Solution Structure of Porcine Delta Sleep-inducing Peptide Immunoreactive Peptide A Homolog of the *Shortsighted* Gene Product*

Gabi Seidel, Knut Adermann‡, Thomas Schindler§, Andrzej Ejchart, Rainer Jaenicke¶, Wolf-Georg Forssmann‡, and Paul Rösch‖

From the Lehrstuhl für Biopolymere, Universität Bayreuth, Universitätsstraße 30, D-95447 Bayreuth, Germany; ‡Niedersächsisches Institut für Peptid-Forschung GmbH, Feodor-Lynen-Straße 31, D-30625 Hannover, Germany; §Lehrstuhl für Biochemie, Universität Bayreuth, Universitätsstraße 30, D-95447 Bayreuth, Germany; and ¶Institut für Biophysik und Physikalische Biochemie, Universität Regensburg, D-93053 Regensburg, Germany

The 77-residue delta sleep-inducing peptide immunoreactive peptide (DIP) is a close homolog of the *Drosophila melanogaster shortsighted* gene product. Porcine DIP (pDIP) and a peptide containing a leucine zipper-related partial sequence of pDIP, pDIP(9–46), was synthesized and studied by circular dichroism and nuclear magnetic resonance spectroscopy in combination with molecular dynamics calculations. Ultracentrifugation, size exclusion chromatography, and model calculations indicated that pDIP forms a dimer. This was confirmed by the observation of concentration-dependent thermal folding-unfolding transitions. From CD spectroscopy and thermal folding-unfolding transitions of pDIP(9–46), it was concluded that the dimerization of pDIP is a result of interaction between helical structures localized in the leucine zipper motif. The three-dimensional structure of the protein was determined with a modified simulated annealing protocol using experimental data derived from nuclear magnetic resonance spectra and a modeling approach based on an established strategy for coiled coil structures. The left-handed super helical structure of the leucine zipper type sequence resulting from the modeling approach is in agreement with known leucine zipper structures. In addition to the hydrophobic interactions between the amino acids at the heptade positions a and d, the structure of pDIP is stabilized by the formation of interhelical i to i′ + 5 salt bridges. This result was confirmed by the pH dependence of the thermal-folding transitions. In addition to the amphipatic helix of the leucine zipper, a second helix is formed in the NH_2-terminal part of pDIP. This helix exhibits more 3_{10}-helix character and is less stable than the leucine zipper helix. For the COOH-terminal region of pDIP no elements of regular secondary structure were observed.

Delta sleep-inducing peptide immunoreactive peptide (DIP)[1]

‖ To whom correspondence should be addressed. Tel.: 49-921-553540; Fax: 49-921-553544; E-mail: paul.roesch@uni-bayreuth.de.

[1] The abbreviations used are: DIP, delta sleep-inducing peptide; pDIP, porcine DIP; hDIP, human DIP; bZIP, basic region/leucine zipper; MD, molecular dynamics; RMSD, root mean square deviation; Fmoc, *N*-(9-fluorenyl)methoxycarbonyl; TBTU, *O*-benzotriazol-1-yl-*N,N,N′,N′*-tetramethyluronium tetrafluoroborate; HPLC, high pres-

is a 77-residue NH_2-terminally acetylated peptide that was originally isolated from porcine brain (pDIP) using polyclonal antibodies against the delta sleep-inducing peptide (DSIP) (1, 2). Although DIP was detected via DSIP-specific antibodies that may recognize the sequence GGDA in DSIP and GGSA in pDIP, it is otherwise not sequence-related to this supposed sleep-inducing peptide. The pDIP sequence contains a putative leucine-zipper motif, a Pro/Glu rich domain, and three potential phosphorylation sites (2). DNA-binding capability of pDIP, however, is not obvious from its sequence, because it lacks the basic region found in the original basic region/leucine zipper (bZIP) DNA-binding domains. In the bZIP family of transcription factors, the leucine zipper acts as a dimerization domain and the upstream basic region as a DNA-binding domain (3, 4). Recently, the human analog of pDIP was characterized by cDNA analysis, showing that human DIP (hDIP) differs from the porcine protein in only four residues. Notably, Arg^{55} of pDIP is changed to a cysteine in hDIP. hDIP was shown to be present in a large number of tissues by reverse transcription-polymerase chain reaction/Southern hybridization (5).

pDIP shares significant homology with the murine TSC-22 protein and the product of the *Drosophila melanogaster shortsighted (shs)* gene, and the regions upstream of the leucine zipper are almost identical. TSC-22 was reported to be present in both the cytoplasmic and the nuclear fraction, and it has been discussed to function as a transcriptional regulator that is activated by transcription growth factor-$\beta 1$ and other growth factors of osteoblastic cells (6). The *Drosophila shs* gene product acts in the decapentaplegic pathway leading to photoreceptor differentiation (7). In contrast to TSC-22, *shs* gene product is a cytoplasmic protein present anterior to the furrow. The evolutionary conservation suggests the region upstream of the leucine zipper to be a distinct functional or structural domain.

The well known leucine zipper motif consists of two α-helices with the same sequential directionality forming a coiled coil. The coiled coil represents one of the most efficient packing modes of helices (8) and serves as a model for studies of protein stability and subunit interactions (9–13). The two-stranded α-helical coiled coil is characterized by a heptade repeat denoted as "abcdefg" where positions **a** and **d** are usually occupied by large hydrophobic amino acids such as Leu, Ile, and Val (14, 15), and positions **e** and **g** by oppositely charged amino acids (16, 17). Residues at the **a** and **d** positions have profound effects on the oligomerization states of coiled coils (18–20).

Much attention has been paid recently to the study of coiled

sure liquid chromatography; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy.

This paper is available on line at http://www.jbc.org

coil domains (21–24) or model α-helical peptides (20, 25, 26). bZIP domains were widely studied with a variety of experimental methods (27, 28). Here, however, we present experimental and computational data on an acetylated full-length leucine zipper protein not containing a DNA binding basic domain to more completely understand structural features of leucine zipper domains in the context of full-length proteins.

MATERIALS AND METHODS

Solid-phase Peptide Synthesis—Peptides were assembled using Fmoc chemistry on an automated peptide synthesizer (model 9050, PerSeptive Biosystems, Wiesbaden, Germany). Fmoc amino acids were purchased from Orpegen (Heidelberg, Germany) and PerSeptive Biosystems. Fmoc-Arg(Pbf) was from Sygena (Liestal, Switzerland). *N,N*-Dimethylformamide (peptide synthesis grade) and polyethylene glycol-polystyrine resins were purchased from PerSeptive Biosystems. TBTU was from Peboc (Llangefni, Wales). Acetonitrile (HPLC grade), acetic anhydride, dichloromethane, *tert*-butylmethyl ether, pyridine, piperidine, 1,2-ethanedithiol, and trifluoroacetic acid were purchased from Merck (Darmstadt, Germany). 1-Hydroxybenzotriazole and diisopropylethylamine were obtained from Fluka (Neu-Ulm, Germany). Solid phase synthesis of pDIP was carried out on a preloaded Fmoc-Val-polyethylene glycol-polystyrene resin (loading 0.19 mmol/g, 0.78 g), while fragment pDIP(9–46) was synthesized using a Fmoc-peptide amide liuber-polyethylene glycol-polystyrene amide resin (loading 0.16 mmol/g, 0.8 g). Acylations were performed in 30 min with a 4-fold excess of Fmoc-L-amino acid in the presence of TBTU/diisopropylethylamine/1-hydroxybenzotriazole (4 eq). Fmoc groups were cleaved by treatment with 20% piperidine in *N,N*-dimethylformamide for 10 min. After deprotection of the terminal amino groups, the peptidyl resins were acetylated with a mixture of dichloromethane/*N,N*-dimethylformamide/acetic anhydride/pyridine (40:40:19:1, volume) in 20 min at 0 °C. Subsequently, the resins were washed with *N,N*-dimethylformamide, 2-propanol, and dichloromethane (3×) and then dried to a constant weight. Resin cleavage and deprotection were carried out with a freshly prepared mixture of trifluoroacetic acid/ethanedithiol/water (94:3:3, volume, 10 ml/g resin) for 2 h. After filtration, the resin was washed with trifluoroacetic acid, and the crude peptide was precipitated by addition of chilled *tert*-butylmethyl ether, washed with *tert*-butylmethyl ether, and lyophilized from 5% acetic acid (crude yields: pDIP, 516 mg, 39.7%; pDIP(9–46), 225 mg, 38.7%). For purification, the dried crude products were dissolved in water (50 ml), loaded onto a Vydac C18 column (20 × 250 mm, 10 μm, 300 Å, The Separations Group, Hesperia, CA) and separated (buffer A: 0.06% trifluoroacetic acid in water; buffer B: 0.05% trifluoroacetic acid in acetonitrile/water, 4:1, linear gradient 20–100% buffer B in 80 min, flow rate 9 ml/min, detection at 230 nm). Pure fractions, detected by analytical HPLC (Vydac C18, 5 μm, 300 Å, 4.6 × 250 mm, flow, 0.8 ml/min, detection at 215 nm) were pooled and lyophilized. pDIP: yield, 66 mg (5.1%, calculated from initial resin loading). Molecular weight (M_r) by electrospray mass spectroscopy [M + 4H]$_{4+}$ 2189.5 (M_r calculated 8754.9). Amino acid composition (after hydrolysis with 6 N HCl at 150 °C for 90 min, 1090 Aminoquant, Hewlett-Packard): Ala 6.15 (6), Arg 3.80 (4), Asx 4.20 (4), Glx 19.50 (19), Gly 2.09 (2), His 1.06 (1), Ile 1.98 (2), Leu 9.92 (10), Lys 5.09 (5), Met 1.99 (2), Phe 1.01 (1), Pro 7.10 (7), Ser 4.69 (5), Thr 2.89 (3), Tyr 1.05 (1), Val 5.01 (5). pDIP(9–46): yield, 32.5 mg (5.6%). Electrospray mass spectroscopy, [M + 2H]$^{2+}$ 2271.5 (M_r calculated 4542.7). Amino acid composition: Ala 2.02 (2), Arg 2.82 (3), Asx 1.99 (2), Glx 10.41 (10), Ile 1.93 (2), Leu 6.11 (6), Lys 2.99 (3), Met 0.96 (1), Pro 0.97 (1), Ser 1.91 (2), Thr 2.07 (2), Tyr 1.10 (1), Val 2.92 (3).

Gel Filtration—Peptides were loaded at 4 μM–2.5 mM on a Superdex 75 column equilibrated in 50 mM potassium phosphate buffer, pH 7, 0.1 M NaSO₄, using a flow rate of 0.5 ml/min. The column was calibrated with bovine serum albumin (67 kDa), ovalbumine (43 kDa), chymotrypsinogen A (25 kDa) ribonuclease A (13.7 kDa), ubiquitin (8.5 kDa), and blue dextran 2000 (2 MDa).

Sedimentation Analysis—Sedimentation experiments were performed at room temperature in a Beckman model E analytical ultracentrifuge equipped with a high-intensity light source, 10" recorder, and AnH-Ti rotor. Measurements under the conditions of the NMR experiments made use of cells with ultra-thin pathlength and schlierenoptics: sedimentation velocity at 68,000 rpm, high-speed sedimentation equilibrium at 30,000 and 24,000 rpm; UV scanning at 240 nm. Initial peptide concentrations were 0.2 and 1.1 mM in 50 mM potassium phosphate buffer, pH 7, and 50 mM potassium phosphate buffer, 100 mM Na₂SO₄, pH 7, respectively. The partial specific volume was calculated from the amino acid composition. The sedimentation coefficient was calculated from the slope in an ln(*r*) *versus* time plot, the molecular mass distribution was fitted assuming a monomer-dimer equilibrium using the „multeq 3b" program kindly provided by Dr. A. Minton. For the determination of the diffusion coefficient, areas and maximal gradients were obtained from 20-fold magnifications of schlieren photographies.

CD Spectroscopy—CD spectra were measured on a Jasco J600 CD spectropolarimeter. A water bath was used to control the cell temperature. Data were collected with a 0.2 nm step resolution, a time constant of 1 s, and a scan speed of 20 nm/min. Wavelength scans were performed at discrete temperatures from 25 to 90 °C in a thermostatically controlled quartz cell of 0.5 or 0.05 cm pathlength, depending on peptide concentration. For each temperature point, spectra were obtained from 250 to 190 nm. Wavelength scans were processed by subtracting buffer scans taken at the same temperature and converting the data set to mean residual ellipticity prior to analysis. 10 mM phosphate buffer was used throughout.

Thermal transition curves were recorded from 5 to 90 or 95 °C at a fixed wavelength of 222 nm. The samples were heated at intervals of 2.5–5 °C. At each temperature the samples were allowed to equilibrate for 15–20 min. The reversibility of the denaturations was verified by slowly cooling the sample from final to initial temperature. Typically, more than 95% of the CD signal was recovered after cooling. The transition curves were normalized to the fraction of the unfolded peptide using the standard equation: $f_u = (\theta_{222} - \theta_n)/(\theta_u - \theta_n)$, where θ_n and θ_u represent the ellipticity values for the fully folded and fully unfolded species, respectively, and θ_{222} is the observed ellipticity at 222 nm. The transition temperature, T_m, of each heat-induced reversible denaturation was determined by fitting the CD signal change at 222 nm (θ_{222}) as a function of temperature to a two-state denaturation process using a least-squares fit program based on equations described elsewhere (29). The conformational stability ΔG^0 was determined as $\Delta G^0 = \Delta G + RT \ln P$, where P is the total protein concentration.

Secondary Structure Predictions—Secondary structure predictions were carried out using the Gibrat *et al.* (30), Levin *et al.* (31), double prediction method (32), self-optimized prediction methods from alignments (33), profile network from Heidelberg (34), secondary structure prediction (35), and PREDATOR (36) algorithms.

NMR Spectroscopy—NMR measurements were carried out on a standard Bruker AMX 600 spectrometer with 3.5 mM pDIP in 50 mM potassium phosphate buffer, pH 3.5, 298 K. Data from the following NMR spectra were employed for the sequence specific assignment of spin systems and the evaluation of nuclear Overhauser enhancement spectroscopy (NOESY) distance constraints: double quantum filtered correlated spectroscopy, total correlation spectroscopy with 40, 80, and 120 ms of mixing time, respectively, and a spin lock field of approximately 12 kHz and NOESY with mixing times of 100 and 200 ms, respectively. Solvent suppression was performed by continuous coherent irradiation prior to the first excitation pulse and during the mixing time in the NOESY experiment. The sweep width in ω_1 and ω_2 was 7246 Hz for all spectra. Quadrature detection was used in both dimensions with the time proportional phase incrementation technique in ω_1 (37). 4 K data points were collected in ω_2 and 512 data points in ω_1. Zero filling of the time domain data resulted in a frequency-domain matrix with 1 and 2 K data points in ω_1 and ω_2, respectively. All two-dimensional spectra were multiplied by a squared sineball function phase shifted by $\pi/4$ for NOESY and for total correlation spectroscopy spectra and by $\pi/8$ for COSY spectra. Base-line correction to the seventh order was used. Data were evaluated on X-window workstations with the NDee program package (Software Symbiose, Inc., Bayreuth, Germany). Chemical shift values are reported relative to external 2,2-dimethyl-2-silapentane sulfonate.

Structure Calculations—Distance information was obtained from two-dimensional NOESY spectra in H₂O/²H₂O (9:1) and in 99.998% ²H₂O. NOESY cross peaks were classified into three categories according to their volume intensity as estimated from the number of contours in NOESY spectra: strong 0.18–0.27 nm; medium, 0.18–0.4 nm; weak, 1.8–5.5 nm. Pseudoatom corrections were used to adjust distances that involved nonstereospecifically assigned protons such as methyl groups or aromatic ring protons (38). Three-dimensional structures were calculated with the X-PLOR 3.1 package (39). The standard protocols for *ab initio* simulated annealing and simulated annealing refinement were applied with some modifications. The initial structure calculations started from an extended template with satisfactory local geometry. For the leucine zipper domain a modeling approach for coiled coil proteins (40) was used. The method draws upon knowledge of the oligomerization state, the helix directionality, and the properties of heptade repeat sequences. Unknown structural parameters are heavily sampled. The
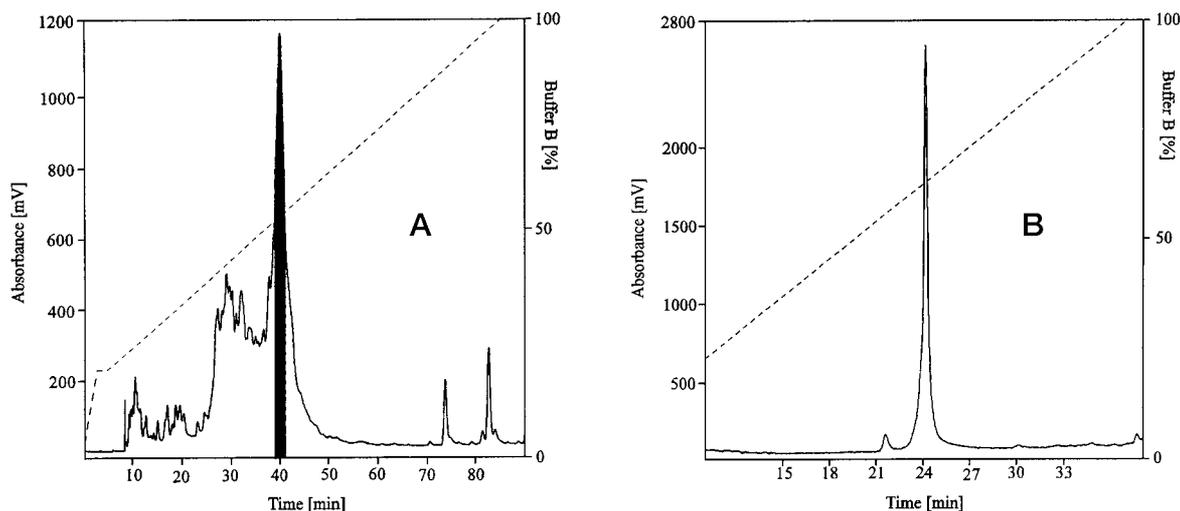
FIG. 1. **Reversed-phase HPLC of crude (preparative run) (A) and purified pDIP (B) (for conditions see "Material and Methods").** Fractions according to the dark area of *panel A* were pooled to the pure product. The small peak at 22 min is due to a single Met(O) formation after long time storage.

coiled coil twist angle, for example, is sampled with an initial range from −35° to 35° in one degree increments. The initial C_α positions were those of a regular α-helix, and the initial separation between the helices was set to 10 Å. For each initial structure, side chain and backbone atoms were grown from the C_α position by applying a protocol similar to those used for the generation of initial coordinates in NMR structure determination (41). Each structure was relaxed with the following simulated annealing protocol: (i) a 5 ps molecular dynamics slow-cooling stage from 500 to 300 K, (ii) a 20 ps constant temperature molecular dynamics (MD) simulation at 300 K, and (iii) 1000 steps of conjugate gradient minimization. During the slow-cooling stage, C_α atoms were held in place with harmonic point restraints that were slowly reduced. The helical hydrogen bond restraints were active during all stages, but no other restraints were applied during the constant temperature MD and energy minimization stages. A time step of 0.5 fs was used for temperatures above 350 K during the slow-cooling stage, otherwise a time step of 1 fs was applied. The coordinates for the residues in the leucine zipper domain thus obtained were used as a reference set in the *ab initio* simulated annealing and the simulated annealing refinement to restrain the main coordinate set. The restraints were incorporated as point restraints in the form of a harmonic potential.

Unrestrained MD—Unrestrained MD calculations were carried out using the parameters for a representative leucine zipper structure and the TIP3P water model (42) that was supplied with the standard X-PLOR force field (39). The overlay was achieved by placing the protein in the center of a cubic water box (6.33 nm) and deleting all solvent molecules closer than 0.26 nm to any heavy atom of the protein. Close nonbonded solute-solvent interactions were removed in two steps. First, 100 cycles of conjugate gradient energy minimization (43) were carried out, keeping the positions of all protein atoms fixed. Second, in 300 cycles of energy minimization, a harmonic potential was used to restrain the protein to its original conformation. During the first 15 ps of the MD calculations, the system was gradually heated to 300 K while coupled to an external water bath (44). The MD calculations were carried out using the Verlet algorithm (45) with a time step of 1 fs. The SHAKE algorithm (46) was used to constrain covalent bond lengths. A dielectric constant of 1.0 was applied with a scaling factor of 0.4 for 1–4 electrostatic interactions. All nonbonded interactions were cut off at a distance of 0.85 nm. During the whole simulation of 200 ps, minimum image periodic boundary conditions were used. Coordinates, energies, and velocities were saved every 0.5 ps for further analysis. Simulations and analyses were performed on Hewlett Packard HP 735 computers. A 1-ps simulation required about 3 h of CPU time.

A second unrestrained MD calculation was carried out with the leucine zipper of GCN4 using the crystal structure (47) (PDB coordinates 2ZTA) as a starting structure. Protons were added to this structure, and an energy minimization was carried out.

RESULTS AND DISCUSSION

Solid-phase Synthesis—In contrast to many other peptides with comparable lengths prepared by solid-phase peptide syn-

thesis, application of well established Fmoc chemistry to the synthesis of pDIP resulted in a crude product containing the desired product as the major component. The analytical reversed-phase HPLC of the crude product shows pDIP as the last eluting compound of the highest absorbance (Fig. 1). Because most other components in the crude product are characterized by significantly shorter retention times, it is most probable that these are by-products were derived from truncations during solid-phase assembly of the peptide. The COOH-terminal Pro/Glu-rich domain, however, caused no difficulties. HPLC from the crude pDIP(9–46) indicated a similar quality of the crude product. Both peptides were obtained in a purity of about 93% after only one preparative HPLC separation according to mass spectrometry and HPLC analysis. pDIP and the leucine zipper-containing fragment are eluted from the analytical HPLC at about 60% buffer B (Fig. 1), demonstrating that the peptides are retarded by significant interactions with the hydrophobic column.

Sequence Analysis—Residues 19 to 40 of pDIP have been proposed to form a leucine zipper (2). The arrangement of the sequence in an α-helical structure may be modeled schematically (Fig. 2). For clarity, the helical wheel representation for the maximal five heptades (Met[9]-Leu[43]) is shown as a helix with a pitch of 3.5 residues/turn as found in coiled coil structures. The interface between the two helices consists of hydrophobic residues with the exception of Asn[31] and Asn[38]. Asparagine pairs are proposed to form an interhelical H-bond at the GCN4 homodimer interface (47). Recently, Lumb and Kim (48) showed that these Asn residues impart specificity for the formation of a two-stranded parallel coiled coil at the expense of stability. Electrostatic interactions between charged residues at positions **e** and **g** are sterically allowed in coiled coil structures (47) (Fig. 2, *curved arrows*), and formation of salt bridges between the two helices could further enhance dimer stability (49, 50).

Protein secondary structure prediction suggests a high amount of helical secondary structure for pDIP, and a long helix starting from the NH_2 terminus and spanning the leucine zipper domain (with the exception of the residues around Asn[30]) is predicted by all of the methods used (Fig. 2). The algorithm by Woolfson and Alber (51) distinguishes between amino acid sequences of dimeric and trimeric coiled coils. This method predicts a preference for dimer formation of the pDIP leucine zipper domain. The Δ-score measuring the difference of

```
         .   10    .   20    .   30    .   40    .   50    .   60    .   70    .
             |        |        |        |        |        |        |
         MDLVKNHLMYAVREEVEILKEQIRELVEKNSQLERENTLLKTLASPEQLEKFQSRLSPEEPAPETPEAPEAPGGSAV

1        HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH-H-HHHHHHHHH--------------------EE
2        HHHHHHHHHHHHHHHHHHHHHHHHHHH---H-HHHHHHHHH----HHHH--HH--T------------T---HE
3        --EEHHHEHHHEHHHHHHHHHHHHHHEHHHHHHHHHHHHH-HT-HHHHHHH-HHT----T----HT--TT-----
4        HHHHHHHHHHHHHHHHHHHHHHHHHH------HHHHHHHHH---HHHHHHHH--------------------HE
5        ---HHHHHHHHHHHHHHHHHHHHHHHH-HHHH-HHHH------HHHHHHHHH--------------------
6        HHHHHHHHHHHHHHHHHHHHHHHHHHHH-HHHHHHHHHHHHHH---HHHHHH--------------------
7        ---HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH----HHHHHHHH------------------

Consensus HHHHHHHHHHHHHHHHHHHHHHHHHHHHH-HHHHHHHHHHHHH---HHHHHHHH--------------------
```


FIG. 2. **Secondary structure of pDIP.** The predictions were carried out using the algorithms by Gibrat (30), Levin (31), and the DPM (32), SOPMA (33), PhD (34), SSPRED (35), and PREDATOR methods (36). The consensus prediction shows the type of secondary structure (*H*, helix; *T*, turn; *E*, sheet; —, no determined element) that is present in more than 50% of the predictions. The hydrophobic positions of the 3–4 repeat are *boxed*. Helical wheel representation of residues 9 to 43 of pDIP (3.5 residues/turn) are shown in the *bottom panel*. View is from the NH₂ terminus and the heptade positions are labeled **a** through **g**. Residues that would be able to form intermonomeric i to i' + 5 ion pairs under the assumption of a leucine zipper type coiled coil structure are connected with *solid arrows*. *Dashed arrows* indicate putative intramonomeric i to $i + 3$ or i to $i + 4$ electrostatic attractions.

dimer and trimer formation propensities of the test sequence was found to be 2.4, characteristic for dimer formation. Sequence analysis thus suggests that pDIP forms a two-stranded coiled coil. Dimerization may be achieved through a leucine zipper in which the hydrophobic interaction is complemented by the interaction of oppositely charged residues. Our subsequent spectroscopic studies were aimed at further defining structural characteristics of pDIP in solution.

Oligomerization State—pDIP fragment-(9–46) eluted from a Superdex 75 gel-filtration column as a symmetric single peak with the molecular weight expected for a dimer (Fig. 3). In contrast, full-length pDIP eluted as a symmetric peak with an apparent molecular mass of 41.2 kDa, closer to a pentameric (43.8 kDa) than to the expected dimeric (17.5 kDa) form, indicating the presence of higher oligomers in solution (Fig. 3).

The state of association of pDIP was further analyzed by sedimentation equilibrium ultracentrifugation, because this technique does not suffer as much from peak shape dependence as does gel filtration. On the contrary, boundary analysis and curve fitting allow the minimal molecular mass and the association constant of assemblies to be quantified. Making use of an ultrathin cell (2 mm pathlength), combined with schlierenoptics, the sedimentation coefficient and diffusion constant were determined under the conditions of the NMR experiments. At c_{pDIP} = 2 mM and 68,000 rpm, no significant trailing of the schlieren peak was detectable; with $s^0_{20,w}$ = 1.09 ± 0.02 S and $D_{20,w}$ = (6.59 ± 0.10) 10^{-7} cm² s⁻¹, the maximum molecular mass is found to be $M_{s,D}$ = 15,720 ± 1300 Da.

Sedimentation equilibrium experiments made use of the meniscus depletion technique (52) applying *A*. Minton's „multeq 3b" program, the $A_{280\ nm}$ *versus r* profiles were fitted assuming a monomer-dimer equilibrium. Fig. 4 shows the result; the fit is

FIG. 3. **Gel filtration analysis.** Elution profiles are shown for pDIP (—) and pDIP(9–46) (– – –). The elution times for the standards are represented with *bars* at the *top*.

perfect for both potassium phosphate buffer in the absence and presence of 100 mM Na₂SO₄. The residuals do not exceed ±0.01. The association constant is K_a = 8.0 10^2 M⁻¹. The minimum molecular mass, extrapolating the ln(*c*) *versus* r^2 profile to zero concentration (meniscus depletion), yields weight average M_r = 8730 ± 1295 in accordance with the calculated monomer.

CD Spectroscopy—To examine helical content and thermal stability of pDIP, CD studies were performed at different temperatures (Fig. 5*B*). The well defined isodichroic point around 202 nm suggests that the denaturation equilibrium can be described on the basis of a two-state structural transition. Minima at 207 and 221 nm and a maximum at 190 nm dem-

Solution Structure of pDIP

FIG. 4. **Sedimentation equilibrium analysis of pDIP.** A sample was run at 23 °C in 50 mM potassium phosphate, 100 mM Na$_2$SO$_4$, pH 7. Rotor speed used was 30,000 rpm. Scans at two different concentrations (○ and ●) were fitted assuming a 2 M ⇌ M$_2$ equilibrium with M = 8755 Da. The *curve* depicts the fit; the *upper frames* depict the residuals.

onstrate that pDIP is largely α-helical as predicted. Spectra for pDIP(9−46) also indicate predominantly α-helical structure (Fig. 5*A*). The helix content for pDIP and pDIP(9−46) was determined according to Price (53) to be 57 and 61%, respectively. The lower relative helical content for pDIP(9−46) (74% expected for the four heptades) could reflect the occurrence of two pairs of Asn at **a** positions. These polar residues are known to be less stabilizing compared with hydrophobic residues at this position in the heptade repeat (48). A similar result was obtained for CD studies of the Max homodimer, a member of the b-HLH-ZIP family (54).

Thermal Denaturation—The peak of negative ellipticity at 222 nm differs significantly in the spectrum of native and denatured pDIP providing a probe for the thermal unfolding reaction that was studied at acidic and neutral pH (Fig. 6*A*). The sigmoidal shape of the curves suggests a two-state transition in both pH regions, implying a cooperative transition. The melting temperature, defined as the temperature where 50% of the peptide is in the unfolded state, increases from 326 to 337 K when the pH is increased from 2.5 to 6.

The melting temperature is concentration dependent, since it decreases with decreasing protein concentration (Fig. 6*B*). This observation is consistent with the idea of unfolding connected to dissociation of the two pDIP subunits. If the folded
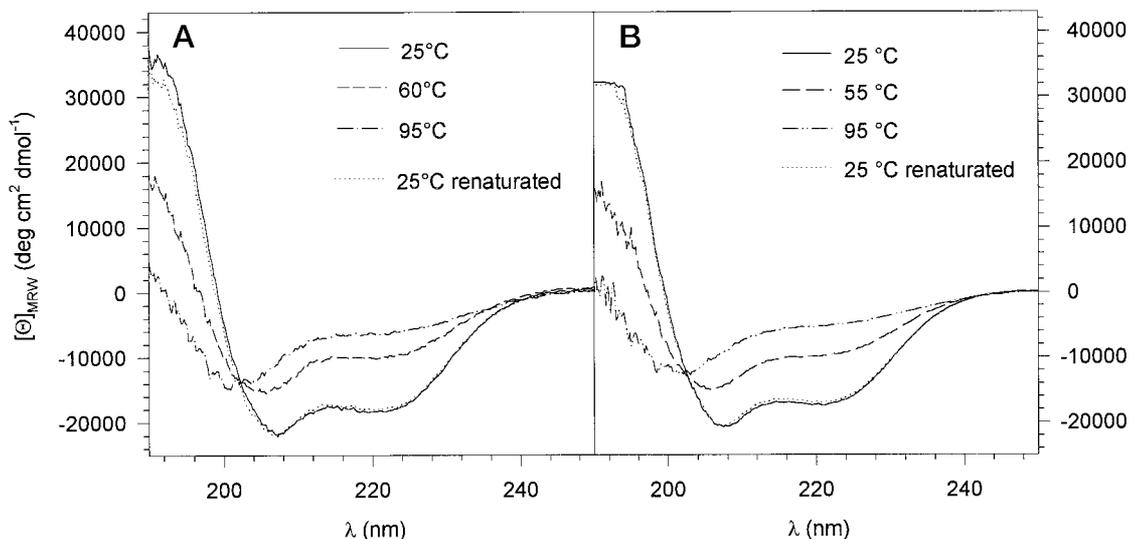
FIG. 5. **Far UV circular dichroism spectra.** *A*, CD spectra of pDIP(9−46) recorded at pH 6 and a peptide concentration of 80 μM with varying temperature. Measured rotations are converted to mean residue ellipticities. *B*, CD spectra of pDIP recorded at pH 2.5 and a peptide concentration of 80 μM with varying temperature. Measured rotations are converted to mean residue ellipticities.
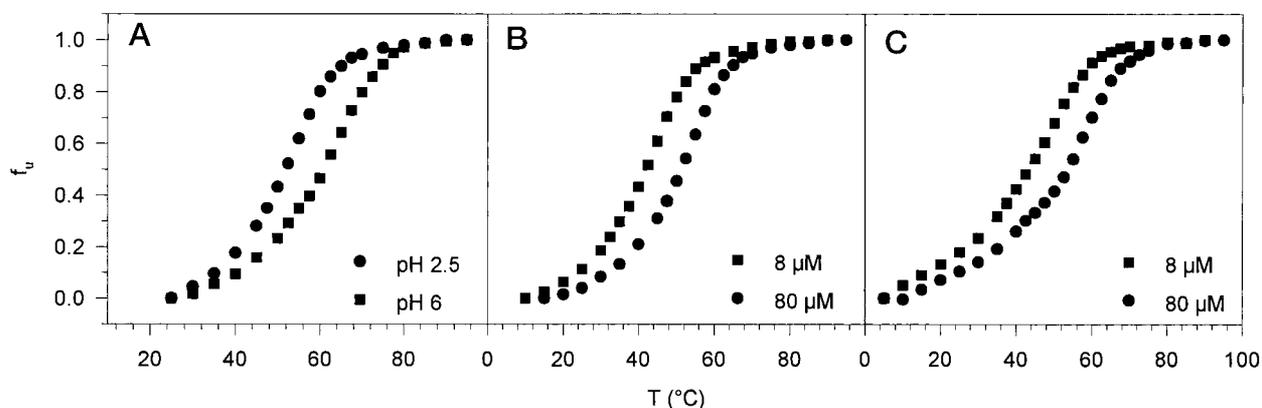
FIG. 6. *A*, thermal unfolding curves for pDIP as a function of pH. The data were recorded at pH 2.5 (●) and pH 6 (■) at a concentration of 80 μM. Thermal unfolding curves as a function of concentration are shown as follows: *B*, data for pDIP at concentrations of 8 μM (■) and 80 μM (●) at pH 6; *C*, data for pDIP(9−46) at concentrations of 8 μM (■) and 80 μM (●) at pH 6.
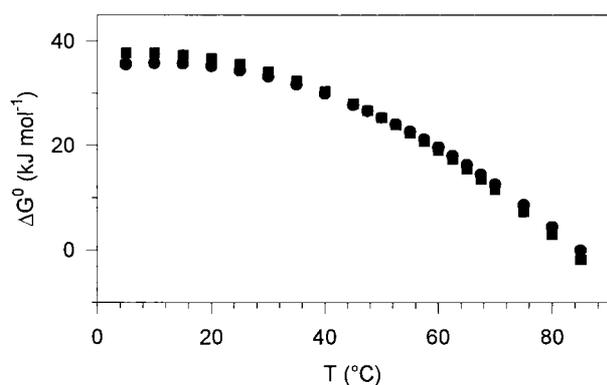
FIG. 7. **Unfolding free energy as a function of temperature.** Data are shown for a protein concentration of 8 μM ($\blacksquare$) and 80 μM ($\bullet$).

monomer state is essentially unpopulated, unfolding can be described with the two-state model and the energetics can be readily determined in a manner similar to that for monomer denaturation:

$$N_2 \rightleftharpoons 2D \qquad \text{(Eq. 1)}$$

where

$$K_U = [D]^2/[N_2] = 2P_t[f_u^2/(1 - f_u)] \qquad \text{(Eq. 2)}$$

with P_t being the total protein concentration and f_u the fraction of the unfolded protein. The two states are the native dimer, N_2, and the denatured monomer, D. If this two-state model for the unfolding reaction of a dimer provides a reasonable description, then identical values of ΔG^0 (value for a protein concentration of 1 M) should follow from experiments performed at different protein concentration. Fig. 7 shows unfolding free energies calculated for two different protein concentrations of pDIP as a function of temperature. The free energies calculated at different protein concentrations assuming the two-state model for the unfolding reaction of a dimeric protein are the same within experimental error, consistent with this model. The conformational stability ΔG^0 of 35 kJ/mol is comparable to the free enthalpy determined for the Arc repressor (40 kJ/mol) (29) or the dimerization domain of transcription factor LFB1 (48 kJ/mol) (55). For physiological protein concentrations, the free enthalpy is lower than 10 kJ/mol; a reasonable value if pDIP acts as a regulatory protein stabilized by ligand interaction.

The pH dependence of the thermal stability of pDIP could be explained with the formation of interhelical salt bridges between ionizable side chains at positions **e** and **g** in the heptade repeat. The interhelical salt bridges of pDIP would thus contribute to the stability of the coiled coil. Recently, Yu *et al.* (56) have also shown that at low ionic strength (10 mM) the coiled coil was much more stable against temperature at neutral pH than at acidic pH, indicating that the Lys-Glu ion pairs present in the coiled coil interface at neutral pH contribute significantly to the stability of the coiled coil, contrasting the results from Lumb and Kim (57).

The results of a study of the concentration dependence of the thermal unfolding transition for pDIP(9–46) are in agreement with that found for the full-length peptide (Fig. 6*C*), evidencing the existence of a dimerization area in the leucine zipper domain. A comparison of the shape, however, shows that the thermal unfolding transition of pDIP(9–46) is biphasic rather than monophasic, indicating a three state transition. This fact may be explained by the existence of native monomers during the denaturation process or by the formation of higher oligomers in the native state of the peptide. The differences in the unfolding reaction of pDIP and pDIP(9–46) indicate an influence of the residues outside the leucine zipper domain on for-

TABLE I
^{1}H chemical shifts and assignments for pDIP at pH 3.5, 298 K relative to external DSS, accuracy $\pm$ 0.01 ppm

Residue	HN	$C_\alpha H$	$C_\beta H$	$C_\gamma H$	Others
Met[1]	8.40	4.31	1.97	2.58; 2.50	
Asp[2]	8.56	4.57	2.82; 2.71		
Leu[3]	8.02	4.21	1.67	1.55	0.89; 0.80 (δ)
Val[4]	7.83	3.88	2.07	0.93; 0.87	
Lys[5]	8.18	4.12	1.75	1.42; 1.32	2.95 (ϵ); 1.65 (δ); 7.53 (ϵNH)
Asn[6]	8.13	4.60	2.78		7.61; 6.90 (γNH$_2$)
His[7]	8.41	4.53	3.30		8.62 (H2); 7.27 (H4)
Leu[8]	8.19	4.21	1.70	1.57	0.89; 0.82 (δ)
Met[9]	8.15	4.31	2.03	2.48; 2.55	
Tyr[10]	7.99	4.21	3.04		7.05 (H2/6); 6.76 (H3/5)
Ala[11]	8.07	4.12	1.45		
Val[12]	7.93	3.88	2.10	1.02	
Arg[13]	8.04	3.90	1.85	1.75; 1.58	7.22 (δNH); 3.13 (δ)
Glu[14]	8.45	4.23	1.95	2.44	
Glu[15]	8.29	4.24	1.99;1.88	2.41	
Val[16]	8.10	3.39	2.17	0.94; 0.85	
Glu[17]	7.59	3.98	2.16	2.48	
Ile[18]	8.04	3.75	1.95	1.66; 1.17; 0.89	0.78 (δ)
Leu[19]	8.59	3.90	1.81	ND	0.92; 0.78 (δ)
Lys[20]	8.27	4.07	ND	1.48	7.54 (ϵNH); 2.92 (ϵ); 1.68 (δ)
Glu[21]	7.94	ND	2.11	ND	
Gln[22]	8.62	4.09	ND	ND	
Ile[23]	7.65	3.76	1.97	1.63; 1.20	0.89 (δ)
Arg[24]	7.96	3.91	1.94	1.80; 1.58	7.36 (δNH); 3.22 (δ)
Glu[25]	ND				
Leu[26]	8.35	3.47	2.11	ND	0.82 (δ)
Val[27]	8.99	3.63	2.15	1.07; 0.93	
Glu[28]	7.84	4.13	1.97	2.20	
Lys[29]	8.26	4.04	1.49	1.31	7.54 (ϵNH); 2.91 (ϵ); 1.76 (δ)
Asn[30]	8.86	4.08	3.26;2.65		7.68; 6.35 (γNH$_2$)
Ser[31]	8.42	4.21	4.06; 4.00		
Gln[32]	ND				
Leu[33]	ND				
Glu[34]	ND	3.94	1.95	2.07	
Arg[35]	7.81	4.07	1.94	1.79; 1.58	7.24 (δNH); 3.22 (δ)
Glu[36]	7.96	4.21	2.01	2.46	
Asn[37]	8.73	4.06	3.18; 2.66		7.64; 6.50 (γNH$_2$)
Thr[38]	7.71	4.34	4.14	1.31	
Leu[39]	7.50	4.21	1.66		0.79 (δ)
Leu[40]	ND				
Lys[41]	8.60	4.11	1.45	1.34	7.53 (ϵNH); 2.96 (ϵ); 1.64 (δ)
Thr[42]	7.94	4.41	4.02	1.12	
Leu[43]	ND				
Ala[44]	8.10	4.35	1.27		
Ser[45]	8.34	4.26	3.94; 3.90		
Pro[46]		4.38	1.85	2.01	3.75; 3.63 (δ)
Glu[47]	8.40	4.26	1.93	2.42	
Gln[48]	7.65	4.08	1.98	2.23	
Leu[49]	8.63	3.76	1.82	1.56	0.64; 0.30 (δ)
Glu[50]	8.04	4.11	2.10	2.47	
Lys[51]	8.10	4.16	1.59	1.46	7.50 (ϵNH); 2.90 (ϵ); 1.64 (δ)
Phe[52]	8.15	4.30	3.30; 3.00		7.25 (H4); 7.16 (H3/5); 7.13
Gln[53]	8.07	4.02	2.31	2.58	7.02 (H2/6); 6.61 (δNH$_2$)
Ser[54]	7.96	4.25	3.77		
Arg[55]	7.78	4.27	1.90	1.75; 1.66	7.23 (δNH); 3.11 (δ)
Leu[56]	7.63	4.18	1.44	ND	0.67 (δ)
Ser[57]	7.89	4.66	3.84; 3.76		
Pro[58]		4.36	2.25; 1.86	1.97	3.66 (δ)
Glu[59]	8.34	4.27	2.04; 1.91	2.44	
Glu[60]	8.22	4.65	2.06; 1.85	2.43	
Pro[61]		4.27	2.03; 1.91	1.97	3.78; 3.64 (δ)
Ala[62]	8.27	4.54	1.31		
Pro[63]		4.39	2.25; 1.85	1.96	3.74; 3.59 (δ)
Glu[64]	8.41	4.32	2.06; 1.93	2.43	
Thr[65]	8.17	4.54	4.11	1.19	
Pro[66]		4.36	2.24; 1.85	1.98	3.80; 3.67 (δ)
Glu[67]	8.34	4.27	2.03; 1.91	2.44	
Ala[68]	8.31	4.55	1.30		
Pro[69]		4.41	2.24; 1.85	1.95	3.79; 3.66 (δ)
Glu[70]	8.30	4.27	2.04; 1.91	2.41	
Ala[71]	8.34	4.52	1.31		
Pro[72]		4.38	2.24; 1.87	1.97	3.77; 3.62 (δ)
Gly[73]	8.48	3.93			
Gly[74]	8.25	3.96			
Ser[75]	8.19	4.42	3.80		
Ala[76]	8.31	4.34	1.35		
Val[77]	7.91	4.11	2.09	0.87	

^{a} ND, not determined.

mation and stability of the dimer.

NMR—The NH–NH region of the NOESY spectrum often can be used to define elements of secondary structure. From the
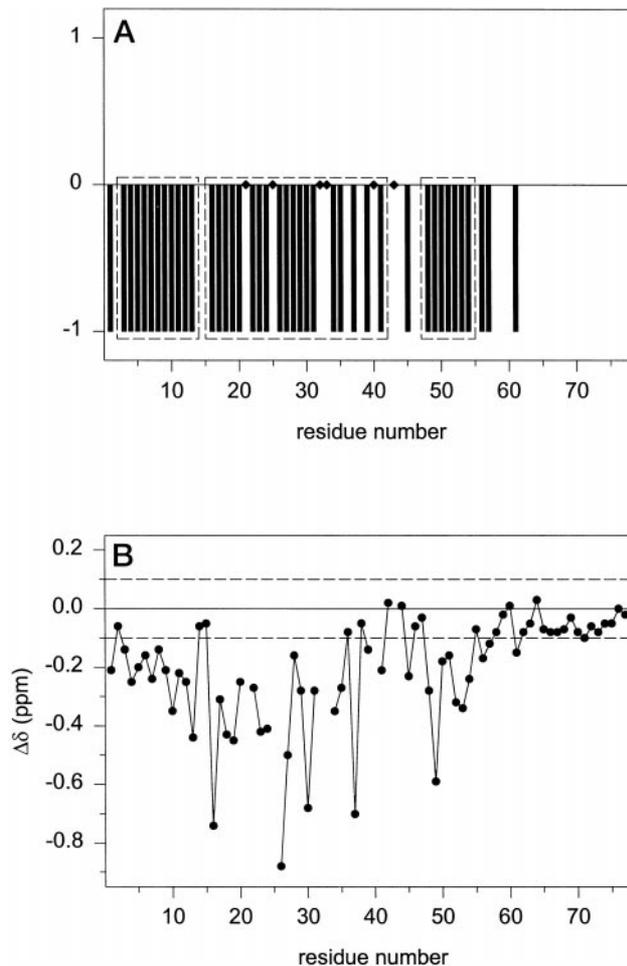
FIG. 8. *A*, chemical shift index of the C_α proton frequencies. Amino acids for which the chemical shift of the C_α proton was not determined are marked by *diamonds* (◆). Stretches of negative chemical shift indices (−1), which are characteristic for α-helical structure, are *boxed*. *B*, chemical shift differences Δδ in ppm of the C_α proton frequencies. *Dashed lines* indicate chemical shift differences of +0.1 ppm and −0.1 ppm, respectively, the values characteristic for the formation of secondary structure.

large α-helix content observed in the CD spectra, an extensive array of interresidue connectivities is expected in the NH–NH region of the NOESY spectrum of pDIP. The sharp cross peaks observed in this region were found, however, to correspond only to the residues NH_2-terminal of the leucine zipper domain. In addition to the absence of narrow interresidue NH–NH cross peaks between amino acids of the leucine zipper domain, the fingerprint region in the double quantum filtered COSY and the total coherence spectroscopy experiments show only 60% of the total number of cross peaks expected for the amino acid sequence of pDIP. Following standard methodology (38) it was possible to assign these resonances to the residues NH_2- and COOH-terminal of the leucine zipper domain.

In addition to the narrow cross peaks, broad resonances of generally small intensity were found in the NOESY spectrum. These resonances correspond to residues from the leucine zipper domain. The lineshape of these cross peaks could result from an exchange between monomeric and dimeric or oligomeric states of pDIP, the life time of these states being intermediate on the NMR time scale. Additionally, the leucine zipper domain may indeed be a rigid region, and thus protons therein would be subject to strong cross-relaxation and poor J-coupling (58). A similar behavior of the two subdomains in BR-LZ containing the DNA-binding domain and the leucine zipper domain of GCN4 was observed earlier (27).

Despite these difficulties, it was possible to assign a large number of backbone resonances starting from some well separated NH resonances in the downfield region. For this assignment, NH–NH NOEs and both sequential $C_\alpha H–NH(i + 1)$ NOEs and medium range $C_\alpha H–NH(i + 3)$ NOEs from the better resolved $C_\alpha H$ resonances were used (59). The strong sequential overlap in the side chain region does not allow tracing of the complete spin system for each backbone assignment. Using the two different assignment procedures for the residues from the leucine zipper domain and the remaining residues, the unambiguous sequence-specific assignment of 72 out of 77 amino acids was possible.

Comparing the chemical shifts of the Asn side chain protons, the differences between Asn[6] in the NH_2-terminal region, and Asn[30] and Asn[37] in the leucine zipper (Table I) are obvious. Whereas the side chain chemical shifts of Asn[6] show random coil values (38), Asn[31] and Asn[38] show an extreme chemical shift separation (0.5–0.6 ppm) of the C_β protons and (1.1–1.3 ppm) of the γNH_2. This observation may be explained by the formation of one or more hydrogen bonds (22, 47).

FIG. 9. **Survey of the short and intermediate range NOEs observed for pDIP.** The *height* of the bars indicate the intensities of the NOEs. *Gray bars* indicate that the NOE could not be observed because of frequency degeneration.

The deviation of C_α proton chemical shifts from their random coil positions are sensitive indicators of secondary structure. Therefore, a secondary structure estimate according to the chemical shift index strategy (60) was used (Fig. 8*A*). C_α proton resonances shifted to high field, relative to the corresponding random coil C_α proton resonances, indicate local α-helical structure. High field shifts more than 0.1 ppm are marked by -1. For interpretation of the chemical shift index only resonances should be taken into account with the same sense chemical shift deviation for a stretch of more than three sequential residues (60). This procedure indicates for pDIP two helical regions, extending from Leu[3] to Arg[13] and from Gln[48] to Ser[54]. Assuming the chemical shift indices of Glu[21], Glu[25], Gln[33], and Leu[40] to be -1 a third helical region from Val[16] to Lys[41] can be predicted. The stretch of negative chemical shift indices is interrupted around Asn[37], one of the two polar residues known to be less stabilizing as compared with hydrophobic residues at this position in the heptade repeat (47, 48, 61).

A more quantitative description can be made (Fig. 8*B*) by using the absolute values of the chemical shift differences. The

high field shifts of the C_α-proton resonances from residues in the leucine zipper are more pronounced than those of the other regions. This fact can be explained with the formation of a dimerization-stabilized helix. Particularly remarkable are the strong highfield shifts of the C_α-proton resonances from some hydrophobic residues at positions **a** and **d**.

No NOEs ($i, i + j, j > 8$) were observed, especially no NOEs between residues NH_2- and COOH-terminal to the leucine zipper domain, indicating that the leucine zipper domain forms a parallel coiled coil. The pattern and the intensities of sequential and medium range NOEs ($i, i + j, j \leq 4$) (Fig. 9) indicate α-helical structure of the peptide between residues Met[1] and Val[12]. In the leucine zipper domain only a very small number of unambiguous NOEs were found.

Calculation and Description of the Structure—The structure was determined by a combinatorial strategy using NMR data supplemented by modeling approaches for coiled coil proteins (40). The assumptions used in the modeling strategy were all confirmed by our previous studies. The sequence of pDIP contains the typical leucine heptade repeats, CD spectra were characteristic of α-helical structure, the C_αH resonances show high field shifts compared with the random coil values, and various results indicate pDIP to form dimers. Modeling of the leucine zipper domain was performed as described under "Materials and Methods." The correlation of structures with a positive starting twist angle is superior to that of structures with a negative starting twist angle (Fig. 10). No transitions between structures with positive starting twist angle and negative starting twist angle were found during the calculations, indicating that the energy barrier between these two types of structure is too high to be overcome by this modeling approach. Positive twist angles correspond to a left-handed conformation of the super helix, which is in agreement with proposals by Crick (8) and with known leucine zipper structures (47). Further structural analyses show that the hydrophobic and electrostatic interactions are stronger in the case of left-handed conformations (Fig. 11). Most of the leucine-leucine distances in these structures are in the range of 5–6 Å when measured as average distances of the C_γ. Identical values are found in the x-ray structure of GCN4 (47). Formation of salt bridges as defined by the distance between ionizable side chains being less than 4 Å (62) is indicated between Glu[34] and Lys[29] and between Glu[36] and Lys[41] (Fig. 2). In the modeled coiled coil structures, the leucine zippers corresponding to a left-handed conformation are stabilized by such intermonomer electrostatic interac-
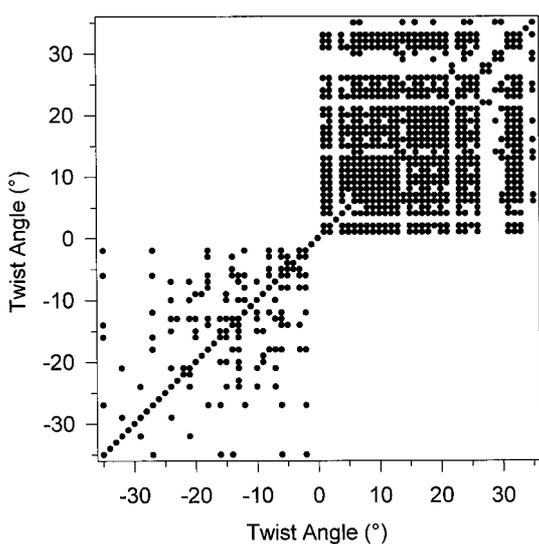
FIG. 10. **Correlation plot for the structures with different initial twist angles.** Negative twist angles correspond to a right-handed super helix; positive twist angles correspond to a left-handed super helix. *Dots* at position (i, j) indicate an RMSD value of less than 1.6 Å between structures i and j.
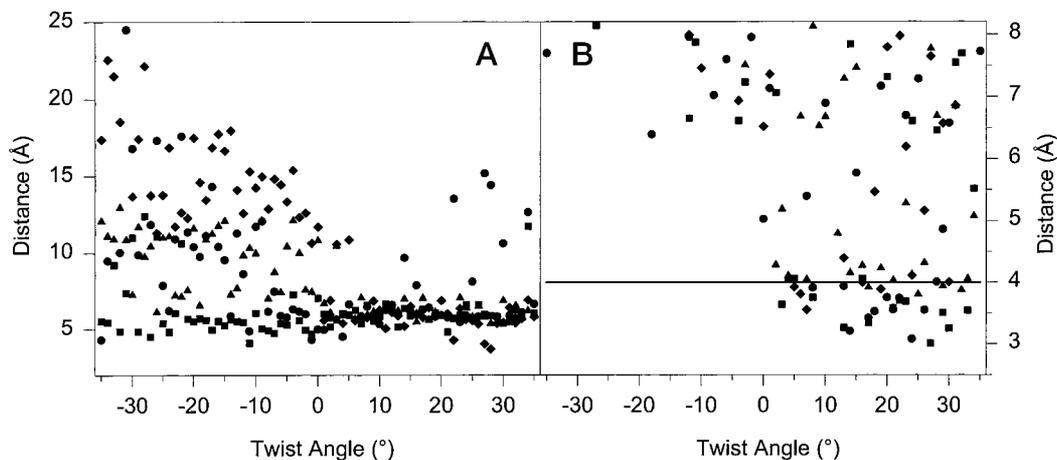
FIG. 11. **Hydrophobic and electrostatic interactions in the modeled leucine zipper.** *A*, distances between the C_γ-atoms of the leucines at position **d** and **d'** (●, Leu[19]-Leu[19]; ■, Leu[26]-Leu[26]; ▲, Leu[33]-Leu[33]; ◆, Leu[40]-Leu[40]) *versus* the helix twist angle. Stronger hydrophobic interactions are formed in structures with a positive starting twist angle. *B*, distances between the ionizable side chains of charged residues at positions **g** and **e** (●, Glu[34]-Lys[29]; ■, Lys[29]-Glu[34]; ▲, Glu[36]-Lys[41]; ◆, Lys[41]-Glu[36]). The *solid line* indicates distances between ionizable side chains of less than 4 Å (61). The formation of intermonomeric i to i' + 5 salt bridges is only possible for left-handed conformations of the super helix.

tions (Fig. 11*B*). The existence of salt bridges stabilizing the α-helical coiled coil is in agreement with the pH dependence of the thermal unfolding process.

The structure calculation of the full-length peptide was performed as described above under "Materials and Methods." The coordinates of a representative structure of the modeling approach were used as reference coordinates to restrain this part of the peptide harmonically to the known structure of the leucine zipper domain. For the residues NH_2- and COOH-terminal to the leucine zipper domain, only the NOEs determined from the 100 ms NOESY spectrum were used as experimental restraints, and no symmetry was imposed on the structure.

A schematic picture of the pDIP structure is shown in Fig. 12, and the energy and root mean square deviation (RMSD) values for 10 structures that resulted from this molecular dynamics calculation procedure are listed in Table II. The positive term for the NOE effective energy is caused by no more than two NOE violations larger than 0.05 nm per structure.

The leucine zipper region, Val^{16}–Ala^{44}, shows the lowest RMSD values due to the harmonic fitting. A second region with low flexibility was found in the NH_2-terminal part of the peptide. Structural analysis with the PROMOTIF program (63) shows this part to form a helical secondary structure, as was already indicated from analysis of the NOE pattern (Fig. 9) and
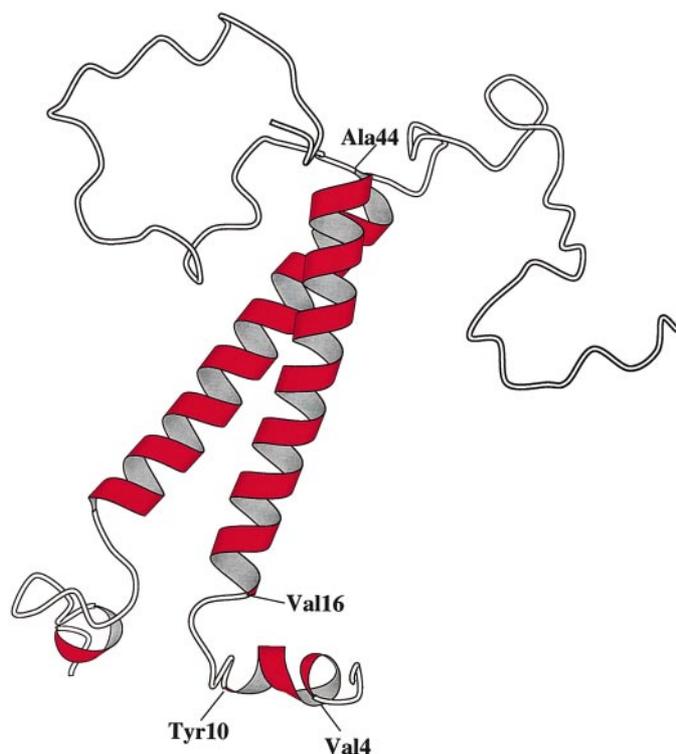
the chemical shift index (Fig. 8). The helix is extending from Val^4 to Tyr^{10} and shows 3_{10}-character in the region from Leu^8 to Tyr^{10} to some extent. The existence of a second helix, amino-terminal to the leucine zipper, is somewhat reminiscent of the arrangement of helices in bZIP proteins. In pDIP, however, this first helix is not basic and does not contain the consensus sequence found in most bZIP proteins, that is N*XX*(A)(A)*XX*(C/S)R (64). Within the bZIP protein family members with a dimerization domain but no intact basic region responsible for the DNA-binding such as CHOP (65) act by dimerization with proteins containing the basic region and inhibiting their DNA-binding activity. The formation of such heterodimers would be a possible mode of action of pDIP.

The COOH-terminal region of the peptide is very flexible and does not show stable elements of regular secondary structure (Fig. 12). Structured parts of pDIP show higher homology within the sequences of DIP, TSC-22 proteins from different mammals, and two *shs* variants (Fig. 13): The NH_2 terminus is highly conserved (80% identity). Residues responsible for the hydrophobic and electrostatic interactions between the two α-helices in the leucine zipper domain are 88% identical. Lys^{29}, Glu^{34}, Glu^{36}, and Lys^{41} (sequence position in pDIP), all involved in salt bridge formation, are conserved in all sequences and the hydrophobic amino acids at positions **a** and **d** show high homology. The evolutionary conservation of these residues underscores their functionality. It is known that the residues at positions **a** and **d** influence the oligomerization state of the

TABLE II
NOE and X-PLOR statistics

E_{total}, total energy; E_{VDW}, van der Waals energy; E_{NOE}, effective NOE energy term; E_{harm}, energy term for the harmonic fitting. The energy values result from the target function used for the MD calculations as described in the text. The values are mean values over 10 refined structures.

Parameter	Value
No. of NOEs	
Total	416
Sequential $|i - j| = 1$	244
Medium range $|i - j| = 2, 3, 4$	162
Long range $|i - j| > 4$	10
Average energy	
E_{tot}	4211 ± 166 kJ mol^{-1}
E_{bonds}	198 ± 19 kJ mol^{-1}
E_{angles}	2067 ± 78 kJ mol^{-1}
$E_{improper}$	412 ± 183 kJ mol^{-1}
E_{vdw}	617 ± 62 kJ mol^{-1}
E_{NOE}	608 ± 53 kJ mol^{-1}
E_{harm}	307 ± 35 kJ mol^{-1}
RMSD from idealized geometry	
Bonds	0.000 ± 0.000 nm
Angles	$0.85 \pm 0.02°$
Impropers	$0.72 \pm 0.16°$
NOE	0.008 ± 0.000 nm
RMSD among backbone structures	
Whole protein (dimer)	0.54 nm
Whole protein (monomer)	0.47 nm
Residues 2–15	0.08 nm
Residues 16–44	0.01 nm

FIG. 12. **MOLSCRIPT representation (69) of the three-dimensional structure of pDIP.**

```
              abcdefgabcdefgabcdefgabcdefg
DIP_PIG     MDLVKNHLMYAVREEVEILKEQIRELVEKNSQLERENTLLKTLASPEQLEKFQSRLSP-EEPAPETPEAPEAPGGSAV--------
DIP_HUMAN   MDLVKNHLMYAVREEVEILKEQIRELVEKNSQLERENTLLKTLASPEQLEKFQSCLSP-EEPAPESPQVPEAPGGSAV--------
CHKTSC      MDLVKSHLMYAVREEVEVLKEQIKELIEKNSQLEQENTLLKTLASPEQLAQFQAQLQTGSPPSSGQAQGTAQQPAQP---------
TSC2_MOUSE  MDLVKSHLMYAVREEVEVLKEQIKELIEKNSQLEQENNLLKTLASPEQLAQFQAQLQTGSPPATTQPQGTTQPPAQPASQGSGSTA
TSC2_RAT    MDLVKSHLMYAVREEVEVLKEQIKELIEKNSQLEQENNLLKTLASPEQLAQFQAQLQTGSPPATTQPQGTTQPPAQPA--------
DROSHS      MDLVKSHLMIAVREEVEVLKERISELMDKINKLELENSILKSNIPQETLQQLQLQLQLAAPPATPAIQAAPAVQSVVA--------
DROSHS_II   MDLVKNHLMYAVREEVEILKEQIRELVEKNNKLELENSILKSNIPQETLQQLQLQLQLAAPPATPAIQAAPAVQSVV---------
            ***** *** ******* ***.* **..*   .** ** .**.   * *. * * .    *.    .
```

FIG. 13. **Amino acid sequence alignment according to the ClustalW program using the standard set of parameters (70).** *Asterisks* below the sequences denote amino acid identity among all proteins. *Dots* mark conservative amino acid replacements. The heptade positions of the leucine-zipper are marked (*abcdefg*). The sequences shown are as follows: pDIP (*DIP_PIG*), hDIP (*DIP_HUMAN*), TSC-22 protein from chicken (*CHKTSC*), mouse (*TSC2_MOUSE*) and rat (*TSC2_RAT*) and two *shs* gene product sequences (*DROSHS* and *DROSHS_II*).
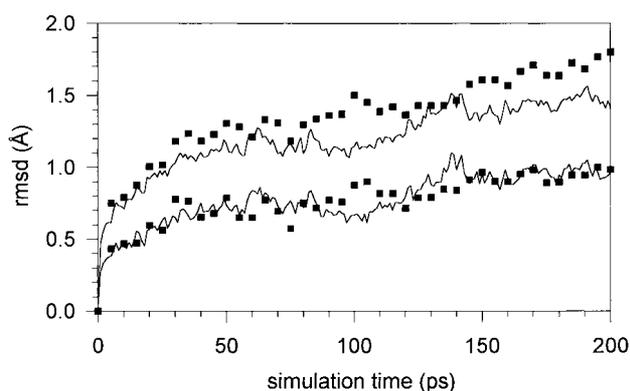
FIG. 14. **RMSD values of current structure *versus* starting structure during 200 ps of unrestrained molecular dynamics calculation.** *Lower trace*, backbone heavy atoms; *upper trace*, all heavy atoms of the leucine zipper of GCN4 (■) and pDIP (—).

coiled coil protein and the distribution of charged residues at positions **e** and **g** control the topology of the mutual helix arrangement (18, 66). Therefore, the leucine zipper domains of the other proteins in the sequence alignment (Fig. 13) also have an intrinsic disposition to form parallel two-stranded coiled coils.

The unstructured COOH-terminal region of pDIP, in contrast, shows only 20% identity within the aligned sequences. Although the proline-rich sequence of pDIP does not show homology to SH$_3$-binding domains (67) it is possible that this region is responsible to protein-protein interaction. Proline-rich polypeptides have highly restricted mobility even before binding. Thus, binding leads to a smaller drop in entropy than it would for a normal more flexible peptide, and hence a larger overall binding energy is achieved (68).

Unrestrained Molecular Dynamics—Unrestrained MD calculations of the leucine zipper of pDIP were performed to compare its structural stability to that of the well known GCN4 leucine zipper structure from the bZIP family. The RMSD values of the backbone and side chain heavy atoms of the structures in the unrestrained molecular dynamics calculation as compared with the starting structure were determined for both structures (Fig. 14). For both proteins, the RMSD values increase early in the calculation as expected due to an equilibration with the solvent. For the remaining time the RMSD values of the backbone atoms remain relatively stable, whereas the RMSD values of the side chain heavy atoms still increase. The RMSD values for the backbone atoms of pDIP and GCN4 leucine zippers are of the same order after the simulation time of 200 ps (0.96 and 0.98 Å, respectively), indicating virtually identical conformational stability of these leucine zippers. This indicates the modeled leucine zipper of pDIP to exhibit the same degree of flexibility as the leucine zippers of the bZIP family.

REFERENCES

1. Monnier, M., Dudler, L., Gächter, R., Maier, P. F., Tobler, H. J., and Schoenenberger, G. A. (1977) *Experientia* (*Basel*) **33,** 548–552
2. Sillard, R., Schulz-Knappe, P., Vogel, P., Raida, M., Bensch, W., Forssmann, W. G., and Mutt, V. (1993) *Eur. J. Biochem.* **216,** 429–436
3. Vinson, C. R., Sigler, P. B., and McKnight, S. L. (1989) *Science* **246,** 911–916
4. Busch, S. J., and Sassone-Corsi, P. (1990) *Trends Genet.* **6,** 36–40
5. Vogel, P., Mägert, H. W., Cieslak, A., Adermann, K., and Forssmann, W. G. (1996) *Biochim. Biophys. Acta* **1309,** 200–204
6. Shibanuma, M., Kuroki, T., and Nose, K. (1992) *J. Biol. Chem.* **267,** 10219–10224
7. Treisman, J. E., Lai, Z. C., and Rubin, G. M. (1995) *Development* **121,** 2835–2845
8. Crick, F. H. C. (1953) *Acta Crystallogr. Sec. D* **6,** 689–697
9. Talbot, J. A., and Hodges, R. S. (1982) *Acc. Chem. Res.* **15,** 224–230
10. Lau, S. Y. M., Taneja, A. K., and Hodges, R. S. (1984) *J. Biol. Chem.* **259,** 13253–13261
11. Hodges, R. S., Semchuk, P. D., Taneja, A. K., Kay, C. M., Parker, J. M. R., and Mant, C. T. (1988) *Pept. Res.* **1,** 19–30
12. Zhou, N. E., Zhu, B. Y., Kay, C. M., and Hodges, R. S. (1992) *Biopolymers* **32,** 419–426
13. Adamson, J. G., Zhou, N. E., and Hodges, R. S. (1993) *Curr. Biol.* **4,** 428–437
14. Cohen, C., and Parry, D. A. D. (1986) *Trends Biochem. Sci.* **11,** 245–248
15. Cohen, C., and Parry, D. A. D. (1990) *Proteins Struct. Funct. Genet.* **7,** 1–15
16. McLachlan, A. D., and Steward, M. (1975) *J. Mol. Biol.* **98,** 293–304
17. Stone, D., Sodek, J., Johnson, P., and Smillie, L. B. (1975) in *Proceedings of the IX Federation of European Biochemical Societies Meeting, Proteins of Contractile Systems* (Biro, E. N. A., ed), pp. 125–136, North Holland, Amsterdam
18. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) *Science* **262,** 1401–1407
19. Harbury, P. B., Kim, P. S., and Alber, T. (1994) *Nature* **371,** 80–83
20. Zhu, B. Y., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *Protein Sci.* **2,** 383–394
21. Fairman, R., Chao, H. G., Mueller, L., Lavoie, T. B., Shen, L., Novotny, J., and Matsueda, G. R. (1995) *Protein Sci.* **4,** 1457–1469
22. Junius, F. K., Mackay, J. P., Bubb, W. A., Jensen, S. A., Weiss, A. S., and King, G. F. (1995) *Biochemistry* **34,** 6164–6174
23. Muhle-Goll, C., Nilges, M., and Pastore, A. (1995) *Biochemistry* **34,** 13554–13564
24. Wendt, H., Berger, C., Baici, A., Thomas, R. M., and Bosshard, H. R. (1995) *Biochemistry* **34,** 4097–4107
25. Kohn, D. W., Kay, C. M., and Hodges, R. S. (1995) *Protein Sci.* **4,** 237–250
26. Nautiyal, S., Woolfson, D. N., King, D. S., and Alber, T. (1995) *Biochemistry* **34,** 11645–11651
27. Saudek, V., Pastore, A., Morelli, M. A. C., Frank, R., Gausepohl, H., Gibson, T., Weih, F., and Rösch, P. (1990) *Protein Eng.* **4,** 3–10
28. Ellenberger, T. E., Brandl, C. J., Struhl, K., and Harrison, S. C. (1992) *Cell* **71,** 1223–1237
29. Bowie, J. U., and Sauer, R. T. (1989) *Biochemistry* **28,** 7139–7143
30. Gibrat, J.-F., Garnier, J., and Robson, B. (1987) *J. Mol. Biol.* **198,** 425–443
31. Levin, J. M., Robson, B., and Garnier, J. (1986) *FEBS Lett.* **205,** 303–308
32. Deleage, G., and Roux, B. (1987) *Protein Eng.* **1,** 289–294
33. Geourjon, C., and Deleage, G. (1994) *Protein Eng.* **7,** 157–164
34. Rost, B., and Sander, C. (1994) *Proteins* **19,** 55–72
35. Mehta, P. K., Hering, J., and Argos, P. (1995) *Protein Sci.* **4,** 2517–2525
36. Frishman, D., and Argos, P. (1996) *Protein Eng.* **9,** 133–142
37. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **113,** 967–974
38. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York
39. Brünger, A. T. (1993) *XPLOR Version 3.1*, Yale University Press, New Haven, CT
40. DeLano, W. L., and Brünger, A. T. (1994) *Proteins* **20,** 105–123
41. Nilges, M., Clore, M., and Gronenborn, A. M. (1988) *FEBS Lett.* **239,** 129–136
42. Jørgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) *J. Chem. Phys.* **79,** 926–935
43. Powell, M. J. D. (1977) *Mathemat. Progr.* **12,** 241–254
44. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNiola, A., and Haak, J. R. (1984) *J. Chem. Phys.* **81,** 3684–3690
45. Verlet, L. (1967) *Physiol. Rev.* **159,** 98–103
46. van Gunsteren, W. F., and Berendsen, H. J. C. (1977) *J. Mol. Phys.* **34,** 1311–1327
47. O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) *Science* **254,** 539–544
48. Lumb, K. J., and Kim, P. S. (1995) *Biochemistry* **34,** 8642–8648
49. Zhou, N. E., Kay, C. M., Hodges, R. S. (1994) *Protein Eng.* **7,** 1365–1372
50. Krylov, D., Mikhailenko, I., Vinson, C. (1994) *EMBO J.* **13,** 2849–2861
51. Woolfson, D. N., and Alber, T. (1995) *Protein Sci.* **4,** 1596–1607
52. Yphantis, D. A. (1964) *Biochemistry* **3,** 297–317
53. Price, N. C. (1995) Circular dichroism in protein analysis. In: Meyers, R. A. (Ed.): *Molecular biology and biotechnology. A comprehensive desk reference*. VCH Publishers, Inc., New York
54. Lavigne, P., Kondejewski, L. H., Houston, M. E., Jr., Sönnichsen, F. D., Lix, B., Sykes, B. D., Hodges, R. S., and Kay, C. M. (1995) *J. Mol. Biol.* **254,** 505–520
55. De Francesco, R., Pastore, A., Vecchio, G., and Cortese, R. (1991) *Biochemistry* **30,** 143–147
56. Yu, H., Monera, O. D., Hodges, R. S., and Privalov, P. L. (1996) *J. Mol. Biol.* **255,** 367–372
57. Lumb, K. J., and Kim, P. S. (1995) *Science* **268,** 436–438
58. Jardetzky, O., and Roberts, G. C. K. (1981) *NMR in Molecular Biology*, Academic Press, New York
59. Englander, S. W., and Wand, A. J. (1987) *Biochemistry* **26,** 5953–5958
60. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry* **31,** 1647–1651
61. Hu, J. C., and Sauer, R. T. (1992) *Nucleic Acids Mol. Biol.* **6,** 82–101
62. Barlow, D. J., and Thornton, J. M. (1983) *J. Mol. Biol.* **168,** 867–885
63. Hutchinson, E. G., and Thornton, J. M. (1996) *Protein Sci.* **5,** 212–220
64. Hurst, H. (1994) *Protein Profile* **1,** 123–168
65. Ron, D., and Habener, J. F. (1992) *Genes Dev.* **6,** 439–453
66. Monera, O. D., Kay, C. M., and Hodges, R. S. (1994) *Biochemistry* **33,** 3862–3871
67. Chen, J. K., Lane, W. S., Brauer, A. W., Tanaka, A., and Schreiber, S. L. (1993) *J. Am. Chem. Soc.* **115,** 12591–12592
68. Williamson, M. P. (1994) *Biochem. J.* **297,** 249–260
69. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
70. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680