Protein Science (2000), *9*:2354–2365. Cambridge University Press. Printed in the USA.

Backbone dynamics of sequence specific recognition and binding by the yeast *Pho4* bHLH domain probed by NMR

JOHN W. CAVE,[1,2] WERNER KREMER,[2,3] AND DAVID E. WEMMER[1,2]

[1]Department of Chemistry, University of California at Berkeley, Berkeley, California 94720
[2]Physical Bioscience Division, Calvin Laboratory, Lawrence Berkeley National Laboratory, Berkeley, California 94720

(RECEIVED May 9, 2000; FINAL REVISION August 9, 2000; ACCEPTED October 3, 2000)

Abstract

Backbone dynamics of the basic/helix-loop-helix domain of *Pho4* from *Saccharomyces cerevisae* have been probed by NMR techniques, in the absence of DNA, nonspecifically bound to DNA and bound to cognate DNA. Alpha proton chemical shift indices and nuclear Overhauser effect patterns were used to elucidate the secondary structure in these states. These secondary structures are compared to the co-crystal complex of *Pho4* bound to a cognate DNA sequence (Shimizu T, Toumoto A, Ihara K, Shimizu M, Kyogou Y, Ogawa N, Oshima Y, Hakoshima T, 1997, *EMBO J 15*: 4689–4697). The dynamic information provides insight into the nature of this DNA binding domain as it progresses from free in solution to a specifically bound DNA complex. Relative to the unbound form, we show that formation of either the nonspecific and cognate DNA bound complexes involves a large change in conformation and backbone dynamics of the basic region. The nonspecific and cognate complexes, however, have nearly identical secondary structure and backbone dynamics. We also present evidence for conformational flexibility at a highly conserved glutamate basic region residue. These results are discussed in relation to the mechanism of sequence specific recognition and binding.

Keywords: helix-loop-helix protein; nonspecific binding; *Pho4*; protein dynamics; sequence-specific binding; spectral density mapping

The PHO system is a set of at least 17 genes in *Saccharomyces cerevisae* that assist in the regulation of phosphate metabolism (Oshima, 1982). They encode several phosphatases, transport proteins, and transcriptional regulators. The PHO4 gene product is a 312 amino acid, repressible transcription factor that regulates the expression of several other genes in the PHO system (Oshima, 1982; Gilliquet & Berben, 1993). Like most transcription factors, it is composed of several domains (Ogawa & Oshima, 1990) and its function and cellular localization are highly regulated. Under phosphate rich conditions, *Pho4* is under the negative control of the cyclin-CDK complex of *Pho80-Pho85* (Schneider et al., 1994; O'Neil et al., 1996). This cyclin/CDK complex phosphorylates *Pho4* and results in localization of *Pho4* in the cytoplasmic space (Kaffman et al., 1998a). In times of phosphate starvation, however, *Pho81* prevents phosphorylation of *Pho4* by the *Pho80-Pho85* complex and allows for translocation across the nuclear membrane, which is mediated by the import receptor *Pse1/Kap121* complex (Kaffman et al., 1998b). When inside the nucleus, *Pho4* serves its role as a positive transcriptional regulator.

Pho4 accomplishes its positive transcriptional control, in part, by recognition and binding of an E-box sequence (5′-*CANNTG*-3′) in the promoter regions of genes (Vogel et al., 1989). Binding to this recognition element occurs through the basic region/helix-loop-helix (bHLH) motif, which is formed by the homodimerization of the C-terminal 65 residues. This bHLH motif is a member of a subclass of HLH protein domain folds (Littlewood & Evan, 1995). HLH domains, first identified by Murre et al. (1989), consist of either homo- or heterodimers. Although higher order association has been observed (i.e., formation of oligomers), dimerization is the most frequent mode of association (Littlewood & Evan, 1995). Each monomer contributes two helices that are separated by a loop region and form a four helix bundle upon dimerization. Another subclass of HLH dimers is further stabilized by the addition of a leucine zipper domain adjacent to the HLH fold (forming an HLH/Z motif).

Most HLH-containing proteins act as transcriptional regulators. Efficient recognition and binding to DNA require that each monomer contribute to a DNA binding domain. Those HLH proteins that do not contain DNA binding domains, however, can heterodimer-

Reprint requests to: David E. Wemmer, Department of Chemistry, University of California at Berkeley, Berkeley, California 94720; e-mail: dewemmer@lbl.gov.

[3]Current address: Institut fuer Biophysik und physikalische Biochemie, Universitaet Regensburg, Universitaetsstrasse 31, D-93053 Regensburg, Germany.

ize with other HLH proteins that do, and thus serve as effective negative regulators (Littlewood & Evan, 1995). DNA binding by HLH proteins occurs through a helical region, containing a number of basic residues, that is adjacent to the HLH fold (forming a bHLH or bHLH/Z motif). The binding of DNA by this basic region is reminiscent of the basic region in the bZIP motif (Ellenberger, 1994).

Detailed characterization of the interactions between bHLH and bHLH/Z domains specifically bound to promoter DNA sequences has been achieved for several proteins using X-ray diffraction (Ferre-D'Amare et al., 1993, 1994; Ellenberger et al., 1994; Ma et al., 1994; Shimizu et al., 1997; Parraga et al., 1998). These structures verified the expectation that dimerization occurs by formation of the four helix bundle (Murre et al., 1989) and that sequence specific recognition of E-box half sites is mediated by the highly conserved residues in the basic region of each monomer. These structural studies have further revealed that residues located on the loops separating the helices of the core HLH fold can contact the DNA phosphodiester backbone and contribute to the stability of the protein-DNA complex.

The structure of the *Pho4*-DNA complex was determined by Shimizu and co-workers using X-ray crystallography (Shimizu et al., 1997). In addition to the many structural features common to the previous co-crystal structures of bHLH or bHLH/Z DNA complexes, the *Pho4* complex also revealed some novel features. Among these was a short helical segment in the loop connecting the two helices forming the HLH core. In all previous studies, the loop region was devoid of secondary structure. The crystal structure also showed several side-chain contacts to base pairs flanking the consensus sequence, which explains how *Pho4* can preferentially recognize and bind particular E-box sequences among various gene promoter regions also containing E-box elements.

Considerable structural information about DNA bound complexes of bHLH or bHLH/Z proteins is available; however, far less is known about the bHLH or bHLH/Z motifs in the absence of DNA sequences that they specifically bind. Fairman et al. (1997) have studied the bHLH motif of *E47* in the absence of DNA using NMR. Much like the solution NMR studies of bZIP domains in the absence of DNA (Saudek et al., 1991; Talanian et al., 1992; Bracken et al., 1999), the basic region of the *E47* bHLH domain appears to lack well-defined secondary structure. For both classes of protein folds, however, there is evidence for nascent or flexible helical secondary structure. Understanding the dynamic behavior of DNA binding domains in the absence of DNA will provide a better understanding of the nature of these domains when transcription factors are not bound to DNA.

To our knowledge, there is no structural information on bHLH or bHLH/Z domains bound to DNA that is devoid of a cognate recognition sequence (that is, nonspecifically bound to DNA). Furthermore, there have only been a few structural studies published describing the interactions of sequence specific DNA binding domains in the presence of nonspeific DNA sequences (Winkler et al., 1993; Gewith & Sigler, 1995; Jen-Jacobson, 1997). The interaction of sequence specific DNA binding domains with nonspecific DNA is an important intermediate step in the process of sequence specific recognition and binding.

In this study, we have used solution NMR techniques to investigate the secondary structure and backbone dynamics of the bHLH domain from *Pho4* bound to a cognate DNA sequence, nonspecifically bound to DNA and free in solution. We report differences in the secondary structure and backbone dynamics for the differently bound states of this bHLH domain. The structural and dynamic differences are discussed in relation to a possible mechanism for the recognition of E-box DNA sequence by *Pho4*.

Results

Using sequence alignments with other known bHLH and bHLH/Z domains (Littlewood & Evan, 1995), *Pho4* wild-type residues D248 to T312 (Fig. 1B) were chosen to be the construct utilized in this present study (this construct is simply referred to as *Pho4* here after). These residues are capable of forming a bHLH motif and E-box sequence specific binding. An amino terminal methionine residue was added to the construct as a consequence of the recombinant bacterial expression system. The cognate target DNA sequence (Fig. 1B) was constructed with an E-box element at its center. The nonspecific DNA target sequence (Fig. 1B) was designed to have the same nucleotide composition as the cognate sequence, but the E-box nucleotides were rearranged to prevent cognate complex formation. Circular dichroism (CD) analysis of DNA titrations with *Pho4* reveals that *Pho4* binds both sequences tightly (Fig. 2A). The cognate complex, however, has a lower dissociation constant (*Kd*) than the nonspecific complex (*Kd* = 1 nM and *Kd* $\approx$ 20 nM for the cognate and nonspecific complexes, respectively). Despite the relatively tight binding with the nonspecific target DNA sequence, polyacrylamide gel electrophoresis mobility shift assays (EMSA) reveal that *Pho4* does not bind this target sequence with any specificity (Fig. 2B). The EMSA experiments also display a higher *Kd* for the cognate complex (*Kd* $\approx$ 50 nM) than was observed with the CD titrations, but this observation is most likely due to the presence of unlabeled nonspecific competitor DNA in the gel mobility shift experiments.

Experimental conditions chosen for the NMR experiments in this work were determined by screening one-dimensional ^{1}H NMR spectra under various pH and temperature in the presence of cognate DNA. These experiments indicated optimal resolution of observable resonances at 35 °C and pH 6.3.

The formation of both nonspecific and cognate complexes was accompanied by the partial precipitation upon addition of the bHLH and DNA components (the DNA sequences used are displayed in Fig. 1). ^{15}N-^{1}H heteronuclear single quantum coherence (HSQC) spectra collected while titrating DNA into a solution with *Pho4* indicates that *Pho4* free and in the complex exchange at a rate that is significantly faster in the nonspecific complex than in the cognate complex. The average ^{15}N peak width in the HSQC spectra was the same for all three samples (21 $\pm$ 2, 22 $\pm$ 2, and 21 $\pm$ 2 Hz for the free form, nonspecific, and cognate complexes, respectively). Some residues, however, significantly deviate from these average linewidths. Much of the basic region and the first four residues of helix 2 (S44 to T47) in the nonspecific complex demonstrate slightly broader lines (24 $\pm$ 2 Hz). Within the basic complex of the cognate complex, residue E12 also had a modestly broader linewidth (26 Hz) and Q13 was visibly broader, but could not be reliably measured.

Portions of the ^{15}N-^{1}H HSQC spectra for *Pho4* in the absence of DNA, in the presence of nonspecific DNA and in the presence of cognate DNA are presented in Figure 3. Of the 62 backbone amide resonances expected, both the unbound sample and the cognate complex contain 61 cross peaks while the sample with the nonspecific complex has 57. The nonspecific complex has 20 and cognate complex has 24 observed backbone HSQC resonances that significantly change relative to the unbound cross-peak chemical shift values (see Fig. 4). The residues that changes the most are

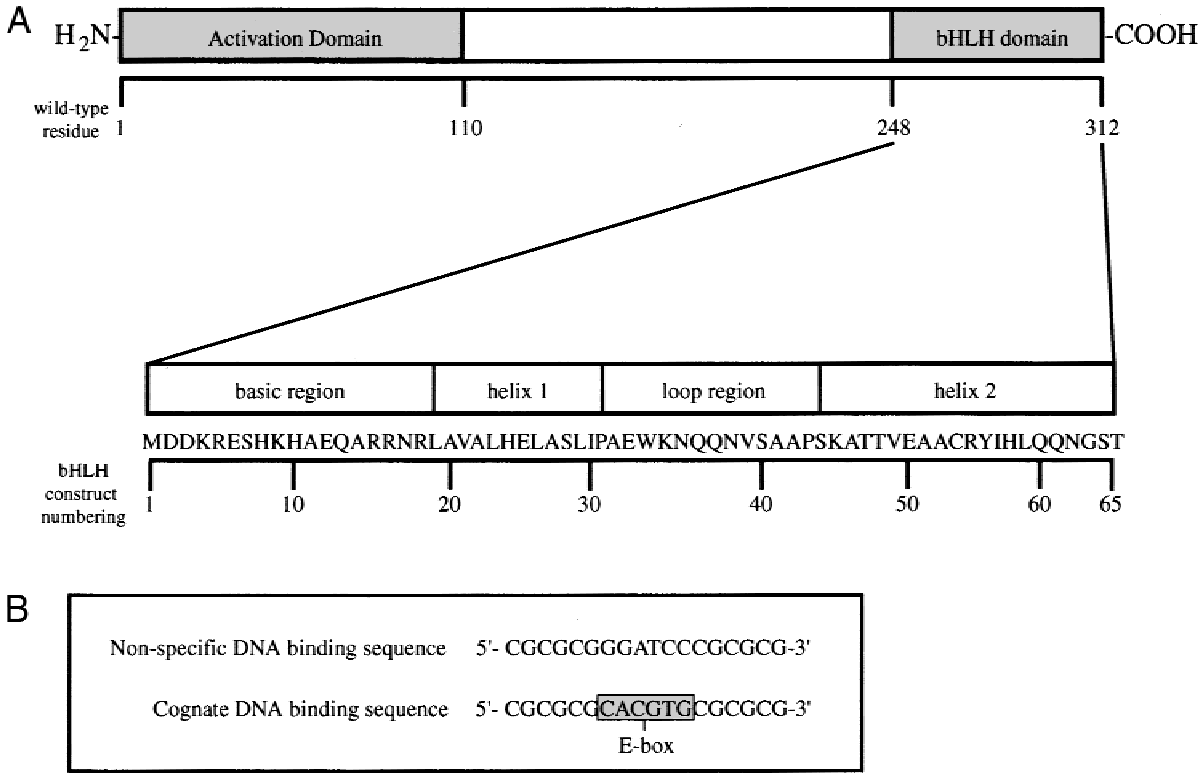

Fig. 1. A: A schematic of the *Pho4* protein is shown, indicating the positions of the transcriptional activation and bHLH DNA binding domains. The expanded region displays the construct and residue numbering system used in the present study. **B:** The oligonucleotide sequences used to form the nonspecific and cognate DNA complexes.

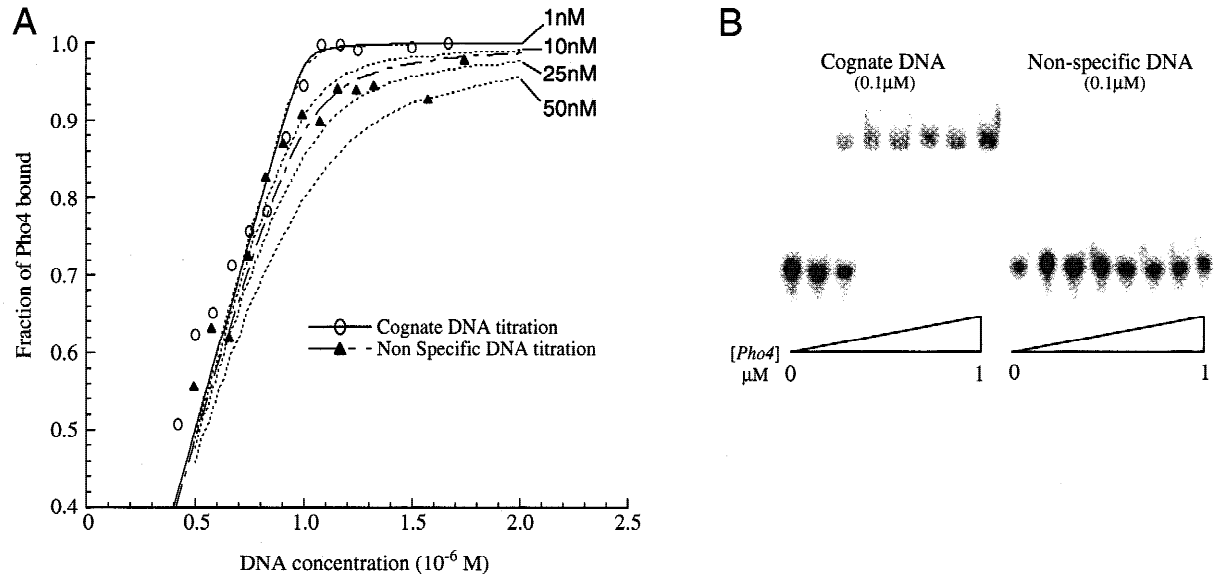

Fig. 2. The binding behavior of the cognate and nonspecific complexes. **A:** The fraction of 1 μM *Pho4* bound as a function of DNA concentration. The fraction of *Pho4* bound was determined by the CD absorbance at 222 nm during the titration. The cognate complex (open circles) has been fit with a curve corresponding to a 1 nM *Kd*, while the nonspecific complex (solid triangles) has been fit with a 20 nM *Kd* curve. Curves representing 1, 10, 25, and 50 nM *Kd* values have also been plotted as a reference (the 1 nM curve overlaps with the cognate complex curve). **B:** EMSA experiments demonstrate that *Pho4* sequence specifically binds the cognate DNA target, while its interaction with the nonspecific target displays no such preference.

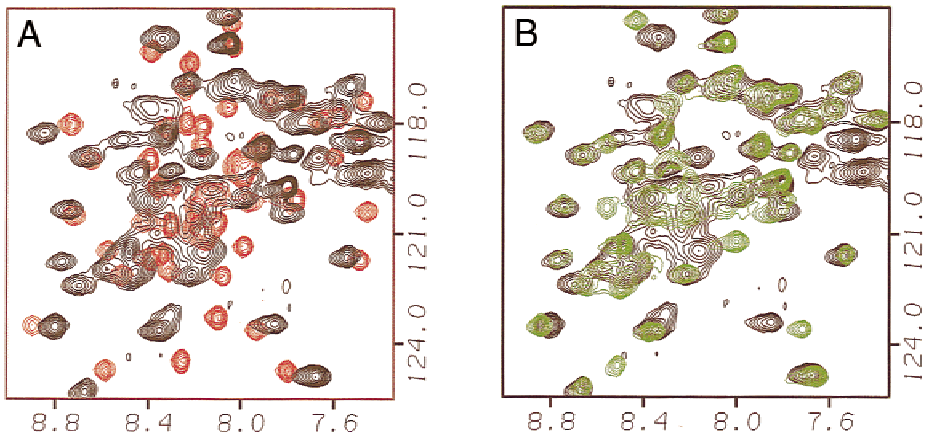

Fig. 3. Overlays of regions from ^{1}H-^{15}N HSQC spectra of the *Pho4* bHLH domain and NDA complexes are shown. **A:** The unbound form of *Pho4* (red contours) and nonspecific complex (black contours) are shown. **B:** The nonspecific complex (black contours) is overlapped with the cognate complex (green contours).

primarily found in the basic region, the loop region, and the N-terminal portion of helix 2. Significant changes in the observed backbone amide ^{15}N and ^{1}H chemical shifts were anticipated based on preliminary CD experiments. CD measurements (Fig. 5) demonstrated that the presence of either nonspecific or cognate DNA results in an approximate 25% increase in the observed helical content. This observed increase in helical content is approximately equivalent to the percentage of the basic region relative to the

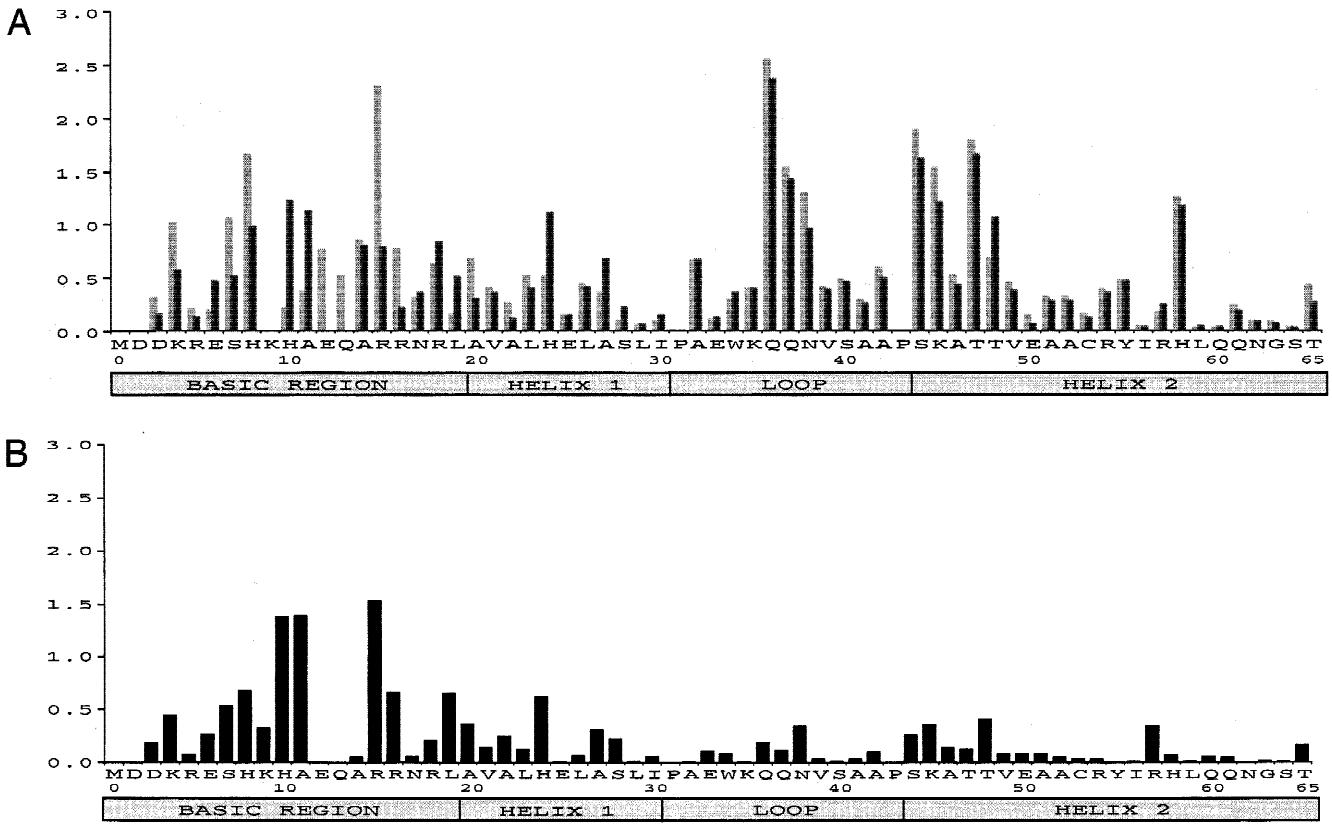

Fig. 4. A: The observed changes (ppm) in the ^{1}H-^{15}N HSQC cross-peak position upon forming the nonspecific complex (dark gray bars) and cognate complex (light gray bars), relative to the free protein cross-peak positions. **B:** The observed changes (ppm) in the ^{1}H-^{15}N HSQC cross-peak position between the nonspecific and cognate complexes. The changes in cross-peak positions were quantified by $[(\Delta^{15}N_{ppm})^2 + (\Delta^{1}H_{ppm})^2)]^{1/2}$.

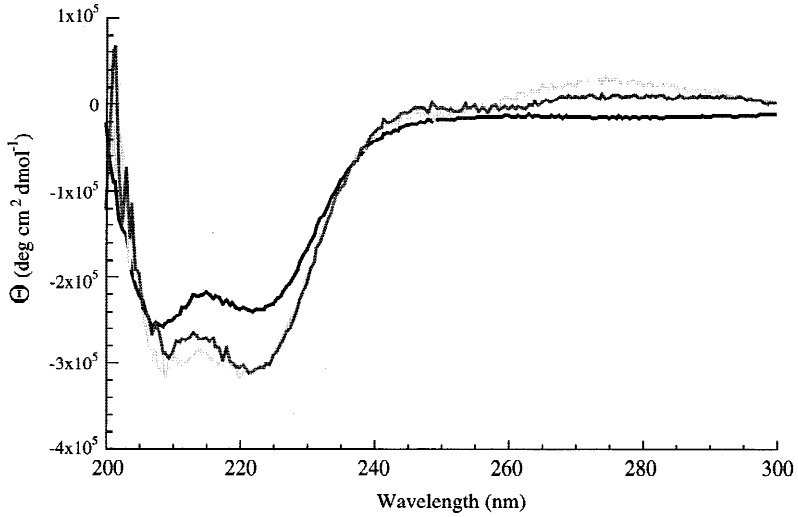

Fig. 5. CD data (in molar ellipticity, Θ) of the free protein (black), the nonspecific complex (darker gray), and cognate complex (lighter gray).

entire *Pho4* construct. This observation also coincides well with previous structural studies of bZIP, bHLH, and bHLH/Z DNA binding proteins (Weiss, 1990; O'Neil et al., 1991; Anthony-Cahill et al., 1992; Ferre-D'Amare et al., 1994; Meierhans et al., 1995; Kunne et al., 1998; Winston et al., 1999). Between the nonspecific and cognate complex, however, the residues that display a significant change in their cross-peak position are exclusive to the basic region.

Backbone resonance assignments

Backbone resonances of the three samples were assigned using double and triple resonance experiments. The assignments for M1, D2, K9, P31, and P43 in the unbound protein were never established. Relative to the unbound form of *Pho4*, most of the backbone amide ^{1}H and ^{15}N resonances of the HLH fold exhibit only modest changes in their chemical shift values in the nonspecific and cognate complexes (see Fig. 4). The basic region in both of the DNA complexes, however, was more challenging to assign. The assignment process for this region was hindered by the poor coherence transfer out to the side chains. The assignment process was assisted by recent studies with a modified *Pho4* construct bound to nonspecific DNA. The modified construct (lacking the last three C-terminal amino acids) has a nearly identical chemical shift dispersion as the original construct, but forms a more soluble complex and provides higher quality data. Assignments from the modified *Pho4* nonspecific complex mapped directly onto the original nonspecific complex. The backbone resonances for residues M1, D2, E12, Q13, P31, and P43 in the original and modified *Pho4* nonspecific complex, however, were never identified. The nonspecific complex assignments facilitated the completion of assignments for the cognate data sets. The only backbone assignments absent in the cognate complex were for residues M1, D2, P31, and P43.

Secondary structure

From crystallographic and NMR studies that have been published with bHLH and bHLH/Z proteins, it was expected that only random coil and helical secondary elements would be present. Using ^{15}N resolved [^{1}H-^{1}H] nuclear Overhauser effect spectroscopy (NOESY)-HSQC data, helical secondary structure can be recognized by unique patterns of ^{1}H-^{1}H NOE contacts, including NOE contacts between amide protons of neighboring residues (*dNN* contacts), contacts between the alpha proton of the i^{th} residue and the backbone amide proton of the $i + 3$ and $i + 4$ residues ($d\alpha N_{i+3}$ and $d\alpha N_{i+4}$, respectively) as well as contacts between the β protons of the i^{th} residue and of the amide protons of the $i + 1$ residue ($d\beta N$).

Figure 6 displays the observed short- and medium-range NOEs observed in ^{15}N resolved [^{1}H-^{1}H]NOESY experiments for all three samples. $d\alpha N_{i+4}$ contacts were not observed in any of the samples. All three samples show a well-defined HLH core. Helix 1 is well defined from residue A20 to I30 while helix 2 shows some variation. In the unbound protein, helix 2 is well defined between residues E50 and Q60, while in both DNA complexes this helix extends to encompass residues A46 to N62. NOE contacts also show that there is a short helix between residues A32 and Q37 (A32 to K35 in the free form).

Within the HLH core, both DNA complexes also exhibit interesting NOE contacts in the loop region (V39 to A42). The *Pho4* co-crystal structure does not indicate that there should be any secondary structure in this region. In addition to the observed *dNN* and $d\beta N$ NOEs, NOE contacts are observed between V39 γ^{1}H and the backbone $(N)^{1}$H of S40 and A41 as well as $d\alpha N_{i+1}$ contacts from V39 to S40 and S40 to A41 and A42.

The basic region demonstrates the most dramatic change in secondary structure between the three samples. The unbound form is almost devoid of inter-residue NOEs. Both DNA complexes show the shorter range NOE patterns consistent with helix formation, but medium-range NOEs, such as $d\alpha N_{i+3}$, were not identified in the basic region for these complexes. Identification of such longer range NOEs, however, was complicated by residues missing or obscured in the ^{15}N resolved [^{1}H-^{1}H]NOESY-HSQC data sets. For the cognate complex, data for residues K9, Q13, and A14 were missing from the spectrum, while residues H8, K9, E12, Q13, and A14 were missing in the nonspecific complex data set. The

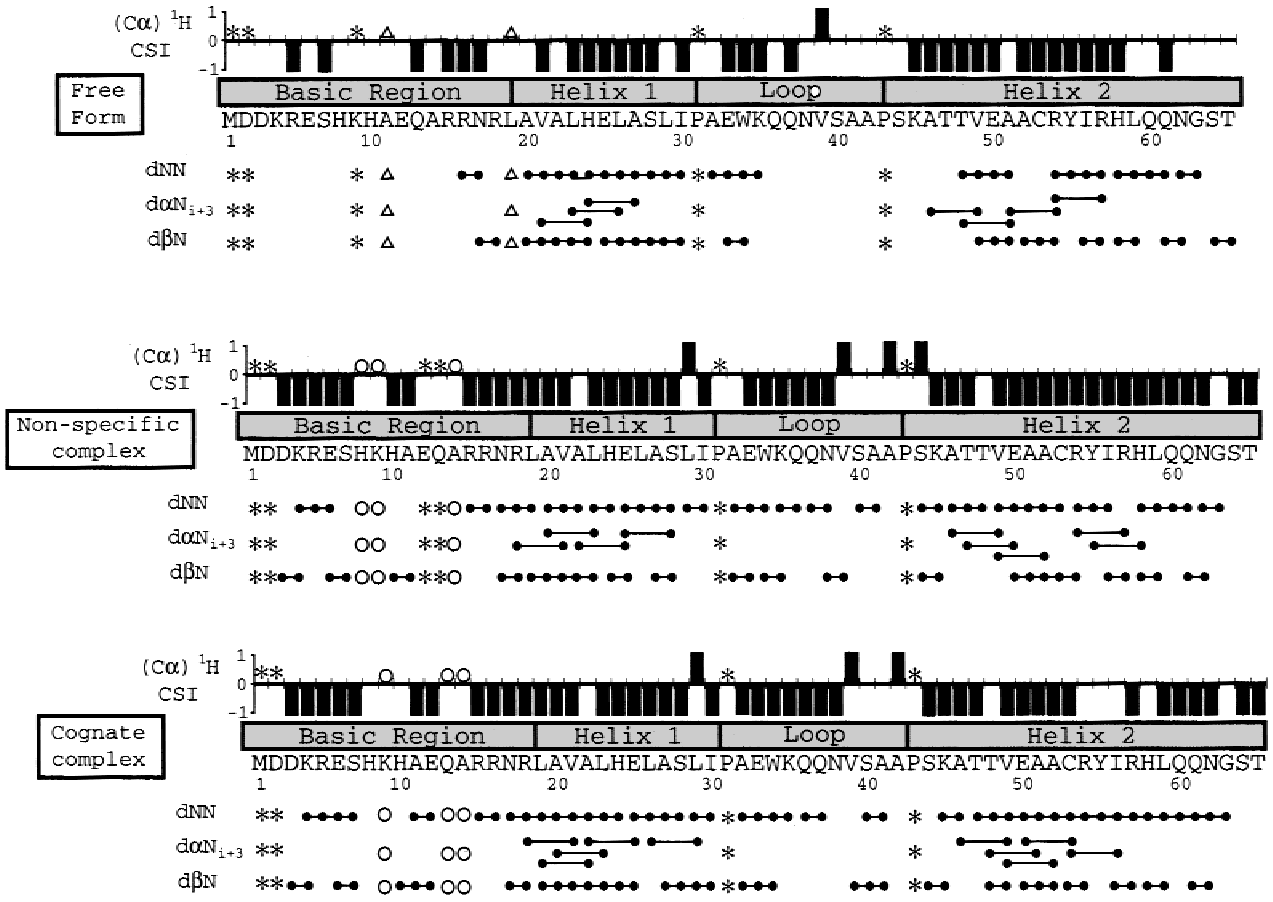

Fig. 6. The observed short- and medium-range NOEs and (Cα)^{1}H CSI for the free protein, nonspecific complex, and cognate complex are shown schematically. For the CSI data, a value of -1 indicates a helical chemical shift and a value of $+1$ indicates a strand-like chemical shift. Asterisks indicate that the information is absent because of a missing backbone amide ^{1}H and/or ^{15}N assignments. Open triangles indicate information that is obscured due to chemical shift degeneracy. Open circles indicate peaks for the given residue within the relevant experiment were not present despite assignment of the backbone amide ^{1}H and ^{15}N chemical shifts and not a result of chemical shift degeneracy.

absence of data for E12 and Q13 in the nonspecific complex is due to missing assignments; Q13 in the cognate complex is most likely missing due to line broadening; the absence of the other residues, however, is unexplained.

The secondary structural characteristics established from the NOE data can be supplemented with the (Cα)^{1}H chemical shift index (CSI). CSI analysis relies on the well-documented observation that the formation of secondary structure generates consistent changes in the observed chemical shift values of backbone nuclei compared to those found in random coils (Wishart & Sykes, 1994). The (Cα)^{1}H CSI information for the unbound form and DNA complexes is displayed in Figure 6. For all three forms, the regions of helical secondary structure, as defined by the NOE data, are also supported by the (Cα)^{1}H CSI values. Within the free form, however, the regions defined to be helical in the CSI analysis are larger than suggested by the NOE data. These regions with helical CSI character, but lacking helical NOE data, would suggest that they are transiently stable (or nascent) helices. Residues V39 and A42, which are the in loop region that forms a turn-like structure in the DNA complexes, have (Cα)^{1}H CSI values that are characteristic of a β-strand conformation. Residues S40 and A41, however, are within the range of random coil chemical shifts.

Backbone dynamics

Backbone ^{15}N T_1, T_2, and [^{1}H]-^{15}N heteronuclear NOE (hetNOE) experiments allow for the determination of the backbone mobility. These ^{15}N relaxation parameters reflect both local and global motions that occur on the nanosecond timescale, although the T_2 experiment can also detect motions at the millisecond timescale. Because either the assignment was lacking, the signal to noise ratio was too low to obtain reliable data, or chemical shift degeneracy, relaxation data were not obtained for residues H8, K9, A14, R15, R18, A20, H24, L29, A42, and A51 in the unbound state; K9, H10, E12 through N17, Q37, K45, T46, T48, E50, R54, R57, and L59 in the nonspecific complex; K9, A14, L19, Q36, Q37, V49, and C53 in the cognate complex.

To properly analyze the experimental relaxation data, the most probable shape and global rotational of the unbound and DNA bound *Pho4* samples had to be considered. The dimensions of the DNA complexes could be readily estimated from the coordinates of the Pho4-DNA co-crystal structure. From these measurements, the $D_{\|}/D_{\perp}$ ratio of the Pho4-DNA complexes is approximately 2.0, and thus there is significant rotational anisotropy in solution. In the absence of DNA significant rotational anisotropy is also expected,

but it is difficult to approximate the dimensions because a significant portion of the basic region appears to be random coil.

The model free analysis, as described by Lipari and Szabo (1982a, 1982b), efficiently allows for the calculation of a residue specific backbone order parameter (S^2) and internal correlation time (τ_e). For proteins with global anisotropic rotation, the determination of S^2 and τ_e becomes cumbersome. Both S^2 and τ_e must be defined with respect to a projection onto each motional axis describing rotational anisotropy and the rotational correlation times for each motional axis must also be known. Alternative protocols to model free analysis for anisotropically rotating proteins have been proposed (Schurr et al., 1994; Bruschweiler et al., 1995), but these alternative protocols are only best applied to systems with $D_{\|}/D_{\perp} \ll 2.0$ (Schurr et al., 1994; Lee et al., 1997). Under the given circumstances, reduced spectral density mapping (Farrow et al., 1995; Ishima & Nagayama, 1995) is the most appropriate method to analyze the experimental relaxation data. This type of analysis is independent of assumptions regarding the shape of the spectral density function and global rotational diffusion that are implicit in the model free methods.

The reduced spectral density data for the three samples are displayed in Figure 7. Those residues for which relaxation data was not obtained are necessarily missing from the reduced spectral density analysis. Examination of Figure 7 shows that the helices of the HLH dimerization fold (residues A20 to I30, A32 to K35, and T48 to Q60) have consistent spectral density values in all three samples. The high $J_{eff}(0)$ and low $J(0.87\omega_H)$ values are representative of residues with a well-ordered backbone conformation. Res-

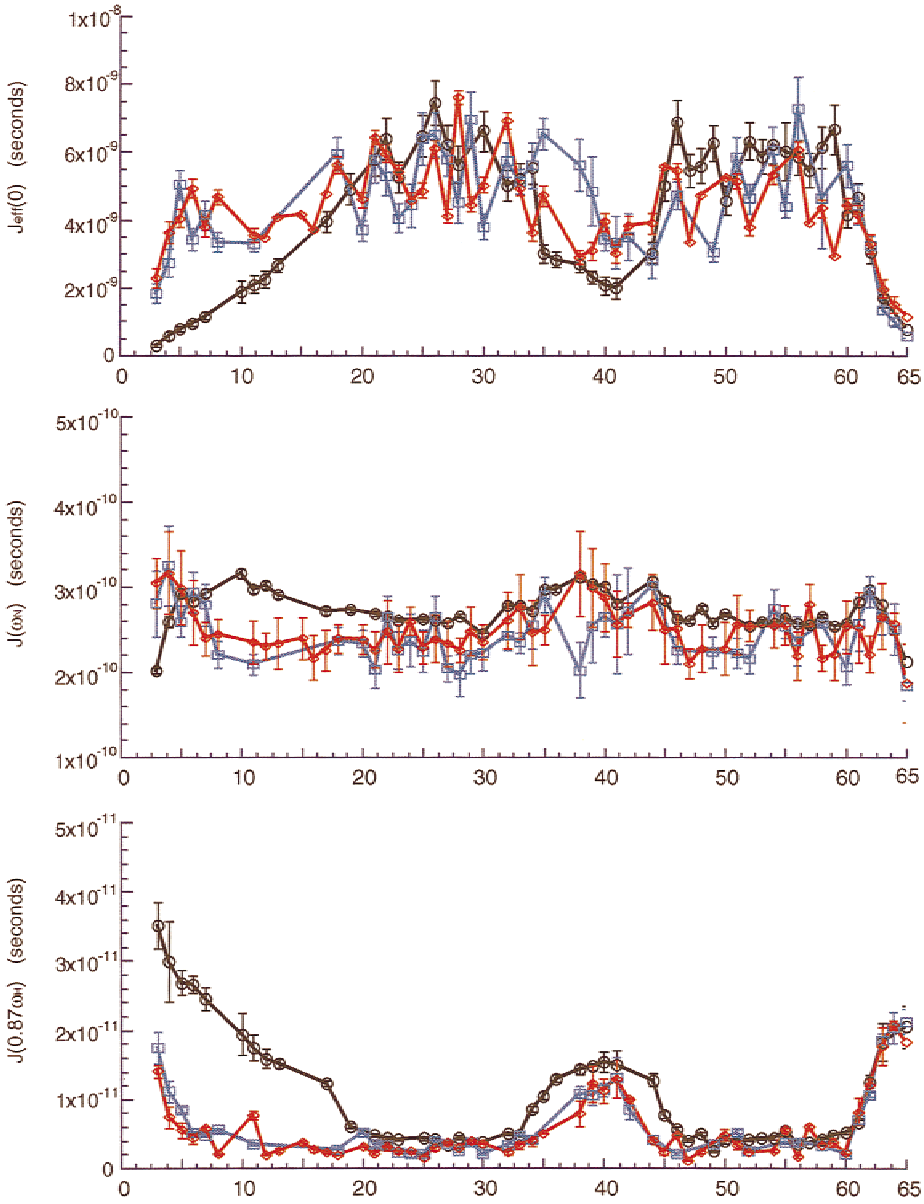

Fig. 7. The reduced spectral density functions maps for the free protein (black), nonspecific complex (blue), and cognate complex (red).

idues Q37 to A42, within the loop region, that connect the short helix (A32 to K35) and helix 2 within the HLH fold, however, possess lower $J_{eff}(0)$ values and larger $J(\omega_N)$ and $J(0.87\omega_H)$ values. These spectral density values indicate an increase in higher frequency nano-picosecond motions and most likely reflect a lower degree of conformational order and an increase in internal mobility. The $J(0.87\omega_H)$ values very clearly indicate that the conformational flexibility of these loop residues decreases upon complexation with DNA, but they remain less ordered than the stable helical regions. This implies that the possible turn structure formed by N38 to A42 revealed in the NOE data must possess considerable flexibility.

The basic region residues display significant changes between the unbound protein and DNA complexes. In the absence of DNA, the spectral density values of the basic region form a gradient that runs from the N-terminus to the start of helix 1. The residues at the N-terminus are dominated by high frequency motions (large $J(0.87\omega_H)$ and small $J_{eff}(0)$ values) and reflect a random coil backbone conformation. The intensity of these high frequency motions progressively diminishes until the junction with helix 1 where the spectral density values reflect those of the well ordered helices of the HLH fold. A very similar backbone dynamic gradient has also been seen with the basic region of the unbound GCN4 bZIP domain (Bracken et al., 1999). This gradient in the backbone motions within the basic region also is consistent with the formation of nascent helical secondary structure by residues nearest helix 1, as suggested by the $(C\alpha)^1H$ CSI data.

Upon complexation with DNA, the spectral density values of the basic region residues become similar to those of the well-ordered HLH helices. This change in the spectral density for these residues is expected based on the CD, $(C\alpha)^1H$ CSI, and NOE data that detail a coil to helix transition associated with DNA binding. Between the nonspecific and cognate DNA complexes, however, there do not appear to be any significant changes in the observed spectral density. This indicates that the backbone order of the nonspecific complex is essentially the same as the cognate complex. Thus, the changes observed the in the HSQC ^1H and ^{15}N chemical shifts between these two DNA bound complexes are not the consequence changes in backbone conformation.

Only the residues D3 and K4 appear to remain unstructured in the basic region of the DNA complexes. The C-terminal residues, G63 to T65, also display similar spectral density characteristics and also appear to be random coil in the all three of the samples.

Discussion

The NMR data presented show that complexation of DNA by the *Pho4* bHLH domain is highlighted by a large conformational change in the basic region. In the absence of DNA, only the residues in the three helices of the HLH dimerization fold are organized in well formed secondary structure. The basic region residues have a backbone dynamics profile that forms a gradient running from the stable helix 1 interface to the random coil N-terminus. The $(C\alpha)^1H$ CSI data and dynamics profile suggest that the basic region residues H10 to L19 exist in a transient helical conformation. The remaining residues of the basic region have backbone conformational flexibility and fast internal motions.

Evidence for nascent α-helical secondary structure within the basic region in the absence of DNA is not entirely unexpected. Solution NMR studies on an E47 bHLH construct in the absence of DNA found evidence for nascent helix formation within the same region (Fairman et al., 1997). Several other groups have also reported evidence from NMR studies for transient and flexible helices within the basic region of bZIP proteins in the absence of DNA (Saudek et al., 1991; Talanian et al., 1992; Bracken et al., 1999). Our findings with *Pho4* are unexpected, however, since solution NMR studies with a nearly identical bHLH *Pho4* construct reported in the crystal structure paper indicated no evidence for helical secondary structure within the basic region under DNA free conditions (Shimizu et al., 1997). It is not clear from that report, however, if that interpretation included transient helices.

Upon binding either nonspecific or cognate DNA, the basic region residues adopt a helical conformation. This transition is well characterized by the CD, $(C\alpha)^1H$ CSI, NOE, and backbone dynamics data. Analysis of the ^1H-^{15}N HSQC cross-peak positions also reveals that residues Q36 to N38 and S44 to T48 exhibit significant changes upon DNA binding as well. The cross-peak positions of residues Q36 to N38 shift as a result of the extension of the helical nature of residues A32 to K35. The NOE and $(C\alpha)^1H$ CSI data show that this short helix in the loop region expands upon DNA binding to include residues Q36 and Q37. The cross-peak changes for residues S44 to T48 are the result of the interaction of S44 and K45 with the phosphodiester backbone of the DNA. This interaction between residues in helix 2 and the DNA is a conserved interaction among bHLH proteins (Garrell & Campuzano, 1991; Littlewood & Evan, 1995).

Although it does not produce a dramatic change in the ^1H-^{15}N HSQC cross-peak positions, the NOE data suggest that loop residues V39 to A42 adopt a backbone conformation that is induced by both nonspecific and cognate DNA binding. The $(C\alpha)^1H$ CSI data also indicate that residues V39 and A42 have an extended conformation, and it is tempting to speculate that residues V39 to A42 may form a turn-like structure. The dynamics data, however, show that these residues do not possess the same backbone order and motions of the well-ordered helical regions. The presence of such backbone mobility allows for the possibility that the observed NOE patterns may be the result of interconverting conformations. Thus, the precise backbone conformation for these residues cannot be ascertained at the present time.

The nonspecifically and cognate bound *Pho4*-DNA complexes have nearly identical secondary structure characteristics and backbone dynamics. The largest changes observed in the chemical shifts of the ^1H-^{15}N HSQC cross-peak positions between the two DNA complexes are confined to those residues, and their adjacent neighbors, that make DNA base specific contacts in the cognate complex. In both DNA complexes, the basic regions and helix 1 form a continuous helix. The spectral density data indicate that this continuous helical region has the same backbone order in the nonspecific complex as it does in the cognate complex. This would suggest that backbone conformational flexibility is not crucial for the nonspecifically bound *Pho4* to recognize and bind its cognate sequence.

The broader ^{15}N linewidths that are observed for the basic region residues of the nonspecific complex in the ^1H-^{15}N HSQC are primarily the result from the nonspecific nature of the target DNA. These residues presumably come into contact with several different combinations of base pairs by directly binding different DNA base pairs or by sliding along the target sequence. The sampling of base pair environments can produce chemical exchange contributions to line broadening. Contributions from backbone conformational motions to the observed line broadening, however, cannot be completely ruled out.

Basic region residues E12 and Q13 in the cognate complex that were observed to have broader linewidths are interesting because E12 is a highly conserved residue in all E-box binding bHLH or bHLH/Z proteins. In crystal structures, this residue contacts the initial CA base step in the E-box sequence. Because of the nature of cognate binding, the observed line broadening should not be from the sampling of multiple DNA base pair environments. Conformational flexibility is the likely source of this line broadening. While it is possible that these motions are inherent in bHLH domains, it is also possible that there are additional elements within the full length *Pho4* protein that stabilize basic region when bound to DNA and they are not present in our current construct. Alternatively, an additional protein may be required for greater stability of the basic region. The binding of the full length *Pho4* protein to the PHO5 promoter has been shown to involve cooperative binding with the transcription factor *Pho2* (Barbaric et al., 1996, 1998). The *Pho2* interaction site within *Pho4* has been demonstrated to be near or overlapping with basic region (Hirst et al., 1994; Shao et al., 1996). This protein–protein interaction may have an effect on the stability on the basic region when bound to DNA.

The absence of medium-range $d\alpha N_{i+3}$ and $d\alpha N_{i+4}$ NOEs, characteristic of an α-helical backbone geometry, in the basic region of both DNA complexes is an interesting observation. The current paradigm is that the basic region in each monomer of bHLH and bHLH/Z domains adopts an α-helical backbone conformation upon DNA binding. The published crystal structure of the *Pho4* cognate complex, however, describes several asymmetric features within the *Pho4* dimer-DNA complex, including the presence of a 3_{10} helix stretching from E6 to A11 in only one of the monomers (using the present residue numbering; Shimizu et al., 1997). Using the Kabsch and Sander analysis of secondary structure elements (Kabsch & Sander, 1983; Fig. 8), one can see that the *Pho4* backbone geometry in the crystal structure deviates from an α-helical conformation in several places in an asymmetric manner. As seen in Figure 8, whereas one of the *Pho4* monomers is α-helical between residues K4 to R18, the other has only one turn of α-helix from Q13 to R16. The solution structure of the *Pho4* bHLH dimer must be symmetric, however, since there is only one set of resonances in the NMR data. The single crystal packing interaction reported between symmetry related molecules cannot account for the extensive asymmetric secondary structure in the *Pho4*-DNA co-crystal complex. Thus, it is not immediately obvious which monomer conformation is expected to be closest to the true solution structure. It is possible that the solution structure of *Pho4* bound either to nonspecific or cognate DNA has each monomer of the *Pho4* dimer in a conformation that is intermediate to those described by the crystal structure. This intermediate conformation may preclude the observation of medium-range $d\alpha N_{i+3}$ and $d\alpha N_{i+4}$ NOEs.

There are no published solution NMR structures of bHLH or bHLH/Z domains free or complexed with DNA to which we can compare our data. Applying the Kabsch and Sander analysis to all of the bHLH and bHLH/Z co-crystal structures accessible in the Protein Data Bank reveals that all of the complexes also possess disruptions in the α-helical backbone conformations. For example, both strands in the Max dimer also show a disruption in the α-helical backbone at the highly conserved glutamate residue in the basic region. Although disruption of the α-helical backbone seems to be a common feature in bHLH and bHLH/Z proteins, there does not appear to be a common pattern in the location of these disruptions in the bHLH or bHLH/Z folds.

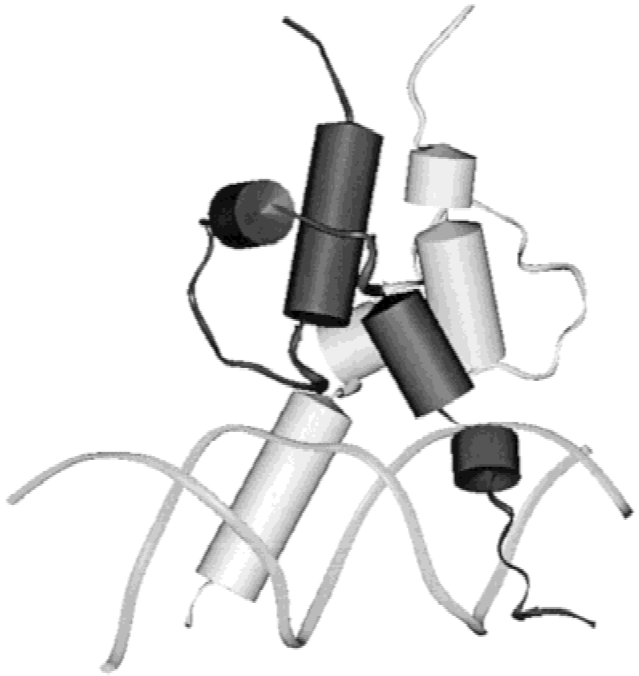

Fig. 8. The Kabsch–Sander secondary structure representation (Kabsch & Sander, 1983) derived from the crystal structure of the complex of *Pho4* bound in a cognate DNA complex (Shimizu et al., 1997; PDB ID 1A0A). α-Helical regions are represented by cylinders. The darker shaded *Pho4* monomer is α-helical for residues Q13 to R16, R18 to L26, A32 to K35, and T48 to Q60. The lighter shaded *Pho4* monomer is α-helical for residues K4 to R18, V21 to H24, K45 to R54, and I56 to Q60. The DNA phosphodiester backbone is represented as a ribbon.

From the NMR data presented in this study, we have been able describe the secondary structure and backbone dynamics changes of *Pho4* that occur when it is bound to nonspecific and cognate DNA sequences. The formation of the nonspecific complex, which presumably is formed first when transcriptionally competent *Pho4* enters the nucleus, induces the random coil basic region to adopt a stable helical backbone conformation. Unlike the α-helices of the HLH fold, it is not clear whether the helical conformation of the basic region is a true α-helix. DNA binding also appears to induce elements of secondary structure in the loop residues V39 to A42. The nature of this secondary structure, absent from the co-crystal structure, is not known as the backbone dynamics indicate that these residues retain a significant degree of conformational flexibility and the observed NOEs may not reflect a single conformation. The formation of the cognate complex involves virtually no change in the secondary structure or backbone dynamics relative to the nonspecific complex. Thus, the nonspecifically bound complex lies directly along the pathway to the formation of the specifically bound complex. It has been proposed that the conformational changes and loss of overall conformational entropy associated with the formation of protein-DNA complexes are compensated by the binding free energy (Johnson et al., 1994; Spolar & Record, 1994). The similar backbone conformation and dynamics of the basic region in both nonspecific and cognate complexes implies that cognate sequence recognition and binding is dominated by protein side chain conformation and dynamics. Although the role side-chain dynamics must be crucial, the observation that residues E12

and Q13 in the basic region of the cognate complex may retain some backbone conformational flexibility suggests that backbone dynamics may have some influence binding site selection. Alternatively, additional proteins may be required for stability of the basic region.

Materials and methods

Cloning, protein purification, and sample preparation

The gene encoding the C-terminal 65 residues encoding the bHLH domain of Pho4 was given to our lab as a generous gift by Dr. E.K. O'Shea (UCSF, San Francisco, California). It was subcloned into pET21d (Novagen, Madison, Wisconsin) using *NcoI* and *EcoRI* restriction endonuclease sites. The plasmid was amplified using DH5*α Escherichia coli* cells (from Gibco/BRL, Rockville, Maryland) and transformed into the BL21(DE3) strain of *E. coli* cells for protein expression.

Protein expression and uniform isotopic enrichment of ^{15}N or $^{15}N/^{13}C$ were accomplished by growing the transformed BL21(DE3) cells on M9 growth medium. The M9 medium used $[^{15}N]NH_4Cl$ and $[^{13}C_6]$d-glucose (from Cambridge Isotope Laboratories, Andover, Massachusetts or Isotec Inc., Miamisburg, Ohio) as its sources of isotopic enrichment. Cells were grown at 37 °C to an OD of 0.6 and induced with isopropyl-β-D-thiogalactopyranoside (IPTG) from Calbiochem (La Jolla, California) so that the final concentration of IPTG was 5 mM. After induction, cells were allowed to grow another 8 h before being harvested. Following centrifugation, the cell pellets were resuspended in 40 mL of 50 mM Tris, pH 6.0, 1 mM ethylenediaminetetraacetic acid (EDTA) and 20 mM dithiothreitol (DTT). The resuspended pellet was sonicated on ice for six 15 s intervals separated by 45 s. The sonicated cellular debris was removed by centrifugation. The resulting supernatant was decanted and run through a gravity flow SP Sephrarose column (Pharmacia Biotech, Sweden). The cation exchange column was rinsed with 40 mL of 0.5, 1.0, 1.5, and 2.0 M NaCl solutions. *Pho4* elutes in the 1.0 M NaCl fraction and was further purified using reversed-phase high-performance liquid chromatography with a C18 column. The purity of the isolated *Pho4* bHLH fragment was determined by electrospray-ionization mass spectrometry. For the unbound form, NMR samples were made by dissolving the lyophilized protein in 30 mM sodium phosphate at pH 6.3 with 10 mM DTT and 1 mM NaN_3. All unbound NMR samples were 500 μL and contained 3–6 mM *Pho4* bHLH dimer.

For the DNA complexes, the DNA sequences were synthesized by Operon Technologies (Alameda, California). The sequences used were 5′-CGCGCGGGATCCCGCGCG-3′ for the nonspecific complex and 5′-CGCGCGCACGTGCGCGCG-3′ for the cognate sequence. To ensure proper annealing of the self-complimentary DNA sequences, the DNA sequences were dissolved in 30 mM sodium phosphate at pH 6.3 and heated to 95 °C before being slowly cooled back to room temperature. The concentrations of the DNA samples were quantified by measuring the absorbance at 260 nm while at 95 °C. The *Pho4* bHLH and DNA complexes were formed by slowly titrating the *Pho4* bHLH into a DNA solution with 30 mM sodium phosphate at pH 6.3, 10 mM DTT, and 1 mM NaN_3. The protein was added so that there was a 1:1 ratio of duplex DNA and *Pho4* bHLH dimer. The solutions of complex were concentrated to a final volume of 420 μL using centrifugation with YM-3 Centricon filters (Millipore, Bedford, Massachusetts). The final volume was adjusted to 500 μL with D_2O and the final concentrations of the protein/DNA complexes were ~0.8 mM for both the nonspecific and cognate samples.

CD measurements

All CD measurements were acquired with an AVIV CD spectrophotometer (Aviv Associates, Lakewood, New Jersey). Measurement of the *Pho4 Kd* values was performed by observing the ellipticity signal at 222 nm (the signal was averaged for 20 s). DNA was titrated into a 1 μM *Pho4* solution that was buffered with 30 mM sodium phosphate at 25 °C. The DNA was titrated from an initial concentration of 500 nM to a final concentration of 2 μM. The data were analyzed using Kaliedagraph (Synergy Software, Reading, Pennsylvania).

The spectra in Figure 5 were acquired at 35 °C with 2 μM of either unbound *Pho4* bHLH dimer or dimer/duplex DNA complex in 30 mM sodium phosphate buffer (pH 6.3) over the wavelength range of 190 to 300 nm. Signals were averaged for 3 s and incremented in 0.5 nm steps.

Polyacrylamide gel electrophoresis mobility shift assays

EMSA experiments were conducted using 8% 29:1 acrylamide-bisacylamide gels with a running buffer of 25 mM Tris-borate, pH 8.0, and 1 mM EDTA. Target DNA sequences were ordered as oligonucleotides that would have a 3 base pair 5′ overhang upon duplex formation. The duplex sequences were then incubated with Klenow fragment, dATP, γ^{32}P dCTP, dGTP, dTTP, and a labeling buffer (Amersham-Pharmacia Biotech) at 37 °C for 1 h. The Klenow fragment incorporated the ^{32}P label by filling in the 5′ overhang regions of the oligonucleotide sequence. The radio-labeled DNA was purified from the unincorporated nucleotides by using size exclusion spin chromatography columns (BioRad, Hercules, California).

In addition to *Pho4*, the labeled DNA (at 0.1 μM) was incubated at room temperature with calf thymus DNA, BSA, and 0.05% NP-40. After an hour, glycerol was added (10% final v/v) and loaded into the gel. Gels were run at 200 V for 15 min and then at 100 V for another 1 h and 45 min. The gels were imaged using phosphorimager (Molecular Dynamics Amersham Pharmacia, Piscataway, New Jersey), and the data were analyzed using Cricket Graph (Cricket Software, Malvern, Pennsylvania).

NMR experiments

All data were collected at 35 °C on either a Bruker AMX-600 or Bruker DRX-500 spectrometer (Bruker Instruments Inc., Karlsruhe, Germany) equipped triple axis or single axis gradient probe, respectively. The ^1H and ^{15}N correlation spectra were examined for suitable spectral dispersion using HSQC experiments (Mori et al., 1995). The triple resonance heteronuclear three-dimensional experiments utilized for backbone assignments were the HNCA (Grzesiek & Bax, 1992), HNCACB (Wittekind & Mueller, 1993), and CBCACONH (Bax & Pochapsky, 1992). The ^{15}N resolved $[^1H\text{-}^1H]$NOESY-HSQC (Talluri & Wagner, 1996) and ^{15}N resolved $[^1H\text{-}^1H]$TOCSY-HSQC experiments (Cavanagh & Rance, 1992) were used to supplement the triple resonance backbone assignment data and gather secondary structure information. The ^{15}N dynamics data were collected using $[^1H]^{15}N$-NOE, T_1 and T_2 experiments (Barbato et al., 1992).

All data processing was done on Silicon Graphics Indigo2 workstations (Silicon Graphics Instruments, Mountain View, California) with Felix 95 or 97 (BIOSYM/Molecular Simulations, San Diego, California). Directly acquired dimensions were treated with a convolution function to minimize the water signal and skewed sinebell apodization functions before Fourier transformations. The transformed, directly acquired dimensions were phased and baseline corrected using third order polynomial correction functions. Indirectly acquired dimensions were apodized with skewed sinebell functions before Fourier transformation and application of phase correction. Any additional baseline adjustments required after transforming the indirectly acquired dimensions were done using the baseline convolution function available in Felix 95 and 97.

Relaxation data analysis

Peak intensities of the processed dynamics spectra were measured using Felix 95 or 97. For the hetNOE data, the ratio of the peak intensities from the saturation experiment relative to reference experiment were calculated to determine the NOE enhancement. Three sets of hetNOE data were collected to determine an average hetNOE enhancement. For the T_1 and T_2 relaxation experiments, peak intensities were transferred to Kaliedagraph (Synergy Software, Reading, Pennsylvania) and plotted as a function of the corresponding relaxation delay. Using a least-squares fit algorithm provided by Kaliedagraph, each plot was fit to the equation $I(\tau) = I_0 \exp(R_1 \text{ or } R_2)\tau$, where I_0 is the initial peak intensity, $I(\tau)$ is the peak intensity at time point τ, and R_1 and R_2 are $1/T_1$ and $1/T_2$, respectively. Reduced spectral density mapping of the relaxation data was preformed according to the protocol of Farrow et al. (1995) and outline of Bracken et al. (1999).

Acknowledgments

We thank Dr. E.K. O'Shea of the University California, San Francisco for providing the PHO4 gene; Dr. David King of the University California Molecular Biology Department, and Howard Hughes Medical Institute for the ESI-MS analysis; Dr. Michael Nohaile, Dr. Corey Liu, and Dr. Jeffery G. Pelton for their assistance in the acquisition and analysis of the NMR data; Dr. A.G. Palmer for his help with Modelfree 4.0 and 4.01.

This work was supported by the Director, Office of Biological & Environmental Research, Office of Energy Research of the U.S. Department of Energy under contract number DE-AC03-76SF00098, and through instrumentation grants from the U.S. Department of Energy, DE FG05-86ER75281, and the National Science Foundation, DMB 86-09305 and BBS 87-20134. D.E.W. would also like to thank the Miller Institute for support during part of this work. J.W.C. was supported by N.I.H. Molecular Biophysics Training Grant GM08295. W.K. acknowledges the support of the Deutsche Forschungsgemeinschaft (DFG).

References

Anthony-Cahill SJ, Benfield PA, Fairman R, Wasserman ZR, Brenner SL, Stafford WF, Altenbach C, Hubbell WL, DeGrado WF. 1992. Molecular characterization of helix-loop-helix peptides. *Science 255*:979–983.

Barbaric S, Munsterkotter M, Goding C, Horz W. 1998. Cooperative Pho2-Pho4 interactions at the PHO5 Promoter are critical for binding of Pho4 to UASp1 and for efficient transactivation by Pho4 at UASp2. *Mol Cell Biol 18*:2629–2639.

Barbaric S, Munsterkotter M, Svaren J, Horz W. 1996. The homeodomain protein Pho2 and the basic-helix-loop-helix protein Pho4 bind DNA cooperatively at the yeast PHO5 promoter. *Nucleic Acids Res 24*:4479–4486.

Barbato G, Ikura M, Kay LE, Pastor RW, Bax A. 1992. Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy: The central helix is flexible. *Biochemistry 31*:5269–5278.

Bax A, Pochapsky SS. 1992. Optimized recording of heteronuclear multidimensional NMR spectra using pulsed field gradients. *J Magn Reson 99*:638–643.

Bracken C, Carr PA, Cavanagh J, Palmer AG. 1999. Temperature dependence of intramolecular dynamics of the basic leucine zipper of GCN4: Implications for the entropy of association with DNA. *J Mol Biol 285*:2133–2146.

Bruschweiler R, Liao X, Wright PE. 1995. Long-range motional restrictions in a multidomain zinc-finger protein from anisotropic tumbling. *Science 268*: 886–889.

Cavanagh J, Rance M. 1992. Suppression of cross-relaxation effects in TOCSY spectra via a modified DIPSI-2 mixing sequence. *J Magn Reson 96*:679–678.

Ellenberger T. 1994. Getting a grip on DNA recognition: structures of the basic region leucine zipper, and the basic region helix-loop-helix DNA binding domains. *Curr Opin Struct Biol 4*:12–21.

Ellenberger T, Fass D, Arnaud M, Harrison SC. 1994. Crystal structure of transcription factor E47: E-box recognition by a basic region helix-loop-helix dimer. *Genes Dev 8*:970–980.

Fairman R, Beran-Steed RK, Handel TM. 1997. Heteronuclear (^{1}H, ^{13}C, ^{15}N) NMR assignments and secondary structure of the basic region-helix-loop-helix domain of E47. *Protein Sci 6*:175–184.

Farrow NA, Zhang O, Szabo A, Torchia DA, Kay LE. 1995. Spectral density mapping using 15N relaxation data exclusively. *J Biol NMR 6*:153–162.

Ferre-D'Amare AR, Pendergrast GC, Ziff EB, Burley SK. 1993. Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature 363*:38–45.

Ferre-D'Amare AR, Pognonec P, Roeder RG, Burley SK. 1994. Structure and function of the b/HLH/Z domain of USF. *EMBO J 13*:180–189.

Garrell J, Campuzano S. 1991. The helix-loop-helix domain: A common motif for bristles, muscles and sex. *Bioessays 13*:493–498.

Gewith DT, Sigler PB. 1995. The basis for half-site specificity explored through a non-cognate steroid receptor-DNA complex. *Nat Struct Biol 2*:386–394.

Gilliquet V, Berben G. 1993. Positive and negative regulators of the Saccharomyces cerevisae "PHO system" participate in several cell functions. *FEMS Microbiol Lett 108*:333–340.

Grzesiek S, Bax A. 1992. Improved 3D triple resonance NMR techniques applied to a 31-kD protein. *J Magn Reson 96*:432–440.

Hirst K, Fisher F, McAndrew PC, Goding CR. 1994. The transcription factor, the Cdk, its cyclin and their regulator: Directing the transcriptional response to a nutritional signal. *EMBO J 13*:5410–5420.

Ishima R, Nagayama K. 1995. Protein dynamics revealed by quasi spectral density function analysis of amide N-15 nuclei. *Biochemistry 34*:3162–3171.

Jen-Jacobson L. 1997. Protein-DNA recognition complexes: Conservation of structure and binding energy in the transition state. *Biopolymers 44*:153–180.

Johnson NP, Lindstrom J, Baase WA, von Hippel PH. 1994. Double-stranded DNA templates can induce a-helical conformation in peptides containing lysine and alanine: Functional implications for leucine zipper and helix-loop-helix transcription factors. *Proc Natl Acad Sci USA 91*:4840–4844.

Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*:2577–2637.

Kaffman A, Rank NM, O'Neill EM, Huang LS, O'Shea EK. 1998a. The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus. *Nature 398*:482–486.

Kaffman A, Rank NM, O'Shea EK. 1998b. Phosphorylation regulates association of the transcription factor Pho4 with its import receptor Pse1/Kap21. *Genes Dev 12*:2673–2683.

Kunne AG, Sieber M, Meierhans D, Allerman RK. 1998. Thermodynamics of DNA binding reaction of transcription factor MASH-1. *Biochemistry 37*: 4217–4223.

Lee LK, Rance M, Chazin WJ, Palmer AG. 1997. Rotational diffusion anisotropy of proteins from simultaneous analysis of 15N and 13C alpha nuclear spin relaxation. *J Biol NMR 10*:287–298.

Lipari G, Szabo A. 1982a. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and validity. *J Am Chem Soc 104*:4546–4559.

Lipari G, Szabo A. 1982b. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J Am Chem Soc 104*:4559–4570.

Littlewood TD, Evan GI. 1995. Transcription factors 2: Helix-loop-helix proteins. *Protein Profile 2*:621–702.

Ma PCM, Rould MA, Weintraub H, Pabo CO. 1994. Crystal structure of MyoD bHLH domain-DNA complex: Perspectives on DNA recognition and implications for transcriptional activation. *Cell 77*:451–459.

Meierhans D, el-Ariss C, Neuenschwander M, Sieber M, Stackhouse JF, Allerman RK. 1995. DNA binding specificity of the basic-helix-loop-helix protein MASH-1. *Biochemistry 34*:11026–11036.

Mori S, Abeygunawardana C, Johnson MO, Vanzijl PCM. 1995. Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J Magn Reson B 108*:94–98.

Murre C, McCaw PS, Baltimore D. 1989. A new DNA binding and dimerization motif in immuglobulin enhancer binding, *daughterless*, *MyoD* and *myc* proteins. *Cell 56*:777–783.

Ogawa N, Oshima Y. 1990. Functional domains of a positive regulatory protein, PHO4, for transcriptional control of the phosphate regulon in *Saccharomyces cerevisae*. *Mol Cell Biol 10*:2224–2236.

O'Neil EM, Kaffman A, Jolly ER, O'Shea EK. 1996. Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex. *Science 271*:209–212.

O'Neil KT, Shuman JD, Ampe C, DeGrado WF. 1991. DNA induced increase in the α-helical content of C/EBP and GCN4. *Biochemistry 30*:9030–9034.

Oshima Y. 1982. Regulatory circuits for gene expression: The metabolism of galactose and phosphate. In: Strathen JN, Jones EW, Broach JR, eds. *The molecular biology of the yeast Saccharomyces: Metabolism and gene expression*. Cold Spring Harbor Laboratory, New York: Cold Spring Harbor. pp 159–180.

Parraga A, Bellsolell L, Ferre-D'Amare AR, Burley SK. 1998. Co-crystal structure of sterol regulatory element binding protein 1a at 2.3 A resolution. *Structure 6*:661–672.

Saudek V, Pasley HS, Gibson T, Gausepohl H, Frank R, Pastore A. 1991. Solution structure of the basic region from the transcriptional activator GCN4. *Biochemistry 30*:1310–1317.

Schneider KR, Smith RL, O'Shea EK. 1994. Phosphate-regulated Inactivation of the kinase PHO80-PHO85 by the CDK inhibitor PHO81. *Science 266*: 122–126.

Schurr JM, Babcock HP, Fujimoto BS. 1994. A test of the model-free formulas. Effects of anisotropic rotational diffusion and dimerization. *J Magn Reson B 105*:211–224.

Shao D, Creasy CL, Bergman LW. 1996. Interaction of *Saccharomyces cerevisae* Pho2 with Pho4 increases the accessibility of the activation domain. *Mol Gen Genet 251*:358–364.

Shimizu T, Toumoto A, Ihara K, Shimizu M, Kyogou Y, Ogawa N, Oshima Y, Hakoshima T. 1997. Crystal structure of PHO4 bHLH domain-DNA complex: Flanking base recognition. *EMBO J 15*:4689–4697.

Spolar RS, Record MT Jr. 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science 263*:777–784.

Talanian RV, McKnight CJ, Rutkowski R, Kim PS. 1992. Minimum length of a sequence specific DNA binding peptide. *Biochemistry 31*:6871–6875.

Talluri S, Wagner G. 1996. An optimized 3D NOESY-HSQC. *J Magn Reson B 112*:200–205.

Vogel K, Horz W, Hinnen A. 1989. The two positively acting regulatory proteins PHO2 and PHO4 physically interact with PHO5 upstream activation regions. *Mol Cell Biol 9*:2050–2057.

Weiss MA. 1990. Thermal unfolding studies of a leucine zipper domain and its specific DNA complex: Implications for a scissor's grip recognition. *Biochemistry 29*:8020–8024.

Winkler FK, Banner DW, Oefner C, Tsernoglou D, Brown RS, Heathman SP, Bryan RK, Martin PD, Petratos K, Wilson KS. 1993. The crystal structure of EcoRV endonuclease and its complexes with cognate and non-cognate DNA fragments. *EMBO J 12*:1781–1795.

Winston RL, Millar DP, Gottesfeld JM, Kent SB. 1999. Characterization of the DNA binding properties of the bHLH domain of Deadpan to single and tandem sites. *Biochemistry 38*:5138–5146.

Wishart DS, Sykes BD. 1994. Chemical shift as a tool for structure determination. *Methods Enzymol 239*:363–393.

Wittekind M, Mueller L. 1993. HNCACB, a high sensitivity 3D NMR experiment to correlate amide proton and nitrogen resonances with the alpha carbon and beta carbon resonances in proteins. *J Magn Reson B 101*:201–205.