Protein Science (1993), *2*, 1461-1471. Cambridge University Press. Printed in the USA.

Bacterial expression and characterization of the CREB bZip module: Circular dichroism and 2D ^{1}H-NMR studies

ZULMA I. SANTIAGO-RIVERA,[1] JOHN S. WILLIAMS,[2] DAVID G. GORENSTEIN,[1] AND OURANIA M. ANDRISANI[2]

[1] Department of Chemistry, School of Science, Purdue University, West Lafayette, Indiana 47907
[2] Department of Physiology and Pharmacology, School of Veterinary Medicine, Purdue University, West Lafayette, Indiana 47907

(RECEIVED April 13, 1993; REVISED MANUSCRIPT RECEIVED June 24, 1993)

Abstract

In this paper we describe the expression and purification from bacteria of the recombinant basic leucine zipper (bZip) domain of the cAMP response element binding protein, $CREB_{327}$. The bZip peptide, $CREB_{259-327}$, purified to near homogeneity, maintains the sequence-specific CRE site recognition demonstrated by in vitro competition assays. Alkylation of the three cysteine residues of $CREB_{259-327}$ was employed to prevent aggregation of the peptide due to cysteine oxidation. The K_d of the purified native and modified $CREB_{259-327}$ for the CRE site was determined by gel retardation assays to be on the order of 10^{-7} M. We employed CD spectroscopy to study the folding properties of the native and modified $CREB_{259-327}$. The CD analyses of the native/modified $CREB_{259-327}$ peptide demonstrated a 20% increase in the α-helical content upon binding to the cAMP response-element. Only a 5% increase in the α-helical content of $CREB_{259-327}$ is observed upon binding to the AP-1 site. This observation contrasts with CREB from the GCN4 protein (Weiss, M.A., et al., 1990, *Nature 347*, 575-578). In addition, the two-dimensional (2D) ^{1}H-NMR studies of the bZip CREB peptide further support the distinct features of the CREB protein, in comparison to GCN4. Analysis by CD and 2D NMR of the dimerization domain of CREB suggests that the distinct DNA binding characteristics of CREB reside in the basic portion of the bZip module.

Keywords: basic leucine zipper domain; cAMP response element; cAMP response element binding protein; CD spectroscopy; two-dimensional ^{1}H-NMR

The transcription factor CREB (cAMP response element binding protein) regulates the expression of a number of cellular genes in response to changes in intracellular cAMP levels, by binding to the cAMP response element, CRE (Roesler et al., 1988). In vitro studies have demonstrated that the CREB protein binds to the palindromic CRE site, TGACGTCA (Montminy & Bilezikjian, 1987; Andrisani et al., 1988, 1989). The structural features of the CREB protein mediating its binding to the CRE site include a bipartite structural motif composed of the leucine zipper, which is preceded by the DNA binding domain, a stretch of basic amino acid residues (Dwarki et al., 1990).

Members of the leucine zipper class of proteins include the AP-1 family (Vogt & Bos, 1990); the yeast transactivator, GCN4 (Hope & Struhl, 1986, 1987); the CAAT box binding transcription factor, C/EBP (Landschulz et al., 1988); and the CREB/ATF family of proteins (Hai et al., 1989). The leucine zipper proteins bind to their target DNA as dimers (Kouzarides & Ziff, 1988; Landschulz et al., 1988). The dimerization domain consists of a region of 35 amino acids having leucine residues spaced with a periodicity of seven amino acid residues (Landschulz et al., 1988). Structural studies of synthetic peptides corresponding to the leucine zipper motif of GCN4 demonstrated that this domain assumes a coiled-coil

Reprint requests to: Ourania M. Andrisani, Department of Physiology & Pharmacology, Hansen Life Science Research Building, Box 101, Purdue University, West Lafayette, Indiana 47907.

structure of parallel helices (O'Shea et al., 1989a,b; Oas et al., 1990). The α-helical structure and parallel coiled-coil arrangement of this domain have been corroborated by two-dimensional (2D) ^{1}H-NMR (Oas et al., 1990) and X-ray crystallographic data (O'Shea et al., 1991).

The DNA binding domain of leucine zipper proteins is predominantly a basic region of approximately 30 amino acid residues located upstream of the leucine zipper motif. CD studies of GCN4 (Weiss, 1990; Weiss et al., 1990) and Jun/Fos (Patel et al., 1990) have demonstrated that the basic region is partially α-helical in solution. Binding to the cognate DNA binding site stabilizes its α-helical conformation. Both GCN4 (Weiss, 1990) as well as Jun/Fos (Bush & Sassone-Corsi, 1990) bind to the CRE and AP-1 DNA sequences with similar affinity. Binding of GCN4 to either DNA site stabilizes the α-helical conformation of the DNA binding domain to the same extent (Weiss et al., 1990).

In contrast to GCN4 and Jun/Fos, which bind to the CRE and AP-1 DNA sites with the same affinity, CREB binds specifically to the CRE site (Andrisani et al., 1988; Bush & Sassone-Corsi, 1990). However, the determinants involved in the selective DNA recognition displayed by CREB and the flexibility exhibited by GCN4 and Jun/Fos are not yet understood. We have successfully expressed the $CREB_{259-327}$ peptide in *Escherichia coli* in order to examine the structural basis of this recognition process. We report in this study the expression and characterization of the $CREB_{259-327}$ peptide. The studies include the conformational transitions observed by CD upon DNA binding to specific and nonspecific DNA sequences, as well as ^{1}H-NMR studies and structure of the leucine zipper region in solution.

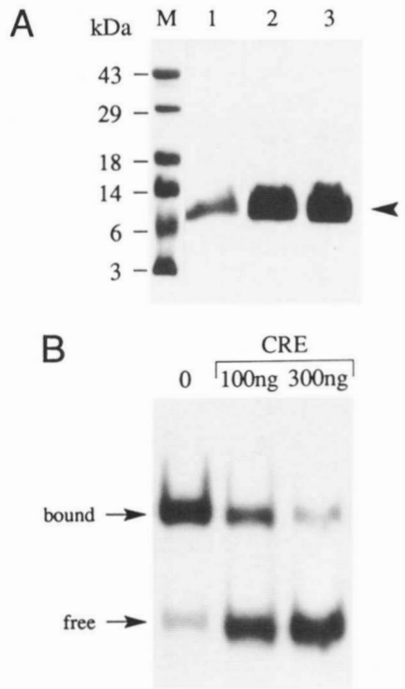

Fig. 1. A: Fifteen percent SDS-PAGE of phenyl superose-eluted fractions. M, molecular weight markers. Lane 1, 5 ng of input sample; lanes 2 and 3, 10 ng of phenyl superose-eluted fractions. Arrow points to the $CREB_{259-327}$ peptide. **B:** Gel retardation/competition assay. Binding/competition reactions were carried out with 10 ng pure $CREB_{259-327}$, in the presence of radiolabeled CRE probe and 100 or 300 ng unlabeled CRE DNA. The reactions were analyzed on 5% native acrylamide gels. Electrophoresis was at 200 V for 1.5 h. Arrows point to the bound and free DNA.

Results

Bacterial expression and purification of $CREB_{259-327}$

A number of bacterial expression vectors are available for the affinity purification of recombinant proteins. However, the limitation of these systems is that the protein or peptide of interest is a fusion molecule containing a number of non-native amino acid residues (Smith & Johnson, 1988; Abate et al., 1990). This limits the use of the recombinant protein for higher resolution studies such as 2D NMR, where the upper molecular weight limit is approximately 10,000 Da. The pT7-7 system offers the advantage that the native protein/peptide can be obtained. The purification procedure that we have developed specifically for the recombinant $CREB_{259-327}$ peptide, employing the pT7-7 vector system (described under Materials and methods), yields approximately 2 mg of pure peptide per liter of bacterial culture.

Figure 1A shows analysis of the final phenyl superose fast performance liquid chromatography (FPLC) eluate on 15% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and demonstrates a single band migrating with the expected molecular weight of 8,450. The activity of the expressed $CREB_{259-327}$ peptide was examined by gel retardation assays (Andrisani et al., 1988). The somatostatin CRE oligonucleotide was employed as the radiolabeled probe. Figure 1B shows the sequence-specific binding of the peptide to the CRE site, as assessed by in vitro competition assays in the presence of increasing amounts of unlabeled CRE DNA.

The sample purity was further verified by reverse-phase high performance liquid chromatography (HPLC) (Fig. 2A) where the single symmetrical peak eluting at 52% acetonitrile corroborates >90% purity. Furthermore, the identity of the purified peptide was verified by amino acid analysis (Table 1) and peptide sequencing. The amino acid content of the recombinant $CREB_{259-327}$ compares with the expected values, except for the absence of one of the three expected methionine residues. Peptide sequencing of the same sample verified the absence of the initiator methionine. The remaining amino acid sequence agrees with the expected amino acid sequence of $CREB_{259-327}$. Thus, the described purification protocol

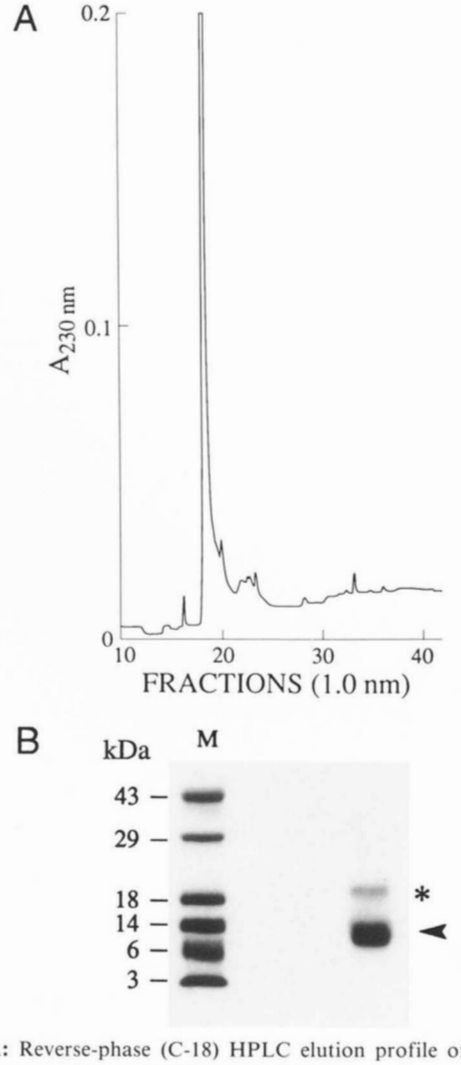

Fig. 2. A: Reverse-phase (C-18) HPLC elution profile of purified $CREB_{259-327}$ peptide, in 0–100% H_2O/CH_3CN gradient, containing 0.1% TFA. **B:** Fifteen percent SDS-PAGE of $CREB_{259-327}$ peptide following lyophilization of the HPLC-eluted peak. Arrow points to monomeric $CREB_{259-327}$ peptide. Asterisk indicates the covalent $CREB_{259-327}$ dimer.

yields $CREB_{259-327}$ peptide >90% pure, which maintains the sequence-specific recognition of the cognate CRE site.

Carboxymethylation of cysteine residues

During the initial characterization of the purified peptide, we repeatedly observed that it precipitated out of solution. Furthermore, denaturing SDS-PAGE analysis of $CREB_{259-327}$ demonstrated the presence of a second band in addition to that of the peptide monomer (Fig. 2B), which migrates with an apparent molecular weight of approximately 17,000 Da, suggestive of the formation of covalent dimers. $CREB_{259-327}$ contains three cysteine residues, one in the DNA binding domain and two in the leucine zipper. Cysteine oxidation could result in covalent dimers formed within as well as between dimer subunits, resulting in the formation of high molecular weight aggregates that precipitate out of solution. We thus undertook the modification of the thiol groups of the cysteine residues in the purified $CREB_{259-327}$ peptide utilizing iodoacetamide. All three residues in the peptide were carboxymethylated as determined by the absence of free thiols with the Ellman's (5,5′-dithio-bis(2-nitrobenzoic) acid) test. Following modification of the cysteines, covalent disulfide dimers are not observed by SDS-PAGE.

Table 1. *$\Delta CREB_{259-327}$ amino acid composition*

Amino acid	Expected	Observed
Alanine	8	7.87
Arginine	9	8.91
Asp/Asn	6	6.14
Cysteine	3	2.93
Glu/Gln	12	12.12
Histidine	1	0.99
Isoleucine	2	1.99
Leucine	8	8.43
Lysine	10	10.40
Methionine	2	1.01
Proline	2	1.99
Serine	1	1.06
Threonine	2	1.99
Tyrosine	2	2.08
Valine	4	4.10

To quantitatively assess the effect of the modification of the cysteine residues on the binding affinity of $CREB_{259-327}$ for the CRE site, we carried out the determination of the apparent dissociation constant, K_d, as described in Williams et al. (1993). The binding reactions were carried out in the presence of increasing amounts of purified $CREB_{259-327}$ peptide. The CRE probe is present at concentrations less than 10 pM. Following gel electrophoresis and autoradiography (Fig. 3A), the bands were excised, and the unbound and bound DNA fractions were quantitated by scintillation counting. The free DNA fraction was plotted against the protein concentration in order to determine the protein concentration at which half the DNA is bound. This concentration is a good approximation of K_d under conditions where concentration of total DNA is much less than the total concentration of the protein (Carey, 1989). We determined that the K_d of the modified and native $CREB_{259-327}$ peptide is 8×10^{-7} M and 5×10^{-7} M, respectively (Fig. 3B). Furthermore, the half maximal point determined by measuring the disappearance of the unbound DNA is identical to the value obtained from the bound DNA, suggesting that complex dissociation did not occur during electrophoresis.

The fact that the cysteine modification did not alter the DNA binding affinity of the expressed CREB peptide suggests that the cysteine residue in the basic region is not one of the main determinants involved in the sequence-specific

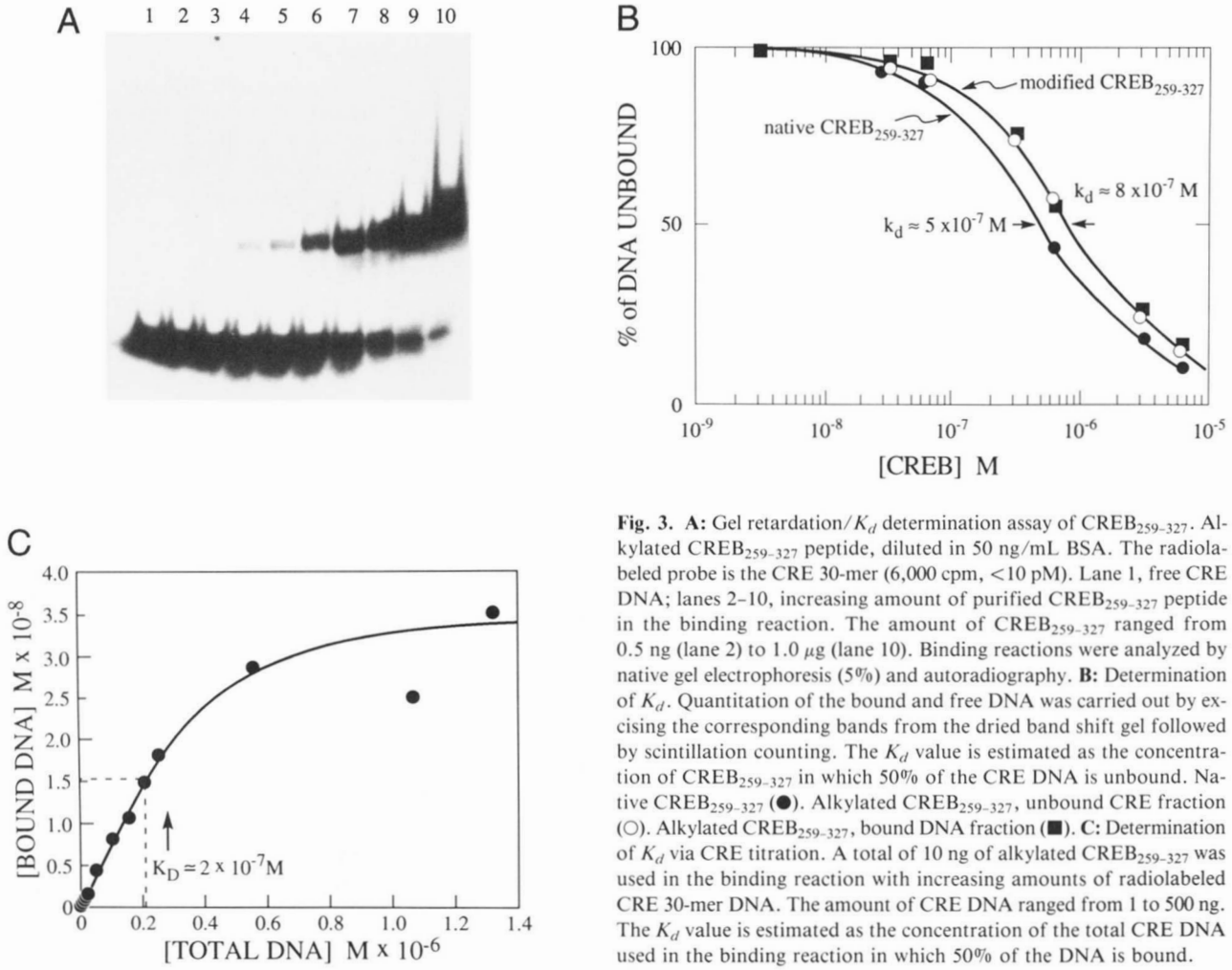

Fig. 3. A: Gel retardation/K_d determination assay of $CREB_{259-327}$. Alkylated $CREB_{259-327}$ peptide, diluted in 50 ng/mL BSA. The radiolabeled probe is the CRE 30-mer (6,000 cpm, <10 pM). Lane 1, free CRE DNA; lanes 2–10, increasing amount of purified $CREB_{259-327}$ peptide in the binding reaction. The amount of $CREB_{259-327}$ ranged from 0.5 ng (lane 2) to 1.0 μg (lane 10). Binding reactions were analyzed by native gel electrophoresis (5%) and autoradiography. **B:** Determination of K_d. Quantitation of the bound and free DNA was carried out by excising the corresponding bands from the dried band shift gel followed by scintillation counting. The K_d value is estimated as the concentration of $CREB_{259-327}$ in which 50% of the CRE DNA is unbound. Native $CREB_{259-327}$ (●). Alkylated $CREB_{259-327}$, unbound CRE fraction (○). Alkylated $CREB_{259-327}$, bound DNA fraction (■). **C:** Determination of K_d via CRE titration. A total of 10 ng of alkylated $CREB_{259-327}$ was used in the binding reaction with increasing amounts of radiolabeled CRE 30-mer DNA. The amount of CRE DNA ranged from 1 to 500 ng. The K_d value is estimated as the concentration of the total CRE DNA used in the binding reaction in which 50% of the DNA is bound.

recognition of DNA. The data support the results obtained previously in our laboratory (Andrisani & Dixon, 1991). We showed, via mutagenesis studies of the DNA binding domain of CREB, that the $Cys_{286} \rightarrow Ala$ substitution reduces the DNA binding potential of CREB by 50% (Andrisani & Dixon, 1991). This is also observed in the difference between the K_d value of the native and alkylated basic leucine zipper (bZip) peptides, shown in Figure 3B.

To assess the active protein fraction in the alkylated bZip sample and thus further confirm the K_d values shown in Figure 3B, we measured the K_d of the alkylated peptide via a DNA titration experiment. The DNA binding assays were carried out using a defined amount of alkylated $CREB_{259-327}$ peptide (10 ng) in the presence of increasing amounts of CRE DNA. The binding reactions were analyzed by gel electrophoresis and autoradiography. The bound and free DNA were quantitated by scintillation counting. The results are shown in Figure 3C. We determined the K_d, via the DNA titration method, to be 2×10^{-7} M, fourfold different from the K_d shown in Figure 3B. Thus, it is concluded that the bZip of $CREB_{259-327}$ is on the order of 10^{-7} M. Furthermore, the analyses shown in Figure 3C support that the preparation of the alkylated bZip peptide is nearly 100% active.

CD studies

The K_d determination studies demonstrated that the modification of the cysteine residues did not alter the affinity of $CREB_{259-327}$ for the CRE site. We have used CD analyses to examine the effect of the modification on the overall peptide conformation. We also examined the overall folding pattern of modified peptide that had been renatured from 6 M guanidine hydrochloride and after removal of the denaturant. The CD spectrum of a 25 μM modified $CREB_{259-327}$ sample is shown in Figure 4A. The spectrum is typical of α-helical type conformations with minima at 207 and 222 nm (Johnson, 1988). The mean residue ellipticity at 222 nm at this concentration is approximately −13,100, which corresponds to an estimated 43% α-helical content at 20 °C (Weiss et al., 1990). Thirty-one of the 71 residues present in $CREB_{259-327}$ (amino acid residues 296–327) are part of the leucine zipper region.

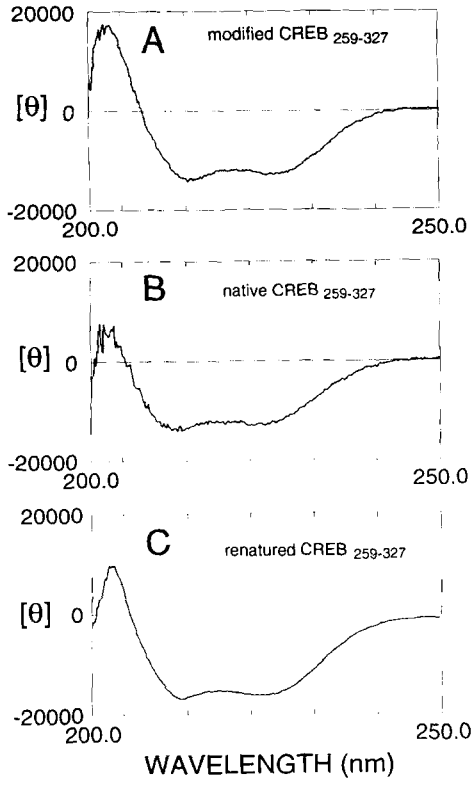

Fig. 4. CD analyses. **A:** CD spectrum of 25 μM modified $CREB_{259-327}$ peptide in 50 mM phosphate, 0.10 M NaCl (pH 7.4) at 20 °C. After carboxymethylation of the cysteine residues, the peptide sample was denatured in 6 M guanidine hydrochloride and dialyzed against 50 mM phosphate buffer, 0.10 M NaCl (pH 7.4). **B:** CD spectrum of 20 μM native $CREB_{259-327}$. **C:** CD spectrum of 20 μM denatured/renatured $CREB_{259-327}$ peptide.

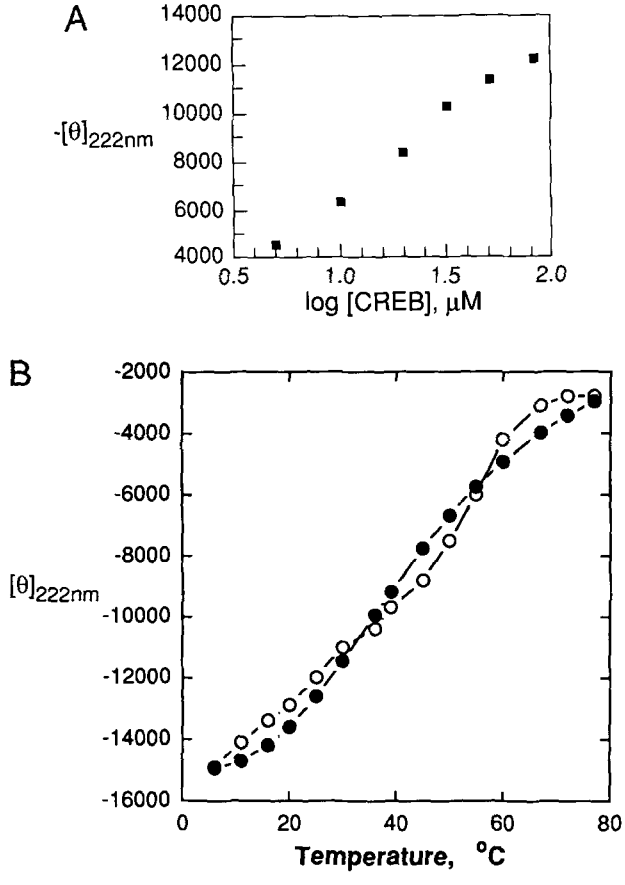

Fig. 5. A: Concentration dependence of the mean residue ellipticity at 222 nm of modified $CREB_{259-327}$ bZip peptide. All CD measurements were made using a 1-mm-pathlength cell in 50 mM phosphate, 0.10 M NaCl (pH 7.4) at 20 °C. Data points represent the average of five scans. **B:** Temperature dependence of the mean residue ellipticity at 222 nm of $CREB_{259-327}$ at a 45 μM concentration. Data points (○) covered the temperature range of 5–80 °C and were obtained at intervals of 5 °C. The best-fit calculated melting curve (●) assumes a single transition.

The observed mean residue ellipticity can be interpreted to indicate that, in the free peptide, the dimerization domain assumes a 100% α-helical conformation, while the DNA binding domain is predominantly unfolded. The total α-helical content observed for $CREB_{259-327}$ compares with that reported previously for the free Jun/Fos heterodimers with similar peptide chain length (Patel et al., 1990), although it is lower than the total α-helical content of the leucine zipper portion of GCN4 (Weiss, 1990; Weiss et al., 1990). Furthermore, the total ellipticity observed for the alkylated $CREB_{259-327}$ compares with that of the native peptide (Fig. 4B), as well as with that of the renatured samples concentrated to micromolar levels (Fig. 4C). These data indicate that protection of the cysteine residues or peptide renaturation does not alter the overall folding of the $CREB_{259-327}$ maintained at micromolar concentrations. The remaining analyses employ the modified $CREB_{259-327}$ peptide.

In order to characterize the folding properties of the $CREB_{259-327}$ peptide further, we examined the CD spectra as a function of peptide concentration and temperature (Fig. 5). Figure 5A shows the concentration dependence observed in the mean residue ellipticity values. The change in mean residue ellipticity observed in Figure 5A is a characteristic property of the leucine zipper family of proteins, where the coil-helix transition is coupled to the monomer-dimer equilibrium. The estimated dimerization constant for the CREB peptide is on the order of 20 μM, which compares with that of the synthetic GCN4 leucine zipper stretch (O'Shea et al., 1989a) as well as GCN4 peptide containing the DNA binding domain (O'Neal et al., 1991).

In addition to the concentration dependence, the coil-helix transition is temperature dependent. The overall melting temperature of the alkylated $CREB_{259-327}$ dimer at a 45 μM concentration is approx. 46 °C (Fig. 5B). This value compares with the T_m of the synthetic GCN4 leucine zipper peptide (O'Shea et al., 1989a). However the T_m of the $CREB_{259-327}$ peptide is 19 °C lower than the T_m of the GCN4 peptide containing the DNA binding do-

main (Weiss, 1990). The T_m obtained for the alkylated $CREB_{259-327}$ is identical to the melting temperature of the native or unmodified $CREB_{259-327}$ sample (data not shown). These data support the idea that cysteine modification has not altered the overall folding and thermodynamic properties of the peptide. Note that the data do not conform to a simple S-shaped melting curve (Fig. 5B). This could be due to a biphasic transition, as has been reported previously for the GCN4-bZip region. The broad transition with T_m of approximately 25 °C may correspond to the melting of the basic region, followed by the higher transition representing dissociation of the dimer.

One of the apparent properties of proteins of the leucine zipper family is the fact that the binding to their cognate DNA sequences induces a coil–helix transition of the basic DNA binding region (Patel et al., 1990; Weiss et al., 1990). Figure 6A shows that binding of the modified $CREB_{259-327}$ to a 30-bp oligonucleotide containing the CRE sequence induces approximately a 20% increase in the total α-helical content of $CREB_{259-327}$. A similar pattern is observed in the native unmodified $CREB_{259-327}$ (data not shown), indicating that modification of the cysteine residue in the DNA binding domain does not alter the recognition and overall folding pattern upon DNA binding. The increase in α-helical content can be attributed to specific DNA–protein interactions. Binding of $CREB_{259-327}$ to a shorter CRE sequence (14 bp) (Fig. 6B) induces the same extent of conformational change, 18%. On the other hand, binding to the 21-bp nonspecific AP-1 DNA sequence induces less than 5% increase in the total α-helical content of $CREB_{259-327}$ (Fig. 6C). This is consistent with the specific binding of $CREB_{327}$ to the CRE and not to the AP-1 site (Andrisani et al., 1988). It is interesting to note the CD results obtained when $CREB_{259-327}$ binds to an oligonucleotide containing the *lac* pseudo-operator sequence (Fig. 6D). This 22-bp oligonucleotide contains the sequence TGAgCGcTCA. The sequence is perfectly palindromic like the CRE site; however, it differs from the CRE sequence (TGACGTCA) by the insertion of the G and C residues on either side of the central CG pair. Binding of the modified $CREB_{259-327}$ peptide to the *lac* pseudo-operator stabilizes the α-helical structure of the DNA region to a greater extent than the AP-1 site, as indicated by the 10% increase in the total helical content.

NMR of the CREB bZip domain

The NH–NH region of the 2D nuclear Overhauser effect spectroscopy (NOESY) spectrum often can be used to define elements of secondary structure. Given the large extent of α-helical content observed in the CD spectra (Fig. 4), it is expected that an extensive array of interresidue connectivities would be observed in the NH–NH region of the spectrum. However, the few crosspeaks that were observed in this region correspond primarily to the aromatic as well as Asn and Gln side-chain connectivities (spectra not shown). In addition to the absence of interresidue crosspeaks indicative of some type of secondary structure in the peptide sample, the fingerprint region in the double quantum filtered COSY and total coherence spectroscopy (TOCSY) experiments showed <60% of the total number of crosspeaks expected for this specific peptide sequence. The absence of the majority of the intraresidue NH-alpha connectivities suggested the possibility of sample degradation. However, denaturing PAGE (Fig. 7A), as well as mass spectrometry analysis (data not shown), confirmed the integrity of the sample.

The NH chemical shifts cover a very narrow chemical shift range. In addition, the line widths were broader than expected for a peptide of this size ($\Delta\nu_{1/2}$ = 15–20 Hz at room temperature). Examination of the NMR spectra of

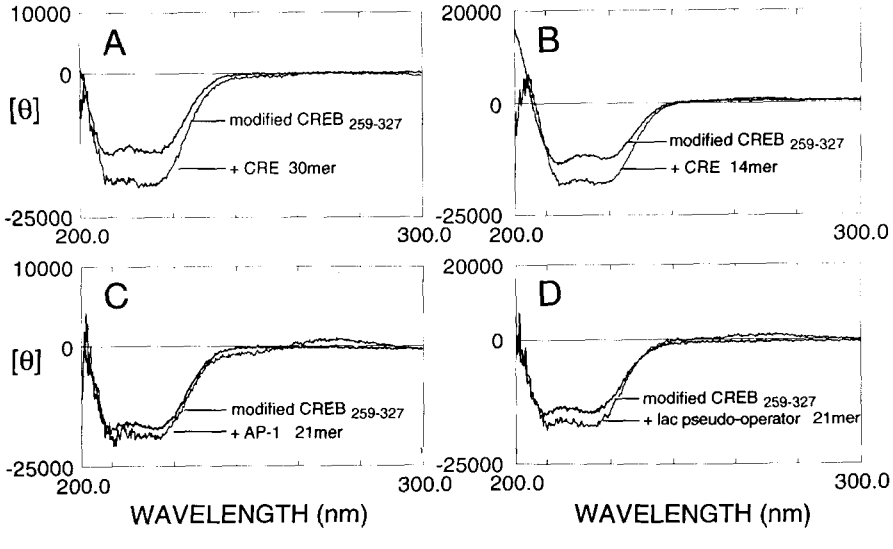

Fig. 6. CD of alkylated $CREB_{259-327}$ bound to DNA. **A:** Induced CD of $CREB_{259-327}$ by a 30-bp CRE oligonucleotide fragment. Peptide concentration, 25 μM; DNA concentration, 16 μM. **B:** Induced CD of $CREB_{259-327}$ by a 14-bp CRE oligonucleotide fragment. Peptide concentration, 25 μM; DNA concentration, 16 μM. **C:** Induced CD of $CREB_{259-327}$ by a 21-bp AP-1 oligonucleotide. Peptide concentration, 24 μM; DNA concentration, 16 μM. **D:** Induced CD of $CREB_{259-327}$ by the 22-bp *lac* pseudo-operator fragment. Peptide concentration, 16 μM; DNA concentration, 9 μM.

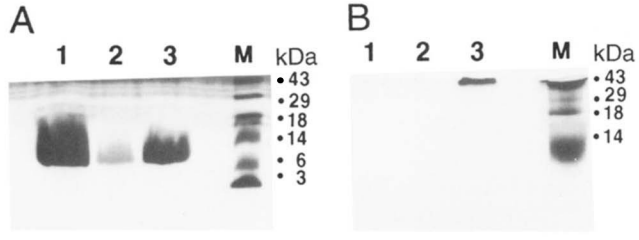

Fig. 7. A: Fifteen percent denaturing polyacrylamide gel electrophoresis. **B:** 15% native gel electrophoresis. Lane 1, 40 μg native $CREB_{259-327}$; lane 2, 20 ng denatured/renatured $CREB_{259-327}$; lane 3, 70 ng modified $CREB_{259-327}$; M, molecular weight markers.

the $CREB_{259-327}$ peptide under different peptide concentration or sample pH did not alter the observed spectral pattern (data not shown).

In order to understand the observed lack of spectral dispersion, we examined the conformation of the peptide by CD after it had been concentrated to millimolar NMR concentration. The CD spectrum of this sample appears to be 90% α-helical, as shown in Figure 8. Thus, concentration of the sample to millimolar levels and dilution to micromolar concentrations has doubled the α-helical content, presumably because of irreversible aggregation of the sample. Surprisingly, analysis of the peptide samples by nondenaturing PAGE (Fig. 7B) revealed that the native (lane 1), denatured/renatured (lane 2), and modified (lane 3) peptide appear to run as high molecular aggregates, i.e., the peptides do not enter the gel matrix. The same samples were analyzed in parallel by SDS-PAGE (Fig. 7A), thus confirming the integrity of the peptide samples. The diversity and size of these aggregated forms can thus explain the unresolved, broad ^{1}H-NMR lines for the bZip peptide. The large extent of α-helical character of the bZip peptide concentrated to millimolar levels can only be explained if, in addition to the leucine zipper region, protein–protein interactions also occur at the basic and linker regions, in the aggregated forms of the peptide.

Secondary structure of the CREB dimerization domain

In order to determine if the differences observed between the CREB bZip domain and other members of the leucine family also involved differences in the properties of the leucine zipper stretch, we further characterized a peptide sequence containing the amino acid residues 292–327 of the CREB leucine zipper and the linker region.

Figure 9 shows the CD spectrum of the CREB leucine zipper at a 68 μM concentration. The mean residue ellipticity is −26,670, which corresponds to an estimated 87% α-helical content. After corroborating the overall folding pattern of the CREB leucine zipper by CD, the sequence was then examined by 2D ^{1}H NMR. Figure 10A shows the NH–NH region of the 200-ms NOESY spectra. Contrary to the absence of connectivities observed for the CREB bZip sequence, the dimerization domain exhibits an extensive number of connectivities of the type (NH to NH) that are typically present in α-helices. The sequential assignment of the peptide spectra was performed following the methodology described by Wüthrich (1986). After completion of the assignment (Table 2), a total of 30 connectivities of the d_{NN} type were identified. In addition, an extensive number of connectivities of the $d_{\beta N}$, $d_{\alpha N}(i, i+3)$, and $d_{\alpha\beta}(i, i+3)$ were resolved (Fig. 10B), typical of a regular α-helix (Wüthrich, 1986). However, the connectivities corresponding to the peptide stretch between residues 31 and 35 were not resolved. Based on the identified assignments, the CREB leucine zipper sequence appears as one continuous helix from residue 1 to residue 25. Because of the symmetry of a coiled-coil α-helical dimer, NMR cannot be used to show that the helix exists as a dimer.

Discussion

The sequence-specific recognition of *cis*-acting elements by DNA binding transcription factors is the most important aspect of gene regulation. The molecular details of how transcription factors scan hundreds of thousands of base pairs and bind to their cognate binding site is largely

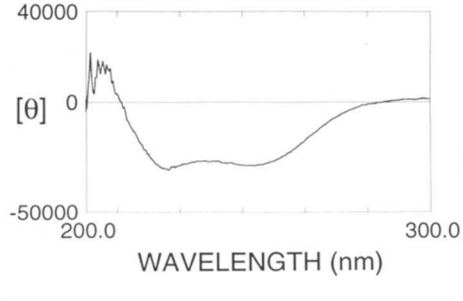

Fig. 8. The CD spectrum of modified $CREB_{259-327}$ peptide, initially concentrated to 3.1 mM, then diluted to 19 μM. Spectrum was acquired within 10 min after dilution.

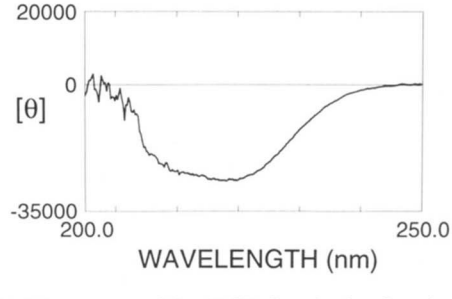

Fig. 9. The CD spectrum of the CREB dimerization domain at a 68 μM concentration. Spectrum was obtained in 20 mM DTT, pH 5.0.

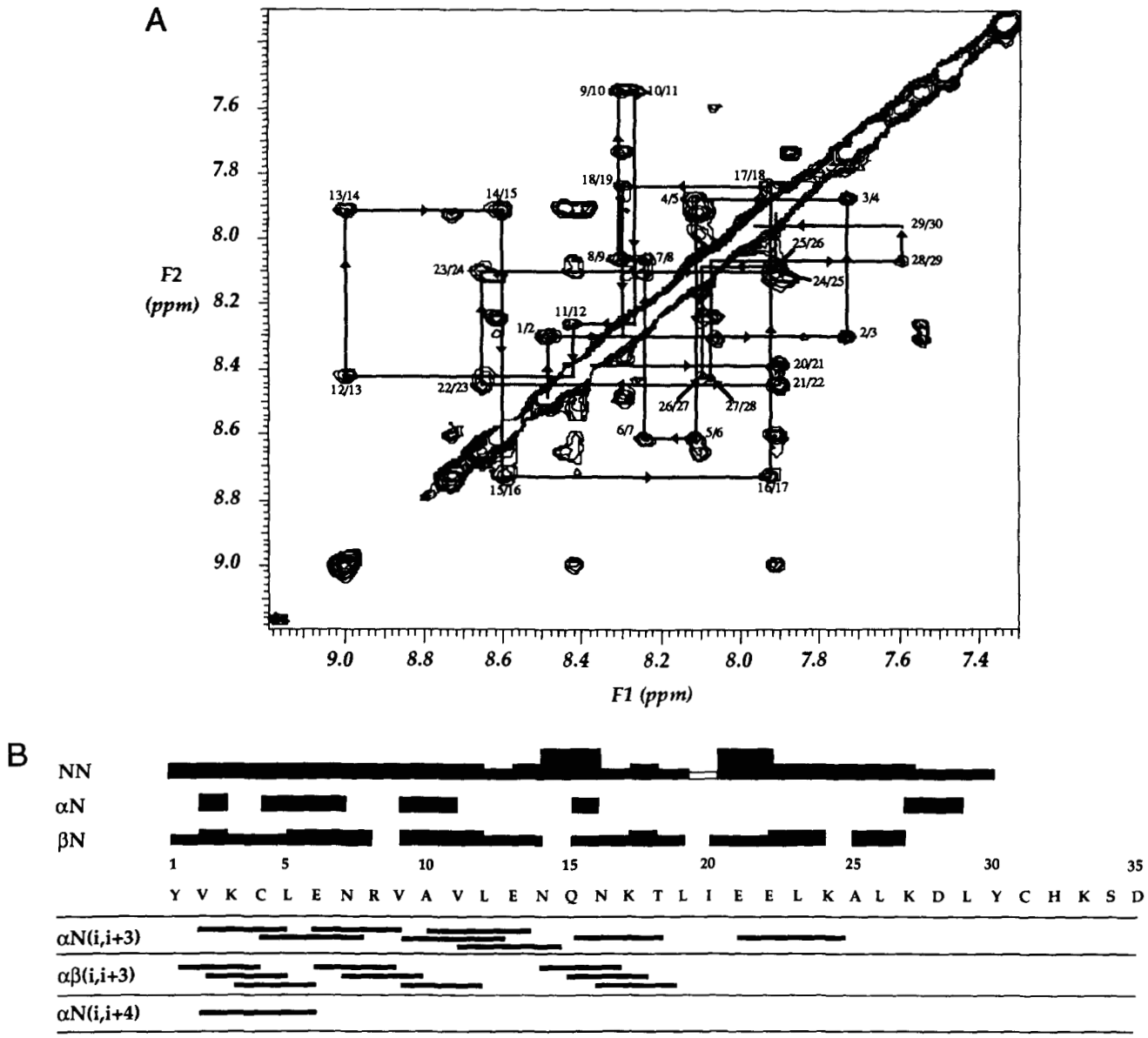

Fig. 10. A: The NH–NH region of a 200-ms NOESY spectrum of CREB dimerization domain. Sequential connectivities of the NH–NH type are indicated by the arrows. Crosspeaks not labeled have not been resolved yet. **B:** NOEs observed in the CREB dimerization domain. Connectivities between residues 31 and 35 could not be assigned unambiguously. Thicknesses of the NN, αN, and βN bars correlate with the intensity of the respective NOEs.

unknown. The leucine zipper family of proteins is composed of three distinct classes of proteins (Johnson & McKnight, 1989): the C/EBP protein, which binds to the CAAT box sequence; the CREB family of proteins, which bind to the palindromic CRE site (TGACGTCA); and the AP-1 family of proteins, which bind equally well to the AP-1 (TGAC/GTCA) and CRE sites. Recent biochemical studies point to the importance of the linker region of the CREB protein in CRE recognition (Andrisani & Dixon, 1991). However, the determinants of specificity governing the distinct binding of CREB to the CRE site are not yet understood. Structural studies examining the AP-1 and CREB class of proteins upon binding to the CRE and/or AP-1 sites are promising for their approach toward deciphering the specificity determinants of this class of leucine zipper proteins. To date, biophysical studies have only been carried out with the GCN4 and Jun/Fos proteins. In this paper we have conducted structural studies to understand how the CREB protein binds to the CRE site. We have successfully expressed in bacteria and purified the bZip domain of CREB. Utilizing the purification protocol described herein, we purified the $CREB_{259-327}$ peptide to near homogeneity (Fig. 1A). The purified peptide maintains the sequence-specific CRE rec-

Table 2. *Chemical shift (ppm) assignment of the CREB dimerization domain*[a]

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Others
Tyr 1	8.47	4.37	2.89, 3.16	6.75, 7.10
Val 2	8.28	3.31	1.95	0.84, 0.96
Lys 3	7.72	4.01	1.79, 1.66	1.45
Cys 4	7.85	4.10	3.04	
Leu 5	8.09	3.87	1.13, 1.23	1.84, 0.60, 0.69
Glu 6	8.59	3.75	1.99, 2.23	2.36
Asn 7	8.23	4.45	2.82, 2.88	
Arg 8	8.04	4.01	1.71, 1.99	3.07, 3.21, 7.14
Val 9	8.29	3.32	2.09	0.75, 0.091
Ala 10	7.53	4.055	1.50	
Val 11	8.24	3.69	2.11	0.89, 1.05
Leu 12	8.41	3.98	1.24, 1.74	0.82, 0.91
Glu 13	8.98	3.99	2.02, 2.25	2.41, 2.68
Asn 14	7.89	4.45	2.64	
Gln 15	8.58	3.99	1.99, 2.13	2.57
Asn 16	8.70	4.25	2.61, 3.12	
Lys 17	7.91	3.91	1.96	1.56
Thr 18	7.83	3.88	4.28	1.16
Leu 19	8.27	4.01	1.80	0.76
Ile 20	8.37	3.60	1.87	0.81
Glu 21	7.89	4.01	2.15, 2.28	
Glu 22	8.43	4.01	2.35, 2.75	
Leu 23	8.64	3.90	1.69, 1.84	1.57, 0.85
Lys 24	8.08	3.75	1.82, 1.52	1.29
Ala 25	7.90	4.06	1.44	
Leu 26	8.07	4.06		
Lys 27	8.41	3.91	2.80	
Asp 28	8.04	4.34	2.59, 3.12	
Leu 29	7.57	4.02	1.29, 1.49	
Tyr 30	7.93	4.36	2.45, 2.64	6.64, 6.93
Cys 31				
His 32				
Lys 33				
Ser 34	8.40	4.38	3.78, 3.83	
Asp 35				

[a] Assignments were referenced to the HOD peak at 4.76 ppm.

ognition, as demonstrated initially by gel retardation/CRE competition assays (Fig. 1B) and subsequently by the CD studies (Fig. 6).

The CD analysis of the modified $CREB_{259-327}$ peptide demonstrated that the overall α-helical character of $CREB_{259-327}$ increases from 43% to 63% upon binding to the CRE site (Fig. 6). The change in α-helical content is comparable to that observed for GCN4 (Weiss et al., 1990) and Jun/Fos (Patel et al., 1990). However, GCN4 binds with the same affinity to both the AP-1 and CRE sites, and identical comformational changes are induced in the protein upon binding to either DNA sequence (Weiss et al., 1990). In contrast to GCN4, CREB binds only to the CRE site (Andrisani et al., 1988). We demonstrate here that, as expected, binding of $CREB_{259-327}$ to the AP-1 site does not increase significantly the α-helical content of the bZip domain (Fig. 6), thus demonstrating that helix stabilization is coupled to specific DNA recognition.

The NMR studies of the GCN4 zipper (Oas et al., 1990) and the GCN4 bZip region (Saudek et al., 1990) are in agreement with the available CD data, and support that the dimerization domain is 100% α-helical in solution, whereas the basic DNA binding region displays the potential to fold in a similar type of structure. Both the CD (Fig. 9) and the 2D NMR data (Fig. 10A) corroborated the fact that the zipper region of the peptide is largely a continuous helix (presumably a dimeric coiled-coil) in aqueous solution. However, in contrast to the structural stability displayed by the GCN4 bZip region, our NMR structural studies of the $CREB_{259-327}$ could not be carried out because of aggregation of the sample. Further examination by CD of the native peptide, as well as modified peptide samples that have been concentrated to millimolar levels and then diluted to micromolar levels, revealed them to be approximately 90% helical in the absence of DNA (Fig. 8). The samples migrate as high molecular weight aggregates in nondenaturing PAGE (Fig. 7). This high degree of α-helical content of $CREB_{259-327}$ can only be explained if additional protein–protein contacts in the aggregated forms of the peptide further stabilize the predominantly unfolded basic region. The high quality of the NMR spectra for the zipper region alone suggests that aggregation of the $CREB_{259-327}$ peptide is not due to the leucine zipper portion of the peptide.

Recent studies (Williams et al., 1993) determined that the native full-length CREB protein displays a K_d of approximately 5×10^{-9} M for the perfectly palindromic CRE. Comparison of the K_d values of the full-length $CREB_{327}$ and bZip $CREB_{259-327}$ shows that $CREB_{327}$ binds to the perfectly palindromic CRE motif with 200-fold higher affinity. These results suggest that residues upstream of the basic region are necessary for proper folding or structural stabilization of the DNA binding domain. Whether any tertiary interactions indeed occur between the CREB amino-terminal residues and the basic region is presently not known.

Materials and methods

Cloning and expression of the $CREB_{259-327}$ peptide

A cDNA fragment corresponding to the bZip domain of $CREB_{327}$, spanning amino acid residues 259–327, was obtained by polymerase chain reaction (Saiki et al., 1988). The cDNA fragment was cloned into the *Eco*RI/*Hind*III sites of vector pT7-7 (Tabor & Richardson, 1987), in frame with the initiator methionine provided by the pT7-7 vector. This cloning arrangement results in the production of a peptide that has 72 amino acid residues, of which 68 are native amino acid residues from the carboxyl-terminus of $CREB_{327}$ and 4 are non-native amino acids at the

N-terminal part of the $CREB_{259-327}$ peptide. The plasmid was transformed into the bacterial strain Bl21(DE3)pLys S (Studier & Moffat, 1985). Bacterial cultures were grown to an optical density $(OD)_{600} = 0.6$, and expression of $CREB_{259-327}$ was induced with 0.5 mM isopropyl-β-thiogalactoside (IPTG) at 37 °C. Bacterial cultures were harvested 3 h after IPTG induction.

Peptides

Purification of the $CREB_{259-327}$ peptide

Approximately 30 g of cells were resuspended in 20 mL of lysis buffer (10 mM phosphate, pH 7.4, 0.5 M NaCl, 1 mM EDTA, 5 mM dithiothreitol [DTT]), sonicated for 1 min (two pulses, 30 s each) on ice, followed by centrifugation at 10,000 rpm for 20 min at 4 °C. The supernatant was heated to 75 °C for 10 min and centrifuged in an SS-34 rotor at 10,000 rpm for 20 min at 4 °C. The $CREB_{259-327}$ peptide was precipitated with 80% ammonium sulfate. The pellet was dissolved in 10 mM phosphate buffer (pH 7.4), 5 mM DTT, and dialyzed against the same buffer. The dialysate was chromatographed onto a hydroxylapatite (HP-Bio-Rad) column equilibrated with 10 mM phosphate buffer (pH 7.4), 0.2 M NaCl, and 5 mM DTT. Approximately 25 mL of resin were used for the sample obtained from a 12-L bacterial culture. The CREB peptide was eluted with three column volumes of 50 mM phosphate (pH 7.4), 0.5 M NaCl, 5 mM DTT, and was loaded onto a 2-mL S-Sepharose column. Following extensive washing (>5 column volumes) with 50 mM phosphate (pH 7.4) containing 0.4 M NaCl and 5 mM DTT, the $CREB_{259-327}$ peptide was eluted with 10 column volumes of the same buffer containing 0.5 M NaCl. The final purification step includes chromatography on an FPLC phenyl superose column (Bio-Rad). The S-Sepharose eluate was adjusted to 1 M ammonium sulfate and loaded onto the FPLC hydrophobic resin. Elution was at 50% of the starting buffer. Peptide purity was corroborated by reverse-phase HPLC. The sample was injected into a Vydac C-18 column using a 0–100% H_2O/CH_3CN gradient containing 0.1% trifluoroacetic acid.

Synthesis and purification of the leucine zipper peptide

The leucine zipper peptide, spanning amino acid residues 292–327 of CREB, was synthesized using the solid-phase method on an Applied Biosystems Model 431A machine using the small scale (0.1 mmol) t-Boc program. The complete peptide was cleaved from the resin with HF/10% anisole at 0 °C for 45 min, extracted into dilute acetic acid, and lyophilized. Analytical reverse-phase HPLC indicated a single major product. Aliquots (30 mg) of crude peptide were purified by reverse-phase chromatography.

K_d determination studies

Gel retardation assays (Andrisani et al., 1988) were utilized for the determination of the K_d of $CREB_{259-327}$ for the CRE site. The binding reactions were incubated at room temperature for 5 min. Electrophoresis was at 200 V on 5% native, low ionic strength polyacrylamide gels for 1 h. The gel was dried and autoradiographed. The free and bound DNA was quantitated by excising and counting the corresponding portions of the dried gel by liquid scintillation counting.

For the K_d determination studies carried out as described in Williams et al. (1993), purified CREB peptide was freshly diluted in the presence of 50 ng/mL of bovine serum albumin (BSA) to prevent nonspecific adsorption. The amount of $CREB_{259-327}$ utilized in the binding reactions ranged from 0.5 ng to 1.0 μg. The CRE 30-mer oligonucleotide was used as the radiolabeled probe (6,000 cpm/reaction, <10 pM). For the K_d determination via DNA(CRE) titration, 10 ng of alkylated bZip peptide were diluted in 1 mg/mL BSA. The amount of radiolabeled CRE 30-mer DNA utilized in the binding reactions ranged from 1 to 500 ng.

Reduction and carboxymethylation of cysteine residues

Purified $CREB_{259-327}$ was reduced in 50 mM phosphate buffer (pH 8.5), 0.5 M NaCl, using a 100-fold molar excess of DTT. Reduction proceeded under argon for 2 h. Iodoacetamide was then added to the sample, maintaining a threefold excess of iodoacetamide to DTT. The sample pH was maintained at 8.5 during the carboxymethylation step (45 min to 1 h) and was then adjusted to pH 5.0 in order to stop the reaction. Modification of the thiol groups was monitored by means of the Ellman's reagent test following Pierce's manufacturing instructions.

CD studies

CD spectra were obtained on a Jasco J600 CD spectropolarimeter. All spectra were run using a 1-mm-pathlength cell and represent the computer average of five scans. Samples were prepared in 50 mM phosphate buffer, 0.1 M NaCl (pH 7.4), at 20 °C. DNA-protein binding reactions were incubated at room temperature for 20 min prior to scanning. The CD melting data was fit to an S-shaped logistic function by nonlinear least-squares program (SigmaPlot). Attempts to fit the data to a biphasic curve were unsuccessful.

NMR studies

Spectra were recorded using either a Varian VXR500S or Varian VXR600S spectrometer. The samples corresponding to the CREB dimerization domain were prepared in

90% H_2O/10% D_2O, 0.10 M deuterated DTT (pH 5.0). The TOCSY experiments (Braunshweiler & Ernst, 1983) were carried out using 35–125 ms mixing time. The mixing time in the NOESY experiments (Kumar et al., 1981) varied between 100 and 300 ms. All experiments were performed with a sweep width of 6,000 Hz, at room temperature (approximately 20 °C) unless specified otherwise. The data were typically acquired with two-dimensional points in the F_2 dimension, 512 complex points, and 32 transients per free induction decay (FID). All spectra were processed using a combination of sine bell and Gaussian functions.

Acknowledgments

We thank Pei-Ling Wang for technical assistance. Part of the work reported was performed by O.M.A. while a Research Scientist in the laboratory of Dr. Jack E. Dixon, Walther Cancer Institute and Department of Biochemistry, Purdue University. We thank him for his support and advice on these studies. This work was supported by start-up funds, an institutional grant from the Purdue Cancer Center, NIH grant R29 DK44533 to O.M.A., and NIH grant AI27744 to D.G.G. The Purdue University Biochemical Magnetic Resonance Laboratory is supported by the NSF Biological Facilities Center on Biomolecular NMR, Structure and Design at Purdue (grants BBS-8614177 and DIR-9000360 from the Division of Biological Instrumentation) and the NIH Designated Aids Center at Purdue (grant AI727713).

References

Abate, C., Luk, D., Gentz, R., Raucher, F.J., & Curran, T. (1990). Expression and purification of the leucine zipper and DNA-binding domains of Fos and Jun: Both Fos and Jun contact DNA directly. *Proc. Natl. Acad. Sci. USA 87*, 1032–1036.

Andrisani, O.M. & Dixon, J.E. (1991). Involvement of lysine residues 289 and 291 of the cAMP-responsive element-binding protein in the recognition of the cAMP-responsive element. *J. Biol. Chem. 266*, 21444–21450.

Andrisani, O.M., Pot, D.A., Zhu, Z., & Dixon, J.E. (1988). Three sequence-specific DNA–protein complexes are formed with the same promoter element essential for expression of the rat somatostatin gene. *Mol. Cell. Biol. 8*, 1947–1956.

Andrisani, O.M., Zhu, Z., Pot, D.A., & Dixon, J.E. (1989). In vitro transcription directed from the somatostatin promoter is dependent upon a purified 43-kDa DNA-binding protein. *Proc. Natl. Acad. Sci. USA 86*, 2181–2185.

Braunshweiler, L. & Ernst, R.R. (1983). Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J. Magn. Reson. 53*, 521–528.

Bush, S.J. & Sassone-Corsi, P. (1990). Fos, Jun and CREB basic-domain peptides have intrinsic DNA-binding activity enhanced by a novel stabilizing factor. *Oncogene 5*, 1549–1556.

Carey, J. (1989). Gel retardation at low pH resolves trp repressor–DNA complexes for quantitative study. *Proc. Natl. Acad. Sci. USA 85*, 975–979.

Dwarki, V.J., Montminy, M., & Verma, I.M. (1990). Both the basic region and the leucine zipper domain of CREB protein are essential for transcriptional activity. *EMBO J. 9*, 225–232.

Hai, T., Liu, F., Coukos, W.J., & Green, M.R. (1989). Transcription factor ATF cDNA clones: An extensive family of leucine zipper proteins able to selectively form DNA-binding heterodimers. *Genes Dev. 3*, 2083–2090.

Hope, I.A. & Struhl, K. (1986). Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell 46*, 885–894.

Hope, I.A. & Struhl, K. (1987). GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. *EMBO J. 6*, 2781–2784.

Johnson, P.F. & McKnight, S.L. (1989). Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem. 58*, 799–839.

Johnson, W.C., Jr. (1988). Secondary structure of proteins through circular dichroism spectroscopy. *Annu. Rev. Biophys. Chem. 17*, 145–166.

Kouzarides, T. & Ziff, E. (1988). The role of the leucine zipper in the Fos–Jun interaction. *Nature 336*, 646–651.

Kumar, A., Wagner, G., Ernst, R.R., & Wüthrich, K. (1981). Buildup rates of the nuclear Overhauser effect measured by two-dimensional proton magnetic resonance spectroscopy: Implications for studies of protein comformation. *J. Am. Chem. Soc. 103*, 3654–3658.

Landschulz, W.H., Johnson, P.F., & McKnight, S.L. (1988). The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science 240*, 1759–1764.

Montminy, M.R. & Bilezikjian, L.M. (1987). Binding of a nuclear protein to the cAMP response element of the somatostatin gene. *Nature 328*, 175–178.

Oas, T.G., McIntosh, L.P., O'Shea, E.K., Dahlquist, W., & Kim, P.S. (1990). Secondary structure of a leucine zipper determined by nuclear magnetic resonance spectroscopy. *Biochemistry 29*, 2891–2894.

O'Neal, K.T., Shuman, J.D., Ampe, C., & DeGrado, W.F. (1991). DNA-induced increase in the α-helical content of C/EBP and GCN4. *Biochemistry 30*, 9030–9034.

O'Shea, E.K., Klemm, J.D., Kim, P.S., & Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled-coil. *Science 254*, 539–544.

O'Shea, E.K., Rutkowski, R., & Kim, P.S. (1989a). Evidence that the leucine zipper is a coiled-coil. *Science 243*, 538–542.

O'Shea, E.K., Rutkowski, R., Stafford, W.F., & Kim, P.S. (1989b). Preferential heterodimer formation by isolated leucine zippers from Fos and Jun. *Science 245*, 646–648.

Patel, L., Abate, C., & Curran, T. (1990). Altered protein conformation on DNA binding by Fos and Jun. *Nature 347*, 572–575.

Roesler, W.J., Vanderbark, G.R., & Hanson, R.W. (1988). Cyclic AMP and the induction of eukaryotic gene transcription. *J. Biol. Chem. 263*, 9063–9066.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., & Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science 239*, 487–491.

Saudek, V., Pastore, A., Morelli, M., Frank, R., Gausepohl, H., Gibson, T., Wih, F., & Roesch, P. (1990). Solution of the DNA-binding domain of the yeast transcriptional activator protein GCN4. *Protein Eng. 4*, 3–10.

Smith, D.B. & Johnson, K.S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene 67*, 31–40.

Studier, F.W. & Moffat, B.A. (1985). Use of bacteriophage T7 RNA polymerase to direct the expression of cloned genes. *J. Mol. Biol. 189*, 113–130.

Tabor, S. & Richardson, C.C. (1987). A bacteriophage T7 polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA 82*, 1074–1078.

Vogt, P.K. & Bos, T.J. (1990). Jun: Oncogene and transcription factor. *Adv. Cancer Res. 55*, 1–31.

Weiss, M.A. (1990). Thermal unfolding studies of a leucine zipper domain and its specific DNA complex: Implications for scissor's grip recognition. *Biochemistry 29*, 8020–8024.

Weiss, M.A., Ellenberger, T., Wobbe, C.R., Lee, J.P., Harrison, S.C., & Struhl, K. (1990). Folding transition in the DNA-binding domain of GCN4 on specific binding to DNA. *Nature 347*, 575–578.

Williams, J.S., Dixon, J.E., & Andrisani, O.M. (1993). Binding constant determination studies utilizing recombinant ΔCREB protein. *DNA Cell Biol. 12*, 183–190.

Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*. Wiley, New York.