

# Coupled Folding and Site-Specific Binding of the GCN4-bZIP Transcription Factor to the AP-1 and ATF/CREB DNA Sites Studied by Microcalorimetry<sup>†</sup>

Christine Berger, Ilian Jelesarov, and Hans Rudolf Bosshard\*

Biochemisches Institut der Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Received June 3, 1996; Revised Manuscript Received September 9, 1996<sup>®</sup>

**ABSTRACT:** The site-specific interaction of the basic leucine zipper protein C62GCN4, which corresponds to the C-terminal sequence 220–281 of the yeast transcription factor GCN4, with the AP-1 and ATF/CREB DNA recognition sites was analyzed by isothermal titration microcalorimetry. Free C62GCN4 is a dimer composed of a C-terminal leucine zipper and a basic, mainly unstructured DNA binding domain. Upon association with the target DNA, C62GCN4 folds to a fully  $\alpha$ -helical dimer [Ellenberger et al. (1992) *Cell* 71, 1223–1237; König and Richmond (1993) *J. Mol. Biol.* 233, 139–154]. The protein-bound AP-1 site is straight, and the protein-bound ATF/CREB site is bent by 20° toward the leucine zipper domain. The coupling between protein folding and DNA association resulting in two conformationally different complexes with C62GCN4 poses interesting thermodynamic problems. The association was strongly exothermic for both DNA target sites. The free energies of binding were indistinguishable in buffers of low salt concentration, and no change of the protonation state of C62GCN4 and/or the DNA target site occurred on formation of the complexes. Both complexes exhibited large and negative heat capacity changes. The empirical correlation between buried nonpolar and polar surfaces and the reduction in heat capacity concomitant to complexation did hold for the reaction with the AP-1 site at low salt concentration. However, in the case of the ATF/CREB site, the change in heat capacity was larger than could be accounted for by the burial of solvent-accessible surface. Potential sources of the extra decrement in the heat capacity could be restrictions in the vibrational modes of polar groups and of bound water molecules at the protein–DNA interface, thought to result from the bending of the ATF/CREB site. In the presence of high concentrations of glutamate and NaCl, the complex with the ATF/CREB site was significantly weaker than the complex with the AP-1 site.

Specific protein–DNA interactions are of fundamental importance to target genetic regulatory proteins to promoter and enhancer regions of transcribed genes. Of the different classes of proteins that control DNA transcription, the bZIP<sup>1</sup> transcription factors contain the simplest,  $\alpha$ -helix based DNA binding motif: a short region of predominantly basic amino acid residues immediately preceding a leucine zipper domain that acts as a dimerization unit. bZIP factors bind as dimers to specific DNA recognition sites and are believed to regulate the efficiency with which RNA polymerase II binds to DNA and initiates transcription (Baxeavanis and Vinson, 1993; Hurst, 1995). The transcriptional activator GCN4 is a homodimeric protein that induces amino acid synthesis in yeast (Hinnebusch, 1984; Hope and Struhl, 1985). The bZIP motif is located near the C-terminus of GCN4. The 60 C-terminal residues are sufficient for dimer formation and site-specific DNA recognition (Hope and Struhl, 1987). GCN4 recognizes the target sequence ATGACTCAT *in vivo*

(Hill et al., 1986), which is called the AP-1 site because it is recognized by the mammalian AP-1 factor (Lee et al., 1987). In addition, GCN4 also binds to the ATF/CREB site ATGACGTCAT (Hai et al., 1989; Dwarki et al., 1990). The AP-1 site has a pseudo-palindromic sequence (ATGAC/G) whereas the ATF/CREB site is a true palindrome (ATGAC). The two sites differ by a central G•C base pair between the two half-site ATGA sequences (Figure 1). The binding affinities of GCN4 for the two recognition sites are very similar (Sellers et al., 1990).

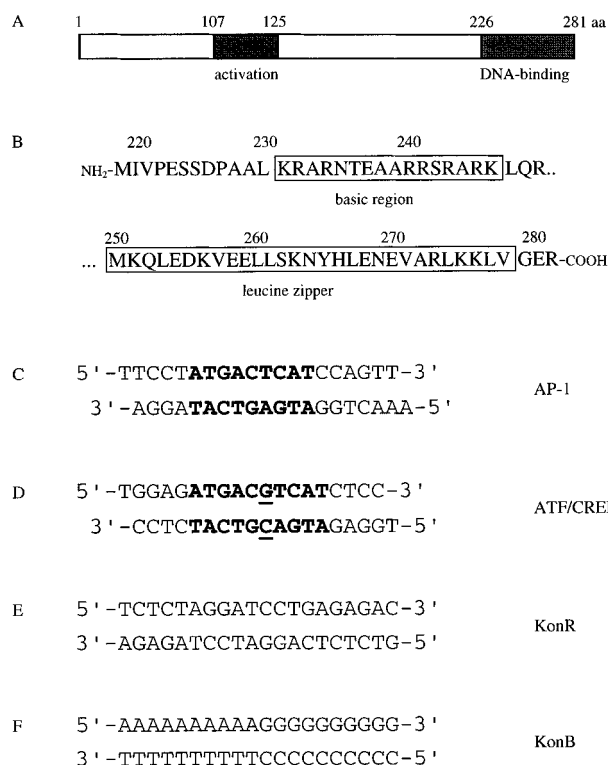
The crystal structures of the complexes of the GCN4-bZIP domain with the AP-1 and the ATF/CREB site have been solved at 3 Å resolution (Ellenberger et al., 1992; König and Richmond, 1993), and the complex with the ATF/CREB site has been refined to 2.2 Å (Keller et al., 1995). The structure of the leucine zipper in the GCN4:ATF/CREB complex is highly similar to the  $\alpha$ -helical coiled coil seen in the crystal structure of the AP-1 complex (Keller et al., 1995). In the ATF/CREB complex, the additional G•C base pair is accommodated by a 20° bend of the DNA toward the leucine zipper and a larger base pair inclination. These conformational differences in the DNA site are the principle reason why the same direct contacts between bases and side chains of GCN4 are seen in both complexes (Ellenberger, 1994). The overall conformation of ATF/CREB in the complex is B-form, but the central region shows features of A-DNA (Keller et al., 1995). In contrast, the DNA in the AP-1 complex is not bent and has little A-like character (Ellenberger et al., 1992). Whether the DNA is already bent in the free ATF/CREB site or becomes bent only on binding

<sup>†</sup> This work was supported in part by the Swiss National Science Foundation.

\* Corresponding author. Phone: +41 1 257 5540; Fax: +41 1 363 7947; E-mail: HRBOSS@BIOC.UNIZH.CH

<sup>®</sup> Abstract published in *Advance ACS Abstracts*, November 1, 1996.

<sup>1</sup> Abbreviations: AP-1, activation protein 1; ATF/CREB, activating transcription factor/cyclic AMP responsive element binding protein; bZIP, basic leucine zipper; C62GCN4, bZIP domain encompassing C-terminal residues 220–281 of GCN4; GCN4, gene control of amino acid synthesis nonderepressible mutant 4; ITC, isothermal titration calorimetry; RP-HPLC, reversed phase high performance liquid chromatography; RT, room temperature; ds, double stranded; ss, single stranded.



The coupled binding and folding reactions and the similar binding affinities for different target sequences pose interesting thermodynamic questions that need to be answered in order to understand in physical terms the macromolecular recognition process between GCN4 and its DNA target sites. What are the enthalpic and entropic changes on binding to the two target sites of which one is bent and the other is straight in the complex? How do the thermodynamic parameters, in particular the change in heat capacity ( $\Delta C_p$ ),<sup>2</sup> correlate with the stereochemical properties of the two

The ATF/CREB site DNA (see Figure 1 for DNA sequences) was synthesized on an Applied Biosystems synthesizer, using standard phosphoramidite chemistry, and purified by RP-HPLC on a Nucleosil 300-5 C8 column employing a gradient from 4% to 40% acetonitrile in 0.1 M triethylammonium acetate, pH 7.0, 1 mL/min, RT. Fractions with DNA were lyophilized, resuspended in water, and precipitated by addition of 0.1 volume of 3 M sodium acetate, pH 4.5, and 2.5 volumes of ethanol. After centrifugation (20000g, 10 min, 4 °C), the pellet was dried under a stream of nitrogen. The self-complementary oligonucleotide was annealed by heating to 85 °C and cooling to RT over a period of 2 h. The AP-1 site DNA and the control oligonucleotides KonR and KonB were custom-synthesized by Microsynth (Balgach, Switzerland), desalted on Sephadex, checked for purity by polyacrylamide gel electrophoresis, precipitated by ethanol, dried, and annealed. Concentrations of oligonucleo-

<sup>3</sup> *Handbook of Biochemistry and Molecular Biology* (1976) 3rd ed., p 187, CRC Press, Cleveland, OH.

tides were determined by absorption measurements using extinction coefficients calculated according to Brown and Brown (1991). In addition, the oligonucleotides were completely digested with venom snake nuclease, purified by RP-HPLC on a Nucleosil 300-5 C8 column, and eluted with a gradient from 0% to 20% methanol in 20 mM  $\text{KH}_2\text{PO}_4$ , pH 5.5, 1.2 mL/min, RT. Peaks were integrated to determine the relative content of the deoxyribonucleosides dA and dG (McLaughlin and Piel, 1984). DNA concentrations determined by the two methods agreed within  $\pm 10\%$ .

**Buffers.** Low salt Tris buffer: 50 mM Tris-HCl, pH 7.5, 10 mM NaCl, 10 mM  $\text{MgCl}_2$ . Low salt Hepes buffer: 160 mM Hepes-NaOH, pH 7.5, 10 mM NaCl, 10 mM  $\text{MgCl}_2$ . High glutamate buffer: 100 mM Hepes-KOH, pH 7.5, 250 mM potassium glutamate, 150 mM NaCl, 20 mM magnesium acetate, 0.5 mM EDTA.

**Native Polyacrylamide Gel Electrophoresis.** Oligonucleotides were analyzed under native conditions on a 20% polyacrylamide gel (Sambrook et al., 1989). Samples of different single strand concentration in Tris buffer were diluted 1:1 with 30% glycerol and loaded onto the gel. Gels were run for 3 h at a constant voltage of 60 V in TBE (44 mM Tris, 44 mM boric acid, 2 mM EDTA, pH 8.3). To visualize the DNA, gels were stained for 10 min in a 5  $\mu\text{g}/\text{mL}$  ethidium bromide solution, the background was destained in deionized water (10 min), and the bands were detected by UV light ( $\lambda = 302 \text{ nm}$ ). Alternatively, fluorescent labeled AP-1 and ATF/CREB were used and bands were visualized by fluorescence emission of a *N,N'*-dimethyl-*N*-(acetyl)-*N'*-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)ethylenediamine group (Turcatti et al., 1995) that was thioester-linked to the phosphodiester backbone near the 5'-end (C. Berger, unpublished).

**CD Spectroscopy.** CD spectra and thermal denaturation curves were measured on a JASCO J-500 C spectropolarimeter in a thermostated cuvette of 1 cm path length. Wavelength scans from 310 to 210 nm were performed at a constant temperature with a 1 nm bandwidth and a 2  $\text{nm min}^{-1}$  scan speed. Temperature was maintained using a JULABO F10 circulating water bath and was monitored by a PT100 temperature probe in physical contact with the DNA solution. Temperature scans were performed by scanning continuously from 6 to 94  $^{\circ}\text{C}$  at a scan rate 1  $^{\circ}\text{C min}^{-1}$ , and the ellipticity at 286 nm was measured at discrete intervals. Reversibility was checked by repeated scans. Temperature gradients were reproducible within 0.2  $^{\circ}\text{C}$ . The total ssDNA concentration was varied between 0.5 and 40  $\mu\text{M}$ . Experiments were performed in low salt Hepes buffer. Approximately 50 equally spaced data points per curve were subjected to analysis.

**Isothermal Titration Calorimetry.** ITC was performed on an OMEGA titration calorimeter (MicroCal, Inc., Northampton, MA). The instrument has been described in detail by Wiseman et al. (1989), and its principle and application have been reviewed elsewhere (Jelesarov et al., 1996). The calorimeter was calibrated with electrically generated heat pulses as recommended by the manufacturer. To improve baseline stability, the temperature of the system was kept about 5  $^{\circ}\text{C}$  below the working temperature with the help of a circulating water bath. Temperature equilibration was allowed for 10–12 h. All solutions were thoroughly degassed by evacuation. Samples of C62GCN4 and DNA were prepared with the same batch of buffer to minimize artifacts due to minor differences in buffer composition. The

reaction cell contained 1.33 mL of 10  $\mu\text{M}$  DNA<sup>4</sup> in buffer. The reference cell contained 0.02% sodium azide in water or high glutamate buffer. The injection syringe was filled with 220 or 350  $\mu\text{M}$  C62GCN4 in buffer and was rotated at 350 rpm during equilibration and experiment. Injections were started after equilibration to baseline stability (rms noise  $< 5 \text{ ncal s}^{-1}$ , baseline drift  $< 10 \text{ ncal min}^{-1}$ ).

A titration experiment consisted of 15 injections, each of 8  $\mu\text{L}$  volume and 12 s duration, with a 300 s interval between injections. The titration data were corrected for small heat changes observed in control titrations of buffer into buffer and C62GCN4 into buffer. Data analysis was carried out with the software provided with the instrument (Wiseman et al., 1989). The *c* value, defined as the concentration of DNA multiplied by the binding constant  $K_A$ , was 250–500. *c* values of this magnitude preclude a precise determination of  $K_A$  (Wiseman et al., 1989). The total heat of binding,  $\Delta H$ , and the association constant,  $K_A$ , were obtained by nonlinear least-squares fitting of the data to a 1:1 binding model (one C62GCN4 dimer per one DNA double strand) or a 2:1 binding model (two C62GCN4 dimers per one DNA double strand) utilizing the Marquardt algorithm. Stoichiometries of the fits to the 1:1 binding model agreed to within  $\pm 15\%$ .

The association constant of the AP-1 dsDNA was estimated by titration of the 5'-strand to the 3'-strand, or vice versa. The concentration of ssDNA was 25  $\mu\text{M}$  in the syringe and 1  $\mu\text{M}$  in the sample cell. Because of the high stability of ds AP-1, only a lower limit for  $K_A$  could be estimated (*c* value  $> 1000$ ).

**Molecular Modeling and Surface Area Calculations.** Coordinates of the crystal structure of the GCN4-bZIP domain complexed with the AP-1 site were obtained from the Brookhaven Protein Data Base (entry 1YSA). Coordinates of the GCN4-bZIP domain complexed with the ATF/CREB site were kindly provided by Dr. W. Keller (file 1DGC). Some partly disordered amino acid residues were deleted to obtain comparable structures corresponding to the sequence Pro<sup>227</sup>–Glu<sup>280</sup> of GCN4. Modeling was performed with the program Insight II, version 2.3.5, from Biosym, San Diego. Solvent-accessible surface area (ASA) was calculated according to the Lee-Richards algorithm (probe radius 1.4 Å) using the program QUANTA, version 4.0, from Molecular Simulation, Inc. The net change in ASA on complexation was calculated as

$$\Delta\text{ASA} = \text{ASA}_{\text{complex}} - (\text{ASA}_{\text{C62GCN4}} + \text{ASA}_{\text{DNA}})$$

The sequence Pro<sup>227</sup>–Arg<sup>249</sup>, which is fully helical in the complex, was modeled in three ways to calculate  $\text{ASA}_{\text{C62GCN4}}$ . In model E, the polypeptide chain was in the fully extended conformation. In model N, a nascent helix was simulated by introduction of two turns of  $\alpha$ -helix separated by extended chain. The average ASA from 5 randomly chosen conformations with two sequentially separated helical turns was used in the calculation of  $\Delta\text{ASA}$  for model N. In model R, the conformation of C62GCN4 was the same as in the complex.  $\text{ASA}_{\text{DNA}}$  was calculated assuming the B-form DNA conformation.

<sup>4</sup> The indicated concentrations refer always to double stranded DNA and dimeric C62GCN4 unless stated otherwise.

Table 1: Thermodynamic Parameters of Binding of C62GCN4 to Different DNA Sites in Low Salt Tris Buffer<sup>a</sup>

<i>T</i> (K)	<i>K</i> <sub>A</sub> × 10 <sup>-7</sup> (L mol <sup>-1</sup> )	Δ <i>G</i> (kJ mol <sup>-1</sup> ) <sup>b</sup>	Δ <i>H</i> (kJ mol <sup>-1</sup> )	<i>T</i> Δ <i>S</i> (kJ mol <sup>-1</sup> ) <sup>c</sup>
C62GCN4 with DNA Site AP-1				
282	5.50 ± 1.30	-41.78 ± 1.06	-117.30 ± 1.45	-75.51 ± 2.51
285	5.00 ± 2.20	-42.01 ± 1.33	-126.15 ± 2.50	-84.14 ± 3.83
292	5.42 ± 0.92	-43.84 ± 0.45	-137.12 ± 0.85	-93.89 ± 1.30
298	2.85 ± 0.72	-42.52 ± 0.89	-153.83 ± 2.02	-111.31 ± 2.91
C62GCN4 with DNA Site ATF/CREB				
282	15.80 ± 11.0	-44.28 ± 2.91	-70.01 ± 0.95	-25.74 ± 3.86
285	11.60 ± 10.0	-44.00 ± 4.70	-73.65 ± 1.84	-29.65 ± 6.54
286	14.90 ± 9.80	-44.79 ± 2.59	-81.22 ± 1.11	-36.43 ± 3.70
292	9.18 ± 6.50	-44.51 ± 5.77	-101.37 ± 3.01	-56.86 ± 8.78
298	3.70 ± 1.50	-43.18 ± 0.94	-117.09 ± 2.01	-73.91 ± 2.95
C62GCN4 with Control DNA KonR				
282	0.24 ± 0.05	-34.39 ± 0.52	-70.63 ± 1.56	-36.24 ± 2.08
290	0.43 ± 0.08	-36.84 ± 0.51	-67.41 ± 1.24	-30.57 ± 1.75
298	0.09 ± 0.03	-33.98 ± 1.0	-63.43 ± 3.68	-29.45 ± 4.68

<sup>a</sup> Numbers are mean ± maximum experimental error. Standard deviations of the mean were always lower than the maximum experimental error. <sup>b</sup> Calculated from *K*<sub>A</sub>: Δ*G* = -RT ln *K*<sub>A</sub>. <sup>c</sup> Obtained from the equation: *T*Δ*S* = Δ*H* - Δ*G*.

Table 2: Thermodynamic Parameters of Binding of C62GCN4 to Different DNA Sites in High Glutamate Buffer<sup>a</sup>

<i>T</i> (K)	<i>K</i> <sub>A</sub> × 10 <sup>-7</sup> (L mol <sup>-1</sup> )	Δ <i>G</i> (kJ mol <sup>-1</sup> ) <sup>b</sup>	Δ <i>H</i> (kJ mol <sup>-1</sup> )	<i>T</i> Δ <i>S</i> (kJ mol <sup>-1</sup> ) <sup>c</sup>
C62GCN4 with DNA Site AP-1				
282	1.61 ± 0.20	-38.91 ± 0.31	-110.16 ± 0.87	-71.25 ± 1.18
290	1.14 ± 0.14	-39.18 ± 0.32	-131.26 ± 1.19	-92.08 ± 1.51
298	0.31 ± 0.04	-37.01 ± 0.59	-138.28 ± 2.93	-101.28 ± 3.52
C62GCN4 with DNA Site ATF/CREB				
282	0.09 ± 0.01	-32.22 ± 0.32	-70.47 ± 1.42	-38.25 ± 1.74
290	0.07 ± 0.02	-32.56 ± 0.81	-72.94 ± 4.52	-40.38 ± 5.33
298	0.03 ± 0.01	-30.91 ± 1.58	-81.02 ± 12.94	-50.11 ± 14.52

<sup>a</sup> Numbers are mean ± maximum experimental error. Standard deviations of the mean were always lower than the maximum experimental error. <sup>b</sup> Calculated from *K*<sub>A</sub>: Δ*G* = -RT ln *K*<sub>A</sub>. <sup>c</sup> Obtained from the equation: *T*Δ*S* = Δ*H* - Δ*G*.

## RESULTS

ITC allows for the simultaneous and direct determination in the same experiment of the binding enthalpy (Δ*H*), the association constant (*K*<sub>A</sub>), and the stoichiometry (*N*) of a binding reaction (Wiseman et al., 1989). Values of *K*<sub>A</sub>, Δ*G*, Δ*H*, and *T*Δ*S* for the association of C62GCN4 with the AP-1 and ATF/CREB site measured in low salt Tris buffer and in high glutamate buffer are listed in Tables 1 and 2, respectively. Titrations were performed in the range 9–25 °C. This small temperature range was chosen because the basic segment of free C62GCN4 shows very little changes in structure below 25 °C (Weiss, 1990). Moreover, this temperature range had to be chosen because the leucine zipper domain of C62GCN4 begins to melt above 30 °C under conditions of the ITC experiments. Such additional conformational changes in free C62GCN4 would have complicated the combined folding and binding reaction we wished to analyze by calorimetric titration.

During an ITC experiment, C62GCN4 is diluted on addition to the reaction cell. This may lead to partial dissociation into monomers, an endothermic reaction (Thompson et al., 1993). The dissociation constant of the C62GCN4 dimer is 10<sup>-7</sup>–10<sup>-8</sup> M (Krylov et al., 1995; Zitzewitz et al., 1995; Wendt et al., 1994). Already after

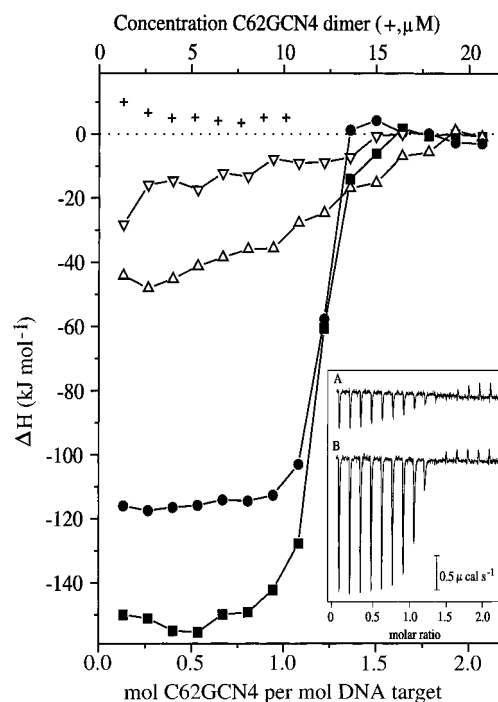
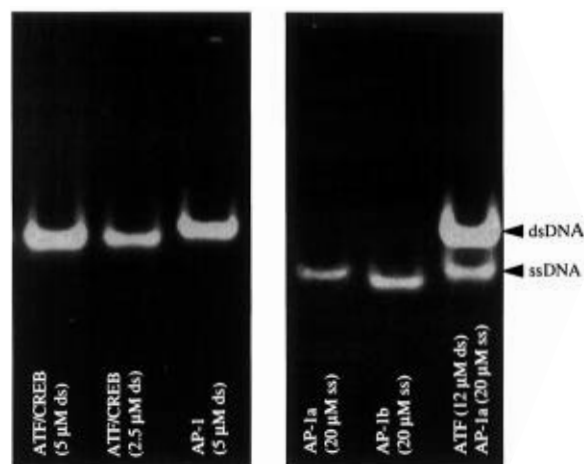


FIGURE 2: Isothermal titration of AP-1 (■), ATF/CREB (●), KonR (Δ), and KonB (▽) with C62GCN4. 10 μM DNA was titrated with 220 μM C62GCN4 added in 15 aliquots of 8 μL each, in low salt Tris buffer, 25 °C. Control titration of C62GCN4 into buffer (+, upper abscissa). Inset: raw data for titration of AP-1 DNA (trace B) and KonR DNA (trace A) with C62GCN4. Conditions as for main figure.

addition of the first aliquot to the reaction cell, the concentration of C62GCN4 was very much higher than *K*<sub>d</sub> and the fraction of monomer must have been very small throughout the course of a titration experiment. This was supported by the control titration of C62GCN4 into buffer, which exhibited small endothermic peaks that were of equal magnitude within the limit of errors (Figure 2). The heat of dissociation of C56GCN4, corresponding to C62GCN4 truncated by 6 N-terminal residues, is about 100 kJ mol<sup>-1</sup> at room temperature (Thompson et al., 1993). Hence, the fraction of dimer dissociating on addition to the reaction cell must have been very small and hidden underneath the peaks seen in the control titration. In addition, the helix content of the free C62GCN4 was determined under conditions of the ITC experiment as described previously for other leucine zippers (Wendt et al., 1995). These control experiments revealed that the helical content of the free C62GCN4 was larger than 95% at the concentrations and in the buffers used in the ITC measurements (CD data not shown).

The ATF/CREB site—but not the AP-1 site—is completely palindromic and may, therefore, have the potential to form hairpins (Figure 1). To control the double stranded nature of the ATF/CREB and AP-1 oligonucleotides, the experiments of Figure 3 were performed. In polyacrylamide gel electrophoresis, the 19-mer ATF/CREB had essentially the same mobility as the 20-mer AP-1, and both moved well behind the 3'- and 5'-single strands of AP-1 (Figure 3A). In the thermal unfolding experiments, the two oligomers had about the same midpoint temperature of unfolding, which was concentration-dependent in accord with a transition from dsDNA to ssDNA (Figure 3B). Furthermore, a lower limit of 2 × 10<sup>-9</sup> M for *K*<sub>d</sub> of AP-1 dsDNA was estimated by ITC. In conclusion, both ATF/CREB and AP-1 were double stranded under the conditions of ITC experiment.

A



B

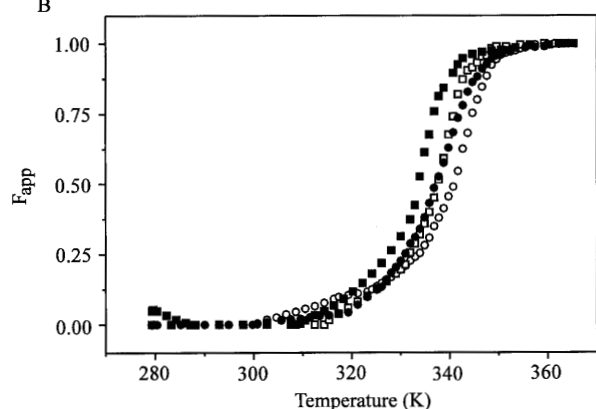


FIGURE 3: Testing the double strand nature of AP-1 and ATF/CREB target DNA. (A) Native polyacrylamide gel electrophoresis of ATF/CREB, AP-1, the 5'-strand (AP-1a), and the 3'-strand (AP-1b) of AP-1. (B) Temperature dependence of the apparent fraction of unfolded DNA in low salt Hepes buffer. AP-1: 0.5  $\mu\text{M}$  (■), 20  $\mu\text{M}$  (□). ATF/CREB: 0.6  $\mu\text{M}$  (●), 16  $\mu\text{M}$  (○). The apparent fraction of ssDNA was calculated from  $F_{\text{app}} = (\theta - \theta_N) / (\theta_U - \theta_N)$ , where  $\theta_N$  and  $\theta_U$  are the average ellipticity in the range 10–30 and 80–90 °C, respectively.

**Free Energy of Binding and Stoichiometry of C62GCN4-DNA Complexes.** All the titration data could be fitted best for a stoichiometry of one dimeric C62GCN4 per double stranded DNA site. Examples are shown in Figure 2. In the case of the ATF/CREB site, but not the AP-1 site, "overtitration" with a large excess of C62GCN4 indicated the possibility of a second binding reaction which, however, was much weaker and slightly endothermic (not shown). The significance of this observation was unclear and it was not studied further. Incidentally, a primary strong and exothermic site and a secondary weak and endothermic site has been described for the *trp* repressor–operator interaction (Ladbury et al., 1994).

In low salt Tris buffer, the association constants of the complexes of C62GCN4 with the AP-1 and the ATF/CREB site were indistinguishable within the limit of errors in the temperature range 9–25 °C (Table 1). At 25 °C,  $K_A$  was  $(2.85 \pm 0.72) \times 10^7 \text{ M}^{-1}$  for the AP-1 site and  $(3.70 \pm 1.50) \times 10^7 \text{ M}^{-1}$  for the ATF/CREB site. Using a gel retardation assay, Weiss et al. (1990) had reported  $K_A = 5 \times 10^7 \text{ M}^{-1}$  for the complex of GCN4 peptide 225–281 with the AP-1 site. Very similar association constants had been obtained

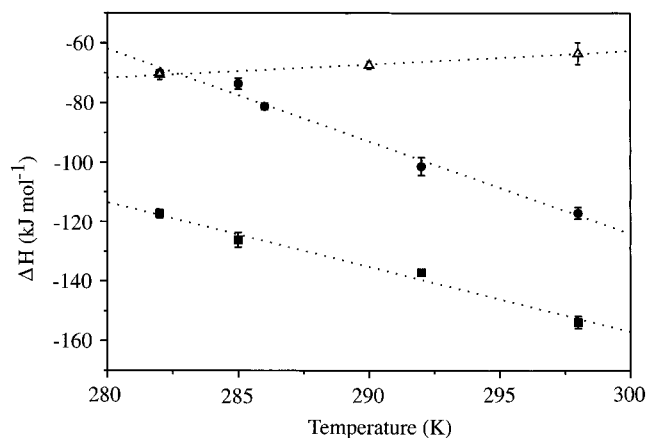


FIGURE 4: Plot of  $\Delta H$  against temperature for the titrations of AP-1 (■), ATF/CREB (●), and KonR (Δ) with C62GCN4 in Tris buffer.  $\Delta C_p$  was calculated by linear regression.

also for the ATF/CREB site (Sellers et al., 1990; König and Richmond, 1993).

In the high glutamate buffer,  $K_A$  of both complexes was lower, but now the C62GCN4:ATF/CREB complex was 15 times less stable than the C62GCN4:AP-1 complex (Table 2).

Control titrations were performed with the oligonucleotides KonR and KonB. KonR had the base composition of the AP-1 site arranged in a computer-randomized sequence. KonB consisted of blocks of ten G·C pairs and ten A·T pairs (Figure 1). Binding of KonR in low salt Tris buffer and high glutamate buffer occurred with  $K_A$  lower than  $10^6 \text{ M}^{-1}$ , a range typical of an unspecific electrostatic protein–DNA interaction (von Hippel and Berg, 1986). KonB was also bound unspecifically in either of the buffers. This agreed with previous work where it had been reported that the presence of high concentrations of salt, including 250 mM glutamate, the dominant intracellular anion in *E. coli* and some eucaryotic cells, optimizes the specific affinity and maximizes the difference between specific and nonspecific binding (Leirimo et al., 1987).

**Enthalpy, Entropy, and Heat Capacity Changes of Association Reactions.** Binding reactions were very strongly exothermic, and  $\Delta H$  for the AP-1 site was more negative than  $\Delta H$  for the ATF/CREB site in both buffer systems (Tables 1 and 2 and Figure 2). To test if the binding reaction was accompanied by protonation/deprotonation, ITC experiments were performed in low salt Hepes buffer. Values of  $\Delta H$  were indistinguishable in Tris and Hepes buffer. The heat of ionization of Tris buffer is  $47.4 \text{ kJ mol}^{-1}$  (Christensen et al., 1976) and of Hepes buffer  $21.0 \text{ kJ mol}^{-1}$  (Morin and Freire, 1991). If a significant change in the state of protonation of C62GCN4 or/and target DNA had accompanied the formation of the complex, this would have led to the observation of apparently different  $\Delta H$  values in the two buffers. Therefore,  $\Delta H$  in Tables 1 and 2 seem to represent the true binding enthalpy. No linkage of proton binding to DNA–protein interaction has been observed also for the TBP-E4 promoter complex (Petri et al., 1995).

$\Delta H$  measured by ITC and the derived  $T\Delta S$  showed a strong temperature dependence, and the values largely compensated to give  $\Delta G$  values that were almost independent of temperature (Tables 1 and 2). Heat capacity changes were calculated assuming that  $\Delta C_p$  was temperature-independent in the narrow range of 9–25 °C. The change of  $\Delta H$  with temperature in low salt Tris buffer is shown in Figure 4.

Linear regression of the data of Figure 4 gave  $\Delta C_p = -2.17 \pm 0.19 \text{ kJ mol}^{-1} \text{ K}^{-1}$  for the AP-1 site and  $\Delta C_p = -3.11 \pm 0.22 \text{ kJ mol}^{-1} \text{ K}^{-1}$  for the ATF/CREB site.  $\Delta C_p$  for the control oligonucleotide KonR was near zero or slightly positive. Small, slightly positive  $\Delta C_p$  values were observed also for the binding of *trp* repressor and *cro* protein to control nonoperator DNA segments (Takeda et al., 1992; Ladbury et al., 1994). In high glutamate buffer, the  $\Delta H$  values for the AP-1 site were only marginally lower and so was  $\Delta C_p$ ,  $-1.76 \pm 0.51 \text{ kJ mol}^{-1} \text{ K}^{-1}$  as compared to  $-2.17 \pm 0.19 \text{ kJ mol}^{-1} \text{ K}^{-1}$ . In stark contrast,  $\Delta C_p$  for the ATF/CREB site was much decreased in high glutamate buffer:  $-0.66 \pm 0.21 \text{ kJ mol}^{-1} \text{ K}^{-1}$  as compared to  $-3.11 \pm 0.22 \text{ kJ mol}^{-1} \text{ K}^{-1}$ .

## DISCUSSION

The results from the ITC experiments of the binding of C62GCN4 to the DNA target sites can be summarized as follows. (i) Complexation of C62GCN4 with the AP-1 and ATF/CREB sites was accompanied by a favorable net change in enthalpy and an unfavorable net change in entropy at all temperatures studied.  $\Delta H$  was more negative for the AP-1 site than for the ATF/CREB site. (ii) By changing the buffer from low salt to high glutamate concentration,  $K_A$  was reduced very significantly for the ATF/CREB site but not for the AP-1 site. (iii) At low salt concentration, the heat capacity changes deduced from the temperature dependence of  $\Delta H$  were large and negative;  $\Delta C_p$  was more negative for the ATF/CREB site than for the AP-1 site.

*Association of the bZIP Domain of GCN4 Is Exothermic.* Strong enthalpy/entropy compensation and large negative heat capacity changes resulting in a strong temperature dependence of  $\Delta H$  and  $\Delta S$ , but not of  $\Delta G$ , are commonly observed in protein folding and in many biological association reactions (reviews by Murphy and Freire, 1992; Spolar and Record, 1994). Therefore, there exists a temperature  $t_H$  where  $\Delta H$  changes sign.  $t_H$  could be calculated assuming that  $\Delta C_p$  is temperature-independent over a larger temperature interval than used in our measurements. Under both buffer condition,  $t_H$  values of AP-1 and ATF/CREB, respectively, were much below 0 °C. Thus, the association of the bZIP domain of GCN4 with its two DNA sites was strongly exothermic over the entire physiological temperature range. In contrast, most other protein–DNA interactions studied so far were endothermic in the range of physiological temperatures (Ha et al., 1989; Ladbury et al., 1994; Foguel and Silva, 1994; Petri et al., 1995), with only few exceptions (Merabet and Ackers, 1995).

*The Magnitude of  $\Delta C_p$  and the Amount of Water-Accessible Surface Buried in the Protein–DNA Complex.* A large negative heat capacity change is thought to be a manifestation of specific recognition based on stereospecific macromolecular interactions between large and highly complementary surfaces (Murphy and Freire, 1992). The magnitude of  $\Delta C_p$  can be correlated with the reduction of solvent-accessible nonpolar and polar surfaces in the macromolecular complex (Gill et al., 1985; Privalov and Gill, 1988; Spolar et al., 1989; Livingstone et al., 1991; Ladbury et al., 1994). Spolar et al. (1992) reported an empirical correlation between  $\Delta C_p$  and the changes in water-accessible nonpolar and polar surface area for protein–protein and protein–DNA complexes. The parameters of their equation seem to describe the protein–DNA interaction better than the parameters

established by Murphy and Freire (1992), which are more suitable for calculation of area changes in protein folding/unfolding processes. This empirical correlation of  $\Delta C_p$  with structure changes has a large uncertainty. Also the calculation of the surface area has a significant error which depends on the quality of the coordinates of the X-ray structure and the choice of the probe radius.

To calculate  $\Delta C_p$ , we determined the changes in buried nonpolar and polar surface areas from the crystal structures of the complexes and the individual, unbound components. The three-dimensional structures of free C62GCN4 and of the free AP-1 and ATF/CREB sites are not precisely known. We assumed the free AP-1 and ATF/CREB sites to be straight double stranded DNA in the canonical B-form. In the free bZIP domain of GCN4, only the leucine zipper is tightly folded whereas the basic DNA binding domain is largely unfolded but exhibits populations of transient (nascent)  $\alpha$ -helical conformations (Weiss et al., 1990; Weiss, 1990; Saudek et al., 1991; O'Neil et al., 1991). On binding to the target DNA, the DNA binding domain folds from a nascent helix into a full  $\alpha$ -helix (Ellenberger et al., 1992; König and Richmond, 1993; Keller et al., 1995). To model a nascent helix, two sequentially separated helical turns were introduced into the segment Pro<sup>227</sup>–Arg<sup>249</sup> (Figure 1B) and the rest of the segment was left in an extended conformation. The surface calculation assuming this nascent helix conformation was called model N (for nascent helix) and was based on the average conformation of five randomly chosen two-turn structures. Surface calculations were also performed for two extreme models. In model R (for rigid body), the unreasonable assumption was made that the free protein had the same conformation as the protein bound to the DNA target site, i.e., that the reaction conformed to the association of rigid bodies and no association-coupled folding took place. At the other extreme, model E (for extended) assumed a fully extended conformation of the segment Pro<sup>227</sup>–Arg<sup>249</sup>. In model E, the maximum burial of solvent-accessible surface had to occur.

Table 3 summarizes the calculated surface changes, and Table 4 compares calculated  $\Delta C_p$  with  $\Delta C_p$  from experiment. The calculation based on model N, which is thought to best mimic the mechanism of association-coupled folding, correlated reasonably well with the experimental  $\Delta C_p$  for the reaction with the AP-1 site:  $\Delta C_p(\text{calculated}) = -2.21 \text{ kJ mol}^{-1} \text{ K}^{-1}$  versus  $\Delta C_p(\text{measured}) = -2.17 \text{ kJ mol}^{-1} \text{ K}^{-1}$ . However, in the case of the ATF/CREB site, even the assumption of folding of a fully extended chain (model E) gave a less negative value of  $\Delta C_p$  than was obtained by experiment:  $\Delta C_p(\text{calculated}) = -2.84 \text{ kJ mol}^{-1} \text{ K}^{-1}$  versus  $\Delta C_p(\text{measured}) = -3.11 \text{ kJ mol}^{-1} \text{ K}^{-1}$ . Thus, the empirical relationship underestimated  $\Delta C_p$  for the ATF/CREB site. We believe that the bent, and hence strained, DNA conformation in the ATF/CREB complex can cause restrictions in vibrational modes of hydrated polar groups at and near to the complex interface. Although the X-ray structure does not reveal to what extent the bending of the DNA at the specific interface decreases the degrees of freedom of the complementary surfaces, it has been argued that unique and precise complementarity, as it is required for correct and tight protein–DNA interactions, imposes severe restrictions on the dynamic flexibility of the interactive surfaces (Ladbury et al., 1994; Cooper et al., 1994). Upon complex formation, the dynamic fluctuations of the interfacial nucleotide bases of the solvent-exposed backbone elements and of surface-

Table 3: Calculation of Solvent-Accessible Surface Areas in C62GCN4:DNA Complexes<sup>a</sup>

	AP-1 site			ATF/CREB site		
	total accessible area (Å <sup>2</sup> )	nonpolar contribution (Å <sup>2</sup> )	polar contribution (Å <sup>2</sup> )	total accessible area (Å <sup>2</sup> )	nonpolar contribution (Å <sup>2</sup> )	polar contribution (Å <sup>2</sup> )
Model R (Rigid Body)						
C62GCN4:DNA complex	14680	7940	6740	14110	7550	6560
C62GCN4 in complex	9240	5820	3420	9500	5980	3520
DNA in complex	7770	3720	4050	7330	3440	3890
net change (complex – components)	–2330	–1600	–730	–2720	–1870	–850
Model E (Extended Chain)						
C62GCN4:DNA complex	14680	7940	6740	14110	7550	6560
C62GCN4 free	11630	7300	4330	11770	7230	4540
DNA free	7460	3420	4040	7120	3250	3870
net change (complex – components)	–4410	–2780	–1630	–4780	–2930	–1850
Model N (Nascent Helix)						
C62GCN4:DNA complex	14680	7940	6740	14110	7550	6560
C62GCN4 free with helices	10780	6750	4030	10880	6730	4150
DNA free	7460	3420	4040	7120	3250	3870
net change (complex – components)	–3560	–2230	–1330	–3890	–2430	–1460

<sup>a</sup> See text for explanation of models R, E, and N; see Materials and Methods for details of calculations.

Table 4: Experimentally Measured and Calculated  $\Delta C_p$ <sup>a</sup>

	exptl	model N <sup>d</sup>	model R <sup>d</sup>	model E <sup>d</sup>
AP-1 site	–2.17 ± 0.19 <sup>b</sup> –1.76 ± 0.51 <sup>c</sup>	–2.21 ± 0.43	–1.72 ± 0.29	–2.77 ± 0.54
ATF/CREB site	–3.11 ± 0.22 <sup>b</sup> –0.66 ± 0.21 <sup>c</sup>	–2.38 ± 0.43	–2.00 ± 0.34	–2.84 ± 0.58

<sup>a</sup> Values are in kJ mol<sup>–1</sup> K<sup>–1</sup>. <sup>b</sup> Low salt Tris buffer. <sup>c</sup> High glutamate buffer. <sup>d</sup> Calculated according to the equation:  $\Delta C_p$  (cal mol<sup>–1</sup> K<sup>–1</sup>) = (0.32 ± 0.04) $\Delta A_{np}$  – (0.14 ± 0.04) $\Delta A_p$ ;  $\Delta A_{np}$  is the nonpolar and  $\Delta A_p$  the polar surface area (Spolar et al., 1992). Surface area values were used from Table 3.

bound water may be restricted. This decrease of the internal freedom of motion of the DNA binding site could have further reduced the heat capacity change (Sturtevant, 1977; Ladbury et al., 1994; Cooper et al., 1994; Petri et al., 1995; Merabet and Ackers, 1995). It should be noted, however, that the importance of vibrational contributions to  $\Delta C_p$  is subject to debate (Ha et al., 1989; Ladbury et al., 1994; Spolar and Record, 1994). The AP-1 recognition site is not bent and does not seem to adopt a tense conformation, which may explain why the heat capacity change for the AP-1 site could be quantitatively described in terms of contributions from changes in water-accessible surface area.

**Conformational Adaptation of the ATF/CREB Site in the Protein–DNA Complex.** High concentrations of glutamate weakened only the complex with the ATF/CREB site to a significant extent. Furthermore,  $\Delta H$  for the association with the ATF/CREB site was consistently less favorable than for the AP-1 site. A comparison of both crystal structures reveals differences in the basic region conformation on the AP-1 site versus the ATF/CREB site. The basic region helices are splayed further apart on the ATF/CREB site. DNA contacts involving invariant Arg<sup>243</sup> of both subunits are different in the two complexes. The changes in protein conformation help accommodate the ~3.4 Å increase in half-site separation and ~35° rotation of half-sites resulting from the additional central base pair in the ATF/CREB sequence (Kim and Struhl, 1995). The bend in the ATF/CREB DNA and underwinding at the center of the site further accommodate the ATF/CREB half-site spacing (König and Richmond, 1993). The different complex stability and the smaller  $\Delta H$  of the association reaction may have been due to the conformational differences in the basic region helices and the DNA site.

An analysis of several crystal structures of protein–DNA complexes has revealed that, in some complexes, the DNA major groove becomes narrower or wider so that it fits the protein secondary structure. Overall, a conformation between A- and B-DNA is typical of protein–DNA complexes (Nekludova and Pabo, 1994; Suzuki and Yagi, 1996). Compared to the protein-bound AP-1 site, the bound ATF/CREB site has greater A-like character as measured by the base pair tilt/inclination values (König et al., 1995). Moreover, the ATF/CREB site is bent by 20° toward the leucine zipper domain. The sequence of the ATF/CREB site shows no intrinsic propensity for bending (Goodsell and Dickerson, 1994), and also a pre-existing A-like conformation is unlikely. An induced bending of the DNA site upon protein binding is energetically unfavorable and may have contributed to a less negative  $\Delta H$  of association (Crothers, 1994; Strauss and Maher, 1994). Furthermore, this distortion of the target site causes an unstacking of the nucleotide bases which may be accompanied by a rather large unfavorable enthalpy change (Petri et al., 1995). One may speculate that, in the low salt buffer, the deformation of the ATF/CREB site was assisted by electrostatic interactions between the positively charged protein residues involved in binding and the negatively charged DNA (Matthew and Ohlendorf, 1985; Erie et al., 1994). This is to say that electrostatic energy may have been used to strain the DNA.

## CONCLUSIONS

The differences seen in the crystal structures of the GCN4:AP-1 and GCN4:ATF/CREB complexes are paralleled by differences in the thermodynamics of the association reactions. Whereas the observed entropy and heat capacity changes of the GCN4:AP-1 complex can be accounted for

by the association-coupled folding of the DNA-binding domain of GCN4 into an  $\alpha$ -helix, the thermodynamics of the GCN4:ATF/CREB complex indicate the occurrence of additional structural events. The origin of the larger than expected heat capacity changes, observed here for the GCN4:ATF/CREB complex and by others for protein—DNA complexes (Ladbury et al., 1994; Petri et al., 1995), remains to be based on a sound theoretical foundation. In particular, the thermodynamic consequences of DNA bending, observed in an increasing number of cases and perhaps characteristic of specific protein—DNA recognition (Paoletta et al., 1994; Pellegrini et al., 1995), deserve further study.

## ACKNOWLEDGMENT

We are very grateful to Drs. Walter Keller and Timothy J. Richmond for providing the GCN4 expression vector system and the *E. coli* strains, and to Dr. Walter Keller for help with protein purification. We thank Dr. Amedeo Caflisch for help with the surface area calculation using the program QUANTA, and Dr. Tom Ellenberger for reading the manuscript and for very helpful comments.

## REFERENCES

- Baxevanis, A. D., and Vinson, C. R. (1993) *Curr. Opin. Genet. Dev.* 3, 278–285.
- Brown, T., and Brown, D. J. S. (1991) in *Oligonucleotides and Analogues: A Practical Approach* (Eckstein, F., Ed.) pp 1–24, IRL Press, Oxford.
- Christensen, J. J., Hansen, L. D., and Izatt, R. M. (1976) *Handbook of proton ionization heats and related thermodynamic quantities*, Wiley, New York.
- Cooper, A., McAlpine, A., and Stockley, P. G. (1994) *FEBS Lett.* 348, 41–45.
- Crothers, D. M. (1994) *Science* 266, 1819–1820.
- Dill, K. A. (1990) *Biochemistry* 29, 7133–7155.
- Dwarki, V. J., Montminy, M., and Verma, J. M. (1990) *EMBO J.* 9, 225–232.
- Ellenberger, T. (1994) *Curr. Opin. Struct. Biol.* 4, 12–21.
- Ellenberger, T. E., Brandl, C. J., Struhl, K., and Harrison, S. C. (1992) *Cell* 71, 1223–1237.
- Erie, D. A., Yang, G., Schultz, H. C., and Bustamante C. (1994) *Science* 266, 1562–1566.
- Foguel, D., and Silva, J. L. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 8244–8247.
- Gill, S. J., Dec, S. F., Olofsson, G., and Wadsö, I. (1985) *J. Phys. Chem.* 89, 3758–3761.
- Goodsell, D. S., and Dickerson, R. E. (1994) *Nucleic Acids Res.* 22, 5497–5503.
- Ha, J.-H., Spolar, R. S., and Record, M. T., Jr. (1989) *J. Mol. Biol.* 209, 801–816.
- Hai, T. W., Liu, F., Coukos, W. J., and Green, M. R. (1989) *Genes Dev.* 3, 2083–2090.
- Hill, D. E., Hope, I. A., Macke, J. P., and Struhl, K. (1986) *Science* 234, 451–457.
- Hinnebusch, A. G. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 6442–6446.
- Hope, I. A., and Struhl, K. (1985) *Cell* 43, 177–188.
- Hope, I. A., and Struhl, K. (1987) *EMBO J.* 6, 2781–2784.
- Hurst, H. C. (1995) in *Protein Profile* (Shetlerline, P., Ed.) Vol. 2, pp 105–168, Academic Press, London.
- Jelesarov, I., Leder, L., and Bosshard, H. R. (1996), *Methods: A Companion to Methods in Enzymology* 9, 533–541.
- Keller, W., König, P., and Richmond, T. J. (1995) *J. Mol. Biol.* 254, 657–667.
- Kim, J., and Struhl, K. (1995) *Nucleic Acids Res.* 23, 2531–2537.
- König, P., and Richmond, T. J. (1993) *J. Mol. Biol.* 233, 139–154.
- Krylov, K., Olive, M., and Vinson, C. (1995) *EMBO J.* 14, 5329–5337.
- Ladbury, J. E., Wright, J. G., Sturtevant, J. M., and Sigler, P. B. (1994) *J. Mol. Biol.* 238, 669–681.
- Lee, W., Mitchell, P., and Tjian, R. (1987) *Cell* 49, 741–752.
- Leirmo, S., Harrison, C., Cayley, D. S., Burgess, R. R., and Record, M. T., Jr. (1987) *Biochemistry* 26, 2095–2101.
- Livingstone, J. R., Spolar, R. S., and Record, M. T., Jr. (1991) *Biochemistry* 30, 4237–4244.
- Matthew, J. B., and Ohlendorf, D. H. (1985) *J. Biol. Chem.* 260, 5860–5862.
- McLaughlin, L. W., and Piel, N. (1984) in *Oligonucleotide Synthesis: A Practical Approach* (Gait, M. J., Eds.) Chapter 5, IRL Press, Oxford.
- Merabet, P. E., and Ackers, G. K. (1995) *Biochemistry* 34, 8554–8563.
- Morin, P. E., and Freire, E. (1991) *Biochemistry* 30, 8494–8500.
- Murphy, K. P., and Freire, E. (1992) *Adv. Protein Chem.* 43, 313–361.
- Nekludova, L., and Pabo, C. O. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 6948–6952.
- O'Neil, K. T., Shuman, J. D., Ampe, C., and DeGrado, W. F. (1991) *Biochemistry* 30, 9030–9034.
- Paoletta, D. N., Palmer, C. R., and Schepartz, A. (1994) *Science* 264, 1130–1133.
- Pellegrini, L., Tan, S., and Richmond, T. J. (1995) *Nature* 376, 490–498.
- Petri, V., Hsieh, M., and Brenowitz, M. (1995) *Biochemistry* 34, 9977–9984.
- Privalov, P. L., and Gill, S. J. (1988) *Adv. Protein Chem.* 39, 191–234.
- Ross, P. D., Howard, F. B., and Lewis, M. S. (1991) *Biochemistry* 30, 6269–6275.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual* (Ford, N., Ed.), Chapter 6, CSH Laboratory Press, New York.
- Sauadek, V., Pasley, H. S., Gibson, T., Gausepohl, H., Frank, R., and Pastore, A. (1991) *Biochemistry* 30, 1310–1317.
- Sellers, J. W., Vincent, A. C., and Struhl, K. (1990) *Mol. Cell. Biol.* 10, 5077–5086.
- Spolar, R. S., and Record, M. T., Jr. (1994) *Science* 263, 777–784.
- Spolar, R. S., Ha, J., and Record, M. T., Jr. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 8382–8385.
- Spolar, R. S., Livingstone, J. R., and Record, M. T., Jr. (1992) *Biochemistry* 31, 3947–3955.
- Strauss, J. K., and Maher, L. J. III (1994) *Science* 266, 1829–1834.
- Sturtevant, J. M. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 2236–2240.
- Suzuki, M., and Yagi, N. (1996) *J. Mol. Biol.* 255, 677–687.
- Takeda, Y., Ross, P. D., and Mudd, C. P. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 8180–8184.
- Thompson, K. S., Vinson, C. R., and Freire, E. (1993) *Biochemistry* 32, 5491–5496.
- Turcatti, G., Vogel, H., and Chollet, A. (1995) *Biochemistry* 34, 3972–3980.
- von Hippel, P. H., and Berg, O. H. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 1608–1612.
- Weiss, M. A. (1990) *Biochemistry* 29, 8020–8024.
- Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C., and Struhl, K. (1990) *Science* 347, 575–578.
- Wendt, H., Baici, A., and Bosshard, H. R. (1994) *J. Am. Chem. Soc.* 116, 6973–6974.
- Wendt, H., Berger, C., Baici, A., Thomas, R. M., and Bosshard, H. R. (1995) *Biochemistry* 34, 4097–4107.
- Wiseman, T., Williston, S., Brandts, J. F., and Lin, L.-N. (1989) *Anal. Biochem.* 179, 131–137.
- Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E., and Matthews, C. R. (1995) *Biochemistry* 34, 12812–12819.