Stabilization of Coiled-Coil Peptide Domains by Introduction of Trifluoroleucine†

Yi Tang,‡ Giovanna Ghirlanda,§ Nagarajan Vaidehi,‡ Jeremy Kua,‡ Daniel T. Mainz,‡ William A. Goddard III,‡ William F. DeGrado,§ and David A. Tirrell*,‡

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, and The Johnson Research Foundation, Department of Biochemistry & Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received September 26, 2000; Revised Manuscript Received December 22, 2000

ABSTRACT: Substitution of leucine residues by 5,5,5-trifluoroleucine at the ***d***-positions of the leucine zipper peptide GCN4-p1d increases the thermal stability of the coiled-coil structure. The midpoint thermal unfolding temperature of the fluorinated peptide is elevated by 13 °C at 30 μM peptide concentration. The modified peptide is more resistant to chaotropic denaturants, and the free energy of folding of the fluorinated peptide is 0.5−1.2 kcal/mol larger than that of the hydrogenated form. A similarly fluorinated form of the DNA-binding peptide GCN4-bZip binds to target DNA sequences with affinity and specificity identical to those of the hydrogenated form, while demonstrating enhanced thermal stability. Molecular dynamics simulation on the fluorinated GCN4-p1d peptide using the Surface Generalized Born implicit solvation model revealed that the coiled-coil binding energy is 55% more favorable upon fluorination. These results suggest that fluorination of hydrophobic substructures in peptides and proteins may provide new means of increasing protein stability, enhancing protein assembly, and strengthening receptor−ligand interactions.

Engineering of stable enzymes and robust therapeutic proteins is of central importance to the biotechnology and pharmaceutical industries. Although protein engineering provides powerful tools for the enhancement of enzymatic activity and protein stability (*1*−*4*), the scope of in vivo engineering methods is limited by the availability of just 20 naturally occurring proteinogenic amino acids (*5*). Increasing success in the incorporation of noncanonical amino acids into recombinant proteins in vivo has allowed the introduction of novel side-chain functionality into engineered proteins (*6*−*10*) and raises prospects of new approaches to the design of peptides and proteins of enhanced activity and/or stability.

Leucine zipper peptides are ideal models for the study of protein secondary and tertiary interactions (*11*−*20*, and references therein). Such peptides assemble into coiled-coil dimers, trimers, and tetramers in order to exclude solvent at the hydrophobic interfaces between adjacent peptide helices, and the relations between sequence and stability have been carefully examined (*21*). In particular, the structure (*13*), dimerization kinetics (*20*), and thermodynamics (*19*) of the model peptide GCN4-p1 have been thoroughly described.

GCN4-p1 constitutes the dimerization domain of bZip, which is a 56-amino acid DNA binding segment (residues 226−281) of the eukaryotic transcription factor GCN4. The N-terminus of bZip contains a DNA recognition domain rich in the basic residues lysine and arginine. The C-terminal subdomain of bZip contains the GCN4-p1 peptide segment and facilitates dimerization of the protein. While direct contact between DNA and the N-terminal subdomain is important to recognition, protein−protein interactions at the C-terminus also contribute to the specificity and affinity of peptide−DNA binding (*22*−*24*).

A

R MKQ***L***EDK VEE***L***LSK NYH***L***ENE VAR***L***KKL VGER

B

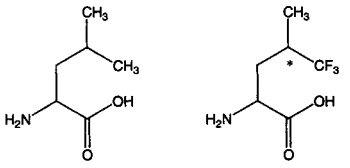

FIGURE 1: (A) Amino acid sequence of GCN4-p1d. The amino acids in the leucine-zipper motif are represented by the heptad *abc**d**efg*. The leucine residues that were replaced by trifluoroleucine are at the ***d***-positions in the heptad repeat (shown in bold). The peptides prepared in this work were *not* acylated at the amino terminus. Amino acid abbreviations: G, glycine; A, alanine; V, valine; L, leucine; I, isoleucine; S, serine; T, threonine; P, proline; D, aspartic acid; N, asparagine; E, glutamic acid; Q, glutamine; K, lysine; R, arginine; C, cystein; M, methionine; H, histidine; W, tryptophan; F, phenylalanine; Y, tyrosine. (B) Comparison between trifluoroleucine (right) and leucine. Trifluoroleucine was used as a mixture of the (2S, 4S) and (2S, 4R) isomers.

We present here a successful attempt to stabilize the coiled-coil forms of both GCN4-p1 (Figure 1A) and GCN4-bZip by substitution of the core leucine residues of the peptides with 5,5,5-trifluoroleucine (Tfl)[1] (Figure 1B). Furthermore,

† This work was supported by a grant from the U.S. Army Research Office to D.A.T. and by NIH Grant GM54616 to W.F.D. Y. Tang is grateful for a graduate research fellowship from the Whitaker Foundation.

* To whom correspondence should be addressed.
‡ California Institute of Technology.
§ University of Pennsylvania.

10.1021/bi0022588 CCC: $20.00
Published on Web 02/07/2001

the DNA binding behavior of the fluorinated form of GCN4-bZip is shown to be identical to that of the wild-type peptide.

The choice of Tfl (Figure 1B) for these studies was based on three factors. First, we imagined that Tfl might behave as a hyper-hydrophobic analogue of leucine, since many organofluorine compounds exhibit lower solubility in water than their hydrocarbon equivalents (*25*, *26*). Second, fluorinated amino acids are nearly isosteric to their natural counterparts, and are equally (or more) inert chemically (*27*). Third, Tfl is accepted as a substrate by the endogenous leucyl-tRNA synthetase (LeuRS) of microbial cells, and is readily incorporated into recombinant proteins *in vivo* (*28*, *29*). When leucine-depleted cultures of *Escherichia coli* are shifted to Tfl-supplemented medium, more than 95% of the leucine positions in a recombinant protein can be substituted by Tfl under appropriate conditions (*30*).

We also discuss in this paper the results of molecular dynamics simulation of the thermodynamic properties of the fluorinated leucine zipper peptide. Simulation with an implicit solvation model confirmed the experimental finding that the Tfl-substituted coiled-coils are indeed more stable than their wild-type counterparts.

MATERIALS AND METHODS

Trifluoroleucine Synthesis. DL-Trifluoroleucine was prepared by a modification of the procedures outlined in Ref *28*. The *N*-acetylated racemic mixture was resolved to L-trifluoroleucine by treatment with porcine kidney acylase (Sigma) to >99% enantiomeric excess (e.e.). The determination of e.e. was accomplished by ^{1}H NMR spectroscopy following derivatization with Mosher's acid, *R*-(+)-methoxytrifluoromethylphenylacetic acid.

Peptide Synthesis and Purification. Peptides were synthesized at the Biopolymer Synthesis Center at the California Institute of Technology (Pasadena, CA 91125). Automated, stepwise solid-phase synthesis was performed on an ABI 433A synthesizer employing Fmoc chemistry. The trifluoroleucine units were incorporated into the peptide with extended coupling cycles. After chain assembly was complete, the peptide was deprotected and removed from the resin support with trifluoroacetic acid in the presence of 1,2-ethanedithiol, thioanisole, and water. Peptides were precipitated into cold methyl *tert*-butyl ether and isolated by centrifugation. Peptide products were purified by preparative C_{18} reversed-phase HPLC using a nonlinear gradient of 0 to 80% elution solution (0.1% TFA/60% acetonitrile/40% H_2O) in 120 min. The peptides examined in this work were *not* acylated at their N-termini, and are designated GCN4-p1d to distinguish them from the acylated variants reported in earlier papers (e.g., ref *11*). The thermal melting temperature of Leu-GCN4-p1d is slightly lower than that reported for the acylated form GCN4-p1 (*11*).

Ultracentrifugation Analysis. Sedimentation equilibrium analysis was performed using a Beckman XLI analytical ultracentrifuge, recording interference data and radial absorbance at 236 and 280 nm at the same time. Initial peptide concentrations ranged between 100 and 300 μM; buffer was 0.01 M sodium phosphate, pH 7.4, containing 0.1 M NaCl. The samples were centrifuged at 35 000, 40 000, 45 000 rpm, until equilibrium was reached. Partial specific volumes were calculated by the residue-weighted average method of Cohn and Edsall (*31*). Solution densities were estimated by using solute concentration-dependent density tables in the *CRC Handbook of Chemistry and Physics*. The data were fit as single species to provide an estimate of the aggregation states. Curve fitting of analytical ultracentrifuge data was done using Igor Pro (Wavemetrics Inc., Oswego WS) with procedures adapted from ref *32* by Dr. James D. Lear.

Spectroscopic Analysis. CD spectra were recorded on an Aviv 62DS spectropolarimeter (Lakewood, NJ) in PBS buffer, pH 7.0. All peptide concentrations were determined by amino acid analysis of a stock solution (2 mg/mL). Experiments were performed in a rectangular cell with path length of 1 mm. Spectra were scanned from 260 to 194 nm with points taken every 1 nm. The temperature of the solution was maintained by a thermostatically controlled cuvette holder (HP model 89101A). Temperature scans were performed from 0 to 100 °C in 1 °C steps. Five scans were performed on a single sample and averaged. Each data point was collected after 30 s of thermal equilibration at the desired temperature.

The thermodynamic data were calculated by fitting the thermal denaturation curves to a monomer−dimer equilibrium according to published procedures (*16*). The data were fit using a standard state of 1 M. Mean residue ellipticity at 222 nm was fit as a function of temperature and total peptide concentration using MLAB software as stated in the reference. The data in the text were obtained from global fits of the thermal denaturation data at peptide concentrations of 3 and 85 μM. All thermodynamic quantities reported in the text are per mole of monomers.

Gel Retardation Assays. Oligonucleotides containing the AP-1 binding site (AP-1, 5′-GTGGAGATGACTCATCTC-CGG-3′), the CREB binding site (CREB, 5′-TGGAGAT-GACGTCATCTCCT-3′) and the nonspecific sequence (NON, 5′-GATCCCAACACGTGTTGGGATC-3′) were synthesized at the Caltech DNA synthesis facility. The oligonucleotides were labeled with γ-[^{32}P]ATP (>6000 Ci/mmol), annealed, and purified. Labeled probes (5000 cpm) were incubated with protein for 30 min at 4 °C in 40 μL of binding buffer [20 mM Tris-HCl, pH 7.0, 50 mM NaCl, 1 mM DTT, 1 mM EDTA, 10% glycerol and 40 μg/mL poly(dl-dC)·poly(dl-dC)]. An aliquot of the reaction volume (10 μL) was analyzed by electrophoresis on 5% polyacrylamide gel. Free and protein-bound DNA were visualized by autoradiography. The relative intensities of the bands were measured by densitometry.

Molecular Dynamics Simulation. Molecular dynamics (MD) simulation of Leu-GCN4-p1d and Tfl-GCN4-p1d was performed using the MPSim program (*33*). The MPSim MD program, which includes the Cell Multipole Method (*34*), provides fast and accurate calculations of nonbond interactions. MPSim describes the energies and forces due to polarization of the continuum solvent using either the Poisson−Boltzmann (PB) (*35*) or Surface Generalized Born (SGB) (*36*) models. For the dynamics, we used the SGB approximation to the PB continuum solvent model to calculate the forces on the Leu-GCN4-p1d dimer due to

[1] Abbreviations: CD, circular dichroism; Tfl, 5,5,5-trifluoroleucine; Hfl, 5,5,5,5′5′,5′-hexafluoroleucine; GuHCl, guanidinium hydrochloride; LeuRS, leucyl-tRNA synthetase; MD, molecular dynamics; PB, Poisson Boltzmann; SGB, Surface Generalized Born; FF, force field.

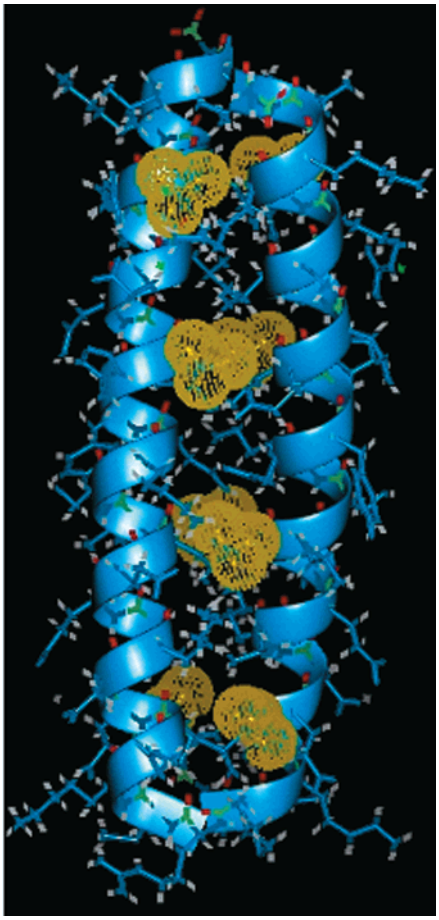

FIGURE 2: Three-dimensional representation of the dimeric form of GCN4-p1d substituted with trifluoroleucine at the four ***d***-positions in the helix. The van der Waals radii of the fluorine atoms are shown as yellow spheres.

solvent polarization; the final energies were then calculated using the PB method, which leads to more accurate energies.

The DREIDING force field (FF) (*37*) was used in all calculations. The recommended van der Waals potentials in DREIDING (having the exponential-six form) are used here. The default in the MSI software (PolyGraf and Cerius^2) is Lennard-Jones 12-6 which we have found to be significantly less accurate. Also we use a dielectric constant of $\epsilon = 1$ (not the distance-scaled ϵ of PolyGraf and Cerius^2). The charges on Tfl-GCN4-p1d and Hfl-GCN4-p1d were derived from first principles quantum mechanics (QM): LMP2 method using the 6-31G** basis set [Jaguar 3.5 (*38*) from Schrodinger Inc., Portland, OR]. These charges were calculated for the tripeptide, Gly-Leu (nF)-Gly, where nF = TfL or HfL. The charges for the natural residues in the protein were taken from the CHARMM FF (*39*).

RESULTS AND DISCUSSION

We prepared four peptides to study the effects of fluorination on coiled-coil stability and DNA binding affinity and specificity. We prepared the wild-type and fluorinated forms for both the GCN4-p1d (Leu-GCN4-p1d and Tfl-GCN4-p1d) and GCN4-bZip (Leu-bZip and Tfl-bZip) peptides. Solid-phase synthesis was used in this initial study to replace quantitatively all four ***d***-position leucines in GCN4-p1d (Figure 2). For Tfl-bZip, the corresponding ***d***-position

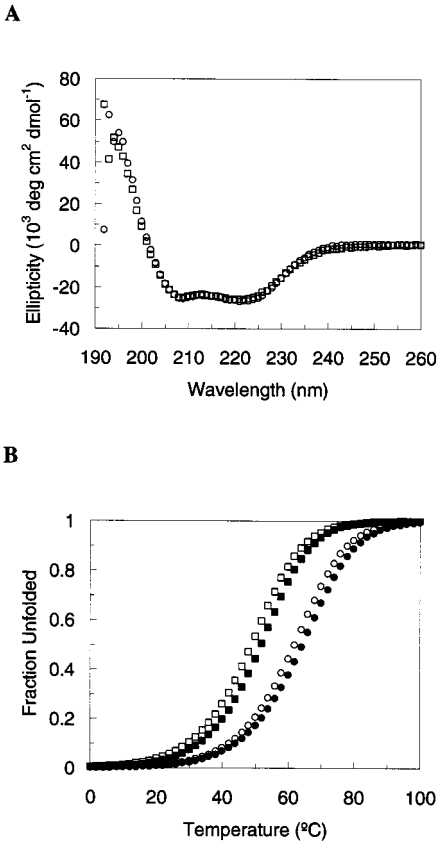

FIGURE 3: (A) CD spectra of Leu-GCN4-p1d (□) and Tfl-GCN4-p1d (○) at 0 °C and 30 μM. The nearly identical spectra and the magnitude of the intensity at 222 nm indicate that the peptides have similar secondary structures and are highly helical. (B) Thermal unfolding profiles for Leu-GCN4-p1d (squares) and Tfl-GCN4-p1d (circles) at 30 μM (open symbols) and 85 μM (closed symbols). The thermal melting temperature, T_m, is defined as the temperature at which 50% of the peptide is unfolded. At 30 μM, the values of T_m for Leu-GCN4-p1d and Tfl-GCN4-p1d are 49 and 62 °C, respectively. The thermal melting profiles were fitted globally to yield the thermodynamic quantities (see text and ref *20*).

leucines in the GCN4-p1d dimerization domain were substituted with Tfl. *N*-Fmoc-5,5,5-trifluoro-L-leucine was used as an equimolar mixture of the 2S,4S- and the 2S,4R− isomers to prepare the fluorinated peptides. None of the four peptides was acylated at the *N*-terminus. The peptides were purified by HPLC, and the molar masses of the purified peptides were verified by mass spectrometry.

Spectroscopic Characterization of Leu-GCN4-p1d and Tfl-GCN4-p1d. Circular dichroism (CD) spectra of Leu-GCN4-p1d and Tfl-GCN4-p1d indicated that both peptides are highly helical as evidenced by intense minima at 222 and 208 nm (Figure 3A). The spectra of the wild-type and fluorinated peptides are essentially coincident, suggesting nearly identical secondary structures; both peptides are highly helical at 0 °C. This observation confirmed our conjecture that replacement of leucine with Tfl would not disrupt interhelical packing and interfere with folding of the coiled-coil structure. Ultracentrifugation indicates that Tfl-GCN4-p1d is predominantly dimeric at the concentrations of interest in this work. Data for Tfl-GCN4-p1d were fit to a monomer−dimer−trimer equilibrium, giving values of K_d of the order of 10^{-8} M and 10^{-14} M^2, respectively, for the monomer-to-dimer and monomer-to-trimer equilibria. In the concentration

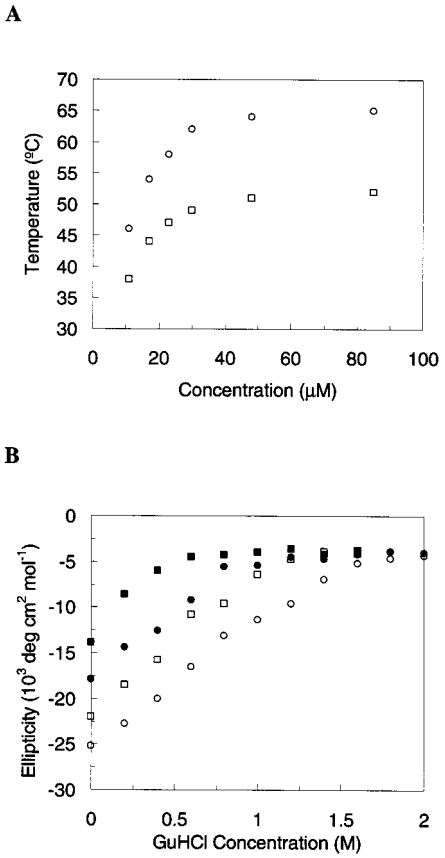

FIGURE 4: (A) Concentration dependence of the thermal melting temperatures of Leu-GCN4-p1d (□) and Tfl-GCN4-p1d (○). The decrease in T_m at decreasing concentration is evidence of self-association. (B) Guanidinium hydrochloride (GuHCl) titration of Leu-GCN4-p1d (squares) and Tfl-GCN4-p1d (circles). Data were recorded at 30 °C (open symbols) and 50 °C (closed symbols). The ellipticity was monitored at 222 nm at a peptide concentration of 30 μM. At each temperature, the fluorinated peptide is less susceptible to GuHCl denaturation than the hydrogenated form.

range of approximately 10−40 μM the peptide is ca. 85% dimeric.

The thermal stabilities of the coiled-coil dimers of Leu-GCN4-p1d and Tfl-GCN4-p1d were examined by CD spectroscopy (Figure 3B). A significant increase in the thermal stability of Tfl-GCN4-p1d as compared to Leu-GCN4-p1d is reflected in an elevation of the thermal denaturation temperature from 48 to 61 °C at a peptide concentration of 30 μM. The 13 °C increment in T_m is remarkable in view of the fact that no increase in the thermal stability of GCN4-p1d has been reported based solely on substitution of the leucine residues at the ***d***-positions. Mutations at the ***d***-positions to other natural amino acids have all resulted in losses in helix stability due to decreases in packing efficiency since such changes are usually of the "large to small" type (*21*). As expected for a monomer−dimer equilibrium, the denaturation curves depend on the peptide concentrations, and their midpoints shift to higher temperature as the concentrations of peptides are increased (Figure 4A). The thermodynamic changes associated with the transition (folded dimer to unfolded monomers) can be calculated from the melting curves by fitting the data to a monomer−dimer equilibrium. The thermodynamic analysis of Tfl-GCN4-p1d is complicated by its heterogeneity as a result of the presence of different Tfl stereoisomers. A more accurate thermodynamic analysis of the Tfl-GCN4-p1d can be obtained when pure stereoisomers of Tfl become available. (During HPLC purification, an additional peptide fraction of the same molecular weight as Tfl-GCN4-p1d was recovered. This fraction has a thermal denaturation temperature slightly lower than that of Tfl-GCN4-p1d. We attribute this result to stereochemical heterogeneity at the γ-position of L-Tfl, consistent with the prediction of the molecular dynamics simulation reported here.)

Global analysis of the thermal unfolding curves at two different concentrations (approximately 85 μM and 3 μM) yielded $\Delta H° = 60.2 \pm 1.1$ kcal mol^{-1}, $T_m = 385.4 \pm 0.4$ K and $\Delta C_p = 530 \pm 40$ cal mol^{-1} K^{-1} for Tfl-GCN4-p1d (1 M standard state); the corresponding values for Leu-GCN4-p1d are 70.3 ± 1.4 kcal mol^{-1}, 365.6 ± 0.4 K and 740 ± 50 cal mol^{-1} K^{-1}, respectively (all thermodynamic data reported are for per mole monomer). Under all conditions, where a direct experimental comparison was possible, Tfl-GCN4-p1d is 0.5−1.2 kcal mol^{-1} more stable than Leu-GCN4-p1d; for example, at 50 °C, K_d for dimerization is 67.8 μM for Leu-GCN4-p1d and 9.8 μM for Tfl-GCN4-p1d.

The stability of Tfl-GCN4-p1d toward denaturation by chaotropic reagents was demonstrated through guanidine hydrochloride (GuHCl) titration experiments (Figure 4B). At each temperature examined, the fluorinated peptide displayed significantly lower susceptibility toward denaturation by GuHCl; in each case, the concentration of GuHCl needed to unfold 50% of the peptide was higher for Tfl-GCN4-p1d than for the wild-type peptide. The free energies of folding at 25 °C can be obtained by globally fitting the GuHCl denaturation curves at two different concentrations for each peptide to a monomer−dimer equilibrium, resulting in extrapolated $\Delta G°$ (in absence of GuHCl) of 8.0 ± 0.1 kcal mol^{-1} for Leu-GCN4-p1d and 8.6 ± 0.1 kcal mol^{-1} for Tfl-GCN4-p1d, corresponding to values of K_d of 1.2 and 0.51 μM, respectively.

DNA Binding Studies. To determine whether protein function can be retained upon fluorination, we compared the affinities and specificities of DNA binding by Leu-bZip and Tfl-bZip. The secondary structures of the two proteins were identical as indicated by CD spectroscopy and the thermal melting temperature of Tfl-bZip was elevated by 8 °C compared to Leu-bZip at 10 μM peptide concentration (data not shown).

The DNA-binding domain of Leu-bZip is conformationally disordered in the absence of specific DNA sequences, while the dimerization domain forms a two-stranded coiled-coil through the leucine zipper motif at concentrations above the monomer-to-dimer equilibrium (*40*). Upon presentation of target DNA sequences, the DNA-binding region folds into a α-helical structure and the peptide binds to the DNA in a "chopstick" model (*41*−*43*). CD analysis of Tfl-bZip revealed that the fluorinated peptide behaves in the same manner as Leu-bZip (Figure 5). Before addition of oligonucleotides containing the target (CREB) binding site, Tfl-bZip is approximately 70% helical, as expected on the basis of the length of its helical dimerization domain. After addition of target DNA, Tfl-bZip is essentially 100% helical, indicating a transition of the DNA-binding region from coil to helix. The similar changes in secondary structure observed for Leu-bZip and Tfl-bZip suggest that fluorination of the

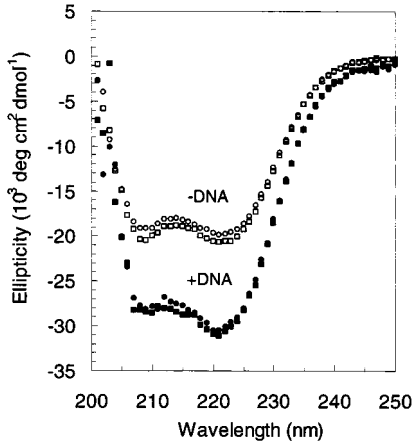

FIGURE 5: CD spectra for Leu-bZip (squares) and Tfl-bZip (circles) with (closed symbols) and without (open symbols) CREB at 0 °C. The peptide helicity, as estimated from the molar ellipticity at 222 nm, is increased from approximately 70% (∼−20 000 10^3 deg cm^2 dmol^{-1}) to nearly 90% (∼−31 000 deg cm^2 dmol^{-1}). The DNA concentration is 5 μM and the protein concentration is 10 μM (PBS buffer, pH 7.4).

zipper domain does not cause qualitative changes in the nature of the association between the peptide and DNA.

The affinity and specificity of binding of the fluorinated peptide were shown to be essentially identical to those of the wild-type peptide on the basis of gel-retardation assays (*44*, *45*) (Figure 6). Leu-bZip binds to the AP-1 and CREB binding sites with nearly equal affinities even though the spacing between the half-sites of these DNAs is different. Densitometric analysis of mobility shift assays revealed that Tfl-bZip binds to both sequences with specificities and affinities (K_d =12.5 ± 0.7 nM for AP-1 and 5.4 ± 0.6 nM for CREB) nearly identical to those of Leu-bZip (K_d = 12.8 ± 1 nM for AP-1 and 4.8 ± 0.5 nM) for CREB. Neither Leu-bZip nor Tfl-bZip recognizes nonspecific (NON) sequences, as shown by the lack of detectable peptide-bound DNA in assays with NON oligonucleotides (Figure 6, bottom).

Molecular Dynamics Simulation. To determine the origins of the stabilizing effect of side-chain fluorination, we carried out molecular dynamics (MD) calculations on Leu-GCN4-p1d and Tfl-GCN4-p1d using the Poisson−Boltzmann (PB) continuum description of the solvent. The PB description of solvation implicitly includes entropic changes in the solvent; thus the calculations lead directly to binding free energies (ΔG^{BE}). The MPSIM MD program and the DREIDING Force Field (FF) (*37*) were used for all calculations. The starting structure for the Leu-GCN4-p1d dimer was that of GCN4-p1 as reported in the RCSB Protein Data Bank; those of the fluorinated dimers were derived from the native dimer structure by replacement of the appropriate methyl hydrogens with fluorines, followed by reoptimization of the structure. Because the γ-carbon of Tfl is asymmetric (Figure 1B), multiple arrangements of adjacent diastereotopic trifluoromethyl groups must be considered (Figure 7). When both Tfl residues at a given ***d***-position are of the (2S,4S) configuration, the two trifluoromethyl groups are relatively close to one another; the fluorinated carbon centers are separated by ca. 6 Å. On the other hand, when two (2S,4R) isomers are juxtaposed, the corresponding carbon−carbon distance increases to about 8 Å. In the remaining configurations (where the two strands carry different isomers), the trifluoromethyl groups are separated by intermediate distances. We performed simulations on all configurations to determine how side-chain stereochemistry affects dimer stability. For simulation of strands containing different stereoisomers of trifluoroleucine, we considered only those cases in which all four trifluoroleucines on one strand have the same stereoconfiguration. For each dimer, we carried out 1 ns of constant temperature (300 K) Nose-Hoover MD with the SGB description of the water solvent. We used counterions for neutralizing the charges on the side chains of Arg, Lys, Asp, and Glu. For the native dimer, the 1 ns MD simulation leads to a structure in good agreement with the experimental structure of GCN4-p1, with an RMS deviation in coordinates of the main-chain atoms of 2.15 Å.

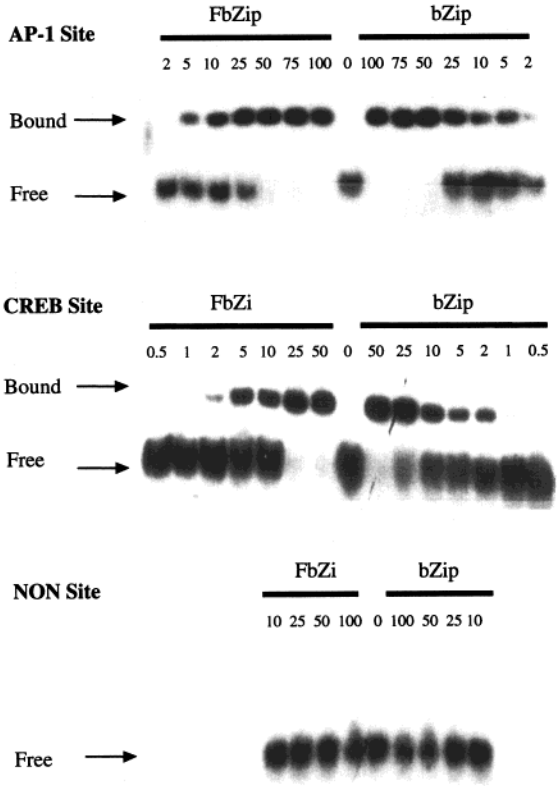

FIGURE 6: Mobility shift assay of Leu-bZip and Tfl-bZip binding to oligonucleotides containing the AP-1 binding site (5′-GTGGAGATGACTCATCTCCGG-3′, top), the CREB binding site (5′-TGGAGATGACGTCATCTCCT-3′, middle) and the nonspecific sequence (NON, 5′-GATCCCAACACGTGTTGGGATC-3′, bottom). The protein concentrations (nM) are indicated above each lane. Similar binding affinities were observed for Leu-bZip and Tfl-bZip. No binding to the nonspecific sequence could be detected.

From the 1 ns trajectory, we calculated the average properties over 800 ps after equilibration. ΔG^{BE} was calculated as the difference in energies of the solvated dimer and the corresponding solvated monomers. The monomers were described with short (50 ps) SGB MD simulations starting with the random coil structure. The dynamics structures were minimized and an average energy of the minimized structures was considered for the calculation (ΔG^{BE}). For the final (ΔG^{BE}), solvation was calculated using PB (which is more accurate than SGB). Table 1 reports the average values of ΔG^{BE} (per monomer) for the native and fluorinated forms of GCN4-p1d.

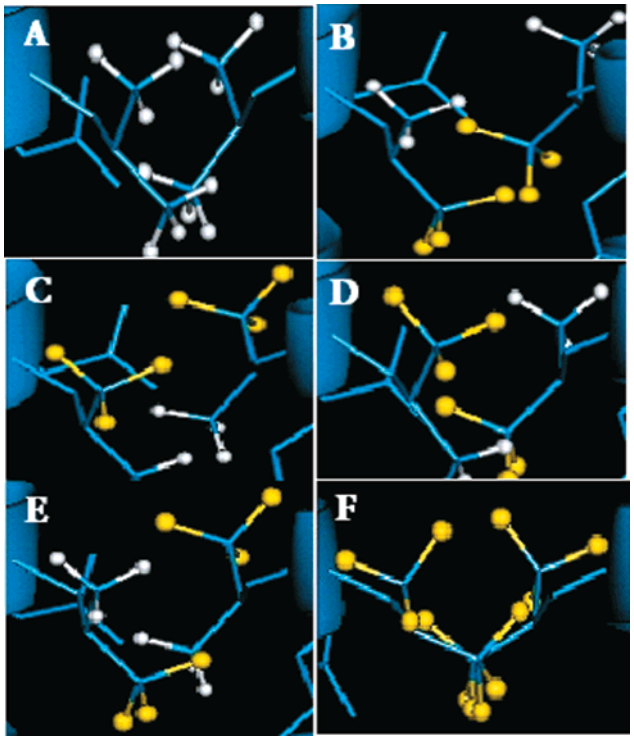

FIGURE 7: Possible configurations of the leucine/trifluoroleucine packing at each ***d***-position in the heptad. Shown in the figure are residues Leu_{22} from both stands. Several combinations of the 4R- and 4S-isomers of trifluoroleucine are possible due to the side-chain stereocenter. Shown here are the various packing possibilities considered in the molecular dynamics simulations. **A**, Leu/Leu; **B**, 4S/4S, "close"; **C**, 4R/4R, "far"; **D**, 4R/4S; **E**, 4S/4R; and **F**, Hfl/Hfl, where both leucines are replaced by hexafluoroleucine.

Table 1: Binding Free Energies (ΔG^{BE}, kcal/mol) of Leu-GCN4-p1d and Fluorinated Peptides[a]

structure	ΔG^{BE}	% increase	% helicity[c]
Leu-GCN4-p1d	65.08	0	90.8
Close (4S, 4S)[b]	93.75	44	84.3
Far (4R, 4R)	98.14	51	79.4
Mixed (4S, 4R)	99.20	52	81.1
Mixed (4R, 4S)	111.15	71	89.3
Tfl-average	100.56	55	83.5
Hfl-GCN4-p1d	77.21	19	78.5

[a] ΔG^{BE} is the difference in energy (averaged over 800 ps of MD after equilibration) of solvated monomers and the solvated dimer each from separate SGB MD calculations (final solvation energies with PB). ΔG^{BE} is quoted per mole of the monomer. Percent increase is the increase in ΔG^{BE} compared to the Leu-GCN4-p1d structure. Also shown is the % helicity of each peptide. [b] Close, Far, Mixed: configuration of the pair of trifluoromethyl groups as illustrated in Figure 7. Tfl-average: the averaged ΔG^{BE} of the four configurations. [c] Helicity quoted here has been calculated as the ratio of the residues with torsion angles ϕ and ψ in the helical region of the Ramachandran plot to the total number of residues in the protein.

The Tfl-GCN4-p1d dimers are predicted to exhibit ΔG^{BE} ca. 55% larger than that of the Leu-GCN4-p1d (calculated relative to the respective random coil monomers). The various stereochemical arrangements lead to increases in binding energies ranging from 44% to 71%, indicating that side-chain configuration may have some differential effect on dimer stability. Similar calculations for the hexafluoroleucine (Hfl) dimer lead to the prediction that such dimers (which were not prepared experimentally in this work) should be significantly less stable than the Tfl dimers but marginally more stable (19%) than the wild-type. To investigate the source of stability of the fluorinated dimers we analyzed the components of the binding energy for each peptide (Table 2). We find that the primary driving forces for stabilizing the Tfl-GCN4-p1d dimers arise from van der Waals (vdW) and hydrogen-bonding interactions. The predicted structures of the Hfl and wild-type monomers are globular, while the Tfl-GCN4-p1d monomer is more extended, with local "kinks" arising from favorable electrostatic interactions between Tfl residues (e.g., between Tfl_5 and Tfl_{12}). The globular structures of Leu-GCN4-p1d and Hfl-GCN4-p1d form more nonlocal hydrogen bonds and more favorable vdW contacts than the more extended Tfl-GCN4-p1d monomer. Hence, the gain in H-bond and vdW energies in forming a dimer is greater for Tfl-GCN4-p1d than for the wild-type or Hfl peptides because the latter monomers must pay the energy cost of extension as a prerequisite for dimerization.

Table 2: Components of ΔG^{BE} (kcal/mol) for Leu-GCN4-p1d and Fluorinated Dimers (quoted for 1 mol of the monomer)

structure	$\Delta G^{valence}$	$\Delta G^{coulomb+solvation}$	ΔG^{vdW}	ΔG^{Hbond}
Leu-GCN4-p1d	−16.12	−16.66	41.82	56.05
close (4S, 4S)[a]	−8.46	−16.64	59.56	59.29
far (4R, 4R)	−9.54	+1.73	42.16	63.80
mixed1 (4S, 4R)	−5.36	−10.55	65.17	49.94
mixed2 (4R, 4S)	−27.96	−1.79	48.29	92.61
Tfl-average	−12.83	−6.82	53.80	66.41
Hfl-GCN4-p1d	−23.06	7.24	36.51	56.19

[a] Close, Far, Mixed: configuration of the pair of trifluoromethyl groups as illustrated in Figure 4. Tfl-average: the averaged ΔG of the four configurations.

Furthermore, consideration of electrostatic (intra- and interpeptide coulomb forces) and solvation interactions suggests that the hydrophobic preference in the dimer for burial of CF_3 is greater than for CH_3. Considering just coulomb and solvation interactions, the driving force for dimerization is predicted to decrease in the order Hfl > Tfl > Leu. It is the balance of desolvation, electrostatics, H-bonding, and vdW forces that leads to the prediction that the Tfl dimers are more stable than the Hfl dimer which in turn is more stable than the native leucine dimer. The average helicity of the dimers is predicted to be 90.8% for Leu-GCN4-p1d, 83.5% for Tfl-GCN4-p1d, and 78.5% for Hfl-GCN4-p1d.

CONCLUSION

The experimental and computational results described here demonstrate that the subtle change from four leucine methyl groups to four trifluoromethyl groups results in a substantial gain in stability of the folded structure of a dimeric coiled-coil peptide. It is remarkable that for a peptide of the size of GCN4-p1d, fluorination results in a coiled-coil structure that is highly resistant to both thermal and denaturant unfolding as compared to the wild-type peptide. Although these studies used solid-phase peptide synthesis to prepare GCN4-p1d, we have also demonstrated that fluorinated peptides produced in vivo exhibit substantial elevation in stability compared to their wild-type analogues (*30*).

Given the ease with which trifluoroleucine can be incorporated in vivo, a wide range of proteins can be prepared in fluorinated form as a means to improve stability. The method may prove to be quite general, as leucine is the most

abundant of the amino acids in cellular proteins (9%) (*46*) and is especially important in determining the structure and stability of hydrophobic protein subdomains. A potential advantage of using fluorination to stabilize proteins is that it is complementary to other existing methods of protein stabilization: Fluorination might therefore serve as a "final push" toward higher stability after other methods, such as rational design (*47*) and directed evolution (*48*), have achieved some initial stabilization. Peptides that rely on hydrophobic side chains to form channels in membranes may also exhibit increased membrane (or interpeptide) association upon fluorination (*49*). In vivo methods for incorporation of Tfl should allow fluorination of enzymes, signaling molecules, protein ligands, etc., and may prove to be of broad utility in the engineering of robust macromolecular assemblies.

ACKNOWLEDGMENT

We thank Dr. James D. Lear for writing the Igor Pro procedures.

REFERENCES

1. Cleland, J. L., and Craik, C. S. (1996) *Protein Engineering*: *Principles and Practice*, Wiley-Liss, New York.
2. Mendel, D., Ellman, J. A., Chang, Z. Y., Veenstra, D. L., and Kollman, P. A. (1992) *Science 256*, 1798.
3. Matthews, B. W. (1995) *Adv. Protein Chem. 46*, 249.
4. Fersht, A. R., and Serrano, L. (1995) *Curr. Opin. Struct. Biol. 3*, 75.
5. Cornish, V. W., Mendel, D., and Schultz, P. G. (1995) *Angew. Chem., Int. Ed. Engl. 34*, 621.
6. van Hest, J. C. M., and Tirrell, D. A. (1998) *FEBS Lett. 428*, 68.
7. Duewel, H., Daub, E., Robinson, V., and Honek, J. F. (1997) *Biochemistry 36*, 3404.
8. van Hest, J. C. M., Kiick, K. L., and Tirrell, D. A. (2000) *J. Am. Chem. Soc. 122*, 1282.
9. Kiick, K. L., van Hest, J. C. M., and Tirrell, D. A. (2000) *Angew. Chem., Int. Ed. 39*, 2148.
10. Sharma, N., Furter, R., and Tirrell, D. A. (2000) *FEBS Lett. 467*, 37.
11. (a) O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1989) *Science 243*, 538. (b) O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) *Science 254*, 539. (c) Lumb, K. J., and Kim, P. S. (1995) *Science 268*, 436. (d) O'Shea, E. K., Lumb, K. H., and Kim, P. S. (1993) *Curr. Biol. 3*, 658.
12. (a) Hodges, R. S. (1996) *Biochem. Cell. Biol. 74*, 133−54. (b) Yu, Y., Monera, O. D., Hodges, R. S., and Privalov, P. L. (1996) *J. Mol. Biol. 255*, 367.
13. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) *Science 262*, 1401.
14. Wendt, H., Berger, C., Baici, A., Thomas, R. M., and Bosshard, H. R. (1995) *Biochemistry 34*, 4097.
15. Krylov, D., Mikhailenko, I., and Vinson, C. (1994) *EMBO J. 13*, 2849.
16. Schneider, J. P., Lear, J. D., and DeGrado, W. F. (1997) *J. Am. Chem. Soc. 119*, 5742.
17. Moran, L. B., Schneider, J. P., Kentsis, A. Reddy, G. A. and Sosnick, T. R. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 10699.
18. Myers, J. K., and Oas, T. G. (1999) *J. Mol. Biol. 289*, 205.
19. Thompson, K. S., Vinson, C. R., and Freire, E. (1993) *Biochemistry 32*, 5491.
20. (a) Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E., and Matthews, C. R. (1995) *Biochemistry 34*, 12812. (b) Zitzewitz, J. A., Ibarra-Molero, B., Fishel, D. R., Terry, K. L., and Matthews, C. R. (2000) *J. Mol. Biol. 296*, 1105.
21. (a) Moitra, J., Szilak, L., Krylov, D., and Vinson, C. (1997) *Biochemistry 36*, 12567. (b) Tripet, B., Wagschal, K., Lavigne, P., Mant, C. T., Hodges, R. S. (2000) *J. Mol. Biol. 300*, 377.
22. Arndt, K., and Fink, G. R. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8516.
23. Aizawa, Y., Sugiura, Y., Ueno, M., Mori, Y., Imoto, K., Makino, K., and Morii, T. (1999) *Biochemistry 38*, 4008.
24. Hockings, S. C., Kahn, J. D., and Crothers, D. M. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 1410.
25. Gough, C. A., Pearlman, D. A., and Kollman, P. (1993) *J. Chem. Phys. 99*, 9103.
26. Hine, J., and Mookerjee, P. K. (1975) *J. Org. Chem. 40*, 292.
27. Kukhar, V. P., and Soloshonok, V. A. (1995) *Fluorine Containing Amino Acids—Synthesis and Properties*, John Wiley & Sons, Chichester.
28. Rennert, O. M., and Anker, H. S. (1963) *Biochemistry 2*, 471.
29. Kothakota, S., Dougherty, M. J., Fournier, M. J., Mason, T. L., and Tirrell, D. A. (1995) *Macromol. Symp. 98*, 573.
30. Tang, Y., Ghirlanda, G., Petka, W. A., Nakajima, T., De Grado, W. F. and Tirrell, D. A. (2001) *Angew. Chem.* (in press).
31. Harding, S. E., Rowe, A. J., and Harton, J. C. (1992) *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, The Royal Society of Chemistry, Cambridge.
32. Brooks, I. S., Soneson, K. K., and Hensley, P. (1993) *Biophys. J. 64*, 244.
33. Lim, K. T., Brunett, S., Iotov, M., McClurg, B., Vaidehi, N., Dasgupta, S., Taylor, S., and Goddard, W. A., III (1997) *J. Comput. Chem. 18*, 501.
34. Ding, H. Q., Karasawa, N., and Goddard, W. A., III (1992) *J. Chem. Phys. 97*, 4309.
35. Tannor, D. J., Marten, B., Murphy, R., Friesner, R. A., Sitkoff, D., Nicholls, A., Ringnalda, M., and Goddard, W. A., III (1994) *J. Am. Chem. Soc. 116*, 11875.
36. Ghosh, A., Rapp, C. S., and Friesner, R. A. (1998) *J. Phys. Chem. B. 102*, 10983.
37. Mayo, S. L., Olafson, B. D., and Goddard, W. A., III (1990) *J. Phys. Chem. 94*, 8897.
38. Greeley, B. H., Russo, T. V., Mainz, D. T., Friesner, R. A., Langlois, J.-M., Goddard, W. A., III, Donnelly, R. E., and Ringnalda, M. N. (1994) *J. Chem. Phys. 101*, 4028.
39. Mackerell, A. D., Wiorkiewicz-Kuczera, J., and Karplus, M. J. (1996) *J. Am. Chem. Soc. 117*, 11946.
40. Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C., and Struhl, K. (1990) *Nature 347*, 575.
41. Ellenberger, T. E., Brandl, C. J., Struhl, K., and Harrison, S. C. (1992) *Cell 71*, 1223.
42. Konig, P., and Richmond, T. J. (1993) *J. Mol. Biol. 233*, 139.
43. Keller, W., Konig, P., and Richmond, T. J. (1995) *J. Mol. Biol. 234*, 657.
44. Metallo, S. J., and Schepartz, A. (1994) *Chem. Biol. 1*, 143.
45. Paolella, D. N., Palmer, C. R., and Schepartz, A. (1994) *Science 264*, 1130.
46. Creighton, T. (1997) *Proteins*: *Structure and Molecular Properties*, W. H. Freeman and Company, New York.
47. DeGrado, W. F., Summa, C. M., Pavone, V., Nastri, F., and Lombardi, A. (1999) *Annu. Rev. Biochem. 68*, 779.
48. Giver, L., Gershenson, A., Freskgard, P. O., and Arnold, F. H. (1995) *Proc. Natl. Acad. Sci. U.S.A. 95*, 12809.
49. Roux, M., Nezil, F., Monck, M., and Bloom, M. (1994) *Biochemistry 33*, 307.

BI0022588