with N10 of the PBD unit that ultimately makes GATC (-73.1 kcal/mol) energetically more favourable over GTAC (-59.6 kcal/mol). In case of GAATC, covalent bond formation of one PBD unit of SJG-136 with one guanine (underlined) causes significant distortion of the DNA structure; this leads to a loss of base pairing and a downstream mismatch that makes second guanine (complementary to cytosine) unfavourable for interstrand cross-linking (175).

Besides the experimental conditions, such as the effect of heat in the electrophoresis that can cause fraying of the cross-linked adducts, formation of secondary structures in oligonucleotides will also be important. Oligonucleotides (Oligo-4, 6 and 7) can self-anneal to form intermolecular duplexes (Figure 5.9) or intramolecular hairpins (Oligo-4, 5 and 7) (Figure 5.10) can also formed. Formation of such structures does not favour cross-linking by SJG-136 and subsequent elongation to form long DNA ladder.

(I)

```
          5'ATCATATATGATCTATACTA3'
3'ATCATATCTAGTATATACTA5'
```

(II)

```
           5'TAATTAATATGTACTA3'
3'ATCATGTATAATTAAT5'
```

(III)

```
           5'TAATTATAGTACATAT3'
3'TATACATGATATTAAT5'
```

Figure 5.9: Self-annealing of oligonucleotides; (I) oligo-4, (II) oligo-6 and (III) oligo-7

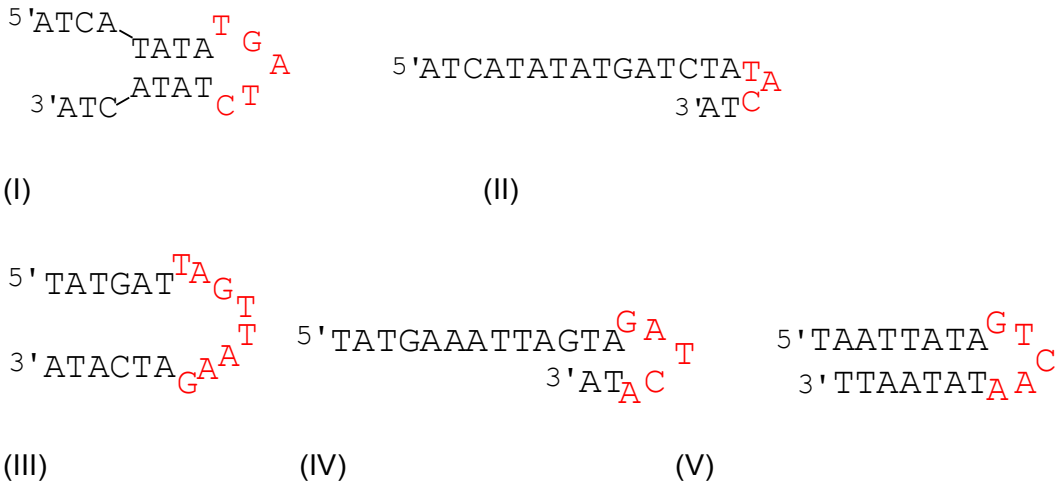


Figure 5.10: Hairpin formation by oligonucleotides; (I) and (II) for oligo-4, (III) and (IV) for oligo-5 and (V) for oligo-7

5.5 Conclusion

The ligation assay in this chapter clearly showed that PBD-dimer (SJG-136) consisting of two PBD units with a propyldioxy linkage has the highest preference for GATC for forming cross-linked DNA adducts. The ligand also binds (though less well) with GTAC, but does not produce any cross-links with GAATC.

6 Hairpin polyamides

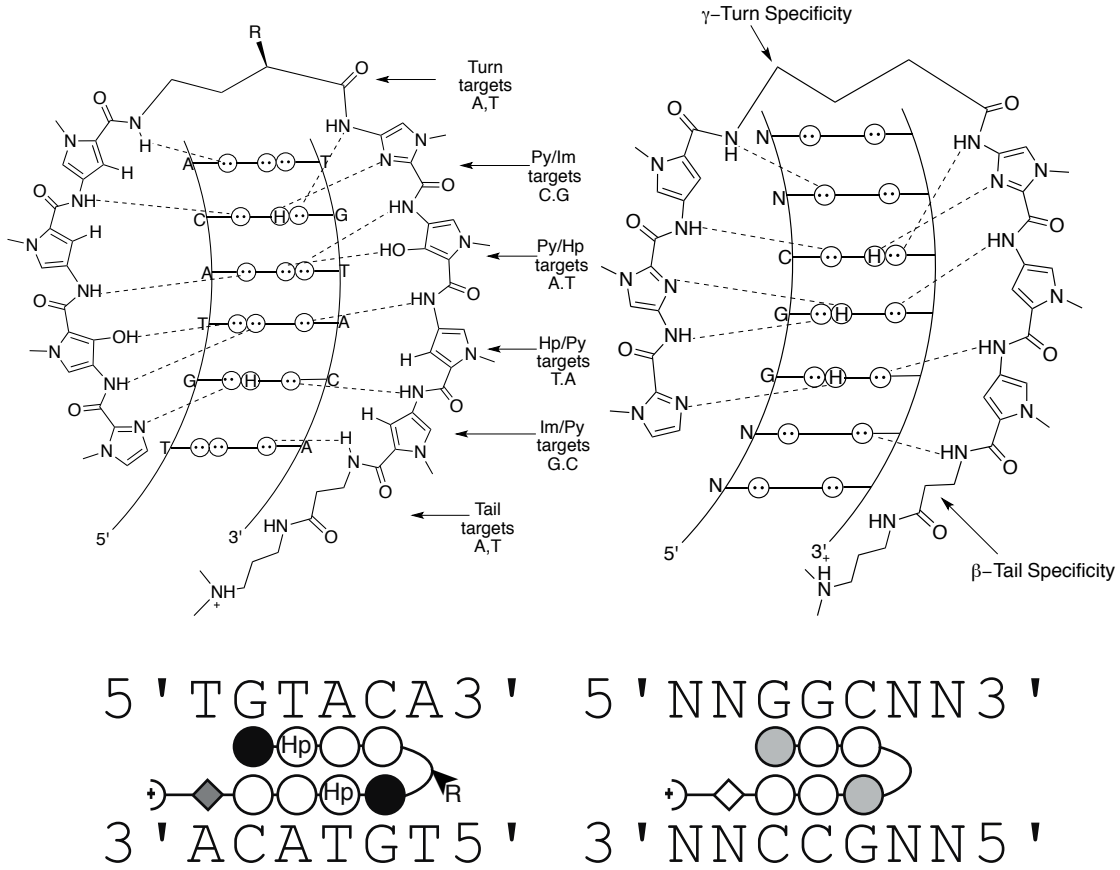
6.1 Introduction

Distamycin A and netropsin serve as the foundation for the development of synthetic polyamide molecules, made up of functional groups that are capable for reading individual Watson-Crick base pairs (Details in Chapter 1, Section 1.2.4.6.3). Polyamides containing pyrrole and imidazole amino acids bind as dimers in an antiparallel orientation in the DNA minor groove. This side-by-side pairing of *N*-methylpyrrole and *N*-methylimidazole results in sequence-specific binding; pairing of imidazole (Im) opposite pyrrole (Py) targets GC base pairs, while the pyrrole-imidazole combination binds to CG; a pyrrole/pyrrole arrangement can bind both A.T and T.A base pairs (151,153). Discrimination between AT and TA can be achieved by introducing hydroxypyrrole (Hp), where hydroxypyrrole/pyrrole and pyrrole/hydroxypyrrole pairs bind specifically to TA and AT respectively (154). β-alanine has been very effective when polyamide chains have more than five consecutive pyrrole or imidazole to adjust the pitch between amide bonds of Py-Im polyamides. Antiparallel pairing of Py/β and β/Py specify either AT or TA while β/Im and Im/β target GC and CG respectively (157,158).

For some DNA sequences, a single polyamide chain is sufficient to bind at a specific site. For example 5'-TGACA-3' can be targeted with ImPyPy, which binds as a homodimer. 5'-TGTTA-3' can be targeted by a mixture of ImPyPy and PyPyPy (distamycin), which binds as an antiparallel heterodimer. However these two polyamides will also simultaneously bind as homodimers, with comparable affinities, at sites 5'-$(A,T)_5$-3' (by distamycin) and 5'-TGACA-3' (by ImPyPy) (297,298). For such cases where different polyamides contact each strand in the minor groove, it was necessary to design motifs that can link the side-by-side polyamides in order to favour heterodimeric binding. The linking of ImPyPy and PyPyPy was attempted with amino acids of varying length - glycine, β-alanine, γ-aminobutyric acid and 5-aminovaleric acid. DNase footprinting experiments showed that γ-aminobutyric acid, known as a 'turn unit', produced the highest affinity and sequence selectivity, suggesting that the six ring polyamide adopted a hairpin structure (299). The hairpin motif bound with at least 2 orders of magnitude greater affinity over the its uncoupled components. Similar results have been reported with an eight ring polyamide, which produced >3600-fold enhancement in affinity for its target site over the unlinked components (300). As well as increasing the binding affinity, the 'turn unit' enforces unambiguous ring-pairing and

so eliminates any slipped binding modes that result from incomplete overlap of the polyamide chains. The uncoupled polyamides contain positive charges at their tails in the antiparallel homodimeric form, whereas the hairpin motif contains only one charge, thereby creating energetically more favourable conditions for binding (299). Incorporation of a chiral amine at the α-position of this turn unit further increases affinity (301).

For a polyamide containing *n* consecutive ring pairings, the binding site size will be *n+2*. Each side-by-side ring pairing recognises a central base pair and the outer base pairs are recognised by the terminal amides. Thus, a six-ring-hairpin polyamide ImImPy-γ-ImPyPy-β-Dp will bind to a 5-bp target 5'-NGGCN-3' (N = A, T, G, C); the 'GGC' is the core sequence recognised by three corresponding Im/Py, Im/Py and Py/Im ring pairings. To investigate the sequence specificity of γ-turn and β-tail amino acids, MPE.Fe(II) footprinting and affinity cleaving experiments (302) showed that adenine and thymine at the 3'-end (underlined) of the sequence 5'-TGGCA-3' and 5'-TGGCT-3' were equally favoured by the γ-turn. The sequence with a CG base pair at 5'-TGGCC had 28-fold reduced affinity, while the sequence with a GC base pair in 5'-TGGCG was bound with >400-fold reduced affinity. The preference of the aliphatic turn unit for AT base pairs presumably arises from hydrogen bond formation to the N3 of adenine or O2 of thymine; an unfavourable steric clash with the exocyclic NH_2 of GC base pairs causes the reduced binding affinity for this base pair. (303). Similar binding preferences were also observed for the β-tail binding at A.T or T.A (underlined); 5'-AGGCT-3' and 5-TGGCT-3' showed >300-fold higher affinity than 5'-GGGCT and 5'-CGGCT (302). Thus, hairpin polyamides can be very useful for recognition of specific sequences following the same pairing rules having significantly higher affinity than their corresponding components (Figure 6.1).

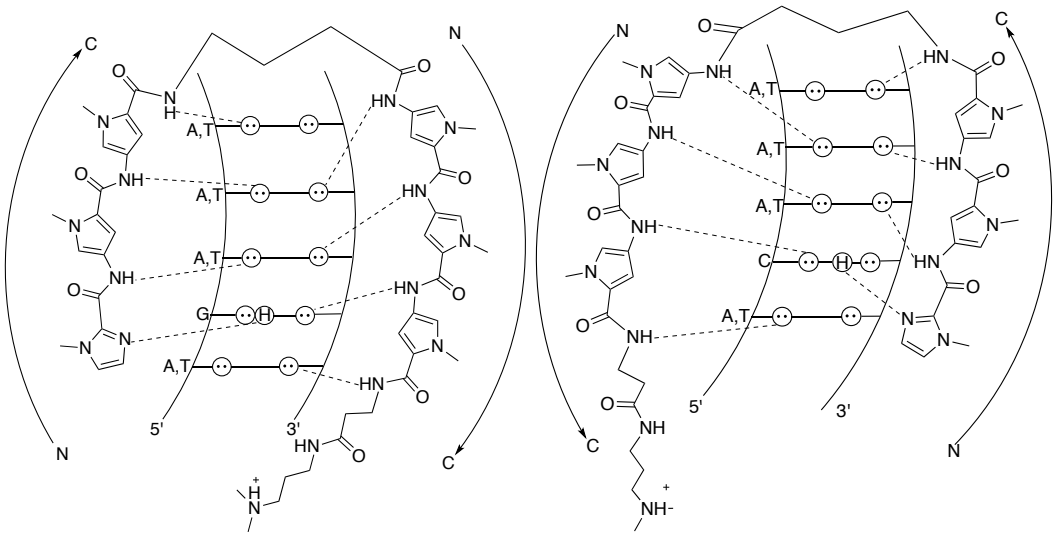


ImImPy-γ-ImPyPy-β-Dp•5'-NNGGCNN-3'

Figure 6.1: Chemical structure of a hairpin polyamide, with the putative hydrogen bonds to the DNA minor groove shown. Pairing rules for the five-membered heterocycles shown to the left (304). On the right diagram, binding models for the complex formed between ImImPy-γ-ImPyPy-β-Dp and 5'-NNGGCNN-3'. Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. N is represented as A or T but can be A, T, G, or C (302). The imidazole and pyrrole rings are represented as shaded and unshaded circles respectively, the curved line represents γ-aminobutyric acid, and β-alanine residue is represented as diamond; 'Hp' denotes hydroxypropyl.

The binding of a hairpin polyamide to DNA could, in principle, occur in two different orientations - *i.e*. with each polyamide chain positioned with its N-terminus located toward the 5'-side of the targeted DNA strand (5'-3') or binding with the C-terminus located toward the 5'-side of binding site (3'-5'). Thus a six-ring hairpin polyamide of ImPyPy-γ-PyPyPy could recognize 5'-TGTTA-3' by placing the N-terminus of each three ring polyamide subunit at the 5'-side of each recognized DNA strand. With the N-terminus on the 3'-side of each recognized strand the target sequence would be 5'-TCTTA-3' (Figure 6.2). MPE.Fe(II) footprinting demonstrated that ImPyPy-γ-PyPyPy-β-Dp bound strongly at 5'-TGTTA-3' while weak interaction at 5'-TCTTA-3', with a 16-fold preference for placing the N-termini of the polyamide subunits at the 5'-end of the targeted DNA strand (305). A positive charge (β-Dp) at the C-terminal is necessary for

enhanced sequence specificity; an uncharged ethoxyamide group showed equivalent affinity for the target site but its specificity was significantly (9-fold) reduced; on the other hand, placing a positive charge at N-terminal binds enables binding to both 5'TGTTA-3' and 5'-TCTTA-3' with equal affinity thus losing the orientational specificity. Thus, for optimal orientation and sequence-specificity, a positive charge at the C-terminus and no substitution at N-terminus are preferred.



ImPyPy-γ-PyPyPy-β-Dp 5'-(A,T)G(A,T)$_3$-3'

ImPyPy-γ-PyPyPy-β-Dp 5'-(A,T)C(A,T)$_3$-3'

Figure 6.2: Potential binding models for ImPyPy-γ-PyPyPy-β-Dp in complex with 5'-WGWWW-3' (left) and 5'-WCWWW-3' (right) (W = A or T). Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines. For the schematic binding models, the imidazole and pyrrole rings are represented as shaded and unshaded spheres respectively, the curved line represents γ-aminobutyric acid, and the β-alanine residue is represented as an unshaded diamond. Shaded arrows represent the orientation of individual polyamide subunits (305).

The recognition of specific sequences by hairpin polyamides has been proved useful, for example, for sequence selective alkylation by hairpin conjugates (306-308), inhibition of gene expression by interfering with normal DNA/protein interactions (309,310) and as molecular probes to identify particular DNA sites (311-313). The DNA alkylator chlorambucil, which alkylates the N3 of adenine in the minor groove, was conjugated with an eight-ring hairpin polyamide (ImPy-β-ImPy-γ-ImPy-β-ImPy-β-Dp).

The polyamide can recognise 5'-WGCWGCW-3' (W = A or T), which is the sequence adjacent to both sides of the TATA box in the HIV promoter regions (314). Thus combination of chlorambucil at the α-position of the γ-turn of the polyamide offers a class of bifunctional agents that can potentially bind to predetermined DNA sequences with high affinity and specificity for subsequent covalent reactions at N3 of adenine by chlorambucil. This hairpin-conjugate showed binding at 5'-AGCTGCT-3' and 5'-TGCAGCA-3'- sites proximal to the TATA box; the equilibrium constants of these sites for hairpin-chlorambucil were similar to those of the unsubstituted polyamide, indicating that the polyamide binding affinity and specificity are not altered by substitution (308). Moreover, alkylations selectively occurred at subnanomolar concentrations only at adenines adjacent to these binding sites, as evident from a thermal cleavage assay, whereas no cleavage occurred even at 10 μM for free chlorambucil. Conjugate formation between chlorambucil and the twelve-ring hairpin polyamide (ImPy-β-ImPy-β-ImPy-γ-ImPy-β-ImPy-β-ImPy-β-Dp-Chlorambucil) can target two similar sites containing AGC trinucleotide repeats - 11 bp (5'-G*C**AGCAGCAGCA*-3') and 12 bp (5'-*AGCAGCAGCAGC*A-3') (the alkylation sites are underlined and the polyamide binding site is shown in italics)(306). Different trinucleotide repeat sequences exist in some regions of genomic DNA and the expansion of this repeat sequence sometimes causes neurological disorders (315,316); the expansion of AGC trinucleotide repeats in the first exon of the Huntingtin (HTT) gene produces HTT protein that selectively disrupts neuronal protein function, ultimately causing Huntington's disease (317-319). Sequence selective binding at these AGC repeats, with simultaneous alkylation, prevents progression of RNA polymerase that ultimately ceases HTT protein production (306). Another conjugate (HP-CPI) composed of hairpin (ImImPyPy-γ-ImPyPyPy) with cyclopropylpyrrolindole (CPI) showed some cell growth inhibition in mammalian cells by alkylating a single adenine flanking the polyamide binding site (307).

Hairpin polyamides have also been successfully applied as molecular probes for the detection of specific double stranded DNA sequences by a fluorescent method. Three eight-ring pyrrole-imidazole polyamides, conjugated with thiazole orange, were reported to bind specifically to the target site without compromising their affinity and specificity (312); the polyamide moiety directs the nonfluorescent conjugate to a specific site of the DNA duplex and delivers the tethered thiazole orange fluorophore to an adjacent site. Upon polyamide binding, the thiazole orange intercalates into the DNA, restricting the conformation of the fluorophore, resulting in substantial fluorescence emission following excitation. All the conjugates fluoresced weakly in the

absence of DNA but showed significant enhancement (>1000-fold) upon addition of 1 equivalent of match DNA and only slightly enhancement with the addition of a mismatch DNA (312). Recently, an eight-ring hairpin polyamide (ImImImPy-γ-PyPyPyPy-β-Dp-NH_2) has been reported to target human telomere sequence (5'-TTAGGG-3'). Thermal melting analysis showed a large increase in T_m (27.7 °C) of a telomeric duplex containing this sequence. Surface plasmon resonance analysis demonstrated efficient binding of the polyamide with this sequence and demonstrated that it could even discriminate between 1 bp mismatch site (5'-TTAAGG-3'); thus the compound has become a good probe for human telomere staining (313).

6.2 Experimental design

The DNA sequence selectivity of hairpin polyamides (PA1 and PA2) was examined individually over a range of concentrations from 10 μM to 0.5 μM with three different DNA fragments- HexAfor, HexBfor and MS2. The compounds were further examined with another two DNA fragments containing potential sites for these compounds- LE1 and LE2 (Figure 6.3), which has been designed and prepared to test the specificity of these two compounds; for these fragments, the ligand concentrations ranged from 3 μM to 0.01 μM. For both footprinting experiments, samples were prepared by mixing 1.5 μl of each ligand (diluted in 10 mM Tris-HCl, pH 7.5 containing 10 mM NaCl) with 1.5 μl of radiolabelled DNA and were incubated at room temperature for 24 hours.

LE1
5′GATCACGTTATAT**GGTG**ATATAT**GGTC**ATATGATCACGTTATAT**GGTG**ATATAT**GGTC**
3′CTAGTGCAATATA**CCAC**TATATA**CCAG**TATACTAGTGCAATATA**CCAC**TATATA**CCAT**

ATATGATC3′
TATACTAG5′

LE2
5′GATCACGTTATAT**GGTG**ATATAT**GGTC**ATATGATCATAT**GACC**ATATAT**CACC**ATATAA
3′CTAGTGCAATATA**CCAC**TATATA**CCAG**TATACTAGTATA**CTGG**TATATA**GTGG**TATATT

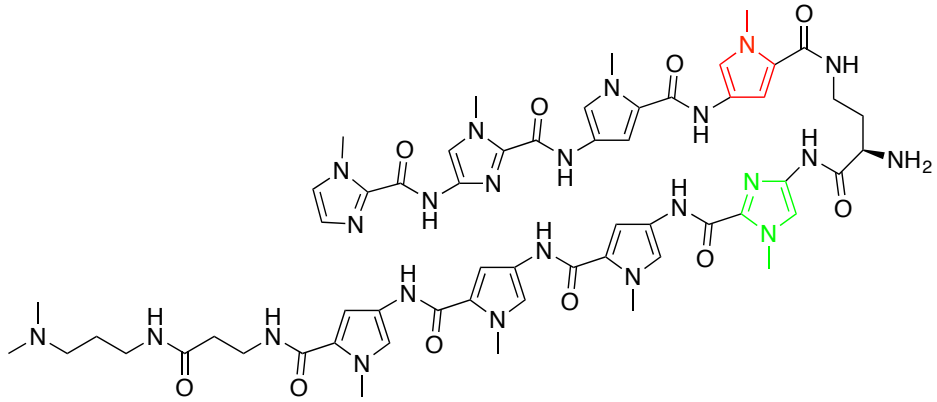
CGTGATCATAT**GACC**ATATAT**CACC**ATATAACGTGATC3′
GCACTAGTATA**CTGG**TATATA**GTGG**TATATTGCACTAG5′

Figure 6.3: Sequences of DNA fragments containing predicted sites of compounds PA1 and PA2. Predicted binding sites (5'-GGTC-3' for PA1 and 5'-GGTG-3' for PA2) are shown in bold letters.

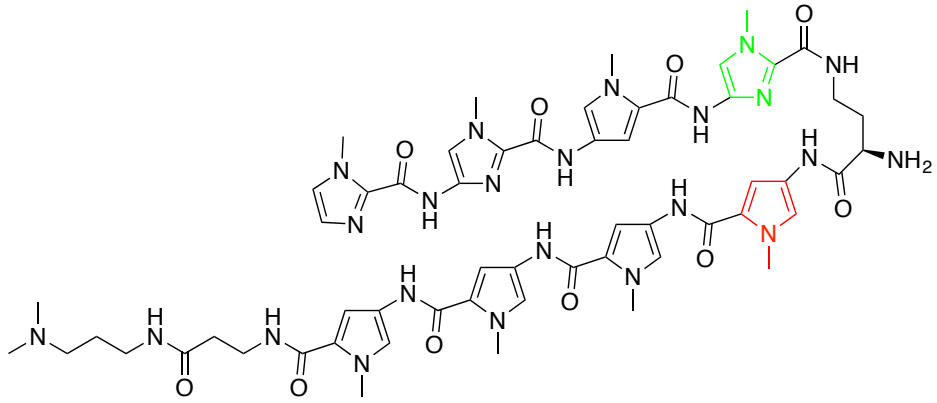
6.3 Chemical structures of PA1 and PA2

The compounds were synthesized by Andrew Fallows in Dr Glenn Burley's lab in University of Leicester. In PA1 and PA2, 'Im' and 'Py' indicate *N*-methylimidazole and *N*-methylpyrrole respectively; 'γ' indicates gamma-aminobutyric acid and 'β' denotes

beta-alanine. β-dimethylaminopropylamine is abbreviated as ‘Dp’. The difference in arrangement of pyrrole and imidazole rings between PA1 and PA2 are shown in colours..



ImImPyPy-γ-ImPyPyPy-β-Dp
(PA1)



ImImPyIm-γ-PyPyPyPy-β-Dp
(PA2)

6.4 Results

6.4.1 With DNA fragments – HexAfor, HexBfor and MS2

Figure 6.4 shows DNase I footprints for the interaction of compounds PA1 and PA2 with HexAfor, HexBfor and MS2; for each DNA fragment, the footprinting gels of both compounds are placed together. The sequences protected from cleavage in these fragments are summarized in Figure 6.5.

For HexAfor with 10 μM of PA2, a footprint covering 14 bases (site 1) is observed at the bottom of the gel. Compound PA1 produces footprints at the same site with 5 μM. Another large footprint protecting 18 bases (site 2) is seen for both compounds with 10 μM; the protections are not very clear and appear as attenuated bands rather than complete protection. Both compounds produce a clear footprint at 5’AGGCCTGC (site

3); for PA1, the footprint is evident at a concentration of 2 μM, while for PA2 the footprint is seen with 10 μM and 5 μM. Protection at 5'ACGTTAAC (site 4) is observed with 2 μM PA1 whereas PA2 protects this site at a concentration of 1 μM. Another large footprint at the top of the gel is produced at 5'ATAGATCTAGAAT (site 5) with 10 μM of PA1 while PA2 produces two separate protections in this region which persist to a concentration of 2 μM. The clearest footprints are at sites 1, 3 and 4 for PA1, but at site 4 for PA2. *For HexBfor*, a large footprint covering 20 bp (site 1) with attenuated bands at the bottom of the gel is produced with 10 μM PA1; the same site is observed with PA2 except three bases at 5'-(upper) end are unprotected. A clear footprint at 5'GACGTCATGAT (site 2) is produced by PA1 with concentrations 10 μM and 5 μM. This site is also seen with PA2 at a concentration of 5 μM. Compound PA1 protects a 16 bp site (site 3) with 10 μM while PA2 produces a shorter site at the same region, protecting 5'ATATACATAT at a concentration of 5 μM. Both compounds protect 5'AGCTTATAA (site 5); for PA1, the protection occurs only with highest concentration, whereas PA2 produces a footprint with 5 μM. Another footprint is produced by 10 μM PA1 at 5'ACGTATACG (site 4); at the same location, PA2 shows a large footprint covering 15 bp with 10 μM ligand that is shortened to 5'GTATACG with 5 μM. Protection is seen at 5'AGCTTATAA (site 5) with 10 μM PA1 while PA2 produced a footprint with 5 μM at the same site leaving the first adenine at 5'-end unprotected. In summary, for this fragment PA1 produces only one clear footprint at site 2, while PA2 produces several clear footprints the strongest of which seems to be at site 3. *With DNA fragment MS2*, both compounds protect 5'GTGTATT (site 1) near the bottom of the gel; ligand PA1 protects with 10 μM whereas PA2 binds at a concentration as low as 1 μM (the lane for 2 μM PA2 is missing). Another footprint is produced at 5'GTTGGAAATCGG (site 2) with 10 μM PA1 and 5 μM PA2. A clear footprint is observed with PA1 at 5'TACGACTAGT (site 3), which persists to 1 μM; this site is protected with 5 μM of PA2. Both compounds bind at 5'TTAACCTTGA (site 4) at concentrations as low as 0.5 μM. A clear footprint at 5'ATACCAGGA (site 5) is produced by 0.5 μM PA2. This site is bound by 10 μM PA1 though the lower part of this footprint is still attenuated with 1 μM ligand. A large footprint (site 6) covering 17 bp is seen with 10 μM of PA1; in this location, PA2 produces two footprints at 5'ACATAAC and 5'TCACGCC at concentrations of 10 μM and 5 μM respectively. Another footprint at the top of the gel (site 7) in the region around 5'AGACCCCACC is produced by 0.5 μM of both ligands. For this fragment PA1 binds best to sites 3, 4 and 7 while PA2 binds best to sites 4, 5 and 7. From the footprints of both compounds, possible hexanucleotide binding sites are summarized in Table 6.1.

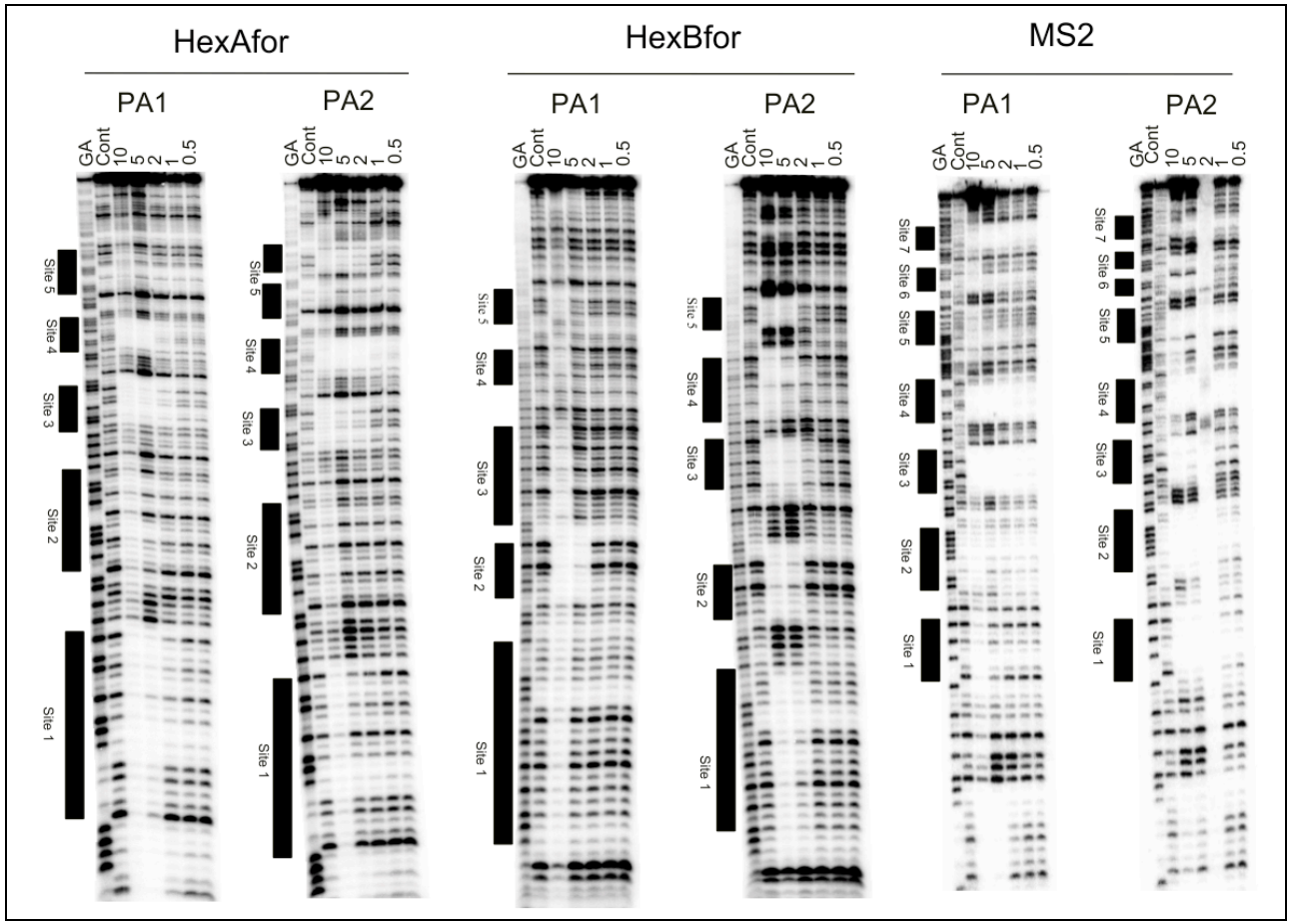


Figure 6.4: DNase I footprinting gels showing the interaction of various concentrations of PA1 and PA2 with HexAfor, HexBfor and MS2. The samples were incubated with the ligand at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands. Filled bars indicate the location of the footprints.

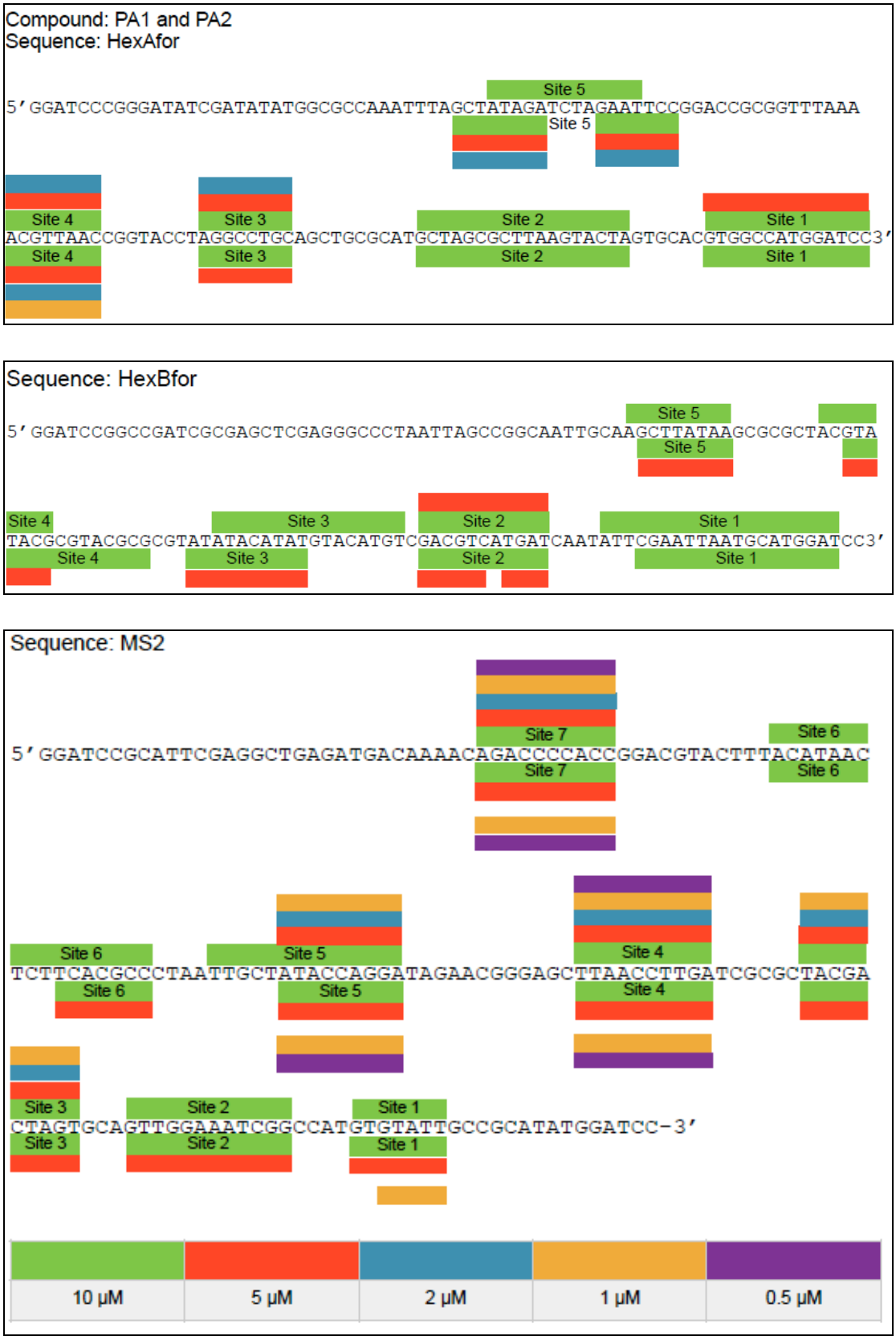


Figure 6.5: Sequence of the footprinting substrates indicating the regions protected by compounds PA1 (boxes above the strand) and PA2 (boxes below the strand) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand. Only the sequence of radiolabelled strand is shown.

Table 6.1: Possible hexanucleotide binding sites of compounds PA1 and PA2 with corresponding concentrations of the compounds required to produce footprints at those sites.

DNA fragment	Binding sites	Compound PA1	Compound PA2
HexAfor	5'-CGTTAA-3' (site 4)	2.0 μM	1.0 μM
	5'-AGGCCT-3' (site 3)	2.0 μM	5.0 μM
	5'-TGGCCA-3' (site 1)	5.0 μM	10 μM
	5'-AGATCT-3' (site 5)	10 μM	2.0 μM
	5'-AGCGCT-3' (site 2)	10 μM	10 μM
	5'-AGTACT-3' (site 2)	10 μM	10 μM
HexBfor	5'-TGACGT-3' (site 2)	5.0 μM	5.0 μM
	5'-AGCTTA-3' (site 5)	10 μM	5.0 μM
	5'-CGTATA-3'(site 4)	10 μM	5.0 μM
	5'-TGGATC-3' (site 1)	10 μM	10 μM
	5'-TGTATA-3' (site 3)	10 μM	5 μM
MS2	5'-GGGTCT-3' (site 7)	0.5 μM	0.5 μM
	5'-TGGGGT-3' (site 7)	0.5 μM	0.5 μM
	5'-AGGTTA-3' (site 4)	0.5 μM	0.5 μM
	5'-TGGTAT-3' (site 5)	1.0 μM	0.5 μM
	5'-AGTCGT-3' (site 3)	1.0 μM	5.0 μM
	5'-AATACA-3' (site 1)	10 μM	1.0 μM
	5'-GGGCGA-3' (site 6)	10 μM	5.0 μM
	5'-TGGAAA-3' (site 2)	10 μM	5.0 μM

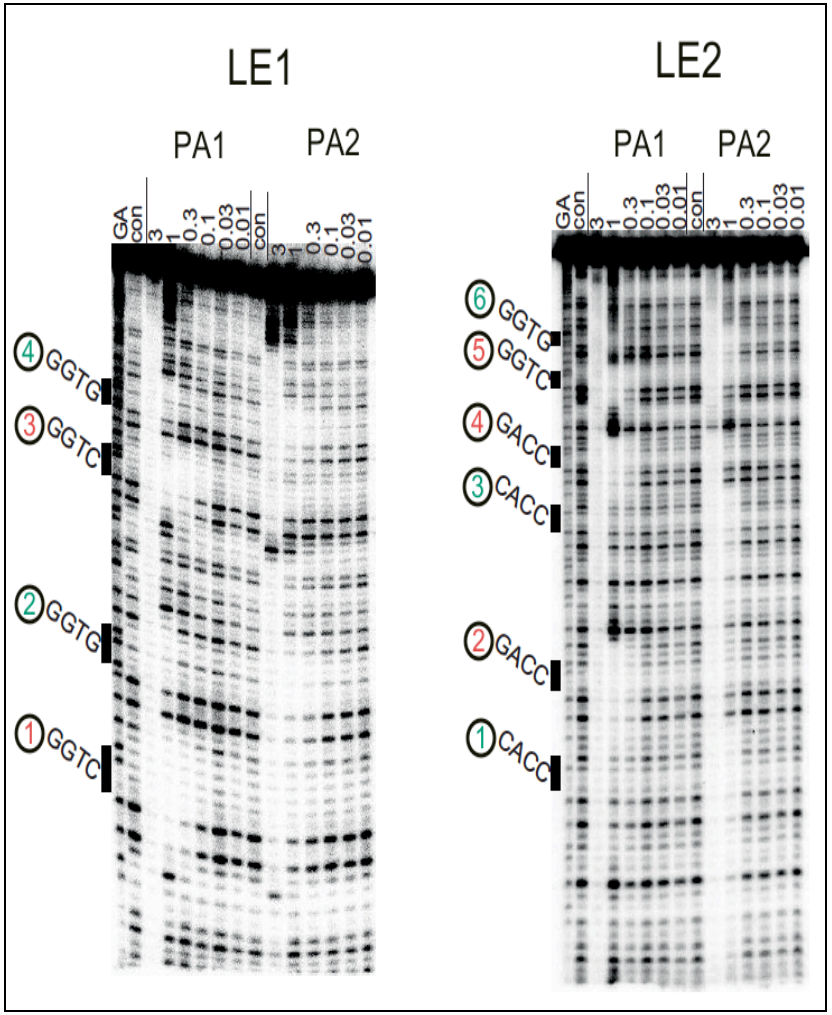
6.4.2 With DNA fragments- LE1 and LE2

On the basis of the usual Py/Im pairing rules, compounds PA1 and PA2 might be predicted to bind to the sites GG(A/T)C and GG(A/T)G respectively. DNA fragments LE1 and LE2 were therefore prepared, containing these potential binding sites. These two fragments are dimer and trimer of the oligonucleotide duplex used for cloning and so contain two and three copies respectively of each of these potential target sequences. Figure 6.6 (left panel) shows DNase I footprints for the interaction of compounds PA1 and PA2 with LE1 and LE2.The sequences protected from cleavage in these fragments are summarized in the right panel of Figure 6.6.

DNA fragment LE1 contains two sites each of 5'-GGTC-3'and 5'-GGTG-3', while LE2 contains three copies of each of these sites, Both PA1 and PA2 both show footprints at all sites with highest concentration (3 μM). However there are clear differences

between the compounds at lower concentrations. In LE1, PA1 protected site 1 and 3 (contain 5'-GGTC-3') at concentrations between 0.1 - 0.3 μM, while the other two sites (2 and 4), containing 5'-GGTG-3' only generated protections at concentrations of 1 μM and above. The converse is seen with PA2, for which all four sites were protected with 1 μM ligand, but footprints are only seen at sites 2 and 4 on decreasing the concentration to 0.3 μM.

Similar results are seen with DNA fragment LE2. PA1 protected sites 2, 4 and 5 (all of which contains the sequence 5'-GGTC-3') at 0.3 μM, while sites 1, 3 and 6 only produced footprints at concentrations of 1 μM and above. Again, the opposite was seen with PA2 for which footprints can be seen at all the sites (1 to 6) with 3 μM ligand but only but sites 1 and 6 (containing 5'-GGTG-3') were protected at 1 μM.



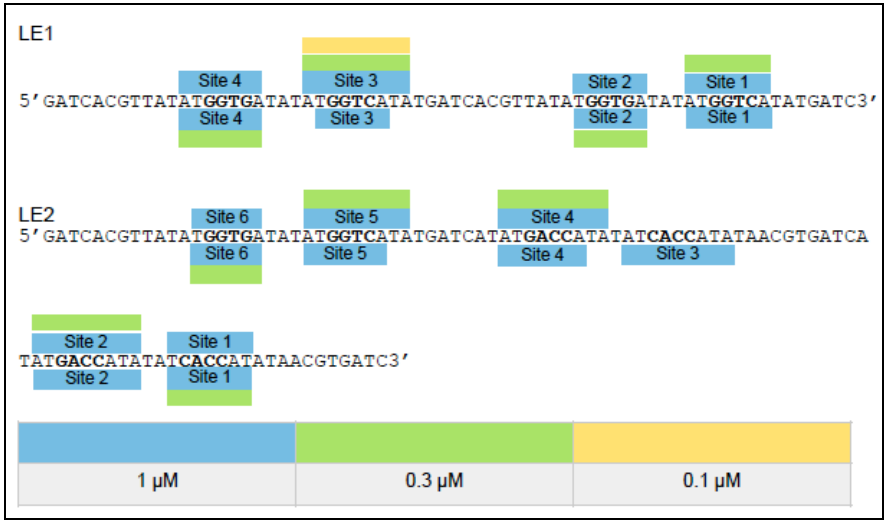


Figure 6.6: *In the left panel*, DNase I footprinting gels showing the interaction of various concentrations of PA1 and PA2 with LE1 and LE2. The samples were incubated with the ligand at room temperature for 24 hours before digesting with DNase I. Ligand concentrations (μM) are shown at the top of each gel lane. Tracks labelled “GA” correspond to markers specific for purines, while “Cont” indicates DNase I cleavage in the absence of added ligands. Filled bars along with numbers in circles indicate the location of the footprints and the binding sites. *In the right panel*, Sequence of the footprinting substrates indicating the regions protected by compounds PA1 (boxes above the strand) and PA2 (boxes below the strand) from DNase I cleavage. Coloured boxes indicate the binding sequences at different concentrations of ligand. Only the sequence of radiolabelled strand is shown.

6.5 Discussion

Compounds PA1 and PA2 are each composed of eight-ring polyamides. The only difference between them is the position of one imidazole-pyrrole pair, adjacent to the γ-turn. In PA1, the pyrrole is on the N-terminal side of the γ-turn, with imidazole on the C-terminal side, while in PA2 these positions are reversed. All the other pyrroles and imidazoles are in identical positions in both compounds.

In principle, the compounds could bind in two different orientations without violating the pairing rules: i) with the N-terminus of the polyamide towards 5'-end or ii) the C-terminus toward the 5'-end. In case of N-terminus at 5'-end, the potential binding sites will be 5'-XGGXCX-3' and 5'-XGGXGX-3' for PA1 and PA2 respectively (X = A or T). For alternate orientation, the binding sites become 5'-XCCXGX-3' and 5'-XCCXCX-3'. White *et al* (305) reported that a six-ring polyamide (ImPyPy-γ-PyPyPy-β-Dp) bound to 5'-TGTTA-3' with 16-fold greater affinity over 5'-TCTTA-3', indicating that the N-terminus of the polyamide subunit is preferentially positioned at the 5'-end of the target DNA strand; the binding of hairpin polyamides towards 5'-end has also been reported in numerous papers (300,302,306,311-313,320). In addition to their strong affinity for their putative binding sites, hairpin polyamides have also been reported to bind to some mismatch sites with variable affinity. For example, the six-ring hairpin polyamide (ImImPy-γ-PyPyPy-β-Dp) was found to bind with 1.8-fold preference for the designated binding site 5'-TGGTA-3' over the single base pair mismatch site 5'-TGTTA-3' while it exhibited 102-fold lower affinity for the double base pair mismatch site 5'-TATTA-3' (321). Thus it can be anticipated that compounds PA1 and PA2 would bind at sites in accordance with the pairing rules, as shown in Figure 6.7, though they might also tolerate the presence of some mismatches.

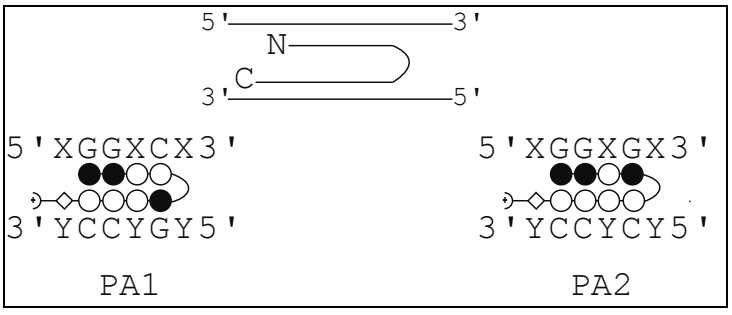


Figure 6.7: Binding models for compounds PA1 and PA2 in complex with (5'XGGXCX3')/(3'YCCYGY5') and (5'XGGXGX3')/(3'YCCYCY5') (X = A or T while Y = T or A) respectively. Imidazole and pyrrole rings are represented as shaded and unshaded spheres respectively, the curved line represents γ-aminobutyric acid, and the β-alanine residue is represented as an unshaded diamond.

DNA fragments LE1 and LE2 contain the potential binding sites, 5'-TGGTCA-3' and 5'-TGGTGA-3', along with two sites that contain single base mismatches from each of these: 5'-TGATCA-3' and 5'-ACGTGA-3'. It is clear that compound PA1 binds with highest affinity at 5'-TGGTCA-3' and the site fully complies with the expected pairing rules. 5'-TGGTGA-3' and 5'-TGATCA-3' consist of 1 bp mismatch sites while 5'-ACGTGA-3' contains 2 bp mismatch sites and are lower affinity sites for which DNase I cleavage was only inhibited with higher concentrations (Figure 6.8). Similarly, compound PA2 targets 5'-TGGTGA-3' with greatest affinity as expected, while footprints at the 1 bp mismatch sites 5'-TGGTCA-3' and 5'-ACGTGA-3' and the 2 bp mismatch site 5'-TGATCA-3' are only seen with higher ligand concentrations.

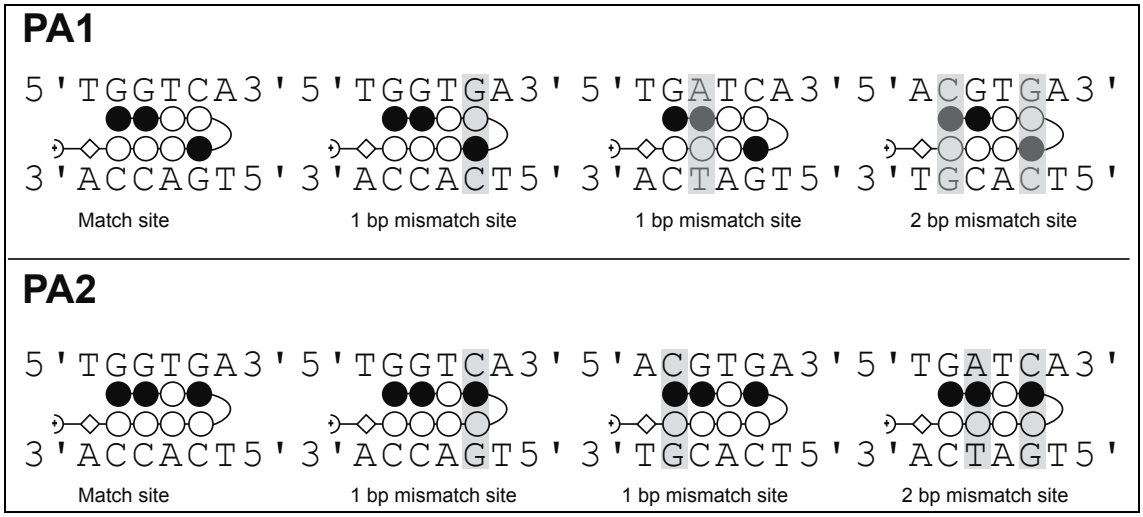


Figure 6.8: Binding models for compounds PA1 and PA2 in complex with match site and mismatch sites in DNA fragments LE1 and LE2. Imidazole and pyrrole rings are represented as shaded and unshaded spheres respectively, the curved line represents γ-aminobutyric acid, and the β-alanine residue is represented as an unshaded diamond. The grey vertical box (s) indicates the position of mismatch site (s).

DNA fragments HexAfor, HexBfor and MS2 offer numerous regions for binding of PA1 and PA2; though none of these fragments contain the preferred sites for PA1 and PA2 5'-(A/T)GG(A/T)C(A/T)-3' and 5'-(A/T)GG(A/T)G(A/T)-3' respectively. However, clear footprints are observed on both fragments, indicating that the compounds must be binding to regions that contain one or more deviations from the expected target sequence. The site bound at lowest concentration (0.5 μM) of PA1 is 5'-GGGTCT-3' (site 7/MS2) that is targeted by pairing of pyrroles and imidazoles (underlined bases) and the terminal thymine with a γ-turn, thus the site can be considered as a single mismatch against one base pair. Binding at 5'-AGGTTA-3' (site 4/MS2) also has a single against one base pair and shows binding at same concentration. It should be noted that the lowest concentration used for these DNA fragments was 0.5 μM, thus a further decrease in concentration might differentiate between these two sites. The

presence of two consecutive guanines seems more favourable for binding as 5'-TGGTAT-3' (site 5/MS2), 5'-AGGCCT-3' (site 3/HexA) and 5'-TGGCCA-3' (site 1/HexA) have greater affinity than the other sites with ligand mismatches against one base pair, such as 5'-AGTACT-3' (site 2/HexA), 5'-AGATCT-3' (site 5/HexA) and 5'-TGTACA-3' (site 3/HexB) (mismatch sites underlined) as evident from their concentration dependence. Schemes showing the potential interaction with these mismatch sites are shown in Figure 6.9. 5'-TGGATC-3' (site 1/HexB), though having mismatches against two base pairs (underlined) was still protected at lower concentrations than some single base pair mismatch sites, further emphasizing the importance of the two imidazole-pyrrole pairs. Additionally, uninterrupted pairing of pyrrole and imidazole gives the highest affinity as seen with 5'-TGGTAT-3' (site 5/MS2) which contains four continuous base pair matches, which showed higher affinity than 5'-AGGCCT-3' (site 3/HexA) having a mismatch against the base pair in the middle of the binding site (mismatch sites underlined).

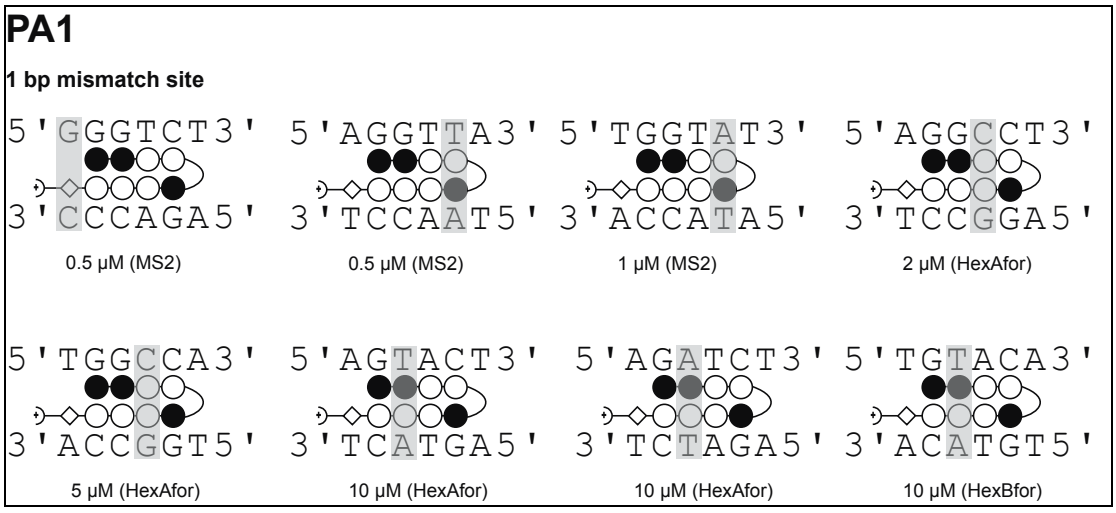


Figure 6.9: Binding models for compound PA1 in complex with 1 bp mismatch sites. Imidazole and pyrrole rings are represented as shaded and unshaded spheres respectively, the curved line represents γ-aminobutyric acid, and the β-alanine residue is represented as an unshaded diamond. The grey vertical box (s) indicates the position of mismatch site (s).

Compound PA2 has high affinity for 5'-CGGTGG-3' (site 7/MS2), as might be expected as the core sequence (GGTG) fully matches the pairing of pyrroles and imidazoles, despite mismatch of γ-turn and β-Dp tail which are both opposite GC base pairs. The ligand also has equally high affinity, producing footprints at low concentrations, with 5'-AGGTTA-3' (site 4/MS2) and 5'-TGGTAT-3' (site 5/MS2) that each contain mismatches against single base pairs (underlined). The affinity is further reduced when number of mismatch sites is increased as evident from sites 5'-TGACGT-3' (site 2/HexB), 5'-AGGCCT-3' (site 3/HexA) and 5'-AGCTTA-3' (site 5/HexB) (Figure 6.10). Both

compounds bind at 5'-ACGTTAAC-3' (site 4/HexA), the site contain more than 2 bp mismatch sites. The differences in affinity and binding at mismatch sites between PA1 and PA2 may be due to the composition of each polyamide subunit. For PA1, the N-terminal subunit contains two imidazoles and two pyrroles; on the other hand, three imidazoles and one pyrrole are present in PA2. It is known that the number of imidazoles in the polyamide as well as their positions can significantly influence the overall structure and their association with the DNA (320). Also Han *et al* (322) reported that a polyamide composed of four imidazoles (ImImImIm) adopted a planar structure while one with four pyrroles (PyPyPyPy) had a more helical structure and suggested that the planarity of the polyamide increases with the number of imidazoles. Thus the higher number of imidazoles in the N-terminal half of the PA2 hairpin might make it harder for the ligand to align with its DNA binding sites.

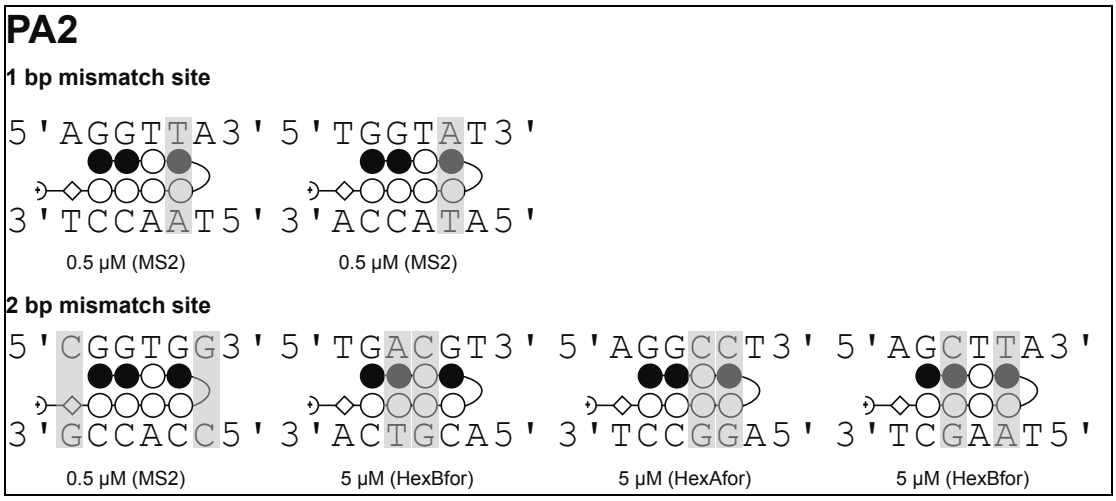


Figure 6.10: Binding models for compound PA2 in complex with 1 bp and 2 bp mismatch sites. Imidazole and pyrrole rings are represented as shaded and unshaded spheres respectively, the curved line represents γ-aminobutyric acid, and the β-alanine residue is represented as an unshaded diamond. The grey vertical box (s) indicates the position of mismatch site (s).

6.6 Conclusion

From the footprinting assay, it is clearly evident that hairpin polyamides can be designed in accordance with the pairing rules to target specific sequences. Both hairpin polyamides PA1 and PA2 are found to identify their putative sites 5'-TGGTCA-3' and 5'-TGGTGA-3' with highest affinity.

7 Conclusion

The experiments presented in this thesis describe the sequence selectivity of two groups of pyrrolobenzodiazepine-conjugates; i) conjugation with pyrrole, imidazole and thiazole, and ii) conjugation with benzofused rings such as benzothiophene, benzofuran and indole. The sequence selectivity of these compounds was examined by DNase I footprinting assay on the DNA fragments HexA, HexB and MS1/MS2. HexA/HexB contain all the symmetrical hexanucleotide sequences (257) while MS1/MS2 have all possible tetranucleotide sequences (256). PBD-conjugates with pyrrole, imidazole and thiazole were then examined with three designed DNA fragments; GXC3, GXC2 and CXG2, which contained putative ligand binding sites (based on the results with the other DNA fragments). Both groups of conjugates were also examined by fluorescence melting on seven oligonucleotide duplexes to determine their effects on stability; each duplex contained a hexanucleotide binding site for the ligand. This research work also developed a ligation assay to determine the preferred binding site of PBD-dimers; here, PBD-dimer, SJG-136 was tested against different combinations of oligonucleotide duplexes containing the potential binding sites GATC, GTAC and GAATC. In addition to the work with the PBD compounds, the sequence selectivity of two hairpin polyamides was also determined by DNase I footprinting.

PBD conjugates with pyrrole, imidazole and thiazole

Conjugation of pyrrolobenzodiazepine with polyamides composed of pyrrole, imidazole and thiazole is an effective way to produce DNA minor groove binders with increased sequence selectivity. The polyamides that consist of three heterocyclic rings (RMH41, RMH43, RMH47 and RMH53) have higher sequence selectivity and greater affinity than those with only two rings (RMH35 and RMH71). All the PBDs with three heterocyclic rings have similar binding preferences and bind best to AT-rich sequences. Among these, PBD-Im-Py-Py (RMH41) and PBD-Thz-Py-Py (RMH43) have highest affinity while placement of thiazole in between the two pyrroles (RMH47) reduces the affinity of the compound. The compound with a greater number of imidazoles (RMH53, PBD-Py-Im-Im) shows some affinity for GC-rich sequences at higher ligand concentration. The preferences of these compounds for binding to AT-rich sequences was further confirmed in DNase I footprinting with designed fragments, in which all the compounds bind at ATATAT regions having guanine at either 5'-end or 3'-end, rather than across the different GC-rich regions.

PBD conjugates with benzofused rings

Conjugation of PBD with benzofused rings clearly demonstrates the effect of these heterocycles on overall sequence selectivity of the compound. These benzofused conjugates show similar binding preferences to the aforementioned polyamide conjugates, but bound at lower concentrations. Benzofuran and benzothiophene can be good choice for conjugation as these heterocycles bind at AT-rich sequences such as TATA, ATAT, AATT, TTAA with a guanine at 5'-end or 3'-end as observed from footprints from PBD-Py-Benzofuran and PBD-Py-Benzothiophene; on the other hand, the combination of indole and PBD did not appear to bind. Addition of a terminal acetate moiety with benzothiophene does not significantly affect the sequence selectivity. On the other hand, the presence of imidazole in the linker (PBD-Im-Benzothiophene-Acetate) causes some interaction with GC-rich sequences at higher concentrations. The presence of pyrrole or imidazole in the linker is necessary as the conjugate composed of PBD and benzothiophene without any heterocycle shows the lowest potency among all conjugates; also can cover shorter sequences only at higher concentrations possibly due to shorter size of the ligand.

In the fluorescence melting studies with different hexanucleotide oligonucleotide duplexes, all the PBD conjugates caused a large increase in melting temperature indicating strong binding of these compounds with the duplexes. The biphasic melting curves demonstrate the formation of both 1:1 (1^{st} adduct) and 2:1 (2^{nd} adduct) ligand-DNA complexes. The binding of one PBD unit (forming the 1^{st} adduct) with one guanine produces a large increase in melting temperature, a further large increase in T_m is seen at higher ligand concentrations consistent with binding to two guanines (formation of the 2^{nd} adduct).

The effect of polyamides on the overall sequence selectivity of PBD-conjugates would be better understood if the footprints and melting profiles of these compounds could be compared with those for the polyamide portion alone; but these compounds were not available. Wells *et al* (167) reported footprinting studies with PBD-conjugates with polypyrrole chain and showed that the binding of the compounds is controlled by the covalent bond formation of PBD unit. Additionally, molecular modelling of PBD-Im-Py-Py was found to conform 1:1 binding mode (164). Our results of footprinting studies with PBD conjugates with pyrrole, imidazole and thiazole with hexanucleotide containing binding sites also indicate binding preferences towards 1:1 mode suggesting the binding is controlled by the PBD unit while the polyamide extremities assist in finding suitable minor groove position.

Ligation assays with PBD dimers

The ligation assay was designed to determine the preferred binding site for SJG-136, a PBD dimer composed of two PBD units joined by propyldioxy linker, though it could be applied to any other agent that produces inter-strand cross-links. The compound was exposed to different oligonucleotide duplexes containing binding sites- GATC, GTAC and GAATC. The duplexes, upon annealing, produce overhangs containing one of these binding sites; the ligand can then form interstrand cross-links between complementary overhangs and the preferred binding site will give the greatest yield of the adducts. The compound SJG-136 was able to produce cross-links at the 4 base pairs sites GATC and GTAC, but no adduct formation was observed with GAATC. Between the 4 bp sites, G<u>A</u>TC was more efficiently cross-linked that GTAC. This can be explained as the more stable adduct forms a hydrogen bond between the adenine (underlined) and the anthranilate ring of PBD. On the other hand, the two PBD units of SJG-136 cannot cross-link between the two guanines of the complementary strands of GAATC, possibly due to distortion of DNA structure. This distortion would arise when one PBD unit forms a covalent link to one guanine; such that the second guanine on the complementary strand cannot be reached by the second PBD unit. The ligation assay thus can successfully differentiate between similar binding sites for SJG-136. The assay could be applicable for determining the preferred binding sites of other types of PBD-dimers; for example, ligands with longer linkers (323) or linker containing polyamides between the two PBD units (324).

Hairpin polyamides

The hairpin polyamides examined in this thesis were found to bind to the predicted sites 5'-(A/T)GG(A/T)C(A/T)-3' and 5'-(A/T)GG(A/T)G(A/T)-3' with highest affinity; binding to these sites are as expected from the well known recognition of A/T or T/A by pyrrole-pyrrole and C/G and G/C by pyrrole-imidazole and imidazole-pyrrole respectively (151,153), while the beta-alanine and hairpin turn unit both prefer A/T base-pairs (302,303). The compounds also bound to other sites at higher concentrations that contain 1 bp or 2 bp mismatch sites. Binding at mismatch sites is concentration dependent and also uninterrupted pairing of heterocyclic rings is more favoured than mismatch at the middle of the binding site. The composition of each polyamide chain also significantly affects binding as large number of imidazoles in one chain conforms planar structure that is less suitable for binding. The sequence selectivity of these hairpin polyamides could be capitalized to synthesize new DNA alkylating agents by conjugating them with pyrrolobenzodiazepine, as a combination of

hairpin polyamide and chlorambucil (306) or cyclopropylpyrrolindole (307) has been reported to have enhanced biological activities than the individual components.

Appendices

Appendix 1: Commercial sequencing of puC18 designed DNA fragments clones

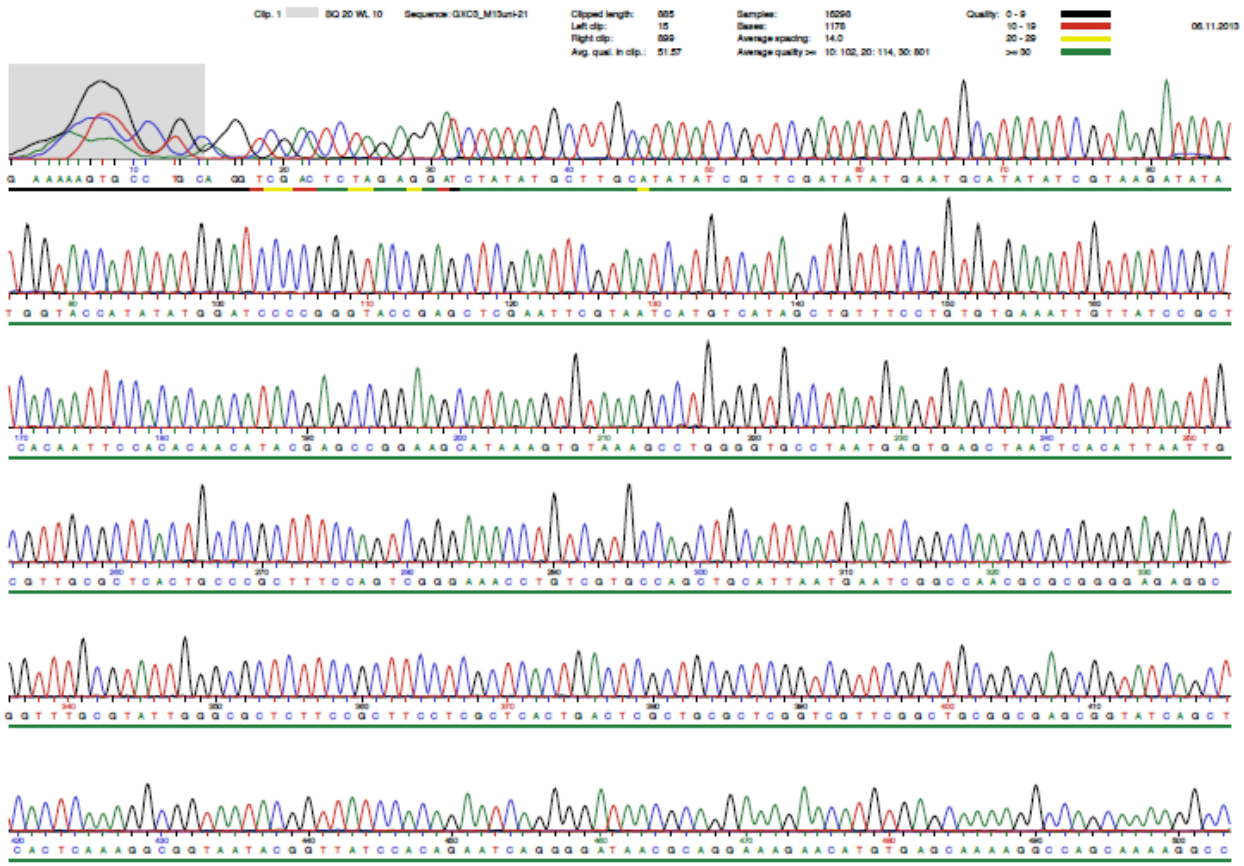


Figure A. Commercial sequencing of **GXC3**

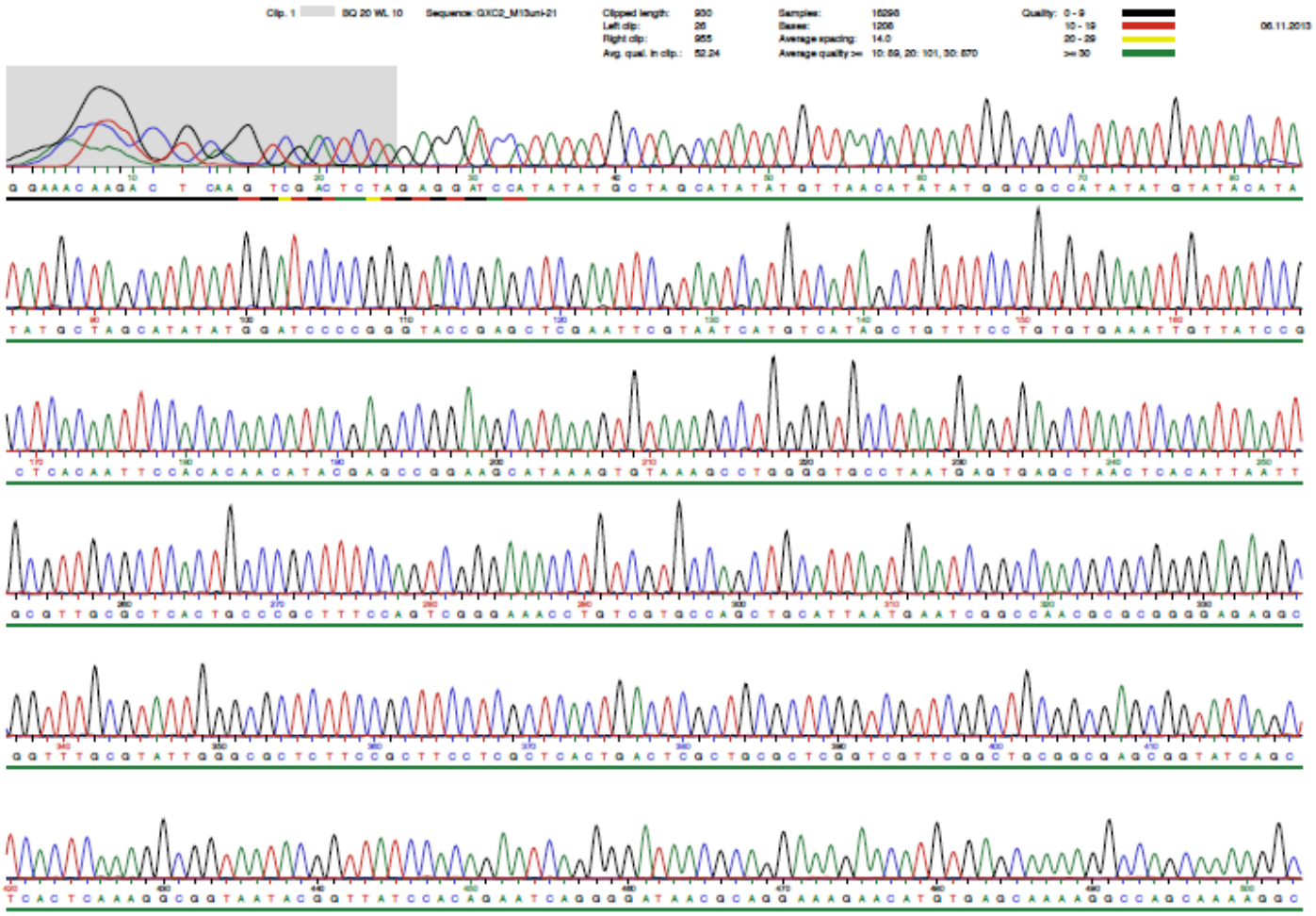


Figure B. Commercial sequencing of **GXC2**

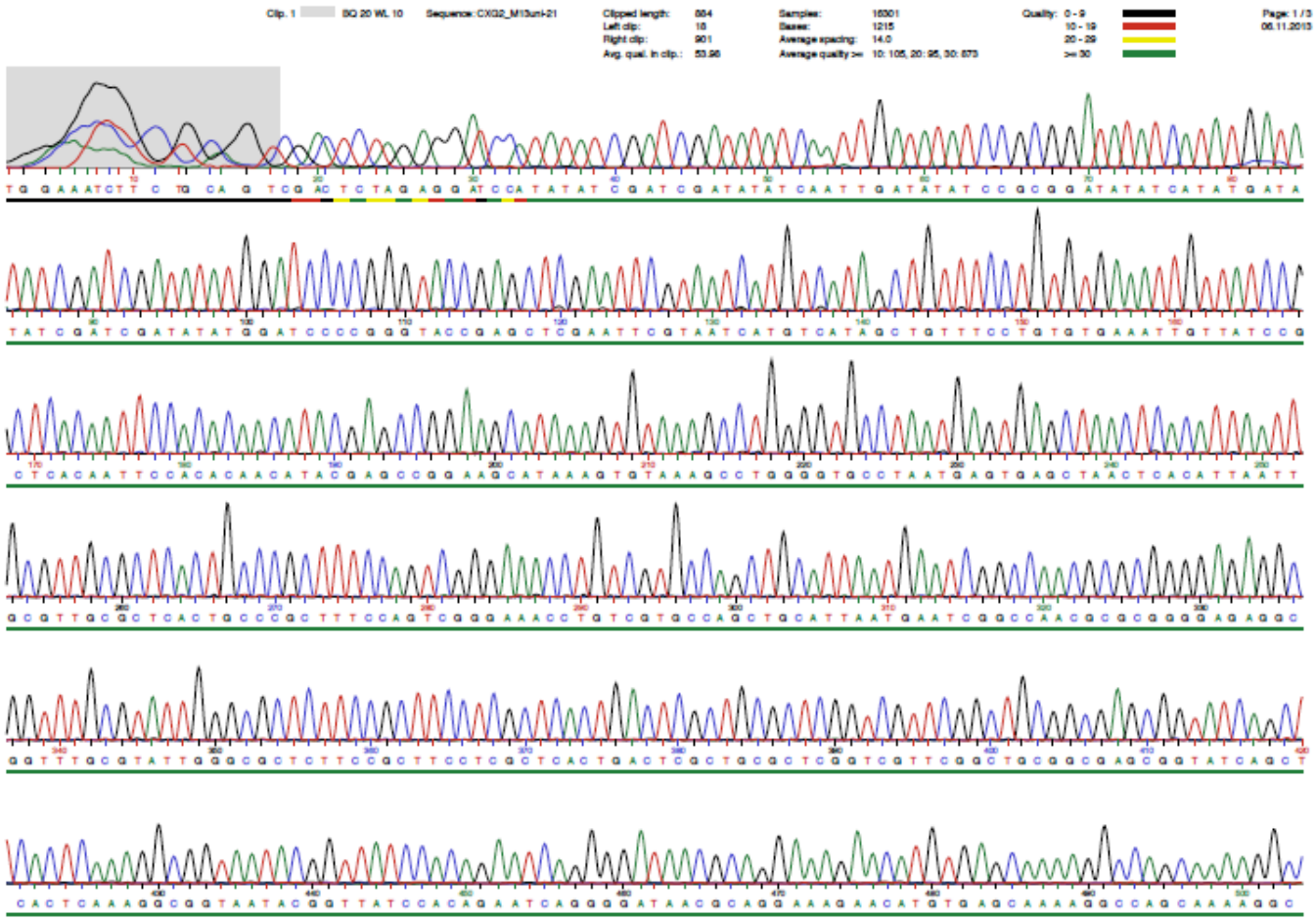


Figure C. Commercial sequencing of **CXG2**

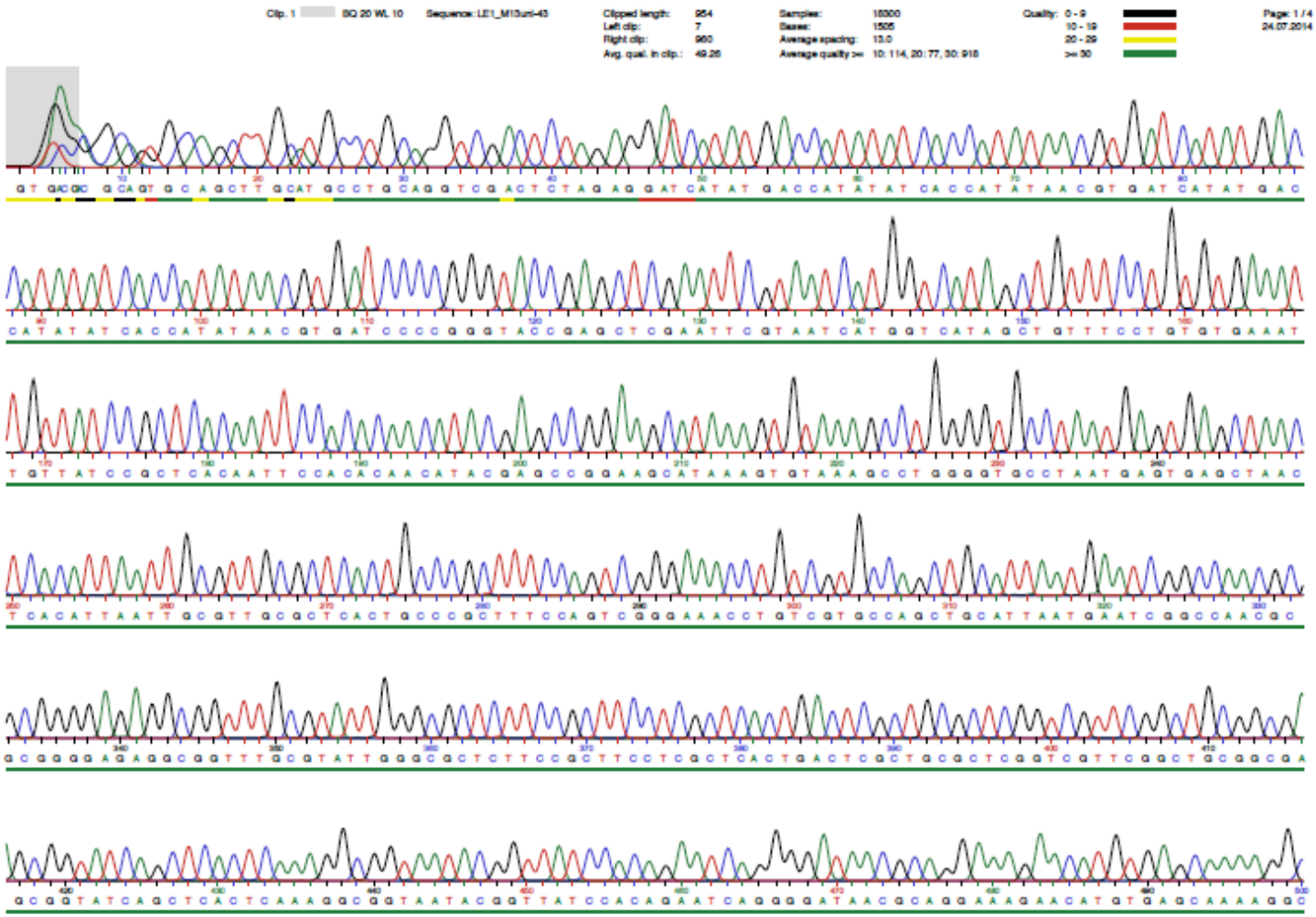


Figure D. Commercial sequencing of **LE1**

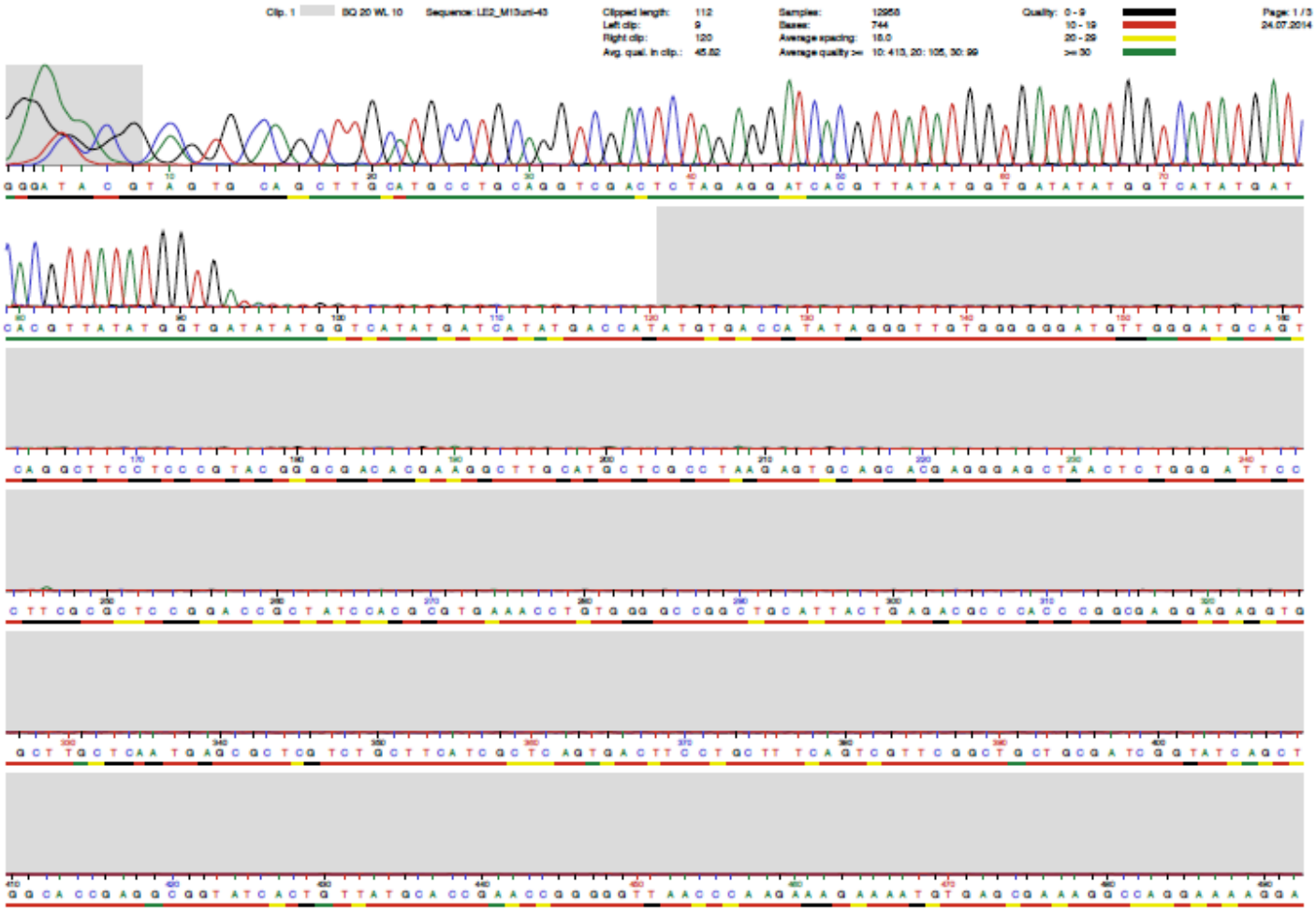


Figure E. Commercial sequencing of **LE2**

Appendix 2:

Melting transitions (T_m) calculated from the fluorescence melting studies (first and second melting curves) of PBD conjugates with pyrrole, imidazole and thiazole (Table I-VIII), and with benzofused rings (Table IX-XVIII)

Table I: Melting transitions (T_m) calculated from the fluorescence melting studies (first melting curves) of **PBD-Im-Py-Py (RMH41)** for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GAATGC and GCTTGC.

Concentration	Melting temperature (T_m) (5'-GGTACC) Control T_m = 41.4 °C		Melting temperature (T_m) (5'-GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'-GTATAC) Control T_m = 36.5 °C		Melting temperature (T_m) (5'-GTTAAC) Control T_m = 37.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	65.9,76.8*	87.0	61.9,75.0*	88.9	64.3	82.6	66.2	83.6
5 µM	65.9,76.8*	87.0	61.9,75.0*	90.0	64.0	82.6	65.2	83.6
2 µM	65.9,76.8*	87.6	75.0	90.0	63.6	83.3	64.3	83.6
1 µM	75.1	-	75.4	-	63.9	82.3	64.5	-
0.5 µM	75.1	-	75.8	-	64.2	-	64.8	-
0.1 µM	-	-	-	-	-	-	-	-

*Melting temperatures in closed proximity are placed in same column

Concentration	Melting temperature (T_m) (5'-GCTAGC) Control T_m = 43.9 °C		Melting temperature (T_m) (5'-GAATGC) Control T_m = 40.7 °C		Melting temperature (T_m) (5'-GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	64.9,73.1*	92.5	72.0	88.2	75.9	91.3
5 µM	64.8,74.0*	92.3	71.6	87.7	75.9	91.8
2 µM	75.4	92.0	72.5	87.7	76.4	91.8
1 µM	75.9	92.0	72.5	-	76.4	-
0.5 µM	76.4	-	72.7	-	76.4	-
0.1 µM	76.4	-	-	-	-	-

Table II: Melting transitions (T_m) calculated from the fluorescence melting studies (second melting curves) of **PBD-Im-Py-Py (RMH41)** for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GAATGC and GCTTGC.

Concentration	Melting temperature (T_m) (5'-GGTACC) Control T_m = 41.7 °C		Melting temperature (T_m) (5'-GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'-GTATAC) Control T_m = 36.6 °C		Melting temperature (T_m) (5'-GTTAAC) Control T_m = 38.1 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	65.7,75.8*	-	64.4,76.4*	-	64.8	-	66.5	-
5 µM	65.7,75.8*	-	64.4,76.4*	-	64.7	-	65.7	-
2 µM	64.4,76.9*	-	64.5,76.4*	-	64.7	-	65.4	-
1 µM	66.7,75.9*	-	64.8,76.4*	-	64.7	-	65.4	-
0.5 µM	66.6,75.9*	-	76.4	-	64.7	-	-	-
0.1 µM	-	-	-	-	-	-	-	-

* Melting temperatures in closed proximity are placed in same column

Concentration	Melting temperature (T_m) (5'-GCTAGC) Control T_m = 44.3 °C		Melting temperature (T_m) (5'-GAATGC) Control T_m = 40.9 °C		Melting temperature (T_m) (5'-GCTTGC) Control T_m = 44.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	74.6	89.9	68.9	-	76.1	92.0
5 µM	74.6	89.9	68.9	-	76.1	-
2 µM	76.4	-	-	-	76.5	-
1 µM	76.6	-	-	-	77.0	-
0.5 µM	76.6	-	-	-	77.0	-
0.1 µM	-	-	-	-	-	-

Table III: Melting transitions (T_m) calculated from the fluorescence melting studies (first melting curves) of **PBD-Thz-Py-Py (RMH43)** for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GAATGC and GCTTGC.

Concentration	Melting temperature (T_m) (5'-GGTACC) Control T_m = 41.4 °C		Melting temperature (T_m) (5'-GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'-GTATAC) Control T_m = 36.5 °C		Melting temperature (T_m) (5'-GTTAAC) Control T_m = 37.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	72.0	89.7	75.4	92.3	64.5	87.0	64.8	86.3
5 µM	72.0	90.3	74.8	92.4	64.6	86.8	64.4	86.8
2 µM	72.0	-	75.4	92.4	64.3	86.6	69.6	86.8
1 µM	72.5	-	75.9	-	64.3	87.0	64.4	87.0
0.5 µM	72.6	-	77.1	-	65.0	86.0	68.0	87.0
0.1 µM	73.2	-	77.6	-	-	-	68.8	-

Concentration	Melting temperature (T_m) (5'-GCTAGC) Control T_m = 43.9 °C		Melting temperature (T_m) (5'-GAATGC) Control T_m = 40.7 °C		Melting temperature (T_m) (5'-GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	73.1	92.5	72.0	90.6	73.9	92.3
5 µM	73.6	92.5	72.0	90.6	73.6	92.3
2 µM	73.8	92.5	71.8	90.6	73.9	92.3
1 µM	73.9	92.2	72.4	90.6	74.4	92.3
0.5 µM	75.9	92.2	72.9	91.3	75.0	92.3
0.1 µM	76.5	-	73.0	-	76.5	-

Table IV: Melting transitions (T_m) calculated from the fluorescence melting studies (second melting curves) of **PBD-Thz-Py-Py (RMH43)** for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GAATGC and GCTTGC.

Concentration	Melting temperature (T_m) (5'-GGTACC) Control T_m = 41.7 °C		Melting temperature (T_m) (5'-GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'-GTATAC) Control T_m = 36.6 °C		Melting temperature (T_m) (5'-GTTAAC) Control T_m = 38.1 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	64.9,72.3*	-	63.9,75.2*	-	64.8	86.8	65.6	86.9
5 µM	64.9,72.3*	-	63.9,76.2*	-	64.8	86.8	65.5	86.9
2 µM	72.5	-	63.9,76.8*	-	64.8	86.8	65.7	-
1 µM	72.9	-	76.8	-	64.8	86.8	67.0	-
0.5 µM	73.4	-	76.9	-	65.2	-	68.8	-
0.1 µM	-	-	-	-	-	-	-	-

Concentration	Melting temperature (T_m) (5'-GCTAGC) Control T_m = 44.3 °C		Melting temperature (T_m) (5'-GAATGC) Control T_m = 40.9 °C		Melting temperature (T_m) (5'-GCTTGC) Control T_m = 44.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	75.3	89.9	72.5	89.9	75.0	89.6
5 µM	75.5	89.9	73.0	89.9	75.0	89.6
2 µM	76.1	-	72.6	89.9	75.6	-
1 µM	76.1	-	72.6	-	75.6	-
0.5 µM	76.8	-	72.6	-	75.9	-
0.1 µM	76.8	-	-	-	76.0	-

Table V: Melting transitions (T_m) calculated from the fluorescence melting studies (first melting curves) of **PBD-Py-Thz-Py (RMH47)** for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GAATGC and GCTTGC.

Concentration	Melting temperature (T_m) (5'-GGTACC) Control T_m = 41.4 °C		Melting temperature(T_m) (5'-GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'-GTATAC) Control T_m = 36.5 °C		Melting temperature (T_m) (5'-GTTAAC) Control T_m = 37.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	63.7	89.0	70.7	91.6	62.4	84.0	64.3	84.0
5 µM	64.0	89.0	70.7	92.0	62.4	84.0	64.3	84.0
2 µM	63.4	89.2	70.7	92.0	62.4	84.0	64.3	84.0
1 µM	63.4	89.3	70.7	92.0	61.7	84.0	63.5	84.0
0.5 µM	61.7	89.3	72.0	92.0	61.8	84.0	63.5	84.0
0.1 µM	68.9	-	72.7	-	62.3	-	63.6	84.0

Concentration	Melting temperature (T_m) (5'-GCTAGC) Control T_m = 43.9 °C		Melting temperature(T_m) (5'-GAATGC) Control T_m = 40.7 °C		Melting temperature (T_m) (5'-GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	69.1	91.8	75.6	92.0	75.0	92.0
5 µM	68.7	91.8	75.6	92.0	74.0	92.0
2 µM	67.8	91.8	75.6	92.0	73.5	92.0
1 µM	66.3	91.5	75.6	92.0	73.5	92.0
0.5 µM	64.5	91.7	76.0	92.0	73.5	92.0
0.1 µM	69.0	-	76.2	-	72.6	-

Table VI: Melting transitions (T_m) calculated from the fluorescence melting studies (second melting curves) of **PBD-Py-Thz-Py (RMH47)** for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GAATGC and GCTTGC.

Concentration	Melting temperature (T_m) (5'-GGTACC) Control T_m = 41.7 °C		Melting temperature (T_m) (5'-GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'-GTATAC) Control T_m = 36.6 °C		Melting temperature (T_m) (5'-GTTAAC) Control T_m = 38.1 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	64.5	-	64.3,73.2*	-	64.5	83.7	65.6	-
5 µM	64.5	-	64.1,71.9*	-	64.1	84.0	65.1	-
2 µM	63.2	-	62.5,72.5*	-	63.6	-	64.9	-
1 µM	63.2	-	62.5,72.7*	-	63.5	-	64.7	-
0.5 µM	63.2	-	62.0,72.7*	-	63.2	-	64.5	-
0.1 µM	-	-	73.6	-	63.4	-	64.5	-

Concentration	Melting temperature (T_m) (5'-GCTAGC) Control T_m = 44.3 °C		Melting temperature(T_m) (5'-GAATGC) Control T_m = 40.9 °C		Melting temperature (T_m) (5'-GCTTGC) Control T_m = 44.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	70.0	90.0	75.7	-	74.3	92.0
5 µM	69.5	90.0	75.7	-	74.0	92.0
2 µM	72.4	90.0	75.7	-	73.4	-
1 µM	72.4	-	76.2	-	73.2	-
0.5 µM	73.4	-	76.1	-	72.7	-
0.1 µM	73.0	-	76.1	-	73.0	-

Table VII: Melting transitions (T_m) calculated from the fluorescence melting studies (first melting curves) of **PBD-Py-Im-Im (RMH53)** for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GAATGC and GCTTGC.

Concentration	Melting temperature (T_m) (5'-GGTACC) Control T_m = 41.4 °C		Melting temperature (T_m) (5'-GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'-GTATAC) Control T_m = 36.5 °C		Melting temperature (T_m) (5'-GTTAAC) Control T_m = 37.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	67.4	88.8	71.0	91.4	63.3	84.5	65.3	83.9
5 µM	67.9	88.8	70.5	91.8	63.6	84.7	65.3	84.3
2 µM	67.9	89.6	71.5	92.0	63.9	84.3	65.3	84.3
1 µM	68.7	89.6	71.4	92.6	64.1	84.3	65.7	84.2
0.5 µM	69.5	-	71.8	-	64.3	84.0	65.1	84.5
0.1 µM	69.5	-	-	-	-	-	-	-

Concentration	Melting temperature (T_m) (5'-GCTAGC) Control T_m = 43.9 °C		Melting temperature (T_m) (5'-GAATGC) Control T_m = 40.7 °C		Melting temperature (T_m) (5'-GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	71.3	92.6	74.7	88.8	72.2	92.1
5 µM	71.1	92.3	74.7	88.8	73.2	92.1
2 µM	71.1	92.5	75.3	89.1	71.9	92.0
1 µM	72.0	92.2	75.7	89.7	71.9	92.0
0.5 µM	72.6	-	76.2	-	72.9	92.0
0.1 µM	-	-	-	-	-	-

Table VIII: Melting transitions (T_m) calculated from the fluorescence melting studies (second melting curves) of **PBD-Py-Im-Im (RMH53)** for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GAATGC and GCTTGC.

Concentration	Melting temperature (T_m) (5'-GGTACC) Control T_m = 41.7 °C		Melting temperature (T_m) (5'-GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'-GTATAC) Control T_m = 36.6 °C		Melting temperature (T_m) (5'-GTTAAC) Control T_m = 38.1 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	69.5	89.1	72.1	90.0	64.2	84.0	66.2	84.5
5 µM	69.5	-	72.1	90.0	64.2	84.0	66.2	-
2 µM	69.5	-	72.4	-	64.2	-	66.2	-
1 µM	69.5	-	72.4	-	64.2	-	66.2	-
0.5 µM	-	-	72.4	-	64.2	-	66.2	-
0.1 µM	-	-	-	-	-	-	-	-

Concentration	Melting temperature (T_m) (5'-GCTAGC) Control T_m = 44.3 °C		Melting temperature (T_m) (5'-GAATGC) Control T_m = 40.9 °C		Melting temperature (T_m) (5'-GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	71.6	90.0	75.8	88.7	72.6	89.5
5 µM	72.7	90.0	76.0	-	72.8	89.5
2 µM	73.0	-	76.2	-	72.8	-
1 µM	73.0	-	76.2	-	72.8	-
0.5 µM	73.0	-	76.2	-	72.8	-
0.1 µM	-	-	-	-	-	-

Table IX: Melting transitions (T_m) calculated from the fluorescence melting studies (first melting curves) of **PBD-Py-Benzofuran (KMR31**) for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GCTTGC and GAATGC.

Concentration	Melting temperature (T_m) (5'-GGTACC) Control T_m = 41.4 °C		Melting temperature (T_m) (5'-GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'-GTATAC) Control T_m = 36.5 °C		Melting temperature (T_m) (5'-GTTAAC) Control T_m = 37.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10 µM	67.6	88.2	69.5	90.4	-	85.5	63.4	85.6
5 µM	67.9	88.8	69.7	90.6	65.2	85.5	63.1	85.6
2 µM	67.9	88.3	70.6	90.6	64.0	86.1	62.8	86.5
1 µM	69.2	88.3	71.3	90.6	64.0	85.4	63.3	86.5
0.5 µM	69.9	-	71.6	90.6	63.9	85.4	64.0	86.5
0.1 µM	70.2	-	72.5	-	64.5	-	64.0	-

Concentration	Melting temperature (T_m) (5'GCTAGC) Control T_m = 43.9 °C		Melting temperature (T_m) (5'GAATGC) Control T_m = 40.7 °C		Melting temperature (T_m) (5'GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10µM	72.1	90.7	73.0	91.4	72.3	91.4
5µM	72.1	91.2	73.4	91.4	73.5	91.8
2µM	72.1	91.1	73.5	90.9	73.3	91.8
1µM	72.6	91.3	73.9	90.9	73.3	91.8
0.5µM	72.9	91.7	73.9	90.9	73.1	91.8
0.1µM	73.3	-	73.9	-	73.4	-

Table X: Melting transitions (T_m) calculated from the fluorescence melting studies (second melting curves) of **PBD-Py-Benzofuran (KMR31**) for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GCTTGC and GAATGC.

Concentration	Melting temperature (T_m) (5'GGTACC) Control T_m = 41.7 °C		Melting temperature (T_m) (5'GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'GTATAC) Control T_m = 36.6 °C		Melting temperature (T_m) (5'GTTAAC) Control T_m = 38.1 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	69.3	-	71.4	-	63.7	85.8	63.9	86.1
5μM	69.6	-	71.8	-	63.6	85.8	64.2	-
2μM	70.1	-	71.8	-	64.2	-	64.2	-
1μM	-	-	71.8	-	-	-	-	-
0.5μM	-	-	71.8	-	-	-	-	-
0.1μM	-	-	-	-	-	-	-	-

Concentration	Melting temperature (T_m) (5'GCTAGC) Control T_m = 44.3 °C		Melting temperature (T_m) (5'GAATGC) Control T_m = 40.9 °C		Melting temperature (T_m) (5'GCTTGC) Control T_m = 44.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	73.1	90.0	74.1	-	73.4	89.8
5μM	73.1	90.0	74.0	-	73.4	89.8
2μM	73.6	-	74.2	-	73.9	-
1μM	73.6	-	74.2	-	73.9	-
0.5μM	73.6	-	74.2	-	73.9	-
0.1μM	-	-	-	-	-	-

Table XI: Melting transitions (T_m) calculated from the fluorescence melting studies (first melting curves) of **PBD-Py-Benzothiophene (KMR32**) for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GCTTGC and GAATGC.

Concentration	Melting temperature (T_m) (5'GGTACC) Control T_m = 41.4 °C		Melting temperature (T_m) (5'GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'GTATAC) Control T_m = 36.5 °C		Melting temperature (T_m) (5'GTTAAC) Control T_m = 37.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	68.2	89.0	68.5	90.1	65.3	85.9	63.8	85.6
5μM	68.7	89.0	70.8	90.1	65.3	85.9	63.8	85.6
2μM	68.9	88.7	71.5	90.5	65.3	85.9	64.1	85.6
1μM	70.1	-	71.3	90.5	64.5	85.6	64.1	85.6
0.5μM	70.4	-	72.1	90.5	64.8	85.6	64.1	85.6
0.1μM	70.4	-	72.6	-	64.8	-	64.1	-

Concentration	Melting temperature (T_m) (5'GCTAGC) Control T_m = 43.9 °C		Melting temperature (T_m) (5'GAATGC) Control T_m = 40.7 °C		Melting temperature (T_m) (5'GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	71.9	91.5	73.5	91.1	73.3	91.1
5μM	72.5	91.5	73.5	91.1	73.1	91.9
2μM	72.5	91.7	73.5	91.0	73.3	91.9
1μM	73.1	91.6	73.7	91.0	73.6	91.9
0.5μM	73.2	91.7	73.9	90.0	73.8	91.9
0.1μM	73.6	91.7	74.3	-	74.4	-

Table XII: Melting transitions (T_m) calculated from the fluorescence melting studies (second melting curves) of **PBD-Py-Benzothiophene (KMR32**) for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GCTTGC and GAATGC.

Concentration	Melting temperature (T_m) (5'GGTACC) Control T_m = 41.7 °C		Melting temperature (T_m) (5'GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'GTATAC) Control T_m = 36.6 °C		Melting temperature (T_m) (5'GTTAAC) Control T_m = 38.1 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	70.1	-	71.5	89.7	63.7	86.1	63.9	-
5μM	70.6	-	71.8	-	63.7	-	64.2	-
2μM	70.6	-	72.5	-	64.4	-	64.2	-
1μM	70.6	-	72.5	-	64.6	-	64.2	-
0.5μM	70.6	-	72.5	-	-	-	-	-
0.1μM	-	-	-	-	-	-	-	-

Concentration	Melting temperature (T_m) (5'GCTAGC) Control T_m = 44.3 °C		Melting temperature (T_m) (5'GAATGC) Control T_m = 40.9 °C		Melting temperature (T_m) (5'GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	73.0	89.4	74.0	-	73.4	89.2
5μM	73.5	89.4	74.0	-	74.0	-
2μM	73.9	-	74.3	-	74.4	-
1μM	73.9	-	74.3	-	74.4	-
0.5μM	74.1	-	74.3	-	74.4	-
0.1μM	-	-	-	-	-	-

Table XIII: Melting transitions (T_m) calculated from the fluorescence melting studies (first melting curves) of **PBD-Py-Benzothiophene-Acetate (KMR33**) for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GCTTGC and GAATGC.

Concentration	Melting temperature (T_m) (5'GGTACC) Control T_m = 41.4 °C		Melting temperature (T_m) (5'GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'GTATAC) Control T_m = 36.5 °C		Melting temperature (T_m) (5'GTTACC) Control T_m = 37.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	69.5	89.5	71.5	92.0	-	86.9	64.9	87.0
5μM	69.5	89.5	71.5	92.0	-	86.9	64.9	87.0
2μM	69.5	89.5	71.9	92.0	65.0	86.9	65.1	87.0
1μM	69.5	89.5	71.9	92.0	65.3	86.9	65.1	87.0
0.5μM	71.4	89.5	73.1	92.1	65.3	87.2	65.1	87.0
0.1μM	71.5	-	74.7	-	65.3	-	65.7	-

Concentration	Melting temperature (T_m) (5'GCTAGC) Control T_m = 43.9 °C		Melting temperature (T_m) (5'GAATGC) Control T_m = 40.7 °C		Melting temperature (T_m) (5'GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	73.3	92.3	73.1	91.2	72.8	92.1
5μM	73.3	92.3	73.1	91.2	72.8	92.1
2μM	73.3	92.3	73.4	90.7	74.4	92.1
1μM	73.3	92.3	73.4	90.2	73.8	92.1
0.5μM	74.0	92.3	74.3	90.2	74.2	92.1
0.1μM	74.8	-	74.1	-	75.7	-

Table XIV: Melting transitions (T_m) calculated from the fluorescence melting studies (second melting curves) of **PBD-Py-Benzothiophene-Acetate (KMR33**) for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GCTTGC and GAATGC.

Concentration	Melting temperature (T_m) (5'GGTACC) Control T_m = 41.7 °C		Melting temperature (T_m) (5'GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'GTATAC) Control T_m = 36.6 °C		Melting temperature (T_m) (5'GTTACC) Control T_m = 37.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	71.1	89.6	73.0	89.7	64.5	86.7	66.0	-
5μM	71.5	-	73.5	89.7	64.2	86.7	65.7	-
2μM	72.0	-	74.0	-	64.2	-	65.9	-
1μM	72.1	-	74.1	-	63.7	-	65.9	-
0.5μM	72.1	-	74.2	-	63.7	-	65.9	-
0.1μM	-	-	-	-	-	-	-	-

Concentration	Melting temperature (T_m) (5'GCTAGC) Control T_m = 44.3 °C		Melting temperature (T_m) (5'GAATGC) Control T_m = 40.9 °C		Melting temperature (T_m) (5'GCTTGC) Control T_m = 44.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	74.5	89.7	73.9	89.9	74.6	89.6
5μM	74.5	89.7	73.9	89.9	74.8	89.6
2μM	75.0	-	73.9	-	75.3	-
1μM	75.1	-	73.9	-	75.3	-
0.5μM	75.3	-	73.9	-	75.8	-
0.1μM	-	-	-	-	-	-

Table XV: Melting transitions (T_m) calculated from the fluorescence melting studies (first melting curves) of **PBD-Im-Benzothiophene-Acetate (KMR173**) for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GCTTGC and GAATGC.

Concentration	Melting temperature (T_m) (5'GGTACC) Control T_m = 41.4 °C		Melting temperature (T_m) (5'GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'GTATAC) Control T_m = 36.5 °C		Melting temperature (T_m) (5'GTTAAC) Control T_m = 37.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	67.5	-	65.4	87.4	63.7	82.9	63.0	87.3
5μM	67.5	-	65.4	87.4	63.7	82.9	61.8	88.1
2μM	66.9	-	64.6	87.8	62.9	82.0	61.4	88.1
1μM	65.6	-	71.9	87.8	62.9	82.8	61.9	87.7
0.5μM	65.6	-	72.0	-	62.6	82.3	63.4	88.1
0.1μM	65.6	-	72.0	-	63.3	-	62.9	-

Concentration	Melting temperature (T_m) (5'GCTAGC) Control T_m = 43.9 °C		Melting temperature (T_m) (5'GAATGC) Control T_m = 40.7 °C		Melting temperature (T_m) (5'GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	72.1	89.0	73.4	88.9	76.4	90.5
5μM	72.1	89.0	73.0	88.9	76.4	90.5
2μM	71.8	89.0	72.8	-	75.5	90.5
1μM	71.8	-	72.8	-	75.5	-
0.5μM	71.8	-	72.1	-	74.5	-
0.1μM	71.8	-	72.1	-	74.5	-

Table XVI: Melting transitions (T_m) calculated from the fluorescence melting studies (second melting curves) of **PBD-Im-Benzothiophene-Acetate (KMR173**) for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GCTTGC and GAATGC.

Concentration	Melting temperature (T_m) (5'GGTACC) Control T_m = 41.7 °C		Melting temperature (T_m) (5'GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'GTATAC) Control T_m = 36.6 °C		Melting temperature (T_m) (5'GTTAAC) Control T_m = 38.1 °C	
	1^{st} T_m	2^{nd} T_m	1^{st} T_m	2^{nd} T_m	1^{st} T_m	2^{nd} T_m	1^{st} T_m	2^{nd} T_m
10μM	67.8	-	66.7	-	62.1	-	-	-
5μM	66.7	-	65.8	-	62.1	-	-	-
2μM	66.4	-	64.8	-	-	-	-	-
1μM	65.9	-	64.8	-	-	-	-	-
0.5μM	65.9	-	65.1	-	-	-	-	-
0.1μM	-	-	-	-	-	-	-	-

Concentration	Melting temperature (T_m) (5'GCTAGC) Control T_m = 44.3 °C		Melting temperature (T_m) (5'GAATGC) Control T_m = 40.9 °C		Melting temperature (T_m) (5'GCTTGC) Control T_m = 44.8 °C	
	1^{st} T_m	2^{nd} T_m	1^{st} T_m	2^{nd} T_m	1^{st} T_m	2^{nd} T_m
10μM	73.0	-	73.8	-	76.8	-
5μM	72.5	-	73.3	-	76.4	-
2μM	72.1	-	73.3	-	75.9	-
1μM	72.1	-	72.9	-	75.9	-
0.5μM	72.7	-	72.9	-	75.9	-
0.1μM	-	-	-	-	-	-

Table XVII: Melting transitions (T_m) calculated from the fluorescence melting studies (first melting curves) of **PBD-Benzothiophene (KMR175**) for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GCTTGC and GAATGC.

Concentration	Melting temperature (T_m) (5'GGTACC) Control T_m = 41.4 °C		Melting temperature (T_m) (5'GGATCC) Control T_m = 42.3 °C		Melting temperature (T_m) (5'GTATAC) Control T_m = 36.5 °C		Melting temperature (T_m) (5'GTTAAC) Control T_m = 37.8 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	62.0	79.1	64.0	80.7	57.2	74.1	58.1	74.6
5μM	62.8	79.1	64.2	80.7	57.2	74.1	58.1	74.6
2μM	62.8	79.1	64.6	81.4	57.9	74.1	57.8	74.6
1μM	63.7	79.1	64.6	81.4	57.3	74.6	57.8	74.6
0.5μM	63.6	-	64.6	81.4	57.3	74.6	58.1	74.6
0.1μM	63.6	-	66.0	-	58.1	-	59.4	-

Concentration	Melting temperature (T_m) (5'GCTAGC) Control T_m = 43.9 °C		Melting temperature (T_m) (5'GAATGC) Control T_m = 40.7 °C		Melting temperature (T_m) (5'GCTTGC) Control T_m = 44.6 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	65.8	82.8	67.4	84.9	66.5	82.9
5μM	65.7	82.5	66.5	85.0	66.5	82.9
2μM	65.9	82.5	67.0	85.0	67.8	83.1
1μM	67.0	82.5	67.4	84.9	67.4	83.1
0.5μM	66.5	82.5	67.4	84.9	67.1	83.3
0.1μM	66.5	-	67.4	-	68.1	-

Table XVIII: Melting transitions (T_m) calculated from the fluorescence melting studies (second melting curves) of **PBD-Benzothiophene (KMR175**) for oligonucleotide duplexes that contain the sequences GGTACC, GGATCC, GTATAC, GTTAAC, GCTAGC, GCTTGC and GAATGC.

Concentration	Melting temperature (T_m) (5'GGTACC) Control T_m = 42.1 °C		Melting temperature (T_m) (5'GGATCC) Control T_m = 42.5 °C		Melting temperature (T_m) (5'GTATAC) Control T_m = 36.6 °C		Melting temperature (T_m) (5'GTTAAC) Control T_m = 37.7 °C	
	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m	1st T_m	2nd T_m
10μM	-	-	57.9,65.2*	-	-	-	-	-
5μM	-	-	58.6,65.3*	-	-	-	-	-
2μM	-	-	-	-	-	-	-	-
1μM	-	-	-	-	-	-	-	-
0.5μM	-	-	-	-	-	-	-	-
0.1μM	-	-	-	-	-	-	-	-

Concentration	Melting temperature (T_m) (5'GCTAGC) Control T_m = 44.1 °C		Melting temperature (T_m) (5'GAATGC) Control T_m = 41.1 °C		Melting temperature (T_m) (5'GCTTGC) Control T_m = 45.0 °C	
	1st T_m	2nd T_m	1st Tm	2nd T_m	1st T_m	2nd T_m
10μM	66.7	-	67.7	-	67.4	-
5μM	66.7	-	67.7	-	67.4	-
2μM	-	-	-	-	-	-
1μM	-	-	-	-	-	-
0.5μM	-	-	-	-	-	-
0.1μM	-	-	-	-	-	-

References

1. (2014) *World Cancer Report 2014*. International Agency for Research on Cancer, WHO.
2. Scagliotti, G.V. and Selvaggi, G. (2006) Antimetabolites and cancer: emerging data with a focus on antifolates. *Expert Opin Ther Pat*, **16**, 189-200.
3. Pommier, Y., Pourquier, P., Fan, Y. and Strumberg, D. (1998) Mechanism of action of eukaryotic DNA topoisomerase I and drugs targeted to the enzyme. *Biochim Biophys Acta*, **1400**, 83-105.
4. Burden, D.A. and Osheroff, N. (1998) Mechanism of action of eukaryotic topoisomerase II and drugs targeted to the enzyme. *Biochim Biophys Acta*, **1400**, 139-154.
5. Hortobagyi, G.N. (1997) Anthracyclines in the treatment of cancer. An overview. *Drugs*, **54 Suppl 4**, 1-7.
6. Hollstein, U. (1974) Actinomycin. Chemistry and mechanism of action. *Chemical reviews*, **74**, 625-652.
7. Chen, J. and Stubbe, J. (2005) Bleomycins: towards better therapeutics. *Nature Reviews Cancer*, **5**, 102-112.
8. Bradner, W.T. (2001) Mitomycin C: a clinical update. *Cancer Treat Rev*, **27**, 35-50.
9. Brion, J.-D. (2003) *Agent alkylants: Generalites Medicaments antitumoraux et perspectives dans le traitment des cancers*, Paris: TEC & DOC.
10. Holland, J.F. and Frei, E. (2003) *Holland-Frei cancer medicine*. 6 ed. Hamilton, ON: BC Decker Inc.
11. Beranek, D.T. (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutat Res*, **231**, 11-30.
12. Drablos, F., Feyzi, E., Aas, P.A., Vaagbo, C.B., Kavli, B., Bratlie, M.S., Pena-Diaz, J., Otterlei, M., Slupphaug, G. and Krokan, H.E. (2004) Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair (Amst)*, **3**, 1389-1407.
13. Shrivastav, N., Li, D. and Essigmann, J.M. (2010) Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. *Carcinogenesis*, **31**, 59-70.
14. Pullman, A. and Pullman, B. (1981) Molecular electrostatic potential of the nucleic acids. *Q Rev Biophys*, **14**, 289-380.
15. Shooter, K.V., Howse, R., Shah, S.A. and Lawley, P.D. (1974) The molecular basis for biological inactivation of nucleic acids. The action of methylating agents on the ribonucleic acid-containing bacteriophage R17. *Biochem J*, **137**, 303-312.
16. Mishina, Y., Duguid, E.M. and He, C. (2006) Direct reversal of DNA alkylation damage. *Chem Rev*, **106**, 215-232.
17. Gates, K.S. (2009) An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem Res Toxicol*, **22**, 1747-1760.
18. Margison, G.P., Santibanez Koref, M.F. and Povey, A.C. (2002) Mechanisms of carcinogenicity/chemotherapy by O6-methylguanine. *Mutagenesis*, **17**, 483-487.
19. Puyo, S., Montaudon, D. and Pourquier, P. (2014) From old alkylating agents to new minor groove binders. *Crit Rev Oncol Hematol*, **89**, 43-61.
20. Povirk, L.F. and Shuker, D.E. (1994) DNA damage and mutagenesis induced by nitrogen mustards. *Mutat Res*, **318**, 205-226.

21. Gobin, Y.P., Dunkel, I.J., Marr, B.P., Brodie, S.E. and Abramson, D.H. (2011) Intra-arterial chemotherapy for the management of retinoblastoma: four-year experience. *Arch Ophthalmol*, **129**, 732-737.
22. Rai, K.R., Peterson, B.L., Appelbaum, F.R., Kolitz, J., Elias, L., Shepherd, L., Hines, J., Threatte, G.A., Larson, R.A., Cheson, B.D. *et al.* (2000) Fludarabine compared with chlorambucil as primary therapy for chronic lymphocytic leukemia. *N Engl J Med*, **343**, 1750-1757.
23. Giraud, B., Hebert, G., Deroussent, A., Veal, G.J., Vassal, G. and Paci, A. (2010) Oxazaphosphorines: new therapeutic strategies for an old class of drugs. *Expert Opin Drug Metab Toxicol*, **6**, 919-938.
24. Walter, R.B., Appelbaum, F.R., Estey, E.H. and Bernstein, I.D. (2012) Acute myeloid leukemia stem cells and CD33-targeted immunotherapy. *Blood*, **119**, 6198-6208.
25. Sukhodub, L.F., Shelkovsky, V.S., Kosevich, M.V., Pyatigorskaya, T.L. and Zhilkova, O.Y. (1986) Nucleic acid base complexes with thioTEPA as revealed by field ionisation mass spectrometry. *Biomed Environm Mass Spectrom*, **13**, 167-170.
26. Musser, S.M., Callery, P.S. and Pan, S. (1988) Identification of the principal adduct formed between DNA and 1,1',1''-phophinothioy-lidinetrisaziridine (thioTEPA). *Proc Am Ass Cancer Res*, **29**, 267.
27. Cohen, N.A., Egorin, M.J., Snyder, S.W., Ashar, B., Wietharn, B.E., Pan, S.S., Ross, D.D. and Hilton, J. (1991) Interaction of N,N',N''-triethylenethiophosphoramide and N,N',N''-triethylenephosphoramide with cellular DNA. *Cancer Res*, **51**, 4360-4366.
28. Teicher, B.A., Holden, S.A., Cucchi, C.A., Cathcart, K.N., Korbut, T.T., Flatow, J.L. and Frei, E., 3rd. (1988) Combination of N,N',N''-triethylenethiophosphoramide and cyclophosphamide in vitro and in vivo. *Cancer Res*, **48**, 94-100.
29. Carter, S.K. and Crooke, S.T. (1979) *Mitomycin C; Current status and new developments*. Academic Press, New York.
30. Franck, R.W. and Tomasz, M. (1990) *Chemistry of Antitumor agents*. Blackie and Son Ltd, Glasgow.
31. Remers, W.A. and Dorr, R.T. (1988) *Alkaloids: Chemical and Biological Perspectives*. Wiley, New York.
32. Andrews, P.A., Pan, S. and Bachur, N.R. (1986) Electrochemical reductive activation of mitomycin C. *Journal of the American Chemical Society*, **108**, 4158-4166.
33. Cowan, A.J., Laszlo, G.S., Estey, E.H. and Walter, R.B. (2013) Antibody-based therapy of acute myeloid leukemia with gemtuzumab ozogamicin. *Front Biosci (Landmark Ed)*, **18**, 1311-1334.
34. Brookes, P. and Lawley, P.D. (1961) The reaction of mono- and di-functional alkylating agents with nucleic acids. *Biochem J*, **80**, 496-503.
35. Tong, W.P. and Ludlum, D.B. (1980) Crosslinking of DNA by busulfan. Formation of diguanyl derivatives. *Biochim Biophys Acta*, **608**, 174-181.
36. Colvin, M. and Chabner, B.A. (1990) In Chabner, B. A. and Collins, J. M. (eds.), *Alkylating agents in cancer chemotherapy: principles and practice*. Lippincott, Philadelphia, pp. 276-313.
37. Audette, R.C., Connors, T.A., Mandel, H.G., Merai, K. and Ross, W.C. (1973) Studies on the mechanism of action of the tumour inhibitory triazenes. *Biochem Pharmacol*, **22**, 1855-1864.
38. Marchesi, F., Turriziani, M., Tortorelli, G., Avvisati, G., Torino, F. and De Vecchis, L. (2007) Triazene compounds: mechanism of action and related DNA repair systems. *Pharmacol Res*, **56**, 275-287.
39. Friedman, H.S., Kerby, T. and Calvert, H. (2000) Temozolomide and treatment of malignant glioma. *Clin Cancer Res*, **6**, 2585-2597.

40. Agarwala, S.S. and Kirkwood, J.M. (2000) Temozolomide, a novel alkylating agent with activity in the central nervous system, may improve the treatment of advanced metastatic melanoma. *Oncologist*, **5**, 144-151.
41. Coomes, M.W. and Prough, R.A. (1983) The mitochondrial metabolism of 1,2-disubstituted hydrazines, procarbazine and 1,2-dimethylhydrazine. *Drug Metab Dispos*, **11**, 550-555.
42. Dunn, D.L., Lubet, R.A. and Prough, R.A. (1979) Oxidative metabolism of N-isopropyl-alpha-(2-methylhydrazino)-p-toluamide hydrochloride (procarbazine) by rat liver microsomes. *Cancer Res*, **39**, 4555-4563.
43. Erikson, J.M. and Prough, R.A. (1986) Oxidative metabolism of some hydrazine derivatives by rat liver and lung tissue fractions. *J Biochem Toxicol*, **1**, 41-52.
44. Rosenberg, B., VanCamp, L., Trosko, J.E. and Mansour, V.H. (1969) Platinum compounds: a new class of potent antitumour agents. *Nature*, **222**, 385-386.
45. Rosenberg, B. and VanCamp, L. (1970) The successful regression of large solid sarcoma 180 tumors by platinum compounds. *Cancer Res*, **30**, 1799-1802.
46. Kelland, L. (2007) The resurgence of platinum-based cancer chemotherapy. *Nat Rev Cancer*, **7**, 573-584.
47. Jamieson, E.R. and Lippard, S.J. (1999) Structure, Recognition, and Processing of Cisplatin-DNA Adducts. *Chem Rev*, **99**, 2467-2498.
48. Kelland, L.R., Mistry, P., Abel, G., Loh, S.Y., O'Neill, C.F., Murrer, B.A. and Harrap, K.R. (1992) Mechanism-related circumvention of acquired cis-diamminedichloroplatinum(II) resistance using two pairs of human ovarian carcinoma cell lines by ammine/amine platinum(IV) dicarboxylates. *Cancer Res*, **52**, 3857-3864.
49. Ishida, S., Lee, J., Thiele, D.J. and Herskowitz, I. (2002) Uptake of the anticancer drug cisplatin mediated by the copper transporter Ctr1 in yeast and mammals. *Proc Natl Acad Sci U S A*, **99**, 14298-14302.
50. Katano, K., Kondo, A., Safaei, R., Holzer, A., Samimi, G., Mishima, M., Kuo, Y.M., Rochdi, M. and Howell, S.B. (2002) Acquisition of resistance to cisplatin is accompanied by changes in the cellular pharmacology of copper. *Cancer Res*, **62**, 6559-6565.
51. Holzer, A.K., Manorek, G.H. and Howell, S.B. (2006) Contribution of the major copper influx transporter CTR1 to the cellular accumulation of cisplatin, carboplatin, and oxaliplatin. *Mol Pharmacol*, **70**, 1390-1394.
52. Mistry, P., Kelland, L.R., Abel, G., Sidhar, S. and Harrap, K.R. (1991) The relationships between glutathione, glutathione-S-transferase and cytotoxicity of platinum drugs and melphalan in eight human ovarian carcinoma cell lines. *Br J Cancer*, **64**, 215-220.
53. Ishikawa, T. (1992) The ATP-dependent glutathione S-conjugate export pump. *Trends Biochem Sci*, **17**, 463-468.
54. Lewis, A.D., Hayes, J.D. and Wolf, C.R. (1988) Glutathione and glutathione-dependent enzymes in ovarian adenocarcinoma cell lines derived from a patient before and after the onset of drug resistance: intrinsic differences and cell cycle effects. *Carcinogenesis*, **9**, 1283-1287.
55. Rabik, C.A. and Dolan, M.E. (2007) Molecular mechanisms of resistance and toxicity associated with platinating agents. *Cancer Treat Rev*, **33**, 9-23.
56. Welsh, C., Day, R., McGurk, C., Masters, J.R., Wood, R.D. and Koberle, B. (2004) Reduced levels of XPA, ERCC1 and XPF DNA repair proteins in testis tumor cell lines. *Int J Cancer*, **110**, 352-361.
57. Koberle, B., Masters, J.R., Hartley, J.A. and Wood, R.D. (1999) Defective repair of cisplatin-induced DNA damage caused by reduced XPA protein in testicular germ cell tumours. *Curr Biol*, **9**, 273-276.

58. Barnham, K.J., Djuran, M.I., Murdoch, P.S., Ranford, J.D. and Sadler, P.J. (1996) Ring-Opened Adducts of the Anticancer Drug Carboplatin with Sulfur Amino Acids. *Inorg Chem*, **35**, 1065-1072.
59. Boulikas, T. and Vougiouka, M. (2003) Cisplatin and platinum drugs at the molecular level. (Review). *Oncol Rep*, **10**, 1663-1682.
60. Kasparkova, J., Vojtiskova, M., Natile, G. and Brabec, V. (2008) Unique properties of DNA interstrand cross-links of antitumor oxaliplatin and the effect of chirality of the carrier ligand. *Chemistry*, **14**, 1330-1341.
61. Takeda, Y., Ohlendorf, D.H., Anderson, W.F. and Matthews, B.W. (1983) DNA-binding proteins. *Science*, **221**, 1020-1026.
62. Kielkopf, C.L., White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E., Dervan, P.B. and Rees, D.C. (1998) A structural basis for recognition of A.T and T.A base pairs in the minor groove of B-DNA. *Science*, **282**, 111-115.
63. Pabo, C.O. and Sauer, R.T. (1984) Protein-DNA recognition. *Annu Rev Biochem*, **53**, 293-321.
64. Neidle, S. (2001) DNA minor-groove recognition by small molecules. *Nat Prod Rep*, **18**, 291-309.
65. Arcamone, F., Penco, S., Orezzi, P., Nicolella, V. and Pirelli, A. (1964) Structure and Synthesis of Distamycin A. *Nature*, **203**, 1064-1065.
66. Pelton, J.G. and Wemmer, D.E. (1988) Structural modeling of the distamycin A-d(CGCGAATTCGCG)2 complex using 2D NMR and molecular mechanics. *Biochemistry*, **27**, 8088-8096.
67. Baguley, B.C. (1982) Nonintercalative DNA-binding antitumour compounds. *Mol Cell Biochem*, **43**, 167-181.
68. Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. and Dickerson, R.E. (1985) The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc Natl Acad Sci U S A*, **82**, 1376-1380.
69. Van Dyke, M.W. and Dervan, P.B. (1983) Methidiumpropyl-EDTA.Fe(II) and DNase I footprinting report different small molecule binding site sizes on DNA. *Nucleic Acids Res*, **11**, 5555-5567.
70. Zaffaroni, N., Lualdi, S., Villa, R., Bellarosa, D., Cermele, C., Felicetti, P., Rossi, C., Orlandi, L. and Daidone, M.G. (2002) Inhibition of telomerase activity by a distamycin derivative: effects on cell proliferation and induction of apoptosis in human cancer cells. *Eur J Cancer*, **38**, 1792-1801.
71. D'Incalci, M. and Sessa, C. (1997) DNA minor groove binding ligands: a new class of anticancer agents. *Expert Opin Investig Drugs*, **6**, 875-884.
72. Cai, X., Gray, P.J., Jr. and Von Hoff, D.D. (2009) DNA minor groove binders: back in the groove. *Cancer Treat Rev*, **35**, 437-450.
73. Khan, G.S., Shah, A., Zia ur, R. and Barker, D. (2012) Chemistry of DNA minor groove binding agents. *J Photochem Photobiol B*, **115**, 105-118.
74. Dorn, A., Affolter, M., Muller, M., Gehring, W.J. and Leupin, W. (1992) Distamycin-induced inhibition of homeodomain-DNA complexes. *EMBO J*, **11**, 279-286.
75. Broggini, M., Ponti, M., Ottolenghi, S., D'Incalci, M., Mongelli, N. and Mantovani, R. (1989) Distamycins inhibit the binding of OTF-1 and NFE-1 transfactors to their conserved DNA elements. *Nucleic Acids Res*, **17**, 1051-1059.
76. Bellorini, M., Moncollin, V., D'Incalci, M., Mongelli, N. and Mantovani, R. (1995) Distamycin A and tallimustine inhibit TBP binding and basal in vitro transcription. *Nucleic Acids Res*, **23**, 1657-1663.
77. Geroni, C., Marchini, S., Cozzi, P., Galliera, E., Ragg, E., Colombo, T., Battaglia, R., Howard, M., D'Incalci, M. and Broggini, M. (2002) Brostallicin, a novel anticancer agent whose activity is enhanced upon binding to glutathione. *Cancer Res*, **62**, 2332-2336.

78. Tagliabue, G., Pifferi, A., Balconi, G., Mascellani, E., Geroni, C., D'Incalci, M. and Ubezio, P. (1993) Intracellular glutathione heterogeneity in L1210 murine leukemia sublines made resistant to DNA-interacting anti-neoplastic agents. *Int J Cancer*, **54**, 435-442.
79. Fedier, A., Fowst, C., Tursi, J., Geroni, C., Haller, U., Marchini, S. and Fink, D. (2003) Brostallicin (PNU-166196)- a new DNA minor groove binder that retains sensitivity in DNA mismatch repair-deficient tumour cells. *British journal of cancer*, **89**, 1559-1565.
80. Hanka, L.J., Dietz, A., Gerpheide, S.A., Kuentzel, S.L. and Martin, D.G. (1978) CC-1065 (NSC-298223), a new antitumor antibiotic. Production, in vitro biological activity, microbiological assays and taxonomy of the producing microorganism. *J Antibiot (Tokyo)*, **31**, 1211-1217.
81. Martin, D.G., Biles, C., Gerpheide, S.A., Hanka, L.J., Krueger, W.C., McGovren, J.P., Mizsak, S.A., Neil, G.L., Stewart, J.C. and Visser, J. (1981) CC-1065 (NSC 298223), a potent new antitumor agent improved production and isolation, characterization and antitumor activity. *J Antibiot (Tokyo)*, **34**, 1119-1125.
82. Ichimura, M., Ogawa, T., Katsumata, S., Takahashi, K., Takahashi, I. and Nakano, H. (1991) Duocarmycins, new antitumor antibiotics produced by Streptomyces; producing organisms and improved production. *J Antibiot (Tokyo)*, **44**, 1045-1053.
83. Takahashi, I., Takahashi, K., Ichimura, M., Morimoto, M., Asano, K., Kawamoto, I., Tomita, F. and Nakano, H. (1988) Duocarmycin A, a new antitumor antibiotic from Streptomyces. *J Antibiot (Tokyo)*, **41**, 1915-1917.
84. Baraldi, P.G., Bovero, A., Fruttarolo, F., Preti, D., Tabrizi, M.A., Pavani, M.G. and Romagnoli, R. (2004) DNA minor groove binders as potential antitumor and antimicrobial agents. *Med Res Rev*, **24**, 475-528.
85. Hopton, S.R. and Thompson, A.S. (2011) Manipulative interplay of two adozelesin molecules with d(ATTAAT)(2)achieving ligand-stacked Watson-Crick and Hoogsteen base-paired duplex adducts. *Biochemistry*, **50**, 4143-4154.
86. Ding, Z.M. and Hurley, L.H. (1991) DNA interstrand cross-linking, DNA sequence specificity, and induced conformational changes produced by a dimeric analog of (+)-CC-1065. *Anticancer Drug Des*, **6**, 427-452.
87. Seaman, F.C. and Hurley, L. (1993) Interstrand cross-linking by bizelesin produces a Watson-Crick to Hoogsten base-pairing transition region in d(CGTATTACG)2. *Biochemistry*, **32**, 12577-12585.
88. Pommier, Y., Kohlhagen, G., Bailly, C., Waring, M., Mazumder, A. and Kohn, K.W. (1996) DNA sequence- and structure-selective alkylation of guanine N2 in the DNA minor groove by ecteinascidin 743, a potent antitumor compound from the Caribbean tunicate Ecteinascidia turbinata. *Biochemistry*, **35**, 13303-13309.
89. Bonfanti, M., La Valle, E., Fernandez Sousa Faro, J.M., Faircloth, G., Caretti, G., Mantovani, R. and D'Incalci, M. (1999) Effect of ecteinascidin-743 on the interaction between DNA binding proteins and DNA. *Anticancer Drug Des*, **14**, 179-186.
90. Jin, S., Gorfajn, B., Faircloth, G. and Scotto, K.W. (2000) Ecteinascidin 743, a transcription-targeted chemotherapeutic that inhibits MDR1 activation. *Proc Natl Acad Sci U S A*, **97**, 6775-6779.
91. Minuzzo, M., Marchini, S., Broggini, M., Faircloth, G., D'Incalci, M. and Mantovani, R. (2000) Interference of transcriptional activation by the antineoplastic drug ecteinascidin-743. *Proc Natl Acad Sci U S A*, **97**, 6780-6784.
92. Erba, E., Bergamaschi, D., Bassano, L., Damia, G., Ronzoni, S., Faircloth, G.T. and D'Incalci, M. (2001) Ecteinascidin-743 (ET-743), a natural marine compound, with a unique mechanism of action. *Eur J Cancer*, **37**, 97-105.

93. Valoti, G., Nicoletti, M.I., Pellegrino, A., Jimeno, J., Hendriks, H., D'Incalci, M., Faircloth, G. and Giavazzi, R. (1998) Ecteinascidin-743, a new marine natural product with potent antitumor activity on human ovarian carcinoma xenografts. *Clin Cancer Res*, **4**, 1977-1983.
94. Janjigian, Y.Y., Lee, W., Kris, M.G., Miller, V.A., Krug, L.M., Azzoli, C.G., Senturk, E., Calcutt, M.W. and Rizvi, N.A. (2010) A phase I trial of SJG-136 (NSC#694501) in advanced solid tumors. *Cancer Chemother Pharmacol*, **65**, 833-838.
95. Hochhauser, D., Meyer, T., Spanswick, V.J., Wu, J., Clingen, P.H., Loadman, P., Cobb, M., Gumbrell, L., Begent, R.H., Hartley, J.A. *et al.* (2009) Phase I study of sequence-selective minor groove DNA binding agent SJG-136 in patients with advanced solid tumors. *Clin Cancer Res*, **15**, 2140-2147.
96. Leimgrub, W., Stefanov, V., Schenker, F., Karr, A. and Berger, J. (1965) Isolation and characterization of anthramycin- a new antitumor antibiotic. *Journal of the American Chemical Society*, **87**, 5791-5793.
97. Fotso, S., Zabriskie, T.M., Proteau, P.J., Flatt, P.M., Santosa, D.A., Sulastri and Mahmud, T. (2009) Limazepines A-F, Pyrrolo[1,4]benzodiazepine Antibiotics from an Indonesian Micrococcus sp. *J Nat Prod*, **72**, 690-695.
98. Dai, J., Carte, B.K., Sidebotton, P.J., Sek New, A.L., Ng, S., Huang, Y. and Butler, M.S. (2001) Circumdatin G. a new alkaloid from the fungus *Aspergillus ochraceus*. *J Nat Prod*, **64**, 125-126.
99. Ookura, R., Kito, K., Ooi, T., Namikoshi, M. and Kusumi, T. (2008) Structure revision of circumdatins A and B, benzodiazepine alkaloids produced by marine fungus Aspergillus ostianus, by X-ray crystallography. *Journal of Organic Chemistry*, **73**, 4245-4247.
100. Rahbaek, L. and Breinholt, J. (1999) Circumdatins DE. F: further fungal benzodiazepine analogues from *Aspergillus ochraceus*. *J Nat Prod*, **62**.
101. Rahbeak, L., Breinholt, J., Frisvad, J.C. and Christophersen, C. (1999) Circumdatin A, B, and C: Three new benzodiazepine alkaloids isolated from a culture of the fungus *Aspergillus ochraceus*. *The Journal of organic chemistry*, **64**, 1689-1692.
102. Zhang, D., Yang, X., Kang, J.S., Choi, H.D. and Son, B.W. (2008) Circumdatin I, a new ultraviolet- a protecting benzodiazepine alkaloid from a marine ioslate of the fungus Exophiala. *J Antibiot (Tokyo)*, **91**, 40-42.
103. Hurley, L.H., Reck, T., Thurston, D.E., Langley, D.R., Holden, K.G., Hertzberg, R.P., Hoover, J.R., Gallagher, G.J., Faucette, L.R., Mong, A.M. *et al.* (1988) Pyrrolo[1,4]benzodiazepine antitumor antibiotics: relationship of DNA alkylation and sequence specificity to the biological activity of natural and synthetic compounds. *Chemical research in toxicology*, **1**, 258-268.
104. Petrusek, R.L., Anderson, G.L., Garner, T.F., Fannin, Q.L., Kaplan, D.J., Zimmer, S.G. and Hurley, L.H. (1981) Pyrrolo[1,4]benzodiazepine antibiotics-proposed structures and characteristics of the in vitro deoxyribonucleic-acid adducts of anthramycin, tomaymycin, sibiromycin, and neothramycin-A and neothramycin-B. *Biochemistry*, **20**, 1111-1119.
105. Barkley, M.D., Cheatham, S., Thurston, D.E. and Hurley, L.H. (1986) Pyrrolo[1,4]benzodiazepine antitumor antibiotics-evidence for 2 forms of tomaymycin bound to DNA. *Biochemistry*, **25**, 3021-3031.
106. Romers, W.A., Mabilia, M. and Hopfinger, A.J. (1986) Conformations of complexes between pyrrolo[1,4]benzodiazepines and DNA segments. *J Med Chem*, **29**, 2492-2503.
107. Antonow, D. and Thurston, D.E. (2011) Synthesis of DNA-interactive pyrrolo[2,1-c][1,4]benzodiazepines (PBDs). *Chemical reviews*, **111**, 2815-2864.
108. Rahman, K.M., Mussa, V., Narayanaswamy, M., James, C.H., Howard, P.W. and Thurston, D.E. (2009) Observation of a dynamic equilibrium between DNA

hairpin and duplex forms of covalent adducts of a minor groove binding agent. *Chem Commun (Camb)*, 227-229.
109. Kohn, K.W. and Spears, C.L. (1970) Reaction of anthramycin with deoxyribonucleic acid. *J Mol Biol*, **51**, 551-572.
110. Rahman, K.M., Corcoran, D.B., Bui, T.T., Jackson, P.J. and Thurston, D.E. (2014) Pyrrolobenzodiazepines (PBDs) do not bind to DNA G-quadruplexes. *PLoS One*, **9**, e105021.
111. Wells, G., Martin, C.R.H., Howard, P.W., Sands, Z.A., Laughton, C.A., Tiberghien, A., Woo, C.K., Masterson, L.A., Stephenson, M.J., Hartley, J.A. *et al.* (2006) Design, synthesis, and biophysical and biological evaluation of a series of Pyrrolobenzodiazepine - Poly(N-methylpyrrole) conjugates. *J Med Chem*, **49**, 5442-5461.
112. Warpehoski, M.A. and Hurley, L.H. (1988) Sequence selectivity of DNA covalent modification. *Chem Res Toxicol*, **1**, 315-333.
113. Hurley, L.H. (1977) Pyrrolo(1,4)Benzodiazepine Antitumor Antibiotics - Comparative Aspects of Anthramycin, Tomaymycin and Sibiromycin. *J Antibiot*, **30**, 349-370.
114. Hurley, L.H. and Petrusek, R. (1979) Proposed structure of the anthramycin-DNA adduct. *Nature*, **282**, 529-531.
115. Boyd, F.L., Stewart, D., Remers, W.A., Barkley, M.D. and Hurley, L.H. (1990) Characterization of a Unique Tomaymycin-D(Cicgaattcicg)2 Adduct Containing 2 Drug Molecules Per Duplex by Nmr, Fluorescence, and Molecular Modeling Studies. *Biochemistry*, **29**, 2387-2403.
116. Baraldi, P.G., Bovero, A., Fruttarolo, F., Preti, D., Tabrizi, M.A., Pavani, M.G. and Romagnoli, R. (2004) DNA minor groove binders as potential antitumor and antimicrobial agents. *Medicinal Research Reviews*, **24**, 475-528.
117. Thurston, D.E., Bose, D.S., Howard, P.W., Jenkins, T.C., Leoni, A., Baraldi, P.G., Guiotto, A., Cacciari, B., Kelland, L.R., Foloppe, M.P. *et al.* (1999) Effect of A-ring modifications on the DNA-binding behavior and cytotoxicity of pyrrolo[2,1-c] [1,4]benzodiazepines. *J Med Chem*, **42**, 1951-1964.
118. Alley, M.C., Hollingshead, M.G., Pacula-Cox, C.M., Wand, W.R., Hartley, J.A., Howard, P.W., Gregson, S.J., Thurston, D.E. and Sausville, E.A. (2004) SJG-136 (NSC 694501), a novel rationally designed DNA minor groove interstrand cross-linking agent with potent and broad spectrum antitumor activity. Part 2: Efficacy evaluations. *Cancer Research*, **64**, 6700-6706.
119. Hertzberg, R.P., Hecht, S.M. and Reynolds, V.L. (1986) DNA sequence specificity of the pyrrolo[2,1-c][1,4]benzodiazepine antitumor antibiotics. Methidiumpropyl-EDTA-iron(II) footpriniting analysis of DNA binding sites for anthramycin and related drugs. *Biochemistry*, **25**, 1249-1258.
120. Hertzberg, R.P., Hecht, S.M., Reynolds, V.L., Molineux, I.J. and Hurley, L.H. (1986) DNA-Sequence Specificity of the Pyrrolo[1,4]Benzodiazepine Antitumor Antibiotics - Methidiumpropyl-Edta-Iron(Ii) Footprinting Analysis of DNA-Binding Sites for Anthramycin and Related Drugs. *Biochemistry*, **25**, 1249-1258.
121. Puvvada, M.S., Forrow, S.A., Hartley, J.A., Stephenson, P., Gibson, I., Jenkins, T.C. and Thurston, D.E. (1997) Inhibition of bacteriophage T7 RNA polymerase in vitro transcription by DNA-binding pyrrolo[2,1-c][1,4]benzodiazepines. *Biochemistry*, **36**, 2478-2484.
122. Graves, D.E., Pattaroni, C., Krishnan, B.S., Ostrander, J.M., Hurley, L.H. and Krugh, T.R. (1984) The reaction of anthramycin with DNA. Proton and carbon nuclear magnetic resonance studies on the structure of the anthramycin-DNA adduct. *J Biol Chem*, **259**, 8202-8209.
123. Graves, D.E., Stone, M.P. and Krugh, T.R. (1985) Structure of the anthramycin-d(ATGCAT)2 adduct from one- and two-dimensional proton NMR experiments in solution. *Biochemistry*, **24**, 7573-7581.

124. Krugh, T.R., Graves, D.E. and Stone, M.P. (1989) Two-dimensional NMR studies on the anthramycin-d(ATGCAT)2 adduct. *Biochemistry*, **28**, 9988-9994.
125. Rao, S.N., Singh, U.C. and Kollman, P.A. (1986) Molecular mechanics simulations on covalent complexes between anthramycin and B DNA. *J Med Chem*, **29**, 2484-2492.
126. Remers, W.A., Mabilia, M. and Hopfinger, A.J. (1986) Conformations of complexes between pyrrolo[1,4]benzodiazepines and DNA segments. *J Med Chem*, **29**, 2492-2503.
127. Zakrzewska, K. and Pullman, B. (1986) A theoretical investigation of the sequence specificity in the binding of the antitumor drug anthramycin to DNA. *Journal of biomolecular structure & dynamics*, **4**, 127-136.
128. Kopka, M.L., Goodsell, D.S., Baikalov, I., Grzeskowiak, K., Cascio, D. and Dickerson, R.E. (1994) Crystal structure of a covalent DNA-drug adduct: anthramycin bound to C-C-A-A-C-G-T-T-G-G and a molecular explanation of specificity. *Biochemistry*, **33**, 13593-13610.
129. Vargiu, A.V., Ruggerone, P., Magistrato, A. and Carloni, P. (2006) Anthramycin-DNA binding explored by molecular simulations. *J Phys Chem B*, **110**, 24687-24695.
130. Jackson, P.J., James, C.H., Jenkins, T.C., Rahman, K.M. and Thurston, D.E. (2014) Computational studies support the role of the C7-sibirosamine sugar of the pyrrolobenzodiazepine (PBD) sibiromycin in transcription factor inhibition. *ACS Chem Biol*, **9**, 2432-2440.
131. Khokhlov, D.N., Brusov, R.V., Grokhovskii, S.L., Nikolaev, V.A., Pis'menskii, V.F., Zhuze, A.L. and Gurskii, G.V. (1995) [A synthetic zinc chelating peptide competes for DNA binding sites with antibiotics, adsorbed in a minor DNA groove]. *Mol Biol (Mosk)*, **29**, 354-364.
132. Cooper, N., Hagan, D.R., Tiberghien, A., Ademefun, T., Matthews, C.S., Howard, P.W. and Thurston, D.E. (2002) Synthesis of novel C2-aryl pyrrolobenzodiazepines (PBDs) as potential antitumor agents. *Chem Commun*, 1764-1765.
133. Fox, K.R. and Waring, M.J. (1984) DNA structural variations produced by actinomycin and distamycin as revealed by DNAase I footprinting. *Nucleic Acids Res*, **12**, 9271-9285.
134. Portugal, J. and Waring, M.J. (1987) Hydroxyl radical footprinting of the sequence-selective binding of netropsin and distamycin to DNA. *FEBS Lett*, **225**, 195-200.
135. Abudaya, A., Brown, P.M. and Fox, K.R. (1995) DNA-Sequence Preferences of Several at-Selective Minor-Groove Binding Ligands. *Nucleic acids research*, **23**, 3385-3392.
136. Coll, M., Frederick, C.A., Wang, A.H. and Rich, A. (1987) A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc Natl Acad Sci U S A*, **84**, 8385-8389.
137. Pelton, J.G. and Wemmer, D.E. (1989) Structural Characterization of a 2-1 Distamycin A.D(Cgcaaattggc) Complex by Two-Dimensional Nmr. *P Natl Acad Sci USA*, **86**, 5723-5727.
138. Uytterhoeven, K., Sponer, J. and Van Meervelt, L. (2002) Two 1 : 1 binding modes for distamycin in the minor groove of d(GGCCAATTGG). *European journal of biochemistry / FEBS*, **269**, 2868-2877.
139. Chen, X., Ramakrishnan, B. and Sundaralingam, M. (1997) Crystal structures of the side-by-side binding of distamycin to AT-containing DNA octamers d(ICITACIC) and d(ICATATIC). *J Mol Biol*, **267**, 1157-1170.
140. Chen, X., Ramakrishnan, B., Rao, S.T. and Sundaralingam, M. (1994) Binding of two distamycin A molecules in the minor groove of an alternating B-DNA duplex. *Nat Struct Biol*, **1**, 169-175.

141. Klevit, R.E., Wemmer, D.E. and Reid, B.R. (1986) 1H NMR studies on the interaction between distamycin A and a symmetrical DNA dodecamer. *Biochemistry*, **25**, 3296-3303.
142. Pelton, J.G. and Wemmer, D.E. (1990) Binding modes of distamycin A with d(CGCAAATTTGCG)2 determined by two-dimensional NMR. *Journal of the American Chemical Society*, **112**, 1393-1399.
143. Leslie, A.G., Arnott, S., Chandrasekaran, R. and Ratliff, R.L. (1980) Polymorphism of DNA double helices. *J Mol Biol*, **143**, 49-72.
144. Youngquist, R.S. and Dervan, P.B. (1985) Sequence-specific recognition of B-DNA by oligo(N-methylpyrrolecarboxamide)s. *Proc Natl Acad Sci U S A*, **82**, 2565-2569.
145. Schultz, P.G. and Dervan, P.B. (1983) Sequence-specific double-strand cleavage of DNA by penta-N-methylpyrrolecarboxamide-EDTA X Fe(II). *Proc Natl Acad Sci U S A*, **80**, 6834-6837.
146. Youngquist, R.S. and Dervan, P.B. (1985) A synthetic peptide binds 16 base pairs of A, T double helical DNA. *J Am Chem Soc*, **109**, 7564-7566.
147. Skamrov, A.V., Rybalkin, I.N., Bibilashvili, R., Gottikh, B.P. and Grokhovskii, S.L. (1985) [Specific protection of DNA by distamycin A, netropsin and bis-netropsins against the action of DNAse I]. *Mol Biol (Mosk)*, **19**, 177-195.
148. Schultz, P.G. and Dervan, P.B. (1983) Sequence-specifi double-strand clevage of DNA by bis(EDTA-distamycin,iron(II)) and EDTA-bis(distamycin).iron(II). *Journal of the American Chemical Society*, **105**, 7748-7750.
149. Youngquist, R.S. and Dervan, P.B. (1985) Sequence-specific recongition of B DNA by bis(EDTA-distamycin)fumaramide. *Journal of the American Chemical Society*, **107**, 5528-5529.
150. Griffin, J.H. and Dervan, P.B. (1986) Specific sequence-specific recognition of right-handed double-helical DNA by (2S,3S)- and (2R,3R)-dihydroxybis(netropsin)succinamide. *Journal of the American Chemical Society*, **108**, 5008-5009.
151. Wade, W.S., Mrksich, M. and Dervan, P.B. (1992) Design of Peptides That Bind in the Minor Groove of DNA at 5'-(a,T)G(a,T)C(a,T)-3' Sequences by a Dimeric Side-by-Side Motif. *Journal of the American Chemical Society*, **114**, 8783-8794.
152. Geierstanger, B.H., Mrksich, M., Dervan, P.B. and Wemmer, D.E. (1994) Design of a G.C-specific DNA minor groove-binding peptide. *Science*, **266**, 646-650.
153. Mrksich, M., Wade, W.S., Dwyer, T.J., Geierstanger, B.H., Wemmer, D.E. and Dervan, P.B. (1992) Antiparallel Side-by-Side Dimeric Motif for Sequence-Specific Recognition in the Minor Groove of DNA by the Designed Peptide 1-Methylimidazole-2-Carboxamide Netropsin. *P Natl Acad Sci USA*, **89**, 7586-7590.
154. White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E. and Dervan, P.B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, **391**, 468-471.
155. Dervan, P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorganic & medicinal chemistry*, **9**, 2215-2235.
156. Kelly, J.J., Baird, E.E. and Dervan, P.B. (1996) Binding site size limit of the 2:1 pyrrole-imidazole polyamide-DNA motif. *P Natl Acad Sci USA*, **93**, 6981-6985.
157. Trauger, J.W., Baird, E.E., Mrksich, M. and Dervan, P.B. (1996) Extension of sequence specific recognition in the minor groove of DNA by pyrrole-imidazole polyamides to 9-13 base pairs. *Journal of the American Chemical Society*, **118**, 6160-6166.
158. Swalley, S.E., Baird, E.E. and Dervan, P.B. (1997) A pyrrole-imidazole polyamide motif for recognition of elevan base pair sequences in the minor groove of DNA. *Chem Eur J*, **3**, 1600-1607.

159. Trauger, J.W., Baird, E.E. and Dervan, P.B. (1998) Recognition of 16 base pairs in the minor groove of DNA by a pyrrole-imidazole polyamide dimer. *Journal of the American Chemical Society*, **110**, 3534-3535.
160. James, P.L., Merkina, E.E., Khalaf, A.I., Suckling, C.J., Waigh, R.D., Brown, T. and Fox, K.R. (2004) DNA sequence recognition by an isopropyl substituted thiazole polyamide. *Nucleic acids research*, **32**, 3410-3417.
161. Hampshire, A.J., Khairallah, H., Khalaf, A.I., Ebrahimabadi, A.H., Waigh, R.D., Suckling, C.J., Brown, T. and Fox, K.R. (2006) DNA sequence recognition by an imidazole-containing isopropyl-substituted thiazole polyamide (thiazotropsin B). *Bioorganic & medicinal chemistry letters*, **16**, 3469-3474.
162. Rao, K.E., Bathini, Y. and Lown, J.W. (1990) Synthesis of Novel Thiazole-Containing DNA Minor Groove Binding Oligopeptides Related to the Antibiotic Distamycin. *Journal of Organic Chemistry*, **55**, 728-737.
163. Farmer, J.D., RUdnicki, S.M. and Suggs, J.W. (1988) Synthesis and DNA crosslinking ability of a dimeric anthramycin analog. *Tetrahydron Lett*, **29**, 5105-5108.
164. Brucoli, F., Hawkins, R.M., James, C.H., Jackson, P.J., Wells, G., Jenkins, T.C., Ellis, T., Kotecha, M., Hochhauser, D., Hartley, J.A. *et al.* (2013) An extended pyrrolobenzodiazepine-polyamide conjugate with selectivity for a DNA sequence containing the ICB2 transcription factor binding site. *J Med Chem*, **56**, 6339-6351.
165. Brucoli, F., Hawkins, R.M., James, C.H., Wells, G., Jenkins, T.C., Ellis, T., Hartley, J.A., Howard, P.W. and Thurston, D.E. (2011) Novel C8-linked pyrrolobenzodiazepine (PBD)-heterocycle conjugates that recognize DNA sequences containing an inverted CCAAT box. *Bioorganic & medicinal chemistry letters*, **21**, 3780-3783.
166. Kotecha, M., Kluza, J., Wells, G., O'Hare, C.C., Forni, C., Mantovani, R., Howard, P.W., Morris, P., Thurston, D.E., Hartley, J.A. *et al.* (2008) Inhibition of DNA binding of the NF-Y transcription factor by the pyrrolobenzodiazepine-polyamide conjugate GWL-78. *Mol Cancer Ther*, **7**, 1319-1328.
167. Wells, G., Martin, C.R., Howard, P.W., Sands, Z.A., Laughton, C.A., Tiberghien, A., Woo, C.K., Masterson, L.A., Stephenson, M.J., Hartley, J.A. *et al.* (2006) Design, synthesis, and biophysical and biological evaluation of a series of pyrrolobenzodiazepine-poly(N-methylpyrrole) conjugates. *J Med Chem*, **49**, 5442-5461.
168. Bose, D.S., Thompson, A.S., Ching, J.S., Hartley, J.A., Berardini, M.D., Jenkins, T.C., Neidle, S., Hurley, L.H. and Thurston, D.E. (1992) Rational Design of a Highly Efficient Irreversible DNA Interstrand Cross-Linking Agent Based on the Pyrrolobenzodiazepine Ring-System. *Journal of the American Chemical Society*, **114**, 4939-4941.
169. Gerratana, B. (2012) Biosynthesis, syntheis, and biological activities of pyrrolobenzodiazepines. *Medicinal Research Reviews*, **32**, 254-293.
170. Armstrong, R.W., Salvati, M.E. and Nguyen, M. (1992) Novel Interstrand Cross-Links Induced by the Antitumor Antibiotic Carzinophilin Azinomycin-B. *Journal of the American Chemical Society*, **114**, 3144-3145.
171. Pratt, W.B. and Ruddon, R.W. (1979) *The anticancer drugs*. Oxford University Press.
172. Hartley, J.A., Berardini, M.D. and Souhami, R.L. (1991) An Agarose-Gel Method for the Determination of DNA Interstrand Cross-Linking Applicable to the Measurement of the Rate of Total and 2nd-Arm Cross-Link Reactions. *Anal Biochem*, **193**, 131-134.
173. Jenkins, T.C., Hurley, L.H., Neidle, S. and Thurston, D.E. (1994) Structure of a Covalent DNA Minor-Groove Adduct with a Pyrrolobenzodiazepine Dimer - Evidence for Sequence-Specific Interstrand Cross-Linking. *J Med Chem*, **37**, 4529-4537.

174. Walton, M.I., Goddard, P., Kelland, L.R., Thurston, D.E. and Harrap, K.R. (1996) Preclinical pharmacology and antitumour activity of the novel sequence selective DNA minor-groove cross linking agent DSB 120. *Cancer Chemoth Pharm*, **38**, 431-438.
175. Rahman, K.M., Thompson, A.S., James, C.H., Narayanaswamy, M. and Thurston, D.E. (2009) The Pyrrolobenzodiazepine Dimer SJG-136 Forms Sequence-Dependent Intrastrand DNA Cross-Links and Monoalkylated Adducts in Addition to Interstrand Cross-Links. *Journal of the American Chemical Society*, **131**, 13756-13766.
176. Hopton, S.R. and Thompson, A.S. (2011) Nuclear magnetic resonance solution structures of inter- and intrastrand adducts of DNA cross-linker SJG-136. *Biochemistry*, **50**, 4720-4732.
177. Gregson, S.J., Howard, P.W., Hartley, J.A., Brooks, N.A., Adams, L.J., Jenkins, T.C., Kelland, L.R. and Thurston, D.E. (2001) Design, synthesis, and evaluation of a novel pyrrolobenzodiazepine DNA-interactive agent with highly efficient cross-linking ability and potent cytotoxicity. *J Med Chem*, **44**, 737-748.
178. Martin, C., Ellis, T., McGurk, C.J., Jenkins, T.C., Hartley, J.A., Waring, M.J. and Thurston, D.E. (2005) Sequence-selective interaction of the minor-groove interstrand cross-linking agent SJG-136 with naked and cellular DNA: footprinting and enzyme inhibition studies. *Biochemistry*, **44**, 4135-4147.
179. Doyle, M., Feuerbaum, E.A., Fox, K.R., Hinds, J., Thurston, D.E. and Taylor, P.W. (2009) Response of Staphylococcus aureus to subinhibitory concentrations of a sequence-selective, DNA minor groove cross-linking pyrrolobenzodiazepine dimer. *J Antimicrob Chemoth*, **64**, 949-959.
180. Howard, P.W., Chen, Z.Z., Gregson, S.J., Masterson, L.A., Tiberghien, A.C., Cooper, N., Fang, M., Coffils, M.J., Klee, S., Hartley, J.A. *et al.* (2009) Synthesis of a novel C2/C2 '-aryl-substituted pyrrolo[2,1-c][1,4]benzodiazepine dimer prodrug with improved water solubility and reduced DNA reaction rate. *Bioorganic & medicinal chemistry letters*, **19**, 6463-6466.
181. Hartley, J.A., Hamaguchi, A., Coffils, M., Martin, C.R.H., Suggitt, M., Chen, Z.Z., Gregson, S.J., Masterson, L.A., Tiberghien, A.C., Hartley, J.M. *et al.* (2010) SG2285, a Novel C2-Aryl-Substituted Pyrrolobenzodiazepine Dimer Prodrug That Cross-links DNA and Exerts Highly Potent Antitumor Activity. *Cancer Research*, **70**, 6849-6858.
182. Gregson, S.J., Howard, P.W., Hartley, J.A., Brooks, N.A., Adams, L.J., Jenkins, T.C., Kelland, L.R. and Thurston, D.E. (2001) Design, synthesis, and evaluation of a novel pyrrolobenzodiazepine DNA-interactive agent with highly efficient cross-linking ability and potent cytotoxicity. *J Med Chem*, **44**, 737-748.
183. Rahman, K.M., James, C.H. and Thurston, D.E. (2011) Effect of base sequence on the DNA cross-linking properties of pyrrolobenzodiazepine (PBD) dimers. *Nucleic Acids Res*, **39**, 5800-5812.
184. Rahman, K.M., Thompson, A.S., James, C.H., Narayanaswamy, M. and Thurston, D.E. (2009) The pyrrolobenzodiazepine dimer SJG-136 forms sequence-dependent intrastrand DNA cross-links and monoalkylated adducts in addition to interstrand cross-links. *J Am Chem Soc*, **131**, 13756-13766.
185. Gregson, S.J., Howard, P.W., Gullick, D.R., Hamaguchi, A., Corcoran, K.E., Brooks, N.A., Hartley, J.A., Jenkins, T.C., Patel, S., Guille, M.J. *et al.* (2004) Linker length modulates DNA cross-linking reactivity and cytotoxic potency of C8/C8' ether-linked C2-exo-unsaturated pyrrolo[2,1-c][1,4]benzodiazepine (PBD) dimers. *J Med Chem*, **47**, 1161-1174.
186. Kung Sutherland, M.S., Walter, R.B., Jeffrey, S.C., Burke, P.J., Yu, C., Kostner, H., Stone, I., Ryan, M.C., Sussman, D., Lyon, R.P. *et al.* (2013) SGN-CD33A: a novel CD33-targeting antibody-drug conjugate using a pyrrolobenzodiazepine dimer is active in models of drug-resistant AML. *Blood*, **122**, 1455-1463.

187. Jurcic, J.G. (2012) What happened to anti-CD33 therapy for acute myeloid leukemia? *Curr Hematol Malig Rep*, **7**, 65-73.
188. Rosenblat, T.L., McDevitt, M.R., Mulford, D.A., Pandit-Taskar, N., Divgi, C.R., Panageas, K.S., Heaney, M.L., Chanel, S., Morgenstern, A., Sgouros, G. *et al.* (2010) Sequential cytarabine and alpha-particle immunotherapy with bismuth-213-lintuzumab (HuM195) for acute myeloid leukemia. *Clin Cancer Res*, **16**, 5303-5311.
189. Jeffrey, S.C., Burke, P.J., Lyon, R.P., Meyer, D.W., Sussman, D., Anderson, M., Hunter, J.H., Leiske, C.I., Miyamoto, J.B., Nicholas, N.D. *et al.* (2013) A potent anti-CD70 antibody-drug conjugate combining a dimeric pyrrolobenzodiazepine drug with site-specific conjugation technology. *Bioconjug Chem*, **24**, 1256-1263.
190. Day, R.S., 3rd, Ziolkowski, C.H., Scudiero, D.A., Meyer, S.A., Lubiniecki, A.S., Girardi, A.J., Galloway, S.M. and Bynum, G.D. (1980) Defective repair of alkylated DNA by human tumour and SV40-transformed human cell strains. *Nature*, **288**, 724-727.
191. Yarosh, D.B., Foote, R.S., Mitra, S. and Day, R.S., 3rd. (1983) Repair of O6-methylguanine in DNA by demethylation is lacking in Mer- human tumor cell strains. *Carcinogenesis*, **4**, 199-205.
192. Phillips, W.P., Jr., Willson, J.K., Markowitz, S.D., Zborowska, E., Zaidi, N.H., Liu, L., Gordon, N.H. and Gerson, S.L. (1997) O6-methylguanine-DNA methyltransferase (MGMT) transfectants of a 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU)-sensitive colon cancer cell line selectively repopulate heterogenous MGMT+/MGMT- xenografts after BCNU and O6-benzylguanine plus BCNU. *Cancer Res*, **57**, 4817-4823.
193. Dolan, M.E., Mitchell, R.B., Mummert, C., Moschel, R.C. and Pegg, A.E. (1991) Effect of O6-benzylguanine analogues on sensitivity of human tumor cells to the cytotoxic effects of alkylating agents. *Cancer Res*, **51**, 3367-3372.
194. Friedman, H.S., Dolan, M.E., Pegg, A.E., Marcelli, S., Keir, S., Catino, J.J., Bigner, D.D. and Schold, S.C., Jr. (1995) Activity of temozolomide in the treatment of central nervous system tumor xenografts. *Cancer Res*, **55**, 2853-2857.
195. Wedge, S.R. and Newlands, E.S. (1996) O6-benzylguanine enhances the sensitivity of a glioma xenograft with low O6-alkylguanine-DNA alkyltransferase activity to temozolomide and BCNU. *Br J Cancer*, **73**, 1049-1052.
196. Esteller, M., Hamilton, S.R., Burger, P.C., Baylin, S.B. and Herman, J.G. (1999) Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. *Cancer Res*, **59**, 793-797.
197. Esteller, M. and Herman, J.G. (2004) Generating mutations but providing chemosensitivity: the role of O6-methylguanine DNA methyltransferase in human cancer. *Oncogene*, **23**, 1-8.
198. Esteller, M., Garcia-Foncillas, J., Andion, E., Goodman, S.N., Hidalgo, O.F., Vanaclocha, V., Baylin, S.B. and Herman, J.G. (2000) Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J Med*, **343**, 1350-1354.
199. Duncan, T., Trewick, S.C., Koivisto, P., Bates, P.A., Lindahl, T. and Sedgwick, B. (2002) Reversal of DNA alkylation damage by two human dioxygenases. *Proc Natl Acad Sci U S A*, **99**, 16660-16665.
200. Aas, P.A., Otterlei, M., Falnes, P.O., Vagbo, C.B., Skorpen, F., Akbari, M., Sundheim, O., Bjoras, M., Slupphaug, G., Seeberg, E. *et al.* (2003) Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature*, **421**, 859-863.
201. Margison, G.P. and Santibanez-Koref, M.F. (2002) O6-alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy. *Bioessays*, **24**, 255-266.

202. Bobola, M.S., Tseng, S.H., Blank, A., Berger, M.S. and Silber, J.R. (1996) Role of O6-methylguanine-DNA methyltransferase in resistance of human brain tumor cell lines to the clinically relevant methylating agents temozolomide and streptozotocin. *Clin Cancer Res*, **2**, 735-741.
203. Middlemas, D.S., Stewart, C.F., Kirstein, M.N., Poquette, C., Friedman, H.S., Houghton, P.J. and Brent, T.P. (2000) Biochemical correlates of temozolomide sensitivity in pediatric solid tumor xenograft models. *Clin Cancer Res*, **6**, 998-1007.
204. Walker, M.C., Masters, J.R. and Margison, G.P. (1992) O6-alkylguanine-DNA-alkyltransferase activity and nitrosourea sensitivity in human cancer cell lines. *Br J Cancer*, **66**, 840-843.
205. Hickman, M.J. and Samson, L.D. (1999) Role of DNA mismatch repair and p53 in signaling induction of apoptosis by alkylating agents. *Proc Natl Acad Sci U S A*, **96**, 10764-10769.
206. Karran, P. (2001) Mechanisms of tolerance to DNA damaging therapeutic drugs. *Carcinogenesis*, **22**, 1931-1937.
207. Stojic, L., Brun, R. and Jiricny, J. (2004) Mismatch repair and DNA damage signalling. *DNA Repair (Amst)*, **3**, 1091-1101.
208. Wilson, D.M., 3rd and Barsky, D. (2001) The major human abasic endonuclease: formation, consequences and repair of abasic lesions in DNA. *Mutat Res*, **485**, 283-307.
209. Fortini, P. and Dogliotti, E. (2007) Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. *DNA Repair (Amst)*, **6**, 398-409.
210. Stucki, M., Pascucci, B., Parlanti, E., Fortini, P., Wilson, S.H., Hubscher, U. and Dogliotti, E. (1998) Mammalian base excision repair by DNA polymerases delta and epsilon. *Oncogene*, **17**, 835-843.
211. Sobol, R.W., Kartalou, M., Almeida, K.H., Joyce, D.F., Engelward, B.P., Horton, J.K., Prasad, R., Samson, L.D. and Wilson, S.H. (2003) Base excision repair intermediates induce p53-independent cytotoxic and genotoxic responses. *J Biol Chem*, **278**, 39951-39959.
212. Sobol, R.W., Prasad, R., Evenski, A., Baker, A., Yang, X.P., Horton, J.K. and Wilson, S.H. (2000) The lyase activity of the DNA repair protein beta-polymerase protects from DNA-damage-induced cytotoxicity. *Nature*, **405**, 807-810.
213. Sobol, R.W., Watson, D.E., Nakamura, J., Yakes, F.M., Hou, E., Horton, J.K., Ladapo, J., Van Houten, B., Swenberg, J.A., Tindall, K.R. *et al.* (2002) Mutations associated with base excision repair deficiency and methylation-induced genotoxic stress. *Proc Natl Acad Sci U S A*, **99**, 6860-6865.
214. Ochs, K., Sobol, R.W., Wilson, S.H. and Kaina, B. (1999) Cells deficient in DNA polymerase beta are hypersensitive to alkylating agent-induced apoptosis and chromosomal breakage. *Cancer Res*, **59**, 1544-1551.
215. Op het Veld, C.W., Jansen, J., Zdzienicka, M.Z., Vrieling, H. and van Zeeland, A.A. (1998) Methyl methanesulfonate-induced hprt mutation spectra in the Chinese hamster cell line CHO9 and its xrcc1-deficient derivative EM-C11. *Mutat Res*, **398**, 83-92.
216. Teo, I.A., Broughton, B.C., Day, R.S., James, M.R., Karran, P., Mayne, L.V. and Lehmann, A.R. (1983) A biochemical defect in the repair of alkylated DNA in cells from an immunodeficient patient (46BR). *Carcinogenesis*, **4**, 559-564.
217. Lawley, P.D. and Phillips, D.H. (1996) DNA adducts from chemotherapeutic agents. *Mutat Res*, **355**, 13-40.
218. Murnane, J.P. and Byfield, J.E. (1981) Irreparable DNA cross-links and mammalian cell lethality with bifunctional alkylating agents. *Chem Biol Interact*, **38**, 75-86.

219. Clingen, P.H., De Silva, I.U., McHugh, P.J., Ghadessy, F.J., Tilby, M.J., Thurston, D.E. and Hartley, J.A. (2005) The XPF-ERCC1 endonuclease and homologous recombination contribute to the repair of minor groove DNA interstrand crosslinks in mammalian cells produced by the pyrrolo[2,1-c][1,4]benzodiazepine dimer SJG-136. *Nucleic Acids Res*, **33**, 3283-3291.
220. McLellan, L.I. and Wolf, C.R. (1999) Glutathione and glutathione-dependent enzymes in cancer drug resistance. *Drug Resist Updat*, **2**, 153-164.
221. Tew, K.D. (1994) Glutathione-associated enzymes in anticancer drug resistance. *Cancer Res*, **54**, 4313-4320.
222. Ranson, H., Prapanthadara, L. and Hemingway, J. (1997) Cloning and characterization of two glutathione S-transferases from a DDT-resistant strain of Anopheles gambiae. *Biochem J*, **324 (Pt 1)**, 97-102.
223. Tang, A.H. and Tu, C.P. (1994) Biochemical characterization of Drosophila glutathione S-transferases D1 and D21. *J Biol Chem*, **269**, 27876-27884.
224. Kauvar, L.M., Morgan, A.S., Sanderson, P.E. and Henner, W.D. (1998) Glutathione based approaches to improving cancer treatment. *Chem Biol Interact*, **111-112**, 225-238.
225. Adler, V., Yin, Z., Fuchs, S.Y., Benezra, M., Rosario, L., Tew, K.D., Pincus, M.R., Sardana, M., Henderson, C.J., Wolf, C.R. *et al.* (1999) Regulation of JNK signaling by GSTp. *EMBO J*, **18**, 1321-1334.
226. Tew, K.D., Monks, A., Barone, L., Rosser, D., Akerman, G., Montali, J.A., Wheatley, J.B. and Schmidt, D.E., Jr. (1996) Glutathione-associated enzymes in the human cell lines of the National Cancer Institute Drug Screening Program. *Mol Pharmacol*, **50**, 149-159.
227. Potapova, O., Haghighi, A., Bost, F., Liu, C., Birrer, M.J., Gjerset, R. and Mercola, D. (1997) The Jun kinase/stress-activated protein kinase pathway functions to regulate DNA repair and inhibition of the pathway sensitizes tumor cells to cisplatin. *J Biol Chem*, **272**, 14041-14044.
228. Kodym, R., Calkins, P. and Story, M. (1999) The cloning and characterization of a new stress response protein. A mammalian member of a family of theta class glutathione s-transferase-like proteins. *J Biol Chem*, **274**, 5131-5137.
229. Voehringer, D.W., Hirschberg, D.L., Xiao, J., Lu, Q., Roederer, M., Lock, C.B., Herzenberg, L.A., Steinman, L. and Herzenberg, L.A. (2000) Gene microarray identification of redox and mitochondrial elements that control resistance or sensitivity to apoptosis. *Proc Natl Acad Sci U S A*, **97**, 2680-2685.
230. Galas, D.J. and Schmitz, A. (1978) Dnaase Footprinting - Simple Method for Detection of Protein-DNA Binding Specificity. *Nucleic acids research*, **5**, 3157-3170.
231. Fox, K.R. and Waring, M.J. (1984) DNA Structural Variations Produced by Actinomycin and Distamycin as Revealed by Dnaase-I Footprinting. *Nucleic acids research*, **12**, 9271-9285.
232. Vandyke, M.W., Hertzberg, R.P. and Dervan, P.B. (1982) Map of Distamycin, Netropsin, and Actinomycin Binding-Sites on Heterogeneous DNA - DNA Cleavage-Inhibition Patterns with Methidiumpropyl-Edta.Fe(Ii). *P Natl Acad Sci-Biol*, **79**, 5470-5474.
233. Vandyke, M.W. and Dervan, P.B. (1983) Chromomycin, Mithramycin, and Olivomycin Binding-Sites on Heterogeneous Deoxyribonucleic-Acid - Footprinting with (Methidiumpropyl-Edta)Iron(Ii). *Biochemistry*, **22**, 2373-2377.
234. Carpenter, M.L., Marks, J.N. and Fox, K.R. (1993) DNA-Sequence Binding Preference of the Gc-Selective Ligand Mithramycin - Deoxyribonuclease-I/Deoxyribonuclease-Ii and Hydroxy-Radical Footprinting at Cccg, Ccgc, Cggc, Gccc and Gggg Flanked by (at)(N) and a(N).T(N). *European Journal of Biochemistry*, **215**, 561-566.
235. Chandler, S.P. and Fox, K.R. (1996) Specificity of antiparallel DNA triple helix formation. *Biochemistry*, **35**, 15038-15048.

236. Gowers, D.M., Bijapur, J., Brown, T. and Fox, K.R. (1999) DNA triple helix formation at target sites containing several pyrimidine interruptions: stabilization by protonated cytosine or 5-(1-propargylamino)dU. *Biochemistry*, **38**, 13747-13758.
237. Wang, Y., Rusling, D.A., Powers, V.E.C., Lack, O., Osborne, S.D., Fox, K.R. and Brown, T. (2005) Stable recognition of TA interruptions by triplex forming oligonucleotides containing a novel nucleoside. *Biochemistry*, **44**, 5884-5892.
238. Cardew, A.S. and Fox, K.R. (2010) In Fox, K. R. (ed.), *Drug-DNA interaction protocols, Methods in Molecular biology*. Humana Press, Vol. 613.
239. Hampshire, A.J., Rusling, D.A., Broughton-Head, V.J. and Fox, K.R. (2007) Footprinting: a method for determining the sequence selectivity, affinity and kinetics of DNA-binding ligands. *Methods*, **42**, 128-140.
240. Laskowski, M. (1971) In Boyer, P. (ed.). Academic Press, London, Vol. 4, pp. 289-311.
241. Suck, D., Lahm, A. and Oefner, C. (1988) Structure Refined to 2a of a Nicked DNA Octanucleotide Complex with Dnase-I. *Nature*, **332**, 464-468.
242. Price, P.A. (1975) Essential Role of Ca2+ in Activity of Bovine Pancreatic Deoxyribonuclease. *J Biol Chem*, **250**, 1981-1986.
243. Lahm, A. and Suck, D. (1991) Dnase I-Induced DNA Conformation 2a Structure of a Dnase I-Octamer Complex. *J Mol Biol*, **222**, 645-667.
244. Weston, S.A., Lahm, A. and Suck, D. (1992) X-Ray Structure of the Dnase I-D(Ggtatacc)(2) Complex at 2.3-Angstrom Resolution. *J Mol Biol*, **226**, 1237-1256.
245. Suck, D. and Oefner, C. (1986) Structure of DNase I at 2.0 A resolution suggests a mechanism for binding to and cutting DNA. *Nature*, **321**, 620-625.
246. Tullius, T.D. and Dombroski, B.A. (1986) Hydroxyl Radical Footprinting - High-Resolution Information About DNA Protein Contacts and Application to Lambda-Repressor and Cro Protein. *P Natl Acad Sci USA*, **83**, 5469-5473.
247. Tullius, T.D. (1988) DNA footprinting with hydroxyl radical. *Nature*, **332**, 663-664.
248. Tullius, T.D., Dombroski, B.A., Churchill, M.E.A. and Kam, L. (1987) Hydroxyl Radical Footprinting - a High-Resolution Method for Mapping Protein DNA Contacts. *Method Enzymol*, **155**, 537-558.
249. Portugal, J. and Waring, M.J. (1987) Hydroxyl Radical Footprinting of the Sequence-Selective Binding of Netropsin and Distamycin to DNA. *Febs Lett*, **225**, 195-200.
250. Cons, B.M.G. and Fox, K.R. (1989) High-Resolution Hydroxyl Radical Footprinting of the Binding of Mithramycin and Related Antibiotics to DNA. *Nucleic acids research*, **17**, 5447-5459.
251. Abu-Daya, A. and Fox, K.R. (1997) Interaction of minor groove binding ligands with long AT tracts. *Nucleic acids research*, **25**, 4962-4969.
252. Fenton, H.J.H. (1894) Oxidation of tartaric acid in presence of iron. *Journal of Chemical Society, Transactions*, **65**, 899-910.
253. Cons, B.M. and Fox, K.R. (1989) High resolution hydroxyl radical footprinting of the binding of mithramycin and related antibiotics to DNA. *Nucleic Acids Res*, **17**, 5447-5459.
254. Abu-Daya, A. and Fox, K.R. (1997) Interaction of minor groove binding ligands with long AT tracts. *Nucleic Acids Res*, **25**, 4962-4969.
255. Hertzberg, R.P. and Dervan, P.B. (1982) Cleavage of Double Helical DNA by (Methidiumpropyl-Edta)Iron(Ii). *Journal of the American Chemical Society*, **104**, 313-315.
256. Lavesa, M. and Fox, K.R. (2001) Preferred binding sites for [N-MeCys(3),N-MeCys(7)]TANDEM determined using a universal footprinting substrate. *Anal Biochem*, **293**, 246-250.

257. Hampshire, A.J. and Fox, K.R. (2008) Preferred binding sites for the bifunctional intercalator TANDEM determined using DNA fragments that contain every symmetrical hexanucleotide sequence. *Anal Biochem*, **374**, 298-303.
258. Mergny, J.L. and Maurizot, J.C. (2001) Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat. *Chembiochem*, **2**, 124-132.
259. Mergny, J.L., Phan, A.T. and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett*, **435**, 74-78.
260. Cahen, P., Luhmer, M., Fontaine, C., Morat, C., Reisse, J. and Bartik, K. (2000) Study by Na-23-NMR, H-1-NMR, and ultraviolet spectroscopy of the thermal stability of an 11-basepair oligonucleotide. *Biophys J*, **78**, 1059-1069.
261. Guo, Q., Lu, M. and Kallenbach, N.R. (1992) Adenine Affects the Structure and Stability of Telomeric Sequences. *Journal of Biological Chemistry*, **267**, 15293-15300.
262. Hardin, C.C., Henderson, E., Watson, T. and Prosser, J.K. (1991) Monovalent cation induced structural transitions in telomeric DNAs: G-DNA folding intermediates. *Biochemistry*, **30**, 4460-4472.
263. Jin, R.Z., Breslauer, K.J., Jones, R.A. and Gaffney, B.L. (1990) Tetraplex formation of a guanine-containing nonameric DNA fragment. *Science*, **250**, 543-546.
264. Duguid, J.G., Bloomfield, V.A., Benevides, J.M. and Thomas, G.J., Jr. (1996) DNA melting investigated by differential scanning calorimetry and Raman spectroscopy. *Biophys J*, **71**, 3350-3360.
265. Thomas, R. (1993) The denaturation of DNA. *Gene*, **135**, 77-79.
266. Darby, R.A., Sollogoub, M., McKeen, C., Brown, L., Risitano, A., Brown, N., Barton, C., Brown, T. and Fox, K.R. (2002) High throughput measurement of duplex, triplex and quadruplex melting curves using molecular beacons and a LightCycler. *Nucleic Acids Res*, **30**, e39.
267. Tyagi, S. and Kramer, F.R. (1996) Molecular beacons: Probes that fluoresce upon hybridization. *Nat Biotechnol*, **14**, 303-308.
268. Yang, X.L., Kaenzig, C., Lee, M. and Wang, A.H. (1999) Binding of AR-1-144, a tri-imidazole DNA minor groove binder, to CCGG sequence analyzed by NMR spectroscopy. *Eur J Biochem*, **263**, 646-655.
269. Yang, X.L., Kaenzig, C., Lee, M. and Wang, A.H.J. (1999) Binding of AR-1-144, a tri-imidazole DNA minor groove binder, to CCGG sequence analyzed by NMR spectroscopy. *Eur J Biochem*, **263**, 646-655.
270. Marmur, J. and Doty, P. (1962) Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J Mol Biol*, **5**, 109-118.
271. Yakovchuk, P., Protozanova, E. and Frank-Kamenetskii, M.D. (2006) Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res*, **34**, 564-574.
272. Breslauer, K.J., Frank, R., Blocker, H. and Marky, L.A. (1986) Predicting DNA duplex stability from the base sequence. *Proc Natl Acad Sci U S A*, **83**, 3746-3750.
273. Owczarzy, R., Vallone, P.M., Gallo, F.J., Paner, T.M., Lane, M.J. and Benight, A.S. (1997) Predicting sequence-dependent melting stability of short duplex DNA oligomers. *Biopolymers*, **44**, 217-239.
274. Garbett, N.C. and Graves, D.E. (2004) Extending natures leads: the anticancer agent ellipticine. *Current Medicinal Chemistry-Anti-cancer Agents*, **4**, 149-172.
275. Canals, A., Purciolas, M., Aymami, J. and Coll, M. (2005) The anticancer agent ellipticine unwinds DNA by interactive binding in an orientation parallel to base pairs. *Acta Crystallographica*, **D61**, 1009-1012.

276. Lisgarten, J.N., Coll, M., Portugal, J., Wright, C.W. and Aymami, J. (2002) The antimalarial and cytotoxic drug cryptolepine intercalates into DNA at cytosine-cytosine sites. *Nat Struct Biol*, **9**, 57-60.
277. Pjura, P.E., Grzeskowiak, K. and Dickerson, R.E. (1987) Binding of Hoechst-33258 to the Minor Groove of B-DNA. *J Mol Biol*, **197**, 257-271.
278. Teng, M., Usman, N., Frederick, C.A. and Wang, A.H.J. (1988) The Molecular-Structure of the Complex of Hoechst-33258 and the DNA Dodecamer D(Cgcgaattcgcg). *Nucleic acids research*, **16**, 2671-2690.
279. Hayakawa, I., Shioya, R., Agatsuma, T., Furukawa, H., Naruto, S. and Sugano, Y. (2004) 4-Hydroxy-3-methyl-6-phenylbenzofuran-2-carboxylic acid ethyl ester derivatives as potent anti-tumor agents. *Bioorg Med Chem Lett*, **14**, 455-458.
280. Kim, S., Salim, A.A., Swanson, S.M. and Kinghorn, A.D. (2006) Potential of cyclopenta[b]benzofurans from Aglaia species in cancer chemotherapy. *Anticancer Agents Med Chem*, **6**, 319-345.
281. Zhou, Z.Z., Zou, M., Zhou, J., Zhou, C.Q., Deng, Y.H., Chen, M.H., Gu, C.P., Jiang, Z.H., Chen, W.H. and Liu, S.W. (2011) Synthesis, cytotoxicities and DNA-binding affinities of benzofuran-3-ols and their fused analogs. *Chem Pharm Bull (Tokyo)*, **59**, 1057-1061.
282. Aleksic, M., Bertosa, B., Nhili, R., Uzelac, L., Jarak, I., Depauw, S., David-Cordonnier, M.H., Kralj, M., Tomic, S. and Karminski-Zamola, G. (2012) Novel Substituted Benzothiophene and Thienothiophene Carboxanilides and Quinolones: Synthesis, Photochemical Synthesis, DNA-Binding Properties, Antitumor Evaluation and 3D-Derived QSAR Analysis. *J Med Chem*, **55**, 5044-5060.
283. Shi, W. and Lowary, T.L. (2011) Structure-activity relationships in glycosylated 2-phenyl-indoles, 2-phenyl-benzo[b]thiophenes and 2-phenyl-benzo[b]furans as DNA binding and potential antitumor agents. *Bioorg Med Chem*, **19**, 1779-1789.
284. Shi, W., Marcus, S.L. and Lowary, T.L. (2011) Cytotoxicity and topoisomerase I/II inhibition of glycosylated 2-phenyl-indoles, 2-phenyl-benzo[b]thiophenes and 2-phenyl-benzo[b]furans. *Bioorg Med Chem*, **19**, 603-612.
285. Schnell, J.R., Ketchem, R.R., Boger, D.L. and Chazin, W.J. (1999) Binding-induced activation of DNA-alkylation by duocarmycin SA: Insights from the structure of an indole derivative-DNA adduct. *Journal of the American Chemical Society*, **121**, 5645-5652.
286. Gupta, S.P., Pandya, P., Kumar, G.S. and Kumar, S. (2012) Indole Derivatives as DNA Minor Groove Binders. *Chemistry of Phytopotentials: Health, Energy and Environmental Perspectives*, 149-153.
287. Thurston, D.E. (1993) In Neidle, S. and Waring, M. J. (eds.), *In molecular aspects of anticancer drug-DNA interactions*. The Macmillan Press Ltd, London, UK, Vol. 1, pp. 54-88.
288. Jenkins, T.C., Hurley, L.H., Neidle, S. and Thurston, D.E. (1994) Structure of a covalent DNA minor groove adduct with a pyrrolobenzodiazepine dimer: evidence for sequence-specific interstrand cross-linking. *J Med Chem*, **37**, 4529-4537.
289. Mountzouris, J.A., Wang, J.J., Thurston, D. and Hurley, L.H. (1994) Comparison of a DSB-120 DNA interstrand cross-linked adduct with the corresponding bis-tomaymycin adduct: an example of a successful template-directed approach to drug design based upon the monoalkylating compound tomaymycin. *J Med Chem*, **37**, 3132-3140.
290. Walton, M.I., Goddard, P., Kelland, L.R., Thurston, D.E. and Harrap, K.R. (1996) Preclinical pharmacology and antitumour activity of the novel sequence-selective DNA minor-groove cross-linking agent DSB-120. *Cancer Chemother Pharmacol*, **38**, 431-438.
291. Smellie, M., Bose, D.S., Thompson, A.S., Jenkins, T.C., Hartley, J.A. and Thurston, D.E. (2003) Sequence-selective recognition of duplex DNA through

covalent interstrand cross-linking: kinetic and molecular modeling studies with pyrrolobenzodiazepine dimers. *Biochemistry*, **42**, 8232-8239.

292. Hartley, J.A., Spanswick, V.J., Brooks, N., Clingen, P.H., McHugh, P.J., Hochhauser, D., Pedley, R.B., Kelland, L.R., Alley, M.C., Schultz, R. *et al.* (2004) SJG-136 (NSC 694501), a novel rationally designed DNA minor groove interstrand cross-linking agent with potent and broad spectrum antitumor activity. Part 1: Cellular pharmacology, in vitro and initial in vivo antitumor activity. *Cancer Research*, **64**, 6693-6699.
293. Pepper, C.J., Hambly, R.M., Fegan, C.D., Delavault, P. and Thurston, D.E. (2004) The novel sequence-specific DNA cross-linking agent SJG-136 (NSC 694501) has potent and selective in vitro cytotoxicity in human B-cell chronic lymphocytic leukemia cells with evidence of a p53-independent mechanism of cell kill. *Cancer Res*, **64**, 6750-6755.
294. Alley, M.C., Hollingshead, M.G., Pacula-Cox, C.M., Waud, W.R., Hartley, J.A., Howard, P.W., Gregson, S.J., Thurston, D.E. and Sausville, E.A. (2004) SJG-136 (NSC 694501), a novel rationally designed DNA minor groove interstrand cross-linking agent with potent and broad spectrum antitumor activity: part 2: efficacy evaluations. *Cancer Res*, **64**, 6700-6706.
295. Hochhauser, D., Meyer, T., Spanswick, V.J., Wu, J., Clingen, P.H., Loadman, P., Cobb, M., Gumbrell, L., Begent, R.H., Hartley, J.A. *et al.* (2009) Phase I Study of Sequence-Selective Minor Groove DNA Binding Agent SJG-136 in Patients with Advanced Solid Tumors. *Clinical Cancer Research*, **15**, 2140-2147.
296. Owczarzy, R., You, Y., Moreira, B.G., Manthey, J.A., Huang, L.Y., Behlke, M.A. and Walder, J.A. (2004) Effects of sodium ions on DNA duplex oligomers: Improved predictions of melting temperatures. *Biochemistry*, **43**, 3537-3554.
297. Mrksich, M. and Dervan, P.B. (1993) Antiparallel side-by-side heterodimer for sequence-specific recognition in the minor groove of DNA by a distamycin/1-methylimidazole-2-carboxamide-netropsin pair. *J Am Chem Soc*, **115**, 2572-2576.
298. Geierstanger, B.H., Jacobsen, J.P., Mrksich, M., Dervan, P.B. and Wemmer, D.E. (1994) Structural and dynamic characterization of the heterodimeric and homodimeric complexes of distamycin and 1-methylimidazole-2-carboxamide-netropsin bound to the minor groove of DNA. *Biochemistry*, **33**, 3055-3062.
299. Mrksich, M., Parks, M.E. and Dervan, P.B. (1994) Hairpin peptide motif. A new class of oligopeptides for sequence-specific recognition in the minor groove of double-helical DNA. *Journal of the American Chemical Society*, **116**, 7983-7988.
300. Trauger, J.W., Baird, E.E. and Dervan, P.B. (1996) Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature*, **382**, 559-561.
301. Herman, D.M., Baird, E.E. and Dervan, P.B. (1998) Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *Journal of the American Chemical Society*, **120**, 1382-1391.
302. Swalley, S.E., Baird, E.E. and Dervan, P.B. (1999) Effect of gamma-turn and beta-tail amino acids on sequence-specific recognition of DNA by hairpin polyamides. *Journal of the American Chemical Society*, **121**, 1113-1120.
303. Woods, C.R., Ishii, T., Wu, B., Kenneth, W.B. and Boger, D.L. (2002) Hairpin versus extended DNA binding of a substituted beta-alanine linked polyamide. *Journal of the American Chemical Society*, **124**, 2148-2152.
304. Dervan, P.B., Doss, R.M. and Marques, M.A. (2005) Programmable DNA binding oligomers for control of transcription. *Curr Med Chem Anticancer Agents*, **5**, 373-387.

305. White, S., Baird, E.E. and Dervan, P.B. (1997) Orientation preferences of pyrrole-imidazole polyamides in the minor groove of DNA. *J Am Chem Soc*, **119**, 8756-8765.
306. Asamitsu, S., Kawamoto, Y., Hashiya, F., Hashiya, K., Yamamoto, M., Kizaki, S., Bando, T. and Sugiyama, H. (2014) Sequence-specific DNA alkylation and transcriptional inhibition by long-chain hairpin pyrrole-imidazole polyamide-chlorambucil conjugates targeting CAG/CTG trinucleotide repeats. *Bioorg Med Chem*, **22**, 4646-4657.
307. Wang, Y., Dziegielewski, J., Chang, A.Y., Dervan, P.B. and Beerman, T.A. (2002) Cell-free and cellular activities of a DNA sequence selective hairpin polyamide-CBI conjugate. *Journal of Biological Chemistry*, **277**, 42431-42437.
308. Wurtz, N.R. and Dervan, P.B. (2000) Sequence specific alkylation of DNA by hairpin pyrrole-imidazole polyamide conjugates. *Chem Biol*, **7**, 153-161.
309. Chiang, S.Y., Burli, R.W., Benz, C.C., Gawron, L., Scott, G.K., Dervan, P.B. and Beerman, T.A. (2000) Targeting the ets binding site of the HER2/neu promoter with pyrrole-imidazole polyamides. *J Biol Chem*, **275**, 24246-24254.
310. Lai, Y.M., Fukuda, N., Ueno, T., Matsuda, H., Saito, S., Matsumoto, K., Ayame, H., Bando, T., Sugiyama, H., Mugishima, H. *et al.* (2005) Synthetic pyrrole-imidazole polyamide inhibits expression of the human transforming growth factor-beta1 gene. *J Pharmacol Exp Ther*, **315**, 571-575.
311. Best, T.P., Edelson, B.S., Nickols, N.G. and Dervan, P.B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc Natl Acad Sci U S A*, **100**, 12063-12068.
312. Fechter, E.J., Olenyuk, B. and Dervan, P.B. (2005) Sequence-specific fluorescence detection of DNA by polyamide-thiazole orange conjugates. *Journal of the American Chemical Society*, **127**, 16685-16691.
313. Guo, C., Kawamoto, Y., Asamitsu, S., Sawatani, Y., Hashiya, K., Bando, T. and Sugiyama, H. (2015) Rational design of specific binding hairpin Py-Im polyamides targeting human telomere sequences. *Bioorg Med Chem*, **23**, 855-860.
314. Gottesfeld, J.M., Neely, L., Trauger, J.W., Baird, E.E. and Dervan, P.B. (1997) Regulation of gene expression by small molecules. *Nature*, **387**, 202-205.
315. Mirkin, S.M. (2007) Expandable DNA repeats and human disease. *Nature*, **447**, 932-940.
316. Cummings, C.J. and Zoghbi, H.Y. (2000) Fourteen and counting: unraveling trinucleotide repeat diseases. *Hum Mol Genet*, **9**, 909-916.
317. Kremer, B., Goldberg, P., Andrew, S.E., Theilmann, J., Telenius, H., Zeisler, J., Squitieri, F., Lin, B., Bassett, A., Almqvist, E. *et al.* (1994) A worldwide study of the Huntington's disease mutation. The sensitivity and specificity of measuring CAG repeats. *N Engl J Med*, **330**, 1401-1406.
318. Reddy, P.H., Williams, M., Charles, V., Garrett, L., Pike-Buchanan, L., Whetsell, W.O., Jr., Miller, G. and Tagle, D.A. (1998) Behavioural abnormalities and selective neuronal loss in HD transgenic mice expressing mutated full-length HD cDNA. *Nat Genet*, **20**, 198-202.
319. Lin, X., Cummings, C.J. and Zoghbi, H.Y. (1999) Expanding our understanding of polyglutamine diseases through mouse models. *Neuron*, **24**, 499-502.
320. Swalley, S.E., Baird, E.E. and Dervan, P.B. (1997) Discrimination of 5'-GGGG-3', 5'-GCGC-3', and 5'-GGCC-3' sequences in the minor groove of DNA by eight-ring hairpin polyamides. *Journal of the American Chemical Society*, **119**, 6953-6961.
321. Pilch, D.S., Poklar, N., Baird, E.E., Dervan, P.B. and Breslauer, K.J. (1999) The thermodynamics of polyamide-DNA recognition: hairpin polyamide binding in the minor groove of duplex DNA. *Biochemistry*, **38**, 2143-2151.
322. Han, Y.W., Matsumoto, T., Yokota, H., Kashiwazaki, G., Morinaga, H., Hashiya, K., Bando, T., Harada, Y. and Sugiyama, H. (2012) Binding of hairpin pyrrole

and imidazole polyamides to DNA: relationship between torsion angle and association rate constants. *Nucleic Acids Res*, **40**, 11510-11517.

323. Rosado, H., Rahman, K.M., Feuerbaum, E.A., Hinds, J., Thurston, D.E. and Taylor, P.W. (2011) The minor groove-binding agent ELB-21 forms multiple interstrand and intrastrand covalent cross-links with duplex DNA and displays potent bactericidal activity against methicillin-resistant Staphylococcus aureus. *J Antimicrob Chemoth*, **66**, 985-996.
324. Rahman, K.M., Rosado, H., Moreira, J.B., Feuerbaum, E.A., Fox, K.R., Stecher, E., Howard, P.W., Gregson, S.J., James, C.H., de la Fuente, M. *et al.* (2012) Antistaphylococcal activity of DNA-interactive pyrrolobenzodiazepine (PBD) dimers and PBD-biaryl conjugates. *J Antimicrob Chemoth*, **67**, 1683-1696.