Prothymosin α: A Biologically Active Protein with Random Coil Conformation†

Klaus Gast,‡ Hilde Damaschun,§ Klaus Eckert,|| Kai Schulze-Forster,|| H. Rainer Maurer,|| Marlies Müller-Frohne,‡ Dietrich Zirwer,§ Jan Czarnecki,⊥ and Gregor Damaschun*,§

Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse 10, D-13122 Berlin, Germany, Institute of Biology, Humboldt-University at Berlin, c/o MDC, Robert-Rössle-Strasse 10, D-13122 Berlin, Germany, Institute of Pharmacy, Free University of Berlin, Kelchstrasse 31, D-12169 Berlin, Germany, and Thymoorgan GmbH, Schiffgraben 23, D-38690 Vienenburg, Germany

Received April 21, 1995; Revised Manuscript Received August 7, 1995⊗

ABSTRACT: Prothymosin is an acidic protein with an unusual amino acid composition. Though its exact function is not yet known, its high evolutionary conservation and wide tissue distribution suggest an essential biological role. Its physical state, which is controversially discussed in previous publications, was investigated using small-angle X-ray scattering, dynamic light scattering, mass spectrometry, and circular dichroism (CD). Our results unequivocally demonstrate that prothymosin is a monomer under physiological conditions. The protein adopts a random coillike conformation but exhibits persistence of direction and curvature. No regular secondary structure is detectable by CD. The Stokes radius, R_S = 3.07 nm, and the radius of gyration, R_G = 4.76 nm, are 1.77 and 3.42 times larger, respectively, than those expected for a compactly folded protein consisting of 109 amino acid residues. A remarkable amount of secondary structure is formed only in the presence of trifluoroethanol at low pH. The finding that a biologically active protein molecule with 109 amino acid residues adopts a random coil conformation under physiological conditions raises the question whether this is a rare or a hitherto-overlooked but widespread phenomenon in the field of macromolecular polypeptides.

Prothymosin α (ProTα)[1] is a highly acidic protein (pI = 3.5) first isolated from rat thymus (Haritos et al., 1984a). Initially, ProTα was called a "thymic hormone", but its high evolutionary conservation (Makarova et al., 1989; Frillingos et al., 1991) and its wide tissue distribution (Haritos et al., 1984b; Economou et al., 1988) are inconsistent with the hormone hypothesis and suggest a more general role. The exact function of ProTα remains unclear, although evidence has been accumulated for its intracellular as well as extracellular role. ProTα contains a nuclear-targeting sequence, is transported into the nucleus (Manrow et al., 1991), and binds to histone protein H1 (Papamarcaki & Tsolas, 1994). The central stretch of amino acids shows some homology to the high mobility group 1 protein (Bustelo et al., 1991). It has been proposed that ProTα plays a similar role in the nucleus and may be required to open chromatin structure (Szabo & Weksler, 1992). Treatment of myeloma cells with ProTα antisense oligomers inhibited cell division (Sburlati et al., 1991). In cancer tissues, an increased concentration of ProTα was reported (Tsitsiloni et al., 1993). These facts as well as the activation of the ProTα gene by the *c*-myc protein (Eilers et al., 1991) indicate an association of ProTα with cell proliferation.

In human blood, the total quantity of ProTα was estimated as 11−14 pmol/mL. About 90% was recovered in the leukocyte fraction (Panneerselvam et al., 1987), and only trace amounts (1−2%) were found in erythrocytes. The extracellular effects are of immunogenic nature and may be mediated by recently described high- and low-affinity binding sites on human peripheral blood mononuclear cells (Cordero et al., 1994). After PHA induction and ProTα treatment, these cells show an enlarged proliferation by an increased IL-2 receptor expression (Cordero et al., 1991). ProTα enhances the cytotoxicity of natural killer cells (Cordero et al., 1992a) and restores depressed allogenic cell-mediated lympholysis and natural killer cell activity in cancer patients (Baxevanis et al., 1993). ProTα normalizes deficient antitumor activity of monocytes from melanoma patients in vitro (Garbin et al., 1994). Also, for the N-terminal fragment 1−28 (thymosin α_1), immunological effects have been reported, e.g., the improvement of depressed lymphokine-activated killer cells from immunodeficient patients (Eckert et al., 1994).

ProTα from calf thymus contains 109 amino acids with 49% aspartic and glutamic acid, whereas aromatic and sulfur amino acids are totally absent. Instead of the sequence, we have shown its remarkable charge distribution in Figure 7.

The physical state of ProTα in solution is a controversial issue. Gel filtration experiments at neutral pH resulted in an apparent molar mass that is 5 times greater than that calculated from the amino acid sequence (11 983 g/mol) (Haritos et al., 1984a, 1989). On the other hand, sedimentation equilibrium measurements under identical conditions gave a molar mass of 12 800 g/mol, a value being nearly that expected for the monomeric ProTα molecule (Haritos et al., 1989).

† Supported by grants from the Deutsche Forschungsgemeinschaft (Da292/1-2 and Da292/1-3) and from the Bundesministerium für Forschung und Technologie (01KB 9104/1) and by a grant of the Fonds der chemischen Industrie to G.D.

‡ Max-Delbrück-Center for Molecular Medicine.
§ Humboldt-University at Berlin.
|| Free University of Berlin.
⊥ Thymoorgan GmbH.

⊗ Abstract published in *Advance ACS Abstracts,* September 15, 1995.

[1] Abbreviations: ProTα, prothymosin α; SAXS, small-angle X-ray scattering; PBS, phosphate-buffered saline; TFE, trifluoroethanol.

0006-2960/95/0434-13211$09.00/0

Watts et al. (1990) have inferred from proton NMR and circular dichroism studies that ProTα at neutral pH might exist as a disordered chain molecule and elutes earlier than expected for a globular protein of the same molar mass in exclusion chromatography at pH 7.

However, Cordero et al. (1992b) have recently reported on results from gel filtration and sodium dodecyl sulfate (SDS) electrophoresis of ProTα. From gel filtration at pH 7, they obtained an apparent molar mass of 59 000 g/mol in accordance with the value found by Haritos et al. (1989). In sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) under denaturing conditions of high temperature, SDS, and β-mercaptoethanol, ProTα exhibited a molar mass of 20 600 g/mol. No conditions were found which resulted in a 12 000 g/mol band. Therefore, Cordero et al. (1992b) conclude that ProTα exists as a stable dimer.

To investigate the conformation of ProTα in solution at neutral pH in terms of geometric parameters which cannot be derived from spectroscopic techniques or hydrodynamic measurements alone as well as to elucidate the state of association of ProTα, we have combined the methods of small-angle X-ray scattering (SAXS), static and dynamic light scattering, and circular dichroic spectrometry (CD). Conformational changes in dependence on pH as well as those effected by the helix-promoting agent trifluoroethanol (TFE) were monitored by changes in the circular dichroic spectra. Our results show that ProTα is a monomeric protein under physiological conditions, having the conformation of a random coil with persistence of direction and curvature. As judged by CD criteria, elements of regular secondary structure are absent at pH 7.4 even in the presence of 2 M NaCl or bivalent cations. However, decreasing of pH to pH = 2.4 induces some secondary structure. This effect can be considerably enhanced by addition of TFE.

MATERIALS AND METHODS

Materials. ProTα was a gift from Thymoorgan Pharmazie & Co. KG (Vienenburg, Germany). It was extracted from calf thymus according to the method of Haritos et al. (1984a). The investigations under physiological conditions were performed in PBS, pH 7.4. The measurements at low pH were done in 10 mM sodium acetate buffer, pH 4.6, and in 10 mM glycine/HCl buffer, pH 2.4, respectively. TFE, NMR grade, was from Aldrich.

Protein Concentrations. Stock solutions of defined protein concentration were produced by the dissolving of a certain amount of freeze-dried protein in an appropriate buffer volume. Protein concentrations after subsequent dilution and filtration were determined spectrophotometrically by the measuring of the absorbances at 215 and 225 nm. From measurements on stock solutions, we have estimated $A_{1\,cm}^{0.1\%}$ (225 nm) = 2.75 and $A_{1\,cm}^{0.1\%}$ (215 nm) = 7.32.

Circular Dichroism. CD spectra were obtained on a J-720 spectrometer (JASCO, Japan) using 1 mm cells if not otherwise stated. The sample temperature was held at 20 °C by a temperature control system (NESLAB, USA). Mean residue ellipticities [Θ] were calculated using a value of 109.6 for the mean residue weight of ProTα. The spectrometer was calibrated with (+)-10-camphorsulfonic acid (Johnson, 1990).

Mass Spectrometry. Mass spectra of ProTα were acquired on a TSQ 700 triple stage quadrupole mass spectrometer (Finnigan MAT, Germany) equipped with an electrospray ion source. ProTα was dissolved in methanol/water (1:1) containing 1% acetic acid (final protein concentration, 4 pmol/μL) and introduced into the ion source by a microsyringe pump (Harvard apparatus) at a flow rate of 0.5 μL/min.

Light Scattering. Static and dynamic light scattering investigations were performed using a spectrometer which has been described previously (Gast et al., 1992). All measurements were done at 20 °C, at a scattering angle of 90°, and at the wavelength 514.5 nm of an argon laser operating at 1 W output power. The solvent and the protein solutions were filtered through 100 nm pore-size filters (Protein Solutions Ltd., U.K.) directly into 100 μL flow-through cells (Hellma, Germany).

The molar mass of the protein was derived from the measured time-integrated scattered intensity using benzene as a reference sample. The homodyne time−autocorrelation functions of the scattered intensity, $G^2(\tau)$, were calculated by a 90-channel multibit multiple-τ correlator and then fed into an on-line-coupled PC equipped with a transputer board, ALV-800 (ALV Laser-Vertriebsgesellschaft mbH, Germany), for data evaluation using the program CONTIN (Provencher, 1982). The translational diffusion coefficients D obtained from the autocorrelation functions were converted into Stokes radii, R_S, via the Stokes−Einstein equation, $R_S = kT/(6\pi\eta_0 D)$, where k is Boltzmann's constant, T is the temperature in Kelvin, and η_0 is the solvent viscosity.

Small-Angle X-ray Scattering. SAXS measurements were performed on an X-ray diffractometer URD6 (Freiberger Praezisionsmechanik, Germany) at 20 °C. The device is equipped with a special slit system for small-angle measurements. The incident radiation is collimated by two slits. A third slit is positioned in front of the sample to shield it from the radiation scattered by the collimator slits. The radiation scattered by the sample is collimated by another pair of slits on the detector arm between which a Soller-slit system is mounted. The slit widths are adjusted so that all sampling points of a scattering curve from particles smaller than 60 nm can be collected. The scattered Cu Kα radiation is measured by a proportional counter in step−scan mode. The absolute scattered intensity was calibrated by scattering experiments using two proteins of known molar mass, namely apocytochrome *c* (Damaschun et al., 1991) and apomyoglobin (Gast et al., 1994). Data handling, including desmearing and calculation of parameters and functions, is carried out by means of a program written by Müller et al. (1986).

RESULTS

Circular Dichroism. The CD spectrum of ProTα in PBS, pH 7.4, is shown in Figure 1. It has a pattern typical of a protein lacking regular secondary structure as it is observed usually in solutions containing high concentrations of guanidinium chloride or urea. In order to check whether the protein can be transformed into a partly folded conformation when strong electrostatic repulsion is reduced by addition of salt, we have performed measurements in PBS, pH 7.4, containing additionally 2 M NaCl. But, the measured spectrum was indistinguishable from that shown in Figure 1. Bivalent cations Mg^{2+}, Ca^{2+}, and Zn^{2+} in 250-fold (Mg^{2+} and Ca^{2+}) or 100-fold (Zn^{2+}) molar excess of protein

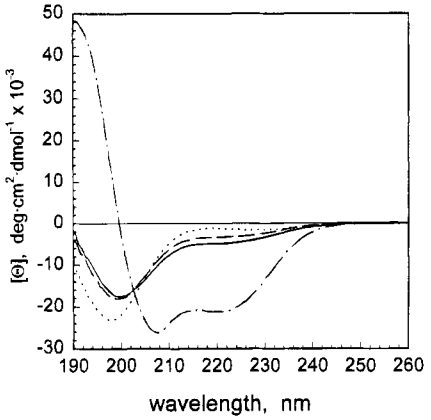

FIGURE 1: CD spectra of ProTα measured at 20 °C in different buffers in cells with pathlength d. - - -, PBS, pH 7.4, $d = 0.1$ mm, $c = 0.89$ g/L; — —, 10 mM sodium phosphate, pH 4.6, $d = 1$ mm, $c = 0.22$ g/L; —, 10 mM glycine/HCl, pH 2.4, $d = 1$ mm, $c = 0.24$ g/L; — • —, 5 mM glycine/HCl, pH 2.4, 50% v/v TFE, $d = 1$ mm, $c = 0.12$ g/L.

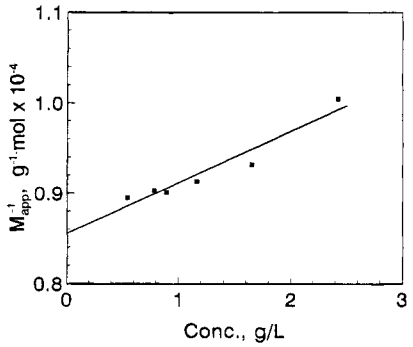

FIGURE 2: Concentration dependence of the reciprocal apparent molar mass of ProTα at 20 °C in PBS, pH 7.4. The straight line results from a linear least-squares fit.

concentration were without any influence on the CD spectra. On the other hand, a tendency to form secondary structure is perceptible when the pH is lowered. This can be deduced from the spectra measured at pH 4.6 and 2.4. The amount of regular secondary structure is considerably increased when TFE is added. This increase is most pronounced at low pH and saturates at about 50% v/v TFE. The CD spectrum of ProTα at pH 2.4 and 50% v/v TFE is also shown in Figure 1. It has a pattern typical of a protein having an appreciable content of helical structure. Taking the mean residue ellipticity at 222 nm, $[\Theta]_{222}$, as a measure for the percentage of helical structure [-1000 deg·cm²·dmol^{-1} for 0% and $-30\,000$ deg·cm²·dmol^{-1} for 100% helical structure (Chen et al., 1974)], we find ~0% at pH 7.4, 8% at pH 4.6, 13% at pH 2.4, and 69% at pH 2.4 in the presence of 50% v/v TFE.

Mass Spectrometry. Electrospray ionization mass spectrometry investigations yielded a single peak with an average molar mass of 11 982.8 g/mol.

Static and Dynamic Light Scattering. Measurements of the excess scattered intensity of protein solutions at finite concentrations yield apparent molar masses, M_{app}. The molar mass can be estimated by a linear extrapolation to 0 protein concentration according to $1/M_{app} = 1/M + 2Bc$, where B is the second virial coefficient. The corresponding data for ProTα in PBS, pH 7.4, are shown in Figure 2. A linear least-squares fit yields $M = (11\,700 \pm 1200)$ g/mol and a second virial coefficient $B = (2.84 \pm 0.30) \times 10^{-3}$ mol·cm³/g². The translational diffusion coefficients measured at protein concentrations between 0.5 and 2.5 g/L have been used for an extrapolation to 0 protein concentration to get $D^0_{20,w}$. From $D^0_{20,w} = (6.99 \pm 0.10) \times 10^{-7}$ cm²/s, we have calculated a Stokes radius, $R_S = (3.07 \pm 0.04)$ nm.

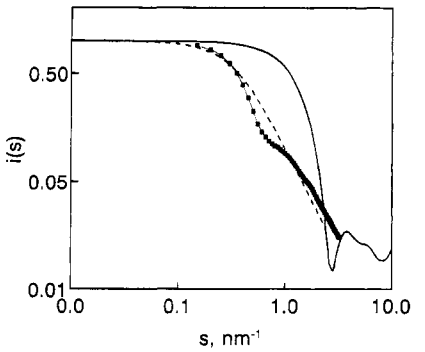

FIGURE 3: Experimental X-ray scattering curves of ProTα (■) (109 amino acids) and cytochrome c (—) (horse heart, 104 amino acids). — —, calculated scattering curve of a Gaussian coil in the Debye approximation (eq 1, $R_G = 4.76$ nm). The average experimental error is about twice the size of the symbols. The same holds for Figures 4 and 5.

Small-Angle X-ray Scattering. SAXS from ProTα in PBS was measured at protein concentrations between 2 and 5 g/L. Each individual scattering curve was desmeared. The molar mass calculated from the scattered intensity, $I(0)$, extrapolated to 0 protein concentration amounts to $M = (12\,000 \pm 1500)$ g/mol. Figure 3 shows the scattering curve for the sample with $c = 4.94$ g/L within the range 0.14 nm$^{-1} \leq s \leq 4.3$ nm^{-1}, where $s = (4\pi/\lambda) \sin \theta$ is the scattering vector. In addition, we show the scattering curves for the compact globular protein cytochrome c (Damaschun et al., 1991) consisting of 104 amino acids and for a random coil which was calculated by means of Debye's formula

$$I(s) = \frac{2}{\langle R^2_G \rangle^2 s^4}[\exp(-\langle R^2_G \rangle s^2) - 1 + \langle R^2_G \rangle s^2] \quad (1)$$

The calculated curve for a random coil was fitted to the experimental data by a least-squares algorithm with the root-mean-square radius of gyration, $\langle R^2_G \rangle^{0.5}$, as the free parameter. The close similarity between the experimental curve and the calculated random coil scattering curve is clearly visible, while the scattering curve typical of a globular protein is significantly different. The radii of gyration, R_G, were also determined for each protein concentration by the Guinier approximation (Figure 4)

$$I(s) = I(0) \exp\left[-\frac{s^2 R^2_G}{3}\right] \quad (2)$$

The resulting radii of gyration were linearly extrapolated to 0 protein concentration. We obtained from the extrapolation $R_G = (4.76 \pm 0.1)$ nm. Furthermore, the scattering curve is shown in the form of a Kratky plot (Figure 5). This scattering curve exhibits the characteristics typical of random coil molecules with persistence of direction and persistence of curvature (Kirste & Oberthür, 1982). Recently, similar scattering curves have been measured by us for acid-denatured apocytochrome c (Damaschun et al., 1991), cold-denatured phosphoglycerate kinase (Damaschun et al., 1993b), and acid-denatured apomyoglobin (Gast et al., 1994)

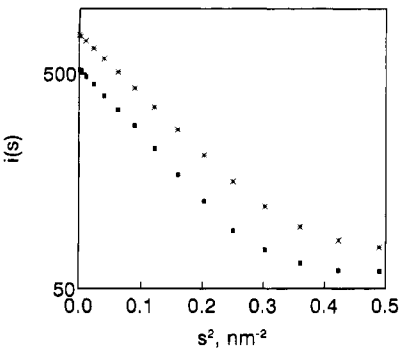

FIGURE 4: Guinier plots of the X-ray scattering curves of ProTα. *, c = 2.65 g/L; ■, c = 4.94 g/L. For clarity, the plots have been arbitrarily shifted along the ln I axis.

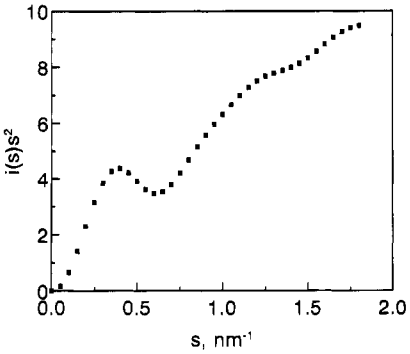

FIGURE 5: Kratky plot of the experimental X-ray scattering curve of ProTα.

and by Schweers et al. (1994) for τ protein from Alzheimer brain tissue.

The persistence length was determined by the following procedure. The scattering curve was approximated for $s <$ 0.9 nm^{-1} by eq 1 by the performing of a nonlinear least-squares fit. For $s >$ 1.2 nm^{-1}, the scattering curve was approximated by that for a thin rod $I \propto s^{-1}$. The persistence length a = (1.87 ± 0.2) nm was calculated from the point of intersection of the two curves s^* via $a = 1.91/s^*$.

DISCUSSION

The results obtained in this study allow for a clear decision regarding the question whether or not ProTα exists as a monomer in solution. Mass spectrometry of ProTα was performed in order to determine the exact molar mass of the protein used for these investigations. The apparent molar mass of 11 982.8 g/mol agrees very well with that calculated from the published amino acid sequence of bovine ProTα (11 983 g/mol). The fact that ProTα is indeed in the monomeric state in solution is undoubtedly demonstrated by the fact that our light scattering and small-angle X-ray scattering results yield a molar mass of 11 700 and 12 000 g/mol, respectively. Light scattering is extremely sensitive to protein association because the measured molar mass is a weight average. Thus, the measured molar mass, M = 11 700 g/mol, gives clear evidence of the monomeric state of the protein. Furthermore, we could not detect any tendency to form multimeric structures within the investigated concentration range between 0.5 and 2.5 g/L. The controversial findings regarding the state of association of ProTα in solution result from the presumption that ProTα adopts a folded conformation. This is most striking in the case of gel filtration experiments where the peak position of ProTα is related to those of folded globular proteins (Haritos et al., 1984a, 1989; Cordero et al., 1992b). The contradictions between different experimental results disappear if one takes into account that ProTα is unfolded under physiological conditions and exhibits a random coil conformation. The first evidence of this remarkable fact came from the NMR and CD studies of Watts et al. (1990), which revealed the absence of and only little secondary structure at neutral and low pH, respectively. However, they have mentioned that their preparative procedure yields denatured products. In our investigations, we used ProTα prepared according to Haritos et al. (1984a), which was biologically active. Additionally, we have performed CD measurements in order to study the influence of salt (2 M NaCl), bivalent ions, and TFE on the conformation of ProTα. Only the presence of TFE leads to a considerable increase in secondary structure. This is most pronounced at low pH.

Normally, proteins need a specific three-dimensional structure in order to perform their particular function. Either ProTα adopts such a specific structure under cellular conditions, for example, caused by interactions with membranes or receptor molecules, or its nonfolded conformation might be advantageous for its interaction with other biological macromolecules. Therefore, it is worthwhile to characterize the unfolded state of ProTα in more detail in terms of geometric dimensions and flexibility. This can be done on the basis of SAXS data and the combination of SAXS and dynamic light scattering results. The radius of gyration and the Stokes radius of monomeric ProTα are R_G = (4.76 ± 0.1) nm and R_s = (3.07 ± 0.04) nm, respectively. Thus, the ϱ-factor, $\varrho = R_G R_s^{-1}$, amounts to ϱ = 1.55 ± 0.05. The ϱ-factor is 0.8 for compact proteins. The experimentally determined value of ϱ = 1.55 is typical of unfolded proteins and very similar to ϱ = 1.51 calculated for a random coil under θ conditions (Burchard et al., 1980).

According to the scaling laws $R_G = 0.29N^{1/3}$ and $R_s = 0.362N^{1/3}$ (Damaschun et al., 1993a), one would expect values of R_G = 1.39 nm and R_s = 1.73 nm for a compactly folded protein with N = 109 amino acids. The experimentally determined values are 3.42 and 1.77 times larger, respectively. In the light of these findings, the results of previous sedimentation equilibrium measurements (Haritos et al., 1989) and gel filtration experiments (Haritos et al., 1984a, 1989; Cordero et al., 1992b) are no longer contradictory. Because the experimental Stokes radius is larger by a factor of 1.77 than expected for a folded protein, ProTα should behave like a protein of the 5.6-fold molar mass in gel filtration experiments. This was indeed observed. Haritos et al. (1989) found a ratio of the relative molar mass from gel filtration experiments to the calculated molar mass of 5.1.

The finding that ProTα adopts a random chain conformation under physiological conditions is corroborated by comparison of the experimental scattering curve with scattering curves of models (Figures3 and 5). The experimental curve is very similar to that for a random coil but totally different from the scattering curve for a compact, folded protein. Figure 5 shows that ProTα is a random coil with persistence of direction. In order to test whether the distribution function of its end-to-end distances can be approximated by a Gaussian distribution, we compare the independently determined values of the persistence length a = (1.87 ± 0.3) nm and of the radius of gyration R_G = (4.76

± 0.1) nm. Benoit and Doty (1953) have shown that, for random coils with Gaussian distribution holds

$$R_G = \frac{1}{3}na^2(1 - 3/n + 6/n^2 - 6/n^3 + e^{-n}6/n^3) \quad (3)$$

$n = La^{-1}$ is the ratio of the length L of the polypeptide chain to its persistence length. From our data, we obtain $n = 19.93 \approx 20$. Therefore, the Benoit–Doty equation yields $R_G = 4.48$ nm. This value is only somewhat smaller than the experimental one.

The mean-square end-to-end distance of a sufficiently large random coil is given by

$$\langle h^2_{ee} \rangle = 6\langle R^2_G \rangle \quad (4)$$

From this relation, we obtain $h_{ee} = 11.66$ nm.

According to Kratky and Porod (1949), for a random coil obeying persistence statistics holds

$$\langle h^2_{ee} \rangle = 2na^2\left(1 - \frac{1}{n} + \frac{e^{-n}}{n}\right) \quad (5)$$

Taking our values of n and a (see above), this relation yields $h_{ee} = 11.5$ nm. This result agrees within the margins of error with that directly calculated from the radius of gyration using eq 4. These calculations demonstrate that the experimentally determined values of the persistence length, the radius of gyration, and the Stokes radius are consistent when the conformation of ProTα is assumed to be that of a random coil with the statistical distribution function of a Gaussian coil with persistence of direction.

The length of one statistical chain segment $l = 2a = 3.74$ nm corresponds to 10 amino acids. The same value was found by us for apocytochrome c at acidic pH ($a = 1.81$ nm), for cold-denatured yeast phosphoglycerate kinase ($a = 1.74$ nm), and for further 11 proteins unfolded by 6 M guanidinium chloride ($a = 1.7$ nm) (Damaschun et al., 1991, 1993b, 1995) and by Schweers et al. (1994) for τ protein ($a = 1.9$ nm). The oscillations of the experimental scattering curve about the Debye scattering curve for the random coil model (Figures 3 and 5) indicate that the ProTα molecules exhibit persistence of curvature in addition to persistence of direction. This is to be expected for a polypeptide chain consisting of L-amino acids (Damaschun et al., 1993a).

In order to give a clear idea of the geometric expansion of the ProTα molecule, we compare in Figure 6 the distribution function of the radii of gyration calculated from the experimental data (Damaschun et al., 1995; Flory, 1988) with the radius of gyration of a compact protein consisting of 109 amino acids and with the radius of gyration of a rodlike molecule having the contour length of the polypeptide chain, respectively.

The calculated radius of gyration of a polypeptide chain with the amino acid sequence of ProTα amounts to $R_G = 4.04$ nm under θ conditions. The experimental value is larger by the factor $\alpha = 1.18$, indicating clearly that repulsive forces between the amino acids prevail over attractive forces under physiological conditions. The Flory parameter C (Flory, 1988) is another index of the polypeptide chain stiffness, and $C = 8.64$ for ProTα.

It is a widespread procedure to apply secondary structure prediction methods when the three-dimensional structure of a protein is unknown or is not accessible for experimental reasons. Secondary structure predictions have also been done for ProTα (Haritos et al., 1985; Schmidt & Werner, 1991) in order to derive structure–function relationships. The average accuracy of such predictions is about 60%, but it depends strongly on the particular protein. Frequently, joint prediction methods are used (Nishikawa & Noguchi, 1991). Taking into account that ProTα is not folded under physiological conditions, the adequate use of prediction methods is questionable. Usually, these prediction methods have been tested or even have been trained with proteins, which are folded into a compact conformation. Nevertheless, we have applied some well-established methods using (a) physicochemical properties, ALB (Ptitsyn & Finkelstein, 1983), (b) statistical information, CF (Chou & Fasman, 1974), GOR (Garnier et al., 1978), SIMPA (Levin & Garnier, 1988), and (c) multilayered neuronal networks incorporating evolutionary information, PHD (Rost & Sander, 1993). The results are divided into three structural classes: helix, extended structure (β-strand), and loop structure. Environmental conditions such as pH and ionic strength can be taken into account in the case of ALB (but these are often set to standard conditions in particular software packages). The results are shown in Figure 7. In the case of ProTα, the coincidence of the results of the different prediction methods at particular positions is poor in general. On average, the content of helical structure is overestimated by all methods compared to that derived from experimental data. The

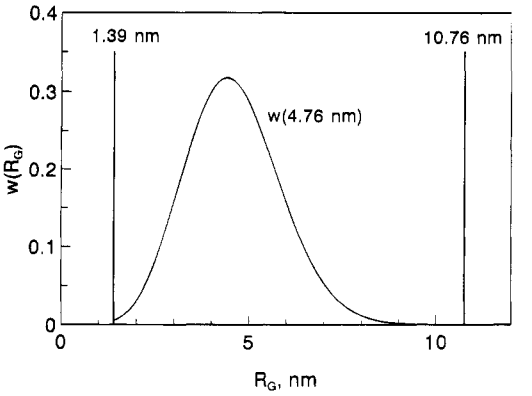

FIGURE 6: Distribution functions of the radius of gyration for ProTα in PBS, pH 7.4, with a radius of gyration $R_G = 4.76$ nm, for a compactly folded protein with 109 amino acids and $R_G = 1.39$ nm, and for a stiff β-chain with 109 amino acids and $R_G = 10.76$ nm.

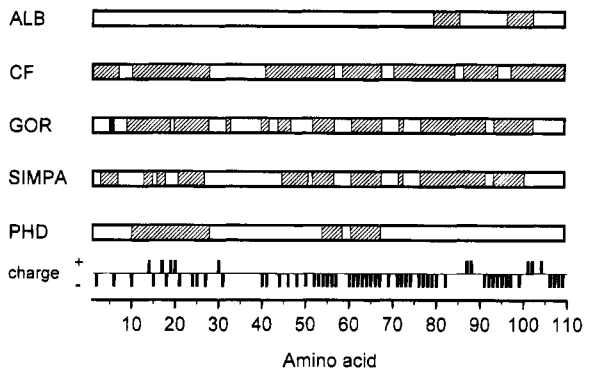

FIGURE 7: Results of secondary structure prediction of five different methods for ProTα. The predicted secondary structure classes are as follows: helix, hatched bars; extended (β-) structure, black bars; loop, empty bars.

present example shows that methods predicting secondary structure should be cautiously applied and that irrelevant results may be obtained. In any case, secondary structure predictions should only be performed in connection with experimental techniques (e.g., CD), which allow for checking, whether the protein forms secondary structure in solution. The predictions for ProTα may be helpful to identify regions of the polypeptide chain, which have a high propensity to form secondary structure under particular conditions such as low pH or in the presence of TFE. Shiraki et al. (1995) have measured the TFE-induced conformational transitions of more than 20 proteins of various secondary structural types. They found for proteins with low helical content in the native state that the helical content in TFE was correlated more with the helical content predicted by secondary structure prediction methods than with the helical content in the native state. Furthermore, it was proposed (Thomas & Dill, 1993; Shiraki et al., 1995) that the effects of TFE on proteins could be explained as mainly weakening the nonlocal hydrophobic interactions and slightly enhancing local helical interactions. Because hydrophobic interactions are extremely weak in ProTα, our results underline the importance of the second aspect. Accordingly, ProTα is an interesting molecule to study general problems of protein folding.

Prothymosin α is an unusual polypeptide with respect to its high content of acidic amino acids, its low content of hydrophobic amino acids, and the absence of aromatic and sulfur-containing amino acids (Haritos et al., 1985). To our knowledge, there exists only one further protein, namely τ protein, which is involved in Alzheimer's disease (Schweers et. al., 1994), consisting of more than 100 amino acids that exhibits the conformation of a random coil in aqueous solution under physiological conditions. Usually, "normal" proteins adopt this conformation only in solvents containing 6 M GuHCl (Damaschun et al., 1995), at extreme pH values (Damaschun et al., 1991), or in the cold-denatured state (Damaschun et al., 1993b). ProTα lacks α-helices, β-hairpins, or any other regular folded structures. Each molecule adopts different conformations in time so that the distance distributions between two atoms result in a Gaussian distribution function. The φ,ψ-angles between more than 90% of the amino acids have values corresponding to that of an extended chain and less than 10% values corresponding to that of a right-handed α-helix (Flory et al., 1988). These two most frequent φ,ψ-values are randomly distributed along the polypeptide chain and are permanently changing in time. Repulsive forces between charged amino acids effect some expansion of the coil compared to the θ state.

α-Helical segments form only at low pH and especially in the presence of TFE, comprising 75 of the 109 amino acids of the molecule at most. Helical secondary structure in the presence of micelles supports the proposal that the TFE-induced structures of polypeptides are not an artifact of their environment but an indication of the conformation that the molecule might adopt in close proximity to the membrane surface and possibly when it is bound to the receptor.

Structures similar to that observed by us for ProTα, viz. α-helical segments jointed by residues with flexible conformation in the presence of TFE and random structures in aqueous milieu, have been described for some peptide hormones. However, these peptide hormones generally consist of fewer amino acids than ProTα. Examples of these hormones are the human corticotropin releasing factor (CRF) (Romier et al., 1993), galanin, neuropeptide y (Arvidsson et al., 1993), parathyroid hormone (Strickland et al., 1993; Bardon & Cuthberson, 1993; Klaus et al., 1991), neuropeptide PACAP (pituitary adenylate cyclase activating polypeptide), secretin, glucagon, growth hormone release factor, vasoactive intestinal peptide (Wray et al., 1993), sarcotoxin IA, cecropin A, cecropin B (Iwai et al., 1993), and calcitonin (Motta et al., 1991). All these peptides interact with membrane-bound receptors as it is also assumed for ProTα (Cordero et al., 1994).

In the last years, an increasing number of reports have substantiated the biological activity of ProTα (Cordero et al., 1991, 1992a; Baxevanis et al., 1993). The ProTα preparation used in this study was found to be effective, as demonstrated by its stimulating effects on peripheral blood lymphocytes and monocytes. ProTα, at a concentration of 500 ng/mL, showed restorative effects on defective IL-2 secretion and lymphokine-activated killer cell activity of peripheral blood lymphocytes (K. Eckert, unpublished results) and augments the defective antitumor activity of monocytes from melanoma patients in vitro (Garbin et al., 1994).

Though the exact mechanism of action of ProTα is not well-understood, we will discuss the conformational properties with respect to the observed biological activity. The characterization of two binding sites with high and low affinity on lymphocytes (Cordero et al., 1994) suggests a receptor-mediated process for the extracellular activities. It was concluded that, for maximum bioactivity, the occupation of almost all receptor molecules is necessary. Interestingly, common high-affinity receptors to IFN-α2 and thymosin α1, being identical with the 28 N-terminal amino acids of ProTα were found on murine thymocytes (Zavʼyalov et al., 1991). In thymosin α1, the sequence 16−23 is responsible for this activity. The carboxyl terminus of ProTα contains a nuclear localization signal (KKQK). Indeed, in COS cells transfected with a ProTα cDNA plasmid, the protein was found in the nucleus (Manrow et al., 1991). For the binding of small peptides or segments of larger polypeptides to proteins, the preferred conformation is an extended, unfolded, or not yet folded chain (Stanfield & Wilson, 1995). Secondary structure is formed only when the peptide is bound to the receptor protein. Therefore, the observed unfolded conformation must not be in contrast to the immunogenic nature of ProTα. There is also evidence for interactions with nuclear proteins (Papmarcaki & Tsolas, 1994). These authors have demonstrated the specific binding of ProTα to histone H1, which is obviously mediated via its central acidic region. To undergo electrostatic interactions with H1, an extended conformation of ProTα in solution could be favorable. While the unfolded state of ProTα may be in accordance with the different interactions discussed above, it might promote the degradation of the protein considerably. ProTα may be protected against proteolytic degradation first by the N-terminal acetylation (Haritos et al., 1984a) and second by a remarkable cluster of so-called stabilizing amino acids at the N-terminus. Following the "N-end rule" proposed by Varshavsky (Bachmair et al., 1986), the identity of its N-terminal residues relates to the half-life of a protein. Ten out of the first 12 N-terminal positions of ProTα are occupied

by amino acid residues belonging to the stabilizing class. Nevertheless, the existence of a more compact conformation due to the interaction with still unknown ligands cannot be excluded.

The high degree of evolutionary conservation of the amino acid sequence of ProTα (Makarova et al., 1989; Frillingos et al., 1991) cannot be due to the requirements of a specific three-dimensional structure. There must exist other reasons which we do not know yet. Random coil structures as well as α-helical segments do not react sensitively to the exchange of individual amino acids. It should be taken into consideration that particular structures of ProTα are formed in contact with receptors, membranes, or components of chromatin. But presently, this is only speculation. The finding that a biologically active protein with 109 amino acids adopts a dynamic random coil conformation under physiological conditions raises the following general question. Are we dealing with a rare or a widespread phenomenon in the field of macromolecular polypeptides? The latter is possibly true. The fact that only few physicochemical data are available might be due to troubles inherent in investigating statistically determined structures like that of ProTα. Taken together, the reported characterization of the protein conformation of ProTα represents an essential step in our understanding of the molecular mechanisms of interaction between ProTα and the immune cells as well as its proposed nuclear role.

ACKNOWLEDGMENT

The authors thank Dr. P. Franke, Institute of Biochemistry of the Free University of Berlin, for performing mass spectroscopy, I. Reim for performing CD and dynamic light scattering measurements, and Mrs. D. Otto for careful processing of the manuscript.

REFERENCES

Arvidsson, K., Landt, T., Langel, U., Bartfei, T., & Ehrenberg, A. (1993) *Biochemistry 32*, 7787−7798.

Bachmair, A., Finley, D., & Varshavsky, A. (1986) *Science 234*, 179−186.

Bardon, J. A., & Cuthberson, R. M. (1993) *Eur. J. Biochem. 215*, 315−321.

Baxevanis, C. N., Reclos, G. J., & Papamichail, M. (1993) *Int. J. Cancer 53*, 264−268.

Benoit, H., & Doty, P. (1953) *J. Phys. Chem. 57*, 958−963.

Burchard, W., Schmidt, M., & Stockmayer, W. H. (1980) *Macromolecules 13*, 1265−1272.

Bustelo, X. R., Otero, A., Gomez-Marquez, J., & Freire, M. (1991) *J. Biol. Chem. 266*, 1443−1447.

Chen, Y.-H., Yang, J. T., & Chau, K. H. (1974) *Biochemistry 13*, 3350−3359.

Chou, P. Y., & Fasman, G. D. (1974) *Biochemistry 13*, 222−245.

Cordero, O. J., Sarandeses, C. S., Lopez, J. L., Cancio, E., Regueiro, B. J., & Nogueira, M. (1991) *Int. J. Immunopharmacol. 13*, 1059−1065.

Cordero, O. J., Sarandeses, C. S., Lopez, J. L., & Nogueira, M. (1992a) *Lymphokine Cytokine Res. 11*, 277−285.

Cordero, O. J., Sarandeses, C. S., Lopez, J. L., & Nogueira, M. (1992b) *Biochem. Int. 28*, 1117−1124.

Cordero, O. J., Sarandeses, C., & Nogueira, M. (1994) *FEBS Lett. 341*, 23−27.

Damaschun, G., Damaschun, H., Gast, K., Gernat, Ch., & Zirwer, D. (1991) *Biochim. Biophys. Acta 1078*, 289−295.

Damaschun, G., Damaschun, H., Gast, K., Misselwitz, R., Zirwer, D., Gührs, K. H., Hartmann, M., Schlott, B., Triebel, H., & Behnke, H. (1993a) *Biochim. Biophys. Acta 1161*, 244−248.

Damaschun, G., Damaschun, H., Gast, K., Misselwitz, R., Müller, J. J., Pfeil, W., & Zirwer, D. (1993b) *Biochemistry 32*, 7739−7746.

Damaschun, G., Gast, K., Müller-Frohne, M., & Zirwer, D. (1995) *Biopolymers* (in press).

Eckert, K., Schmitt, M., Garbin, F., Wahn, U., & Maurer, H. R. (1994) *Int. J. Immunopharmacol. 16*, 1019−1025.

Economou, M., Seferiadis, K., Frangou-Lazaridis, M., Horecker, B. L., & Tsolas, O. (1988) *FEBS Lett. 233*, 342−346.

Eilers, M., Schirm, S., & Bishop, J. M. (1991) *EMBO J. 10*, 133−141.

Flory, P. J. (1988) *Statistical Mechanics of Chain Molecules*, Oxford University Press, Oxford, U.K.

Frillingos, S., Frangou-Lazaridis, M., Seferiadis, K., Hulmes, J. D., Pan, Y. C., & Tsolas, O. (1991) *Mol. Cell. Biochem. 108*, 85−94.

Garbin, F., Eckert, K., Büttner, P., Garbe, C., & Maurer, H. R. (1994) *Anticancer Res.* (in press).

Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol. 120*, 97−120.

Gast, K., Damaschun, G., Misselwitz, R., & Zirwer, D. (1992) *Eur. Biophys. J. 21*, 357−361.

Gast, K., Damaschun, H., Misselwitz, R., Müller-Frohne, M., Zirwer, D., & Damaschun, G. (1994) *Eur. Biophys. J. 23*, 297−305.

Haritos, A. A., Goodall, G. J., & Horecker, B. L. (1984a) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1008−1011.

Haritos, A. A., Tsolas, O., & Horecker, B. L. (1984b) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1391−1393.

Haritos, A. A., Blacher, R., Stein, S., Caldarella, J., & Horecker, B. L. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 343−346.

Haritos, A. A., Yialouris, P. P., Heimer, E. P., Felix, A. M., Hannappel, E., & Rosemeyer, M. A. (1989) *FEBS Lett. 244*, 287−290.

Iwai, H., Nakajima, Y., Natori, S., Arata, Y., & Shimada, I. (1993) *Eur. J. Biochem. 217*, 639−644.

Johnson, W. C., Jr. (1990) *Proteins: Struct., Funct., Genet. 7*, 205−214.

Kirste, R. G., & Oberthür, R. C. (1982) in *Small-angle X-ray Scattering* (Glatter, O., & Kratky, O., Eds.) pp 397−431, Academic Press, London, New York.

Klaus, W., Diekmann, T., Wray, V., & Schomburg, D. (1991) *Biochemistry 30*, 6936−6942.

Kratky, O., & Porod, G. (1949) *Rec. Trav. Chim. Pays-Bas 68*, 1106−1122.

Levin, J. M., & Garnier, J. (1988) *Biochim. Biophys. Acta 955*, 283−295.

Makarova, T., Grebenshikov, N., Egorov, C., Vartapetian, A., & Bogdanov, A. (1989) *FEBS Lett. 257*, 247−250.

Manrow, R. E., Sburlati, A. R., Hanover, J. A., & Berger, S. L. (1991) *J. Biol. Chem. 266*, 3916−3924.

Motta, A., Pastore, A., Goud, N. A., & Morelli, M. A. C. (1991) *Biochemistry 30*, 10444−10450.

Müller, J. J., Zalkova, T. N., Damaschun, G., Misselwitz, R., Serdyuk, I. N., & Welfle, H. (1986) *Stud. Biophys. 112*, 151−162.

Nishikawa, K., & Noguchi, T. (1991) *Methods Enzymol. 202*, 31−44.

Panneerselvam, C., Haritos, A. A., Caldarella, J., & Horecker, B. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4465−4469.

Papamarcaki, T., & Tsolas, O. (1994) *FEBS Lett. 345*, 71−75.

Provencher, S. W. (1982) *Comput. Phys. Commun. 27*, 229−242.

Ptitsyn, O. B., & Finkelstein, A. V. (1983) *Biopolymers 22*, 15−25.

Romier, C., Bernassau, J.-M., Cambillau, C., & Darbon, H. (1993) *Protein Eng. 6*, 149−156.

Rost, B., & Sander, C. (1993) *J. Mol. Biol. 232*, 584−599.

Sburlati, A. R., Manrow, R. E., & Berger, S. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 253−257.

Schmidt, G., & Werner, D. (1991) *Biochim. Biophys. Acta 1088*, 442−444.

Schweers, O., Schönbrunn-Hanebeck, E., Marx, A., & Mandelkow, W. (1994) *J. Biol. Chem. 269*, 24290−24297.

Shiraki, K., Nishikawa, K., & Goto, Y. (1995) *J. Mol. Biol. 245*, 180−194.

Stanfield, R. L., & Wilson I. A. (1995) *Curr. Opin. Struct. Biol. 5*, 103−113.

Strickland, L. A., Bozzato, R. P., & Kronis, K. A. (1993) *Biochemistry 32*, 6050−6057.

Szabo, P., & Weksler, M. E. (1992) *Clin. Immunol. Immunopathol. 65*, 195−200.

Thomas, P. D., & Dill, K. A. (1993) *Proteins Sci. 2*, 2050−2065.

Tsitsiloni, O. E., Stiakakis, J., Koutselinis, A., Gogas, J., Markopoulos, Ch., Yialouris, P., Bekris, S., Panoussopoulos, D., Kiortsis, V., Voelter, W., & Haritos, A. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 9504−9507.

Watts, J. D., Cary, P. D., Sautiere, P., & Crane-Robinson, C. (1990) *Eur. J. Biochem. 192*, 643−651.

Wray, V., Kokoschke, C., Nokihara, K., & Naruse, S. (1993) *Biochemistry 32*, 5832−5841.

Zav`yalov, V. P., Navolotskaya, E. V., Abramov, V. M., Galaktionov, V. G., Isaev, T. S., Kaurov, O. A., Kozhich, A. T., Maiorov, V. A., Prusakov, A. N., Vasilenko, R. N., & Volodina, E. Ya. (1991) *FEBS Lett. 278*, 187−189.

BI9508911