Electrostatic control of half-site spacing preferences by the cyclic AMP response element-binding protein CREB

Jin Kim Montclare, Leslie S. Sloan and Alanna Schepartz*

Department of Chemistry, Yale University, PO Box 208107, New Haven, CT 06520-8107, USA

Received May 22, 2001; Revised June 15, 2001; Accepted July 1, 2001

ABSTRACT

Basic region leucine zipper (bZIP) proteins represent a class of transcription factors that bind DNA using a simple, dimeric, \( \alpha \)-helical recognition motif. The cAMP response element-binding protein (CREB) is a member of the CREB/ATF subfamily of bZIP proteins. CREB discriminates effectively in vivo and in vitro between the 10 bp cAMP response element (ATGACGTCAT, CRE) and the 9 bp activating protein 1 site (ATGACTCAT, AP-1). Here we describe an alanine scanning mutagenesis study designed to identify those residues within the CREB bZIP element that control CRE/AP-1 specificity. We find that the preference of CREB for the CRE site is controlled in a positive and negative way by acidic and basic residues in the basic, spacer and zipper segments. The CRE/AP-1 specificity of CREB is increased significantly by four glutamic acid residues located at positions 24, 28, 35 and 41; glutamic acid residues at positions 10 and 48 contribute in a more modest way. Specificity is decreased significantly by two basic residues located at positions 21 and 23; basic residues at positions 14, 18, 33 and 34 and V17 contribute in a more modest way. All of the residues that influence specificity significantly are located on the solvent-exposed face of the protein–DNA complex and likely participate in interactions between and among proteins, not between protein and DNA. The finding that the CRE/AP-1 specificity of CREB is dictated by the presence or absence of charged residues has interesting implications for how transcription factors seek and selectively bind sequences within genomic DNA.

INTRODUCTION

Basic region leucine zipper (bZIP) proteins represent a class of transcription factors that bind DNA using a simple, dimeric, \( \alpha \)-helical recognition motif (1,2). Each monomer of the folded, dimeric motif contains \( \sim 60 \) amino acids of contiguous \( \alpha \)-helix. Each helix is comprised of three regions that each perform a unique function. Towards the N-terminus is a region rich in basic residues that contact DNA (this segment is often called the ‘basic region’). Near the C-terminus is a region rich in non-polar residues that stabilize the protein dimer (this segment is called the ‘zipper segment’). Connecting the basic and zipper segments is a six residue spacer that sets the register between the two attached functional regions (3–7). High resolution structures of several bZIP–DNA complexes show the zipper segments of two protein monomers assembled into a parallel coiled coil that lies perpendicular to the DNA helix axis and diverges near the DNA to position the two basic segments into adjacent major grooves (3–6,8–10). Although the helical structure of the coiled coil is maintained in the presence and absence of DNA at high protein concentrations, the basic segment exists as a structured \( \alpha \)-helix only in the presence of specific DNA (11–14).

The cAMP response element-binding protein (CREB) is a member of the CREB/ATF subfamily of bZIP proteins. CREB is expressed in a variety of cells, including brain, heart, liver, testis, kidney and spleen, where its primary role is regulation of transcription in response to cAMP (15–19). Increased cAMP levels lead to phosphorylation of CREB at Ser133 by cAMP-dependent protein kinase A. Once phosphorylated, CREB interacts with the co-activator CREB-binding protein (CBP); subsequent interactions between CBP and the general transcriptional machinery up-regulate transcription of cAMP-dependent genes (15,20,21). Thus, CREB is a transcriptional end-point of the cAMP signal transduction cascade.

Like other members of the CREB/ATF subfamily, CREB effectively discriminates between two closely related DNA sequences. One, the 10 bp cAMP response element (CRE), is comprised of two ATGAC half-sites arranged as an inverted repeat (ATGACGT CAT). The other, the 9 bp activating protein 1 site, is composed of two ATGA half-sites separated by a single CG base pair (ATGACT CAT) (22–25). Despite the similarity between these two DNA sequences, CREB displays a strong half-site spacing preference (\( \Delta G = 2.5 \text{kcal} \cdot \text{mol}^{-1} \)) for the CRE site over the AP-1 site (26). Not all bZIP proteins display a half-site spacing preference; for example, GCN4 binds approximately equally to the CRE and AP-1 sites (4,5,27,28).

Crystallographic analysis of the specific DNA complexes formed by CREB and GCN4 homodimers and the Fos–Jun heterodimer identified five conserved residues within the basic region (N–AA–S/CR) that directly contact DNA bases (Fig. 1;
3–5, 10, 28). Taken together, these bZIP–DNA structures provide a clear image of how this set of conserved basic and hydrophobic residues recognize the sequence ATGTA with high affinity. Yet they provide little or no information on precisely which residues are responsible for discriminating between the CRE and AP-1 sites. Previous biochemical work has localized the residues responsible for half-site spacing specificity (CRE/AP-1 specificity) to the basic and spacer segments (23, 24, 29–33).

Here we describe an alanine scanning mutagenesis study designed to identify those residues within the CREB bZIP element that govern CRE/AP-1 specificity. We find that the preference of CREB for the CRE site is controlled in a positive and negative way by discrete acidic and basic residues in the basic, spacer and zipper segments. Specificity is increased significantly by four glutamic acid residues located at positions 24, 28, 35 and 41. Conversion of any one of these residues to alanine diminished CRE/AP-1 specificity by at least 24, 28, 35 and 41. Conversion of any one of these residues to alanine improved CRE/AP-1 selectivity by at least 1.3 kcal·mol−1. Glutamic acid residues at positions 10 and 48 contribute, but less significantly; conversion of either of these residues to alanine diminished CRE/AP-1 specificity by at least 0.6 and −0.8 kcal·mol−1. CRE/AP-1 specificity is decreased significantly by two basic residues located at positions 21 and 23. Conversion of either of these residues to alanine improved CRE/AP-1 selectivity by ΔΔG > 2.3 kcal·mol−1. Basic residues at positions 14, 18, 33 and 34 and V17 also contribute, but less significantly; conversion of any one of these residues to alanine increased CRE/AP-1 specificity by between ΔΔG = −1.0 and −1.3 kcal·mol−1. Most of the residues that change specificity significantly are located on the solvent-exposed face of the protein–DNA complex and likely participate in electrostatic interactions between and among proteins, not between protein and DNA. The finding that the CRE/AP-1 specificity of CREB is controlled by charged residues has interesting implications for how transcription factors seek out and selectively bind to precise sequences within genomic DNA and provide guidelines for the design of exceptionally selective artificial transcription factors.

**MATERIALS AND METHODS**

**Peptides**

A peptide containing residues 271–341 of CREB (B70) and alanine derivatives thereof were overexpressed in *Escherichia coli* strain BL21(DE3)plysS from plasmids prepared from pAED-Δ350 by Kunkel mutagenesis (34). All peptides were purified by FPLC using an SP-Sepharose column (Pharmacia) and were ≥98% pure as judged by HPLC. The identity of each peptide was confirmed by amino acid analysis using homoserine and norleucine as internal standards and by ESI mass spectrometry. Aliquots of each peptide were stored at concentrations of between 5 and 50 µM in 5 mM DTT and 10 mM Tris (pH 7.4) at −20°C. Stock peptide concentrations were as follows: B70, 1 µM; Q6A, 220 nM; E9A, 40 nM; E10A, 1 µM; R13A, 400 nM; K14A, 200 nM; R15A, 200 nM; E16A, 10 µM; V17A, 1 µM; R18A, 200 nM; L19A, 180 nM; M20A, 10 µM; K21A, 200 nM; R23A, 180 nM; E24A, 200 nM; R27A, 200 nM; E28A, 200 nM; R31A, 1 µM; K32A, 1 µM; K33A, 5 µM; K34A, 1 µM; E35A, 5 µM; V37A, 1 µM; C39A, 40 nM; E41A, 200 nM; N42A, 400 nM; E48A, 40 nM. Several peptides (B70, E9A, R13A, R15A, E16A, L19A, M20A, R23A, E24A, E28A, C39A, N42A and E48A) aggregated upon long-term storage. These peptides were treated with 10 mM DTT, heated in a 55°C water bath for 10 min and frozen at −20°C before use.

**DNA**

The oligodeoxyribonucleotides CRE24 and AP-123 were synthesized on the 0.2 µmol scale using a Millipore Expedite DNA/RNA Synthesizer Model 8909 (35). CRE24A d(AGTGGAGATGACGTACATCTCGTGC) and its complement CRE23B d(GCAGCAGATGACGTATCTCCACT) contain the CRE target site, whereas AP-123A d(AGTGGA-GATGACTACATCTCGTGC) and its complement AP-123B d(GCAGCAGATGAGATCCTCCACT) contain the AP-1 site. Oligodeoxyribonucleotides were labeled on the 5′-end with T4 polynucleotide kinase and [γ-32P]ATP and annealed to the complementary strand by heating the mixture to 95°C for 2 min and slowly cooling to room temperature.

** Electrophoretic mobility shift assays**

Binding reactions were performed in a buffer containing 1.4 mM KH2PO4, 4.3 mM Na2HPO4, 2.7 mM KCl, 137 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1% NP-40, 0.4 mg/ml acetylated BSA and 5% glycerol (pH 7.4) (23, 24). Briefly, a set of serially diluted peptides were incubated with <50 pM [γ-32P]-labeled duplex DNA for 30 min at room temperature and then applied to a non-denaturing 8% polyacrylamide gel (80:1 acrylamide:N,N′-methylene bisacrylamide) prepared in 1× TBE. The temperature of the gel and the running buffer (5× TBE) were maintained at 25°C using a circulating, temperature controlled water bath. The samples were electrophoresed at 500 V for 30 min. The gels were analyzed using a Storm 840 Phosphorimager (Molecular Dynamics). Amounts of free and bound DNA were analyzed using the program Kaleidagraph (Abelbeck Software). Dissociation constants (Kd) were determined by fitting the data to the Langmuir equation $F = c × [1/(1 + K_d/[peptide])]$. In these equations $F$ = (c.p.m. in protein–DNA complex)/(c.p.m. in protein–DNA complex + c.p.m. in free DNA), [peptide] is the total peptide concentration and $c$ is an adjustable parameter representing the maximum value of $F$ ($c$ ≤ 1, for many peptides $c$ was defined as 1) (23, 36, 37). Values reported represent the average ± standard error of at least three independent trials. Error bars on the plots represent the standard error for each data point. ΔG values were calculated from the relationship $\Delta G = -RT\ln K_{eq}$, where R is the universal gas constant (1.987 × 103 kcal·mol−1·K−1) and T is the temperature in Kelvin.

**RESULTS**

**Quantitative analysis of CRE/AP-1 specificity**

To identify which residues within CREB contribute to CRE/AP-1 selectivity, we prepared 26 CREB variants in which a single residue in the bZIP element (residues 271–341) was replaced by alanine. The positions of alanine substitution were located throughout the basic and spacer segments and in the N-terminal region of the zipper segment, as illustrated schematically in Figure 1. The affinity of each CREB variant for a duplex oligonucleotide containing a CRE (ATGACGT CAT)
or AP-1 (ATGACTCAT) target site (CREB271–341, GCN4 (225–281), c-Fos (139–193) and c-Jun (263–317). Residues that contact bases directly are in italic; residues changed to alanine in this study are underlined; leucines within the zipper region are in bold. Residues which when changed to alanine caused enhanced CRE/AP-1 specificity are in blue; residues which when changed to alanine caused diminished CRE/AP-1 specificity are in red.

Figure 1. The bZIP element sequences of B70 (CREB271–341), GCN4 (225–281), c-Fos (139–193) and c-Jun (263–317). Residues that contact bases directly are in italic; residues changed to alanine in this study are underlined; leucines within the zipper region are in bold. Residues which when changed to alanine caused enhanced CRE/AP-1 specificity are in blue; residues which when changed to alanine caused diminished CRE/AP-1 specificity are in red.

Figure 2. Binding isotherms illustrating the CRE/AP-1 specificities of B70 and variants thereof. Plots show the fraction of CRE24 (open symbols) and AP-123 (closed symbols) bound (F) as a function of peptide concentration at 25°C. Curves represent the best fit of the data to the equation C·{1/(1 + K_{app}/[peptide])}, where K_{app} and C are adjustable parameters. The equation used to determine values for K_{app} is independent of whether the bZIP peptide forms a dimer before or after the DNA binding event (58). Each point represents the average of at least three independent determinations. Error bars denote the standard error.
specificity of $-2.5 \pm 0.2$ kcal·mol$^{-1}$, in good agreement with the value reported previously (26; Fig. 4 and Table 1). Interestingly, 22 of the 26 CREB variants displayed higher DNA affinity than the wild-type CREB sequence (B70).

**Variants displaying decreased CRE/AP-1 selectivity**

Four CREB variants, E24A, E28A, E35A and E41A, displayed a diminished ability to discriminate between the CRE and AP-1 sites when compared to B70. Each of these variants exhibited a specificity value at least 1.0 kcal·mol$^{-1}$ lower than that exhibited by B70 itself (Figs 2 and 3B and Table 1). Although many CREB variants bound DNA with affinities higher or lower than B70, in all other cases the change in affinity for CRE/A was roughly paralleled the change in affinity for AP-1.$^{23,24}$ It is notable that two other variants containing alanine in place of glutamic acids (E9A and E41A) displayed levels of CRE/AP-1 specificity that were comparable to B70. This finding suggests that the altered specificities displayed by E24A, E28A, E35A and E41A were not due to changes in overall peptide charge but rather to the specific replacement of these four glutamic acid residues by alanine.

The four glutamic acid residues implicated by our experiments as important for specificity are not localized to one small region of the bZIP element, but rather are scattered throughout the basic, spacer and zipper segments (Fig. 3A). Two, at positions 24 and 28, are located in the basic segment, one, at position 35, is located in the spacer segment and one, at position 41, is located in the N-terminal region of the zipper segment (Fig. 5). Interestingly, none of these single site alanine substitutions eliminated CRE/AP-1 specificity entirely. Comparison of the specificity values exhibited by E24A, E28A, E35A and E41A (Table 1) suggested that the four glutamic acids do not function additively, that is, the sum of the specificities ($\Delta G$) displayed by each variant exceeded the specificity of B70 (38,39). Loss of any one glutamic acid residue diminished but did not abolish specificity, suggesting that CREB specificity is encoded in a somewhat redundant manner.

**Variants displaying increased CRE/AP-1 selectivity**

Two CREB variants, K21A and R23A, displayed an enhanced preference for the CRE site over the AP-1 site when compared to B70. Each exhibited a specificity value that was nearly double that of B70, $-4.8 \pm 0.3$ and $-5.0 \pm 0.5$ kcal·mol$^{-1}$ for K21A and K23A, respectively (Figs 2 and 3B and Table 1). In contrast to the lack of proximity between those residues whose conversion to alanine lowered CRE/AP-1 specificity, these two basic residues are proximal in sequence as well as structure. In fact, they flank the most highly conserved DNA contact residue within the bZIP motif, N22 (Figs 1, 3A and 5).

**The interplay between affinity and specificity**

It is useful to consider the changes in CRE and AP-1 affinity that result in the observed changes in selectivity. First consider the four variants displaying diminished selectivity, E24A, E28A, E35A and E41A. In three cases (E24A, E28A and E41A) CRE/AP-1 specificity is reduced in the variant peptide by virtue of enhanced affinity for the AP-1 site (Table 1). For example, the E24A–AP-1 complex is 3.9 kcal·mol$^{-1}$ more stable than the B70–AP-1 complex, whereas the E28A–CRE complex is only 2.0 kcal·mol$^{-1}$ more stable than the B70–CRE complex. Similarly, the E28A–AP-1 complex is 3.0 kcal·mol$^{-1}$ more stable than the B70–AP-1 complex, whereas the E28A–CRE complex is only 2.0 kcal·mol$^{-1}$ more stable than the B70–CRE complex. Surprisingly, the trend also holds for a residue in the bZIP zipper segment that is located outside the expected range of DNA contact. The E41A–AP-1 complex is 2.2 kcal·mol$^{-1}$ more stable than the B70–AP-1 complex, whereas the E41A–CRE complex is only 1.0 kcal·mol$^{-1}$ more stable than the B70–CRE complex. The one exception to this trend is the E35A variant. E35A displays diminished CRE/AP-1 specificity solely by virtue of decreased (1.0 kcal·mol$^{-1}$) affinity for the CRE site. We conclude that the glutamic acid residues located at positions 24, 28 and 41 contribute to CRE/AP-1 specificity by lowering the stability of CREB–AP-1 complexes; these three residues select against an interaction with the AP-1 site. In contrast, the glutamic acid residue at position 35 selectively increases the stability of CREB–CRE complexes.

Next we consider the two variants that displayed enhanced CRE/AP-1 selectivity, K21A and R23A. In these cases the specificity of the variant is increased by virtue of enhanced affinity for the CRE site (Table 1). For example, the K21A–CRE complex is 3.6 kcal·mol$^{-1}$ more stable than the B70–CRE complex, whereas the K21A–AP-1 complex is only 1.1 kcal·mol$^{-1}$ more stable than the B70–AP-1 complex. Similarly, the R23A–CRE complex is 4.8 kcal·mol$^{-1}$ more stable than the

![Figure 3](A) A helical net illustrating the basic spacer and partial zipper segments of B70. Residues shaded yellow are highly conserved among bZIP proteins and contact bases directly in all known bZIP–DNA structures. Residues shaded blue illustrate positions where alanine substitution diminished CRE/AP-1 specificity; residues shaded red illustrate positions where alanine substitution diminished CRE/AP-1 specificity. The residue shaded black is the first leucine in the zipper region. (B) The specificity values of all CREB variants compared to B70 (colored black). The shaded area behind the bars represents the region within 1 kcal·mol$^{-1}$ of the specificity of B70 (2.5 kcal·mol$^{-1}$). Color coding is identical to that in (A).
2.5 kcal·mol$^{-1}$ more stable than the B$_{70}$–AP-1 complex. For $K_{\text{eq}}$ is the equilibrium dissociation constant, Table 1. The phenomenon of CRE/AP-1 specificity was first identified in 1987. The basic residues located at positions 21 and 23 in wild-type CREB modulate CRE/AP-1 specificity. We conclude that the basic residues located for the AP-1 site. Various proposals to explain CRE/AP-1 selectivity have been advanced since 1987. One early proposal suggested that the coiled coil in bZIP proteins that display CRE/AP-1 specificity (such as CREB/ATF proteins) might be rigid enough to orient the basic segment helices to specifically recognize one site and not the other. Proteins that do not display CRE/AP-1 specificity (such as GCN4) were envisioned to be more flexible and attract attention because of the remarkable similarity between the two DNA sequences in question. The CRE and AP-1 target sites share the same consensus half-site and differ by only a single, central G·C base pair. Coupled with the similarity between the two oligonucleotides is a remarkable similarity among bZIP proteins, whether these proteins discriminate between the two sites (such as CREB/ATF proteins) (22,25) or not (such as GCN4) (4,5,10,28).

**DISCUSSION**

The molecular mechanism that accounts for the CRE/AP-1 specificities of bZIP proteins has been a focus of research since this DNA recognition motif was first identified in 1987 (1,3,40–42). The phenomenon of CRE/AP-1 specificity attracted attention because of the remarkable similarity between the two DNA sequences in question. The CRE and AP-1 target sites share the same consensus half-site and differ by only a single, central G·C base pair. Coupled with the similarity between the two oligonucleotides is a remarkable similarity among bZIP proteins, whether these proteins discriminate between the two sites (such as CREB/ATF proteins) (22,25) or not (such as GCN4) (4,5,10,28).

**Previous proposals for CRE/AP-1 specificity**

Various proposals to explain CRE/AP-1 selectivity have been advanced since 1987. One early proposal suggested that the coiled coil in bZIP proteins that display CRE/AP-1 specificity (such as CREB/ATF proteins) might be rigid enough to orient the basic segment helices to specifically recognize one site and not the other. Proteins that do not display CRE/AP-1 specificity (such as GCN4) were envisioned to be more flexible and permit recognition of both sites (41,43,44). This proposal was

### Table 1. Equilibrium dissociation constants of CREB–CRE and CREB–AP-1 complexes at 25°C determined by electrophoretic mobility shift analysis

<table>
<thead>
<tr>
<th>Peptide</th>
<th>$\Delta G_{\text{CREB}}$ (kcal·mol$^{-1}$)</th>
<th>$\Delta G_{\text{AP-1}}$ (kcal·mol$^{-1}$)</th>
<th>$\Delta G_{\text{CREB}}$ ($\Delta G_{\text{mut}} - \Delta G_{B70}$) (kcal·mol$^{-1}$)</th>
<th>$\Delta G_{\text{AP-1}}$ ($\Delta G_{\text{mut}} - \Delta G_{B70}$) (kcal·mol$^{-1}$)</th>
<th>$\Delta G_{\text{spec}}$ (kcal·mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B70</td>
<td>-21.5 ± 0.1</td>
<td>-19.0 ± 0.2</td>
<td>-2.5</td>
<td>-2.9</td>
<td>-2.5 ± 0.2</td>
</tr>
<tr>
<td>Q6A</td>
<td>-24.0 ± 0.2</td>
<td>-21.9 ± 0.1</td>
<td>-2.5</td>
<td>-2.9</td>
<td>-2.1 ± 0.3</td>
</tr>
<tr>
<td>E9A</td>
<td>-25.0 ± 0.2</td>
<td>-21.9 ± 0.3</td>
<td>-3.5</td>
<td>-3.1</td>
<td>-3.1 ± 0.1</td>
</tr>
<tr>
<td>E10A</td>
<td>-22.4 ± 0.2</td>
<td>-20.5 ± 0.3</td>
<td>-1.9</td>
<td>-1.9</td>
<td>-1.9 ± 0.1</td>
</tr>
<tr>
<td>R13A</td>
<td>-23.8 ± 0.5</td>
<td>-20.7 ± 0.1</td>
<td>-3.1</td>
<td>-3.1</td>
<td>-3.1 ± 0.5</td>
</tr>
<tr>
<td>K14A</td>
<td>-24.6 ± 0.4</td>
<td>-21.1 ± 0.3</td>
<td>-3.1</td>
<td>-3.1</td>
<td>-3.1 ± 0.6</td>
</tr>
<tr>
<td>R15A</td>
<td>-24.7 ± 0.4</td>
<td>-21.5 ± 0.1</td>
<td>-3.2</td>
<td>-3.2</td>
<td>-3.2 ± 0.3</td>
</tr>
<tr>
<td>E16A</td>
<td>-23.4 ± 0.1</td>
<td>-20.8 ± 0.1</td>
<td>-1.9</td>
<td>-1.8</td>
<td>-1.8 ± 0.1</td>
</tr>
<tr>
<td>V17A</td>
<td>-23.6 ± 0.1</td>
<td>-19.8 ± 0.2</td>
<td>-2.1</td>
<td>-0.8</td>
<td>-3.8 ± 0.3</td>
</tr>
<tr>
<td>R18A</td>
<td>-24.3 ± 0.3</td>
<td>-20.8 ± 0.3</td>
<td>-2.8</td>
<td>-1.8</td>
<td>-3.5 ± 0.1</td>
</tr>
<tr>
<td>L19A</td>
<td>-25.2 ± 0.1</td>
<td>-22.4 ± 0.4</td>
<td>-3.7</td>
<td>-3.4</td>
<td>-2.8 ± 0.2</td>
</tr>
<tr>
<td>M20A</td>
<td>-23.7 ± 0.2</td>
<td>-21.6 ± 0.2</td>
<td>-2.2</td>
<td>-2.6</td>
<td>-2.1 ± 0.2</td>
</tr>
<tr>
<td>K21A</td>
<td>-25.1 ± 0.6</td>
<td>-20.1 ± 0.2</td>
<td>-5.0</td>
<td>-5.0</td>
<td>-5.0 ± 0.5</td>
</tr>
<tr>
<td>R23A</td>
<td>-26.3 ± 0.4</td>
<td>-21.5 ± 0.1</td>
<td>-4.8</td>
<td>-4.8</td>
<td>-4.8 ± 0.3</td>
</tr>
<tr>
<td>E24A</td>
<td>-24.2 ± 0.1</td>
<td>-22.9 ± 0.1</td>
<td>-1.7</td>
<td>-3.9</td>
<td>-1.3 ± 0.1</td>
</tr>
<tr>
<td>R27A</td>
<td>-23.6 ± 0.3</td>
<td>-20.4 ± 0.1</td>
<td>-2.1</td>
<td>-1.4</td>
<td>-3.2 ± 0.3</td>
</tr>
<tr>
<td>E28A</td>
<td>-23.5 ± 0.2</td>
<td>-22.0 ± 0.2</td>
<td>-3.0</td>
<td>-1.5</td>
<td>-1.5 ± 0.1</td>
</tr>
<tr>
<td>R31A</td>
<td>-21.4 ± 0.1</td>
<td>-18.3 ± 0.2</td>
<td>+0.7</td>
<td>+0.7</td>
<td>-3.1 ± 0.2</td>
</tr>
<tr>
<td>K32A</td>
<td>-20.8 ± 0.3</td>
<td>-18.0 ± 0.2</td>
<td>+0.7</td>
<td>+1.0</td>
<td>-2.8 ± 0.5</td>
</tr>
<tr>
<td>K33A</td>
<td>-21.6 ± 0.1</td>
<td>-18.1 ± 0.1</td>
<td>-0.1</td>
<td>+0.9</td>
<td>-3.5 ± 0.1</td>
</tr>
<tr>
<td>K34A</td>
<td>-23.4 ± 0.2</td>
<td>-19.9 ± 0.6</td>
<td>-1.9</td>
<td>-0.9</td>
<td>-3.5 ± 0.4</td>
</tr>
<tr>
<td>E35A</td>
<td>-20.5 ± 0.1</td>
<td>-19.0 ± 0.1</td>
<td>+1.0</td>
<td>0</td>
<td>-1.5 ± 0.05</td>
</tr>
<tr>
<td>V37A</td>
<td>-22.0 ± 0.3</td>
<td>-19.3 ± 0.3</td>
<td>-0.5</td>
<td>-0.3</td>
<td>-2.7 ± 0.3</td>
</tr>
<tr>
<td>C39A</td>
<td>-25.1 ± 0.1</td>
<td>-23.0 ± 0.03</td>
<td>-3.6</td>
<td>-4.0</td>
<td>-2.1 ± 0.1</td>
</tr>
<tr>
<td>E41A</td>
<td>-22.5 ± 0.1</td>
<td>-21.2 ± 0.2</td>
<td>-1.0</td>
<td>-2.2</td>
<td>-1.3 ± 0.1</td>
</tr>
<tr>
<td>N42A</td>
<td>-23.3 ± 0.2</td>
<td>-20.8 ± 0.3</td>
<td>-1.8</td>
<td>-1.8</td>
<td>-2.5 ± 0.1</td>
</tr>
<tr>
<td>E48A</td>
<td>-24.2 ± 0.3</td>
<td>-22.5 ± 0.3</td>
<td>-2.7</td>
<td>-3.5</td>
<td>-1.7 ± 0.3</td>
</tr>
</tbody>
</table>

All values represent the average ± SE of at least three independent determinations. Values of $\Delta G$ were calculated from the equation $\Delta G = -RT\ln K_{\text{eq}}$, where $K_{\text{eq}}$ is the equilibrium dissociation constant, $R = 1.987 \times 10^3$ kcal·mol$^{-1}$·K$^{-1}$ and $T$ is the temperature in Kelvin.
supported by the observation that certain bZIP peptide models containing unnatural dimerization domains displayed high levels of CRE/AP-1 specificity (36,37). However, extensive experimentation has shown that, irrespective of its role in peptide models, the coiled coil in bZIP proteins is not the primary determinant of CRE/AP-1 specificity (23,24,45).

A second proposal to explain CRE/AP-1 selectivity focused on potential differences in the flexibility of the spacer segment, as opposed to the coiled coil. Indeed, replacement of the six residue spacer segment of GCN4 (which displays low CRE/AP-1 specificity) with that of CRE-BP1 (which displays high CRE/AP-1 specificity) modestly increased specificity (23). Similar results were observed when three residues in the spacer segment of GCN4, L247, R249 and M250, were replaced with the corresponding residues K247, Y249 and V250 from CREB (46). These studies, along with those on bZIP peptide models containing non-natural dimerization domains (described above), indicate that subtle sequence-dependent conformational adjustments in the spacer can alter CRE/AP-1 selectivity in certain cases (36,37,47). However, subsequent ‘segment swapping’ experiments demonstrated clearly that for CREB the primary determinants of selectivity were located within the basic segment, not the spacer segment (23).

A third proposal to explain the origins of CRE/AP-1 specificity focused instead on differential DNA bending (or bendability) within the nucleic acid portion of the bZIP–DNA partnership. This proposal was based on experiments that revealed the existence of an intrinsic bend in the CRE site. This bend was removed, and the DNA straightened, upon binding of the CRE site by the CREB/ATF protein CRE-BP1 (21,26,48–51). No detectable conformational adjustments occurred in the DNA upon binding of GCN4, providing an attractive explanation for how GCN4 binds both sites with equal affinity. However, significant bending of both sites occurred upon binding of the CREB/ATF proteins CRE-BP1 and CREB (26,48,49), suggesting that the differential cost of DNA bending could dictate CRE/AP-1 specificity. Although numerous experiments support differential DNA bending as a mechanism for achieving specificity (52–54), subsequent studies of variant CRE/BP1–GCN4 proteins demonstrated that DNA bending and specificity were unrelated in this case (24).
in the CREB–CRE crystal structure (Fig. 5). E24 on one monomer (E295) contributes two hydrogen bonds; one from the side chain to a non-bridging phosphate oxygen between G₁ and C₁, and another from the main chain carbonyl to the side chain of R27 (R298). E24 on the second monomer (referred to as E’24) contributes a hydrogen bond from the side chain to the εN of R’27. The side chain of E28 (E299) donates a hydrogen bond to the side chain on R31 (R302). Finally, E41 (E312) accepts a hydrogen bond from the phenolic proton of Y’36 (Y307) on the other protein chain. Changing each of these residues to alanine lowers CRE/AP-1 specificity by selectively increasing affinity for the AP-1 site (Table 1).

It is notable that most of these interactions are among and between proteins and not nucleic acid; both E28 and E41 are located on the outside of the protein–DNA complex and point away from the DNA. It is also notable that the Y36F CREB variant forms a less stable bZIP dimer in the absence of DNA away from the DNA. It is also notable that the Y36F CREB located on the outside of the protein–DNA complex and point between proteins and not nucleic acid; both E28 and E41 are due to alanine lowers CRE/AP-1 specificity by selectively increasing affinity for the AP-1 site (Table 1).

Residues that select for the CRE site: E35

Of the four residues that contribute to CRE/AP-1 specificity, only one, E35 (E306), does so by selectively increasing affinity for the CRE site. Changing this residue to alanine lowers CRE/AP-1 specificity solely by virtue of a decreased (1.0 kcal·mol⁻¹) affinity for the CRE sequence. In the CREB–CRE structure the side chain of E35 is also located on the solvent-exposed face of the protein–DNA complex and donates a hydrogen bond to the side chain on R31 (R302) (3). This intramolecular contact may also be important in orienting CREB for optimal interaction with the CRE target site. Since mutation of E35 to alanine exhibited no effect on the energetics of the CREB–AP-1 complex, we predict that the side chain of E35 does not contact R31 in the CREB–AP-1 complex. These data emphasize that there may be subtle but energetically different differences between the structures of CREB–AP-1 and CREB–CRE.

Residues that select against the CRE site: K21 and R23

The two basic residues that select against the CRE site both contact DNA in the CREB–CRE crystal structure. The side chain of R23 (R294) donates a hydrogen bond to a non-bridging oxygen between G₁ and T₁ within the consensus DNA site. Interestingly, K21 (K292) donates a hydrogen bond to a non-bridging phosphate between G’₁ and G’₄ outside the consensus 10 bp target site (Fig. 5) (3). Our data suggest that one or both of these contacts are not optimal for increasing affinity for the CRE sites. The fact that specificity can be enhanced by changing one of these residues to alanine provides clear guidelines for the design of non-natural molecules that possess this property.

The role of Mg²⁺

One of the most interesting details of the CREB–CRE complex as determined by crystallography was the appearance of a hexahydrated Mg²⁺ ion bound in a cavity between the basic region and DNA. Subsequent experimental work indicated that the presence of 10 mM Mg²⁺ increased the stability of CREB285–341–CRE by at least 14-fold (3,55). Mg²⁺ did not enhance binding of CREB to the AP-1 site. Although the experiments reported herein were performed in a buffer that lacked Mg²⁺ and with a smaller CREB domain (B₁₋₃₀ contains residues 271–341 of CREB), we find that CRE/AP-1 specificity is decreased (to ΔΔGₐₚ = −1.8 kcal·mol⁻¹), not increased, by addition of 10 mM Mg²⁺ (Fig. 4). Although the affinity of B₁₋₃₀ for the CRE site was increased 2-fold by addition of 10 mM Mg²⁺, a value consistent with the 14-fold effect reported earlier, we find that affinity for the AP-1 site was increased more. This result is consistent with a recent report describing no effect of Mg²⁺ on the affinity of CREB for several non-consensus CRE sequences (55). In conclusion, we find no evidence that Mg²⁺ increases the CRE/AP-1 specificity of CREB.

CONCLUSION

There is considerable current interest in understanding the molecular mechanisms used by proteins to locate their target sites within a large genome (56–61) and also on how to design non-natural transcription factors able to regulate gene expression in a prescribed manner (62–73). Here we show that the selectivity of a simple natural transcription factor is controlled in both a positive and negative way by electrostatic interactions. Notably, many of these interactions are between and among proteins and not between protein and DNA (at least not directly). A similar situation describes the role of E44 in integration host factor, which specifies T:A/A:T discrimination without directly contacting DNA (74). The finding that the CRE/AP-1 specificity of CREB is controlled by charged residues has interesting implications about how transcription factors seek and selectively bind to precise sequences within genomic DNA. This result emphasizes the ease with which specificity can be altered through the formation of different heterodimers or through combinatorial interactions with cellular or viral accessory factors (56,59,75–78). In addition, these findings provide additional criteria useful for the design of exceptionally selective artificial transcription factors (66,73).

ACKNOWLEDGEMENTS

We thank Richard Brennan for providing the coordinates of the CREB–CRE complex. We thank the National Science foundation for predoctoral fellowships to both J.K.M. and L.S.S. This work was supported by the NIH.

REFERENCES


