Study of a noncovalent *trp* repressor: DNA operator complex by electrospray ionization time-of-flight mass spectrometry

NOELLE POTIER,^{1,4} LYNDIA J. DONALD,² IGOR CHERNUSHEVICH,^{1,5} AYEDA AYED,²
WERNER ENS,¹ CHERYL H. ARROWSMITH,³ KENNETH G. STANDING,¹
AND HARRY W. DUCKWORTH²

¹Time-of-Flight Laboratory, Department of Physics, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada

²Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada

³Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, 610 University Ave.,
Toronto, Ontario M5G 2M9, Canada

(RECEIVED November 3, 1997; ACCEPTED February 6, 1998)

Abstract

Electrospray ionization time-of-flight mass spectrometry (ESI-TOF MS) has been used to study noncovalent interactions between the *trp* apo-repressor (TrpR), its co-repressor tryptophan and its specific operator DNA. In 5 mM ammonium acetate, TrpR was detected as a partially unfolded monomer. In the presence of a 21-base-pair DNA possessing the two symmetrically arranged CTAG consensus sequences required for specific TrpR binding, a homodimer-dsDNA complex with a 1:1 stoichiometry was observed. Co-repressor was not needed for the complex to form under our experimental conditions. Collision induced dissociation (CID-MS) revealed that this complex was very stable in the gas phase since dissociation was achieved only at energies that also broke covalent bonds. We saw no evidence for the presence of the six water molecules that mediate the interaction between the protein and the DNA in the crystal structure. To check the binding specificity of the TrpR for its target DNA, a competitive experiment was undertaken: the protein was mixed with an equimolar amount of three different DNAs in which the two CTAG sequences were separated by 2, 4, and 6 bp, respectively. Only the DNA with the correct consensus spacing of 4 bp was able to form stable interactions with TrpR. This experiment demonstrates the potential of ESI-MS to test the sequence-specificity of protein-DNA complexes. The interactions between the TrpR-DNA complex and 5-methyl-, L- and D-tryptophan were also investigated. Two molecules of 5-methyl- or L-tryptophan were bound with high affinity to the TrpR-DNA complex. On the other hand, D-tryptophan appeared to bind to the complex with poor specificity and poor affinity.

Keywords: electrospray ionization; noncovalent; protein-DNA interaction, time-of-flight mass spectrometry, *trp* repressor

Over the last ten years, electrospray ionization mass spectrometry (ESI-MS) has demonstrated a high potential for the characterization of biomolecules by allowing the determination of molecular masses up to about 150 kDa with an accuracy of 0.01% (Fenn et al., 1990; Mann & Wilm, 1995). If experimental parameters

such as pH, buffers, temperature, and ion kinetic energy are carefully controlled, then noncovalent interactions between molecules can also be preserved during the ionization-desorption process. This technique has been successfully used to study protein-protein (Light-Wahl et al., 1994; Fitzgerald et al., 1996; Loo et al., 1997; Potier et al., 1997), protein-ligand (Ganem et al., 1991; Baca & Kent, 1992; Wu et al., 1997), antibody-antigen or DNA-drug interactions (Gale & Smith, 1995), and more recently, to examine protein-DNA complexes (Cheng et al., 1996a, 1996b). Although they cannot provide the detailed structural information obtained from crystallography or NMR, MS measurements give complementary information on the binding stoichiometry, specificity and stability of the complexes. Problems with ESI-MS include uncertainty about whether the ions observed reflect specific interactions found in solution, or simply artefactual nonspecific associations arising during the electrospray process. There are also solubility

Reprint requests to: Harry W. Duckworth, Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada; e-mail: hduckwth@cc.umanitoba.ca.

⁴Present address: Laboratoire de Spectrométrie de Masse Bio-Organique, UMR 7509, CNRS-Université Louis Pasteur, Faculté de Chimie, 1, rue Blaise Pascal, F 67008 Strasbourg, France.

⁵Present address: SCIEX, 71 Four Valley Drive, Concord, Ontario L4K 4V8, Canada.

Abbreviations: NH₄OAc, ammonium acetate; ESI, electrospray ionization; MS, mass spectrometry; TOF, time of flight; CD, circular dichroism; CID, collision induced dissociation.

problems because the choice of buffers for ESI is limited, and there are difficulties related to the m/z range of the mass spectrometer analyzer. Nevertheless, rigorous control experiments in the above studies have established that the ESI-MS measurements can give a good picture of conditions in solution provided that the electrospray parameters are set appropriately, but it is important to compare ESI-MS results with those obtained by established techniques on well characterized systems.

Escherichia coli *trp* apo-repressor (TrpR) is a DNA-binding protein that is involved in the regulation of tryptophan biosynthesis. The apo-repressor and its complexes with DNA and tryptophan have been well characterized by various techniques including X-ray crystallography (Otwinski et al., 1988; Lawson & Carey, 1993), high resolution NMR (Ramesh et al., 1994; Zhang et al., 1994) and electrophoretic mobility shift assay (Carey, 1988; Carey et al., 1991). The apo-repressor is a homodimer of calculated mass 24,448 Da. Each subunit consists of six helices (A to F), four of which (A, B, C, F) interact with the respective helices of the second subunit to form a stable structural core, while helices D and E form the helix-turn-helix DNA-binding domain. Functional operators are palindromes with a 4 bp CTAG sequence in each operator half-site. At high salt concentration (>100 mM) two molecules of the co-repressor L-tryptophan bind in a noncooperative manner to TrpR, forming a holorepressor complex (Arvidson et al., 1986).

We report here the use of ESI-MS for the characterization of noncovalent complexes between the *trp* apo-repressor, its co-repressor L-tryptophan and three operator DNA fragments. We chose the *E. coli* *trp* repressor-operator system as a model because the complex is so well characterized, and because it offers the opportunity to study a DNA binding protein whose affinity for DNA is regulated by a cofactor. In addition, there has been some controversy about the presence in solution of the six water molecules that were observed to mediate the DNA-protein interaction in the crystal structure (Yoderian & Arvidson, 1994), and it was hoped that MS measurements would shed further light on this matter.

Results and discussion

Characterization of the free protein

TrpR is a homodimer under the experimental conditions used for both X-ray crystallography (Zhang et al., 1987) and high resolution NMR (Zhao et al., 1993). At the low salt concentrations required for the electrospray measurements, however, the structure of TrpR is considerably different. Figure 1 shows the ESI mass spectrum obtained for the TrpR (10 μ M) in 5 mM NH_4OAc pH 6.0. Three series of multiply-charged ions can be detected. From the major series, a molecular mass of $12,224.3 \pm 0.3$ Da was measured, which is in excellent agreement with the expected mass of the TrpR monomer (calculated mass from the sequence: 12,224.0 Da). The two minor series correspond to TrpR dimer (measured mass: $24,448.2 \pm 0.4$ Da) and trimer (measured mass: $36,671.7 \pm 1.7$ Da). Wide distributions of charge states (7+ to 17+) are obtained for the multimers, strongly suggesting that these result from nonspecific associations. No protein-DNA complex is observed if ds DNA1 (the consensus operator sequence) is added to the protein under these conditions.

A closer look at the ESI spectrum for the monomer reveals two envelopes of charge states centered at 7+ and 10+. A bimodal distribution is usually indicative of different conformational states

Fig. 1. Characterization of the free protein. ESI m/z spectrum of *trp* repressor (10 μ M) in 5 mM NH_4OAc , pH 6.0. The declustering voltage V_c was 80 V.

in solution (Chowdhury et al., 1990; Loo et al., 1991; Light-Wahl et al., 1994; Smith & Zhang, 1994). Therefore, it appears that at least some of the protein is not in the native state. The presence of more than one conformation of the protein monomer, as well as the multiply-charged dimers and trimers, indicates that normal electrospray conditions are not ideal for study of the protein.

CD and fluorescence measurements were undertaken to determine whether the TrpR is folded under the conditions used for electrospray. Duplicate samples were used to compare the behavior of TrpR in 5 mM NH_4OAc and in 50 mM sodium phosphate, 150 mM NaCl, pH 7.6, conditions at which TrpR is known to be folded and stable. The mean residue ellipticity at 222 nm in the NH_4OAc buffer was 20% less than that in the phosphate plus NaCl buffer, indicating some loss of alpha helical structure (data not shown). However, fluorescence emission spectra of TrpR in both buffers were identical, with an emission maximum at 320 nm upon excitation at 280 nm. This indicates that the environments of the two tryptophans in TrpR, W19, and W99 are nonpolar, although it does not prove that such environments are identical under the two buffer conditions.

When the TrpR was prepared in a mixture of 5 mM NH_4OAc /MeOH:1/1, 1% acetic acid, the ESI spectrum showed a single envelope of charge states (5 to 16 protonations with a maximum at 11+). The CD spectrum of this sample showed a further small decrease in the α -helical content but was not judged to be fully unfolded when compared to the CD spectrum in a strongly denaturing solvent, 9 M urea.

The substantial amount of secondary structure retained under electrospray conditions indicates that the protein was not fully denatured, but that it was not in a native state either, and was probably not functionally active. That the protein is already destabilized at room temperature by the low salt buffer is shown by the CD results, and the electrospray conditions may further aggravate the destabilization.

Characterization of the protein-DNA complex

Figure 2A shows the spectrum of the TrpR-DNA complex obtained by adding an equimolar amount of the consensus operator se-

Fig. 2. Characterization of the TrpR-DNA complex. ESI m/z spectra of the *trp* repressor (20 μ M) in presence of (A) 10 μ M DNA1 and (B) 30 μ M DNA1 in 5 mM NH_4OAc , pH 6.0. A molecular mass of $37,266 \pm 8$ Da was measured for the TrpR dimer-DNA1 complex, which has the expected stoichiometry of [protein dimer: dsDNA] = [1:1]. The declustering voltage was 100 V.

quence ds DNA1 to the free protein at high salt concentration, and then dialyzing against 5 mM NH_4OAc (see Materials and methods). The signal corresponding to the free protein has decreased greatly compared to Figure 1, and a new, narrow envelope of multiply-charged ions is detected ($m/z = 3,388.3$; $m/z = 3,727.8$ and a small peak at $m/z = 4,141.9$), yielding a molecular mass of $37,266 \pm 8$ Da. This compares well to the mass 37,260 Da expected for a [TrpR dimer:ds DNA1] = [1:1] noncovalent complex. When the TrpR protein is incubated with a threefold excess of ds DNA1 (Fig. 2B), the mass spectrum shows the same narrow charge envelope of ions from a TrpR dimer-DNA complex of 1:1 stoichiometry and new peaks due to ss DNA. No major peaks corresponding to a complex of TrpR with ss DNA or complexes of other stoichiometries are observed. These results agree exactly with the model proposed by Joachimiak et al. (1983), the crystal structure (Otwinski et al., 1988), and the NMR solution structure (Zhang et al., 1994). (For review, see Youderian & Arvidson, 1994.) Unlike the free TrpR protein, and the unbound ds DNA, the TrpR-DNA1 complex is very stable in 5 mM NH_4OAc and can be kept at 4 °C for several weeks (data not shown).

There is a small amount of monomer protein at low m/z (see also Fig. 1). Some very minor peaks are also visible in the spectra in Figure 2. The minor peaks at $m/z = 4,978.3$ and 5,334.8 could be assigned as the 15+ and 14+ ions from a double complex, with two (TrpR dimer-ds DNA) units. Minor peaks at $m/z = 3,639.8$

and 3,970.1, which are more noticeable in Figure 2B, probably represent the 12+ and 11+ ions from a complex plus one strand of ss DNA. These peaks were present when there was excess DNA, and our interpretation of them is that the DNA is acting as a nonspecific "adduct" in the same way as Na^+ (see below). The excess DNA (not involved in the protein-DNA complex) appears as a single strand, presumably because the excess duplex DNA dissociates under our experimental conditions. Cheng et al. (1996b) also noticed dissociation of the duplex DNA that was not part of their DNA-transcription factor PU.1 complex.

Up to three sodium adducts can be detected despite the extensive dialyses performed in NH_4OAc . No such adducts are observed with the free protein, suggesting that they are associated with the DNA part of the complex. Because of the high polarity of the DNA phosphate backbone, alkali metal cations have a tendency to bind very tightly to oligonucleotides, leading to broad peaks and to erroneous mass measurements (Stults & Marsters, 1991; Potier et al., 1994). However, the resolution of the TOF instrument used here is sufficient to separate the sodium adducts and thus to provide a very precise assignment for the molecular mass of the complex. The appearance of only a few charge states for the complex is in contrast to the broad charge distribution envelope observed for the free protein, and is consistent with a relatively compact or "native" structure for the complex. Such a narrow charge distribution is commonly observed in ESI-MS when the samples are in a nondenaturing buffer solution and are very compact in the gas phase (Light-Wahl et al., 1994; Potier et al., 1997).

Stability of the TrpR-DNA complex

Collision Induced Dissociation (CID-MS) experiments have been proposed as a method of evaluating the stability of noncovalent complexes by mass spectrometry (Light-Wahl et al., 1994; Schwartz et al., 1995). The dissociation can be achieved by increasing the internal energy of the ions entering the mass spectrometer using the declustering voltage (V_c). In our system, increasing the declustering voltage from 100 V to 200 V results in the partial dissociation of the TrpR dimer-DNA1 noncovalent complex to the free TrpR monomer, a point illustrated by comparison of the relative intensities of the two species in Figures 2 and 3. Interestingly, a significant portion of the protein-DNA complex remains intact.

Fig. 3. Stability of the TrpR-DNA complex. ESI m/z spectrum of the TrpR dimer-DNA1 (10 μ M) using a declustering voltage of 200 V. Covalent fragmentations at -150 Da are observed simultaneously with dissociation of the noncovalent complex.

Ions corresponding to the loss of 150 Da ($m/z = 3,374.9$ and $m/z = 3,361.6$ for the 11+ charged ion; $m/z = 3,712.6$ and $m/z = 3,697.1$ for the 10+ charged ion) are also observed (Fig. 3), likely due to breakage of covalent bonds of the protein or of the DNA in the complex. Similar CID experiments performed on the free protein do not produce any subsidiary peaks, suggesting that the 150 Da fragments arise from the DNA, and could actually correspond to depurination events (149 Da). It will be interesting to see if this kind of fragmentation is a feature of other protein-DNA complexes.

Specificity of the TrpR-DNA complex

The specificity of TrpR for its target DNA was investigated in a competitive experiment involving a mixture of three different oligonucleotides (DNA1, DNA2, and DNA3), varying in sequence and length (see Materials and methods). The only difference between the three oligonucleotides was the size of the spacer between the two symmetrically arranged CTAG binding sites. The mass differences between the three oligonucleotides (>650 Da) are large enough that we could readily detect a mixture of the three protein-DNA complexes if they were present in solution.

When the TrpR protein is mixed with an equimolar mixture of the three DNAs, only one complex is observed; it has a molecular mass of 37,265 Da (Fig. 4) corresponding to a TrpR dimer and ds DNA1 which has the correct, consensus spacing of 4 bp. No signal corresponding to free DNA1 is evident, indicating complete conversion into the TrpR-DNA1 complex. There is also a small amount of monomer protein (see also Fig. 2). The two other oligonucleotides, DNA2 and DNA3, appear as free species in the mass spectrum, again as ss DNAs. Moreover, neither DNA2 nor DNA3 is able to bind to the TrpR when mixed separately with the protein (data not shown), indicating that no stable interactions can be observed under our conditions unless the DNA has the proper spacing between the CTAG consensus sequences. This observation is consistent with those obtained by other techniques (Carey et al., 1991; Czernik et al., 1994). Our result shows that significant differences in TrpR binding affinity for different DNAs can be distinguished by ESI-MS and further supports our belief that the mass spectra obtained under our conditions reflect solution specific associations.

Fig. 4. Specificity of the TrpR-DNA complex. ESI mass spectrum of the *trp* repressor (20 μ M) in presence of a mixture of DNA1 (10 μ M), DNA2 (10 μ M), and DNA3 (10 μ M) in 5 mM NH_4OAc , pH 6.0. The declustering voltage was set to 100 V. The calculated molecular weights (MW) are for ds DNA.

Fig. 5. Characterization of the protein-ligand complex. ESI m/z spectrum of *trp* repressor (10 μ M) in 5 mM NH_4OAc , pH 6.0 in presence of 3×10^{-4} M 5-methyl-D-L-tryptophan. $N = 1, 2, \dots, 4$ displays the number of methyl-tryptophans attached to the TrpR monomer. The declustering voltage V_c was 80 V. The inset shows a deconvoluted mass spectrum obtained from the m/z spectrum.

Interaction between the protein, the DNA, and tryptophan

The tryptophan analogue, 5-methyl-L-tryptophan binds 4.5 times better to the TrpR protein than does L-tryptophan (Marmorstein et al., 1987) and should stabilize the homodimer (Arvidson et al., 1986). Unfortunately, only the racemic mixture was available, so samples were prepared in NH_4OAc buffer with 3×10^{-4} M 5-methyl-D-L-tryptophan (i.e., 30-fold excess compared to the protein concentration). As shown in Figure 5, complexes of TrpR monomer with up to five ligand molecules are detected. The free unfolded monomer remains the major species, however, and the relative abundances of the monomer and the aggregates (dimer and trimer) appear unchanged. The charge state distribution of the TrpR-5-methyl-tryptophan complexes is no different from that of the free protein, indicating that the ligand is unable to stabilize the native protein conformation under these conditions. These observations, as well as the requirement of a large concentration of ligand to detect any binding, strongly suggest that these bound tryptophans are simply adducts, resulting from nonspecific associations.

With the DNA-protein complex the situation is very different. Figure 6 displays the mass spectra obtained after the addition of 0, 1, 2, and 10 equivalents of L-tryptophan (0, 10, 20, and 100 μ M) to the TrpR-DNA1 complex at 10^{-5} M in 5 mM NH_4OAc . The signal corresponding to the TrpR dimer-DNA1 complex has been replaced with new ions, with mass differences equal to the mass of L-tryptophan. After the addition of two equivalents of L-tryptophan, the major species is a noncovalent complex formed by one TrpR dimer, one ds DNA, and two L-tryptophan molecules, as expected for the holo-repressor-DNA complex (measured mass: 37,675 Da; expected mass: 37,668 Da). The two equivalents of L-tryptophan are fully bound at a total corepressor concentration of 2×10^{-5} M, one order of magnitude less than the concentration used with the free protein (Fig. 5). Whereas TrpR dimer has been shown to contain only two binding sites for L-tryptophan (Arvidson et al., 1986), up to five adducts can be detected by ESI-MS when an excess of L-tryptophan is present in solution. It is not possible to define where these extra tryptophans bind from the ESI-MS data alone. However, despite the excess of the added ligand, the most abundant species still corresponds to the binding of two L-tryptophan molecules to the protein-DNA1 complex, indicating that the latter

Fig. 6. Study of the affinity of L-tryptophan for the TrpR-DNA complex. ESI mass spectra of the TrpR-DNA1 complex ($10 \mu\text{M}$) in presence of (A) 0, (B) $1 \mu\text{M}$, (C) $20 \mu\text{M}$, and (D) $100 \mu\text{M}$ of L-tryptophan. Two charge states ($10+$ and $11+$) of the ternary TrpR-DNA1-L-tryptophan complexes are shown. $N = 0, 1, 2, \dots, 4$ corresponds to the number of L-tryptophan bound to the protein-DNA complex. All spectra were acquired in 5 mM NH_4OAc , pH 6.0 at $V_c = 60 \text{ V}$.

is the specific complex. The additional peaks probably arise from nonspecific aggregation on the droplet surface or binding to low affinity sites (see below). Their pattern is reminiscent of the nonspecific multi-adducts observed for the free protein. Similar results were obtained by using 5-methyl-tryptophan instead of L-tryptophan.

In these ligand binding experiments, the interface conditions had to be slightly modified from those used for the apo-repressor-DNA complex. In particular, the declustering voltage (V_c) was reduced from 100 V (value giving the best signal to noise ratio for the protein-DNA complex) to 60 V in order to prevent ligand dissociation in the interface. This lower declustering voltage gave the best balance between an adequate desolvation (leading to sharp peaks and thus precise mass measurements) and the preservation of the noncovalent interactions holding all the components together. A further reduction of the internal energy of the ions by decreasing the declustering voltage to 40 V or the capillary heating to 50°C only led to a broadening of the peaks, a characteristic of incomplete desolvation (data not shown). Increasing the declustering voltage above 100 V resulted in the disappearance of the peaks corresponding to the ternary complex in favor of the TrpR dimer-ds DNA1 complex. The ternary holo-repressor-DNA complex appears, therefore, to be more sensitive to gas phase collisions than the apo-repressor-DNA1 complex.

Our observation of a stable, specific, ternary holo-repressor-DNA complex in the gas phase in the absence of "stable bound"

internal water molecules is in seeming contradiction to results obtained from crystallographic data (Otwinowski et al., 1988; Lawson & Carey, 1993). On the other hand, it is consistent with the conclusions of Youderian and Arvidson (1994), who argue on the basis of biological data, that these waters may indeed be present to mediate the DNA-protein interaction in the solid, but not in solution.

However, it is likewise not completely clear how well the present results in the gas phase reflect the state of the complex in solution. A loss of water molecules might occur in the $\sim 1 \text{ ms}$ flight time in the atmosphere-vacuum interface (pressure $\sim 2.5 \text{ torr}$) or in the high vacuum of the mass spectrometer, if the complex is flexible enough. The lifetime of the L-tryptophan ligand in the ternary complex, as determined by NMR spectroscopy, is $\sim 300 \text{ ms}$ at 45°C (Lee et al., 1995). Likewise, the upper limit for the lifetime of "stable bound" internal water molecules of proteins and protein complexes is 1 ms (Denisov & Halle, 1994; Dötsch & Wider, 1995). Thus it would be possible to rationalize our observation of a tryptophan-containing complex with the escape of bound waters, if it is assumed that the relative probabilities for loss of tryptophan or water in the gas phase are similar to the probabilities for exchange in solution. We note that in the MS study of a specific complex of PU.1 with DNA (Cheng et al., 1996b), no stable bound water molecules were observed in the gas phase despite the presence of many water-mediated interactions seen in the protein-DNA co-crystal (Kodandapani et al., 1996). On the other hand, Chung et al. (1997) have some evidence that solvent waters involved in the DNA gyrase-novobiocin complex are retained in ESI MS. The question obviously requires further study.

Comparison of the affinity and specificity of L- and D-tryptophan

The affinity and specificity of the binding of L-tryptophan ($K_D = 14.6 \mu\text{M}$) were evaluated by comparison with those of its D-enantiomer ($K_D = 343 \mu\text{M}$) (Marmorstein et al., 1987). Figure 7 displays the mass spectra obtained after the addition of 0, 1, 2, and 10 equivalents of D-tryptophan to the protein-DNA1 complex under the same experimental conditions as those used for L-tryptophan binding. These spectra reveal that D-tryptophan also binds to the TrpR dimer-DNA1 complex. Complexes containing up to five D-tryptophan molecules are detected after the addition of 10 equivalents of D-tryptophan (10^{-4} M). However, in contrast to the L-tryptophan data, there is no preference for the incorporation of two D-tryptophan molecules, indicating that all of the ligand molecules bind to the complex with a similar affinity, a characteristic of nonspecific associations.

Figure 8 displays the average number of tryptophan molecules bound to the complex (as determined by ESI-MS) versus the concentration of L- and D-tryptophan in solution. To produce this figure, the average number of tryptophans bound was calculated by integrating all peaks corresponding to the incorporation of one or more tryptophans for each charge state. This graph further displays the different behavior of the two enantiomers. The curve obtained for L-tryptophan displays two distinct parts. The first part confirms the presence of two high affinity binding sites since the first two molecules bind quantitatively to the protein-DNA complex. The second part of the curve has a much shallower slope, reflecting the lower affinity of sites for additional tryptophans. The curve obtained for D-tryptophan displays a lower slope throughout, reflecting the weaker binding affinity of D-tryptophan. A total ligand concentration of $7 \times 10^{-5} \text{ M}$ (as compared to $2 \times 10^{-5} \text{ M}$ for

Fig. 7. Study of the affinity of D-tryptophan for the TrpR-DNA complex. ESI mass spectra of the TrpR-DNA1 complex ($10\ \mu\text{M}$) in presence of (A) 0, (B) $10\ \mu\text{M}$, (C) $20\ \mu\text{M}$, and (D) $100\ \mu\text{M}$ of D-tryptophan. Two charge states ($10+$ and $11+$) of the ternary TrpR dimer-DNA1-D-tryptophan complexes are shown. $N = 0, 1, 2, \dots, 5$ corresponds to the number of D-tryptophan bound to the protein-DNA complex. All spectra were acquired in $5\ \text{mM}\ \text{NH}_4\text{OAc}$, pH 6.0 at $V_c = 60\ \text{V}$.

L-tryptophan) is required to achieve an average number of two bound D-tryptophans.

Why mass spectrometry?

The interactions between TrpR, three DNAs, and three ligands have been successfully studied by ESI-MS. The results show many similarities, but also some differences, from previously reported data for the TrpR system in solution and in crystalline form. Although TrpR was shown to be a stable homodimer by NMR and X-ray crystallography, it is detected as a monomer by ESI-MS. Thus, the requirement for low ionic strength volatile buffers provides a severe limitation of ESI-MS for the characterization of this DNA-binding protein in the absence of DNA. The high number of charge states obtained in the mass spectrum as well as the inability to bind its operator DNA suggest that the apo-protein is in a partially unfolded state under the low salt conditions used. Site-specific binding of tryptophan is not observed with free TrpR, indicating that the ligand is not able to stabilize the native protein conformation under these conditions. Thus, the noncovalent complexes between the free protein and the tryptophan appear to arise from a nonspecific aggregation of tryptophan with the unfolded protein. The use of this buffer may also be responsible for the denatured DNA (observed as single stranded in the mass spectra). It is well known that the duplex stability in solution depends on

Fig. 8. The average number of ligand molecules bound to the TrpR-DNA1 complex was determined for L- and D-tryptophan as a function of the ligand concentration. An average number of two ligands bound is observed for a total ligand concentration of $2 \times 10^{-5}\ \text{M}$ for L-tryptophan and $7 \times 10^{-5}\ \text{M}$ for D-tryptophan.

temperature and ionic strength (Light-Wahl et al., 1993; Cheng et al., 1996b).

Only when the DNA and the protein are mixed under conditions in which they are both native and active and later dialyzed against NH_4OAc was a stable noncovalent complex detected. Our findings establish that the observed protein-DNA ions are not the result of artefactual nonspecific associations, but rather reflect the macromolecular complexes present in solution (for review, see Loo, 1997). First, no complex was detected using a denatured protein, showing that the tertiary structure of the protein was required for formation of the complex. Second, only the complex with the expected stoichiometry was obtained $[\text{TrpR dimer:dsDNA}] = [1:1]$. Indeed, if the observed noncovalent ions had resulted from random aggregation, one would expect to detect a mixture of species having $[\text{protein monomer:DNA}] = [1:1], [1:2], [2:1], [2:2]$, etc. Third, the specific addition of two L-tryptophans to the protein-DNA complex confirms the dynamic nature of the complex (Lee et al., 1995), even under such stringent conditions that the free protein by itself is probably inactivated. Finally, the TrpR protein binds only to DNA with a specific consensus sequence, containing two CTAG boxes linked by a spacer of 4 bp. Thus, we conclude that the protein-DNA complexes detected by ESI-MS did not arise from a simple aggregation of the species during the passage to the gas phase (i.e., the ions were not only due to nonspecific electrostatic interactions that might be stabilized in vacuum), but are derived from associations formed in solution.

From these results as well as recent reports from other groups (Loo et al., 1997), a relationship between the stability of noncovalent complexes in the ESI process and the type of interaction involved in the complex formation seems to emerge. In the case of a leucine-zipper peptide (Li et al., 1993, 1994) or of interactions between acyl coenzyme A-binding protein and acyl CoA analogs (Robinson et al., 1996), ESI data did not reflect the relative abundance of the noncovalent interactions in solution, presumably because these interactions were mostly hydrophobic. On the other hand, Cheng et al., (1996a) reported that mild interface conditions were not necessary for the observation of a gene V protein-DNA complex because of the substantial electrostatic component of the

interactions. In the case of the *trp* repressor, the protein-DNA complex appears to be very stable in the gas phase since dissociation could not be achieved except at energies causing breakage of a covalent bond. Such covalent fragmentations in noncovalent complexes have already been described (Feng, 1995; Potier et al., 1997). Although no general rule can yet be proposed, this result suggests that ESI-MS should be broadly applicable to the study of protein-DNA interactions, provided that carefully controlled experiments are performed.

In these measurements we have used a time-of-flight mass analyzer for the observation of the noncovalent complexes. This instrument has the advantage of an unlimited (in principle) m/z range, while most commercial instruments of other types (e.g., quadrupole mass filters) are restricted to $m/z < 4,000$. Although most of the ions observed in the present case had m/z values within this limited range, our negative results (for example, the absence of random multimers in the spectrum of the TrpR-DNA complex) would be rather unconvincing without the ability to search large m/z values. Moreover, instrumentation with an extended m/z range will certainly be required for the study of larger protein-DNA complexes, because most of the DNA-binding proteins are active as multimeric entities, leading to decreases in the number of available protonations due to coulombic restrictions. Such a need for high m/z range analyzers has been discussed (Light-Wahl et al., 1994; Loo, 1995), especially for characterization of multimeric proteins.

Materials and methods

Electrospray mass spectrometry

All studies were performed using an ESI-TOF mass spectrometer constructed at the University of Manitoba (Verentchikov et al., 1994; Krutchinsky et al., 1995). Samples were dissolved in 5 mM NH_4OAc pH 6.0 at a final concentration of 10^{-5} M and were continuously infused into the ion source at a flow rate of 0.4 $\mu\text{L}/\text{min}$ using a Harvard Model 11 syringe pump (Harvard Apparatus). To preserve the noncovalent complexes, relatively mild interface conditions were used. The declustering voltage (V_c), which controls the kinetic energy of the ions in the interface, ranged from 40 V to 100 V, and the capillary temperature was set to 100 °C to optimize desolvation and minimize thermal denaturation of the sample. Harsher interface conditions (V_c up to 250 V) were used to perform collision induced dissociation (CID) of the complexes. Data were acquired in the positive mode and calibration was performed using the multiply-charged ions produced by a separate injection of substance P dissolved in a mixture of water/methanol: 1/1 with 1% acetic acid. Some data were also acquired in the negative mode, but these were of lower quality and did not yield additional information.

Chemicals

Substance P was purchased from Sigma. Ammonium acetate (NH_4OAc , 99.999%), L-, D-, and 5-methyl-D,L-tryptophan were obtained from Aldrich. The concentrations of the tryptophans were determined assuming extinction coefficients of $5.26 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$ for 5-methyl-tryptophan and $5.47 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$ for L- and D-tryptophan (Marmorstein et al., 1987). All samples were dialyzed using 12-14000 MWCO dialysis membrane (Spectra/Por).

Sample preparation

The *trp* apo-repressor (TrpR) was purified as previously described (Paluh & Yanofsky, 1986; Arrowsmith et al., 1990), and lyophilized in 0.5 mL aliquots in elution buffer (10 mM NaPO_4 pH 7.6, 0.1 mM EDTA, 0.45 M NaCl). The protein was recovered by adding 0.5 mL of water and the concentration was measured precisely from the absorbance at 280 nm using an extinction coefficient of $1.5 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ (Joachimiak et al., 1983).

Self-complementary oligonucleotides were synthesized on a 200 nmole scale and purified by HPLC (Integrated DNA Technologies). The DNA was dissolved in nanopure water at 1 nmol/ μL as measured by absorbance at 260 nm. Prior to mass spectrometric analysis, an aliquot was mixed with an equal volume of 2 M NH_4OAc and annealed by heating to 80 °C and slowly cooling to 4 °C. The three oligonucleotides used in this work are shown below, with the TrpR recognition sites underlined:

DNA1: 5'CCGTACTAGTTTAACTAGTTACG 3'

MW (calc,ds) = 12,812 Da

DNA2: 5'CCGTACTAGTACTAGTTACG 3'

MW (calc,ds) = 11,578 Da

DNA3: 5'CGTACTAGTTTTAACTAGTTACG 3'

MW (calc,ds) = 13,469 Da. (1)

Individual protein (80 μM in elution buffer) and DNA (50 μM in 1 M NH_4OAc) samples were mixed to give 50 μL solutions of the desired ratios and assembled in "waterbugs" (Orr et al., 1995). These mixtures were dialyzed first against 90 mM NaCl, 20 mM NaH_2PO_4 , pH 6.0 buffer for 2 h and then against 100 mL of 10 mM NH_4OAc pH 6.0, six times over 48 h. All dialyses were done at 4 °C. Prior to mass analysis, the samples were diluted to a final concentration of 10 μM in 5 mM NH_4OAc , pH 6.0.

5-Methyl-, L-, and D-tryptophan were dissolved in 5 mM NH_4OAc at a concentration of 10^{-3} M. The ternary noncovalent complexes (TrpR-DNA-ligands) were formed by adding 1-10 molar equivalents of ligand to the protein-DNA1 complex in 5 mM NH_4OAc .

TrpR samples for CD and fluorescence were prepared in parallel to those prepared for mass spectrometry. Two identical samples were dialyzed first against 50 mM sodium phosphate, 150 mM NaCl, pH 7.6, buffer conditions that favor the dimer. After 4 h, one sample was removed to 10 mM NH_4OAc and prepared exactly as for ESI. After 48 h, both samples were recovered for experiments. The sample in 10 mM NH_4OAc was split into two aliquots; one was diluted with an equal volume of nanopure water, the other was diluted with an equal volume of 2% acetic acid in methanol. ESI and CD spectra were acquired at the same time. CD spectra were acquired in the near UV region in a 0.1 cm path quartz cuvette using a Jasco 500 spectropolarimeter. Fluorescence measurements were carried out in a 1 cm path quartz cuvette using a Fluoro IV fluorimeter (Gilford Instruments). Samples were excited at 280 nm and emission spectra collected in the wavelength range of 500 to 300 nm.

Acknowledgments

We are grateful to Victor Spicer for technical assistance.

This work was supported by NSERC (Canada) research grants to HWD and KGS, by an NSERC Collaborative Project Grant to HWD, KGS, and WE and by the National Cancer Institute of Canada with funds from the

Canadian Cancer Society to CHA. NP was funded by a grant from "L'association français pour la recherche thérapeutique." CHA is a Research Scholar of the Medical Research Council of Canada.

References

- Arrowsmith CH, Pachter R, Altman RB, Iyer SB, Jardetzky O. 1990. Sequence-specific ^1H NMR assignments and secondary structure in solution of *Escherichia coli trp* repressor. *Biochemistry* 29:6332-6341.
- Arvidson DN, Bruce C, Gunsalus RP. 1986. Interaction of the *Escherichia coli trp* aporepressor with its ligand, L-tryptophan. *J Biol Chem* 261:238-243.
- Baca M, Kent SBH. 1992. Direct observation of a ternary complex between the dimeric enzyme HIV-1 protease and a substrate-based inhibitor. *J Am Chem Soc* 114:3992-3993.
- Carey J. 1988. Gel retardation at low pH resolves *trp* repressor-DNA complexes for quantitative study. *Proc Natl Acad Sci USA* 85:975-979.
- Carey J, Lewis DEA, Lavoie TA, Yang J. 1991. How does *trp* repressor bind to its operator? *J Biol Chem* 266:24509-24513.
- Cheng X, Harms AC, Goudreau PN, Terwilliger TC, Smith RD. 1996a. Direct measurement of oligonucleotide binding stoichiometry of gene V protein by mass spectrometry. *Proc Natl Acad Sci USA* 93:7022-7027.
- Cheng X, Morin PE, Harms AC, Bruce JE, Ben-David Y, Smith RD. 1996b. Mass spectrometric characterization of sequence-specific complexes of DNA and transcription factor PU.1 DNA binding domain. *Anal Biochem* 239:35-40.
- Chowdhury SK, Katta V, Chait BT. 1990. Probing conformational changes in proteins by mass spectrometry. *J Am Chem Soc* 112:9012-9013.
- Chung EW, Rostom A, Robinson CV. 1997. Investigating non-covalent protein-ligand interactions using nanoflow electrospray. In: *Proceedings of the 45th ASMS Conference on Mass Spectrometry and Allied Topics*. Palm Springs, CA: ASMS. p 359.
- Czernik PJ, Shin DS, Hurlburt BK. 1994. Functional selection and characterization of DNA binding sites for *trp* repressor of *Escherichia coli*. *J Biol Chem* 269:27869-27875.
- Denisov VP, Halle B. 1994. Dynamics of the internal and external hydration of globular proteins. *J Am Chem Soc* 116:10324-10325.
- Dötsch V, Wider G. 1995. Exchange rates of internal water molecules in proteins measured using pulsed field gradients. *J Am Chem Soc* 117:6064-6070.
- Feng R. 1995. Unusually strong binding of a noncovalent gas-phase spermine-peptide complex and its dramatic temperature dependence. In: *Proceedings of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics*. Atlanta, GA: ASMS. p 1264.
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. 1990. Electrospray ionization for mass spectrometry of large biomolecules. *Mass Spectrom Rev* 9:37-70.
- Fitzgerald MC, Chernushevich I, Standing KG, Whitman CP, Kent SBH. 1996. Probing the oligomeric structure of an enzyme by electrospray ionization time-of-flight mass spectrometry. *Proc Natl Acad Sci USA* 93:6851-6856.
- Gale DC, Smith RD. 1995. Characterization of noncovalent complexes formed between minor groove binding molecules and duplex DNA by electrospray ionization-mass spectrometry. *J Am Soc Mass Spectrom* 6:1154-1164.
- Ganem B, Li YT, Henion JD. 1991. Observation of non-covalent enzyme-substrate and enzyme-product complexes by ion-spray mass spectrometry. *J Am Chem Soc* 113:7818-7819.
- Joachimski A, Kelly RL, Gunsalus RP, Yanofsky C, Sigler PB. 1983. Purification and characterization of *Trp* aporepressor. *Proc Natl Acad Sci USA* 80:668-672.
- Kodandapani R, Pio F, Ni C-Z, Piccialli G, Klemsz M, McKercher S, Maki RA, Ely KR. 1996. A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. *Nature* 380:456-460.
- Krutchinsky AN, Chernushevich IV, Spicer V, Ens W, Standing KG. 1995. Collisional focusing of ions in the electrospray TOF mass spectrometer. In: *Proceedings of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics*. Atlanta, GA: ASMS. p 126.
- Lawson CL, Carey J. 1993. Tandem binding in crystals of a *trp* repressor/operator half-site complex. *Nature* 366:178-182.
- Lee W, Revington M, Farrow NA, Nakamura A, Utsunomiya-Tate N, Miyake Y, Kainosho M, Arrowsmith CH. 1995. Rapid corepressor exchange from the *trp*-repressor/operator complex: An NMR study of [ul- ^{13}C / ^{15}N]-L-tryptophan. *J Biomolec NMR* 5:367-375.
- Li YT, Hsieh YL, Henion JD, Ocain TD, Schiehsler GA, Ganem B. 1994. Analysis of the energetics of gas-phase immunophilin ligand complexes by ion spray mass spectrometry. *J Am Chem Soc* 116:7487-7493.
- Li YT, Hsieh YL, Henion JD, Senko MW, McLafferty FW, Ganem B. 1993. Mass spectrometric studies on noncovalent dimers of leucine zipper peptides. *J Am Chem Soc* 115:8409-8413.
- Light-Wahl KJ, Schwartz BL, Smith RD. 1994. Observation of the noncovalent quaternary associations of proteins by electrospray ionization mass spectrometry. *J Am Chem Soc* 116:5271-5278.
- Light-Wahl KJ, Springer DL, Winger BE, Edmonds CG, Camp DG II, Thrall BD, Smith RD. 1993. Observation of a small oligonucleotide duplex by electrospray ionization mass spectrometry. *J Am Chem Soc* 115:803-804.
- Loo JA. 1995. Observation of large subunit protein complexes by electrospray ionization mass spectrometry. *J Mass Spectrom* 30:180-183.
- Loo JA. 1997. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom Rev* 16:1-23.
- Loo JA, Hu P, McConnell P, Mueller WT, Sawyer TK, Thanabal V. 1997. A study of Src SH2 domain protein-phosphopeptide binding interactions by electrospray ionization mass spectrometry. *J Am Soc Mass Spectrom* 8:234-243.
- Loo JA, Ogorzalek-Loo RR, Udseth HR, Edmonds CG, Smith RD. 1991. Solvent-induced conformational changes of polypeptides probed by electrospray-ionization mass spectrometry. *Rapid Commun Mass Spectrom* 5:101-105.
- Mann M, Wilm M. 1995. Electrospray mass spectrometry for protein characterization. *Trends Biochem Sci* 20:219-224.
- Marmorstein RQ, Joachimski A, Sprinzl M, Sigler PB. 1987. The structural basis for the interaction between L-tryptophan and the *Escherichia coli trp* aporepressor. *J Biol Chem* 262:4922-4927.
- Orr A, Ivanova VS, Bonner WM. 1995. "Waterbug" dialysis. *Biotechniques* 19:204-206.
- Otwinski Z, Schevitz RW, Zhang RG, Lawson CL, Joachimski A, Marmorstein RQ, Luisi BF, Sigler PB. 1988. Crystal structure of *trp* repressor/operator complex at atomic resolution. *Nature* 335:321-329.
- Paluh JL, Yanofsky C. 1986. High level production and rapid purification of the *E. coli trp* repressor. *Nucleic Acids Res* 14:7851-7860.
- Potier N, Barth P, Tritsch D, Biellmann JF, Van Dorsselaer A. 1997. Study of non-covalent enzyme-inhibitor complexes of aldose reductase by electrospray mass spectrometry. *Eur J Biochem* 243:274-282.
- Potier N, Van Dorsselaer A, Cordier Y, Roch O, Bischoff R. 1994. Negative electrospray ionization mass spectrometry of synthetic and chemically modified oligonucleotides. *Nucleic Acids Res* 22:3895-3903.
- Ramesh V, Frederick RO, Syed SEH, Gibson CF, Yang J-C, Roberts GCK. 1994. The interactions of *Escherichia coli trp* repressor with tryptophan and with an operator oligonucleotide. *Eur J Biochem* 225:601-608.
- Robinson CV, Chung EW, Kragelund BB, Knudsen J, Apelin RT, Poulsen FM, Dobson CM. 1996. Probing the nature of noncovalent interactions by mass spectrometry. A study of protein-CoA ligand binding and assembly. *J Am Chem Soc* 118:8646-8653.
- Schwartz BL, Bruce JE, Anderson GA, Hofstadler SA, Rockwood AL, Smith RD, Chilkoti A, Stayton PS. 1995. Dissociation of tetrameric ions of non-covalent streptavidin complexes formed by electrospray ionization. *J Am Soc Mass Spectrom* 6:459-465.
- Smith DL, Zhang Z. 1994. Probing noncovalent structural features of proteins by mass spectrometry. *Mass Spectrom Rev* 13:411-429.
- Stults JT, Marsters JC. 1991. Improved electrospray ionization of synthetic oligodeoxynucleotides. *Rapid Commun Mass Spectrom* 5:359-363.
- Verentchikov AN, Ens W, Standing KG. 1994. Reflecting time-of-flight mass spectrometer with an electrospray ion source and orthogonal extraction. *Anal Chem* 66:126-133.
- Wu Q, Gao J, Joseph-McCarthy D, Sigal GB, Bruce JE, Whitesides GM, Smith RD. 1997. Carbonic anhydrase-inhibitor binding: From solution to the gas phase. *J Am Chem Soc* 119:1157-1158.
- Yoderian P, Arvidson DN. 1994. Direct recognition of the *trp* operator by the *trp* holorepressor—A review. *Gene* 150:1-8.
- Zhang H, Zhao D, Revington M, Lee W, Jia X, Arrowsmith C, Jardetzky O. 1994. The solution structures of the *trp* repressor-operator DNA complex. *J Mol Biol* 238:592-614.
- Zhang RG, Joachimski A, Lawson CL, Schevitz RW, Otwinowski Z, Sigler PB. 1987. The crystal structure of *trp* aporepressor at 1.8 Å shows how binding tryptophan enhances DNA affinity. *Nature* 327:591-596.
- Zhao D, Arrowsmith CH, Jardetzky O. 1993. Refined solution structures of the *Escherichia coli trp* holo- and aporepressor. *J Mol Biol* 229:735-746.