REVIEW ARTICLE

Chemical nucleases as probes for studying DNA–protein interactions

Athanasios G. PAPAVASSILIOU

Differentiation Programme, European Molecular Biology Laboratory, Postfach 10.2209, Meyerhofstrasse 1, D-69012 Heidelberg, Germany

INTRODUCTION

Over the past fifteen years much research has been devoted to the mode(s) by which the linear DNA macromolecule communicates with cellular proteins in the course of genetic programming, thus determining the patterns of growth and development in living organisms. A plethora of proteins have been shown to interact non-covalently with DNA either as structural elements, exemplified by the histones of eukaryotic chromatin, or as components of the complex cellular machineries operating in the essential biological processes of DNA replication, transcription and recombination. DNA–histone interactions are largely sequence neutral and rely mostly on electrostatic attractions between basic amino acids and the negatively charged sugar-phosphate backbone, whereas the DNA binding of proteins necessary for the conservation and selective reading of genetic information (e.g. transcription factors) is highly sequence-dependent. The principal basis for such sequence discrimination is direct contact between the polypeptide chain and the exposed edges of the base pairs, in both the major and minor grooves of B-DNA. These contacts may involve either hydrogen bonding or van der Waals, ionic and hydrophobic interactions between amino acid residues of the protein and specific atoms of the purine and pyrimidine rings of the DNA double helix (for a recent review see [1]). Furthermore, these direct interactions are supplemented by the sequence-dependent bendability or deformability of DNA, which limits the energetically favourable configurations of a particular binding site and thereby imposes additional sequence-dependent constraints on the binding affinity [1]. An emerging feature of specific DNA–protein binding is that several parameters governing DNA structure and conformational microheterogeneity [e.g. deoxyribose conformation and phosphate–phosphate distances (groove size)] can be significantly distorted by interacting proteins.

Although the ultimate fine molecular description of DNA–protein complexes requires crystallographic and/or n.m.r. analyses, the development and application of a veritable armoury of enzymic and chemical probes that are capable of breaking (either directly or indirectly) the backbone of DNA has made it possible to obtain quite detailed structural and dynamic pictures of numerous DNA–protein systems. These range from chromatin fibres via RNA polymerase-promoter to a broad spectrum of transcription factor–DNA complexes. Depending on the particular probe and the experimental set-up, various kinds of information about a DNA–protein complex can be gleaned. In the traditional protection footprinting approach (the term footprinting being evocative of the visual result of the experiment), the difference in DNA reactivity towards the probe in the presence or absence of bound protein under 'single-hit' conditions (i.e. roughly one nick or modification event per DNA molecule), is assessed. Decreased reactivity may be a result of direct contacts (proximity) with the protein concerned, whereas enhanced reactivity reflects either allosteric effects caused by a protein-induced rearrangement of the DNA conformation or preferential partitioning of the attacking reagent into a volume adjacent to the DNA. Several proteins generate strong footprints when studied using protection footprinting analysis. However, proteins with low affinity for DNA often produce very weak footprint patterns (bands in the region of the recognition sequence), due to a high level of signal from uncomplexed DNA. Since the data from these weak footprints are difficult to interpret, it is desirable to augment the footprint signal over the noise of the digestion pattern of unbound DNA. Provided that the sequence-specific protein, once bound to its binding site, does not exchange rapidly between DNA molecules, this can be achieved easily by separating the probed, complexed and naked DNA species using electrophoresis on a native (mobility-shift [2,3]) polyacrylamide gel (coupled protection footprinting/mobility-shift electrophoresis assay, Figure 1). Then the DNA can be eluted from the corresponding areas of the gel and analysed on a high-resolution (sequencing) gel to reveal the protected domain (Figure 1). In an interference footprinting analysis, the DNA is first subjected to limited reaction with the probe and the population of modified DNA molecules that have retained protein-binding capacity is obligatorily resolved by the mobility-shift gel electrophoresis assay, eluted, and analysed on a sequencing gel (Figure 1). This procedure infers regions and features of the DNA that are critical for DNA–protein complex formation. In both experimental strategies the DNA molecule is radioactively labelled on one end of one of its strands (Figure 1), so that the chemical reactions developed for DNA sequencing can be used to identify DNA–protein contacts at each individual nucleotide [4].

CHEMICAL NUCLEASES AS FOOTPRINTING TOOLS: GENERAL CONSIDERATIONS

As noted above, a variety of reagents have been used to define the sequence-specific contacts of a protein on DNA by footprinting analysis. Although the most widespread reagents are dimethyl sulphate (DMS) and pancreatic deoxyribonuclease I (DNase I), chemical nucleases, a group of synthetic compounds exhibiting efficient nucleolytic activity, are increasingly employed in footprinting applications and in this perspective will be the focus of the present review (photochemical footprinting probes will not be considered here).

Chemical nucleases are redox-active co-ordination complexes that nick the phosphodiester backbone of DNA (but also RNA) under physiological pH and temperature by oxidatively de-

Abbreviations used: DMS, dimethyl sulphate; DNase I, deoxyribonuclease I; MPE, methidiumpropyl-EDTA; OP, orthophenanthroline (1,10-phenanthroline).

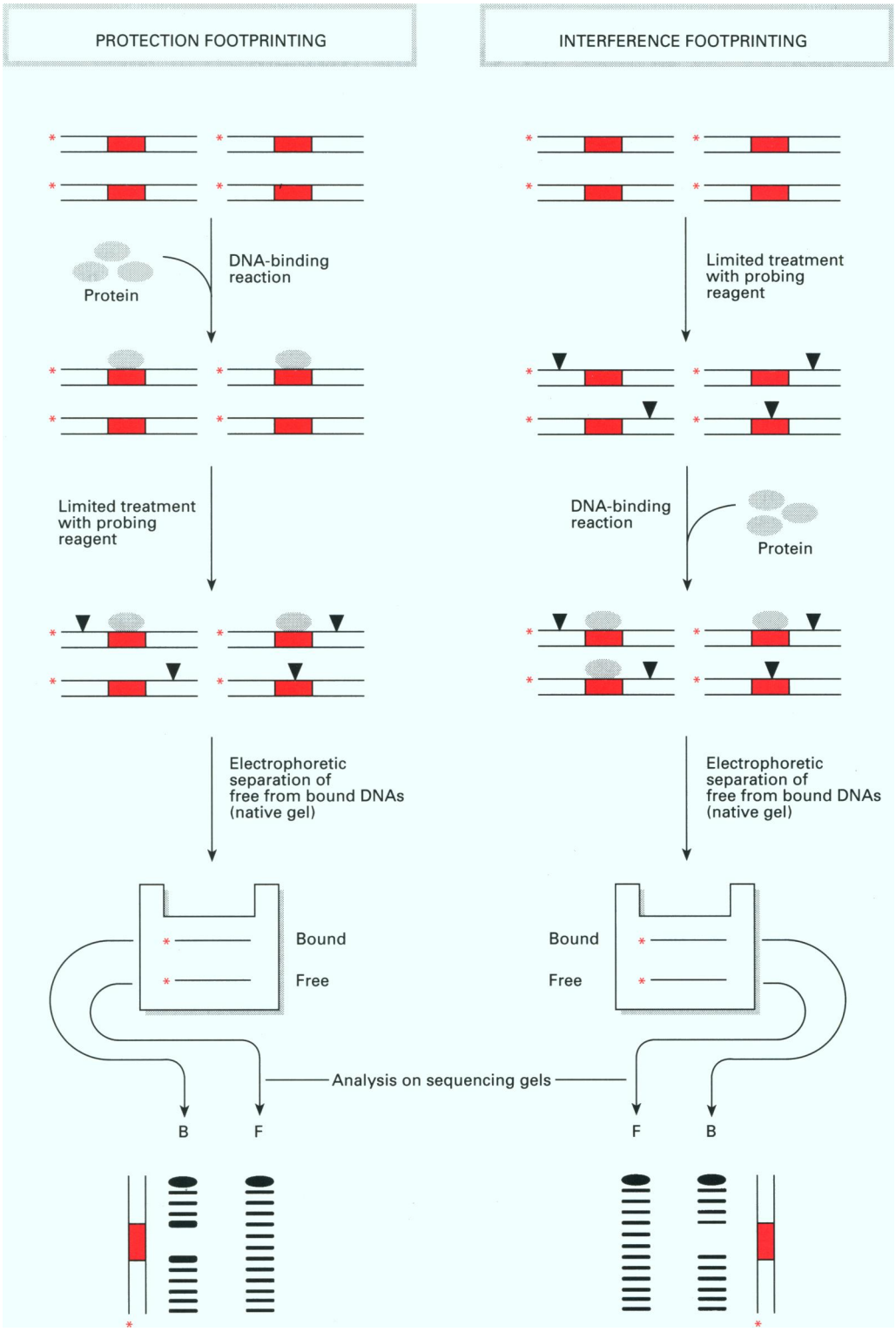

Figure 1 Schematic outline of the experimental approach followed in a coupled protection footprinting/mobility-shift electrophoresis assay and in an interference footprinting analysis

DNA fragments are depicted as bars with the protein-binding site indicated by a red box. Asterisks denote a radioactive label at one end of a DNA strand. Protein is symbolized by a grey oval. DNA is subjected to modification or single-strand nicking (denoted by the black triangles) either before (interference footprinting) or after the DNA-binding reaction (protection footprinting). This treatment is performed in a manner that only allows one 'hit' per DNA molecule, or for large molecules about one 'hit' per 300 nucleotides. The bound and non-bound (free) DNA fractions are separated by the mobility-shift electrophoresis assay, eluted and analysed on a sequencing gel (chemically modified DNA must be subjected to further chemical treatment leading to strand scission at modified nucleotides, prior to denaturing electrophoresis). The digested sites are visualized as autoradiograph bands, each of which corresponds to an affected position on the DNA template. By comparing band intensities from free and complexed DNA probed in this manner, it is possible to deduce the attachment site for the protein, since interacting sites exhibit altered reactivities. Calibration of the sequencing gel is facilitated by co-electrophoresing a Maxam–Gilbert sequencing ladder of the terminally labelled DNA used in the assay. At the bottom of both panels the anticipated

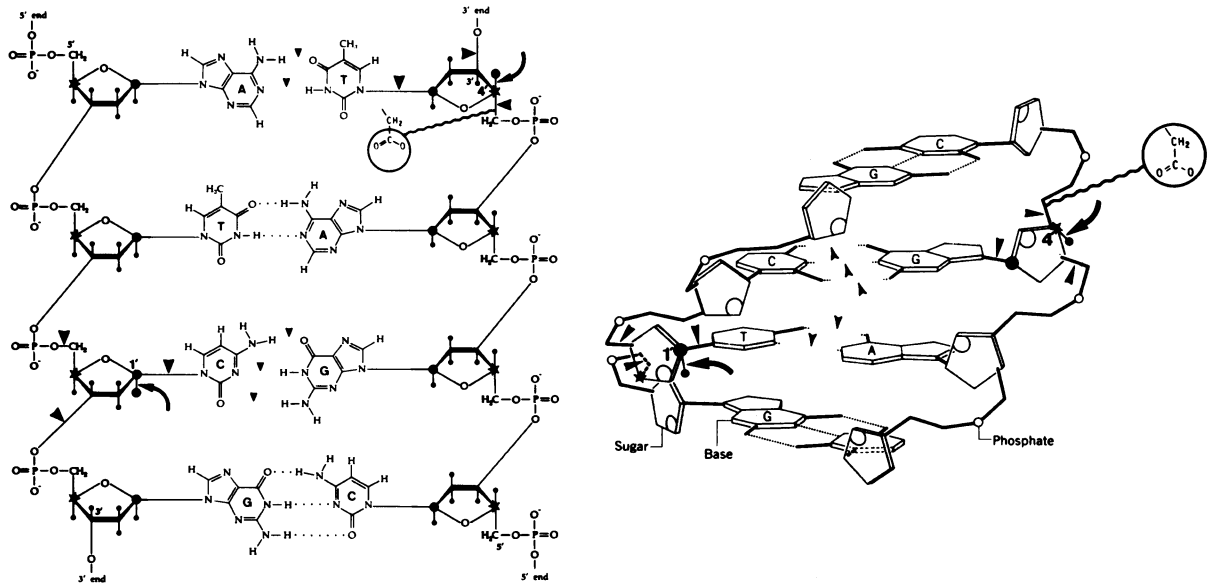

Figure 2 Chemical nucleases attack the deoxyribose moiety of DNA

A portion of a double-stranded DNA molecule is depicted on the left, with the positions of various atomic groups on the polynucleotide strands being distorted in order to make them visible in one plane. The deoxyribose moieties and phosphate groups connect to form the backbone of each strand, and a base attaches to each deoxyribose. The four different bases are represented by the letters A, T, G and C; dots denote hydrogen bonds between them (base pairing). The five carbon atoms in each deoxyribose ring are assigned numbers 1′ (linked to a base nitrogen) to 5′ (esterified to the phosphate group). Synthetic chemical nucleases primarily attack the C1′ and C4′ atoms of the deoxyribose ring (indicated by large black spheres and black stars, respectively), possibly via C-H insertion (curved arrows) and abstraction of the H1′ and H4′ hydrogen atoms (corresponding small black spheres). Heavy and light arrowheads point to the chemical bonds broken after oxidative attack at representative C1′ and C4′ deoxyribose atoms. In both cases the products include the liberated base and 5′-phosphorylated termini. Attack at the C1′ deoxyribose atom generates a 3′-phosphorylated end, while C4′ oxidation produces a 3′-phosphoglycolate end (shown in the projection originating from that position). The right part of the figure depicts all the above in the context of a schematic diagram of a short section of an actual DNA double helix (base pairs are tilted so that the plane of the bases lies perpendicular to the page).

grading the deoxyribose moiety [5–7] (Figure 2). The kinetic scheme for the nucleolytic activity by a chemical nuclease involves reversible formation of weak or strong complexes with DNA, followed by the scission reaction which is funnelled through highly reactive and non-selective oxidative intermediates generated near the surface of the DNA. A putative exception to this kinetic scheme is ferrous-EDTA (see below) which, because of its net negative charge, has been proposed to act in a non-DNA-bound form as a generator of highly reactive diffusible hydroxyl radicals. Unlike real enzymes, where binding and catalysis are integrally interconnected, the binding of a chemical nuclease directs the cleavage event but does not activate the redox chemistry responsible for scission which proceeds equally efficiently free in solution or on the DNA molecule [8,9]. This chemistry is activated by reducing agents (e.g. thiol or ascorbic acid) in the presence of molecular oxygen or H_2O_2. Studies on the mechanism of DNA scission by synthetic chemical nucleases have demonstrated that the primary stable products generally include free bases, 5′-phosphorylated termini, and 3′-phosphorylated termini plus 3′-phosphoglycolates at variable ratios (Figure 2). The 3′-phosphomonoester ends are most likely derived from oxidative attack on the C1′ hydrogen atom of the deoxyribose (possibly either via C-H insertion or an oxygen rebound mechanism), while the 3′-phosphoglycolate ends presumably arise from attack on the C4′ hydrogen atom of the deoxyribose ring [10–14] (Figure 2). Since the chemical nucleolytic activity produces DNA fragments bearing 5′- and 3′-phosphorylated termini, sequencing gels can be accurately calibrated with the Maxam–Gilbert sequencing reactions [4].

The deoxyribose-directed reactivity sharply distinguishes chemical nucleases from other chemical modification reagents which are employed in biochemical studies of DNA–protein complexes. For instance, DMS, diethyl pyrocarbonate, OsO_4 and $KMnO_4$ attack predominantly atoms or bonds within the heterocyclic ring(s) of the various bases, and ethylnitrosourea reacts with the phosphates of the DNA backbone. In addition to being hazardous compounds, these reagents do not cleave the phosphodiester backbone without subsequent alkali treatment (e.g. piperidine as in Maxam–Gilbert sequencing reactions), while protection footprinting experiments on DNA–protein complexes are not feasible using ethylnitrosourea due to the harsh conditions required for DNA ethylation (50% ethanol, 50 °C; [15]). By contrast, chemical nucleases are far less toxic reagents which cause DNA-strand scission directly under physiological conditions at all sequence positions irrespective of the base moiety attached to the deoxyribose oxidized. In this context, chemical nucleases have several advantages over the 'classical' chemical (but also enzymic) footprinting probes, such as more convenience in handling their solutions, better definition of protein-protected sequences [16], the potential for revealing major- and minor-groove interactions [13] and augmented sensitivity to protein-triggered alterations in DNA structure [17]. These features will be extensively discussed in the following sections. A putative drawback of chemical nucleases over the

results are presented (B, bound DNA; F, free DNA). Residues protected by the protein or residues whose modification/deletion interferes with protein binding can be identified as missing (or under-represented) bands, and residues exhibiting enhanced reactivity as a consequence of protein binding by increased intensities. As a result of the 'single-hit' kinetics followed during the probing step, most of the DNA molecules remain uncleaved (thick band on top of the gels).

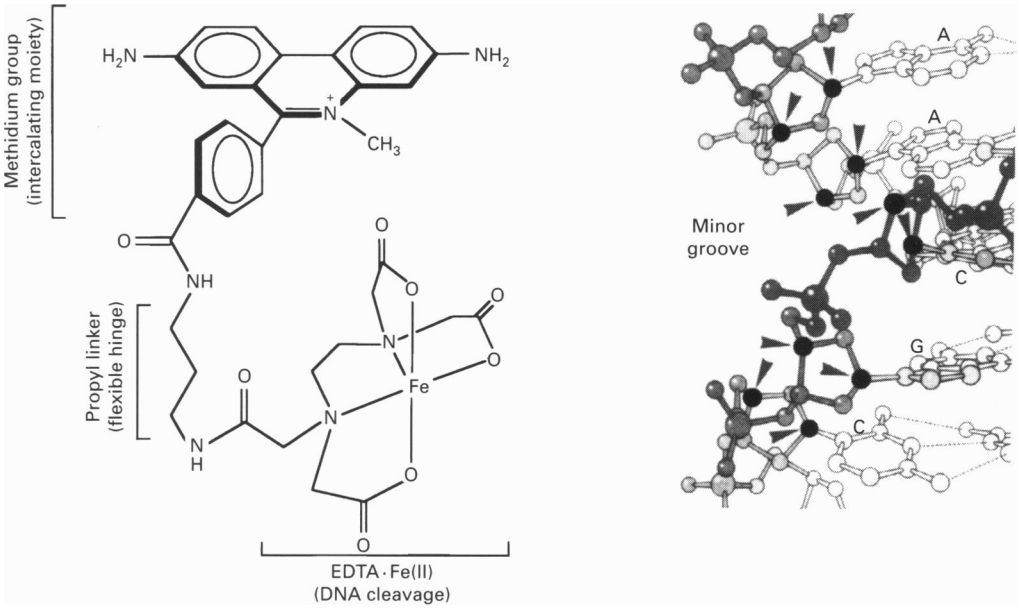

Figure 3 The synthetic chemical nuclease methidiumpropyl-EDTA·Fe(II) and its targeting sites in DNA

The chemical structure of methidiumpropyl-EDTA·Fe(II) [(MPE·Fe(II)] is depicted on the left, indicating the different structural units of the molecule and their function when interacting with DNA. A three-dimensional model of a portion of a 'half' DNA double helix is shown on the right, illustrating the topography of the MPE·Fe(II)-attacked C1′ and C4′ deoxyribose atoms within the minor groove (black spheres marked by arrowheads). C1′ is involved in the deoxyribose-base linkage (A, C and G symbolize the various bases attached to the sugar-phosphate backbone). DNA-strand cleavage is most likely achieved via the generation of diffusible hydroxyl radicals near the deoxyribose ring.

chemical modification reagents mentioned above is that induction of direct cleavage of the DNA backbone precludes their use in an analysis of the effects of superhelical density on DNA–protein interactions.

The bis(1,10-phenanthroline)–cuprous ion complex was the first synthetic co-ordination complex shown to accomplish DNA-strand scission [5,18] followed by derivatives of ferrous-EDTA [6,19,20], various metalloporphyrins [21,22] and ruthenium complexes of 4,7-diphenyl-1,10-phenanthroline [23]. In addition to their profitable use as footprinting tools for exploring DNA–protein interactions, research with the above chemical nucleases has developed in several other directions including the study of small drug molecules bound to DNA ([24] and references therein), characterization of sequence-dependent conformational variability in DNA and RNA [25–27] and the design of site-specific (DNA-targeted) scission reagents for the dissection of chromosomal DNA [8,28].

Biochemical correlates of three chemical nucleases elaborated as footprinting probes will be presented, with emphasis on their mechanism of scanning protein–DNA complexes, versatility in probing diverse aspects of DNA–protein interactions, comparative 'use-profile' and advantages over conventional footprinting reagents. The chronological application of these chemical nucleases as footprinting tools will be followed in their presentation.

Methidiumpropyl-EDTA·iron(II)

The synthetic footprinting reagent methidiumpropyl-EDTA·iron(II) [MPE·Fe(II)] was designed by Hertzberg and Dervan [6] by tethering a well-characterized DNA-intercalating ligand, methidium (the methyl analogue of ethidium), via a hydrocarbon (propyl) chain to the metal chelator ethylenediaminetetra-acetic acid (EDTA) (Figure 3). The intercalator part of MPE 'delivers' the iron/oxygen chemistry to the DNA double helix: EDTA forms a stable complex with ferrous ion [Fe(II)] which is positioned in the minor groove and in the presence of dioxygen (in the form of O_2, or better, H_2O_2) induces single-strand breaks in DNA [12]. From the analyses of DNA-cleavage products by MPE·Fe(II), the correspondence of one released base per nicking event and the nature of the DNA polymer termini (5′-phosphoryl and roughly equal proportions of 3′-phosphoryl and 3′-phosphoglycolic acid) suggest that DNA-strand scissions are due to oxidation (abstraction of a hydrogen atom) of the C1′–H and C4′–H bonds of deoxyribose rings (Figure 2), most likely through production of diffusible hydroxyl radicals (·OH) near the deoxyribose in the minor groove of DNA [11,12] (Figure 3). The radical that is left behind on a deoxyribose residue eventually decomposes, resulting in the loss of the sugar and its attached base from the DNA backbone and hence the introduction of a single-stranded gap into the DNA molecule at the site of hydroxyl radical attack (Figure 2). The MPE·Fe(III) product of the reaction is reduced back to MPE·Fe(II) by added thiol (typically dithiothreitol) or ascorbate, establishing a catalytic cycle and thus a continuous source of active metal ion [12].

Since MPE·Fe(II) induces DNA backbone scissions proximal to the site of methidium intercalation [12], a bound protein is expected to inhibit the cleavage reaction either by interfering with the ability of the flat methidium group to intercalate into the DNA, or by sterically blocking access of the diffusible hydroxyl radical to the DNA backbone. Whatever the case, protection from MPE·Fe(II)-mediated cleavage implies that at least a portion of the DNA-binding protein occupies the minor groove. Because, however, MPE·Fe(II) intercalates into DNA biasing it toward scission of double-helical regions, a 'pseudoprotection'

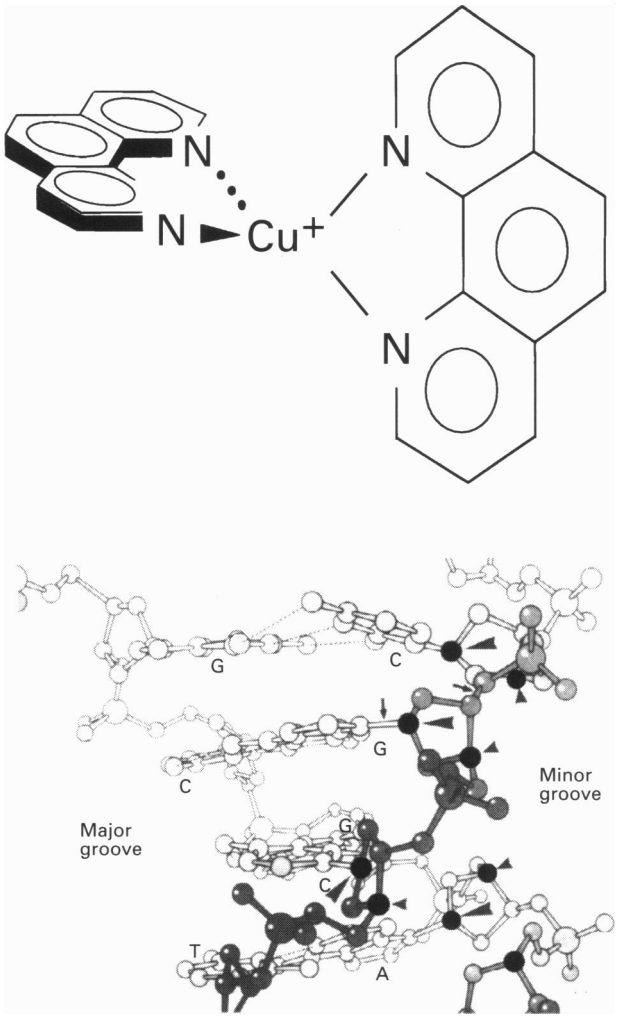

Figure 4 The synthetic chemical nuclease bis(1,10-phenanthroline)·Cu(I) and its targeting sites in DNA

The chemical structure of the 2:1 1,10-phenanthroline·Cu(I) [$(OP)_2Cu^+$] tetrahedral co-ordination complex is shown on the top. A three-dimensional model of a portion of the DNA double helix is depicted on the bottom, revealing the location of the $(OP)_2Cu^+$-attacked C1′ and C4′ deoxyribose atoms within the minor groove [black spheres indicated by long (C1′ atoms) and short (C4′ atoms) arrowheads]. G, C, T and A represent the different bases attached to the sugar-phosphate backbones and broken lines denote base pairing. Small arrows point to the bonds cleaved after oxidative attack at the corresponding C1′ deoxyribose atom, leading to a single-stranded gap into the DNA molecule at this position.

of sites that become single-stranded after protein binding can arise that is actually due to the inability of the reagent to cut in these sites. An artifactual protection will also be observed if protein binding triggers conformational alterations in double-stranded regions, which interfere with intercalation of the methidium moiety. MPE·Fe(II) cleaves DNA in a non-selective manner (the generated hydroxyl radical reacts with essentially every deoxyribose at about the same rate; [29]) and with only little sequence specificity (originating from the 2-fold preference of the DNA-intercalating methidium group for GC motifs). All nucleotides are therefore probed in a naked DNA molecule and the boundaries of the footprint are precisely delineated at the DNA sequence level. While it is possible to terminate the MPE·Fe(II)-mediated cleavage by the addition of the iron-sequestering compound bathophenanthroline disulphonate [30,31], this reagent may disrupt protein binding to DNA and hence abolish detection of DNA–protein complexes in coupled MPE·Fe(II) protection footprinting/mobility-shift assays [32] (Figure 1). Therefore, the reaction is preferably stopped by directly applying the footprinting mixture on to the native polyacrylamide gel.

Although MPE·Fe(II) is considered as the chemical version of DNase I (similar to the enzymic probe it does not often cleave within the footprint, even in places where the DNA backbone is known to be exposed to solvent [33]), its much smaller dimensions and virtual sequence neutrality have made this reagent an attractive alternative to DNase I for determining the location, size and relative importance of the binding sites of proteins on native DNA. It should be noted, however, that such analysis assumes that the agent used to cleave DNA or to score protein-binding positions on the linear polymer does not itself perturb the DNA–protein equilibrium. While this is probably true for most cleaving agents used in protection footprinting experiments, it may not be the case for studies carried out with MPE·Fe(II). In as much as the rate of DNA cleavage by this iron complex is very slow compared with that of DNase I and other chemical nucleases ([24]; usually leaving > 75% of the DNA intact), high levels of the reagent are generally required to obtain a sufficient amount of deoxyoligonucleotide products for mapping analysis. When the concentration of MPE·Fe(II) becomes comparable with the concentration of protein-binding sites, an increased possibility exists that the cleaving agent can perturb the equilibrium between the protein and DNA, and therefore the stability of the complex may be severely affected. Nevertheless, the variety of different systems analysed by MPE·Fe(II) footprinting in several laboratories clearly demonstrate the validity of this now established general footprinting methodology.

MPE·Fe(II) has been widely used not only in standard protection and interference footprinting applications (e.g. [32,34,35]) (Figure 1), but also to investigate nucleosome positioning in the vicinity of eukaryotic genes in studies of chromatin structure in nuclei, taking advantage of the preferential cleavage of inter-nucleosomal linker DNA by this reagent [30].

Precise and rigorous descriptions of MPE·Fe(II) footprinting protocols can be found in [31,32].

Bis(1,10-phenanthroline)·copper(I)

Another metal-containing footprinting reagent is based on the complex of copper(I) [Cu(I)] with 1,10-phenanthroline (ortho-phenanthroline, OP). The 2:1 1,10-phenanthroline·Cu(I) complex [$(OP)_2Cu^+$; Figure 4] is an efficient chemical nuclease which nicks the phosphodiester backbone of DNA under physiological conditions by oxidation of the deoxyribose moiety [5,18]. The chemistry of B-DNA scission proceeds by the kinetic scheme summarized in Figure 5. The first step is the formation of the 1,10-phenanthroline–cupric ion co-ordination complex, under conditions that favour the 2:1 stoichiometry [$(OP)_2Cu^{2+}$]. The DNA-cleavage process is initiated by adding a thiol (usually 3-mercaptopropionic acid because of its poor chelating properties) or ascorbic acid to the aerobic reaction mixture containing the target DNA [36]. Under these conditions, the 2:1 cupric complex [$(OP)_2Cu^{2+}$] is reduced to the 2:1 cuprous complex [$(OP)_2Cu^+$] which is in turn oxidized by molecular oxygen to generate H_2O_2 via a superoxide intermediate. H_2O_2 is an essential co-reactant for the chemical nuclease activity and can be formed endogenously from the oxidation of the diffusible cuprous complex by O_2 as indicated (Figure 5), or it may be added

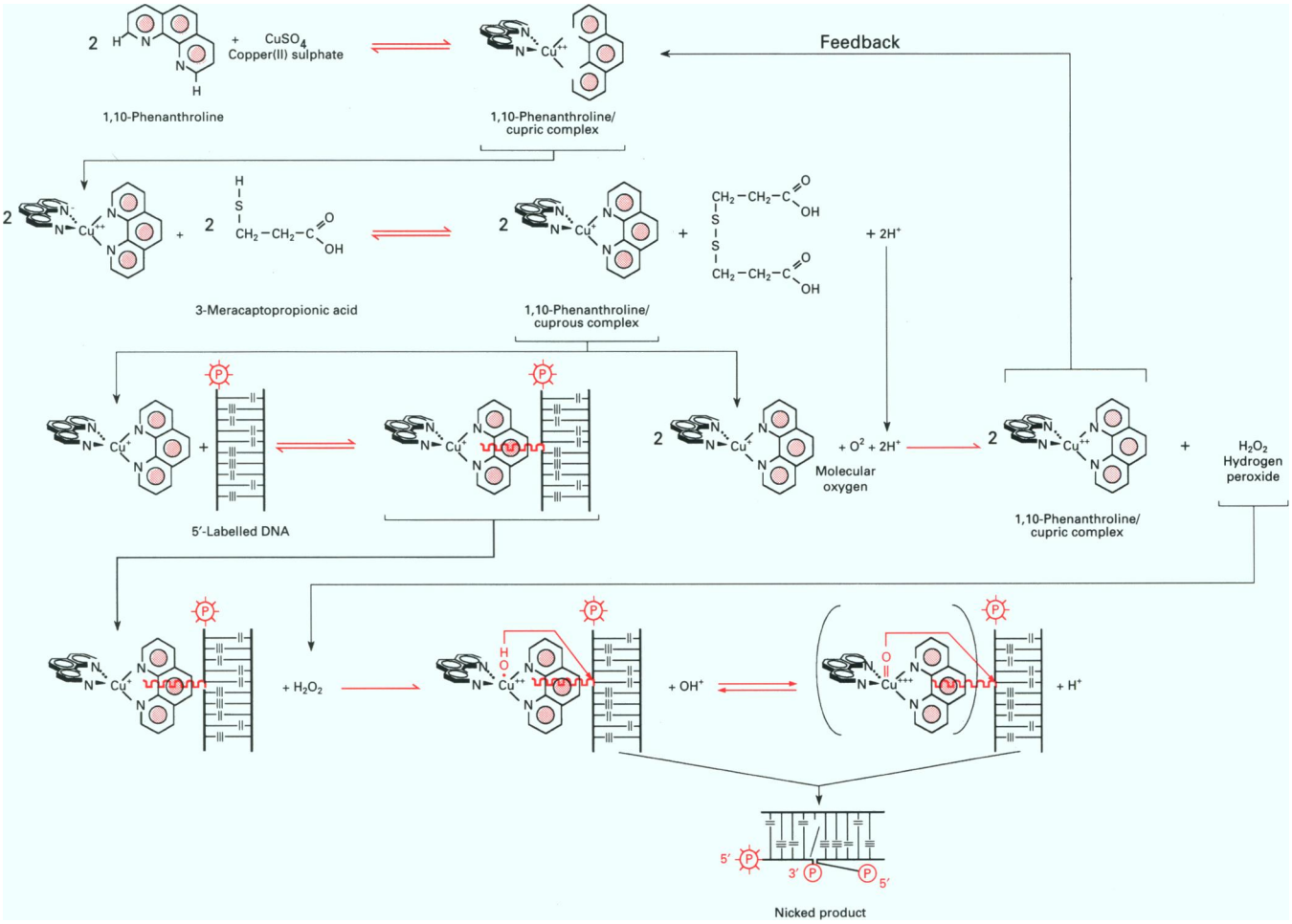

Figure 5 Schematic representation of the kinetic mechanism for the nuclease activity of bis(1,10-phenanthroline)·Cu(I)

The $(OP)_2Cu^+$ complex is formed near the surface of the DNA by the reduction of the $(OP)_2Cu^{2+}$ complex. The *in situ* one electron reductant can be either thiol (as the 3-mercaptopropionic acid used in this case) or ascorbic acid. A non-covalently bound adduct between the $(OP)_2Cu^+$ complex and DNA (represented by the curly red line) explains the high efficiency of the reaction and the sensitivity to substituents on the phenanthroline ligand. Reaction of the DNA-bound $(OP)_2Cu^+$ with H_2O_2 [which can be produced from the oxidation of the diffusible $(OP)_2Cu^+$ complex by molecular oxygen as indicated here or it may be added exogenously] proceeds via a copper-oxo species (depicted either as a hydroxyl radical co-ordinated to the cupric ion or as a copper-oxene structure), whose existence is consistent with the fact that H_2O_2 is relatively inert as an outer-sphere oxidant [82] and is entirely compatible with the co-ordination chemistry of the $(OP)_2Cu^{2+/+}$ couple [42]. This short-lived species then attacks deoxyribose linkages along the DNA polymer, culminating in scission of the phosphodiester backbone and generation of DNA fragments terminating in 5′ and 3′ phosphates. The reaction pathway is obligatory because the DNA–$(OP)_2Cu^{2+}$ complex is not reducible.

exogenously [37]. The tetrahedral cuprous complex, present at the steady-state concentration defined by the experimental conditions (feedback mechanism in Figure 5), binds reversibly (in a non-intercalative manner) to the minor groove of DNA to form a central, non-covalent intermediate through which the reaction is funnelled ([8,13,18,38] and references therein). This mode of binding is responsible for the structural and sequence-dependent variability of the nucleolytic activity (see below), enhancing the reagent's selectivity more than the fairly non-specific intercalative binding of the methidium group of MPE·Fe(II). The DNA-associated cuprous complex is subjected *in situ* (i.e. on the DNA surface) to one-electron oxidation by H_2O_2 to generate a short-lived, highly reactive DNA-bound copper-oxo species, which can be represented either as a hydroxyl radical co-ordinated to the cupric ion or as a copper-oxene structure [10,37] (Figure 5). This species then attacks the H1′ [and to a lesser extent (10–20%) the H4′] deoxyribose proton of nucleotides, which is deeply recessed in the floor of the minor groove [39] (Figure 4), initiating a series of reactions culminating in cleavage of the phosphodiester backbone [10,14] (Figure 5). The stability of the essential non-covalent DNA–$(OP)_2Cu^+$ complex and the orientation and proximity of the reactive copper-oxo species relative to the C1′ deoxyribose hydrogen (which are dependent on the shape and particular configuration of the minor groove attacked; Figure 4) will therefore govern the efficiency and kinetics of strand scission at any given sequence position. In this respect, B-DNA is the most preferred helical substrate, A-DNA is digested at about one-quarter of the rate, and Z-DNA is not detectably cleaved under conditions where B-DNA is readily cut [7,37,40]. Stable products in the above reaction scheme include 5′- and 3′-phosphorylated ends, the free base, the deoxyribose oxidation product 5-methylene-2-furanone and minor amounts (10–20%, arising from oxidative attack on the C4′ H of the deoxyribose) of 3′-phosphoglycolate termini [10,13,14] (Figures 2 and 5). The nuclease activity of $(OP)_2Cu^+$ is efficiently quenched by the ligand 2,9-dimethyl-OP (neocuproin), which by sequestering all

the available copper in an inert and stable cuprous complex, destroys the ability of the $(OP)_2Cu$ complex to cycle between the oxidized and reduced forms in the thiol oxidation and DNA-nicking reactions [41,42].

The nucleolytic activity of $(OP)_2Cu^+$ presents several advantages as a footprinting reagent relative to protection analyses utilizing the 'classical' probes DNase I or DMS. First, the $(OP)_2Cu^+$ chelate is a small molecule (compared with DNase I) that permits cleavage closer to the edge of the DNA sequence protected by protein binding, and therefore a more precise definition of it. Secondly, since the scission chemistry involves attack on the deoxyribose moiety, $(OP)_2Cu^+$ is able to cut at almost every sequence position, regardless of the base linked to the deoxyribose oxidized. However, the intensity of cutting does depend on local primary sequence, with the nucleotide 5′ to the cutting site appearing to be the major determinant in governing cleavage efficiency [25–27,43]. Analysis of the digestion patterns of a variety of DNAs revealed that A and T are equally preferred, G is slightly favoured relative to A and T, and C is slightly disfavoured as a cutting site with respect to A and T [25]. This lack of intrinsic preference for A and T contrasts with the resistance of runs of consecutive A or T residues or TpA islands (and phosphodiester bonds in their immediate vicinity) to the endonucleolytic activity of DNase I [44]. Thirdly, because $(OP)_2Cu^+$ binds to the minor groove of DNA it will reveal minor-groove interactions (DMS will not, because its primary targeting sites on DNA are the N7 atoms of guanine residues which protrude into the major groove). The complex will also detect binding in the major groove when access to its minor-groove binding site is sterically blocked, or if the interaction of the protein in the major groove grossly distorts the geometry of the minor groove of the recognition sequence. In as much as binding of the co-ordination complex should be restricted to three base pairs in the minor groove [17] (Figure 4), $(OP)_2Cu^+$ is far more sensitive to local, protein-induced conformational changes than DNase I, which by possessing an extended minor-groove binding site, may be unable to sense. In addition, cleavage by DNase I first requires bending of the DNA molecule to the appropriate conformation [45]. This major-groove-centred and localized bend results in a widening of the minor groove [45] and therefore a possible rearrangement of the bound protein on its recognition sequence with various effects on the footprinting pattern. Furthermore, since the probability of cleavage at a particular base step is likely to be a measure of the ability of that base step and its adjacent neighbours to adopt the required conformation, the flexibility of the DNA in the immediate vicinity of the cleavage site will dictate the selectivity of cutting by DNase I, leading to reduced or enhanced susceptibility. None of the above considerations apply to the DNA-cleaving activity of $(OP)_2Cu^+$. Notably, unlike MPE·Fe(II), $(OP)_2Cu^+$ can be used in minimal amounts which do not perturb any DNA–protein interactions to be monitored.

A novel feature of $(OP)_2Cu^+$ as a footprinting tool is its ability to 'report' protein-induced changes in DNA [17]. This unique feature of the reactivity of $(OP)_2Cu^+$ may arise from the hydrophobic cation's affinity for protein surfaces and/or its intercalation (formation of a pi complex) into underwound DNA [8,9,46]. For instance, single-stranded regions produced at the active site of *Escherichia coli* RNA polymerase of the kinetically competent open transcription complex are efficiently cleaved on the template strand by $(OP)_2Cu^+$ [17,38,47]. Thus $(OP)_2Cu^+$ has been used to detect the series of intermediates formed during the initiation of transcription with the *lac* UV-5 promoter [17,48]. The inference that DNA downstream of the positions of nucleotide triphosphate incorporation is double-stranded relies on the observation of cutting on both strands. Scission sites on the template strand upstream of the site of nucleotide triphosphate incorporation are not associated with any nicking on the non-template strand. Methylation of cytosine at N4 with DMS has demonstrated that this region is unpaired [49]. Detection of these strained intermediates appears to be exclusive for the $(OP)_2Cu^+$ complex. Neither DNase I nor MPE·Fe(II) can detect RNA polymerase-induced changes at the start of transcription [8,46]. Moreover, unlike base-specific reagents such as DMS, diethyl pyrocarbonate and $KMnO_4$, $(OP)_2Cu^+$ detects all nucleotides in these single-stranded regions due to its cleavage neutrality.

Besides its serving as a reliable footprinting reagent in conventional protection or interference assays (e.g. [41]) (Figure 1), $(OP)_2Cu^+$ has been employed in an elegant protection footprinting method that incorporates mobility-shift technology prior to the chemical nuclease reaction [41]. In as much as the structural, thermodynamic and functional properties of DNA–protein complexes are not altered by entrapment in a polyacrylamide gel matrix [50], the small size and ready diffusibility of all reaction components in solid supports and the insensitivity of the reaction to inhibition by organic radical scavengers [even low concentrations of free acrylamide monomers and glycerol in gels interfere with the action of the MPE·Fe(II) and ferrous-EDTA reagents] permit the coupling of $(OP)_2Cu^+$ footprinting with the mobility-shift electrophoresis assay [41,51]. In this method, the DNA-binding reaction is performed and electrophoresed as usual and the entire mobility-shift gel is immersed in the footprinting reaction mixture. Footprints are obtained after direct elution of the free and bound deoxyoligonucleotide products from the gel and analysis on a sequencing gel (Figure 6). The use of DNase I and DMS in similar coupled assays (e.g. [52,53]) is subjected to the limitations imposed by their respective mechanisms of cutting/probing DNA (see above) and the requirement for subsequent treatment with base to effect cleavage (DMS).

The major advantage of this combined $(OP)_2Cu^+$ footprinting procedure arises from the topography of treatment: preformed DNA–protein complexes are exposed to the chemical nuclease while embedded in the gel matrix, i.e. not prior, but subsequent, to a mobility-shift experiment. This characteristic of the technique makes it ideal for protection analysis of transiently stable or extraordinarily labile complexes (e.g. [54]). At least three factors account for the latter. The first is that the background cleavage is greatly reduced by the separation of unbound DNA from the DNA–protein complex(es) pool before probing (Figure 6). The second factor is the so called 'gel-caging effect' [2,3]. The polyacrylamide matrix may provide 'cage-like' compartments that cause an increase in the effective concentration of the interacting components by slowing their diffusional escape after dissociation, leading to enhancement of the bimolecular (and thus concentration-dependent) reassociation reaction. It is also possible that the phenomenon arises from other factors, such as an influence of the high acrylamide percentage (5–10%) usually employed in mobility-shift gels on the thermodynamic activity of water (if bound water molecules are released from DNA or protein in forming the complex, then a lowered concentration of water in the gel will result in an augmented binding affinity) [55]. Whatever the mechanism(s), the increase in stability of the complex contributed by the gel leads to a more efficient blockage of the access of the $(OP)_2Cu^+$ chelate to the protein-binding DNA domain. The third factor comes from the nature and site of action of the cupryl intermediate through which the reaction is funnelled, and acts synergistically with the previous one. Because this highly reactive oxidative species is relatively short-lived and attacks DNA in the immediate vicinity of its minor-

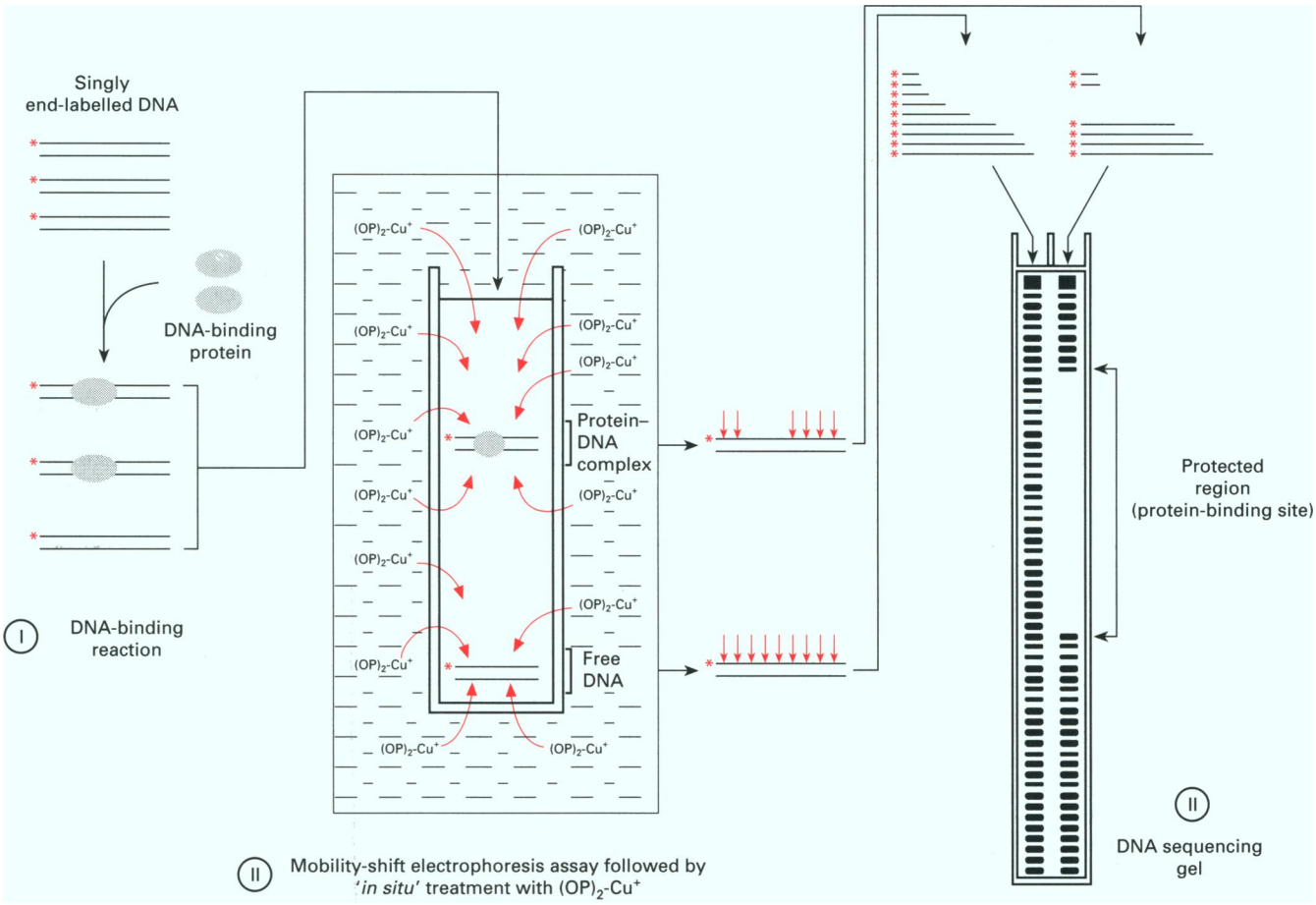

Figure 6 Rationale in the coupled mobility-shift electrophoresis/*in situ* bis(1,10-phenanthroline)·Cu(I) footprinting methodology

(**I**) Singly end-labelled DNA fragments (indicated as bars with the asterisks denoting the radioactive label) bearing a protein-binding site, and crude or partly purified cellular extracts containing the specific DNA-binding protein (represented by grey ovals) are mixed and incubated under previously established conditions. (**II**) Following equilibration of the DNA-binding reaction, the free and protein-complexed DNA fractions are resolved by native PAGE. The gel is subsequently immersed into a freshly prepared chemical reaction mixture and the gel-embedded free and protein-bound DNA species are subjected to nucleolytic attack by the gel matrix-diffused $(OP)_2Cu^+$ molecules (red arrows; nicks are introduced at a statistical and low level, i.e. less than one cut per DNA molecule). After quenching the nucleolytic activity with 2,9-$(CH_3)_2$-OP, the two DNA fragment populations are located by autoradiographic exposure of the wet gel, excised, and recovered from the gel matrix by standard techniques. (**III**) Comparable radioactivity from the two fractions is then analysed by denaturing PAGE and autoradiography. In the lane containing the free DNA sample, radioactive bands varying in intensity (reflecting the reactivity of each position, see text) will appear at all sequence positions. In the lane containing the sample that was derived from the complexed DNA fraction, radioactive bands will appear at all sequence positions except those contacted by the DNA-binding protein (protected region, demarcated by arrowheads connected with a line). Enhanced cleavage may be observed at the boundaries of the footprinted site, indicating protein-induced alterations in the conformation of adjacent DNA regions that rendered them more accessible to the chemical nuclease. The intense band at the top of each gel lane corresponds to full-length uncut DNA, representing the 'single-hit' kinetics conditions followed.

groove binding site without formation of diffusible radicals (Figure 5), it will be unable to achieve a fast equilibrium distribution along the DNA polymer. Consequently, protein-binding sites exposed during multiple dissociation events will, most of the time, escape the nucleolytic attack and hence remain intact. Since footprints of multiple complexes with defined stoichiometry/physical arrangement of their components and a vastly different kinetic character can be obtained simultaneously from crude nuclear extracts (e.g. [56,57]), the coupled assay obviates the need for pure protein preparations.

Interestingly, $(OP)_2Cu^+$ footprinting *in situ* following mobility-shift electrophoresis assays can also be applied to the analysis of RNA–protein complexes [58,59], whereas an alternative *in situ* approach utilizing Southwestern blots as the solid support has been recently described [60].

A thorough presentation of the experimental procedures followed in the entire gamut of $(OP)_2Cu^+$ footprinting applications exemplified by specific paradigms is given in [46,61]. For a detailed, step-by-step protocol of the gel $(OP)_2Cu^+$ mapping technique that should be of immediate value in typical laboratory situations, practitioners are referred to the corresponding chapter of a recently published methods book [62].

Ferrous-EDTA (hydroxyl radical)

The DNA-cleaving properties of the 'Fenton-like' (or Haber–Weiss) reaction [63–65] of the ferrous-EDTA complex [EDTA·Fe(II)] in the presence of an oxidant (H_2O_2) were exploited by Tullius and Dombroski for protection footprinting applications [29]. Three components are essential in this innocuous footprinting reaction mixture: $[Fe^{2+}(EDTA)]^{2-}$, H_2O_2 and sodium ascorbate. In the process of the reaction (schematically depicted in Figure 7) an electron from EDTA·Fe(II) reduces and breaks the O–O bond in H_2O_2, generating EDTA·Fe(III), the hydroxide ion and the powerfully

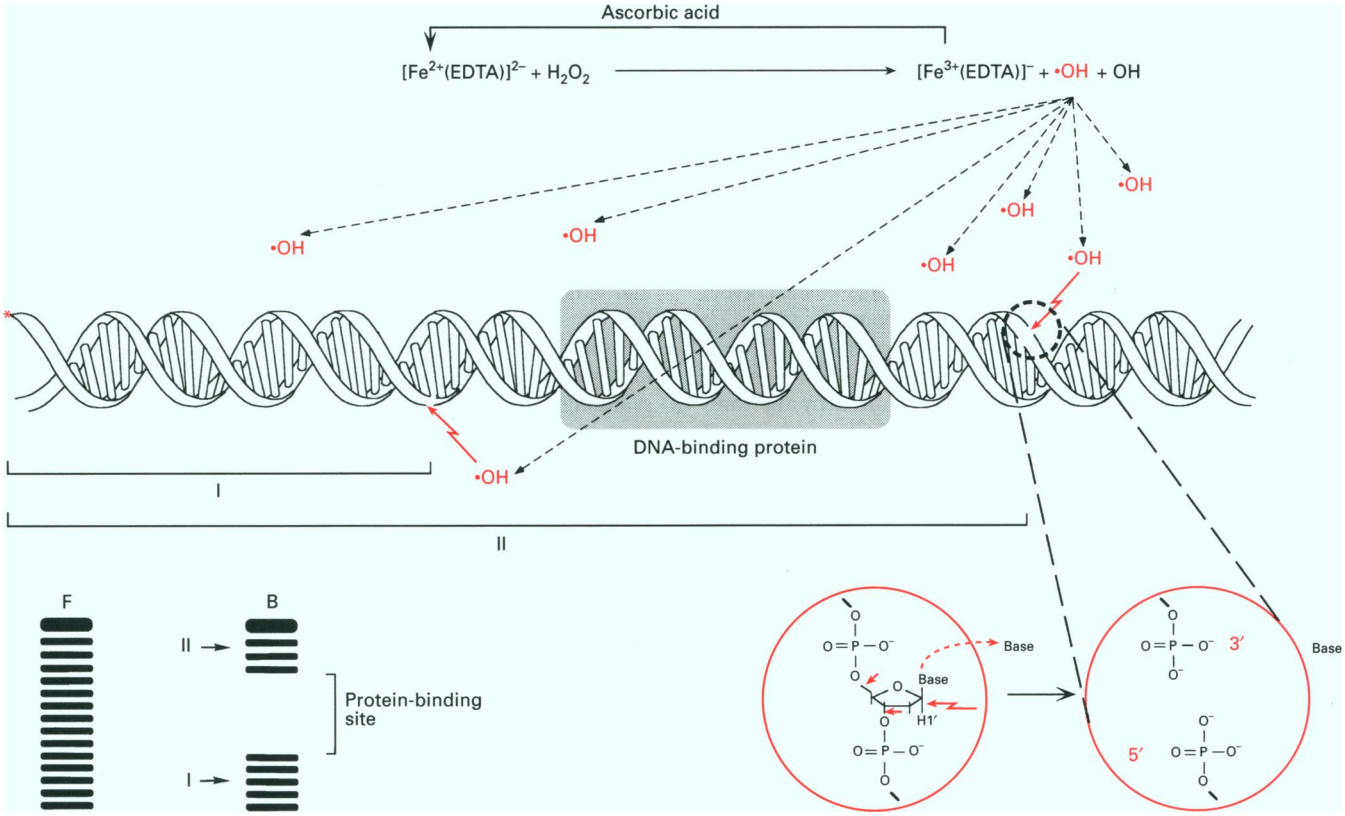

Figure 7 Strategy for the use of hydroxyl radical for studies of DNA–protein complexes

The EDTA·Fe(II) complex reacts with H_2O_2 in acidic medium with concomitant oxidation of Fe(II) and intermediate formation of highly reactive hydroxyl radicals (·OH) (Fenton reaction). Sodium ascorbate reduces Fe(III) to Fe(II), thus establishing a catalytic cycle. Because both the DNA and the metal complex are negatively charged, EDTA·Fe(II) is electrostatically repelled and therefore does not bind to the DNA molecule; thus the ultimate probe of protein binding is the hydroxyl radical. The hydroxyl radical diffuses through solution and attacks the DNA backbone (red zigzag arrows) most likely by abstracting a deoxyribose hydrogen atom (H1′ in this case, illustrated in the left projection at the lower right-hand part of the schematic), resulting in the loss of the sugar and base at that position in the DNA strand (red arrowheads and dashed arrow). A single-stranded nick is therefore produced into the DNA molecule, with the 3′- and 5′-phosphate moieties originally connected to the deoxyribose that was oxidized remaining intact (right projection at the lower right-hand part of the schematic). The singly end-labelled (indicated by an asterisk) DNA fragment is allowed to react for a length of time sufficient to statistically obtain about one scission event per DNA molecule ('single-hit' kinetics). Each cleavage point thus corresponds to a specific length of labelled DNA fragment as analysed by electrophoresis in a sequencing gel (lower left-hand part in the schematic), after elution of the free and protein-complexed DNA species from a mobility-shift gel (Figure 1, protection footprinting). F, free DNA; B, DNA recovered from the bound DNA fraction; Protein-binding site, missing (or under-represented) DNA fragments. The intense band at the top of the two gel lanes corresponds to full-length uncleaved DNA, reflecting the 'single-hit' kinetics conditions followed.

reactive hydroxyl radical (·OH), an exceedingly small and uncharged molecule of fleeting existence in solution. The ascorbate ion serves to reduce the Fe(III) product of the Fenton reaction back to Fe(II), thereby creating a catalytic cycle (Figure 7) and permitting micromolar concentrations of EDTA·Fe(II) to be effective in nicking DNA. A consequence of this scheme is that the concentration of the three chemical species may be varied to optimize the generation of the hydroxyl radical under different solution conditions (e.g. to compensate for the presence of organic radical traps such as glycerol in the DNA-binding buffer). This, however, can be technically challenging as even slightly elevated H_2O_2 concentrations have always the risk of adversely perturbing the DNA-binding ability of certain proteins, possibly by oxidizing sulphydryl groups that are necessary for the protein's 'active' structure. Similarly, the stability of the DNA–protein complex may be affected by the amount of EDTA used to produce the hydroxyl radical, if the protein(s) involved contains metal cofactors co-ordinated to cysteine residues (e.g. zinc fingers; a low level of EDTA does not seem to affect this DNA–metal protein interaction, since the exchange rate of zinc in the DNA–protein complex is probably much slower than it is in the free protein [66]). The hydroxyl radical initiates DNA-strand cleavage (in both single- and double-stranded DNA) most likely by abstraction (oxidation) of the C1′ and C4′ (and possibly C3′) hydrogen atoms from the deoxyribose moieties arrayed along the surface of DNA [11,12,66] (Figure 2). The subsequent breakdown of the deoxyribose-centred radical intermediate causes removal of the sugar and its linked base, giving as final products two DNA chains terminated by a 5′-phosphate and by 3′-phosphate and 3′-phosphoglycolate groups (at a ratio dependent on H_2O_2 concentration; [12]) respectively (Figures 2 and 7). Other minor products are formed as well, and thorough characterization of these is currently underway [67].

Although the chemistry of DNA cleavage is the same as with the MPE·Fe(II) reagent, the mechanism of image generation is different. Because EDTA·Fe(II) is negatively charged (Figure 7), it is subject to electrostatic repulsion by the polyanionic DNA molecule and therefore does not bind directly to it [68]. Instead,

it sits in solution at some distance from the DNA, reacting with H_2O_2 and bombarding DNA with the diffusible hydroxyl radical (Figure 7). The neutral hydroxyl radical, not the metal complex, is thus the probe of protein binding, eliminating the problem of putative alterations in DNA conformation induced by association of the metal species with the DNA molecule. However, this view has been challenged by studies claiming that, even in this case, DNA scission may proceed through very weak binding of EDTA·Fe(II) or formation of an intermediate iron-oxo complex or iron–DNA adducts [69,70]. Nonetheless, EDTA·Fe(II) footprints can be thought of to a large extent as maps of the solvent accessibility of the DNA surface in the presence of interacting protein (well-defined areas of low-cutting frequency), reflecting intimate protein–DNA backbone (deoxyribose) contacts [20, 68,71] (Figure 7). This interpretation is supported by the footprinting results obtained with λ repressor- and Cro protein-O_R1 operator complexes, revealing a correspondence between the deoxyriboses protected from EDTA·Fe(II) cleavage by the proteins [20] and the DNA–protein contacts determined by ethylation interference and X-ray crystallographic techniques [72]. Besides steric blockage of a cutting site, the rate of cleavage by hydroxyl radical (and therefore the footprinting pattern) may be influenced by protein-induced changes in the conformation (pucker) of the deoxyribose as a result of their effect on the stability of the initial radical generated at the site of attack.

As a consequence of: (i) its tiny dimensions (it is the smallest chemical probe available, comparable in size to a water molecule); (ii) subtle sequence variability in its nucleolytic activity (the site of attack on the nucleotides is the deoxyribose and not the heterocyclic base, so that an even cleavage pattern is obtained in mixed-sequence DNA [12,29,73]); and (iii) reaction rates that approach the diffusion limit (DNA cleavage does not depend on equilibrium binding of the reagent, hence the time course of a process can be followed for times as short as ~ 30 s; [65,74]), the hydroxyl radical produces tight footprints with very high resolution, which are usually smaller than those obtained with the DNA-binding probes DNase I, MPE·Fe(II), and 1,10-phenanthroline·Cu(I). For instance, positions on the 'backside' of the DNA–λ repressor complex that are exposed to the solvent but are inaccessible to the large DNase I molecule react readily with the hydroxyl radical [20]. Instead of the complete 'blank' 30-nucleotide footprint yielded by DNase I, the footprint generated by hydroxyl radical exhibits small patches of protection only where the protein directly covers the DNA backbone, giving a 'sharp' view of the structure and symmetry of the DNA–protein complex. For the same reasons, hydroxyl radical is more suitable than DNase I and the aforementioned chemical nucleases as a probe for measuring thermodynamic parameters and cooperativity of protein–DNA binding by footprint titration experiments (i.e. derivation of individual-site binding isotherms for the interaction of a protein with multiple DNA recognition sequences by quantitative analysis of hydroxyl radical footprints generated at various protein concentrations) [66].

A disadvantage of its small size is the rather small difference between the rates of cleavage of protein-bound DNA and naked DNA. As a result, the footprint obtained is often more subtle in appearance (i.e. not easily discernible by eye) than those produced by other probes and cannot be evaluated visually with accuracy (a representative paradigm of this type of signal pattern is shown in Figure 8). A more precise and objective method is to subtract the cutting pattern of the free DNA from that of the complexed DNA so that the effects of protein binding can be seen more clearly. This involves calculating the probability of cleavage at each phosphodiester bond, which is related to the amount of radioactivity (or intensity) in the corresponding band in the digestion pattern, as measured by scanning the autoradiograph with one- or two-dimensional densitometers [75,76]. Computer-aided summation of all probabilities of cleavage in a lane provides the average number of cuts in a DNA strand, while the probabilities of cutting at each band in the complexed and free DNA are used to calculate a difference probability along the entire length of the protein-binding site and surrounding DNA. In the derived difference probability plot the difference in the extent of cleavage at each nucleotide is represented by a bar, with negative values showing nucleotide residues that are protected and values close to zero indicating a reactivity similar to that of the uncomplexed DNA [75,76]. The advanced technology of phosphor-screen imagers (providing a substitute for X-ray film autoradiography) and the current generation of two-dimensional scanning densitometers, combined with the development of more sophisticated data analysis procedures [67] will permit previously impossible experiments to be performed and justify the high degree of accuracy required for this kind of quantitative manipulation.

Because of its uniform cutting at each backbone position, the hydroxyl radical is also a suitable reagent for studying the helical periodicity of DNA molecules that contain regulatory sequences (i.e. promoters) and therefore mapping the three-dimensional relationship of protein-binding sites along them [29,68]. In as much as the measured rate of cutting reflects the reactivity or accessibility of the target atom or bond in each nucleotide toward attack by the hydroxyl radical, this information could be used as an aid in orienting a protein domain on its binding site, in such a way that the observed pattern of reactivity is generated [66]. Thus if both strands exhibit a footprint at the same region it indicates that the protein wraps around the DNA, while modulation of the band intensity with a regular phasing according to the helix repeat is indicative of protein binding to one side of the DNA (in the latter case the complementary strand should exhibit the same pattern but with an offset of two or three bases, as a consequence of the helicity of the DNA). Due to the moderately large reductions in cutting (i.e. discontinuities in the cutting pattern) arising from a narrowing of the minor groove, hydroxyl radical can also be employed for the detection of 'inherent' or protein-induced DNA bending at a particular DNA sequence [20,68,77].

Finally, in a novel interference approach, DNA pretreated with EDTA·Fe(II) (and thus containing, on average, fewer than one randomly placed single-nucleoside gap per DNA fragment) can be used to study the effects of missing nucleosides in DNA–protein recognition, and hence to determine in just one experiment at which positions in a DNA molecule the base moieties make energetically important contacts with a sequence-specific DNA-binding protein (missing nucleoside experiment; [66,78]). Accordingly, the missing nucleoside approach is superior to methylation interference assays which yield information for only a subset of base-specific contacts (and, unfortunately, only contacts with purines) [15], as well as to the related 'missing contact' method [79] which assesses the contribution to protein binding of each member of a base pair independently and requires additional chemical reactions (Maxam–Gilbert sequencing chemistry; [4]). Furthermore, because the missing nucleoside is a centre of enhanced flexibility in the DNA molecule, the effect of DNA deformability on the DNA–protein interaction can be studied. Although this technique permits an ordering of the strength of interaction of the DNA-binding protein with the different nucleotides within the recognition sequence (thus obliterating the need for arduous and time-consuming mutational analyses), the conclusions are only strictly applicable to the binding of the protein to gapped DNA [8,78].

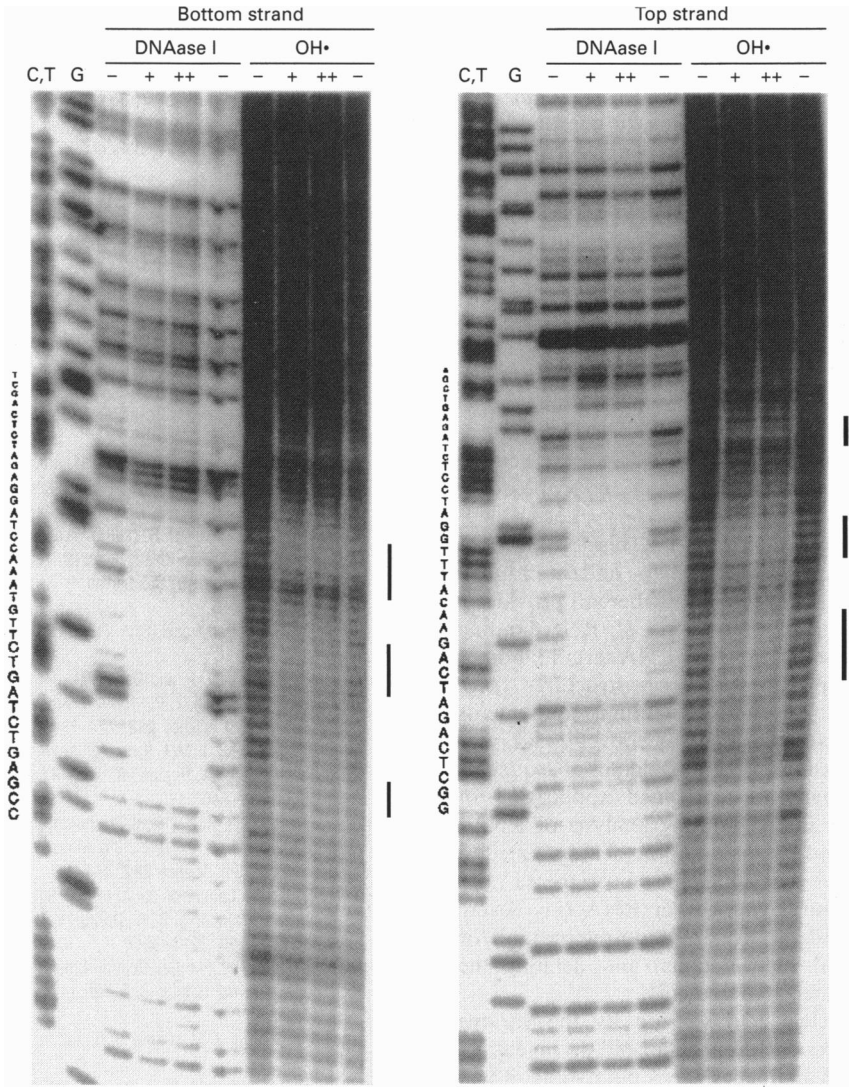

Figure 8 A typical example illustrating the subtle appearance of hydroxyl radical footprints

A DNA fragment containing a binding site for the glucocorticoid receptor was ^{32}P-labelled on either the top or bottom strand and incubated with increasing concentrations (+, ++) of the DNA-binding domain of the glucocorticoid receptor over-expressed and purified from *Escherichia coli* cells, or left without protein (—). The two reactions were subsequently subjected to either DNase I or hydroxyl radical (OH·) treatment. Nucleotides protected from hydroxyl radical cleavage are indicated by solid bars. Lanes G and C,T contain probe DNA digested with guanosine-specific or cytidine- and thymidine-specific reagents respectively. The nucleotide sequence in the footprint region is shown on the left of each panel. (Reprinted from Garabedian, M. J., Labaer, J., Liu, W.-H. and Thomas, J. R. (1993) in Gene Transcription: A Practical Approach (Hames, B. D. and Higgins, S. J., eds), pp. 243–293 [83] by permission of Oxford University Press.)

A comprehensive description of the experimental protocols employed in hydroxyl radical footprinting approaches illustrated for a variety of systems and including quantitative treatment of the data obtained can be found in [66,67].

CONCLUDING REMARKS AND FUTURE PROSPECTS

It has hopefully emerged from the preceding presentation that chemical nucleases are to the field of DNA–protein interactions what affinity labels and binding/active-site-targeted inhibitors are to protein and enzyme systems. They currently provide a quiver of sophisticated footprinting strategies for analysing the topography of protein-binding sites and 'outlining' the equilibria and energetics of DNA–protein assemblies in solution and in solid supports.

The nucleotide-level resolution of chemical nuclease mapping approaches allows the description of DNA–protein complexes of vastly different character at a level of detail not afforded by other more widely used footprinting technologies. In this regard, chemical nucleases have already enriched the palette available to researchers wishing to explore DNA structure/function relationships, and further modifications of their DNA-scission properties will probably render these reagents even more efficient and broadly applicable. For instance, a route of future investigation is the possibility of hydrolytic cleavage by the respective co-ordination complexes with a non-redox-active metal atom. Conformational anomalies (e.g. bends or kinks) in the bound DNA could enhance the relatively slow kinetics of their chemical mechanism by introducing bond strain in the phosphodiester backbone. In the same vein, derivatives of 1,10-phenanthroline could be designed with profound impacts on association of the co-ordination complex with DNA containing various structural perturbations. Similarly, certain electron-donating groups incorporated into the chemical nuclease structure could drastically

influence the DNA-cleavage event by altering the efficiency with which the modified $(OP)_2Cu^+$ complex is oxidized by molecular oxygen producing H_2O_2 [61] (Figure 5).

A parallel research avenue involves the covalent attachment of chemical nucleases to DNA-binding proteins (or conformationally stable, chemically synthesized peptides derived from their DNA-binding domains) via various linkers (reviewed in [8,9,46]). The fast progress in the development of intriguing procedures for synthesizing such chemical nuclease-modified proteins, in conjuction with the aforementioned structural manipulations of chemical nucleases, should endow the field with a panoply of novel nucleolytic species bearing a wide spectrum of specificities and efficiencies in cleaving DNA. These powerful tools could complement other approaches in studies aiming at inferring the precise orientation of a DNA-binding protein with respect to its DNA-recognition sequence. Furthermore, the scission patterns obtained by protein targeting of the chemical nuclease activity offer a test for models of protein–DNA interactions proposed by crystallographic or model-building studies.

The nucleolytic activity of MPE·Fe(II), bis(1,10-phenanthroline)·Cu(I) and EDTA·Fe(II) on single- and/or double-stranded RNA molecules ([9] and references therein) provides a viable means of mapping the interactions of RNA-binding proteins with their cognate sequences (e.g. RNA termini, homopolymeric sequences, or specific stem–loop structures [80]), circumventing the nucleotide/secondary structure requirements and the experimental constraints imposed by conventional ribonucleases [59]. Since little is known about the details of sequence-specific RNA–protein recognition, combined application of different chemical nucleases in footprinting analysis of RNA–protein complexes will be (and has already been demonstrated to be [81]) invaluable toward revealing protein-triggered and functionally essential structural changes in RNA (i.e. severe perturbations of single-stranded bulge and loop regions and/or unwinding of double-helical stem regions) and defining the binding surfaces of RNA-associated proteins.

As the chemistry of nucleic acid scission by the various chemical nucleases is understood in more detail and the inventiveness of developing improved analogues keeps flourishing, it will be possible to pinpoint more precisely the stereochemistry of DNA–protein and RNA–protein complexes and to illuminate the 'functional' dynamics of these systems. This knowledge will enable us to advance finer structural and kinetic models for DNA/RNA–protein interactions in cellular extracts or inside cells, and in this way to start tackling chemical problems underlying the biological expression of genetic information. Such problems include the role exerted by local DNA polymorphism in the way a protein locates its recognition sequence within a large and otherwise monotonous linear DNA molecule; the assembly mechanism of complex, and sometimes transient, three-dimensional nucleoprotein structures in which genetically distant elements are brought into close spatial proximity to direct the transcription, replication, or recombination machineries; the determination of rate-limiting steps (and how they are relieved) in promoter/enhancer occupation by multiprotein species assembled along pathways defined by distinct developmental and physiological stages of the cell; and the elucidation of the RNA–protein recognition code(s) operating in cellular processes such as spliceosome and ribosome biogenesis, mRNA transport and loading of the translational apparatus. Gaining insight into these 'daunting' questions will undoubtedly spawn the next generation of breakthroughs.

I am grateful to Sigrid Lazzaro and Petra Riedinger of the EMBL PhotoLab for invaluable help with computer drawings.

REFERENCES

1 Travers, A. A. (1993) DNA–Protein Interactions, pp. 52–86, Chapman & Hall, London
2 Fried, M. and Crothers, D. M. (1981) Nucleic Acids Res. **9**, 6505–6525
3 Garner, M. M. and Revzin, A. (1981) Nucleic Acids Res. **9**, 3047–3060
4 Maxam, A. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 560–564
5 Sigman, D. S., Graham, D. R., D'Aurora, V. and Stern, A. M. (1979) J. Biol. Chem. **254**, 12269–12272
6 Hertzberg, R. P. and Dervan, P. B. (1982) J. Am. Chem. Soc. **104**, 313–314
7 Pope, L. E. and Sigman, D. S. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 3–7
8 Sigman, D. S. (1990) Biochemistry **29**, 9097–9105
9 Sigman, D. S. and Chen, C. B. (1990) Annu. Rev. Biochem. **59**, 207–236
10 Pope, L. E., Reich, K. A., Graham, D. R. and Sigman, D. S. (1982) J. Biol. Chem. **257**, 12121–12128
11 Henner, W. D., Grunberg, S. M. and Haseltine, W. A. (1983) J. Biol. Chem. **258**, 15198–15205
12 Hertzberg, R. P. and Dervan, P. B. (1984) Biochemistry **23**, 3934–3945
13 Kuwabara, M. D., Yoon, C., Goyne, T. E., Thederahn, T. B. and Sigman, D. S. (1986) Biochemistry **25**, 7401–7408
14 Goyne, T. E. and Sigman, D. S. (1987) J. Am. Chem. Soc. **109**, 2846–2848
15 Siebenlist, U. and Gilbert, W. (1980) Proc. Natl. Acad. Sci. U.S.A. **77**, 122–126
16 Van Dyke, M. W. and Dervan, P. B. (1983) Nucleic Acids Res. **11**, 5555–5567
17 Spassky, A. and Sigman, D. S. (1985) Biochemistry **24**, 8050–8056
18 Sigman, D. S. (1986) Acc. Chem. Res. **19**, 180–186
19 Schultz, P. G., Taylor, J. S. and Dervan, P. B. (1982) J. Am. Chem. Soc. **104**, 6861–6863
20 Tullius, T. D. and Dombroski, B. A. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 5469–5473
21 Ward, B., Skorobogaty, A. and Dabrowiak, J. C. (1986) Biochemistry **25**, 6875–6883
22 Groves, J. T. and Farrell, T. P. (1989) J. Am. Chem. Soc. **111**, 4998–5000
23 Barton, J. K. (1986) Science **233**, 727–734
24 Nielsen, P. E. (1990) J. Mol. Recog. **3**, 1–25
25 Yoon, C., Kuwabara, M. D., Law, R., Wall, R. and Sigman, D. S. (1988) J. Biol. Chem. **263**, 8458–8463
26 Veal, J. M. and Rill, R. L. (1988) Biochemistry **27**, 1822–1827
27 Veal, J. M. and Rill, R. L. (1989) Biochemistry **28**, 3243–3250
28 Dervan, P. B. (1986) Science **232**, 464–471
29 Tullius, T. D. and Dombroski, B. A. (1985) Science **230**, 679–681
30 Cartwright, I. L., Hertzberg, R. P., Dervan, P. B. and Elgin, S. C. R. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 3213–3217
31 Goodwin, G. H. (1990) in Gel Electrophoresis of Nucleic Acids: A Practical Approach (Rickwood, D. D. and Hames, B. D., eds.), 2nd edn., pp. 225–247, Oxford University Press, Oxford
32 Landolfi, N. F., Yin, X.-M., Capra, J. D. and Tucker, P. W. (1989) Biotechniques **7**, 500–504
33 Tullius, T. D., Dombroski, B. A., Churchill, M. E. A. and Kam, L. (1987) Methods Enzymol. **155**, 537–558
34 Sawadogo, M. and Roeder, R. G. (1985) Cell **43**, 165–175
35 Van Dyke, M. W., Roeder, R. G. and Sawadogo, M. (1988) Science **241**, 1335–1338
36 Veal, J. M., Merchant, K. and Rill, R. L. (1991) Nucleic Acids Res. **19**, 3383–3388
37 Marshall, L. E., Graham, D. R., Reich, K. A. and Sigman, D. S. (1981) Biochemistry **20**, 244–250
38 Thederahn, T. B., Spassky, A., Kuwabara, M. D. and Sigman, D. S. (1990) Biochem. Biophys. Res. Commun. **168**, 756–762
39 Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V. and Kopka, M. L. (1982) Science **216**, 475–485
40 Sigman, D. S. and Spassky, A. (1989) in Nucleic Acids and Molecular Biology (Eckstein, F. and Lilley, D. M. J., eds.), vol. 3, pp. 13–27, Springer-Verlag, Berlin
41 Kuwabara, M. D. and Sigman, D. S. (1987) Biochemistry **26**, 7234–7238
42 Tamilarasan, R., McMillin, D. R. and Liu, F. (1989) in Metal–DNA Chemistry (Tullius, T. D., ed.), pp. 48–58, American Chemical Society Symposium Series **402**, Washington DC
43 Yoon, C., Kuwabara, M. D., Spassky, A. and Sigman, D. S. (1990) Biochemistry **29**, 2116–2121
44 Lomonossoff, G. P., Butler, P. J. G. and Klug, A. (1981) J. Mol. Biol. **149**, 745–760
45 Suck, D., Lahm, A. and Oefner, C. (1988) Nature (London) **332**, 464–468
46 Sigman, D. S., Kuwabara, M. D., Chen, C.-H. B. and Bruice, T. W. (1991) Methods Enzymol. **208**, 414–433
47 Spassky, A., Rimsky, S., Buc, H. and Busby, S. (1988) EMBO J. **7**, 1871–1879
48 Spassky, A. (1986) J. Mol. Biol. **188**, 99–103
49 Kirkegaard, K., Spassky, A., Buc, H. and Wang, J. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 2544–2548
50 Revzin, A., Ceglarek, J. A. and Garner, M. M. (1986) Anal. Biochem. **153**, 172–177
51 Kakkis, E. and Calame, K. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 7031–7035
52 Papavassiliou, A. G. (1993) Methods Mol. Cell. Biol. **4**, 95–104
53 Papavassiliou, A. G. (1993) Nucleic Acids Res. **21**, 757–758

54 Papavassiliou, A. G. and Silverstein, S. J. (1990) J. Biol. Chem. **265**, 9402–9412
55 Crothers, D. M., Gartenberg, M. R. and Shrader, T. E. (1991) Methods Enzymol. **208**, 118–146
56 Law, R., Kuwabara, M. D., Briskin, M., Fasel, N., Hermanson, G., Sigman, D. S. and Wall, R. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 9160–9164
57 Flanagan, W. M., Papavassiliou, A. G., Rice, M., Hecht, L. B., Silverstein, S. and Wagner, E. K. (1991) J. Virol. **65**, 769–786
58 Del Angel, R. M., Papavassiliou, A. G., Fernandez-Tomas, C., Silverstein, S. J. and Racaniello, V. R. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 8299–8303
59 Papavassiliou, A. G. (1993) Anal. Biochem. **214**, 331–334
60 Polycarpou-Schwarz, M. and Papavassiliou, A. G. (1993) Methods Mol. Cell. Biol. **4**, 22–26
61 Mazumder, A. (1993) in Footprinting of Nucleic Acid–Protein Complexes (Revzin, A., ed.), pp. 45–73, Academic Press, San Diego, CA
62 Papavassiliou, A. G. (1994) in DNA-Protein Interactions: Principles and Protocols (Kneale, G. G., ed.), pp. 43–78, Humana Press, Totowa, NJ
63 Fenton, H. J. H. (1894) J. Chem. Soc. **65**, 899–910
64 Udenfriend, S., Clark, C. T., Axelrod, J. and Brodie, B. B. (1954) J. Biol. Chem. **208**, 731–739
65 Walling, C. (1975) Acc. Chem. Res. **8**, 125–131
66 Dixon, W. J., Hayes, J. J., Levin, J. R., Weidner, M. F., Dombroski, B. A. and Tullius, T. D. (1991) Methods Enzymol. **208**, 380–413
67 Bashkin, J. S. and Tullius, T. D. (1993) in Footprinting of Nucleic Acid–Protein Complexes (Revzin, A., ed.), pp. 75–106, Academic Press, San Diego, CA
68 Tullius, T. D. (1987) Trends Biochem. Sci. **12**, 297–300
69 Jezewska, M. J., Bujalowski, W. and Lohman, T. M. (1989) Biochemistry **28**, 6161–6164
70 Jezewska, M. J., Bujalowski, W. and Lohman, T. M. (1990) Biochemistry **29**, 5220
71 Tullius, T. D. (1988) Nature (London) **332**, 663–664
72 Jordan, S. R. and Pabo, C. O. (1988) Science **242**, 893–899
73 Celander, D. W. and Cech, T. R. (1990) Biochemistry **29**, 1355–1361
74 Rush, J. D., Maskos, Z. and Koppenol, W. H. (1990) Methods Enzymol. **186**, 148–156
75 Drew, H. R. and Travers, A. A. (1985) J. Mol. Biol. **186**, 773–790
76 Rhodes, D. (1989) in Protein Function: A Practical Approach (Creighton, T. E., ed.), pp. 177–198, Oxford University Press, Oxford
77 Burkhoff, A. M. and Tullius, T. D. (1987) Cell **48**, 935–943
78 Hayes, J. J. and Tullius, T. D. (1989) Biochemistry **28**, 9521–9527
79 Brunelle, A. and Schleif, R. F. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 6673–6676
80 Nagai, K. (1992) Curr. Opin. Struct. Biol. **2**, 131–137
81 Frankel, A. D., Mattaj, I. W. and Rio, D. C. (1991) Cell **67**, 1041–1046
82 Vu, D. T. and Stanbury, D. M. (1987) Inorg. Chem. **26**, 1732–1736
83 Garabedian, M. J., Labaer, J., Liu, W.-H. and Thomas, J. R. (1993) in Gene Transcription: A Practical Approach (Hames, B. D. and Higgins, S. J., eds.), pp. 243–293, Oxford University Press, Oxford