Characterization of the Low pH Solution Structure and Dynamics of the Region 4 of *Escherichia coli* RNA Polymerase σ^{70} Subunit

Jarosław Poznański,* Krystyna Bolewska, Igor Zhukov, and Kazimierz L. Wierzchowski*

Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawińskiego 5a, 02-106 Warszawa, Poland

Received June 2, 2003; Revised Manuscript Received August 29, 2003

ABSTRACT: Solution structure of the region 4 of σ^{70} subunit of *Escherichia coli* RNA polymerase, whose 4.2 subregion is involved in specific recognition of the −35 element of cognate promoters, has not been yet studied. Using multinuclear NMR spectroscopy, we have assigned recently all the backbone and aliphatic side-chain ^{13}C resonances for a recombinant His_6-tagged protein containing the whole region 4 and a part of region 3.2 of σ^{70} in aqueous solution at pH 2.8 (Poznański, J., Zhukov, I., Bolewska, K., and Wierzchowski, K. L. (2001) *J. Biomol. NMR* 20, 181−2). The protein proved to be sufficiently soluble and did not aggregate only in the protonated state. In this paper, the structure and dynamics of this state at pH 2.8 have been extensively examined using CD and NMR spectroscopy. Both analysis of CD spectra and NMR observables (secondary chemical shifts of the $^{13}C\alpha$, ^{13}CO, and $^{1}H\alpha$ nuclei and of vicinal $^{3}J_{HNH\alpha}$ coupling constants) indicated that a significant amount of helical structure remained in the protonated protein. The amount of this structure increased upon deprotonation of carboxylic amino acids, as shown by pH titration CD experiments. 2,2,2-Trifluoroethanol induced an even more extensive build up of this structure. Distribution along the protein sequence of the secondary shifts and $^{3}J_{HNH\alpha}$ couplings demonstrated partition of the helical secondary structure into three helices located similarly as in the crystal structures of the homologous region 4 of the σ^{A} subunit of *Thermus aquaticus* RNA polymerase (Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, A., Weinman, O., Trester-Zedlitz, M. L., and Darst, S. A. (2002) *Mol. Cell* 9, 527−39) and σ^{70} of the *Thermus thermophilus* RNA polymerase (Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., and Yokoyama, S. (2002) *Nature 417,* 712−9.). Spectral density analysis of NMR relaxation parameters, R_1 and R_2, and $\{^{1}H\}-^{15}N$ heteronuclear NOEs indicated that backbone fluctuations in the whole region embracing the three helices and intervening nonhelical sequences are severely restricted on the nanosecond time scale as compared with the N- and C-terminal protein segments. Inspection of the side-chain contacts stabilizing the crystal structures well explains the observed folding and solution properties of $\sigma^{70}{}_4$ protein in its protonated state.

The primary sigma factor σ^{70} of *Escherichia coli* (EC σ^{70})[1] RNA polymerase regulates the initiation of the transcription of genes involved in the general metabolism of cells (*1*). When bound to the core of the RNA polymerase (α_2, β, β'), it helps to recognize the −35 and −10 regions of cognate promoters through specific interactions with the 4.2 and 2.4 subdomains, respectively. Some parts of the 4.2 subdomain have also been identified as binding sites for antisigma and transcription activation factors (*2*). Alignment of the 4.2 subdomain sequence with the HTH consensus sequence (3) and subsequent homology modeling studies (*2, 4*) have suggested that the structure of this region is similar to that of a typical HTH DNA binding motif. On our part, we have overexpressed the whole region 4 of σ^{70} (called $\sigma^{70}{}_4$) and assigned, sequence-specifically, ^{15}N and ^{13}C backbone and aliphatic side-chain resonances in the NMR spectra of this protein in acidic aqueous solution (*5*). Recently, crystal structures of corresponding regions in σ subunits of RNA polymerases $\sigma^{A}{}_4$ of *Thermus aquaticus* (*6*) and $\sigma^{70}{}_4$ of *Thermus thermophilus* (ref *7*, abbreviated there as CD) have been solved to a high resolution and shown to be consistent with earlier suggestions. In this paper, we report results of our further heteronuclear NMR and CD studies on the solution structure of the $\sigma^{70}{}_4$ recombinant polypeptide. *N*-His_6-tagged $\sigma^{70}{}_4$ dissolves in water without aggregation only below pH 3, where it is almost fully protonated. Preliminary chemical shift analysis suggested that the protonated protein could retain locally the helical secondary structure (*5*). Therefore, a number of NMR parameters including secondary chemical shifts, $^{3}J_{HNH\alpha}$ coupling constants, $\{^{1}H\}-^{15}N$ heteronuclear NOE, and longitudinal (R_1) and transverse (R_2) relaxation rates, known to be sensitive to the structure surrounding the corresponding nuclei (*8*), were determined and used to evaluate the propensity of various backbone regions of the protonated protein to retain locally a stable conformation. By all the criteria applied, in two subregions of $\sigma^{70}{}_4$, 4.1 and 4.2, three helices were found

* Corresponding authors. E-mail: (J.P.) jarek@ibb.waw.pl; (K.L.W.) klw@ibb.waw.pl.

[1] Abbreviations: EC $\sigma^{70}{}_4$, region 4 of σ^{70} subunit of *Escherichia coli* RNA polymerase; TT $\sigma^{70}{}_4$, region 4 of σ^{70} subunit of *Thermus thermophilus* RNA polymerase; $\sigma^{A}{}_4$, region 4 of σ^{A} subunit of *Thermus aquaticus* RNA polymerase; HLHTH, helix−loop−helix−turn−helix motif; TFE, 2,2,2-trifluoroethanol.

10.1021/bi034943v CCC: $25.00
Published on Web 10/30/2003

to be significantly populated, while the intervening residues forming a putative loop and turn remained in an extended β-conformation. Comparison of the conformational and dynamic properties of $\sigma^{70}{}_4$ in solution with the crystal structures of $\sigma^{A}{}_4$ (*6*) and TT $\sigma^{70}{}_4$ (*7*) has shown that the three helical regions forming this structure are partially retained in the ensemble of unfolded states.

MATERIALS AND METHODS

Materials. Restriction enzymes and T_4 DNA ligase were purchased from Amersham. Plasmids pHTT7f1-σ and pAED4 were obtained from Drs. W. Werel and P. S. Kim, respectively; plasmid pET-15b and His-Bind resin were purchased from Novagen. Ammonium sulfate (^{15}N, 98%) and D-glucose-$^{13}C_6$ (^{13}C, 98%) were from Cambridge Isotope Laboratories, Inc. All chemicals used were of the highest purity grade.

Preparation of Recombinant $\sigma^{70}{}_4$ Protein. (1) Plasmid Construction. The Xho I/Hind III fragment of the rpoD gene, corresponding to the C-terminal residues 528−613 of $\sigma^{70}{}_4$, was first cloned into pAED4 plasmid, cut off from the latter using Xho I and BamH I restriction enzymes, and recloned into pET-15b plasmid within a frame containing a sequence coding for the $(His)_6$−Tag and thrombin cleavage site. The sequence of the gene was confirmed by the dideoxy sequencing method.

(2) Gene Expression and Protein Purification. E. coli cells, strain BL21(DE3), were transformed with pET-15b plasmid bearing the $\sigma^{70}{}_4$ gene fragment. The transformed cells were grown from overnight cultures at 37 °C in Luria broth (500 mL) containing 100 μg/mL ampicilin, induced at $A_{600} = 0.7$ with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 1.5 h, and harvested by centrifugation. The overexpressed protein was found accumulated in the form of inclusion bodies. The cells (ca. 2.0 g) were resuspended in 10 mL of buffer A (20 mM Tris·HCl, pH 7.9/500 mM NaCl/5 mM imidazole), sonicated, and centrifuged at 4 °C. The pelleted cells were washed with buffer A by centrifugation. The pellet was resuspended in 7 mL of buffer B (20 mM Tris·HCl, pH 7.9/500 mM NaCl/5 mM imidazole/6 M urea) and incubated at room temperature for 1 h and again centrifuged at 4 °C. The supernatant was loaded on a 1 mL column containing His·Bind resin, but the column was washed first with 15 mL of buffer B and then with 7.5 mL of 25 mM imidazole buffer B. The His_6-tagged protein was eluted with 4 mL of 0.5 M imidazole in buffer B. The yield was ca. 25 mg per 1 L of the cell culture. For NMR investigations, the protein was uniformly labeled isotopically (*9*), and the yields of the ^{15}N-labeled and ^{15}N, ^{13}C double-labeled protein were, respectively, 18 and 15 mg per 1 L of cell culture. All protein samples were found homogeneous in polyacrylamide gel electrophoresis and electrospray mass spectrometry. The last method and NMR sequencing indicated a lack of the N-terminal methionine. The experimentally determined mass of the unlabeled protein $M_w = 12\,033.3$ was close to the calculated average one of 12 032.5. Similar agreement was found between experimental and calculated masses of labeled proteins.

The His_6-tag from the recombinant protein was removed by thrombin cleavage that, owing to solution properties of the protein, had to be performed under acidic pH conditions that are far from optimal for this protease: 20 mM MES buffer, pH 5.0 at 0.75 mg/mL protein concentration (1:500, w/w ratio of thrombin to the fusion protein), and incubation time 20 h at room temperature. The yield of the reaction was less than 50%. The cleavage product was separated from the remaining His-tagged protein using the Ni^{2+}-NTA agarose batch procedure under denaturing conditions in buffer B. The effluent containing denatured $\sigma^{70}{}_4$ was dialyzed against 1 mM HCl, lyophilized, dissolved again in buffer B, and mixed with His·Bind resin (0.5 mL per 10 mg of protein) for 1 h. After the removal of the beads by centrifugation, the solution was dialyzed against 1 mM HCl. Protein concentrations were determined according to Bradford (*10*). Polyacrylamide gel electrophoresis of protein samples was performed according to ref *11*.

CD Spectroscopy. The CD spectra were recorded using an Aviv Model 202 spectropolarimeter equipped with a HP 89100A temperature controller. All measurements were carried out in a 1 mm path-length cell. The protein concentration was determined spectrophotometrically at 274 nm assuming a molar extinction coefficient of 1400 (*12*). CD spectra were deconvoluted into contributions from various structural forms with the aid of the CDNN program (*13*).

Mass Spectrometry. Mass spectra for natural abundance, ^{15}N enriched, and ^{13}C, ^{15}N double enriched proteins were determined on an electrospray (ESI−MS) quadrupole time-of-flight (Q-Tof) Micromass spectrometer, with a resolution of approximately 10 000. The spectrometer was calibrated with a bovine pancreatic trypsin inhibitor (BPTI) probe.

NMR Spectroscopy. ^{1}H, ^{13}C, and ^{15}N chemical shifts of $\sigma^{70}{}_4$ were determined for ^{15}N, ^{13}C doubly enriched protein at 3 mM concentration in aqueous solution (pH 2.8) containing 10% D_2O, at 298 K. All 2-D and 3-D heteronuclear spectra were processed with the aid of NMRPipe (*14*) and analyzed using the XEASY (*15*) programs. Resolutions in the indirect dimensions were increased by linear prediction and $\pi/3$ shifted squared sine-bell weight function multiplication followed by zero-filling.

Resonance Assignments. The ^{15}N−^{1}H HSQC resonances were assigned to a given type of residue using the C(CO)NH (*16*), HN(CO)CA (*17*), and CBCA(CO)NH (*18*) experiments. Because of extensive overlap of the carbon signals, sequential assignment was done simultaneously by a combination of the gradient versions of the CBCA(CO)NH (*18*), HNCACB (*19*), HNCO (*20*), and (HCA)CO(CA)NH (*21*) experiments. Hα resonances were assigned on the basis of the HA(CO)NH (*22*) experiment. Most of the ^{1}H aliphatic side-chain resonances were assigned on the basis of a combination of H(CCO)NH (*22*), ^{15}N−NOESY−HSQC (*23*) ($\tau_m = 150$, 250 ms), and ^{13}C−^{1}H HSQC spectra. To avoid errors, each sequential assignment trace was followed using Cα, Cβ, and CO resonances and verified by the ^{15}N−NOESY−HSQC experiment.

$^3J_{HNH\alpha}$ Coupling Constants. Values of $^3J_{HNH\alpha}$ were evaluated from a series of *J*-modulated HSQC experimental data (*24*) according to the following relation: $V = V_o \exp(-t/T_2) \cos(2\pi Jt)$, where t is the evolution time (40, 50, 60, 70, 80, 90, 100, 120, 160 ms), T_2 is the HN transverse relaxation time, and V and V_o are signal volumes measured (at t) and estimated (at $t = 0$), respectively. The $^3J_{HNH\alpha}$ vicinal coupling constants, ^{15}N relaxation parameters, and $\{^1H\}$−^{15}N heteronuclear NOEs (see next) have been deposited in the Bio-

MagResBank database (http://www.bmrb.wisc.edu) under the same accession number of 4870 at which chemical shift data were deposited previously (*5*).

Relaxation Measurements. Relaxation measurements were performed at 298 K for ^{15}N uniformly labeled protein dissolved in water (pH 2.8) at a single 11.4 T magnetic field (500 MHz spectrometer). The pulse sequences, used for determination of the longitudinal (R_1) and transverse (R_2) relaxation rates, were analogous to those previously reported by Farrow et al. (*25*). For R_2 measurements, a Carr−Parcell−Meiboom−Gill (CPMG) 180° pulse train with a refocusing delay of 650 μs was used during evolution. Delays between proton π pulses, used for the suppression of cross-correlation effects between ^{1}H and ^{15}N nuclei (*26*), were 5 and 10 ms in R_1 and R_2 measurements, respectively. The recycle delay was kept as long as 2.5 s. A total of 1024 and 128 complex data points in time domains were collected in the hypercomplex mode. ^{15}N decoupling during acquisition was performed as 3.2 kHz GARP pulse scheme (*27*). The relaxation rates were measured using eight delays both for R_1 (0.01, 0.05, 0.13, 0.25, 0.41, 0.61, 0.85, 1.13 s) and for R_2 (0.01, 0.03, 0.05, 0.09, 0.13, 0.17, 0.21, 0.25 s). {^{1}H}−^{15}N heteronuclear NOEs were measured according to the dynamic−progressive saturation manner with seven delays (0.00, 0.06, 0.18, 0.42, 0.90, 1.80, 3.30 s) using the standard pulse sequence (gNnoe) from the ProteinPack (Varian Inc., Palo Alto, CA) software.

Data Transformation. To avoid problems with calibration and to diminish experimental uncertainties, all relaxation experiments were performed as the pseudo-3-D experiments (*28*). The specific processing script for transforming such data under the NMRPipe program (*14*) was used. Spectra were processed by applying a square cosine-bell weighting function in both time dimensions followed by a zero-filling procedure up to 2000 and 1000 complex points. The baseline correction was applied during processing along the F2 dimension.

Determination of Relaxation Parameters. Relaxation analysis was performed for 94 out of the 102 backbone amide groups characterized by nonoverlapping cross-peaks in the ^{1}H−^{15}N HSQC spectrum, using the method of spectral density mapping (*29*). This method provides values of $J(0)$, $J(\omega_N)$, and $J(0.87\omega_H)$ from R_1, R_2, and the {^{1}H}−^{15}N cross relaxation rates.

Computational Methods. Sequential alignment of $\sigma^A{}_4$ and $\sigma^{70}{}_4$ domains was conducted using version 1.8 of the CLUSTAL-W program (*30*) and the Blosum30 (*31*) scoring matrix. Structural analysis was performed with the help of the MolMol program (*32*).

Statistical Treatment. For more precise analysis of the distribution of NMR observables (CSI and $^3J_{HNH\alpha}$) along the sequence, a moving average smoothing (*33*) procedure was applied with a five residue window roughly corresponding to the length of one helical turn.

RESULTS

Properties of the His$_6$-Tagged $\sigma^{70}{}_4$ in Solution. The recombinant 107 residue protein contained the C-terminal fragment of σ^{70}, residues from 528 to 613 (comprising the whole region 4 and a part of region 3.2), and a 21 residue sequence carrying the thrombin cleavage site and N−His$_6$−Tag (cf. Figure 1). The protein was purified by metal-ion affinity chromatography under denaturing conditions (6 M urea), as described in the Materials and Methods. Attempts at its renaturation by stepwise removal of the denaturing agent through dialysis against 50 mM Tris/HCl and acetate buffers of pH from 4 to 8 failed because the protein precipitated. Only dialysis against 1 mM HCl allowed the protein to be retained in solution in a nonaggregated form at a concentration up to ca. 5 mM. Under these pH conditions, all basic amino acid residues and most of the acidic ones are protonated (calculated pI = 9.79). Lyophilized protein could be redissolved in water, but when the pH was increased above 4, it started to aggregate, as indicated by the broadening and disappearance of signals from the ^{15}N−HSQC NMR spectrum. The aggregation could be reversed by the addition of HCl to pH ≈3. In acidic solution, the protein was found to be chemically stable for a prolonged time.

```
EC                MGSSHHHHHH SSGLVPRGSH MLELPLDSAT 536

EC     TESLRAATHD VLAGLTAREA KVLRMRFGID MNTDYTLEEV 576
       :*.*. *      *: *: ***  **::* *:   . ::***** 
TAQ    SEELEKA--- -LSKLSEREA MVLKLRKGLI DGREHTLEEV 401
TT     SEELEKA--- -LSKLSEREA MVLKLRKGLI DGREHTLEEV 386

EC     GKQFDVTRER IRQIEAKALR KLRH-PSRSE VLRSFLDD   613
       *  *.***** ***** **** **:   **:.  **.**:   
TAQ    GAYFGVTRER IRQIENKALR KLKYHESRTR KLRDFLE-   438
TT     GAFFGVTRER IRQIENKALR KLKYHESRTR KLRDFLD-   423
```

FIGURE 1: Sequence of the recombinant *E. coli* (EC) $\sigma^{70}{}_4$ and its alignment with sequences of the *T. aquaticus* (TAQ) $\sigma^A{}_4$ (*6*) and *T. thermophilus* (TT) $\sigma^{70}{}_4$ domain of the cognate RNA polymerase σ subunit (*7*). The His$_6$-tag and thrombin cleavage sequence are in italics. In the alignment, symbols *, :, and . indicate identical residues, homologues, and polar replacements, respectively. Regions 4.1 and 4.2 of EC $\sigma^{70}{}_4$ (*1*) are underlined and double underlined, respectively. Helical regions in the crystal structures of $\sigma^A{}_4$ and TT $\sigma^{70}{}_4$ domains, named H1, H2, and H3, are marked by a gray background.

Untagged $\sigma^{70}{}_4$ (calculated pI = 8) exhibited a stronger tendency for aggregation relative to its tagged form under similar pH and concentration conditions. Therefore, the tagged form, called $\sigma^{70}{}_4$ further in the text, was selected for CD and NMR investigations.

CD Investigations. CD spectra of $\sigma^{70}{}_4$ recorded at a number of pH values are shown in Figure 2. The spectra were deconvoluted (*13*), and contributions from different conformational forms thus obtained were plotted as a function of pH in the inset to Figure 2. It is evident from these plots that at pH 2.7 the dominant contributions to the spectrum come from extended β (ca. 0.38) and random coil (0.31) conformations, while those due to β-turn (0.2) and α-helical (0.11) secondary structures are much smaller. Previously (*5*), the content of the latter form was overestimated. Upon increasing the pH up to 4.7, the content of the α-helical form grows at the cost of the extended β-ones, whereas contributions from the other two remain practically invariant. This observation indicates the occurrence of a simple pH-dependent two-state conformational equilibrium between some extended β and α-helical forms. The transition between these two states manifests itself in the spectra by the presence of an isoeliptic point at 204 nm. Therefore, from these plots, the corresponding two-state equilibrium constant, pK_{trans} = 3.94 (±0.13), was estimated. The value of pK_{trans} suggests that this transition is controlled by protonation/deprotonation

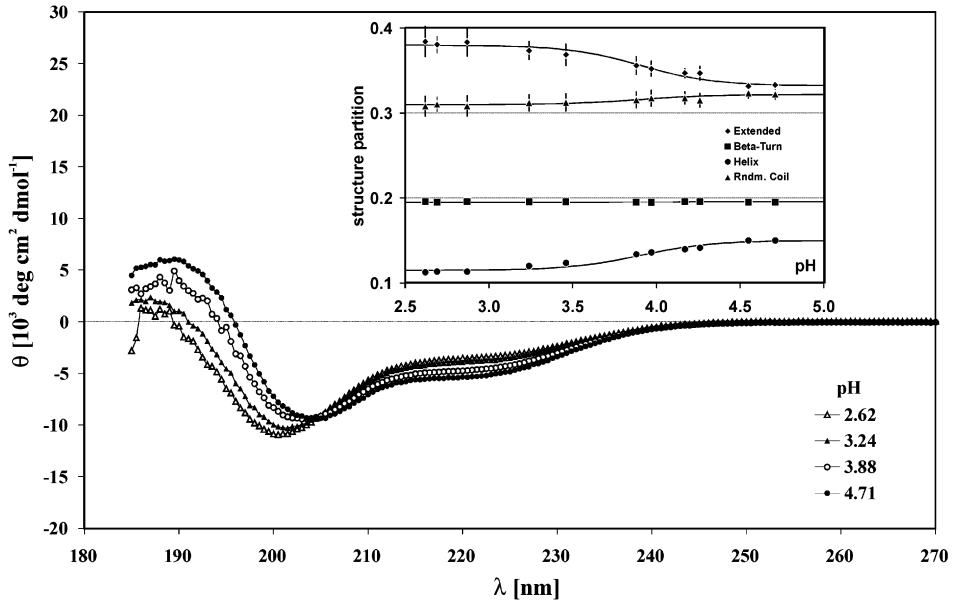

FIGURE 2: CD spectra of σ^{70}_4 in aqueous solution at indicated pH values, 25 °C. In the inset, the estimated populations of different secondary structure elements contributing to the observed spectra are plotted as a function of pH (solid lines drawn through the data points correspond to the two-state transition with $pK_{trans} = 3.94$).

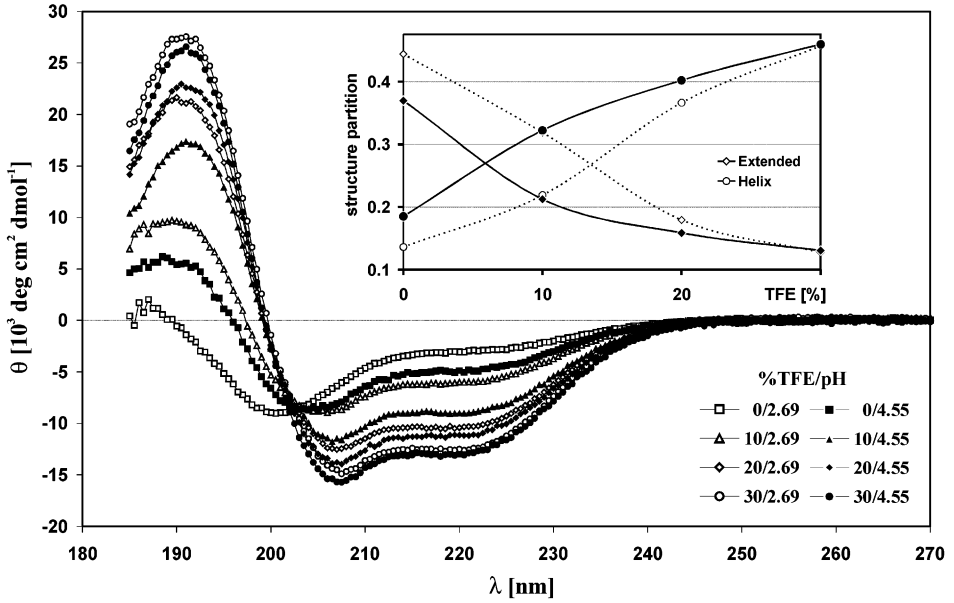

FIGURE 3: TFE-induced changes in the CD spectra of σ^{70}_4 at pH 2.69 (open symbols) and 4.55 (filled symbols); the content of TFE (%, v/v) is indicated in the script window. In the inset, the estimated populations of α-helical and extended β-secondary structure elements are plotted as a function of TFE content (practically invariant contributions from random coil and β-turn forms omitted).

equilibria involving carboxylic groups of Glu and Asp and the C-terminal carboxyl present in the protein.

The CD spectrum of the protonated form of σ^{70}_4 was only slightly sensitive to temperature: between 5 and 45 °C, only a general decrease in optical activity, accompanied by a small blue shift of the minimum at 200 nm, was observed (not shown).

The presented data allow us to conclude that protonated σ^{70}_4 in aqueous solution is partially unfolded and retains some residual α-helical structure since CD is sensitive only to ordered secondary structures. This conclusion is strongly supported by changes induced in the CD spectra of σ^{70}_4 (cf. Figure 3) on the stepwise addition of TFE. Owing to specific solvation of peptide groups (*34*, *35*), TFE induces the closure of individual hydrogen bonds (*36*) and thus propagates context-dependent formation of nativelike secondary structures in polypeptides (*37*). The spectra recorded at pH 2.69 and 4.55 in solutions containing 0−30% (v/v) of TFE form a family with an isoeliptic point at 203 nm indicative of the occurrence of a conformational transition induced by TFE. Indeed, systematic changes in the calculated contributions to the experimental spectra from α-helical and β-extended forms (contributions from the random coil and β-turn form proved practically invariant), plotted as a function of TFE content (cf. inset to Figure 3), clearly demonstrate that they are mainly due to the build-up of α-helical structures up to 45%, almost at the sole cost of the extended β-forms. TFE induced also large systematic changes in the chemical shift patterns of $^{15}N-^{1}H$ HSQC NMR spectra (not shown), the structural interpretation of which was not yet attempted.

NMR Investigations. Detailed insight into the structure of partially folded and unfolded proteins can be obtained with the help of high-resolution NMR experiments providing information on the distribution of the φ, ψ dihedral angles and backbone dynamics (*8*). In particular, the deviations of the chemical shifts of $^{13}C\alpha$, ^{13}CO, and $^{1}H\alpha$ nuclei from random coil values, termed secondary chemical shifts, and the vicinal $^{3}J_{HNH\alpha}$ coupling constants, both primarily determined by the backbone φ, ψ dihedral angles, are used for the estimation of local secondary structure propensities in unfolded or partially folded proteins. Analysis of measured ^{15}N longitudinal (R_1) and transverse (R_2) magnetization relaxation rates, as well as the heteronuclear $\{^{1}H\}-^{15}N$ steady-state NOEs, provides clues to backbone dynamics. Some connectivity patterns, inferred from NOESY experiments, can also be useful. We applied all these approaches to further characterize secondary structures contributing to the CD spectra of protonated σ^{70}_4.

Chemical Shifts. Using the experimental chemical shift values of $^{13}C\alpha$, ^{13}CO, and $^{1}H\alpha$ from the NMR spectra of σ^{70}_4 in aqueous solution at pH 2.8 and respective reference values for random-coil chemical shifts (*38*−*42*) (corrected for sequence-dependent contributions (*43*)), corresponding secondary chemical shifts, $\Delta\delta(i)$, were calculated. Plots of the $\Delta\delta(i)$ values for $^{13}C\alpha$, ^{13}CO, and $^{1}H\alpha$ nuclei as a function of residue number in the σ^{70}_4 sequence shown in Figure 4 indicate that in some regions of the 4.1 and 4.2 subregions, ^{13}CO and $^{13}C\alpha$ resonances form a consistent pattern of downfield shifts, whereas $^{1}H\alpha$ are shifted upfield; all the three descriptors are thus indicative of a high propensity of these backbone regions to populate the helical conformation (*8*). These regions are helically folded in the crystal state of the homolgous domains of *T. aquaticus* (*6*) and *T. thermophilus* (*7*).

The population of those helically folded regions in the protonated σ^{70}_4 was evaluated from the measured distribution of the residual secondary chemical shift values $\Delta\delta(i)$ (Figure 4) based on the concept of the chemical shift index (CSI) (*39*). Assuming for the α-helical structure average values of secondary chemical shifts $\Delta\delta_o(i)$ for $^{13}C\alpha$, ^{13}CO, and $^{1}H\alpha$ equal to 2.8, 2.3, and −0.4 ppm, respectively (*41*, *42*), the population of the *i*th residue in this conformation can be expressed as $p_\alpha(i) = \Delta\delta(i)/\Delta\delta_o(i)$. The data obtained were smoothed by a moving average method (*33*) with a five residue window, using the expression:

$$\tilde{p}_\alpha(i) = \max\left(0, \frac{1}{5}\sum_{|j-i|<3} p_\alpha(j)\right)$$

In view of the consistency of the $\Delta\delta(i)$ distribution pattern for $^{13}C\alpha$, ^{13}CO, and $^{1}H\alpha$ (cf. Figure 4a−c), population analysis of local structural propensities could be done simultaneously for the three structure descriptors in the form of their geometric average:

$$\hat{p}_\alpha(i) = \sqrt[3]{\prod_{C\alpha,CO,H\alpha} \tilde{p}_\alpha(i)}$$

The calculated values of the $\hat{P}_\alpha(i)$ function, plotted in Figure 4d, clearly identified in σ^{70}_4 the two most populated helical regions embracing the L551−M561 and F580−A594 sequences, corresponding to the H1 and H3 crystal state helices, respectively, and a much less populated third region of T572−E575, corresponding to the H2 helix in the crystal structure. The average partition of helical structure in protonated σ^{70}_4 estimated in this way was found to be equal to 0.11, in perfect agreement with the population of this structure obtained from the analysis of the CD spectra (see infra).

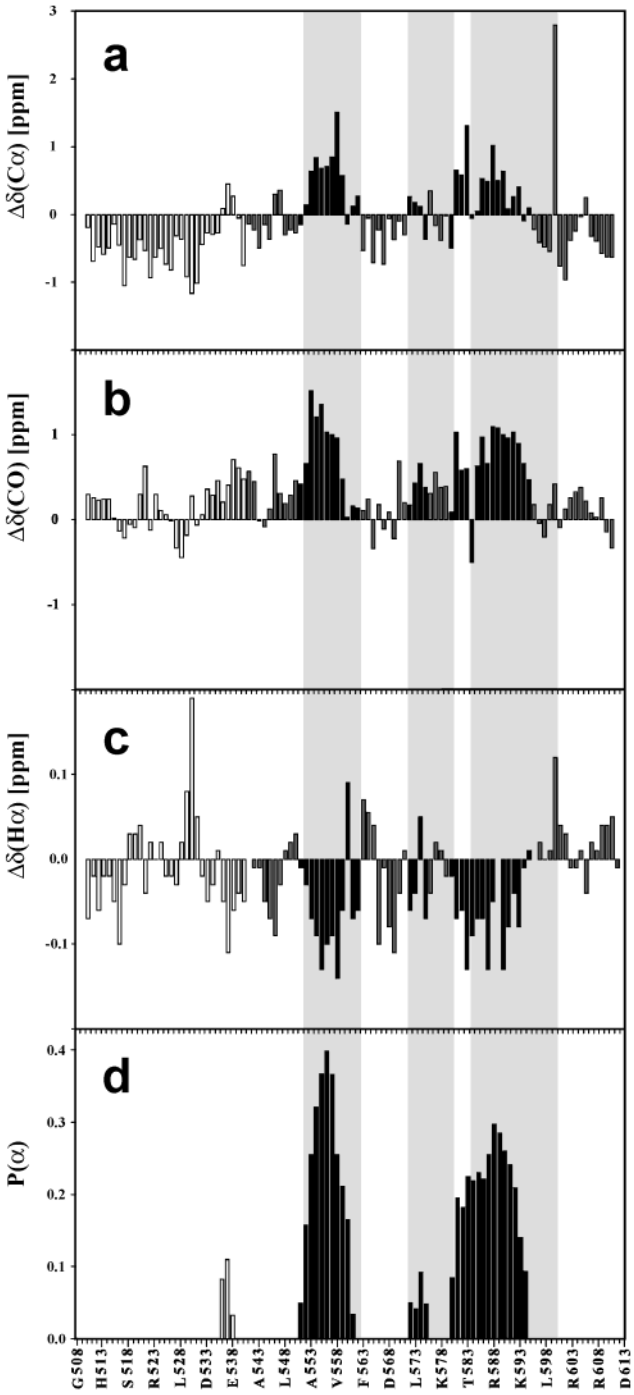

FIGURE 4: Secondary chemical shifts, $\Delta\delta$(i), corrected for the sequence-dependent contributions (*43*), for (a) $^{13}C\alpha$, (b) ^{13}CO, and (c) $^{1}H\alpha$ nuclei as a function of residue number in protonated σ^{70}_4 at pH 2.8 and 25 °C. (d) Population of α-helical regions in σ^{70}_4 estimated as the residue-dependent geometric mean of the $\Delta\delta(C\alpha)$, $\Delta\delta(CO)$, and $\Delta\delta(H\alpha)$ descriptors averaged over five residue windows; black bars identify helical regions, and outlined ones correspond to the N-terminal part of σ^{70}_4 preceding the proper domain 4 sequence; helical regions found in the crystal structure of TT σ^{70}_4 (*7*) are indicated by gray vertical strips.

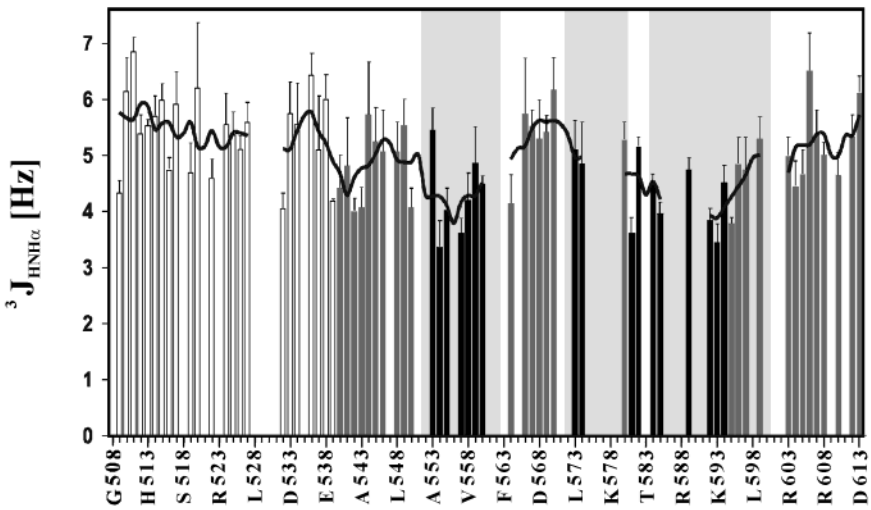

FIGURE 5: Vicinal coupling constants $^3J_{HNH\alpha}$ of protonated σ^{70}_4 at pH 2.8 and 25 °C. Coupling constants averaged over the five residue windows are depicted as solid curves over the bars. Black bars identify helical regions, and outlined ones correspond to the N-terminal part of σ^{70}_4 preceding the proper domain 4 sequence; gray vertical strips mark helical regions found in the crystal structure of TT σ^{70}_4 (*7*).

Comparison of the three helical regions in the solution structure of protonated σ^{70}_4 and in the crystal structure of TT σ^{70}_4 (cf. gray strips in Figure 4d) indicates that the H1 and H3 helices are located similarly. In the protonated σ^{70}_4, the ends of helix H1 are considerably frayed, as indicated by the distribution of $\hat{P}_\alpha(i)$ values; the longer H3 helix seems to be more compact than helix H1 and does not extend up to P601 toward the C-end, as observed in the crystal structures, and propagates back into the turn 581DVT region; the third H2 helix is less populated and much shorter. However, both location and extension of the loop region between the H1 and the H2 helices in the solution structure are similar to that seen in the crystal structures. The backbone conformation of the loop in solution seems to be generally similar, as suggested by negative values of $\Delta\delta(C\alpha)$ in this region (cf. Figure 4a), compatible with a β-type structure.

$^3J_{HNH\alpha}$ Coupling Constants. The measured values of vicinal $^3J_{HNH\alpha}$ coupling constants for protonated σ^{70}_4 at pH 2.8 are presented in the form of a histogram in Figure 5 (some of the couplings could not be determined owing to the signal's overlap). They exhibit relatively low variability along the protein sequence in the range of 3.5−6.5 Hz. To make this distribution pattern more meaningful, measured values of coupling constants were averaged over the five residue window (cf. solid boundaries over the bars in Figure 5), which roughly corresponds to one helical turn.

Interpretation of the vicinal coupling data can be performed with a reference to the mean $^3J_{HNH\alpha}$ values predicted for α-helical and β-strand conformations from the population of torsion angles in a database of 85 high-resolution protein structures (*44*) equal to 5.2 Hz (with residues at the termini of helical structures included in calculations) and 8.5 Hz, respectively. The calculated coupling constants for a majority of residues found in the random coil state in the database are in the range of 7.1−7.7 Hz. In unfolded protein states in solution, conformational averaging would result in intermediate values of the observed $^3J_{HNH\alpha}$ couplings.

The $^3J_{HNH\alpha}$ values determined for σ^{70}_4 at pH 2.8 are well below those expected for residues in the random coil state and thus point to a significant population of the helical φ, ψ dihedral angles in the ensemble of unfolded states. Despite some gaps in the $^3J_{HNH\alpha}$ data, the two most populated α-helical regions, H1 and H3, proposed on the basis of the secondary chemical shifts, could be easily recognized in the regions of consistently lower $^3J_{HNH\alpha}$ values between 3.5 and 5 Hz. The elevated values of the coupling constant, $^3J_{HNH\alpha} > 5$ Hz, in the I565−E574 region corresponding to a putative loop between the two helices indicate a larger contribution from extended conformations.

Table 1: Structural Crosspeaks from ^{15}N Edited NOESY Spectrum (250 ms Mixing Time) of σ^{70}_4 at pH 2.8[a]

	NOE between atoms				
L	548	HN	A	546	HA
G	**550**	**HN**	**L**	**548**	**QQD**
G	550	HN	L	548	QB
T	552	HN	E	555	QB
T	552	HN	E	555	QG
R	554	HN	T	552	HB
A	556	HN	L	551	QQD
F	**563**	**HN**	**I**	**565**	**QG2**
G	564	HN	R	562	HA
G	**564**	**HN**	**M**	**567**	**QB**
I	565	HN	F	563	QB
I	565	HN	F	563	HA
M	567	HN	I	565	HA
D	**570**	**HN**	**I**	**565**	**QG2**
T	572	HN	E	575	QG
L	**573**	**HN**	**D**	**581**	**HA**
E	**574**	**HN**	**D**	**581**	**HA**
G	577	HN	E	575	HA
Q	579	HN	G	577	QA
D	581	HN	G	577	QA

[a] In bold, those that do not agree within the 6 Å limit with the crystal structure of TT σ^{70}_4 (*7*).

NOESY Data. Analysis of the ^{15}N−NOESY−HSQC spectra allowed us to assign per residue more than three sequential cross-peaks. In the HLH region, 20 relatively strong $HN_{i,i+1}$ correlations were found. However, in the H3 region, the existence of an efficient $HN_{i,i+1}$ magnetization transfer could be neither confirmed nor falsified. Additionally, a total number of 20 structural cross-peaks was found and unequivocally assigned (cf. Table 1). Unfortunately, because of very low dispersion of Hα and amide HN resonances and the presence of slow conformational ex-

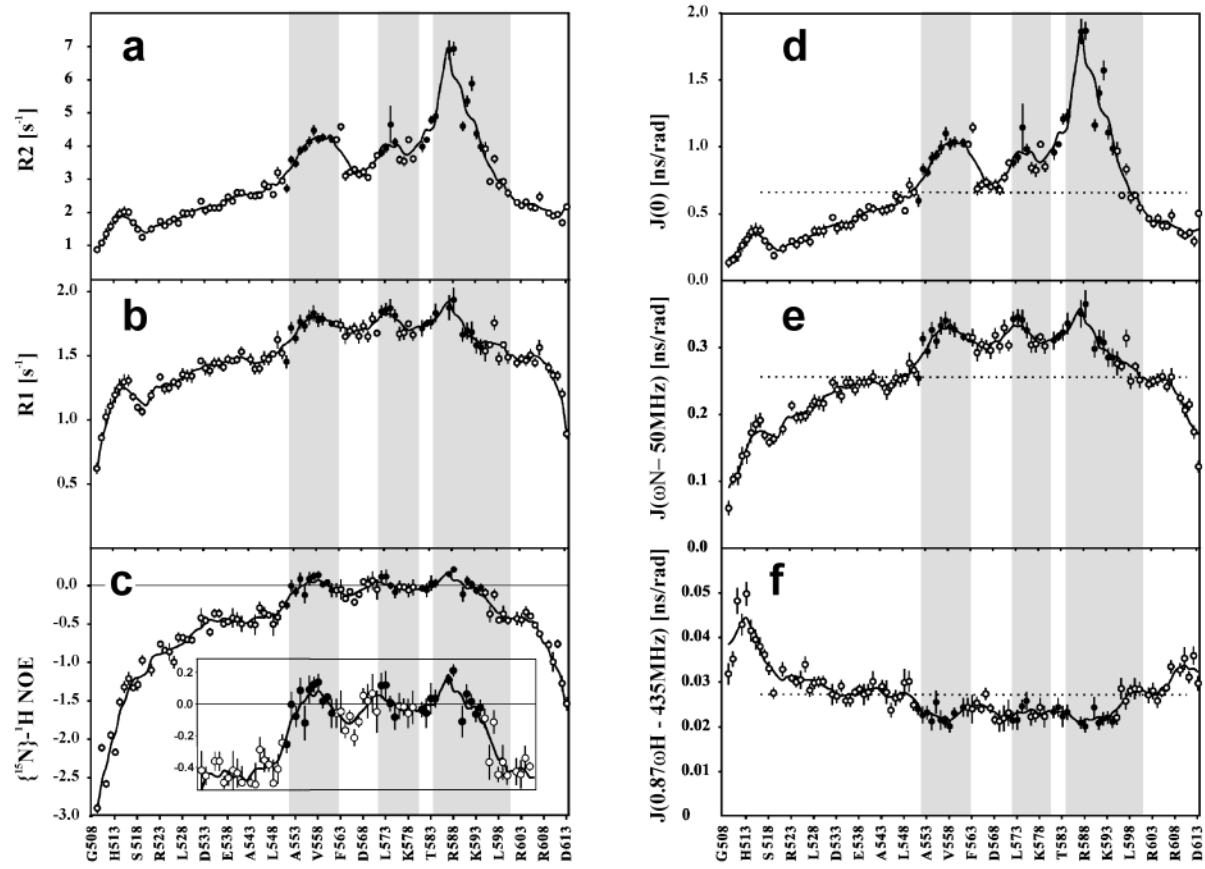

FIGURE 6: (a) Longitudinal (R_2) relaxation rates, (b) transverse (R_1) relaxation rates, and (c) heteronuclear {^{1}H}−^{15}N NOEs (in inset, magnified central fragment) for the amide backbone nitrogens in aqueous σ^{70}_4, at pH 2.8 and 25 °C, and corresponding calculated spectral density values of $J(\omega)$ function sampled at (d) 0, (e) 50, and (f) 465 MHz. The horizontal dotted lines in panels d−f show the mean values of 0.67, 0.26, and 0.028 ns rad^{-1}, respectively. Filled and open circles correspond to α-helical and extended β-structures, respectively; gray vertical strips mark helical regions found in the crystal structure of TT σ^{70}_4 (*7*).

change, the $i,i+3$ connectivity pattern that would support the postulated helical structures could not be discerned.

Inspection of the assigned structural cross-peaks shows that they locate mainly at the termini of putative helices, and in particular, in the I565−E574 loop region. This observation will be discussed later in connection with the crystal structure of TT σ^{70}_4.

Relaxation Analysis. Longitudinal (R_1) and transverse (R_2) magnetization relaxation rates for amide ^{15}N nuclei and the heteronuclear {^{1}H}−^{15}N steady-state NOEs were measured at a single 11.4 T magnetic field for 94 out of 102 backbone amide ^{15}N nuclei observed in the ^{1}H−^{15}N HSQC spectrum of uniformly ^{15}N-labeled σ^{70}_4. The R_1 and R_2 relaxation rates and {^{1}H}−^{15}N NOEs thus obtained exhibit a characteristic variation along the σ^{70}_4 sequence (Figure 6a−c), indicative of greatly differentiated dynamics of the protein backbone in the acid-unfolded state. The patterns of distribution of the relaxation parameters across the whole region embracing residues T552−L598 are similar to each other and characterized by the presence of three maxima corresponding to relatively slower backbone motions. Location of the three maxima coincides closely with the three H1−H3 helical regions deduced from the $\Delta\delta$ and $^3J_{HNH\alpha}$ data (cf. Figures 4 and 5). Residues exhibiting weakly positive NOEs (cf. inset in Figure 6c) and elevated R_1 values in these subregions of the T552−L598 sequence are motionally restricted on a subnanosecond time scale; simultaneously, they are characterized by the largest R_2 values indicative of slower motions with a possible contribution from chemical exchange processes on the micro- to millisecond time scale. Large negative values of {^{1}H}−^{15}N NOE and distinctly lower R_1 and R_2 values for N- and C-terminal sequences indicate increased flexibility of the polypeptide chain in these regions (*45*). Only the His_6-tag fragment seems to be somewhat more restricted motionally.

The model-free approach (*46*) is not applicable for the description of backbone dynamics of unfolded proteins because of their nonisotropic tumbling in solution (*47*−*49*). In such a case, quantitative interpretation of the ^{15}N relaxation data can be performed by direct analysis of the spectral density function $J(\omega)$ using spectral density mapping (*50*). Such analysis has been successfully applied in investigations on the conformational dynamics of a number of unfolded proteins (*8*, *51*−*53*). Following this approach, the spectral densities $J(0)$, $J(\omega_N)$, and $J(0.87\omega_H)$, sampled at frequencies 0, $\omega_N = 50$, and $0.87\omega_H = 465$ MHz, respectively, were calculated from the measured {^{1}H}−^{15}N NOEs and R_1 and R_2 relaxation rates. They are plotted as a function of residue number in panels d−f of Figure 6.

$J(0)$ is more sensitive to R_2; therefore, increased values of this function reflect the occurrence of slower motions on the nanosecond time scale with possible contribution from chemical exchange processes on the micro- to millisecond time scale, while low values point to a dominant contribution from rapid pico- to nanosecond internal motions. Indeed, the distribution pattern of the calculated $J(0)$ values (Figure 6d)

closely matches that of R_2 (Figure 6a). The whole fragment embracing the H1−H3 three helical subregions and a part of the intervening loop between H1 and H2 helices, approximately F563−Y571, exhibits significantly elevated $J(0)$ values over the average value of $J(0) = 0.67$ ns/rad. The maxima of the $J(0)$ function at the middle part of the H1 and H3 helical sequences reflect a slow conformational exchange of the residues in the interface region. However, most of the F563−Y571 loop sequence is characterized by the lowest $J(0)$ values in the whole region in question. Also, there is an evident minimum in the $J(0)$ plot between H2 and H3 helical regions. The flanking sequences are characterized by distinctly lower $J(0)$ values, compatible with the dominance of fast pico- to nanosecond motions. Also, $J(\omega_N)$ values are distinctly larger in this region and similarly distributed as in the case of $J(0)$ (cf. Figure 6d). The $J(0.87\omega_H)$ values most sensitive to fast pico- to nanosecond internal motions proved to be the smallest, exhibiting an inverted distribution over this region (Figure 6f), as compared with the distribution of $J(0)$ and $J(\omega_N)$ (cf. panels d and e of Figure 6). In other words, contribution of these motions to the backbone dynamics is minimal in those sequence regions in which the contribution from slow motions becomes the largest.

The slow backbone dynamics of the whole fragment embracing 4.1 and 4.2 subregions of protonated $\sigma^{70}{}_4$ in solution, and in particular, of the three identified helical fragments, strongly supports the notion that the helical secondary structures become formed cooperatively and exhibit a high propensity to fold into a relatively stable tertiary structure similar to that of the HLHTH DNA binding motif seen in the crystal structures of homologous σ_4 domains.

DISCUSSION

The presented results of CD and NMR experiments and their analysis demonstrated that the ensemble of acid-unfolded states of $\sigma^{70}{}_4$ in aqueous solution retains some of the secondary structural features shown to be present in the crystal structures (*6*, *7*). Characteristic for this structure, the HLHTH folding pattern is clearly reflected in sequential distribution of secondary chemical shifts and vicinal coupling constants, both indicative of preferential population of the α-region of the φ, ψ space. Two of the three helical regions, corresponding to the H1 (L551−M561) and H3 (F580−A594) helices, are particularly highly populated (∼30−40%). The third one between T572 and V575, corresponding to the H2 helix, is less populated (∼10%), reduced to one initial helical turn and followed by a short turn sequence GKQ. The H3 helix is elongated toward the N-terminus by one turn, into an interhelical putative turn region (DVT), and shortened at its C-terminal end.

The whole fragment, embracing the H1, H2, and H3 helices and the intervening loop/turn sequences, is displayed by significantly higher than the average values of the spectral density functions $J(0)$ and $J(\omega N)$ and by smaller than the average values of $J(0.87\omega H)$, indicating a reduced backbone flexibility on the pico- to nanosecond time scale and large contributions from slow motional processes on the micro- to millisecond time scale in the helical regions. The greatly increased values of the $J(0)$ function, with the maxima

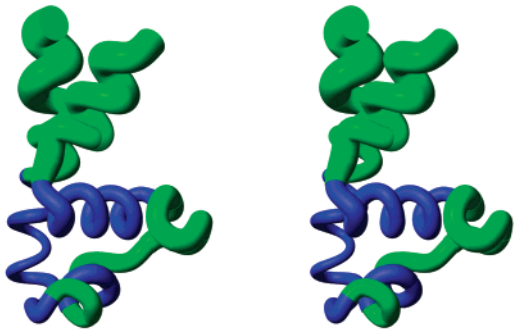

FIGURE 7: Sausage model of $\sigma^{70}{}_4$ based on the crystal structure of TT $\sigma^{70}{}_4$ (*7*). The radius of the tube following the Cα trace of the backbone is proportional to $1/J(0)$, taken as a rough measurement of the backbone conformational mobility. The three helical regions determined from NMR secondary chemical shifts and confirmed by relaxation data for the acid-unfolded protein are marked in blue; unstructured regions in solution are in green.

coincident with middle segments of the helical regions, provide evidence for slow conformational exchange processes involving cooperative formation/dissociation of the helices and hydrophobic contacts between their interfaces. The relaxation data thus strongly suggest that in the acid-unfolded state of $\sigma^{70}{}_4$, the two, H1 and H3, helices tend to be similarly oriented and remain in close contact with one another, as in the crystal structures. Therefore, results of this analysis were included in a sausage model (shown in Figure 7), based on the crystal structure of TT $\sigma^{70}{}_4$, where the radius of the tube following the Cα trace is proportional to the low-frequency $J(0)$ correlation time. This model illustrates in which sequence regions the ensemble-averaged structure of acid-unfolded $\sigma^{70}{}_4$ largely retains the HLHTH-like fold characteristic for the neutral form of the protein. The two helices crossing each other, H1 and H3, marked blue form a universal framework for the spatial organization of the HLHTH DNA binding motif. The N- and C-terminal regions, helical in the crystal state, are unstructured in solution.

The sequences of the $\sigma^{70}{}_4$ domains of *E. coli* and *T. thermophilus* in the helical regions differ one with another only by a few homologous amino acid replacements (cf. Figure 1). In the crystal structures, these amino acids have fully solvent-exposed side chains. Thus, their replacement by homologous amino acids should not disturb the backbone organization. This observation justified the construction of the sausage model on the basis of the TT $\sigma^{70}{}_4$ Cα backbone trace.

By a similar token, the interactions responsible for stabilizing the helical motifs retained in the acid unfolded structure of EC $\sigma^{70}{}_4$ should be generally similar to those between the helices forming the HLHTH fold in the crystal structures. Inspection of interhelical side chain contacts (defined by interatomic distances in the range of 2.5−5 Å) in these structures indicated that the main characteristic feature of these contacts are interspersed hydrophobic clusters and salt bridges, particularly numerous at the H1/H3 interface. The helical structures of the two 4 regions are also locally stabilized by a number of putative $i,i+3$ salt bridges.

The pattern of interactions involved in stabilization of the crystal structures helps to rationalize the conformational and solution properties of $\sigma^{70}{}_4$ in its acid-unfolded state. In this state, deprotonation of carboxylic groups, involved in salt bridging, can be expected to increase the electrostatic

stabilization of the secondary and tertiary structures of σ^{70}_4. The growing contribution from helical structures to the CD spectra of σ^{70}_4 on an increase of the pH of the solution, above the value corresponding to $pK_{trans} = 3.94$, is surely a manifestation of this process. On the other hand, the significant partition of the extended β-structures at pH < 4, accompanied by a dramatic decrease in the protein solubility at its higher concentration, strongly suggest a possibility of formation of an extensive network of intermolecular/intramolecular salt bridges and local hydrophobic clusters facilitating protein aggregation and precipitation, observed under such conditions. The expected structure of fully folded σ^{70}_4 helps also in understanding the greatly increased content of helical forms in its unfolded form in the presence of TFE in solution. It is well-known that solvation by TFE of partially unfolded proteins in mixed aqueous solutions helps to stabilize their residual nativelike, hydrogen-bonded folds by decreasing the solute−water interactions involving H-bonding, charge interactions, and polarizability effects. Hence, the folded motifs become more compact owing to an increased number of intramolecular H-bonds. Usually, maximal stabilization of the conformational ensemble is attained at 30% of TFE (*36*), and a further increase in its content does not influence the observed partition. One can thus expect that partial replacement of water by TFE in the solvation sphere of σ^{70}_4 should drive the equilibrium between unfolded and HLHTH-folded forms in favor of the latter. It is thus highly probable that the buildup of α-helical structures up to 45% content at pH 4.55 in 30% aqueous TFE solution of σ^{70}_4, deduced from analysis of the respective CD spectrum (Figure 3), reflects a large shift in this equilibrium.

The scanty structural NOESY data (Table 1) deserve a comment in connection with the conclusion drawn from the interpretation of the spectral density functions. Most of the unambiguously assigned 20 structural cross-peaks are localized in the regions terminating the adjacent helices. This seems to indicate the occurrence of dynamic and spatial constraints imposed on the corresponding pairs of residues by interaction of adjacent helices, H1/H2 and H2/H3, leading to the stabilization of the whole HLHTH fold. This region is less rigid than the helices themselves, allowing spatial organization of the H1LH2 motif. Several cross-peaks lying in the F563−D570 putative extended loop region cannot be explained by the crystal structure (*7*). In the loop region, as much as eight out of nine residues of σ^{70}_4 are different than in σ^{A}_4 and TT σ^{70}_4 (cf. Figure 1); in particular, none of the three basic residues of σ^{70}_4 is present in the former two σ_4 regions, and the two carboxylic amino acids are located in different regions of the loop. It is thus understandable that conformation of the loop in acid-unfolded σ^{70}_4 also can be expected to be different than in the two crystal structures. It is worth pointing out that in both crystal structures this loop region is also more flexible, as shown by the increased values of the *B* factor for TT σ^{70}_4 (*7*) and lack of order in σ^{A}_4 (*6*).

In this connection, it has to be stressed that in the whole family of σ^{70} proteins the largest divergence between sequences is observed for this loop region (*1*). Our own alignment (not shown) of over 100 protein sequences exhibiting more than 50% identity with the region 4 of σ^{70} demonstrated that the largest sequential heterogeneity was found for the region corresponding to D566−Y571 of σ^{70}_4, with the only consensus of a small residue (G, S, or N) at 568 location. We are thus tempted to conclude that the lack of sequence conservation in the L loop region, and hence its conformational flexibility, is important for functional interactions of the σ_4 domain with the β-subunit flap-tip helix in RNA polymerase holoenzyme (*54*, *55*) and with different transcription regulation factors mapped in the binding hydrophobic cleft of this domain (*2*). This notion is supported by the most recent model of the specific binding of the T_4 anti-σ factor (AsiA) by this domain, formed by the C-terminal helices of the 4.1 and 4.2 subregions, according to which upon binding AsiA must disrupt the interaction of σ_4 with the RNA polymerase β-subunit flap domain (*56*).

Finally, we would like to relate the results of this work with those of the most extensive studies thus far on multiphase acid-induced unfolding of apomyoglobin (ref *8* and references cited). Structural and dynamic NMR studies have provided strong evidence that (i) in the molten globule intermediate at pH 4 (*57*), three highly populated helices are retained, forming a compact rigid core of nativelike topology and (ii) in the fully protonated state of apomyoglobin at pH 2.3, two of these helices, although not exactly similar, still have appreciable helical content and exhibit restricted backbone motions (*58*). Kinetic unfolding experiments at pH 2.7 (*59*) and refolding experiments at neutral pH (*60*) have suggested a strong similarity between the acid unfolding and the refolding pathways. Moreover, in a recent fully atomistic simulation of acid-induced unfolding of apomyoglobin (*61*), cooperativity in secondary and tertiary structure formation was observed. These findings provide strong support to the presented interpretation of the experimental data for acid-unfolded σ^{70}_4 in terms of a residual nativelike average structure of the HLHTH motif, stabilized by concomitant formation of the secondary helical structure and tertiary hydrophobic/hydrogen bonded polar contacts between the helices, undergoing a transition to a more nativelike form on the deprotonation of carboxylates and/or replacement in the solvation layer of the high-dielectric aqueous environment by a more low-dielectric TFE environment, strengthening the network of H-bonds responsible for stabilizing the secondary structures.

ACKNOWLEDGMENT

The authors thank Mrs. Teresa Rak for technical assistance in the preparation of the σ^{70}_4 recombinant protein, Mgr. Jacek Olędzki in mass spectrometric characterization of its purity, Mgr. Magdalena Guzowska for help in the analysis of NMR spectra, and Prof. Andrzej Ejchart for helpful discussions of the relaxation data. Mass spectra of the protein were measured in the Laboratory of the Biological Mass Spectrometry (IBB PAS). Clear-sighted comments of one of the referees, which helped to improve quality of this paper, are greatly acknowledged.

REFERENCES

1. Gross, C., Lonetto, M., and Losick, R. (1992) Bacterial σ factors, in *Transcriptional Regulation* (McKnight, S., and Yamamoto, K., Eds.) pp 129−176, Cold Spring Harbor Laboratory Press, Woodbury, NY.
2. Lonetto, M. A., Rhodius, V., Lamberg, K., Kiley, P., Busby, S., and Gross, C. (1998) Identification of a contact site for different transcription activators in region 4 of the *Escherichia coli* RNA polymerase σ 70 subunit, *J. Mol. Biol. 284*, 1353−65.

3. Gardella, T., Moyle, H., and Susskind, M. M. (1989) A mutant *Escherichia coli* σ 70 subunit of RNA polymerase with altered promoter specificity, *J. Mol. Biol. 206*, 579−90.
4. Reddy, B. V., Gopal, V., and Chatterji, D. (1997) Recognition of promoter DNA by subdomain 4.2 of *Escherichia coli* σ 70: a knowledge based model of −35 hexamer interaction with 4.2 helix−turn−helix motif, *J. Biomol. Struct. Dyn. 14*, 407−19.
5. Poznański, J., Zhukov, I., Bolewska, K., and Wierzchowski, K. L. (2001) ^{1}H, ^{15}N, and ^{13}C resonance assignments for the whole region 4 of *Escherichia coli* RNA polymerase σ 70 subunit, *J. Biomol. NMR.* 20, 181−2.
6. Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L., and Darst, S. A. (2002) Structure of the bacterial RNA polymerase promoter specificity σ subunit, *Mol. Cell. 9*, 527−39.
7. Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., and Yokoyama, S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution, *Nature 417*, 712−9.
8. Dyson, H. J., and Wright, P. E. (2003) Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance, *Adv. Protein Chem. 62*, 311−40.
9. McIntosh, L. P., Grifey, R. H., Muchmore, D. C., Nielson, C. P., Redfield, A. G., and Dahlquist, F. W. (1987) Proton NMR measurements of bacteriophage T_4 lysozyme aided by nitrogen-15 isotopic labeling: Structural and dynamic studies of large proteins, *Proc. Natl. Acad. Sci. U.S.A. 84*, 1244−8.
10. Bradford, M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein−dye binding, *Anal. Biochem. 72*, 248−54.
11. Schagger, H., and von Jakov, G. (1987) Tricine-sodium dodecyl sulfate−polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, *Anal. Biochem. 166*, 368−79.
12. Fasman, G. D., Ed. (1976) *Handbook of Biochemistry and Molecular Biology*, 3rd ed., Vol. 1, CRC Press, Cleveland, OH.
13. Bohm, G., Muhr, R., and Jaenicke, R. (1992) Quantitative analysis of protein far-UV circular dichroism spectra by neural networks, *Protein Eng. 5*, 191−5.
14. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−93.
15. Bartels, C., Xia, T., Billeter, M., Güntert, P., and Wüthrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules, *J. Biomol. NMR 6*, 1−10.
16. Gardner, K. H., Konrat, R., Rosen, M. K., and Kay, L. E. (1996) An (H)C(CO)NH-TOCSY pulse scheme for sequential assignment of protonated methyl groups in otherwise deuterated ^{15}N−^{13}C-labeled proteins, *J. Biomol. NMR 8*, 351−6.
17. Bax, A., and Ikura, M. (1991) An efficient 3-D NMR technique for correlating the proton and ^{15}N backbone amide resonances with the α-carbon of the preceding residue in uniformly ^{15}N/^{13}C enriched proteins, *J. Biomol. NMR 1*, 99−104.
18. Grzesiek, S., and Bax, A. (1992) Correlating Backbone Amide and Side Chain Resonances in Larger Proteins by Multiple Relayed Triple Resonance NMR, *J. Am. Chem. Soc. 114*, 6291−3.
19. Wittekind, M., and Mueller, L. (1993) HNCACB, a High-Sensitivity 3-D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the α- and β-Carbon Resonances in Proteins, *J. Magn. Reson., Ser. B 101*, 201−5.
20. Muhandiram, D. R., and Kay, L. E. (1994) Gradient-Enhanced Triple-Resonance 3-D NMR Experiments with Improved Sensitivity, *J. Magn. Reson., Ser. B 103*, 203−16.
21. Löhr, F., and Rüterjans, H. (1995) A new triple-resonance experiment for the sequential assignment of backbone resonances in proteins, *J. Biomol. NMR 6*, 189−97.
22. Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in ^{13}C/^{15}N-Enriched Proteins by Isotropic Mixing of ^{13}C Magnetization, *J. Magn. Reson., Ser. B 101*, 114−9.
23. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domains of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques *J. Biomol. NMR 4*, 845−58.
24. Billeter, M., Neri, D., Otting, G., Qian, Y. Q., and Wuthrich, K. (1992) Precise vicinal coupling constants $^{3}J_{HNHa}$ in proteins from nonlinear fits of J-modulated [^{15}N,^{1}H]-COSY experiments, *J. Biomol. NMR 2*, 257−74.
25. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelton, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation, *Biochemistry 33*, 5984−6003.
26. Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A., and Torchia, D. (1992) Pulse schemes for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T_1 and T_2 values in proteins, *J. Magn. Reson. 97*, 359−75.
27. Shaka, A. J., Baker, P. B., and Freeman, R. (1985) Computer-optimized decoupling scheme for wideband applications and low-level operation, *J. Magn. Reson. 64*, 547−52.
28. Zhukov, I., and Ejchart, A. (1999) Factors improving the accuracy of determination of ^{15}N relaxation parameters in proteins, *Acta Biochim. Pol. 46*, 665−71.
29. Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Kay, L. E. (1995) Spectral density function mapping using ^{15}N relaxation data exclusively, *J. Biomol. NMR 6*, 153−62.
30. Higgins, D., Thompson, J., Gibson, T., Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice, *Nucleic Acids Res. 22*, 4673−80.
31. Henikoff, S., and Henikoff, J. G. (1992) Amino acid substitution matrixes from protein blocks, *Proc. Natl. Acad. Sci. U.S.A. 89*, 10915−19.
32. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−5.
33. Yaffee, R. A., and McGee, M. (2000) *An Introduction to Time Series Analysis and Forecasting: with Applications of SAS and SPSS*, Academic Press, London.
34. Fioroni, M., Diaz, M. D., Burger, K., and Berger, S. (2002) Solvation phenomena of a tetrapeptide in water/trifluoroethanol and water/ethanol mixtures: a diffusion NMR, intermolecular NOE, and molecular dynamics study, *J. Am. Chem. Soc. 124*, 7737−44.
35. Diaz, M. D., Fioroni, M., Burger, K., and Berger S. (2002) Evidence of complete hydrophobic coating of bombesin by trifluoroethanol in aqueous solution: an NMR spectroscopic and molecular dynamics study, *Chemistry 8*, 1663−9.
36. Jaravine, V. A., Alexandrescu, A. T., and Grzesiek, S. (2001) Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide, *Protein Sci. 10*, 943−50.
37. Lawrence, J. R., and Johnson, W. C. (2002) Lifson−Roig nucleation for α-helices in trifluoroethanol: context has a strong effect on the helical propensity of amino acids, *Biophys. Chem. 101−102*, 375−85.
38. Spera, S., and Bax, A. (1991) Empirical correlation between protein backbone conformation and Cα and Cβ ^{13}C nuclear magnetic resonance chemical shifts, *J. Am. Chem. Soc. 113*, 5490−2.
39. Wishart, D. S., and Sykes, B. D. (1994) The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data, *J. Biomol. NMR 4*, 171−80.
40. Schwarzinger, S., Kroon, G. J., Foss, T. R., Wright, P. E., and Dyson, H. J. (2000) Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMR view, *J. Biomol. NMR 18*, 43−8.
41. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) ^{1}H, ^{13}C, and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects, *J. Biomol. NMR 5*, 67−81.
42. Merutka, G., Dyson, H. J., and Wright, P. E. (1995) "Random coil" ^{1}H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG, *J. Biomol. NMR 5*, 14−24.
43. Schwarzinger, S., Kroon, G. J., Foss, T. R, Chung, J., Wright, P. E, and Dyson, H. J. (2001) Sequence-dependent correction of random coil NMR chemical shifts, *J. Am. Chem. Soc. 123*, 2970−8.
44. Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M., and Dobson, C. M. (1996) Analysis of main chain

torsion angles in proteins: prediction of NMR coupling constants for native and random coil conformations, *J. Mol. Biol. 255*, 494−506.
45. Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease, *Biochemistry 28*, 8972−9.
46. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. theory and range of validity, *J. Am. Chem. Soc. 104*, 4546−59.
47. Alexandrescu, A. T., and Shortle, D. (1994) Backbone dynamics of a highly disordered 131 residue fragment of staphylococcal nuclease, *J. Mol. Biol. 242*, 527−46.
48. Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A., and Kay, L. E. (1995) Spectral density function mapping using ^{15}N relaxation data exclusively, *J. Biomol. NMR 6*, 153−62.
49. Wong, K. B., Clarke, J., Bond, C. J., Neira, J. L., Freund, S. M., Fersht, A. R., and Daggett, V. (2000) Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding, *J. Mol. Biol. 296*, 1257−82.
50. Peng, J. W., and Wagner, G. (1992) Mapping of spectral density functions using heteronuclear NMR relaxation measurements, *J. Magn. Reson. 98*, 308−32.
51. Meekhof, A. E., and Freund, S. M. (1999) Separating the contributions to ^{15}N transverse relaxation in a fibronectin type III domain, *J. Biomol. NMR 14*, 13−22.
52. Ran, X., Miao, H. H., Sheu, F. S., and Yang, D. (2003) Structural and dynamic characterization of a neuron-specific protein kinase C substrate, neurogranin, *Biochemistry 42*, 5143−50.
53. Garcia, P., Serrano, L., Durand, D., Rico, M., and Bruix, M. (2001) NMR and SAXS characterization of the denatured state of the chemotactic protein CheY: implications for protein folding initiation, *Protein Sci. 10*, 1100−12.
54. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., and Darst, S. A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme−DNA complex, *Science 296*, 1285−90.
55. Murakami, K. S., Masuda, S., and Darst, S. A. (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution, *Science 296*, 1280−4.
56. Simeonov, M. F., Bieber Urbauer, R. J., Gilmore, J. M., Adelman, K., Brody, E. N., Niedziela-Majka, A., Minakhin, L., Heyduk, T., and Urbauer, J. L. (2003) Characterization of the interactions between bacteriophage T_4 AsiA protein and RNA polymerase, *Biochemistry 42*, 7717−26.
57. Eliezer, D., Chung, J., Dyson, H. J., and Wright, P. E. (2000) Native and non-native secondary structure and dynamics in the pH 4 intermediate of apomyoglobin, *Biochemistry 39*, 2894−901.
58. Yao, J., Chung, J., Eliezer, D., Wright, P. E., and Dyson, H. J. (2001) NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding, *Biochemistry 40*, 3561−71.
59. Jamin, M., Yeh, S. R., Rousseau, D. L., and Baldwin, R. L. (1999) Submillisecond unfolding kinetics of apomyoglobin and its pH 4 intermediate, *J. Mol. Biol. 292*, 731−40.
60. Tsui, V., Garcia, C., Cavagnero, S., Siuzdak, G., Dyson, H. J., and Wright, P. E. (1999) Quench-flow experiments combined with mass spectrometry show apomyoglobin folds through an obligatory intermediate, *Protein Sci. 8*, 45−49.
61. Onufriev, A., Case, D. A., and Bashford, D. (2003) Structural details, pathways, and energetics of unfolding apomyoglobin, *J. Mol. Biol. 325*, 555−67.

BI034943V