Hairpin Formation within the Human Enkephalin Enhancer Region. 2. Structural Studies

ARTICLE *in* BIOCHEMISTRY · NOVEMBER 1994

Impact Factor: 3.02 · DOI: 10.1021/bi00205a035 · Source: PubMed

Hairpin Formation within the Human Enkephalin Enhancer Region. 2. Structural Studies†

Cynthia T. McMurray,*,‡,§ Nenad Juranić,‡ Subramanian Chandrasekaran,∥ Slobodan Macura,‡ Ying Li,∥ Robert L. Jones,⊥ and W. David Wilson∥

Department of Biochemistry and Molecular Biology and Department of Pharmacology, Mayo Foundation, Mayo Graduate School, Rochester, Minnesota 55905, Department of Chemistry, Georgia State University, Atlanta, Georgia 30303, and Chemistry Department, Emory University, Atlanta, Georgia 30303

Received April 21, 1994; Revised Manuscript Received July 20, 1994®

ABSTRACT: Receptor-mediated induction of the human proenkephalin gene has been mapped to an imperfect palindrome located between −104 and −86, upstream of the transcriptional start site. Several lines of evidence suggest that receptor-mediated transcription of proenkephalin involves a reversible conformational change from duplex to a hairpin state of the enhancer [McMurray, C. T., Wilson, W. D., & Douglass, J. O. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 666]. To determine the structure that would form if such a conformational change took place, we have synthesized two 23-bp oligonucleotides, d(GCTGGCG-TAGGGCCTGCGTCAGC) and d(GCTGACGCAGGCCCTACGCCAGC), whose sequences are identical to the top and the bottom strands of the native enhancer. We have found that each oligonucleotide strand exists primarily as a hairpin structure over a wide range of oligonucleotide concentrations and a wide range of temperatures (0–45 °C). The assignment of each imino proton was carried out using 1D and 2D nuclear Overhauser effects (NOE) and by comparison with the spectra of hairpins containing single base substitutions. The hairpin structure for each oligonucleotide contains a 3-member loop, a 10-bp stem, and two mismatched pairs. The hairpin that forms from the top strand of the enhancer and contains two GT mispaired bases creates an alternative binding site for the cyclic adenosine monophosphate element binding protein (CREB), a transcription factor that binds to and regulates the human proenkephalin gene. Circular dichroism and ^{31}P NMR indicate that, despite the presence of mismatched pairs, each oligonucleotide hairpin adopts a B-form conformation with no unusual bending or kinking. The structure of the hairpin may explain the effect on expression of point mutations within the enhancer.

The enkephalin gene is one of many neuroendocrine peptide genes in which gene transcription is induced as a result of receptor activation and second-message production (Mayer, 1993; Douglass et al., 1991; Walton & Rehfuss, 1990; Goodman, 1990). Receptor-mediated induction of proenkephalin has been mapped to a short element located between −104 and −86, upstream of the transcriptional start site (Comb et al., 1986). This element has been shown to act as a transcriptional enhancer and to mediate a response to cyclic AMP (cAMP[1])-, phorbol ester-, and Ca^{2+}-dependent intracellular signaling pathways (van Nguyen et al., 1990; Comb et al., 1986, 1988). Deletion analysis in this region has indicated that two elements, termed CRE-1 and CRE-2, work in concert to constitute a cAMP-responsive enhancer (Spiro et al., 1993; Comb et al., 1986, 1988). Several proteins are known to bind within the proenkephalin enhancer region (Chu et al., 1994; Spiro et al., 1993; Konradi et al., 1993; Kobierski et al., 1991; Sonnenberg, 1989; Comb et al., 1988). While mutational analysis within the human proenkephalin enhancer is consistent with the binding of a CREB-like or AP-1-like protein within CRE-2, several independent point mutations that alter gene expression do not affect the binding of known transcriptional proteins (Comb et al., 1988). The results may be due to the binding of an uncharacterized protein. Alternatively, the expression data from the point mutations are consistent with the effect of altering hairpin stability in the enhancer (Spiro et al., 1993; McMurray et al., 1991).

CRE-1 and CRE-2 are contained within a 23-base pair, nearly palindromic sequence. We have previously shown that the enhancer 23-bp duplex can spontaneously and reversibly dissociate into two stable hairpin structures that contain mismatched base pairs (McMurray et al., 1991). The top strand of the proenkephalin duplex forms a hairpin with two non-Watson–Crick GT pairs (GT hairpin); the bottom strand of the duplex forms with two non-Watson–Crick AC mismatched pairs (AC hairpin) (Figure 1A,B, II and IV). Both hairpin arms and, specifically, the mismatched base pairs are implicated in the regulation of proenkephalin expression. Proenkephalin transcription is regulated by the cAMP response element binding protein (CREB), among other proteins (Chu, 1994; Spiro et al., 1993; Konradi et al., 1993; Huggenvik et al., 1991). *Invivo*, CREB-induced activation of proenkephalin requires both CRE-1 and CRE-2, even though a binding site for CREB is found only in CRE-2 (Spiro et al., 1993). Despite the presence of a CRE-2 binding site in the linear duplex enhancer, CREB binds with high-affinity binding to an

† This work was supported by National Science Foundation Grant IBN 9222848 to C.T.M.

* Author to whom correspondence should be addressed.
‡ Department of Biochemistry and Molecular Biology, Mayo Graduate School.
§ Department of Pharmacology, Mayo Graduate School.
∥ Georgia State University.
⊥ Emory University.

® Abstract published in *Advance ACS Abstracts,* September 1, 1994.

[1] Abbreviations: bp, base pair; cAMP, cyclic adenosine monophosphate; CRE, cAMP response element; CREB, CRE binding protein; CREB-341, full-length CREB; AC hairpin, 23-bp oligonucleotide strand of the enkephalin enhancer that forms a hairpin containing two AC mismatches; GT hairpin, 23-bp oligonucleotide strand of the enkephalin enhancer that forms a hairpin containing two GT mismatches; NOE, nuclear Overhauser effect; EDTA, ethylenediaminetetraacetic acid; Pipes, piperidine-*N*,*N*′-bis(2-ethanesulfonic acid); Pipes-00, 10 mM Pipes and 0.1 mM EDTA; T_m, the midpoint of the melting transition.

0006-2960/94/0433-11960$04.50/0

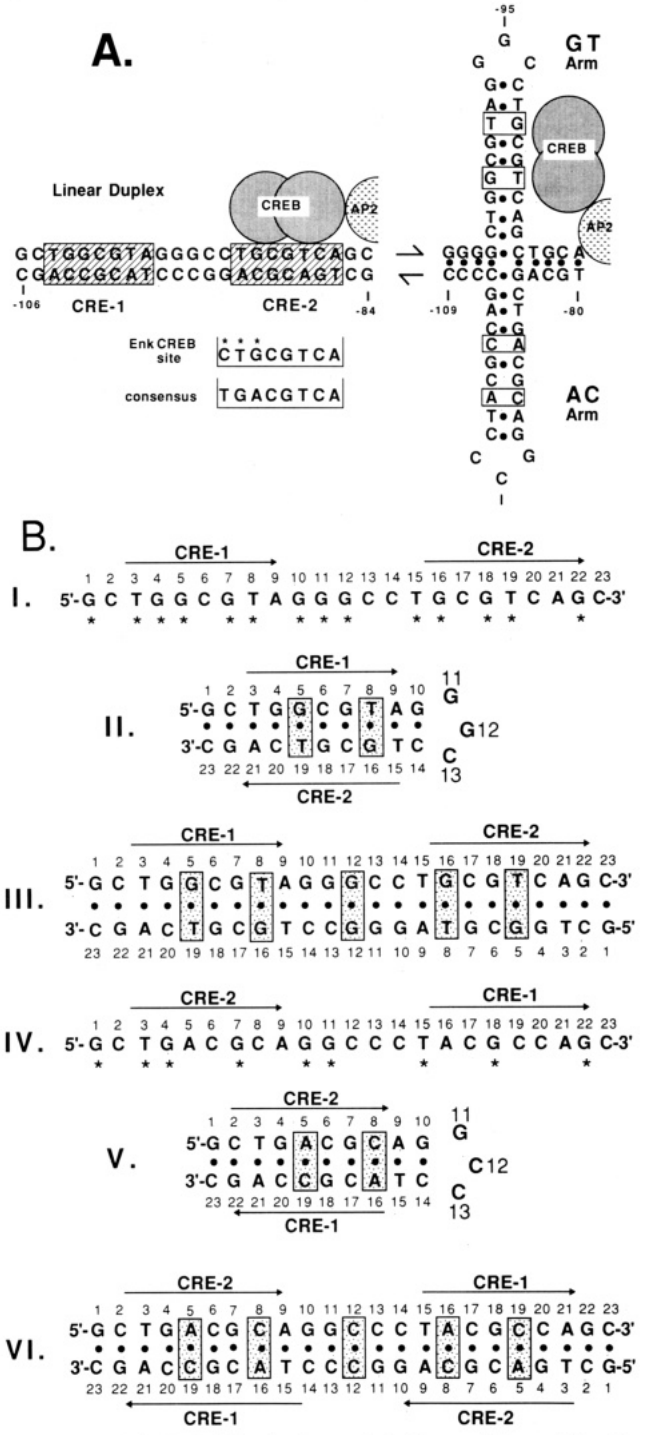

FIGURE 1: (A) Model for duplex to hairpin transition within the human proenkephalin enhancer. (Left) The linear duplex weakly binds CREB at a site that differs from the consensus binding site in three nucleotides (*). (Right) CREB strongly binds to the GT hairpin arm formed from the top strand of the enhancer. The CREB binding site in the hairpin contains two GT mismatched pairs. The bottom strand forms a hairpin with mismatched AC pairs. (B) Summary of possible structures for the complementary oligonucleotides that compose the human enhancer in the absence of their partner strands. CRE-1 and CRE-2 indicate cAMP-dependent regulatory regions within the enhancer. Boxes indicate mismatched base pairs. I and IV represent the oligonucleotide strands that form the top (I) or the bottom (IV) of the human enkephalin enhancer. Stars indicate nucleotides with observable imino protons. II and V are potential GT (II) and AC (V) hairpin structures, and III and VI are potential GT (III) and AC (VI) homoduplex structures.

alternative binding site in the double-stranded stem of the GT hairpin that forms from the enhancer (Figure 1A). Formation of the CREB binding site requires both CRE-1 and CRE-2 (Spiro et al., 1993; McMurray et al., 1991), explaining, at least in part, the requirement of both CREs in transcriptional regulation by CREB (Figure 1A). The double-stranded binding site that is created in the stem of the hairpin arm has a sequence identical to that found in the linear duplex, with slight modifications contributed by the presence of the mismatched base pairs (Spiro et al., 1993). High-affinity CREB binding to the hairpin is not dependent on hairpin conformation, but requires the specific sequence that is found only in the double-stranded stem of the hairpin. The T in one of the GT mismatches determines the hairpin binding specificity of CREB-341 (Spiro et al., 1993). CREB binding to the hairpin is sequence specific and reversible (Spiro et al., 1993).

The bottom hairpin that forms from the enhancer appears to operate as a proton switch that determines the overall stability of the potential cruciform. We have shown that, even in the presence of its hydrogen-bonded partner strand, a synthetic enkephalin enhancer duplex will completely convert into two stable hairpin structures, the stabilities of which are determined by the local pH (McMurray et al., 1991). Lowering the pH from 8.0 to 6.0 results in a 30 °C increase in the T_m (°C) for the AC hairpin. At pH 6.0, the AC hairpin alone is more stable than the duplex enhancer under low ionic strength conditions (McMurray et al., 1991). The pK_a of AC hairpin protonation is 7.2, and van't Hoff analysis indicates that two protons bind within the AC hairpin at approximately 1 pH unit below the pK_a (McMurray et al., 1991). Although the site of protonation has not been precisely determined, it appears that a likely location is the N1 of adenine (Patel et al., 1984b; Hunter et al., 1986; Gao & Patel, 1987; Kalnik et al., 1988). Protonation enables the formation of two additional hydrogen bonds to stabilize the AC hairpin structure. At the same time, decreasing the pH destabilizes the duplex. Thus, small, local changes of pH can determine whether the enkephalin enhancer exists as a duplex or as a hairpin structure *in vitro*. Stabilization of the hairpin form of the enhancer, at the expense of duplex formation, is driven, in part, by proton transfer to the AC hairpin and by protein binding to the GT hairpin.

DNA secondary structure within the enhancer region appears to play an active role in the cAMP-inducible activation of the human enkephalin gene (Figure 1A). Formation of the secondary structure provides a strong CREB binding site for the lifetime of the hairpin, and the regulation mechanism appears to utilize DNA mispairings that facilitate reversible gene regulation. Therefore, the binding site, the orientation of the mispaired bases, and the local dynamics of the hairpin loop are important features of hairpin regulation. In this initial report, we describe the hydrogen-bonding scheme in the hairpin binding site, the assignment of imino protons, and the conformation of the GT and AC hairpin arms that form from the native enkephalin enhancer. We determine the sequence of the CREB binding site within the GT hairpin, verify the presence of the mismatched base pairs in both the GT and AC hairpins, and determine the size of the hairpin loops. Circular dichroism and ^{31}P NMR indicate that, despite the presence of mismatched pairs, each oligonucleotide hairpin adopts a B-form conformation with no unusual bending or kinking. Due to the importance of the GT mispairings and loop region in stabilizing CREB binding, we evaluate the local dynamics of the loop bases and the mispaired bases by thermal melting studies. The structural features that we observe for the GT hairpin may explain several aspects of CREB binding to the hairpin arm and the results of enhancer mutations on enhancer-dependent expression.

MATERIALS AND METHODS

Oligonucleotide Synthesis. The oligonucleotide 5′-(GCTGGCGTAGGGCCTGCGTCAGC)-3′ and its complement, 5′-(GCTGACGCAGGCCCTACGCCAGC)-3′, were synthesized using standard solid state synthesis techniques by Oligos Etc. (Wilsonville, OR). Oligomers were purified by gel electrophoresis, and the sequence was verified by m13 DNA sequencing (Sanger, 1975). To remove all traces of solvent impurities resulting from the synthesis and purification, the oligomers were purified further by passage through 3–6 Bio-Rad, (Richmond, CA) Econo-Pac 10DG disposable 10 mL desalting columns. The purified oligomers were lyophilized, resuspended in distilled water, and stored at −20 °C until use. The purity of each strand was checked by ^{1}H NMR.

Formation and Purification of Duplex from Single Strands. The extinction coefficients for each oligomer were calculated using the molar absorptivity of each nucleotide, taking into account nearest neighbor interactions (Cantor & Tinoco, 1965; Cantor et al., 1970). The per nucleotide extinction coefficient of the GT strand is 9091 M^{-1} cm^{-1}; the extinction coefficient for the AC strand is 9026 M^{-1} cm^{-1}. Concentrations were determined by absorbance spectroscopy as described (McMurray et al., 1991). For all experiments, duplexes were made by equimolar mixing of the single strands. The oligomer duplexes were separated from any excess single strands by electrophoresis on 12% polyacrylamide gels. The band corresponding to the duplex form was identified by UV shadowing to avoid exposure to any type of dye molecule that could affect the experimental results. Oligomer duplexes were cut out of the gel, electroeluted, and further purified on a Sephadex G-25 column buffered in 10 mM Hepes and 0.1 mM EDTA (pH 7.0).

5′-^{32}P-End-Labeling. Oligonucleotides were labeled at their 5′-termini according to standard methods (Tabor & Struhl, 1993). Free label was removed from the labeled oligonucleotides by ammonium acetate precipitation (Chory & Baldwin, 1993).

Gel Electrophoresis. Nondenaturing polyacrylamide gels were made according to standard methods described by Chory and Baldwin (1993), with slight modifications in the sample and running buffers. Oligonucleotides were separated on 30 cm × 14 cm × 0.7 mm nondenaturing 12–15% polyacrylamide gels buffered in 10 mM Pipes and 0.1 mM EDTA under the appropriate pH conditions, ranging from pH 5.5 to 7.0. Samples were diluted into the appropriate buffer containing 10% sucrose. Due to the possibility that the marker dye could differentially stabilize either the duplex or hairpin forms of the oligonucleotides, no dye was added to the sample before loading on the gel. Sample buffers containing bromophenol blue and xylene cyanol were loaded into separate wells to monitor mobility. Each gel was run under a constant current of 20–25 mA until the desired separation was achieved. The running buffer in all cases was identical to the buffer in the gel: 10 mM Pipes and 0.1 mM EDTA (pH 7.0). After separation, each gel was dried and exposed to autoradiographic film.

Thermal Melting Analysis and Calorimetry. Hairpin or duplex samples were diluted to 1.0×10^{-7} to 7.2×10^{-6} M strand in Pipes-00 at pH 7.0 in a final volume of 1 mL. Thermal melting analysis was performed as described in the accompanying paper (Gacy & McMurray, 1994).

The T_m data in the concentration range between 9.6×10^{-8} and 7.2×10^{-6} M strand were obtained using absorbance measurements. To obtain higher concentration points (1.2×10^{-4} M strand), we utilized the results of calorimetry studies that yield a precise T_m value at high concentrations and the results of ^{1}H NMR melting analyses (9.6×10^{4} M strand). Differential scanning calorimetric experiments were performed and analyzed as described previously (Chou et al., 1987; Sturtevant, 1987; Privalov & Potekhin, 1986; Biltonen & Freire, 1978; Lysko et al., 1981) with an MC-2 scanning calorimeter and associated computer programs purchased from Microcal, Inc. A DSC curve determined with Pipes buffer in both sample and reference cells was used as the instrumental baseline and normally was used to calculate the apparent ΔC_p values by the appropriate subtraction from data determined at each temperature when a Pipes solution containing the oligonucleotides was in the sample cell.

Previous studies on biological systems have used the difference between the heat capacity in a reference state (all bases hydrogen bonded for the systems in this study) and the observed heat capacity. This difference can be shown (Privalov & Potekhin, 1986) to be virtually identical with the difference calculated directly from the apparent ΔC_p curve. Our best splines connection of the regions before and after the transition was nearly the same as a linear extrapolation. We note, as have others (Sturtevant, 1987; Privalov & Potekhin, 1986; Biltonen & Freire, 1978; Lysko et al., 1981), that construction of the extrapolation between pre- and posttransition baselines is somewhat arbitrary, but we attempted to minimize this by using the linear extrapolation whenever possible. Connection of the heat capacity curve before and after the transition region by linear extrapolation provides estimates for the heat capacities of the reference state at temperatures in the transition region. Subtraction of these values from the corresponding apparent ΔC_p values and division of the result at each temperature by the moles of base in the cell yield the "excess" heat capacities, which we define as ΔC_p (kcal/mol of base). After sample and reference solutions were placed in the cell, the calorimeter was cooled to 5 °C and allowed to sit for at least 30 min before a scan was started at 100 °C/h. High-frequency noise was removed by numerical smoothing with a polynomial fit over short temperature ranges. The sensitivity and accuracy of the instrument were determined periodically by passing current through a heater mounted on the sample cell, and heat capacity measurements were found to have a precision in the 1% range.

^{1}H NMR. One-dimensional ^{1}H NMR experiments were performed on a Bruker AMX-500 spectrometer. For exchangeable proton NMR studies, ~2.5 mM oligomer strands were dissolved in 0.5 mL of phosphate buffer (7.5 mM NaH_2PO_4 and 1 mM EDTA) with $^1H_2O/^2H_2O$ (9:1) in a 5 mm NMR tube. The pH value of the sample was adjusted by directly adding HCl or NaOH stock solution into the NMR tube and measured inside the tube by a microelectrode. Since DNA is stable in its absence, no NaCl was added to the samples to minimize any potential homoduplex formation. One-dimensional NMR spectra for exchangeable protons were obtained using either the 1-1 or the 1-3-3-1 solvent suppression pulse sequence (Hore, 1983), with the carrier frequency set at the water resonance. The sweep width varied from 8000 to 10 000 Hz with a recycle delay of 1 s. Sodium 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$ was used as an internal reference. 1D NOE difference spectra for imino proton assignment were obtained by alternately placing the decoupler on and off imino proton resonances of interest at 0 or 5 °C. The decoupler power was adjusted to give an irradiated signal saturation of approximately 90% in 2 s. A total of 1024 scans was averaged for each experiment. The off-resonance FID was subtracted from the on-resonance FID, and the resultant FID was Fourier transformed to generate the difference spectra.

Two-dimensional ^{1}H NMR cross-relaxation spectroscopy (NOESY) of exchangeable imino protons was performed using the standard three-pulse sequence (Macura & Ernst, 1980), with the third pulse replaced by the 1-3-3-1 water suppression pulse sequence (Boelens et al., 1985). Oligonucleotide samples were purified, lyophilized, and resuspended to obtain a 2.5 mM solution of the GT or AC strand in 90% H_2O/10% D_2O, 7.5 mM sodium phosphate, and 0.1 mM EDTA (pH 7.0) at 0–80 °C. Pure absorption NOESY spectra were obtained by time proportional phase incrementation (Marion & Wüthrich, 1983). To prevent undesirable water magnetization inversion in the phase incrementation cycle, a soft 90 deg pulse was applied to water resonance prior to the first pulse and after the second pulse. The phases of the soft pulses were opposite those of the corresponding hard pulses. A total of 512 free induction decays with 2K data points and 32 scans each (3 s repetition time) was collected with a spectral width of 8000 Hz in both dimensions. Spectra were processed with Gaussian filters in both directions (GB1 = 0.1, LB1 = –3, GB2 = 0.05, LB2 = –3), and a fifth-order polynomial baseline correction was applied to spectra in both dimensions.

^{31}P NMR. Phosphorus NMR spectra were obtained on a JEOL GX 270 at 109.25 MHz under the following experimental conditions: typically 3000–4000 scans, 45° flip angle, 4 s pulse repetition, broad-band proton decoupling, 2000 Hz spectral width, and 4 Hz line broadening applied before Fourier transformation. Samples of 0.5 mL contained in a 5 mm NMR tube were inserted into a 10 mm tube containing phosphate buffer in D_2O and trimethyl phosphate (TMP) as a reference.

RESULTS

Possible Structures of the Strands That Compose the Human Enkephalin Enhancer. To examine the structure of the enkephalin enhancer region, we synthesized two 23-bp oligonucleotides, each of which contains the sequence of one strand of the enkephalin enhancer duplex from residues –106 to –84 upstream of the transcriptional start site. Each 23-bp strand contains CRE-1, CRE-2, and flanking sequences in single-stranded form, shown in Figure 1B. Structures I and IV represent the single-stranded form of either the top (I) or the bottom strand (IV) of the enkephalin duplex. Our analysis was limited to structures formed within this 23-bp region since deletion analysis (Comb et al., 1986) has shown that only this region is necessary for the gene to respond to the receptor occupation and to increases in intracellular cAMP *in vivo*.

Inspection of the DNA sequence in the enhancer region reveals that CRE-1 and CRE-2 are contained within a 23-bp imperfect palindrome. Therefore, in the absence of the hydrogen-bonded partner, each 23-bp oligonucleotide has the potential to form two additional solution structures other than the denatured strands (I and IV). Both strands I and IV can form intramolecular or intermolecular hydrogen bonds to give rise to individual hairpin structures (Figure 1B, II and V) or to homoduplex structures (Figure 1B, III and VI), respectively. A special feature of each hydrogen-bonded structure is that each hairpin or homoduplex forms with mismatched base pairs. Strand I, the top strand of the enkephalin enhancer, forms two GT pairs in the hairpin structure (II) and four GT mismatch pairs and one GG mismatched pair in the homoduplex form (III). Strand IV, partner to strand I in the native enkephalin duplex, forms two AC pairs in the hairpin form (V) and four AC and one CC pair in the homoduplex form (VI). For these reasons, we will refer to I as the GT strand, II as the GT hairpin, III as the GT homoduplex, IV as the AC strand, V as the AC hairpin, and VI as the AC homoduplex in subsequent discussion.

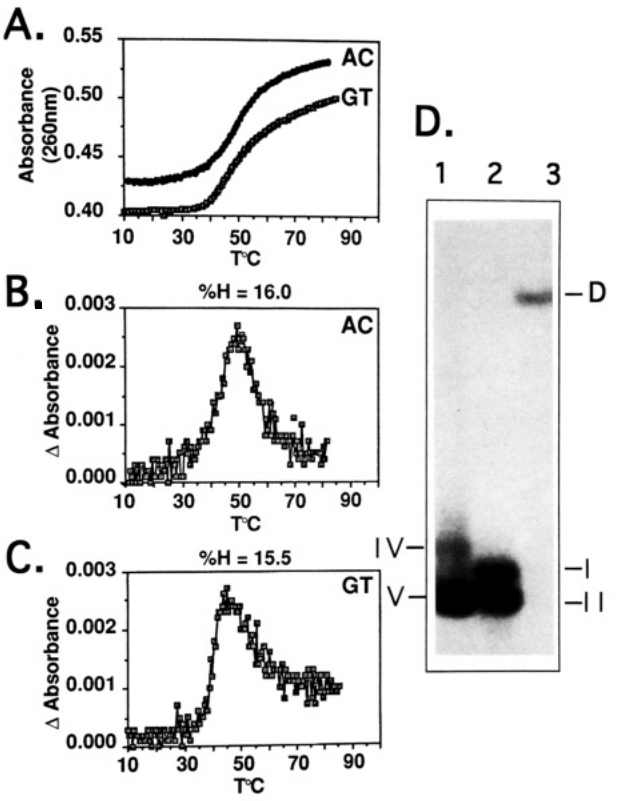

FIGURE 2: Thermal melting and gel electrophoresis analyses of the GT and AC oligonucleotides. (A) Plot of absorbance at 260 nm versus temperature for the GT and AC hairpins in 10 mM Pipes and 0.1 mM EDTA, pH 7.0 (Pipes-00). Hairpin samples were diluted to 2.4×10^{-6} M strand in Pipes-00 to a final volume of 1 mL. (B and C) First derivative of the data in A for the (B) AC hairpin and the (C) GT hairpin. (D) 20% polyacrylamide gel electrophoresis analysis of the ^{32}P-labeled (1) AC oligonucleotide, (2) the GT oligonucleotide, and (3) the duplex formed from the AC + GT oligonucleotides. I, II, IV, and V refer to the forms shown in Figure 1B. D is duplex.

Structures I and IV can be distinguished from hydrogen-bonded structures by a thermal melting analysis. The base unstacking that occurs during hydrogen bond breaking can be observed as a cooperative increase in absorbance (hyperchromicity) with increasing temperature. The results of the thermal melting experiments utilizing GT and AC oligonucleotides are shown in Figure 2. For both oligonucleotides, increases in temperature result in a significant increase in absorbance (Figure 2A), indicative of the large degree of base unstacking found in hydrogen-bonded structures upon melting. Both melting transitions are broad, similar to the profile observed for most hairpin to strand transitions (Cantor & Schimmel, 1980). The first-derivative plots of the GT and AC melting curves are shown in Figure 2B,C. In 10 mM Pipes and 0.1 mM EDTA at pH 7.0 (Pipes-00), the midpoint for thermal melting (T_m) of the GT strand transition is 45 °C, and the T_m for the AC strand transition is 48.5 °C. (The percent hyperchromicity for the melting of each strand is 15–16%.) The results of the thermal melting analysis clearly show that both the GT and AC strands are not found in a linear strand form (I and IV), but exist in a hydrogen-bonded state.

To distinguish between hairpin (II and V) and duplex (III and VI) forms of the GT and AC oligonucleotides, we examined the concentration dependence of the T_m. Because it is an intramolecular event, the T_m of a hairpin to strand melting transition is independent of strand concentration, in contrast to duplex melting. The results of the thermal melting analysis as a function of concentration are shown in Table 1. The melting data reveal that the T_m's of both the AC and GT

Table 1: Concentration Dependence of Oligonucleotide Melting Transition Midpoints[a]

[strand] (M)	GT	AC	duplex
9.6×10^{-8} [b]	44.3	48.5	48.2
7.2×10^{-6} [b]	44.6	48.1	56.7
1.2×10^{-4} [c]	44.7	48.4	63.2
9.6×10^{-4} [d]	45.0[d]	48.0[d]	

[a] Oligonucleotides strands, and duplexes were prepared as described in Materials and Methods. Concentrated stocks were diluted to the appropriate concentration in Pipes-00 buffer. The value for the midpoint of the thermal melting transition is reported as (T_m, °C). M refers to moles of oligonucleotide strand/liter. [b] T_m (°C) measurement was performed by monitoring absorbance versus temperature, as described in Materials and Methods. [c] T_m (°C) measurement was determined from calorimetry experiments (Hopkins, Hamilton, and McMurray, manuscript in preparation). [d] T_m (°C) measurement was determined from the temperature dependence of the area under the imino proton peaks.

melting transitions are independent of concentration over a 1000-fold range. For the thermal melting experiments that monitor either absorbance or enthalpy, the melting profiles for either the GT or AC oligonucleotide at all concentrations were nearly identical in hyperchromicity and shape. In contrast, the T_m of the native enkephalin enhancer duplex that comprises both 23-bp oligonucleotides is distinctly dependent on the concentration of the individual strands. The independence of the T_m for both the AC and GT oligonucleotides conclusively demonstrates an intramolecular equilibrium and confirms that the melting transitions cannot arise from a structure that involves multiple strands.

Gel electrophoresis confirms that little to no duplex is present in either the GT or AC oligonucleotide samples at intermediate concentrations. Figure 2D shows the results of a representative gel electrophoretic separation for both the ^{32}P-labeled GT and ^{32}P-labeled AC strand forms. As a control, we included the gel-purified, ^{32}P-labeled native enkephalin enhancer duplex of 23 bp, shown in Figure 2D, (lane 3). The duplex form is easily distinguished from the strand forms by its greatly reduced mobility. For both the GT (lane 2) and AC strands (lane 1), we observed no band that comigrates with the native enkephalin duplex. The GT and AC oligonucleotides are seen as doublets (Figure 2D), with the fastest migrating forms (II and V) corresponding to the hairpin conformation and the slower migrating forms (I and IV) corresponding to the linear strand conformation (McMurray et al., 1991). Thus, under low concentration and low ionic strength conditions, neither the GT nor the AC strand appears to adopt a homoduplex structure, but exists as either a strand or a hairpin. The faster migrating hairpin bands in Figure 2 (II and V) are present in larger amounts relative to the denatured strand forms (I and IV). Therefore, the results of the gel electrophoresis experiments indicate that the hairpin is the preferred structure for both the GT and AC strands in solution at 30 °C (the gel temperature). Taken together with the thermal melting results, we conclude that, at moderate concentrations and under low ionic strength conditions, each strand primarily adopts a hairpin conformation (form II and V) in solution below 45 °C.

^{1}H NMR Analysis of the Temperature and Concentration Dependencies of the Imino Protons of GT and AC Hairpins. To answer questions concerning the mismatched base pairs, loop regions, and local dynamics of the GT and AC hairpin bases, we monitored the imino protons of the GT and AC oligonucleotides by ^{1}H NMR. Figure 3 shows the 500 MHz ^{1}H NMR spectra of the imino protons of the GT 23-mer oligonucleotide (part b) and the AC 23-mer oligonucleotide (part a) at 2.0 mM strand at 25 °C. We expected a maximum of 9 observable imino resonances for the AC hairpin and a maximum of 14 imino protons for the GT hairpin (Figure 1A). Although early melting protons are no longer observed in the imino proton spectrum at 25 °C (Figure 3), we observed all expected resonances at or below this temperature (see Figure 4). The fact that we observe imino resonances in water at neutral pH provides direct evidence that the imino protons of the two hairpins are in slow exchange.

Although the predominant conformation is the hairpin, each oligonucleotide forms some homoduplex under NMR conditions. Thus, we evaluated the hairpin/homoduplex equilibrium to determine the conditions that minimize homoduplex formation. At 15 °C, dilution to 0.2 mM induced a distinct sharpening of the resonances for both the GT and AC oligonucleotides (not shown). This observation suggests that dilution reduces exchange between more than one conformation. Dilution below 0.2 mM resulted in no further change in the imino spectrum, indicating that, under our buffer conditions, the equilibrium is shifted largely toward the hairpin conformation.

To monitor the local dynamics of individual hairpin bases, we measured the temperature dependence of the imino protons by the ^{1}H NMR spectrum. The temperature dependence of the imino protons of the GT (part A) and AC (part B) oligonucleotides is shown in Figure 4. As expected, we found that, at 0.2 mM, the T_m (45 °C) of the GT oligonucleotide was identical to the T_m that was measured under lower concentration conditions (Table 1). The temperature dependence of the imino ^{1}H NMR spectra in water shows that the two T resonances (T3 and T15) broaden as the temperature is increased before most of the Watson–Crick G imino resonances. These data indicate that the AT pairs in the hairpin are exchanging more rapidly with solvent than the GC pairs, probably due to melting. We observed premature line broadening (at low temperature) for resonances G1, G11, and G12 (Figure 4A), which correspond to the stem terminal base pair and the loop residues (see assignments below). These data indicate that the protons of loop and terminal bases exchange rapidly with solvent well before the T_m is reached. Rapid exchange of G1, G11, and G12 resonances is consistent with their positions in the hairpin structure, end-fraying at the stem terminus, and increased dynamics for unpaired bases within the loop. Above 25 °C, all remaining resonances of the GT oligonucleotide broaden and lose intensity only as the T_m is approached. Even under concentration and ionic strength conditions under which the hairpin predominates, imino protons that correspond to peaks T15 and G10 display two broad lines. These data suggest that T15 and G10 undergo exchange between two hairpin conformations (Figure 4A). Protons T15 and G10 correspond to base pairs that form an interface between the loop and one of the GT mismatched pairs. Thus, the top of the hairpin stem can exist in two different states.

Similar to the GT oligonucleotide, the AC oligonucleotide melts as a single form; all resonances appear to lose intensity only as the T_m (48 °C) is approached (Figure 4B). The T_m °C is identical to that found at lower concentrations (Table 1).

Resonance Assignment of Imino Protons in the Stem of the GT and AC Hairpins. The assignments of the 14 peaks that correspond to the imino protons of the GT hairpin were confirmed using 1D and 2D nuclear Overhauser effects (NOE) (Table 2). The 2D NOE (NOESY) spectrum for the GT oligonucleotide (in 10 mM phosphate, pH 7.0) at 15 °C is shown in Figure 5A. We were able to completely assign the imino protons in the stem of the GT hairpin. The strategy for assignment began with the assumption that the collection

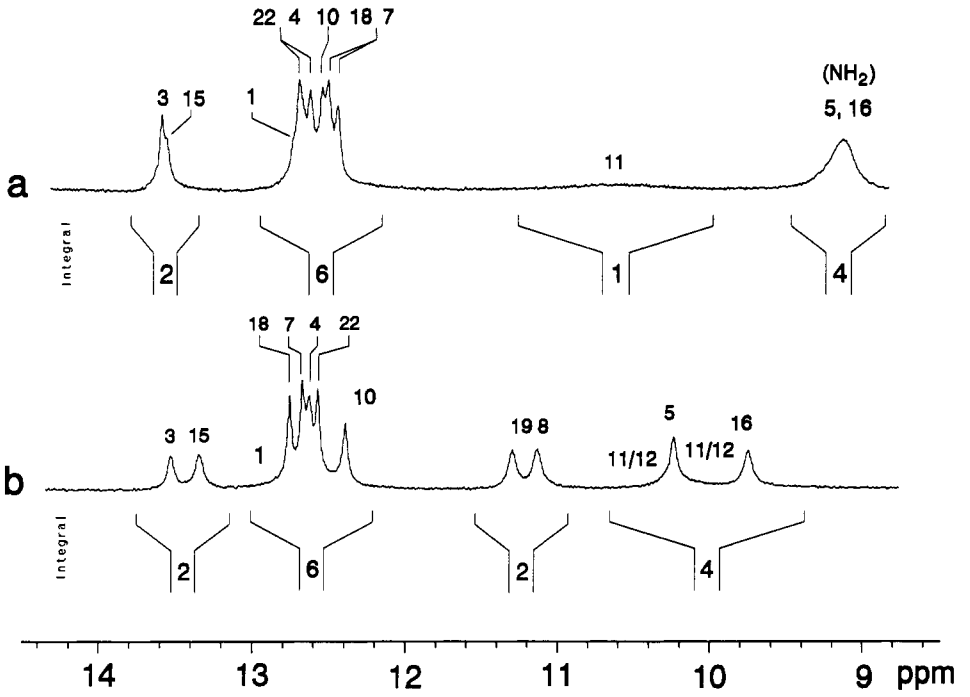

FIGURE 3: 500 MHz ^{1}H NMR spectrum of the exchangeable protons of the GT and AC oligonucleotides. Imino proton spectra for the AC oligonucleotide (a) and the GT oligonucleotide (b) at 25 °C. The concentration of each oligonucleotide is 2.0×10^{-3} M strand. Oligomer strands were dissolved in 0.5 mL of phosphate buffer (7.5 mM NaH_2PO_4 and 1 mM EDTA) with $^1H_2O/^2H_2O$ (1:9).

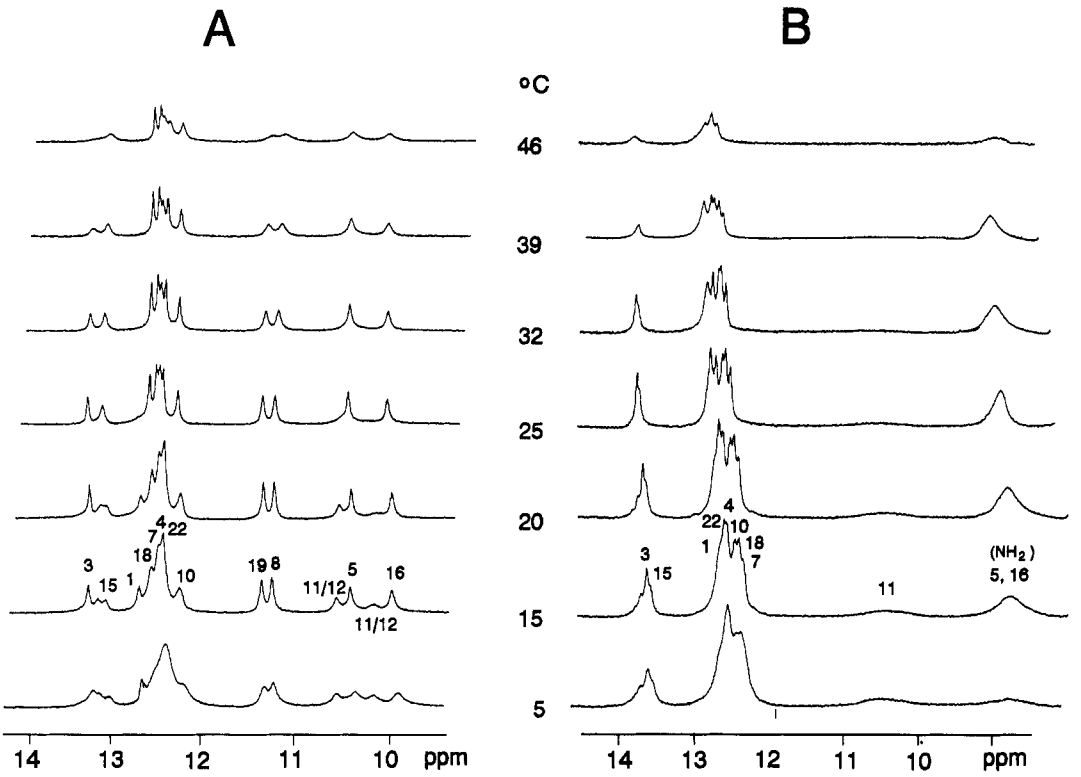

FIGURE 4: Temperature dependence of the 500 MHz ^{1}H NMR spectrum for imino protons of the GT and AC oligonucleotides. Imino proton spectra for the GT oligonucleotide (A) and the AC oligonucleotide (B) in 7.5 mM NaH_2PO_4 and 1 mM EDTA at a concentration of 2.0×10^{-4} M strand. Oligomer strands were dissolved in 0.5 mL of phosphate buffer with $^1H_2O/^2H_2O$ (1:9).

of resonances near 13.5 ppm corresponds to the Watson–Crick T imino protons, and the cluster of resonances near 13.0 ppm is assigned to the Watson–Crick G imino protons, as previously observed [Wüthrich (1986) and references therein]. Since there are only two Watson–Crick AT pairs that would form in the GT hairpin (Figure 1B), we expect that T3 should display NOE connectivity only to G4 and G22 (within the Watson–Crick G cluster) if structure II is correct. The NOESY connectivities (Figure 5A) confirm that the most downfield resonances in the GT spectrum correspond to T3 and T15, respectively. T3 displays the expected crosspeaks to G4 and G22 (that overlap at 15 °C), and T15 displays crosspeaks to T8, G16, and G10. As expected, T15 forms crosspeaks to T8 and G16 that fall in the upfield region of the spectrum, corresponding to mispaired or unpaired imino protons, and to A9 amino resonances (Wüthrich, 1986). At 15 °C, T15 of the hairpin exists in two slowly exchanging conformations. We know that these are two hairpin conformations of T15 since the relative areas of both T15 resonances are unaltered by dilution under identical buffer and tem-

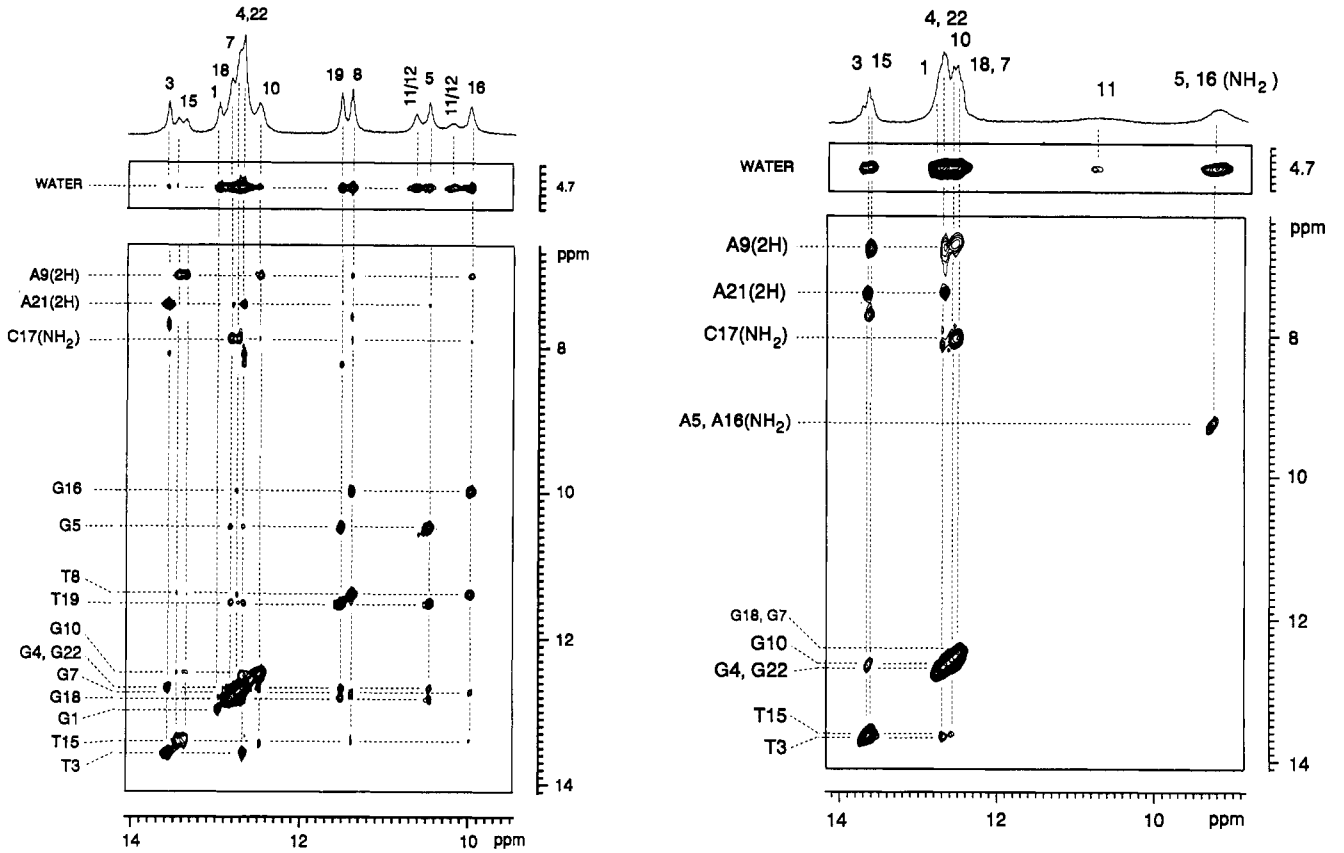

FIGURE 5: (A, left) 500 MHz NOESY spectrum of the exchangeable protons of the GT oligonucleotide at 15 °C. 2D NOESY (mixing time, 200 ms) of the GT hairpin in 7.5 mM phosphate buffer, pH 7.0. Oligomer strands were dissolved in 0.5 mL of phosphate buffer (7.5 mM NaH_2PO_4 and 1 mM EDTA) with $^1H_2O/^2H_2O$ (1:9). Water indicates the region of the ^1H NMR spectrum for which water/imino interactions are observable. (B, right) 500 MHz NOESY spectrum of the exchangeable protons of the AC oligonucleotide at 15 °C. Same as part A for the AC hairpin.

perature conditions. The Watson–Crick G imino protons are assigned as G1, G18, G7, G22, G4, and G10 from downfield to upfield. At 15 °C (Figure 5A), G4 and G22 overlap.

In the upfield imino region of the GT spectrum, we observed six peaks between 9 and 12 ppm. This region of the spectrum corresponds to protons that are not hydrogen bonded or are engaged in nonstandard base pairing (Wüthrich, 1986), verifying the presence of mismatched or unpaired imino protons. The NOESY crosspeaks conclusively identify the two GT mismatched pairs. For example, NOESY crosspeaks at 10.2 and 11.3 ppm correspond to the T8/G16 pairing. The NOESY spectrum of the imino protons confirms that the predominant form of the GT oligonucleotide is a hairpin. Under the conditions of the NOESY experiment, some homoduplex exists in the sample, as evidenced by a weak G10/G11 (12.45/12.7 ppm) crosspeak that is visible in the spectrum only at low contour levels (Figure 5A). Consistent with their assignments, loop residues (G11 and G12) display increased local dynamics relative to stem protons, made evident by their line broadening at low temperature (Figure 4). Thus, NMR and melting analyses together confirm that the GT oligonucleotide forms structure II, a hairpin that contains the 10-bp stem and a 3-bp loop (Figure 1A).

Figure 5B shows the NOESY spectrum for the AC hairpin oligonucleotide (in 10 mM phosphate, pH 7.0) at 20 °C. The NOESY crosspeaks allowed complete assignment of imino protons within the AC hairpin stem. We observed both T3 and T15 near 13.6 ppm, as expected. Through proximity to the T residues, we were able to assign all resonances within the Watson–Crick G cluster (Figure 5B, Table 2). The sole upfield imino resonance at 10.6 ppm was assigned as G11 of the hairpin. We assign the 10.6 ppm peak as G11 and not as a protonated form of adenine, since imino protons within a protonated AC pair exchange with solvent too rapidly to be observed by NMR (Patel et al., 1984b; Hunter et al., 1986; Gao & Patel, 1987; Kalnik et al., 1988). The position of G11 in the AC hairpin is similar to that observed for G11 of the GT hairpin. Consistent with the presence of AC mispaired bases, we observe a strong peak near 9.10 ppm that has an integrated area corresponding to four protons. On the basis of previous assignments of AC mispaired bases in duplexes (Patel et al., 1984b; Hunter et al., 1986; Gao & Patel, 1987;

Table 2: Chemical Shifts of the Exchangeable Imino Protons of the GT and AC Hairpins at 15 °C[a]

base	GT (δ)	AC (δ)
1	12.97	12.80
3	13.57	13.66
4	12.68	12.71
5	10.47	
7	12.73	12.46
8	11.39	
10	12.49	12.57
11	(10.63)	10.60
12	(10.20)	
15	13.46	13.60
	13.37	
16	9.98	
18	12.82	12.52
19	11.52	
22	12.68	12.67

[a] The GT sequence is d(5′-GCTGGCGTAGGGCCTGCGTCAGC-3′); the AC sequence is d(5′-GCTGACGCAGGCCCTACGCCAGC-3′). Purified oligomer samples were dissolved in water, lyophilized, and resuspended in 7.5 mM phosphate (pH 7.0), 1.0 mM EDTA, and (85:15) H_2O/D_2O. Measurements were performed at 15 °C. The δ refers to chemical shift. Numbers in parentheses are for loop protons (see text).

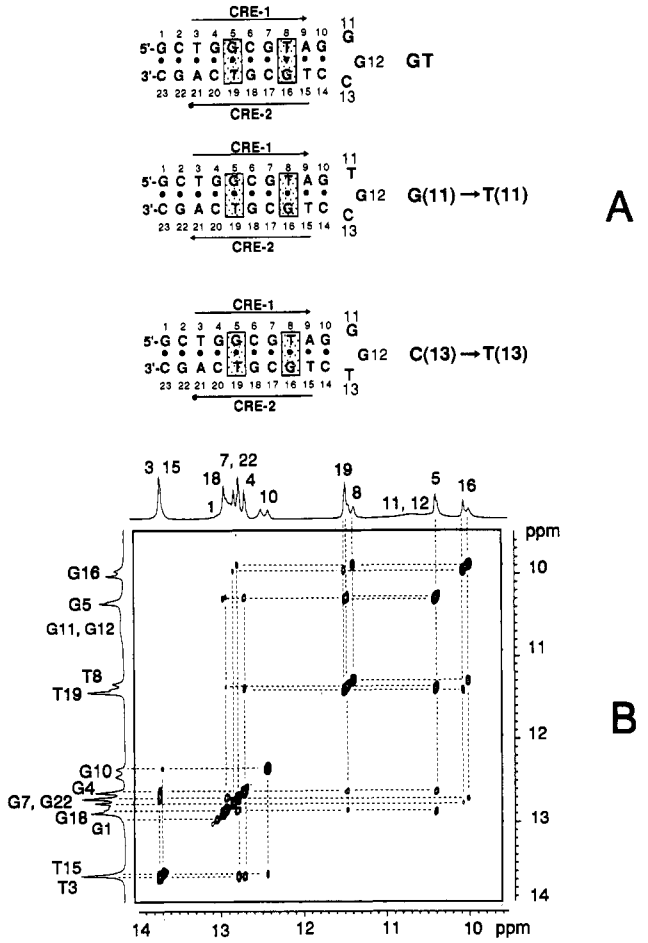

FIGURE 6: (A) Summary of structures for substituted GT hairpins. The substitutions are located within the loops of the hairpins at position 11 (G(11)–T(11)) or at position 13 (C(13)–T(13)). (B) 500 MHz ¹H NMR NOESY spectrum of the G(11) to T(11) loop substitution oligonucleotide at 15 °C, 2.0×10^{-3} M strand, and mixing time 200 ms. Oligomer strands were dissolved in 0.5 mL of phosphate buffer with $^1H_2O/^2H_2O$ (1:9).

Kalnik et al., 1988), we assign this peak to the amino protons of adenine that arise from its association with the N3 of cytosine.

Loop Resonances. To determine the local dynamics and orientation of the loop and to aid in the assignment of loop resonances, we synthesized two derivatives of the GT hairpin that contained loop substitutions, shown in Figure 6A. G(11)–T(11) contains a G to T substitution at position 11; C(13)–T(13) contains a C to T substitution at position 13. We expected to observe a maximum of 14 imino protons for the GT and the G(11)–T(11) oligonucleotide and a maximum of 15 imino protons for the C(13)–T(13) oligonucleotide.

To check the conformation of the oligonucleotide derivatives, we examined the concentration and temperature dependencies of each oligonucleotide. Similar to the GT oligonucleotide, we observed that 10-fold dilution of T(11) resulted in a single conformation of the imino protons that did not change upon further dilution. Thus, at high concentration and low temperature, the G(11)–T(11) oligonucleotide also forms some homoduplex similar to the GT oligonucleotide. The NOESY spectrum for G(11)–T(11) at 15 °C is shown in Figure 6B. At low temperature and high concentration, the NOESY spectrum for G(11)–T(11) is a mixture of duplex and hairpin, and both conformations are observed as crosspeak doublets. For example, G16 occurs as two peaks, both of which display crosspeaks to T8 and G7 (Figure 6B). Since a mixture of duplex and hairpins is observed in the NOESY spectrum, we

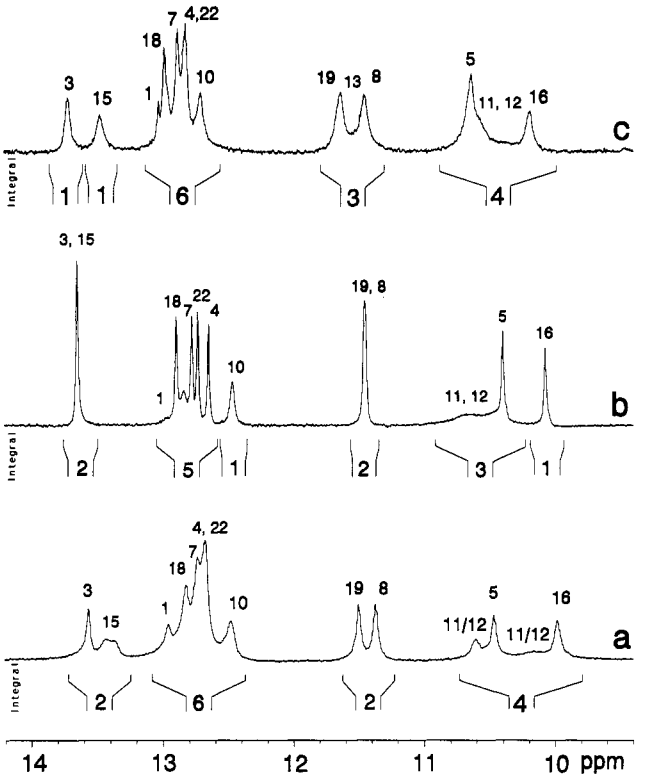

FIGURE 7: 500 MHz ¹H NMR analysis for G(11)–T(11) and the C(13)–T(13) loop substitution oligonucleotides. 1D spectrum for the imino protons of (a) the GT hairpin, (b) the G(11)–T(11) hairpin, and (c) the C(13)–T(13) hairpin at 15 °C in 7.5 mM NaH_2PO_4 and 1 mM EDTA. The samples were dissolved in 0.5 mL of phosphate buffer with $^1H_2O/^2H_2O$ (1:9) at a concentration of 2.0×10^{-4} M strand. Numbers refer to the bases in the hairpins.

compared the 1D spectra at low and high concentrations to determine which peak was the hairpin. Comparison of the 1D spectrum of the T(11) derivative at high and low concentrations (compare Figure 6B and Figure 7b) indicates that the hairpin conformation is located downfield of the same proton in the duplex conformation. Despite the presence of the doublets, the NOESY spectrum confirms that the relative position of each stem proton is similar to that of the GT stem (Figure 6B). However, a G11 to T11 change in the loop alters the environment of several stem protons, as observed by the differences in chemical shifts of individual stem resonances (see below). Thus, alterations in the loop have pronounced effects within the stem of the hairpin (Figure 7).

Since dilution to 0.2 mM results in a single hairpin conformation, and stem resonances were assigned by the 2D spectrum, we examined the 1D spectrum of each oligonucleotide under low concentration conditions (0.2 mM) at 20 °C to examine the hairpin loop dynamics. The 1D spectra at 0.2 mM for the hairpin conformations of the GT, G(11)–T(11), and C(13)–T(13) oligonucleotides are shown in Figure 7. As expected, the integrated intensities of the entire imino region are consistent with the maximum number of expected protons for each oligonucleotide (Figure 7). Loop resonances were assigned as upfield peaks that did not correspond to stem mismatched pairs. From the integrated area in the upfield region, we observed a total of two, two, and three such protons, respectively, for the GT, T(11), and T(13) oligonucleotides. (For each oligonucleotide, four of the upfield protons are unequivocally assigned to the two GT mismatched pairs in the stem; thus, the remaining resonances must arise from loop residues.) For the GT hairpin, two protons at a chemical shift of 10.65 ppm are observed, in addition to the stem GT pairs that are assigned to G(11) and G(12). Similarly, G(11)–

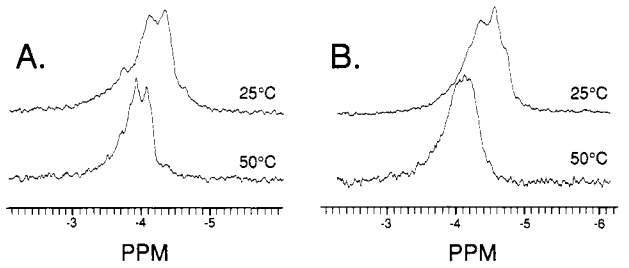

FIGURE 8: Conformation of the GT and AC hairpins. (A) ^{31}P NMR spectra for the GT hairpin at 20 and 50 °C and (B) ^{31}P NMR spectra of the AC hairpin at 20 and 50 °C in phosphate buffer at pH 7.0.

T(11) and C(13)–T(13) contain two broad resonances between 10.4 and 10.8 ppm that are assigned to a loop structure. For C(13)–T(13), a T for C substitution at the 3′ end of the loop gives rise to an additional resonance that is shifted downfield relative to the other loop resonances, overlapping with the stem T imino protons at 11.3 ppm. The 1D spectrum reveals that loop resonances corresponding to G or T imino protons are broad lines that cannot be distinguished by their chemical shift. For each hairpin, the 2D spectrum reveals the absence of diagonal peaks for loop resonances. The loss of the diagonal peaks signifies fast relaxation of these protons, so that they are lost during the NOE mixing time. Chemical exchange with water is responsible for their relaxation. We observe crosspeaks to water in the 2D NOE spectrum, even though diagonal peaks are lost in the noise (see Figure 5B for the GT hairpin). Thus, loop resonances are exchanging with water rapidly relative to stem protons. The water crosspeaks in the 2D spectrum confirm that the upfield, nonstem imino protons are hairpin loop imino protons pointing out of the loop into solvent. However, the lack of NOESY correlations prevents assignment of individual loop resonances at this stage.

Conformation of the Hairpins. To examine whether mismatching of base pairs induced alterations, kinking, or unusual conformational perturbations of the phosphodiester backbone within the enhancer, we recorded the ^{31}P NMR spectra of both the GT and AC hairpins. Figure 8 shows the ^{31}P spectrum for GT (part A) and AC (part B) in 10 mM phosphate (pH 7.0) at 25 and 50 °C. Although variations in the torsional angles of the phosphodiester backbone of DNA exist at DNA mispairings (Roongta et al., 1990), all 23 phosphorus resonances lie between chemical shifts −3.8 and −4.6 ppm relative to trimethyl phosphate (TMP) as an internal standard. At 50 °C, the envelope of resonances shifts downfield, as expected, roughly 0.3 ppm. Thus, the mismatched base pairs do not confer any large alteration in the conformation of the phosphodiester backbone of the DNA, and all phosphorus resonances lie within the limits of the B-form conformation (Dickerson, 1983; Dickerson & Drew, 1981; Saenger, 1988; Roongta et al., 1990).

Circular dichroism studies confirm that the conformation of all strands and the duplex is the B-form (not shown). The CD spectrum between 310 and 240 nm is typical of B-form nucleic acids, displaying a local maximum near 270 nm, a crossover point near 260 nm, and a local minimum near 250 nm (Bush, 1974; Cantor & Schimmel, 1980).

DISCUSSION

Although enhancers are known to modulate the rate of transcription, the exact mechanism by which they operate is unknown. A general feature that has been noted for many enhancers is palindromic character. Palindromic enhancers are over-represented in the regulatory region of genes, at either the 5′ or the 3′ end (Muller & Fitch, 1982; Wells, 1988). In some cases, palindromic enhancers serve as binding sites for protein dimers (Mitchell & Tjian, 1989). However, many enhancers are contained within stretches of palindromic sequences that are much larger than those required for protein binding. We have recently shown that an important transcriptional enhancer binding protein, CREB, binds to the hairpin form of the proenkephalin enhancer. CREB belongs to a superfamily of leucine zipper proteins that bind as dimers to their DNA recognition sites (Brindle & Montminy, 1992; Goodman, 1990; Landshutz et al., 1988; Vinson et al., 1989). Although a potential binding site exists within the duplex form of the enhancer, CREB preferentially binds to a hairpin form of the enhancer duplex (Spiro et al., 1993). The hairpin binding site forms from the top strand of the enhancer and requires a stretch of palindromic sequence that includes both CRE-1 and CRE-2 (see Figure 1A).

CREB binding requires a specific DNA recognition site, CGTCA (Brindle & Montminy, 1992; Ferreri & Montminy, 1994). ^{1}H NMR studies confirm that the GT hairpin contains a binding site for CREB, CTGCGTCA, located within its 10-bp double-stranded stem. The only sequence difference between the CREB binding site found in the hairpin and that found in the duplex is provided by the presence of the two GT mismatched base pairs (see Figure 1) that are also required for strong CREB binding to the hairpin. Since the CREB protein recognizes a specific sequence provided by the GT pairs in the hairpin (Spiro et al., 1993), the orientation of the GT pairs is an important feature of protein recognition. Previous studies of DNA duplexes containing GT mispaired bases suggest that G and T residues are stacked and form two hydrogen bonds similar to an AT pair (Early et al., 1978; Patel et al., 1982, 1984a). In the GT hairpin, the NMR data reveal that the two GT mismatched pairs within the stem binding site are not identical. The G5/T19 pair displays resonances similar to those of previously characterized GT wobble pairs that are hydrogen bonded (Early et al., 1978; Patel et al., 1982, 1984a). The chemical shifts of both the G5 and T19 imino protons within the stem are located downfield of the expected shift for the free imino protons or for free, but stacked imino protons (Early et al., 1978; Haasnoot et al., 1979; Patel et al., 1982, 1984a). However, the G16 resonance of the T8/G16 pair that is critical in protein recognition is located roughly 0.5 ppm upfield of G5. These data suggest that the T8/G16 pair is stacked but may not be paired under the conditions of the NOESY. According to Tinoco rules (Aboula-ela et al., 1985), the presence of GT mismatched pairs has no negative effect on helix-forming ability, and GT pairs appear nearly as stable as AT pairs. ^{31}P NMR and circular dichroism results confirm that, despite the presence of mismatched pairs, both oligonucleotide hairpins form normal B-form helices. Thus, a protein binding site formed from mispaired bases on the GT arm is similar in conformation to a site that contains AT Watson–Crick pairing. Since the GT pairs are required for strong CREB binding to the hairpin (Spiro et al., 1993), and since both GT pairs form a stacked pair, it is likely that CREB interacts with chemical groups that lie in the major groove of the GT pair. Recent studies using templates in which the T is replaced with C, ^{Me}C, or U reveal that the methyl group of the thymine in the GT pair is crucial in stabilizing CREB binding (C. Spiro, unpublished results). Indeed, the crystal structure of a another leucine zipper protein, GCN4, complexed to DNA has revealed that interactions of methyl groups in T residues with alanines are critical for the interaction of GCN4 with its DNA binding site (Ellenberger et al., 1992).

Hairpin loop residues appear to play a role in stabilizing the CREB interactions with the proenkephalin hairpin (Spiro et al., 1993). However, at present, there are several possible models by which the loop may influence CREB binding. The loop may stabilize CREB interactions through direct hydrogen bonding with loop bases. [1]H NMR and thermal melting results confirm that the GT hairpin contains a 3-member loop with G, G, and C residues from 5′ to the 3′ side. The lack of NOESY peaks of the loop residues with neighboring imino residues suggests that these bases may point outward, possibly providing donor–acceptor residues for interaction with CREB. Substitutions of other nucleotides within the loop at positions (G)–96 or (C)–94 reduce CREB's binding affinity (Spiro et al., 1993). The [1]H NMR studies indicate that loop bases are unpaired; thus, several interactions between CREB and loop bases are possible. Since the G and the C cannot be substituted without sacrificing binding affinity, the data suggest that specific moieties on the G and C are important. Substitution of an A in position –96 completely abolishes CREB binding, while the identical T substitution only diminishes binding. Taken together, these data suggest that the CREB interaction may involve the carbonyl group or the imino proton of G.

Alternatively, stacking in the substituted loop may be altered such that the hairpin binding site becomes unstable. Thermal melting studies (M. Gacy, unpublished results) reveal that the –96 G to A substitution does have a destabilizing effect on hairpin structure. Interestingly, substitution of loop bases results in significant changes in the chemical shifts of T15, G10, T19, G22, G4, T8, G5, and G16 within the stem of the hairpin. Thus, point mutations in the loop can alter hairpin stem conformation and, consequently, the conformation of the CREB binding site. The NMR data reveal that there are two conformations of bases G10 and T15 located at the top of the hairpin stem. These bases lie between the T8/G16 mismatched pair and the loop. The [1]H NMR results also indicate that the T8/G16 bases are not constrained by hydrogen bonding. Thus, these bases may have more conformational flexibility relative to other stem base pairs, possibly influencing the two hairpin conformations that we observe for G10 and T15. In previous studies of hairpins, Williamson and Boxer (1989a,b) reported that the phosphate backbone at the top of the hairpin containing a 5-member loop was in intermediate exchange between two conformations. The two conformations were due to an alteration of the phosphate backbone of the top base in the stem (equivalent to the relative position of G10 of the GT hairpin) from a trans-gauche (t,g) to a g,g^- orientation (Williamson & Boxer, 1989a,b; Saenger, 1984). The authors suggest that there are two energetically equivalent stacking arrangements among the bases at the top of the loop, in which three of the five loop residues stacked either on the 5′ or the 3′ side of the stem (Williamson & Boxer, 1989a,b). Although the GT hairpin contains only a 3-member loop, the existence of two conformations at the top of the GT stem may represent an exchange between the stacking of loop bases at either the 5′ or the 3′ side of the stem. Since the loop residues play a role in stabilizing CREB binding, it is possible that the CREB protein may preferentially bind and stabilize one of these conformations.

The NMR, thermal melting, and gel electrophoresis data together confirm that the GT and AC oligonucleotides form B-form hairpins that each contain a 3-member loop and a 10-bp stem. Considering the hairpin structure obtained from NMR analysis, we compared the results of enhancer point mutations on the expression of the proenkephalin gene. Previously, mutations at positions –90, –91, and –92 could not be explained in terms of protein binding to the duplex

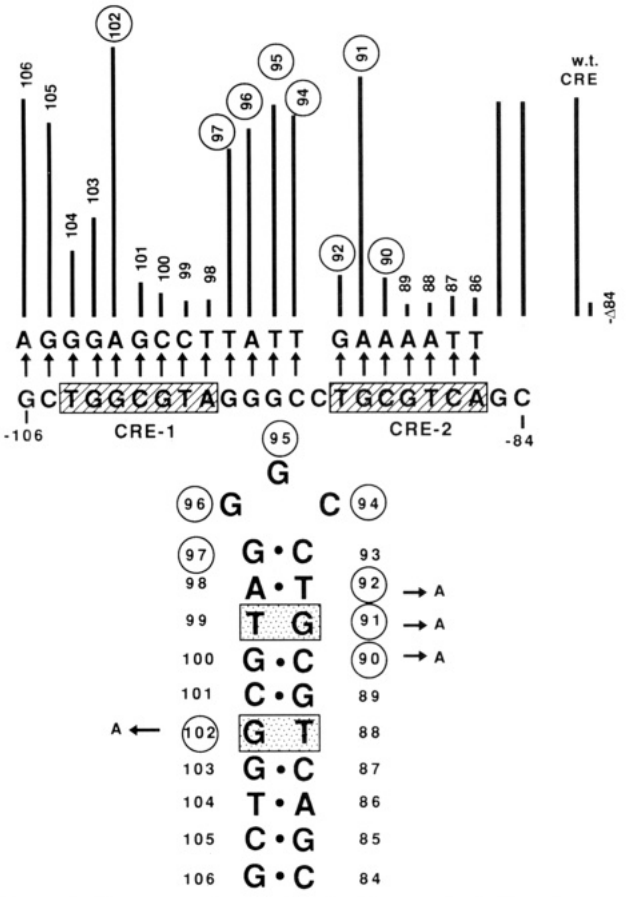

FIGURE 9: Summary of expression for the proenkephalin enhancer. Data were taken from Comb et al. (1988) and Collins-Hicok (unpublished). Bars represent the activity measured in transfected cells from a series of single point mutations within the enhancer. Mutant enhancers were recombined with a bacterial reporter gene to form a series of synthetic recombinant genes. The bacterial chloramphenicol acetyl transferase (CAT) gene was used as the reporter; activity is measured as CAT activity, the percent of [^{14}C]-chloramphenicol that is acetylated by the protein product of the reporter gene.

form of the proenkephalin enhancer (Comb et al., 1988). Figure 9 shows the striking correlation between the structure that we have determined and the *in vivo* expression supported by mutant enkephalin enhancers (Comb et al., 1988; Collins-Hicok, unpublished results). The point mutations that have no effect on protein binding but that increase expression (including –91) are those mutations that potentially stabilize a cruciform structure by converting either one of the non-Watson–Crick pairs in the hairpin stem to an AT or GC pair (Figure 9). Those point mutations (including –90 and –92) that eliminate enhancer function without eliminating protein binding sites are base changes that would create additional mismatched base pairs within the stem and therefore destabilize the hairpin. Thus, all expression from point mutations is consistent with the predicted effects of alteration of a CREB-like binding site and/or alterations of hairpin stability. Although the sequence for the CREB binding site is unaltered by loop mutations, binding studies reveal that the consequence of certain loop mutations is to alter (diminish) the affinity of CREB (Spiro et al., 1993). The [1]H NMR studies reveal that point substitutions within the loop of the hairpin (see Figure 9) outside the consensus binding site clearly affect the conformation of the hairpin. It may be possible that the hairpin binding site comprises an optimum sequence for binding and an optimum conformation for binding. Further structural studies will be necessary to elucidate the conformation of the hairpin arms and to precisely define structure of the CREB

binding site. Both the sequence and the conformation of the hairpin arms from the proenkephalin enhancer have the potential to profoundly influence binding and, consequently, the overall rate of transcription.

ACKNOWLEDGMENT

The authors thank H. Cheng, C. Spiro, M. Gacy, L. Lin, and X. Wu for critical comments and discussion; Ms. Marlené Boyd for her excellent contribution of figure preparation and text proofreading; and the Grainger Analytical NMR Facility and the Kresge Foundation for support.

REFERENCES

Aboula-ela, F., Koh, D., & Tinoco, I., Jr. (1985) *Nucleic Acids Res. 13*, 4811–4824.
Biltonen, R. L., & Freire, E. (1978) *CRC Crit. Rev. Biochem. 5*, 85–124.
Bloomfield, V. A., Tinoco, D. M., & Crothers, I., Jr. (1974) in *Physical Chemistry of Nucleic Acids*, Harper & Row, New York.
Boelens, R., Scheek, R. M., Dijkstra, K., & Kaptein, R. J. (1985) *J. Magn. Reson. 62*, 378–386.
Borden, K. L. B., Jenkins, T., Skelly, J. V., Brown, T., & Lane, A. N. (1992) *Biochemistry 31*, 5411–5422.
Brindle, P. K., & Montminy, M. M. (1992) *Curr. Opin. Genet. Dev. 2*, 199–204.
Bush, C. A. (1974) in *Basic Principles of Nucleic Acid Chemistry* (Ts'o, P. P. P., Ed.) Vol. 2, p 91, Academic Press, New York.
Cantor, C. R., & Tinoco, I., Jr. (1965) *J. Mol. Biol. 13*, 65–77.
Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry: Techniques for the study of biological structure and function*, pp 409–432, Freeman & Co., San Francisco, CA.
Cantor, C. R., Warshaw, M. M., & Shapiro, H. (1970) *Biopolymers 9*, 1059–1077.
Chory, J., & Baldwin, A. (1993) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K., Eds.) p 2, Greene/Wiley Interscience, New York.
Chou, W. Y., Marky, L. A., Zaunczkowski, D., & Breslauer, K. J. (1987) *J. Biomol. Struct. Dyn. 5*, 345–359.
Chu, H. C., Tan, Y., Kobierski, L., Balsam, L. B., & Comb, M. J. (1994) *Mol. Endocrinol. 8*, 59–68.
Comb, M., Birnberg, N. C., Seasholtz, A., Herbert, E., & Goodman, H. M. (1986) *Nature (London) 323*, 353–356.
Comb, M., Mermod, N., Hyman, S. E., Pearlberg, J., Ross, M. E., & Goodman, H. M. (1988) *EMBO J. 7*, 3793–3805.
Dickerson, R. E. (1983) *J. Mol. Biol. 166*, 419–441.
Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–768.
Douglass, J. O., Iadarola, M. J., Hong, J. S., Garrett, J. E., & McMurray, C. T. (1991) *NIDA Res. Monogr. 111*, 133–148.
Early, T. A., Olmstead, J., III, Kearns, D. R., & Lezius, A. G. (1978) *Nucleic Acids Res. 21*, 8269–8279.
Ellenberger, T. E., Brandl, C. J., Struhl, K., & Harrison, S. C. (1992) *Cell 71*, 1223–1237.
Ferreri, K., & Montminy, M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1210–1213.
Gacy, A. M., & McMurray, C. T. (1994) *Biochemistry* (preceding article in this issue).
Gao, X., & Patel, D. J. (1987) *J. Biol. Chem. 262*, 16973–16984.
Goodman, R. H. (1990) *Annu. Rev. Neurosci. 13*, 111–127.
Haasnoot, A. G. C., den Hartog, J. H. J., deRooij, J. F. M., van Boom J. H., & Altona, C. (1979) *Nature (London) 281*, 235–236.
Hore, P. J. (1983) *J. Magn. Reson. 55*, 283–300.
Huggenvik, J. I., Collard, M. W., Stofko, R. E., Seasholtz, A. F., & Uhler, M. D. (1991) *Mol. Endocrinol. 5*, 921–930.
Hunter, W. N., Brown, T., Anand, N. N., & Kennard, O. (1986) *Nature (London) 320*, 552–555.
Kalnik, M., Kouchakdjian, M., Li, F. L., Swann, P. F., & Patel, D. (1988) *Biochemistry 27*, 100–108.
Kobierski, L. A., Chu, H. M., Tan, Y., & Comb, M. J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10222–10226.
Konradi, C., Kobierski, L. A., Nguyen, T. V., Heckers, S., & Hyman, S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7005–7009.
Landschulz, W. H., Johnson, P. F., & McKnight, S. L. (1988) *Science 240*, 1759–1764.
Lysko, K. A., Carlson, R., Snow, J., Taverna, R., & Brandts, J. F. (1981) *Biochemistry 20*, 5570–5576.
Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Mayer, E. A. (1993) *Biol. Signals 2*, 57–76.
McMurray, C. T., Wilson, W. D., & Douglass, J. O. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 666–670.
Mitchell, P. J., & Tjian, R. (1989) *Science 245*, 371–378.
Muller, U. R., & Fitch, W. M. (1982) *Nature 298*, 582–585.
Patel, D. J., Kozlowski, S. A., Marky L. A., Broka, C., Rice, J. A., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry 21*, 428–436.
Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984a) *Biochemistry 23*, 3207–3217.
Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984b) *Biochemistry 23*, 3218–3226.
Privalov, P. L., & Potehkin, T. (1986) *Methods Enzymol. 131*, 4–51.
Roongta, V. A., Jones, C. R., & Gorenstein, D. G. (1990) *Biochemistry 29*, 5245–5258.
Saenger, W. (1988) in *Principles of Nucleic Acid Structure*, p 194, Springer-Verlag, New York.
Sanger, F., & Coulson, A. R. (1975) *J. Mol. Biol. 94*, 441–450.
Sonnenberg, J. L., Rauscher, F. J., Morgan, J. I., & Curran, T. (1989) *Science 246*, 162–1625.
Spiro, C., Richards, J. P., Chandrasekaran, S., Brennan, R. G., & McMurray, C. T. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 4606–4610.
Sturtevant, J. M. (1987) *Annu. Rev. Phys. Chem. 38*, 463–488.
Tabor, S., & Struhl, K. (1993) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K., Eds.) p 2, Greene/Wiley Interscience, New York.
van Nguyen, T. V., Kobierski, L., Comb, M., & Hyman, S. E. (1990) *J. Neurosci. 10*, 2825–2833.
Vinson, C. R., Sigler, P. B., & McKnight, S. L. (1989) *Science 246*, 911–916.
Walton, K. M., & Rehfuss, R. P. (1990) *Mol. Neurobiol. 4*, 197–210.
Wells, R. D. (1988) *J. Biol. Chem. 263*, 1095–1098.
Williamson, J. R., & Boxer, S. G. (1989a) *Biochemistry 28*, 2819–2831.
Williamson, J. R., & Boxer, S. G. (1989b) *Biochemistry 28*, 2831–2836.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, Wiley & Sons, New York.