Sequential Compaction of a Random Copolymer of Hydrophilic and Hydrophobic Amino Acid Residues

Anton S. Karnoup and Vladimir N. Uversky*

Institute of Protein Research, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia

Received January 14, 1997; Revised Manuscript Received September 5, 1997[⊗]

ABSTRACT: The behavior of a high molecular weight (48.8×10^3) synthetic random copolymer composed of Glu, Leu, and Trp residues in molar ratio 82.5:16:1.5 under compaction and unfolding conditions has been studied using circular dichroism, fluorescence spectroscopy, viscometry, and size-exclusion chromatography (FPLC). It has been shown that the decrease of the pH value leads to two-stage compaction of the copolymer molecules. This means the formation of an intermediate conformational state. Structural properties and stability of the revealed conformations (compact state and less compact intermediate) of the copolymer toward temperature and urea-induced unfolding have also been investigated. The compact conformation of the macromolecules is characterized by high resistance toward urea action (at 21.5 °C), whereas the transition from intermediate (less compact) state to coil-like conformation can easily be accomplished by the addition of small amounts of urea (at 21.5 °C). Melting of the compact state in the presence of urea is a two-stage process as well. This fact also confirms the existence of a folding intermediate.

1. Introduction

Every known protein, under the appropriate environmental conditions, folds into its native unique three-dimensional (3D) structure. Although each given protein has its quite definite primary structure, it is thought that natural polypeptides originated as random copolymers constructed of amino acid residues (peptide units) and were just "edited" in the course of biological evolution on the principle of natural selection to acquire and refine their various unique 3D structures and functional properties[1−3] (see also references in ref 3). Therefore, to understand general physico-chemical principles of self-organization of protein molecules, it is important to study model polypeptides with random amino acid sequences. Significant progress has been made in this field to date. Theoretical studies showed that the basic features of the compact 3D structure of typical globular proteins may perfectly be inherent in random amino acid sequences.[1−6] Numerous investigations of more or less simplified theoretical models of polypeptide chain folding have revealed great similarities between the model statistical copolymers and real proteins.[1−13] This suggestion has been confirmed in part by experimental studies of random copolymers of glutamic acid (Glu, polar residue) with leucine (Leu, nonpolar residue) with average molecular masses up to $\sim 10 \times 10^3$ g/mol (~80 amino acid residue (aar)).[14−17] The possibility of isolation by gel filtration of a nonaggregating fraction from each of the copolymers investigated has been shown. Interestingly, the amino acid composition of these non-aggregating fractions was nearly constant ($17 \pm 2\%$ Leu) regardless of the amino acid composition of the parent copolymers (which varied in the experiments from 9% to 40% Leu).[14−17] The ability of these fractions to undergo a coil−helix transition and, eventually, to acquire compact globular conformation(s) (nearly as compact as that of real globular proteins) upon pH decrease, was demonstrated. It was also established that the coil−globule transitions in these synthetic random copolymers definitely do not belong to an "all-or-none" type.[15−17] In accord with these findings, Rao *et al.*[18] reported the existence of "collapsed" globular structures (with overall 46% helix content) in a statistical copolymer of glutamic acid, lysine, and alanine (in molar ratio 3:3:4; $M \sim (5-10) \times 10^3$, i.e., ~50−95 aar) in aqueous phosphate buffer at pH 7. It should be emphasized here that the above experiments were performed not for single amino acid sequences but for heterogeneous conformational mixtures, i.e., for ensembles of the polypeptide chains of a given average length and composition (for "sequence-space soup"[2]). Random sequence polypeptides with strictly determined primary structure *within each given pool (sample)* of macromolecules can be obtained by combinatorial mutagenesis[19,20] (see also references in ref 20). In this approach, a library of synthetic genes is generated. Each gene contains a set of randomly distributed codons and an epitope-encoding tail in its protein-encoding region. The library is then transformed into cells, and the proteins produced by cells are identified by immunoblotting to the epitope tag (this tag may subsequently be cleaved after isolation of a "random" protein). Using this approach, Sauer *et al.*[19] (see also ref 20) have clearly shown that cooperatively folded "proteins" with compact globular structures can easily be recovered from a library containing 80-residue "proteins" predominantly composed of glutamine (40%), leucine (28−30%), and arginine (15−18%). Sauer *et al.* have reported recently[20] that one "random protein" from a library composed of random combinations of 16 of the naturally occurring amino acids has a cooperatively folded β-structure and appears to have a well-packed core by the criterion of near-UV CD. The above findings suggest that a significant amount of structural information is carried by the simple pattern of polar and non-polar residues along a polypeptide chain.

Besides, it has been shown that not only copolymers of polypeptide nature but also such homopolymers as poly(*N*-isopropylamide) (PNIPAM) and poly(*N*-isopropylmethacrylamide) (PNIPMAM) can undergo reversible coil−globule transitions (induced by temperature).[21]

* Correspondence address: Institute of Protein Research, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia. Tel./Fax: 7(095)924-0493. E-mail: uversky@sun.ipr.serpukhov.su.

[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1997.

S0024-9297(97)00033-8 CCC: $14.00

The ability to attain compact globular structure (with the degree of compactness similar to that of native globular proteins) and an "all-or-none" character of the coil−globule transition has been established for the fractions of PNIPAM and PNIPMAM with molecular masses ~(10−20) × 10^3 g/mol.

Results summarized above definitely show that (1) the compact (globular) state with pronounced secondary structure can be inherent not only in proteins with unique primary structure but also in random sequences of polar and nonpolar residues and (2) great similarity exists in the denaturation and renaturation behavior between proteins, synthetic polypeptides, and some other macromolecules, including even homopolymers.

However, it is necessary to emphasize here that it is clearly a mistake to quietly substitute the globular state of a native protein by the homopolymer globule (or by the globule of synthetic polypeptide comprising a few different amino acid residuces). For instance, it was established that a contracted molecule of poly(*N*-isopropylacrylamide) contains about 74% water within the "globule",[22] while there are typically no water molecules inside the compact structure of the native globular protein (probably due to the specific amino acid sequence that allows such space-filling). This is one (but not the only!) example showing that compact polymer conformation cannot be considered as a structural analogue of the native protein globule. On the other hand, it was shown that the native state is not the only compact globular conformation of globular proteins−at definite conditions they can be transformed into the molten globule state (see below and ref 3 for a recent review). Comparison of structural properties of this intermediate state with those of the polymer globule shows that it is difficult to exclude that the protein molten globule is relatively well modeled by the compact form of a synthetic polypeptide. Indeed, molten globule, being an equilibrium intermediate state of many proteins, is realized at mild denaturing conditions. Structural properties of a protein molecule in this state are intermediate between those of the native and unfolded protein. The rigid tertiary structure in the molten globule state is either absent or strongly reduced, but the protein molecule still has a number of important features of the native secondary structure and the overall native architecture.[3] This state is relatively compact (the hydrodynamic volume of a protein molecule in this state is 1.4−1.5 times larger than that of the native protein)[3,23−25] and has globular structure,[26−29] but there is a large amount of water inside the protein molecule in the molten globule state.[30,31] Interestingly, it was shown quite recently that the molten globule is not the only relatively compact equilibrium intermediate state of small globular proteins and a premolten globule state was discovered.[32−39] A protein molecule in this premolten globule state is less compact than in the molten globule but considerably more compact than in the unfolded state. It preserves a substantial part of secondary structure but has no globular structure.[32−39] It was suggested that both molten globule[40−44] and premolten globule[32−39] states represent equilibrium counterparts of kinetic intermediates accumulated during refolding of a globular protein from the completely unfolded state.

In this paper, we study the behavior of a high molecular weight synthetic random copolymer of hydrophilic and hydrophobic amino acid residues (Glu, Leu, and Trp in molar ratio 82.5:16:1.5), similar to those investigated earlier[14−17] but with the higher molecular mass (fraction with $M \sim 48.8 \times 10^3$, *i.e.*, ~380 aar), under compaction and unfolding conditions. A few interesting and previously unobserved features have been noted in the folding/unfolding behavior of this polypeptide.

2. Materials and Methods

2.1. Materials. γ-Benzyl L-glutamate (γ-OBzl-L-Glu) and L-leucine (L-Leu) were from Fluka. L-Tryptophan (L-Trp) was from Reanal. Before use, dioxane (Reakhim) was boiled (bp 100−102 °C) with metallic Na and then distilled. Benzene (Reakhim) was distilled (bp 80−81 °C) just before use. Petroleum ether was used as received from Fluka. TBA (Fluka) was distilled (bp 88−90 °C) prior to use. DCM and methanol were used as received from Reakhim. Phosgene ($COCl_2$) was synthesized by reaction of carbon tetrachloride (Reakhim) with oleum (sulfuric acid fuming 27−30% SO_3) (Reakhim) and purified as described in ref 45. Hydrogen fluoride was used as received from Fluka. Diethyl ether was used as received from Reakhim. DMS and *m*-cresol were from Aldrich Chemical Co. Blue Dextran 2000 ($M = 2 \times 10^6$) and a Gel Filtration Calibration Kit were from Pharmacia. Acrylamide (Fluka) was used without preliminary purification. Urea (Reakhim) was purified by recrystallization. The denaturant concentration in solution was measured by refractive index. The components of the buffer solutions were used without preliminary purification. All aqueous solutions were prepared in double-distilled water and passed through 0.2 μm Nylon 66 filters (Rainin Instruments Co.).

2.2. Equipment. Phosgene was generated with the use of a laboratory-made $COCl_2$ unit (see ref 45). Melting points were measured on a VEB Analytik Dresden melting point apparatus and are uncorrected. Removal of the OBzl protecting groups from Glu residues of the copolymer was performed with the use of a laboratory-made HF-unit (see ref 45). pH measurements were performed using a Mettler Delta 320 pH meter. "Visking" dialysis tubes from Serva were used for preparative desalting and removal of polypeptides with molecular masses less than 10 × 10^3. Preparative and analytical chromatographic procedures were done using columns containing Toyopearl HW-60F gel (Toyo Soda), which were connected to an FPLC System (Pharmacia). A Biotronik Lc 5001 amino acid analyzer was employed for determination of the amino acid composition of the copolymer. Infrared (IR) spectra were recorded on a Specord 75 IR spectrometer (Carl Zeiss). Samples for IR spectrometry were either solutions or suspensions in DCM, dioxane, or benzene in CaF_2 cuvettes. UV spectra were recorded on a Specord UV−vis spectrophotometer (Carl Zeiss). Circular dichroism measurements were done with a JASCO-600 spectropolarimeter (Japan Spectroscopic Co.). Fluorescence measurements were performed with an Aminco SPF-1000cs corrected spectrofluorometer (American Instrument Corp.). All the spectral instruments were equipped with a temperature-controlled holder.

2.3. Synthesis of Amino Acid *N*-Carboxyanhydrides. NCAs of γ-OBzl-L-Glu, L-Leu, and L-Trp were prepared in accordance with a general Fuchs−Farthing procedure.[46−48] A typical procedure for the synthesis of NCAs (given here in detail for NCA of γ-OBzl-L-Glu) was as follows.[45] A 1000 mL two-neck round-bottom flask was charged with dried and finely divided γ-OBzl-L-Glu (12.60 g, 53.2 mmol) and dry dioxane (500 mL). The suspension was stirred and heated to 35 °C. Dry $COCl_2$ was slowly bubbled through the suspension for 2.5 h. During that time, the initial suspension became almost completely clear. After overnight stirring, dioxane, HCl, and excess $COCl_2$ were distilled off at 20 °C and 20 mmHg. The dry residue (viscous oil, in the case of crude NCA of L-Leu) was taken up with DCM, which was then distilled off as before. During this step, repeated five times to remove all traces of HCl, the product crystallized as a white (yellowish in the case of the crude NCA of L-Leu; pinkish in the case of the crude NCA of L-Trp) powder. Then the product was again dissolved in DCM and passed through a Whatman GF/C borosilicate glass filter (Whatman) under nitrogen to remove the small

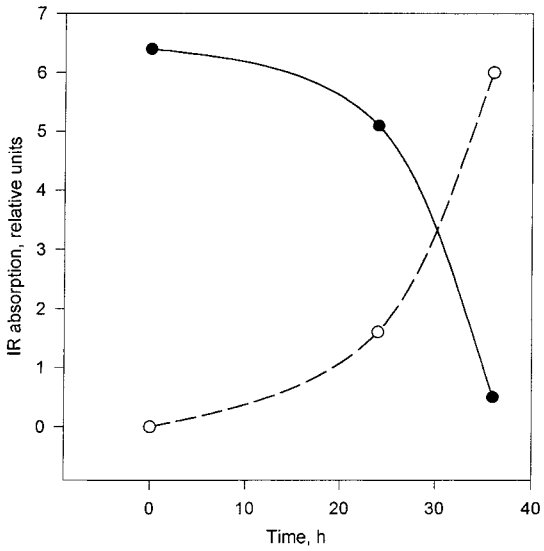

Figure 1. IR monitoring of the copolymerization reaction: anhydride carbonyl stretching vibrations at 1792 cm^{-1} (filled circles and the solid line), and the amide I band (peptide bond, α-helical conformation) at 1645 cm^{-1} (hollow circles and the dashed line). The copolymerization was carried out in benzene (see text).

amount of precipitate. The clear solution was transferred to a 500 mL round-bottom flask and concentrated to ~50 mL. Then petroleum ether (~150 mL) was added and the mixture was placed in the freezer (−20 °C) overnight to allow complete precipitation of the NCA. The product was filtered under nitrogen, washed with petroleum ether, and dried in vacuum under P_2O_5 to give 12.74 g (48.4 mmol, 91% yield) of the NCA in the form of fine white crystals: mp 94−96 °C; IR 1850, 1785, 1720 cm^{-1}.

The NCAs of L-Leu and L-Trp were prepared according to the same procedure and afforded IR spectra almost identical to that of NCA of γ-OBzl-L-Glu. NCAs of L-Leu and L-Trp were recrystallized three times from the DCM/petroleum ether (1/4) mixture. For the NCA of L-Leu: phosgenation time 2.5 h; white, with slightly yellowish shade, fine crystals; yield 30%; mp 74−76 °C; IR 1855, 1795 cm^{-1}. For the NCA of L-Trp: phosgenation time 1.5 h; pinkish fine powder; yield 55%; mp 121−125 °C; IR 1855, 1787 cm^{-1}; UV 275 (major band), 280, 291 (narrow sharp peak) nm.

The dry NCAs were stored in the freezer (−76 °C) in hermetically sealed containers. The NCA of L-Trp was prepared immediately before copolymerization.

2.4. Preparation of the Copolymer. A 500 mL round-bottom flask equipped with a drying tube was charged with the NCA of γ-OBzl-L-Glu (2.20 g, 8.36 mmol), the NCA of L-Leu (0.59 g, 3.75 mmol), and the NCA of L-Trp (0.028 g, 0.12 mmol). Benzene (300 mL) was added via cannula, and after dissolution of the NCAs, under stirring, TBA (39 μL, 0.16×10^{-3} mmol) was added (the initial molar ratio NCAs/TBA was 75). The pH of the solution was nearly constant (~7.5) during the copolymerization. The copolymerization was monitored by IR spectroscopy until complete disappearance of the band corresponding to the anhydride carbonyl stretching vibrations (1792 cm^{-1}), accompanied by a simultaneous increase in intensity of the amide I band (1645 cm^{-1}) (see Figure 1) assigned to the α-helical conformation of the growing polypeptide chain. No amide I vibrations at ~1620 cm^{-1} (usually assigned to a β-sheet conformation) were seen. After ~1.5 days, the reaction mixture was concentrated under reduced pressure, methanol (~100 mL) was added, and the precipitate was filtered off, washed with methanol, and dried in a vacuum oven (40 °C) overnight.

The polypeptide was deprotected by a standard low-HF procedure in a DMS:HF:*m*-cresol mixture (65:25:10).[23] The resulting substance was washed with diethyl ether, dried, dissolved in aqueous phosphate buffer (0.02 M K phosphate/0.2 M KCl, pH 7.0), and dialyzed against the same buffer (3 days at 4 °C, 10-fold change of buffer) to remove all low molecular mass impurities and short polypeptides. The solution was then acidified to pH 5.2 by HCl, and precipitate was removed by centrifugation (at 30000 rpm for 90 min). The supernatant was acidified to pH 2.0 by HCl, the resulting suspension was centrifuged again, the precipitate was redissolved in K phosphate buffer (pH 7.0), and the solution was fractionated on a preparative Toyopearl HW-60F column (95 × 1.6 cm i.d.) equilibrated with the same buffer. Then the fraction nonaggregating at pH 5.20 was isolated from the triple copolymer in accordance with ref 14, but using the Toyopearl HW-60F column.[45] The molecular mass (M) of the isolated fraction was determined from the value of its averaged Stokes radius in the completely unfolded state (when $[\Theta]_{220} = 0$ deg·cm^2·$dmol^{-1}$) of macromolecules. The Stokes radius of the unfolded copolymer (R_s^U) was measured by FPLC (see further) on an analytical Toyopearl HW-60F column (24 × 1 cm i.d.) in K phosphate buffer, pH 6.50, containing 10 M urea (or 8 M GdmCl) at 24 °C. By this method, we obtained an M value of 48.8×10^3 g/mol. The amino acid composition of the isolated fraction of the copolymer was determined by amino acid analysis after hydrolysis of an aliquot of the copolymer by 6 M HCl for 72 h at 110 °C. The number of Trp residues was determined from comparison of the UV-absorption spectra of the solution of nonstructured copolymer (pH 7.0) with those of L-Trp (at known concentrations).[17,45] The nonaggregating (at pH 5.20) fraction of the copolymer had a Glu:Leu:Trp ratio of 82.5:16:1.5.

2.5. Preparation of Copolymer Solutions. We used aqueous solutions of the copolymer in 0.02 M K phosphate (or citrate)/0.2 M KCl buffer. Prior to measurements, all the copolymer solutions with the appropriate pH and denaturant concentration values were incubated at a desired temperature. Copolymer concentrations were 0.05−0.1 mg/mL for FPLC experiments, 0.02 mg/mL for fluorescence measurements, and 1.0 mg/mL for CD measurements. The pH of the solutions was measured with an accuracy of ±0.01. The denaturant concentration in the solutions was measured with an accuracy of ±0.01.

2.6. Size-Exclusion Chromatography (FPLC). The Toyopearl HW-60F column (24 × 1 cm i.d.) was calibrated by a set of eight native proteins with known Stokes radii.[24] The R_s value (in Å) for the copolymer under each given condition was calculated from the respective availability coefficient,[49] K_{av}, determined by the expression

$$K_{av} = \frac{V_{el} - V_0}{V_t - V_0} \tag{1}$$

where V_{el} is the respective elution volume (mL), V_0 is the void volume of the column (elution volume of Blue Dextran 2000) (mL), and V_t is the whole geometrical volume of the column (mL). Our experimentally obtained dependence of R_s on K_{av}[49,50] is

$$R_s = 215.082[\log(K_{av})]^{1/2} - 80.319 \tag{2}$$

The Stokes radii of protein molecules of a given molecular mass M in native (R_s^N) and completely unfolded (R_s^U) states are described by the empirical equations[24]

$$\log(R_s^N) = 0.369 \log(M) - 0.254 \tag{3}$$

and

$$\log(R_s^U) = 0.533 \log(M) - 0.682 \tag{4}$$

based on the experimental data of Tanford.[51]

In a typical chromatographic experiment, an aliquot of the copolymer solution (10 −30 μL) in K-citrate buffer with the desired pH and denaturant concentration values was loaded on the analytical Toyopearl HW-60F column equilibrated with the same buffer. The flow rate was 0.5 mL/min. The eluted copolymer was detected by UV absorption at 226 or 275 nm. FPLC measurements were done at 23 ± 1 °C.

2.7. Intrinsic Viscosity Measurements. Intrinsic viscosity values $[\eta]$ were determined as

$$[\eta] = \lim_{C\to 0} \eta_{red} = \lim_{C\to 0} \frac{\eta - \eta_o}{\eta_o C} \quad (5)$$

where η is the viscosity of the polymer solution of concentration C and η_o is the viscosity of the solvent. Reduced viscosity η_{red} values were measured with a Zimm rotational viscosimeter (Bureau of Biological Instrumentation, Pushchino, Russia) and were calculated as

$$\eta_{red} = \frac{t - t_o}{t_o C} \quad (6)$$

where t and t_o correspond to the time of one turn of the rotor in the polymer solution of concentration C or in the solvent, respectively.

2.8. Fluorescence and CD Measurements. To decrease the probability of energy transfer from Trp to Trp, the wavelength of the excitation beam in all fluorescence measurements was 296 nm.[17] Quartz cuvettes with the path length of 1 cm were used. Experiments on quenching of the fluorescence of Trp residues were performed as follows. Samples of the copolymer in K-citrate buffer (2 mL each) with adjusted pH values were incubated at 24 °C for 3 h, and their emission spectra were recorded. Then acrylamide (10 μL) was added to the samples to its final concentration 0.4 M. The samples were incubated for 10 more min at 24 °C, and their emission spectra were recorded again. The results were expressed in terms of efficiency of quenching of the fluorescence, E (%),

$$E = 100(1 - I/I_0) \quad (7)$$

where I_0 is the fluorescence intensity (at λ_{max}) of a sample in the absence of acrylamide, and I is the fluorescence intensity of a sample in the presence of acrylamide.

For far-UV CD measurements, quartz cuvettes with the path length of 0.148 mm were used. Calculations of the percentage of α, β, and unordered structure were performed using the obtained CD spectra and empirical dependence given in ref 52. Data fitting, construction of the best-fit lines, and deconvolution of the experimental curves were done using the complex Regression program.[53]

3. Results and Discussion

3.1. pH-Induced Compaction. Changes in Far-UV CD Spectra and Secondary Structure Content. Figure 2 presents far-UV CD spectra of the copolymer at various pH values, ranging from pH 6.00 (secondary structure is virtually absent) to pH 4.23 (a spectrum corresponding to the most pronounced secondary structure). Two isodichroic points (where the spectrum lines intersect) are seen in the picture, one at ∼203 nm, and the other at ∼204 nm. It is seen also that the transition between these two isodichroic points takes place in a pH range from 5.14 to 5.27. It is well-known that the presence of an isodichroic point (see, e.g., ref 54) indicates that one deals with a transition between two states. The presence of two isodichroic points (as in our case) reflects the existence of at least three discrete states. Figure 3A depicts the pH dependence of the molar ellipticity $[\Theta]_{220}$ (open circles, solid line), which is in good agreement with the above suggestion: a bend at about pH 5.15 discloses the biphasic character of the curve. Deconvolution of the pH dependence of $[\Theta]_{220}$ confirmed that this is a superposition of at least two sigmoidal curves (dotted lines in Figure 3A), one of them with a semitransition point at pH 4.89 and the other with a semitransition point ($pH_{1/2}$) at pH 5.53. $[\Theta]_{220}$ is nearly constant at pH higher than ∼5.75 and lower than ∼4.3. Additional confirmation of the existence of

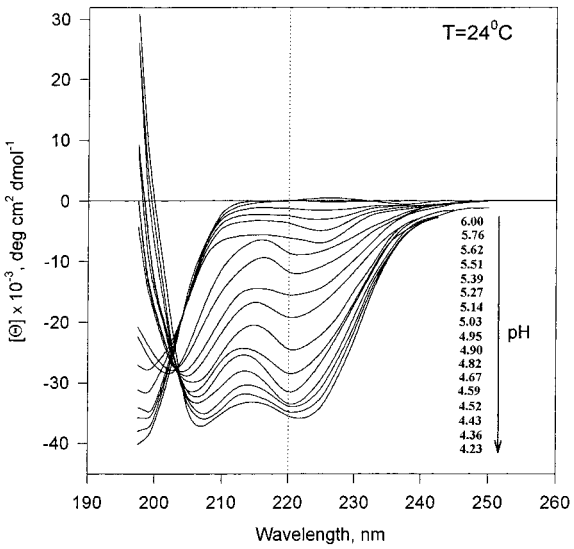

Figure 2. Far-UV CD spectra of the random copolymer Glu:Leu:Trp (82.5:16:1.5, fraction with $M \sim 48.8 \times 10^3$ g/mol) at different pH values. Measurements were done in 0.02 M K-citrate/0.2 M KCl buffer, at copolymer concentration ∼1 mg/mL.

an intermediate state can be obtained from the analysis of changes in the shape of CD spectra (e.g., from the $[\Theta]_{222.5}/[\Theta]_{207.5}$ ratio). Figure 3B shows the pH dependence of the ratio of molar ellipticities at 222.5 and 207.5 nm, respectively. Two phases of the $[\Theta]_{222.5}/[\Theta]_{207.5}$ change upon a decrease in pH are clearly distinguishable in the picture. The $[\Theta]_{222.5}/[\Theta]_{207.5}$ ratio is essentially constant at pH values higher than ∼5.8, lower than ∼4.6, and within the narrow interval of pH from 5.14 to 5.39. Semitransition points corresponding to each part of this complex double-sigmoidal curve are at pH 4.84 and pH 5.59, respectively. A curve of this type is indicative of existence of at least three general conformational states for polypeptide macromolecules. Figure 4 displays the pH dependence of the secondary structure content calculated from CD spectra given in Figure 2. It is seen that, upon a pH decrease, the α-helix content grows first slowly and monotonously in the pH range from ∼5.7 to ∼5.1, and then more rapidly in the pH range from ∼5.1 to ∼4.3 (up to ∼80%), whereas the β-structure content jumps up sharply to ∼20% in the narrow pH range from ∼5.3 to ∼5.1 and then remains nearly constant. As for the remaining structure (unordered and/or not regular) content (calculated as $100 - f_\alpha - f_\beta$, where f_α and f_β are the contents of α- and β-structure, respectively), its pH dependence is also biphasic (see Figure 4, open circles, dashed line).

These results taken together allow us to conclude that a pH decrease provokes two-stage growth of the coplymer's secondary structure content. In other words, accumulation of an equilibrium intermediate conformer takes place. Interestingly, all changes in β-structure content occur during the first stage, i.e., in the course of formation of the intermediate state.

3.2. pH-Induced Compaction As Monitored by SEC, Viscometry, and Quenching of Fluorescence of Trp Residues. SEC permits one to evaluate hydrodynamic dimensions of different conformers of protein[24,35,55,56] and polymer (of nonpeptide nature)[21] molecules with high accuracy at low concentrations (∼0.1−0.001 mg/mL). It was shown also[24,56] that a chromatographic (FPLC or HPLC) column can safely be applied to studying transitions between native, molten globule, and unfolded states of proteins. Here we

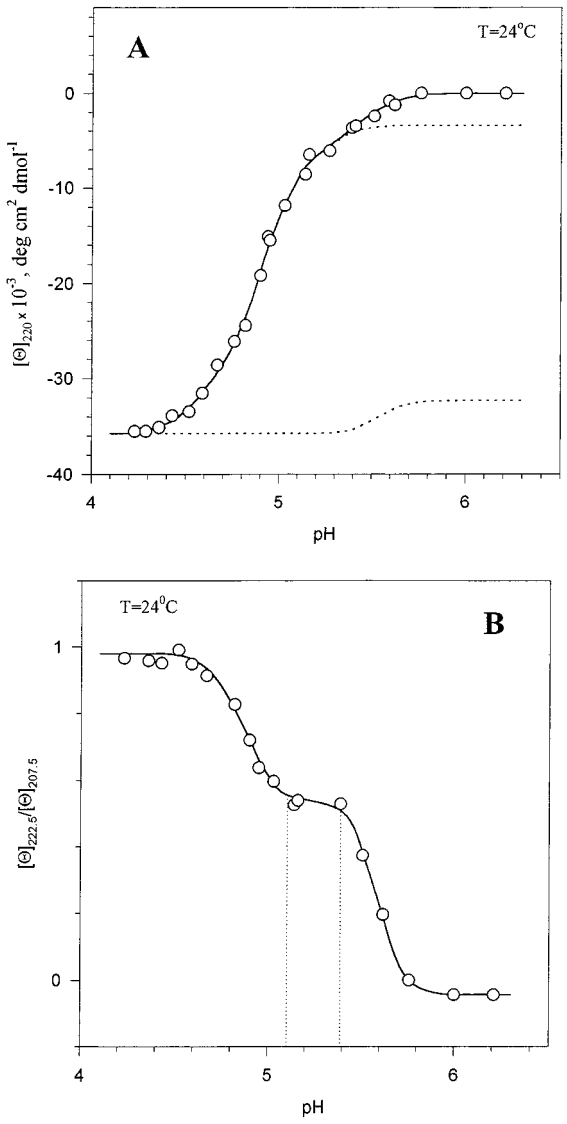

Figure 3. pH dependences of (A) molar ellipticity $[\Theta]_{220}$, and (B) the $[\Theta]_{222.5}/[\Theta]_{207.5}$ ratio for the random copolymer Glu:Leu:Trp (82.5:16:1.5, fraction with $M \sim 48.8 \times 10^3$ g/mol). Solid lines drawn through the symbols are computer-generated best-fit lines. Each double-sigmoidal line is the superposition of two simpler sigmoidal lines (shown by dotted curves in (A)). Here, as well as in all the other cases, semitransitional pH values ($pH_{1/2}$) were extracted from the curves deconvoluted in a manner described in ref 53.

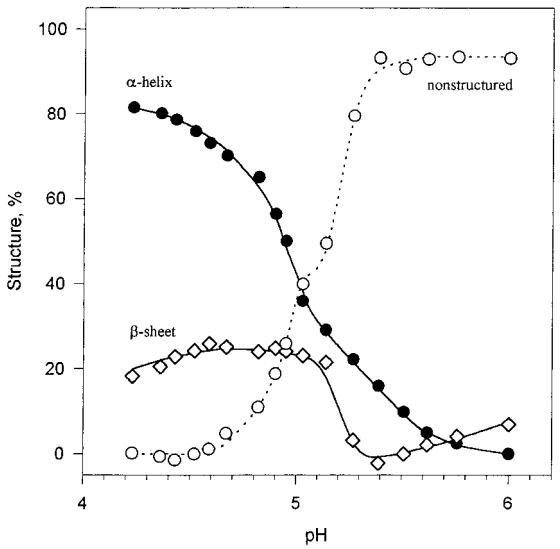

Figure 4. pH dependence of the secondary structure content of the random copolymer Glu:Leu:Trp (82.5:16:1.5, fraction with $M \sim 48.8 \times 10^3$ g/mol) calculated using CD spectra given in Figure 2. The pH course of the α-helix percentage is shown by filled circles and a solid line. The pH course of the β-structure percentage is shown by open rectangles and a solid line. The dotted line and open circles present the percentage of nonstructured parts of the averaged polypeptide chain in the course of pH change.

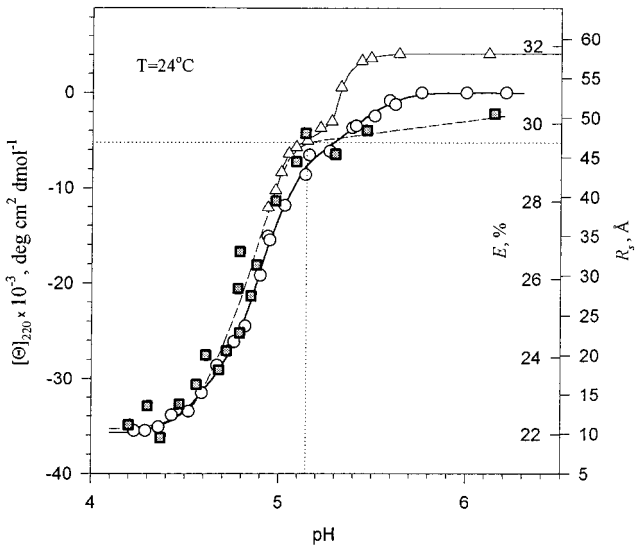

Figure 5. pH dependences of molar ellipticity $[\Theta]_{220}$ (open circles, bold solid line) (see also Figure 3B), averaged molecular Stokes radius (R_s) (open triangles, thin solid line), and efficiency of quenching of Trp fluorescence (E) (shaded squares, dashed line) for the random copolymer Glu:Leu:Trp (82.5:16:1.5, fraction with $M \sim 48.8 \times 10^3$ g/mol) at room temperature.

applied SEC (FPLC) to studying averaged dimensions of our copolymer molecules in different conformational states. Unfortunately, correct measurements of molecular size (Stokes radius, R_s) were possible only at pH values not lower than ~4.9, because a further pH decrease posed complications in chromatographic behavior of the copolymer molecules due to the appearance of interactions of the polypeptide with the column matrix. However, it is possible to measure R_s correctly at pH ≥ 4.9, and the corresponding part of the pH dependence of R_s, shown in Figure 5 by open triangles, exhibits at least biphasic character of the curve: a short plateau at about pH 5.15 is observed. The semitransition point of the completed first transition detected by SEC is at pH 5.31.

It is known that the value of intrinsic viscosity $[\eta]$ of a macromolecule in solution is clearly related to its molecular dimensions.[29] Indeed, globular proteins in their native[51] or molten globule[3,24,25] states have compact structures, which is reflected in an intrinsic viscosity of 3−4 mL/g. Complete unfolding of a protein molecule by strong denaturants (such as urea or GdmCl) leads to the considerable increase in its hydrodynamic volume and, as a consequence, to the increase of the $[\eta]$ value.[51] It was even established that the value of intrinsic viscosity for random coils depends on the molecular weight of the macromolecule.[51] It means that to determine hydrodynamic dimensions of the polymer molecule in different conformational states (random coil, intermediate state, and compact state), the intrinsic viscosity values should be measured at different experimental conditions. Results of these investigations are summarized in Table 1. One can see that the intrinsic viscosity of the copolymer molecule measured at neutral pH is somewhat smaller than that in the presence of

Table 1. Structural Properties of a Random Copolymer Glu:Leu:Trp (82.5:16:1.5), $M = 48.8 \times 10^3$ g/mol, in Different Conformational States

conformer no.	conformational state	R_s, Å[a]	$[\eta]$, mL/g[b]	$[\Theta]_{220}$, (deg cm^2)/dmol[b]	secondary structure: α-helix	secondary structure: β-structure	relative accessibility of Trp residues for quencher[b]
1	*random coil in "good"[c] solvent*, pH > 5.7, 10 M urea	65.6	36.7	0	0	0	*well* accessible
	random coil in "poor"[c] solvent pH > 5.7, 0 M urea	58.1	26.9	0	∼2	∼4	*well* accessible
2	*intermediate state* pH 5.1−5.2, 0 M urea	47.0	13.5	-8.5×10^3	∼28.5	∼18	*well* accessible
3	*compact state* pH < 4.3, [urea] < 9.5 M	29.9[d]	3.7	-35.5×10^3	∼80	∼20	*poorly* accessible

10 M urea. This increase in $[\eta]$ values reflects the normal swelling of an unfolded polypeptide chain in good solvent.[35,51] Table 1 shows that the copolymer intrinsic viscosity decreases with the decrease of pH value: $[\eta]$ is equal to 26.9, 13.5, and 3.7 mL/g in the random coil (pH > 5.7), intermediate state (pH 5.1−5.2), and compact state (pH 4.2−4.3), respectively (see Table 1). It is necessary to emphasize that further acidification of copolymer solution resulted in the considerable increase of the $[\eta]$ value (data not shown), reflecting association and aggregation of polymer molecules. This observation is in good agreement with earlier investigations of analogous copolymers of lower molecular weight.[15−17] It means that the second conformational transition, induced in a random copolymer of hydrophobic and hydrophilic amino acid residues, is not attributed to the association of polymer molecules and, most probably, describes an intramolecular process.

It is well-known that a correlation exists between the compactness of a protein (or polypeptide) molecule and the accessibility of its Trp residues to molecules of a quencher of Trp fluorescence.[57] Here we employed a method of quenching of Trp fluorescence by acrylamide (neutral quencher). To this end, intensities of Trp fluorescence in the absence and presence of 0.4 M acrylamide were compared. The efficiency of quenching of fluorescence, E (see Experimental Section), virtually does not change in the course of a pH decrease from neutral values to about pH 5.15 ($E \approx 30\%$) but drops abruptly (by ∼8%) in the pH range from ∼5.15 to ∼4.3 and then becomes constant again (see Figure 5, filled squares). Such behavior of Trp fluorescence in the course of pH change may point to the occurrence of at least one conformational state, whose Trp residues are less accessible to the quencher than those of the unfolded polypeptide. It should be emphasized that the beginning of E decrease (upon pH decrease) coincides well with the beginning of the second conformational transition (pH ∼ 5.1−5.4) detected by far-UV CD and R_s measurements (see Figure 5). pH values at semitransition points obtained by three different physical methods (CD, SEC, quenching of fluorescence) are close. The calculated mean semitransitional pH values ($pH_{1/2}^{calc}$) are 5.45 ± 0.14 for the first conformational transition and 4.87 ± 0.02 for the second one. *It should also be noted that the pH-induced transitions just described are fully reversible.*

Thus, we can conclude from the above findings that the process of pH-induced compaction of the random copolymer under investigation consists of at least two distinct stages; i.e., there exist at least three distinct averaged conformational states for macromolecules of our copolymer: unfolded (pH > 5.7), intermediate (pH 5.1−5.2), and compact (pH < 4.3) states. The properties of these states are summarized in Table 1 (for details see below). Further, we tried to characterize these three averaged conformational states in more detail.

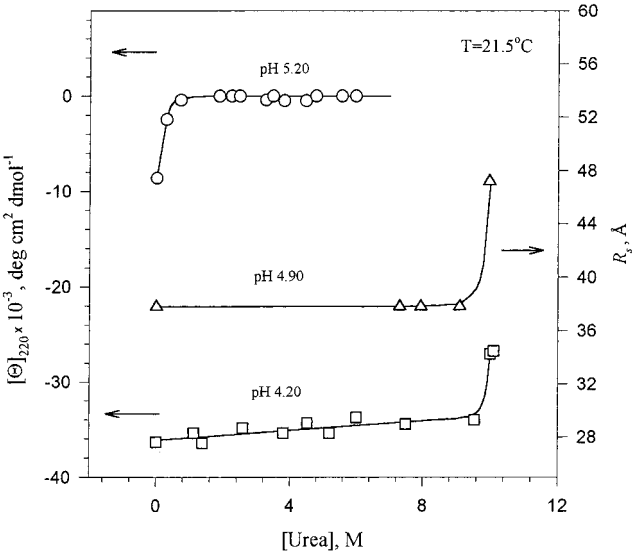

Figure 6. Urea-induced unfolding/refolding of the compact (pH 4.20; squares) and intermediate (pH 5.20; circles) states of the random copolymer Glu:Leu:Trp (82.5:16:1.5, fraction with $M \sim 48.8 \times 10^3$ g/mol) monitored by changes in molar ellipticity $[\Theta]_{220}$ at room temperature. Triangles present urea-induced unfolding of the copolymer molecules at pH 4.90 (i.e., at the semitransition point of the transition between the intermediate and compact states) monitored by measurements of the molecular Stokes radius (R_s).

3.3. Comparative Stability Studies. Effect of Urea. Studies on the stability of the polymer conformers toward denaturant (urea, GdmCl) action were carried out at 22 °C. These studies showed that the secondary structure of the intermediate state (pH 5.20) can be destroyed by the addition of relatively small amounts (∼0.7 M) of urea (Figure 6, circles). In contrast, the compact state (pH 4.20) is highly resistant to large concentrations of the denaturing agent, and its secondary structure only starts to disintegrate at extremely high denaturant concentrations (10 M urea) (Figure 6, squares). Interestingly, the copolymer molecules at pH 4.90 (i.e., at the semitransition point of the transition between intermediate and compact states) behave similarly, when unfolded by urea (or GdmCl) with the detection of molecular sizes by R_s (Figure 6, triangles): the macromolecules unfold in 10 M urea (or 8 M GdmCl) merely to the dimensions characteristic of the intermediate state (47 Å) (see Table 1, entry 2, and Figure 5, at pH ∼ 5.15). Such strong resistance of the compact state to urea unfolding correlates well with the experimental results of Scholtz *et al.*[58] (plotted in Figure 7 by filled circles). Figure 7 shows the dependence of semitransitional urea concentration on the logarithm of the number of amino acid residues (N) in helical polypeptides of the general sequence Ac-Tyr-(Ala-Glu-

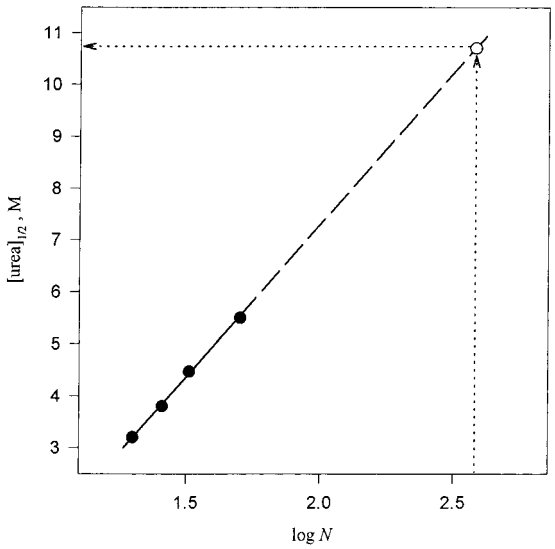

Figure 7. Dependence of semitransitional urea concentration ($[urea]_{1/2}$) on the logarithm of the number of amino acid residues (N) for helical polypeptides of the general sequence Ac-Tyr-(Ala-Glu-Ala-Ala-Lys-Ala)$_k$-Phe-NH_2. The dependence was constructed using the experimental data given in ref 38 (filled circles). Our extrapolation to higher N values performed for the random copolymer Glu:Leu:Trp (82.5:16:1.5, fraction with $M \sim 48.8 \times 10^3$ g/mol) is shown by the dashed line, an open circle, and arrows. This extrapolation gives us the $[urea]_{1/2}$ value of 10.7 M for our ~380-residue polypeptide.

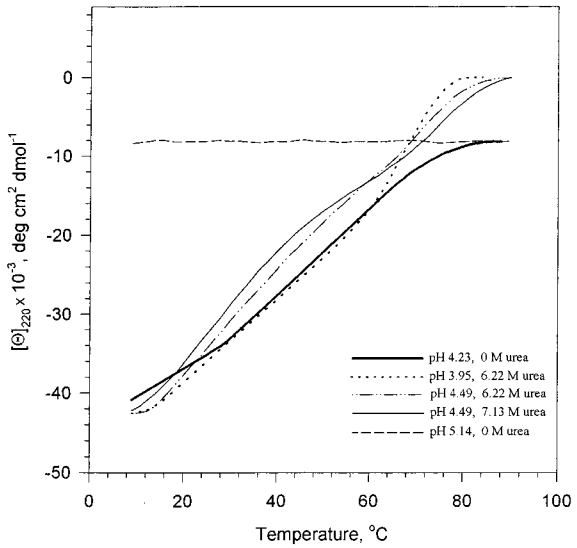

Figure 8. Changes in molar ellipticity $[\Theta]_{220}$ of the copolymer molecules at pH 5.14 and 0 M urea (dashed line), pH 4.23 and 0 M urea (bold solid line), pH 3.95 and 6.22 M urea (dotted line), pH 4.49 and 6.22 M urea (dashed-and-dotted line), and pH 4.49 and 7.13 M urea (thin solid line) upon heating/cooling.

Ala-Ala-Lys-Ala)$_k$-Phe-NH_2 (see ref 18). The extrapolation to higher N for our approximately 380-residue polypeptide gave us the semitransitional value of urea concentration ($[urea]_{1/2}$) of 10.7 M (an open circle in Figure 7). *It should be emphasized here that the described denaturant-induced transitions are completely reversible as well.*

3.4. Comparative Stability Studies. Effect of Temperature. Figure 8 depicts melting of the copolymer's secondary structure monitored by changes in far-UV CD at different pH and urea concentration values. It is necessary to emphasize here that the secondary structure content of the intermediate state itself (pH 5.14, 0 M urea; dashed line in Figure 8) remains virtually constant within the whole temperature interval explored. In contrast, the secondary structure of the compact state decreases significantly upon heating. As can be seen from Figure 8, the temperature-induced reduction in the $[\Theta]_{220}$ value of the compact state (pH 4.23) in the absence of urea (bold solid line in Figure 8) continues only up to the value typical of the intermediate state ($[\Theta]_{220} \approx -8.5 \times 10^3$ deg·cm·dmol^{-1}), and occurs in a wide temperature interval (the transition begins at ~30 °C and ends at ~70 °C). When urea is added to the solution of the compact copolymer (pH ~ 4.0−4.5), the picture changes drastically: now two, instead of one, temperature-induced transitions take place. The latter transition corresponds to the complete unfolding of the polypeptide chain (*intermediate state → unfolded state*), and occurs in the temperature interval from ~66 to ~88 °C. It is seen also that as the urea concentration rises, the steepness (cooperativity) of the temperature-induced transition between the compact and intermediate states increases somewhat, whereas the steepness of the transition between the intermediate and completely unfolded states is somewhat reduced. *It should be emphasized, as before, that the temperature-induced transitions just described are fully reversible.*

Thus, the data on temperature-induced unfolding/refolding of the copolymer macromolecules in the compact state also confirm the existence of an equilibrium intermediate between the compact and fully unfolded states.

4. Conclusions

From the data presented in this paper, a conclusion can be drawn that pH-induced compaction of the high molecular mass random copolymer Glu:Leu:Trp (82.5:16:1.5) is at least a two-stage process. General structural properties of the observed conformers (see Table 1) have been revealed by a few different physical methods. The intermediate state is realized during pH-induced compaction of the unfolded polypeptide (at pH ~ 5.1−5.2, 0 M urea), or during temperature-induced unfolding of the compact state (e.g., at pH < 4.3, 0 M urea, and high temperature (>80 °C)). Copolymer molecules in this state are partially collapsed (R_s decreases from ~58 to ~47 Å). At the same time, their Trp residues and those of the copolymer in the unfolded state are equally accessible for the quencher molecules. At these conditions, the polypeptide has a relatively well-developed secondary structure (~30% α-helices and ~20% β-structure). At pH < 4.3, the compact state is observed. As was shown earlier,[15−17] the copolymer in this state should be almost as compact as native globular proteins. This is in good agreement with our results on intrinsic viscosity ($[\eta]$ = 3.7 mL/g) and acrylamide quenching of Trp fluorescence. Finally, the polypeptide in the compact state has a highly ordered secondary structure (~80% α-helices and ~20% β-structure).

It has been shown that the compact state of the copolymer is extremely resistant to the denaturing action of urea, while the less compact intermediate state can be totally unfolded by small amounts of urea. We also have shown that, in the absence of urea, a temperature rise (10−90 °C) does not induce any changes in secondary structure content of the copolymer macromolecules in the intermediate state, whereas the compact state can be unfolded by temperature, but not completely: only to the less compact intermediate state.

The presence of urea imparts complex character to the melting curve: melting (as well as cooling) takes place in two stages. Thus, melting of the compact state in the presence of urea confirms the existence of a folding intermediate. In our opinion, the compact state of the copolymer might be considered as an analogue of the molten globule state[3] of proteins (see above), and the less compact intermediate state of our copolymer might be viewed as a putative analogue of the premolten globule state[3,32−39] of proteins (see above).

Acknowledgment. We are grateful to Dr. V. N. Medvedkin, Profs Yu.V. Mitin, and O. B. Ptitsyn for numerous helpful discussions and valuable comments during this work. We thank Dr. L. F. Markova for the help in performing amino acid analyses. This work was supported in part by the Russian Foundation for Basic Research (Grant No. 93-04-6635).

References and Notes

(1) Ptitsyn, O. B. In *Conformation in Biology*; Srinivasan, R., Sarma, R. H., Eds; Adenine Press: New York, 1983; p 49.
(2) Lau, K. F.; Dill, K. A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 638.
(3) Ptitsyn, O. B. *Adv. Protein Chem.* **1995**, *47*, 83.
(4) Finkelstein, A. V. In *Structure and Biosynthesis of Protein, issue 3*; Spirin, A. S., Ed.; Scientific Centre for Biological Research: Pushchino: Russia, 1988; p 10.
(5) Finkelstein, A. V.; Ptitsyn, O. B. *Prog. Biophys. Mol. Biol.* **1987**, *50*, 171.
(6) Ptitsyn, O. B.; Finkelstein, A. V. *Q. Rev. Biophys.* **1980**, *13*, 339.
(7) Sasai, M. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 8438.
(8) Brass, A.; Fiszer, A.; Clamp, M. *Biochem. Soc. Trans.* **1995**, *23*, 715.
(9) Galzitskaya, O. V.; Finkelstein, A. V. *Protein Eng.* **1995**, *8*, 883.
(10) Shakhnovich, E. I.; Gutin, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7195.
(11) Dinner, A. R.; Sali, A.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8356.
(12) Mirny, L. A.; Abkevich, V.; Shakhnovich, E. I. *Folding Design* **1996**, *1*, 103.
(13) Ptitsyn, O. B.; Volkelstein, M. V. *Biomol. Struct. Dyn.* **1986**, *4*, 137.
(14) Bychkova, V. E.; Gudkov, A. T.; Miller, W. G.; Mitin, Yu. V.; Ptitsyn, O. B.; Shpungin, I. L. *Biopolymers* **1975**, *14*, 1739.
(15) Anufrieva, E. V.; Bychkova, V. E.; Krakovyak, M. G.; Pautov, V. D.; Ptitsyn, O.B. *FEBS Lett.* **1975**, *55*, 46.
(16) Bychkova, V. E.; Semisotnov, G. V.; Ptitsyn, O. B.; Gudkova, O. V.; Mitin, Yu.V.; Anufrieva, E. V. *Mol. Biol. (USSR)* **1980**, *14*, 278.
(17) Semisotnov, G. V.; Zikherman, K. Kh.; Kasatkin, S. B.; Ptitsyn, O. B.; Anufrieva, E. V. *Biopolymers* **1981**, *20*, 2287.
(18) Rao, S. P.; Carlstrom, D. E.; Miller, W. G. *Biochemistry* **1974**, *13*, 943.
(19) Davidson, A. R.; Lumb, K. J.; Sauer, R. T. *Nature Struct. Biol.* **1995**, *2*, 856.
(20) Sauer, R. T. *Folding Design* **1996**, *1*, R27.
(21) Tiktopulo, E. I.; Uversky, V. N.; Lushchik, V. B.; Klenin, S. I.; Bychkova, V. E.; Ptitsyn, O. B. *Macromolecules* **1995**, *28*, 7519.
(22) Meewes, M.; Ricka, J.; Desilva, M.; Nyffenegger, R.; Binkert, T. *Macromolecules* **1991**, *24*, 5811.
(23) Ptitsyn, O. B. In *Protein Folding*; Creighton, T. E., Ed.; Freeman: New York, 1992; p 243.
(24) Uversky, V. N. *Biochemistry* **1993**, *32*, 13288.
(25) Uversky, V. N. *Int. J. Bio-Chromatogr.* **1994**, *1*, 103.
(26) Eliezer, D.; Chiba, K.; Tsuruta, H.; Doniach, S.; Hodson, K. O.; Kihara, H. *Biophys. J.* **1993**, *65*, 912.
(27) Kataoka, M.; Hagihara, Y.; Mihara, K.; Goto, Y. *Proteins: Struct. Funct. Genet.* **1993**, *229*, 591.
(28) Kataoka, M.; Kuwajima, K.; Tokunaga, F.; Goto, Y. *Protein Sci.* **1997**, *6*, 422.
(29) Semisotnov, G. V.; Kihara, H.; Kotova, N. V.; Kimura, K.; Amemiya, Y.; Wakabayashi, K.; Serdyuk, I. N.; Timchenko, A. A.; Chiba, K.; Nikaido, K.; Ikura, T.; Kuwajima, K. *J. Mol. Biol.* **1996**, *262*, 559.
(30) Finkelstein, A. V.; Shakhnovich, E. I. *Biopolymers* **1989**, *28*, 1681.
(31) Kharakoz, D. P.; Bychkova, V. E. *Biochemistry* **1997**, *36*, 1882.
(32) Uversky, V. N.; Semisotnov G. V.; Ptitsyn O. B. *Biofizika (Rusian)* **1993**, *38*, 37.
(33) Palleros, D. R.; Shi, L.; Reid, K. L.; Fink, A. L. *Biochemistry* **1993**, *32*, 4214.
(34) Uversky, V. N.; Ptitsyn, O. B. *Biochemistry* **1994**, *33*, 2782.
(35) Uversky, V. N.; Ptitsyn, O. B. *J. Mol. Biol.* **1996**, *255*, 215.
(36) Uversky, V. N.; Ptitsyn, O. B. *Mol. Biol. (Russian)* **1996**, *30*, 1124.
(37) Uversky, V. N.; Ptitsyn, O. B. *Mol. Biol. (Russian)* **1996**, *30*, 1135.
(38) Zhang, Y.-L.; Zhou, J.-M.; Tsou, C.-L. *Biochem. Biophys. Acta* **1996**, *1295*, 239.
(39) Uversky, V. N. *Mol. Biol.*, in press (in Russian).
(40) Dolgikh, D. A.; Kolomiets, A. P.; Bolotina, I. A.; Ptitsyn, O. B. *FEBS Lett.* **1984**, *165*, 88−92.
(41) Semisotnov, G. V.; Rodionova, N. A.; Kutyshenko, V. P.; Ebert, B.; Blank, J.; Ptitsyn, O. B. *FEBS Lett.* **1987**, *224*, 9.
(42) Ptitsyn, O. B.; Pain, R. H.; Semisotnov, G. V.; Zerovnik, E.; Razgulyaev, O. I. *FEBS Lett.* **1990**, *262*, 20.
(43) Goldberg, M. E.; Semisotnov, G. V.; Friguet, B.; Kuwajima, K.; Ptitsyn, O. B.; Sugai, S. *FEBS Lett.* **1990**, *263*, 51.
(44) Ptitsyn, O. B.; Semisotnov, G. V. In *Conformations and Forces in Protein Folding*; Nall, B. T.; Dill, K. A.; Eds.; AAAS Press: Washington, DC, 1991; p 155.
(45) Karnoup, A. S. Diploma Thesis, Institute of Protein Research, Russian Academy of Sciences, Pushchino, Russia 1995.
(46) Blacklock, T.; Hirschmann, R.; Veber, D. In *The Peptides. Analysis, Synthesis, Biology*, Udenfriend, S., Meienhofer, J., Eds; Academic Press: New York, 1993; Vol. 9, Part C, pp 39−102.
(47) Greenstein, J.; Winitz, M. *Chemistry of the Amino Acids*; John Wiley & Sons, Inc.: New York, 1961.
(48) Farthing, A. C. *J. Chem. Soc.* **1950**, 3213.
(49) Osterman, L. A. *Chromatography of Proteins and Nucleic Acids*; Nauka: Moscow, 1985; pp 111, 149−151.
(50) Laurent, T. C.; Killander, J. *J. Chromatogr.* **1964**, *14*, 317.
(51) Tanford, C. *Adv. Protein Chem.* **1968**, *23*, 121.
(52) Venyaminov, S. Yu.; Gudkov, A. T.; Gogia, Z. V.; Tumanova, L. G. *Absorption and Circular Dichroism Spectra of Individual Proteins from Escherichia Coli Ribosomes* (communication); Institute of Protein Research, Russian Academy of Sciences, Scientific Centre for Biological Research: Pushchino, Russiam, 1981.
(53) Sidorov, I. A. *REGRESSION* 1996, ver. 6.3.
(54) Katamari, Y. O.; Konno, T.; Kataoka, M.; Akasaka, K. *J. Mol. Biol.* **1996**, *259*, 512.
(55) Ackers, G. K. *J. Biol. Chem.* **1967**, *242*, 3247.
(56) Corbett, R. J. T.; Roche, R. S. *Biochemistry* **1984**, *23*, 1888.
(57) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; Plenum Press: New York and London, 1983.
(58) Scholtz, J. M.; Barrick, D.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 185.

MA9700338