



ELSEVIER

Biochimica et Biophysica Acta 1432 (1999) 324–332



www.elsevier.com/locate/bba

## Structure and stability of recombinant protein depend on the extra N-terminal methionine residue: S6 permutein from direct and fusion expression systems

Vladimir N. Uversky<sup>a,b,\*</sup>, Ziedulla Kh. Abdullaev<sup>c</sup>, Alexander S. Arseniev<sup>c</sup>, Eduard V. Bocharov<sup>c</sup>, Dmitry A. Dolgikh<sup>c</sup>, Ramil F. Latypov<sup>d</sup>, Tatyana N. Melnik<sup>d</sup>, Konstantin S. Vassilenko<sup>d</sup>, Mikhail P. Kirpichnikov<sup>c</sup>

<sup>a</sup> Institute for Biological Instrumentation, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia

<sup>b</sup> Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA 95064, USA

<sup>c</sup> Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 117871 Moscow, Russia

<sup>d</sup> Institute of Protein Research, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia

Received 27 January 1999; received in revised form 25 March 1999; accepted 21 April 1999

### Abstract

Two permuted variants of S6 ribosomal protein were obtained in direct and fusion expression systems, respectively. The product of direct expression contained the extra N-terminal methionine residue. The structural properties and conformational stability of these permuteins were compared using 1-D <sup>1</sup>H-NMR, circular dichroism, intrinsic fluorescence, differential scanning calorimetry and resistance to urea-induced unfolding. A pronounced difference in all the parameters studied has been demonstrated. This means that the structure of recombinant protein can be sensitive to peculiarities of the expression and purification procedures, leading particularly to the presence or absence of the Met at the first position in the target protein sequence. © 1999 Elsevier Science B.V. All rights reserved.

**Keywords:** Permuted protein; Recombinant protein; Protein structure; Conformational stability

### 1. Introduction

Extremely rapid development of recombinant DNA techniques during the last decades has generated cloned genes for numerous proteins. At present, this technology has attracted special attention in the search for a way to obtain large quantities of protein of interest. The crucial step in this direction will be

the expression of a foreign gene in a heterologous cell. Though the large variety of expression systems dealing with different cells and organisms has been elaborated, the most frequently used approach is the expression of a target gene in *Escherichia coli* [1]. In addition to the great number of obvious advantages of this system, it has some important drawbacks, including the propensity of foreign protein to form inclusion bodies during intracellular expression [2], and rapid degradation of foreign proteins in the new host if they are not protected from intracellular proteases [3]. It is a widespread belief that these problems can be solved if the target protein is syn-

\* Corresponding author. Present address: Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA 95064, USA. Fax: +1-831-459-2935; E-mail: uversky@hydrogen.ucsc.edu

thesized not from its own N-terminus, but as a fusion protein (i.e. as a hybrid with another protein attached to the amino-terminus) [3]. The other attractive feature of this method is that the fusion protein can be more easily purified [3].

In the present paper, we attempt to answer the question whether the difference in the expression system and in the purification procedure can affect the structural properties and conformational stability of recombinant protein. To this end, the gene coding the permuted amino acid sequence of ribosomal protein S6 from *Thermus thermophilus* was cloned into two different vectors (ensuring direct and fusion expression) and expressed in *E. coli*. The structural properties and conformational stability of the obtained products were compared by several physicochemical methods.

## 2. Materials and methods

### 2.1. Protein expression and isolation

The gene of wild-type S6 protein was cloned into expression vector pGEMEX-1 (Promega), while coding the permuted S6 protein was cloned into the same vector and also into vector pMAL-c (New England Biolabs). Both these vectors were expressed in *E. coli*. Wild-type and permuted proteins were isolated as described elsewhere [4].

As it has been reported earlier [4,5], we have constructed the gene coding the circularly permuted amino acid sequence of *T. thermophilus* ribosomal protein S6. The goal of such permutation was to obtain the protein with an unusual  $\alpha\beta\alpha\beta\beta$  fold which was suspected of being the target topology for the de novo protein albebetin [6,7]. The following changes were induced in *wtS6* in order to have *cpS6* protein: (1) first four C-terminal residues were removed together with its N-terminal Met and Arg; (2) the polypeptide chain was cut between Asn-13 and Leu-14; and (3) the old termini were linked through an Ala-Ser-Thr-Thr-Pro-Gly loop.

To have large quantities of proteins, the gene of wild-type S6 protein was cloned into expression vector pGEMEX-1 (Promega), whereas the gene coding the *cpS6* sequence was cloned into the two different vectors, pMAL-c and pGEMEX-1. The first vector

expresses the target product as a fusion with the maltose-binding protein which can be then split off. The second vector expresses *cpS6* directly from the T7 phage promoter (in this case, an additional ATG codon corresponding to the N-terminus Met residue was included into the gene). Nucleotide sequences of the genes were confirmed by DNA sequencing. The genes were expressed in *E. coli*. As the result, two forms of the permuted protein *cpS6f* and *cpS6d* (here symbols *f* and *d* correspond to the fusion and direct expression) were obtained. A detailed description of the design, genetic, cloning and expression procedures as well as the process of protein purification were presented in our previous papers [4,5]. HPLC and PAGE were used to analyze the protein purity, which was no less than 95% according to these methods. N-terminal sequence analysis has shown that *cpS6d* protein contained the extra methionine residue at the first position, while the *cpS6f* variant began from Leu.

### 2.2. Circular dichroism studies

Circular dichroism studies were carried out using Jasco-600 spectropolarimeter (Japan Spectroscopic, Tokyo, Japan), equipped with a temperature-controlled holder. The protein concentration for near and far UV CD measurements was 0.8 mg/ml. The cell pathlength was 0.19 and 10 mm for far and near UV CD measurements, respectively.

### 2.3. Microcalorimetry measurements

Calorimetric measurements were done with a new precise differential adiabatic scanning microcalorimeter SCAL-1 (Scal, Pushchino, Russia), equipped with a newly designed calorimetric unit with cylindrical glass cells (0.34 ml) providing high chemical resistance and excellent dynamic characteristics. The protein concentration for microcalorimetric measurements was 2.1 mg/ml; the scanning rate was 0.9 K/min. The van 't Hoff and calorimetric enthalpies of melting were calculated as described in [8,9].

### 2.4. $^1\text{H-NMR}$ measurements

$^1\text{H-NMR}$  spectra were recorded at 30°C in a 50 mM aqueous solution of ammonium citrate, pH

7.9, containing 10% D<sub>2</sub>O, in a UNITY-600 spectrometer (Varian) at a working proton frequency of 600 MHz.

### 2.5. Fluorescence measurements

Fluorescence measurements were carried out using the Aminco (SPF-1000cs) corrected spectrofluorimeter (American Instrument, Silver Spring, MD, USA) equipped with a temperature-controlled holder. Fluorescence was excited at 290 nm. Protein concentration was 0.01 mg/ml. Cell pathlength was 10.0 mm. Measurements were carried out at 25°C.

### 2.6. Urea-induced unfolding

Urea-induced unfolding of permuted protein was studied by the characteristic red shift of the tryptophan fluorescence spectrum maximum [10]. The unfolding curve was analyzed in accordance with type I analysis [11,12], or type II analysis [13–15]. Type I analysis was done assuming an all-or-none mechanism of the transition between the initial (native) and final (unfolded) states. In this case it is possible to estimate the free energy of the initial state relative to that of the unfolded one,  $\Delta G_{\text{fold}}$ , under the given conditions as in [11]:

$$\Delta G_{\text{fold}} = -RT \ln K_{\text{eq}} \quad (1)$$

where  $K_{\text{eq}}$  is the equilibrium constant of the transition. The value of  $\Delta G_{\text{fold}}$  within the transition region, where it can be measured, varies linearly with the denaturant concentration [12]:

$$\Delta G_{\text{fold}} = \Delta G_{\text{fold}}^{\text{H}_2\text{O}} + m [\text{denaturant}]. \quad (2)$$

The parameter  $m$  reflects the dependence of the free energy on the denaturant concentration, [denaturant]; whereas the values of  $\Delta G_{\text{fold}}$  are usually extrapolated linearly to determine its value in the absence of denaturant,  $\Delta G_{\text{fold}}^{\text{H}_2\text{O}}$ .

In type II analysis the value of  $\Delta \nu^{\text{eff}}$  (which is the difference in the number of denaturant molecules 'bound' to one protein molecule in its two states) is determined from the unfolded curve by equation [13]

$$\Delta \nu^{\text{eff}} = 4a_t \left( \frac{\partial \Theta}{\partial a} \right) a = a_t \quad (3)$$

where  $a_t$  is the denaturant activity at midpoint, and  $\Theta$  is the fraction of molecules in the unfolded state.

## 3. Results

It was already noted that there is a difference in the *cpS6d* and *cpS6f* amino acid sequences, namely *cpS6d* has an extra Met residue at the first position (see Section 2). The presence of zero formyl-methio-

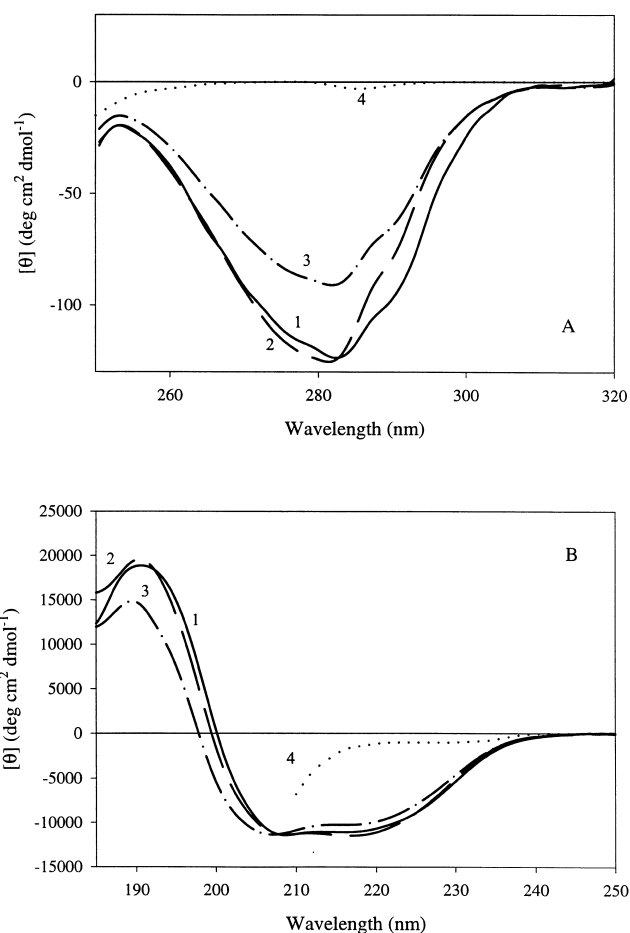


Fig. 1. Near (A) and far UV CD spectra (B) of ribosomal S6 protein from *Thermus thermophilus* (1) and its permuted variants *cpS6d* (2) and *cpS6f* (3) obtained in direct and fusion expression systems, respectively. Spectra of completely unfolded by 8 M GdmCl protein are also shown (4). Measurements were carried out at 25°C in 20 mM sodium phosphate buffer, pH 7.4. Protein concentration was 1.0 mg/ml. Cell pathlength was 10.0 and 0.19 mm, for near and far UV CD measurements, respectively.

nine is a common feature of a recombinant protein expressed in *E. coli* [16]. In the case of *cpS6f*, this methionine is split off together with the maltose-binding protein. On the other hand, the attached maltose-binding protein can affect the in vivo folding and the structure of *cpS6f*. The structural perturbations in the target protein can be retained after the digestion of maltose-binding protein. Both these factors may obviously lead to the appearance of the structural divergence between the studied proteins. Data presented below were addressed to answer the questions: Is there any difference between *cpS6d* and *cpS6f* structural properties and conformational stability? And if so, why do the products of different expression systems behave differently?

### 3.1. Near UV CD spectra

A rigid asymmetric environment of aromatic amino acid residues gives rise to a pronounced CD spectrum in the near UV region [17]. The studied proteins had almost identical aromatic amino acid content (Trp-1/1 (*wtS6/cpS6*), Tyr-5/5 (*wtS6/cpS6*) and Phe-2/1 (*wtS6/cpS6*), sole Phe-97 was substituted by Gly during mutagenesis. Fig. 1A represents the near UV CD spectra of two permuted variants (*cpS6f* and *cpS6d*) and *wtS6* protein. The spectrum of protein completely unfolded by 8 M GdmCl is also shown for comparison. All three proteins are characterized by intense spectra, reflecting the fact that they have a rigid tertiary structure. However, Fig. 1 clearly shows that there is a remarkable difference between *cpS6f* and *cpS6d*. It can be seen that spectra of permutein from the direct expression system and wild-type protein are close in their intensities. Some discrepancy between the long-wave parts of the spectra can be attributed to the appearance of a permutation-induced difference in the Trp environment. On the other hand, the *cpS6f* spectrum has a shape similar to that of *cpS6d*, while its intensity is about 1.35-times lower. This suggests that the structure of permutein from the fusion expression system is somewhat disturbed as compared with that of *cpS6d* and *wtS6*.

### 3.2. Far UV CD spectra

Fig. 1B shows the far UV CD spectra of the proteins and gives an additional confirmation of the

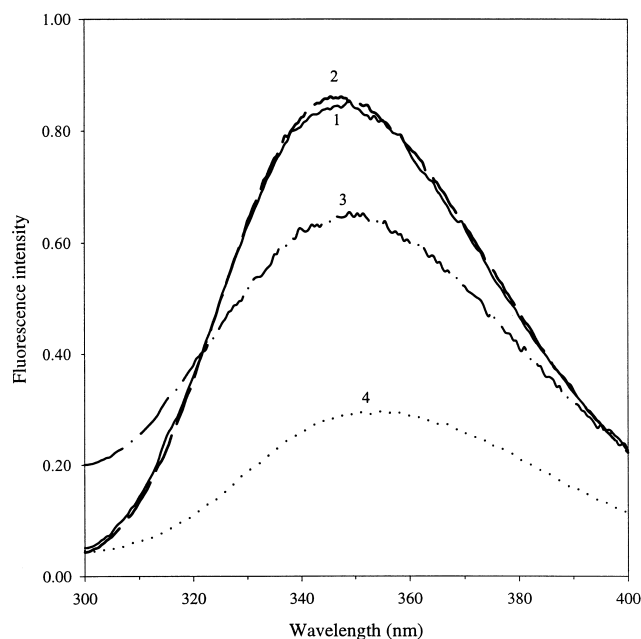


Fig. 2. Tryptophan fluorescence spectra measured for *wtS6* (1) protein, its permuted variants *cpS6d* (2) and *cpS6f* (3), as well as for the completely unfolded by 8 M GdmCl protein (4). Measurements were carried out at 25°C in 20 mM sodium phosphate buffer, pH 7.4. Protein concentration was 0.01 mg/ml. Cell pathlength was 10.0 mm. Fluorescence was excited at 290 nm.

above suggestion. It is seen that the spectra of *cpS6d* and *wtS6* coincide almost completely, assuming the identity of the secondary structure content in these two proteins. Analysis of the *cpS6f* far UV CD spectrum shows that the secondary structure of this protein is slightly more disordered than that of the wild-type protein and S6 from the direct expression system. This conclusion follows from the pronounced shift of the intersection point to a shorter wavelength, an overall decrease of spectrum intensity and the increase of  $[\theta]_{208}/[\theta]_{222}$  ratio (the value of this parameter is 1.07, 1.04 and 1.16 for the wild-type, *cpS6d* and *cpS6f*, respectively) [17].

### 3.3. Tryptophan fluorescence spectra

The position and intensity of the tryptophan fluorescence spectrum is sensitive to the surroundings of this chromophore group [18]. For example, it has been established that the fluorescence spectrum of the solvated tryptophan residue has a maximal value in the vicinity of 350 nm, whereas its embedding into

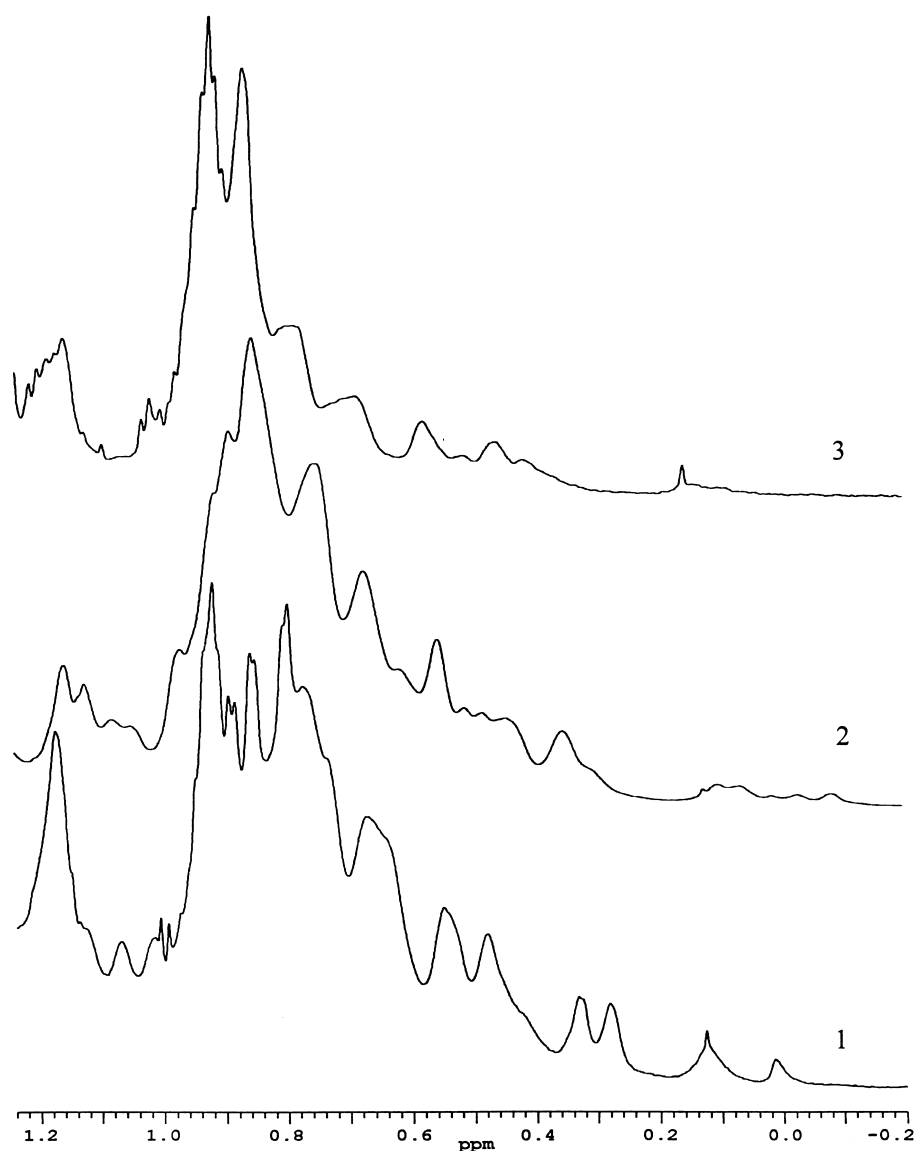


Fig. 3.  $^1\text{H}$ -NMR spectra of *wtS6* (1), *cpS6d* (2) and *cpS6f* (3) proteins. Measurements were carried out at 30°C in 50 mM aqueous solution of ammonium citrate, pH 7.9, containing 10%  $\text{D}_2\text{O}$ .

the non-polar environment leads to the characteristic blue shift of the maximum position [7,19–21]. This observation is commonly used for investigation of structural transformations of a protein molecule. Fig. 2 represents the Trp fluorescence spectrum of wild-type S6 protein as well as that of its two permuted forms. The spectrum of the completely unfolded by 8 M GdmCl S6 protein is also shown for comparison. It is seen that the fluorescence spectrum of wild-type S6 protein is relatively red shifted ( $\lambda_{\text{max}} = 345$  nm). This observation is consistent with

the assumption that Trp residues of the protein are relatively exposed to solvent. Nevertheless, the unfolding of S6 protein is accompanied by a pronounced long-wave shift of the fluorescence spectrum maximum (from 345 to 355 nm) and by about a 3-fold decrease in fluorescence intensity. Fig. 2 clearly shows that the spectra of *cpS6d* and *wtS6* concur, reflecting the fact that the Trp residues of these two proteins have very similar environments. It should be emphasized that the fluorescence behavior of *cpS6f* is quite different from that of *wtS6* and

*cpS6d*: the spectrum is characterized by a visible shift of maximum towards a longer wavelength ( $\sim 348$  nm) and by an approximately 1.3-fold intensity decrease. Both these facts verify the existence of some perturbations in the microenvironment of the Trp residues in S6 permutin from the fusion expression system and show their increased accessibility to the solvent.

### 3.4. $^1\text{H-NMR}$ spectroscopy data

The  $^1\text{H-NMR}$  spectra of wild-type S6 protein and its two permuted forms exhibited a number of specific features of native globular proteins. Particularly, the dispersion of chemical shifts for the amide and  $\text{C}\alpha$ -protons points in favor of a pronounced secondary structure in all three proteins. The methyl regions of the  $^1\text{H-NMR}$  spectra of these proteins are presented in Fig. 3. Analysis of these spectra suggests that the values of the chemical shifts for methyl groups and their dispersion in all the forms of S6 protein are characteristic of proteins with a globular tertiary structure, where the aromatic residues and methyl groups are close to each other. On the other hand, the increased half-width of the signals from the

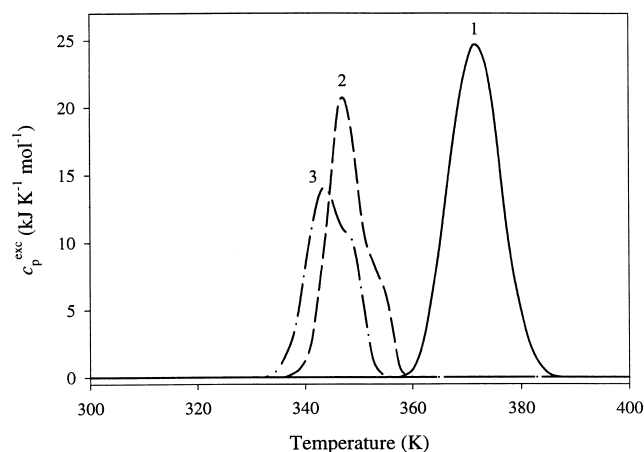


Fig. 4. Excess heat capacity curves measured for *wtS6* (1), *cpS6d* (2) and *cpS6f* (3) proteins in 20 mM sodium phosphate buffer, pH 7.4. The protein concentration was  $\sim 2.1$  mg/ml.

polypeptide backbone protons (e.g. from  $\text{C}\alpha$ -protons) may result from the increased flexibility of the tertiary structure in the millisecond range, or from protein association. Finally, Fig. 3 shows that all three proteins are characterized by different spectra. This is a very important finding as it allows conclude that the structures of *cpS6d* and *cpS6f* are different.

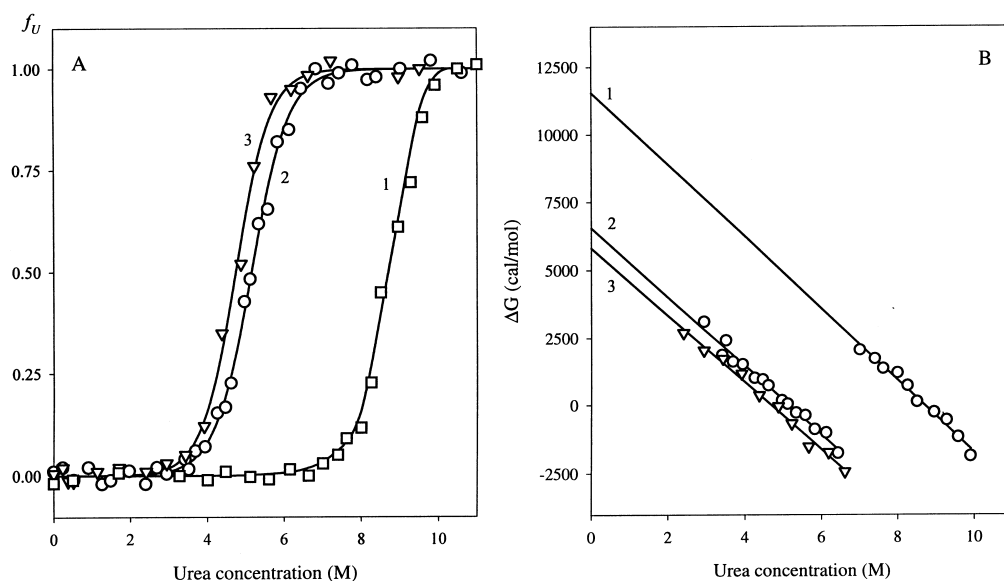


Fig. 5. Urea-induced unfolding of *wtS6* (1), *cpS6d* (2) and *cpS6f* (3) proteins in 20 mM sodium phosphate buffer, pH 7.4. (A) Normalized unfolding curves determined from the denaturant-induced changes of Trp fluorescence spectrum of the proteins as the dependence of parameter  $A$  on urea concentration. Protein concentration was 0.01 mg/ml. Fluorescence was excited at 290 nm. (B)  $\Delta G_{\text{fold}}$  for *wtS6*, *cpS6d* and *cpS6f* proteins as a function of urea concentration. The experimental data were calculated from the data presented in A using Eq. 1 of the text. The thermodynamic parameters estimated from these data are given in Table 1.

### 3.5. DSC data

Effect of temperature on *wtS6*, *cpS6d* and *cpS6f* structures was studied by means of differential scanning calorimetry, a method which is one of the most powerful approaches in protein stability investigations [5,6,22]. Fig. 4 collects the temperature dependencies of excess heat capacity (calorimetric curves) measured for all three proteins. The considerable difference in the calorimetric behavior of these proteins can be seen. Indeed, *wtS6* is characterized by an extremely high resistance to the temperature ( $T_m = 372$  K) which is expected for protein from thermophilic organisms. On the other hand, Fig. 4 shows that permutation causes considerable destabilization of the protein molecule ( $T_m = 347$  and  $342$  K for *cpS6d* and *cpS6f*, respectively). Interestingly, calorimetric analysis confirmed the considerable effect of the expression method on the protein structure, *cpS6f* was shown to be less stable than *cpS6d*.

### 3.6. Urea-induced unfolding

Studies of protein stability to the urea- or GdmCl-induced unfolding are frequently used to characterize natural, mutated and artificial proteins. Such analysis yields important information on the conformational stability of a protein in the absence of denaturant,  $\Delta G_{fold}^{H_2O}$  (see above).

We analyzed the unfolding behavior of S6 protein and its permuted variants from the urea-induced changes in the Trp fluorescence spectra. It follows from Fig. 2 that the unfolding of any of the proteins studied is accompanied by the pronounced change of the fluorescence spectrum. The unfolding curves were obtained as the dependence of parameter  $A$  value

( $A = I_{370}/I_{320}$ , where  $I_{370}$  and  $I_{320}$  correspond to the intensity values at 370 and 320 nm, respectively) on the urea concentration, which is sensitive to changes in position and shape of the fluorescence spectrum [23]. The main advantage of this approach is that we could decrease considerably the protein concentration (from  $\sim 2.0$  mg/ml used for DSC measurements up to 0.01 mg/ml), thus avoiding uncertainties connected with protein association.

Fig. 5A depicts the normalized urea-unfolding curves for *wtS6*, *cpS6d* and *cpS6f* proteins. It is necessary to underline that the unfolding data perfectly fitted within the frame of the two-state model for all three proteins. This allowed us to perform the usual analysis suggested by Tanford [11] and developed by Pace [12] (see Section 2).

Results of such analysis for S6 proteins are presented in Fig. 5B (as  $\Delta G_{fold}$  versus [urea] dependencies) and in Table 1. Fig. 5B shows that the data points for all three proteins fitted well by straight lines, reflecting the validity of the all-or-none approximation for the description of unfolding of these proteins. It was established that typical value of  $m$  is about 1 kcal/(mol M) for urea-induced unfolding of native natural proteins, whereas that of  $\Delta G_{fold}^{H_2O}$  is between 5 and 10 kcal/mol [24]. As follows from Table 1, *wtS6*, *cpS6d* and *cpS6f* are extremely stable to urea-induced unfolding ( $[urea]_{1/2} > 4.5$  M,  $\Delta G_{fold}^{H_2O} > 5$  kcal/mol and  $m > 1$  kcal/(mol M)).

Additional confirmation of this suggestion was obtained by type II analysis of the unfolding data. It has been shown that the steepness of urea- or guanidinium chloride-induced unfolding curve depends strongly on whether a given protein has a unique tertiary structure (i.e. it is native) or it is already denatured and exists as a molten globule [14,15].

Table 1  
Thermodynamic parameters of urea-induced unfolding of *wtS6*, *cpS6d* and *cpS6f* proteins

Protein	[Urea] <sub>1/2</sub> (M)	$\Delta G_{fold}^{H_2O}$ (kcal mol <sup>-1</sup> )	$m$ (kcal mol <sup>-1</sup> M <sup>-1</sup> )	$\Delta v^{eff}$
<i>wtS6</i>	8.68 ± 0.04	11.55 ± 0.05	1.33 ± 0.05	12.5 ± 0.5
<i>cpS6d</i>	5.18 ± 0.04	6.55 ± 0.05	1.27 ± 0.05	10.8 ± 0.5
<i>cpS6f</i>	4.69 ± 0.04	5.81 ± 0.06	1.24 ± 0.05	9.6 ± 0.5
Natural protein	–	$5 < \Delta G_{fold}^{H_2O} < 10$	$m > 1$	N → U <sup>a</sup> : 8.9 ± 2.0 MG → U <sup>b</sup> : 3.4 ± 2.0

<sup>a</sup>Calculated for the natural protein of 11 kDa by Eq. 4 [15], see the text.

<sup>b</sup>Calculated for the natural protein of 11 kDa by Eq. 5 [15], see the text.

To this end, the corresponding value of the difference in the number of denaturant molecules ‘bound’ to one protein molecule in its two states,  $\Delta v^{\text{eff}}$  (see Section 2), should be determined. Then this quantity should be compared with  $\Delta v_{\text{N} \rightarrow \text{U}}^{\text{eff}}$  and  $\Delta v_{\text{MG} \rightarrow \text{U}}^{\text{eff}}$  values corresponding to the native-coil and molten globule-coil transitions in globular protein of a given molecular mass. These two parameters can be estimated from [15]:

$$\log \Delta v_{\text{N} \rightarrow \text{U}}^{\text{eff}} = 0.97 \log M - 0.07, \quad (4)$$

$$\log \Delta v_{\text{MG} \rightarrow \text{U}}^{\text{eff}} = 0.89 \log M - 0.4, \quad (5)$$

where  $M$  is the protein molecular mass in kDa. For native-coil and molten globule-coil transitions in natural proteins of 11 kDa (as in S6 protein) these equations give  $\Delta v_{\text{N} \rightarrow \text{U}}^{\text{eff}} = 8.9 \pm 2.0$  and  $\Delta v_{\text{MG} \rightarrow \text{U}}^{\text{eff}} = 3.4 \pm 2.0$ , respectively. The value of standard deviation ( $\pm 2.0$ ) in these cases was determined from the corresponding figures in [15]. The  $\Delta v^{\text{eff}}$  values calculated for S6 proteins (see Table 1) show that the proteins studied have a rigid native-like tertiary structure and are highly resistant to the denaturant action, as  $\Delta v^{\text{eff}} > 8.9$  for all three.

On the other hand, the data presented in Fig. 5 and Table 1 are consistent with the conclusion that there is a considerable difference in the unfolding behavior of *cpS6d* and *cpS6f* proteins. Table 1 shows that all the unfolding parameters of *cpS6f* ( $[\text{urea}]_{1/2}$ ,  $m$ ,  $\Delta G_{\text{fold}}^{\text{H}_2\text{O}}$  and  $\Delta v^{\text{eff}}$ ) are substantially lower than those of *cpS6d*. This means that the expression method can affect the conformational stability of recombinant protein.

#### 4. Discussion

The main result of the present study is to ascertain that *two forms of one protein* obtained from an *identical gene* expressed in *two different vectors* are quite different in their structural properties and conformational stability. Identical purification procedures without denaturing steps [4] were applied for both products of the gene, coding the permuted S6 ribosomal protein sequence. Both proteins were investigated in the same conditions (including identical buffers). This means that the found divergence in structure and stability of *cpS6d* and *cpS6f* proteins

are definitely due to the difference in the expression and purification methods. The cardinal distinction between the two approaches is that the *cpS6f* protein was obtained after splitting off the maltose-binding protein from the fusion product of the pMAL-c vector, whereas *cpS6d* was produced in the direct expression system. Let us try to understand the driving forces for the difference.

First of all, one could expect the existence of some effect of the attached maltose-binding protein on the *in vivo* folding and, as a consequence, on the structure of permuted S6 protein. The *cpS6f* can retain these structural peculiarities, even after the digestion of maltose-binding protein. However, we have established that the urea-induced unfolding of wild-type protein, as well as that of its two permuted variants is reversible. This means that the effect of maltose-binding protein can be excluded. In other words, the observed difference between *cpS6d* and *cpS6f* proteins cannot be explained by the discrepancy between their *in vivo* folding. Secondly, the expression methods induced a distinction in the N-terminal part of the protein amino acid sequence. The direct expression system requires the inclusion of Met at the first N-terminal position and that is why we had to introduce this residue into the gene of *cpS6d*. NH<sub>2</sub>-terminal sequence analysis has proved the divergence in the N-terminal sequences of both proteins, namely, *cpS6d* protein contained an extra methionine residue at the first position, while the *cpS6f* variant began from Leu. Thus, we assume that the difference in structural properties and conformational stability between *cpS6f* and *cpS6d* can be due to the presence or absence of methionine residue at the first position.

The same conclusion was recently drawn for a couple of other proteins [25–32]. In the case of  $\alpha$ -lactalbumin, the Ca<sup>2+</sup> binding constant of recombinant protein was shown to be one order of magnitude lower than that of the authentic protein isolated from milk [25]. Recombinant  $\alpha$ -lactalbumin possessed a remarkable decrease in conformational stability [25–32]. It was suggested that this was primarily due to excess conformational entropy of the N-terminal Met in the unfolded state, and also due to exposure of hydrophobic surface on unfolding [27]. Destabilizing effect of zero methionine was also shown for recombinant apo-myoglobin [28] and hen egg white lysozyme [29]. Contrary to these

cases, the attachment of methionine appears to stabilize S6 permutin structure. The same behaviour was described for recombinant ribonuclease A [30]. On the other hand, neither structural properties, nor conformational stability of recombinant apolipoprotein AIV [31] and calbinding D9k [32] were affected by the attachment of the extra methionine residue. The fact that there is difficult to predict whether protein structure will be stabilized, destabilized or unaffected by an extra N-terminal methionine shows that there are no general rules in the structural consequence of this modification, and that details of this effect have not yet been well understood.

The data presented in our paper unambiguously show that the structural properties and conformational stability of recombinant protein can depend on the peculiarities of the protein expression and purification procedure. We assume that the more disturbed the tertiary structure of recombinant protein, the more the effect of expression and purification. This point should be taken into account especially in the cases when one deals with mutated and artificial proteins.

### Acknowledgements

This work was supported in part by grants from the Russian Foundation for Basic Research and the International Science and Technology Center, Moscow.

### References

- [1] G.T. Yarranton, A. Mountain, in: A.R. Rees, M.J.E. Sternberg, R. Wetzel (Eds.), *Protein Engineering. A Practical Approach*, IRL Press, Oxford, 1992, pp. 303–325.
- [2] G.H. Schein, *Biotechnology* 7 (1989) 1141–1149.
- [3] R.W. Old, S.B. Primrose, *Principles of Gene Manipulation. An Introduction to Genetic Engineering*, 5th edn., Blackwell Science, Cambridge, 1994.
- [4] R.F. Latypov, Z.Kh. Abdullaev, A.Ya. Badretdinov, E.V. Bocharov, T.N. Mel'nik, I.Yu. Afasizheva, A.S. Arsen'ev, D.A. Dolgikh, V.N. Uverskii, A.V. Finkelstein, M.P. Kirpichnikov, *Mol. Biol. (Moscow)* 32 (1998) 109–116.
- [5] Z.Kh. Abdullaev, R.F. Latypov, A.Ya. Badretdinov, D.A. Dolgikh, V.N. Uversky, A.V. Finkelstein, M.P. Kirpichnikov, *FEBS Lett.* 414 (1997) 243–246.
- [6] D.A. Dolgikh, A.N. Fedorov, V.V. Chemeris, B.K. Chernov, A.V. Finkelstein, A.A. Shulga, Yu.B. Alakhov, M.P. Kirpichnikov, O.B. Ptitsyn, *Dokl. Acad. Nauk. SSSR* 320 (1991) 1266–1270.
- [7] A.N. Fedorov, D.A. Dolgikh, V.V. Chemeris, B.K. Chernov, A.V. Finkelstein, A.A. Shulga, Yu.B. Alakhov, M.P. Kirpichnikov, O.B. Ptitsyn, *J. Mol. Biol.* 225 (1992) 927–931.
- [8] P.L. Privalov, *Adv. Protein Chem.* 33 (1979) 167–241.
- [9] P.L. Privalov, S.A. Potekhin, *Methods Enzymol.* 131 (1986) 4–51.
- [10] L. Stryer, *Science* 162 (1968) 526–540.
- [11] C. Tanford, *Adv. Protein Chem.* 23 (1968) 121–282.
- [12] C.N. Pace, *Methods Enzymol.* 131 (1986) 266–280.
- [13] K.C. Aune, C. Tanford, *Biochemistry* 8 (1969) 4586–4590.
- [14] O.B. Ptitsyn, V.N. Uversky, *FEBS Lett.* 341 (1994) 15–18.
- [15] V.N. Uversky, O.B. Ptitsyn, *Folding Design* 1 (1996) 117–122.
- [16] K. Marcker, F. Sander, *J. Mol. Biol.* 8 (1964) 354–360.
- [17] A.J. Adler, N.J. Greenfield, G.D. Fasman, *Methods Enzymol.* 27 (1973) 675–735.
- [18] G. Weber, F.W.J. Teale, in: H. Neurath (Ed.), *The Proteins*, Vol. 3, Academic Press, New York, 1966, pp. 445–452.
- [19] G.M. Barenboim, A.N. Domanski, K.K. Turoverov, *Luminescence of Biopolymers and Cells*, Plenum Press, New York, 1969.
- [20] E.A. Burshtein, *Luminescence of Protein Chromophores. Model Investigations. Itogi Nauki i Tekhniki (Biofizika)*, Vol. 6) Moscow, 1976.
- [21] E.A. Permyakov, *Luminescence Spectroscopy of Proteins*, CRC Press, London, 1993.
- [22] O.B. Ptitsyn, *Adv. Protein Chem.* 47 (1995) 83–229.
- [23] K.K. Turoverov, S.Yu. Khaitlina, G.P. Pinaev, *FEBS Lett.* 62 (1976) 4–7.
- [24] T.E. Creighton, *Proteins – Structure and Molecular Properties*, W.H. Freeman, New York, 1993, pp. 289–291.
- [25] D.B. Veprintsev, *Thermodynamic and structural aspects of bivalent ion binding by proteins. Candidates dissertation (in Russian)*, Pushchino, 1997.
- [26] N. Ishikawa, T. Chiba, L.T. Chen, A. Shimizu, M. Ikeguchi, S. Sugai, *Protein Eng.* 11 (1998) 333–335.
- [27] T.K. Chaudhuri, K. Horii, T. Yoda, M. Arai, S. Nagata, T.P. Terada, H. Uchiyama, T. Ikura, K. Tsumoto, H. Kataoka, M. Matsushima, K. Kuwajima, I. Kumagai, *J. Mol. Biol.* 285 (1999) 1179–1194.
- [28] M.S. Hargrove, S. Krzywda, A.J. Wilkinson, Y. Dou, M. Ikeda-Saito, J.S. Olson, *Biochemistry* 33 (1994) 11767–11775.
- [29] S. Mine, T. Ueda, Y. Hashimoto, T. Imoto, *Protein Eng.* 10 (1997) 1333–1338.
- [30] D.A. Schultz, R.L. Baldwin, *Protein Sci.* 1 (1992) 910–916.
- [31] N. Duverger, A. Murry-Brelier, M. Latta, S. Reboul, G. Castro, J.F. Mayaux, J.C. Fruchart, J.M. Taylor, A. Steinmetz, P. Deneffe, *Eur. J. Biochem.* 201 (1991) 373–383.
- [32] J. Kordel, S. Forsen, W.J. Chazin, *Biochemistry* 28 (1989) 7065–7074.