Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 2435–2443

BIOORGANIC & MEDICINAL CHEMISTRY

Effects of Nucleic Acids and Polyanions on Dimer Formation and DNA Binding by bZIP and bHLHZip Transcription Factors

Jennifer J. Kohler and Alanna Schepartz*

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

Received 23 February 2001; accepted 29 May 2001

Abstract—A large fraction of known transcription factors form 2:1 complexes with DNA. In our studies of the assembly of such ternary (protein–protein–DNA) complexes formed by bZIP and bHLHZip proteins, we found that the proteins recognize DNA as monomers. Here we show that protein monomer–DNA complexes are favored at high DNA concentrations. Further, we show that, due to fast rates of association with protein monomers, DNA and other polyanions accelerate the rate of protein dimer formation. Finally, we find that DNA-assisted formation of protein dimers provides a mechanism by which dimeric transcription factors can rapidly discriminate between specific and nonspecific sites.

Introduction

Many transcription factors form dimeric complexes with DNA, in which two protein monomers interact with each other while each recognizes DNA. For these proteins, one region of each protein monomer directly contacts DNA, while another region facilitates dimerization with a second protein monomer. The rate and mechanism of assembly of such protein–protein–DNA complexes will influence the identity and order of addition of other transcriptional components. As a result, the way that dimeric transcription factors bind DNA is an important determinant of the composition, and perhaps the function, of the supramolecular complex that forms at each promoter and is responsible for initiation of transcription. Consequently, improved insight into the factors governing assembly of dimeric transcription factor–DNA complexes will enhance our understanding of gene regulation and guide attempts to manipulate the level of transcription of individual genes.

DNA binding by dimeric proteins requires the assembly of a ternary complex (one DNA and two protein components). As ternary collisions are relatively unlikely, assembly probably occurs via a two-step process. For proteins that form dimeric complexes with DNA, there exist two limiting pathways (Fig. 1) that may describe the route of complex assembly. The protein can dimerize first, then associate with DNA (the route we refer to as the dimer pathway), or can follow a pathway in which two monomers bind DNA sequentially and assemble their dimerization interface while interacting with DNA (the monomer pathway). Since the two pathways constitute two halves of a thermodynamic cycle, the equilibrium affinity of the protein for DNA will be equivalent whether binding proceeds through the monomer pathway, the dimer pathway, or some combination of the two.

We[1–3] and others[4–9] have found that many dimeric transcription factors bind DNA by following the monomer pathway. The bZIP family of transcription

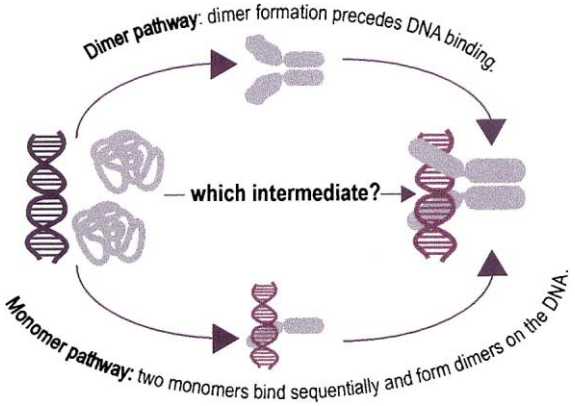

Figure 1. Two pathways to describe binding of dimeric proteins to DNA. In the dimer pathway, proteins first form dimers, which then go on to bind DNA. Along the monomer pathway, a monomer–DNA complex is formed first, followed by recruitment of the second monomer to form the final complex.

*Corresponding author. Tel.: +1-203-432-5094; fax: +1-203-432-3486; e-mail: alanna.schepartz@yale.edu

0968-0896/01/$ - see front matter
PII: S0968-0896(01)00221-8

factors has been particularly well studied. Members of this class that have been shown to bind DNA through the monomer pathway include GCN4,[5] ATF-2,[2] and the heterodimer Fos–Jun.[3] bZIP proteins contain the simplest dimeric DNA-recognition motif known to date:[10,11] each monomer contacts the DNA major groove using an otherwise isolated α-helix that contains an abundance of basic residues (the basic region) (Fig. 2A). The dimerization interface is a parallel coiled coil formed by α-helical extensions of the two basic region helices.[10,11] Although considerable evidence suggests that many, if not all, bZIP proteins bind DNA through the monomer pathway,[2,3,5,7] the compulsory intermediate along this pathway, the DNA–protein monomer complex, has been rarely and only fleetingly observed.[1,12,13] Moreover, in all previous cases, the DNA–monomer complex observed contained only basic region peptides that lacked the ability to form a stable, ternary complex.

In this work, we extend our studies of the assembly pathways of dimeric transcription factors in three areas. First, we use fluorescence polarization and fluorescence resonance energy transfer (FRET) to provide clear evidence for a DNA–protein monomer complex. Second, we show, for proteins that use the monomer pathway to bind DNA, that a variety of negatively charged polymers can substitute for DNA to accelerate the rate of dimer formation. Finally, experiments presented here show that the monomer pathway allows a bHLHZip protein to locate its DNA target site expeditiously while in the presence of long, genomic DNA sequences.

Results

Fluorescence polarization provides evidence for a bZIP monomer–DNA complex at high DNA concentration

We used fluorescence polarization and a uniquely labeled bZIP peptide to monitor the relative size of the peptide–DNA complexes formed at varying concentrations of a specific DNA target site. A peptide comprising the bZIP element of ATF-2 (A_{71}) was covalently labeled at its unique cysteine with iodoacetamidofluorescein to generate A_{71}^{SFlu}. The polarization of fluorescence emanating from A_{71}^{SFlu} exhibits a dependence on concentration that reports on the oligomerization state of the labeled peptide.[14] At low peptide concentrations, A_{71}^{SFlu} exists predominantly as a monomer and exhibits low fluorescence polarization. At higher concentrations, A_{71}^{SFlu} forms a dimer that displays higher fluorescence polarization. We measured the fluorescence polarization of A_{71}^{SFlu} alone and with increasing concentrations of target site DNA, CRE_{24} (Fig. 3). As expected, we observed low fluorescence polarization for 25 nM A_{71}^{SFlu} in the absence of DNA. Addition of 5–125 nM DNA increased the fluorescence polarization, and the maximal polarization value was similar to that observed for the A_{71}^{SFlu} dimer (polarization = 0.20).[14] Maximum polarization was observed at a concentration of 50 nM CRE_{24}, which corresponds precisely to a stoichiometry of two A_{71}^{SFlu} molecules to one CRE_{24} molecule. However, further increases in DNA concentration (500 nM CRE_{24}) decreased the fluorescence polarization to approximately the level observed for the A_{71}^{SFlu} monomer in the absence of DNA (polarization = 0.16).

The dependence of $A71^{SFlu}$ fluorescence polarization on DNA concentration can be explained in the following way (Fig. 4). At low DNA concentrations (less than 125 nM CRE_{24}), mostly bZIP dimer–DNA complexes are formed. These complexes have both a high molecular

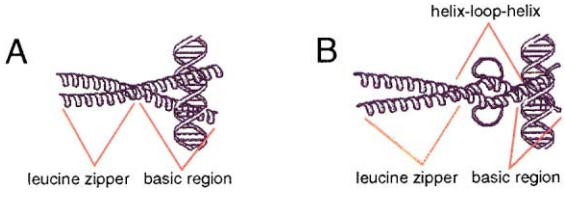

Figure 2. Comparison of (A) bZIP[11] and (B) bHLHZip[22] proteins bound to their DNA target sites.

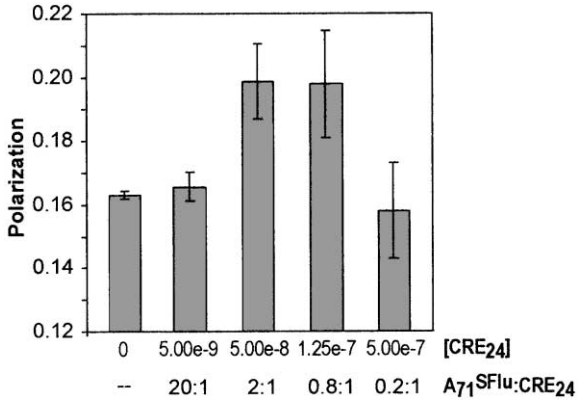

Figure 3. Fluorescence polarization of A_{71}^{SFlu}. 25 nM A_{71}^{SFlu} was incubated along with varying concentrations of CRE_{24}. The fluorescence polarization of the samples is shown for various concentrations of CRE_{24}. Shown below the CRE_{24} concentrations are the corresponding ratios of A_{71}^{SFlu} to CRE_{24}. Error bars represent the standard deviation of five independent measurements.

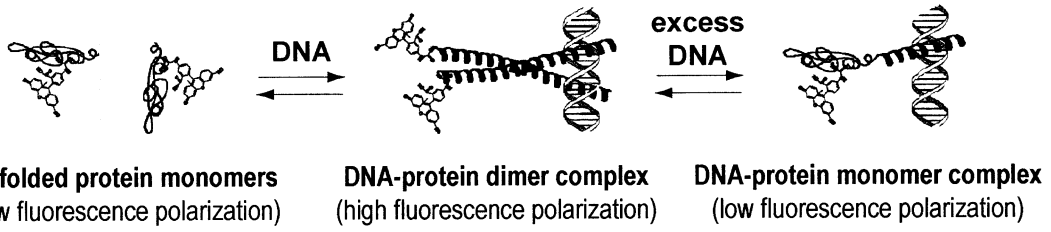

Figure 4. A model for the biphasic dependence on of A_{71}^{SFlu} polarization on CRE_{24} concentration.

weight (about 33 kDa) and a structured protein-protein interface, two features that combine to produce high polarization. At high DNA concentrations, however, the effects of mass action shift the equilibrium to favor the formation of bZIP monomer–DNA complexes.[15] These complexes have a lower molecular weight (about 24 kDa) and the leucine zipper, to which the fluorophore is attached, is likely to be unstructured. These two factors combine to produce a lower fluorescence polarization for the complex.

FRET provides evidence for a bZIP monomer–DNA complex at high DNA concentration

We extended our studies of bZIP monomer–DNA complexes to the heterodimeric transcription factor, AP1. AP1 is comprised of two bZIP proteins, Fos and Jun, that prefer to bind DNA as the heterodimer Fos–Jun, rather than as the homodimers Fos–Fos or Jun–Jun (reviewed in ref 16). Fos and Jun form a more stable dimer than ATF-2 and, consequently, were of interest for our studies of DNA recognition by protein monomers. A peptide comprising the bZIP element of Fos was labeled at a unique C-terminal cysteine with iodoacetamidotetramethylrhodamine to produce $FosC^{SRho}$. Similarly, a peptide comprising the bZIP element of Jun was covalently labeled at a unique C-terminal cysteine with iodoacetamidofluorescein to produce $JunC^{SFlu}$. Formation of the $FosC^{SRho}$–$JunC^{SFlu}$ complex places the two fluorophores in close proximity, with a separation of about 38 Å,[2] such that fluorescence resonance energy (FRET) between the two is observed (Fig. 5).[3] FRET, and consequently dimer formation, can therefore be quantified by monitoring the decrease in fluorescein emission observed at 520 nm. Mixing low (20 nM each) concentrations of $JunC^{SFlu}$ and $FosC^{SRho}$ modestly quenched $JunC^{SFlu}$ fluorescence (about 12% of the maximal quenching), indicating that, as expected based on the equilibrium dissociation constant of the Fos–Jun dimer,[3] only 3–4 nM of the dimer formed at this concentration (Fig. 6A and B). The addition of calf thymus DNA caused further quenching, indicating that dimer formation was promoted by the DNA. The observed quenching is concentration dependent but does not scale linearly with DNA concentration: as DNA concentration increases, the amount of quenching increases, then decreases again at even higher DNA concentration (Fig. 6B).

Three lines of evidence reported previously[3] support our conclusion that the observed quenching of fluorescein fluorescence results from formation of a Fos–Jun heterodimer and not nonspecific interactions. First, the extent of quenching depends on the presence of both $FosC^{SRho}$ and $JunC^{SFlu}$. Fos forms homodimers that possess low thermodynamic stability ($K_d \approx 6$ μM).[17–19] Consistent with this K_d value, little quenching is observed when fluorescein-labeled Fos is mixed with $FosC^{SRho}$.[3] Second, the extent of DNA-induced quenching depends on addition of polymeric DNA. Experiments performed with 20 nM $JunC^{SFlu}$ in the presence of 100 nM $FosC^{SRho}$ show that while 500 nM bp calf thymus DNA increases both the rate and amount of $JunC^{SFlu}$ quenching observed, the non-polymeric DNA equivalent, 1 μM dNTPs, has no such effect.[3] Finally, quenching depends on the ability to bind DNA. Removing the basic region of the Jun peptide, generating $JunLZ^{SFlu}$, produced a peptide that forms a dimer with $FosC^{SRho}$, but cannot bind DNA. As such, DNA is not expected to template formation of $JunLZ^{SFlu}$–$FosC^{SRho}$ dimers. Accordingly, the rate and amount of $JunLZ^{SFlu}$ quenching observed in the presence of $FosC^{SRho}$ was unaffected by the addition of DNA.[3] Taken together, these control experiments suggest that the observed decreases of $JunC^{SFlu}$ fluorescence are not due to nonspecific quenching and that quenching of $JunC^{SFlu}$ fluorescence can be monitored to measure dimer formation.

The extent of dimer formation was quantified by measuring the decrease in fluorescein emission at 515 nm. Fluorescence resonance energy transfer to rhodamine was verified by a small fluorescence increase at 580 nm. However, as observed by others,[20,21] the increase in rhodamine emission (signal at 580 nm) was small relative to the decrease in fluorescein emission and resulted

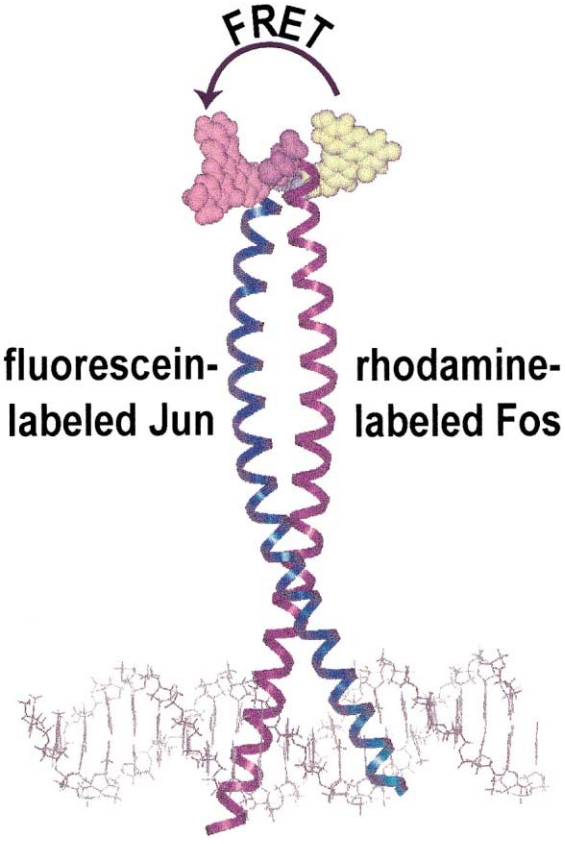

Figure 5. The Fos–Jun–DNA complex. Coordinates for Fos, Jun and DNA were derived from the crystal structure of the complex, with additional amino acids modeled in alpha-helical conformation. The C-terminal cysteines were further modified with acetamidofluorescein and acetamidotetramethylrhodamine. Fluorophores are shown oriented away from the leucine zipper. Fos is shown as a purple ribbon with its appended cysteine in purple CPK representation. The rhodamine label is shown in pink CPK representation. Jun is shown as a blue ribbon with its appended cysteine in blue CPK representation. The fluorescein label is shown in yellow CPK representation. DNA is shown in grey.

both from resonance transfer of energy from the fluorescein donor and from direct absorption of the 490 nm incident light by rhodamine. In both equilibrium and kinetic fluorescence measurements the decrease in emission at 515 nm, rather then the increase at 580 nm, was used quantify $JunC^{SFlu}$–$FosC^{SRho}$ complex formation.

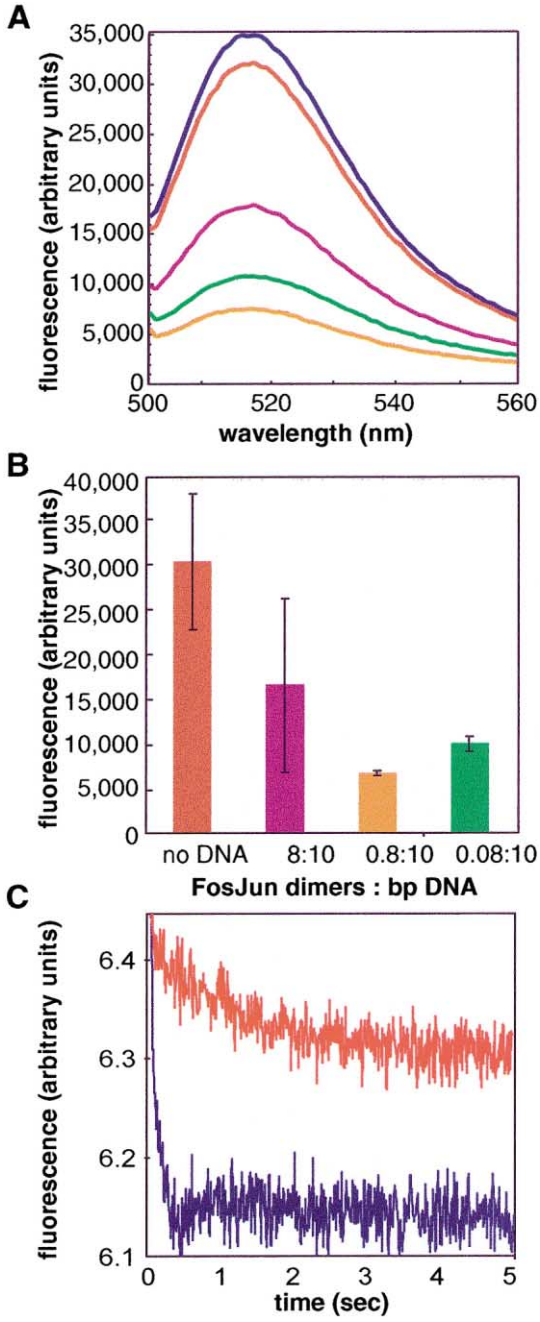

Figure 6. (A) Fluorescence of $JunC^{SFlu}$ and $FosC^{SRho}$ in the presence of calf thymus DNA. Spectra showing 20 nM $JunC^{SFlu}$ alone (blue), with 20 nM $FosC^{SRho}$ (orange), with 20 nM $FosC^{SRho}$ and 50 nM base pairs calf thymus DNA (purple), with 20 nM $FosC^{SRho}$ and 500 nM base pairs calf thymus DNA (yellow), and with 20 nM $FosC^{SRho}$ and 5 μM base pairs calf thymus DNA (green). (B) Histogram illustrating the amount of fluorescence at 520 nm for various concentrations of calf thymus DNA. The amount of DNA is indicated as a ratio of the concentration of potential target sites to concentration of $JunC^{SFlu}$ and $FosC^{SRho}$. Data shown represent the average of three trials; error bars indicate the standard deviation. (C) Stopped-flow fluorescence measurement of the rate of dimer formation. Shown is the time-dependent fluorescence change observed at 520 nm. Fluorescein quenching observed upon rapid mixing of 20 nM $JunC^{SFlu}$ and 100 nM $FosC^{SRho}$ (red trace), and 20 nM $JunC^{SFlu}$, 100 nM $FosC^{SRho}$ and 50 nM $AP1_{23}$ DNA (blue trace).

The dependence of FRET on DNA concentration, like the dependence of fluorescence polarization on DNA concentration, can be explained by invoking the existence of a bZIP monomer–DNA species. In the absence of DNA, few Fos–Jun dimers are formed and little $JunC^{SFlu}$ quenching occurs. As the DNA concentration increases up to a ratio of 0.8 protein dimers per 10 base pairs of DNA (the approximate length of a bZIP binding site), more of the proteins bind DNA as dimers and $JunC^{SFlu}$ quenching increases to its maximal value. At the highest DNA concentration, a ratio of 0.08 protein dimers per 10 base pairs of DNA, the effects of mass action promote formation of bZIP monomer–DNA complexes. In these complexes, the bZIP monomers are not associated with one another and, consequently, little $JunC^{SFlu}$ quenching is observed (Fig. 6B).

Evidence that negatively charged polymers can substitute for DNA to accelerate the rate of bZIP dimer formation

Formation of a bZIP monomer–DNA complex is consistent with our earlier finding that Fos and Jun bind DNA through the monomer pathway.[3] We found that the presence of polymeric nucleic acid (in the form of calf thymus DNA) increased the rate at which Fos–Jun dimers form. This data, combined with numeric simulations, allowed us to conclude that Fos and Jun bind DNA through a pathway in which two monomers bind DNA sequentially, first forming the bZIP monomer–DNA complex, then recruiting the second monomer to form the final, bZIP dimer–DNA complex. Similarly, Rentzeperis et al. found that different types of nucleic acids and other polyanions increased the rate at which Arc repressor dimers formed.[8] They also interpret their results as evidence that Arc repressor dimers form on the DNA, rather than prior to binding DNA. Here we conduct experiments with a variety of charged and uncharged polymers to ask whether a specific DNA target site is a prerequisite for the polyanion-dependent acceleration of Fos–Jun dimer formation.

In a three-way mixing experiment, $FosC^{SRho}$, $JunC^{SFlu}$, and the polymer of interest were rapidly combined and the time-dependent change in $JunC^{SFlu}$ fluorescence was recorded. In a control experiment, the two peptides were mixed together and buffer substituted for the polymer. In the absence of any polymer, the quenching of $JunC^{SFlu}$ fluorescence was described by a single exponential with a relaxation rate of 0.91 s^{-1} (Table 1). This decay rate is indicative of the rate of $FosC^{Srho}$–$JunC^{SFlu}$ dimer formation.[3] As observed previously,[3] the addition of calf thymus DNA caused quenching to occur more rapidly and produced a curve best described by two rate constants. The smaller of these two rate constants (0.81 s^{-1}) is similar to that observed in the absence of DNA, while the larger one is not and increases with increasing DNA concentration. Furthermore, the amplitude associated with the fast phase increases with increasing DNA concentration while the

amplitude associated with the slow phase decreases, suggesting that at higher DNA concentrations more of the binding occurs through the monomer pathway. We have argued[3] that the slower phase represents dimer formation occurring off of the DNA, while the faster phase represents DNA-assisted dimer formation. At the highest calf thymus DNA concentration tested here (5000 nM base pairs), all of the fluorescence quenching is due to the fast phase.

We tested a number of other polymers, charged and uncharged, to determine if they accelerated $FosC^{SRho}$–$JunC^{SFlu}$ dimer formation (Fig. 6C and Table 1). We found that short DNA oligomers containing the AP-1 target site ($AP1_{23}$), to which the Fos–Jun dimer binds, efficiently accelerated the rate of Fos–Jun dimer formation; the presence of 50 nM $AP1_{23}$ was sufficient to cause all of the observed fluorescence quenching to occur through the fast phase. Other polyanions also increased the rate of dimer formation: the presence of either poly-L-glutamic acid or heparin sulfate increased the rate of fluorescence quenching. In contrast, both the polycation poly-L-lysine, and the neutral polymer polyethylene glycol, had little effect on the rate of fluorescence quenching. The presence of negatively charged polymer appears to be sufficient to neutralize the charge of the basic DNA contact regions and, thereby, to bring two dimers together and to template formation of the dimer interface.

Evidence that the monomer pathway allows a bHLHZip protein to locate its DNA target site rapidly in the presence of long, genomic DNA sequences

DNA binding by the bHLHZip (basic region, helix–loop–helix, leucine zipper) family of transcription factors exploits a structural simplicity similar to that of the bZIP family (Fig. 2B).[22] bHLHZip Proteins also use isolated α-helices to recognize the DNA major groove, and also form dimers through parallel alignment of their dimerization elements, in this case, both a helix–loop–helix motif and a leucine zipper. The bHLHZip proteins Max[2] and MyoD[6] have been observed to bind DNA via the monomer pathway. We have previously studied the effect of nonspecific DNA polymers on the rate at which Max binds to its specific DNA target site.[2]

Table 1. Effect of polymers on the rate of Fos–Jun dimer formation

Polymer	τ_1 (s^{-1})	τ_2 (s^{-1})	A_1:A_2
None	0.91	—	100:0
Polyanions			
50 nM bp CT DNA	0.81	11.9	34:66
5000 nM bp CT DNA	—	30.3	0:100
50 nM $AP1_{23}$	—	14.0	0:100
140 μM poly-L-glutamic acid[a]	0.25	3.20	63:37
150 μM heparin sulfate	0.28	6.60	46:54
Polycations			
140 μM poly-L-lysine	0.79	—	100:0
Neutral polymers			
200 μM polyethylene glycol	0.41	—	100:0

[a]Concentrations given in monomer subunits.

We found that the presence of an oligonucleotide containing a mutated target site had little effect on the rate of specific DNA binding by Max. However, when Max monomers were covalently linked to one another, their rate of specific DNA binding was slowed by the addition of the mutated target site. It appears that covalently linked Max monomers are kinetically trapped at the mutated target site and cannot quickly access the specific consensus sequence. Here we study the effects of other nonspecific nucleic acids on the rate of DNA binding by Max to assess the advantage of monomer binding for target site recognition in the context of long, genomic pieces of DNA.

We used two Max constructs. $Max_{22–113}$ consisted of the monomeric form of Max; this form binds DNA by following the monomer pathway (Fig. 7A). $Max^{SS}_{22–105}$ is a covalent dimer of Max; the two monomers are linked through a disulfide bond and, consequently, are forced to bind DNA by following the dimer pathway (Fig. 7B). We incubated 6 nM of $Max_{22–113}$ and 2 nM of $Max^{SS}_{22–105}$ in the absence and presence of a nucleic acid competitor, either 500 nM (bp) calf thymus DNA or 22.7 μg/mL HeLa nuclear extracts. The reactions were then challenged by the addition of a radioactively labeled specific DNA, $Ebox_{22}$, and the rate of binding to the specific site monitored over time by quantifying the formation of the specific complexes, $(Max_{22–113})_2$–$Ebox_{22}$ and $Max^{SS}_{22–105}$–$Ebox_{22}$.[2] These experiments assayed the kinetic specificity of Max in the presence of large, genomic DNA fragments, which are more similar to the nucleic acid content in vivo.

Neither calf thymus DNA nor HeLa nuclear extracts had any effect on the rate at which $Max_{22–113}$ (which binds through the monomer pathway) bound to its specific target site (Fig. 8). We defined the relative rate as the rate with which $Max_{22–113}$ bound $Ebox_{22}$ in the absence of competitor divided by the rate with which $Max_{22–113}$ bound $Ebox_{22}$ in the presence of competitor. For $Max_{22–113}$, the relative rate was 1.1 whether the competitor was calf thymus DNA (Fig. 8A) or HeLa nuclear extracts (Fig. 8B). In contrast, both calf thymus DNA and HeLa nuclear extracts slowed the rate at which $Max^{SS}_{22–105}$ (which binds through the dimer pathway) bound to its specific target site. The relative rate was 3.8 when the competitor was calf thymus DNA (Fig. 8A) and 4.0 with HeLa nuclear extracts (Fig. 8B). We find that the competitors calf thymus DNA and HeLa nuclear extracts both act in essentially the same fashion as a mutated DNA target site: they kinetically trap covalently linked Max dimers, but have little effect on the rate with which Max monomer forms stable complexes with its target site.

Discussion

Assembly of a ternary complex composed of a dimeric transcription factor and DNA requires formation of three macromolecular interfaces—one protein–protein interface between the two monomer subunits and two protein–DNA interfaces formed between individual

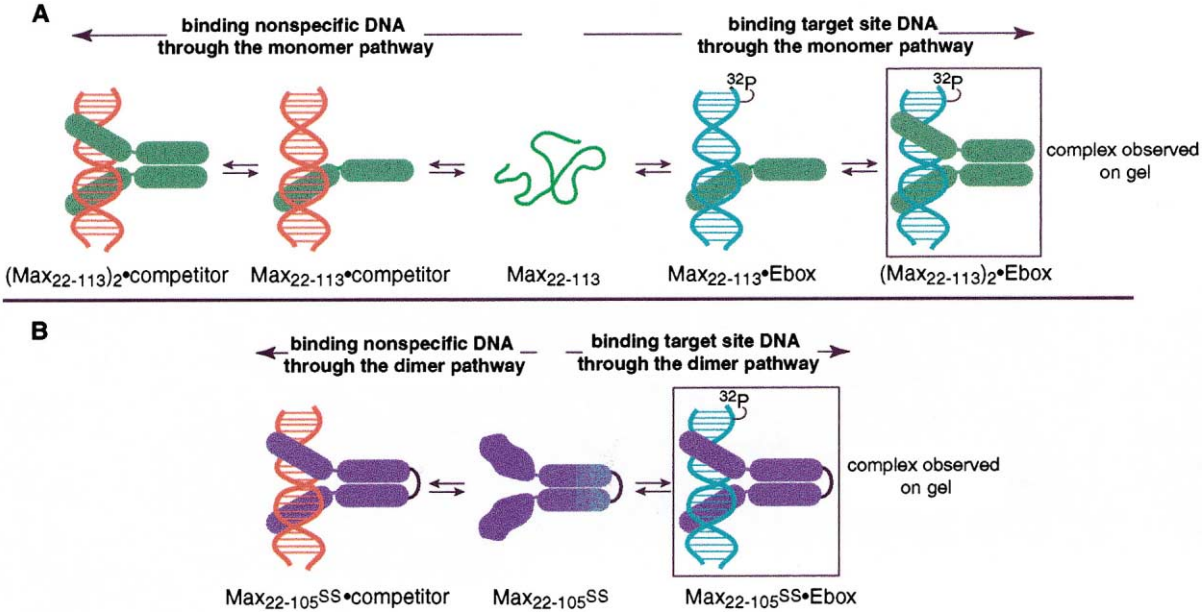

Figure 7. Design of the competition experiment. This assay compared the rates of Max binding, through the monomer and dimer pathways, to specific target site DNA (^{32}P-Ebox$_{22}$, shown in light blue) in the presence and absence of competitor (shown in red). This Scheme illustrates the constructs used to compare the two pathways. (A) The protein Max$_{22-113}$ (shown in green) binds DNA through the two-step monomer pathway. (B) The construct Max$^{SS}_{22-105}$ (shown in purple) contains two covalently linked monomer subunits. Thus, DNA binding by this protein must occur after dimerization, in fashion resembling the second step of the dimer pathway. For both Max$_{22-113}$ and Max$^{SS}_{22-105}$, the rate of binding to the radioactively-labeled specific DNA target site was monitored by gel electrophoresis.

protein monomers and DNA. Here, and in previous work,[1–3,14] we studied DNA binding by proteins belonging to the bZIP and bHLHZip families of dimeric transcription factors. In all cases studied, we found that protein monomers recognized and formed complexes with DNA. The DNA provided a template for the formation of protein dimers and accelerated the rate of dimer formation. Here we provide further evidence for formation of the monomer–DNA intermediate, an important species along the monomer pathway.

The preference for DNA recognition prior to dimer formation may result from the relative kinetics of Coulombic electrostatic interactions, compared to hydrophobic interactions, and from the relative importance of each type of interaction in the formation of protein–DNA and protein–protein contacts. The interfaces formed between protein monomers and DNA are dominated by charge–charge interactions between bridging phosphates in the DNA backbone and (typically) arginine and lysine side chains on the protein. In general, charge–charge interactions are less numerous in the protein–protein interface between individual transcription factor monomers, an interface composed mainly of hydrophobic residues. The attractive force between two molecules of unlike charge depends on the inverse of the separation distance squared and therefore operates at significant distances to guide and accelerate binding events. Thus, association of two macromolecules will be accelerated when those molecules are oppositely charged[23,24] as in the case of basic protein monomers and acidic DNA. The importance of these charge–charge interactions is underscored by our finding that nonspecific nucleic acids and polyanions substitute for DNA target sites to accelerate the rate of dimer formation.

Electrostatic-induced acceleration is most pronounced when the two macromolecules present opposite charges in complementary surface patterns.[24] For the situation of protein monomers binding DNA, recognition of negatively-charged DNA is accomplished by positively-

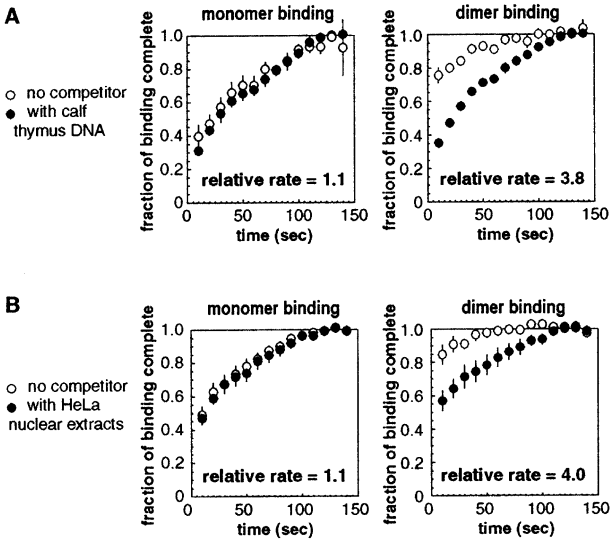

Figure 8. Results of the competition experiment. The relative rates at which Max$_{22-113}$ (monomer binding) and Max$^{SS}_{22-105}$ (dimer binding) bound to ^{32}P-Ebox$_{22}$ were measured in the absence (○) and presence (●) of nucleic acid competitors. Competitors used were (A) 500 nM base pairs calf thymus DNA, and (B) 22.7 μg/mL HeLa nuclear extracts.

charged, but unfolded, proteins that are unlikely to present a fully organized and complementary DNA binding surface. However, another explanation of the kinetic advantage of DNA recognition prior to dimer formation relies on the unfolded state of protein monomers.[25] This explanation takes note of the fact that an unstructured protein has a greater range of motion and, therefore, may have a greater capture radius than its folded counterpart. As a result, the unfolded protein will bind its DNA target more weakly, but will be able to contact that target from a larger distance. Upon locating its target site, the protein monomer is 'reeled in' to simultaneously complete binding and folding.[25] This mechanism for target site binding predicts that binding by unfolded monomers will result in (1) an increased rate of complex formation due to the greater capture radius, and (2) improved kinetic discrimination due to guidance of binding events by the folding funnel.

Our experiments with Max showed that DNA binding by protein monomers provided the most rapid route to ternary complex formation. We also found that improved kinetic discrimination resulted. In a system in which protein monomers recognized DNA prior to forming dimers, the addition of nucleic acid competitor had little effect on the overall rate of ternary complex formation. In contrast, when the same system was engineered such that dimer formation preceded DNA binding, nucleic acid competitors significantly slowed formation of the ternary complex. We interpret the improved kinetic discrimination as reflecting the longer lifetime of complexes formed from dimeric protein and nonspecific DNA partners when compared to those complexes formed from protein monomers and nonspecific DNA. Since the rate of transcription initiation is believed to be under kinetic control,[26,27] formation of transcription factor–DNA complexes with high levels of specificity even at times prior to equilibrium will likely provide maximal fidelity in transcriptional responses.

We observe that oligomerization of DNA-binding subunits upon DNA binding, rather than prior to DNA binding, improves kinetic discrimination and decreases the time necessary to achieve equilibrium. This result has important implications for the design of DNA-binding molecules intended for artificial gene regulation. Recent efforts have focused on tethering together individual DNA recognition units and have produced designed proteins capable of recognizing long (18 and 30 base pair) target sites with high affinity.[28–31] With binding sites of this length, contemporary designed transcription factors approach the goal of being able to uniquely target any site within the human genome.[32]

High specificity has proved to be a more elusive quality to attain through design, although recent progress has been made by careful tailoring of the linkers between individual recognition units.[33] Unfortunately, even when achieving specific recognition, many of the highest affinity molecules are also distinguished by their requirement for long (from 48 to greater than 600 h) equilibration times in order to specifically locate the preferred DNA target site in the presence of competitor.[28,29] Future design efforts may consider ways to couple DNA binding by individual recognition elements[34,35] to their oligomerization (as suggested in ref 36). Such artificial transcription factors would be kinetically trapped only at their specific DNA target sites. They would be expected to require shorter equilibration times and would be less likely to produce unintended transcriptional responses. Indeed, improved RNA recognition by peptide dimers has been achieved by coupling RNA recognition with folding of the dimerization interface.[37]

The improved kinetic specificity we observe is likely to be most apparent for larger protein–nucleic acid interfaces, particularly those that span both sides of a DNA helix (the DNA target sites studied here range from 6 to 10 base pairs). For smaller DNA recognition elements, other factors may influence the kinetics of specific recognition. For instance, preorganization of small DNA binding elements allows for fast and specific DNA recognition.[35,38]

The concept that protein monomers possess a more efficient DNA search mechanism than the corresponding protein dimers may provide insight into the storage mechanism for transcription factors that are constitutively expressed. The presence of high DNA concentrations in the nucleus will, through the effects of mass action, favor formation of complexes composed of protein monomers bound to DNA. Interestingly, many genes are regulated by the effects of heterodimeric transcription factors in which one member of the pair is ubiquitously expressed and the other member's expression patterns correlate with a specific cellular state. (The bHLHZip proteins Max and Myc provide an example, see ref 39) Constitutive expression of only one member of a heterodimer may allow it to bind DNA quickly, but would force it to remain as a kinetically unstable species, due to the absence of the dimerizing partner. Transient expression of an inducible transcription factor might result in capture of the monomer–DNA species and lead to formation of the stable, ternary complex. In this case, the intermediate along the DNA binding pathway would be a monomer–DNA complex rather than a dimeric transcription factor, as assumed previously. Attempts to precisely modulate the effects of transcription factors may be aided by a knowledge of the intermediates relevant to in vivo complex formation.

Conclusions

bZIP proteins have been shown to form complexes with DNA consisting of a single protein monomer bound to DNA. These monomer—DNA complexes are the intermediate species along the pathway proposed to account for formation of the final, dimer—DNA complex. The proposed monomer binding pathway allows for rapid formation of dimer—DNA complexes. Polyanions other than DNA can accelerate the rate of dimer formation but polycations and neutral polymers fail to do so. Finally, proteins that bind DNA through a monomer—DNA intermediate are able to rapidly find their

specific DNA target site in the presence of nonspecific nucleic acids.

Experimental

DNA

The DNA duplex CRE_{24} contained the following sequence: 5′-AGTGGAGATGACGTCATCTCGTGC-3′ and its complementary strand. The DNA duplex $AP1_{23}$ contained the following sequence: 5′-AGTGGAGATG-ACTCATCTCGTGC-3′ and its complementary strand. The DNA duplex $Ebox_{22}$ contained the following sequence 5′-GTGTAGGCCACGTGACCGGGTG-3′ and its complementary strand. Oligonucleotides were synthesized on a 0.2 μmol scale at the HHMI Biopolymer/Keck Foundation Biotechnology Resource Laboratory (Yale University School of Medicine, New Haven, CT). Crude oligonucleotides were purified by denaturing gel electrophoresis. For electrophoretic mobility shift assays, one strand of the DNA duplex was labeled at the 5′ end with γ-^{32}P-ATP. Complementary strands were annealed together to form duplex DNA.

Calf thymus DNA was purchased from Gibco. Poly-L-glutamic acid, heparin sulfate, poly-L-lysine, and polyethylene glycol were purchased from Sigma.

Peptides

A_{71}, JunC, and FosC are peptides comprising the bZIP elements of the proteins ATF-2, Jun and Fos, respectively. All peptide constructs have mutations of their DNA-contact cysteines to serine and introduce C-terminal cysteines for covalent modification with fluorophores. A_{71}, JunC and FosC were expressed and purified as described previously.[3,14] The peptides were fluorescently labeled at their unique cysteines using either 5-iodoacetamidofluorescein or 5-iodoacetamidotetramethylrhodamine (Molecular Probes) and then purified by HPLC, as described previously.[3,14]

$Max_{22–113}$ and $Max^{SH}_{22–105}$ were expressed in *Escherichia coli* and purified as described previously.[2] Covalent dimers of Max ($Max^{SS}_{22–105}$) were produced by cystine formation between two $Max^{SH}_{22–105}$ monomers and were purified as described previously.[2]

Fluorescence polarization measurements

Fluorescence polarization measurements were conducted on a Panvera Beacon instrument equipped with fluorescein filters (488 nm excitation, 535 nm emission). Samples (100 μL) were prepared in 1×PBS buffer (1.4 mM KH_2PO_4, 4.3 mM Na_2HPO_4, 2.7 mM KCl 137 mM NaCl, pH 7.4) and then transferred to 6×50 mm Beacon disposable test tubes. Polarization was measured for 1 min. All sample readings were performed in triplicate and corrected for background fluorescence by measuring the fluorescence polarization of identical samples lacking A^{SFlu}_{71}. Data shown represent the average of five independent experiments; error bars indicate standard deviation.

Equilibrium fluorescence measurements

Fluorescence emission scans were measured on a PTI QuantaMaster C-60 Spectrofluorimeter. Samples were prepared in 1X PBS buffer, then incubated at 25 °C for 20 min before transfer to a 1 cm pathlength cuvette (Hellma) for acquisition of fluorescence spectra. The excitation wavelength was set to 490 nm (4 nm slit width) and the fluorescence emission was measured at 1 nm intervals from 500 to 600 nm (4 nm slit width) with a 1 s sampling time. The spectra shown represent the averages of three scans of individual samples. The absorbance of buffer and/or nucleic acid components at 490 nm was always less than 0.005 absorbance units. Fluorescence resonance energy transfer to rhodamine was verified by a fluorescence increase at 580 nm. However, as observed by others,[20,21] the rhodamine emission (signal at 580 nm) was small relative to the fluorescein emission and resulted both from resonance transfer of energy from the fluorescein donor and from direct absorption of the 490 nm incident light. In both equilibrium and kinetic fluorescence measurements the decrease in emission at 520 nm, rather then the increase at 580 nm, was used quantify $FosC^{SRho}$–$JunC^{SFlu}$ complex formation. We note that the absorption spectrum of tetramethylrhodamine changes significantly upon conjugation to FosC, with a loss in intensity at 555 nm and an additional peak appearing at 520 nm. Similar changes in absorption spectra upon conjugation of tetramethylrhodamine to proteins have been observed[40] and will likely lead to increased direct excitation of tetramethylrhodamine (due to the increased absorbance at 520 nm) as well as decreased transfer from fluorescein (due to the decreased absorbance at 555 nm).

Stopped-flow fluorescence measurements

Stopped-flow fluorescence experiments were performed on a KinTek SF-2001 stopped-flow spectrophotometer. Fluorescein was excited at 495 nm using a 6 nm slit width and a Xenon short arc lamp light source (Ushio, Japan). The time-dependent change in fluorescein emission was monitored at 520 nm, using a 10 nm bandpass filter (Corion; Franklin, MA). $JunC^{SFlu}$ and $FosC^{SRho}$ were each incubated in 1X PBS and loaded into two 5 mL syringes through their individual loading ports. A third (0.5 mL) syringe was loaded with 1X PBS, nucleic acid, or other polymer. Three-way mixing was initiated by injecting samples through a microjet mixer at a flow rate of 6 mL/s before the reaction mixture reached the 20 μL observation cell. A high-torque stepper motor drive with dead time of less than 1 msec was used to control the injection. The final concentration of $JunC^{SFlu}$ was 20 nM and that of $FosC^{SRho}$ was 100 nM. Data shown represent the average of three to five individual traces. The averaged kinetic traces were fit to a single (Observed fluorescence = A_1*exp(−k_1*time) + equilibrium fluorescence) or double exponential (Observed fluorescence = A_1*exp(−k_1*time) + A_2*exp(−k_2*time) + equilibrium fluorescence) using a nonlinear least squares algorithm to determine the rates (R_1 and R_2) and amplitudes (A_1 and A_2).

Electrophoretic mobility shift experiments

6 nM $Max_{22–113}$ and 2 nM $Max^{SS}_{22–105}$ were incubated together for 30 min at 25 °C in 1X Binding Buffer (1X PBS, 1 mM EDTA, 0.1% NP-40, 5% glycerol). Binding was initiated by the addition of ^{32}P-$Ebox_{22}$ to a final concentration of 2 nM. The course of the reaction was followed by loading aliquots on a running gel at 10 s intervals. In the competition experiment, the proteins were incubated together with either 500 nM calf thymus DNA (concentration in base pairs) or with 22.7 μg/mL HeLa nuclear extracts (Promega). The competition binding reaction was then challenged by the addition of ^{32}P-$Ebox_{22}$ to a final concentration of 2 nM. The progress of the binding reaction was analyzed at 10 second intervals by gel electrophoresis. Two complexes were visible on both sets of gels: $Max^{SS}_{22–105}$–^{32}P-$Ebox_{22}$ and $(Max_{22–113})_2$–^{32}P-$Ebox_{22}$. The amount of each complex formed was normalized to equal a value of 1 at equilibrium. The fraction of equilibrium binding was calculated at each time point for each protein–DNA complex and was plotted versus time. Data shown represents an average of four independent trials.

Acknowledgements

This work was supported by the NIH (GM 52544). J.J.K. was supported by a NSF graduate research fellowship and by a predoctoral fellowship provided by Wyeth-Ayerst Research through the American Chemical Society, Division of Medicinal Chemistry. We thank Tanya Schneider for the clone encoding A71, Lin Chen and Steven Harrison for supplying the clones used to generate Fos and Jun expression vectors, and Satish Nair and Stephen Burley for the clones encoding $Max_{22–113}$ and $Max^{SH}_{22–105}$. We thank Stacey Rutledge for comments on the manuscript. Fluorescence polarization data was acquired in the laboratory of Paul Sigler.

References and Notes

1. Metallo, S. J.; Schepartz, A. *Nat. Struct. Biol.* **1997**, *4*, 115.
2. Kohler, J. J.; Metallo, S. J.; Schneider, T. L.; Schepartz, A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11735.
3. Kohler, J. J.; Schepartz, A. *Biochemistry* **2001**, *40*, 130.
4. Kim, B.; Little, J. W. *Science* **1992**, *255*, 203.
5. Berger, C.; Piubelli, L.; Haditsch, U.; Bosshard, H. R. *FEBS Lett.* **1998**, *425*, 14.
6. Wendt, H.; Thomas, R. M.; Ellenberger, T. *J. Biol. Chem.* **1998**, *273*, 5735.
7. Wu, X. L.; Spiro, C.; Owen, W. G.; McMurray, C. T. *J. Biol. Chem.* **1998**, *273*, 20820.
8. Rentzeperis, D.; Jonsson, T.; Sauer, R. T. *Nat. Struct. Biol.* **1999**, *6*, 569.
9. Chinenov, Y.; Henzl, M.; Martin, M. E. *J. Biol. Chem.* **2000**, *275*, 7749.
10. Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* **1992**, *71*, 1223.
11. Konig, P.; Richmond, T. J. *J. Mol. Biol.* **1993**, *233*, 139.
12. Park, C.; Campbell, J. L.; Goddard, W. A. *J. Am. Chem. Soc.* **1996**, *118*, 4235.
13. Cao, W.; Liu, L.; Klein, D. E.; Wei, L. Y.; Lai, L. H. *Thermochim. Acta* **2000**, *360*, 47.
14. Schneider, T. L.; Schepartz, A. *Biochemistry* **2001**, *40*, 2835.
15. Bray, D.; Lay, S. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13493.
16. Karin, M.; Liu, Z. G.; Zandi, E. *Curr. Opin. Cell Biol.* **1997**, *9*, 240.
17. O'Shea, E. K.; Rutkowski, R.; Stafford, W. F.; Kim, P. S. *Science* **1989**, *245*, 646.
18. Cohen, D. R.; Curran, T. *Oncogene* **1990**, *5*, 929.
19. Rauscher, F. J.; Voulalas, P. J.; Franza, B. R.; Curran, T. *Genes Dev.* **1988**, *2*, 1687.
20. Parkhurst, K. M.; Brenowitz, M.; Parkhurst, L. J. *Biochemistry* **1996**, *35*, 7459.
21. Perkins, T. A.; Wolf, D. E.; Goodchild, J. *Biochemistry* **1996**, *35*, 16370.
22. Ferre-D'Amare, A. R.; Prendergast, G. C.; Ziff, E. B.; Burley, S. K. *Nature* **1993**, *363*, 38.
23. Schreiber, G.; Fersht, A. R. *Nat. Struct. Biol.* **1996**, *3*, 427.
24. Selzer, T.; Albeck, S.; Schreiber, G. *Nat. Struct. Biol.* **2000**, *7*, 537.
25. Shoemaker, B. A.; Portman, J. J.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8868.
26. Li, X. Y.; Virbasius, A.; Zhu, X. C.; Green, M. R. *Nature* **1999**, *399*, 605.
27. Kuras, L.; Struhl, K. *Nature* **1999**, *399*, 609.
28. Kim, J.-S.; Pabo, C. O. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2812.
29. Kamiuchi, T.; Abe, E.; Imanishi, M.; Kaji, T.; Nagaoka, M.; Sugiura, Y. *Biochemistry* **1998**, *37*, 13827.
30. Beerli, R. R.; Segal, D. J.; Dreier, B.; Barbas, C. F. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14628.
31. Moore, M.; Choo, Y.; Klug, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 1432.
32. Dervan, P. B. *Science* **1986**, *232*, 464.
33. Moore, M.; Klug, A.; Choo, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 1437.
34. Zondlo, N. J.; Schepartz, A. *J. Am. Chem. Soc.* **1999**, *121*, 6938.
35. Chin, J. W.; Schepartz, A. *J. Am. Chem. Soc.* **2001**, *123*, 2929.
36. Pomerantz, J. L.; Wolfe, S. A.; Pabo, C. O. *Biochemistry* **1998**, *37*, 965.
37. Campisi, D. M.; Calabro, V.; Frankel, A. D. *Embo J.* **2001**, *20*, 178.
38. Baliga, R.; Baird, E. E.; Herman, D. M.; Melander, C.; Dervan, P. B.; Crothers, D. M. *Biochemistry* **2001**, *40*, 3.
39. Littlewood, T. D.; Evan, G. I. *Protein Profile* **1994**, *1*, 639.
40. Ravdin, P.; Axelrod, D. *Anal. Biochem.* **1977**, *80*, 585.