Mapping Protein Domains Involved in Macromolecular Interactions: A Novel Protein Footprinting Approach†

Ewa Heyduk and Tomasz Heyduk*

Department of Biochemistry and Molecular Biology, St. Louis University, Medical School, 1402 South Grand Boulevard, St. Louis, Missouri 63104

Received March 4, 1994; Revised Manuscript Received June 8, 1994®

ABSTRACT: A novel direct approach, analogous to DNA footprinting, for mapping protein domains involved in macromolecular interactions is presented in this paper and applied to cAMP receptor protein (CRP) interactions with the allosteric ligand (cAMP) and DNA. In this approach, a protein–macromolecule complex is subjected to a nonspecific cleavage by Fe–EDTA. The cleavage products are resolved by SDS–PAGE and transferred to a PVDF membrane. Transferred polypeptides are visualized by immunostaining with antibodies specific to the N-terminal peptide of the protein. The mobility of the bands visualized in such a way is directly proportional to the distance of the cleavage sites from the N-terminus, and thus the positions of the sites protected from cleavage by a bound macromolecule can be determined. Thus, protein domains involved in macromolecular interactions can be mapped. In the case of CRP, the cleavage conditions were established which resulted in, on the average, less than one cleavage event/protein molecule and which preserved satisfactory levels of protein and DNA activity. When applied to CRP–DNA interactions, the protein footprinting approach correctly identified domains of CRP that were known to be involved in the recognition of DNA. The obtained results showed also that the binding of CRP to the DNA binding site perturbed the region of CRP involved in intersubunit interactions. An allosteric ligand (cAMP) appeared to perturb the same region of CRP. This stresses out the importance of intersubunit interactions in cAMP modulation of protein DNA binding affinity. The protein footprinting methodology presented in this paper should be broadly generalizable to any protein–macromolecule system.

A great number of biological functions of proteins are mediated via interactions with other macromolecules. In such cases, one of the essential steps in understanding those functions is to identify the regions of a protein molecule which participate in the protein–macromolecule complex formation. When a three-dimensional structure of a complex is not known, this task is quite difficult. The two most often used methods to map protein domains participating in macromolecular interactions involve chemical modification methods and mutagenesis. In the former case, the mapping is accomplished by studying the protection against chemical modification (including proteolysis) of the protein [for example, Burnens et al. (1987)] or by chemical cross-linking of interacting macromolecules [for example, Gentry and Burgess (1993)]. The chemical methods are often very laborious, and the identification of protected (or cross-linked) regions of the protein is difficult. In mutagenesis approaches, the loss of protein function (for which the protein–macromolecule interaction is essential) is studied in response to alteration of protein structure by site-directed or deletion mutagenesis [for example, Zhou et al. (1993) and Lesley and Burgess (1989)]. This suffers from low resolution of mapping (deletion mutagenesis), and it is often necessary to assume without the experimental proof that the observed effects of a mutation are caused by a direct perturbation of the protein–macromolecule interaction rather than by an indirect effect on the protein conformation. Thus, a simple methodology to map protein contact domains is lacking.

DNA footprinting is a highly successful approach to study DNA domains involved in macromolecular interactions (Galas & Schmitz, 1978). In this approach, DNA is end-labeled with a radioactive probe, and the complex of the labeled DNA and a macromolecule is subjected to a limited degradation by nuclease. The products of the cleavage are separated by electrophoresis. The relative mobility of end-labeled DNA fragments is directly related to the distance of the nuclease cleavage site to the end of the DNA molecule. Thus, the precise position of a macromolecule binding site on DNA can be easily determined by observing which nuclease cleavage sites become protected in the presence of a macromolecule. The success of the DNA footprinting approach is largely due to the simplicity of identifying the cleavage sites in the DNA molecule by electrophoretic separation of end-labeled DNA molecules (Galas & Schmitz, 1978). Jue and Doolittle (1985) were first to note and experimentally confirm that a similar approach to identifying cleavage sites in the case of proteins is feasible as well. However, specific end-labeling of native proteins is difficult, and proteolytic enzymes are unsuitable for high-resolution footprinting since they usually attack only a very limited number of sites in a native protein. Also, the resistance to proteolytic degradation varies greatly between proteins. It has been known for some time that the radicals generated by a metal/H_2O_2 system can degrade polypeptide chains (Phelps et al., 1961). Recently, several groups have shown a specific cleavage of a protein by a covalently or noncovalently attached Fe–EDTA complex (Rana & Meares, 1990, 1991a; Hoyer et al., 1990; Schepartz & Cuenoud, 1990; Ermacora et al., 1992). The mechanism of this cleavage reaction was studied using a model peptide system, and it was shown that the Fe–EDTA complex can be an effective, nonspecific chemical protease (Platis et al., 1993). The cleavage of polypeptides by the Fe–EDTA complex can involve

† This work was partially supported by a grant from the American Cancer Society (NP-72785).

* Corresponding author. Phone: (314) 577-8152; Fax: (314) 577-8156; E-mail address: heyduk@sluava.slu.edu.

® Abstract published in *Advance ACS Abstracts,* July 15, 1994.

0006-2960/94/0433-9643$04.50/0

both diffusible hydroxy radicals (Hoyer et al., 1990; Platis et al., 1993) as well as a Fe–EDTA–peroxo species (Rana & Meares, 1991b). The Fe–EDTA complexes will also cleave DNA, and they have been successfully used for footprinting protein binding sites on DNA (Tulius & Dombroski, 1986). Thus, we reasoned that Fe–EDTA could be also used to footprint DNA (or other macromolecule) binding sites on proteins. Using this nonspecific chemical protease rather than proteolytic enzymes (or specific chemical cleavages) could greatly increase the resolution of protein footprinting methodology and make it applicable to any protein. By combining the chemical protease cleavage of a protein–macromolecule complex with the electrophoretic identification of the cleavage site positions (Matsudaira et al., 1985; Bazari et al., 1985; Arruda et al., 1992), we established a relatively simple and general protein footprinting approach.

In this paper, we describe the protein footprinting methodology and the results of its application to CRP (the bacterial cAMP receptor protein) interaction with DNA. CRP, a transcription activator of many genes in *Escherichia coli*, binds to DNA in a cAMP-modulated fashion [for the recent review, see Kolb et al. (1993)]. The crystal structures of cAMP–CRP (Weber & Steitz, 1987) and the cAMP–CRP–DNA complex (Schultz et al., 1991) are known, which permitted critical evaluation of our protein footprinting approach. The results obtained showed excellent agreement between the information obtained from protein footprinting experiments and the crystal structure of a protein–DNA complex. Also, the results showed that the interaction of CRP with cAMP and DNA perturbed the subunit–subunit interface of the protein. This points to intersubunit interactions being an important element of the transmission of information from cAMP to the DNA binding domain of CRP.

MATERIALS AND METHODS

Materials. CRP was purified as described previously (Heyduk & Lee, 1989). The concentration of the protein was determined using the absorption coefficient of 20 400 M^{-1} cm^{-1} at 278 nm (Takahashi et al., 1980). A 42 base pair (bp) DNA fragment corresponding to ICAP consensus sequence of CRP binding site (Ebright et al., 1989) was chemically synthesized. The single-stranded complementary oligonucleotides were purified on denaturing polyacrylamide gel and were hybridized as described previously (Heyduk & Lee, 1990). The concentration of the DNA fragment was determined by UV absorption at 260 nm, assuming that 1 OD unit corresponds to 50 μg of ds DNA. Alkaline phosphatase-conjugated anti-rabbit goat IgG, BCIP (5-bromo-4-chloro-3-indolyl phosphate), and NBT (nitro blue tetrazolium) were from Bio-Rad.

Preparation of Antibodies to N-Terminal Peptide of CRP. A peptide, Val-Leu-Gly-Lys-Pro-Glu-Thr-Asp-Cys, corresponding to a squence of the first eight N-terminal residues of CRP with additional Cys residues at its carboxy end was coupled to thyroglobulin. The coupling was performed through the peptide Cys residue and the thyroglobulin Lys residues using a bifunctional cross-linker, 3-maleimidobenzoic acid *N*-hydroxysuccinimide (Doolittle, 1986). Rabbits were injected with 2 mg of conjugated peptide in complete Freund adjuvant boosted with 2 mg 28 days later and 1 mg 44 and 59 days later in incomplete Freund adjuvant. The collected serum was stored at −70 °C and used for immunostaining without further processing.

To verify the specificity of our antiserum, we produced a CRP derivative in which the eight N-terminal amino acids used to generate antibodies were removed. This was achieved by limited acid hydrolysis of a single Asp-Pro bond (Piszkiewicz et al., 1970) between residues 8 and 9. This derivative was not detected on a Wester blot (when 50 ng was loaded).

Footprinting Reactions. All reactions were performed at room temperature in 10 mM MOPS buffer (pH 7.2) containing 250 mM NaCl and 10 mM $MgCl_2$. Typically, 50 μL of 5 μM CRP was subjected to nonspecific cleavage by adding freshly prepared $FeSO_4$, ascorbate, EDTA, and H_2O_2 to a final concentration of 1, 20, 2, and 1 mM, respectively. Later experiments showed that the footprinting reactions can be carried out with much lower concentrations of reactants. The optimal conditions seemed to be 0.2, 5, 2, and 1 mM for $FeSO_4$, ascorbate, EDTA, and H_2O_2, respectively. However, for the sake of consistency, all the experiments presented in this paper were performed in the initially established conditions, i.e., higher concentrations of reagents. The cAMP and DNA were 100 and 5 μM, respectively, when present in the reaction mixture. The reaction was allowed to proceed for 10 min (unless time dependence of the reaction was studied) and was stopped by adding 0.25 vol of 4x sample buffer [16% SDS, 48% glycerol, 200 mM Tris (pH 6.8), 8% 2-mercaptoethanol, 0.04% bromophenol blue]. Samples were stored at −70 °C until they were loaded on the gel. Control experiments showed that the addition of a sample buffer and freezing effectively stops the cleavage reaction.

SDS–PAGE Electrophoresis and Transfer to Membrane. The cleavage products were resolved on the discontinuous Tricine–SDS–PAGE gel system described by Schagger and von Jagow (1987), which was shown to resolve polypeptides down to about 2 kDa. A 16.5% separating and 10% stacking gels were used. The electrophoresis was performed in 16 cm long slab gels using S-600 electrophoresis apparatus (Hoeffer) at room temperature. Protein samples were allowed to enter the gel at 15–30 mA, and then the gel was run at constant 40–60 mA current. Using this system, when the gel was run until bromophenol blue ran out of the gel, one residue difference in size for a polypeptide of approximately 20 kDa gave a difference of about 0.2 mm in mobility. This resolution could be improved by running longer gels or by tailoring the running time for the range of polypeptide sizes of interest.

After electrophoresis, the samples were electroblotted to a PVDF membrane (MSI). The transfer was performed in 25 mM Tris, 192 mM glycine, and 15% methanol buffer (pH 8.3) for 1.5 h at room temperature in a Genie electrophoretic blotter (Idea Scientific). The lane containing molecular weight standards (low range molecular weight standards from Promega) was stained with 0.1% Coomassie Blue in 50% methanol/10% acetic acid and destained with 10% acetic acid/50% methanol. The remainder of the membrane was blocked (2 h to overnight) with 5% nonfat milk in 50 mM Tris (pH 7.4) and 150 mM NaCl (TBS). After being washed three to four times for 5 min with TBS buffer containing 0.05% Tween 20 (TBST buffer), the membrane was incubated for 2 h with constant shaking with a 1:1000 dilution of CRP N-terminal peptide-specific antisera in TBST buffer. After being washed with TBST as described above, the membrane was incubated for 2 h with 1:3000 dilution of alkaline phosphatase-conjugated secondary antibody. After being washed with TBST, the bands were visualized by incubation with a BCIP/NBT substrate mixture (0.15 and 0.30 mg/mL, respectively) in 100 mM Tris (pH 9.5) and 0.5 mM $MgCl_2$. Once the color has developed to the desired intensity, the reaction was stopped by washing the membrane with distilled water, and the membrane was dried.

Assignment of Bands. The positions of cleavage sites corresponding to bands on a Western blot of resolved cleavage products were identified by their molecular weights. Relative mobility of molecular weight markers was used to obtain the standard curves of $\log(M_w)$ vs relative mobility. Low range molecular weight markers from Promega were used with the following molecular weights (in thousands): 31.0, 20.4, 19.7, 16.9, 14.4, 8.1, 6.2, and 2.5. The distance from the arbitrary position at the top of a blot was taken as a measure of the relative mobility. Standard curves obtained in such way were nonlinear and were fitted to third degree polynomial. The molecular weights of the cleavage products were determined from these standard curves and were used to calculate the cleavage site positions. These were rounded to the nearest 5 (for example, if the calculated cleavage position was 173, it would be reported as 175). This rounding was performed to take into account the resolution limitations of SDS–PAGE, as discussed later in this manuscript.

SDS–PAGE has been known to provide a remarkably good correlation between mobility and polypeptide size (Weber & Osborn, 1975). However, there are also examples where the correlation between size and mobility was poor (Hames, 1990). To check if CRP or its fragments do not display anomalous gel mobility in SDS–PAGE, three N-terminal fragments of CRP cleaved at known positions were prepared by proteolytic digestion: 1–116 fragment [by subtilisin digestion (Tsugita et al., 1982)], 1–136 fragment [by chymotrypsin digestion (Angulo & Krakow, 1985)], and 1–171 fragment [by V8 protease digestion (Angulo & Krakow, 1985)]. Their electrophoretic mobility in comparison with molecular weight standards was then determined. The mobility of CRP and its N-terminal fragments showed an excellent correlation with the standard curve obtained with molecular weight markers (not shown). Since no anomalous mobility of CRP was observed, the molecular weight of cleavage products could be used to determine positions of cleavage sites.

Determination of Degree of Protein and DNA Degradation. To determine the percentage of the intact protein and DNA molecules remaining after the exposure to Fe–EDTA, both macromolecules were individually incubated with Fe–EDTA, and at different time intervals samples were withdrawn and mixed with gel electrophoresis sample buffer to stop the reaction. Samples were then loaded on 15% SDS–PAGE gel (protein samples) or denaturing 8 M urea acrylamide gel (DNA samples). After electrophoresis, the protein gels were stained with Coomasie Blue and the DNA gels were stained with ethidium bromide. The bands of intact protein and of intact DNA were quantitated with a video densitometer (Biomed Instruments, Inc.), and the results were expressed as the intensity of a given band/intensity of the band for 0 time (in percent).

Effect of Footprinting Reaction Conditions on the Activity of CRP and DNA. To estimate the extent of overall damage of CRP and DNA due to the exposure to Fe–EDTA, the activity of CRP and DNA exposed to footprinting reaction conditions was determined by measuring their ability to participate in specific protein–DNA complex formation. Samples of CRP (or DNA) were incubated with Fe–EDTA, and at various times the fraction of CRP (or DNA) molecules remaining competent to form the specific protein–DNA complex was measured by stoichiometric binding titrations. The binding reactions were performed in 50 mM Tris (pH 7.8), 100 mM KCl, 1 mM EDTA, and 1 mM 2-mercaptoethanol buffer containing 100 μM cAMP. The formation of a complex was followed using the fluorescence anisotropy

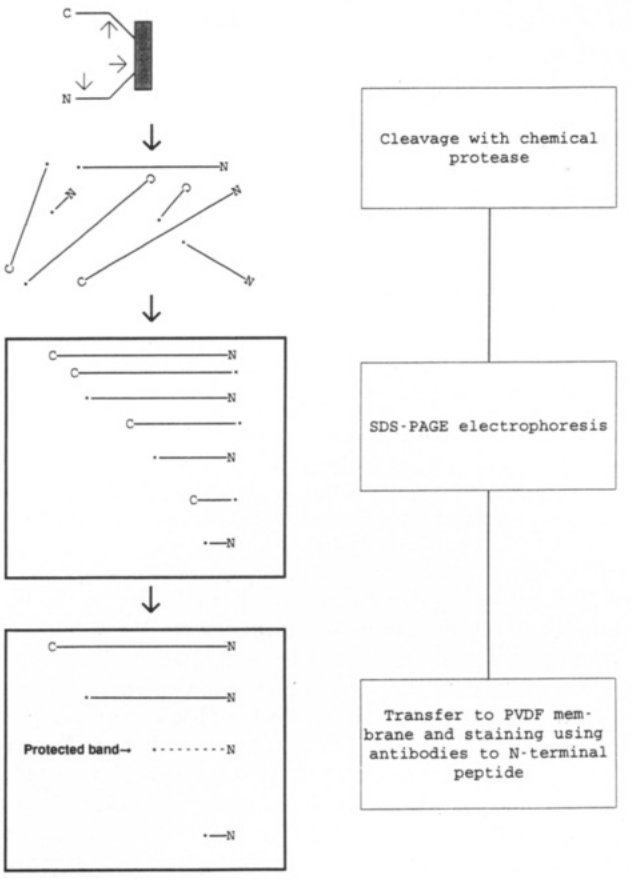

FIGURE 1: General scheme of protein footprinting method. For illustration purposes, the hypotetical protein–macromolecule complex is considered. When this complex is subjected to limited cleavage with Fe–EDTA [the model protein used in this example can be cleaved at three sites, and one of the sites is within the domain where the macromolecule binds to the protein; in the case of a "real" protein, all solvent-exposed peptide bonds should be attacked (Platis et al., 1993)]. The products of the limited cleavage in which the protein is cleaved once or not at all (seven in this example) are resolved according to their size on SDS–PAGE and electroblotted to a PVDF membrane. The membrane is stained using immunodetection with antibodies specific to the N-terminal peptide of the protein. The mobility of visualized bands (four in this example) is directly proportional to a distance of the cleavage site from the N-terminus of the protein. Thus, by comparing with the electrophoretic mobility of appropriate standards, the positions of the cleavage sites in protein primary structure can be determined as well as the positions of sites which are protected by protein–macromolecule complex formation.

change accompanying the complex formation between CRP and a DNA fragment labeled with a fluorescence probe. Details of this approach are published elsewhere (Heyduk & Lee, 1990).

RESULTS

Figure 1 shows the general scheme of the protein footprinting method. The protein–macromolecule complex is subjected to limited cleavage with the chemical protease (for illustration purposes in the example shown in Figure 1, it is assumed that the protein can be cleaved at three sites, and one of these sites is within a domain where the macromolecule binds to the protein). The products of this limited cleavage are resolved by SDS–PAGE and transferred to a PVDF membrane. The membrane is stained with antibodies specific to the N-terminal peptide of the protein. The mobility of the visualized bands (4 in this example) is directly proportional to a distance of the cleavage site from the N-terminus of the protein. Thus, by comparing with the electrophoretic mobility of appropriate standards, the positions of the cleavage sites in the protein

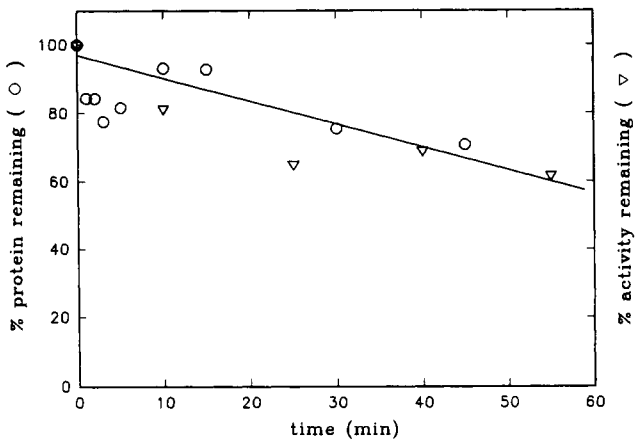

FIGURE 2: Percentage of uncut CRP remaining (○) and CRP activity remaining (▽) as a function of time of exposure to Fe–EDTA. The experiments were performed as described in Materials and Methods.

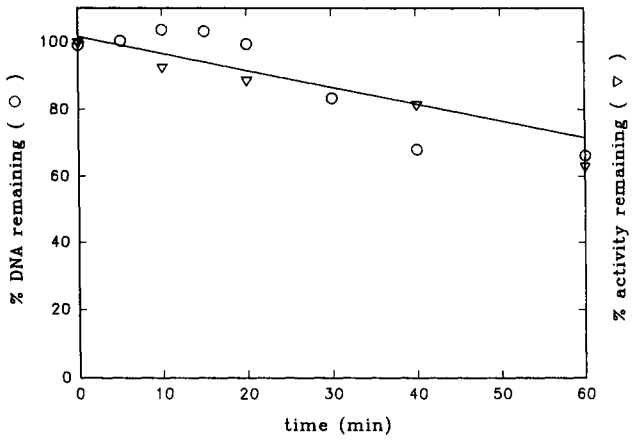

FIGURE 3: Percentage of uncut ICAP DNA remaining (○) and DNA activity remaining (▽) as a function of time of exposure to Fe–EDTA. The experiments were performed as described in Materials and Methods.

primary structure can be determined as well as the positions of the sites which are protected in a protein–macromolecule complex. The application of this methodology to CRP and its complexes with cAMP and DNA is described below.

Cleavage of CRP with Fe–EDTA Complex. The conditions of protein cleavage with Fe–EDTA have been reported so far only for Fe–EDTA covalently or noncovalently bound to the protein (Rana & Meares, 1990, 1991a; Hoyer et al., 1990; Schepartz & Cuenoud, 1990; Ermacora et al., 1992). In the initial experiments, we observed that CRP can be cleaved to a number of polypeptide fragments by free Fe–EDTA in the presence of H_2O_2 and sodium ascorbate. The metal ions were necessary for the cleavage (both Fe^{2+} and Fe^{3+} worked equally well when ascorbate was present), whereas in the absence of H_2O_2 and/or sodium ascorbate, only limited cleavage occurred. In the presence of radical scavengers (ethanol), very little cleavage was observed (not shown). These initial experiments led to establishing the optimal concentrations of the reagents (see Materials and Methods).

Protein cleavage by Fe–EDTA could be used for protein footprinting when two conditions would be satisfied: (1) most of the protein molecules are subjected to not more than a single cut; (2) the protein should not undergo extensive inactivation before any cleavage occurs. Multiple cuts in the protein should be avoided to make certain that the majority of the cleavage products result from the attack of Fe–EDTA on the native protein and not on the cleavage product of the previous cut. The first cleavage event could induce protein denaturation which in turn could lead to exposure to reagents and cleavage at secondary, internal sites. Brenowitz et al. (1986) have analyzed the relationship between the number of cuts per molecule and the amount of uncut macromolecule remaining. They found that the "single-cut" conditions can be satisfied if the large portion of the macromolecule remained uncut. Figure 2 shows the percentage of uncut CRP remaining as a function of exposure to Fe–EDTA. Even after 1 h incubation, about 60% of CRP remained uncut. This amount of uncut protein would still satisfy single-cut cleavage conditions (Brenowitz et al., 1986). Thus, after 1 h incubation with Fe–EDTA, the probability that CRP molecules would be subjected to multiple cuts is small. Additional confirmation that secondary cleavages are not significant was obtained from the analysis of the changes in the pattern of CRP cleavage products with time. The intensity of all bands increased with time, and no lag periods for appearance of any bands (which would be very characteristic of secondary cleavages of the denatured protein) were observed (not shown).

The highly reactive radicals generated by the Fe–EDTA system will not only cleave polypeptides but may also result in chemical modification of protein side chains (Stadtman, 1993). Under unfavorable circumstances, the protein could undergo extensive inactivation through side-chain modification before any fragmentation occurs, making the protein footprinting with Fe–EDTA impractical. In experiments with a model peptide system, the oxidative modification of amino acid side chains by Fe–EDTA was not observed (Platis et al., 1993). This suggests that it should be possible to cleave a polypeptide chain without first causing severe chemical alteration of its structure. This is consistent with our experiments where the degree of CRP fragmentation was compared with the extent of its inactivation (Figure 2). The activity of the protein decreased as a function of the exposure to Fe–EDTA. This decrease was more or less comparable to the rate of protein fragmentation. Since the inactivation of the protein is a result of both the cleavage and the chemical modification of sides chains, the results shown in Figure 2 suggest that not every cleavage event results in protein inactivation, and also no significant preferential inactivation over cleavage in the case of CRP was observed. Cleavage of ICAP DNA under the same conditions showed the same overall characteristics as in the case of CRP (Figure 3). Thus, in our model system of CRP and ICAP DNA, the cleavage of the protein by Fe–EDTA appeared to be suitable for footprinting.

Footprinting CRP Domains Involved in DNA Recognition. Figure 4 shows the results obtained when the protein footprinting methodology described above was applied to the CRP protein. Cleavage of the free CRP with Fe–EDTA and visualization of the bands with immunostaining procedure using antibodies specific to the N terminus revealed a ladder of discrete bands (Figure 4). The observation of discrete bands suggests that only a fraction of all peptide bonds was cleaved, consistent with the notion that only solvent-exposed regions of the protein should be susceptible to the cleavage. The intensity of cleavage products bands decreased as their molecular weight decreased (Figure 4A). Virtually no cleavage products smaller than about 5 kDa could be detected. Thus, about 70% of the protein (from about residue 50 to the C terminus) is amenable to the analysis by protein footprinting, when the detection with the antibodies specific to the N terminus is employed. The absence of small molecular weight bands in the blot of CRP cleavage products could be a result of any of the following: (1) it may reflect the relative solvent inaccessibility of the N-terminal domain of the protein; (2)

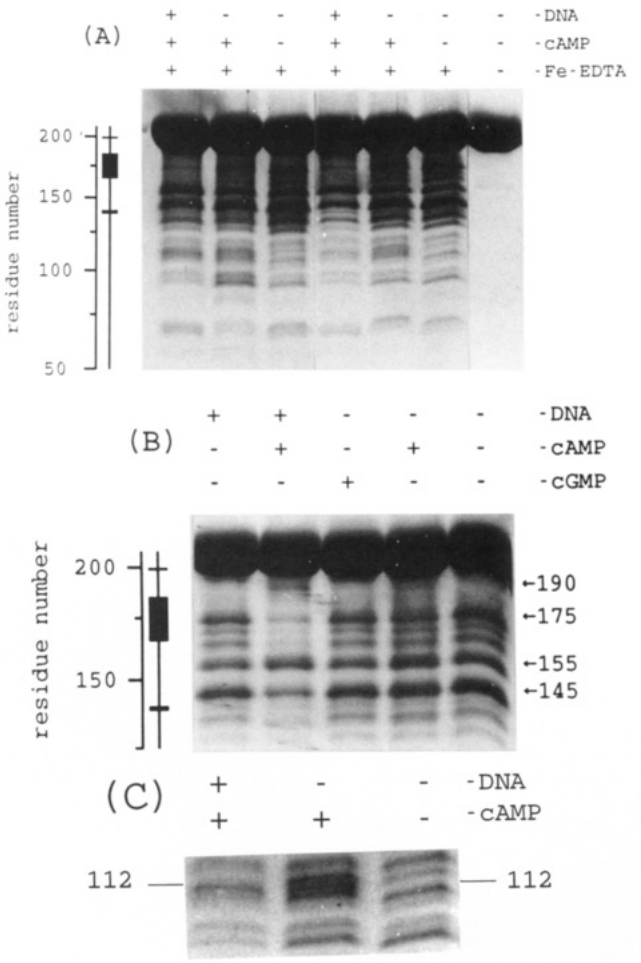

FIGURE 4: Protein footprinting analysis of CRP interactions with cAMP and DNA. The footprinting reactions, gel electrophoresis, electroblotting, immunodetection, and band assignments were performed as described in Materials and Methods. The concentration of DNA (42 bp ICAP DNA), cAMP, and cGMP when present were 5, 100, and 100 μM, respectively. Lines representing the predicted positions of bands corresponding to cleavages at residues from 50 to 209 and from 125 to 209 are shown in A and B, respectively. In the same scale, the regions of CRP identified by X-ray crystallography to be in direct contact with DNA (Schultz et al., 1991) are shown (thick blocks on the line parallel to the residue number line). These regions are residue 199, residues 166–188, and residues 139–140. (A) Two independent repeats of the same experiment are shown to illustrate the reproducibility of the band pattern. The blots were overexposed for better visualization of bands in the 50–100-residue region. Only the portion of the blot down to bands corresponding to the cleavage at residue 50 is shown; essentially no bands with molecular bands smaller than about 5000 Da could be detected. (B) The region corresponding to the major protection due to DNA is shown. To achieve better visualization of the protection pattern, the gel was run for a longer time, and the staining of the blot was optimized. The numbers and arrows indicate the assignment of the cleavage sites for several major bands in this region. (C) The region where cAMP-induced changes are observed (residues 95–120). A position of a new band appearing in the presence of cAMP is marked.

the cleavage products may be progressively less retained by the membrane as their molecular weight decreases, a known problem of electroblotting small peptides (Lin & Kamamatsu, 1983); (3) the sensitivity of immunodetection of smaller peptides may be decreased, since the segment of the peptide which is actually recognized by the antibody becomes a significant portion of a small molecular weight fragment, and thus the binding of small fragments to the membrane would be more likely to interfere with the interaction of the peptide with the antibody. Although more experimental data would be necessary to understand these phenomena, we feel that options 2 and 3 described above would be largely responsible for the difficulty in detecting any small molecular weight cleavage products. For example, when the cleavage is performed in 8 M urea, which should destroy the native structure of the protein, we also could not detect any small molecular weight cleavage products (not shown).

Due to the already discussed electroblotting and immunostaining artifacts, any correlation between the intensity of the cleavage product bands and the protein structure should be approached with the extreme caution. However, one interesting observation could be made. As can be seen in Figure 4A, the overall intensity and frequency of the bands in the region corresponding to the C-terminal domain (residues 134–209) of the protein seem to be much higher than in the region corresponding to the N-terminal domain (residues 1–133). The average thermal factor for C-terminal domain is about 27 Å^2, whereas for the N-terminal domain it is only about 12 Å^2 (Weber & Steitz, 1987). This suggests that the C-terminal domain is either less compact or more flexible than the N-terminal domain. This could be a reason of the greater susceptibility of this domain to chemical cleavage observed in this paper and would point out that the dynamic properties of the protein can also be important for the accessibility of a peptide bond to the cleavage with Fe–EDTA.

Formation of the CRP–cAMP complex induced only small changes in the CRP cleavage pattern (Figure 4A, discussed in the next section). When CRP formed a complex with a 42-bp DNA fragment (ICAP), profound changes in the cleavage pattern occurred (Figure 4A,B). The most obvious change was the pronounced protection of the bands corresponding to cleavage at residues from about 163 to 182 (Figure 4B). This corresponds, within the experimental error, to the helix–turn–helix motif of CRP, which provides almost all contact sites for the interaction with a specific DNA binding site and which becomes buried in the protein–DNA interface when the complex is formed (Schultz et al., 1991). The localization of protein cleavage sites protected by DNA is schematically shown in Figure 5A and compared with the localization of CRP residues determined by X-ray crystallography to be in contact with DNA (Figure 5B). Other DNA-induced changes in the CRP footprint included a moderate protection in the region of 135–145 and around residue 195, and a change in the pattern of bands in the region 95–120 was seen (Figure 4A,C). And finally, a band corresponding to a cleavage at residue 190 became hypersensitive in the DNA–CRP complex (Figures 4B and 5). Although these smaller changes were not as profound as in the case of the 163–182 region, they were reproducible and significantly larger than any variability of the footprints due to electroblotting and staining artifacts. Thus, we believe they are real. These smaller changes in the CRP footprint induced by DNA can also be correlated with the crystal structure of the CRP–DNA complex (Schultz et al., 1991). The moderate protection observed in the region around 135–145 is consistent with the protein–DNA backbone phosphate contact in which residues 139 and 140 participate (Schultz et al., 1991). The DNA-induced hypersensitivity around position 190 also seems to fit into the picture of the CRP–DNA complex. This region of CRP corresponds to a solvent-exposed CRP fragment between the last DNA contact of the F helix of the helix–turn–helix motif (Lys 188) and a DNA phosphate backbone contact of His 199 (Schultz et al., 1991) (Figure 5). Thus, this polypeptide segment is not involved in protein–DNA interactions, but it is in between two segments that directly interact with DNA residues. It seems that

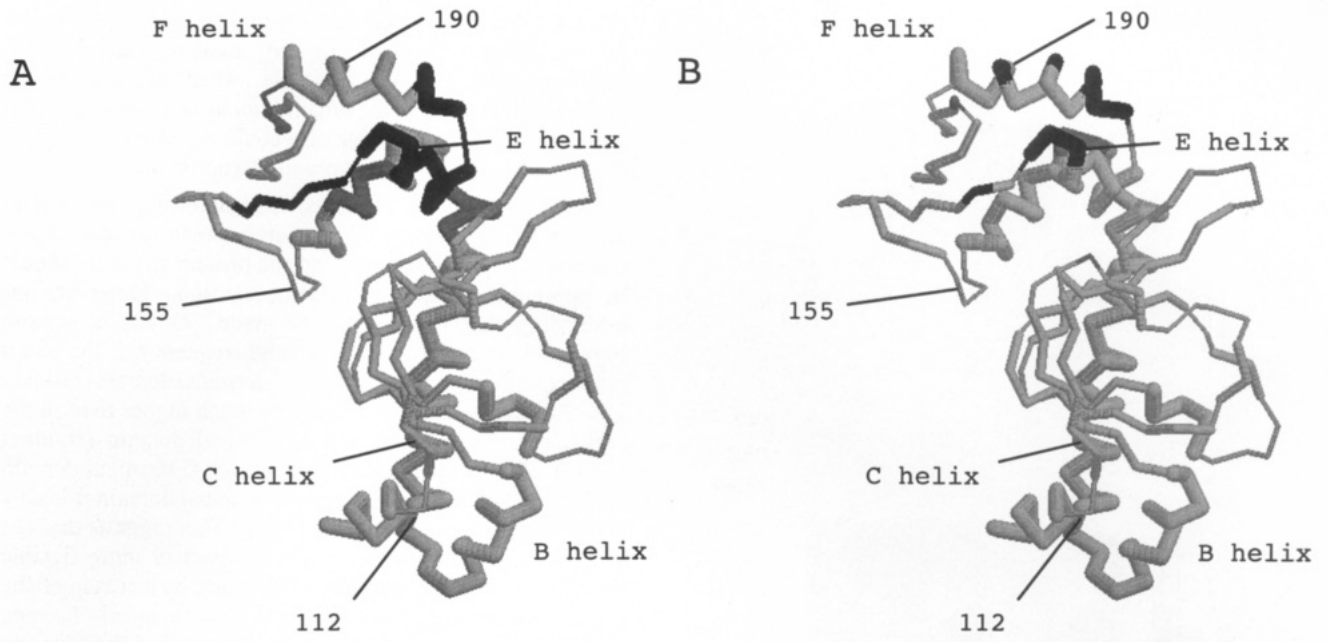

FIGURE 5: Schematic representation of cleavage site localization in the three-dimensional structure of CRP. The figure was drawn using Raswin program (R. Sayle, BioMolecular Structure Group, Glaxo Research and Development, Greenford, UK). Coordinates of the CRP–DNA complex crystallographic structure (Schultz et al., 1991) were used and were obtained from Brookhaven Protein Data Bank (accession code 1CGP). Only one subunit of the CRP dimer is shown. The secondary structure elements are depicted by: α helices, thick cylinders; β sheets, medium cylinders; other, thin cylinders. (A) Cleavage sites protected by DNA. The protected residues within the helix–turn–helix motif are colored black. The protected residues outside the helix–turn–helix motif are shown in dark gray. The protected residues ± 2 residues were colored to take into account a minimal error of band assignment using SDS–PAGE (see the text). (B) Residues forming direct contacts with DNA as determined by X-ray crystallography (Schultz et al., 1991). The residues within the helix–turn–helix motif are colored black. The residues outside the helix–turn–helix motif are shown in dark gray.

anchoring two ends of this loop through the interaction with DNA causes either stress or a change in its dynamic properties such that it became more susceptible to cleavage. Interestingly, the DNA between the sites of contacts with His 199 and Lys 188 is distorted (Schultz et al., 1991), which could be a source of the strain in the peptide connecting residues 188 and 199. The protection seen around position 195 can be caused by nearby residue 199 making a direct contact with DNA (Schultz et al., 1991) (Figure 5).

All DNA-induced changes result from the specific protein–DNA complex formation. This statement is based on the following observations: (1) in the presence of the equivalent amount (in base pairs) of calf thymus DNA, no changes in CRP footprint were observed (now shown); (2) in the absence of cAMP, which is required for high-affinity DNA binding of CRP (Fried & Crothers, 1984), DNA had no effect on the cleavage pattern of CRP (Figure 5B); (3) the protection of the 163–182 region was not changed with the increase of DNA concentration after all CRP is converted into the CRP–DNA complex (Figure 6). Under the salt conditions employed in the experiment illustrated in Figure 6, ICAP DNA fragment binding is stoichiometric (Heyduk et al., 1993). Figure 6 also shows that the intensity of the band that corresponds to a cleavage in the region of CRP, which is not involved in DNA binding (a cleavage around residue 155), was essentially not influenced by the presence of DNA at different concentrations. The results presented in Figure 6 suggests that it may be possible to use the protein footprinting methodology for quantitative measurements of protein–macromolecule interactions.

Conformation Changes of CRP. The binding of cAMP to CRP induces a conformational transition in the protein [reviewed recently by Kolb et al. (1993)], so it was of interest to examine if any information regarding this transition could be obtained from the protein footprinting data. Based on the mutagenesis data, it has been proposed that this change may involve reorientation of the hinge region, which could cause the recognition helices of the DNA binding domain to assume a configuration compatible with the spacing of two successive major grooves of DNA (Kim et al., 1992; Kolb et al., 1993). No changes in the presence of cAMP were observed in the CRP footprint in the DNA recognition domain of CRP (residues 165–188) (Figure 4B) and in the hinge region (130–139) (Figure 4A). This would suggest that cAMP-induced alterations of the CRP structure in these regions do not involve major changes in solvent accessibility of amino acid residues. The only significant and resproducible change in the cleavage pattern in the presence of cAMP occurred in the region between residues 95 and 120 (Figure 4C). In the case of free CRP, six bands, could be detected in this region corresponding to cleavages at residues 119, 114, 109, 105, 98, and 94 (Figure 4C). In the presence of cAMP, a new band corresponding to a cleavage at position 112 appeared, and there were some small changes in the intensity of the remaining bands. These changes were not considered because they were not always reproducible. The appearance of a new band (cleavage at 112) was specific for cAMP. In the presence of cGMP, which binds CRP as well as cAMP (Takahashi et al., 1980) but does not activate CRP (Fried & Crothers, 1984), no changes in CRP occurred, and the 112 band was not observed (not shown). Also, when CRP formed a complex with DNA, the footprint in the 95–120 region of CRP was changed (Figure 4C). The intensity of the band corresponding to a cleavage at residues 112 increased, and the intensity of the remaining bands decreased in such a way that the 112 band became the most intense band in this region of the footprint (Figure 4C).

DISCUSSION

The footprinting pattern of CRP in a complex with DNA showed an excellent correlation with the data available from

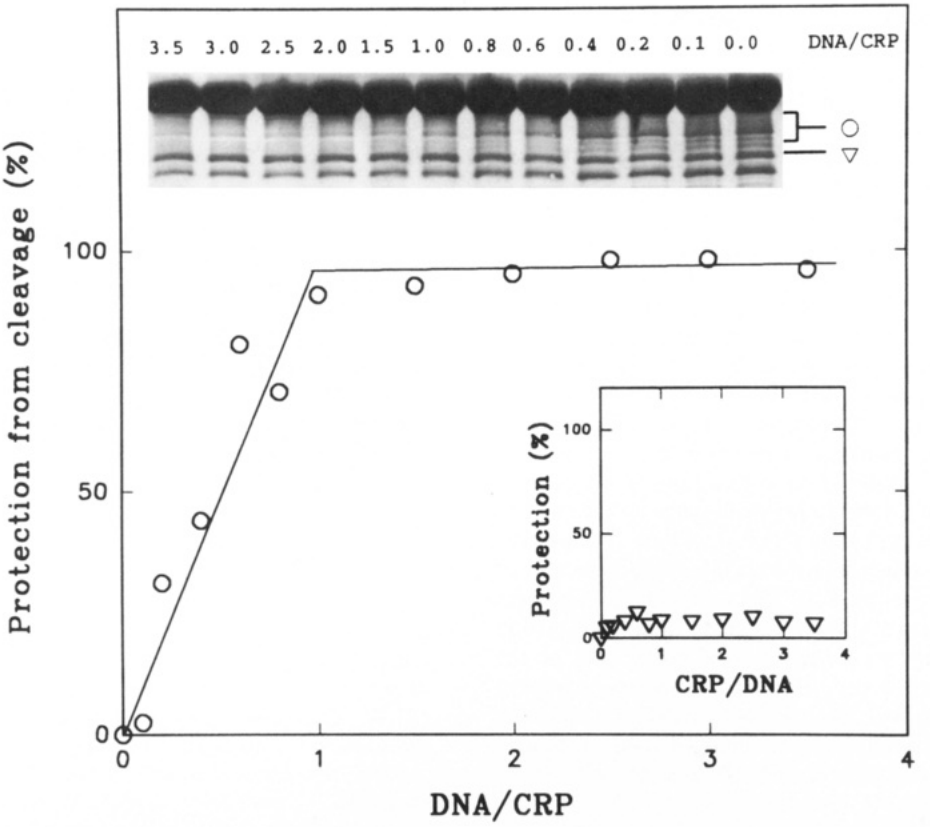

FIGURE 6: Quantitative protein footprinting analysis of CRP titration with ICAP 42-bp DNA analyzed by protein footprinting. The footprinting at different ratios of DNA to CRP was performed as described in Materials and Methods. The data are presented as protection (in percent), which corresponds to a difference in the intensity of a band at a given DNA concentration and the intensity of the band in the absence of DNA divided by the intensity in the absence of DNA. The intensity of the bands was obtained from two-dimensional scanning of the blot with a video densitometer (Biomed Instruments). (○) Protection of the 163–182 region. (▽) Protection of the 155 band. (Inset) A photograph of the blot that was used to obtain the protection data. Only a portion of the blot down to the cleavage at position about 130 is shown. The symbols on the right side of the blot show the regions that were quantified by scanning.

the three-dimensional structure of the CRP–DNA complex (Schultz et al., 1991) (Figure 5). The major protection from the cleavage was observed at residues 163–182. This is almost exactly the helix–turn–helix motif of CRP (residues 166–188) that becomes buried in the protein–DNA interface. Also, the smaller effects of DNA on the CRP footprint could be explained by the features of the CRP–DNA complex identified by X-ray crystallography. Thus, not only the major DNA binding domain of the protein could be identified but also the finer details of the CRP–DNA complex could be detected.

The DNA-induced changes in the CRP footprint in the 95–120 residue region cannot be explained by any direct effect of DNA on CRP since no protein–DNA contacts were observed in this region (Schultz et al., 1991). The simplest explanation would be that these changes in the footprint correspond to DNA-induced conformational changes in the CRP molecule. The same region of the protein was also affected by cAMP binding. Residues 95–120 map in the N-terminal portion of a long C-helix and C-terminal portion of the B-helix (Weber & Steitz, 1987) (Figure 5). This region is involved in the subunit–subunit interaction of CRP (Weber & Steitz, 1987). Proteolytic fragments of CRP, 1–116 and 1–136, are still dimeric with slightly decreased stability as compared to the native protein (Heyduk et al., 1992; Cheng et al., 1993). This suggests that the N-terminal portion of the C-helix and the C-terminal portion of the B-helix contribute the most to the dimerization free energy. Thus, the observation that this region was affected by cAMP and DNA binding stresses out the importance of subunit–subunit interactions in the transmission of information from the cAMP binding sites to the DNA binding domain. Whatever the exact nature of cAMP-induced conformational transition in CRP might be, it seems to involve a perturbation of subunit–subunit interactions in the N-terminal domain.

The successful footprinting of DNA binding domains in CRP strongly suggests that the protein footprinting methodology described in this paper could be readily applied to other proteins. The interpretation of the protein footprinting results may be often not as straightforward as in the case of the CRP–DNA complex for which the X-ray structure is available. One should remember that the changes in the intensity of bands in the footprint reflect only the changes in the solvent exposure of the protein and are not automatically a signature of a binding site for the macromolecule. The binding-induced conformational changes in the protein can also produce changes in the footprint such as protection, hypersensitivity, or the appearance of new bands. The last two effects can be easily interpreted as a result of a conformational change in the protein. Additional data (studies on mutants proteins, cross-linking, etc.) might be necessary to be able to interpret protection data with absolute confidence.

To adapt the approach described here to a different system, it should be only necessary to obtain antibodies to the N- or (and) C-terminal peptides of the protein under study, to perform control experiments to establish conditions for the single-cut cleavage, and to determine that the protein is not extensively modified before significant cleavage occurs. Previously, de Arruda et al. (1992) have successfully used a combination of specific chemical modification of methionine residues, cleavage, and immunodetection to "footprint" the

actin binding site on villin. A nonspecific cleavage with the Fe–EDTA complex in our approach offers the possibility of a greater density of sampling of the protein surface.

The resolution of the footprinting is mostly determined by the resolution and accuracy with which the molecular weight of the cleavage products could be determined by SDS–PAGE. The accuracy of this determination can be improved by the use of molecular weight standards derived from the protein under study (as we have done in the case of CRP), by analyzing the mobility of cleavage product bands as a function of acrylamide concentration, and by using the immunodetection with the antibodies to both N and C terminus of the protein. By assigning the cleavage sites using both N- and C-terminal detection, the small variations of the correlation between the size of the peptide and its electrophoretic mobility would be averaged, and large deviations could be detected. Analyzing the electrophoretic mobility as a function of acrylamide concentration (Ferguson, 1964) will help to detect any anomalously migrating bands (Hames, 1990). The protein footprinting methodology described in this paper requires only small amounts of the protein. The complex, multiprotein systems can be conveniently studied since the immunodetection selects for observation only the desired component of the system. This approach can be applied to not only mapping protein contact sites but also the topography of membrane proteins, protein conformational changes, and structural aspects of protein unfolding may be studied as well.

ACKNOWLEDGMENT

We would like to thank Dr. Guang-Zuan Cai for his patient and expert help in immunological techniques and Drs. Claudette Klein and Duane Grandgenett for helpful comments about the manuscript.

REFERENCES

Angulo, J. A., & Krakow, J. S. (1985) *Arch. Biochem. Biophys. 236*, 11–16.

Bazari, W. L., Matsudaira, P., Wallek, M., Smeal, T., Jakes, R., & Ahmed, Y. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4986–4990.

Brenowitz, M., Senear, D., Shea, M. A., & Ackers, G. K. (1986) *Methods Enzymol. 30*, 133–181.

Burnens, A., Demotz, S., Corradin, G., Binz, H., & Bosshard, H. R. (1987) *Science 235*, 780–783.

Cheng, X., Gonzales, M., & Lee, J. C. (1993) *Biochemistry 32*, 8130–8139.

de Arruda, M. V., Bazari, H., Wallek, M., & Matsudaira, P. (1992) *J. Biol. Chem. 267*, 13079–13085.

Doolittle, R. F. (1986) *Of urfs and orfs: a primer on how to analyze derived amino acid sequences*, p 87, University Science Books, Mill Valley, CA.

Ebright, R. H., Ebright, Y., & Gunasakera, A. (1989) *Nucleic Acids Res. 17*, 10295–10305.

Ermacora, M. R., Delfino, J. M., Cuenoud, B., Schepartz, A., & Fox, R. O. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6383–6387.

Ferguson, K. A. (1964) *Metabolism 13*, 985–1002.

Fried, M. G., & Crothers, D. M. (1984) *J. Mol. Biol. 172*, 241–262.

Galas, D. J., & Schmitz, B. (1978) *Nucleic Acids Res. 5*, 3157–3170.

Gentry, D. R., & Burgess, R. R. (1993) *Biochemistry 32*, 11224–11227.

Hames, B. D. (1990) in *Gel Electrophoresis of Proteins. A Pratical Approach* (Hames, B. D., & Rickwood, D., Eds.) IRL Press, Oxford.

Heyduk, E., Heyduk, T., & Lee, J. C. (1992) *Biochemistry 31*, 3682–3688.

Heyduk, T., & Lee, J. C. (1989) *Biochemistry 28*, 6914–6924.

Heyduk, T., & Lee, J. C. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1744–1748.

Heyduk, T., Lee, J. C., Ebright, Y. W., Blatter, E. E., Zhou, Y., & Ebright, R. H. (1993) *Nature 364*, 548–549.

Hoyer, D., Cho, H., & Schultz, P. G. (1990) *J. Am. Chem. Soc. 112*, 3249–3250.

Jue, R. A., & Doolittle, R. F. (1985) *Biochemistry 24*, 162–170.

Kolb, A., Busby, S., Buc, H., Garges, S., & Adhya, S. (1993) *Annu. Rev. Biochem. 62*, 749–785.

Kim, J., Adhya, S., & Garges, S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9700–9704.

Lesley, S. A., & Burgess, R. R. (1989) *Biochemistry 28*, 7728–7734.

Lin, W., & Kasamatsu, H. (1983) *Anal. Biochem. 128*, 302–311.

Matsudaira, P., Jakes, R., Cameron, L., & Atherton, E. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6788–6792.

Phelps, R. A., Neet, K. E., Lynn, L. T., & Putnam, F. W. (1961) *J. Biol. Chem. 236*, 96–105.

Piszkiewicz, D., Landon, M., & Smith, E. L. (1970) *Biochem. Biophys. Res. Commun. 40*, 1173–1178.

Platis, I. E., Ermacora, M. R., & Fox, R. O. (1993) *Biochemistry 32*, 12761–12993.

Rana, T. M., & Meares, C. F. (1990) *J. Am. Chem. Soc. 112*, 2457–2458.

Rana, T. M., & Meares, C. F. (1991a) *J. Am. Chem. Soc. 113*, 1859–1861.

Rana, T. M., & Meares, C. F. (1991b) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10578–10582.

Schagger, H., & von Jagow, G. (1987) *Anal. Biochem. 166*, 368–279.

Schepartz, A., & Cuenoud, B. (1990) *J. Am. Chem. Soc. 112*, 3247–3249.

Schultz, S. C., Shields, G. C., & Steitz, T. A. (1991) *Science 253*, 1001–1007.

Stadtman, E. R. (1993) *Annu. Rev. Biochem. 62*, 797–821.

Takahashi, M., Blazy, B., & Baudras, A. (1980) *Biochemistry 19*, 5124–5130.

Tsugita, A., Blazy, B., Takahashi, M., & Baudras, A. (1982) *FEBS Lett. 144*, 304–308.

Tulius, T. D., & Dombroski, B. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5469–5473.

Weber, I. T., & Steitz, T. A. (1987) *J. Mol. Biol. 198*, 311–326.

Weber, K., & Osborn, M. (1975) *The Proteins*, 3d ed., Vol. 1, p 179, Academic Press, New York.

Zhou, Y., Zhang, X., & Ebright, R. H. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6081–6085.