Luebke, K. J., & Dervan, P. B. (1989) *J. Am. Chem. Soc. 111*, 8733–8735.
Lyamichev, V. I., Mirkin, S. M., & Frank-Kamenetskii, M. D. (1986) *J. Biomol. Struct. Dyn. 3*, 667–669.
Mirkin, S. N., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Fillipov, S. A., & Frank-Kamenetskii, M. D. (1987) *Nature 330*, 495–497.
Mooren, M. M., Pulleyblank, D. E., Wijmenga, S. S., Bloomers, M. J., & Hilbers, C. W. (1990) *Nucleic Acids Res. 18*, 6523–6529.
Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645–650.
Perronault, L., Asseline, U., Rivalle, C., Thuong, N. T., Bisagni, E., Giovannangeli, C., Le Doan, T., & Helene, C. (1990) *Nature 344*, 358–360.
Posvic, T. J., & Dervan, P. B. (1989) *J. Am. Chem. Soc. 111*, 3059–3061.
Plum, G. E., Park, Y. W., Singleton, S. F., Dervan, P. B., & Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9436–9440.
Praseuth, D., Perronault, L., LeDoan, T., Chassignol, M., Thuong, N., & Helence, C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1349–1353.
Rajagopal, P., & Feigon, J. (1989) *Biochemistry 28*, 7859–7870.
Senior, M., Jones, R. A., & Breslauer, K. J. (1988) *Biochemistry 27*, 3879–3885.
Shimizu, M., Hanvey, J. C., & Wells, R. D. (1989) *J. Biol. Chem. 264*, 5944–5949.
Sklenar, V., & Feigon, J. (1990) *Nature 345*, 836–838.
Strobel, S. A., & Dervan, P. B. (1990) *Science 249*, 73.
Strobel, S. A., & Dervan, P. B. (1991) *Nature 350*, 172–174.
Sun, J. S., Francois, J. C., Montenay-Garestier, T., Saison-Behmoras, T., Roig, V., Thuong, N. T., & Helene, C. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9198–9202.
Van de Ven, F. J., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1–38.
Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., & Wohlrab, F. (1988) *FASEB J. 2*, 2939–2949.

DNA-Induced Increase in the α-Helical Content of C/EBP and GCN4†

Karyn T. O'Neil,‡ Jon D. Shuman,§ Christophe Ampe,‖ and William F. DeGrado*,‡

Central Research and Development Department, The DuPont Company, Wilmington, Delaware 19880-0328, Department of Embryology, Carnegie Institution of Washington, Baltimore, Maryland 21210, and Department of Biochemistry and Biophysics, Yale University, New Haven, Connecticut 06511

Received March 25, 1991; Revised Manuscript Received June 7, 1991

ABSTRACT: Leucine zipper proteins comprise a recently identified class of DNA binding proteins that contain a bipartite structural motif consisting of a "leucine zipper" dimerization domain and a segment rich in basic residues responsible for DNA interaction. Protein fragments encompassing the zipper plus basic region domains (bZip) have previously been used to determine the conformational and dynamic properties of this motif. In the absence of DNA, the coiled-coil portion is α-helical and dimeric, whereas the basic region is flexible and partially disordered. Addition of DNA containing a specific recognition sequence induces a fully helical conformation in the basic regions of these fragments. However, the question remained whether the same conformational change would be observed in native bZip proteins where the basic regions might be stabilized in an α-helical conformation even in the absence of DNA, through interactions with portions of the protein not included in the bZip motif. We have now examined the DNA-induced conformational transition for an intact bZip protein, GCN4, and for the bZip fragment of C/EBP with two enhancers that are differentially symmetric. Our results are consistent with the induced helical fork model wherein the basic regions are largely flexible in the absence of DNA and become fully helical in the presence of the specific DNA recognition sequence.

The leucine zipper class of DNA binding proteins comprises a group of structurally related proteins including numerous factors involved in transcriptional modulation and oncogenesis (Johnson & McKnight, 1989; Lanschultz et al., 1988). These proteins bind specifically to DNA as dimers and frequently recognize dyad-symmetric DNA sequences using a bipartite "bZip" motif (Vinson et al., 1989) consisting of a basic region and a leucine zipper. Genetic and biochemical evidence (Lanschultz et al., 1988; Hope & Struhl, 1987; Kouzarides & Ziff, 1989; Sassone-Corsi et al., 1988; Gentz et al., 1989; Turner & Tjian, 1989) indicates that the leucine zipper domain is responsible for forming homo- and heterodimers while the DNA binding activity resides in the basic region, which is immediately N-terminal to the leucine zipper and approximately 25 residues in length. This sequence contains numerous positively charged residues as well as neutral residues that appear essential for sequence-specific interactions with the base pairs (O'Neil et al., 1990; Nakabeppu & Nathans, 1989; Sellers & Struhl, 1989; Nueberg et al., 1989; Shuman et al., 1990). Biophysical studies (O'Shea et al., 1989a,b) indicate that the leucine zipper forms a two-stranded coiled-coil, a 2-fold symmetric dimer of parallel α-helices (Crick, 1953) often found in fibrous proteins (Cohen & Parry, 1990). This structure shows an exact seven-residue repeat, matching the

† C.A. is supported by an Anna Fuller Fund fellowship; additional funding was provided by National Institutes of Health Grant GM39546. J.D.S. is supported by a fellowship stipend from the Leukemia Society; additional funding was provided by the Howard Hughes Medical Research Institute and the Carnegie Institution of Washington.
‡ The Dupont Company.
§ Carnegie Institution of Washington.
‖ Yale University.

0006-2960/91/0430-9030$02.50/0

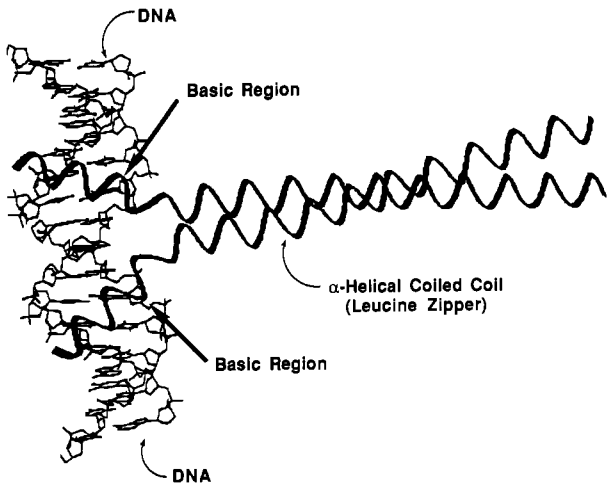

FIGURE 1: Computer-generated model for the site-specific complex between a bZip fragment and its recognition sequence. The model was generated as described previously (O'Neil et al., 1990).

period of leucine residues in the leucine zipper (Lanschultz et al., 1988; O'Shea et al., 1989a). Recent NMR and X-ray scattering studies (Saudek, et al., 1990; Oas et al., 1990; Rasmussen et al., 1991) have shown that the α-helical conformation is continuous throughout the entire zipper, and are also consistent with the symmetry of the α-helical coiled-coil.

Two similar models, the scissors grip (Vinson et al., 1989) and the induced helical fork (O'Neil et al., 1990), have been proposed to describe how the basic regions project from the leucine zipper and mediate a specific interaction with DNA (Figure 1). In both models, the basic regions form α-helices that are continuations of the leucine zipper α-helices, and diverge from the bifurcation point at the N-terminus of the leucine zipper. In a complex with DNA, the basic region helices bind in the major groove of DNA and track in opposite directions to cover the recognition site. The primary difference between the two models is that the induced helical fork model predicts which residues are involved in sequence-specific contacts while the scissors grip model does not explicitly address this point. Also, the scissors grip model predicts that the basic region helix is kinked at a conserved Asn to allow continuous interaction of the basic region helix with the major groove while the induced helical fork model predicts that the basic region helix is more smoothly curved.

The induced helical fork model predicts that a conformational change in the basic region should accompany binding to DNA; in the absence of DNA, electrostatically repulsive interactions between positively charged residues should drive the basic region into a more extended, flexible conformation. Thus, the α-helical conformation would be induced only when the structure was bound to DNA. The scissors grip model also considered the possibility that some portion of the basic region might be flexible in the absence of DNA (Vinson et al., 1989).

Considerable experimental data now support the induced helical fork and scissors grip models. For example, NMR and CD studies of fragments of GCN4, Fos, and Jun have shown that the leucine zipper portion is helical and dimeric (at micromolar concentrations) in the presence and absence of DNA but that the basic region is fully helical only in the presence of its specific DNA binding site (O'Neil et al., 1990; Shuman et al., 1990; Talanian et al., 1990; Weiss et al., 1990; Weiss, 1990; Saudek et al., 1990). Interestingly, O'Neil et al. (1990) found that the α-helical conformation of the basic region was induced only by DNA containing the target site and binding to nonspecific DNA induced considerably less α-helix content. In addition, chemical footprinting (Gartenberg et al., 1990) and DNA-cleavage studies (Oakley & Dervan, 1990) are consistent with models in which the basic regions are oriented in the major groove as in the induced helical fork and scissors grip models.

The goals of the present paper are 2-fold. We first wished to determine whether the conformational change observed for bZip fragments would also occur in intact, native leucine zipper proteins. In a native protein, it is possible that there are regions outside the bZip motif that are involved in stabilizing the basic region in an α-helical conformation, even in the absence of DNA. We therefore conducted CD measurements on intact GCN4 in the presence and absence of its DNA recognition site. A second goal is to determine the extent of conformational change in the basic region that occurs when peptides bind to asymmetric DNA sites. For example, while C/EBP is known to recognize an inverted repeat within some cellular enhancer regions (e.g., TTGCGCAA; Shuman et al., 1990), it also interacts with a considerably less symmetrical site within certain viral enhancers (e.g., TTGCAAAA; Roman et al., 1990). Thus, on the basis of earlier studies showing that the basic regions were fully helical only in the presence of specific DNA, it appeared possible that both copies of the basic regions in the dimer would be fully helical in the complex with symmetrical sites but only one would be α-helical in the complex with the asymmetric site. To test this possibility, we examined the CD spectrum of a 62-residue fragment of C/EBP with oligonucleotides bearing either the symmetric or the asymmetric sites. We show that a fully helical conformation is induced in the basic region of intact GCN4 as well as in the C/EBP fragment irrespective of the symmetry of the binding site.

Materials and Methods

Native GCN4 was purified as described (Gartenberg et al., 1990) except that a final size exclusion chromatography step using a Superose 12 column (Pharmacia/LKB, Piscataway, NJ) was added. The eluting buffer for the size exclusion column was 100 mM NaCl, 100 mM Tris, and 0.2 mM EDTA, pH 8.0. C/EBP62 was purified as described previously (Shuman et al., 1990); the peptide encompasses residues 281–340 of the native protein. All protein and peptide concentrations were determined by amino acid analysis using a Beckman System Gold, ion-exchange system with ninhydrin detection (Beckman Instruments, Somerset, NJ). Oligonucleotides were synthesized on an ABI oligonucleotide synthesizer (ABI, Foster City, CA) using the standard synthesis and cleavage protocol described by the manufacturer. Oligonucleotides for CD measurements were purified by reverse-phase HPLC using either a Rainin Pure DNA column (Rainin Instruments, Woburn, MA) or an ion-exchange FPLC on a Mono Q column (Pharmacia/LKB). For gel mobility experiments, the oligonucleotides were further purified on a 10% polyacrylamide–urea gel (Maniatis, 1982).

CD Measurements. All measurements were made either on a Jasco J500 spectropolarimeter interfaced to an IBM/PC using the software provided by Jasco or on an Aviv 60DS spectropolarimeter. Both instruments were calibrated as described by Johnson (1990). Measurements were made for approximately 20 μM peptide or protein solutions in 10 mM MOPS/150 mM NaCl at pH 7.5. DNA concentration was determined spectrophotometrically for the single strands by using 1 ODu $\approx$33 μg/mL. Proteins had the following sequences: TRE, GTGCACTGGAT**GAGTCA**TAGGCACTG (O'Neil et al., 1990); C/EBP1, CAGTGCC**TTTGCGCAA**CCAGTGCAC (Shuman et al., 1990); C/EBP2, CAGTGTGG**TTTTGCAA**GAGGAAGC (Roman et al.,

1990); random, GTGCACTGGATGTTAGCTAGGCACTG.

After equal molar amounts of complementary strands were mixed in 150 mM NaCl, the oligo mixture was annealed by heating at 70 °C for 10 min followed by slow-cooling in a beaker of water from the 70 °C bath. For CD experiments, the peptide:DNA ratio was always less than 1:1.5 to be sure that all of the peptide was bound in the DNA complex. Due to excessive adsorption of C/EBP62 to the walls of the cells used for CD measurements, all peptide concentrations were determined directly from the CD cell by transfer of an aliquot directly to a tube used for hydrolysis. Amino acid analysis was then used to quantitate the amount of peptide in solution. A base-line run was performed after the cell was briefly rinsed with H_2O. The cell was washed with 2% SDS between each pair of peptide and base-line runs. Subtraction of the spectra for DNA from the DNA/protein complex was performed as described previously (O'Neil et al., 1990).

Gel Electrophoresis. Oligonucleotides were phosphorylated at the 5′ terminus with ^{32}P using polynucleotide kinase and [γ-^{32}P]ATP (6000 Ci/mmol)(New England, Nuclear, Billerica, MA). Labeled oligonucleotides were annealed to the unlabeled complementary oligonucleotide. Binding reactions between labeled DNA (TRE site, 11 nM; C/EBP sites, 100 nM) and protein or peptide (11 or 38 nM GCN4, and 100 or 360 nM C/EBP62) were incubated at 30 °C for 30 min and electrophoresed as described previously (O'Neil et al., 1990).

Results and Discussion

GCN4. The CD spectrum of native GCN4 in the absence of DNA displays a maximum near 190 nm and a minimum at 204 nm with a broad shoulder centered around 222 nm characteristic of proteins with partial helical content (Figure 2A). As illustrated in Figure 2B, the ellipticity of GCN4 at 222 nm is concentration-dependent; the data are well described by a monomer–dimer equilibrium with a dimerization constant of 11.8 ± 3.5 μM. The mean residue ellipticity of the dimer at 222 nm, $[\theta_{222}]$, is −10 800 ± 200 deg cm^2 dmol^{-1}, and the monomer ellipticity is −6000 deg cm^2 dmol^{-1}. The lower helicity observed at low concentrations probably arises from incomplete helix formation in the coiled-coil portion of the protein that is responsible for dimerization.

Addition of DNA to predominantly dimeric GCN4 changes both the overall shape of the CD spectrum and the magnitude of $[\theta_{222}]$ (Figure 2A). In the presence of DNA containing a specific TRE recognition site, $[\theta_{222}]$ of GCN4 changes to −13 500 ± 900 deg cm^2 dmol^{-1}, and the lower wavelength minimum shifts to 207 nm. The enhanced magnitude of the ellipticity at 222 nm in the presence of DNA (beyond the maximum value observed for the dimer; see above) is −2700 ± 900 deg cm^2 dmol^{-1}. When the estimate of Greenfield and Fasman (1969) is applied, this corresponds to approximately 20 residues becoming fully helical in the presence of the specific DNA sequence. This value correlates well with the length of the GCN4 basic region (20–25 residues). As discussed previously, minimal changes occur in the near-UV spectrum associated with the nucleotide bases, suggesting that the DNA has not undergone major conformational changes (O'Neil et al., 1990).

In comparison, addition of a random sequence of DNA induces a much smaller increase in the magnitude of $[\theta_{222}]$. This smaller decrease is not the result of incomplete binding of GCN4 since addition of more DNA did not alter the measurement. This finding is consistent with the previous results of O'Neil et al. (1990) using a minimally designed bZip peptide in which addition of nonspecific DNA induced a small increase in α-helical content while addition of the specific

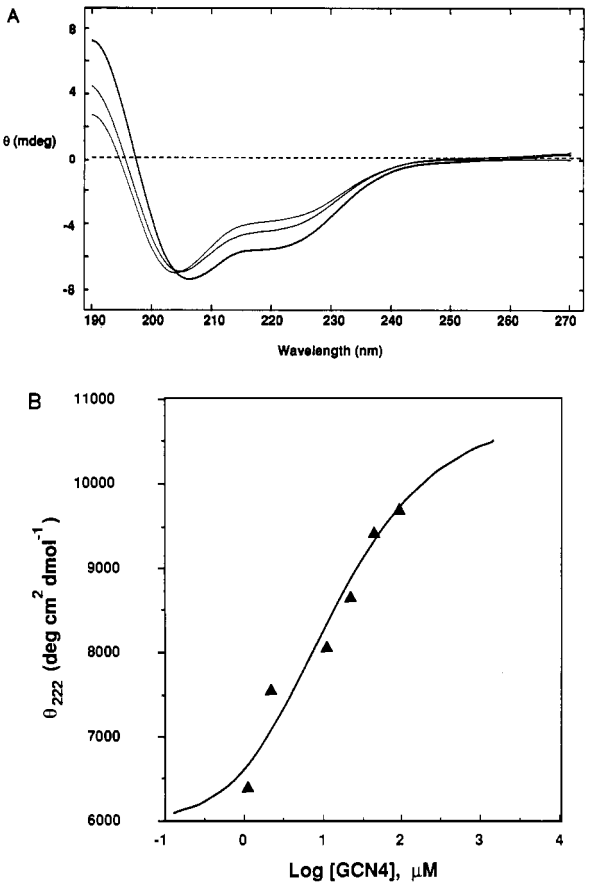

FIGURE 2: (A) Induced CD for GCN4 by either TRE site-specific or nonspecific DNA. Spectra were recorded for 19 μM protein (9.5 μM dimer) and 15 μM DNA duplex concentration in 10 mM MOPS/150 mM NaCl, pH 7.5, at room temperature. Data were collected in 1 nM steps with an 18-s integration time and smoothed by using Aviv software (Aviv Associates, Lakewood, NJ). The top curve is the spectrum for GCN4 in the absence of DNA, the center curve is the induced CD spectrum with nonspecific DNA, and the bottom curve is the spectrum for the specific DNA complex with GCN4. The spectra for the DNA have been subtracted from the spectra of the DNA protein complexes as described previously (O'Neil et al., 1990). (B) Concentration dependence of $[\theta_{222}]$ for GCN4. CD measurements were made for various protein concentrations in 10 mM MOPS/0.15 M NaCl, pH 7.5, in either a 0.01- or a 0.1-cm path-length cell. Data points, were averaged over ~5 min and base-line-corrected and have been fit by using MLAB software running on an IBM/PC (Civilized Software, Inc., Bethesda, MD).

DNA recognition sequence induced a fully helical conformation in the basic region (O'Neil et al., 1990). Such an induced conformational change in a large protein is not without precedent. Tan and Richmond (1990) have recently described a similar DNA-induced conformational change in the yeast transcriptional activator PRTF, based on changes in proteolysis patterns when the protein is bound specifically to DNA.

To confirm that GCN4 bound the TRE sequence with high affinity and in a 1:1 complex, a gel mobility shift assay was performed. Addition of GCN4 to the labeled oligonucleotide resulted in specific complex formation as evidenced by appearance of a new band with lower mobility (Figure 3).

C/EBP. The CD spectrum of C/EBP62 in the absence of DNA shows similar concentration-dependent behavior that is well described by a monomer–dimer equilibrium with a value of $[\theta_{222}]$ equal to −22 000 and −7000 deg cm^2 dmol^{-1} for the dimer and monomer, respectively, and an equilibrium constant of 17 μM (data not shown). The ellipticity of the dimer correlates well with previous studies indicating that the leucine zipper (~38 residues) is helical while the basic region (~25

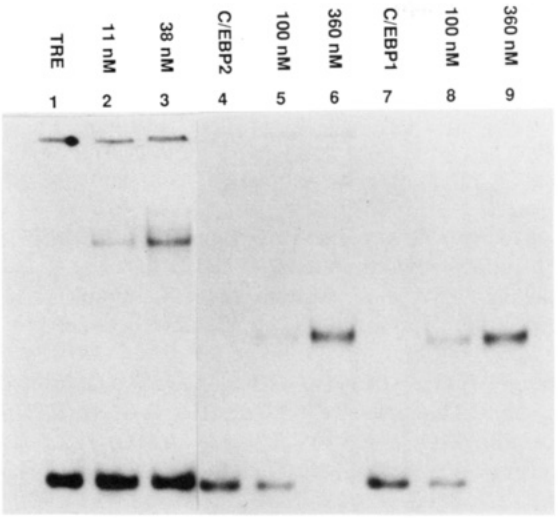

FIGURE 3: Gel shift mobility assay of GCN4 and C/EBP62 with the specific oligos used for CD experiments. Lanes 1–3, TRE site oligo (12 nM) plus indicated amount of GCN4; lanes 4–6, C/EBP2 site oligo (60 nM) plus indicated amount of C/EBP62; lanes 7–9, C/EBP1 site oligo (60 nM) plus indicated amount of C/EBP62. Oligonucleotides were labeled at the 5′-OH with polynucleotide kinase. After incubation under the conditions described under Materials and Methods, the complexes were loaded onto a 5% polyacrylamide gel. After electrophoresis, the gel was fixed with 5% acetic acid and autoradiographed.

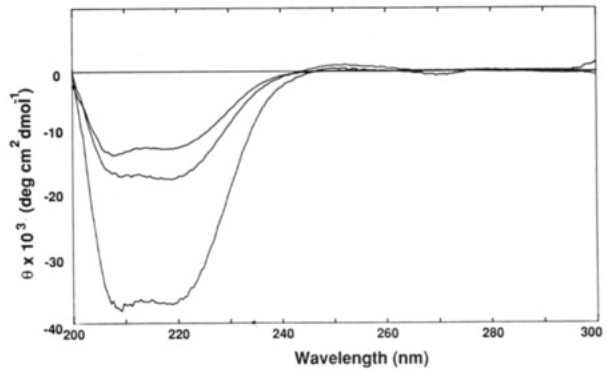

FIGURE 4: Induced CD for C/EBP62 by either C/EBP1 site-specific or nonspecific DNA. Spectra were recorded for approximately 24 μM peptide (12 μM dimer) and 20 μM DNA duplex concentration in 10 mM MOPS/150 mM NaCl, pH 7.5, at room temperature. Data were collected and smoothed on a Jasco J-500 interfaced to an IBM/PC (Jasco, Inc., Easton, MD) using software provided by Jasco. The top curve is the spectrum of C/EBP62 in the absence of DNA, the center curve is the induced CD spectrum with nonspecific DNA, and the bottom curve is the spectrum for the C/EBP1 site-specific complex with C/EBP62.

residues) is disordered (O'Neil et al., 1990; Shuman et al., 1990; Weiss et al., 1990a; Weiss, 1990). The CD spectra of C/EBP62 in the absence and presence of either specific or nonspecific DNA are shown in Figure 4. Addition of the palindromic DNA recognition sequence induces a large increase in the magnitude of $[\theta_{222}]$ to $-36\,600 \pm 1000$ deg cm^2 dmol^{-1}. This change is consistent with a conformational transition in the basic region in which the entire peptide becomes fully helical. Similarly, addition of the viral enhancer sequence to C/EBP62 induces a large conformational change ($\theta_{222} = -42\,000 \pm 2000$ deg cm^2 dmol^{-1}). Ellipticities as large as $-42\,000$ deg cm^2 dmol^{-1} have been observed for long α-helices (Johnson & Tinoco, 1972). As found for intact GCN4 (see above) and the minimally designed bZip peptide (O'Neil et al., 1990), addition of nonspecific DNA induced a slight increase in the helicity of the basic region (Figure 4).

Table I: Ammonium Sulfate Induced Increase in the Helical Content of C/EBP62

$[NH_4(SO_4)_2]$ (mM)	C/EBP62 (deg cm^2 dmol^{-1})	$[NH_4(SO_4)_2]$ (mM)	C/EBP62 (deg cm^2 dmol^{-1})
0	−11 800 (0)[a]	313.4	−23 300 (42)
10	−15 700 (14)	576.3	−29 900 (66)
88	−19 500 (28)	804.3	−36 400 (89)
166	−20 300 (31)	specific complex	−39 400[b]

[a] Percent of the maximal increase based on the difference in $[\theta_{222}]$ for the protein and the specific DNA/protein complex. [b] The mean of the values obtained for the two different DNA sites.

Interaction with Divalent Anions. The binding of leucine zipper proteins to DNA is largely electrostatic in nature. For instance, the Fos/Jun complex dissociates at concentrations of NaCl greater than 0.5 M (Turner & Tjian, 1989). We therefore examined whether high concentrations of divalent anions such as sulfate might stabilize the α-helical conformation of the basic region of C/EBP62. Indeed, the helical content of the bZip fragment increased with increasing concentrations of ammonium sulfate and approached the value observed for the specific complex near 1.0 M ammonium sulfate (Table I).

CONCLUSIONS

Previous work (O'Neil et al., 1990) showed that an α-helical conformation was induced in the basic regions of model bZip peptides when bound to their specific DNA binding sites. The data described herein show that this result is not a characteristic limited to the model peptides but is also observed for fragments of C/EBP as well as for native GCN4. Interestingly, these proteins also showed a considerably smaller induction of α-helical content when they bound nonspecifically to DNA or sulfate. Thus, it is possible to piece together a plausible series of conformational changes that accompany DNA binding. (i) In the absence of DNA, the basic regions of bZip proteins adopt flexible nonhelical or nascent helical (Saudek et al., 1986) conformations that minimize repulsive intramolecular electrostatic interactions among their positively charged side chains, and adopt long stretches of α-helices relatively rarely. (ii) As the protein approaches DNA, it is electrostatically attracted, resulting in a nonspecific interaction between the basic region and the phosphodiester backbone. In the presence of the strongly negative electrostatic potential of the DNA and the concomitantly high local ionic strength, helix formation in the basic region becomes less unfavorable and the α-helical conformation is more populated. (iii) When the recognition sequence is found, the native α-helical conformation in the basic region is stabilized by interactions with the bases and the phosphodiester backbone.

ACKNOWLEDGMENTS

We thank R. Hoess for helpful discussions; many of the experiments have been done in his lab, and we appreciate the time and expertise contributed. We also thank Steve McKnight and Tom Seitz for support and helpful discussions. We thank D. Janvier for oligonucleotide synthesis and amino acid analysis and G. Koukedis for technical assistance.

REFERENCES

Cohen, C., & Parry, D. A. D. (1990) *Proteins: Struct., Funct., Genet. 7*, 1–15.

Crick, F. H. C. (1953) *Acta Crystallogr. 6*, 689–697.

Gartenberg, M. R., Ampe, C., Steitz, T. A., & Crothers, D. M. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6034–6038.

Gentz, R., Rausher, F. J., III, Abate, C., & Curran, T. (1989) *Science 243*, 1695-1699.
Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 8*, 4108-4116.
Hope, I. A., & Struhl, R. (1987) *EMBO J. 6*, 2781-2784.
Johnson, P. F., & McKnight, S. L. (1989) *Annu. Rev. Biochem. 58*, 799-839.
Johnson, W. C., Jr. (1972) *J. Am. Chem. Soc. 94*, 4389-4390.
Johnson, W. C., Jr. (1990) *Proteins: Struct., Funct., Genet. 7*, 205-214.
Kouzarides, T., & Ziff, E. (1989) *Nature 340*, 568-571.
Landschultz, W. H., Johnson, P. F., & McKnight, S. L. (1988) *Science 240*, 1759-1764.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, pp 173-177, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Nakabeppu, Y., & Nathans, D. (1989) *EBMO J. 12*, 3833-3841.
Neuberg, M., Adamkiewics, J., Hunter, J. B., & Muller, R. (1989) *Nature 341*, 243-245.
Oakley, M. G., & Dervan, P. B. (1990) *Science 248*, 847-850.
Oas, T. G., McIntosh, L. P., O'Shea, E. K., Dahlquist, F. W., & Kim, P. S. (1990) *Biochemistry 29*, 2891-2894.
O'Neil, K. T., Hoess, R. H., & DeGrado, W. F. (1990) *Science 249*, 774-778.
O'Shea, E. K., Rutkowski, R., & Kim, P. S. (1989a) *Science 243*, 538-542.
O'Shea, E. K., Rutkowski, R., Stafford, W. F., III, & Kim, P. S. (1989b) *Science 245*, 646-648.
Rasmussen, R., Benvegnu, D., O'Shea, E. K., Kim, P. S., & Alber, T. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 561-564.
Roman, C., Platero, J. S., Shuman, J., & Calame, K. (1990) *Genes Dev. 4*, 1404-1415.
Sassone-Corsi, P., Ransone, L. J., Lamph, W. W., & Verma, I. M. (1988) *Nature 336*, 692-695.
Saudek, V., Pastore, A., Morelli, M. A. C., Frank, R., Gausepohl, H., Gibson, T., Weih, F., & Roesch, P. (1990) *Protein Eng. 4*, 3-10.
Sellers, J. W., & Struhl, K. (1989) *Nature 341*, 74-76.
Shuman, J. D., Vinson, C. R., & McKnight, S. L. (1990) *Science 249*, 771-774.
Talanian, R. V., McKnight, C. J., & Kim, P. S. (1990) *Science 249*, 769-771.
Tan, S., & Richmond, T. J. (1990) *Cell 62*, 367-377.
Turner, R., & Tjian, R. (1989) *Science 243*, 1689-1694.
Vinson, C. R., Sigler, P. B., & McKnight, S. L. (1989) *Science 246*, 911-916.
Weiss, M. A. (1990) *Biochemistry 29*, 8020-8024.
Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C., & Struhl, K. (1990) *Nature 347*, 575-578.

Refolding and Assembly of Penicillin Acylase, an Enzyme Composed of Two Polypeptide Chains That Result from Proteolytic Activation†

Christopher D. Lindsay‡ and Roger H. Pain*
Department of Biochemistry and Genetics, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, U.K.
Received April 3, 1991; Revised Manuscript Received July 10, 1991

ABSTRACT: The in vitro folding and assembly of penicillin acylase (EC 3.5.1.11) (PA) to active enzyme has been studied. PA is a large bacterial protein (M_r = 86 000) comprising two peptides, α and β, produced by proteolytic processing and activation of a 92-kDa precursor. Proteins that result from proteolytic processing are characteristically difficult if not impossible to refold. Different factors that affect folding and assembly of PA, including pH, ionic strength, and temperature, have been studied. Yields of 60% can be obtained, based on recovery of enzyme activity, together with another 20% of folded and associated monomer with conformation closely similar to that of the active enzyme but with the active site not formed. Evidence is presented for in vitro assembly proceeding via initial folding of the N-terminal α-peptide with subsequent collapse of the transiently folded β-chain on to the surface of the former. A slow process of rearrangement follows association in vitro. Competition experiments support the proposal that the linker endopeptide in the precursor serves to increase the probability of productive collision between folded α- and β-peptides. The effect of raised temperature is to interfere with the folding of the α-peptide, thus preventing proper folding of the precursor. This finding accounts for the basis of the temperature regulation of PA production in vivo.

Proteins that derive from precursors by proteolytic processing and activation do not in general unfold reversibly (Anfinsen, 1967). Cathepsin D (Lah et al., 1984) and pepsin (Ahmed & McPhie, 1978) each originate from precursors that will fold in vitro as well as in vivo, but the enzymes that have lost an N-terminal activation peptide are metastable. Once unfolded, it has not been possible to regain any enzyme activity for either enzyme. Chymotrypsin, like insulin, consists of peptide chains held together by disulfide bonds as well as noncovalent interactions. Incubation with protein disulfide isomerase leads to disulfide band exchange and ultimately inactivation and aggregation (Givol et al., 1965). These proteins, unlike their precursors, are therefore not in their lowest accessible free energy states. One of the tenets of the thermodynamic hypothesis is that the information required to specify the folding of a globular protein is contained in the primary sequence and the lack of reversibility has been ascribed tentatively to loss of information residing in the amino acid sequence (Anfinsen, 1967).

Penicillin acylase (EC 3.5.1.11) is a large (M_r = 86 000) monomeric enzyme unique among bacterial enzymes in being activated by proteolytic processing of a single polypeptide

† This work was supported under the European Community Biotechnology Action Programme (contract no. BAP 0042 UK).
‡ Present address: CDE Porton Down, Salisbury, Wilts. SP4 0JQ, UK.

0006-2960/91/0430-9034$02.50/0