Structure-Based Design of a Leucine Zipper Protein with New DNA Contacting Region†

Takashi Morii,*,‡,§ Shin-ichi Sato,‡ Masaki Hagihara,‡ Yasuo Mori,|| Keiji Imoto,|| and Keisuke Makino‡

Institute of Advanced Energy, Kyoto University, and PRESTO, Japan Science and Technology Corporation, Uji, Kyoto 611-0011, Japan, and Department of Information Physiology, National Institute for Physiological Sciences, Okazaki, Aichi 444-8585, Japan

Received October 1, 2001; Revised Manuscript Received December 19, 2001

ABSTRACT: We have employed a structure-based design to construct a small folding domain from the F-actin bundling protein villin that contains the amino acids necessary for the DNA binding of the basic leucine zipper protein GCN4 and have compared its DNA binding with GCN4. The monomeric motif folds into a stable domain and binds DNA in a rigid-body mechanism, while its affinity is not higher than that of the basic region peptide. The addition of the leucine zipper region to the folded domain restored its sequence-specific DNA binding comparable to that of GCN4. Unlike the monomeric folded domain, its leucine zipper derivative undergoes a conformational change upon DNA binding. CD spectral and thermodynamic studies indicate that the DNA-contacting region is folded in the presence or absence of DNA and suggest that the junction between the DNA-contacting and the leucine zipper regions transits to a helix in the presence of DNA. These results demonstrate that the structural transition outside the direct-contacting region, which adjusts the precise location of the DNA-contacting region, plays a critical role in the specific complex formation of basic leucine zipper proteins.

DNA binding proteins frequently undergo conformational changes upon associating with specific DNA sequences. These conformational changes include both the global quaternary rearrangements of the domains and the folding transitions of local residues (*1*). Most of the local folding transitions occur on the major groove of DNA with a flexible domain changing its structure into an α helix, which serves as a critical DNA recognition element in many DNA binding proteins (*2*). These flexible domains are folded into a stable α-helical domain only upon binding to the specific DNA sequence. Such coupling of the α-helix formation to the site-specific protein−DNA interactions can be represented by the example of the yeast transcription factor GCN4 (*1*, *3*−*5*) that has been proposed as a model for the "induced-fit" binding. GCN4 is one of a large family of DNA binding proteins that is featured by the presence of a basic leucine zipper (bZIP)[1] motif (*6*−*8*). When a GCN4 dimer binds a specific DNA sequence, its leucine zipper region assembles into a parallel coiled-coil, while its direct DNA-contacting region, so-called "the basic region", and the hinge region spanning between them undergo a structural transition from a nascent helix to a complete α helix (*9*, *10*).

Several groups have reported the sequence-specific DNA binding and structural transition of monomeric basic region peptides (*11*, *12*) or their artificially assembled dimers (*13*−*19*). Although these model studies allowed for detailed thermodynamic analyses of the basic region−DNA interactions, all of the model peptides exhibited distinct structural transitions at the direct DNA-contacting region. The stable specific-DNA binding complexes of these peptides formed only at low temperatures, perhaps because of large entropic losses of the induced-fit binding at higher temperature. Such a thermodynamically unfavorable binding mechanism of the basic region raises the question as to whether the structural transition at the basic region itself plays a physiological role in a high degree of sequence discrimination by the basic leucine zipper and basic helix−loop−helix families of transcription factors (*20*, *21*).

Structure-based approaches have been successfully used to design novel DNA or RNA binding domains (*22*−*25*). A stably folded small domain containing α helices would provide an ideal scaffold for a folded variant of the basic region peptide that allows us to study the role of structural transition at the protein−DNA interface. A miniature protein designed by grafting the recognition residues of GCN4 onto the avian pancreatic peptide efficiently formed a specific DNA complex at the physiological ionic strength (*24*, *25*). The specific binding of the peptide was observed at 4 °C, associating with a structural transition; however, a dimer of the peptide showed poor specificity (*24*). We report here that a structure-based design using a small domain from the F-actin bundling protein villin (*26*−*28*) affords a compact domain containing the amino acids necessary for the DNA binding of the basic leucine zipper protein GCN4. The folded

† This work was supported in part by the Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan to T.M. (Grant No. 12680590) and to K.M. (Grant No. 11101001).

* To whom correspondence should be addressed. Email: t-morii@iae.kyoto-u.ac.jp.

‡ Kyoto University.

§ PRESTO, Japan Science and Technology Corporation.

|| National Institute of Physiological Sciences.

[1] Abbreviations: bZIP, basic leucine zipper; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; BSA, bovine serum albumin; CD, circular dichroism.

10.1021/bi015820i CCC: $22.00
Published on Web 01/24/2002

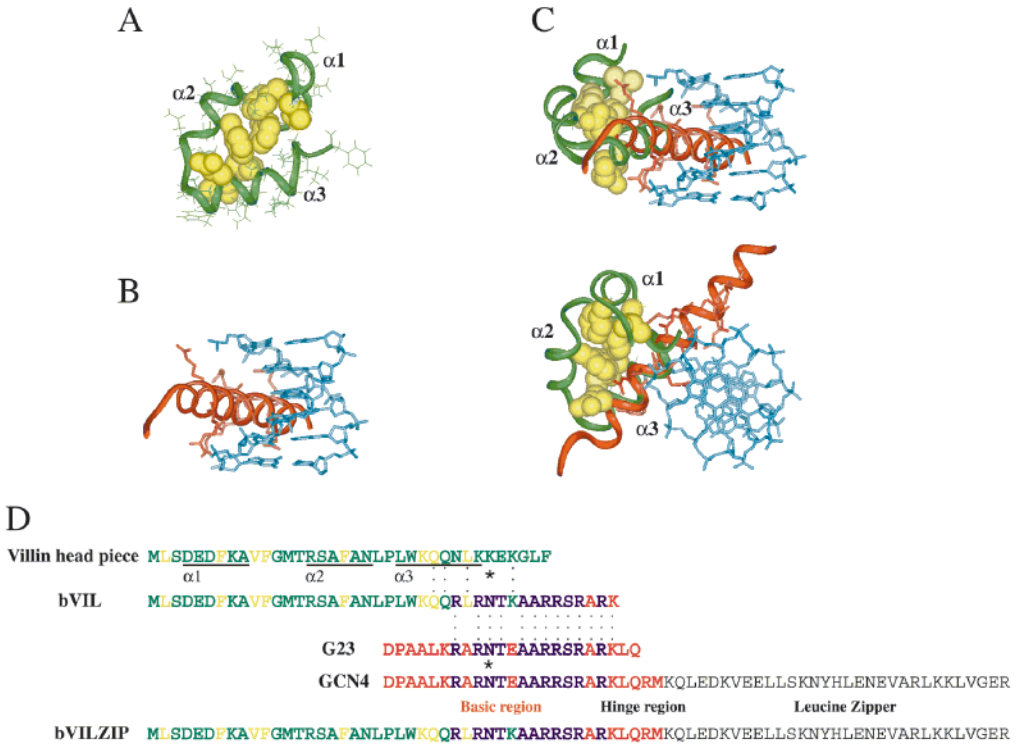

FIGURE 1: Strategy for the structure-based design of folded DNA binding region. The villin headpiece domain (*27*) (A) and an X-ray crystal structure for the complex between GCN4 basic region and AP1 (*9*) (B) are shown in the ribbon representations. The backbone of villin headpiece domain is colored in green, and GCN4 is shown in red. The side chains necessary to form a hydrophobic core of villin headpiece are indicated by CPK presentation (yellow). (C) The superimposition of the backbone atoms of the villin third helix and the GCN4 basic region are shown from different angles. (D) Amino acid sequences for villin, bVIL, GCN4, G23, and bVILZIP are aligned with Lys71 (*) of villin and Asn235 (*) of GCN4. The three helical regions of villin are underlined. Amino acid residues necessary for the hydrophobic core formation of villin, and for the specific DNA recognition by GCN4, are shown in yellow and blue, respectively.

variant of the GCN4 basic region binds the GCN4 binding sequence with little conformational change, while its affinity and selectivity are not higher than that of the basic region peptide. The addition of the leucine zipper to the folded basic region variant restored its sequence-selective DNA binding, which accompanies a structural transition most likely at the hinge between the DNA-contacting and dimerization regions. The structural transition at the hinge region, rather than that at the protein−DNA interface, would play a critical role in locating the DNA-contacting regions to optimize the sequence-specific DNA binding by a basic leucine zipper protein.

MATERIALS AND METHODS

Preparation of the Proteins. Modeling studies of proteins were performed by using the structures of the GCN4−AP1 complex (*9*) and the villin headpiece subdomain (*27*) on a Discover/Insight II package (MSI). Synthetic oligonucleotides encoding bVIL were cloned into the plasmid pET3b. Plasmids encoding bVILZIP and GCN4 were prepared with p4LZ'91 (*29*). The resulting plasmids were transformed into *Escherichia coli* BL21(DE3)-pLysS cells, and the proteins were purified by using Mono S cation exchange and Resource RPC reversed-phase chromatography (Amersham Pharmacia). Molecular weights were confirmed by a MALDI-TOF mass spectrometer Voyager DE-STR (PerSeptive Biosystems) ((bVIL) MW calcd, 4953.8; found, 4954.1; (bVILZIP) MW calcd, 9173.7; found, 9173.8). Amino acid sequences for bVIL and bVILZIP are shown in Figure 1D.

Gel Mobility Shift Assays of Proteins. The binding reaction was performed in the presence of the indicated amount of the protein with 20 pM 5′-^{32}P-labeled oligonucleotide duplex in binding mixture containing 20 mM Tris-HCl (pH 7.5), 100 mM KCl (4 mM for G23 and bVIL), 2 mM $MgCl_2$, 1 mM EDTA, 100 μg/mL BSA, 0.1% NP-40, and 6% sucrose at the indicated temperatures as described previously (*11*). The increase of the mobility-shifted band was quantitated by a Storm phosphor imager (Amersham Pharmacia). The fraction of unbound DNA was obtained by dividing the intensity of the mobility-shifted band with the sum of the intensities of both the bound and unbound bands. The gel mobility shift competition assay was carried out as follows. The binding reaction was performed in the presence of bVIL (1.5 μM) and the indicated amount of competitor oligonucleotide with 10 pM 5′-^{32}P-labeled HS in a binding mixture containing 20 mM Tris-HCl (pH 7.5), 4 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 100 μg/mL acetylated BSA, 0.1% NP-40, and 6% sucrose and analyzed as described previously (*16*). Nucleotide sequences of the oligonucleotides used in the present study are AP1 (5′-CGGATGACTCATTTTTTTC-3′), CRE (5′-CGGATGACGTCATTTTTTTTC-3′), HS (5′-CGGATGACACTGCTTTTTTTC-3′), and NON (5′-GATCCCCCCAACACCTGCTGCCTGA-3′).

Temperature-Dependent DNA Binding Analyses Using the Surface Plasmon Resonance. A 5′-biotinylated oligonucleotide duplex solution was injected over a streptavidin-coated sensor chip (SA5; Pharmacia Biosensor) until a suitable level (∼100 RU) was achieved on BIAcore 2000 (BIAcore Inc.). Tris-HCl buffer (20 mM Tris-HCl, 200 mM KCl, 10 mM $MgCl_2$, 2 mM EDTA, and 0.005% Tween20 (pH 7.5)) was used both as flow buffer and as a sample preparation buffer.

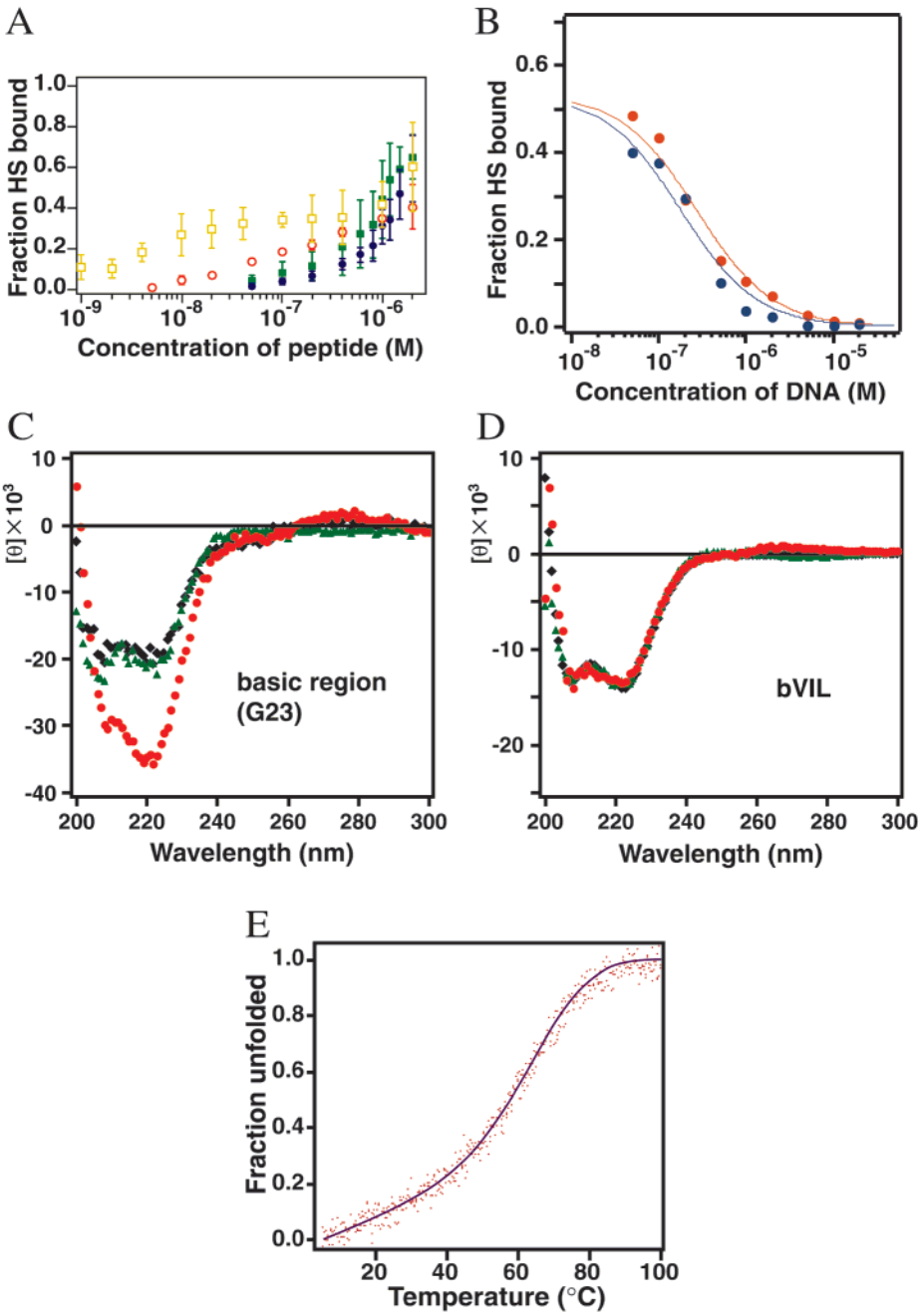

FIGURE 2: DNA binding and circular dichroismic spectral analyses of the GCN4 basic region peptide (G23) and bVIL. (A) Semilogarithmic plots showing the fraction of HS bound to G23 (open symbols) and bVIL (filled symbols) at 4 °C (squares) and 20 °C (circles). (B) Gel shift competition analysis of the relative affinities of bVIL to HS (filled circles in blue) NON (filled circles in red) oligonucleotides. Semilogarithmic plots show the fraction of ^{32}P-labeled HS bound to bVIL as a function of added competitor DNA. The solid curves are the best fit binding titration isotherm obtained from a nonlinear least-squares algorithm as described in Materials and Methods. Circular dichroism difference spectra for the GCN4 basic region peptide (G23) (C) and bVIL (D) show the helical conformation of peptides in the presence of recognition DNA. Spectra in the absence of DNA (filled squares in black) and in the presence of AP1 (filled circles in red) and NON (filled triangles in green). (E) The temperature-unfolding curve of bVIL was obtained by monitoring mean residue ellipticity at 222 nm at the concentration of peptide 100 μM from 4 to 100 °C.

Bound protein was eluted from the DNA by three repeats of a short pulse (5 μL) of 3 M KCl. The amount of protein-bound DNA was obtained at the equilibrium state of association reaction (*19*). Analysis of the data was performed using the evaluation software supplied with the instrument (BIAevaluation, version 3.0).

Circular Dichroismic Spectral Studies. CD experiments were performed on a J-725 CD spectrometer using a 0.1 cm path length cell, as described previously (*11*, *16*). Samples contained 20 mM Tris-HCl (pH 7.5), 100 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 0.1% NP-40, 40 μM bVIL, or G23 monomer (8 μM bVILZIP or GCN4 monomer), and 50 μM oligonucleotide duplex (5 μM for bVILZIP or GCN4) when present. Temperature scans were performed with 20 or 100 μM protein by scanning continuously from 4 to 100 °C for bVIL and 4 to 80 °C for bVILZIP at 222 nm. The thermal unfolding curves were baseline-corrected by subtraction of a buffer-only control experiment. The data have been normalized to fraction unfolded by using the [θ] values at 4 °C for fully folded and values at 100 °C for fully unfolded,

respectively. In all cases, the thermal unfolding was over 95% reversible, and the T_m values did not change after repeated thermal unfolding.

RESULTS

Design of a Folded Basic Region Peptide. Our design strategy utilizes a structurally well-characterized small folded domain of the villin headpiece subdomain (*26*−*28*) as a scaffold for a new DNA binding region (Figure 1A). The F-actin bundling protein villin contains an autonomously folding 35-residue subdomain at its C-terminus. This 35-residue folding domain, termed as HP-35, forms a novel three-helix structure and is thermostable with a midpoint of the thermal unfolding transition (T_m) of 70 °C (*26*). The "HP-35-fold" consists of three short helices surrounding a tightly packed hydrophobic core (*27*, *28*). By using the NMR structure of villin headpiece subdomain (*27*) (Figure 1A) and the crystal structure of the GCN4−AP1 complex (*9*) (Figure 1B), we aligned two helices, one from the basic region of GCN4 and the other from the C-terminal helix of HP-35, to locate the amino acid residues required for the hydrophobic core formation of HP-35 within the basic region helix (Figure 1C). When Lys71 of HP-35 was superimposed on Asn235 of GCN4 (Figure 1D), the first and second helices of HP-35, including the hydrophobic core, were placed outside the DNA major groove without disturbing the positions of amino acid residues necessary for the sequence-selective binding of GCN4 and without causing obvious steric clashes to the DNA phosphate backbone (Figure 1C). The resulting monomeric variant of the GCN4 basic region (bVIL) and its leucine zipper derivative bVILZIP were prepared from plasmids encoding the respective amino acid sequences (Figure 1D).

Structure and DNA Binding of the Monomeric bVIL Peptide. DNA binding of bVIL was tested by titration of the gel shift and compared with that the GCN4 basic region peptide (G23). G23 formed a binding complex with the half-site of GCN4 binding sequence (HS:5′-ATGAC-3′) more efficiently than bVIL did at 4 °C (Figure 2A) but appeared to behave in a biphasic fashion that was not easily interpreted. The complex between G23 and HS was unstable at 20 °C, as reported previously (*11*, *16*). While the complexes between bVIL and HS retained similar stability at 4 and 20 °C with dissociation constants of ca. 2 μM, bVIL bound more weakly at either temperature than G23 did. However, bVIL failed to establish a distinct selectivity to HS by the gel shift competition experiments using a nonrelated sequence (NON) (Figure 2B).

The CD spectrum of bVIL showed two minima at 208 and 222 nm (Figure 2D), and its mean residue ellipticity at 222 nm (−14 000) was virtually identical to that reported for HP-35 (*26*), indicating that bVIL had a folded structure similar to that of HP-35. The conformational stability of bVIL was measured by monitoring the CD signal at 222 nm as the temperature increased from 4 to 100 °C. Analysis of the thermal unfolding behavior for bVIL afforded a reversible melting curve (Figure 2E) with a T_m of 66 °C, which was slightly lower than that reported for HP-35 (*26*). The structure of bVIL in the presence of DNA was next studied by comparing circular dichroism differential spectra. Each difference spectrum of bVIL (40 μM) in the presence of AP1

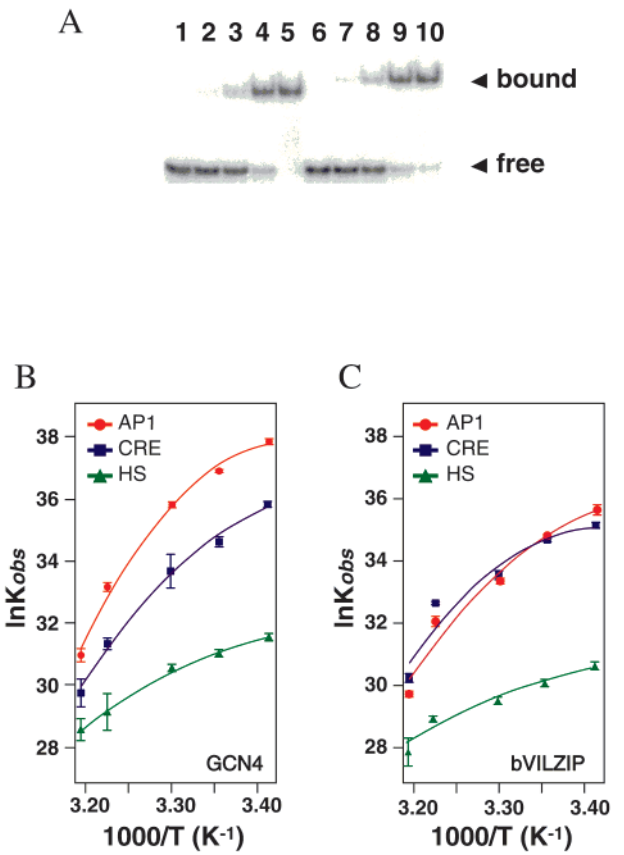

FIGURE 3: DNA binding of bVILZIP. (A) Autoradiograph showing the titration of gel shifts for GCN4 and bVILZIP to the AP1 DNA. In lanes 1 and 6, no protein was added to the reaction. In lanes 2−5 and 7−10, the concentrations of proteins were as follows: 0.1, 0.4, 1, and 8 nM (lanes 2−5, GCN4) and 0.4, 1, 4, and 20 nM (lanes 7−10, bVILZIP), respectively. The van't Hoff plots for binding of GCN4 (B) and bVILZIP (C) to the AP1 (red circles), CRE (blue squares) and HS (green triangles) sequences. The best fit for each plot was obtained by a nonlinear function (*30*, *31*), as described in the text.

(5′-ATGACTCAT-3′) or NON remained unchanged from that in the absence of DNA (50 μM) (Figure 2D). The same results were obtained when 4 μM bVIL and 50 μM DNA were used for the measurements. This observation was in sharp contrast to the case of G23 (Figure 2C) that exhibited a large increase in the intensity of helical band with AP1 but not with NON (*13*, *14*, *17*). These results suggest that bVIL folds in a stable structure and binds DNA with a rigid-body mechanism.

DNA Binding of the bVILZIP Dimer. DNA binding of bVILZIP, the bVIL domain with a leucine zipper region (Figure 1D), was next compared to that of GCN4. Gel mobility shift assays revealed that bVILZIP formed a distinct slow-migrating band with AP1 in an affinity comparable to GCN4 (Figure 3A). The apparent equilibrium binding constants for the specific (AP1 and CRE:5′-ATGACGTCAT-3′) and nonspecific DNA (HS) complex formation by GCN4 or bVILZIP were measured at various temperatures ranging from 20 to 40 °C (Table 1) and were shown as the van't Hoff plot (Figure 3B,C). The plots were nonlinear, and the data were fitted to the second-order function of van't Hoff equation (*30*, *31*).

Thermodynamic parameters obtained from the van't Hoff plot analysis of the binding constants for GCN4 and bVILZIP

Table 1: Summary of Thermodynamic Parameters for GCN4 and bVILZIP Binding to the Specific and Nonspecific DNA

proteins	DNA	temp (K)	ΔG^0obs[a,b]	ΔH^0obs[b,c]	$-T\Delta S^0$obs[b,c]	ΔC_p^0,obs[b,d]
GCN4	AP1	293	-22.1 ± 0.1	-7.2 ± 1.3	-14.9 ± 1.2	-5.3 ± 0.2
		310	-20.5 ± 0.1	-97.7 ± 2.3	77.2 ± 2.4	
	CRE	293	-20.9 ± 0.1	-20.7 ± 8.1	-0.2 ± 8.2	-3.2 ± 0.6
		310	-19.3 ± 0.1	-76.0 ± 1.5	56.7 ± 1.4	
	HS	293	-18.4 ± 0.1	-11.7 ± 0.8	-6.7 ± 0.7	-1.5 ± 0.4
		310	-18.0 ± 0.4	-38.1 ± 6.6	20.2 ± 6.9	
bVILZIP	AP1	293	-20.8 ± 0.1	-17.2 ± 2.7	-3.6 ± 2.6	-3.2 ± 0.3
		310	-19.8 ± 0.1	-72.7 ± 1.7	52.9 ± 1.8	
	CRE	293	-20.5 ± 0.1	-0.8 ± 0.7	-19.7 ± 0.6	-3.9 ± 0.1
		310	-20.1 ± 0.1	-67.2 ± 1.0	47.1 ± 1.0	
	HS	293	-17.9 ± 0.1	-10.9 ± 5.3	-7.0 ± 5.3	-1.2 ± 0.7
		310	-17.9 ± 0.1	-30.9 ± 6.7	13.0 ± 6.7	

[a] The values for ΔG^0obs correspond to the equilibrium binding constant measurements at the indicated temperature. The equilibrium binding constants are the average of at least three determinations. [b] The values were obtained as described in Materials and Methods. [c] Units for ΔG^0obs, ΔH^0obs, and $-T\Delta S^0$obs are kcal mol^{-1}. [d] Units for ΔC_p^0,obs are kcal mol^{-1} K^{-1}.

are summarized in Table 1. As reported previously (*32*), GCN4 preferentially bound AP1 over CRE with one order higher affinity at 20 °C (Table 1 and Figure 3B), whereas bVILZIP bound AP1 and CRE in an almost equal affinity (Table 1 and Figure 3C). Both GCN4 and bVILZIP apparently discriminated the HS from AP1 and CRE. At 37 °C, bVILZIP selectively bound CRE over HS by $\Delta\Delta G = -2.2$ kcal mol^{-1}, which was comparable to the selectivity of GCN4 between AP1 and HS ($\Delta\Delta G = -2.5$ kcal mol^{-1}). With increasing temperature from 20 to 40 °C, the stability of the GCN4−AP1 complex decreased by 1.6 kcal mol^{-1}, which was larger than the decrease observed for the bVILZIP−CRE complex (0.4 kcal mol^{-1}). In contrast, the stabilities of nonspecific complexes of GCN4 and bVILZIP were similar each other within that temperature range (Table 1).

DNA Binding Mode of the bVILZIP Dimer. The leucine zipper dimer bVILZIP bound the specific sequence in an affinity comparable to GCN4 at physiological temperature. We next analyzed the structural aspects bVILZIP in the absence or presence of DNA. The thermal melting curve of GCN4 represented two regions: a region from 4 to 40 °C, displaying a gradual increase of the fraction unfolded, and a region showing a sharp transition centered at 50 °C that would correspond to the melting of coiled-coil (Figure 4A). The thermal melting curve of bVILZIP revealed a more cooperative nature than that of GCN4, and bVILZIP has a higher secondary structure content than GCN4 at the temperature range from 4 to 40 °C, which in turn suggests that the bVIL domain in bVILZIP is stabilized by the hydrophobic core formed in the similar manner for the monomeric bVIL or HP-35.

The intensity of helical band for bVILZIP was higher than that of GCN4 in the absence of DNA (Figure 4B,C). Because the CD signals for both bVILZIP and GCN4 remained constant above the monomer concentration of 5 μM, both proteins were expected to form dimers exclusively in the condition used for CD measurements. The helical residues of free GCN4 dimer are mostly assignable to the leucine zipper region (*5*, *8*). Thus, the observed difference in the intensity of helical band indicates that helical regions other than the leucine zipper coiled-coil would exist in the bVILZIP dimer. In the presence of the specific DNA sequence AP1, the helicity of GCN4 was much higher than that in the free dimeric form because of the helix formation at the basic and hinge regions (*5*, *8*) (Figure 4B). The helicity

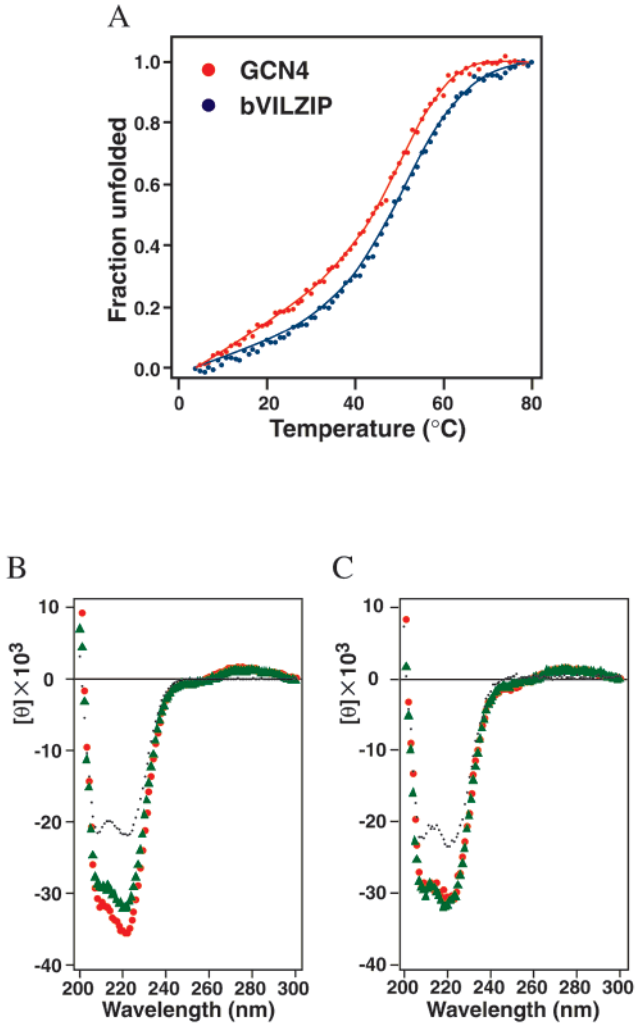

FIGURE 4: Circular dichroismic spectral analyses of GCN4 and bVILZIP. (A) The temperature-unfolding curve of bVILZIP was obtained by monitoring mean residue ellipticity at 222 nm at the concentration of 20 μM peptide. Circular dichroism difference spectra for the GCN4 (B) and bVILZIP (C) show the helical conformation of peptides in the presence of recognition DNAs. Spectra in the absence of DNA (black dots) and in the presence of AP1 (red circles) and HS (green triangles) at 20 °C.

of bVILZIP also increased in the presence of AP1 (Figure 4C) but to a degree smaller than that observed for GCN4. GCN4 in the HS complex contained less helical region than that in the fully helical AP1 complex (Figure 4B), indicating

that the stability of its DNA complex paralleled to the fraction folded of GCN4 (Figures 3B and 4B). In contrast, the fractions folded of bVILZIP remained similar for both the AP1 and HS complexes, despite the fact that its AP1 complex was much more stable than the HS complex (Figures 3C and 4C).

Additional information for the structural transition of bVILZIP was obtained by comparing the heat capacity changes (ΔC_p^0,obs) for DNA complexes of GCN4 and bVILZIP calculated by assuming that this was temperature-independent in the range of 20−40 °C (Table 1). The ΔC_p^0,obs for the GCN4−AP1 complex formation (−5.3 kcal mol^{-1}) is characterized by a larger negative value, as previously reported (*32*), than that for bVIL (−3.2 kcal mol^{-1}). On the other hand, the ΔC_p^0,obs values for the nonspecific complex of GCN4 and bVILZIP, −1.5 and −1.2 kcal mol^{-1}, respectively, are similar each other. Because an apparent negative ΔC_p^0,obs results from coupling of conformational changes in either of the macromolecules to the binding process (*1*, *34*), the conformational change associates with specific complex formation of GCN4 appears to be greater than that of bVILZIP. These results taken together indicate that the bVIL domain of bVILZIP is most likely in the folded state with or without DNA and suggest that bVILZIP changes its structure at a region other than the DNA-contacting or the leucine zipper regions upon the DNA complex formation. The most probable region responsible for the conformational change of bVILZIP is the hinge region that spans the bVIL domain and the leucine zipper.

DISCUSSION

The autonomously folding subdomain from villin headpiece provides a stable scaffold for design of a novel basic leucine zipper protein. The leucine zipper derivative bVILZIP discriminates DNA in the selectivity comparable to GCN4, and the stability of the specific DNA complex of bVILZIP is more tolerant to increasing temperatures than is that of GCN4. The bVILZIP dimer undergoes a structural transition on the DNA binding, while the extent of transition is smaller than that of GCN4. When a similar strategy is applied to design a monomeric DNA binding domain bVIL lacking the leucine zipper region, it folds into a stable domain and exhibits a rigid body DNA binding. A miniature protein designed by grafting the recognition residues of GCN4 onto the avian pancreatic peptide efficiently formed a specific DNA complex at the physiological ionic strength (*24*, *25*). The specific binding of the peptide was observed only at 4 °C associating with a structural transition, and a dimer of the peptide showed poor specificity (*24*). Unlike the induced-fit binding miniature protein, bVIL does not show DNA affinity or selectivity higher than the GCN4 basic region peptide. The stability of the DNA complex of bVIL is more tolerant to the temperature increase than that observed for the basic region peptide G23. Though the thermostability could be attributed to the folded structure of bVIL, there is no favorable change in the free energy of DNA binding. In view of the entropic penalty that associates with the induced-fit binding, the more rigid bVIL should bind more tightly than the basic region peptide does. The observed weak binding of bVIL indicates that the amino acid residues necessary for DNA recognition are preorganized in a way different from the proper geometry in the monomeric bVIL folding domain. The unfolded miniature protein derived from the avian pancreatic peptide, on the other hand, could optimize the interaction between the nucleic acid bases and amino acid residues by the induced-fit mechanism.

GCN4 in the dimeric form suffers from the severe electrostatic repulsion between two basic regions in the absence of DNA. Such an unfavorable orientation between the two DNA-contacting regions would be reinforced in the dimer configuration of bVILZIP, which explains well the result that bVILZIP is not completely folded in the absence of DNA. Several lines of evidence suggest that the bVIL portion of bVILZIP is folded in the presence or absence of DNA: (i) comparison of the melting behavior of bVILZIP and GCN4 from 4 to 40 °C (Figure 4A) reveals that bVILZIP does not contain a nascent helical portion that corresponds to the basic region of GCN4, (ii) the structures of specific and nonspecific DNA complexes of bVILZIP are almost the same, and (iii) the heat capacity changes for DNA complex formation indicate that the conformational changes associated with specific complex formation of bVILZIP are smaller than that of GCN4.

Which region of bVILZIP changes the structure during the specific DNA complex formation? In the case of the GCN4 dimer, both the DNA-contacting and hinge regions have been regarded as a nascent helix in the absence of DNA (*5*, *8*, *35*). Because the DNA-contacting region of bVILZIP is considered to be stabilized with the villin-fold, residues encompassing the hinge region would be deformed into a nonhelical conformation in the bVILZIP dimer to relieve the unfavorable repulsion between two DNA contacting regions. Association with either the specific or nonspecific DNA neutralizes the charge repulsion between two bVIL domains, thereby facilitating the helix formation around the hinge region. In fact, it has been suggested that the structural adaptability of the hinge region is important for the discrimination between the AP1 and CRE sequences by GCN4 (*36*, *37*) and for the specific DNA binding by other bZIP proteins (*38*). The artificially assembled dimers of the basic region peptides lacking the hinge region also preferentially bind CRE over AP1 (*14*, *15*, *17*). The result that bVILZIP binds equally to AP1 and CRE would reflect that the flexibility or the local structures at the hinge region of bVILZIP are different from those of GCN4, despite that both proteins share the same amino acid sequences at the region.

It is also possible that slight changes in the register of the helices of the DNA contacting and leucine zipper regions of bVILZIP would introduce strain within the hinge region upon DNA binding. The structural deformability of the hinge region would facilitate each DNA-contacting region to locate at the DNA major groove of respective half-site, thereby avoiding the risks of nonspecific trapping that might occur for the protein with a completely folded structure. The structural transition of the DNA-protein interface by itself would affect for the sequence-selectivity in a similar manner, as observed for the monomeric DNA binding by basic region peptides, though the entropic penalty for the structural organization will decrease the specific binding affinity at physiological temperatures. Such a disadvantage of entropic penalty might be less emphasized for GCN4, or for other bZIP proteins, because the structural continuity of leucine zipper region stabilizes the helical conformation of the basic region.

Our results indicate that the structural transition outside the direct-contacting region, which adjusts the precise location of the direct-contacting region, plays a critical role in the specific complex formation of bZIP proteins. The induced-fit of the direct DNA-contacting region by itself may not be a prerequisite for the sequence-specific DNA recognition of by proteins as long as the protein region adjacent to the direct-contacting region is flexible. Such a mechanism would be applicable for other macromolecular recognition, as the conformational changes away from the direct contacting site are often observed for many protein complexes. In addition, the thermostable bVIL motif would be quite useful for the design of novel DNA binding proteins, especially for those that are functional in vivo.

ACKNOWLEDGMENT

We thank Dr. P. S. Kim for his generous gift of the plasmid p4LZ'91.

REFERENCES

1. Spolar, R. S., and Record, T. (1994) *Science 263*, 777−784.
2. Pabo, C. O., and Sauer, R. T. (1992) *Annu. Rev. Biochem. 61*, 1053−1095.
3. Frankel, A. D., and Kim, P. S. (1991) *Cell 65*, 717−719.
4. Alber, T. (1992) *Curr. Opin. Genet. Dev. 2*, 205−210.
5. Weiss, M. A., Ellenberger, T., Wobble, C. R., Lee, J. P., Harrison, S. C., and Struhl, K. (1990) *Nature 347*, 575−578.
6. Hinnebusch, A. G. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6442−6446.
7. Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988) *Science 240*, 1759−1764.
8. O'Neil, K. T., Shuman, J. D., Ampe, C., and DeGrado, W. F. (1991) *Biochemistry 30*, 9030−9034.
9. Ellenberger, T. E., Brandl, C. J., Struhl, K., and Harrison, S. C. (1992) *Cell 71*, 1223−1237.
10. König, P., and Richmond, T. J. (1993) *J. Mol. Biol. 233*, 139−154.
11. Morii, T., Yamane, J., Aizawa, Y., Makino, K., and Sugiura, Y. (1996) *J. Am. Chem. Soc. 118*, 10011−10017.
12. Park, C., Campbell, J. L., and Goddard, W. A., III (1996) *J. Am. Chem. Soc. 118*, 4235−4239.
13. Talanian, R. V., McKnight, C. J., and Kim, P. S. (1990) *Science 249*, 769−771.
14. Talanian, R. V., McKnight, C. J., Rutkowski, R., and Kim, P. S. (1990) *Biochemistry 31*, 6871−6875.
15. Cuenoud, B., and Schepartz, A. (1993) *Science 259*, 510−513.
16. Morii, T., Saimei, Y., Okagami, M., Makino, K., and Sugiura, Y. (1997) *J. Am. Chem. Soc. 119*, 3649−3655.
17. Ueno, M., Murakami, A., Makino, K., and Morii, T. (1993) *J. Am. Chem. Soc. 115*, 12575−12576.
18. Ueno, M., Sawada, M., Makino, K., and Morii, T. (1994) *J. Am. Chem. Soc. 116*, 11137−11138.
19. Aizawa, Y., Sugiura, Y., Ueno, M., Mori, Y., Imoto, K., Makino, K., and Morii, T. (1999) *Biochemistry 38*, 4008−4017.
20. Johnson, N. P., Lindstrom, J., Baase, W. A., and von Hippel, P. H. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 4840−4844.
21. Lajmi, A. R., Lovrenic, M. E., Wallace, T. R., Thomlinson, R. R., and Shin, J. A. (2000) *J. Am. Chem. Soc. 122*, 5638−5639.
22. Pomerantz, J. L., Sharp, P. A., and Pabo, C. O. (1995) *Science 267*, 93−96.
23. McColl, D. J., Honchell, C. D., and Frankel, A. D. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 9521−9526.
24. Zondlo, N. J., and Schepartz, A. (1999) *J. Am. Chem. Soc. 121*, 6938−6939.
25. Chin, J. W., and Schepartz, A. (2001) *J. Am. Chem Soc. 123*, 2929−2930.
26. McKnight, C. J., Doering, D. S., Matsudaira, P. T., and Kim, P. S. (1996) *J. Mol. Biol. 260*, 126−134.
27. McKnight, C. J., Matsudaira, P. T., and Kim, P. S. (1997) *Nat. Struct. Biol. 3*, 180−184.
28. Vardar, D., Buckley, D. A., Frank, B. S., and McKnight, C. J. (1999) *J. Mol. Biol. 294*, 1299−1310.
29. O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1989) *Science 243*, 538−542.
30. Takeda, Y., Ross, P. D., and Mudd, C. P. (1994) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8180−8184.
31. Tanha, J., and Lee, J. S. (1997) *Nucleic Acids Res. 25*, 1442−1449.
32. Berger, C., Jelesarov, I., and Bosshard, H. R. (1996) *Biochemistry 35*, 14984−14991.
33. Krylov, D., Barchi, J., and Vinson, C. (1998) *J. Mol. Biol. 279*, 959−972.
34. Murphy, K. P., and Freire, E. (1992) *Adv. Protein Chem. 43*, 313−361.
35. Saudek, V., Pasley, H. S., Gibson, T., Gausepohl, H., Frank, R., and Pastore, A. (1991) *Biochemistry 30*, 1310−1317.
36. Kim, J., Tzamarias, D., Ellenberger, T., Struhl, K., and Harrison, S. C. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 4513−4517.
37. Metallo, S. J., and Schepartz, A. (1994) *Chem. Biol. 1*, 143−151.
38. Sera, T., and Schultz, P. G. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2920−2925.

BI015820I