Biochemical and Biophysical Research Communications **257,** 1–5 (1999)
Article ID bbrc.1998.0103, available online at http://www.idealibrary.com on IDEAL®

BREAKTHROUGHS AND VIEWS

Electrospray Ionization Mass Spectrometry: A Promising New Technique in the Study of Protein/DNA Noncovalent Complexes

Timothy D. Veenstra[1]

Environmental and Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99352

Received December 21, 1998

With the emergence of electrospray ionization mass spectrometry (ESI-MS), mass spectrometry is no longer restricted to the study of small, stable molecules, but has become a viable technique to study large biomolecules as well as noncovalent biomolecular complexes. ESI-MS has been used to study noncovalent interactions involving proteins with metals, ligands, peptides, oligonucleotides, and other proteins. An area where ESI-MS holds significant promise is in the study of protein/DNA interactions. The most common technique employed to study protein/DNA interactions is the electrophoretic gel mobility shift assay (EMSA). Although this technique has and will continue to provide excellent results, ESI-MS has shown the ability to provide detailed results not easily obtainable by EMSA. In this review I will discuss some of the protein/DNA noncovalent interactions that have been measured using ESI-MS, and contrast the results obtained by ESI-MS to those obtained by EMSA. © 1999 Academic Press

Key Words: **electrospray ionization mass spectrometry; DNA; protein; noncovalent.**

The function of a vast majority of biomolecules depends on their specific, noncovalent interactions with other molecules. For example, proteins interact with other proteins, peptides, small molecules, nucleic acids and oligonucleotides, lipids, and polysaccharides. A technique emerging as a new and exciting method by which to study biomolecular noncovalent interactions is electrospray ionization mass spectrometry (ESI-MS) (1). For a comprehensive list of the many different noncovalent complexes observed by ESI-MS, the reader is directed to an excellent review by Joseph Loo (2). In this review I will focus on the study of protein/DNA interactions by ESI-MS and discuss the results to those obtained using solution phase techniques, in particular the electrophoretic gel mobility shift assay (EMSA) (3,4).

It is well known that mass spectrometry studies ions in the gas-phase. Therefore to study molecules by MS, the molecules in solution must first be elevated into the gas-phase. In the study of protein noncovalent interactions this is done via electrospray ionization (ESI) (1). The ESI process creates a series of singly or multiply charged gas-phase ions from solution by creating a fine spray of highly charged droplets in the presence of a strong electric field. This technique produces gas phase ions less than 1 eV above their ground state energy (5), therefore no unwanted molecular fragmentation occurs and intact weakly bound noncovalent interactions are able to survive the electrospray process. To create gas-phase ions, the sample solution is sprayed from the tip of a metal needle to which voltage is applied. Heat or a dry gas is applied to the charged droplets to cause to solvent to evaporate. As the droplets decrease in size the charge density on the surface of the droplets increases, which results in the repulsion of the ion from the droplet into the gas-phase. The charged ions are then focused into the mass analyzer via a series of lenses and travel to the detector (6). The data are presented as a relative population of the various charge states of the molecule separated according to their mass-to-charge ratio (m/z), not their molecular mass. The molecular mass of the species is calculated from these multiply charged peaks by software algorithms included within the instrument software.

One of the most intense areas of biochemical research deals with the regulation of gene transcription, in which protein transcription factors bind to specific

[1] E-mail: timothy.veenstra@pnl.gov.

0006-291X/99 $30.00

DNA sequences ultimately resulting in the expression of the protein product of that gene. The most common technique used to examine protein/DNA interactions is the EMSA (3,4). This technique typically requires radioactively labeling a DNA probe which is incubated with the protein of interest. The mixture is separated on a polyacrylamide gel which is subsequently exposed to a radiographic film. The presence of a positive protein/DNA interaction is visualized by the presence of a radiographic band on the film. In situations where the protein sample is heterogeneous (for example when cell nuclear extracts are used), a super-shifted EMSA is performed using an antibody directed against the protein of interest to confirm its interaction with the target DNA. The ability to detect specific protein/DNA interactions by ESI-MS provides many advantages over this technique. ESI-MS provides an accurate record of the identity and stoichiometry of the complex negating the need to identify the protein components via an indirect technique such as a super-shifted EMSA. In cases where protein/DNA associations require a co-factor ESI-MS analysis can measure the stoichiometry of the co-factor in the complex. Whereas specific protein/DNA complex formation can be determined within a few hours by ESI-MS, an EMSA can require several days to achieve a comparable result. Since ESI-MS provides direct evidence of the formation of a complex based on the observation of a species, corresponding to the expected molecular mass of the complex, no modification of the DNA or protein is required.

EXAMPLES OF PROTEIN/DNA NONCOVALENT COMPLEXES STUDIED BY ESI-MS

1. Gene V Protein

It is only as recently as three years ago that the first specific noncovalent protein/DNA complex was observed using ESI-MS (7) (see Fig. 1). The complex involved the interaction between the gene V protein and a series of different oligonucleotides. In solution the gene V protein exists as a homodimer and binds single-stranded DNA molecules (8,9). The stoichiometry of the binding of the gene V homodimer to DNA has been shown in solution such that eight DNA bases are required for the binding of a single gene V protein homodimer. Therefore increasing the length of the single stranded DNA oligonucleotide will result in the binding of multiple gene V protein dimers. Similar to the results obtained by EMSA and NMR, the ESI-MS study showed that a single protein dimer bound to a single molecule of $d(pT)_{13}$ (7). Increasing the length of the oligonucleotide (i.e., $d(pT)_{15}$ and $d(pT)_{18}$) resulted in the detection of a second gene V protein homodimer bound to the DNA strand by ESI-MS. This result was consistent with previous results, obtained with the protein in solution, showing that each gene V protein monomer requires three to five DNA bases to bind (9). The study also compared the relative affinity of the gene V protein for d(pT) versus d(pA) DNA strands. The gene V protein had been shown to have a two orders of magnitude greater affinity for d(pT) DNA strands than d(pA), in solution (10,11). In binding competition experiments, in which a mixture of $d(pT)_{13}$ and $d(pA)_{14}$ in a molar ratio of 2:100 was mixed with the gene V protein, the ESI-MS study found the ratio of protein/$d(pT)_{13}$ complex to the protein/$d(pA)_{14}$ complex was 8:1 even though the concentration of $d(pA)_{14}$ was fifty times greater than that of $d(pT)_{13}$. No evidence of the gene V protein binding to double stranded d(pA)/d(pT) DNA strands was seen in the ESI-MS data, again supporting solution phase studies which show the preference of the protein to bind to single stranded DNA. The results of this initial ESI-MS study of protein/DNA interactions were invaluable because they showed that the characteristics of this complex in the gas-phase were similar to those observed in solution. Therefore this study validated the further use of ESI-MS to study protein/DNA complexes.

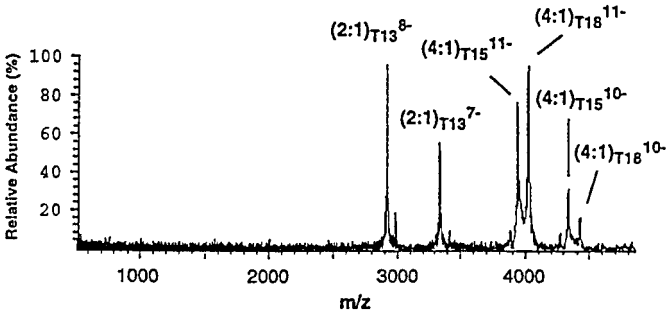

FIG. 1. ESI-MS spectrum for a mixture of 30 μM gene V protein and 5 μM each of $d(pT)_{13}$, $d(pT)_{15}$, and $d(pT)_{18}$. Complexes of the gene V protein and oligonucleotides are indicated by the ratio of monomers of protein:oligonucleotides, respectively. The subscript notes the length of the oligonucleotide in the complex while the superscripted number notes the observed charge state. Reprinted from (7) with permission from the National Academy of Sciences.

2. PU.1 DNA Binding Domain

From this initial example of a protein bound to a single stranded DNA molecule, the next example of a protein/DNA complex studied by ESI-MS involved a protein transcription factor bound to it's target sequence with a double-stranded DNA molecule (12). The system chosen was the DNA-binding domain (DBD) of the PU.1 transcription factor (PU.1 DBD) and a double stranded DNA containing the specific target sequence (i.e., GGAA) recognized by the PU.1 DBD (13). When a mixture of the PU.1 DBD and a 17 base pair (bp) double stranded DNA containing the wild type target sequence were analyzed by ESI-MS, the results showed a noncovalent complex with a 1:1 protein to

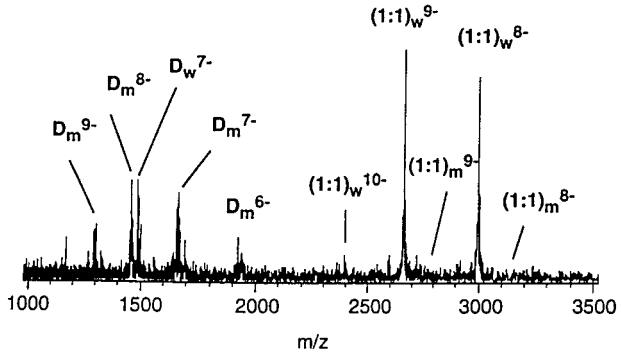

FIG. 2. ESI-MS spectrum of a mixture of the PU.1 DBD (5 μM) with a 17 bp wild type DNA (15 μM) and 19 bp mutant DNA (20 μM). The abbreviations are as follows: $(1:1)_w$, PU.1 DBD bound to the 17 bp wild type DNA in a 1:1 stoichiometry; $(1:1)_m$, PU.1 DBD bound to the 19 bp mutant DNA in a 1:1 stoichiometry; D_w, double stranded 17 bp wild type DNA; D_m, double stranded 19 bp mutant DNA. Reprinted from (12) with permission from Academic Press.

DNA stoichiometry (12). These results were consistent with those obtained by EMSA, however the stoichiometry of the complex could not be absolutely confirmed by EMSA (12). A mixture of the PU.1 DBD with the 17 bp wild type DNA as well as a 19 bp mutant DNA was also analyzed by ESI-MS (Fig. 2). Only complexes corresponding the the PU.1 DBD bound to the wild type DNA were observed in the ESI-MS spectrum.

3. trp Repressor

From these intial studies showing the efficacy of using ESI-MS to study protein/DNA interactions and confirming the results obtained by this technique were similar to those obtained by solution, sprang studies which used ESI-MS to probe the structure of transcription complexes which bind co-factors. One of these studies examined the binding of the trp repressor binding to its target DNA (14). This studied evaluated the ability of the trp repressor to bind to a double stranded DNA oligonucleotide containing two symmetrically arranged CTAG sequences separated by two, four, and six base pairs. The results showed that the protein bound to its target DNA as a homodimer and was able to form a stable interaction when the spacing between the consensus sequences was four base pairs, in agreement with solution phase studies. The results also showed that the co-repressor, tryptophan, was not required for the formation of the protein homodimer/DNA complex. The study further showed that while 5-methyl tryptophan and L-tryptophan were able to bind to the protein/DNA complex with high affinity, D-tryptophan was bound to the complex with poor specificity and affinity (14). This study illustrates the use of ESI-MS in evaluating the DNA sequence and co-factor specificity of protein/DNA complexes.

4. Vitamin D Receptor

The ability of ESI-MS to provide information not readily obtainable using other techniques is illustrated by the study of the binding of the DNA-binding domain of the vitamin D receptor (VDR DBD) to a double stranded DNA oligonucleotide containing the vitamin D response element (VDRE) from the mouse osteopontin gene (mOP) (15). The VDR is a nuclear hormone receptor which has been shown to activate the transcription of more than sixty genes in response to 1,25-dihydroxyvitamin D_3, the active metabolite of vitamin D (16,17). The VDR DBD is composed of two zinc-finger metal-binding domains located within the first 110 residues of the protein (18). The purpose of the ESI-MS study was to evaluate the role of zinc ions (Zn^{2+}) on the noncovalent interaction between the VDR DBD and the mOP. Studies of the glucocorticoid receptor (GR) and the human transcription factor SP1 by EMSA have shown that both of these proteins required Zn^{2+} for binding to their respective target DNA sequences (19,20). These studies also showed that high Zn^{2+} concentrations caused the protein/DNA complex to dissociate, however no definite reason could be determined from the available data.

The importance of the Zn^{2+} concentration in maintaining the VDR DBD/mOP complex was evaluated by ESI-MS. The negative ion ESI multiply charged spectrum of the mOP gene containing the VDRE in the presence of the VDR DBD shows two ion series between m/z 1700-2500 and 2400-3500 representing the protein bound to the DNA as a monomer and dimer, respectively. When the VDR DBD and mOP were incubated in the presence of EDTA, to remove Zn^{2+} from the protein, no complex was observed showing the necessity of Zn^{2+} in the formation of a stable complex (15). Addition of 100 μM Zn^{2+} to the complex solution increased the amount of detectable protein dimer/DNA complex relative to monomeric protein/DNA complex (Fig. 3A). Addition of Zn^{2+} to 200 μM final concentration, however, severely reduced the amount of protein/target DNA complex detected (Fig. 3C). The dependency of the complex association on the Zn^{2+} concentration was confirmed using cadmium (Cd^{2+}), an often used isomorphous replacement for Zn^{2+} (19,21). The addition of magnesium or calcium to a final concentration of 300 μM did not any dissociation of the VDR DBD/mOP complex, showing the effect of Zn^{2+} on the complex was not simply due to a change in the ionic strength of the solution. The results show that the VDR DBD, studied by ESI-MS, behaved similarly to the GR and human SP1 transcription factor at various concentrations of Zn^{2+}.

The utility of using ESI-MS to study the VDR DBD/mOP complex was evidenced by its ability to determine the precise mechanism by which elevated Zn^{2+} or Cd^{2+} concentrations caused dissociation of the protein/DNA

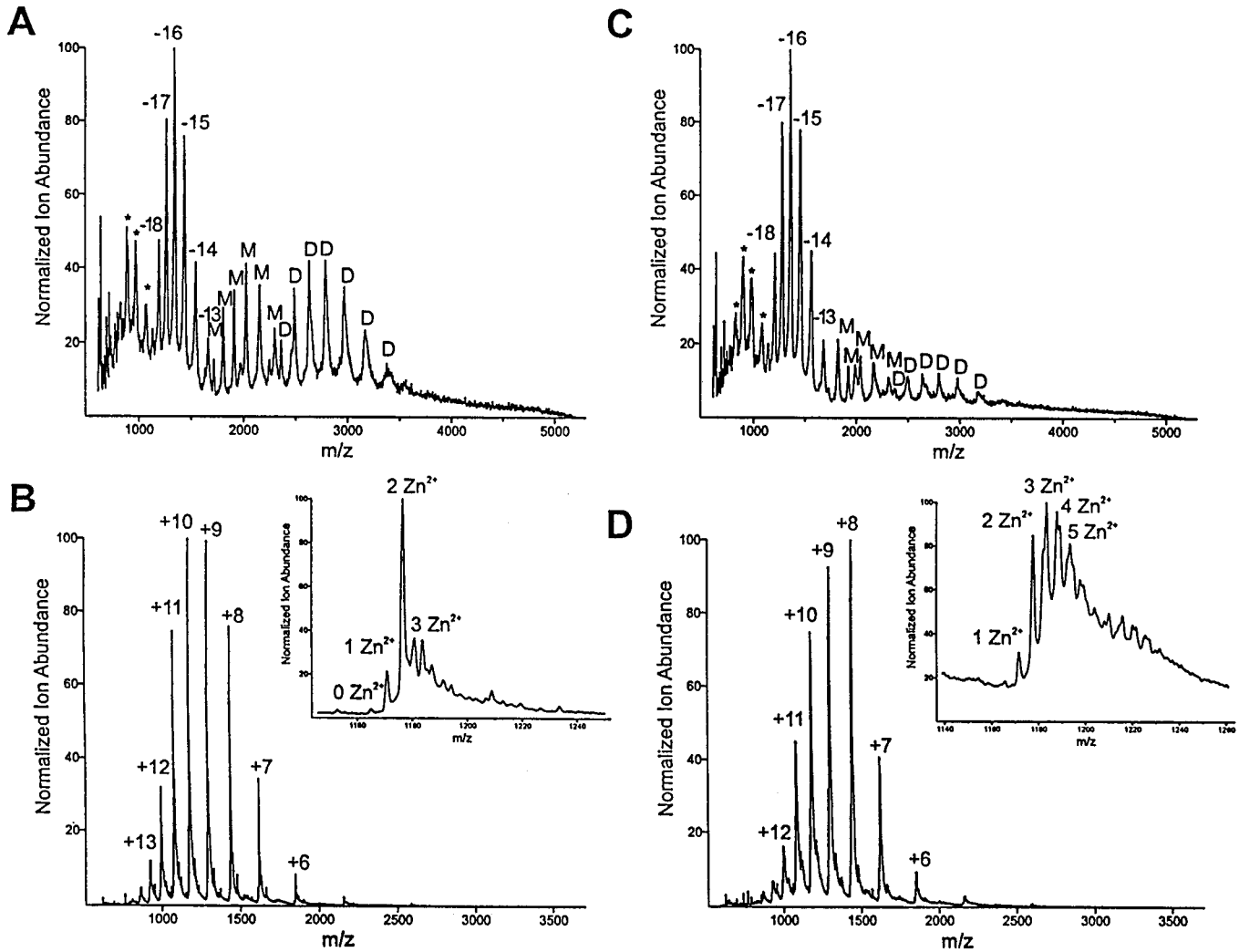

FIG. 3. Dependency of the Zn^{2+} concentration upon the binding of the VDR DBD to mOP DNA. (A) ESI-MS spectrum of the VDR DBD in solution with the double stranded DNA from the mOP gene in the presence of 100 μM Zn^{2+}. (B) Spectrum of the VDR DBD in the presence of 100 μM Zn^{2+}. (C) Spectrum of the VDR DBD/mOP complex in the presence of 200 μM Zn^{2+}. (D) Spectrum of the VDR DBD in the presence of 200 μM Zn^{2+}. Insets in (B) and (D) show expanded m/z region for the +11 charge state and the number of Zn^{2+} bound to the VDR DBD. VDR DBD/mOP complexes with the protein bound as a monomer or dimer are labeled with an M or D, respectively. Double stranded and single stranded DNA are labeled with their charge state or an asterisk (*), respectively. Reprinted from (15) with permission from Nature America.

complex. The VDR DBD has been shown by circular dichroism, inductively coupled plasma MS, and ESI-MS to be able to bind at least five Zn^{2+}; two with high affinity within the zinc-finger domains and at least three with low affinity at undetermined sites (18). The number of Zn^{2+} bound to the VDR DBD at the same concentrations used in the ESI-MS study of the protein bound to the mOP was measured. The results showed that with no extra Zn^{2+} added to the sample, the protein was predominantly bound by two Zn^{2+}. Addition of 100 μM Zn^{2+} also showed the predominant form of the protein was bound by two Zn^{2+} (Fig. 3B). Addition of 200 μM Zn^{2+}, however, resulted in the uptake of as many as six Zn^{2+} by the VDR DBD (Fig. 3D). Similar to the results obtained using Zn^{2+}, the uptake of four or more Cd^{2+} by the VDR DBD was responsible for the dissociation of the complex. This ESI-MS suggests that the binding at the first two sites, presumably within the Zn^{2+}-finger domains, is required for DNA-binding by the VDR DBD, however, binding to additional lower-affinity metal-binding sites is responsible for dissociation of the protein/DNA complex. As mentioned, the transcription factor SP1 and the GR have also been shown to dissociate from their target DNA sequences at elevated Zn^{2+} or Cd^{2+} concentrations, however, the reasons for this could not be determined by EMSA, since this technique is unable to differentiate between the number of ions bound to the complex due its low mass resolution. The accuracy and high mass resolution of ESI-MS enabled the direct

observation of the cause of the dissociation of the VDR DBD/mOP complex at high metal ion concentrations.

DISCUSSION

As mentioned in the introduction, ESI-MS offers some advantages over the EMSA in studying protein/DNA interactions. These advantages are generally related to the direct identification of the complex components and the mass resolution of ESI-MS. It would not be fair, however, not to mention the advantages of the EMSA as well. The amount of material required to perform an EMSA is generally less, however, the material required for ESI-MS analysis is still within the subnanomolar range. EMSA can also be performed on very crude protein preparations, while ESI-MS studies to date have used pure protein samples. The major drawback of ESI-MS compared to the EMSA is the instrument cost. While the equipment necessary to perform an EMSA is accessible to a majority of laboratories, the cost of the equipment necessary to perform ESI-MS studies is generally limited to well-funded laboratories or MS user facilities. Since only a limited number of protein/DNA complexes studied by ESI-MS it is still not certain how many complexes are amenable to study by this method. Although there may be disadvantages in using ESI-MS to study protein/DNA interactions, the specific information it provides can be a great reward for using it.

The potential of ESI-MS in the study of protein/DNA and other types of protein noncovalent interactions is only beginning to be realized. Since ESI-MS is a relatively new technique in this field, it has had to establish itself as a viable method which provides accurate results which are similar to those obtained in solution. Typically this has been done by confirming results obtained using other solution phase techniques. There are, however, a growing number of studies which have used ESI-MS to answer important biological questions which could not be sufficiently answered using other techniques. With the mass resolution provided by ESI-MS, this techniques is ideally suited to study protein/DNA interactions particularly in identifying heterodimer protein transcription factors which are similar in mass. ESI-MS will also play an important role in deciphering the effect of co-factors on transcription since it is able to measure small changes in mass caused by the binding of small molecules and metal ions to the transcription complex.

REFERENCES

1. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) *Science* **246,** 64–71.
2. Loo, J. A. (1997) *Mass Spectrom. Rev.* **16,** 1–23.
3. Fried, M., and Crothers, D. M. (1981) *Nucl. Acids Res.* **9,** 6505–6525.
4. Garner, M. M., and Revzin, A. (1981) *Nucl. Acids Res.* **9,** 3047–3060.
5. Kebarle, R., and Tang, L. (1993) *Anal. Chem.* **22,** 972–986.
6. Siuzdak, G. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 11,290–11,297.
7. Cheng, X. H., Harms, A. C., Goudreau, P. N., Terwilliger, T. C., and Smith, R. D. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 7022–7027.
8. Bayer, G., and McPherson, A. (1984) *Biochemistry* **23,** 340–349.
9. Kansy, J. W., Clack, B. A., and Gray, D. M. (1986) *J. Biomol. Struct. Dynam.* **3,** 1079–1110.
10. Alma, N. C. M., Harmsen, B. J. M., de Jong, E. A. M., Van der Ven, J., and Hilbers, C. W. (1983) *J. Mol. Biol.* **163,** 47–62.
11. Bulsink, H., Marmsen, B. J. M., and Hilbers, C. W. (1985) *J. Mol. Struct. Dynam.* **3,** 227–247.
12. Cheng, X., Morin, P. E., Harms, A. C., Bruce, J. E., Ben-David, Y., and Smith, R. D. (1996) *Anal. Biochem.* **239,** 35–40.
13. Wakylyk, B., Hahn, S. L., and Giovane, A. (1993) *Eur. J. Biochem.* **211,** 7–18.
14. Potier, N., Donald, L. J., Chernushevich, I., Ayed, A., Ens, W., Arrowsmith, C. H., Standing, K. G., and Duckworth, H. W. (1998) *Protein Sci.* **7,** 1388–1395.
15. Veenstra, T. D., Benson, L. M., Craig, T. A., Tomlinson, A. J., Kumar, R., and Naylor, S. (1998) *Nature (Biotechnology)* **16,** 262–266.
16. DeLuca, H. F. (1988) *FASEB J.* **2,** 2043–3053.
17. Pike, J. W. (1991) *Annu. Rev. Nutr.* **11,** 189–216.
18. Craig, T. A., Veenstra, T. D., Naylor, S., Tomlinson, A. J., Johnson, K. L., Macura, S., Juranic, N., and Kumar, R. (1997) *Biochemistry* **36,** 10,482–10,491.
19. Freedman, L. P., Luisi, B. F., Korszun, Z. R., Basavappa, R., Sigler, P. B., and Yamamoto, K. R. (1988) *Nature* **334,** 543–546.
20. Thiesen, H. J., and Bach, C. (1991) *Biochem. Biophys. Res. Commun.* **176,** 551–557.
21. Hanas, J. S., and Gunn, C. G. (1996) *Nucl. Acids Res.* **24,** 924–930.