Works published in the Serie Universitaria cover the following specialities: Architecture and Town Planning; Fine Arts; Biology; Agricultural Sciences; Social Sciences; Journalism; Law; Economics; Phylosophy; Phisics; Geology; History; Engineering; Literature and Philology; Mathematics; Medicine, Pharmacy and Veterinary; Music; Chemistry; Theology.

Each group of specialities has its own colour on every cover.

The course summarized in this publication was given by its author on the 20th thru the 22nd of November, 1989 at the Universidad Internacional Menéndez Pelayo (UIMP) in Cuenca.

There is a limited edition of 450 copies of this volume, available free of charge.

Fundación Juan March

248 Course on DNA - Protein Interaction / Miguel Beato

SERIE UNIVERSITARIA

Fundación Juan March

Course on DNA - Protein Interaction

Miguel Beato

Fundación Juan March

Serie Universitaria

248

Course on DNA - Protein Interaction

Methodological Approach Illustrated with Steroid Hormone Receptors

Miguel Beato
Institut für Molekularbiologie und Tumorforschung
Philipps Universität
E.-Mannkopff-Str.2
D-3550 Marburg
F.R.G.

Fundación Juan March
Castelló, 77. Teléf. 435 42 40
28006-Madrid

The course summarized in this publication was given by its author on the 20th thru the 22nd of November, 1989 at the Universidad Internacional Menéndez Pelayo (UIMP) in Cuenca.

Depósito legal: M-6.871/1990
ISBN: 84-7075-400-9
Impresión: Ediciones Peninsular. Tomelloso, 27. 28026 Madrid

CONTENT

PREFACE

The idea to write this booklet arose as I was planing a course on "Protein-DNA Interactions" to be given in Cuenca, Spain, in the fall of 1989. As an orientation for the students, a short outline of the themes to be treated in the course was required. In the process of writing these abstracts I thought that it will not be much more work to write a few chapters containing a more detailed description of the topics to be discussed. Now, after having expended three months with this exercise, I realized how mistaken I was.

The focus of the course, and therefore of the book, is on methodic aspects, and specially on considerations about the utility or applicability of individual methods. It is thought to serve as an orientation for students without much experience in molecular biology, but interested in gene regulation. The discussion of the advantages and limitations of the individual techniques should help in deciding suitable strategies to tackle different experimental problems. Once the more adequate methods have been selected, the student will have to consult detailed protocols, of which a great variety is now available as laboratory guides.

The treatment of the subject is highly personal, and no attempt has been made to give a complete description of all available methods. In particular, the discussion of the individual topics is influenced by my own experience with the study of gene regulation by steroid hormones and their receptors. Most of the examples that illustrate the use of various techniques have been extracted from this field, in which I have been working for over twenty years. I am aware that the steroid hormone receptors represent a special class of DNA-binding regulatory proteins, that may not be representative of the behavior of other DNA-binding proteins or transcription factors. However, because the family of nuclear receptors is one of the best characterized and most intensively studied group of DNA regulatory proteins, I think that we can learn some basic rules on gene regulation by describing

their behavior in some detail. Of course, when other regulatory proteins are to be studied, one has to keep in mind the possibility of basic differences in the underlaying mechanism of DNA recognition and transcription control.

As always when writing this type of books, the author has only contributed a part, possibly a minor part, to it. The experiments that serve as the skeleton of the book were performed during the past ten years by numerous collaborators, many of which are no longer with me. It will be impossible to thank all of them here, and I will just name only a few to which I am particularly grateful. During his thesis work and as a young postdoctoral fellow, Claus Scheidereit, who is now in the Max Plank Institut for Molecular Genetics in West Berlin, pioneered the development of techniques for analyzing binding of steroid receptors to DNA, and also helped me writing one of the first methodic chapters in this field. More recently Georges Chalepakis and Mathias Truss, both chemistry graduate students, have developed and adapted numerous techniques for detailed analysis of DNA-protein interaction, and Benjamin Piña, a postdoctoral fellow from Mallorca, has introduced the nucleosome reconstitution techniques and demonstrated how the wrapping of the double helix around the histone octamer determines its accessibility for binding of regulatory proteins. During the preparation of the manuscript I had numerous discussion with Emily P. Slater and Robert Haché, both of which also made very stimulating suggestions, and improved the language.

Finally, I want to thank Andrés González, from the Fundación Juan March, who not only supported the initial idea of the course and the book, but also has been a continuing help during the editorial process.

Marburg, October 12, 1989.

INTRODUCTION

Molecular complementary, or more precisely surface complementarity between molecules, has long been recognized as the basic mechanism of biological information, initially in enzymology and later in the receptor field. Precise complementarity between substrates and the active sites of enzymes, or between ligands and the binding sites of their receptors, account for the specificity of biological reactions and for the selectivity of signal transduction. Similar mechanisms operate in the process of transmission, conversion, recognition and utilization of genetic information. This information is stored in the precise nucleotide sequence of nucleic acids and is transferred to the progeny of the organism by a replication process that involves surface complementarity between two nucleic acid molecules. Utilization of this information in the process of transcription and translation is also mediated by surface complementarity between nucleic acid molecules, base paired by hydrogen bonding. The selection of the particular section of the genetic information that is to be expressed in a particular framework of time and space is determined again by surface complementarity, but in this case, proteins are the reading heads that scan nucleic acid sequences.

This book deals with the methodology that has been developed in trying to understand how proteins are able to identify sequence information stored in nucleic acids, how they discriminate among apparently very similar sequences, and how their interactions with the nucleic acids influences the expression of the information they contain. Initial insight into these aspects of biology was gained in the study of prokaryotic systems, in particular, bacteriophages, but most of the examples discussed in this book originate from the more complex genetic systems of eukaryotic organisms. The overwhelming amount of experimental data derived from these studies, clearly shows that, in addition to the structural information required for making proteins, nucleic acids contain an enormous amount of regulatory information, that determines the space-temporal pattern of expression of the structural information. It is the precise program of regulatory

information that, when decoded by DNA-binding regulatory proteins, converts the linear array of genetic messages contained into the DNA of a fertilized egg into the three dimensional structure of the organism. As the regulatory proteins are also encoded in the structural information of the DNA, the whole process of organismic development can be envisaged as a self controlling morphogenetic cascade, **the information cycle** (Figure 1). The structural information is encoded in the DNA structure in form of the genetic code that includes signals for initiation and termination of transcription and translation, as well as the information required to process the RNA. The first period of molecular biology has been concerned with hte elucidation of these mechanisms. The question that is now discussed is whether such thing as a **regulatory code** exists, that could help in deciphering the regulatory information stored in the nucleotide sequence of the DNA. I will deal briefly with this theme at the end of the book.

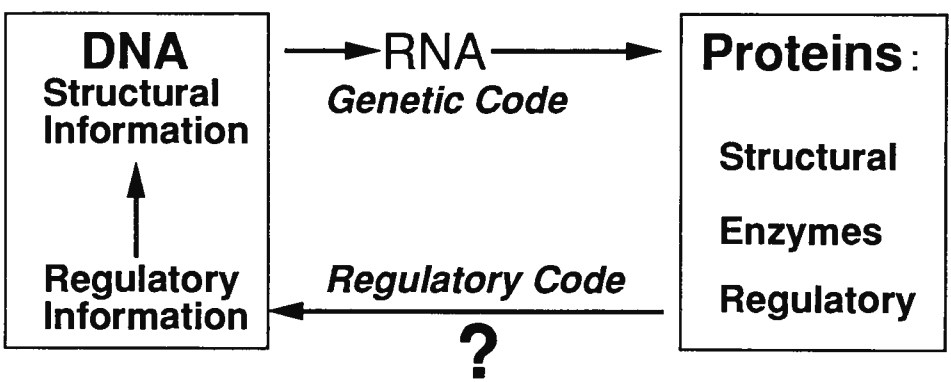


Figure 1. The Information Cycle.

Schematic representation of the information cycle, illustrating the connection between structural and regulatory information.

In the following I will focus on the interaction between eukaryotic regulatory proteins and their recognition sequences on DNA, illustrated by the steroid hormone/thyroid hormone receptor family. The emphasis will be on methodological aspects, and more precisely on the potential of the different methods and the type of information they provide rather than on experimental recipes. Undoubtedly the selection of the system will bias the strategical considerations, and this books does not attempt to cover exhaustively the addressed topic. In particular physico-chemical methods of analysis that yield information of the highest possible resolution, such as X-ray analysis and NMR-technologies, will not be treated in this book. Only molecular biology techniques will be discussed and the book represents a personal view of the subject, determined by the experimental experience of the author. Nevertheless, this discussion could be useful as an example of the type of problems encountered when dealing with DNA-binding regulatory proteins, and as an illustration of the possible strategies to approach their solution.

1. Chapter

DNA-BINDING REGULATORY PROTEINS: AN OVERVIEW

Since the original discovery of the *lac* repressor and its role in gene regulation, an enormous number of DNA-binding regulatory proteins have been described and the DNA sequences to which they bind have been identified and characterized, both in prokaryotic and eukaryotic organisms. Several DNA-binding proteins from bacteria and bacteriophages have been crystallized and their structure is known in great detail. In some cases cocrystalls of the protein bound to its target DNA sequences have been obtained that yield information of the three dimensional structure of the specific complexes of protein and DNA. From a comparison of the structures found, some preferred mechanisms by which regulatory proteins recognized information stored in DNA are becoming clear (Struhl 1989). In the following I will briefly survey these general concepts. Not included in this considerations are DNA restriction endonucleases, as they represent enzymes rather than regulatory proteins.

The question can be considered from two different points of view, the protein structures that are able to recognize DNA information, and the DNA sequences itself. Finally, I will briefly addressed questions related to the statistics and thermodynamic aspects of DNA sequence recognition.

1. The proteins structural motifs.

Several prokaryotic DNA-binding proteins, including *cro* and the cAMP receptor protein of *E. coli* as well as the *lac* repressor the *lambda* repressor, the bacteriophage 434 repressor, and the *trp* repressor of *E. coli*, exhibit a characteristic **Á-helix turn Á-helix** motif in their DNA-binding domains (for a review see Brennan and Matthews 1989). This helix-turn-helix structure (Fig. 1.1) was originally discovered during the analysis of the X-ray structure of the lambda phage cro repressor protein (Anderson et al 1981) and of the E. coli catabolite gene activator protein-cAMP complex (McKay and Steitz 1981). It is known that one of the Á-helixes has

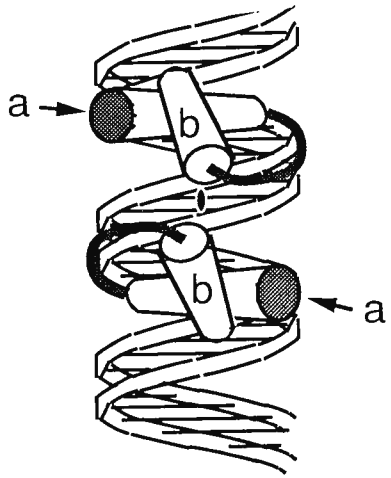


Figure 1.1. Helix-Turn-Helix Motif.

Schematic representation of the two Á-helices that form the DNA-binding motif of several prokaryotic and eukaryotic regulatory proteins. The a-helix has an amphipathic structure. The hydrophylic side of this helix contacts the mayor groove of the DNA double helix, and its hydrophobic side interacts with the b- helix that contributes to its correct positioning. The dyad symmetry axis is indicated.

an amphipathic nature, with the hydrophylic region contacting the major groove of the target DNA sequence, whereas the hydrophobic region is held in position by an interaction with the second Á-helix. A similar structure, has been identified in the homeo-domain of the homeotic *Drosophila* gene *Antennapedia* (Otting et al 1988), and this structure is conserved in other eukaryotic DNA-binding regulatory proteins, including the homeotic genes of *Drosophila*, *Xenopus* and mice, and certain transcription factors (Struhl 1989).

Another typical DNA-binding structure, the so-called **"zinc finger"**, was originally discovered in the *Xenopus* transcription factor TFIIIA (Miller et al 1985) and has later been found in a great variety of other proteins belonging to very different functional classes (Klug and Rhodes 1987; Berg 1988). In this motif, a zinc ion is tetrahedrally coordinated to two pairs of cysteins or to one pair of cysteins and one pair of histidines (Fig. 1.2). Here again an amphipatic Á-helix seems to be the structure that enters the major groove of the DNA target sequence, and the function of the metal ion is to stabilize this structure. How exactly the interaction with the DNA takes place is unknown, but some hypothetical models will be discussed in context of the steroid hormone receptors.

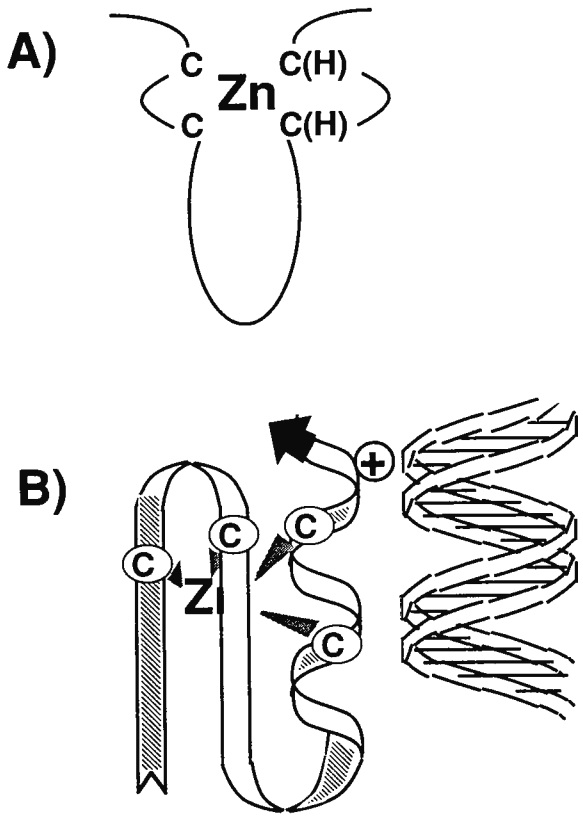


Figure 1.2. The Zinc-Finger Motif.
A) Schematic representation of the general structure of the zinc finger, with the Zn ion tetrahedrally coordinated to either four cysteins or to two cysteines and two histidines.
B) Hypothetical model of the secondary structure of the zinc finger, with the first two cysteins participating in a ß-sheet structure and the second pair of cysteins forming part of an Á-helix. This Á-helix has an amphipathic nature and will contact the major groove of the DNA through its hydrophylic side. The hydrophobic site will contact the ß-sheet and stabilize the structure.

During the last two years a third protein structural motif has been detected in a variety of DNA binding regulatory proteins and oncogenes the so-called **"leucine zipper"** (Fig. 1.3). This structure was originally discovered in the transcription factor C/EBP, and is composed of a series of repeated leucines with a characteristic 7 amino acid spacing (Landschulz, Johnson and McKnight 1988). Although this structure does not directly bind to DNA it appears to stabilized a dimeric structure of the corresponding proteins that keeps an adjacent basic domain in the conformation needed for DNA binding. The exact mechanism of dimerization is still a matter of debate though the original idea, namely the interdigitation of the hydrophobic leucine side chains in the two complementary Á-helixes, has the attraction of simplicity. The interaction can take place among identical monomers or, as in the case of the oncogenes *fos* and *jun*, among two polypeptides encoded in different genes (Curran and Franza 1988). Possibly, variants of the leucine zipper motif, in which other, hydrophobic or ionic, interactions between Á-helices take place could stabilize dimerization.

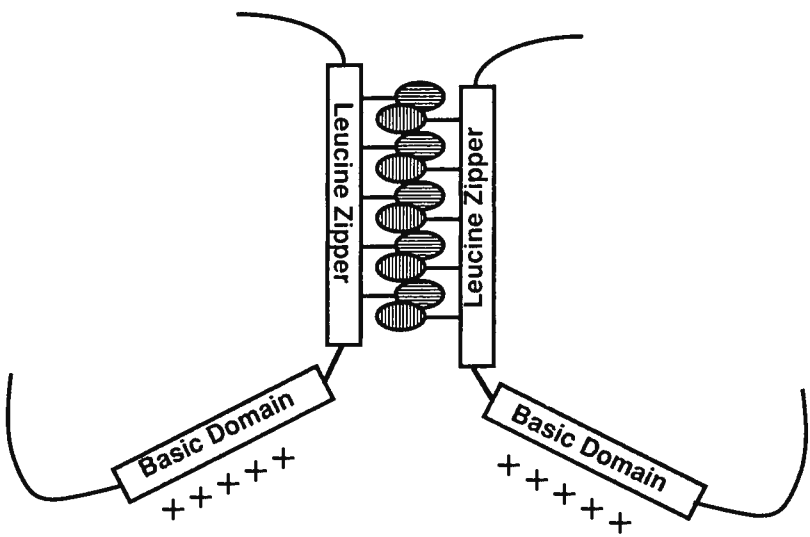


Figure 1.3. The Leucine Zipper Motif.
Schematic representation of an hypothetical leucine zipper with five intertwined leucine side chains on each polypeptide. The proximity of the leucine zipper to the basic domain that interacts with DNA is indicated.

Very recently, another protein motif, the helix-loop-helix motif, has been identified as a common structure of an apparently heterogeneous group of DNA binding regulatory proteins (Murre, McCaw and Baltimore 1989). This group of proteins includes the Myc/MyoD family and myogenin that are involved in myoblast differentiation, E12 and E47 that bind to the immunoglobulin enhancer, and *achaete-scute* and *daughterless*, two genes involved in *Drosophila* neurogenesis. In addition to DNA binding, the helix-loop-helix motif also confers to these proteins the ability to form homo or heterodimers in a highly selective way (Murre et al 1989). As some of these proteins also exhibit a leucine zipper motif, interfamily dimerization is possible and seems a convenient and plausible strategy two generate the combinatorial complexity needed to account for the immense variety of tissue-specific genetic programs (Abel and Maniatis 1989).

In addition to these four frequent motifs, there are other DNA-binding proteins which DNA binding domain does not exhibit any obvious conserved structure. To this class belong several transcription factors and oncogenes, such as nuclear factor I, GAL4, etc. However, as a general rule, one can say that Á-helixes seem to be appropriate structures to interact specifically with the DNA double helix.

A common feature of many DNA-binding regulatory proteins is that they are composed of two, and occasionally four, equal subunits, and that they exhibit a dyad symmetry axis. As we will see below, this quaternary structure reflects the fact that the DNA sequences recognized by most regulatory proteins are palindromic in nature and also contain a dyad symmetry axis. This could be a simple way to improve sequence discrimination.

-The DNA regulatory elements

A large number of DNA sequences has been identified as specific binding sites for DNA-binding regulatory proteins. With the exception of a few transcription factors, such as the SP1 (McKnight and Tjian 1986), most of these sites exhibit a more or

less perfect palindromic structure with a dyad symmetry axis. As mentioned above, this palindromic structure is probably recognized by a dimer of the DNA binding protein. In most cases studied the number of conserved base pairs on each half of the palindrome is small, between three and six, and, although the spacing can vary considerably, the centers of the half palindromes are often separated by a turn of the double helix (approximately 10 base pairs). This suggests that the two monomers of the DNA binding proteins interact with the same side the double helix.

Close inspection of the binding sites for different proteins and examination of the degree of conservation of the individual positions, shows that most palindromes are functionally asymmetric, in that one half of the palindrome is usually more stringently conserved than the other. This asymmetry suggests that binding of the proteins takes places in two phases, with the more conserved half of the palindrome being recognized first.

Very often the imperfect palindromes are repeated either in clusters or, less frequently, as separated copies interspersed with binding sites for other regulatory proteins. In some cases, as in the SV40 enhancer, the particular array of individual binding sequences for different factors is rather complex and some sites may even overlap. This can result in cooperative binding of different factors, but can also lead to competition among different factors for binding to overlapping sequences. I will mention some examples of this behavior in connection with steroid hormone receptors.

-**DNA sequence recognition**.

In principle there are two ways, not mutually exclusive, how proteins can recognize information stored on the DNA double helix. Either they can directly "read" the linear sequence of base pairs along the helix, or they can recognize structural peculiarities of the double helix, themselves determined by the nucleotide sequence. In most cases, probably, both mechanism are used simultaneously with the contribution of each of then varying from system to system.

As initially noticed by Seeman et al (1976), the nucleotide sequence of a double helix can be directly read from the major groove without disturbing its structure. Each base pair can be distinguish in terms of the array of hydrogen bond donor and acceptor sites (Figure 1.4). Each base pair can undergo three hydrogen bonds with side chains of amino acids, and all that is needed to distinguish among the four different base pairs is a minimum of two hydrogen bonds per base pair. In addition the 5'-methyl groups of thymines are a salient feature that proteins can recognize in the major groove of the double helix. Their asymmetric position makes then a diagnostic feature that can be use to distinguish a A/T from a T/A base pair. This distinction is not possible through the minor groove. The simplified "stick-figure" representation (Woodbury and von Hippel 1981) shown in Figure 1.4 can be easily used to schematically represent a particular sequence of base pairs in terms of its potential to interact with the amino acid side chains of proteins (Fig. 1.5.).

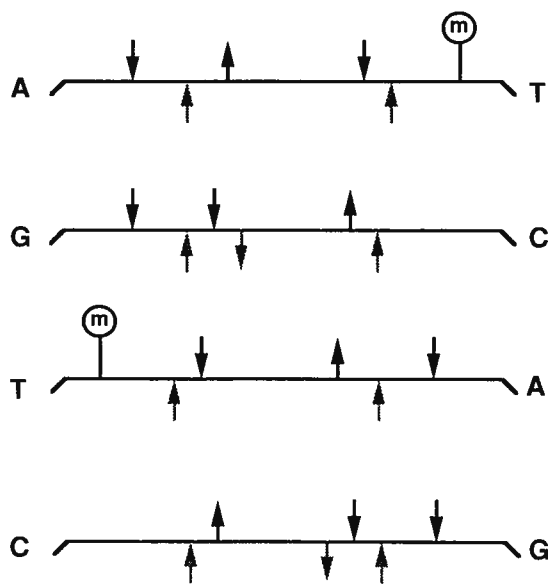


Figure 1.4. Schematic Representation of Individual Base Pairs
Simplified stick representation of the four base pairs. Hydrogen bond donor positions are indicated by an arrow pointing out of the plane of the base pair; hydrogen bond acceptor positions are indicated by arrows pointing towards the plane of the base pair. The 5'-methyl group of thymine is indicated. Major groove pointing upwards.

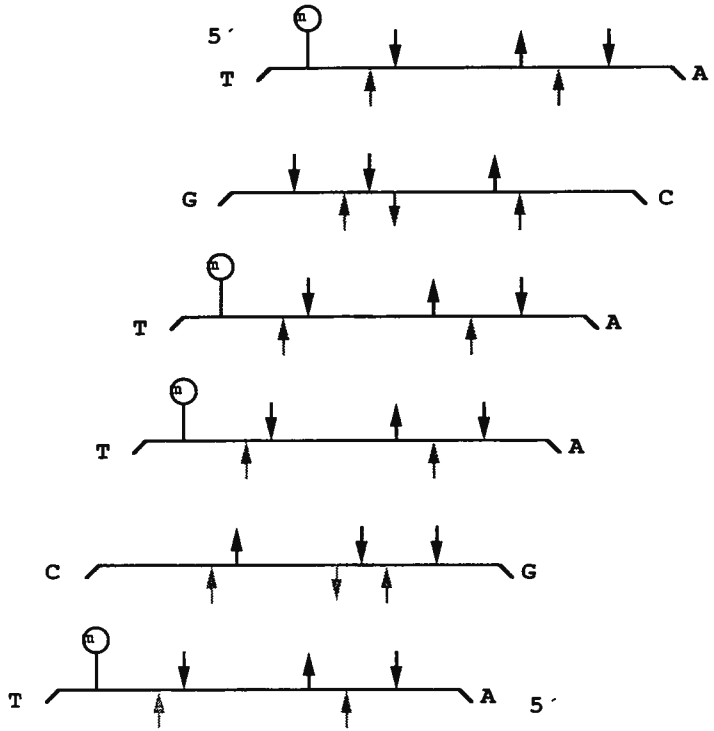


Figure 1.5. Stick Figure of a TGTTCT Motif.
Simplified representation of the TGTTCT motif according to Fig.1.4. The lateral displacement of individual base pairs indicates the torsion of the double helix.

Analysis of the crystal structure of many different oligonucleotides has revealed a great degree of variability in the precise structure of the repeating units along the DNA. Till present there is no evidence for the existence within DNA regulatory regions of strctures other than the B double helix, but many variations of this standard structure do occur. This structural inhomogeneities are determined by the nucleotide sequence and therefore could serve to indirectly recognize a DNA sequence. In this type of recognition the protein distinguishes the particular path of the helix backbone without directly recognizing specific nucleotides. This type of recognition has been postulated for the *trp* repressor (Matthews 1988) and for the interaction between DNA and the histone octamer in nucleosomes (Drew and Travers 1985).

Before closing this introductory chapter I would like to mention a few statistical and thermodynamic considerations concerning the problem of binding site specification. For a detailed treatment of this matter, see the review by von Hippel (1979). There are four types of base pairs in double stranded DNA (A-T, T-A, G-C and C-G) that can be seen as "information elements". A binding site is defined as a precise sequence of base pairs. Statistical considerations can be used to determine the minimum length of base pairs, n, required to specify a recognition site within a genome of size N base pairs. Ideally, this minimal length should be such that the expected frequency with which this site occurs randomly in the genome is less than unity. If the overall base composition is random, with A,T,G and C equally represented, then the probability of a particular base to occur at a certain position is 0.25 x $2N$ (the factor 2 is due to the fact that a binding site in double stranded DNA can occur in either strand). Let us calculate the random probability of the sequence TGTTCT to occur in a typical eukaryotic genome where $N=10^9$ bp.

$$P_n \times 2N=(0.25)^6 \times 2.10^9=4.88 \times 10^5$$

Thus, an hexanucleotide motif will occur randomly some 488.000 fold in an eukaryotic genome. Its information content is not sufficient for a precise binding site specification, as the random sequences will compete for the functional target site and trap the corresponding protein in non-productive interactions. However, if the same motif is duplicated and, as often observed, makes up one half of a palindrome: AGAACA(N)TGTTCT, the expected frequency is:

$$P_n \times 2N=(0.25)^{12} \times 2.10^9=119.2$$

This value is in a probability region that allows more precise specification of the binding site. Though this calculation has been dramatically oversimplified, it provides a feeling for the length of nucleotides required for specific sequence recognition. A plot of the probability against the sequence length permits to read easily the expected frequency for a particular sequence (Fig. 1.6).

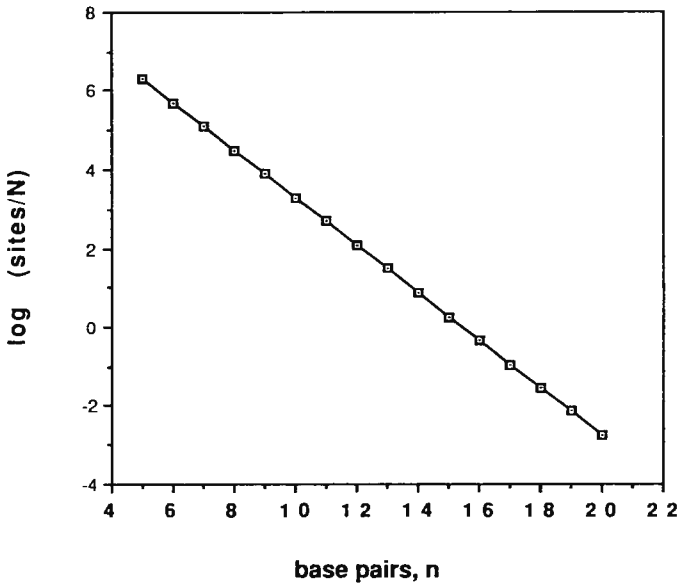


Figure 1.6. Expected Frequency in the Eukaryotic Genome for Sites of Different Length (n).
A genome size, N, of 10^9 base pairs has been assumed. Calculation was as indicated in the text.

These considerations assume that site recognition is absolute and that binding only occurs to the specific site. In reality this is not the case and proteins have a finite and not negligible affinity for non-specific DNA (von Hippel and Berg 1986). It is the difference between specific and non-specific binding affinities that determines the degree of saturation of a particular site at a given concentration of protein (Fig. 1.7). This is an important consideration, as it explains why only in a certain range of protein concentration one observes saturation of the specific site: the so-called "specificity window" (Berg and von Hippel 1987).

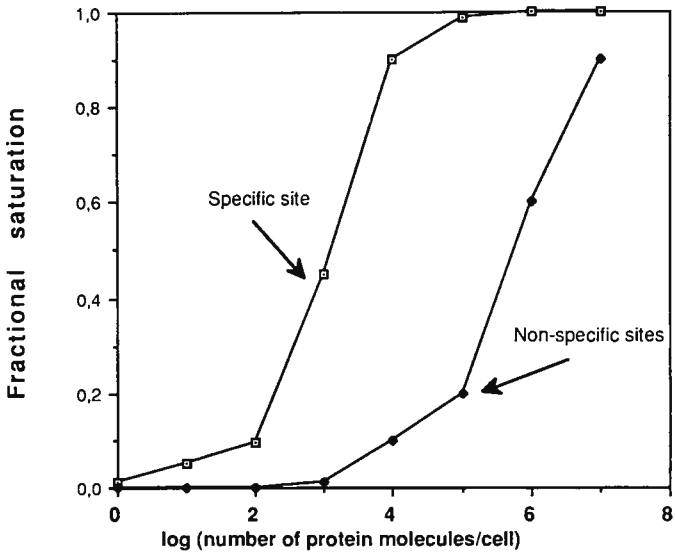


Figure 1.7. Saturation of Specific and Non-specific Binding Sites at Different Concentrations of Intracellular Protein.

For these calculation a difference of 1000 between the affinities for specific and non-specific sites was assumed.

As mentioned above, the affinity of a protein for a particular DNA sequence is mainly determined by two types of interactions: hydrogen bonding and electrostatic interactions. The former is primarily responsible for sequence recognition, but the ionic forces contribute significantly to the free energy of binding. As prior to binding or after dissociation of the protein the hydrogen bond donors and acceptors of the protein and the DNA are involved in hydrogen bonding with water molecules, the average contribution of each correctly formed protein to nucleic acid hydrogen bond is relatively small, in the order of ~-0.5 kcal/mole. Assuming two hydrogen bonds per recognized base pair a 12 bp site will be bound with a free energy of ~-12 kcal/mole. This is a relatively small number and probably a higher contribution to sequence recognition derives from the lost of favourable free energy that occurs at wrong positions. When a functional group of the protein is faced with an "incorrect" base pair at least one hydrogen bond, that was broken in removing the

water solvent, is not replaced. This would lead to an unfavourable contribution of as much as ~+5 kcal/mole. Therefore, incorrect complementarities are highly destabilizing.

Non-specific electrostatic interactions result from the displacement of condensed counterions from the DNA phosphate groups by positively charged amino acids side chains. In the case of the *lac* repressor these forces account for about 40% of the free energy of binding to the *lac* operator (Rezvin and von Hippel 1977). This type of ionic forces are also largely responsible for the non-specific binding of proteins to DNA of random sequence (Fig. 1.7).

Finally, neither the proteins nor the DNA are rigid structures. Both can respond to mutual binding by structural changes that can contribute significantly to the free energy of binding. An example of induced changes in protein structure upon binding to DNA is the *lac* repressor, whereas the *E. coli* cAMP binding protein bends its target DNA (Liu-Johnson, Gartenberg and Crothers 1986). The extent to which this type of induced conformational changes contribute to the stability of specific DNA binding varies extensively with the experimental system and can not be calculated on theoretical terms.

Chapter 2

METHODOLOGY.

The methodological repertoire used for the analysis of protein-DNA interactions was originally developed for the study of bacterial regulatory proteins, and only during the past decade these methods have been adapted to the more complexed eukaryotic systems. Two developments have been essential in this field: the discovery and availability of restriction enzymes with defined sequence specificity and the introduction of the nuclease protection assays by Galas and Schmitz (1978). Later, numerous modifications of this principle, based on the use of different enzymes or chemical reagents have been reported. More recently, a significant improvement was introduced when the band retardation technique (Garner and Revzin 1981; Fried and Crothers 1981) was combined with chemical interference procedures, thus allowing the detection of specific interactions in crude extracts.

In this chapter the different methods will be briefly described, starting with those methods that provide low resolution information (less than nucleotide level) and progressing to the high resolution assays. Most of the discussion will be based on the use on cell-free binding conditions, but at the end I will refer to the genomic footprinting techniques that have been developed for studying protein-DNA interactions in the "intact" cell. Finally I will briefly mention the use of specific protein-DNA interactions for the purification of regulatory proteins and the computer aided molecular modeling of proteins and DNA for the visualization of their interaction.

I. **Low resolution** *in vitro* **methods.**

In a first approximation to the detection of binding sites for a particular protein within a complex genomic region one needs information about the approximate location of the potential binding sites. This information can be obtained by a variety of procedures most of which make use of end-labelled restriction fragments.

I.1. Nitrocellulose Filter Binding.

One of the oldest techniques used for studying protein-DNA interactions in prokaryotic systems is based on the ability of nitrocellulose filters to retain proteins but not double-stranded DNA. Therefore, when a mixture of different labelled DNA fragments is incubated with a putative DNA-binding protein and filtrated through nitrocellulose, only those fragments will be retained that form a stable complex with the protein in question. This method, that was originally developed for the study of the lac-repressor binding to DNA (Riggs, Suzuki and Bourgeois 1970), has since been successfully used in many other systems. The mixture of fragments that is retained on the filter can be analyzed by agarose gel electrophoresis and autoradiography (Fig. 2.1). If several experiments are performed with different restriction enzymes, it is possible to narrow down the region responsible for protein binding to a few nucleotides.

The method is amenable to quantitative analysis, either by cutting the individual bands and counting their radioactivity in a liquid scintillator or by densitometric evaluation. The assay can be performed in a few minutes with many different samples and is also useful to follow a particular binding activity during a purification procedure. In fact the method can be combined with site directed mutagenesis to evaluate the contribution of the individual nucleotides to the binding reaction.

The main limitation of this technique is its poor reproducibility due to its sensitivity to assay conditions, in particular the filtration rate. The results are also dependent on the ionic strength of the binding buffer and on the length of the DNA fragments. Whereas small fragments (<200bp) may create problems due to inefficient retention of the complexes, fragments of up to 5 kb have been successfully used (Geisse et al. 1982). The extent of non-specific DNA binding can also be a problem, in particular when crude preparations of protein and large DNA fragments are used. Addition of non-radioactive competitor DNA may help to solve this problem. Nevertheless, when a long stretch of DNA is to be analyzed for specific protein binding, the filter

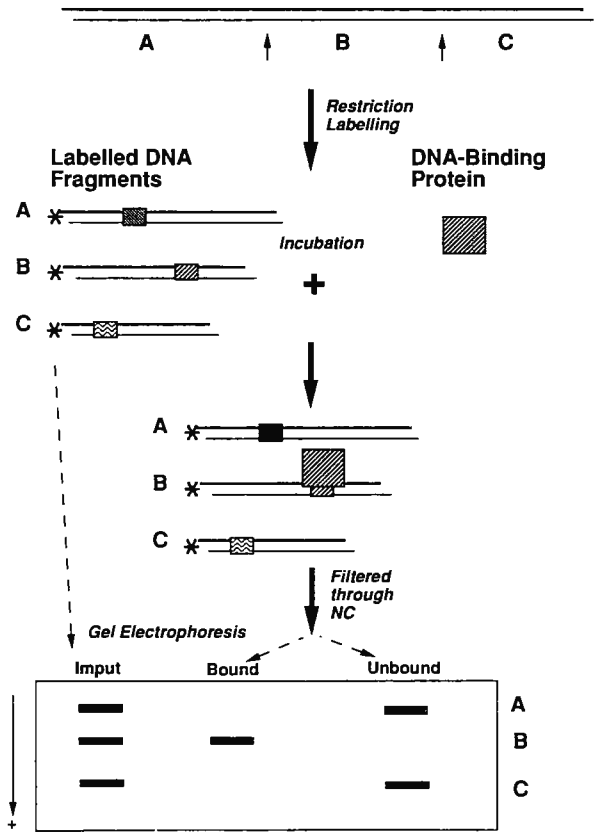


Figure 2.1. Nitrocellulose Filter Binding Assay.
A restriction digest of a DNA molecule that yields three different fragments is shown. Fragment B contains a binding site for the protein of interest and forms a complex upon incubation that is retained on the nitrocellulose filter.

assay is still one of the methods of choice. In particular, because this method is fast and yields quantitative information, it can be used to determined association constants and for kinetic studies (Riggs, Bourgeois and Cohn 1970; Schauer et al 1989)

I.2. Gel Retardation or Mobility Shift Assay.

The gel retardation or band-shift assay is based on the different electrophoretic mobility of free DNA fragments versus complexes of DNA and protein (Figure 2.2A). As the filter assay, this method usually requires the use of a radioactive labelled DNA fragment, although very recently a protocol using non-radioactive detection methods has been reported (Suske, Gross and Beato 1989). The gel retardation assay has the great advantage over the filter binding assay that the individual protein-DNA complexes are directly identified after gel electrophoresis and can be further analyzed. In addition it allows to detect binding of different individual proteins, or more than one protein to the same DNA fragment, since the resulting complexes often exhibit different electrophoretic mobilities (Fig. 2.2B). The separation of the complexes from free DNA can be accomplished in polyacrylamide gels (Garner and Revzin 1981; Fried and Crothers 1981), in agarose gels (Strauss and Varshavsky 1984) or in composite gels containing both polyacrylamide and agarose (Topol, Ruden and Parker 1985). The ionic strength of the electrophoresis buffer can be selected to enhance the stringency of the protein-DNA complex formation. Similarly, the amount of competitor DNA added has to be carefully determined: too little competitor would result in nonspecific binding, whereas too much competitor DNA could also compete with specific binding.

The band shift assay is extremely sensitive, because the protein-DNA complexes are stabilized during the assay by at least two different factors: the low ionic strength of the buffer and the so-called "caging effect". Within the polyacrylamide matrix, the DNA and the proteins can not diffuse away after dissociation and have therefore a higher probability to reassociate (Fried and Crothers 1981; Garner and Revzin 1981). For this reason the

mobility shift assay is more sensitive than the nitrocellulose filter binding assay, in which the complex can dissociate during washing of the filter (Hendrickson and Schleif 1984).

A limitation of this method is that only relatively small DNA fragments can be used, because in larger fragments unspecific binding creates problems and the mobility is not so dramatically influenced by protein binding. In general, when a larger fragment is used on needs competition with specific shorter oligonucleotides to demonstrate that the interaction is taking place with a particular nucleotide sequence. In principle, the method can also be used to analyze binding of a protein to a mixture of different small DNA fragments that can be separated electrophoretically. Those fragments that do not bind the protein migrate as free DNA whereas the particular fragment containing specific binding sequences, will exhibit an altered electrophoretic mobility (Fig. 2.2C).

A very important advantage of the band shift assay is that it can be used with crude cell extracts, provided sufficient non-radioactive competitor DNA is added to suppress unspecific protein binding to the radioactive fragment (Schauer, Chalepakis and Beato 1988). To identify the nature of the protein componenet of the retained complex, sepecific monoclonal antibodies can be used (Kristie and Roizman 1986). The antiboides can either interfere with DNA binding and prevent the formation of a complex, or can bind to the DNA-bound protein generating a ternary complex that migrates even more slowly. When the band shift assay is combined with protection or interference methods, it can generate high resolution information even when only crude preparations of binding protein and even whole nuclear extracts are used. In protection experiments the complexes of protein and unmodified DNA are digested with nucleases or treated with dimethyl sulfate (see bellow) prior to electrophoretic separation of the free and complexed DNA. In interference assays, the DNA is chemically modified prior to incubation with the protein preparations. After electrophoresis, the bands corresponding to free and complexed DNA are isolated and analyzed in sequencing gel (Maxam and Gilbert 1980). Interference experiments are usually performed with dimethyl sulfate (Ogata and Gilbert 1978), but recently potassium

permanganate (Truss, Chalepakis and Beato 1989) and hydroxyl radical interference have been successfully used (Chalepakis, and Beato 1989).

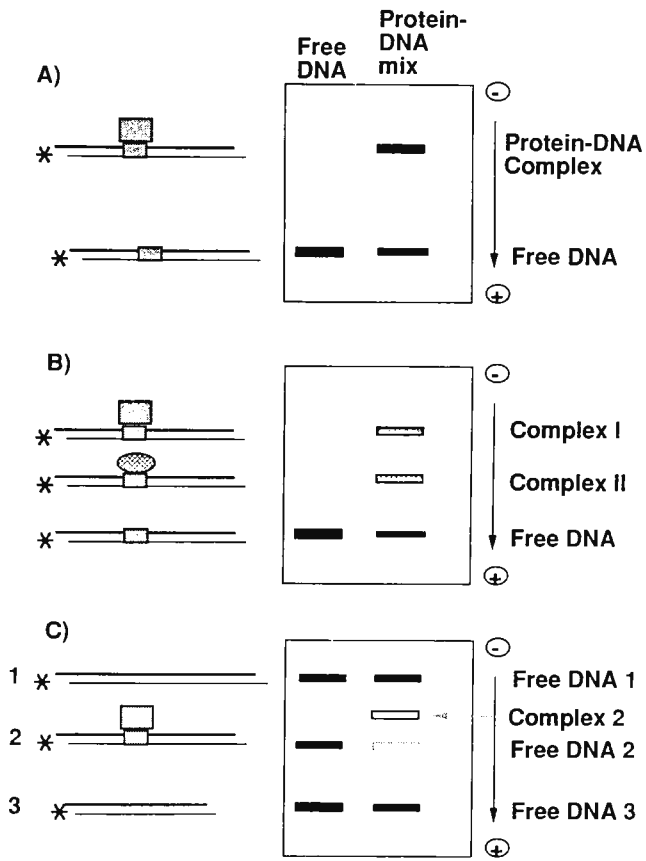


Figure 2.2. Gel Retardation Assay.
A) Simple gel retardation assay. Those fragments forming a complex with proteins are retained and migrate more slowly into the gel.
B) Two different complexes of the same DNA fragment with two proteins of different molecular weights can be separated upon electrophoresis.
C) Out of a mixture of DNA fragments those containing a binding site for a protein form a complex that migrates more slowly, whereas the other fragments are unaffected.

Another very practical use of the band shift assay relates to DNA bending. The mobility of DNA in acrylamide gels depends on the secondary structure of the double strand and, in particular, on its curvature. The binding of a protein to a linear fragment of DNA has a differential influence on the migration the complex depending on the position of the binding site realtive to the ends of the DNA fragment (Zinkel and Crothers 1987). Protein-induced bending of DNA can be determined using circularly permutated fragments and an electrophoretic mobility assay.

The band shift assay can be used not only to study binding of proteins to linear DNA fragments, but also to close circular DNA. It can be used to detect preferential binding of a protein to different topoisomers of a DNA minicircle (Nordheim and Messe 1988; Truss and Beato 1989). It has been also used to detect nucleosome reconstitution and to measure specific binding of regulatory proteins to labelled nucleosomes (Linxweiler W and Hörz W 1984; Piña, Bruggemeier and Beato 1989). Recently this method has been used to study protein binding to DNA in intact cells after transfection of a radioactive labelled fragment (see bellow and Harel-Bellan et al. 1989).

A variant of the band retardation assay, that has been called "reverse mobility shift assay", is based on the use of radioactive proteins synthesized in vitro and unlabelled DNA (Hope and Struhl 1985). This method was developed to study the GCN4 protein of yeast after transcription and translation of cloned cDNA in vitro. Here, the ^{35}S-labeled protein is detected and careful control of nonspecific binding is essential. However, competition experiments with appropiated oligonucleotides enable to determine specific binding precisely. The great power of this technique is that it permits to study any mutant of the particular protein, simple by modifying the cDNA and producing the correponding protein in vitro (Hope and Struhl 1986). Using protein mutants of different sizes it is possible to determine the stoichiometry of DNA binding and to detect the formation of homodimers or heterodimers.

I.3. Avidin-Biotin complex DNA assay.

A very simple DNA binding assay using biotinylated DNA oligonucleotides has been described for the measurement of thyroid hormone receptor binding (Glass et al 1987). This so-called ABCD-assay was developed to allow the detection of receptor labelled with [^{125}I]-ligand when complexed to DNA. Thus, contrary to the

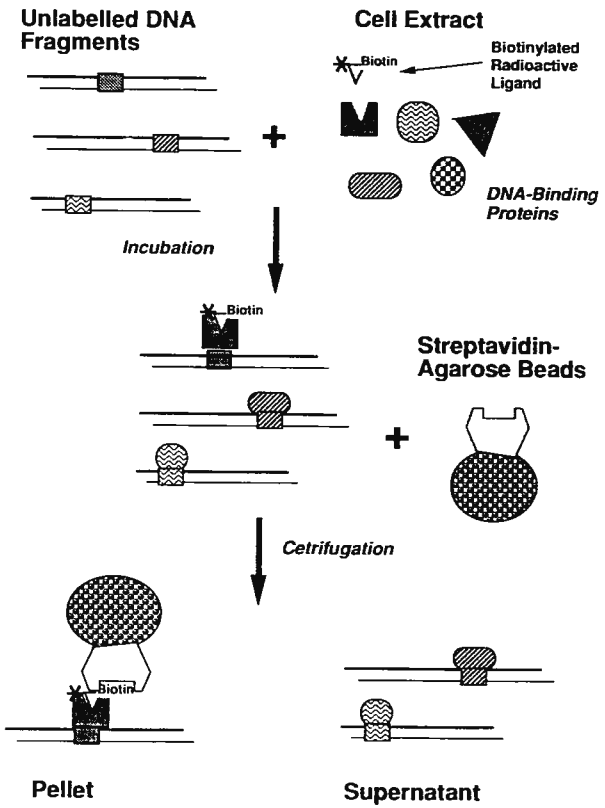


Figure 2.3. Avidin-Biotin Complex DNA Assay.
A biotinylated radioactive ligand is used to label a protein and follow its binding to an unlabeled DNA fragment. Isolation is based on precipitation by streptavidin-Agarose beads.

usual methods, the DNA is not radioactively labelled. After incubation with the biotinylated DNA the protein-DNA complexes are precipitated from solution using streptavidin conjugated to agarose beads (Fig. 2.3.). Using mutant oligonucleotides and different competitor DNA fragments this assay can yield very valuable information even when only crude protein preparation are available. As with the filter assay, the main limitation here is that the complexes are not physically identified, and can therefore not be further analyzed. In addition, only relatively short DNA fragments ca be used, and therefore the analysis has to be focused to a particular DNA sequence. Whenever the protein of interest can be selectively labelled to high specific activity, the ABCD-assay should yield valuable information. Like the filter binding assay the ABCD-assay can be used for the determination of equilibrium and kinetic constants.

I.4. Immunological techniques.

Whenever antibodies to the protein of interest are available, immunological techniques can be used for separating specific protein-DNA complexes from free DNA (Fig. 2.4). In principle this method can be used for the same purposes as the filter binding assay, but introduces the additional specificity of the antibodies as a selective tool for identifying the appropriate complexes. Therefore, the immunological assay is less sensitive to unspecific artefacts due to binding of contaminating proteins to the radiolabeled DNA. For this reason, it can also be used with relatively crude protein preparations (Willmann and Beato 1986). Since the complexes of protein and DNA are selectively isolated, they can be analyzed by high resolution methods, as described for the band shift assay (Willmann and Beato 1986).

Another use of this method is as a control that a given footprint (see bellow) obtained with a crude protein preparation is really due to binding of the protein against which the antibodies are directed (Scheidereit et al 1983). In order to be use for this purpose the corresponding antibodies must have a negative influence on the DNA binding activity of the protein, or they have to be couple to an insoluble carrier and used to remove the protein from the solution.

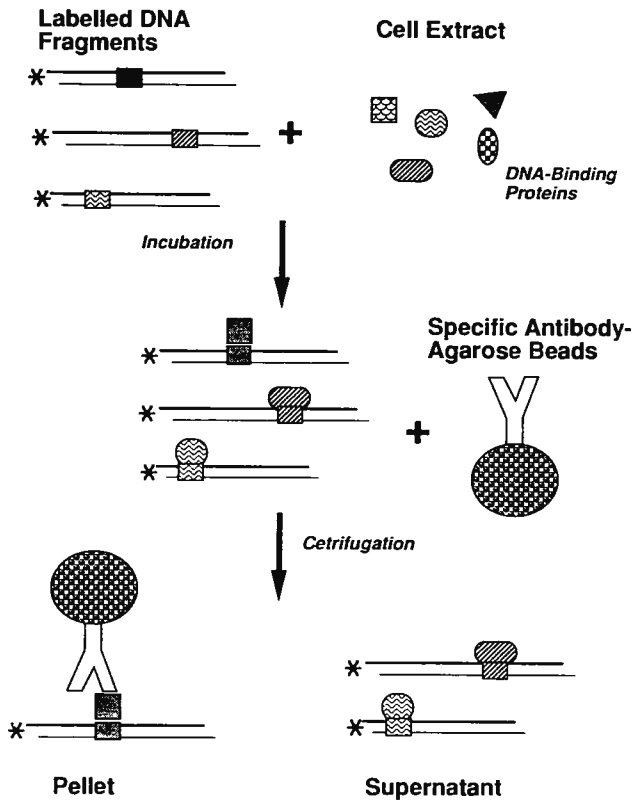


Figure 2.4. Immune-Isolation of Protein-DNA Complexes.

A mixture of labelled DNA fragments is incubated with a protein extract, and the complexes containing the protein of interest are isolated by precipitation with specific antibodies linked to Agarose beads.

I.5. Random Selection Method.

When the protein to be analyzed for DNA binding specificity is available in reasonable purity and sufficient amount, and there is no information on its DNA binding properties, the random selection method can be used to determine sequence specificity. The protein can be attached to an insoluble matrix (cellulose or agarose) and used as an affinity reagent to isolated oligonucleotides of the appropriate sequence. To this end, a mixture of short random sequence oligonucleotides with appropriate cloning ends is synthesized and applied to the affinity column together with carrier DNA. After extensive washing, the specifically bound oligonucloetides can be eluted with high salt and identified by cloning and sequencing (Oliphant, Brandl and Struhl 1989). This procedured has been tested with the yeast regulatory protein GCN4, and enables to generate a nucleotide-use matrix with apprpriate weighting of any particular position (Oliphant, Brandl and Struhl 1989).

I. 6. Protein Blotting (South-Western Blot).

An alternative procedure to detect proteins with affinity for a particular DNA sequence is the so-called South-Western blotting technique (Bowen et al 1980). This method is based on the electrophoretic separation of the proteins in SDS-polyacrylamide gels, followed by blotting onto nitrocellulose paper and incubation of the paper with radiolabeled DNA fragments (Fig. 2.5). If after washing the membrane a radioactive band is detected, it suggests the existence of a polypeptide of the corresponding molecular weight and with high affinity for the labelled DNA. The specificity of the binding can be tested by using different radiolabeled DNA fragments, or by competing the signal with non-radioactive DNA of various sequence composition. With this method polypeptides binding to regulatory regions of different genes have been identified (Cote, Wang and Chiu 1985; White et al 1985; Miskimins et al 1985; Fainsod et al 1986; Silva et al 1987).

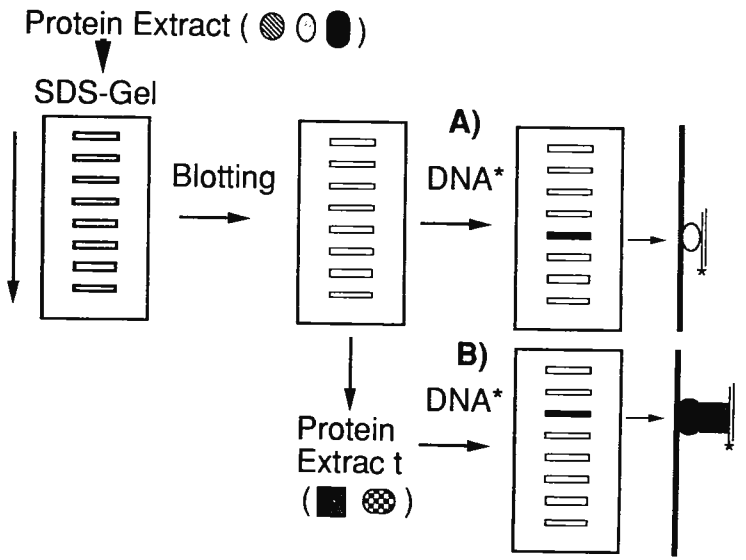


Figure 2.5. South-Western Blotting.
The polypeptides in the protein extract are separated by SDS-polyacrylamide gel electrophoresis, blotted to nitrocellulose and incubated directly with a radioactive DNA fragment (A), or incubated first with another protein extract, followed by the radioactive DNA (B).

The limitations of this procedure derive from the use of denaturing electrophoretic conditions, that lead to the uncertainty of whether the renaturation of the protein structure, required for DNA binding, will take place during blotting and incubation. Thus, a negative result with this method is not very informative. Nevertheless, this procedure may be useful as a screening technique when a positive result has been obtained, for instance with partially protein preparation. It will allow to screen many different cell or tissue extracts in a search for the existence of a particular polypeptide with the desired binding specificity.

Recently this method has been used to detect indirect protein-DNA interactions with a somehow complicated but refined protocol (Hübscher 1987; Matsumo et al 1989). In this protocol the nitrocellulose filter is first incubated with an appropriate extracts containing the potential DNA-binding factor, and subsequently incubated with a radiolabeled DNA fragment. If there is a polypeptide in the mixture of electrophoresed proteins with affinity for a specific DNA-binding protein a radioactive band will show up in the corresponding position. This method could be useful to detect protein mediators for DNA binding of regulatory proteins such as the Adeno E1A gene product, that themselves have no affinity for DNA (Moran and Mathews 1987; Lillie and Green 1989).

II. **High Resolution in vitro Methods.**

The methods described above can under certain conditions, yield information about the DNA sequences implicated in binding of a particular protein, but in order to get high resolution information sequencing gels must be used (Maxam and Gilbert 1980). There are two basic strategies to obtained this type of information: protection and interference. Protection assays determine the influence of protein binding on the reactivity of DNA molecules, whereas interference assays detect the influence of DNA modifications on the subsequent binding of particular proteins. Protection assays often make use of enzymes, usually

nucleases, that attack the DNA more or less randomly, but can also be performed with chemical reagents. Interference techniques usually involve prior chemical modification of the DNA at selected positions.

In the following I will first describe protection assays using nucleases. I will then move to protection against chemical modification at different positions, and will then addressed the binding interference techniques. At the end I will briefly discuss techniques that use UV-light to cross-link proteins to DNA, although this method will be describe in more detail in the section on genomic footprinting (see bellow).

II.A. **Protection assays.**

II. A. 1. Exonuclease III Protection Assay.

This method occupies an intermediate position between low and high resolution techniques, as it only provide information on the 3'-limits of regions contacted by proteins. The method takes advantage of the exonucleolytic 3'-5' progressive degradation of double stranded DNA by E. coli exonuclease III (Richardson, Lehman and Kornberg 1964). The enzyme enters both 3'-ends of a DNA fragment and progresses until enzyme molecules moving in opposite direction meet at the center of the DNA fragment. At this point digestion stops because of steric hindrance and because the enzyme requires double stranded DNA as substrate. When a fragment labelled at one of the 5'-ends is used, a series of bands corresponding to the half size fragment are detected in the autoradiogram of the subsequent sequencing gel. If, however, a protein is tightly bound to a particular region of the DNA it prevents progression of the enzyme, and a band corresponding to the 3'-limit of the region protected on the labelled strand is seen in the autoradiogram (Fig. 2.6). The experiment can be repeated with the labelled at the other strand, and in this way the region protected by the particular protein can be delimited. The method was originally developed for studying DNA binding of highly purified SV40 T antigen (Shalloway, Kleinberger and Livingston 1980), but it can be used to study specific protein binding in crude extracts (Wu 1985; von der Ahe et al 1985). The

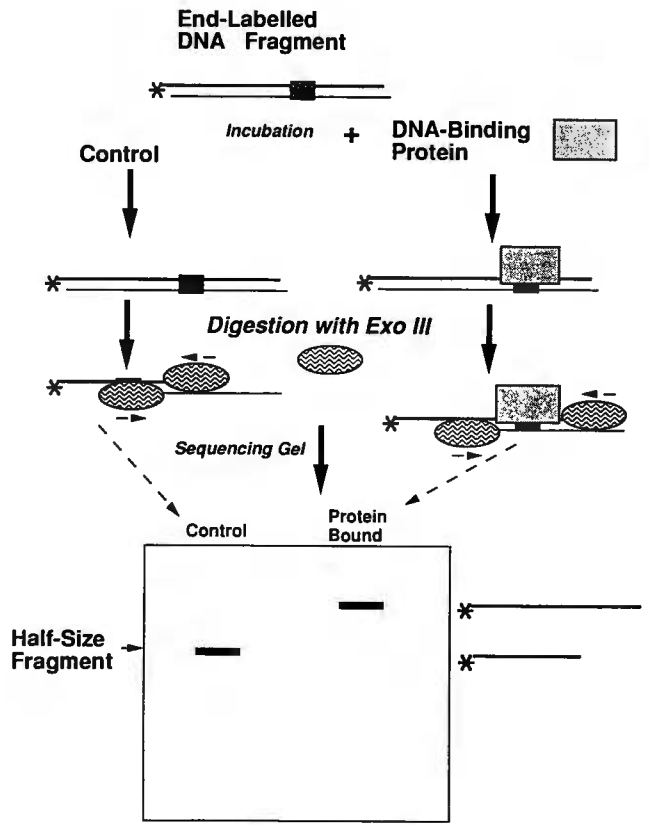


Figure 2.6. Exonuclease III Protection Assay.

An end-labelled DNA fragment is subjected to exonuclease III digestion in the absence (Control) or in the presence of a DNA-binding protein. The digestion products are analyzed in a sequencing gel.

procedure has also been employed for studying chromatin structure (Wu 1984) and nucleosome positioning (Linxweiler and Hörz 1985). As we will see bellow the method can also be used to study the binding of proteins to chromatin in situ in the intact cell (Cordingley, Riegel and Hager 1987).

One of the major advantages of the exonuclease III assay is its very high sensitivity. Since the presence of a protein bound to DNA generates a positive signal, e.g. a new band on the autoradiogram, the method can be used even when the concentration of the particular protein is low (Chan and Lebowitz 1983; von der Ahe et al 1985). Contrary to other protection methods in which a band is weakened upon protein binding, only a small percentage of the labelled DNA molecules have to carry a bound protein in order to generate a visible exonuclease stop. This high sensitivity allows the use of crude extracts if care is taken to prevent non-specific DNA binding by the addition of unlabeled competitor DNA.

The limitations of the procedure are due to the intrinsic properties of the exonuclease. Unclear results are obtained when the binding site is located far away from the enzyme entry site, or distal to the half size region. In addition, very often sequence-dependent pausing of the enzyme at certain positions is observed that generates a series of bands in the control autoradiogram. If the protein-induced stop coincides with one of these control signals, it can be overlooked. Another limitation is the need for 5'-protruding ends as entry sites for the enzyme. The efficiency with which the enzyme enters blunt ends and, in particular, 3'-protruding ends, is very low. This problem can be overcome by digesting the DNA fragment with the Klenow fragment of DNA polymerase in the absence of nulceotides, to remove a few nucleotides from the 3'-ends of the substrate DNA prior to addition of exonuclease (Scheidereit and Beato 1987).

In principle, a similar procedure can be used with 3'-end labelled DNA fragments and the 5'-exonuclease of the bacteriophage lambda (Camier et al 1985), but this assay has not found general application.

II. A. 2. Protection Against Restriction Enzymes.

Provided appropriate restriction enzymes are located over the region of interest, binding of protein to this region can be detected by analyzing its influence on the efficiency of specific DNA cleavage. Strictly speaking, this procedure is again not a high resolution method, since it does not yield the limits of the protected region, but it has the advantage that it can be used even intact cells. In combination with the kinetic technique described by Smith and Birnstiel (1976), it can provide very valuable information. This method has also been used to study the organization of DNA into nucleosomes (Cordingley, Riegel and Hager 1987). A variation of the method, that account to an interference assay, is to study the influence of restricting the DNA with a selected enzyme on the ability to bind a particular protein.

II. A. 3. DNase I footprint.

The most widespread method for the identification of nucleotide sequences recognized by DNA-binding proteins is the DNA-footprinting procedure (Galas and Schmiz 1978). The procedure is based on the use of end-labelled DNA fragments that are digested with limiting concentrations of DNase I, so as to produce less than one cut per DNA molecule. The resulting collection of fragments is separated in sequencing gels and visualized by autoradiography. Ideally a homogeneous ladder of bands, corresponding to cuts at the individual phosphodiester bonds, should be seen in the absence of any bound protein. After binding of sequence-specific protein, the bound DNA region should be protected against nuclease attack and the corresponding bands are missing in the autoradiogram, generating the "footprint" of the protein on the DNA (Fig. 2.7). A sequencing reaction should be run in parallel, that will allow to precisely position the limits of the footprint at the nucleotide level. The experiments should be performed with both strands, because the limits of the footprint are different, and knowledge of their precise position in each strand may help to develop models for the nature of the interaction. Also some proteins only protect one of the strands of the DNA double helix.

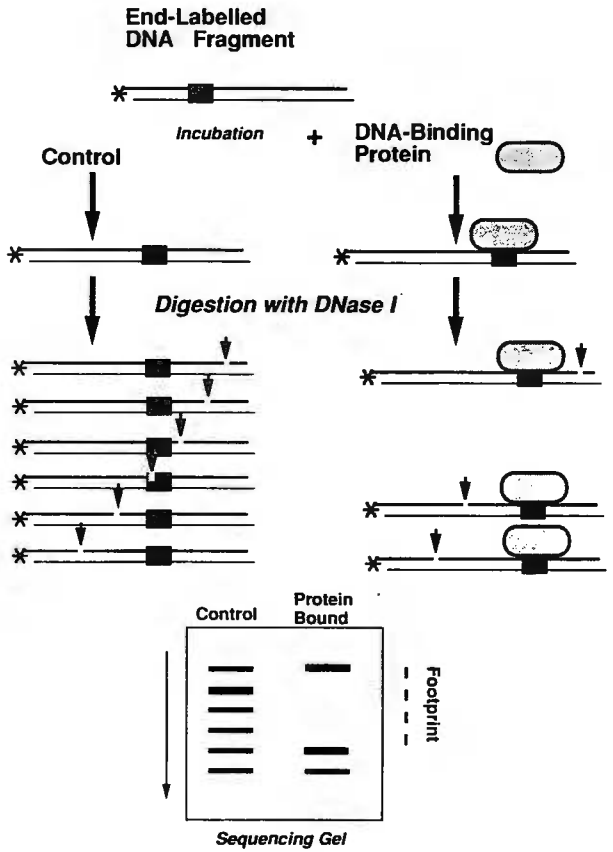


Figure 2.7. DNase I Protection Assay.

An end-labelled DNA fragment is subjected to DNase I digestion in the absence (Control) or in the presence of a DNA-binding protein. The digestion products are analyzed in a sequencing gel. The "footprint" generated by the bound protein is indicated.

In practice, DNase I exhibits considerable sequence specificity and certain positions are preferentially cut, leading to a non-homogeneous ladder of bands. This may complicate the identification of the precise limits of the footprint. However, varying the conditions of nuclease digestion (divalent cations), or even using another nuclease (DNase II or micrococcal nuclease) may help to overcome this problem. Often the identification of the protected region is facilitated by a hypersensitivity to nuclease attack of the phosphodiester bonds flanking the protein binding region. The reason for this hypersensitivity is not well understood, but could result from hydrophobic interactions between the DNA-bound protein and the nuclease, or from deformations of the DNA secondary structure.

The DNaseI footprint method can be used to quantitate the extent of DNA binding to a particular site and yields useful thermodynamic isotherms (Brenowitz et al 1986). It can also be used to measure the rate of dissociation of a protein from its cognate DNA sequence by following the disappearance of the footprint after addition of an excess of competitor non-labelled DNA (Sawadogo and Roeder 1980).

The original method was developed for use with highly purified proteins, but in the meantime conditions have been worked out that allow the use of more crude protein preparations. For this, the main prerequisite is to reduce unspecific binding by addition of competitor DNA, by raising the ionic strength or by any other mean. A very convenient procedure is to separated the protein-DNA complexes from free DNA after nuclease digestion by the band shift assay (Fig. 2.8). The procedure has the additional advantage that the observed protection signal is strong because every DNA molecule in the retarded band contains a bound protein. Other procedures to separate complexes from free DNA, such as the use of specific protein antibodies, have also been employed (Willmann and Beato 1986).

The main limitation of the classical DNaseI footprint procedure is its relatively low sensitivity. Since the signal produce by a bound protein is a negative one, e.g. a reduction of

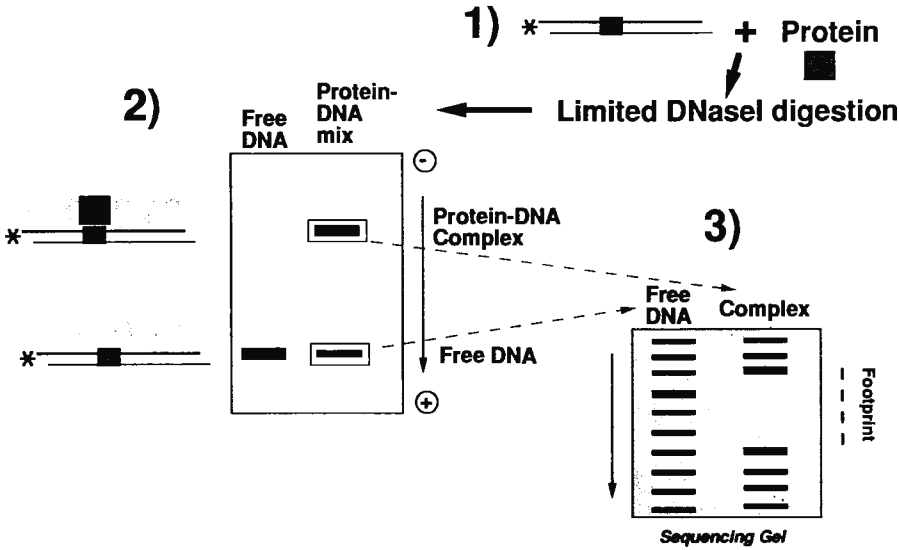


Figure 2.8. Combination of the Band Retardation Assay and DNase I Footprinting.
1) An end-labelled DNA fragment is incubated with or without a DNA-binding protein and subjected to limited DNase I digestion.
2) The digestion products are applied to a non-denaturing gel in which free DNA and protein-DNA complexes ca be separated.
3) The corresponding bands are eluted from the non-denaturing gel, phenol extracted, and applied to a sequencing gel. The "footprint" is indicated.

the intensity of a radioactive band, clear results are only obtained when at least 50% of the DNA molecules are specifically bound. This situation is sometimes difficult to reach, in particular when the protein to be analyzed is present only in low concentration. In this case the combination with the band shift assay may be mandatory.

A reagent that has been occasionally used for footprinting is the antitumor protein antibiotic neocarzinostatin, that cleaves double stranded DNA preferentially at T residues and, with lower efficiency, at A residues (Hatayama et al 1978). The cleavage reaction requires 2-mercaptoethanol and is very inefficient with single stranded DNA. Neocarzinostatin can be used as an alternative to DNaseI both with DNA and chromatin (Ross et al 1979).

The nuclease footprinting technique can also be performed with unlabeled DNA, and the results can be visualized by indirect end-labelling after blotting and hybridization with appropriate radioactive probes. A similar method can be used to study binding of proteins to close circular DNA molecules, a procedure initially reported as "supercoil footprinting (Gralla 1985). This is important because negative supercoiling, or more generally changes in DNA topology, affects the secondary structure of DNA and, therefore, can have an influence of the binding properties of a particular protein. The indirect labelling method has been widely used to study the structure of chromatin within intact cell nuclei (Wu 1980; Nedospasov and Georgiev 1980). The use of this procedure for identifying DNA-binding proteins in situ will be discussed in the section on genomic footprinting (see bellow).

II. A. 4. Chemical Footprinting.

Instead of enzymes chemicals can be used to cleave the phosphodiester bonds of DNA. Methidiumpropyl-EDTA-iron$^{(II)}$ (MPE), for instance, is an intercalating agent that cleaves the DNA with less specificity than most nucleases, yielding a more homogeneous ladder of bands (Cartwright et al 1983; Herzberg and Dervan 1984). This chemical has been widely used for studies of chromatin structure, but is also useful for footprint analysis. The regions protected by DNA-binding protein against MPE attack are usually shorter than those protected against nucleases, and correspond to positions of intimate contact between the protein and the DNA. Thus, the information obtained by MPE footprinting is of higher resolution and stringency (Sawadogo and Roeder 1980; Kownin, Bateman and Paule 1987).

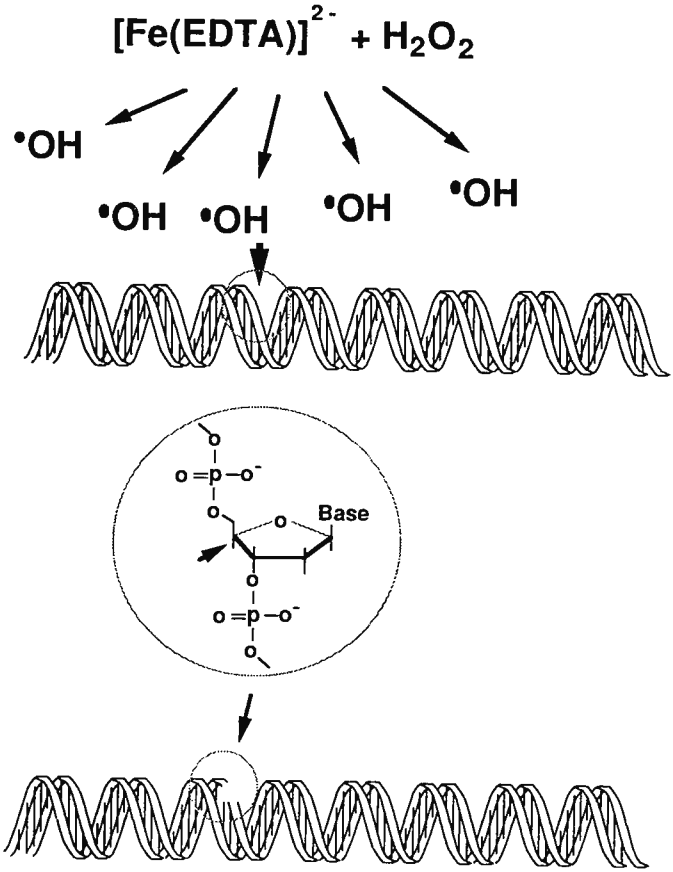


Figure 2.9. Hydroxyl Radical Footprinting.
Mechanism of hydroxyl radical attack on the desoxyribose moiety of the DNA backbone.

More recently, hydroxyl radicals have been used in footprinting analysis (Tullius and Dombroski 1986). Hydroxyl radicals attack the C-4 position of desoxyribose and cleave the DNA backbone (Fig. 2.9). This reagent is very useful for detecting contacts between proteins and the backbone of the double helix, and yields a signal even when no direct interaction with the nucleotide bases take place (Chalepakis, Postma and Beato 1988). Because hydroxyl radicals are very small molecules, the information obtained with this method is of high resolution.

II. A. 5. Methylation protection.

The footprint methods described above are characterized by the use of enzymes or chemical with little or no sequence specificity. To study the participation of individual bases or their reactive groups in the binding reaction one makes use of chemicals that selectively modify certain bases. Most of these base-specific reagents have been developed for the chemical sequencing of DNA. The most popular among these is dimethyl sulfate (DMS) that, under appropriate conditions, reacts with the N-7 position of guanine residues in the major groove and with the N-3 position of adenines in the minor groove of the double helix (Maxam and Gilbert 1980). In practice the modification of the guanines is easier to detect after pyrimidine cleavage and has been widely used. When a protein is specifically bound to DNA and contacts a particular guanine residue through the major groove the reactivity of the corresponding N-7 position towards DMS is usually markedly reduced (Ogata and Gilbert 1978). Therefore, if treatment with DMS is performed in the presence of the binding protein, followed by pyrimidine cleavage, the corresponding band in the autoradiogram is missing or protected (Fig. 2.10). This is indicative of an intimate contact between an amino acid side chain of the protein and the guanine residue, and also shows that the protein binds to the major groove of the double helix. In fact this procedure permits to distinguish among two proteins that bind to the same region of the DNA and generate a similar DNaseI footprint (von der Ahe et al 1986). If adenine residues are protected, and artefacts can be excluded (Beato and Scheidereit 1984), a contact of the protein with the minor groove of the double helix can be postulated.

Very often hyperreactivity of certain purines is observed in the regions flanking a protein bound to DNA (Siebenlist and Gilbert 1980; Scheidereit and Beato 1984). The reason for this phenomenon is not understood, but as with the hypersensitivity to DNaseI, it could be due to either the formation of hydrophobic pockets in the proximity of the DNA-bound protein, or to changes

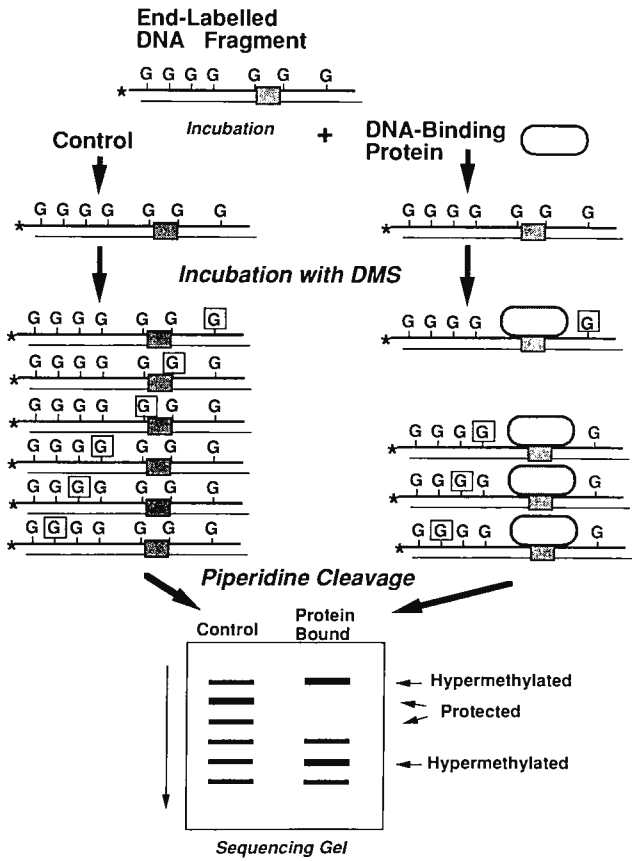


Figure 2.10. Dimethyl Sulfate (DMS) Protection Assay. An end-labelled DNA fragment is subjected to limited methylation by DMS in the absence (Control) or in the presence of a DNA-binding protein. After cleavage with piperidine, the samples are applied to a sequencing gel.

in DNA secondary structure induced by protein binding. In some cases this hypermethylation is diagnostic of a weak interaction between the protein and cryptic DNA binding sites (Scheidereit and Beato 1984).

As with DNaseI footprinting the main limitation of methylation protection is the negative nature of the signal induced by protein binding, and, therefore, its low sensitivity. Sensitivity can be increased by combination with the band retardation assay as described for DNaseI protection (see fig. 2.8). However, the methylation interference assay (see below) is the method of choice.

Methylation protection is also useful for quantitative analysis of the relative affinity of a protein for a particular site on the DNA, as the signals of the individual guanines can be easily quantified densitometrically (Scheidereit and Beato 1984). This procedure has also been used for genomic footprinting as isolated nuclei or even whole cells can be treated with DMS under define conditions (see bellow).

Several other reagents have been used for protection studies and are also useful for determining secondary structure of DNA. For instance, diethylpyrocarbonate (DEPC) also alkylates the N-7 position of guanines, but it is larger than DMS and, due to steric hindrance, reacts only poorly with double stranded B-DNA. In Z-DNA the guanine residues are more exposed and DEPC reactivity increases (Herr 1985, Johnston and Rich 1985). Therefore, this reagent is useful in detecting changes in DNA structure that may be induced by protein binding (Furlong and Lilley 1986; Scholten and Nordheim 1986).

II. B. **Interference Assays.**

Interference assays are based on the negative influence of modifying certain chemical groups of the DNA on the subsequent binding of protein. They make use of the same type of radioactive end-labelled fragments and frequently the same chemicals as used

for the protection assays, but the modification of the DNA is performed prior to protein binding. After incubation with the DNA-binding protein, free and complexed DNA molecules are separated (by the band shift assay, by filtration through nitrocellulose or by means of immunological techniques) and the extent of modification at individual positions analyzed in sequencing gels. If modification at a certain position interferes with protein binding, the corresponding band is underrepresented in the population of protein-complexed DNA molecules, and a in the population of free DNA molecules.

One general limitation of the interference technique is that the DNA-fragment to be studied should contain a single binding site for a particular protein. If the fragment contains more than one binding site and the different sites do not exhibit high cooperativity, modification at functionally significant positions of a particular site will not be detected as reduced bands in the autoradiogram, because the protein can still bind to the other unmodified sites within the fragment.

In the following I will first describe interference procedures that use base-specific reagents and, therefore, permit conclusions as to the participation of individual bases or their reactive groups on protein binding. At the end, I will mention those interference procedure that use reagents reacting with the backbone of the double helix. The interference technique can be used with enzymes, in particular with restriction enzymes, by analyzing the influence of restriction on protein binding, but I will not deal with this obvious type of procedures. I will not discuss either procedures that utilize mutagenesis of individual bases, followed by a conventional analysis of the influence of these changes on protein-binding. Though this procedure has been widely used it does not represent a new binding technique.

II. B. 1. Methylation Interference.

The most widespread interference technique uses DMS treated DNA. If methylation at the N-7 position of a single guanine residue is sufficient for preventing binding of a protein, the

corresponding band will be missing from the subsequent autoradiogram (Fig. 2.11. and Scheidereit and Beato 1984). It is

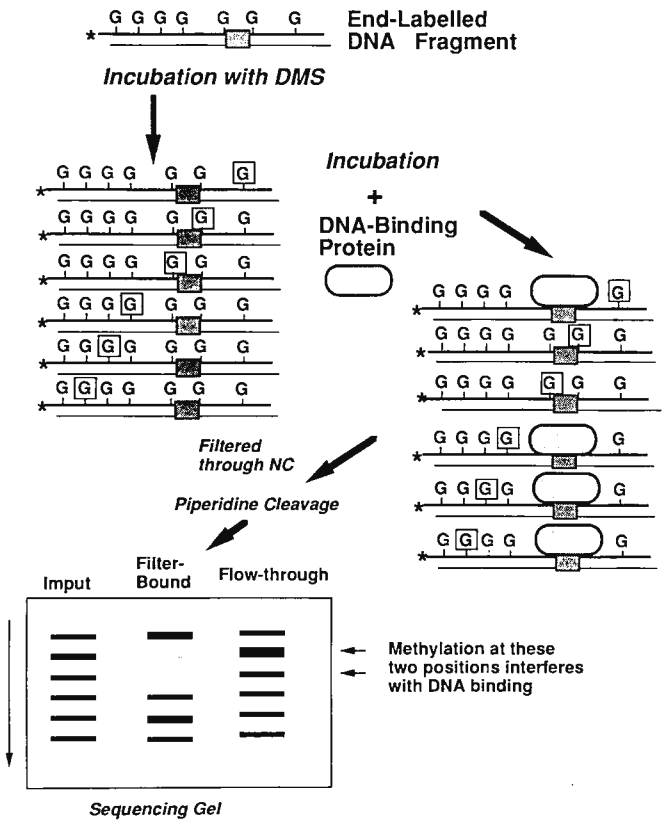


Figure 2.11. Methylation Interference Assay.
An end-labelled DNA fragment is methylated with limiting amounts of DMS and incubated with a DNA-binding protein. The free DNA and the DNA-protein complexes are separated by nitrocellulose filtration (or band retardation), cleaved with piperidine and analyzed in sequencing gels.

important to note that the interference assay is more stringent than the methylation protection assay, and that some positions yield a signal in the protection assay because they are contacted by the protein, but do not show up in the interference assay because a methyl group at this position is compatible with binding (Scheidereit and Beato 1984).

One advantage of the methylation interference procedure is that chemical modification of the DNA is performed prior to addition of the protein. Thus, the protein is not exposed to DMS that could modify certain amino acid side chains and inhibit DNA binding. In addition the modified DNA is still normally double stranded and can be used for band shift assays without particular precautions. This combination of methylation interference and band retardation assay is now a standard technique for studying specific binding of DNA regulatory proteins. Like other methods based on band retardation, it can be used with crude protein preparations and still yields high resolution information.

II. B. 2. Permanganate Interference.

Very often the regulatory sequences recognized by DNA binding proteins contain conserved thymine residues (Ivarie 1987). A classical way to study interactions of proteins with T-residues has been the use bromdesoxyuridine substituted DNA followed by UV-irradiation (Ogata and Gilbert 1977). This treatment leads to cross-linking of the protein to DNA (see bellow), that can be detected in different way. The method is tedious as requires the synthesis of modified oligonucleotides. Recently, we have developed an interference technique based on the controlled modification of single stranded DNA with potassium permanganate, that permits to detect specific contacts with the 5'-methyl groups of thymines in double stranded DNA (Truss, Chalepakis and Beato 1989). At neutral pH permanganate selectively attacks the C5/C6 double bond of thymines and leads to opening of the pyrimidine ring, a reaction called glycolization (Rubin and Schmid 1980; Howgate, Jones and Tittensor 1988). The modified DNA can be renatured and used for binding interference in combination with

the band shift assay. Modification of those positions where the protein contacts the methyl group of thymine leads to inhibition of binding, and the corresponding band is absent from the protein-DNA complexes. This method is simple and specific, and yields very valuable information.

II. B. 3. Missing Contact Probing.

An interference procedure that can be applied to any particular base is the so-called missing contact probing (Brunelle and Schleif 1987). This method is based on the modification of the individual bases by the reagents used in chemical DNA sequencing, such as hydrazine, that lead to dramatic changes in the structure of the bases. Following modification, a classical interference assay can be performed, and those positions where a contact with the base is important for DNA binding will be underrepresented in the protein-DNA complexes. The main disadvantage of this technique is that the extensive chemical modification of one base, can lead to changes in the secondary structure of the DNA, that themselves could influence protein binding. Nevertheless this method is useful for the analysis of contacts between proteins and cytosine or adenine residues which reactive groups are not easily studied by more specific reagents.

II. B. 4. Ethylation Interference.

The interference methods mentioned above are all directed to analyze the interaction of DNA-binding proteins with specific bases. The possibility also exists to investigate the contacts between the protein and the phosphate-sugar backbone of the double helix. Contacts with the phosphate groups are usually important for the free energy of binding and are in principle independent of the particular base sequence. One method to analyze contacts with the phosphate groups is to alkylate then with ethylnitrosourea to form phosphotriester containing an ethyl group. If a protein contacts a particular phosphate group, its ethylation will interfere with binding (Siebenlist and Gilbert 1980). This procedure has been successfully used to analyze contacts of RNA polymerase and other DNA-binding proteins with the double helix (Siebenlist, Simpson and Gilbert 1980). One disadvantage of this

procedure is that it does not yield information on whether the protein approaches the double helix from the major or the minor groove. Since, however, the concentration of negative charges is higher in the minor groove, ionic contacts with the backbone of the DNA usually take place over the minor groove of the helix.

II. B. 5. Hydroxyl Radical Interference.

As mentioned above, hydroxyl radicals react with the sugar moiety of the DNA backbone (Tullius and Dombrowski 1986). Usually this reagents has been used as a footprinting tool, but recently we have combined hydroxyl radical modification and the band shift to generate interference data (Chalepakis and Beato 1989). In addition, since hydroxyl radicals cleave the DNA backbone, this method yields information on the influence of single DNA nicks on protein binding.

II. C. Cross-linking.

As mentioned above, DNA containing bromodesoxyuridine (BrdU) instead of thymine forms covalent complexes with specifically bound proteins after irradiation with UV light of a wave length around 300 nm (Ogata and Gilbert 1977). The efficiency of cross-linking depends on the proximity between amino acid side chains and the bromide group. Usually the introduction of BrdU into the DNA binding site for a protein does not interfere with binding as the bromide atom has a van der Waals radius similar to that of the methyl group. In fact, in many cases the affiinity of a protein for BrdU-substituted DNA is slightly higher than for natural DNA. When UV-light of appropriate wave length (254 nm) is used, cross-linking can also be obtained with unsubstituted DNA, though with low efficiency. This property is the base for in vivo photo-cross-linking in intact cells in particular with UV-lasers (Hockenschmith et al 1986).

A very important application of the cross-linking methods is the identification of the molecular weigth of DNA-binding proteins. For this purpose, the covalent complex with radioactive oligonucleotides, are submitted to DNase digestion until only a short piece of labeled DNA remains bound to the protein. This

complex can than be analyzed in SDS-gel electrophoresis to get an estimate of the molecular weight of the protein. As the efficiency of cross-linking is usually low (bellow 5-10%), only a single cross-linking event per DNA molecule is observed. Thus, if a particular protein binds to the DNA as a dimer, only the monomer form will be detected in SDS-gels after cross-linking. To improve the efficiency of photo-cross-linking DNA molecules substituted with the thymidine analog azido-dU instead of BrdU can be used (Evans and Haley 1987). With this technique, the frequency of cross-linking events is high enough to detect two proteins bound to a single DNA molecules (Roux et al 1989).

The UV cross-linking technique can be employed with crude mixture of proteins, provided sufficient competitor DNA is added to prevent nonspecific binding. In that case, the complex have to be isolated by the band shift assay after irradiation. Alternatively, a binding reaction can be electrophoresed first, and the gel resulting gel submitted to UV irradiation. In this way, the retarded band of interest can be cut off the gel and analysed directly.

Other less specific procedures to obtain cross-linking between proteins and DNA are based on the use of chemical reagents such as glutaraldehyde or formaldehyde . These reagents are often used for the electron microscopic analysis of complexes between DNA and nucleic acids.

III. Purification strategies based on DNA binding.

The ability of regulatory proteins to bind to DNA has been widely used for their isolation by binding to random DNA attached to a matrix (Alberts and Herrick 1970). There have been isolated reports on the use of site-specific DNA-affinity chromatography for the isolation of well characterized bacterial regulatory proteins (Herrick 1980). More recently, specific oligonucleotide affinity chromatography has been developed as a powerful method for purification of several DNA-binding proteins, including the transcription factors SP1, AP1 and NF1 (Briggs et al 1986; Jones et al 1987; Kadonaga et al 1987; Lee, Mitchel and Tjian 1987). Use of this procedure depends on the specificity of DNA binding and

requires a high degree of binding selectivity, that means a large difference in the affinity for specific and non-specific DNA sequences. Those proteins that exhibit relatively high affinity for random DNA, such as the steroid hormone receptors, have not been successfully purified by this strategy, though random DNA-cellulose chromatography is a standard purification step (Wrange, Carlstedt-Duke and Gustafsson 1979; Westphal and Beato 1980).

Instead of oligonucleotides attached to an insoluble matrix, biotinylated DNA molecules can be used for the isolation of DNA-binding proteins in a procedure similar to the ABCD binding assay (see above). The bitinylated oligonucletides can be incuabted with a mixture of proteins in the presence of an excess of competitor DNA, and the complexes containing specifically bound proteins can be removed from the misture ba incubation with a matrix coated with streptavidin. The proteincan then be eluted from the matrix by high salt. In this way a single matrix can be used for many different cicles and with different oligonucleotides.

It is also possible to isolated proteins bound to specific oligonucleotides without the use of an insoluble support. For this purpose the technique of preparative band retardation can be used. After defining the conditions for the formation of specific protein-DNA complexes (see section I.2.), a large volume incubation can be applied to a preparative gel, and the position of the retarded complex can be identified by short autoradiographic exposure of the wett gel. The complex can be cut off the gell and the protein component of the complex can be further purified by conventional techniques. In this way, small amounts of specifically bound proteins can be isolated, sufficient for microsequencing and cloning (Meisterernst et al 1988b).

The DNA binding properties of proteins can also be used to directly clone the corresponding cDNA from expression libraries (Singh et al 1988). This strategy enables the cloning of the corresponding protein withou the need for previous purification or specific antiboides. Provided the DNA binding domain of the protein of interest is able to bind specifically and with high affinity to DNA after expression in E. coli, a recombinant clone expressing the corresponding protein can be identified by probing the replica filters with a radioactive oligonucleotide of the

appropriate nucleotide sequence (Singht et al 1988; Staudt et al 1988). The potential of this method for the future identification and analysis of regulatory proteins can not be overstimated.

IV. In vivo methods.

Experiments that measure binding of regulatory proteins to labelled DNA fragments in extracts, are very useful for determining the potential of a particular protein to interact with particular DNA sequences, but they may not reflect the precise behavior of these same proteins in situ, in the intact cell nucleus. Within the cell, the concentrations of the protein and DNA are vastly different from those used in the reaction tube, and there is a large excess of random DNA. In addition, the DNA is tightly packaged within the cell nucleus in form of chromatin, and many other proteins can interfere with binding of a particular protein to its cognate DNA sequences, either by competing for these sites or by protein-protein interactions. Thus, ultimately the ideal method of analysis should enable to look at protein-DNA interactions within the intact cell without perturbing its integrity. During the past few years attempts to developed such methods have made considerable progress though the ideal procedure is not available yet.

Initial reports were based on the use of UV-light (Becker and Wang 1984; Gilmour and Lis 1984) or the technique of "genomic footprinting" (Church and Gilbert 1984; Jackson and Felsenfeld 1985; Nick and Gilbert W 1985). Both methods can be combined (Selleck and Majors J 1987), and have been refined to yield clear signals by use of the polymerase chain reaction technique (Saluz and Jost 1989). Essentially the method consist in modifying the DNA in situ, either physically (with UV light), chemically (with DMS), or enzymatically (with nucleases) under mild conditions, followed by isolation of the DNA and detection of the modified nucleotides in sequencing gels by the technique of indirect end-labelling (Nedospasov and Georgiev 1980). The main problem is the damaged caused to the cell by the DNA modification procedure. In that respect the most promising procedures are those based on the use of UV-lasers, as they involved very short pulses (nanoseconds) and the wave length of the incident light can be narrowly selected (Hockenschmith et al 1986; Angelov et al 1988). The chemical

modification with DMS can be performed with intact cells, as they can be easily made permeable for DMS. In the enzymatic procedures, cell nuclei have to be isolated first, in order for the nucleases to gain access to the genomic DNA. In some cases this protocol may be sufficient to dissociate a labile protein from its binding site on the DNA (Becker et al 1986). As for the method of detection, blotting of the sequencing gels followed by indirect end-labelling has been widely used (Church and Gilbert 1984), though alternative procedures that avoid blotting have also been developed (Jackson and Felsenfeld 1985). Recently, linear amplification by the polymerase chain reaction has been introduced as a major simplification of the procedure, that also increases its sensitivity considerably (Saluz and Jost 1989). In the future the combination of this method with the use of UV-laser pulses may allow to follow the kinetic of the interaction of different proteins with relevant regions of the DNA in vivo.

V. Computer modelling.

The powerful methods of computer graphics have enable the representation of macromolecules in the three-dimensional space, and its manipulation according to theoretical and experimental constrains. Interactive model building of proteins using the molecular graphic program FRODO (Jones 1978) and of nucleic acids using the MIDAS software (Langridge et al 1981; Bash et al 1983) have been widely used for real time visualization of different protein-DNA interactions (Scheidereit et al 1986). The advantage of this representation is that, provided a reasonable model exists for the relevant protein domain, it enable certain predictions as to the effect on binding affinity of individual mutations in the protein or the DNA. This predictions can be tested experimentally using site-directed mutagenesis, and the results can be incorporated as additional constrains into the modelling process. In this way a progressive refinement can be achieve leading to more precise predictions, that can also be experimentally tested. Before cocrystalls of the protein and DNA are available, this type of strategy may be a very suitable approach to understanding the molecular details of the interaction. Once X-ray data become available, the computer modeling is again a powerful tool for visualization and refinement of the interacting macromolecular surfaces.

Chapter 3.

SPECIFIC EXAMPLES: STEROID HORMONE RECEPTORS

Among the systems that have been widely used to study transcriptional control in animal cells, steroid modulation of gene expression has occupied a preeminent position during the past few years. In contrast to many other control mechanism, gene regulation by steroid hormones is dependent on the addition of small and well defined ligands some of which act as agonist and others as antagonists. There is a whole variety of extensively characterized cellular systems that respond to administration of steroid hormones in a cell-specific and hormone-specific way, and induction as well as repression of particular genes is observed. In addition the receptors for steroid hormones were identified as soluble intracellular proteins already in the sixties and have been intensively studied ever since. Thus, in spite of their minute concentration, these DNA binding regulatory proteins were among the first to be purified.

In the early sixties, Peter Karlson formulated the hypothesis that steroid hormones may exert their action on metabolic pathways by inducing the transcription of the genes for the corresponding enzymes (Karlson 1961). This hypothesis was based on the observation that the insect hormone ecdysone induced the formation of puffs in giant chromosomes (Clever and Karlson 1960). Thus, the idea that steroid hormones modulate gene expression is almost 30 years old. During this 30 years an overwhelming amount of evidence has accumulated supporting the validity of Karlson's hypothesis. In the following I will briefly summarized what is known about the molecular mechanism of gene regulation by steroid hormones, focussing on the interaction of the steroid hormone receptors with the cellular genome and its consequences for gene expression. The interested reader can consult recent reviews on this highly actual topic (Evans and Hollenberg 1988; Berg 1989; Beato 1989).

I. Structural and functional domains.

Analysis of the steroid hormone receptors by the techniques of protein chemistry already suggested that they are composed of at least three structural and functional domains (Wrange and Gustafsson 1978; Carlstedt-Duke et al 1987). During the past five years the cDNAs encoding the receptors for virtually all the steroid hormones have been cloned from different organisms. After the initial cloning of the glucocorticoid (Miesfeld et al 1984; Hollenberg et al 1985; Weinberger et al 1985; Miesfeld et al 1986), the estrogen (Walter et al 1985; Greene et al 1986; Krust et al 1986; Kumar et al 1986; White et al 1987) and the progesterone receptors (Jeltsch et al 1986; Loosfelt et al 1986; Conneely et al 1987; Gronemeyer et al 1987), it was recognized that the hormone receptors are organized according to a general scheme with a central short domain responsible for binding to DNA, a carboxy-terminal domain that binds the hormone ligand and an amino-terminal domain with less well defined function (Figure 3.1).

The DNA binding domain.

The central DNA binding domain is the region of these proteins that serves to classify then as members of the nuclear receptor superfamily (Fig. 3.2). The domain is made of some 70 amino acids and is very well conserved among all steroid hormone receptors. In particular there is an array of perfectly conserved cysteins that has the potential to form two so-called zinc fingers, each of them containing a tetrahedrally coordinated zinc ion (Berg 1989). The individual cysteine residues involved in zinc coordination have been identified by site directed mutagenesis and physico-chemical methods (Severne et al 1988; Freedman et al 1988). Besides the cysteine residues, many other hydrophobic and basic amino acids are conserved in this domain, particularly at the knuckle of the zinc finger. Using the technique of "finger swapping", that consists in exchanging individual zinc fingers among different receptors, it was found that although both fingers are important for DNA binding of the receptors, the first finger appears to determine target specificity in that it discriminates

between a glucocorticoid and an estrogen responsive DNA element (Green et al

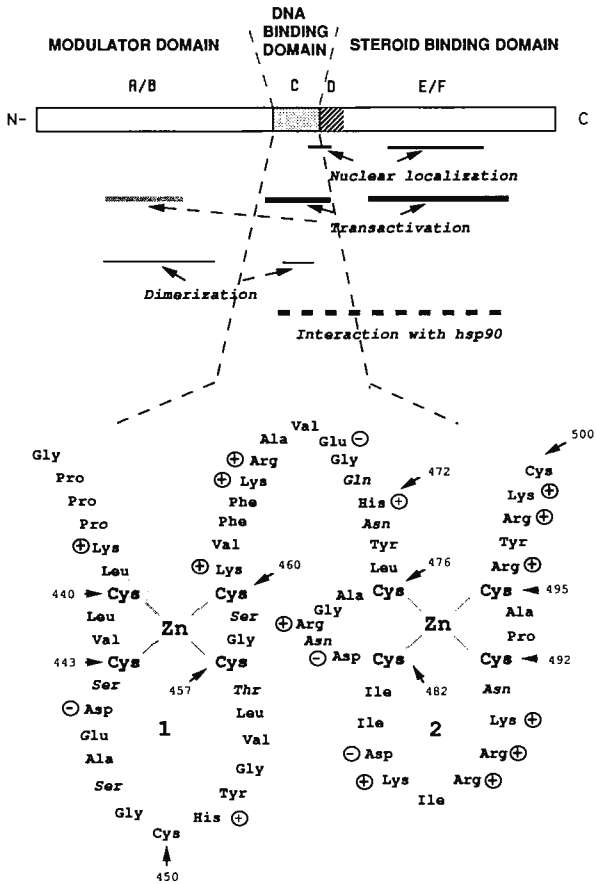
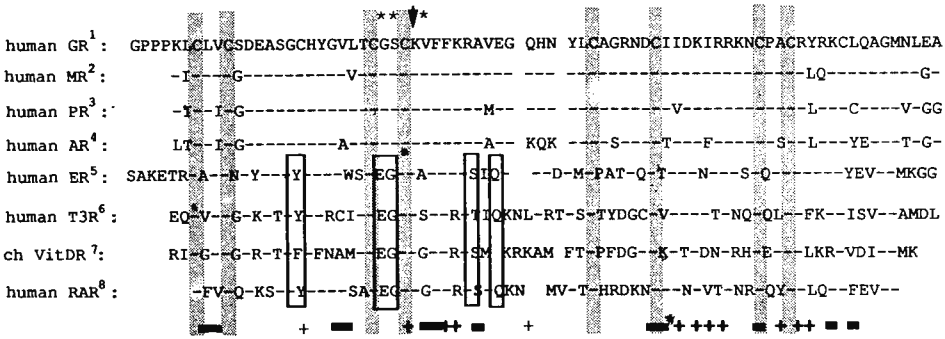


Figure 3.1. General Structure of the Nuclear Receptors. Schematic representation of the domain structure of the nuclear receptors is shown at the top. The different functions attributed to different regions of the protein are indicated underneath. The primary amino acid sequence of the DNA binding domain of the glucocorticoid receptor is shown at the bottom, organized as two hypothetical zinc fingers.

1988). In fact, detailed analysis based on individual amino acid exchanges, has allowed to identified three amino acid residues at the knuckle of the first finger that are decisive for target specificity (Mader et al 1989; Danielsen, Hinck and Ringold 1989; Umesono and Evans 1989). The family of nuclear receptors can be divided into two subfamilies according to the amino acid residues found at these three positions (Beato 1989): the subgroup of glucocorticoid, progesterone, mineralocorticoid and androgen receptors exhibits the structure Cys-**Gly-Ser**-Cys-Lys-**Val**, whereas the estrogen, vitamin D, thyroid hormone, and retinoic acid receptors exhibit the structure Cys-**Glu-Gly**-Cys-Lys-**Ala**(Gly/Ser). In addition to these relevant residues, there are residues at the knuckle of the second zinc finger that are also important for DNA binding specificity, since exchanges in this region generate ambiguous receptors able to interact with more than one responsive element (Danielsen, Hinck and Ringold 1989; Umesono and Evans 1989).

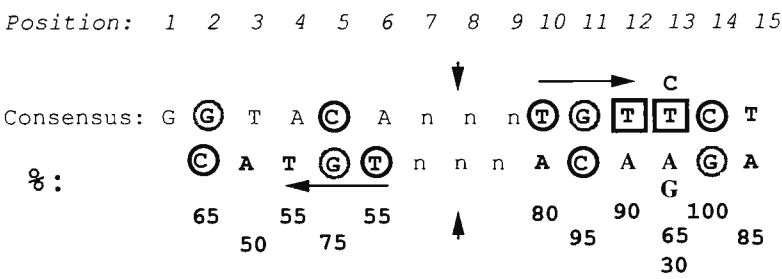


1) Hollenberg et al 1985; 2) Arriza et al 1987; 3) Misrahi et al 1987;
4) Chang et al 1988; Lunahn et al 1988; 5) Walter et al 1985;
6) Sap et al 1986; Weinberger et al 1986; 7) McDonell et al 1987;
8) Giguere et al 1987; Petkovich et al 1988.

Figure 3.2. Amino Acid Sequence of the DNA Binding Domain of Nuclear Receptors.
The amino acid sequence of the DNA binding domain of the human glucocorticoid receptor is shown at the top. The sequence of the receptors for mineralocorticoid, progesterone, androgen, estrogen, thyroid hormone, vitamin D and retinoic acid, are shown underneath. Only the non-conserved positions are indicated.

These similarities and differences in the structure of the DNA binding domains reflect the relative homology of the DNA regulatory elements for the different steroid hormone receptors (Beato 1989). A comparison of DNA sequences known to mediate glucocorticoid induction of various promoters yields a consensus 15-mer, the GRE, with an imperfect palindromic structure: GGTACAnnnTGTTCT (Beato 1989). This 15-mer is protected against DNaseI in the presence of the purified glucocorticoid receptor (Payvar et al 1983; Scheidereit et al 1983; Cato et al 1984; Karin et al 1984; Renkawitz et al 1984; Moore et al 1985; Slater et al 1985; DeFranco and Yamamoto 1986; Miksicek et al 1986; Danesch et al 1987; Hecht et al 1988) and two guanine residues in each strand have been shown to be protected against methylation by dimethyl sulfate after receptor binding (Scheidereit and Beato 1984). Similar experiments with the estrogen receptor and estrogen inducible genes, have led to the identification of a 13 bp palindrome GGTCAnnnTGACC as the estrogen responsive element, ERE (Klein-Hitpass et al 1986; Schip et al 1986; Seiler-Tuyns et al 1986; Martinez, Givel and Wahli 1987). In both, GRE and ERE, the right half of the palindrome is better conserved indicating that the binding site is asymmetric (Fig. 3.3). The main difference between the half-palindromes in the GRE and ERE resides in the third position as the fourth position is often a C in GRE (Karin et al 1984; Slater et al 1985), and the sixth position is not directly contacted by the glucocorticoid receptor (Truss, Chalepakis and Beato 1989). In the third position most known GREs exhibit a T and the 5'-methyl group of this thymine is contacted by the glucocorticoid receptor (Truss, Chalepakis and Beato 1989). In the ERE this third position is variable; in most cases an A is found, but a a C is observed (Slater et al 1989), and a G has also been described (Berry, Nunez and Chambon 1989). It has been reported that two nucleotide exchanges are sufficient to convert a GRE into an ERE (Klock, Strähle and Schütz 1987), but indeed exchanging the A at the third position of the ERE by a T converts it to a GRE (Truss, Chalepakis and Beato, unpublished). It is conceivable that the amino acids at the knuckle of the first zinc finger that distinguishes the glucocorticoid and the estrogen

receptors are responsible for the recognition of this difference in the third position of the half palindrome.



ERE consensus

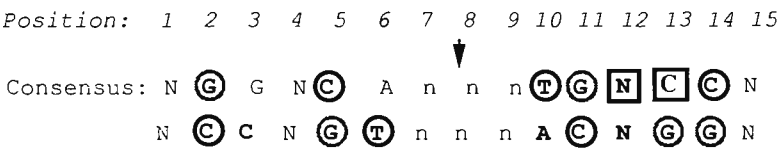


Figure 3.3. Consensus Sequence for the Glucocorticoid Responsive Element (GRE) and the Estrogen Responsive Element (ERE). The numbers underneath the GRE indicated the percentage conservation of the indicated base calculated out of 25 GREs.

The steroid binding domain.

The carboxy-terminal region of the nuclear receptors is also relatively well conserved and is responsible, among other things, for binding of the ligand (Weinberger et al 1985; Giguere et al 1986; Kumar et al 1987; Hollenberg and Evans 1988). In the unliganded receptor this domain appears to block transactivation, as deletion of the C-terminal half leads to constitutive activation of transcription (Godowski et al 1987). A few of the amino acid residues involved in binding of natural and synthetic ligands have been identified by mutational and chemical analysis (Carlstedt-Duke et al 1988; Smith et al 1988). In addition to its function in ligand binding, the carboxy terminal domain is also involved in other functions of the receptor such as nuclear localization, dimerization, binding to hsp90 and transactivation (Beato 1989).

Near the second zinc finger a short amino acid sequence, similar to that found in the T antigen of SV40 (Kalderon et al 1984a and b) appears to be responsible for intranuclear localization of the receptor even in the absence of ligand (Picard and Yamamoto 1987). Another, less well defined nuclear translocation signal, overlaps with the steroid-binding region and is dependent on binding of the hormone (Picard and Yamamoto 1987).

The palindromic nature of the hormone responsive elements already suggested that the receptors bind to DNA as homodimers (Scheidereit and Beato 1984). This idea has been confirmed both by protein biochemistry (Wrange) and in DNA binding studies (Kumar and Chambon 1988; Tsai et al 1988). As for the nuclear translocation function, it appears that there exists two different dimerization regions. One is weak, overlaps with the DNA-binding domain, and is independent of hormone binding , whereas the other is stronger, overlaps with the ligand binding domain and is dependent on binding of the agonist (Kumar and Chambon 1988). Though the ability of the different receptors to forms homodimers differs considerably, the same behavior has been found for estrogen, progesterone and glucocorticoid receptors.

In addition to its potential to dimerize, the steroid hormone receptors have a tendency to associate with the ubiquitous heat shock protein hsp90. This interaction apparently maintains the unliganded receptor in its inactive conformation, and binding of the hormone dissociates the receptor from hsp90 allowing dimerization and DNA binding (Sanchez et al 1987; Denis et al 1988; Howard and Distelhorst 1988; Joab et al 1988; Pratt et al 1988; Rexin, Busch and Gehring 1988). Close to the DNA binding domain, there is a region that has been postulated to be responsible for the interaction of the receptor with hsp90 (Pratt et al 1988), though it seems that other regions of the receptor, in particular the ligand binding domain, could also be involved in binding to this protein.

The carboxy terminal region contains also a transactivation function that is dependent on hormone binding and can be inhibited by antagonist ligands. This transactivation function appears to be distinct from other known motives from transcription factors, as

is not encoded in a single exon and seems to have a complex structure (Webster et al 1989). In addition to this hormone-dependent transactivation function there is also a constitutive transactivation activity that overlaps with the DNA-binding domain. This activity is usually masked by the steroid binding domain and is responsible for the fact that removal of the ligand binding region leads to a constitutively active receptor. A very small receptor fragment containing only the DNA-binding domain and a few flanking amino acid residues is able to activate a GRE-carrying promoter (Godowski et al 1987).

The amino terminal domain.

The most variable region among the nuclear receptors is the N-terminal domain preceding the DNA-binding domain. In some cases, like in the vitamin D receptor (McDonnell et al 1987; Baker et al 1988; Burmester, Maeda and DeLuca 1988), this region is virtually absent, whereas in other receptors, for instance the progesterone receptor (Loosfelt et al 1986; Misrahi et al 1987), the N-terminal region represents more than half the mass of the protein. The variability of this domain correlates with its antigenic potential. Most of the receptor antibodies described thus far are directed against the N-terminal half of the protein (Greene et al 1980; Carlstedt-Duke et al 1982; Moncharmont et al 1982; Westphal et al 1982; Logeat et al 1983; Gametchu and Harrison 1984; Milgrom 1985). The function of this domain is not well defined, though the existence of hormone-resistant variants of the glucocorticoid receptor lacking this part of the molecule, clearly underlines its significance in the intact cell (Mugele et al 1985).

In the case of the glucocorticoid receptor, a transactivation function has been mapped to the N-terminal domain (Giguère et al 1986; Godowski PJ, Picard D and Yamamoto 1988; Hollenberg and Evans 1988). Using the progesterone receptor, a relevance of the N-terminal domain for transactivation of certain promoters has been reported, whereas acting on other promoters deletion of the N-terminal region does not exhibit a detrimental phenotype (Tora et al 1988). Thus, it seems that this region of the receptor protein exhibits a more subtle function, that is only required in combination with particular promoters.

II. Regulation of MMTV-transcription.

II.1. DNA binding experiments

Induction of mouse mammary tumour virus (MMTV) transcription by glucocorticoid hormones has been widely used in the past to study gene regulation by hormones (Ringold 1979). Sequences relevant for hormonal induction have been localized to the long terminal repeat (LTR) region of the proviral DNA in gene transfer experiments (Buetti and Diggelmann 1981; Huang et al 1981; Hynes et al 1981; Lee et al 1981), more precisely between -200 and -50 upstream of the transcription start point (Hynes et al 1983; Payvar et al 1983). The same sequences have identified as binding sites for the glucocorticoid receptor in DNA-binding experiments in vitro (Payvar et al 1981; Geisse et al 1982; Pfahl 1982; Payvar et al 1983; Scheidereit et al 1983). In filter binding experiments with restriction fragments from MMTV-proviruses, a preferential binding to those fragments containing the LTR region was detected (Geisse et al 1982). Using 5'-deletion mutants two regions containing binding sites for the glucocorticoid receptor were identified, one located between -202 and -137 and the other between -137 and -50 (Figure 3.4A). These two binding regions coincided with sequence elements essential for mediating hormone inducibility of an adjacent promoter (Scheidereit et al 1983).

In a MMTV-LTR cloned from GR mice a strong DNaseI footprint between -192 and -163, and three shorter footprints between -124 and -71 are found in binding experiments with the purified glucocorticoid receptor from rat liver (Figure 3.4B. and Scheidereit et al 1983). All these four protected subregions contain the hexanucleotide motif 5'-TGTTCT-3' in the sense strand, but other similarities can be detected between the individual binding sites. In methylation protection studies, the G in the sense strand and the G in the antisense strand within the hexanucleotide motif, have been shown to be in intimate contact with the receptor protein (Figure 3.5A and Scheidereit and Beato 1984). In methylation interference experiments with the promoter

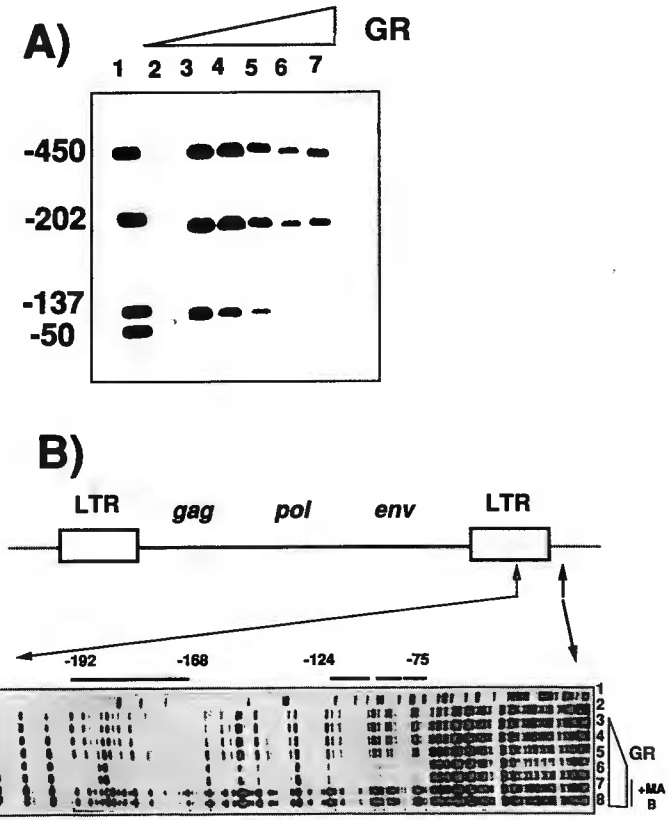


Figure 3.4. Binding of the Glucocorticoid Receptor to the MMTV-LTR.
A) Nitrocellulose filter binding assay with different 5'-deletion mutants. Lane 1, input DNA; Lane 2, control in the absence of receptor; lanes 3-7, 20 ng receptor in the presence of increasing concentrations of calf thymus DNA as competitor.
B) DNase I footprint. Lane 1, guanine-specific sequence reaction; lanes 2 to 6, increasing concentrations of receptor; lanes 7 and 8, same concentration of receptor as lane 6, but preadsorbed to antireceptor monoclonal antibody.

distal binding site of the LTR region, three of the four contacted G-residues yielded an interference signal, indicating that methylation at these positions prevented binding of the glucocorticoid receptor (Figure 3.5B). Within this strong binding site, between -192 and -163, and imperfect palindromic structure is detected and the symmetrically positioned G-residues are contacted by the receptor (Figure 3.6). Similar results were obtained in binding studies with the promoter region of the human metallothionein IIA gene, another glucocorticoid inducible gene (Karin et al 1984). These data suggested that in strong binding sites a dimer of the receptor in head-to-head orientation binds to the imperfect palindromic structure (Scheidereit et al 1986).

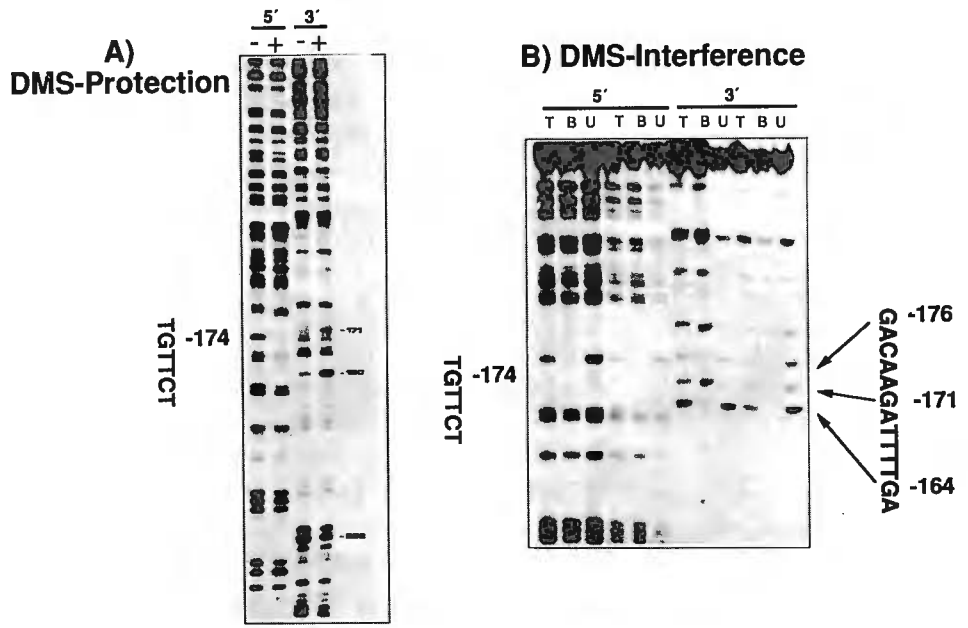


Figure 3.5. Protection and Interference with Dimethyl Sulfate.
A) Methylation protection by the glucocorticoid receptor, of the MMTV-LTR. **5′**: Labelling of the upper or sense strand. **3′**: labelling of the lower or antisense strand. **-** : methylation in the absence of receptor. **+** : methylation in the presence of receptor.
B) Interference by DMS methylation with binding of the glucocorticoid receptor. **T** : input DNA. **B** : DNA bound to nitrocellulose filter. **U** : free DNA. The positions of the relevant G-residues are indicated.,

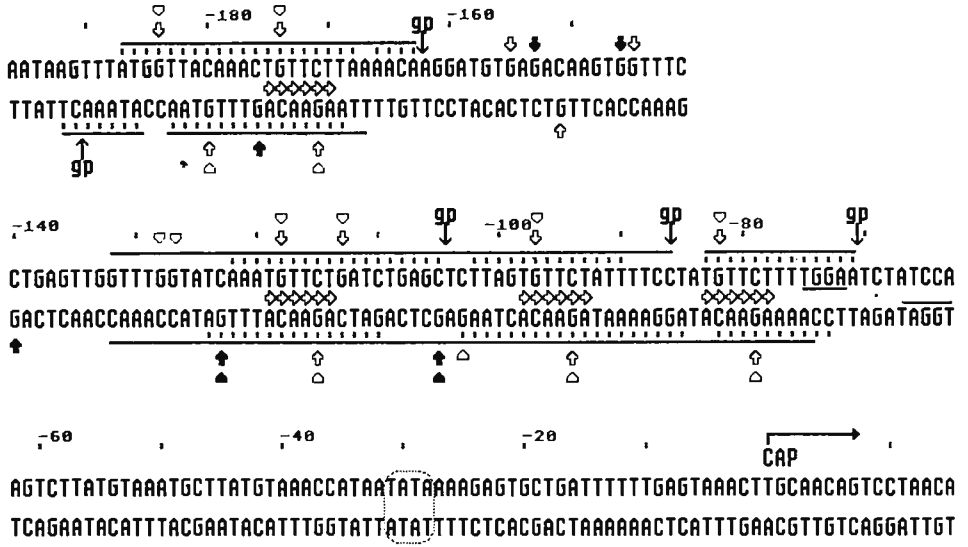


Figure 2.6. Nucleotide Sequence of the MMTV-LTR.
The nucleotide sequence of the promoter region of the MMTV-LTR is indicated. The numbers refer to distance from the initiation site (CAP). The DNaseI footprints generated by the glucocorticoid and the progesterone receptors are indicated by dotted and continuous lines respectively. The G-residues contacted by both receptors ares indicated by arrows (glucocorticoid) and arrow heads (progesterone). Also shown are the exonuclease III stops generated by both receptors.

We tested the possibility that the sequences recognized by the glucocorticoid receptor could also serve as binding sites for other steroid hormone receptors. Using progesterone receptor purified from rabbit uterus, we detected exonuclease III protection signals similar to those obtained with the glucocorticoid receptor (Figure 3.7 and von der Ahe et al 1985). The relevance of these DNA binding results was confirmed in gene transfer experiments, as the GRE of MMTV also mediates progesterone induction of an adjacent promoter (Cato et al 1986). These data demonstrated for the first time that a single DNA sequences could mediate induction by different DNA binding proteins, a concept that is now generally accepted. From these results the concept of a general hormone responsive element (HRE) was derived.

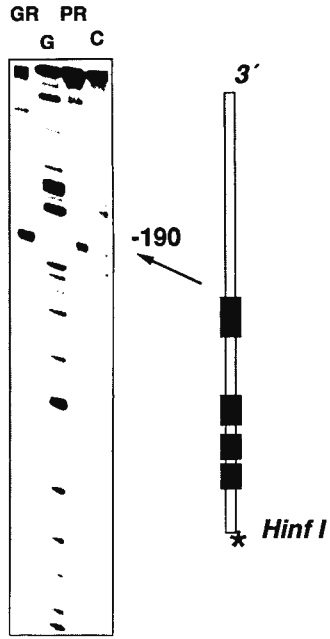


Figure 3.7. Exonuclease III Protection of the MMTV-LTR by glucocorticoid and Progesterone Receptor.
An MMTV-LTR fragment 5'-end labelled at the Hinf I site was subjected to exonuclease III digestion in the absence of protein (C) or in the presence of either glucocorticoid (GR) or progesterone receptor (PR). G : guanine-specific sequencing reaction.

More precise analysis using methylation protection and DNaseI footprinting revealed subtle differences in the way glucocorticoid and progesterone receptors recognize the DNA sequence of the HRE (von der Ahe et al 1986; Chalepakis et al 1988). In particular, the DNaseI footprint generated by the progesterone receptor in the promoter proximal region is larger and several additional G-residues are protected, that are not contacted by the glucocorticoid receptor. These differences in DNA binding are reflected in clear differences observed upon gene transfer with mutations that affect different regions of the HRE (Chalepakis et al 1988). In addition, the influence of DNA topology on the response of HRE-containing plasmids to glucocorticoids and progestins is markedly different, suggesting that both hormone receptors act by somehow different mechanism (Piña et al 1989). In both cases, however, a cooperative interaction between receptor molecules bound to the different sites of the HRE is needed for optimal hormonal response. Footprint experiments with hydroxyl radicals reveal contacts between both receptors and the DNA sequences flanking the DNaseI footprints, suggesting non-specific interactions of the protein with the DNA backbone outside of the conserved sequences (Chalepakis, Postma and Beato 1988).

Recently, we have found that the estrogen receptor is also able to bind to the MMTV-HRE, though no response to estrogens is observed in gene transfer experiments with the corresponding reporter plasmids (E.P. Slater and M. Beato unpublished). It seems that estrogen receptor molecules bound to the HRE are not able to undergo the cooperative interaction required for transcriptional activation. Understanding the difference between productive and non-productive receptor binding to DNA, may help to reveal the mechanism of transactivation.

II.2. Transcriptional activation: Role of nuclear factor I.

Adjacent to the promoter proximal binding site for the hormone receptors the MMTV-LTR contains a sequence recognized by the transcription factors NFI/CTF (Nowock et al 1985). In gene transfer experiments, this NFI binding sites appears to be essential for hormone-induced expression of the MMTV promoter (Miksicek, Borgmeyer and Nowock 1987). In cells carrying minichrosomes with the MMTV-LTR binding of NFI to its cognate sequence in the promoter is only observed after treatment with glucocorticoid hormones (Cordingley, Riedel and Hager 1986). These observations prompted us to analyze the influence of receptor binding to the HRE upon the interaction of NFI with its binding site in the MMTV promoter. In experiments using either NFI purified from pig liver (Meisterernst et al 1988a) or cloned NFI translated in vitro (Meisterernst et al 1988b), we have been unable to detect any cooperativity or synergism between hormone receptor and NFI in respect to DNA binding (U. Bruggemeier and M. Beato, unpublished). Thus, direct protein-protein interactions do not appear to play a key role in the facilitation of NFI binding to the MMTV promoter following hormone treatment.

Very recently we have developed a cell-free transcription assay in which initiation of transcription from the MMTV promoter is stimulated up to 10-fold by preincubation with partially purified progesterone receptor from rabbit uterus (Kalff, Gross and Beato 1989). In this system the transcription efficiency is highly dependent on NFI binding to its cognate sequence in the promoter, as mutation of the corresponding site decreases correct

initiation by at least 10-15 fold. However, these mutated templates can still respond to preincubation with the hormone receptor to the same extent as the wild type promoter (Kalff, Gross and Beato 1989). Therefore, the observed we detect in vitro is mediated by an interaction of the hormone receptors with components of the transcriptional machinery other than NFI.

II.3. Constitutive repression: Role of nucleosome phasing.

A peculiarity of the MMTV promoter is that it remains virtually silent in the absence of hormone and, therefore, the effect of hormone treatment on MMTV transcription is very dramatic. In the cell-free transcription system, however, the MMTV promoter is transcribed efficiently even in the absence of added hormone receptor. Thus, a repression mechanism operating in vivo can not be reproduced under the conditions of cell-free transcription. As the MMTV-LTR is precisely organized into nucleosomes (Richard-Foy and Hager 1987), this could explain the lack of activity of the MMTV-promoter in vivo. After hormone treatment the organization of MMTV into chromatin changes, in that a DNaseI hypersensitive region appears over the promoter region (Zaret and Yamamoto 1984; Cordingley, Riedel and Hager 1986). Again, these experiments suggest that changes in chromatin organization could explain the accessibility of the MMTV-promoter to NFI after treatment with hormone.

To explore these questions we have initiated nucleosome reconstitution experiments with relevant regions of the MMTV promoter. When incubated with H1-free polynucleosomes isolated from rat liver nuclei, these DNA fragments build mononucleosomes in which the DNA double helix follows a precise path on the surface of the histone octamer (Piña, Bruggemeier and Beato 1989). The path of the double helix, as determined by DNaseI digestion and hydroxyl radical attack, permits to predict the accessibility of individual sites for the corresponding proteins. According to these predictions, the binding sites for hormone receptors are largely accessible from the outside of the nucleosome, whereas the binding site for NFI exhibits its major groove masked towards the interior of the nucleosome (Piña, Bruggemeier and Beato 1989). In band-shift assay and footprint experiments with purified hormone

receptors and NFI these predictions were confirmed. Thus, the structure of a nucleosome on the MMTV promoter could be responsible for the observed transcriptional repression in the absence of hormone. According to this idea, binding of the hormone receptors to the nucleosomally organized MMTV promoter will initiate a series of structural changes leading to the exposure of the NFI binding site and to the formation of a stable transcription complex. A positive effect of the hormone receptor on the subsequent transcriptional activation seems probable in view of the results of cell-free transcription experiments. Thus, in the case of the MMTV promoter the effect of the hormone, mediated through the receptor, will be dual: on the one side inhibition of the nucleosome mediated repression, and on the other side activation of transcription through interaction with the transcriptional machinery. The generality of this dual function of the hormone, remains to be established.

IV. Conclusions and Outlook.

This brief description of the methodology employed for analyzing protein-DNA interactions, and the few examples extracted from the field of steroid hormone receptors, should have provided an impression of the great insight gained in our understanding of DNA-binding proteins during the last years. Nowadays, there is not a single issue of the leading journals of molecular biology that is not crowded with DNase I footprints, band retardations assays or methylation interference results. Through the use of these methods we have learned in great detail the mechanisms by which very different regulatory proteins recognize DNA sequences, and this knowledge is increasing at a very fast cadence. Though these methods have become part of the standard arsenal of molecular biologists, they are continuously being improved, refined, and adapted to always higher sensitivity. In particular, the routine availability of synthetic oligonucleotides and the development of the PCR-technology, have unfolded new possibilities in this field. This powerful methodology will certainly expand the use of in vivo DNA footprinting assays. In combination with fotocross-linking induced by nanosecond pulses of UV-lasers, this technique will allow to follow the time kinetics of the formation of protein-DNA complexes on promoter and enhancer DNA-sequences within the intact cells. We will also gain insight into the precise organization of nucleosomes in chromatin and will learn how these structural features modulate gene expression.

Ultimately, a precise understanding of the molecular mechanisms governing protein-DNA interactions, will required the crystallization of functional complexes of both macromolecules. This is still a gigantic task, that has only be solved for very few examples, mostly of prokaryotic origin. For very small proteins, or for small DNA-binding domains of larger proteins, 2D-NMR techniques, in combination with powerful computer modelling programs, will be of great help for understanding protein-DNA interactions in solution. But even if the progress in structural analysis speeds up during the next few years, we will still need dynamic methods that are able to provide information

about the situation in more complex biological systems. In particular, functional in vitro assays able to distinguish between productive and non-productive protein-DNA interactions are of enormous advantage. Progress in the development of cell-free transcription systems has been slow during the last decade, but is expected to be much faster once the relevant general transcription factors have been characterized and cloned. Thus, there is every reason to expect that the field of protein-DNA interaction will continue to expand during the last decade of the century and will sharpen our understanding of the molecular mechanisms underlaying modulation of gene expression and cell differentiation.

In closing this brief review I would like to come byk to the question formulated in the introudction, namely whether some kind of DNA regulatory code can be postulated in analogy to the genetic code. I think that this question can not be conclusively answered today. Although some general structural features of the DNA binding domains of regulatory proteins have been identified, and in particular the Á-helical nature of the recognition site, the molecular mechanism of nucleotide sequence discrimination is still not sufficiently clear. Before a defintive answer can be given, more precise structural information derived from the analysis of co-crystalls of protein and DNA is needed. However, the general impression is that nothing as "simple" as the genetic code will come out of these studies. After all, this should not be a surprise, as the aparent simplicity of the genetic code is only possible because of the adaptor role of tRNA and the precise selection of the correct amino acids by aminoacyl-tRNA synthetases. In dealing with the function of regulatory proteins we are at the very interface between the nucleic acid world and the protein world, and their interaction takes place without the help of a translator entity. Nevertheless, there is no reason to doubt that we will lern the answer to this question before the end of the mlilennium.

REFERENCES

Abel T and Maniatis T (1989)
Action of leucine zippers (News and Views)
Nature 341, 24-25

Alberts B and Herrick G (1970)
DNA-cellulose chromatography
Methods Enzymol 11, 198-217

Anderson WF, Ohlendorf DH, Takeda Y and Matthews BW (1989)
Structure of the cro repressor from bacteriophage lambda and its interaction with DNA.
Nature 290, 754-758

Angelov D, Stefanovsky VY, Dimitrov SI, Russanova VR, Keskinova E and Pashev IG (1988)
Protein-DNA crosslinking in reconstituted nucleohistones, nuclei and whole cells by picosecond UV laser irradiation.
Nucl Acids Res 16, 4525-4538

Baker AR, McDonnell DP, Hughes M, Crisp TM, Mangelsdorf DJ, Haussler MR, Pike JW, Shine J and O'Malley BW (1988)
Cloning and expression of full-length cDNA encoding human vitamin D receptor.
Proc Natl Acad Sci USA 85, 3294-3298

Bash P, Pattberaman N, Huang C, Ferrin TE and Langridge R (1983)
Van der Waal's surfaces in molecular modelling: Implementation with real time computer graphics.
Science 222, 1325-1327

Beato M (1989)
Gene regulation by steroid hormones.
Cell 56, 335-344

Becker PB, Gloss B, Schmid W, Strähle U and Schütz G (1986)
In vivo protein-DNA interactions in a glucocorticoid response element require the presence of the hormone.
Nature 324, 686-688

Becker MM and Wang J (1984)
Use of light for footprinting DNA in vivo.
Nature 309, 682-687

Berg JM (1988)
Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins.
Proc Natl Acad Sci USA 85, 99-102

Berg JM (1989)
DNA binding specificity of steroid receptors.
Cell 57, 1065-1068

Berg OG and von Hippel PH (1987)
Selection of DNA binding sites by regulatory proteins.
Statistical-mechanical theory and application to operators and promoters.
J Mol Biol 193, 723-750

Berry M, Nunez AM and Chambon P (1989)
Estrogen-responsive element of the human pS2 gene is an imperfect palindromic sequence.
Proc Natl Acad Sci USA 86, 1218-1222

Bowen B, Steinberg J, Laemmli UK and Weintraub H (1980)
The detection of DNA-binding proteins by protein blotting.
Nucl Acids Res 8, 1-20

Brennan RG and Matthews BW (1989)
The helix-turn-helix DNA binding motif.
J Biol Chem 264, 1903-1906

Brenowitz M, Senear DF, Shea MA and Ackers GK (1986)
"Footprint" titrations yield valid thermodynamic isotherms.
Proc Natl Acad Sci USA 83, 8462-8466

Briggs MR, Kadonaga JT, Bell SP and Tjian R (1986)
Purification and biochemical characterization of the promoter-specific transcription factor, Sp1.
Science 234, 47-52

Brunelle A and Schleif RF (1987)
Missing contact probing of DNA-protein interactions.
Proc Natl Acad Sci USA 84, 6673-6676

Buetti E and Diggelmann H (1981)
Cloned MMTV virus DNA is biologically active in transfected mouse cells and its expression is stimulated by glucocorticoid hormones.
Cell 23, 335-345

Burmester Jk, Maeda N and DeLuca HF (1988)
Isolation and expression of rat 1,25-dihydroxyvitamin D3 receptor cDNA.
Proc Natl Acad Sci USA 85, 1005-1009

Camier S, Gabrielsen O, Baker R and Sentenac A (1985)
A split binding site for the transcription factor Ó on the $tRNA_3^{Glu}$ gene.
EMBO J 4, 491-500

Carlstedt-Duke J, Okret S, Wrange Ö and Gustafsson JA (1982)
Immunological analysis of the glucocorticoid receptor: identification of a third domain separated from the steroid-binding and DNA binding domains.
Proc Natl Acad Sci USA 79, 4260-4264

Carlstedt-Duke J, Strömstedt PE, Wrange Ö, Berman T, Gustafsson JA and Jörnvall H (1987)
Domain structure of the glucocorticoid receptor protein.
Proc Natl Acad Sci USA 84, 4437-4440

Carlstedt-Duke J, Strömstedt PE, Persson B, Cederlund E, Gustafsson JA and Jörnvall H (1988)
Identification of hormone-interacting amino acid residues within the steroid-binding domain of the glucocorticoid receptor in relation to other steroid hormone receptors.
J Biol Chem 263, 6842-6846

Cartwright IL, Herzberg RP, Dervan PB and Elgin SCR (1983)
Cleavage of chromatin with methidium-propyl-EDTA-iron(II).
Proc Natl Acad Sci USA 80, 3213-3217

Cato ACB, Geisse S, Wenz M, Westphal HM and Beato M (1984)
The nucleotide sequences recognized by the glucocorticoid receptor in the rabbit uteroglobin gene region are located far upstream from the initiation of transcription.
EMBO J 3, 2731-2736

Cato ACB, Heitlinger E, Ponta H, Klein-Hitpass L, Ryffel GU, Bailly A, Rauch C and Milgrom E (1988)
Estrogen and progesterone receptor-binding sites on the chicken vitellogenin II gene: synergism of steroid hormone action.
Mol Cell Biol 8, 5323-5330

Cato ACB, Henderson D and Ponta H (1987)
The hormone response element of the mouse mammary tumour virus DNA mediates the progestin and androgen induction of transcription in the proviral long terminal repeat region.
EMBO J 6, 363-368

Cato ACB, Miksicek R, Schütz G, Arnemann J and Beato M (1986)
The hormone regulatory element of mouse mammary tumour virus mediates progesterone induction.
EMBO J 5, 2237-2240

Cato ACB , Skroch P, Weinmann J, Butkereitis P and Ponta H (1988)
DNA sequences outside the receptor-binding sites differentially modulate the responsiveness of the mouse mammary tumour virus promoter to various steroid hormones.
EMBO J 7, 1403-1410

Cato ACB and Weinmann J (1988)
Mineralocorticoid regulation of transfected mouse mammary tumour virus DNA in cultured kidney cells.
J Cell Biol 106, 2119-2125

Chalepakis G and Beato M (1989)
Hydroxyl radical interference: A new method for the study of protein-DNA interactions"
Nucl Acids Res 17, 1783

Chalepakis G, Postma JPM and Beato M (1988)
A model for hormone receptor binding to the mouse mammary tumour virus regulatory element based on hydroxyl radical footprinting.
Nucl Acids Res 16, 10237-10247

Chan PT and Lebowitz J (1983)
The coupled used of "footprinting" and exonuclease III methodology for RNA polymerase binding and initiation. Application for the analysis of three tandem promoters at the control region of colicin E1.
Nucl Acids Res 11, 1099-1116

Church GM and Gilbert W (11984)
Genomic footprinting
Proc Natl Acad Sci USA 81, 1991-1995

Clever U and Karlson P (1960)
Induktion von Puff-Veränderungen in Speicheldrüsenchromosomen von Chironomus tentans durch Ecdysone.
Exp Cell Res 20, 623-626

Conneely OM, Dobson ADW, Tsai MJ, Beattie WG, Toft DO, Huckaby CS, Zarucki T, Schrader WT and O'Malley BW (1987)
Sequence and expression of a functional chicken progesterone receptor.
Mol Endo 1, 517-525

Cordingley MG, Riegel AT and Hager GL (1987)
Steroid-dependent interaction of transcription factors with the inducible promoter of mouse mammary tumour virus in vivo.
Cell 48, 261-270

Cote GJ, Wang Z and Chiu J-F (1985),
Á-Fetoprotein gene DNA-binding proteins.
Arch. Biochem. Biophys. 243, 320-324

Curran T and Franza BR Jr (1988)
Fos and jun: the AP-1 connection.
Cell 55, 395-397

Danesch U, Gloss B, Schmid W, Schütz G and Renkawitz R (1987)
Glucocorticoid induction of the rat tryptophan oxygenase gene is mediated by two widely separated glucocorticoid-responsive elements.
EMBO J 6, 625-630

Danielsen M, Hinck L and Ringold GM (1989)
Two amino acids within the knuckle of the first zinc finger specify DNA response element activation by the glucocorticoid receptor.
Cell 57, 1131-1138

DeFranco D and Yamamoto KR (1986)
Two different factors act separately or together to specify functionally distinct activities at a single transcriptional enhancer.
Mol Cell Biol 6, 993-1001

Denis M, Poellinger L, Wikström AC and Gustafsson JA (1988)
Requirement of hormone for thermal activation of the glucocorticoid receptor to a DNA-binding state.
Nature 333, 686-688

Drew HR and Travers AA (1985)
DNA bending and its relation to nucleosome positioning
J Mol Biol 186, 773-790

Evans RK and Haley BE (1987)
Synthesis and biological properties of 5-azido-2'-deoxyuridine 5'-triphosphate, a photoactive nucleotide suitable for maiing light-sensitive DNA
Biochemistry 26, 269-276

Evans RM and Hollenberg SM (1988)
Zinc fingers: gilt by association.
Cell 52, 1-3

Fainsod A, Bogarad LD, Ruusala T, Lubin M, Crothers DM and Ruddle FH (1986)
The homeo domain of a murine protein binds 5' to its own homeo box.
Proc Natl Acad Sci USA 83, 9532-9536

Freedman LP, Luisi BF, Korszun ZR, Basavappa R, Sigler PB and Yamamoto KR (1988)
The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain.
Nature 334, 543-546

Fried M and Crothers DM (1981)
Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis
Nucl Acids Res 9, 6505-6525

Furlong JC and Lilley DMJ (1986)
Highly selective modification of cruciform loops by diethyl pyrocarbonate.
Nucleic Acid Res 14, 3995-4007

Galas DJ and Schmiz A (1978)
DNase footprint: A simple method for the detection of Protein-DNA binding specificity.
Nucl Acids Res 5, 3157-3170

Gametchu B and Harrison RW (1984)
Characterization of a monoclonal antibody to the rat liver glucocorticoid receptor.
Endocrinology 114, 274-279

Garner MM and Revzin A (1981)
Gel electrophoretic method for quantifying the binding of proteins to specific DNA regions. Application to components of the E. coli lactose operon regulatory system.
Nucl Acids Res 9, 3047-3060

Geisse S, Scheidereit C, Westphal HM, Hynes NE, Groner B and Beato M (1982)
Glucocorticoid receptors recognize DNA sequences in and around murine mammary tumour virus DNA.
EMBO J 1, 1613-1619

Giguere V, Hollenberg SM, Rosenfeld MG and Evans RM (1986)
Functional domains of the human glucocorticoid receptor.
Cell 46, 645-652

Gilmour DS and Lis JT (1984)
Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes.
Proc Natl Acad Sci USA 81, 4275-4279

Glass CK, Franco R, Weinberger C, Albert VR, Evans RM and Rosenfeld MG (1987)
A c-erb-A binding site in rat growth hormone gene mediates trans-activation by thyroid hormone
Nature 329, 738-741

Godowski PJ, Picard D and Yamamoto KR (1988)
Signal transduction and transcriptional regulation by glucocorticoid receptor-LexA fusion proteins.
Science 241, 812-816

Godowski PJ, Rusconi S, Miesfeld R and Yamamoto KR (1987)
Glucocorticoid receptor mutants that are constitutive activators of transcriptional enhancement.
Nature 325, 365-368

Gralla JD
Rapid "footprinting" on supercoiled DNA
Proc Natl Acad Sci USA 82, 3078-3081

Green S and Chambon P (1986),
Carcinogenesis: A superfamily of potentially oncogenic hormone receptors.
Nature 324, 615-617

Green S, Kumar V, Theulaz I, Wahli W and Chambon P (1988)
The N-terminal DNA-binding "zinc finger" of the oestrogen and glucocorticoid receptors determines target gene specificity.
EMBO J 7, 3037-3044

Greene GL, Fitch FW and Jensen EV (1980)
Monoclonal antibodies to estrophilin: Probes for the study of estrogen receptor.
Proc Natl Acad Sci USA 77, 157-161

Greene GL, Gilna P, Waterfield M, Baker A, Hort Y and Shine J (1986)
Sequence and expression of human estrogen receptor complementary DNA.
Science 231, 1150-1154

Gronemeyer H, Turcotte B, Quirin-Stricker C, Bocquel MT, Meyer ME, Krozowski Z, Jeltsch JM, Lerouge JM, Garnier JM and Chambon P (1987)
The chicken progesterone receptor: sequence, expression and functional analysis.
EMBO J 6, 3985-3994

Harel-Bellan A, Brini AT, Ferris DK, Robin P and Farrar WL (1989)
In situ detection of a heat-shock regulatory element binding protein using a short synthetic enhancer sequence.
Nucl Acids Res 17, 4077-4087

Hecht A, Berkenstam A, Strömstedt PE, Gustafsson JA and Sippel AE (1988)
A progesterone responsive element maps to the far upstream steroid dependent DNase hypersensitive site of chicken lysozyme chromatin.
EMBO J 7, 2063-2073

Hendrickson A and Schleif RF (1984)
Regulation of the E. coli L-arabinose operon studied by gel electrophoresis DNA binding assay
J Mol Biol 174, 611-628

Herr W (1985)
Diethyl pyrocarbonate: a chemical probe for secondary structure in negatively supercoiled DNA.
Proc. Natl. Acad. Sci. U.S.A. 82, 8009-8013

Herrick G (1980)
Site-specific DNA-affinity chromatography of the lac repressor.
Nucl Acids Res 3721-3728

Hertzberg RP and Dervan PB (1984)
Cleavage of DNA with methidiumpropyl-EDTA-Iron(II): Reaction conditions and product analysis.
Biochemistry 23, 3934-3945

Hockensmith JW, Kubasek WL, Vorachek WR and von Hippel PH (1986)
Laser cross-linking of nucleic acids to proteins. Methodology and first applications to the phage T4 DNA replication system
J Biol Chem 261, 3512-3518

Hollenberg SM and Evans RM (1988)
Multiple and cooperative trans-activation domains of the human glucocorticoid receptor.
Cell 55, 899-906

Hollenberg SM, Giguere V, Segui P and Evans RM (1987)
Colocalization of DNA-binding and transcriptional activation functions in the human glucocorticoid receptor.
Cell 49, 39-46

Hollenberg SM, Weinberger C, Ong ES, Cerelli G, Oro A, Lebo R, Thompson EB, Rosenfeld MG and Evans RM (1985)
Primary structure and expression of a functional human glucocorticoid receptor cDNA.
Nature 318, 635-641

Hope IA and Struhl K (1985)
GCN4 protein, synthesized in vitro, binds HIS3 regulatory sequences: implications for general control of amino acid biosynthetic genes in yeast.
Cell 43, 177-188

Hope IA and Struhl K (1986)
Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast.
Cell 46, 885-894

Howard KJ and Distelhorst CW (1988)
Evidence for intracellular association of the glucocorticoid receptor with the 90 kD heat shock protein.
J Biol Chem 263, 3474-3481

Howgate P, Jones AS and Tittensor JR (1988)
The permanganate oxydation of thymidine.
J Chem Soc (C) 275-279

Huang AL, Ostrowski MC, Berard D and Hager G (1981)
Glucocorticoid regulation of the Ha-MuSV p21 gene conferred by sequences from mouse mammary tumour virus.
Cell 27, 245-255

Hübscher, U. (1987)
Double replica southwestern.
Nuleic Acids Res. 15, 5486

Hynes NE, Kennedy N, Rahmsdorf U and Groner B (1981)
Hormone responsive expression of an endogenous proviral gene of MMTV after molecular cloning in gene transfer into cultured cells.
Proc Natl Acad Sci USA 78, 2038-2042

Hynes NE, van Ooyen AJJ, Kennedy N, Herrlich P, Ponta H and Groner B (1983)
Subfragments of the large terminal repeat cause glucocorticoid responsive expression of mouse mammary tumour virus and of an adjacent gene.
Proc Natl Acad Sci USA 80, 3637-3641

Ivarie R (1987)
Thymin methyls and DNA-protein interactions.
Nucleic Acid Res 15, 9975-9983

Jackson PD and Felsenfeld G (1985)
A method for mapping intranuclear protein-DNA interactions and its application to a nuclease hypersensitive site
Proc Natl Acad Sci USA 82, 2296-2300

Jantzen C, Fritton HP, Igo-Kemenes T, Espel E, Janich S, Cato ACB, Mugele K and Beato M (1987)
Partial overlapping of binding sequences for steroid hormone receptors and DNaseI hypersensitive sites in the rabbit uteroglobin gene region.
Nucleic Acid Res 15, 4535-4552

Jeltsch JM, Krozowski Z, Quirin-Stricker C, Gronemeyer H, Simpson RJ, Garnier JM, Krust A, Jacob F and Chambon P (1986)
Cloning of the chicken progesterone receptor.
Proc Natl Acad Sci USA 83, 5424-5428

Joab I, Radanyi C, Renoir M, Buchou T, Catelli MG, Binart N, Mester J and Baulieu EE (1988)
Common non-hormone binding component in the non-transformed chick oviduct receptors of four steroid hormones.
Nature 308, 850-853

Johnston BH and Rich A (1985)
Chemical probes of DNA conformation: detection of Z-DNA at nucleotide resolution.
Cell 42, 713-724

Jones TA (1978)
A graphics model building and refinement system for macromolecules.
J Appl Cryst 11, 268-272

Jones KA, Kadonaga JT, Rosenfeld PJ, Kelly TJ and Tjian R (1987)
A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication.
Cell 48, 79-89

Kadonaga JT, Carner KR, Masiarz FR and Tjian R (1987)
Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain.
Cell 51, 1079-1090

Kalderon D, Roberts BL, Richardson WD and Smith AE (1984)a
A short amino acid sequence able to specify nuclear location.
Cell 39, 499-509

Kalderon D, Richardson WD, Markham AF and Smith AE (1984)b
Sequence requirements for nuclear localization of SV40 large T.
Nature 311, 33-38

Kalff M, Gross B and Beato M (1989)
Progesterone receptor activates cell-free transcription at the MMTV promoter.
Nature

Karin M, Haslinger A, Holtgreve A, Richards RI, Krauter P, Westphal HM and Beato M (1984)
Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionein IIA.
Nature 308, 513-519

Karlson P (1961)
Biochemische Wirkungsweise der Hormone.
Dtsch med Wschr 86, 668-671

Klein-Hitpass L, Schorp M, Wagner U and Ryffell GU (1986)
An estrogen-responsive element derived from the 5'-flanking region of Xenopus vitellogenin A2 gene functions in transfected human cells.
Cell 46, 1053-1061

Klock G, Strähle U and Schütz G (1987)
Oestrogen and glucocorticoid responsive elements are closely related but distinct.
Nature 329, 734-736

Klug A and Rhodes d (1987)
"Zinc fingers": a novel protein motif for nuclei acid recognition
TIBS 12, 464-469

Kristie TM and Roizman B (1986)
Á-4, the major regulatory protein of herpes simplex virus type 1, is stably and specifically associated with promoter-regulatory domains of a genes and/or selected viral genes.
Proc Natl Acad Sci USA 83, 3218-3322

Krust A, Green S, Argos P, Kumar V, Walter P, Bornert J-M and Chambon P (1986)
The chicken estrogen receptor sequence: homology with v-erbA and the human estrogen and glucocorticoid receptors.
EMBO J 5, 891-897

Kumar V, Green S, Stack G, Berry M, Jin JR and Chambon P (1987)
Functional domains of the human estrogen receptor.
Cell 51, 941-951

Kumar V and Chambon P (1988)
The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer.
Cell 55, 145-156

Kumar V, Green S, Staub A and Chambon P (1986)
Localization of the oestradiol-binding and putative DNA-binding domains of the human oestrogen receptor.
EMBO J 5, 2231-2236

Krust A, Green S, Argos P, Kumar V, Walter P, Bornert J-M and Chambon P (1986)
The chicken estrogen receptor sequence: homology with v-erbA and the human estrogen and glucocorticoid receptors.
EMBO J 5, 891-897

Langridge R, Ferrin TE, Kuntz ID and Connolly ML (1981)
Real time colour graphics in studies of molecular interactions
Science 211, 661-666

Lee W, Mitchel P and Tjian R (1987)
Purified transcription factor AP-1 interacts with TPA-inducible enhancer element.
Cell 49, 741-752

Lee F, Mulligan R, Berg O and Ringold G (1981)
Glucocorticoid regulated expression of dihydrofolate reductase cDNA in mouse mammary tumour virus chimeric plasmids.
Nature 294, 228-232

Linxweiler W and Hörz W (1984)
Reconstitution of mononucleosomes: characterization of distinct particles that differ in the position of the histone core.
Nucl Acids Res 12, 9395-9413

Linxweiler W and Hörz W (1985)
Reconstitution experiments show that sequence-specific histone-DNA interactions are the basis for nucleosome phasing on mouse satellite DNA.
Cell 42, 281-290

Liu-Johnson HM, Gartenberg MR and Crothers DM (1986)
The DNA bending domain and bend angle of E. coli CAP protein.
Cell 47, 995-1005

Logeat F, Vu Hai MT, Fournier A, Legrain P, Buttin G and Milgrom E (1983)
Monoclonal antibodies to rabbit progesterone receptor: cross reaction with other mammalian progesterone receptors.
Proc Natl Acad Sci USA 80, 6456-6459

Mader S, Kumar V, Verneuil H de and Chambon P (1989)
Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid-responsive element.
Nature 338, 271-274

Martinez E, Givel F and Wahli W (1987)
The estrogen-responsive element as an inducible enhancer: DNA sequence requirements and conversion to a glucocorticoid responsive element.
EMBO J 6, 3719-3727

Matsuno K, Suzuki T, Takiya S and Suzuki Y (1989)
Complex formation with the fibroin gene enhancer through a protein-protein interaction analyzed by a modified DNA-binding assay.
J Biol Chem 264, 4599-4604

Matthews BW (1988)
Protein-DNA interaction. No code for recognition.
Nature 335, 294-295

Maxam A and Gilbert W (1980)
Sequencing end-labelled DNA
Methods Enzymol 65, 499-560

McDonnell DP, Mangelsdorf DJ, Pike JW, Haussler MR and O'Malley BW (1987)
Molecular cloning of complementary DNA encoding the avian receptor for vitamin D.
Science 235, 1214-1217

McKay DB and Steitz T (1981)
Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA.
Nature 290, 744-749

McNight S and Tjian R (1986)
Transcriptional selectivity of viral genes in mammalian cells.
Cell 46, 795-805

Meisterernst M, Gander I, Rogge L and Winnacker EL (1988)
A quantitative analysis of nuclear factorI/DNA interactions.
Nucleic Acids Res 16, 4419-4435

Meisterernst M, Rogge L, Donath C, Gander I, Lottspeich F, Mertz R, Dobner T, Föckler R, Stelzer G and Winnacker EL (1988)
Isolation and characterization of the porcine nuclear factor I (NFI) gene.
FEBS Letters 236, 27-32

Miesfeld R, Okret S, Wikström A-C, Wrange Ö, Gustafsson J-A and Yamamoto KR (1984)
Characterization of a steroid hormone receptor gene and mRNA in wild-type and mutant cells.
Nature 312, 779-781

Miesfeld R, Rusconi S, Godowski PJ, Maler BA, Okret S, Wikström AC, Gustafsson JA and Yamamoto KR (1986)
Genetic complementation of a glucocorticoid receptor deficiency by expression of cloned receptor cDNA.
Cell 46, 389-399

Miksicek R, Borgmeyer U and Nowock J (1987)
Interaction of the TGGCA-binding protein with upstream sequences is required for efficient transcription of mouse mammary tumor virus.
EMBO J 6, 1355-1360

Miksicek R, Heber A, Schmid W, Danesch U, Posseckert G, Beato M and Schütz G (1986)
Glucocorticoid responsiveness of the transcriptional enhancer of Moloney murine sarcoma virus.
Cell 46, 283-290

Milgrom E (1985)
Monoclonal antibodies to steroid hormone receptors.
J Steroid Biochem 28, 389-415

Miller J, McLachlan AD and Klug A (1985),
Repetitive zinc-binding domains in the protein transcription factorIIIA from Xenopus oocytes.
EMBO J 4, 1609-1614

Miskimins WK, Roberts MP, McClelland A and Ruddle FH (1985)
Use of protein blotting procedure and a specific DNA probe to identify nuclear proteins that recognize the promoter region of the transferring receptor gene.
Proc. Natl. Acad. Sci. USA 82, 6741-6744

Misrahi M, Atger M, d'Aurio L, Loosfelt H, Meriel C, Fridlansky F, Guichon-Mantel A, Galibert F and Milgrom E (1987)

Complete amino acid sequence of the human progesterone receptor deduced from cloned cDNA.
Biochem Biophys Res Comm 143, 740-748

Moore DD, Marks AR, Buckley DI, Kapler G, Payvar F and Goodman HM (1985)
The first intron of the human growth hormone gene contains a binding site for glucocorticoid receptor.
Proc Natl Acad Sci USA 82, 699-702

Murre C, McCaw PS and Baltimore D (1989)
A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD1, and myc proteins.
Cell 56, 777-783

Murre C, Schonleber McCaw P, Vaessin H, Caudy M, Jan LY, Jan YN, Cabrera CV, Buskin JN, Hauschka D, Lassar AB, Weintraub H and Baltimore D (1989),
Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence.
Cell 58, 537-544

Nedospasov SA and Georgiev GP (1980)
Non-random cleavage of SV40 DNA in the compact minichromosomes and free in solution by micrococcal nuclease.
Biochem Biophys Res Comm 92, 532-539

Nick H and Gilbert W (1985)
Detection in vivo of protein-DNA interactions with the lac operon of E. coli.
Nature 313, 795-798

Nordheim A and Meese K (1988)
Topoisomer gel retardation: detection of anti-Z-DNA antibodies bound to Z-DNA within supercoiled minicircles.
Nucleic Acid Res 16, 21-37

Nowock J, Borgmeyer U, Püschel A, Rupp AW and Sippel AE (1985)
The TGGCA protein binds to the MMTV-LTR, the adenovirus origin of replication, and BK virus.
Nucl Acid Res 13, 2045-2062

Ogata RT and Gilbert W (1977)
Contacts between the lac repressor and thymines in the lac operator.
Proc Natl Acad Sci USA 74, 4973-4976

Ogata RT and Gilbert W (1978)
An amino-terminal fragment of the lac-repressor binds specifically to the lac operator.
Proc Natl Acad Sci USA 75, 5851-5854

Oliphant AR, Brandl CJ and Struhl K (1989)
Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: Analysis of yeast GCN4 protein
Mol Cell Biol 9, 2944-2949

Otting G, Qian Y, Müller M, Affolter M, Gehring W and Wütrich K (1988)
Secondary structure determination for the Antennapedia homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif.
EMBO J 7, 4305-4309

Payvar F, DeFranco D, Firestone GL, Edgar B, Wrange Ö, Okret S, Gustafsson JA and Yamamoto KR (1983)
Sequence-specific binding of glucocorticoid receptor to MTV DNA at sites within and upstream of the transcribed region.
Cell 35, 381-392

Payvar F, Wrange Ö, Carlstedt-Duke J, Okret S, Gustafsson JA and Yamamoto KR (1981)
Purified glucocorticoid receptors binds selectively in vitro to a cloned fragment whose transcription is regulated by glucocorticoids in vivo.
Proc Natl Acad Sci USA 78, 6628-6632

Pfahl M (1982)
Specific binding of the glucocorticoid-receptor complex to the mouse mammary tumor proviral promoter region.
Cell 31, 475-482

Picard D, Salser SJ and Yamamoto KR (1988)
A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor.
Cell 54, 1073-1080

Picard D and Yamamoto KR (1987)
Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor.
EMBO J 6, 3333-3340

Piña B, Haché RJG, Arnemann J, Chalepakis G, Slater EP and Beato M (1989)
Hormonal induction of transfected genes depends on DNA topology
Mol Cell Bio, in press

Piña B, Bruggemeier U and Beato M (1989)
Nucleosome positioning modulates accessibility of regulatory proteins to the MMTV promoter.
Cell

Pratt WB, Jolly DJ, Pratt DV, Hollenberg SM, Giguere V, Cadepond FM, Schweizer-Groyer G, Catelli MG, Evans RM and Baulieu EE (1988)
A region in the steroid binding domain determines formation of the non-DNA binding, 9S glucocorticoid receptor complex.
J Biol Chem 263,

Renkawitz R, Schütz G, Ahe D von der and Beato M (1984)
Identification of hormone regulatory elements in the promoter region of the chicken lysozyme gene.
Cell 37, 503-510

Revzin MT and von Hippel PH (1977)
Direct measurement of association constants for the binding of E. coli lac repressor to non-operator DNA.
Biochemistry 16, 4769-4776

Rexin M, Busch W and Gehring U (1988)
Chemical cross-linking of heteromeric glucocorticoid receptors.
Biochemistry 27, 5593-5601

Richard-Foy H and Hager GL (1987)
Sequence specific positioning of nucleosomes over the steroid-inducible MMTV promoter.
EMBO J 6, 2321-2328

Richardson CC, Lehman IR and Kornberg A (1964)
A DNA phosphatase-exonuclease from E. coli II. Characterization of the exonuclease activity.
J Biol Chem 239, 251-258

Riggs AD, Bourgeois S and Cohn M (1970)
The lac repressor-operator interaction III. Kinetic studies.
J Mol Biol 53, 401-417

Riggs AD, Suzuki H and Bourgeois S (1970)
The lac-repressor-operator interactions. I Equilibrium studies
J Mol Biol 48, 67-83

Ringold G (1979)
Glucocorticoid regulation of mouse mammary virus gene expression
Biochim biophys Acta 560, 487-508

Ross W, Landy A, Kikuchi Y and Nash HA (1979)
Interaction of int protein with specific sites on lambda PPatt DNA.
Cell 18, 297-307

Roux E, Strubin M, Hagenbüchle O and Wellauer PK (1989)
The cell-specific transcription factor PTF1 contains two different subunits that interact with DNA
Genes & Develop 3, 1616-1624

Rubin CM and Schmid CW (1980)
Pyrimidine-specific chemical reactions useful for DNA-sequencing.
Nucl Acids Res 8, 4613-4619

Saluz H and Jost JP (1989)
A simple high-resolution procedure to study DNA methylation and in vivo DNA-protein interactions on a single-copy gene level in higher eukaryotes.
Proc Natl Acad Sci USA 86, 2602-2606

Sanchez ER, Meschinchi S, Schlesinger MJ and Pratt WB (1987)
Demonstration that the 90-kilodalton heat shock protein is bound to the glucocorticoid receptor in its 9S non-deoxyribonucleic acid binding form.
Mol Endocrin 1, 908-912

Schauer M, Chalepkais G, Willmann T and Beato M (1989)
Binding of hormone accelerates the kinetics of glucocorticoid and progesterone receptor binding to DNA.
Proc Natl Acad Sci USA 86, 1123-1127

Scheidereit C and Beato M (1984)
Contacts between receptor and DNA double helix within a glucocorticoid regulatory element of mouse mammary tumor virus
Proc Natl Acad Sci USA 81, 3029-3034

Scheidereit C and Beato M (1987)
Steroid receptor binding to DNA sequences.
In: "Steroid Hormones. A Practical Approach" (ed B Green and RE Leake) IRL Press Oxford, 179-204

Scheidereit C, Geisse S, Westphal HM and Beato M (1983)
The glucocorticoid receptor binds to defined nucleotide sequences near the promoter of mouse mammary tumour.
Nature 304, 749-752

Scheidereit C, Westphal HM, Carlson C, Bosshard H and Beato M (1986)
Molecular model of the interaction between the glucocorticoid receptor and regulatory elements of inducible genes
DNA 5, 383-391

Schip F van het, Strijker R, Samallo J, Gruber M and Geert AB (1986)
Conserved sequence motifs upstream from the co-ordinately expressed vitellogenin and apoVLDLII genes of chicken.
Nucl Acid Res 14, 8669-8680

Scholten PM and Nordheim A (1986)
Diethyl pyrocarbonate: a chemical probe for DNA cruciforms.
Nuclei Acid Res 14, 3981-3993

Seeman MC, Rosenberg JM and Rich A (1976)
Sequence-specific recognition of double helical nucleic acids by proteins.
Proc Natl Acad Sci USA 73, 804-808

Seiler-Tuyns A, Walker P, Martinez E, Merillat A-M, Givel F and Wahli W (1986)
Identification of estrogen-responsive DNA sequences by transient expression experiments in a human breast cancer cell line.
Nucl Acid Res 14, 8755-8770

Selleck SB and Majors J (1987)
Photofootprinting in vivo detects transcription-dependent changes in yeast TATA boxes.
Nature 325, 173-177

Severne Y, Wieland S, Schaffner W and Rusconi S (1988)
Metal binding "finger" structures in the glucocorticoid receptor defined by site-directed mutagenesis.
EMBO J 7, 2503-2508

Sawadogo M and Roeder RG (1980)
Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region.
Cell 43, 165-175

Shalloway D, Kleinberger T and Livingstone DM (1980)
Mapping of SV40 DNA replication origin region binding sites for the SV40 T antigen by protection against exonuclease III digestion.
Cell 20, 411-422.

Siebenlist U and Gilbert W (1980)
Contacts between Escherichia coli RNA polymerase and an early promoter of phage T7
Proc Natl Acad Sci USA 77, 122-126

Siebenlist U, Simpson RB and Gilbert W (1980)
E. coli RNA polymerase interacts homologously with two different promoters.
Cell 20, 269-281

Silva CM, Tully DB, Petch LA, Jewel CM and Cidlowski JA (1987)
Application of a protein-blotting procedure to the study of human glucocorticoid receptor interactions with DNA.
Proc Natl Acad Sci USA 84, 1744-1748

Singh H, LeBowitz JH, Baldwin AS and Sharp PA (1988)
Molecular cloning of and enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA.
Cell 52, 425-423

Slater EP, Cato ACB, Karin M, Baxter JD and Beato M (1988)
Progesterone induction of metallothionein-IIA gene expression.
Mol Endocrin 2, 485-491

Slater EP, Mugele K, Redeuihl G, Wiehle RD, Espel E, Theis K, Suske G and Beato M.(1989)
The uteroglobin promoter contains a non-canonical estrogen reponsive element
Mol Endocrin, in press

Slater E, Rabenau O, Karin M, Baxter JD and Beato M (1985),
Glucocorticoid receptor binding an activation of a heterologous promoter in response to dexamethasone by the first intron of the human growth hormone gene.
Mol Cell Biol 5, 2984-2992

Staudt LM, Clerc RG, Singh H, LeBowitz JH, Sharp PA and Baltimore D (1988)
Molecular cloning of lymphoid-specific cDNA encoding a protein that binds to the regulatory octamer DNA motif.
Science

Strauss F and Varshavsky A (1984)
A protein binds to a satellite DNA repeat at three specific sites that would brought into mutual proximity by DNA folding in the nucleosome.
Cell 37, 889-901

Struhl K (1989)
Helix-turn-helix, zinc-finger, and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins.
TIBS 14, 137-140

Suske G, Gross B and Beato M (1989)
Non-radioactive method to visualize specific DNA-protein interactions in the band shift assay.
Nucle Acids Res 17, 4405

Topol J, Ruden DM and Parker CS (1985)
Sequences required for in vitro transcriptional activation of a Drosophila hsp70 gene.
Cell 42, 527-537

Tora L, Gronemeyer H, Turcotte B, Gaub MP and Chambon P (1988)
The N-terminal region of the chicken progesterone receptor specifies target gene activation.
Nature 333, 185-188

Truss M, Piña B and Beato M (1989)
DNA topology influences binding of steroid hormone receptors to the hormone responsive region of the mouse mammary tumor promoter.
In press

Truss M, Chalepkais G and Beato M (1989)
A method for detecting contacts between regulatory proteins and thymines in DNA illustrated with steroid hormone receptors and their responsive elements.
In press

Tsai SY, Carlstedt-Duke J, Weigel NL, Dahlman K, Gustafsson JA, Tsai MJ and O'Malley BW (1988)
Molecular interactions of steroid hormone receptor with its enhancer element: evidence for receptor dimer formation.
Cell 55, 361-369

Tullius TD and Dombroski BA (1986)
Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to lambda repressor and cro protein.
Proc Natl Acad Sci USA 83, 5496-5473

Umesono K and Evans RM (1989)
Determinants of target gene specificity for steroid/thyroid hormone receptors.
Cell 57, 1139-1146

von der Ahe D , Janich S, Scheidereit C, Renkawitz R, Schütz G and Beato M (1985)
Glucocorticoid and progesterone receptors bind to the same sites in two hormonally regulated promoters.
Nature 313, 706-709

von der Ahe D , Renoir JM, Buchou T, Baulieu EE and Beato M (1986)
Receptors for glucocorticosteroid and progesterone recognize distinct features of a DNA regulatory element.
Proc Natl Acad Sci USA 83, 2817-2821

von Hippel PH (1979)
On the molecular bases of the specificity of the interaction of proteins with genome DNA.
In: "Biological Regulation and Development" (ed. RF Goldberger), Plenum, New York, 279-347

von Hippel PH and Berg OG (1986)
On the specificity of DNA-protein interactions.
Proc Natl Acad Sci USA 83, 1608-1612

Walter P, Green S, Greene G, Krust A, Bornert J-M, Jeltsch J-M, Staub A, Jensen E, Scrace G, Waterfield M and Chambon P (1985)
Cloning of the human estrogen receptor cDNA.
Proc Natl Acad Sci USA 82, 7889-7893

Webster NJG, Green S, Tasset D, Ponglikitmongkol M and Chambon P (1989)
The transcriptional activation function located in the hormone-binding domain of the human estrogen receptor is not encoded in a single exon.
EMBO J 8, 1441-1446

Weinberger C, Hollenberg SM, Rosenfeld MG and Evans RM (1985)
Domain structure of human glucocorticoid receptor and ist relationship to the v-erbA oncogene.
Nature 318, 670-672

Weinberger C, Hollenberg SM, Ong ES, Harmon JM, Brower ST, Cidlowski J, Thompson EB, Rosenfeld MG and Evans RM (1985)
Identification of human glucocorticoid receptor complementary DNA clones by epitope selection.
Science 228, 740-742

Westphal HM and Beato M (1980)
The activated glucocorticoid receptor of rat liver. Purification and physical characterization.
Eur J Biochem 106, 395-403

Westphal HM, Moldenhauer G and Beato M (1982)
Monoclonal antibodies to the rat liver glucocorticoid receptor.
EMBO J 1, 1467-1471

White R, Lees JA, Needham M, Ham J and Parker M (1987)
Structural organization and expression of the mouse estrogen receptor.
Mol Endocrin 1, 735-744

White BA, Preston GM, Lufkin TC and Bancroft C (1985)
Detection of two chromatin proteins which bind specifically to the 5'-flanking region of the rat prolactin gene.
Mol Cell Biol 5, 2967-2974

Willmann T and Beato M (1986)
Steroid-free glucocorticoid receptor binds specifically to mouse mammary tumour virus DNA.
Nature 324, 688-691

Woodbury CPJr and von HIppel PH (1981)
Relaxed sequence specificities of EcoRI endonucleases and methylase: Mechanisms, possible practical applications and uses in defining protein-nucleic acids recognition mechanisms.
In: "The Restricting Enzymes" (ed J Chirikjian, Elsevier, Amsterdam, 181-207

Wrange Ö, Carlstedt-Duke J and Gustafsson JA (1979)
Purification of the glucocorticoid receptor from rat liver cytosol
J Biol Chem 254, 9284-9290

Wrange Ö and Gustafsson JA (1978)
Separation of the hormone- and DNA-binding sites of the hepatic glucocorticoid receptor by means of proteolysis.
J Biol Chem 253, 856-865

Wu C (1980)
The 5'-end of Drosophila heat shock genes in chromatin are hypersensitive to DNaseI.
Nature 286, 854-861

Wu C (1984)
Two protein-binding sites in chromatin implicated in the activation of heat-shock genes.
Nature 309, 229-234

Wu C (1985)
An exonuclease protection assay reveals heat-shock element and TATA-box DNA-binding proteins in crude nuclear extracts.
Nature 317, 84-87

Zaret KS and Yamamoto KR (1984)
Reversible and persistent changes in chromatin structure accompanying activation of a glucocorticoid-dependent enhancer element.
Cell 38, 29-38

Zhang RG, Joachimiak A, Lawson CL, Schevitz RW, Otwinowski Z and Sigler PB (1987)
The crystal structure of trp aporepressor at 1.8 A shows how binding tryptophan enhances DNA affinity.
Nature 327, 591-597

Zinkel SS and Crothers DM (1987)
DNA bend direction by phase sensitive detection.
Nature 328, 178-181

LIST OF PARTICIPANTS

Juan Andrés Asturias Ortega	- Dpto. Microbiología. Facultad de Biología. Universidad de León.
Pilar Avila Martín	- Dpto. Biología Molecular. Facultad de Medicina. Universidad de Cantabria. Santander.
José Mª Casasnovas Suelves	- Dpto. Química Macromolecular. C.S.I.C. Barcelona.
Jorge Eduardo Domínguez Reyes	- Centro de Biología Molecular. Facultad de Ciencias. Universidad Autónoma de Madrid.
José Juan Gaforio Martínez	- Hospital Virgen de las Nieves. Servicio Análisis Clínicos. Granada.
Marta García Ramírez	- Dpto. Química Macromolecular. Universidad Politécnica de Barcelona.
Rafael Giraldo Suárez	- Centro de Investigaciones Biológicas. C.S.I.C. Madrid.
Sergio González Crespo	- Unidad de Bioquímica. Facultad de Farmacia. Universidad de Barcelona.
Pedro José González Martínez	- Centro de Biología Molecular. C.S.I.C. Universidad Autónoma de Madrid.
Antonio Jiménez Ruiz	- Dpto. Biología Molecular C-X. Universidad Autónoma de Madrid.
Mª Soledad López de Haro Rubio	- Centro de Biología Molecular. C.S.I.C. Universidad Autónoma de Madrid.
Antonio Martínez Laborda	- Dpto. Genética y Microbiología. Facultad de Biología. Universidad de Murcia.
Mª Luisa Martínez Pomares	- Centro de Biología Molecular. C.S.I.C. Universidad Autónoma de Madrid.
Emilia Matallana Redondo	- Dpto. Bioquímica y Biología Molecular. Facultades de Ciencias. Universidad de Valencia.
Lluis Montoliu José	- Dpto. Genética Molecular. CID-C.S.I.C. Barcelona.
Jesús Murillo Martínez	- Dpto. Microbiología. E.T.S.I. Agrónomos. Universidad Politécnica de Madrid.
Inmaculada Ponte Marull	- Centro de Investigación y Desarrollo. C.S.I.C. Barcelona.
Joaquín Royo Cárcamo	- Dpto. Bioquímica. E.T.S.I. Agrónomos. Universidad Politécnica de Madrid.
Eva Mª Sanz Merino	- Dpto. Inmunología. Centro de Investigaciones Biológicas. Madrid.

Juana Mª Sempere de Gancedo - Instituto de Investigaciones Biomédicas. C.S.I.C. Madrid.

Manuel Serrano Marugán - Centro de Biología Molecular. C.S.I.C. Universidad Autónoma de Madrid.

Fundación Juan March

SERIE UNIVERSITARIA

PUBLISHED TEXTS — Green Series

(Mathematics, Physics, Chemistry, Biology, Medicine)

2 Mulet, A:
Estudio del control y regulación, mediante un calculador numérico, de una operación de rectificación discontinua.

4 Santiuste, J. M.:
Combustión de compuestos oxigenados.

5 Vicent López, J. L.:
Películas ferromagnéticas a baja temperatura.

7 Salvá Lacombe, J. A.:
Mantenimiento del hígado dador in vitro en cirugía experimental.

8 Plá Carrera, J.:
Estructuras algebraicas de los sistemas lógicos deductivos.

11 Drake Moyano, J. M.:
Simulación electrónica del aparato vestibular.

19 Purroy Unanua, A.:
Estudios sobre la hormona Natriurética.

20 Serrano Molina, J. S.:
Análisis de acciones miocárdicas de bloqueantes Beta-adrenérgicos.

22 Pascual Acosta, A.:
Algunos tópicos sobre teoría de la información.

25 I Semana de Biología:
Neurobiología.

26 I Semana de Biología:
Genética.

27 I Semana de Biología:
Genética.

28 Zugasti Arbizu, V.:
Analizador diferencial digital para control en tiempo real.

29 Alonso, J. A.:
Transferencias de carga en aleaciones binarias.

30 Sebastián Franco, J. L.:
Estabilidad de osciladores no sinusoidales en el rango de microondas.

39 Blasco Olcina, J. L.:
Compacidad numerable y pseudocompacidad del producto de dos espacios topológicos.

44 Sánchez Rodríguez, L.:
Estudio de mutantes de saccharomyces cerevisiae.

45 Acha Catalina, J. I.:
Sistema automático para la exploración del campo visual.

47 García-Sancho Martín, F. J.:
Uso del ácido salicílico para la medida del pH intracelular.

48 García Garcia, A.:
Relación entre iones calcio, fármacos ionóforos y liberación de noradrenalina.

49 Trillas, E. y Alsina, C.:
Introducción a los espacios métricos generalizados.

50 Pando Ramos, E.:
Síntesis de antibióticos aminoglicosídicos modificados.

51 Orozco, F. y López-Fanjul, C.:
Utilización óptima de las diferencias genéticas entre razas en la mejora.

52 Gallego Fernández, A.:
Adaptación visual.

55 Castellet Solanas, M.:
Una contribución al estudio de las teorías de cohomología generalizadas.

56 Sánchez Lazo, P.:
Fructosa 1,6 bisfosfatasa de hígado de conejo: modificación por proteasas lisosomales.

57 Carrasco Llamas, L.:
Estudio sobre la expresión genética de virus animales.

59 Alfonso Rodríguez, C. N.:
Efectos magneto-ópticos de simetría par en metales ferromagnéticos.

63 Vidal Costa, F.:
A la escucha de los sonidos cerca de $T\lambda$ en el 4_{He} líquido.

65 Andréu Morales, J. M.:
Una proteína asociada a membrana y sus subunidades.

66 Blázquez Fernández, E.:
Desarrollo ontogénico de los receptores de membrana para insulina y glucagón.

69 Vallejo Vicente, M.:
Razas vacunas autóctonas en vías de extinción.

76 Martín Pérez, R. C.:
Estudio de la susceptibilidad magnetoeléctrica en el Cr_2O_3 policristalino.

80 Guerra Suárez, M.ª D.:
Reacción de amidas con compuestos organoalumínicos.

82 Lamas de León, L.:
Mecanismo de las reacciones de iodación y acoplamiento en el tiroides.

84 Repollés Moliner, J.:
Nitrosación de aminas secundarias como factor de carcinogénesis ambiental.

86 II Semana de Biología:
Flora y fauna acuáticas.

87 II Semana de Biología:
Botánica.

88 II Semana de Biología:
Zoología.

89 II Semana de Biología:
Zoología.

91 Viétez Martín, J. M.:
Ecología comparada de dos playas de las Rías de Pontevedra y Vigo.

92 Cortijo Mérida, M. y García Blanco, F.:
Estudios estructurales de la glucógeno fosforilasa b.

93 Aguilar Benítez de Lugo, E.:
Regulación de la secreción de LH y prolactina en cuadros anovulatorios experimentales.

95 Bueno de las Heras, J. L.:
Empleo de polielectrolitos para la floculación de suspensiones de partículas de carbón.

96 Núñez Álvarez, C. y Ballester Pérez, A.:
Lixiviación del cinabrio mediante el empleo de agentes complejantes.

101 Fernández de Heredia, C.:
Regulación de la expresión genética a nivel de transcripción durante la diferenciación de artemia salina.

103 Guix Pericas, M.:
Estudio morfométrico, óptico y ultraestructural de los inmunocitos en la enfermedad celíaca.

105 Llobera i Sande, M.:
Gluconeogénesis «in vivo» en ratas sometidas a distintos estados tiroideos.

106 Usón Finkenzeller, J. M.:
Estudio clásico de las correcciones radiactivas en el átomo de hidrógeno.

107 Galián Jiménez, R.:
Teoría de la dimensión.

111 Obregón Perea, J. M.ª:
Detección precoz del hiporitoidismo congénito.

115 Cacicedo Egües, L.:
Mecanismos moleculares de acción de hormonas tiroideas sobre la regulación de la hormona tirótropa.

121 Rodríguez García, R.:
Caracterización de lisozimas de diferentes especies.

122 Carravedo Fantova, M.:
Introducción a las orquídeas españolas.

125 Martínez-Almoyna Rullán, C.:
Contribución al estudio de la Manometría ano-rectal en niños normales y con aletraciones de la continencia anal.

127 Marro, J.:
Dinámica de transiciones de fase: Teoría y simulación numérica de la evolución temporal de aleaciones metálicas enfriadas rápidamente.

129 Gracia García, M.:
Estudio de cerámicas de interés arqueológico por espectroscopia Mössbauer.

131 García Sevilla, J. A.:
Receptores opiáceos, endorfinas y regulación de la síntesis de monoaminas en el sistema nervioso central.

132 Rodríguez de Bodas, A.:
Aplicación de la espectroscopia de RPE al estudio conformacional del ribosoma y el tRNA.

136 Aragón Reyes, J. L.:
Interacción del ciclo de los purín nucleóticos con el ciclo del ácido cítrico en músculo esquelético de rata durante el ejercicio.

139 Genís Gálvez, J. M.:
Estudio citológico de la retina del camaleón.

140 Segura Cámara, P. M.:
Las sales de tiazolio ancladas a soporte polimérico insoluble como catalizadores en química orgánica.

141 Vicent López, J. L.:
Efectos anómalos de transporte eléctrico en conductores a baja temperatura.

143 Nieto Vesperinas, M.:
Técnicas de prolongación analítica en el problema de reconstrucción del objeto en óptica.

145 Arias Pérez, J.:
Encefalopatía portosistémica experimental.

147 Palanca Soler, A.:
Aspectos faunísticos y ecológicos de carábidos altoaragoneses.

150 Vioque Cubero, B.:
Estudio de procesos bioquímicos implicados en la abscisión de la aceituna.

151 González López, J.:
La verdadera morfología y fisiología de Azoyobacter: células germinales.

152 Calle García, C.:
Papel modulador de los glucocorticoides en la población de receptores para insulina y glucagón.

154 Alberdi Alonso, M.ª T.:
Paleoecología del yacimiento del Neógeno continental de Los Valles de Fuentidueña (Segovia).

156 Gella Tomás, F. J.:
Estudio de la fosforillasa kinasa de hígado y leucocitos: purificación, características y regulación de su actividad.

157 Margalef Mir, R.:
Distribución de los macrofitos de las aguas dulces y salobres del E. y NE. de España y dependencia de la composición química del medio.

158 Álvarez Fernández-Represa, J.:
Reimplantación experimental de la extremidad posterior en perros.

161 Tomás Ferré, J. M.ª:
Secreción y reutilización de trifosfato de adenosina (ATP) por sinaptosomas colinérgicos.

163 Ferrándiz Leal, J. M.:
Estudio analítico del movimiento de rotación lunar.

164 Rubió Lois, M.; Uriz Lespe, M.ª J. y Bibiloni Rotger, M.ª A.:
Contribución a la fauna de esponjas del litoral catalán. Esponjas córneas.

165 Velasco Rodríguez, V. R.:
Propiedades dinámicas y termodinámicas de superficies de sólidos.

166 Moreno Castillo, I.:
Ciclo anual de zooplancton costero de Gijón.

168 Durán García, S.:
Receptores insulínicos en hipotálamo de rata: localización subcelular y mecanismo(s) de regulación.

169 Martínez Pardo, R.:
Estudio del mecanismo secretor de hormona juvenil en oncopeltus fasciatus.

171 García Jiménez, J.:
Fusariosis del gladiolo: un estudio preliminar.

173 Fernández Aláez, C.:
Análisis estructural en sabinares de la provincia de León.

174 Furio Egea, J.:
Citokinas en agrios. Actividades endógenas, efectos fisiológicos y aplicaciones.

180 Moreno Rodríguez, J. M.:
Estudios ecológicos en jarales (cistion laurofilii): Variación anual de algunos factores del entorno y manifestaciones fenológicas.

182 Pons Vallés, M.:
Estudios espectroscópicos de fosfolípidos polimerizables.

183 Herrero Ruiz de Loizaga, V. J.:
Estudio de reacciones químicas por haces moleculares. Aplicación a la reacción $C_2H_5Br + K$ $Brk + C_2H_5$.

193 Martín García, V. S.:
Utilización sintética en química orgánica de metales pesados como catalizadores. Oxidación asimétrica.

195 Badía Sancho, A.:
Receptores presinápticos en el conducto deferente de rata.

196 Estévez Toranzo, A.:
Supervivencia de patógenos bacterianos y virales de peces en sistemas de cultivo.

197 Lizarbe Iracheta, M.ª A.:
Caracterización molecular de las estructuras de colágeno.

203 López Calderón, I.:
Clonación de genes de «Saccharomyces cerevisiae» implicados en la reparación y la recombinación.

211 Ayala Serrano, J. A.:
Mecanismo de expresión de la PBP-3 de «E. coli»: Obtención de una cepa hiperproductora de la proteína.

240 **Genetic Strategies in Development.**
Symposium in honour of Antonio García Bellido. Lectures by S. Ochoa, S. Brenner, G. S. Stent, E. B. Lewis, D. S. Hogness, E. H. Davidson, J. B. Gurdon y F. Jacob.

244 **Course on Genome Evolution.**
Organized by E. Viñuelas. Lectures by R. F. Doolittle, A. M. Weiner/N. Maizels, G. A. Dover, J. A. Lake, J. E. Walker, J. J. Beintema, A. J. Gibbs, W. M. Fitch, P. Palese, G. Bernardi y J. M. Lowenstein.

246 **Workshop on tolerance: Mechanisms and implications.**
Organized by P. Marrack and C. Martínez-A. Lectures by H. von Boehmer, J. W. Kappler, C. Martínez-A., H. Waldmann, N. Le Douarin, J. Sprent, P. Matzinger, R. H. Schwartz, M. Weigert, A. Coutinho, C. C. Goodnow, A. L. DeFranco y P. Marrack.

247 **Workshop on pathogenesis-related proteins in plants**
Organized by V. Conejero and L. C. Van Loon. Lectures by L. C. Van Loon, R. Fraser, J. F. Antoniw, M. Legrand, Y. Ohashi, F. Meins, T. Boller, V. Conejero, C. A. Ryan, D. F. Klessig, J. F. Bol, A. Leyva and F. García-Olmedo.

Established in 1955,
the Juan March Foundation is a non-profit organization
dedicated to cultural,
scientific and charitable objectives.
It is one of the most important in Europe,
considering its endowment and activities.

The Juan March Foundation has collaborated
significantly in the area of science and research by
funding awards, grants and scholarships
for scientific training,
supporting research studies
and publishing research results.

Since 1970, the Foundation
has dedicated special attention to the field of Biology
by means of various programmes,
such as the Plan on Molecular Biology
and its Applications (1981-88)
and at present the Programme of
International Meetings on Biology,
designed to actively promote the relationship
between Spanish biologists and their international colleagues