INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Dimethylsulfate (DMS) protection footprinting of the interaction of cruciform DNA with a human cruciform binding protein (CBP)

Fiona Robinson

Department of Biochemistry

McGill University, Montréal

August, 1999

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Master's of Science.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Cruciforms are an alternative secondary structure which may be adopted by DNA containing inverted repeats, under conditions of adequate torsional strain. Inverted repeats are distributed, in a non-random fashion, throughout the genomes of prokaryotes and eukaryotes. Mounting evidence suggests that they are involved in the initiation of DNA replication. A structure-specific cruciform DNA binding protein (CBP) has previously been enriched from HeLa cells, and demonstrated to be a member of the 14-3-3 family of proteins. This thesis reports the dimethylsulfate (DMS) protection footprinting of this protein on a stable cruciform, with the goal of testing a model proposed for this interaction. The footprint obtained was not clear or reproducible enough to allow this verification, however, it does support previously identified regions of binding on the cruciform DNA. Possible explanations for the nature of the footprint obtained, and suggestions which may allow the achievement of verification of the model, are discussed.

RÉSUMÉ

La structure cruciforme constitue une forme alternative de l'ADN contenant des répétitions inversées. Cette structure survient une force de torsion adéquate. Les répétitions inversées sont distribuées non-aléatoirement dans le génome des procaryotes et eucaryotes. De plus en plus de preuves suggèrent qu'elles sont impliquées dans l'initiation de la réplication de l'ADN. Préalablement, une protéine s'associant spécifiquement à cet ADN de structure cruciforme (CBP) a été enrichie de cellules HeLa. Cette protéine appartient à la famille des protéines 14-3-3. Cette

thèse a pour but de tester un modèle expliquant l'intéraction entre CBP et l'ADN cruciforme en utilisant un essai de protection contre le diméthylsulfate (DMS). Les résultats obtenus sont ambigus et peu reproductibles, ce qui ne permet pas la vérification du modèle. Toutefois, ils confirment les régions de contact entre la protéine et l'ADN. Des raisons possibles expliquant la qualité des résultats obtenus et des alternatives expérimentales pouvant permettre la vérification du modèle sont présentées.

1. INTRODUCTION

The interactions of proteins and nucleic acids is the intersection between the genetic information and its implementation governing life processes. These interactions therefore present an intriguing field of study to the biochemist, to unlock the secrets behind the regulation of the transmission of this information. One level of regulation is the definition of nucleic acid-protein binding partners. There are two critical parameters of the nucleic acids which define their suitability for binding to specific proteins: sequence and structure. The presentation of a strictly defined series of functional groups, such as hydrogen bond donors and acceptors, delineated by the base sequence of a nucleic acid molecule, is a determinant that is easy to understand in terms of the limited set of polypeptides with which it can form energetically favourable interactions. Dictations governed by secondary structure present a greater challenge to our understanding. One example of a nucleic acid secondary structure which dictates binding of a very limited set of proteins is the cruciform. This thesis investigates the nature of the interactions between a model cruciform and a binding activity which we have termed the cruciform binding protein (CBP), *via* protection footprinting.

1.1 Inverted Repeats and Cruciforms

A palindrome is a sequence of DNA featuring dyad symmetry. This means that the sequence, if read in one direction, for example 5' to 3', is identical to that of the complementary strand also read in the 5' to 3' direction [1]. In the following example the centre of symmetry is located between the TTT and the AAA:

5' GAATTTAAATTC 3'

3' CTTAAATTTAAG 5'

Palindromes are also referred to as inverted repeats (IRs). Their symmetry provides the possibility of an alternative secondary structure in which intra-strand rather than inter-strand base pairing and hydrogen bonding yields a hairpin rather than the usual linear double helix. When a palindromic sequence is flanked by non-symmetrical sequence the resulting structure resembles a cross and is therefore referred to as a "cruciform" (Figure 1.1) [1]. Though their existence had been proposed as early as 1955 [2], it was not until the 1980s that cruciforms became accepted as an alternative secondary structure for DNA containing IRs [3], [4].

IRs are found, distributed in a non-random fashion, in the genomes of prokaryotes and eukaryotes (reviewed in [5]). They have been shown to be associated with regions involved in the control of transcription and replication of DNA, including the origins of replication of prokaryotes, viruses, and eukaryotes, including mammals. This suggests a functional role for these sequences, and their alternative secondary structures, in the control of these processes. They may exert a regulatory effect in their linear form as a binding site for protein dimers, at the RNA level through the formation of attenuation and termination facilitating hairpins, or through the extruded cruciform structure of the DNA. Extensive physical and biochemical studies into the existence of cruciforms *in vivo* (reviewed in [5]) have culminated in the development of cruciform-specific antibodies [6, 7]. These antibodies have been used to demonstrate the presence of 0.6×10^5 to 3×10^5 cruciforms per cell (human and monkey) with a discrete and dynamic nuclear localization during S phase [8, 9].

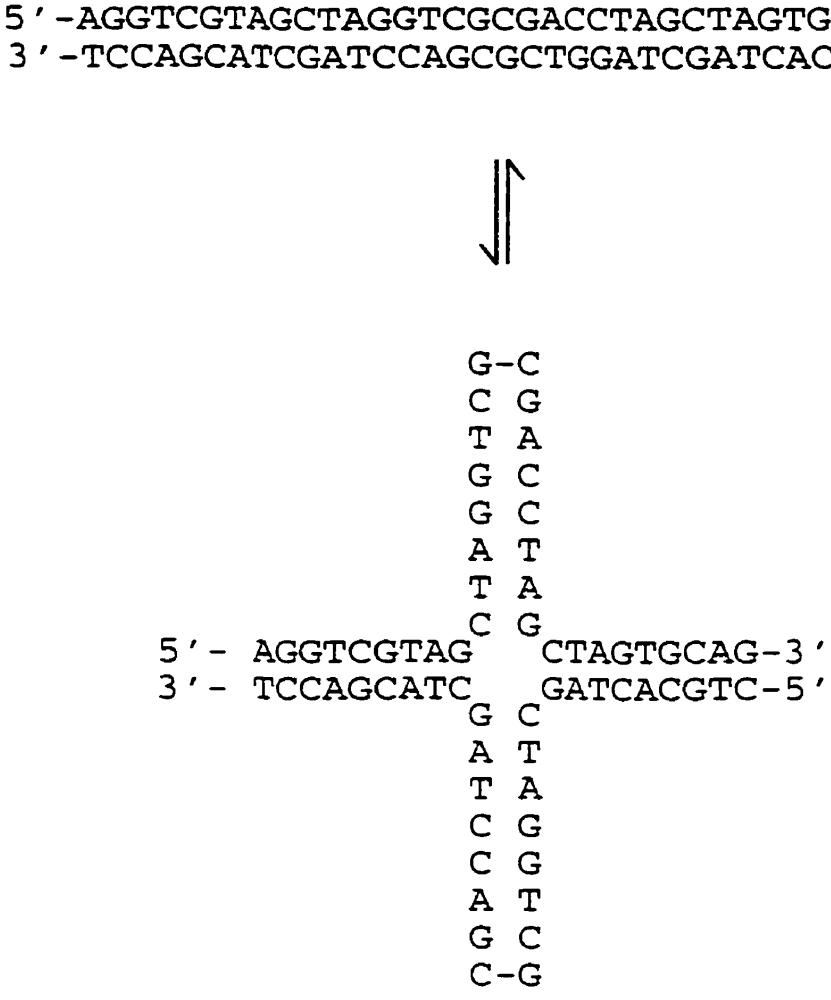


Figure 1.1 The cruciform as an alternative secondary structure for DNA containing an IR. Under appropriate conditions of torsional strain DNA containing an IR, or palindromic sequence, can extrude into a cross-shape or cruciform, which exploits the self-complementarity of the IR to form intrastrand hydrogen bonds. This extrusion is reversible.

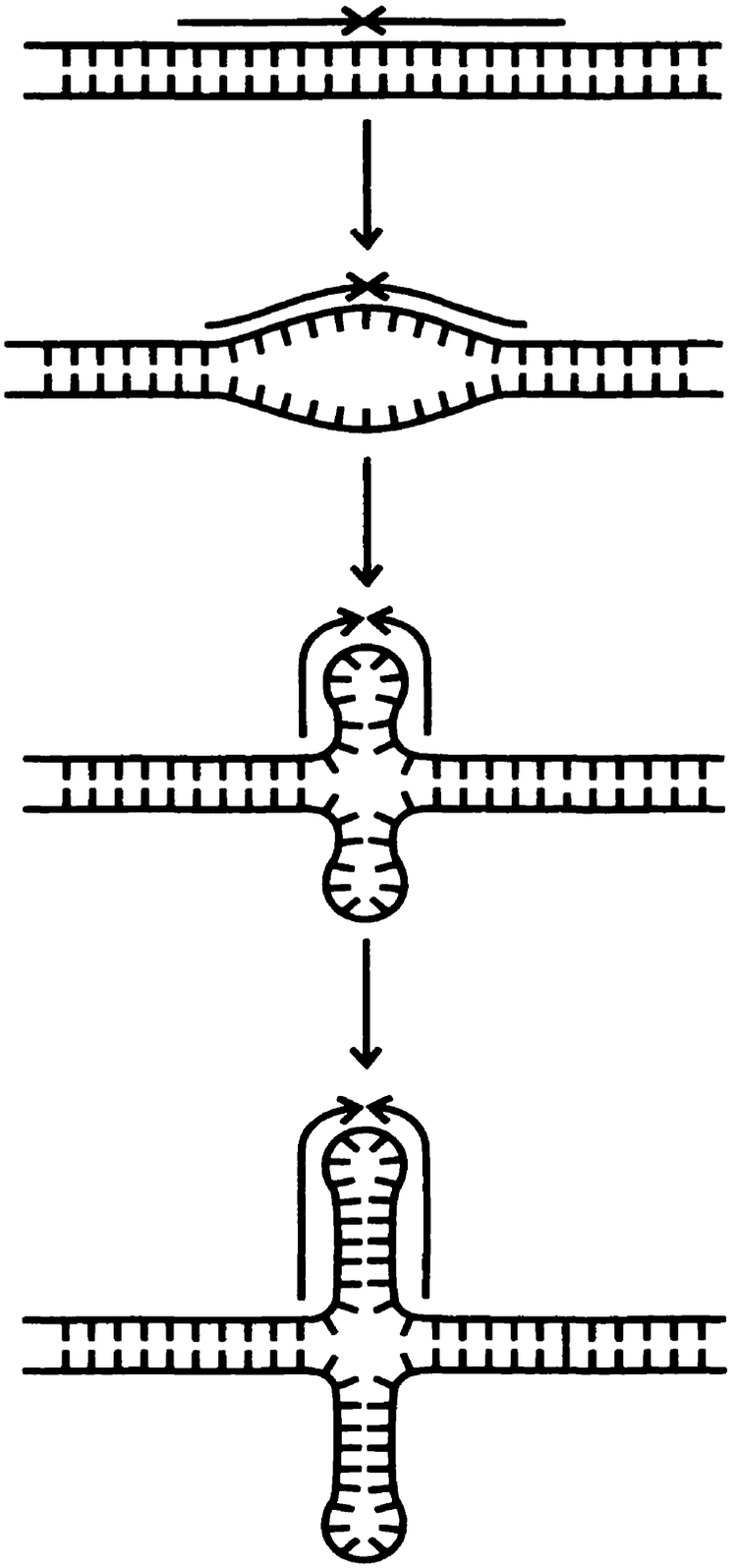
1.2 Characteristics of Cruciforms

1.2 a) Cruciform extrusion

There exist two types of cruciform extrusion: S (salt dependent) and C (named after the ColE1 plasmid in which it was first observed). C-type cruciform formation occurs in the absence of salt and will not be discussed here. S-type cruciform formation is that which is believed to occur under physiological conditions. The central 10-bp of the palindrome melt and intrastrand annealing is nucleated [1]. This is followed by extrusion of the entire palindrome (Figure 1.2). The rate of initial unwinding depends on the temperature, ionic strength and superhelical density of the DNA [1].

Unlike Holliday junctions, in which no areas of incomplete base pairing have been detected [10], cruciforms feature incomplete pairing and stacking of 3-4 bases at the tips of the arms formed by the hairpin loops, which otherwise adopt a B-helix conformation [1]. This less than maximal pairing and stacking makes the cruciform less stable than the corresponding linear structure of the sequence. How then can cruciforms exist? In negatively supercoiled DNA cruciform extrusion absorbs energy resulting from torsional strain. For every 10.5-bp which extrude, one negative supercoil is absorbed from the covalently closed DNA [11]. σ_c, the critical superhelical density, is the threshold of negative supercoils present in the DNA up to which a cruciform does not extrude. If one more negative supercoil than the σ_c is introduced into covalently closed DNA, then an IR present will extrude into a cruciform. This parameter is temperature dependent, decreasing with increasing temperature, since the melting apart of the DNA necessary for nucleation requires less torsional strain. Conversely, the re-linearization of a cruciform requires the introduction of one negative supercoil into the DNA for every 10.5-bp

Figure 1.2 The process of S-type cruciform extrusion. S-type cruciform extrusion involves the initial melting of the central 10-bp of the palindrome, nucleation of intrastrand annealing, and subsequent extrusion of the entire palindrome. Reproduced, with permission from Academic Press, copyright 1994, from [1].



of its length; an energy requiring process. As a result longer cruciforms are more stable than shorter ones. Similarly, once a cruciform has formed, due to σ_c having been exceeded, it may continue to exist at levels of superhelicity well below σ_c, if the energy needed for the introduction of negative supercoils concomitant to linearization is not available ([1] and references therein).

1.2 b) Cruciform structure

A considerable amount of study has been devoted, in the past decade, to the structure of the four-way junction, predominantly with an interest in Holliday junctions (reviewed in [10] and [12]). It has been concluded that the DNA may adopt a stacked-X or a more extended, unstacked conformation [13 , 14]. The principal determinant of which structure is adopted, in solution, is the concentration of cations, especially Mg^{2+} [10]. Both conformations have been evidenced in X-ray crystal structures [15] and deduced from theoretical studies [16].

In the absence of cations the four-way junction is more extended and has an open central region, with a near square arrangement of the four arms ([10], [12], [17] and references therein). In the presence of cations the junction structure is based upon pairwise helical stacking, with a rotation not unlike the opening of a pair of scissors. This allows for an increase in the extent of base-pair stacking, while minimizing steric and electrostatic hindrance [10]. Within the parameters of the stacked-X structure there exists the possibility of a number of conformers based upon the choice of co-axial stacking partners. Though a particular conformer appears to be preferred by a given structure, recent work has shown that the system is far from static and an equilibrium between the conformers generally exists in solution ([12] and references therein). It has also recently become clear that different cruciform binding proteins have different preferences for the conformation

of the DNA, and several actually distort the structure upon binding (See Section 1.7).

1.3 The 21/29 "Stable" Cruciform System

In order to study the behaviour of cruciforms and their interaction with other molecules, a model cruciform system has been developed [6, 7, 18]. Such a cruciform must be simple to prepare and isolate, stable in the systems in which we wish to study it, and accurately represent the defining characteristics of naturally occurring cruciforms. Cruciform formation *in vivo* depends critically upon the local degree of torsional strain present in the DNA (See Section 1.2). All naturally formed cruciforms can revert to their original linear form should the torsional conditions change. Such a dynamic system is undesirable for our studies as it would be difficult to maintain the extruded form of the cruciform. Therefore, rather than the simple extrusion of a palindrome which occurs *in vivo*, our system involves the heteroduplexing of two pieces of DNA possessing unrelated palindromes [18]. The plasmid pRGM21 consists of the 200-bp fragment generated from the *Hind*III/*Sph*I double digest of the wild-type SV40 origin of replication, including the 27-bp palindrome, inserted into the *Hind*III/*Sph*I site of the plasmid pBR322. The *Sph*I/*Xma*III fragment encompassing this sequence was subsequently inserted into the *Eag*I (*Xma*III)/*Hind*III site of the pBluescript-KS(+) vector. This plasmid will be referred to as pBS/RGM21. pRGM29 is identical to pRGM21, except that a 26-bp unrelated palindrome replaces that found in pRGM21. The same series of operations as for pBS/RGM21, yielded pBS/RGM29 from pRGM29.

The 21/29 cruciform is formed by first digesting both plasmids with *Hind*III and *Sph*I, liberating 200-bp fragments which are identical in sequence with the

exception of their unrelated palindromes. Denaturation, under basic and high salt conditions, of all four strands followed by a slow renaturation allows the reannealing of the entirely complementary strands and those unmatched only in the palindromic region with approximately equal efficiency [6], [7]. Therefore, approximately half of the resulting 200-bp fragments are heteroduplexes and "stable" cruciforms (Figure 1.3). Because the extruded palindromic regions are unrelated and not complementary, it is much more energetically favourable for the strands to remain paired in the arms of the cruciform than return to the linear form, and thus the cruciform formation is effectively irreversible. It should be noted that this process generates two complementary cruciforms (Figure 1.3) which co-migrate in 4-8% polyacrylamide gel electrophoresis (PAGE). The cruciform may be separated from the linear 200-bp fragment by PAGE since the more bulky cruciform structure is retarded with respect to the linear fragment [19].

This model cruciform is particularly suited to *in vitro* manipulation due to its stability and ease of preparation. It meets the enzymatic susceptibility criteria of cruciforms [7, 18], demonstrating S1 and mung bean nuclease cleavability at the tips of the loops, resistance to DNaseI cleavage in the elbow regions and recognition and restriction by the four-way junction specific T7 endonuclease I [7], [20]. It is also a more accurate model of *in vivo* cruciforms than those formed by the annealing of four separate oligonucleotides used by many investigators [10], because it possesses the areas of partial base pairing and stacking found at the tips of the naturally occurring secondary structures.

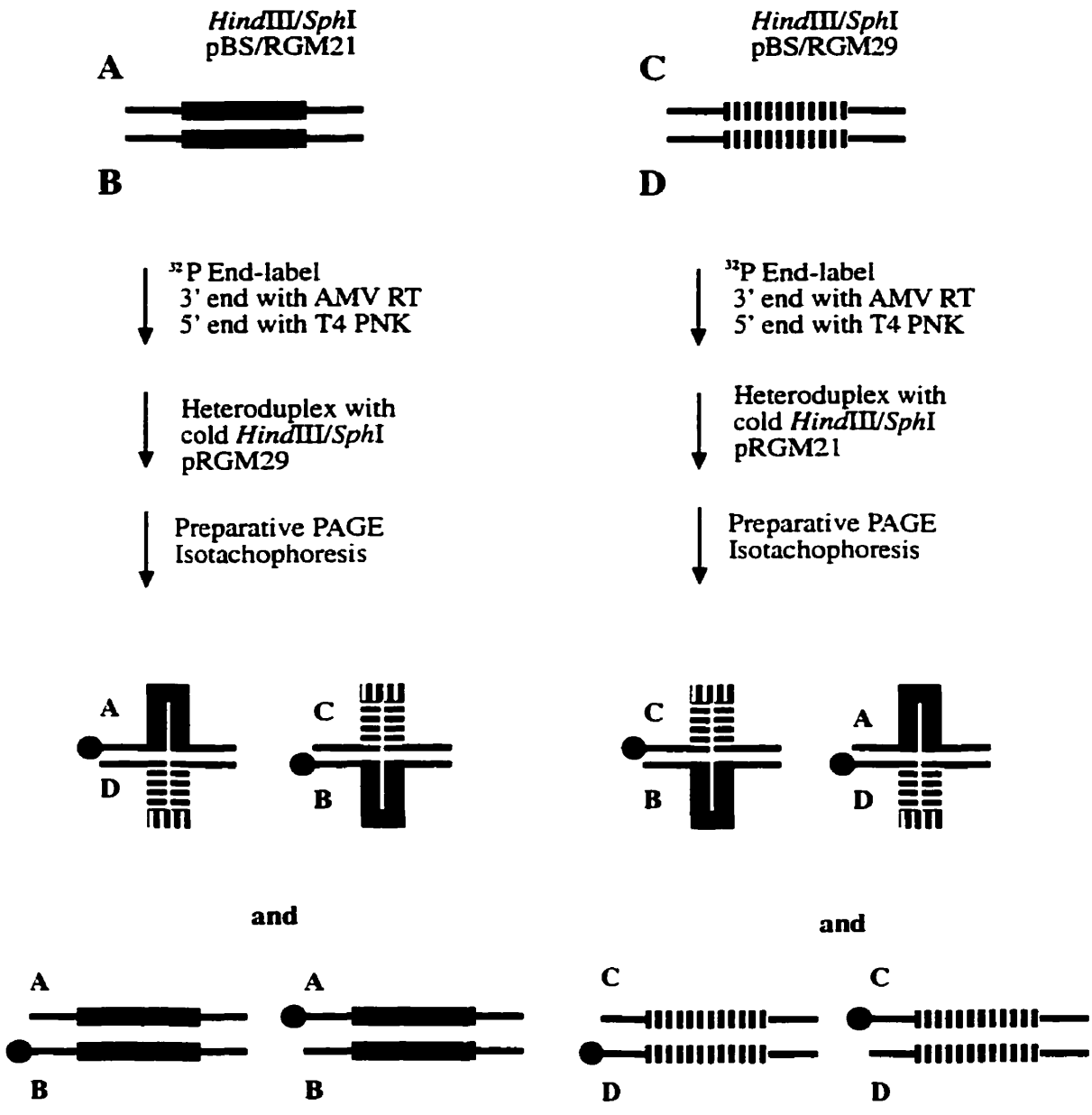


Figure 1.3 Production of the four "stable" cruciforms of the 21/29 system. See sections 1.3 and 2.1 a) for details of the heteroduplexing of the four 200-bp strands of the 21/29 system to yield four irreversible cruciforms. In the re-annealing process linear homoduplexes and cruciform heteroduplexes form with approximately equal efficiency. The thick and dashed lines represent the palindromes (of unrelated sequence) in the pBS/RGM21 and pBS/RGM29 plasmids, respectively. The large black dots denote radioactive end-labels, which allow individual experimental examination of the four strands. AMV RT is avian myeloblastosis virus reverse transcriptase and T4 PNK is T4 polynucleotide kinase.

1.4 Cruciforms and Replication

1.4 a) Possible modes of cruciform involvement in DNA replication

A number of modes, direct and indirect, have been proposed for the involvement of cruciform structures in the initiation of DNA replication (reviewed in [5]). One mode of involvement suggests an indirect role for cruciforms affecting the extent of superhelical density of the DNA, which is known to influence the binding of specific proteins involved in replication initiation [21], [22]. Alternatively, local absorption of torsional strain, resulting from DNA unwinding for replication, by cruciform formation may be involved in the selection of a particular site as the dominant replication origin, if multiple initiation events occur over a region of DNA [23]. Another mode of involvement proposes that the incompatibility of cruciforms, which tend to be associated with origins of replication, with nucleosome assembly helps to make the DNA available for the binding of initiation factors ([5] and references therein). There exists also the possibility that cruciforms themselves interact with a protein or proteins and thus may play a more direct role in the initiation of replication. *In vitro* and *in vivo* control of transcription by the binding of cruciform specific proteins to these secondary DNA structures has been established [24-28]. The activity of RNA polymerase also affects the level of torsional strain in DNA, causing an increase in the number of negative supercoils upstream from the transcription site [29]. This may create a situation favourable to the extrusion of cruciforms and constitute a link between transcription and the initiation of replication.

1.4 b) Direct involvement of cruciforms in DNA replication

Evidence for the involvement of cruciforms in nucleoprotein complexes associated with the initiation of DNA replication exists in a number of systems from the single stranded (ss) bacteriophage to mammals (reviewed in [5]). The ss

bacteriophages ϕX174 and G4 require the binding of a protein to a stem-loop structure [30, 31], which may form from an IR a central non-symmetrical sequence, as one of the initial steps in the assembly of the replication machinery. A similar requirement for a cruciform is seen in the case of the double-stranded (ds) plasmid pT181 [32, 33]. In addition, recognition of a stem-loop structure by a ribonucleoprotein appears to be instrumental in the initiation of mitochondrial DNA synthesis [34, 35]. Analysis of origin-enriched sequences cloned from replicating monkey cells (CV-1) [36], [37] similarly obtained libraries of early-replicating human DNA, and known prokaryotic and viral replication origins, demonstrated an enrichment for IRs ([5] and references therein). Introduction of the cruciform-specific antibodies, mentioned above [6, 7], into cells carrying out replication demonstrated a temporal correspondence between the two maxima of cruciform occurrence in S-phase with those of DNA replication [38], [39]. The cruciform population peaks immediately preceding the DNA synthesis peak [8]. The introduction of these antibodies into replicating cells also influences the levels of replication, resulting in a 2- to 11-fold increase in the relative copy number of low-copy genetic elements. This suggested that the stabilization of cruciforms, caused by antibody binding, resulted in multiple initiations at a single origin site [38]. Taken together these data suggest that the formation of cruciforms is cell-cycle regulated and important for the replication of mammalian DNA, perhaps providing attachment sites for important proteins.

1.5 CBP

1.5 a) Discovery and characterization of CBP

A cruciform binding, structure-specific, sequence-independent activity, termed cruciform binding protein (CBP), has been enriched from HeLa cell extracts and characterized by our laboratory [40-42]. This activity was enriched from nuclear extracts of cells in the logarithmic phase of the cell growth cycle by eluting component proteins from a DEAE-Sephadex (weak anion exchanger) column with a linear salt (potassium acetate) gradient, followed by loading of fractions demonstrating cruciform binding activity onto an Affi-Gel Heparin (which mimics nucleic acids) column. The unbound flow-through of this column contained all the cruciform binding activity and was subjected to glycerol gradient sedimentation, from which the active fractions were retained [40]. The cruciform binding activity was assayed using electrophoretic mobility shift assays (EMSAs) with two cruciforms of unrelated sequence as substrates. Competition assays demonstrated that CBP binds to cruciforms but not to linear DNA of the same sequence, does not bind to ss DNA, but does bind weakly to Y-shaped DNA [40]. The activity of CBP was distinct from that of high mobility group protein 1 (HMG1), a highly abundant protein which binds many DNA structures including cruciforms [43]. Cruciforms bound to CBP have a different mobility from those shifted by HMG1 in EMSAs, and Western analysis of the glycerol gradient fractions showed that the CBP activity (66 kDa) does not co-sediment with HMG1 (28 kDa) [40].

1.5 b) CBP is a member of the 14-3-3 family of proteins

Further analysis of CBP identified it as a member of the 14-3-3 family of proteins [42]. Microsequence analysis of several polypeptides purified by virtue of their cruciform binding activity, showed 100% homology with the ε, β, γ and ζ 14-

3-3 isoforms, and no homology with any other protein families. 14-3-3 purified from sheep brain was shown to possess cruciform-specific DNA binding activity, and the presence of the ε, β and ζ isoforms of 14-3-3 in the nucleus was demonstrated by immunofluorescence. Western analysis of the proteins isolated by their cruciform binding activity confirmed the presence of the β, γ and ε and possibly the ζ isoforms. 14-3-3 proteins have been shown to be part of the transcriptional complex in *Arabidopsis* [44] and maize [45], but there have been no previous reports of DNA binding by these largely cytoplasmic proteins. A recent report of 14-3-3 interaction with p53 has also placed them in the nucleus [46].

1.6 14-3-3

The 14-3-3 family of proteins (reviewed in [47-53]) was first identified in 1967 [54]. The first activity attributed to them was a role in the synthetic pathways of serotonin and dopamine in the brain [55]. They have since been implicated in a wide variety of cellular functions including exocytosis [56], apoptosis [57], cell cycle regulation [58], and signal transduction, where they have been shown to interact with different proteins from a number of pathways (reviewed in [48], [50], [52], [53]).

1.6 a) Structure of 14-3-3 and amino acid conservation between isoforms and species

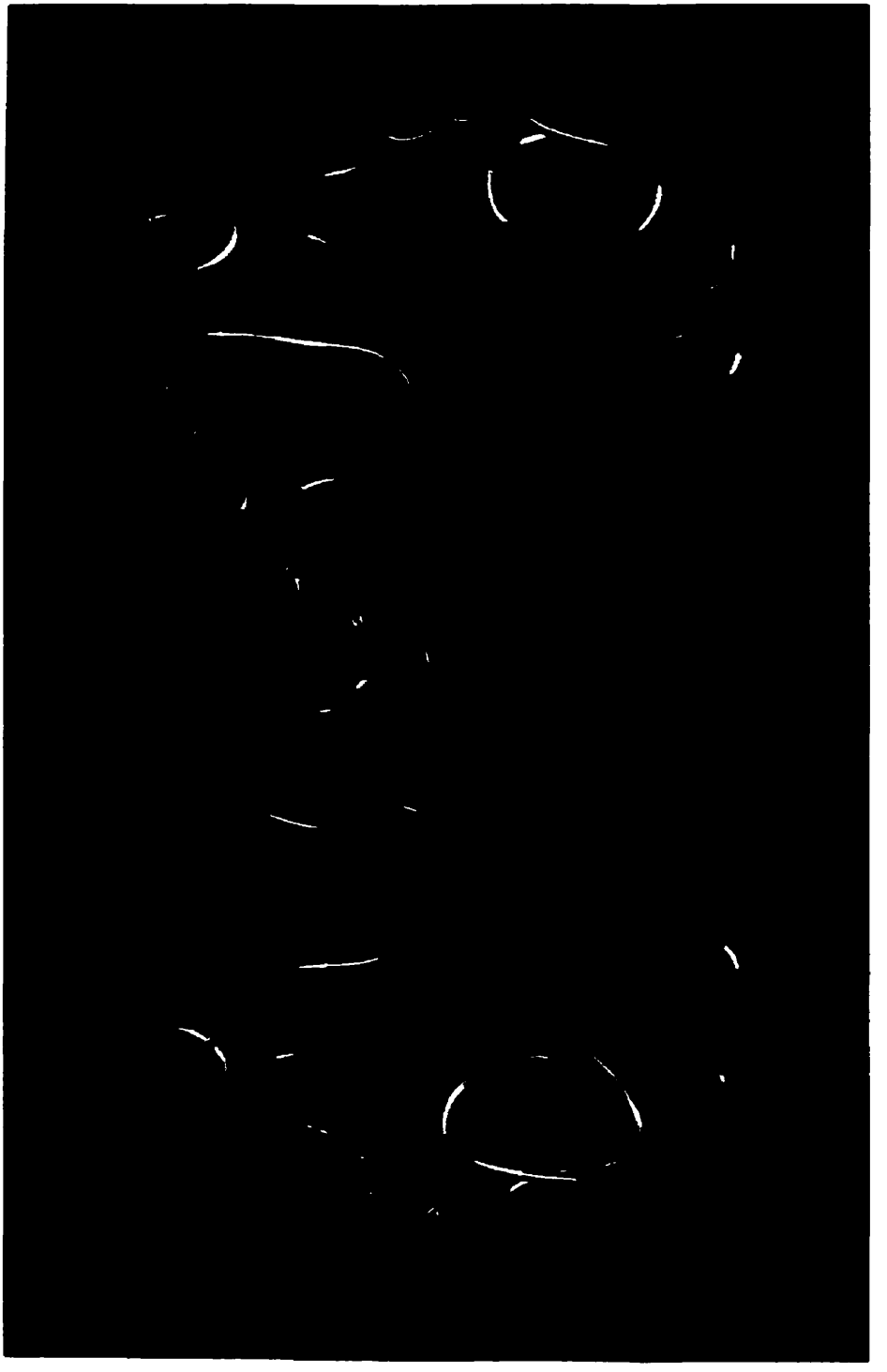
The 14-3-3 family consists of at least seven mammalian isoforms: β (and its phosphorylated form α), ε, γ, ζ (and its phosphorylated form δ), τ and σ, and have been found distributed in all tissues of all eukaryotes studied to date [49]. There is a very high degree of amino acid conservation between isoforms and

across species [59] suggesting a fundamentally important function for these proteins. The crystal structure of 14-3-3 [60], [61] shows the protein adopting a saddle shaped dimer conformation with a large amphipathic groove (approximately 35x35x20 Å, Figure 1.4). Examination of the distribution of residue conservation about this structure, both between isoforms and species, showed that the N-terminal dimerisation region and that lining the channel are conserved to the highest extent, with those on the outside of the saddle structure showing the highest variability [60, 61]. Upon this basis, the possibility of heterodimerisation was proposed, and subsequently demonstrated [62]. This prompted the suggestion of an adapter protein role for 14-3-3, in which different isoforms could interact with different proteins, and heterodimerisation could bring them together.

1.6 b) 14-3-3 as an adapter protein

Two groups have recently reported putative consensus binding motifs for 14-3-3 binding partners: RSXpSXP (where pS denotes phosphoserine) [63] and RXY/FXpSXP [64], which are common to partner proteins with many diverse functions in the cell. The co-crystal structure of the ζ isoform of 14-3-3 with a polypeptide containing the former motif localized its binding site to the interior of the channel, near the C-terminus, with two polypeptides binding to a dimer [64]. More recently, the presence of an overlapping but distinct site for the binding of non-phosphorylated peptides has been demonstrated [65]. This information has allowed a greater insight into the way in which the 14-3-3 family of proteins may act as adapter molecules, interacting with a variety of apparently unrelated proteins and potentially bridging their functions. It is very exciting to consider the possible link that this family of proteins may provide between such processes as signal

Figure 1.4 Crystal structure of the 14-3-3 τ homodimer. Ribbon representation of the structure, as deduced by X-ray crystallography, of the 14-3-3 τ homodimer, reproduced with permission, from Nature [60], copyright 1995, Macmillan Magazines Ltd. The structure, made up of 18 α-helices, 9 in each monomer, adopts a saddle shape with a cleft of approximately 35x35x20 Å. The N-terminal regions provide the dimer interface. Regions of highest amino acid conservation (blue) line the cleft, whereas regions of higher variability (red) are found on the outside of the structure. Green represents an intermediate degree of conservation.



transduction and the control of DNA replication, in light of its cruciform binding activity.

1.7 Other Cruciform Binding Proteins

A number of other proteins which bind cruciform DNA in a structure-dependent manner have been found in organisms that range from bacteria and bacteriophages to eukaryotes and their viruses, and new ones are constantly coming to light [66]. However, the elucidation of these structure-specific interactions is far from complete, and it does not appear that they will converge to a single common binding mode. The majority of information available is for junction-resolving enzymes (reviewed in [67] and [68]), particularly RuvC, T7 endonuclease I and T4 endonuclease VII. X-ray crystal structures have been solved for the former two ([69] and [70], respectively). The CBP activity discovered in our lab has been demonstrated to be devoid of any nuclease activity [40].

1.7 a) T4 endonuclease VII, T7 endonuclease I and RuvC

T4 endonuclease VII was the first enzyme discovered to bind and cleave branched DNA in a structure specific manner [71]. It exhibits binding affinity and cleavage activity with a number of DNA structural substrates including Holliday junctions, cruciforms, Y-junctions, heteroduplex loops, single-stranded overhangs, curved DNA, abasic sites and single base mismatches ([70] and references therein). Its primary function is believed to be the resolution of branchpoints in the process of packaging the viral DNA into the bacteriophage head [72]. T7 Endonuclease I performs this same function in the bacteriophage T7, as well as cleaving the DNA of the host cell [20]. It has a very high binding specificity for branched duplex DNA, but also cleaves ss DNA [73]. The RuvC protein appears to be the major junction resolving enzyme in *E. coli* and is important for homologous

recombination. It is believed to work in conjunction with the RuvARuvB complex to cleave Holliday junctions as a late step in the recombination process ([67] and references therein). RuvA also recognizes and binds four-way junction DNA [15, 74, 75].

T4 endonuclease VII, T7 endonuclease I and RuvC tend to bind their target DNA as dimers, and are speculated to interact with the phosphate backbone of the DNA *via* clefts lined with basic amino acid residues [67, 70]. Beyond these similarities, however, they appear to differ substantially in amino acid sequence, tertiary and quaternary structure, and in the details of the way in which they bind and cleave DNA. Footprinting experiments suggest that the RuvC protein binds to a more open DNA structure with unstacked bases at the cross-over point [76], while T4 endonuclease VII appears to bind a fully stacked form of the junction, as is predicted by the stacked-X model [77]. T7 endonuclease I contacts all four strands at the base of the junction [20], while T4 endonuclease VII occupies only two of the four strands, also at the base of the junction, and its cleavage sites are related by a centre of inversion [77].

An interesting feature common to these three proteins is the observation that, upon protein binding, the junction DNA structure is distorted to a more open structure which has been proposed to be important for cleavage by the enzymes (reviewed in [67]). It has been suggested that junction DNA recognition by T4 endonuclease VII may involve the angle between the segments of DNA found on either side of a branchpoint, which is expected to be about 120° in a stacked-X structure [78]. The X-ray crystal structure of the enzyme appears to support this theory [70].

1.7 b) HMG proteins

HMG proteins are a class of eukaryotic proteins which bind to cruciform DNA, as well as to negative supercoils, crossovers and the axially kinked cis-platinated DNA [79]. The first member of the family to be discovered, and the best characterized to date, is the abundant chromosomal protein HMG1. The HMG box is a 70-80 amino acid sequence found in the HMG family proteins, and many other DNA binding proteins ranging from components of chromatin architecture to transcription factors (reviewed in [80]). Though many HMG box containing proteins bind to linear ds DNA in a sequence specific manner, all HMG domains possess the ability to bind four-way junctions ([81] and references therein). Interestingly, HMG boxes, including those mediating structure-specific binding, bend their target DNA upon binding [82]. It is two of these HMG boxes which mediate the structure-specific cruciform binding of the HMG1 protein [83]. Mutational studies which caused major unfolding of the protein suggest that the cruciform-specific binding may be a property of a primary structure element of the HMG box [84, 85]. Despite considerable investigation, the function of HMG1 remains unclear.

Only very recently have the first footprinting experiments of cruciform DNA bound to HMG1 made available detailed information on the points of physical contact between the two [86]. They show extensive protection of three of the elbows of the junction, and lesser protection of the fourth, indicating an asymmetric binding mode. There is also evidence for the ability of the HMG1 protein to convert the four-way junction DNA from the stacked-X to a more open conformation, upon binding [81]. In contrast to many of the endonucleases which bind DNA junctions with full affinity under the cation conditions most conducive to the stacked-X conformer, HMG1 binding is inhibited by increasing amounts of

Mg^{2+} ions [81]. This emphasizes the differences between the various cruciform binding proteins and supports the idea of distinct binding modes and biological roles.

1.7 c) Other proteins

Another protein with a low level of sequence homology to the HMG-class of proteins has been cloned from *Ustilago maydis*: HMP1 [87]. It does not possess an HMG box or homology to any other known cruciform binding proteins, nor does it cleave DNA, and may be a member of a new family of such proteins. Human p53 has also been shown to bind specifically to cruciforms, with the interaction being predominantly with the junction of the DNA structures [88]. This binding increases the rate of resolution by the bacteriophage enzymes T4 endonuclease VII and T7 endonuclease I [88]. Neither of these interactions have been characterized in detail to date.

1.8 Footprinting

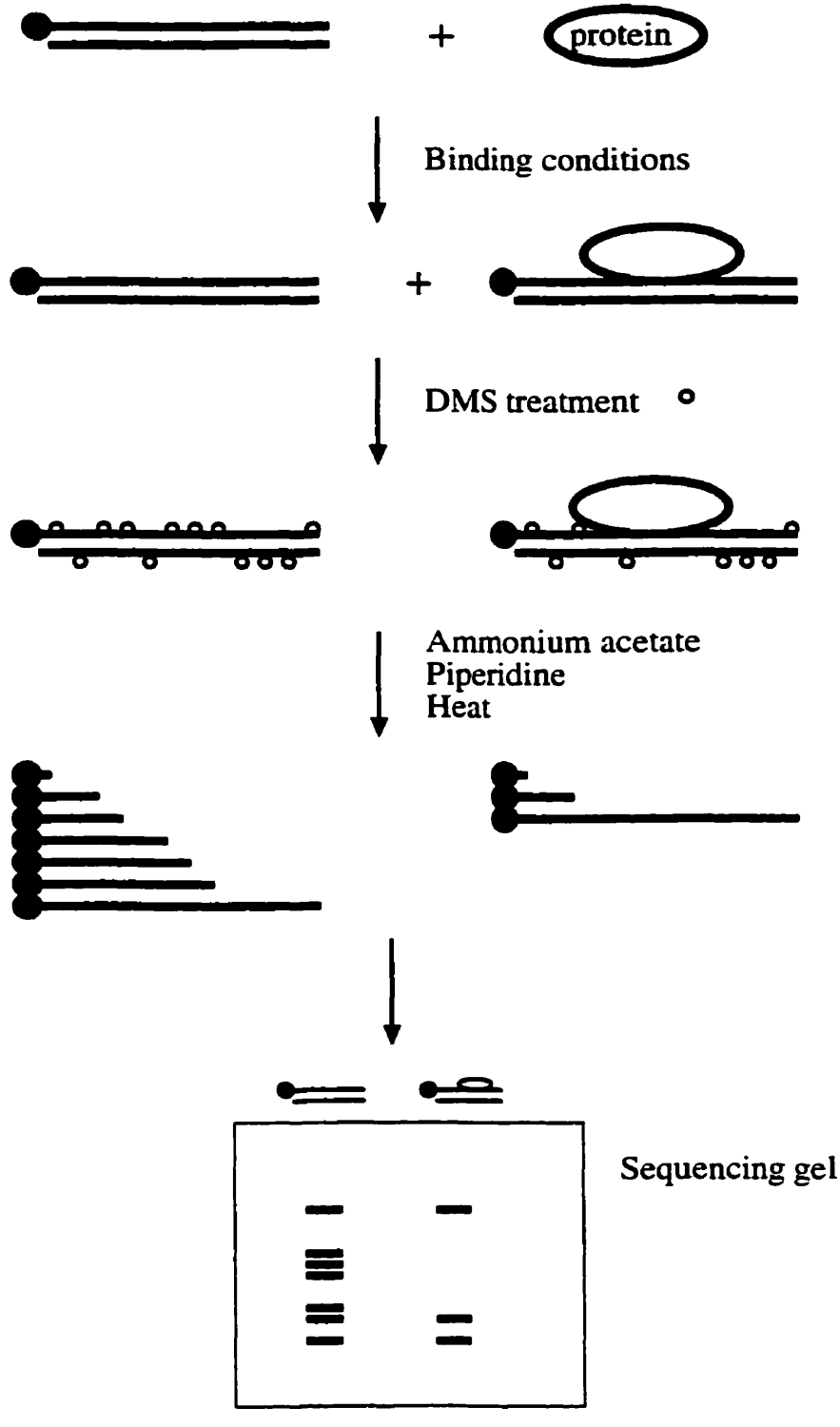
Arguably, the most accurate description of the binding of a protein to DNA is obtained from an X-ray crystal structure. A well resolved structure elucidates the relationship between the protein and DNA, providing information about interatomic distances and allowing detailed analysis of the forces governing the interaction. The main theoretical disadvantage of the crystal structure is that it depicts, by necessity, a solid static system which is far from the solvated and dynamic situation in a living system such as a cell. The practical disadvantage of the technique is the sometimes extreme difficulties involved in the preparation of a crystal, of the protein-DNA complex, of high enough quality for X-ray analysis. A much more amenable, and also very informative, technique is DNA footprinting. (RNA footprinting is also a rapidly developing technique.) Footprinting uses indirect

methods to investigate the interaction of a protein with its target DNA, providing information on the points of contact between the two, the relative importance of these points, the overall mode of binding and sometimes even more intricate details of the system. The experiments are of two main types: protection and interference. Essentially, the former consists of identifying the sites of protection from modification of the DNA by its interaction with the protein, whereas the latter identifies the DNA sites essential to the interaction by the fact that their modification precludes protein binding. The information obtained from the two is thus complementary. The modifications employed cause, or can be followed by, specific cleavage upon reaction of the DNA with appropriate chemical agents. The positions of these cleavages can be seen by subjecting the resulting fragments to electrophoresis, in parallel with a Maxam-Gilbert sequencing ladder of the same DNA, on a denaturing sequencing gel [89, 90]. The utility of footprinting was established in the late 1970's [91] and it remains a much used technique which is constantly being improved.

1.8 a) Protection footprinting

Protection footprinting (Figure 1.5) involves, first, binding the protein to the DNA of interest, and then exposing the complex to the modifying agent. The amount of agent to be used must be adjusted such as to obtain "single-hit kinetics", that is there must be a high ratio of DNA molecules which have been modified once to those which have been modified more than once. This occurs when approximately 70% of the DNA is not modified at all. The result is a population of DNA molecules which have been randomly modified at all sites except those which were protected by the presence of the protein, and an increased confidence that any differences in the amount of a fragment is in fact due to protein-DNA interactions [92]. The modification agents take advantage of the specific reactivities of the

Figure 1.5 The protection footprinting experiment, with dimethylsulfate (DMS) as the probe. The black circle denotes a radioactive end-label, required for visualization of the final products on the sequencing gel. A portion of the DNA is combined with the protein under conditions conducive to binding, and a portion is kept free from protein. Both are then treated with the probe, in this case DMS, which reacts with specific sites on the DNA. Where the protein is bound the DNA is protected from reaction with the probe. The DNA is then cleaved at the sites of modification, in this case by treatment with ammonium acetate and piperidine, with heating. The end-labeled fragments of the free and protein-bound DNA are then compared by separation on a denaturing sequencing gel, and visualized with an autoradiogram.


protein
Binding conditions
DMS treatment
Ammonium acetate
Piperidine
Heat
Sequencing gel

various functional groups of the nucleotide base, sugar and phosphate moieties [90]. Cleavage is then obtained uniquely at the sites of modification, and therefore not at the sites of protection. Comparison of the distribution of the resulting fragments with those obtained from an identical treatment of naked DNA presents the region of protection as bands of decreased, sometimes to near zero, intensity. Occasionally the presence of the protein, particularly at the extremities of the region of interaction of the DNA, will influence the local environment in such a way as to enhance the reactivity of the DNA with the modifying agent. This is seen as an increase in band intensity and is also a source of valuable information. Protection and enhancement can also be indicative of changes in the conformation of the DNA resulting from protein binding, without necessarily indicating direct contact with the protein at that point [89].

1.8 b) Interference footprinting

The interference footprinting protocol differs from that of the protection footprint primarily in the order in which the steps are carried out. In this case the DNA is exposed to the modifying agent prior to complexing with the protein. The subsequent binding reaction yields populations of free DNA and DNA bound to protein, defined by whether or not the modification interferes with the binding. Separation of the two populations is usually carried out by nondenaturing PAGE or nitrocellulose filtration [90]. This step may also be used to decrease the background in a protection experiment. The subsequent cleavage of the separated populations results in fragments in the "bound" group corresponding to positions of modification with no effect on protein binding, and in the "free" group those which prevent it. Comparison with a sequencing ladder allows for precise location of these sites on the DNA sequence [89]. The missing contact and missing nucleoside methods are variations on the interference experiment theme. In the case of the

missing contact experiment [93], the modification which is performed is the removal of a base (depurination or depyrimidination), whereas in the case of the missing nucleoside experiment [94] the modification is the removal of the sugar and the base.

1.8 c) Commonly used footprinting agents

A summary of some of the major modifying agents used for protection and interference footprinting is presented in Table 1.1. There are, of course, numerous other agents that may be employed. In some cases modifications of the major methodologies allow alterations in selectivity and therefore applications of the approaches. For instance alkylation with ethylnitrosourea (ENU) rather than dimethylsulfate (DMS) allows examination of the sugar phosphate backbone protection rather than just that of the guanine and adenine bases [95]. Since the information available from the various techniques is often complementary, the best strategy is usually to do a series of experiments using different modifying agents. Certain methods, such as DNaseI digestion, have been refined for the acquisition of quantitative information, such as individual-site kinetic progress curves, about the relationship between the protein and its target DNA on a millisecond time-scale [96]. Another quantitative application of DNaseI footprinting allows determination of the co-operativity of binding of more than one DNA binding protein [97]. Photofootprinting has progressed through the use of γ-rays [98] and most recently synchrotron generated X-rays [99]. In an interesting new approach, multiple-hit footprinting, rather than the single-hit kinetics described above, has been introduced to characterize conformer population distributions and reactivity rate constants in systems where protein binding involves conformational changes of the DNA [100]. In addition to furthering our understanding of protein-nucleic acid interactions

Table 1.1 Summary of some of the most commonly used modifying agents in protection and interference footprinting.

Agent	Protection Interference or Both	Reaction Target	Modification	Reference
Dimethyl sulfate (DMS)	Both	G and A	Methylation	[101]
DNaseI	Protection	Phosphodiester backbone	Cleavage	[102]
1,10-Phenanthroline copper	Both	Deoxyribose	Cleavage	[103]
Hydroxyl radical	Both	Deoxyribose	Cleavage	[104]
Potassium permanganate	Both Primarily used to detect single-stranded DNA	Single-stranded T	Ring opening of the base	[105]
Ultraviolet photons	Protection	Primarily adjacent pyrimidines	Photodimerisation	[106]
Exonuclease III	Protection	Phosphodiester bond	Cleavage	[107]

footprinting techniques have proved very fruitful in the study of other ligands such as small molecules with pharmacological potential [108].

1.8 d) *In vivo* footprinting

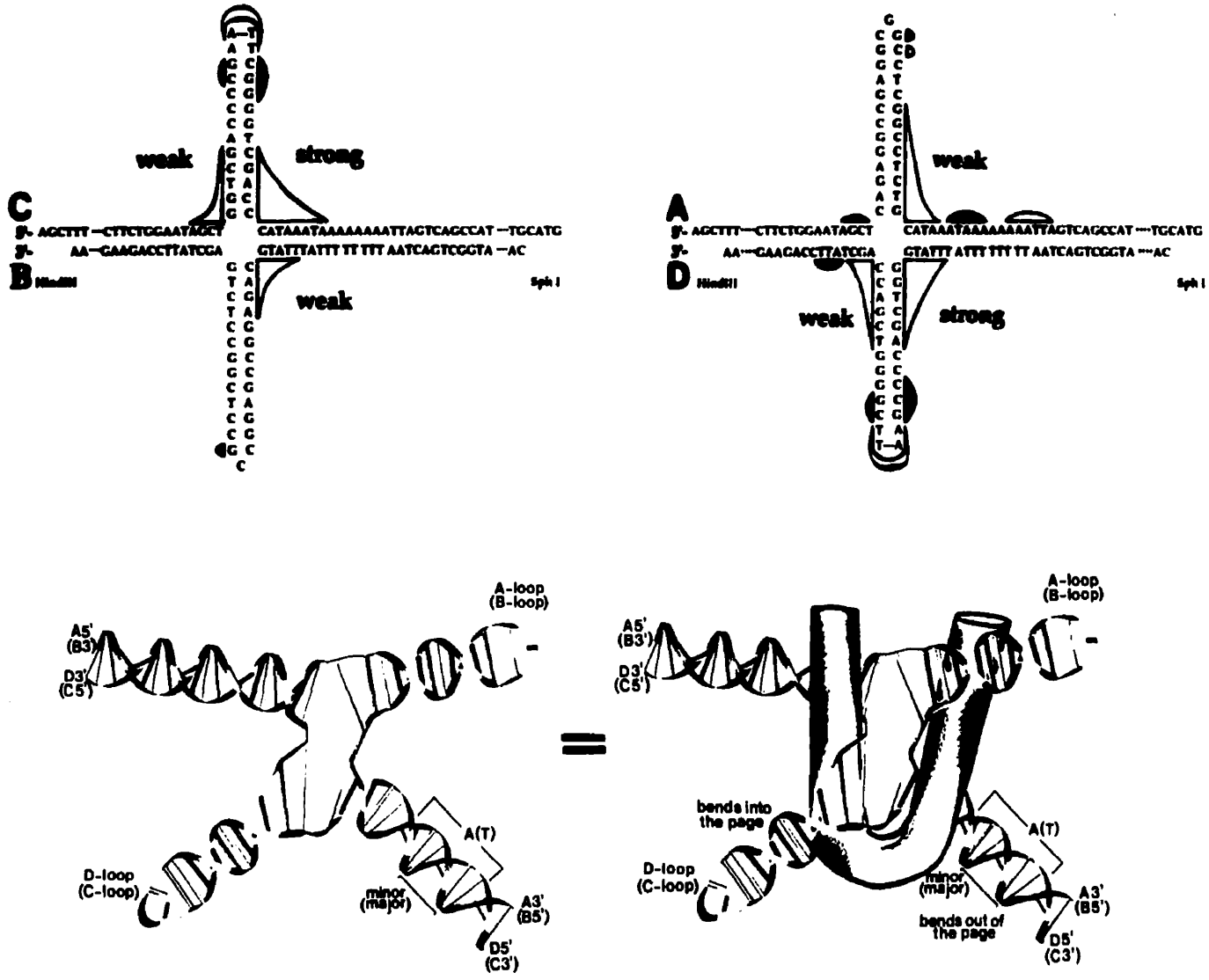
Another important variation of the footprinting technique is the *in vivo* experiment, sometimes referred to as genomic footprinting ([109], reviewed in [110-112]). There are many aspects of the biological systems we study, about which we do not know enough to be able to accurately mimic them in an *in vitro* experiment. As a result, the pertinence of *in vitro* data to the physiological situation is often unclear, and it is important, wherever possible, to do additional work in the context of living cells. The principles of *in vivo* footprinting are exactly the same as for the *in vitro* experiment. The modifying agent is applied to whole cells (or, sometimes, to isolated nuclei) and the footprinting is necessarily a protection assay. Many of the same reagents may be used, provided that they enter the cell without causing excessive damage, and without an impractical loss of their reactivity. For this reason small chemicals such as DMS, bromoacetaldehyde, potassium permanganate, osmium tetroxide, and hydroxyl radicals are amongst the probes of choice. Enzymes such as exonuclease III, DNaseI and micrococcal nuclease have also been used, but require cell permeabilization or nucleus isolation steps. Photofootprinting is also very amenable to the *in vivo* approach. Cross-linking the DNA-protein complexes with formaldehyde or *via* UV-irradiation is sometimes used to improve the stability of the interactions, and thus clarify the footprint [110]. The recent development and refinement of the ligation-mediated polymerase chain reaction (LMPCR) has greatly improved the sensitivity of genomic footprinting (reviewed in [113]).

1.9 Hydroxyl Radical Footprinting of the CBP-Cruciform Interaction

The detailed investigation of the interaction of CBP with cruciform DNA was initiated with hydroxyl radical protection footprinting [41]. These experiments examined the pattern of backbone cleavage of the 21/29 heteroduplex, described above (See Section 1.3) by the radical in the presence and absence of the CBP-enriched fraction of HeLa cell extracts (Figure 1.6, upper panel). They showed that the protein binds in an asymmetric fashion to the elbows of the junction portion of the cruciform, and causes distortion of the DNA structure as a result of binding. This constitutes a novel type of interaction of a protein with cruciform DNA [41]. The patterns of protection and enhancement on the individual strands allowed the construction of a model of the cruciform-protein complex (Figure 1.6, lower panel).

The most striking characteristic of these patterns is that three of the elbow regions are protected while one remains relatively unprotected from hydroxyl radical attack. Protection is also seen at the tip of one of the cruciform arms, flanked by hypersensitive regions. The model evokes an overall structure of the CBP similar to that seen in the 14-3-3 crystal structure [60, 61], (compare Figure 1.6, lower panel to Figure 1.4) to explain this pattern. The protein wraps around three of the four elbow regions, and induces structural changes in one stem-loop arm which may bring it close enough to the protein to be protected, or result in decreased reactivity without direct interactions. While there is some evidence for differences in the fine structure of the two complementary 21/29 cruciforms, the overall patterns of reactivity and induced interactions are very similar. The cruciform itself is modeled as a distorted tetrahedral structure (Figure 1.6, lower panel) [41]. Though there are some differences in the geometry of the structures,

Figure 1.6 Modeling of the CBP-cruciform interaction from hydroxyl radical footprinting. Upper panel: Enhancement and protection of the nucleotides of the two cruciforms of the 21/29 system to hydroxyl radical attack by the presence of CBP. Outlined areas indicate regions of protection, filled areas indicated regions of enhanced reactivity. Lower panel: The model of the cruciform as a distorted tetrahedron, and the mode of binding of CBP proposed from the hydroxyl radical footprinting data. Reproduced, with permission from Oxford University Press, from [41].


C
B
A
D
weak
strong
weak
weak
strong
Sph I
Sph I
A-loop
(B-loop)
A5'
(B3')
D3'
(C5')
D-loop
(C-loop)
A(T)
minor
(major)
A3'
(B5')
D5'
(C3')
=
bends into
the page
bends out of
the page

this tetrahedron is similar in crucial aspects to the stacked-X structure generally seen in other studies employing the same ionic conditions ([41] and See Section 1.2 b)). The co-crystal structure of the Cre protein and one of its four-way DNA junction substrates demonstrates that there are other instances of the bound four-way junction deviating somewhat from both the principal solution-structure models [114].

1.9 a) Inversion of the orientation of presentation of the two complementary cruciforms

Upon close examination, an intriguing feature of the protection and enhancement patterns may be discerned (Figure 1.6, upper panel). Considering that, with the exception of the palindromes, the sequence of strand A is identical to that of strand C and similarly strand B corresponds to strand D, it might be expected that the correspondence of the footprinting pattern be between strands A and C, and B and D. However, the converse is true. The two elbows protected in the AD cruciform are on the D strand whereas in the BC cruciform they are on the C strand. Thus the correspondence is between complementary rather than corresponding strands. Diagramatically, the protection and enhancement patterns may be superimposed by rotating the schematic of one cruciform (Figure 1.6, upper panel) 180 ° about the branch axis. This rotation would align the 3' end of strand B with the 5' end of strand A (Figure 1.6, lower panel). Because these two strands are complementary they will present oppositely either a major or minor groove at any position along their length (Figure 1.7). This suggests, therefore, that CBP is contacting a major groove in one cruciform, and at the same position, a minor groove in the other. Interestingly, no such inversion of protection patterns is seen when hydroxyl radical footprinting is carried out of these same cruciforms in the presence of the 2D3 anti-cruciform antibody [115]. It would appear, then, that this

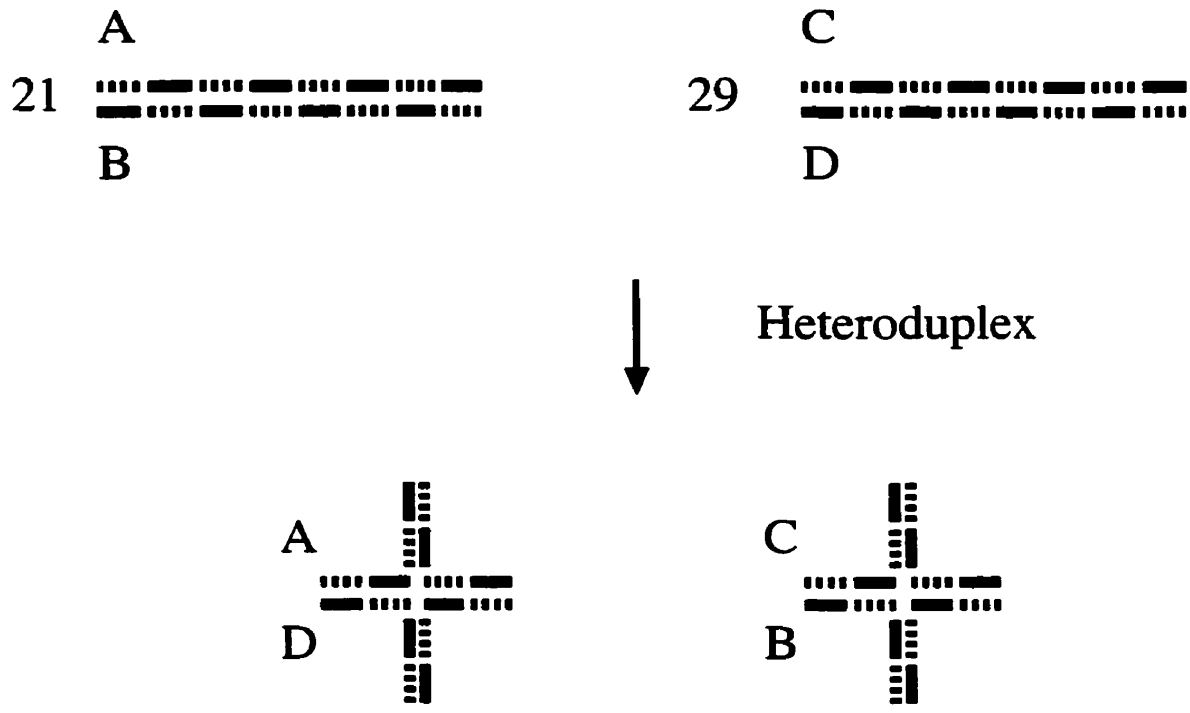


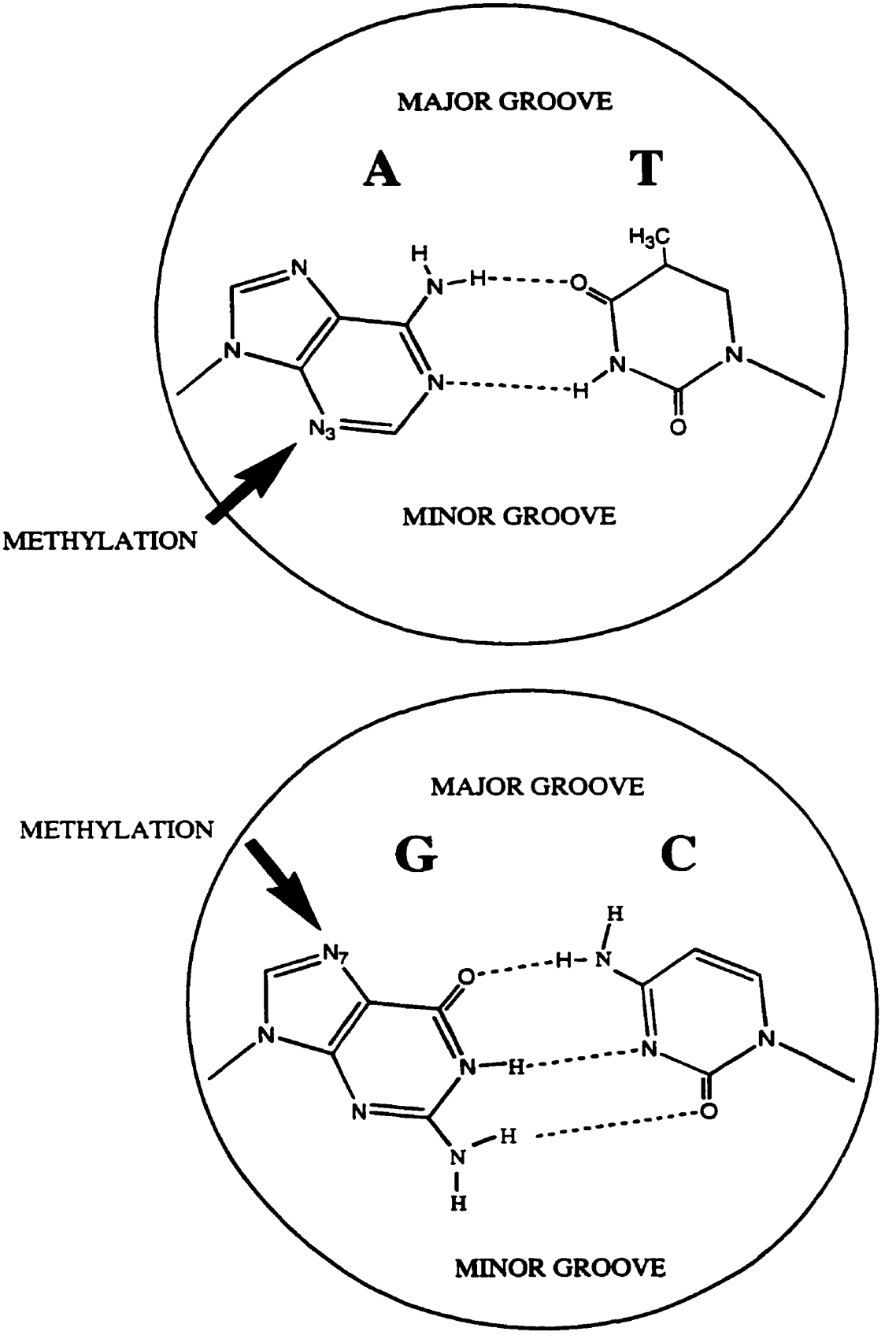
Figure 1.7 Schematic representation of the relative positions of major and minor grooves on linear and cruciform DNA. Dashed lines represent minor grooves, solid lines represent major grooves. This representation is purely schematic and does not attempt to accurately represent the locations of the major and minor grooves of the 200-bp fragments used in this study. Complementary strands (for example A and B) present, oppositely, major or minor grooves along their lengths. Therefore, upon heteroduplexing to form cruciforms, they will present, at identical positions on the cruciform, opposite grooves, to a binding protein.

is a characteristic peculiar to the recognition of cruciforms by CBP. An investigation of this trait could provide useful information about the mode of structure-specific binding employed by this protein, and perhaps also about its biological implications. The observed inversion of major/minor groove presentation of the cruciforms to CBP provides an opportunity for the verification of the model of this interaction proposed from the hydroxyl radical footprinting study.

1.10 DMS Protection Footprinting to Verify the Proposed Model

DMS is a methylating agent whose base specificity earned it a place as one of the reagents in the original Maxam-Gilbert sequencing technique [101]. This reagent methylates the N-7 of guanine in the major groove and the N-3 of adenine in the minor groove (Figure 1.8). The resulting positive charge causes an instability which, under basic conditions, leads to opening of the ring structure of the purine, making it susceptible to displacement and β-elimination by piperidine [116]. Its groove specificity makes the probe particularly useful for the determination of the major/minor groove presentation of DNA to any protein with which it interacts [117]). Regions of adenine protection indicate contact of the minor groove with the protein, whereas proximity to the major groove is demonstrated by protection of guanine bases. This provides precisely the specificity required to test the model proposed by the hydroxyl radical footprinting for the cruciform-CBP interaction. The inversion of cruciform BC with respect to cruciform AD when bound to the CBP, manifested by the major/minor groove presentation to the CBP, would yield a signature footprint with DMS. Positions of adenine protection (minor groove presentation to the protein) on strand A would

Figure 1.8 Sites of DMS methylation. The two principal sites of DNA methylation by DMS are the N3 of adenine, *via* the minor groove, and the N7 of guanine *via* the major groove. This methylation activates the DNA towards cleavage at this point, by piperidine. Reproduced, with permission from Academic Press, copyright 1995, from [117].


MAJOR GROOVE
A
T
METHYLATION
MINOR GROOVE
MAJOR GROOVE
METHYLATION
G
C
MINOR GROOVE

correspond to positions of guanine protection (major groove presentation) on strand C, and vice versa. Thus, DMS protection footprinting of the four strands would allow verification of the major/minor groove presentation of the two cruciforms to the CBP, and thus provide support for, or refute, the model proposed by the hydroxyl radical footprinting. This knowledge would provide further insight into the nature of the interactions between the protein and DNA and perhaps some new clues about this mode of structure-specific binding.

2. MATERIALS AND METHODS

2.1 Cruciform-CBP System

2.1 a) DNA substrates

The plasmids pBS/RGM21 and pBS/RGM29 were used for heteroduplex formation ([6],[7], [18] Figure 1.3). The plasmids were amplified in 1161 *E. coli* bacteria and purified using QIAGEN plasmid purification kits (QIAGEN Inc.). Strand D of pBS/RGM29 was selected for the focus of these experiments. pBS/RGM21 and pBS/RGM29 were independently doubly-digested with *Sph*I and *Hind*III. The pBS/RGM21 DNA was precipitated from the digest with 0.3 M sodium acetate, 0.005 % linear polyacrylamide and ethanol on dry-ice followed by centrifugation. The pBS/RGM29 digest was extracted with phenol, iso-amyl alcohol/$CHCl_3$ (1:24),$CHCl_3$. The final aqueous fraction was passed through a microcon-50 microconcentrator (Amicon, Inc.). 3' end-labeling was achieved by AMV reverse transcriptase (Roche Molecular Biochemicals) with [α-^{32}P]-dATP (Mandel Scientific Company Ltd.). Labeled DNA was separated from free nucleotides on a G-50 Sephadex column (Pharmacia Biotech). The labeled DNA was precipitated with 0.3M sodium acetate, 0.005 % linear polyacrylamide and ethanol on dry-ice followed by centrifugation. The digested pBS/RGM21 plasmid was resuspended in a small volume (20-50 µl) of water and combined with the pellet of the digested, labeled pBS/RGM29 at a ratio of labeled to cold DNA of 2:1, to increase the proportion of labeled heteroduplex. Together they were lyophilized and then resuspended in 0.5 M NaOH, 1.5 M NaCl. After 5 min at room temperature this mixture was placed at 68 °C for 2 h to overnight. It was then diluted to 0.01 M NaOH, 0.03 M NaCl by the addition of 50 x 10 mM Tris, pH

7.6, 1mM EDTA (TE) buffer and then precipitated with 0.11 M NaCl, 0.005 % linear polyacrylamide and ethanol on dry-ice, followed by centrifugation. The pellet was resuspended in a small volume (50-100 µl) of water and the homoduplex and heteroduplex separated *via* 4 % PAGE in 1 x 0.09 M Tris-borate, 0.002 M EDTA (TBE). A wet exposure autoradiogram was used to locate and excise the homoduplex and the slower running heteroduplex (Figure 3.1). The DNA was separated from the gel slices by isotachophoresis [118], with the omission of sodium dodecyl sulfate (SDS) from all buffers, and quantitated by comparing band intensities on an ethidium bromide stained polacrylamide gel, to those of a quantitative molecular weight marker (*Hae*III digested pBluescript-KS(-)).

2.1 b) Preparation of the CBP-enriched fraction

The CBP-enriched fraction was prepared as previously described [119], [40] by Dr. M.T. Ruiz. Total HeLa cell extracts were fractionated on a DEAE-sephadex column, and cruciform binding fractions were applied to a heparin column from which the flow-through contained CBP. The fraction used for all experiments reported herein was the unbound flow-through from the Affi-Gel Heparin Gel column (BioRad). The total protein concentration is 5 µg/µl in Buffer B (0.01 M KH_2PO4, pH 7.4, 0.15 M NaCl, 2.5 mM EDTA, 1 mM PMSF, 2 µg/ml aprotinin, $1x10^{-7}$ M pepstatin A, 5 % glycerol).

2.1 c) Electrophoretic Mobility Shift Assays (EMSAs)

Cruciform binding activity was assayed by combining end-labeled cruciform and the CBP-enriched fraction in 20 mM TrisHCl, pH 7.5, 1 mM EDTA, 1 mM dithiothreitol (DTT), 3 % glycerol, 0.1 mg/ml poly-deoxy-inosinic-deoxy-cytidylic acid (polydIdC) [120] on ice for 20 min, then adding SDS-free loading dye and separating the species by 4 % or 8 % PAGE, in 1 x TBE buffer at 180 V

for 2 and 5.5 h, respectively. Analytical gels were dried and used to expose Kodak X-OMAT XAR-5 (Kodak) film at -70 °C. Wet exposures performed at room temperature were used to locate and excise slices containing species of interest from preparative gels. Controls for these EMSAs consisted of an identical reaction mixture containing an equivalent volume of Buffer B in place of the CBP-enriched fraction, and one in which the homoduplex DNA was combined with the CBP-enriched fraction under identical conditions. For preparative EMSAs these controls were subjected to the same electrophoresis and isotachophoresis steps as the protein containing reactions.

2.1 d) DMS methylation

Reactions were performed in 1.5 ml eppendorf tubes. Reagents for DMS methylation were from the Maxam-Gilbert Oligonucleotide Sequence Analysis kit (Merck) with the exception of ethanol (Commercial Alcohols Inc.), high performance liquid chromatography (HPLC) H_2O (Baxter) which was used for all steps requiring water, and 20 mM ammonium acetate, 0.1 mM EDTA, pH 7 used for G>A specificity. For a sequencing (non-footprinting) reaction, DNA corresponding to the desired amount of radioactivity (usually about 50000 dpm as determined by Cerenkov counting without a scintillant, prepared as above) was lyophilized and resuspended in DMS-buffer (50 mM sodium cacodylate, pH 7.0, 1 mM EDTA). Calf thymus carrier DNA was added to a final concentration of 35 µg/µl. Pure DMS was added; kit conditions suggest a final concentration of 0.5 %, but this resulted in high levels of over-methylation of the DNA, and 0.1 and 0.05 % were found to be more suitable for the substrates in this study (0.05 % was used for experiments omitting the stop-reagent). Methylation reactions were placed at 20 °C for 4 min. The desired temperature was obtained by using a heating block in a 4

°C room, or by adjusting a large beaker full of tap water with ice. Following 4 min of reaction with DMS, stop reagent was added to final concentrations of 0.2 M β-mercaptoethanol (β-ME) and 0.3 M sodium acetate, pH 7.0, followed immediately by addition of 1 ml ethanol, and the tubes were plunged on dry-ice. After at least 20 min on dry-ice the tubes were centrifuged at 14000 rpm for 30 min at room temperature. The supernatant was decanted, the pellet resuspended in 90 µl water and re-precipitated with 0.3 M sodium acetate and ethanol. This second pellet was lyophilized, and then 100 µl 20 mM ammonium acetate, 0.1 mM EDTA, pH 7, were added and the reaction incubated at 90 °C for 15 min (this step imparts the G>A specificity). 90 µl water (or TE, pH 7.6, as specified in the Results) were added, followed by 10 µl piperidine, and the reactions returned to 90 °C for 30 min. Piperidine was then removed by lyophilizing overnight, followed by two washes with 50 µl water, each removed by lyophilization. The final pellet was resuspended in 50 µl water and passed through a micron-10 microconcentrator (Amicon, Inc.). The radioactivity of the eluate was counted in a liquid scintillation counter (Beckman) and aliquots of equivalent counts were lyophilized, resuspended in 2 µl loading dye (95 % deionized formamide, 20 mM EDTA, 0.05 % xylene cyanol, 0.05 % bromophenol blue), boiled for 5 min, and kept on ice until loaded. Reaction products were resolved on a denaturing (7 M urea) 8 % polyacrylamide sequencing gel in 0.5 x TBE buffer using the LKB Macrophor system (Pharmacia). One of the glass plates of this apparatus has a circulating water system *via* which the temperature of the gel can be precisely controlled. This was necessary because

the inverted repeat in the DNA fragment forms secondary structures resulting in band compression on a conventional sequencing apparatus (data not shown). This problem is avoided by maintaining a high enough gel temperature (60-65 °C) throughout the run. In order to ensure maximum quality of the sequencing/footprinting gels the urea (GIBCO or BDH) and ammonium persulfate (APS, GIBCO) were stored in a dessicator, and a fresh 10 % APS solution was prepared for each gel. The acrylamide was not heated during dissolution of urea and care was taken not to pre-run the gel for more than 1.5 h. The gel was pre-run at constant voltage, between 2500 V and 3000 V in 0.5 x TBE buffer, and following sample loading the run was conducted under the same conditions until the xylene cyanol front had traveled 28-30 cm. A wet exposure of the gel, at -70 °C, to Kodak X-OMAT XAR-5 allowed visualization of the results.

2.1 e) DMS footprinting

DMS footprinting followed essentially the same protocol as that outlined above for DMS methylation, with a few modifications. The DMS treatment followed the binding reaction outlined above (Section 2.1 c)), and was therefore in the binding conditions rather than DMS-buffer. No calf thymus DNA was added, as the polydIdC of the binding reaction provides carrier DNA. The DMS reaction was stopped by the addition of neat β-ME. A titration experiment showed that a final β-ME concentration of 1.9 M (10-fold greater than that suggested by the kit conditions) resulted in the most efficient quenching of the reaction and this was used for most experiments. A final β-ME concentration of 50 mM was used for the experiment with minimal β-ME and the experiment with minimal DMS (0.05 %) did not require a stop reagent. In those experiments without a preparative

polyacrylamide gel, the DMS-treated binding reaction was precipitated with ethanol and then the rest of the DMS sequencing protocol was followed as described above (See Section 2.1 d)). When a polyacrylamide gel was used to separate the various species, the stopped reaction was loaded onto a 4 % or 8 % gel as rapidly as possible, and the gels run as described for the EMSAs (See Section 2.1 c)). Addition of glycerol to a final concentration of 6.25 % to the polyacrylamide gel had no effect on the band shift pattern (data not shown). For the minimal β-ME experiment, the preparative gel was pre-run for 2 h with 0.001 % (w/v) thioglycolate in the 1 x TBE buffer. Fresh 1 x TBE with 0.001 % thioglycolate buffer was used for the actual run of the gel. Wet exposures performed at room temperature were used to locate and excise slices containing species of interest from preparative polyacrylamide gels, and the DNA was eluted from the gel by isotachophoresis, omitting SDS from all buffers [118]. The eluted fractions were pooled (when multiple gel slices of the same species were excised from different lanes) and concentrated with a microcon-50 microconcentrator (Amicon, Inc.). The protocol described above was then followed from the 20 mM ammonium acetate, 0.1 mM EDTA, pH 7 step onwards. If the results of the denaturing sequencing gel were unclear, the remainder of the samples were extracted with phenol, iso-amyl alcohol/$CHCl_3$ (1:24), $CHCl_3$ and then passed through a microcon-10 microconcentrator (Amicon, Inc.), and then subjected to denaturing sequencing gel electrophoresis.

2.1 f) Band quantitation

Image files for the quantitation of the sequencing gel band intensities resulting from DMS treatment experiments were generated by scanning autoradiograms with the Millipore BioImage system, or by using a phosphoimaging

screen and the FUJIX Bio-Imaging Analyzer. Quantitation of the bands was done with Millipore BioImage software. Pairs of bands that were too close together to resolve, in the 3' region beyond the cruciform structure itself, were quantitated as a single band. Each lane was normalized for the amount of radioactivity loaded by comparing the average of the three bands directly 3' of the region of interest. The log of the ratio of the normalized values was then plotted against band position. This is a useful method for the examination of footprinting data, as enhancement of reactivity at a given position results in a positive value, whereas reduced reactivity (protection) gives a negative value [115]. In order to estimate the significance of the deviations plotted, the same analysis was carried out on two independent sequencing gels of the same sequencing reaction of the heteroduplex DNA (Figure 3.5 B).

2.2 Positive Control

To ensure the reliability of the DMS footprinting method and reagents, the protocol was carried out on a system for which the DMS footprint is known: nuclear factor I (NF-I) and its target DNA from adenovirus type 5 ([121] and H. Zorbas, Personal Communication).

2.2 a) DNA substrate

The pTAd5 plasmid, into which the NF-I binding site from adenovirus has been inserted, is as previously described [121], with the removal of a 178-bp *Sal*I fragment (obtained from H. Zorbas, University of Munich, Germany). The plasmid was digested with *Sal*I, and dephosphorylated with calf intestinal phosphatase (New England Biolabs) making the 5' ends available for labeling. Digestion with *Ava*I releases a 96-bp fragment containing one of the labelable ends and the NF-I binding site, and a 6-bp fragment containing the other labelable end.

T4 polynucleotide kinase (New England Biolabs) was used to label the 5' ends with [γ-^{32}P]-dATP (Mandel Scientific Company Ltd.). This results in a mixture of plasmid, 6-bp fragment, and 96-bp fragment; however, it is not necessary to separate the latter from the two former because the unlabeled plasmid is not detected in the subsequent experiments, and the 6-bp fragment is too small to interfere in any of the methods used.

2.2 b) EMSAs

Binding of the NF-I protein to its target DNA, the 96-bp fragment, was assayed by EMSA. The DNA preparation containing the end-labeled 96-bp fragment was incubated with the protein (provided at a concentration of 30-40 fmol/µl in 2 M KCl by H. Zorbas, University of Munich, Germany) in 25 mM HEPES/KOH, pH 7.9, 150 mM NaCl, 0.1 mg/ml polydIdC for 30 min at 20 °C. SDS-free loading dye was added and the samples were loaded onto a 4 % polyacrylamide gel, and subjected to electrophoresis for 1.5 h in 1 x TBE buffer at 180 V. The gel was dried and placed with a Kodak X-OMAT XAR-5 film at -70 °C. Protein-DNA molar ratios of 50- to 1000-fold were assayed for optimum binding. 1000-fold excess was selected for footprinting experiments.

2.2 c) DMS footprinting

Following the 20 min incubation of the binding reaction, with a 1000-fold molar excess of protein with respect to DNA, it was diluted 1:5 in the DMS-buffer. The DMS methylation protocol described above, with a final DMS concentration of 0.5%, was then carried out with a few modifications. Following the second precipitation of the DNA with ethanol, it was extracted with phenol, iso-amyl alcohol/$CHCl_3$ (1:24), $CHCl_3$, and re-precipitated before lyophilization. The piperidine step was carried out in TE, pH 7.6.

3. RESULTS

3.1 Choice and Preparation of Substrate DNA

This research sought to compare the DMS protection footprints of the four strands of the two model cruciforms formed by the 21/29 system (Figure 1.3), upon binding of the CBP present in an enriched fraction from HeLa cell extracts, in order to verify the model of CBP-cruciform interaction proposed from hydroxyl radical footprinting [41]. The relationship between the regions of protection of the adenine and guanine bases between the four strands would allow an interpretation of the major/minor groove presentation by the two cruciforms to the protein, which would support or refute the hypothesis of an inversion of orientation between the two (See Section 1.10). Information may be obtained about each of the strands individually by specific end-labeling with ^{32}P. In this way, despite the presence of all four strands, only the strand of interest is visualized on an autoradiogram. In order to establish the best conditions for this investigation, one of the four strands was selected. Considering that only the guanines and adenines will give signals, the strand with the greater concentration of these bases in the region shown to be contacted by CBP by hydroxyl radical footprinting was chosen. This was strand D (See Figure 1.6 upper panel).

Labeling strand D at the 3' end with α-^{32}P-dATP, following *Hind*III/*Sph*I cleavage of the pBS/RGM29 plasmid (See Section 2.1 a)), and heteroduplexing with cold digested pBS/RGM21, yielded three labeled species: the large fragment of the plasmid, the heteroduplex and the homoduplex, which were easily separated on a 4 % polyacrylamide gel (Figure 3.1). The latter two were excised following a wet exposure. Following isotachophoresis, the purified homoduplex and

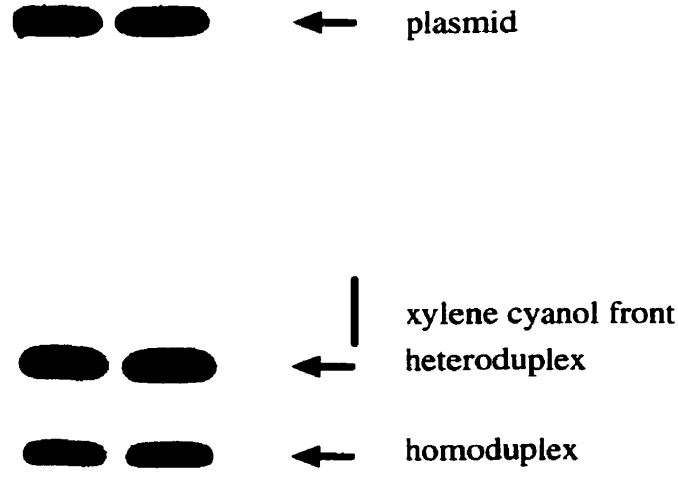


Figure 3.1 Preparation of end-labeled homoduplex and heteroduplex DNAs. Wet exposure autoradiogram of a preparative 4% polyacrylamide gel used to locate and excise the 3' end-labeled homoduplex and heteroduplex, showing the positions of all labeled species relative to the xylene cyanol front. The species were generated by *Hind*III/ *Sph*I cleavage of the pBS/RGM21 and pBS/RGM29 plasmids, 3' end-labeling of strand D, and heteroduplexing to generate the cruciform (See Sections 1.3 and 2.1 a) for details.)

heteroduplex DNAs were quantitated by comparing their band intensities to those of quantitative markers in an ethidium bromide stained polyacrylamide gel.

3.2 Saturation of CBP with Cruciform

The CBP-enriched fraction used for this investigation had a total protein concentration of 5 μg/μL. In order to estimate the proportion of this protein which possessed cruciform binding activity, we performed a titration of CBP with cruciform (heteroduplex) DNA. Band-shift reactions were carried out using a constant amount of CBP-enriched fraction and increasing amounts of cruciform DNA. The point at which the free heteroduplex (cruciform DNA) band became visible indicated saturation of CBP, i.e., all protein capable of binding the cruciform DNA had done so and any additional cruciform remained free in solution. The results of a representative experiment are presented in Figure 3.2. The two main shifted bands (sometimes accompanied by a much fainter third, more retarded, band) are characteristic of the band shift reaction between the CBP-enriched fraction and cruciform DNA [40]. Typically, 30 ng of cruciform DNA were required to saturate the CBP activity in 1 μL of the CBP-enriched fraction. At an average molecular weight of 660 Da per base pair, approximately 132,000 g represent 1 mole of the 200-bp cruciform. 30 ng, therefore, corresponds to approximately 200 fmol. Assuming that each cruciform is bound by one dimer of 14-3-3, at a molecular weight of 66 kDa [122], this corresponds to 15 ng of cruciform binding activity in 5 μg of total protein.

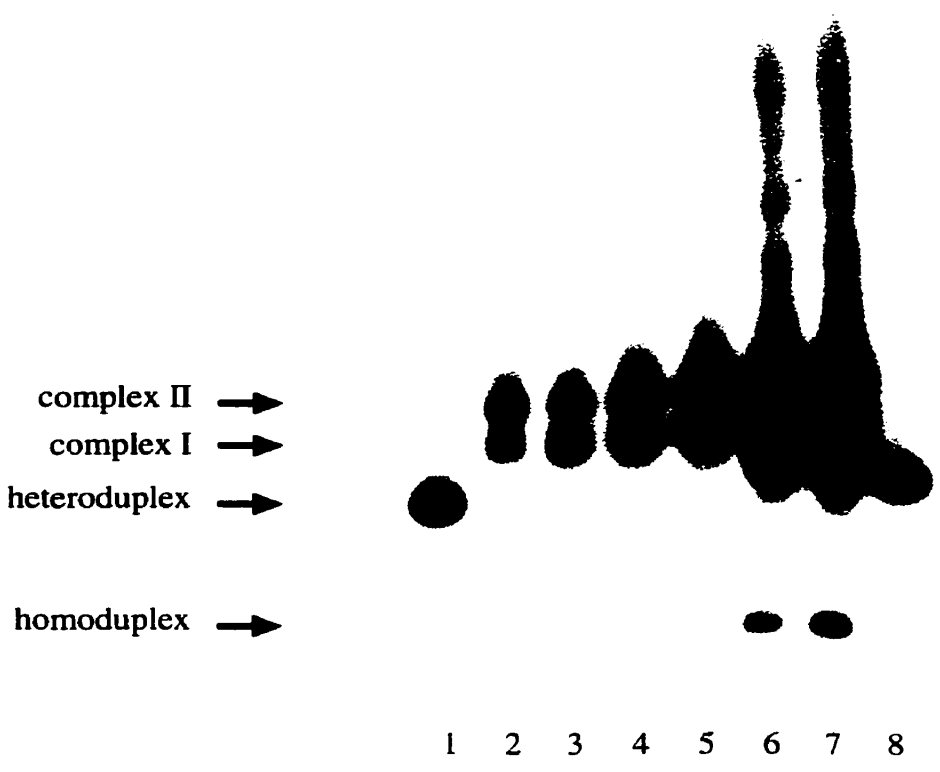


Figure 3.2 Saturation of CBP with cruciform. Titration of a constant amount of CBP-enriched fraction with increasing amounts of end-labeled cruciform (heteroduplex) DNA. Lanes 1 and 8 contain cruciform DNA alone; lanes 2-7 contain 1ul of CBP-enriched fraction (5ug total protein) with 3, 5, 10, 20, 40 and 60 ng of labeled cruciform, respectively, combined under binding conditions (See Section 2.1 c)).

3.3 Saturation of Cruciform with CBP

The ratio of CBP-enriched fraction to heteroduplex DNA required to bind all the cruciform DNA was also determined (Figure 3.3). A constant amount of cruciform DNA was combined in a binding reaction with increasing amounts of the CBP-enriched fraction (lanes 2-5) and compared to the migration of free cruciform not exposed to protein (lane 1). The point at which the free cruciform band is no longer visible marks 100% binding of the cruciform DNA by the CBP present. This is important since the footprinting method involves the comparison of the pattern of DMS reactivity of the DNA in the presence and absence of the protein. If a significant portion of the DNA exposed to the protein is not in fact bound, it will result in a background signal, due to the random methylation and cleavage of these end-labeled molecules, which could mask a footprint. This titration was repeated for each new preparation of radiolabeled DNA to give the most reliable basis for each footprinting experiment. The saturation conditions were then scaled up for the footprinting experiments.

Typically, 2 µL of the CBP-enriched fraction, corresponding to approximately 10 µg of total protein, gave 100 % shifting of 12 ng of cruciform DNA. That is, approximately 1 µg total protein of the CBP-enriched fraction, or 3 ng of CBP, based on the 15 ng / 5 µg total protein calculated above (See Section 3.2), is required to bind 1 ng of cruciform. This estimation of the amount of CBP-enriched fraction required to bind 1 ng of cruciform differs from that calculated from the saturation of CBP with cruciform, above, by a factor of five. There are a number of reasons why the 15 ng / 5 µg total protein calculated above (See Section 3.2), is rather imprecise: the molecular weight of the species involved are

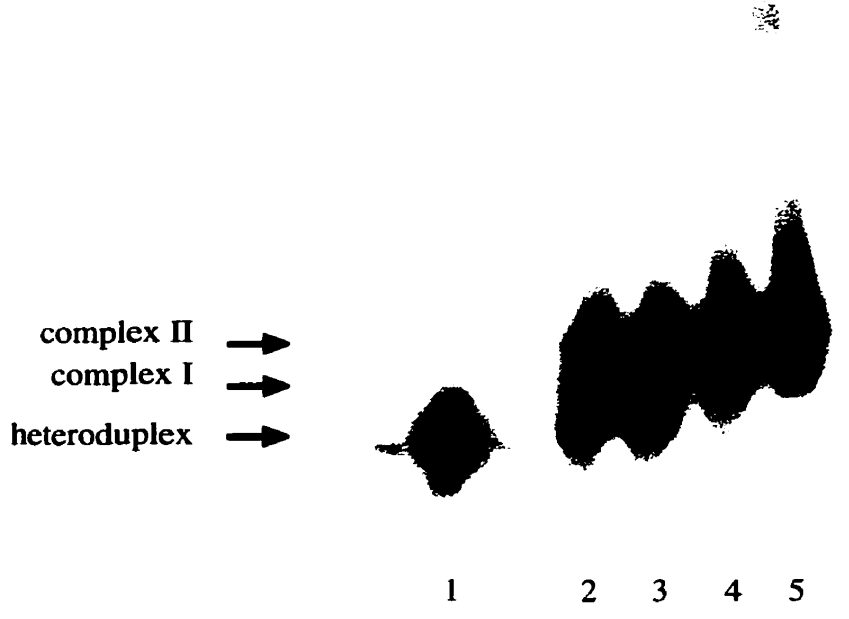


Figure 3.3 Saturation of cruciform with CBP. Titration of a constant amount of end-labeled cruciform with increasing amounts of CBP-enriched fraction. Lane 1 contains cruciform DNA alone; lanes 2-5 contain 12 ng of labeled cruciform with 0.5, 1, 2, and 3ul CBP-enriched fraction (5ug/ul total protein), respectively, combined under binding conditions (See Section 2.1 c)).

estimations and not exact, the DNA quantitation by comparative ethidium bromide band intensity has limited precision, and the possibility exists that a very small population of proteins other than the 14-3-3 dimer, may contribute to the shifting of the cruciform from its free running position on the polyacrylamide gel (but not in a large enough quantity to give a signal on the autoradiogram). Such proteins would contribute to the number of cruciform molecules required to saturate the cruciform binding activity and would lead to an over-estimation of the amount of CBP present. Despite the potential sources of error in this calculation, the result does give a useful estimation of the proportion of total protein in the CBP-enriched fraction which has cruciform binding activity.

Another possible reason for the 5-fold difference in the estimation of the amount of CBP-enriched fraction required to completely bind 1 ng of cruciform DNA, is the design of the experiments themselves. The titration yielding saturation of CBP with cruciform is better suited to the estimation of the amount of active CBP in the CBP-enriched fraction than a titration of the saturation of cruciform DNA with CBP. The point at which free cruciform is first seen in the saturation of CBP experiment (Figure 3.2, lane 6) indicates that all the proteins with the capability to do so are bound to cruciform. However, the point at which the free cruciform DNA is no longer seen in the saturation of cruciform experiment (Figure 3.3, lane 4) does not preclude the presence of proteins free in solution, available to bind more cruciform. Therefore, this titration may easily over-estimate the amount of CBP required to bind a given amount of cruciform, and is therefore less suitable for quantitative estimations.

3.4 Determination of $[DMS]_f$ Yielding Single-Hit Kinetics

Single-hit kinetics of the modifying agent with the DNA is essential to successful protection footprinting (See Section 1.8 a)). The kit conditions for DMS methylation (See Section 2.1 d)) resulted in such extensive over-reaction of the DNA that the unreacted band was barely visible on the autoradiogram. Decreasing the duration of DMS treatment to as little as 40 sec, from 4 min, did not eliminate this problem (data not shown). A titration of the final concentration of DMS employed demonstrated a significant increase in the amount of unreacted DNA when the DMS was decreased 3-fold (Figure 3.4, compare lanes 1 and 2), and a ladder of bands of even intensity when it was decreased 5-fold (lane 3). This corresponds to a final DMS concentration of 0.1 %, which was used for the majority of subsequent experiments (exceptions are mentioned below).

3.5 DMS Reactivity of Homoduplex versus Heteroduplex

Before attempting to assay the effect of the presence of CBP on the DMS reactivity of the DNA, the effect of the cruciform structure on this pattern was investigated. G>A sequencing under the modified kit conditions (See Section 2.1 d)) showed no consistent difference between the reactivities of the two (Figure 3.5 A). A possible reason for this inconsistency might be the fact that following piperidine cleavage, the piperidine is removed by resuspending the DNA in water, and then lyophilizing it. This may create a slightly acidic environment, which, when combined with the heat of the lyophilizer can cause non-specific purine cleavage (H. Zorbas, Personal Communication). Since differences were often difficult to confidently distinguish by examination of the DMS sequencing autoradiograms alone, quantitative assessment of the band intensities was also made

Figure 3.4 Titration of final DMS concentration to establish single-hit kinetics. DMS methylation of homoduplex DNA was carried out under identical conditions (See Section 2.1 d)), but varying the final DMS concentration between full-strength, as defined by kit conditions, = 0.5 % (lane 1), 1/3 = 0.17% (lane 2) and 1/5 = 0.1% (lane 3). The unreacted band is indicated at the top of the gel. The solid line highlights the bands representing the larger DNA fragments which are under-represented, and the dashed line highlights the bands representing the smaller DNA fragments which are over-represented, in the case of over-reaction best seen in lane 1). A final DMS concentration of 0.1% (lane 3) gave the most uniform ladder of DNA fragments. Asterisks indicate markings resulting from the scanning of this film in two sections, due to its length.

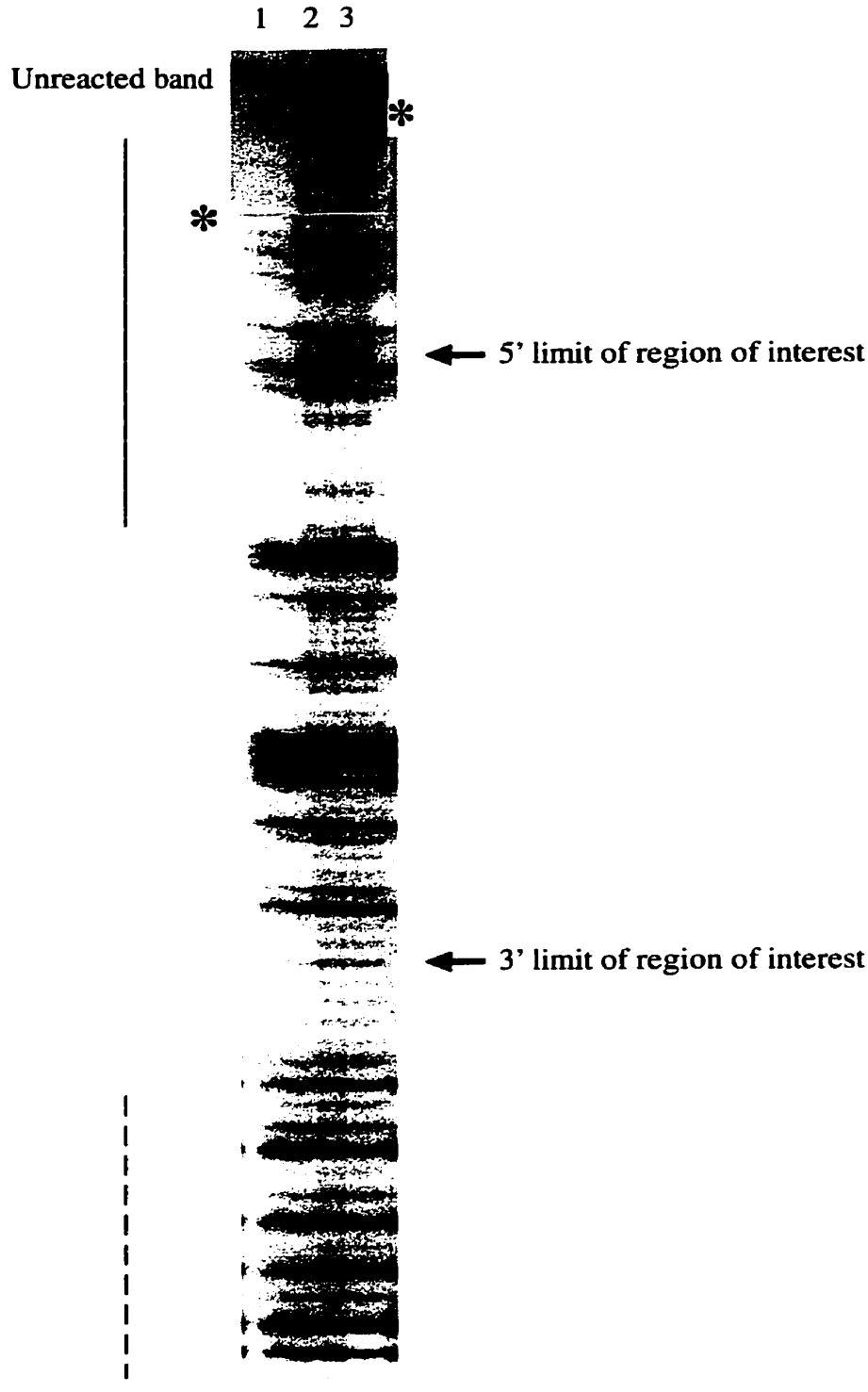
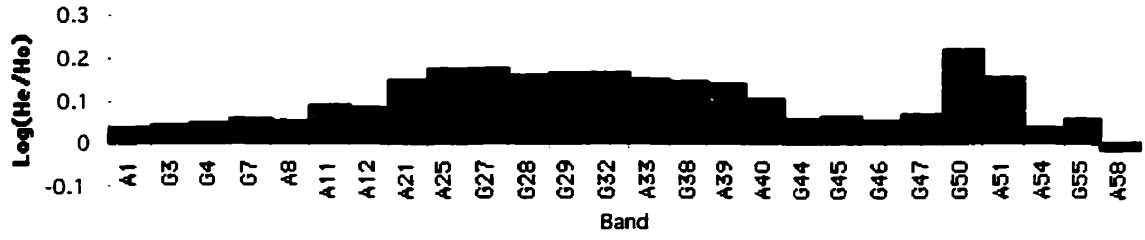
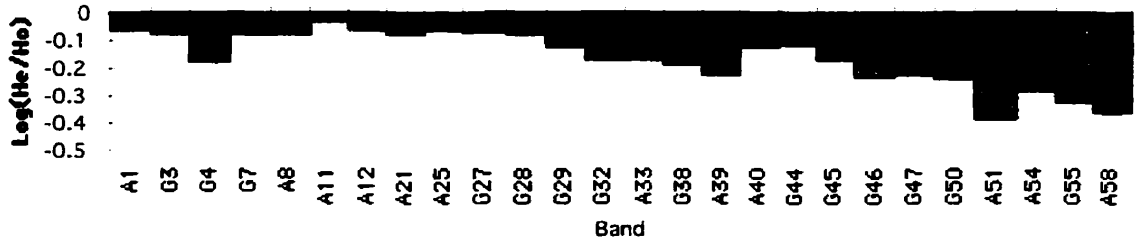

1 2 3
Unreacted band
5' limit of region of interest
3' limit of region of interest

Figure 3.5 DMS reactivity of homoduplex versus heteroduplex DNAs. A (i) and **(ii)** Quantitative histograms (See Section 2.1 f)) of the relative DMS reactivity of the heteroduplex (He) with respect to the homoduplex (Ho) DNA, from two independent experiments. Positive log values indicate enhanced reactivity in the heteroduplex, negative values indicate decreased reactivity. **B** A control quantitation comparing the signals from two independent gels (1 and 2) of the same reaction (See Section 2.1 f)). **C** Numbering system for the bases of Strand D of the cruciform DNA, indicating also the restriction enzymes used to cleave the DNA.

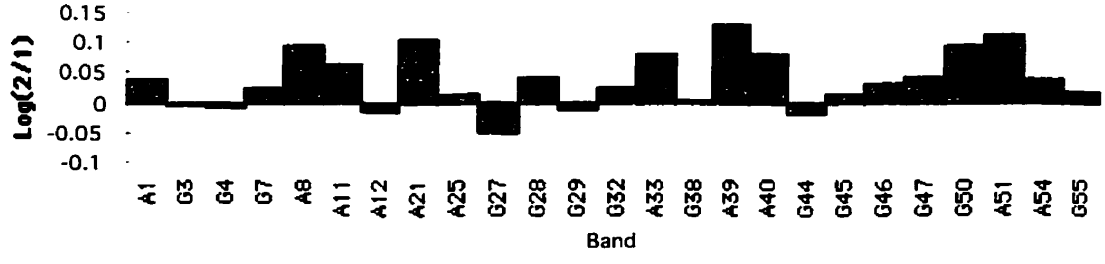
A (i)


Log(Hc/Ho)
0.3
0.2
0.1
0
-0.1
A1 G3 G4 G7 A8 A11 A12 A21 A25 G27 G28 G29 G32 A33 G38 A39 A40 G44 G45 G46 G47 G50 A51 A54 G55 A58
Band

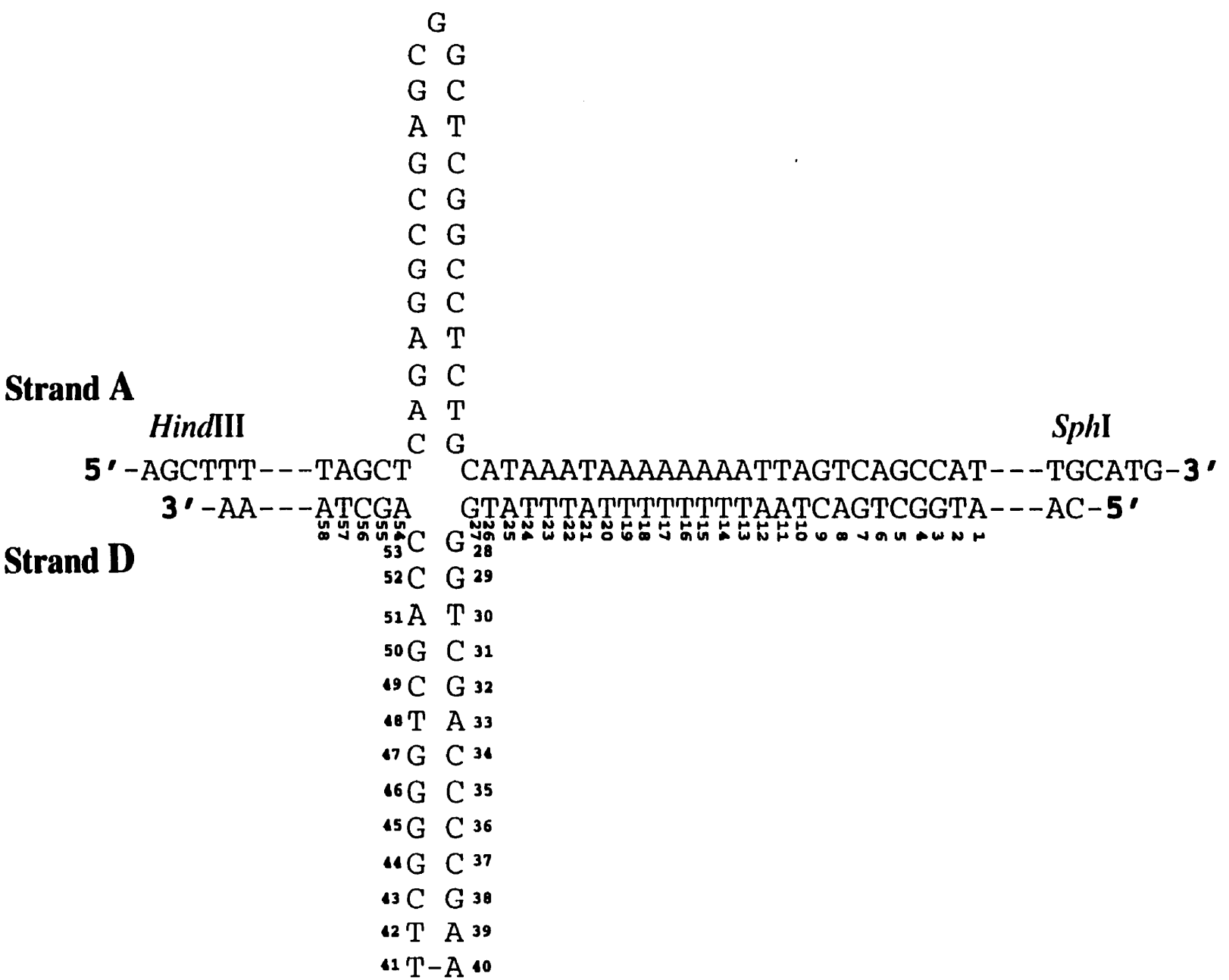
(ii)


Log(Hc/Ho)
0
-0.1
-0.2
-0.3
-0.4
-0.5
A1 G3 G4 G7 A8 A11 A12 A21 A25 G27 G28 G29 G32 A33 G38 A39 A40 G44 G45 G46 G47 G50 A51 A54 G55 A58
Band

(B)


Log(2/1)
0.15
0.1
0.05
0
-0.05
-0.1
A1 G3 G4 G7 A8 A11 A12 A21 A25 G27 G28 G29 G32 A33 G38 A39 A40 G44 G45 G46 G47 G50 A51 A54 G55
Band

C


Strand A
HindIII
SphI
5′-AGCTTT---TAGCT CATAAATAAAAAAAAATTAGTCAGCCAT---TGCATG-3′
3′-AA---ATCGA GTATTTATTTTTTTTTAATCAGTCGGTA---AC-5′
Strand D

(See Section 2.1 f)). This also permitted normalization for the total radioactivity loaded into each lane, a factor which can make faint differences difficult to distinguish by eye. In order to assess the significance of the quantitative differences presented in the histograms, an identical quantitation was carried out on the bands resulting from the DMS treatment of the heteroduplex, run on two separate sequencing gels, and therefore yielding entirely independent autoradiograms (Figure 3.5 B). Figure 3.5 C presents a numbering system for the bases of strand D which will be used to facilitate the description of the differences in DMS reactivity observed.

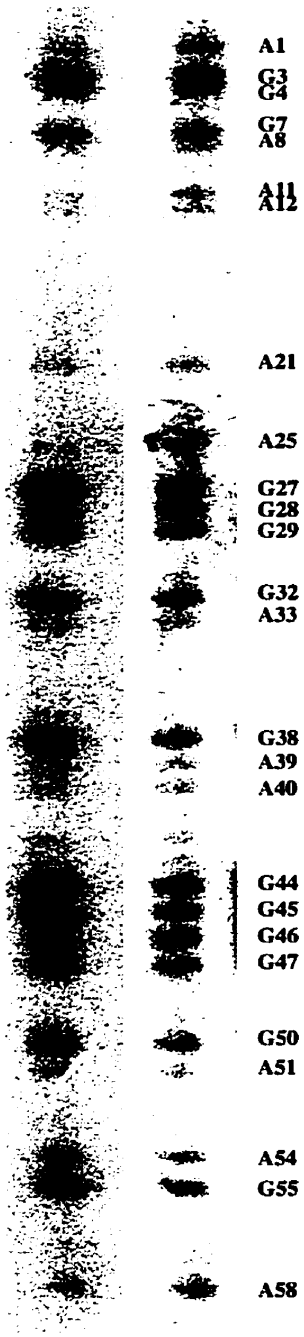
To avoid the problem of non-specific purine cleavage, the piperidine step itself was carried out in TE, pH 7.6, rather than water. pH measurements of volumes of TE and water, equivalent to those used for the piperidine reaction and subsequent washes of the DNA, showed that the pH did not drop below 7.4. Under these conditions, some differences were evident between the DMS reactivity of the homoduplex and heteroduplex DNAs (Figure 3.6). An increase in the reactivity of the majority of adenines proved a reproducible trend, although the extent of enhancement varied between independent experiments. A slightly decreased reactivity of all the guanines located 3' of the tip of the cruciform appears to be less significant. Figure 3.7 summarizes the differences in DMS reactivity at the different points on the cruciform structure. The comparison of homoduplex and heteroduplex DMS reactivity was carried out with heteroduplex DNA which was prepared for the DMS reaction either by lyophilization (Figure 3.6 B (i)) or by precipitation with sodium acetate and ethanol, since it has been suggested that lyophilization may affect the heteroduplex structure, causing reversion to homoduplex [123]. However, ethidium bromide staining and autoradiography of the heteroduplex run on a polyacrylamide gel following lyophilization showed no

Figure 3.6 DMS reactivity of homoduplex versus heteroduplex DNA, with piperidine cleavage in TE. **A** DMS sequencing gels of the homoduplex and heteroduplex indicating the adenine and guanine bases in the cruciform region. The piperidine cleavage was carried out in the presence of TE buffer, pH 7.6, rather than water (See Section 3.5). **(ii)** is composed of scans of two different films, due to the differences in amounts of radioactivity loaded. **B** Quantitative histograms of the autoradiograms presented in A. **(i)** and **(ii)** refer to two independent experiments.

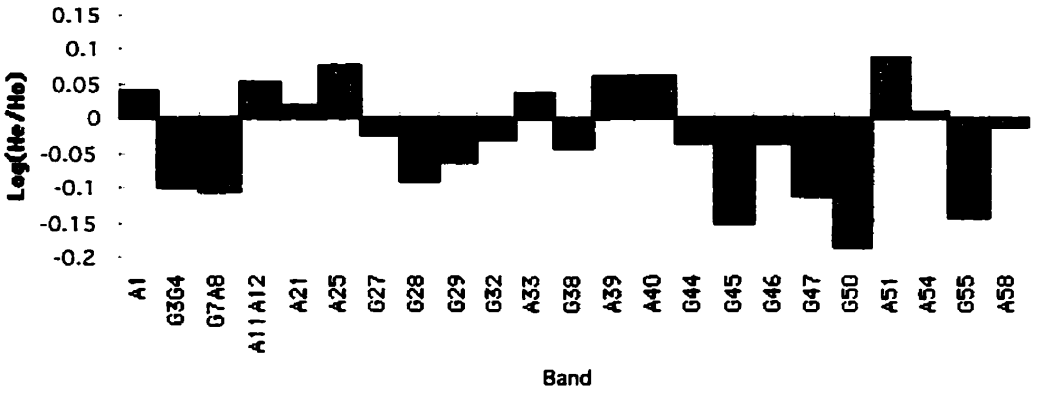
A (i)


Ho
He
A1
G4
G7
A8
A11
A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
G44
G45
G46
G47
G50
A51
A54
G55
A58

A (ii)


Ho
He
A1
G3
G4
G7
A8
A11
A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
G44
G45
G46
G47
G50
A51
A54
G55
A58

B (i)


0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
Log(He/Ho)
A1
G3G4
G7A8
A11A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
G44
G45
G46
G47
G50
A51
A54
G55
A58
Band

(ii)

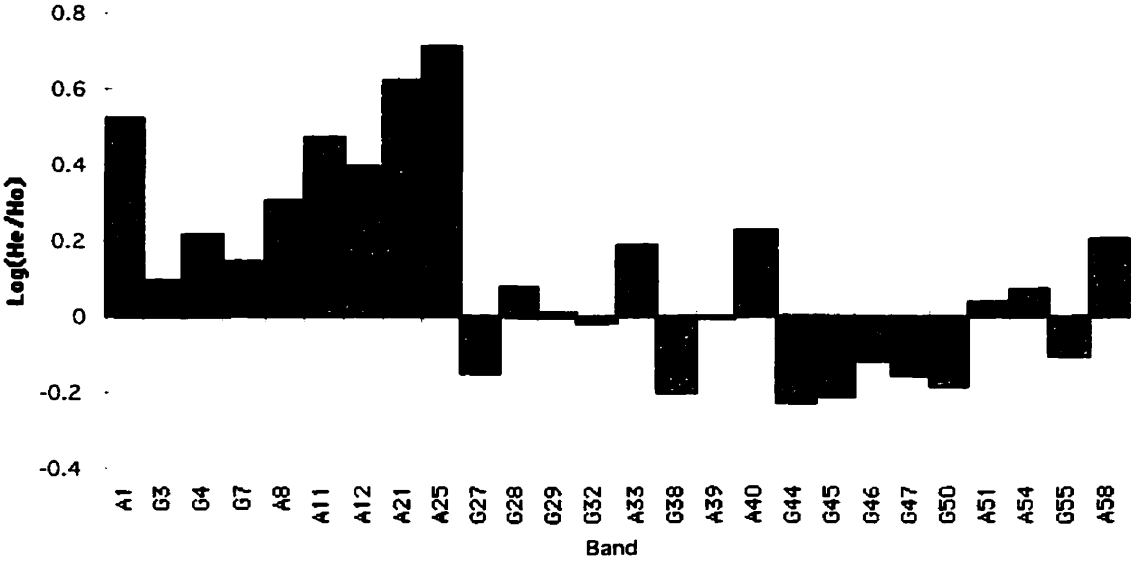
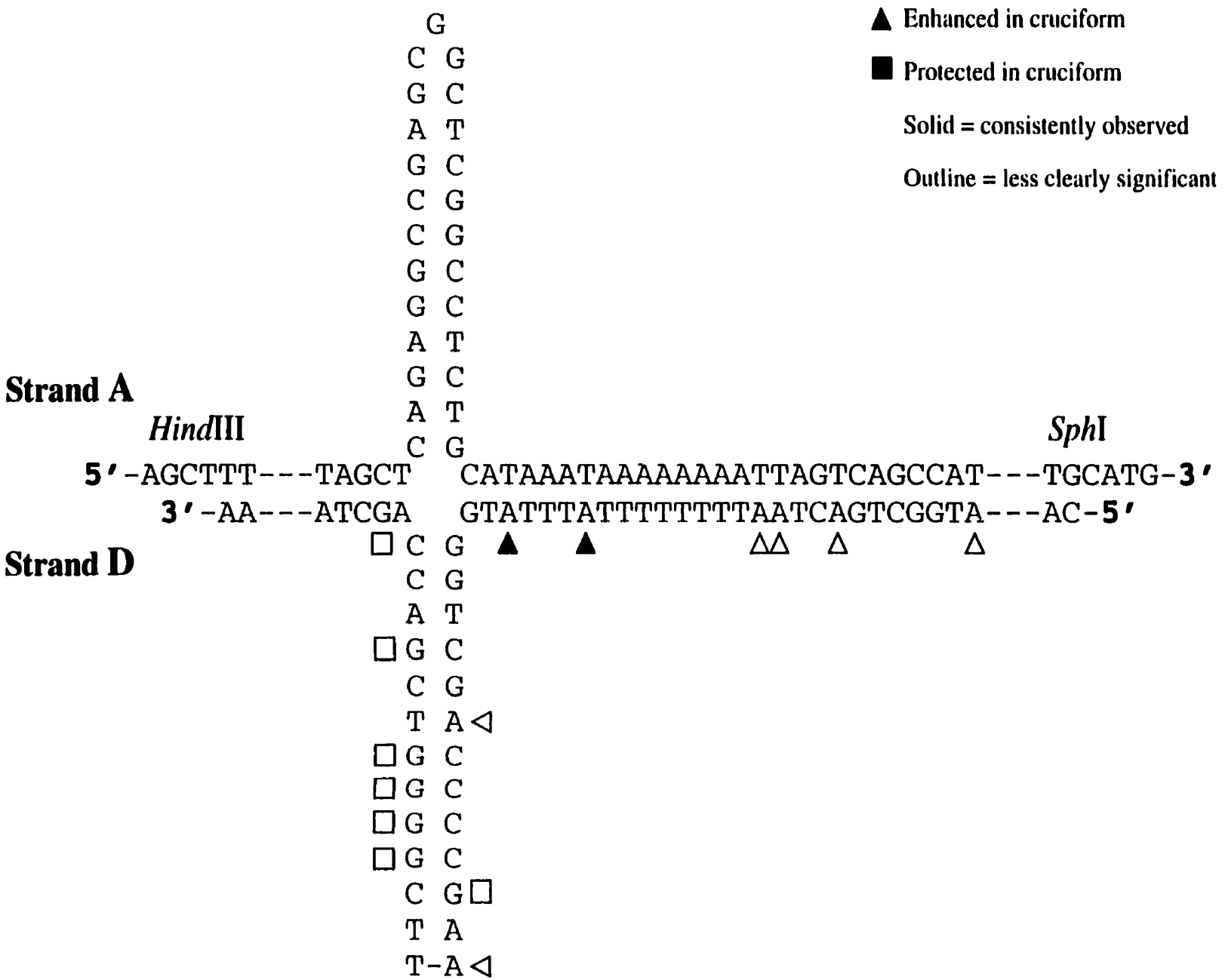

0.8
0.6
0.4
0.2
0
-0.2
-0.4
Log(He/Ho)
A1
G3
G4
G7
A8
A11
A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
G44
G45
G46
G47
G50
A51
A54
G55
A58
Band

Figure 3.7 Summary of differences in DMS reactivity of the cruciform and linear DNA. Location of sites of enhanced and reduced DMS reactivity on the cruciform DNA, relative to the corresponding linear DNA, based upon the quantitative histograms in Figure 3.6.


▲ Enhanced in cruciform
■ Protected in cruciform
Solid = consistently observed
Outline = less clearly significant
Strand A
HindIII
SphI
5′-AGCTTT---TAGCT CATAAATAAAAAAAATTAGTCAGCCAT---TGCATG-3′
3′-AA---ATCGA GTATTTATTTTTTTTAATCAGTCGGTA---AC-5′
Strand D

observable effects on the structure (data not shown). All subsequent footprinting experiments were conducted relative to the heteroduplex and not the homoduplex.

3.6 Footprinting without a Preparative PAGE Step

Since the titration experiments allowed the establishment of conditions under which 100 % of the cruciform is bound by the protein, and no free cruciform is available to obscure the signal, we decided to try footprinting without a preparative PAGE step. A preliminary experiment with a mock binding reaction, using Buffer B (See Section 2.1 b)) instead of the CBP-enriched fraction, was performed to establish the reagent concentrations that would give the desired extent of reaction under the binding conditions, as opposed to the sequencing conditions. A sequencing ladder of even intensity bands was obtained with final concentrations of DMS and β-ME (as the stop reagent) of 0.1 % and 1.9 M, respectively (data not shown). These conditions were used for the majority of the footprinting experiments (exceptions are mentioned below).

Four independent such experiments were carried out. In each case, an analytical sample was removed from the binding reaction prior to treatment with DMS to determine the extent of binding, and analyzed by PAGE (Figure 3.8). Despite the fact that 100 % of the cruciform appeared to be bound, no clear, strong footprint was distinguishable. Comparison of the histograms generated from the quantitation of the footprinting autoradiogram bands (in some cases two independent gels were run of the same reaction products) also showed little in the way of a clear, consistent footprint (Figure 3.9). T42 and C43, between the G tetrad and the cruciform tip, gave an unusually strong signal upon protein binding in two of the four experiments (For example Figure 3.9 A and B), but were

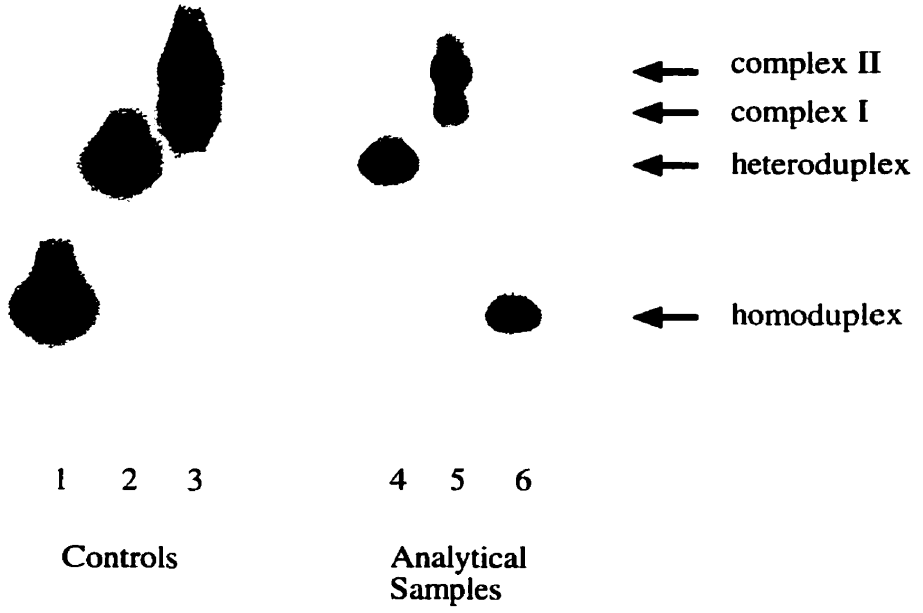
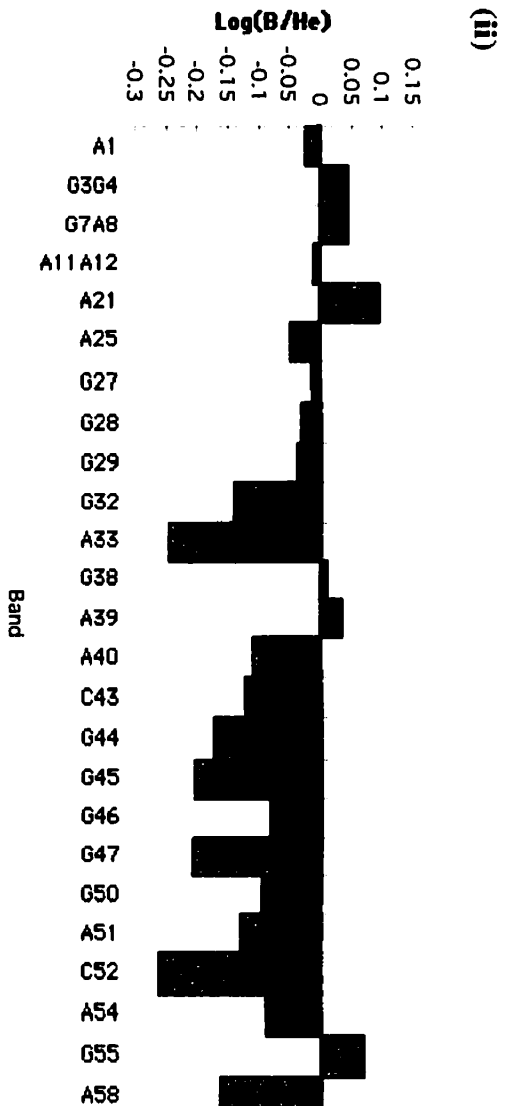
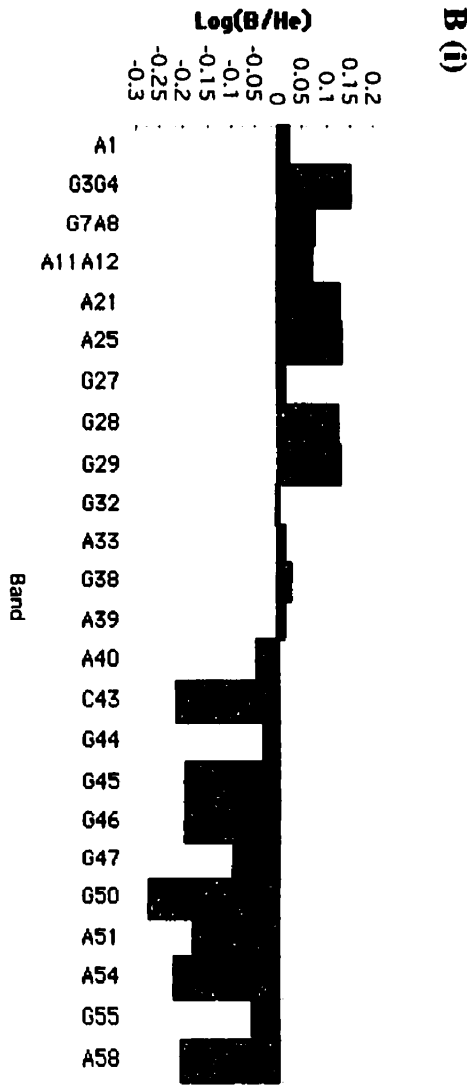
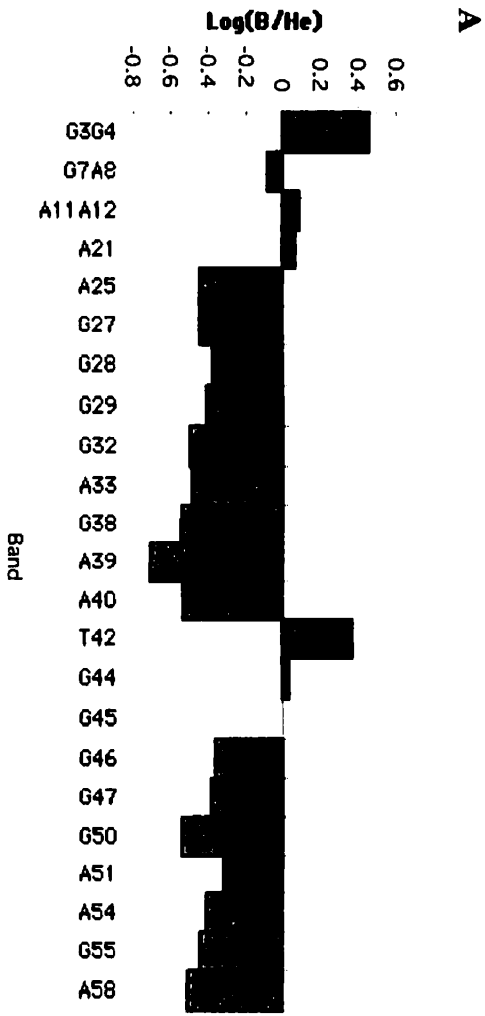


Figure 3.8 Analysis of the extent of binding of the cruciform in footprinting experiments without a preparative PAGE step. Comparison of the analytical samples (lanes 4-6) taken from a footprinting experiment, prior to DMS treatment, to a control EMSA (lanes 1-3). Since all the cruciform is shifted to the characteristic cruciform-CBP complexes, these reactions were not loaded onto a preparative PAGE following DMS treatment.

Figure 3.9 Footprinting without a preparative PAGE step. Representative quantitative histograms of the relative DMS reactivity of the cruciform DNA in the presence of saturating amounts of CBP (B), with respect to that of free cruciform (He). **A** and **B** represent independent experiments, **(i)** and **(ii)** denote independent footprinting gels of the same reactions. Note that the y-axis scales of all the histograms are not identical.


(ii)
Log(B/He)
-0.3
-0.25
-0.2
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
Band
A1
G3G4
G7A8
A11A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
C43
G44
G45
G46
G47
G50
A51
C52
A54
G55
A58


B (i)
Log(B/He)
-0.3
-0.25
-0.2
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
0.2
Band
A1
G3G4
G7A8
A11A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
C43
G44
G45
G46
G47
G50
A51
A54
G55
A58


A
Log(B/He)
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
Band
G3G4
G7A8
A11A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
T42
G44
G45
G46
G47
G50
A51
A54
G55
A58

unaffected, or slightly protected (Figure 3.9 B) in the other two. The GAA at the tip of the cruciform (bases 38-40) was quite consistently of lower intensity than for the free heteroduplex DNA, as was the AG (bases 54 and 55, particularly 55) at the 3' elbow, and A58. Less consistently observed is an enhancement of G3 and G4, possible enhancement of A21, and protection of A33. Comparison of the histograms with the control histogram (Figure 3.5 B) emphasizes the uncertainty of the significance of many of the observed variations. In some cases the independent gels of the same reaction products gave contradictory results. Figure 3.10 summarizes the observed differences in DMS reactivity on the cruciform structure.

3.7 Footprinting with a Preparative PAGE Step

The possibility that one of the two shifted bands would give a clearer footprint than the combination of the two together prompted the addition of a preparative PAGE step to the footprinting protocol. Using a 4 % polyacrylamide gel to separate the various species following binding of DNA and protein, DMS treatment and quenching with β-ME, did not improve the clarity of any possible footprint present (data not shown). Proposing that the 4 % polyacrylamide may not adequately separate the species, an 8 % preparative polyacrylamide gel was substituted.

A number of experiments under these conditions demonstrated that the EMSA pattern following the methylation reaction was not identical to that without DNA methylation (Figure 3.11). Though bands with similar electrophoretic migrations to those of the non-methylated controls were present, several other species were also observed, suggesting that the DMS treatment resulted in an alteration of the protein-DNA complex(es), reducing our confidence that we could

Figure 3.10 Summary of differences in DMS reactivity of the cruciform DNA in the presence and absence of CBP, from experiments without a preparative PAGE step. Location of sites of enhanced reactivity and protected from DMS, on the cruciform DNA in the presence of CBP, relative to the absence of CBP, based upon the quantitative histograms in Figure 3.9.

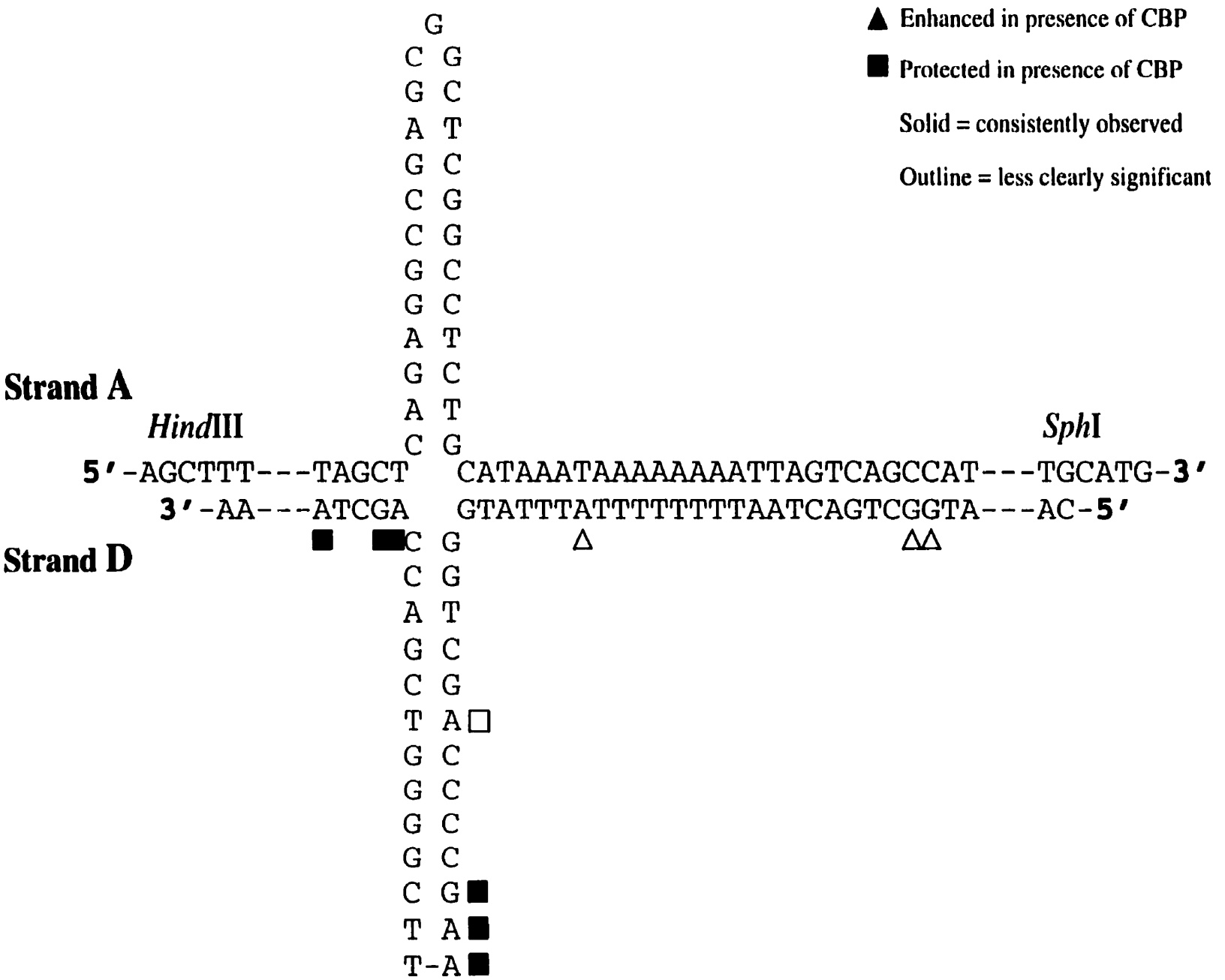
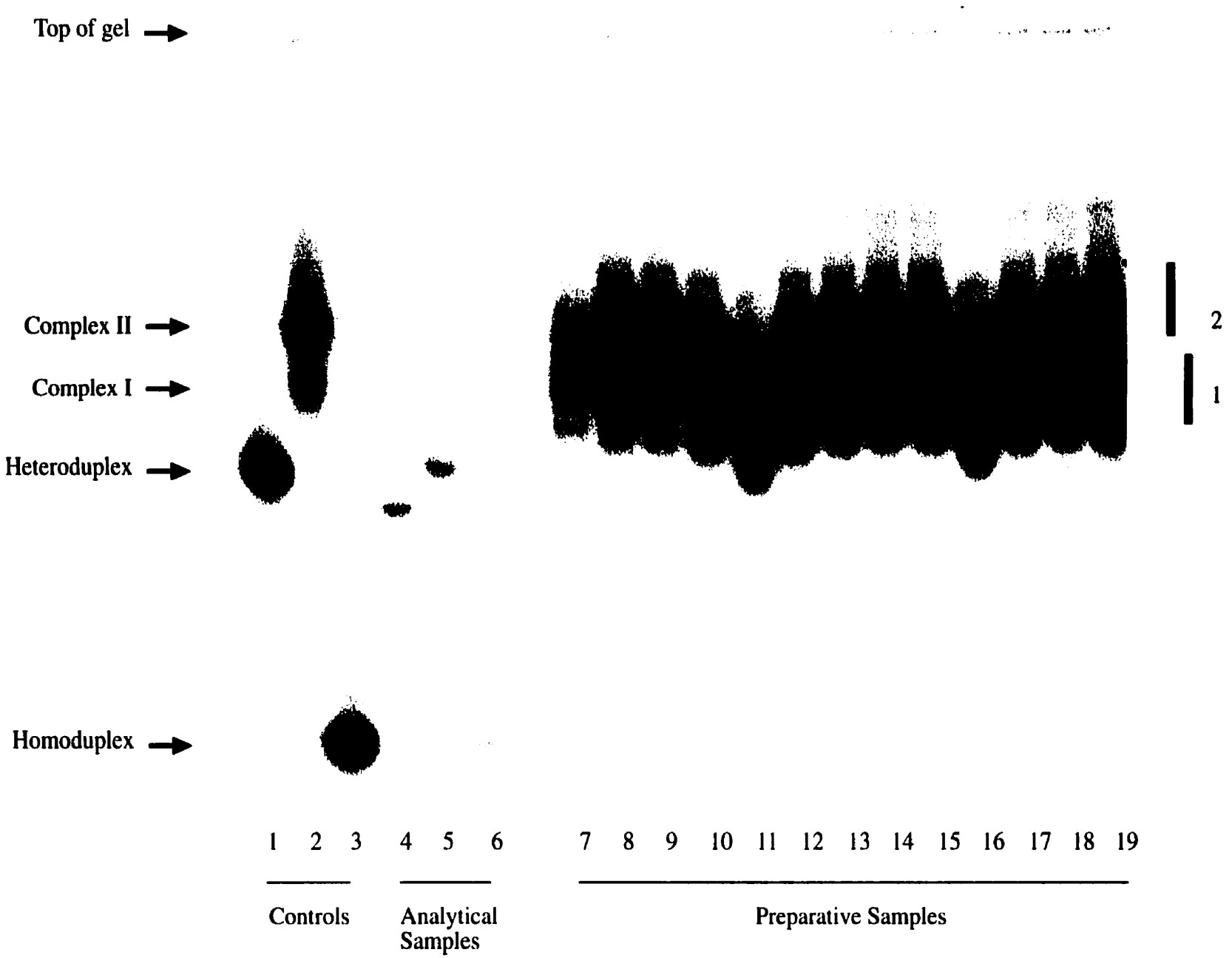

▲ Enhanced in presence of CBP
■ Protected in presence of CBP
Solid = consistently observed
Outline = less clearly significant
Strand A
HindIII
SphI
5′-AGCTTT---TAGCT CATAAATAAAAAAAATTAGTCAGCCAT---TGCATG-3′
3′-AA---ATCGA GTATTTATTTTTTTTAATCAGTCGGTA---AC-5′
Strand D

Figure 3.11 8% preparative polyacrylamide gel for footprinting. Wet exposure autoradiogram of a preparative 8% polyacrylamide gel of a footprinting experiment. Lanes 1-3 are controls of free cruciform, cruciform shifted 100 % by CBP, and linear homoduplex DNA, respectively. Lanes 4-6 are analytical samples taken from the footprinting reactions, following addition and quenching of DMS: lane 4 is cruciform DNA plus Buffer B (See Section 2.1 b)), lane 5 is cruciform DNA plus a saturating amount of CBP-enriched fraction, and lane 6 is linear homoduplex with an equal amount of CBP-enriched fraction as lane 5. Lanes 7-19 are the same reaction as in lane 5, divided over several preparative lanes due to volume, from which gel slices were excised for the completion of the footprinting procedure. Bars to the right of the gel indicate the gel slices excised. The apparent deviation in position of the band in lane 4 from that in lane 1 results from a slight distortion of the gel, and was not observed in other experiments. Incomplete mixing of the reaction mixture prior to loading onto the gel may account for the aberrant migration of the species in lanes 11 and 16.


Top of gel
Complex II
Complex I
Heteroduplex
Homoduplex
2
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Controls
Analytical Samples
Preparative Samples

isolate the same complexes from the preparative lanes as are found in the controls. Furthermore, the EMSA pattern of the analytical samples taken following methylation (Figure 3.11, lanes 4-6) is not identical to the pattern of the rest of the reaction mixture run on the preparative gel (lanes 7-19).

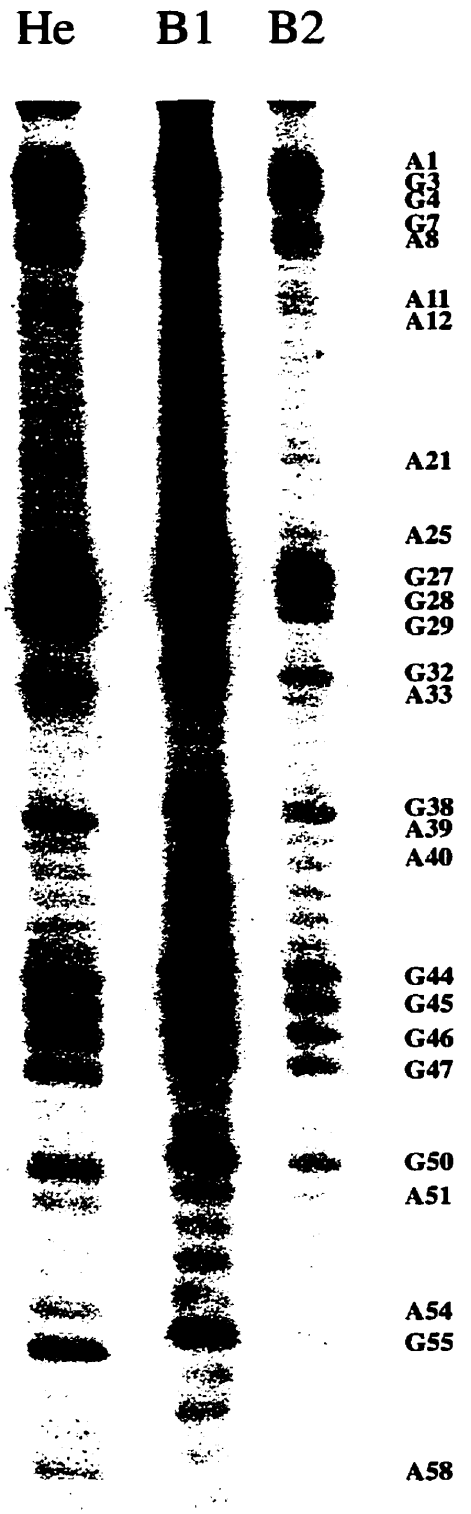
Two independent experiments were analyzed for a footprint by excising the two most significant bands from the preparative polyacrylamide gel (Figure 3.11, bars 1 and 2). No clear footprint is visible from the footprinting autoradiograms (Figure 3.12 A), therefore the majority of information is taken from the quantitative histograms (Figure 3.12 B and C). Both experiments show a great deal of similarity between the patterns of the two complexes analyzed. The most striking alterations in reactivity are the enhancement of the reactivity of the GAA (bases 38-40) at the tip of cruciform, and enhancement of the AG (bases 54 and 55) in the 3' elbow. Less significant enhancements include A's 11, 12, 21 and 25, and the G tetrad (bases 44-47) on the 3' side of the cruciform stem. Protection of G27 and G28, and enhancement of G29 in the 5' elbow are also observable, but the significance is not clear. Figure 3.13 summarizes the observed differences in DMS reactivity on the cruciform structure, and compares them to those observed from experiments without a preparative PAGE step (See Section 3.6). Some correspondence is observed. Common points of modified reactivity upon CBP binding are the AG (bases 54 and 55) in the 3' elbow, the GAA (bases 38-40) at the tip of the cruciform, and A21 (Figure 3.13).

3.8 Titrations to Minimize DMS and β-ME

The β-ME concentration used to quench the DMS reaction is high enough to cause large scale reduction of the CBP and other proteins in the CBP-enriched

Figure 3.12 Results of footprinting experiments including an 8 % preparative polyacrylamide gel step. **A** Footprinting gel comparing the DMS reactivity of the cruciform (He) to the two major complexes excised from the preparative gel (B1 and B2, see Figure 3.11). Note that unequal amounts of radioactivity were loaded onto the three lanes. **B** Quantitative histograms of the autoradiogram presented in A. **C** Quantitative histograms of the autoradiogram of an independent experiment. **(i)** and **(ii)** represent B1 and B2, respectively. Note that the y-axis scale is not identical for all the histograms.

A


He
B1
B2
A1
G3
G4
G7
A8
A11
A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
G44
G45
G46
G47
G50
A51
A54
G55
A58

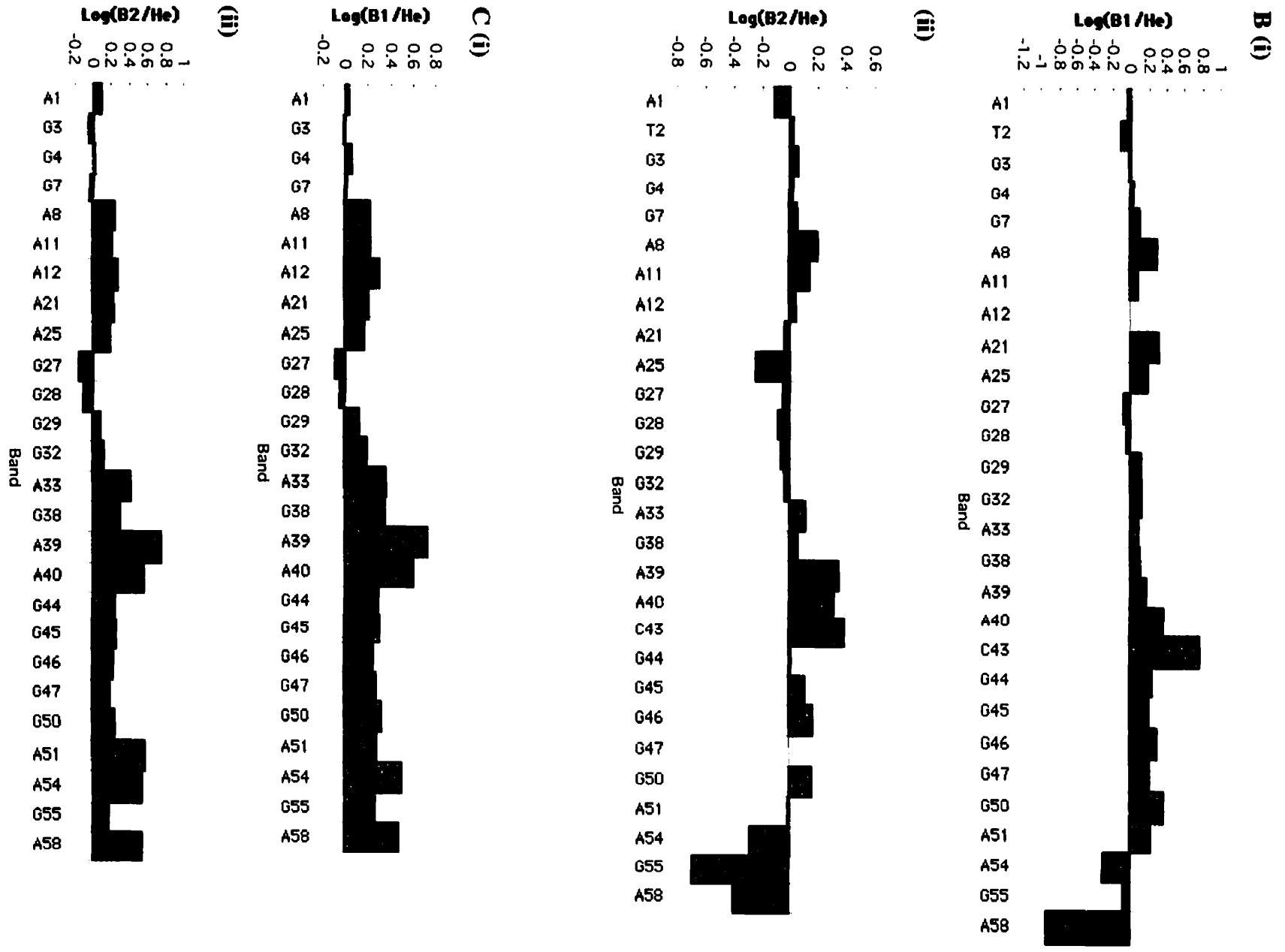
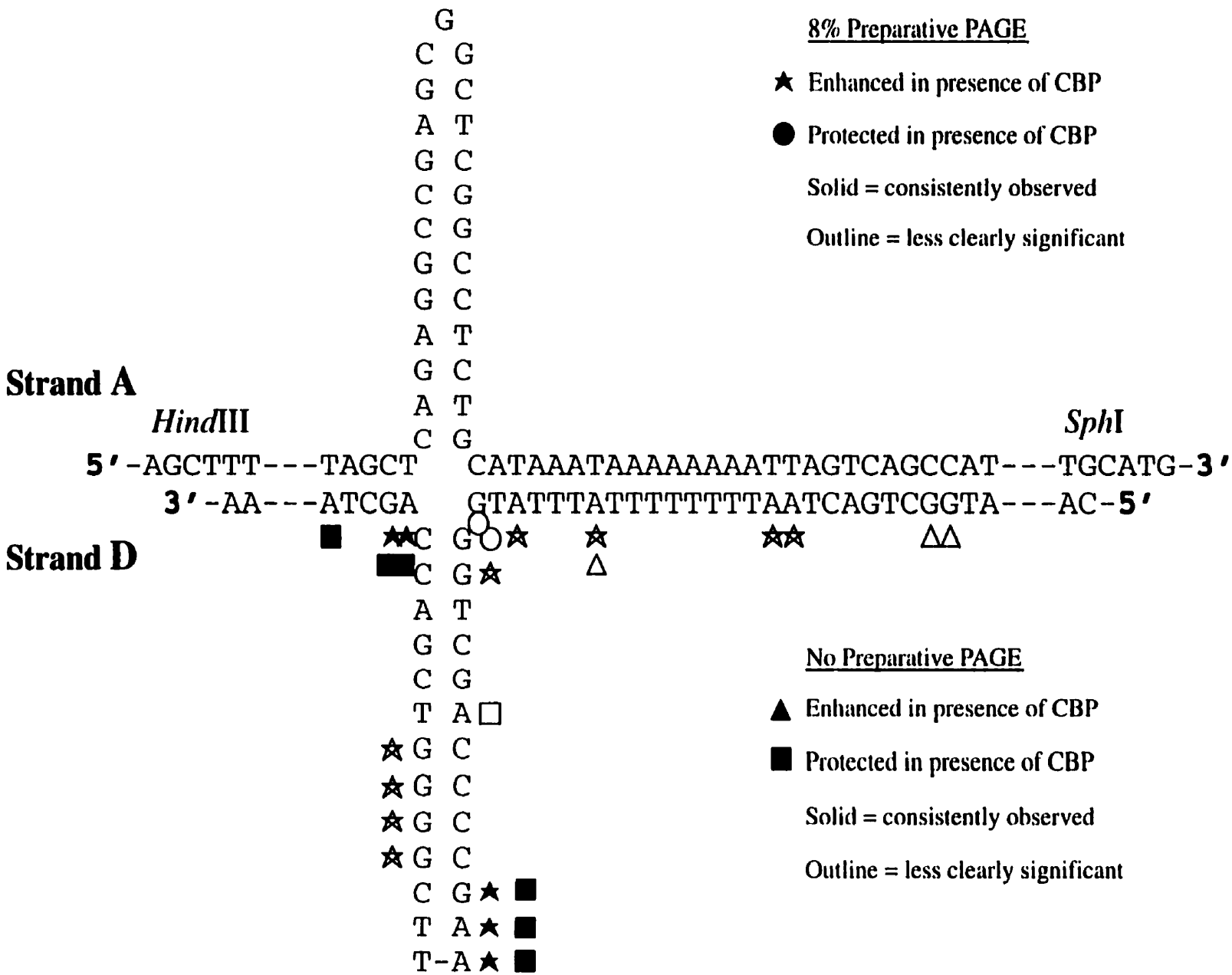

B (i)
Log(B1/He)
-1.2 -1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
A1 T2 G3 G4 G7 A8 A11 A12 A21 A25 G27 G28 G29 G32 A33 G38 A39 A40 C43 G44 G45 G46 G47 G50 A51 A54 G55 A58
Band
(ii)
Log(B2/He)
-0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6
A1 T2 G3 G4 G7 A8 A11 A12 A21 A25 G27 G28 G29 G32 A33 G38 A39 A40 C43 G44 G45 G46 G47 G50 A51 A54 G55 A58
Band
C (i)
Log(B1/He)
-0.2 0 0.2 0.4 0.6 0.8
A1 G3 G4 G7 A8 A11 A12 A21 A25 G27 G28 G29 G32 A33 G38 A39 A40 G44 G45 G46 G47 G50 A51 A54 G55 A58
Band
(ii)
Log(B2/He)
-0.2 0 0.2 0.4 0.6 0.8 1
A1 G3 G4 G7 A8 A11 A12 A21 A25 G27 G28 G29 G32 A33 G38 A39 A40 G44 G45 G46 G47 G50 A51 A54 G55 A58
Band

Figure 3.13 Comparison of the DMS footprints obtained with and without a preparative PAGE step. Location of the sites of enhanced reactivity and protection from DMS on the cruciform DNA in the presence of CBP, relative to the absence of CBP, based upon the quantitative histograms presented in Figures 3.9 (no preparative PAGE) and 3.12 (8 % preparative PAGE).


8% Preparative PAGE
Enhanced in presence of CBP
Protected in presence of CBP
Solid = consistently observed
Outline = less clearly significant
Strand A
HindIII
SphI
5'-AGCTTT---TAGCT CATAAATAAAAAAAATTAGTCAGCCAT---TGCATG-3'
3'-AA---ATCGA GTATTTATTTTTTTTAATCAGTCGGTA---AC-5'
Strand D
No Preparative PAGE
Enhanced in presence of CBP
Protected in presence of CBP
Solid = consistently observed
Outline = less clearly significant

fraction [124]. The environment in the polyacrylamide is very oxidative due to the presence of considerable amounts of reactive by-products of acrylamide polymerization, including oxidative radicals [125]. Transferring the binding mixture from a highly reducing to a highly oxidative environment likely causes significant disruption of the tertiary and quaternary structures of the proteins. This may be responsible for the unexpected appearance of the preparative EMSAs. Two approaches were taken to address this problem: (a) reducing the amount of β-ME and introducing the free radical scavenger thioglycolate into the preparative gel system, and (b) reducing the DMS to the point where no quenching reagent is required.

Maintaining a final DMS concentration of 0.1 %, the methylation of homoduplex DNA was quenched with decreasing amounts of β-ME (Figure 3.14). As little as 100 mM β-ME (Figure 3.14, lane 5) proved sufficient to stop the reaction, yielding sufficient unreacted DNA and a ladder of bands of even intensities indicating the desired single-hit kinetics (Figure 3.14).

Since DMS methylates water in an aqueous solution, producing methanol (H. Zorbas, Personal Communication) it can, if the reaction is allowed to proceed for long enough, exhaust its methylating capabilities and no stop reagent needs to be added. In the context of footprinting, methylation exhaustion must not occur at the expense of single-hit kinetics, therefore a titration of the DMS was carried out. Progressively lower final concentrations of the DMS were used to methylate the DNA in the binding reaction conditions, reactions were left at room temperature for 15 min to mimic the time required to load them onto the preparative polyacrylamide gel, and then the sequencing reactions were completed (See Section 2.1 d)). Under

Figure 3.14 Titration to minimize β-ME. Final β-ME concentrations of 1.9, 1, 0.5, 0.3 and 0.1 M, lanes 1-5, respectively, were used to quench the DMS reaction with homoduplex DNA. In each case a strong unreacted band, and a ladder of bands of even intensities throughout the region of interest, were obtained. The image is fainter in the region above the asterisks because of an overlap of two films during the exposure, due to the length of the region visualized. Note that less radioactivity was loaded into lane 5 than the others.

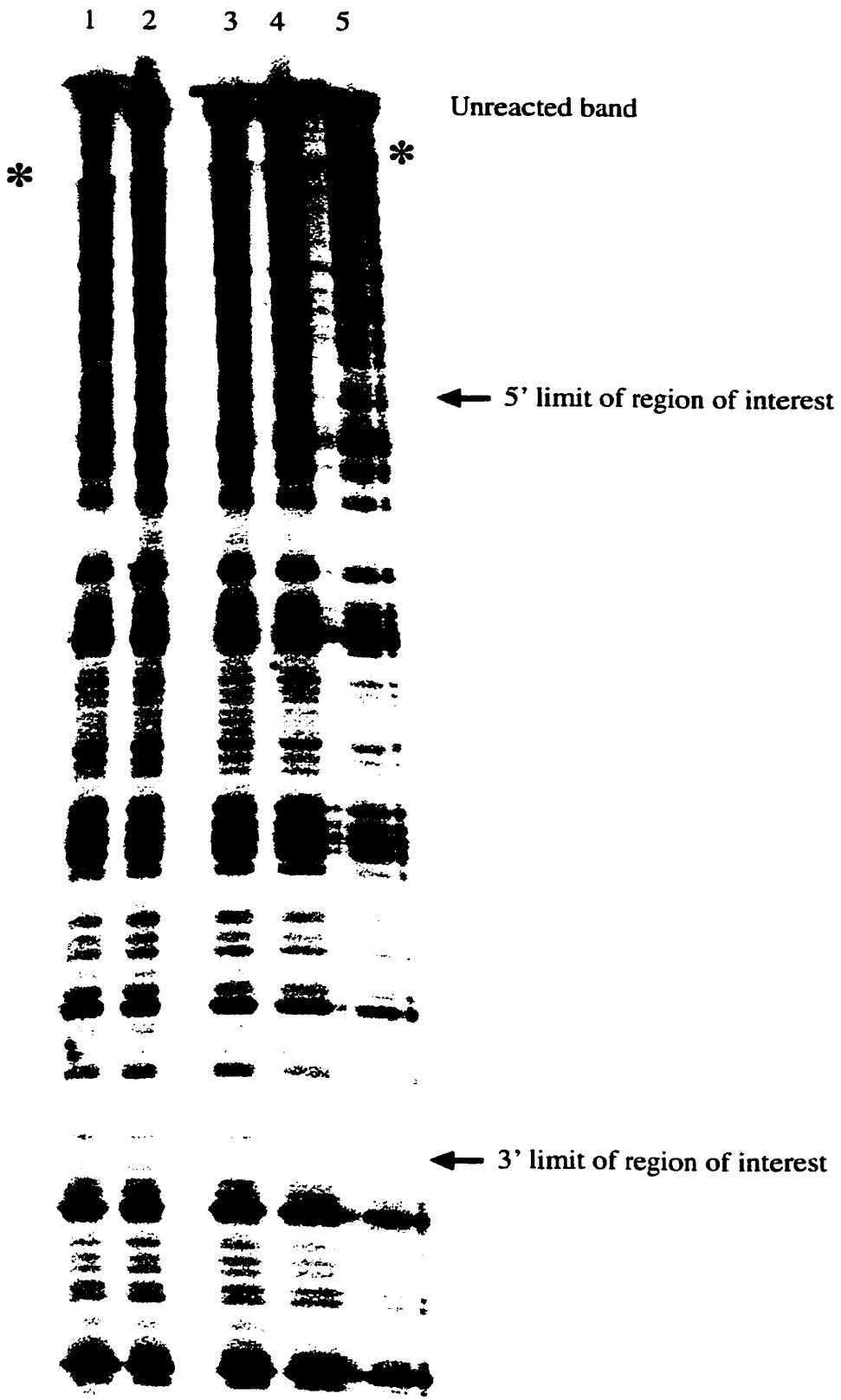
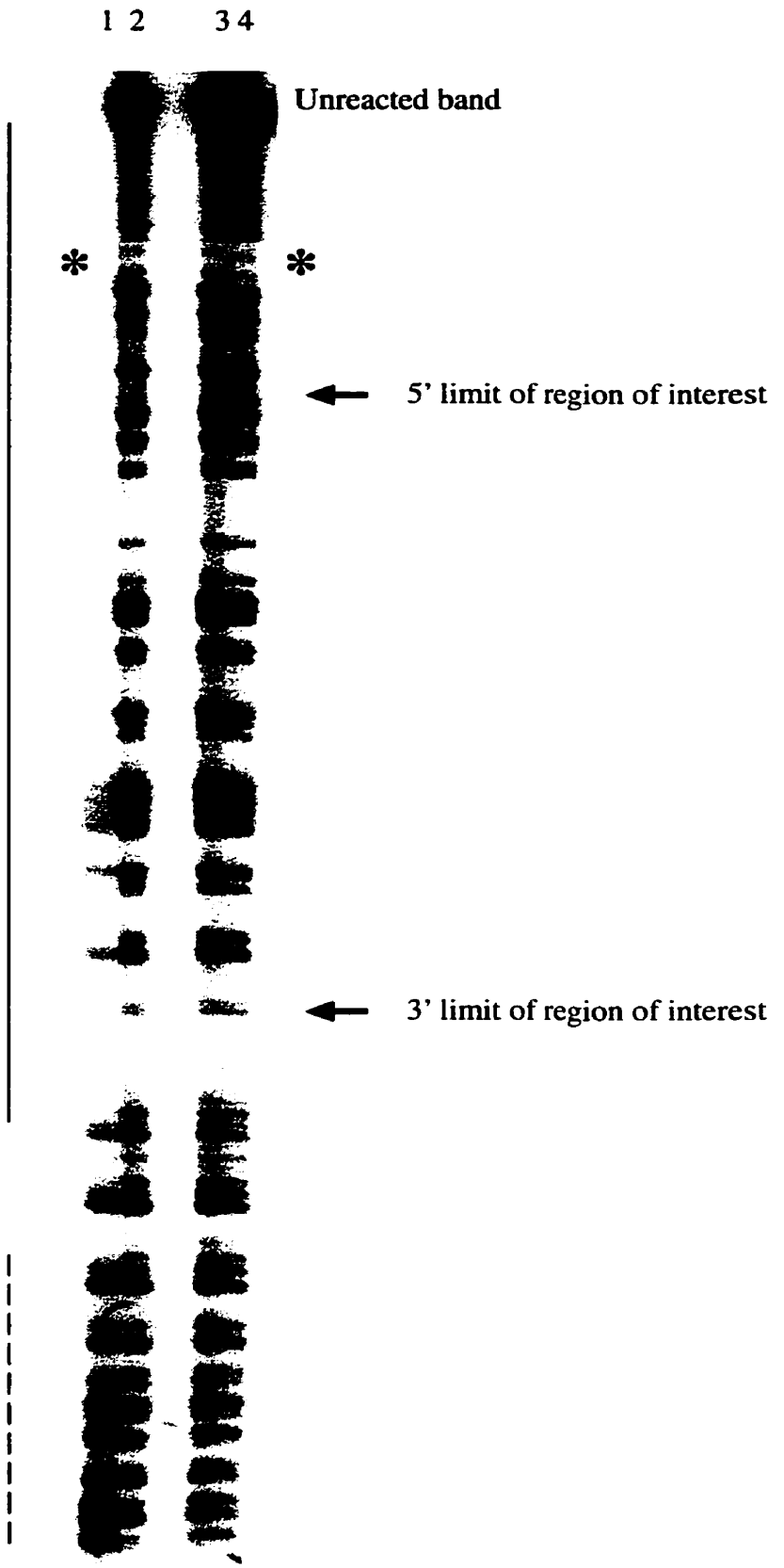

1
2
3
4
5
Unreacted band
5' limit of region of interest
3' limit of region of interest

Figure 3.15 Titration to minimize DMS. Final DMS concentrations of 0.5 (lane 1), 0.1 (lane 2), 0.05 (lane 3) and 0.01% (lane 4) were used to sequence the cruciform DNA. Note the absence of unreacted band in lane 1. The solid line highlights the bands representing the larger DNA fragments, which are under-represented, and the dashed line highlights the bands representing the smaller DNA fragments, which are over-represented, in the case of over-reaction. A final DMS concentration of 0.05 % (lane 3) gave the best ladder of bands of even intensities in the region of interest. The region of the image marked by the asterisks is lighter due to the overlap of two films during exposure, necessary due to the size of the region visualized.

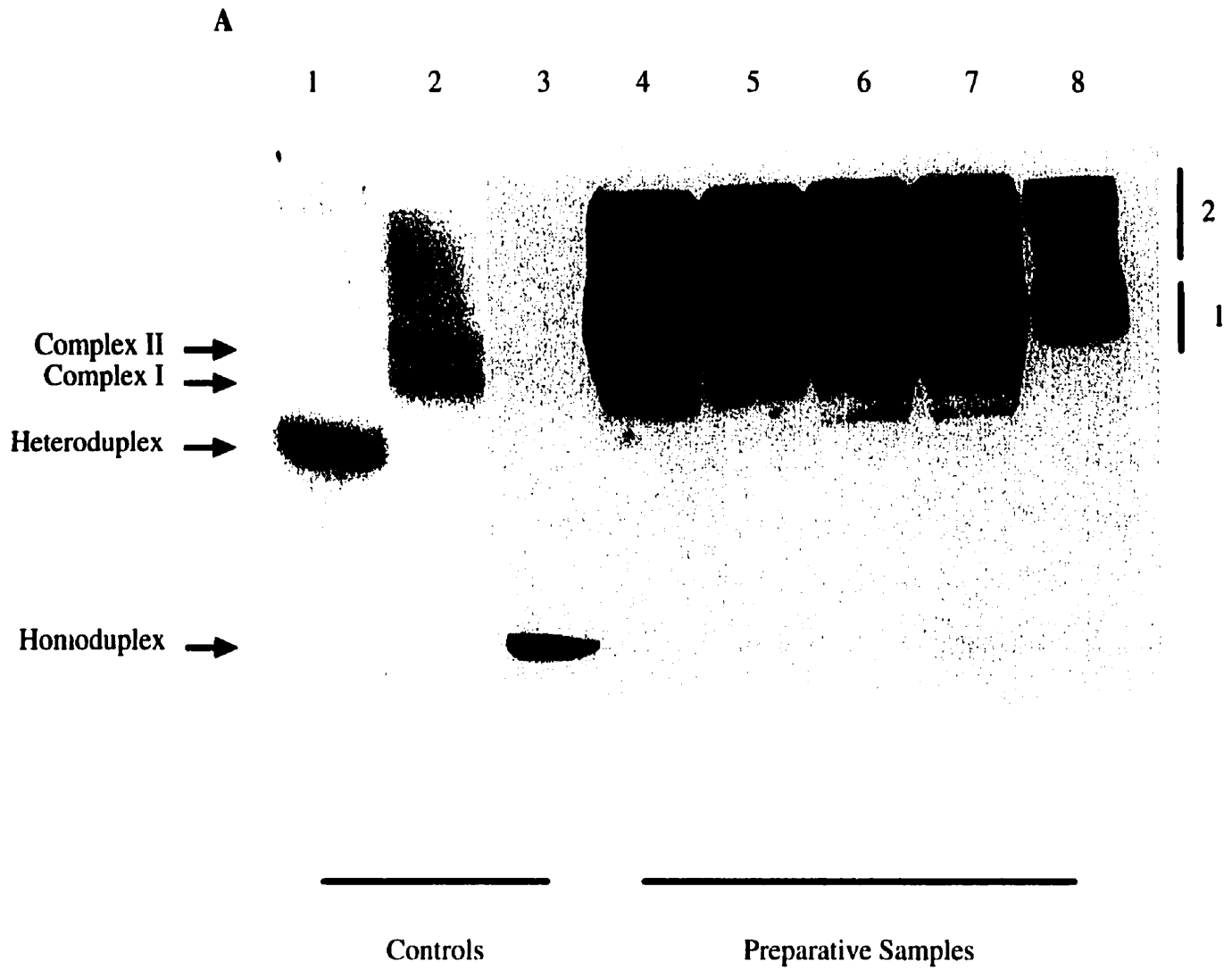

1 2 3 4
Unreacted band
*
*
5' limit of region of interest
3' limit of region of interest

these conditions the best ladder of even intensity bands was obtained with 0.05 % DMS (Figure 3.15, lane 5).

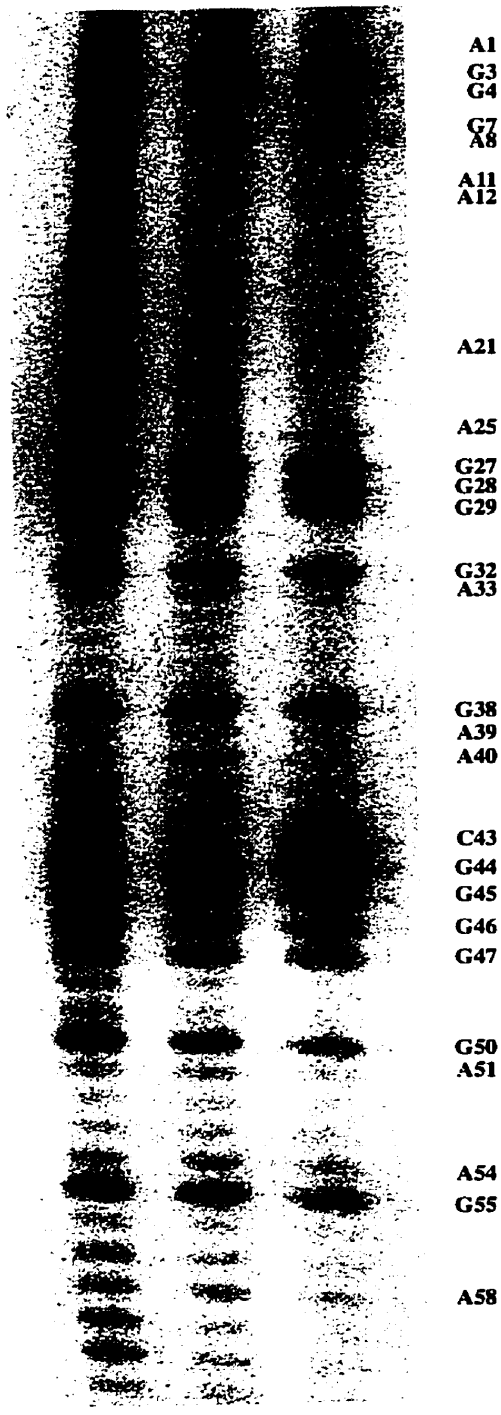
3.9 Footprinting with Minimal β-ME and Thioglycolate as a Free Radical Scavenger

50 mM β-ME was used to quench the methylation reaction and the preparative polyacrylamide gel was pre-run and run in the presence of thioglycolate (See Section 2.1 e)) to scavenge free radicals remaining from the polymerization process. Even under these conditions the preparative binding reaction EMSA is not identical to that of the controls (Figure 3.16 A). The footprinting autoradiogram shows no clear footprint for either complex, the most noticeable difference being an increased reactivity of C43, 5' of the G tetrad on the 3' side of the cruciform stem, in the most retarded complex. The quantitative histograms (Figure 3.16 C) for the two complexes are largely similar, and both show considerable enhancement of A39 at the tip of the cruciform, and some protection of A's 21 and 25. Less clearly significant is a protection of the 3 G's at the 5' elbow (bases 27-29) and G32 and A33, as well as some possible protection of the G tetrad (bases 44-47). Figure 3.17 summarizes the observed differences in DMS reactivity on the cruciform structure, and compares them to those observed from experiments with and without a preparative PAGE step (See Sections 3.6 and 3.7). Again there is some correspondence of sites of modified reactivity, particularly in the elbow and tip regions of the cruciform. However, the modification is not consistent, i.e. a site of enhancement in one experiment is often a site of protection in another.

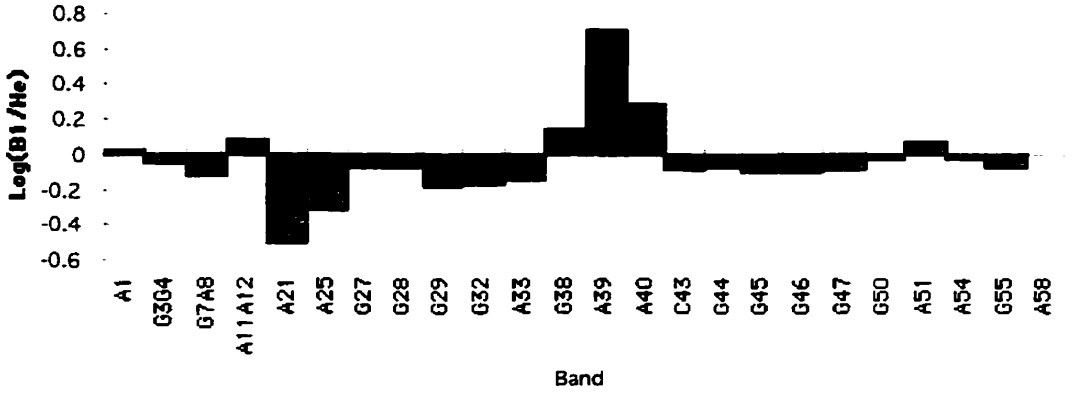
Figure 3.16 Footprinting with minimal β-ME, and an 8% preparative polyacrylamide gel scavenged with thioglycolate. A footprinting experiment was conducted with 50 mM β-ME as the quenching agent, and the products were separated on an 8 % preparative polyacrylamide gel with thioglycolate in the buffer to scavenge the oxidative by-products of polymerization. **A** A wet exposure autoradiogram of the preparative gel showing controls (lanes 1-3), and preparative samples (lanes 4-8). The bars to the right of the gel correspond to the gel slices excised. This figure was generated from two scans of the same film, because of the differences in radioactivity in the preparative versus control lanes. **B** Footprinting autoradiogram of the cruciform DNA (He) and the two excised complexes, see A (B1 and B2). The slight distortion of the bands in lane B2 is due to a very small bubble in the gel. **C** Quantitative histograms of the autoradiogram in B. **(i)** and **(ii)** represent complexes B1 and B2, respectively. Note that the scales of the y-axis are not identical for the two plots.


A
1
2
3
4
5
6
7
8
2
1
Complex II
Complex I
Heteroduplex
Homoduplex
Controls
Preparative Samples

B


He
B1
B2
A1
G3
G4
G7
A8
A11
A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
C43
G44
G45
G46
G47
G50
A51
A54
G55
A58

C (i)


0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
Log(B1/He)
A1
G3G4
G7A8
A11A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
C43
G44
G45
G46
G47
G50
A51
A54
G55
A58
Band

(ii)

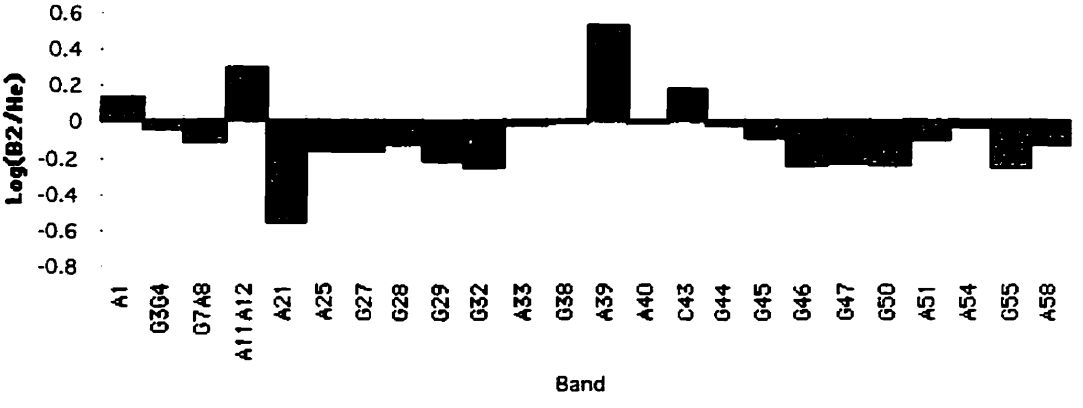
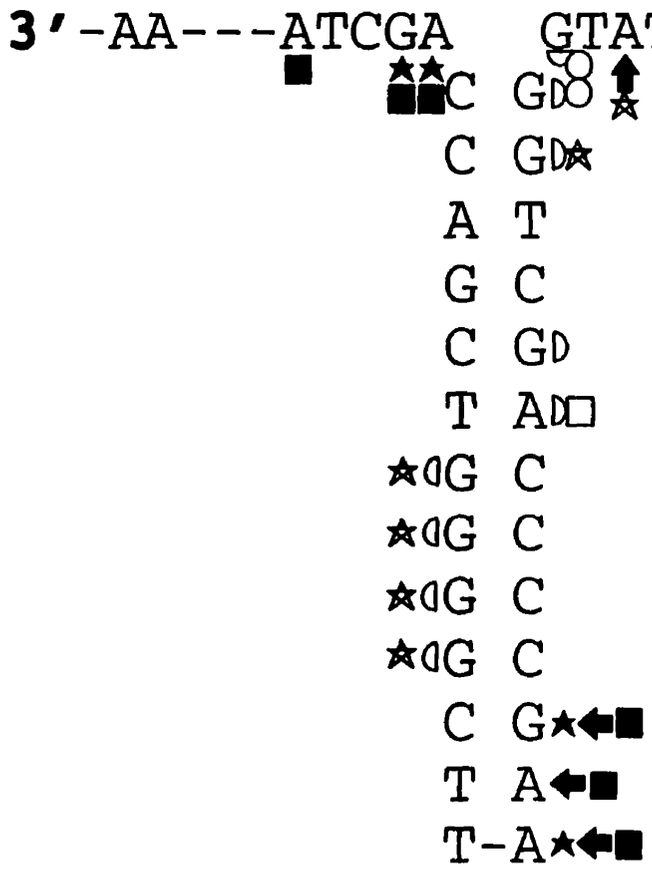

0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
Log(B2/He)
A1
G3G4
G7A8
A11A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
C43
G44
G45
G46
G47
G50
A51
A54
G55
A58
Band

Figure 3.17 Comparison of DMS footprints obtained with and without a preparative PAGE step, and using minimal β-ME and the thioglycolate scavenger. Location of sites of enhanced reactivity, and protection from DMS on the cruciform DNA in the presence of CBP, relative to the absence of CBP, based upon the histograms presented in Figures 3.9 (no preparative PAGE), 3.12 (8% preparative PAGE), and 3.16 (50mM β-ME, thioglycolate). Only strand D is presented.



50mM B-ME, Thioglycolate

- Enhanced in presence of CBP
- Protected in presence of CBP
- Solid = consistently observed
- Outline = less clearly significant

8% Preparative PAGE

- Enhanced in presence of CBP
- Protected in presence of CBP
- Solid = consistently observed
- Outline = less clearly significant

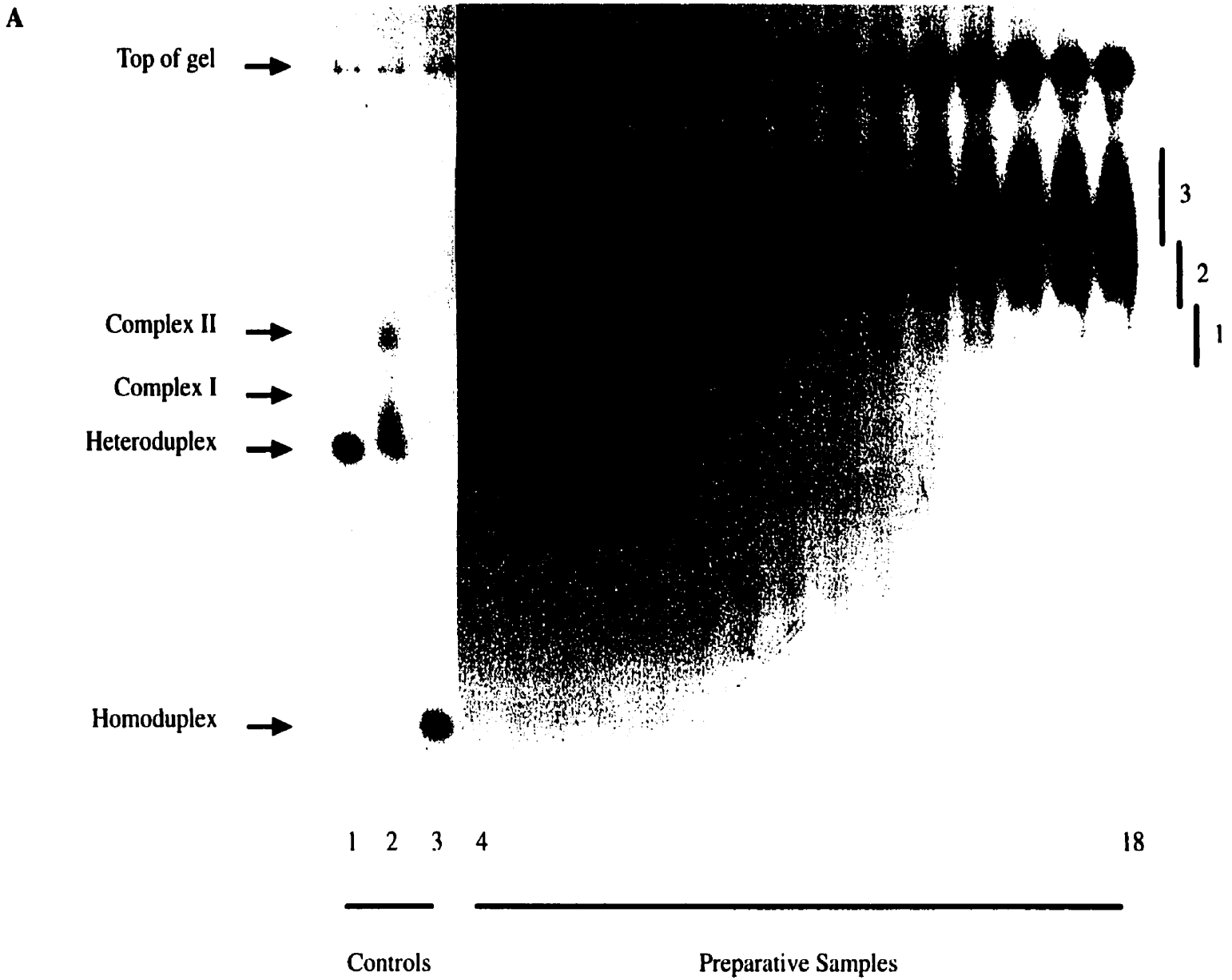
No Preparative PAGE

- Enhanced in presence of CBP
- Protected in presence of CBP
- Solid = consistently observed
- Outline = less clearly significant

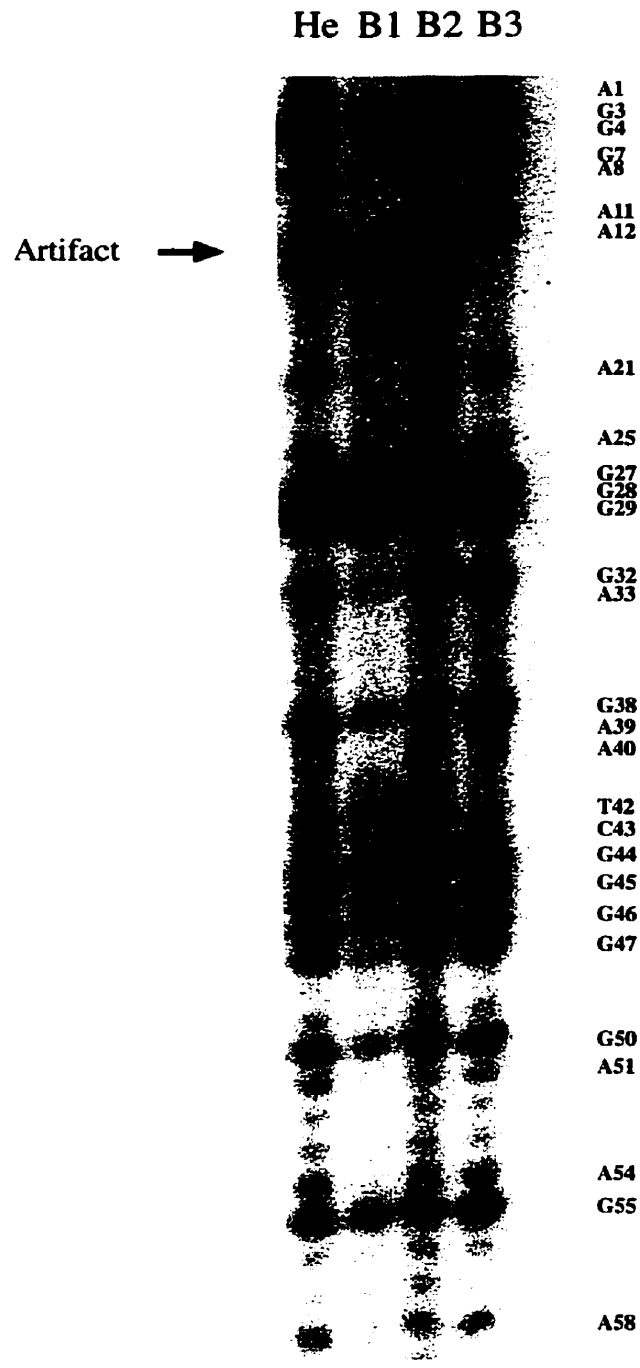
3.10 Footprinting with Minimal DMS and no β-ME

As a final attempt to ensure that the reducing conditions created by the β-ME and the oxidative conditions of the polyacrylamide gel were not responsible for the aberrant preparative EMSA patterns, a final DMS concentration of 0.05 % was used for the methylation, and no stop reagent was added. The piperidine reaction was conducted in the presence of TE rather than water to prevent non-specific purine cleavage during subsequent washes. The preparative polyacrylamide gels were pre-run with regular running buffer to remove reactive side products of acrylamide polymerization (H. Zorbas, Personal Communication). Much the same pattern was observed on this preparative EMSA (Figure 3.18 A) as that from the experiment using 1.9 M β-ME to stop the methylation and in which no attempt was made to neutralize the oxidative conditions of the preparative gel (Figure 3.11). In this experiment, without β-ME, even the fastest moving, much less intense, band was excised and analyzed for a footprint, as well as the two more retarded complexes (Figure 3.18 A). The three histograms are similar (Figure 3.18 C), showing increased reactivity of the three G residues at the 5' elbow junction (bases 27-29), and particularly the slowest migrating complex shows decreased reactivity of the GAA (bases 38-40) at the tip of the cruciform. Enhancement is also seen in the region of the G tetrad (bases 44-47). Protection of G55 in the 3' elbow appears, to varying degrees, in all three complexes. The large variations in the intensities of A11 and A12 are likely caused by an artifact, an intense band not seen in other experiments which ran just ahead of A12 on the footprinting gel, in a position that does not correspond to either a G or an A (Figure 3.18 B). The less clearly significant trends seen in these histograms include protection of A33, and

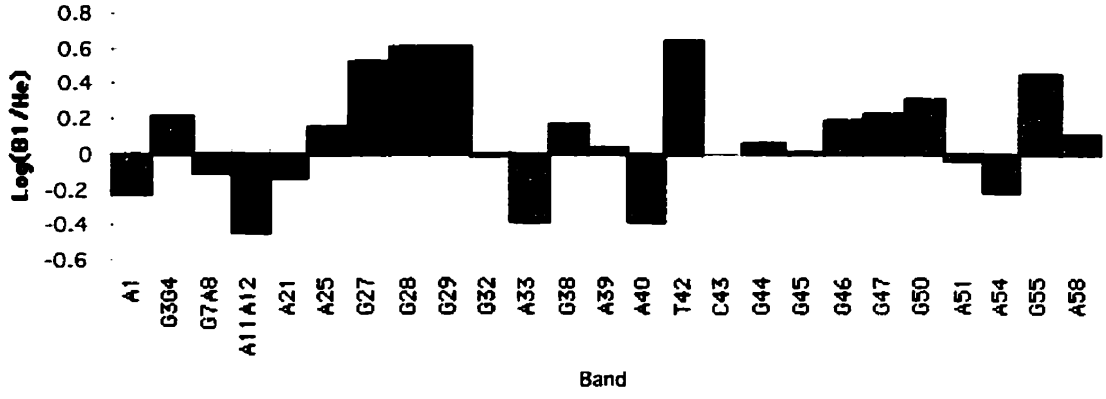
Figure 3.18 Footprinting with minimal (0.05%) DMS and no quenching reagent. A Wet exposure autoradiogram of the 8 % preparative gel used to separate the species of a footprinting reaction, showing controls (lanes 1-3) and preparative samples (lanes 4-18). The image was generated from two separate scans of the same film, due to the difference in radioactivity present in the different lanes. The gel was pre-run for 2 h in 1 x TBE buffer to eliminate oxidative by-products of polymerization. Due to the large amount of protein used in the control, very little of Complex I is observed (lane 2). Numbered bars to the right of the gel correspond to the slices excised. **B** Footprinting autoradiogram of cruciform (He) DNA and the complexes excised from the gel in A (B1, B2 and B3). The four lanes do not contain equal amounts of radioactivity. **C** Quantitative histograms of the autoradiogram in B. **(i)** corresponds to complex B1, **(ii)** to B2 and **(iii)** to B3. Note that the scales of the y-axis are not identical in all the histograms.


A
Top of gel
Complex II
Complex I
Heteroduplex
Homoduplex
3
2
1
1 2 3 4 18
Controls
Preparative Samples

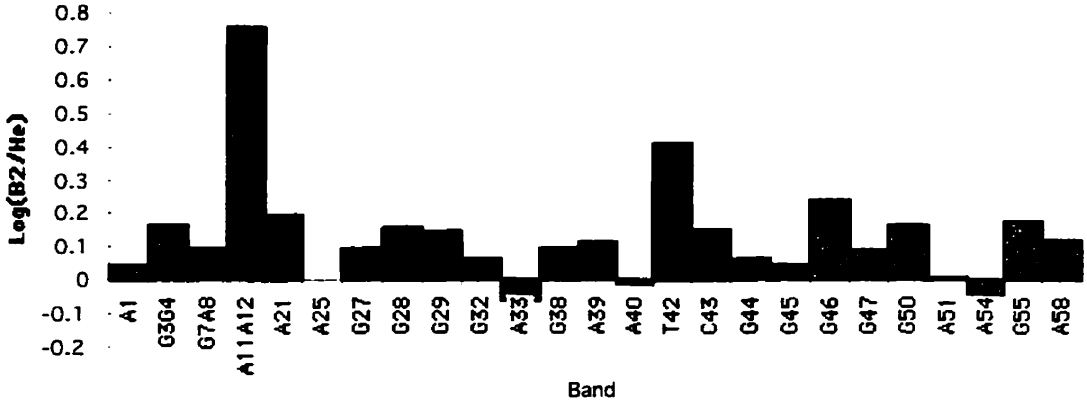
B


He B1 B2 B3
Artifact
A1
G3
G4
G7
A8
A11
A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
T42
C43
G44
G45
G46
G47
G50
A51
A54
G55
A58

C (i)


0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
Log(B1/He)
A1
G3G4
G7A8
A11A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
T42
C43
G44
G45
G46
G47
G50
A51
A54
G55
A58
Band

(ii)


0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
Log(B2/He)
A1
G3G4
G7A8
A11A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
T42
C43
G44
G45
G46
G47
G50
A51
A54
G55
A58
Band

(iii)

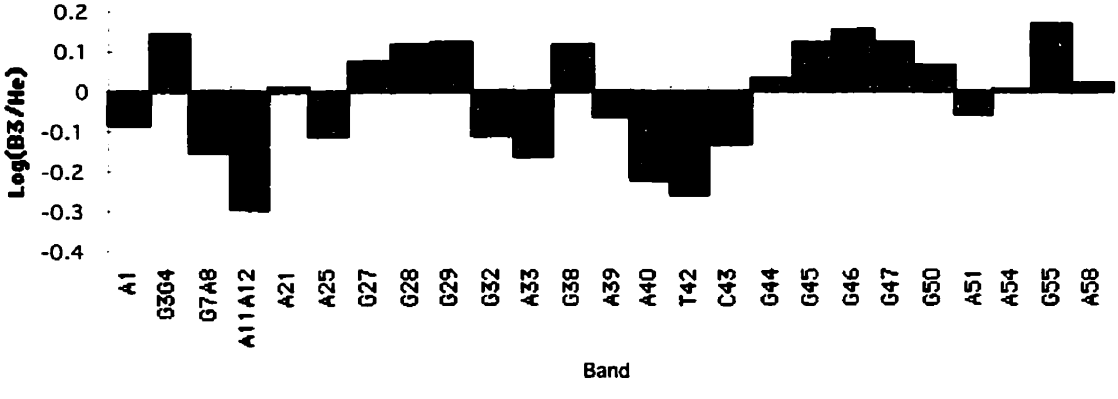
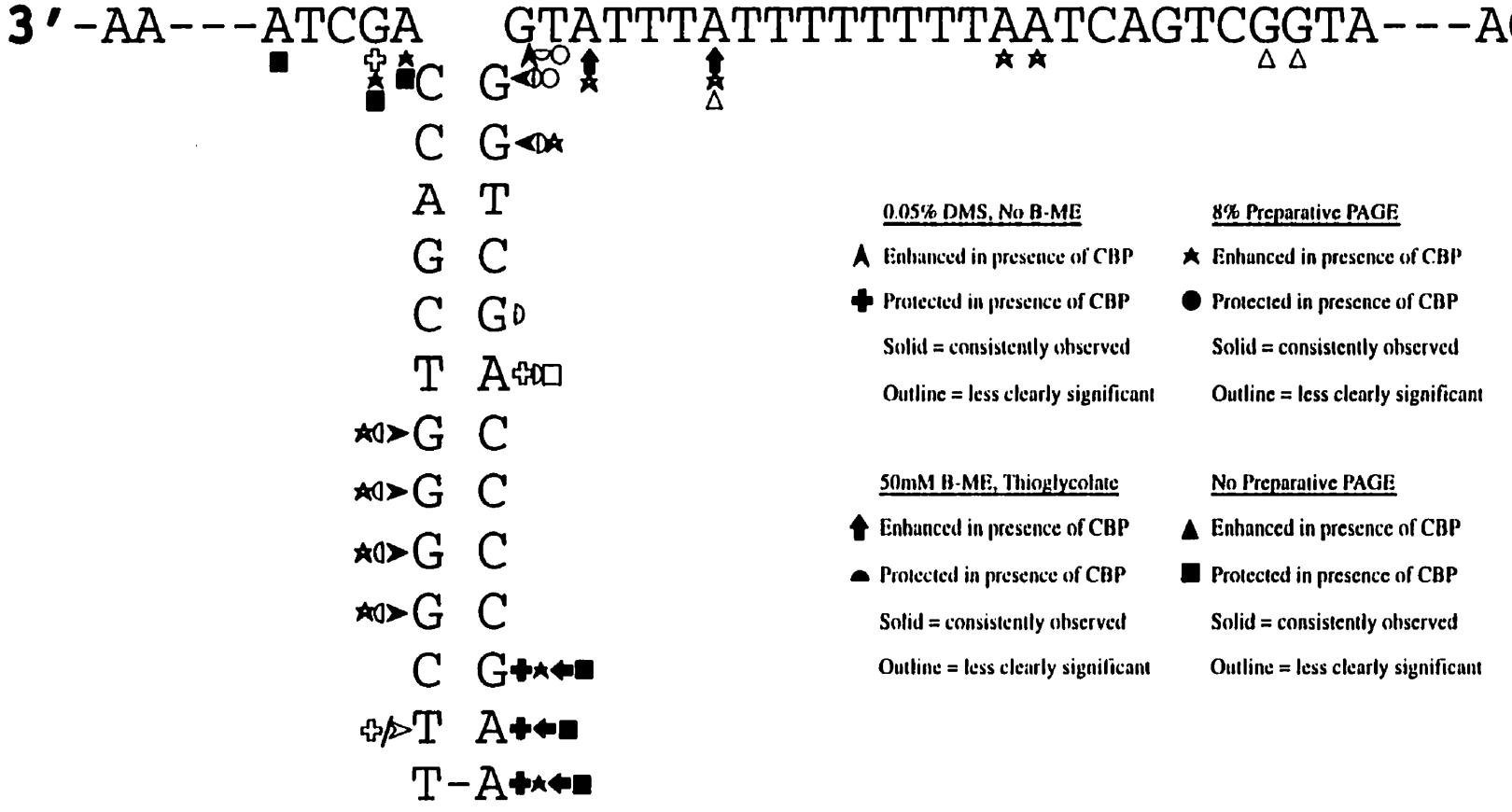

0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
Log(B3/He)
A1
G3G4
G7A8
A11A12
A21
A25
G27
G28
G29
G32
A33
G38
A39
A40
T42
C43
G44
G45
G46
G47
G50
A51
A54
G55
A58
Band

Figure 3.19 Comparison of the DMS footprints obtained with and without a preparative PAGE step, and using minimal β-ME and the thioglycolate scavenger, or minimal DMS and no β-ME. Location of sites of enhanced reactivity and protection from DMS on the cruciform DNA in the presence of CBP, relative to the absence of CBP, based upon the quantitative histograms in Figure 3.9 (no preparative PAGE), 3.12 (8 % preparative gel), 3.16 (50mM β-ME, thioglycolate) and 3.18 (0.05% DMS, no β-ME). Only strand D is presented. The solid arrowhead, outlined half circle and outlined full circle presented horizontally in the 5' elbow region all pertain to the more 5' of the three G's in the elbow.


3′-AA---ATCGA GTATTTATTTTTTTTTAATCAGTCGGTA---AC-5′
C G
C G
A T
G C
C G
T A
G C
G C
G C
G C
C G
T A
T-A
0.05% DMS, No B-ME
Enhanced in presence of CBP
Protected in presence of CBP
Solid = consistently observed
Outline = less clearly significant
8% Preparative PAGE
Enhanced in presence of CBP
Protected in presence of CBP
Solid = consistently observed
Outline = less clearly significant
50mM B-ME, Thioglycolate
Enhanced in presence of CBP
Protected in presence of CBP
Solid = consistently observed
Outline = less clearly significant
No Preparative PAGE
Enhanced in presence of CBP
Protected in presence of CBP
Solid = consistently observed
Outline = less clearly significant

enhancement of T42 in the two lower complexes (Figure 3.18 C (i) and (ii)), but protection of this same T in the slowest migrating species (Figure 3.18 C (iii)). Figure 3.19 summarizes the observed differences in DMS reactivity on the cruciform structure, and compares them to those observed from the previously described experiments (See Sections 3.6, 3.7 and 3.9). This compilation of the trends of modifications of DMS reactivity upon CBP binding demonstrates a significance of a number of sites: the two elbow regions, A21, the G tetrad and the GAA (bases 38-40) at the tip of the cruciform. Figure 3.20 provides an alternative presentation of the recurring sites of modification of DMS reactivity upon protein binding. It is clear that there are regions of the cruciform which are reproducibly affected by the presence of CBP, however in no case is the modification consistently an enhancement or a protection. This limits the information which may be obtained from these results.

3.11 Positive Control: NF-I on its Target DNA

To ensure that the DMS footprinting technique was being carried out correctly and that none of the reagents were somehow "erasing" the footprint, a positive control was conducted. NF-I protein and its target DNA (See Section 2.2) were used for this control. The DNA was end-labeled, quantitated, and sequenced. EMSA titrations allowed determination that a 1000-fold molar excess of protein with respect to DNA substrate, yielded nearly 100 % binding of the DNA (data not shown). DMS footprinting under these conditions (without a preparative PAGE step) showed the protection of two guanines in the centre of the binding site for the protein (Figure 3.21). This corresponds exactly to the position of the footprint obtained with DNaseI and hydroxyl radicals, and to that expected for the DMS footprint ([121] and H. Zorbas, Personal Communication).

Figure 3.20 Sites of modification of DMS reactivity in the presence of CBP. Asterisks denote nucleotides of the cruciform DNA which were repeatedly observed to be enhanced or protected from DMS attack by the presence of CBP. This figure was made by simplifying the information presented in Figure 3.19. There were no sites for which only enhancement or only protection were repeatedly observed.

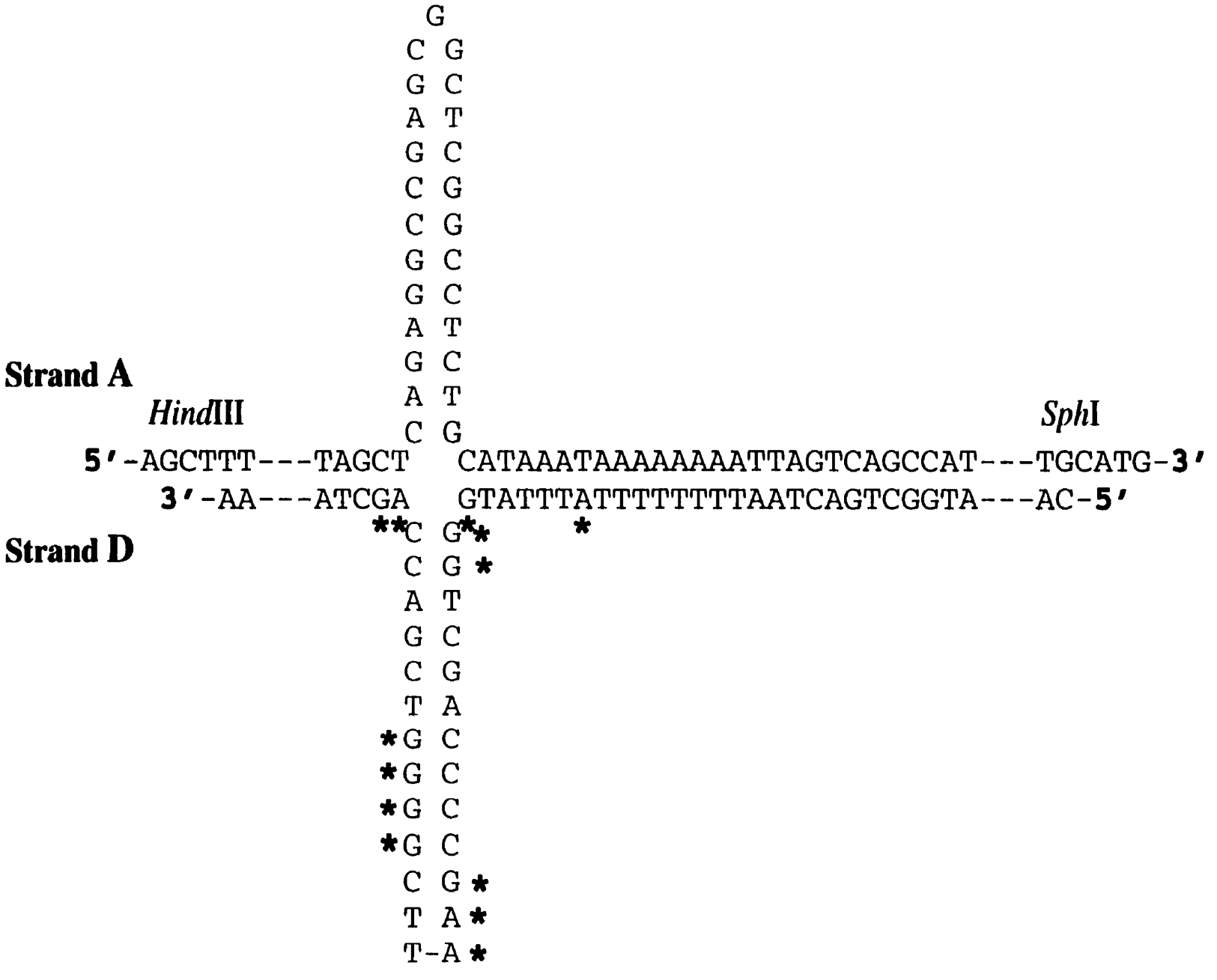
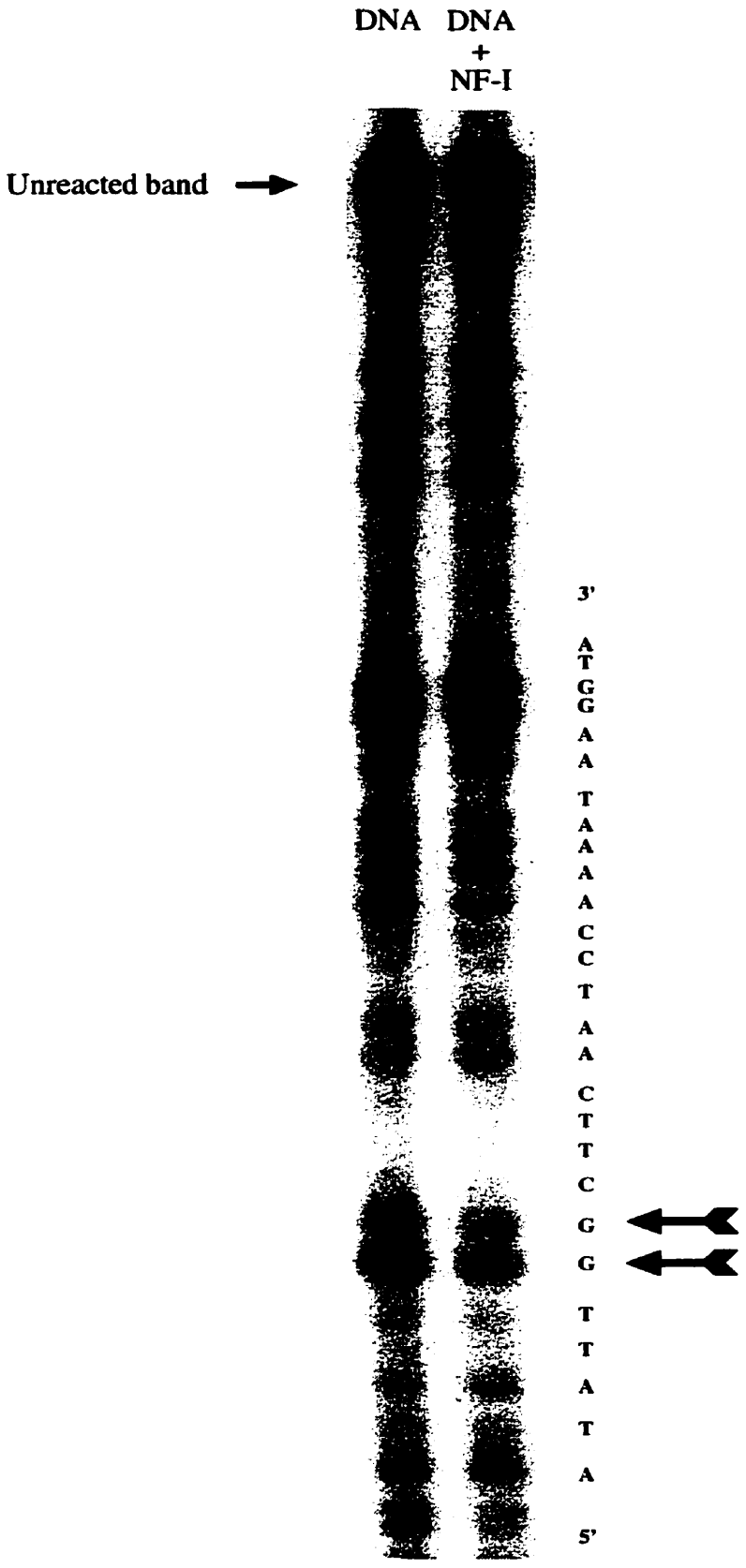

Strand A
HindIII
SphI
5′-AGCTTT---TAGCT
CATAAATAAAAAAAATTAGTCAGCCAT---TGCATG-3′
3′-AA---ATCGA
GTATTTATTTTTTTTAATCAGTCGGTA---AC-5′
Strand D

Figure 3.21 DMS protection footprinting of NF-I on its target DNA - the positive control. DMS footprinting autoradiogram of the interaction of NF-I with its target DNA from type 5 adenovirus (See Sections 2.2 and 3.11 for details), showing the base sequence for the binding region. The two large arrows mark the clear protection of two G bases. These are precisely the two bases predicted to be protected by the binding of this protein.


DNA
DNA + NF-I
Unreacted band
3'
ATGGAAATAAAAACCCTAAACTTTCGGTTTATTA
5'

4. DISCUSSION

Hydroxyl radical footprinting of the interaction of CBP with cruciform DNA allowed the proposal of a model for the structure of the bound DNA and the mode of binding (Figure 1.6, lower panel) [41]. This model suggests that there is an inversion in the binding orientation of the two complementary cruciforms of the 21/29 system with CBP (See Section 1.9 a)). No such inversion was noted for the interaction between the same cruciforms and a cruciform-specific antibody, also studied by hydroxyl radical footprinting [115]. It appears, then, that this inversion is particular to the CBP-cruciform interaction and its verification and further study provide an avenue towards a better understanding of this structure-specific binding, revealing elements important to its biological role. Protection DMS footprinting was selected to pursue this investigation (See Section 1.10).

4.1 Appearance of the CBP-Cruciform EMSA

Combining the CBP-enriched fraction with labeled cruciform, under conditions conducive to binding, results in the formation of two principal protein-DNA species separable on a 4 % polyacrylamide gel, and sometimes a third, fainter, more retarded band (Figures 3.2 and 3.3). HPLC profiles of the products of tryptic digestion of two polypeptides eluted from PAGE purified CBP-cruciform complexes suggest the possibility that the faster migrating of the two major species is a degradation product of the slower [42]. Microsequence analysis supports this theory as peptides from both were found to have 100 % homology to the ε, and the β and/or ζ isoforms of 14-3-3 [42]. The more retarded, much fainter band, could result from a subpopulation of the protein that has undergone a conformational

variation that slows its migration through the gel matrix, but does not eliminate its cruciform binding activity.

4.2 Comparative DMS Reactivity of the Homoduplex and Heteroduplex DNA

The pattern of DMS methylation, and subsequent cleavage, of the heteroduplex DNA was observed to be somewhat different from that of the linear homoduplex (Figures 3.6 and 3.7). The amplitude of the differences varied between experiments; in fact it appeared that the background non-specific cleavage which may result from the slightly heated and acidic conditions following piperidine treatment in water rather than TE, pH 7.6, may have been enough to mask any differences. However, there was a general trend of increased reactivity of the adenine bases. DMS methylation of adenines occurs at the N-3 through the minor groove (Figure 1.8). A molecular mechanical computer modeling study of a four-way junction predicts a widening of the minor groove of the structure [16] and this feature has been proposed as one of the keys in recognition by its binding partners [81]. Such a widening of the avenue of attack for the DMS could be responsible for the increased adenine reactivity.

The alterations in DMS reactivity observed are different from those seen upon comparison of hydroxyl radical reactivity of the homoduplex and heteroduplex DNAs [41]. The hydroxyl radical study found reduced strand cleavage of most bases in the region of the junctions of the cruciforms. ss DNA can scavenge radicals and could decrease the effective concentration of the cleaving agent in the vicinity of these bases [126]. However, the fact that the ss regions of DNA at the tips of the cruciform were unaffected suggests the additional involvement of some other structural feature in the modified reactivity of these

bases in the heteroduplex DNA. Since DMS methylation does not involve radicals, regions of single-stranded character would not be expected to have this scavenging effect on DMS reactivity with adenines and guanines. It appears that the other structural factors believed to contribute to the altered hydroxyl radical susceptibility do not have a significant effect on DMS methylation. The DMS molecule is significantly larger than the hydroxyl radical: their molecular volumes are 108Å^3 and 23Å^3, respectively.[1] As a result of this difference in steric bulk, small changes in the structure of the DNA may not have an observable effect on the ability of DMS to access potential sites of methylation.

4.3 Summary of DMS Footprints Observed

As mentioned at the beginning of the Discussion, the goal of this research was to perform protection DMS footprinting of the binding of CBP to cruciform DNA in order to test the model of binding proposed from the hydroxyl radical footprinting study [41]. A comparison of Figure 3.20 with Figure 1.6, upper panel, demonstrates that the DMS footprinting experiments presented herein do provide evidence for the binding of CBP to the cruciform. The sites of recurring modification of DMS reactivity correspond to the regions which are seen to be protected or enhanced in the hydroxyl radical footprinting [41]. The DMS experiments repeatedly show variations in reactivity: (a) of the adenines and guanines located in the junctions of the cruciform, which are protected from hydroxyl radical attack; (b) at the tip of the cruciform and on either side of the stem

[1] Molecular volumes were calculated by Graeme Day at the Centre for Theoretical and Computational Chemistry, Department of Chemistry, University College London, using the Gaussian 98 electronic structure package [127]. The molecular volume was calculated as the volume inside an envelope of electron density of 0.001e/bohr^3. Geometry optimizations and molecular volume calculations were performed using RHF/6-31G* and UHF/6-31G* for DMS and the hydroxyl radical, respectively.

adjacent to the tip, which are sites of protection and enhancement, respectively, in the hydroxyl radical experiments; and (c) in the AT tract of the 5' arm of the cruciform for which the hydroxyl radical experiments give evidence of contact and/or structural alteration by CBP. However, the signals obtained from DMS footprinting are neither as clear nor as reproducible as those obtained with the hydroxyl radical technique. The affected bases are protected in some experiments, but not in all, or protected in some and enhanced in others (Figure 3.19). Therefore, while these experiments do provide support for the regions of bases influenced by the binding of CBP, they do not provide information which is clear and reproducible enough to support or refute the inversion of major/minor groove presentation by the two complementary 21/29 cruciforms. This verification would require the precise comparison of the protection or enhancement of each adenine and guanine, of all four strands, to determine the sites of major and minor groove presentation (G protection indicating major groove contact by the protein, and A protection indicating minor groove contact). Such a comparison cannot be confidently made if the signals from each base are not highly reproducible. For this reason, the experiments were not repeated using DNA specifically end-labeled on the other three strands of the two cruciforms.

4.4 DMS Reactivity of Cytosines and Thymines

When DNA is exposed to DMS, the principal sites of methylation are the N-7 of guanine and, to a lesser extent, the N-3 of adenine (Figure 1.8, [116]). However, there are other minor products of DMS methylation (reviewed in [128]), and alterations in the structure of the DNA affect the availability of potential methylation sites [129]. In particular, regions of ss DNA are marked by the reactivity of cytosine, as the N-3 normally involved in hydrogen bonding to the

complementary strand becomes available for methylation [116]. This is seen as the appearance of bands on the autoradiogram, at positions corresponding to cytosine in the sequence. This reactivity has been developed into a technique for the detection of regions of ss DNA and the study of RNA structure [130].

Some reactivity of both thymines and cytosines was observed in the experiments reported herein. In addition to the low general background signal, the specific presence of a band corresponding to cleavage at T42 (for example Figure 3.18 B) has been seen. Also, several instances of C43 reactivity (for example Figure 3.16 B) are observed. These two nucleotides are located between the G tetrad and the tip of the cruciform arm (See Figure 3.5 C), suggesting that this particular region may be prone to structural perturbations. The fact that T reactivity was seen, with equal intensity, in the free cruciform as well as the shifted complexes, suggests that it is not a result of protein binding. There were also occasions when this band was seen upon DMS treatment of the homoduplex (data not shown). No clear explanation has been proposed for the appearance of thymine cleavage products in DMS reactions with DNA [128], but they have been observed in a number of other DMS studies (for example [131], [132]).

In most of the footprinting autoradiograms of this study, a band, of varying intensities, is seen corresponding to cleavage at C43 from both the free cruciform and the complexed DNA. In Figure 3.16 B, the evidence for C43 methylation is most strongly seen in the DNA from the upper-most complex, suggesting that the binding of the protein may influence the degree of hydrogen bonding, and therefore cytosine N-3 availability, at that point in the DNA. A recurring trend of modification of the reactivity of this base was observed upon protein binding, however, both enhancement and protection were seen with approximately equal frequency (See quantitative histograms). This suggests that C43 is within the

region of the DNA contacted or affected by the protein, as observed in the hydroxyl radical footprinting experiments [41], and the extent to which it engages in hydrogen bonding with its partner, G38, may be influenced by protein binding. This is not an unexpected trend considering that there is incomplete pairing and stacking of 3-4 bases at the tips of the arms formed by the hairpin loops of cruciforms [1], which is precisely the location of these nucleotides.

4.5 Possible Explanations for the Lack of Clear, Reproducible DMS Footprint

There are two principal explanations for the lack of a clear and reproducible DMS footprint from a protein-DNA interaction, for which there is strong evidence: the binding is either too transient or too "loose" (i.e., the contact of the protein with the DNA is not close enough) to prevent DMS methylation of the DNA within the bound region [133].

4.5 a) Transient protein-DNA association

If the interaction of a protein with DNA has high rates of association and dissociation, and the actual binding is transient, then there would be adequate occasion for the DMS to methylate the DNA within the binding region, during periods of dissociation, and no protection would be observed. The CBP-cruciform interaction does not appear to involve a particularly transient association. The hydroxyl radical footprinting experiments were carried out using this system [41] at room temperature for reaction times of 5 min (H. Zorbas, Personal Communication). These conditions were conducive to a clear footprint. The DMS experiments, outlined herein, were carried out at 20 °C for 4 min. These conditions would, if anything, provide less opportunity for association and dissociation.

Therefore, it is unlikely that a transient nature of the interaction is responsible for the absence of a clear DMS footprint.

4.5 b) "Loose" protein-DNA interaction

The other possible explanation for the lack of a footprint is that the presence of the protein does not preclude DMS attack of the bases, that is, the contact of the protein with the bases is not close enough to prevent DMS penetration. Hydroxyl radical footprinting assays exclusively the protection of the backbone of the DNA and does not indicate the extent to which the bases are contacted [104]. Therefore, the hydroxyl radical footprinting patterns reported [41] do not imply protection of the bases in those regions of the DNA. DMS interference footprinting and hydroxyl radical missing-contact analysis of the CBP-cruciform interaction gave evidence for no essential contacts between the protein and any bases of the cruciform DNA [41]. This observation is in concurrence with the sequence-independent, structure-specific nature of this interaction [40]. An interaction based upon structure recognition cannot require specific base contacts and retain its strictly structure-dependent nature. It is, however, possible for the binding of a protein to consistently protect particular bases from attack, without contact with those bases being essential for the binding of the protein. The essential interactions may be with a very limited and specific portion of the DNA, but the steric bulk of the protein may result in protection of a much larger region. Alternatively, it is possible that the essential interactions are with the backbone of the DNA and that the orientation of the protein is such that it does not closely contact any of the bases, allowing DMS methylation of all the adenines and guanines with approximately equal facility. This may explain the lack of consistent, clear DMS footprint on the cruciform DNA despite strong evidence of protein binding. An example of the importance of protein interactions with the DNA backbone is seen in the X-ray

crystal structures of the RuvA protein complexed with its cruciform DNA substrate. They show that protein contacts with the sugar-phosphate backbone of the DNA are key to this interaction [15], [75].

Another possibility, related to the potential "looseness" of protein-DNA binding, is that the tightness of binding may be affected by DMS methylation. Though methylation interference studies showed that methylation of any base can occur without interfering with the formation of the cruciform-CBP complex [41], it is possible that DMS treatment of the DNA/protein mixture (which is the case for protection but not interference footprinting experiments in which the DNA alone is treated with the footprinting probe) may affect the closeness of the interaction. Perhaps methylation of the protein could result in a change in its conformation, such that it still binds and shifts the DNA in an EMSA, but the interaction may be looser than with unmethylated protein. This would result in an increased access of the footprinting probe to the whole DNA sequence.

4.6 Possible Explanations for the Aberrant Appearance of the Preparative EMSAs Used for Footprinting

Repeatedly, it was observed that the preparative polyacrylamide gels on which the DNA species were separated following DMS treatment of the DNA-protein binding reaction, did not resemble the control reactions (Figures 3.11, 3.16 A and 3.18 A). This made it difficult to be confident that the species analyzed for a footprint were indeed the cruciform DNA bound to a single dimer of CBP, since they did not migrate in the expected position for such a complex. There are three principal potential explanations for this observation: (a) the effect of the contrasting reducing and oxidative environments of the DMS quenching and the polyacrylamide gel; (b) protein-protein interactions due to presence of many proteins with the

potential to bind to 14-3-3; and (c) methylation of the DNA and/or protein(s) resulting in protein-protein or protein-DNA interactions that create complexes with different electrophoretic mobilities than the controls.

4.6 a) Reducing/oxidizing environments

The conditions proposed to be optimal for footprinting experiments, on the basis of a preliminary experiment with a mock binding reaction, using Buffer B (See Section 2.1 b)) instead of the CBP-enriched fraction, involved the addition of β-ME to a final concentration of 1.9M to quench the DMS reaction (See Section 3.6). In addition to the desired effect of eliminating further DNA methylation by the DMS, this level of β-ME would create a highly reducing environment which would be expected to drastically affect the conformation of any proteins present. β-ME is usually used at concentrations of 5 mM to 100 mM as a protein reducing agent [124]. In addition, the polymerization of polyacrylamide, catalyzed by free radicals from APS, results in the presence of oxidative by-products [125]. Loading a protein mixture, which has been reduced by 1.9 M β-ME, onto such an oxidative environment would be expected to result in a very rapid and non-specific oxidation of the proteins. This could either cause complexing of different proteins to the CBP bound to the cruciform, or alter the conformation of CBP itself, and thus the migration of the DNA-protein complex in the gel. To circumvent this problem we tried first reducing the amount of β-ME and introducing the free radical scavenger, thioglycolate, into the gel running system. The preparative EMSA still contained unexpected bands (Figure 3.16 A). We then reduced the final concentration of DMS to the point (0.05 %) that it exhausted its methylating capabilities within the chosen reaction time, and no quenching reagent was necessary. This, combined

with pre-running the gel in its regular running buffer to remove the oxidative by-products of acrylamide polymerization, should have eliminated the reduction/oxidation conditions proposed to be potentially responsible for the aberrant bands. The presence of these bands even under these conditions (Figure 3.18 A) indicates that the contrasting reducing and oxidative conditions were most likely not responsible for their occurrence.

4.6 b) Potential 14-3-3 binding partners present in the CBP-enriched fraction

Another potential explanation for the presence of the unexpected bands in the preparative EMSAs, especially those of higher molecular weight, could be protein-protein interactions between CBP and the potential binding partners of 14-3-3 present in the CBP-enriched fraction. In 1 μL of CBP-enriched fraction, there is 5 μg of total protein, only approximately 15 ng of which is active CBP (See Section 3.2). This means that there are many other proteins present, some of which may have an affinity to bind to members of the 14-3-3 family (reviewed in [48], [50], [52], [53]), of which CBP is one. This could result in non-covalent association with non-CBP proteins, or oligomerization of CBP, in addition to the expected CBP binding of the cruciform. These interactions would not necessarily be disrupted by electrophoresis on a native polyacrylamide gel and would result in bands at positions other than those expected for the simple CBP-cruciform complexes. However, if this is the case, then these bands should also be present in control binding reactions involving the same final concentrations of protein and DNA, even if they are performed on an analytical rather than preparative scale. This is not the case (for example Figure 3.11), therefore this cannot be the explanation for the EMSA appearance.

4.6 c) Possible effects of DMS methylation on protein-protein or protein-DNA interactions

The fact that control binding reactions performed under conditions identical to the preparative reactions, with the exception of the DMS treatment, do not exhibit the unexpected band appearance (Figures 3.11, 3.16 A and 3.18 A) suggests an involvement of the methylation process. The proteins present in the binding reaction are also susceptible to methylation [134] and therefore their interactions with one another, and with the DNA, also as a result of DNA methylation, may be altered.

(i) An effect of DNA methylation on protein binding

It is possible for DNA methylation to influence the affinity with which proteins bind. The majority of research into the effect of DNA methylation on protein binding has focused on the methylation of cytosine in CpG islands, particularly with respect to trasncriptional silencing functions (Reviewed in [135-137]). However, Wang *et al.* have reported an increase in binding of the REB1 protein to its (linear) substrate DNA upon methylation of a particular adenine, in DMS interference experiments [138]. They suggested that the protein may bind preferentially to DNA bearing a slight distortion, and that this particular methylation could stabilize that distortion. A subsequent NMR study supported this hypothesis [139].

In the case of the system reported herein, a simple increase in the binding affinity of CBP for cruciform DNA would not explain the observation of species of different molecular weights. Instead, the binding of a protein to the methylated DNA, which does not bind unmethylated DNA, could be suggested as a potential explanation for the band patterns observed in the preparative EMSAs. However, no such binding was observed for homoduplex DNA identically treated with DMS,

in the presence of the CBP-enriched fraction (data not shown). Therefore, any such protein would have to bind specifically to methylated cruciform, or depend on the prior binding of CBP to the DNA, to be capable of binding itself. Possible candidates for a protein or proteins that might bind specifically to methylated DNA, with a requirement for the cruciform structure and/or the presence of CBP, would include proteins involved in the repair of methylation damage to DNA.

Repair of alkylation damage to DNA has been shown to involve both the base excision repair (BER) and the nucleotide excision repair (NER) pathway, with the BER being of prime importance for *N*-methyl purines (reviewed in [140]). The majority of research on this pathway has been conducted with *E. coli*, however variations of this repair system are believed to exist in all cells. Though the majority of methylation adducts are formed at the N-7 of guanine it is the alkylation of the N-3 of adenine which constitutes the greater threat to the survival of the cell, and therefore elicits the stronger repair response. A number of methylpurine-DNA glycosylases (MPG proteins), the enzymes which carry out the first step in the BER pathway, have been cloned from mammalian cells (reviewed in [140]). Study of the mouse MPG protein indicated that one of its principal functions is the protection of the cell from damage due to purine alkylation [141]. It is possible, therefore, that a protein involved in the BER pathway is present in the CBP-enriched fraction and can bind to the methylated DNA. However, to explain our observations, this binding would have to be cruciform and/or CBP-dependent. The mammalian MPG proteins have not been adequately characterized to allow speculation on the likelihood of such a dependence.

(ii) An effect of protein methylation on DNA-binding activity

Conversely, methylation of a protein, or proteins, present in the CBP-enriched fraction could result in an altered affinity for the DNA resulting in binding

that would not occur without DMS treatment. This phenomenon also would have to be cruciform-specific and/or CBP-dependent to explain the observations made in these experiments. There exists also the possibility that, in the context of methylation, CBP facilitates binding of a protein to the DNA, and then dissociates itself. This could cause the observed shifts in DNA position that do not correspond to the controls, but without leaving a footprint in the CBP-binding region of the DNA. These explanations, though not impossible, seem unlikely and the research reported herein does not support the drawing of a conclusion.

(iii) An effect of protein methylation on protein-protein interactions

Perhaps the most likely explanation of the bands seen in the preparative EMSAs is that the methylation results in an alteration in the protein-protein interactions in the binding reaction mixture. In addition to methylating DNA, DMS does methylate proteins [134]. In the cell, methylation of proteins is usually carried out by methyltransferases that use *S*-adenosylmethionine as the source of methyl groups (reviewed in [142] and [143]). Nucleophilic oxygen, nitrogen and sulfur atoms provide the sites of methylation on the polypeptide backbone, nine of the 20 common amino acid side chains, and other side chains specifically if they are located at the amino or carboxy terminus of the polypeptide [143]. These modifications can result in a number of significant changes in their capacities to mediate interactions with other molecules. For instance the conversion of glutamate to the glutamate methyl ester eliminates one negative charge. Conversely, the addition of three methyl groups to lysine results in the establishment of a fixed positive charge. Methylation of some amino acids, such as arginine, may disrupt their hydrogen-bonding capacities (reviewed in [143] and [144]).

Such changes in the interaction capacities of the amino acids could result in associations between proteins which would not occur in the absence of methylation.

In fact, this is thought to be a key mechanism for the biological effects of protein methylation, which include modulation of the interactions of signaling proteins, a role in the metabolism of damaged proteins, affecting membrane association of otherwise soluble proteins, and regulation of substrate affinity of certain RNA-binding proteins (reviewed in [142], [143], [145] and [144]). In the case of the CBP-enriched fraction/cruciform DNA binding reaction mixture, methylation of CBP, or other proteins present in the fraction, could result in an association of proteins with the DNA-bound CBP that would not occur without DMS treatment. This would result in the formation of complexes of higher, or different, molecular weights than those observed in the absence of methylation. There is also evidence for an interplay between methyltransferases and demethylating enzymes, the latter providing a candidate for a class of proteins that may bind specifically to other proteins following methylation [143].

4.7 Suggestions which May Make Examination of the Putative Inversion of the Orientation of the Two 21/29 Cruciforms Possible

4.7 a) Further purification of CBP

The interference of other proteins, whether methylation-dependent or not, in the DMS protection footprinting experiments, could be decreased or eliminated by further purification of CBP. The presence of approximately 15 ng of active CBP in 5 μg of total protein (in 1 μL of the CBP-enriched fraction, See Section 3.2) underscores just how much of the protein present is not that which we wish to study. Many other proteins, and possibly inactive forms of CBP, may be present in the preparation used for these experiments. While this proved to be adequate for

the hydroxyl radical footprinting studies [41], it is likely that a CBP-fraction of greater purity would be necessary for more successful DMS footprinting attempts.

Toker *et al.* have developed a protocol for the purification of 14-3-3 from sheep brain using a combination of anion-exchange and hydrophobic chromatography steps [146]. They start with homogenization of the source tissue in the presence of protease inhibitors, followed by centrifugation of the homogenate. The supernatant is applied to a DEAE-cellulose (a weak anion exchanger) column, the column washed extensively with a Tris-based buffer (20 mM Tris/Cl pH 7.5, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, hereafter referred to as Buffer A), and then proteins eluted with a linear NaCl gradient (0 - 0.5 M). Two peaks of 14-3-3 result and may be pooled separately. The NaCl content is increased to 2.5 M and the pooled fractions loaded onto a phenyl-Sepharose CL-4B (a hydrophobic gel with no ionic properties) column, equilibrated with 2.5 M NaCl in Buffer A, and the proteins eluted with a linearly decreasing NaCl gradient (2.5 - 0 M). Active fractions are then pooled and dialyzed against Buffer A and loaded onto a Mono Q (a strong anion exchanger) column from which, following washing with the buffer used for dialysis, proteins are eluted using a biphasic NaCl gradient: 0 - 0.6 M NaCl, followed by 0.6 - 1.0 M. This yields a single peak containing several isoforms of 14-3-3, but no other proteins as detected by silver staining [146]. The application of this process to the CBP-enriched fraction, using EMSAs to assay for cruciform binding activity in each of the steps rather than the protein kinase C inhibition used by Toker *et al.*, should yield a purer fraction of 14-3-3 with cruciform binding activity. Alternatively, the process could be applied directly to cellular extracts. Such a fraction was not available when the studies reported herein were undertaken.

Further purification of the 14-3-3 isoforms in the final fraction obtained from this protocol was achieved, by the same group, using reverse-phase HPLC [122]. This allowed complete separation of the different isoforms. If applied to the CBP-enriched fraction it could provide a means to explore the isoform specificity, if any, of the cruciform binding activity of this protein. Furthermore, the resulting, highly pure, cruciform binding protein could provide an ideal subject for further studies, footprinting and other.

4.7 b) Affinity chromatography

An alternative purification approach, for 14-3-3, would be to use an affinity column with a commercially available pan anti-14-3-3 antibody, such that all 14-3-3 isoforms might be separated from other proteins in a mixture. However, the tendency of 14-3-3 to interact, non-covalently, with a wide variety of proteins (see reviews [48], [50], [52], [53]) suggests that at least some of these interactions may be favoured by the same conditions as the 14-3-3-antibody interaction. A satisfactory separation of 14-3-3 isoforms from the proteins with which they interact would, therefore, be doubtful. One might also suggest the purification of CBP by passing the CBP-enriched fraction over an affinity column that uses cruciform DNA, affixed to the column matrix, to select cruciform binding proteins from any mixture. This was attempted, in our laboratory, and the production of the amount of cruciform necessary for such an endeavour proved impractical (A. Todd, Unpublished Results).

4.7 c) Recombinant 14-3-3

Since CBP has been demonstrated to be a member of the 14-3-3 family of proteins, [42] another approach to the detailed characterization of the cruciform-protein interaction would be to use purified recombinant 14-3-3. Care would have to be taken to use a mammalian cell line, since recombinant 14-3-3 prepared from

bacterial cells does not bind cruciform DNA (A. Todd, Unpublished Results). Nuclear extracts prepared from HeLa and CV-1 cells transfected with plasmids expressing the cDNA of myc-tagged ε, γ, and ζ isoforms of 14-3-3 do possess cruciform binding activity (A. Todd and F. Robinson, Unpublished Results, data not shown). Attribution of this activity to 14-3-3 could be confidently made if super-shifting of the DNA were observed upon the addition of an anti-myc antibody to the binding reaction. This was not successfully achieved with these preparations. There are two possible explanations for this result: (a) that the cruciform binding activity does not have a myc tag, or (b) that the myc tag on the protein binding to the cruciform is subsequently unavailable for antibody recognition. We have not yet determined which is the case.

A disadvantage of working with recombinant 14-3-3, rather than purifying to homogeneity the activity in the CBP-enriched fraction, is that we do not know which combination of the 14-3-3 isoforms possess cruciform binding activity. Microsequencing, Western and other analyses (See Section 1.5 b)) showed that CBP is a member of the 14-3-3 family of proteins, and demonstrated the presence of the ε, β, γ and possibly ζ isoforms in the cruciform binding activity [42]. However, we do not know in what combination, and whether as homodimers or heterodimers, these isoforms act. This makes it difficult to select which isoforms to work with, and in which ratios. It would perhaps be more efficient to further elucidate this point using the chromatographic purification scheme outlined above, and then work with the appropriate recombinant protein(s), if it is more convenient.

The other unknown factor in this study is the post-translational modification state of the cruciform binding 14-3-3. The fact that the recombinant 14-3-3 purified from bacterial cells does not exhibit cruciform binding activity (A. Todd,

Unpublished Results) suggests that a post-translational modification carried out in mammalian, but not bacterial, cells may be important. It is not known how this modification might affect the partitioning of the protein possessing the cruciform binding activity in the purification steps discussed above. The other possible explanation for the lack of activity of the bacterially produced recombinant 14-3-3, is that the bacteria may not achieve the correct folding of the protein ([147, 148] and references therein). For both of these reasons, it would be very important to select purification fractions on the basis of cruciform binding activity, and not other characteristics of the 14-3-3 proteins, and to use mammalian cells for the production of recombinant proteins.

4.7 d) 1,10-Phenanthroline copper footprinting as an alternative strategy

The compound 1,10-phenanthroline-copper (OP-Cu) is a nuclease which may provide an alternative footprinting strategy for the determination of the major/minor groove presentation, by the two complementary 21/29 cruciforms, to CBP [149]. The tetrahedral coordination complex $(OP)_2Cu^{2+}$ binds to the minor groove of B-DNA and, upon addition of hydrogen peroxide, is oxidized to a species which attacks the deoxyribose moiety and results in cleavage of the phosphodiester bond [89]. As such it is a good probe for the protection of the minor groove by proteins, or other ligands. The cleavage is sequence-independent and would therefore provide information about the groove presentation of the cruciform DNA to CBP at all positions of interaction, not just adenines and guanines as with DMS. In light of the possibility that CBP may not closely contact any bases of the cruciform DNA (See Section 4.5 b)), another advantage of this method is the fact that it cleaves a component of the sugar-phosphate backbone and does not require base contacts to provide information about the interaction [149].

The hydroxyl radical footprinting of the CBP-cruciform interaction demonstrated that it does indeed form close enough contacts with the backbone to protect it from hydroxyl radical attack [41]. As the chemistry of OP-Cu strand scission is similar to that employed in hydroxyl radical footprinting, the success of the latter approach suggests that the CBP-cruciform interaction could also influence the OP-Cu reactivity of the DNA.

Another advantage of the OP-Cu technique is that it can be carried out within the matrix of the polyacrylamide gel used to separate free and protein-bound DNA [150]. By excising the species of interest (following a wet exposure of the gel) and treating only the separated gel fragments with OP-Cu, single electrophoretic species may be footprinted. The risk of footprinting chemicals affecting association of the protein-DNA complex with other proteins, or otherwise affecting the migration of the species in the gel, is minimized.

The binding specificity of OP-Cu for B-DNA constitutes a potential problem for the use of this probe to investigate the CBP-cruciform interaction. The correct geometry in the minor groove of B-DNA is essential to the binding of the OP-Cu; if it is significantly distorted, the complex cannot bind. Though the precise structure of the 21/29 cruciforms is not known, recent theoretical and crystallographic studies do demonstrate a predominantly B-form DNA structure in both the stacked-X and the open conformations of four-way junctions [16], [15], [86], [75], [151]. Cruciforms differ from Holliday junctions, with which the majority of studies have been conducted, in that they feature incomplete pairing and stacking of 3-4 bases at the tips of the arms formed by the hairpin loops [1]. Whether or not the deviation from normal B-DNA structure would be enough to prevent the useful employment of OP-Cu footprinting could be simply determined by comparing the reactivity of the naked cruciform DNA to the corresponding linear

DNA. A lack of cleavage in the regions of interest, in the absence of protein, would preclude OP-Cu protection footprinting for further study of cruciform DNA-protein interactions.

The other situation in which OP-Cu footprinting would fail to provide useful information would be if CBP contacts the cruciform DNA exclusively through the major groove, making probing of the minor groove futile. Although this is certainly possible, the minor groove has been proposed to be instrumental in the binding of the HMG proteins to their DNA substrates ([81] and references therein), and shown to be important to that of the RuvA [15] and Cre proteins [114], [151] with their respective substrates.

5. CONCLUSIONS

DMS footprinting of the CBP-cruciform interaction supports the model of protein interaction sites on the DNA proposed from hydroxyl radical footprinting [41]. However, the DMS footprint is not clear or reproducible enough for determination of the major/minor groove presentation of the two complementary 21/29 cruciforms to CBP. Therefore, the fine structure of the model remains untested. Further purification of CBP, exploiting the protocols available for the purification of other members of the 14-3-3 protein family, would yield a preparation better suited to further studies. The OP-Cu footprinting technique provides an alternative, perhaps preferable, approach to the procurement of the major/minor groove presentation information, which would allow an evaluation of the current model for the CBP-cruciform interaction, and a more complete understanding of this unique structure-specific binding.

6. ACKNOWLEDGMENTS

I would like to express my appreciation to my supervisor, Dr. Maria Zannis-Hadjopoulos, for giving me the opportunity to do this research. I would also like to thank Dr. G. B. Price of the McGill Cancer Centre and Dr. H. Zorbas of the Genzentrum of the University of Munich for all their support and input into this project. I am grateful to the Genzentrum of the University of Munich, and all its members, for their support during my stay there. I would like to thank all my co-workers and colleagues for their assistance and friendship over the past two years. I would particularly like to thank Dr. Andrea Todd and Pedro Collazo-Rodriguez for all the time that they have taken to teach me and discuss with me. I must also express my appreciation to Dr. Marcia T. Ruiz, of the McGill Cancer Centre, for the preparation of the CBP-enriched fraction, and Graeme Day, of the Centre for Theoretical and Computational Chemistry, University College London, for doing the molecular modeling mentioned in the Discussion. I must express a special word of thanks to my family, friends and The Knockouts for their constant support throughout this endeavor. Finally I wish to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for supporting me for the past two years, the Canderel Fund of the McGill Cancer Centre for supporting the presentation of this work at a conference and contributing to my stay in Munich, and the Cancer Research Society for funding this research.

7. REFERENCES

1. Sinden, R. R. (1994) *DNA Structure and Function.* Academic Press, San Diego.

2. Platt, J. R. (1955) Possible separation of inter-twisted nucleic acid chains by transfer-twist, *Proceedings of the National Academy of Science of the United States of America. 71*, 181-3.

3. Panayotatos, N. & Wells, R. D. (1981) Cruciform structures in supercoiled DNA, *Nature. 289*, 466-70.

4. Lilley, D. M. (1980) The inverted repeat as a recognizable structural feature in supercoiled DNA molecules, *Proceedings of the National Academy of Sciences of the United States of America. 77*, 6468-72.

5. Pearson, C. E., Zorbas, H., Price, G. B. & Zannis-Hadjopoulos, M. (1996) Inverted repeats, stem-loops, and cruciforms: significance for initiation of DNA replication, *Journal of Cellular Biochemistry. 63*, 1-22.

6. Frappier, L., Price, G. B., Martin, R. G. & Zannis-Hadjopoulos, M. (1987) Monoclonal antibodies to cruciform DNA structures, *Journal of Molecular Biology. 193*, 751-8.

7. Frappier, L., Price, G. B., Martin, R. G. & Zannis-Hadjopoulos, M. (1989) Characterization of the binding specificity of two anticruciform DNA monoclonal antibodies, *Journal of Biological Chemistry. 264*, 334-41.

8. Ward, G. K., McKenzie, R., Zannis-Hadjopoulos, M. & Price, G. B. (1990) The dynamic distribution and quantification of DNA cruciforms in eukaryotic nuclei, *Experimental Cell Research. 188*, 235-46.

9. Ward, G. K., Shihab-el-Deen, A., Zannis-Hadjopoulos, M. & Price, G. B. (1991) DNA cruciforms and the nuclear supporting structure, *Experimental Cell Research. 195*, 92-8.

10. Lilley, D. M. & Clegg, R. M. (1993) The structure of the four-way junction in DNA, *Annual Review of Biophysics & Biomolecular Structure. 22*, 299-328.

11. Benham, C. J. (1982) Stable cruciform formation at inverted repeat sequences in supercoiled DNA, *Biopolymers. 21*, 679-96.

12. Lilley, D. M. J. (1997) All change at Holliday junction, *Proceedings of the National Academy of Sciences of the United States of America. 94*, 9513-5.

13. Clegg, R. M., Murchie, A. I. & Lilley, D. M. (1994) The solution structure of the four-way DNA junction at low-salt conditions: a fluorescence resonance energy transfer analysis, *Biophysical Journal. 66*, 99-109.

14. Duckett, D. R., Murchie, A. I., Diekmann, S., von Kitzing, E., Kemper, B. & Lilley, D. M. (1988) The structure of the Holliday junction, and its resolution, *Cell. 55*, 79-89.

15. Hargreaves, D., Rice, D. W., Sedelnikova, S. E., Artymiuk, P. J., Lloyd, R. G. & Rafferty, J. B. (1998) Crystal structure of *E.coli* RuvA with bound DNA Holliday junction at 6 A resolution, *Nature Structural Biology. 5*, 441-6.

16. von Kitzing, E., Lilley, D. M. & Diekmann, S. (1990) The stereochemistry of a four-way DNA junction: a theoretical study, *Nucleic Acids Research. 18*, 2671-83.

17. Shlyakhtenko, L. S., Potaman, V. N., Sinden, R. R. & Lyubchenko, Y. L. (1998) Structure and dynamics of supercoil-stabilized DNA cruciforms, *Journal of Molecular Biology. 280*, 61-72.

18. Nobile, C. & Martin, R. G. (1986) Stable stem-loop and cruciform DNA structures: isolation of mutants with rearrangements of the palindromic sequence at the simian virus 40 replication origin, *Intervirology. 25*, 158-71.

19. Gough, G. W. & Lilley, D. M. (1985) DNA bending induced by cruciform formation, *Nature. 313*, 154-6.

20. Parsons, C. A. & West, S. C. (1990) Specificity of binding to four-way junctions in DNA by bacteriophage T7 endonuclease I, *Nucleic Acids Research. 18*, 4377-84.

21. Cozarelli, N. R. & Wang, J. C. E. (1990) *DNA Topology and its Biological Effects.* Cold Spring Harbour Press, New York.

22. Wang, J. C. & Liu, L. F. (1990) In *DNA Topology and its Biological Effects.* (Cozarelli, N. R. & Wang, J. C., Eds) pp. 321-40, Cold Spring Harbour Press, New York.

23. DePamphilis, M. L. (1993) Eukaryotic DNA replication: anatomy of an origin, *Annual Review of Biochemistry. 62*, 29-63.

24. Tremethick, D. J. & Molloy, P. L. (1986) High mobility group proteins 1 and 2 stimulate transcription in vitro by RNA polymerases II and III, *Journal of Biological Chemistry. 261*, 6986-92.

25. Tremethick, D. J. & Molloy, P. L. (1988) Effects of high mobility group proteins 1 and 2 on initiation and elongation of specific transcription by RNA polymerase II in vitro, *Nucleic Acids Research. 16*, 11107-23.

26. Watt, F. & Molloy, P. L. (1988) High mobility group proteins 1 and 2 stimulate binding of a specific transcription factor to the adenovirus major late promoter, *Nucleic Acids Research. 16*, 1471-86.

27. Singh, J. & Dixon, G. H. (1990) High mobility group proteins 1 and 2 function as general class II transcription factors, *Biochemistry. 29*, 6295-302.

28. Waga, S., Mizuno, S. & Yoshida, M. (1990) Chromosomal protein HMG1 removes the transcriptional block caused by the cruciform in supercoiled DNA, *Journal of Biological Chemistry. 265*, 19424-8.

29. Liu, L. F. & Wang, J. C. (1987) Supercoiling of the DNA template during transcription, *Proceedings of the National Academy of Sciences of the United States of America. 84*, 7024-7.

30. Soeller, W., Abarzua, P. & Marians, K. J. (1984) Mutational analysis of primosome assembly sites. II. Role of secondary structure in the formation of active sites, *Journal of Biological Chemistry. 259*, 14293-300.

31. Hiasa, H., Sakai, H., Komano, T. & Godson, G. N. (1990) Structural features of the priming signal recognized by primase: mutational analysis of the phage G4 origin of complementary DNA strand synthesis, *Nucleic Acids Research. 18*, 4825-31.

32. Gennaro, M. L., Iordanescu, S., Novick, R. P., Murray, R. W., Steck, T. R. & Khan, S. A. (1989) Functional organization of the plasmid pT181 replication origin, *Journal of Molecular Biology. 205*, 355-62.

33. Noirot, P., Bargonetti, J. & Novick, R. P. (1990) Initiation of rolling-circle replication in pT181 plasmid: initiator protein enhances cruciform extrusion at the origin, *Proceedings of the National Academy of Sciences of the United States of America. 87*, 8560-4.

34. Wong, T. W. & Clayton, D. A. (1985) In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis, *Cell. 42*, 951-8.

35. Wong, T. W. & Clayton, D. A. (1985) Isolation and characterization of a DNA primase from human mitochondria, *Journal of Biological Chemistry. 260*, 11530-5.

36. Zannis-Hadjopoulos, M., Persico, M. & Martin, R. G. (1981) The remarkable instability of replication loops provides a general method for the isolation of origins of DNA replication, *Cell. 27*, 155-63.

37. Kaufmann, G., Zannis-Hadjopoulos, M. & Martin, R. G. (1985) Cloning of nascent monkey DNA synthesized early in the cell cycle, *Molecular & Cellular Biology. 5*, 721-7.

38. Zannis-Hadjopoulos, M., Frappier, L., Khoury, M. & Price, G. B. (1988) Effect of anti-cruciform DNA monoclonal antibodies on DNA replication, *EMBO Journal. 7*, 1837-44.

39. McAlear, M., Ward, G. K., Mckenzie, R., Price, G. B. & Zannis-Hadjopoulos, M. (1989) Biphasic activation of DNA replication, *Molecular Genetics (Life Science Advances). 8*, 11-14.

40. Pearson, C. E., Ruiz, M. T., Price, G. B. & Zannis-Hadjopoulos, M. (1994) Cruciform DNA binding protein in HeLa cell extracts, *Biochemistry. 33*, 14185-96.

41. Pearson, C. E., Zannis-Hadjopoulos, M., Price, G. B. & Zorbas, H. (1995) A novel type of interaction between cruciform DNA and a cruciform binding protein from HeLa cells, *EMBO Journal. 14*, 1571-80.

42. Todd, A., Cossons, N., Aitken, A., Price, G. B. & Zannis-Hadjopoulos, M. (1998) Human cruciform binding protein belongs to the 14-3-3 family, *Biochemistry. 37*, 14317-25.

43. Bianchi, M. E., Beltrame, M. & Paonessa, G. (1989) Specific recognition of cruciform DNA by nuclear protein HMG1, *Science. 243*, 1056-9.

44. Wang, J., Goodman, H. M. & Zhang, H. (1999) An Arabidopsis 14-3-3 protein can act as a transcriptional activator in yeast, *FEBS Letters. 443*, 282-4.

45. De Vetten, N. C., Lu, G. & Feri, R. J. (1992) A maize protein associated with the G-box binding complex has homology to brain regulatory proteins, *Plant Cell. 4*, 1295-307.

46. Waterman, M., Stavridi, E. S., Waterman, J. & Halazonetis, T. D. (1998) Atm-dependent activation of p53 involves dephosphorylation and association with 14-3-3 proteins, *Nature Genetics. 19*, 175-178.

47. Aitken, A., Collinge, D. B., van Heusden, B. P., Isobe, T., Roseboom, P. H., Rosenfeld, G. & Soll, J. (1992) 14-3-3 proteins: a highly conserved, widespread family of eukaryotic proteins, *Trends in Biochemical Sciences. 17*, 498-501.

48. Aitken, A. (1996) 14-3-3 and its possible role in co-ordinating multiple signalling pathways, *Trends in Cell Biology. 6*, 341-347.

49. Aitken, A., Jones, D., Soneji, Y. & Howell, S. (1995) 14-3-3 proteins: biological function and domain structure., *Biochemical Society Transactions. 23*, 605-11.

50. Morrison, D. (1994) 14-3-3: modulators of signaling proteins?, *Science. 266*, 56-7.

51. Marais, R. & Marshall, C. (1995) 14-3-3 proteins: structure resolved, functions less clear, *Structure. 3*, 751-3.

52. Burbelo, P. D. & Hall, A. (1995) 14-3-3 proteins. Hot numbers in signal transduction, *Current Biology. 5*, 95-6.

53. Aitken, A. (1995) 14-3-3 proteins on the MAP, *Trends in Biochemical Sciences. 20*, 95-7.

54. Moore, B. W. & Perez, V. J. (1967) Specific acidic proteins of the nervous system, In *Physiological and Biochemical Aspects of Nervous Integration.* (Carlson, F. D., Ed pp. 343-59, Prentice Hall, Woods Hole, MA.

55. Ichimura, T., Sugano, H., Kuwano, R., Sunaya, T., Okuyama, T. & Isobe, T. (1991) Widespread distribution of the 14-3-3 protein in vertebrate brains and bovine tissues: correlation with the distributions of calcium-dependent protein kinases, *Journal of Neurochemistry. 56*, 1449-51.

56. Roth, D., Morgan, A. & Burgoyne, R. D. (1993) Identification of a key domain in annexin and 14-3-3 proteins that stimulate calcium-dependent exocytosis in permeabilized adrenal chromaffin cells, *FEBS Letters. 320*, 207-10.

57. Zha, J., Harada, H., Yang, E., Jockel, J. & Korsmeyer, S. J. (1996) Serine phosphorylation of death agonist BAD in response to survival factor results in binding to 14-3-3 not BCL-X(L), *Cell. 87*, 619-28.

58. Conklin, D. S., Galaktionov, K. & Beach, D. (1995) 14-3-3 proteins associate with cdc25 phosphatases, *Proceedings of the National Academy of Sciences of the United States of America. 92*, 7892-6.

59. Wang, W. & Shakes, D. C. (1996) Molecular evolution of the 14-3-3 protein family, *Journal of Molecular Evolution. 43*, 384-98.

60. Xiao, B., Smerdon, S. J., Jones, D. H., Dodson, G. G., Soneji, Y., Aitken, A. & Gamblin, S. J. (1995) Structure of a 14-3-3 protein and implications for coordination of multiple signalling pathways, *Nature. 376*, 188-91.

61. Liu, D., Bienkowska, J., Petosa, C., Collier, R. J., Fu, H. & Liddington, R. (1995) Crystal structure of the zeta isoform of the 14-3-3 protein, *Nature. 376*, 191-4.

62. Jones, D. H., Ley, S. & Aitken, A. (1995) Isoforms of 14-3-3 protein can form homo- and heterodimers in vivo and in vitro: implications for function as adapter proteins, *FEBS Letters. 368*, 55-8.

63. Muslin, A. J., Tanner, J. W., Allen, P. M. & Shaw, A. S. (1996) Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine, *Cell. 84*, 889-97.

64. Yaffe, M. B., Rittinger, K., Volinia, S., Caron, P. R., Aitken, A., Leffers, H., Gamblin, S. J., Smerdon, S. J. & Cantley, L. C. (1997) The structural basis for 14-3-3-phosphopeptide binding specificity, *Cell. 91*, 961-971.

65. Petosa, C., Masters, S. C., Bankston, L. A., Pohl, J., Wang, B., Fu, H. & Liddington, R. C. (1998) 14-3-3 zeta binds a phosphorylated Raf peptide and an unphosphorylated peptide via its conserved amphipathic groove, *Journal of Biological Chemistry. 273*, 16305-10.

66. Marsischky, G. T., Lee, S., Griffith, J. & Kolodner, R. D. (1999) Saccharomyces cerevisiae MSH2/6 complex interacts with Holliday junctions and facilitates their cleavage by phage resolution enzymes, *Journal of Biological Chemistry. 274*, 7200-6.

67. White, M. F., Giraud-Panis, M. J., Pohler, J. R. & Lilley, D. M. (1997) Recognition and manipulation of branched DNA structure by junction-resolving enzymes, *Journal of Molecular Biology. 269*, 647-64.

68. Suck, D. (1997) DNA recognition by structure-selective nucleases, *Biopolymers. 44*, 405-21.

69. Ariyoshi, M., Vassylyev, D. G., Iwasaki, H., Nakamura, H., Shinagawa, H. & Morikawa, K. (1994) Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from *E. coli*, *Cell. 78*, 1063-72.

70. Raaijmakers, H., Vix, O., Toro, I., Golz, S., Kemper, B. & Suck, D. (1999) X-ray structure of T4 endonuclease VII: a DNA junction resolvase with a novel fold and unusual domain-swapped dimer architecture, *EMBO Journal. 18*, 1447-58.

71. Kemper, B. & Janz, E. (1976) Function of gene 49 of bacteriophage T4. I. Isolation and biochemical characterization of very fast-sedimenting DNA, *Journal of Virology. 18*, 992-9.

72. Kemper, B. & Brown, D. T. (1976) Function of gene 49 of bacteriophage T4. II. Analysis of intracellular development and the structure of very fast-sedimenting DNA, *Journal of Virology. 18*, 1000-15.

73. Sadowski, P. D. (1971) Bacteriophage T7 endonuclease. I. Properties of the enzyme purified from T7 phage-infected Escherichia coli B, *Journal of Biological Chemistry. 246*, 209-16.

74. Parsons, C. A., Tsaneva, I., Lloyd, R. G. & West, S. C. (1992) Interaction of *Escherichia coli* RuvA and RuvB proteins with synthetic Holliday junctions, *Proceedings of the National Academy of Sciences of the United States of America. 89*, 5452-6.

75. Roe, S. M., Barlow, T., Brown, T., Oram, M., Keeley, A., Tsaneva, I. R. & Pearl, L. H. (1998) Crystal structure of an octameric RuvA-Holliday junction complex, *Molecular Cell. 2*, 361-72.

76. Bennett, R. J. & West, S. C. (1995) Structural analysis of the RuvC-Holliday junction complex reveals an unfolded junction, *Journal of Molecular Biology. 252*, 213-26.

77. Parsons, C. A., Kemper, B. & West, S. C. (1990) Interaction of a four-way junction in DNA with T4 endonuclease VII, *Journal of Biological Chemistry. 265*, 9285-9.

78. Bhattacharyya, A., Murchie, A. I., von Kitzing, E., Diekmann, S., Kemper, B. & Lilley, D. M. (1991) Model for the interaction of DNA junctions and resolving enzymes, *Journal of Molecular Biology. 221*, 1191-207.

79. Ner, S. S., Travers, A. A. & Churchill, M. E. (1994) Harnessing the writhe: a role for DNA chaperones in nucleoprotein-complex formation, *Trends in Biochemical Sciences. 19*, 185-7.

80. Bustin, M. & Reeves, R. (1996) High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function, *Progress in Nucleic Acid Research & Molecular Biology. 54*, 35-100.

81. Pohler, J. R. G., Norman, D. G., Bramham, J., Bianchi, M. E. & Lilley, D. M. (1998) HMG box proteins bind to four-way DNA junctions in their open conformation, *EMBO Journal. 17*, 817-26.

82. Pil, P. M. & Lippard, S. J. (1992) Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin, *Science. 256*, 234-7.

83. Bianchi, M. E., Falciola, L., Ferrari, S. & Lilley, D. M. (1992) The DNA binding site of HMG1 protein is composed of two similar segments (HMG boxes), both of which have counterparts in other eukaryotic regulatory proteins, *EMBO Journal. 11*, 1055-63.

84. Teo, S. H., Grasser, K. D., Hardman, C. H., Broadhurst, R. W., Laue, E. D. & Thomas, J. O. (1995) Two mutations in the HMG-box with very different structural consequences provide insights into the nature of binding to four-way junction DNA, *EMBO Journal. 14*, 3844-53.

85. Teo, S. H., Grasser, K. D. & Thomas, J. O. (1995) Differences in the DNA-binding properties of the HMG-box domains of HMG1 and the sex-determining factor SRY, *European Journal of Biochemistry. 230*, 943-50.

86. Hill, D. A., Pedulla, M. L. & Reeves, R. (1999) Directional binding of HMG-I(Y) on four-way junction DNA and the molecular basis for competitive binding with HMG-1 and histone H1, *Nucleic Acids Research. 27*, 2135-44.

87. Dutta, S., Gerhold, D. L., Rice, M., Germann, M. & Kmiec, E. B. (1997) The cloning and overexpression of a cruciform binding protein from Ustilago maydis, *Biochimica et Biophysica Acta. 1352*, 258-66.

88. Lee, S., Cavallo, L. & Griffith, J. (1997) Human p53 binds Holliday junctions strongly and facilitates their cleavage, *Journal of Biological Chemistry. 272*, 7532-9.

89. Revzin, A. E. (1993) *Footprinting of Nucleic Acid-Protein Complexes.* Academic Press, Inc., San Diego.

90. Wissmann, A. & Hillen, W. (1991) DNA contacts probed by modification protection and interference studies, *Methods in Enzymology. 208*, 365-79.

91. Schmitz, A. & Galas, D. J. (1979) The interaction of RNA polymerase and lac repressor with the lac control region, *Nucleic Acids Research. 6*, 111-37.

92. Brenowitz, M., Senear, D. F., Shea, M. A. & Ackers, G. K. (1986) Quantitative DNase footprint titration: a method for studying protein- DNA interactions, *Methods in Enzymology. 130*, 132-81.

93. Brunelle, A. & Schleif, R. F. (1987) Missing contact probing of DNA-protein interactions, *Proceedings of the National Academy of Sciences of the United States of America. 84*, 6673-6.

94. Hayes, J. J. & Tullius, T. D. (1989) The missing nucleoside experiment: A new technique to study recognition of DNA by protein, *Biochemistry. 28*, 9521-27.

95. Buning, H., Baeuerle, P. A. & Zorbas, H. (1995) A new interference footprinting method for analysing simultaneously protein contacts to phosphate and guanine residues on DNA, *Nucleic Acids Research. 23*, 1443-4.

96. Hsieh, M. & Brenowitz, M. (1996) Quantitative kinetics footprinting of protein-DNA association reactions, *Methods in Enzymology. 274*, 478-92.

97. Hochschild, A. (1991) Detecting cooperative protein-DNA interactions and DNA loop formation by footprinting, *Methods in Enzymology. 208*, 343-61.

98. Hayes, J. J., Kam, L. & Tullius, T. D. (1990) Footprinting protein-DNA complexes with gamma-rays, *Methods in Enzymology. 186*, 545-9.

99. Sclavi, B., Woodson, S., Sullivan, M., Chance, M. & Brenowitz, M. (1998) Following the folding of RNA with time-resolved synchrotron X-ray footprinting, *Methods in Enzymology. 295*, 379-402.

100. Tsodikov, O. V., Craig, M. L., Saecker, R. M. & Record, M. T., Jr. (1998) Quantitative analysis of multiple-hit footprinting studies to characterize DNA conformational changes in protein-DNA complexes: application to DNA opening by Esigma70 RNA polymerase, *Journal of Molecular Biology. 283*, 757-69.

101. Maxam, A. M. & Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages, *Methods in Enzymology. 65*, 499-560.

102. Galas, D. J. & Schmitz, A. (1978) DNAse footprinting: a simple method for the detection of protein-DNA binding specificity, *Nucleic Acids Res. 5*, 3157-70.

103. Sigman, D. S., Graham, D. R., D'Aurora, V. & Stern, A. M. (1979) Oxygen-dependent cleavage of DNA by the 1,10-phenanthroline-cuprous complex. Inhibition of Escherichia coli DNA polymerase I., *Journal of Biological Chemistry. 254*, 12269-272.

104. Tullius, T. D., Dombroski, B. A., Churchill, M. E. & Kam, L. (1987) Hydroxyl radical footprinting: a high-resolution method for mapping protein-DNA contacts, *Methods in Enzymology. 155*, 537-58.

105. Hayatsu, H. & Ukita, T. (1967) The selective degradation of pyrimidines in nucleic acids by permanganate oxidation, *Biochemical and Biophysical Research Communications. 29*, 556-561.

106. Becker, M. M. & Wang, J. C. (1984) Use of light for footprinting DNA in vivo, *Nature. 309*, 682-7.

107. Riley, D. & Weintraub, H. (1978) Nucleosomal DNA is digested to repeats of 10 bases by exonuclease III, *Cell. 13*, 281-93.

108. Bailly, C. & Waring, M. J. (1995) Comparison of different footprinting methodologies for detecting binding sites for a small ligand on DNA, *Journal of Biomolecular Structure and Dynamics. 12*, 869-98.

109. Ephrussi, A., Church, G. M., Tonegawa, S. & Gilbert, W. (1985) B lineage-specific interactions of an immunoglobulin enhancer with cellular factors in vivo, *Science. 227*, 134-40.

110. Saluz, H. P. & Jost, J. P. (1993) Approaches to characterize protein-DNA interactions in vivo, *Critical Reviews in Eukaryotic Gene Expression. 3*, 1-29.

111. Becker, P. B., Weih, F. & Schutz, G. (1993) Footprinting of DNA-binding proteins in intact cells, *Methods in Enzymology. 218*, 568-87.

112. Santocanale, C. & Diffley, J. F. (1997) Genomic footprinting of budding yeast replication origins during the cell cycle, *Methods in Enzymology. 283*, 377-90.

113. Hornstra, I. K. & Yang, T. P. (1993) In vivo footprinting and genomic sequencing by ligation-mediated PCR, *Analytical Biochemistry. 213*, 179-93.

114. Gopaul, D. N., Guo, F. & Van Duyne, G. D. (1998) Structure of the Holliday junction intermediate in Cre-loxP site-specific recombination, *EMBO Journal. 17*, 4175-87.

115. Steinmetzer, K., Zannis-Hadjopoulos, M. & Price, G. B. (1995) Anti-cruciform monoclonal antibody and cruciform DNA interaction, *Journal of Molecular Biology. 254*, 29-37.

116. Lawley, P. D. & Brookes, P. (1963) Further studies on the alkylation of nucleic acids and their constituent nucleotides, *Biochemical Journal. 89*, 127-38.

117. Yang, J. & Carey, J. (1995) Footprint phenotypes: structural models of DNA-binding proteins from chemical modification analysis of DNA, *Methods in Enzymology. 259*, 452-68.

118. Ofverstedt, L. G., Hammarstrom, K., Balgobin, N., Hjerten, S., Pettersson, U. & Chattopadhyaya, J. (1984) Rapid and quantitative recovery of DNA fragments from gels by displacement electrophoresis (isotachophoresis), *Biochimica et Biophysica Acta. 782*, 120-6.

119. Pearson, C. E., Frappier, L. & Zannis-Hadjopoulos, M. (1991) Plasmids bearing mammalian DNA-replication origin-enriched (ors) fragments initiate semiconservative replication in a cell-free system, *Biochimica et Biophysica Acta. 1090*, 156-66.

120. Elborough, K. M. & West, S. C. (1988) Specific binding of cruciform DNA structures by a protein from human extracts, *Nucleic Acids Research. 16*, 3603-16.

121. Zorbas, H., Rogge, L., Meisterernst, M. & Winnacker, E. L. (1989) Hydroxyl radical footprints reveal novel structural features around the NF I binding site in adenovirus DNA, *Nucleic Acids Research. 17*, 7735-48.

122. Toker, A., Sellers, L. A., Amess, B., Patel, Y., Harris, A. & Aitken, A. (1992) Multiple isoforms of a protein kinase C inhibitor (KCIP-1/14-3-3) from sheep brain. Amino acid sequence of phosphorylated forms, *European Journal of Biochemistry. 206*, 453-61.

123. Svaren, J., Inagami, S., Lovegren, E. & Chalkley, R. (1987) DNA denatures upon drying after ethanol precipitation, *Nucleic Acids Research. 15*, 8739-54.

124. Wingfield, P. T. (1995) Use of protein folding reagents, in *Current Protocols in Protein Science*. (Coligan, J. E., Dunn, B. M., Ploegh, H. L., Speicher, D. W. & Wingfield, P. T., eds) pp. A.3A.1-A.3A.4, John Wiley and Sons, Inc., New York.

125. Struhl, K. (1998) Isolation of proteins for microsequence analysis, in *Current Protocols in Molecular Biology*. (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds) pp. 10.19.1-10.19.12, Jonh Wiley and Sons Inc., New York.

126. Prigodich, R. V. & Martin, C. T. (1990) Reaction of single-stranded DNA with hydroxyl radical generated by iron(II)-ethylenediaminetetraacetic acid, *Biochemistry. 29*, 8017-9.

127. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J., J.A., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong,

M. W., Andres, J. L., Gonzalez, C., Head-Gordon, M., Replogle, E. S. & Pople, J. A. (1998) Gaussian 98 in , Gaussian, Inc., Pittsburgh.

128. Beranek, D. T. (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents, *Mutation Research. 231*, 11-30.

129. Kirkegaard, K., Buc, H., Spassky, A. & Wang, J. C. (1983) Mapping of single-stranded regions in duplex DNA at the sequence level: single-strand-specific cytosine methylation in RNA polymerase-promoter complexes, *Proceedings of the National Academy of Sciences of the United States of America. 80*, 2544-8.

130. Peattie, D. A. & Gilbert, W. (1980) Chemical probes for higher-order structure in RNA, *Proceedings of the National Academy of Sciences of the United States of America. 77*, 4679-82.

131. Luetke, K. H. & Sadowski, P. D. (1998) Determinants of the position of a Flp-induced DNA bend, *Nucleic Acids Research. 26*, 1401-7.

132. Luetke, K. H., Zhao, B. P. & Sadowski, P. D. (1997) Asymmetry in Flp-mediated cleavage, *Nucleic Acids Research. 25*, 4240-9.

133. Kim, S. W., Ahn, I. M. & Larsen, P. R. (1996) In vivo genomic footprinting of thyroid hormone-responsive genes in pituitary tumor cell lines, *Molecular and Cellular Biology. 16*, 4465-77.

134. Siebenlist, U., Simpson, R. B. & Gilbert, W. (1980) *E. coli* RNA polymerase interacts homologously with two different promoters, *Cell. 20*, 269-81.

135. Baylin, S. B. (1997) Tying it all together: epigenetics, genetics, cell cycle, and cancer, *Science. 277*, 1948-9.

136. Laird, P. W. (1997) Oncogenic mechanisms mediated by DNA methylation, *Molecular Medicine Today. 3*, 223-9.

137. Counts, J. L. & Goodman, J. I. (1995) Alterations in DNA methylation may play a variety of roles in carcinogenesis, *Cell. 83*, 13-5.

138. Wang, H., Nicholson, P. R. & Stillman, D. J. (1990) Identification of a Saccharomyces cerevisiae DNA-binding protein involved in transcriptional regulation, *Molecular & Cellular Biology. 10*, 1743-53.

139. Davis, D. R. & Stillman, D. J. (1997) Altered structure of the DNA duplex recognized by yeast transcription factor Reb1p, *Nucleic Acids Research. 25*, 668-74.

140. Bouziane, M., Miao, F., Ye, N., Holmquist, G., Chyzak, G. & O'Connor, T. R. (1998) Repair of DNA alkylation damage, *Acta Biochimica Polonica. 45*, 191-202.

141. Engelward, B. P., Dreslin, A., Christensen, J., Huszar, D., Kurahara, C. & Samson, L. (1996) Repair-deficient 3-methyladenine DNA glycosylase homozygous mutant mouse cells have increased sensitivity to alkylation-induced chromosome damage and cell killing, *EMBO Journal. 15*, 945-52.

142. Aletta, J. M., Cimato, T. R. & Ettinger, M. J. (1998) Protein methylation: a signal event in post-translational modification, *Trends in Biochemical Sciences. 23*, 89-91.

143. Clarke, S. (1993) Protein methylation, *Current Opinion in Cell Biology. 5*, 977-83.

144. Kim, S., G.H., P. & Paik, W. K. (1998) Recent advances in protein methylation: Enzymatic methylation of nucleic acid binding proteins, *Amino Acids. 15*, 291-306.

145. Parish, C. A. & Rando, R. R. (1996) Isoprenylation/methylation of proteins enhances membrane association by a hydrophobic mechanism, *Biochemistry. 35*, 8473-7.

146. Toker, A., Ellis, C. A., Sellers, L. A. & Aitken, A. (1990) Protein kinase C inhibitor proteins. Purification from sheep brain and sequence similarity to lipocortins and 14-3-3 protein, *European Journal of Biochemistry. 191*, 421-9.

147. Gething, M. J. (1997) Protein folding. The difference with prokaryotes, *Nature. 388*, 329, 331.

148. Netzer, W. J. & Hartl, F. U. (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes, *Nature. 388*, 343-9.

149. Sigman, D. S., Kuwabara, M. D., Chen, C. H. & Bruice, T. W. (1991) Nuclease activity of 1,10-phenanthroline-copper in study of protein-DNA interactions, *Methods in Enzymology. 208*, 414-33.

150. Kuwabara, M. D. & Sigman, D. S. (1987) Footprinting DNA-protein complexes in situ following gel retardation assays using 1,10-phenanthroline-copper ion: Escherichia coli RNA polymerase-lac promoter complexes, *Biochemistry. 26*, 7234-8.

151. Guo, F., Gopaul, D. N. & van Duyne, G. D. (1997) Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse, *Nature. 389*, 40-6.