Effect of the Extra N-terminal Methionine Residue on the Stability and Folding of Recombinant α-Lactalbumin Expressed in Escherichia coli

Tapan K. Chaudhuri¹, Katsunori Horii², Takao Yoda¹, Munehito Arai¹ Shinji Nagata³, Tomoki P. Terada¹, Hidefumi Uchiyama⁴, Teikichi Ikura¹ Kouhei Tsumoto², Hiroshi Kataoka³, Masaaki Matsushima⁵ Kunihiro Kuwajima¹* and Izumi Kumagai²

1Department of Physics Graduate School of Science University of Tokyo, Tokyo 113-0033, Japan
2Department of Biomolecular Engineering, Tohoku University, Sendai 980-8579, Japan
3Department of Biotechnology Faculty of Agriculture & Life Sciences, University of Tokyo Tokyo, 113-8658, Japan
4Department of Biological Science and Technology Science University of Tokyo 2941 Yamazaki, Noda, Chiba 278-8510, Japan
5Rational Drug Design Laboratories, Fukushima 960-1242, Japan

The structure, stability, and unfolding-refolding kinetics of Escherichia coli expressed recombinant goat α-lactalbumin were studied by circular dichroism spectroscopy, X-ray crystallography, and stopped-flow measurements, and the results were compared with those of the authentic protein prepared from goat milk. The electric properties of the two proteins were also studied by gel electrophoresis and ion-exchange chromatography. Although the overall structures of the authentic and recombinant proteins are the same, the extra methionine residue at the N terminus of the recombinant protein remarkably affects the native-state stability and the electric properties. The native state of the recombinant protein was 3.5 kcal/mol less stable than the authentic protein, and the recombinant protein was more negatively charged than the authentic one. The recombinant protein unfolded 5.7 times faster than the authentic one, although there were no significant differences in the refolding rates of the two proteins. The destabilization of the recombinant protein can be fully interpreted in terms of the increased unfolding rate of the protein, indicating that the N-terminal region remains unorganized in the transition state of refolding, and hence is not involved in the folding initiation site of the protein. A comparison of the X-ray structures of recombinant α-lactalbumin determined here with that of the authentic protein shows that the structural differences between the proteins are confined to the N-terminal region. Theoretical considerations for the differences in the conformational and solvation free energies between the proteins show that the destabilization of the recombinant protein is primarily due to excess conformational entropy of the N-terminal methionine residue in the unfolded state, and also due to less exposure of hydrophobic surface on unfolding. The results suggest that when the N-terminal region of a protein has a rigid structure, expression of the protein by E. coli, which adds the extra methionine residue, destabilizes the native state through a conformational entropy effect. It also shows that differences in the electrostatic interactions of the N-terminal amino group with the side-chain atoms of Thr38, Asp37, and Asp83 bring about a difference in the pKₐ value of the N-terminal amino group between the proteins, resulting in a greater negative net charge of the recombinant protein at neutral pH.

Keywords: recombinant goat α-lactalbumin; extra N-terminal methionine residue; protein folding; X-ray crystallographic study; conformational entropy
Introduction

The N-terminal sequence of a recombinant protein expressed in Escherichia coli is known to start with formyl-methionine (Marcker & Sanger, 1964), which is in most cases subsequently processed by deformylase enzyme (Adams, 1968; Takeda & Webster, 1968), and removed by methionine aminopeptidase to finally produce the N-terminal methionine-free recombinant protein. However, removal of the N-terminal methionine does not always take place, and about half of *E. coli*-expressed proteins contain the extra N-terminal methionine residue, because the aminopeptidase action depends on the nature of the penultimate amino acid residue (Morschell et al., 1990). Therefore, the effect of the N-terminal methionine residue, when present, on the structure, stability and folding of *E. coli*-expressed recombinant proteins should be an important issue in biophysical and molecular biological studies that use such recombinant proteins, although this has not been taken seriously in most cases.

The biological and physicochemical properties of the methionylated proteins expressed in *E. coli* may differ from those of the authentic proteins that do not have the N-terminal methionine. For example, recombinant hen egg-white lysozyme contains the N-terminal methionine residue (Miki et al., 1987; Mine et al., 1997) and has lower solubility and stability than the authentic form (Imoto et al., 1987). Similarly, recombinant apomyoglobin expressed in *E. coli* contains the extra N-terminal methionine residue and is less stable than the authentic protein (Hargrove et al., 1994). On the other hand, the presence of the extra N-terminal methionine or the extension or truncation of the N-terminal residues does not interfere with the native-state stability in certain other globular proteins (Kordal et al., 1989; Duverger et al., 1991). In recombinant ribonuclease A, the extra N-terminal methionine is even known to stabilize the native structure (Schultz & Baldwin, 1992; Aronsson et al., 1995). However, details of the effects of the extra N-terminal methionine residue on the structure, stability, and folding of the proteins have not yet been well understood.

**α-Lactalbumin** is a milk Ca$^{2+}$-binding protein, which consists of 123 amino acid residues and has a molecular mass of 14,200 Da. The three-dimensional structure of *α*-lactalbumin from several mammalian species, including goat, cow, guinea pig, and human, has been determined by X-ray crystallographic analysis (Acharya et al., 1991; Pike et al., 1996), and it is very similar to the structure of c-type lysozyme, a homologous protein. *α*-Lactalbumin has been used actively as a model protein in studies of protein folding (Sugai & Ikeguchi, 1994; Kuwajima, 1989, 1996; Vanderheeren & Hansens, 1994; Uchiyama et al., 1995; Schulman & Kim, 1996; Arai & Kuwajima, 1996; Schulman et al., 1997; Wilson et al., 1996; Shimizu et al., 1996; Balbach et al., 1996; Katsumata et al., 1996; Kataoka et al., 1997; Kuhlman et al., 1997; Wu & Kim, 1997; Pfeil, 1998; Ikeguchi et al., 1998), because this protein readily adopts a molten globule state, which is known to be identical with a folding intermediate (Kuwajima, 1989, 1996; Pitsyn, 1995), under a variety of conditions, including those at a low pH, at a moderate concentration of guanidine hydrochloride (GdnHCl), and in the absence of Ca$^{2+}$ and other salts (Kuwajima, 1989, 1996). Recombinant *α*-lactalbumin expressed in *E. coli*, though containing the extra N-terminal methionine, has often been used in these studies of protein folding. A recent study has, however, shown that like recombinant hen egg-white lysozyme, recombinant bovine *α*-lactalbumin is less stable than the authentic protein, although the lactose synthase regulatory activities of the recombinant and authentic proteins have been shown to be identical with each other (Ishikawa et al., 1998).

Here, we show that *E. coli*-expressed recombinant goat *α*-lactalbumin is destabilized by the presence of the extra N-terminal methionine residue by as much as 3.5 kcal/mol and has a more negative electric net charge than the authentic protein. It is concluded that the destabilization of the recombinant protein is primarily brought about by an extra conformational entropy of the methionyl residue in the unfolded state and that the more negative charge of the recombinant protein is caused by a decrease in the $\Delta G$ value of the N-terminal amino group. Because the N-terminal methionine remarkably destabilizes recombinant *α*-lactalbumin, the role of the N terminus in the folding of this protein has also been investigated by stopped-flow circular dichroism (CD) studies of the unfolding and refolding kinetics of the recombinant and authentic proteins. The destabilization of the recombinant protein is shown to be entirely interpreted in terms of an increase in the unfolding rate, indicating that the N terminus is not involved in the folding initiation site of *α*-lactalbumin.

Results

Structure of folded recombinant goat *α*-lactalbumin

The recombinant wild-type protein was expressed in *E. coli* as inclusion bodies with a high yield (15 mg per litre of culture). The protein was solubilized in 8 M urea and refolded in a redox buffer in the absence of urea at pH 8.5 and 4°C. The process of refolding was monitored by reversed-phase HPLC (Uchiyama et al., 1995), and the folded protein was purified (see Materials and Methods). The peptide and aromatic CD spectra of the recombinant protein were measured under native conditions (0 M GdnHCl at pH 8.0) at 25°C, and compared with those of authentic goat *α*-lactalbumin (Figure 1(a) and (b)). There is no significant difference in the CD spectra between the proteins in the aromatic and peptide regions, so that the secondary and tertiary structures of the
two proteins are essentially identical with each other. This conclusion is confirmed by the X-ray crystallographic structure of recombinant goat α-lactalbumin (see below). The results thus indicate that the folded recombinant protein is correctly folded into the native structure. A study has also shown that the lactose synthase regulatory activity of the folded recombinant protein is the same as that of authentic α-lactalbumin (Uchiyama et al., 1995).

Equilibrium unfolding

The GdnHCl-induced equilibrium unfolding transition of the folded recombinant protein was studied by the peptide and aromatic CD spectra, and the results were compared with those of authentic goat α-lactalbumin. Figure 2 shows the unfolding transition curves of the two proteins measured by the CD ellipticities at 222 and 270 nm, and these ellipticities are expressed by the apparent fractional extent \( F_{\text{app}} \) of unfolding as a function of GdnHCl concentration.

From Figure 2, the unfolding transition curves measured at 222 and 270 nm are coincident with each other in authentic and recombinant α-lactalbumin, indicating that the unfolding transitions of the two proteins are well represented by a two-state mechanism, in which only the native (N) and the fully unfolded (U) states are populated in the transition zone as:

\[
N \rightleftharpoons K \rightarrow U
\]

Here \( K_U \) is the equilibrium constant of unfolding and relates to the free energy change, \( \Delta G_U \), of the unfolding transition as:

\[
K_U = \exp(-\Delta G_U/RT)
\]

where \( R \) and \( T \) are the gas constant and the absolute temperature, respectively, and \( \Delta G_U \) is assumed to be linearly dependent on GdnHCl concentration (C) as:

\[
\Delta G_U = \Delta G_U^{\text{H}_2\text{O}} - mC = m(C_m - C)
\]

where \( \Delta G_U^{\text{H}_2\text{O}} \) is the \( \Delta G_U \) in the absence of the denaturant, \( C_m \) is the C at the midpoint of the unfolding transition, and \( m \) represents the dependence of \( \Delta G_U \) on C and is a measure of the cooperativity of the transition (Pace, 1986). From equations (1) to (3), the transition curve expressed
by \( F_{\text{app}} \) is given as a function of \( C \) as:

\[
F_{\text{app}}(C) = \frac{\exp[-m(C - C_m)/RT] + m(C - C_m)/RT}{1 + \exp[-m(C - C_m)/RT]}
\]

(4)

The values of \( m \), \( C_m \), and hence \( \Delta G_{\text{U}}^{\text{H,O}} \), for recombinant and authentic \( \alpha \)-lactalbumin were calculated from the data of Figure 2 by the non-linear least-squares method. The unfolding parameters \( m \), \( C_m \), and hence \( \Delta G_{\text{U}}^{\text{H,O}} \), thus obtained are summarized in Table 1. The continuous lines in Figure 2 are the curves theoretically drawn with the parameter values of Table 1, and show excellent agreement between theory and the experimental data.

Figure 2 also shows that the unfolding transition of the recombinant protein occurs at a remarkably lower concentration of GdnHCl (\( C_m \approx 2.7 \) M) than the transition of authentic \( \alpha \)-lactalbumin (\( C_m = 3.2 \) M). The difference in \( \Delta G_{\text{U}} \) (\( \Delta \Delta G_{\text{U}} \)) is \(-3.5\) kcal/mol at 0 M GdnHCl and \(-1.9\) kcal/mol at 3.2 M GdnHCl, which is the \( C_m \) for the authentic protein (Table 1). Therefore, the folded recombinant protein is remarkably less stable than authentic \( \alpha \)-lactalbumin, although their native structures are practically identical as evidenced by the CD spectra and X-ray structural analysis.

### Table 1. Equilibrium unfolding transition parameters of goat \( \alpha \)-lactalbumin

<table>
<thead>
<tr>
<th>Name of protein</th>
<th>( \Delta G_{\text{U}}^{\text{H,O}} ) (kcal/mol)</th>
<th>( m ) (kcal/mol M)</th>
<th>( C_m ) (M)</th>
<th>( \Delta \Delta G_{\text{U}}^{\text{H,O}} ) (kcal/mol)</th>
<th>( \Delta \Delta G_{\text{U}} ) (kcal/mol) at 3.2 M GdnHCl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authentic goat ( \alpha )-lactalbumin</td>
<td>13.8 ± 0.7</td>
<td>4.4 ± 0.2</td>
<td>3.15 ± 0.01</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Recombinant goat ( \alpha )-lactalbumin</td>
<td>10.4 ± 0.5</td>
<td>3.9 ± 0.2</td>
<td>2.67 ± 0.01</td>
<td>-3.5</td>
<td>-1.9</td>
</tr>
</tbody>
</table>

The values of \( m \), \( C_m \), and hence \( \Delta G_{\text{U}}^{\text{H,O}} \), for recombinant and authentic \( \alpha \)-lactalbumin were calculated from the data of Figure 2 by the non-linear least-squares method. The unfolding parameters \( m \), \( C_m \), and hence \( \Delta G_{\text{U}}^{\text{H,O}} \), thus obtained are summarized in Table 1. The continuous lines in Figure 2 are the curves theoretically drawn with the parameter values of Table 1, and show excellent agreement between theory and the experimental data.

Figure 2 also shows that the unfolding transition of the recombinant protein occurs at a remarkably lower concentration of GdnHCl (\( C_m \approx 2.7 \) M) than the transition of authentic \( \alpha \)-lactalbumin (\( C_m = 3.2 \) M). The difference in \( \Delta G_{\text{U}} \) (\( \Delta \Delta G_{\text{U}} \)) is \(-3.5\) kcal/mol at 0 M GdnHCl and \(-1.9\) kcal/mol at 3.2 M GdnHCl, which is the \( C_m \) for the authentic protein (Table 1). Therefore, the folded recombinant protein is remarkably less stable than authentic \( \alpha \)-lactalbumin, although their native structures are practically identical as evidenced by the CD spectra and X-ray structural analysis.

### Gel electrophoresis and ion-exchange chromatography

In order to investigate further differences between recombinant and authentic \( \alpha \)-lactalbumin, the electrophoretic and ion-exchange chromatographic behavior of the two proteins were investigated. Figure 3(a) shows electrophoretic patterns in a non-denaturing polyacrylamide gel at pH 9.4. It can be seen that the electrophoretic mobility of the recombinant protein is significantly greater than that of the authentic protein. Figure 3(b) shows the elution profiles of recombinant and authentic \( \alpha \)-lactalbumin in an anion-exchange HPLC using a RESOURCE\textsuperscript{TM} Q column (Pharmacia Biotech) with a linear gradient from 0 M to 0.5 M NaCl in the presence of 10 mM NaH\(_2\)PO\(_4\)-Na\(_2\)HPO\(_4\) buffer (pH 7.0). The retention time is longer for the recombinant protein (22.9 minutes) than for the authentic one (19.6 minutes). Both of these results indicate that the recombinant protein is more negatively charged. These differences in the electric properties of the two proteins, however, disappear in the U state in 8 M urea. The electrophoretic mobilities and the chromatographic retention times of the proteins were found to be identical in the presence of 8 M urea (data not shown). Therefore,
the difference in the electric charge between the proteins must be caused by the structural folding of the proteins into the native structure.

SDS/polyacrylamide gel electrophoresis was also carried out for the recombinant and authentic proteins using 15% acrylamide in the resolving gel (Figure 3(c)). The electrophoretic mobilities of the two proteins are the same within the experimental error, indicating that there is no significant difference in the molecular mass between the proteins.

**N-terminal sequence and mass spectrometric analyses**

In order to identify any differences in the amino acid sequence, we performed N-terminal sequencing and mass spectrometric analysis of the recombinant and authentic proteins. The N-terminal sequences of the first five residues of the two proteins have shown that recombinant \( \alpha \)-lactalbumin contains an additional methionine residue. The results of the mass spectrometric analysis indicate that the difference in mass between the recombinant and authentic proteins is 133 (Figure 4), which is nearly equal to the mass of a single methionine residue (131.19), confirming the presence of the extra methionine residue in the recombinant protein. Therefore, the only chemical difference that brings about the difference in the electric charge between the two proteins in the N state is the presence or absence of the extra methionine residue at the N terminus, and this difference may also lead to the remarkable difference in stability between the proteins.

**Methionine-free recombinant \( \alpha \)-lactalbumin**

In order to directly investigate the effect of the extra methionine residue on the electric properties and stability of the recombinant protein, methionine-free recombinant \( \alpha \)-lactalbumin was prepared by cyanogen bromide (CNBr) cleavage. Because there is no methionine residue in authentic goat \( \alpha \)-lactalbumin, only the extra N-terminal methionine of the recombinant protein is expected to be removed by the CNBr cleavage. The removal of the methionine was confirmed by N-terminal sequencing and mass spectrometric analysis (data not shown). The absence of other cleavage products was confirmed by SDS/polyacrylamide gel electrophoresis. The near and far-UV CD spectra of the methionine-free recombinant protein overlap with those of the authentic and original recombinant proteins (data not shown). The electrophoretic mobility in the native gel and the retention time for the anion-exchange chromatography were found to be identical with those of the authentic protein (data not shown). The stability of the methionine-free recombinant protein against the GdnHCl-induced unfolding was investigated, and the equilibrium unfolding transition of the methionine-free protein is shown in Figure 2. The unfolding transition curve coincides well with that of the authentic protein, and gives the same \( C_m \) and \( \Delta G^{\text{U}}_{\text{H}_2\text{O}} \) values. As a control, the authentic protein was also subjected to the conditions of CNBr cleavage, and it was confirmed that the unfolding transition of the protein was not affected by the cleavage conditions (data not shown). These results thus clearly indicate that the observed destabilization and the difference in the electric charge of the recombinant protein is solely due to the presence of the extra N-terminal methionine residue.

**Kinetics of refolding and unfolding**

The above results indicate that the presence of the extra methionine residue at the N terminus of the recombinant protein decreases the relative stability of the N state by as much as 3.5 kcal/mol. Thus, it appears that both recombinant and authentic goat \( \alpha \)-lactalbumin are useful for investigating the role of the N-terminal residue in the kinetic folding of \( \alpha \)-lactalbumin. The kinetic unfolding and refolding reactions of the recombinant and authentic proteins were investigated by stopped-flow CD measurements. The unfolding and refolding reactions were induced by concentration jumps of GdnHCl from 1.0 to 5.4 M and from 5.5 to 0.5 M, respectively. The reactions were monitored by the ellipticity change at 225 nm at pH 7.0 and 25°C. The kinetic progress curves for unfolding and refolding are shown in Figure 5(a) and (b), respectively, and the data were fitted by the non-linear least-squares method with the equation:

\[
A(t) = A(\infty) + \Delta A_{\text{obs}} \sum \alpha_i \exp(-k_i t)
\]

where \( A(t) \) and \( A(\infty) \) are the observed values of the ellipticity at time \( t \) and infinite time, respectively, \( \Delta A_{\text{obs}} \) is the observed total amplitude \( [A(0) - A(\infty)] \), and \( k_i \) and \( \alpha_i \) are the apparent first-
order rate constant and fractional amplitude, respectively, of the \(i\)th kinetic phase.

The kinetic progress curves for unfolding for both the recombinant and authentic proteins were well fitted to a single-exponential equation, and the apparent rate constants and the amplitudes for the two proteins are presented in Table 2. The unfolding reaction of recombinant \(\alpha\)-lactalbumin is 5.7-times faster than that of the authentic protein, while there are no significant differences in the rate constants for the triphasic refolding reactions of the two proteins. Thus, it appears that the N-terminal end of goat \(\alpha\)-lactalbumin is not essential for the kinetic folding of this protein (see Discussion).

### X-ray crystallographic study

In order to further investigate the differences in the folded structure between recombinant and authentic goat \(\alpha\)-lactalbumin, an X-ray crystallographic analysis of the recombinant protein was performed, and the structure was compared with that reported for the authentic protein structure. The crystallographic data are summarized in Table 4. The space group of the crystal of the recombinant protein was altered to \(P2_12_12\) from \(P2_1\) in which the authentic protein was packed (Pike et al., 1996). The number of protein molecules in the asymmetric unit was one, although there were two (Mol A and Mol B) in the authentic protein crystal. The final \(R\) and free \(R\) factors were 0.191 and 0.278 in the resolution range of 8.0 to 2.0 Å. The overall error was estimated at 0.19 Å by a Luzzati plot (Luzzati, 1952). As the space group is altered in the recombinant protein crystal, the N-terminal methionine may affect the molecular packing in the crystal. However, the interactions between the two independent authentic molecules (Mol A and Mol B) were found to be very similar to the interactions between the symmetry-related recombinant molecules (Figure 6).

The structural differences between the recombinant and the authentic proteins are shown in Figure 7, which represents the distances between the C\(^\alpha\) atoms of the two molecules. The root-mean-square deviations of the main-chain atoms are 0.55 Å between the recombinant protein molecule

### Table 2. Kinetic unfolding parameters of goat \(\alpha\)-lactalbumin

<table>
<thead>
<tr>
<th>Name of protein</th>
<th>(k_1) (s(^{-1}))</th>
<th>(\Delta A_{\theta_{125}}) (deg cm(^2) dmol(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authentic goat (\alpha)-lactalbumin</td>
<td>1.26 ± 0.01</td>
<td>~8384</td>
</tr>
<tr>
<td>Recombinant goat (\alpha)-lactalbumin</td>
<td>7.18 ± 0.08</td>
<td>~8056</td>
</tr>
</tbody>
</table>

### Table 3. Kinetic refolding parameters of goat \(\alpha\)-lactalbumin

<table>
<thead>
<tr>
<th>Name of protein</th>
<th>(k_1) (s(^{-1}))</th>
<th>(\Delta A_{\theta_{125}}) (deg cm(^2) dmol(^{-1}))</th>
<th>(k_2) (s(^{-1}))</th>
<th>(\Delta A_{\theta_{125}}) (deg cm(^2) dmol(^{-1}))</th>
<th>(k_3) (s(^{-1}))</th>
<th>(\Delta A_{\theta_{125}}) (deg cm(^2) dmol(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authentic goat (\alpha)-lactalbumin</td>
<td>0.11 ± 0.05</td>
<td>64.28</td>
<td>1.3 ± 1.1</td>
<td>145</td>
<td>4.9 ± 0.3</td>
<td>2282</td>
</tr>
<tr>
<td>Recombinant goat (\alpha)-lactalbumin</td>
<td>0.09 ± 0.04</td>
<td>70.8</td>
<td>1.3 ± 0.4</td>
<td>335</td>
<td>5.7 ± 0.4</td>
<td>2234</td>
</tr>
</tbody>
</table>
and Mol A, and 0.63 Å between the recombinant molecule and Mol B. These values are larger than the root-mean-square deviation between Mol A and Mol B (0.27 Å). From Figure 7, we can see that the intermolecular interactions remarkably affect the structure of the N-terminal and loop regions of the protein, especially between residues 105 and 110, but that the overall structures of the recombinant and authentic proteins are essentially identical, supporting previous observations of the same CD spectra of the proteins in solution.

The structures around the N termini of the recombinant protein and the two molecules of the authentic protein are shown in Figure 8, and we may see structural differences that give rise to the differences in the electric properties and stability between them. The N-terminal amino group strongly interacts with the side-chain atoms of Thr38 and Gln39 in Mol A (Figure 8(a)) and Thr38 in Mol B through hydrogen bonds and/or salt bridges (Figure 8(b)). A similar interaction can also be observed in the recombinant protein, in which the N-terminal amino group is bound to the side-chain of Gln39 by a hydrogen bond (Figure 8(c)), but this interaction may be significantly stronger than the corresponding interaction in the authentic protein (see Discussion). It can also be seen from Figure 8(c) that the methionine side-chain of the recombinant protein is directly in contact with the side-chain of Gln2, and that the orientation of the methionine side-chain is fixed by the hydrogen bonds between the N-terminal amino group and the side-chain of Gln39, and between the main-
effect of the N-terminal methionine on the ther-
et al. Only recently, Ishikawa et al. have thoroughly investigated by equilibrium unfolding and kinetic unfolding-refolding studies as well as their crystallographic packings are different. The side-chain of Glu1 is folded into the inside of the recombinant molecule and interacts with the amino group of Lys11, and this conformation is similar to that in Mol A, but the Glu1 side-chain is shifted further from the amino group of Lys11 in the recombinant protein. The corresponding conformation of the Glu1 side-chain of Mol B is affected by the positively charged His107 side-chains of the symmetry-related molecules (Mol A and Mol B) in the crystal of the authentic protein. Without the presence of these positive charges near Glu1 of Mol B, the conformation of the side-chain is similar to that of the recombinant molecule and Mol A of the authentic protein.

Discussion

The present results show that the presence of the additional N-terminal methionine residue in recombinant α-lactalbumin expressed in E. coli remarkably decreases the stability of the native protein and increases the apparent net negative charge in the native state. Many proteins expressed in E. coli have the N-terminal methionine residue, although whether or not the methionine residue is present depends on the next residue in the recombinant protein (Miller et al., 1987). Thus, the effect of the N-terminal methionine on the structure, stability, and other properties of an E. coli-expressed protein is important when we use the recombinant protein in biophysical and molecular biological studies. Such effects of the N-terminal methionine residue have, however, been ignored so far in most cases. As far as we are aware, the present study is the first of its kind in which the effect of the N-terminal methionine residue is thoroughly investigated by equilibrium unfolding and kinetic unfolding-refolding studies as well as by the CD and X-ray crystallographic analyses. Only recently, Ishikawa et al. (1998) have reported the effect of the N-terminal methionine on the thermal unfolding of bovine α-lactalbumin, and their results and the present results should be complementary to each other.

Because the N-terminal methionine affects the native state stability of α-lactalbumin, the effects of the methionine on the kinetics of refolding and unfolding of the protein were investigated. The results of our study show that the recombinant protein unfolds 5.7 times faster than the authentic one, whereas the rates of refolding remain the same. The results should provide an insight into the role of the N terminus in the folding mechanism for α-lactalbumin.

We will first consider here the structural aspects that explain differences in the native state stability and the electric properties between the recombinant and authentic proteins on the basis of our CD and X-ray crystallographic data. We will then discuss the mechanism of folding for goat α-lactalbumin on the basis of the kinetic unfolding and refolding data.

Stabilities of recombinant and authentic α-lactalbumin

Structure around the N terminus

The results show that although recombinant and authentic α-lactalbumin follow the two-state unfolding transition (Figure 2), the recombinant protein is less stable than the authentic one by as much as 3.5 kcal/mol in the absence of GdnHCl. In order to understand this stability difference (ΔΔG_u), we determined the X-ray structure of the recombinant protein, and this structure was compared with the X-ray structure of the authentic protein determined by Pike et al. (1996). The overall structures of the two proteins are essentially identical with each other, being consistent with the identical CD spectra of the proteins, and the structural differences between the proteins have been found to be localized in the N-terminal and the 105-110 loop regions (Figure 7). Because the structural differences in the 105-110 loop region, which is very flexible in the N state, are likely to be caused by a difference in the crystallographic packing between the proteins (Acharya et al., 1991; Harata & Muraki, 1992; Pike et al., 1996), we have concentrated our attention on the structural differences in the N-terminal region and investigated any interactions that are present in the authentic protein but missing in the recombinant one. Our data, however, show that there are no such interactions identified in the X-ray structures. From Figure 8, it can be seen that the N-terminal amino group of Asp83 of the recombinant protein was not clearly seen in the electron density map, and the B-factors of the side-chain atoms were high. Therefore, the model coordinates could not be explicitly determined. Certain distances are shown in Å, and the residues are shown by the one-letter code.
Glu1 of the authentic protein is hydrogen-bonded with two side-chain oxygen atoms of Thr38 and Gln39 in Mol A and with a side-chain oxygen atom of Thr38 in Mol B. A similar hydrogen bond is also observed in the recombinant protein between the N-terminal amino group and Gln39, and the length of the hydrogen bond is smaller than that in the authentic protein, suggesting that the hydrogen bond is even stronger in the recombinant protein (Figure 8(c)). Although the hydrogen bond between the N-terminal amino group and Thr38 is missing in the recombinant protein, there is an alternative hydrogen bond between the main-chain amido group of Glu1 and the carboxyl group of Asp37. The degrees of the packing interactions of the side-chain atoms are also very similar in the N-terminal regions of the two proteins. The side-chains are closely packed in both proteins. Furthermore, contributions of electrostatic interactions around the N termini to the destabilization of the recombinant protein will be shown to be negligibly small, although they are related to the difference in the electric net charge between the proteins (see below). Therefore, the observed destabilization cannot be interpreted in terms of the presumed interactions missing in the native structure of the recombinant protein.

Conformational entropy of the methionine residue and solvation free energies

If the destabilization of the recombinant protein cannot be simply explained by the interactions identified in the X-ray structures of the recombinant and authentic proteins, what makes the recombinant protein less stable? At this point, it should be noted that the N-terminal residues of both the recombinant and authentic proteins are involved in a rigid structure, so that all the atoms of the residues can be traced in the electron density maps of the proteins by X-ray crystallographic analysis. The B-factors of the backbone atoms of the N-terminal methionine residue of the recombinant protein were found to range from 31 to 35 Å². The values are much larger than those of the residues buried inside the protein molecule (8-15 Å²), but are smaller than those of the fully exposed residues in flexible loop regions. This means that the presence of the additional methionine residue in the recombinant protein destabilizes the native state through an entropic effect, which arises from an additional conformational entropy of the methionine residue in the U state. Because the structure around the N terminus is rigid in the N state of the recombinant protein, the additional methionine residue leads to an increase in entropy on unfolding. Thus, the free energy change of unfolding (ΔG_U), which is the difference in the free energy between the N and U states, decreases, and hence the N state of the recombinant protein is destabilized.

The increase in the conformational entropy of the methionine residue on unfolding has been estimated at 20 cal/(mol K) by Oobatake & Ooi (1993) from an analysis of hydration and heat stability effects on the unfolding of 14 globular proteins, and this corresponds with the free energy change of −5.9 kcal/mol at 25°C. This value is close to but lower than the observed difference (ΔG_U = −3.5 kcal/mol) in ΔG_U between the recombinant and authentic proteins. We have, however, ignored the contribution of the hydration free energy, ΔG_w, and the enthalpic contribution of the conformational unfolding, ΔH_u, which mostly arises from the van der Waals interaction energy, to the ΔΔG_U (Oobatake & Ooi, 1993). These contributions are expected to be proportional to the change in the accessible surface area of the methionine residue on unfolding (Oobatake & Ooi, 1993) and may explain the above difference between the expected contribution of the conformational entropy (−TΔS_u) and the observed ΔΔG_U. The values of ΔG_w and ΔH_u of the N-terminal methionine residue were calculated by the method described by Oobatake & Ooi (1993), and they were −1.2 and 3.3 kcal/mol for ΔG_w and ΔH_u, respectively, so that the free energy change of unfolding of the methionine residue (ΔG_u) was estimated at −3.8 kcal/mol (see equation (8)), which was in good agreement with the observed ΔΔG_U (see Materials and Methods). The contribution of other residues to the ΔΔG_U was also estimated, and it was less than 1 kcal/mol (see Materials and Methods), confirming that the increase in the conformational entropy of the N-terminal methionine residue on unfolding is a dominant factor determining the ΔΔG_U.

In the above argument of ΔΔG_U, however, we have implicitly assumed that the U state is fully unfolded in both the recombinant and authentic proteins. Thus, if there is a difference in the U-state structure between the proteins, such a difference may also contribute to the ΔΔG_U. In fact, the m value of the equilibrium unfolding transition is found to be smaller for the recombinant protein (Table 1). Lower values of m are usually thought to be due to less exposure of hydrophobic surface on unfolding. Because the native structure is essentially identical between recombinant and authentic z-lactalbumin, the less exposure of hydrophobic surface must be due to a difference in the U-state structure, and the U state of the recombinant protein less exposes the hydrophobic surface than that of the authentic one. Similar effects of hydrophobic replacements of amino acid residues on the U-state structure have also been reported in staphylococcal nuclease (Shortle, 1996). The less difference in solvent exposed hydrophobic surface means a smaller difference in ΔG_U. Therefore, this may also be a factor determining the ΔΔG_U between the recombinant and authentic proteins.

Comparison with other proteins

Ishikawa et al. (1998) have also recently reported the difference in stability between recombinant and
authentic bovine α-lactalbumin using thermal denaturation measurements of the proteins. They have shown that the destabilization of the recombinant protein is caused by an entropic effect because the enthalpy change of the thermal unfolding is the same for the two proteins, and their result is fully consistent with our proposal regarding the destabilization of the recombinant protein described above. Although Ishikawa et al. (1998) have attributed the destabilization of the recombinant protein to a weakening of the apparent Ca\(^{2+}\)-binding strength, this interpretation seems to be nothing more than a rephrasing of the destabilization of the protein because the apparent Ca\(^{2+}\)-binding strength of α-lactalbumin is known to be linked to the N \(\equiv\) U equilibrium of the apo protein (Hiraoka & Sugai, 1985). Our X-ray structural data show that there is no essential difference in the structure of the Ca\(^{2+}\)-binding site between the authentic and recombinant proteins, indicating that the weakening of the apparent Ca\(^{2+}\)-binding strength of the recombinant protein is caused by a destabilization of its apo form.

There have been several other reports of the effect of additional residues at the N terminus on the native-state stability of recombinant proteins, and a comparison of these with the present results will provide insight into a rule relating to the effects of an extra methionine residue in the proteins. Hargrove et al. (1994) have observed that the recombinant apomyoglobin expressed in E. coli is less stable than the authentic protein. They have also shown that the N terminus of recombinant apomyoglobin contains an extra methionine residue and that the structure around the N terminus is rigid. Polyhistidine tags in the N and C-terminal regions of Arc repressor (Milla et al., 1993, 1995) have little effect on the stability and folding of the protein, whereas the polyhistidine tags of CspA alter the folding behavior by interacting with the wild-type portion of the protein (Reid et al., 1998). The X-ray crystallographic structures of the Arc repressor (Raumann et al., 1994) and CspA (Goldstein et al., 1990) have shown that the structure around the N-terminal residue in CspA is rigid, whereas that of Arc repressor is flexible. The N-terminal region of staphylococcal nuclease is flexible (Hynes & Fox, 1991), and it has been reported that a 19-residue pro-peptide in the N-terminal region of the nuclease does not significantly destabilize the N state of the recombinant protein (pro staphylococcal nuclease; Suciu & Inouye, 1996). Therefore, these studies together with our study strongly suggest that when the structure around the N-terminal residue of a protein is rigid, the addition of extra residues at the N terminus destabilizes the N state of the protein. On the other hand, when the structure is flexible, the extra residues do not interfere with the native-state stability. From these experimental results, we can thus conclude that when the N-terminal region of a protein has a rigid structure, expression of this protein by E. coli, which adds the extra methionine residue, destabilizes the N state, but that when the N-terminal region is flexible, expression of the protein by E. coli does not interfere with the native-state stability.

**Electric properties of authentic and recombinant α-lactalbumin**

The results of the electrophoresis and ion-exchange chromatography show that the recombinant protein is more negatively charged than the authentic one. It is understood, however, that the side-chain of a methionine residue does not ionize at neutral pH, so that there is no difference in the number of ionizable groups between the authentic and recombinant proteins. In fact, our electrophoresis and ion-exchange chromatography data show that there is no difference in the electric charge between the proteins in the presence of 8 M urea. This means that some of the ionizable groups that have a pK\(_a\) near 7.0 experience a change in pK\(_a\) due to the structural folding of the protein. There are two such ionizable groups, the imidazole group of histidine and the N-terminal amino group, which have intrinsic pK\(_a\) values of 6.5 and 8.0, respectively. If we compare the X-ray structures of the two proteins, no significant differences are observed near the histidine side-chains. However, there is a noticeable difference in the structures around the N-terminal amino groups. The N-terminal amino group is hydrogen-bonded to the oxygen atom of Thr38 and is closer to the side-chains of Asp37 and Asp83 in the authentic protein (Figure 8), and both of these may increase the pK\(_a\) value of the N-terminal amino group through electrostatic interactions. A study of the pH-dependence of the unfolding transition of authentic bovine α-lactalbumin has shown that the N-terminal amino group of the protein has an abnormally high pK\(_a\) value (pK\(_a\) = 8.9) in the N state, which is normalized on unfolding from the N to the molten globule state (Kuwajima et al., 1981).

It should also be mentioned that the ΔpK\(_a\) of the N-terminal group between the recombinant and authentic proteins leads to a difference in the native-state stability between the proteins, but this stability difference is expected to be much smaller than the ΔΔG\(_U\) estimated from equation (6) at pH 7.0. The stability difference (ΔΔG\(_U\)(ΔpK\(_a\))) due to the ΔpK\(_a\) is known to be given by:

\[
\Delta \Delta G_U(\Delta \text{p}K_a) = RT \ln[(1 - K_a(\text{rec})/[H^+])/(1 - K_a(\text{auth})/[H^+])] 
\]

where K\(_a\)(rec) and K\(_a\)(auth) are the dissociation constants of the N-terminal amino groups of the recombinant and authentic proteins, respectively, and [H\(^+\)] is the hydrogen-ion concentration (Tanford, 1970). If we assume that the pK\(_a\)(rec) and pK\(_a\)(auth) are 8.0 and 8.9, respectively, the above equation gives a ΔΔG\(_U\)(ΔpK\(_a\)) of 0.06 kcal/mol at pH 7, and this is negligibly small compared with the observed ΔΔG\(_U\). Thus, the ΔpK\(_a\) of the
N-terminal amino group reasonably interprets the differences in the electric properties between the proteins observed by electrophoresis and ion-exchange chromatography, but it is not sufficient for interpreting the stability difference between the proteins.

**Folding of goat \(\alpha\)-lactalbumin**

Because the presence of the N-terminal methionine residue in the recombinant protein changes the thermodynamic stability of the native state, this system is useful for investigating the role of the N terminus in the folding of \(\alpha\)-lactalbumin. We thus investigated the refolding and unfolding kinetics of the proteins by stopped-flow CD measurements. The results show that the rate of unfolding of the recombinant protein is faster than that of the authentic protein (Table 2), whereas the refolding rates are very similar in the two proteins (Table 3). This shows that the stability difference is caused by the enhanced unfolding rate of the recombinant protein, and this is interpreted in terms of the difference in the free energy of the unfolding transition (\(\Delta\Delta G_U\)) and the difference in the activation free energy (\(\Delta G^*\)) of unfolding. The \(\Delta\Delta G_U\) is known to be given by the ratio of the unfolding rate constants as:

\[
\Delta\Delta G_U = -RT\ln\left(\frac{k_u(\text{rec})}{k_u(\text{auth})}\right)
\]

where \(k_u(\text{rec})\) and \(k_u(\text{auth})\) represent the unfolding rate constants for the recombinant and authentic proteins, respectively. Because \(k_u(\text{rec})\) is 5.7 times larger than \(k_u(\text{auth})\) at 5.4 M GdnHCl, \(\Delta\Delta G_U\) is estimated to be 1.0 kcal/mol, and this value is nearly identical with the estimated to be 1.0 kcal/mol, and this value is interpreted in terms of the rate constants as:

\[
\Delta\Delta G_U = -RT\ln\left(\frac{k_u(\text{rec})}{k_u(\text{auth})}\right)
\]

Materials and Methods

**Chemicals**

GdnHCl was of a specially prepared reagent grade for biochemical use from Nacalai Tesque, Inc. (Kyoto). The concentration of GdnHCl was determined from the refractive index at 589 nm with an Atago 3T refractometer (Pace, 1986). Cyanogen bromide (CNBr) was purchased from Nacalai Tesque Inc. (Kyoto). Authentic goat \(\alpha\)-lactalbumin was prepared from fresh goat milk by the method described (Kuwajima et al., 1980). A Resource\textsuperscript{TM}-Q anion exchange column was purchased from Pharmacia Biotechnology, Inc. (Sweden) and a \(\mu\) BONDASPHERE \(5 \mu\) C4 300 Å reversed-phase column was supplied by Nihon Waters Ltd (Japan).

**Expression and purification of recombinant goat \(\alpha\)-lactalbumin**

The expression system of goat \(\alpha\)-lactalbumin and the procedures for the refolding and purification of the protein have been reported by Kumagai et al. (1990) and recently improved by Uchiyama et al. (1995) utilizing a T7 promoter (Studier & Moffatt, 1986). In brief, the protein expressed in E. coli BL21(DE3) as inclusion bodies was solubilized in 8 M urea containing 20 mM Tris-HCl (pH 8.0) and first purified using a DEAE-Sepharose FF column. The eluted protein was reduced by 50 mM dithiothreitol and dialyzed against 20 mM Tris-HCl (pH 8.0) at 4 °C to remove urea. Refolding of the reduced \(\alpha\)-lactalbumin was performed as described (Sawano et al., 1992), with slight modifications, in a solution containing 20 % (v/v) glycerol, 20 mM Tris-HCl (pH 8.0), 1 mM CaCl\(_2\), 6 mM glutathione, 0.6 mM oxidized glutathione, 3.3 mM \(\alpha\)-lactalbumin at 15 °C for more than 20 hours. The refolding process was monitored by the appearance of a sharp peak on a reversed-phase HPLC chromatogram detected by UV-absorbance at 215 nm using a C4 column with a linear gradient elution of 28 %-52 % acetonitrile in the presence of 0.1 % (v/v) trifluoroacetic acid at a flow rate of 0.5 ml per minute. The refolded protein was then purified by DEAE-Sepharose FF and phenyl-Sepharose CL column chromatographies as described by Lindahl & Vogel (1984). Concentrations of authentic and recombinant goat \(\alpha\)-lactalbumin were determined spectrophotometrically using an extinction coefficient of \(E_280 = 20.1\) for both (Kuwajima et al., 1980). No free cysteinyl residues were detected in the folded recombinant protein by thiol content analysis (Eillman, 1959; Riddle et al., 1979).

**Preparation of methionine-free recombinant goat \(\alpha\)-lactalbumin**

The methionine-free protein was prepared according to the method described by Kim et al. (1997) with slight modifications. Recombinant goat \(\alpha\)-lactalbumin was dissolved in 70% (v/v) formic acid and treated with 100 mM CNBr (50-100-fold molar excess over the protein concentration) for 24 hours in the dark at room temperature. The cleaved product was diluted ten times with water and dialyzed against 10 mM HCl, then dialyzed against 10 mM Tris-HCl (pH 8.5) containing 1 mM CaCl\(_2\). Finally, the protein solution was purified on a Q-Sepharose FF column, which had been equilibrated with 20 mM Tris-HCl (pH 8.5) containing 1 mM CaCl\(_2\) and eluted with a linear gradient of NaCl from 0 M to 0.5 M. The mobilities and retention times of the eluted product were compared with an authentic sample.
fractions were checked by native PAGE and anion-exchange HPLC, and compared with those of the authentic protein under the same conditions. The mass of the methionine-cleaved protein was determined by mass spectrometric analysis, and the removal of the N-terminal methionine residue was confirmed by the N-terminal sequence analysis. The concentration of the CNBr-cleaved protein was calculated using the same extinction coefficient as that given above.

**Mass spectrometric analysis**

Mass spectrometric analyses of the authentic, recombinant and methionine-free proteins were carried out by the MALDI-TOF-MS mass spectroscopic method. Sinapinic acid mix protein samples were used as the matrix, and the spectra were taken in Reflex (Bruker).

**N-terminal sequence analysis**

N-terminal sequencing of recombinant, authentic, and CNBr-cleaved proteins were carried out using an automated Applied Biosystem sequencer model 477a equipped with a model 120A on-line PTH amino acid analyzer. In this study we analyzed the first five residues in the proteins.

**Equilibrium CD measurements**

Equilibrium CD spectra were taken on a Jasco J-720 spectropolarimeter using an optical cuvette with a path length of 1.00 mm for measurements in the peptide region and 10.0 mm for measurements in the aromatic region. The CD spectra of the protein were measured in 50 mM sodium cacodylate, 50 mM NaCl (pH 7.0) containing 1 mM CaCl$_2$. The solutions for the GdnHCl-induced equilibrium unfolding studies were prepared in the same buffer containing various concentrations of GdnHCl. The mean residue ellipticity was calculated as a function of GdnHCl concentration at 25°C by taking 113 as the mean residue mass. The protein concentration in the equilibrium measurements was 0.15-0.2 mg/ml.

The apparent fractional extent ($F_{app}$) of unfolding was calculated by:

$$F_{app} = \frac{\theta_{obs} - \theta_N}{\theta_U - \theta_N}$$  

where $\theta_{obs}$ is the observed ellipticity, and $\theta_N$ and $\theta_U$ are the ellipticities in the native (N) and the fully unfolded (U) states, respectively. The $\theta_N$ and $\theta_U$ values are assumed to linearly depend on the GdnHCl concentration (C) as $\theta_N = \theta_0 + \alpha_1 C$ and $\theta_U = \theta_0 + \alpha_2 C$. The N state baseline was calculated from the ellipticity values between 0.5 and 2 M and between 0.4 and 1.8 M GdnHCl, and the U state baseline was from the values between 4.5 and 6.2 M and between 3.8 and 6.2 M GdnHCl for the authentic and recombinant proteins, respectively.

**Kinetic measurements**

Refolding and unfolding reactions of the authentic and recombinant proteins were induced by GdnHCl concentration jumps, which were performed by a stopped-flow CD apparatus (UNISOKU Inc., Japan) installed in the cell compartment of the J-720 spectropolarimeter (Arai & Kuwajima, 1996). All kinetics were measured in the presence of 1 mM CaCl$_2$, 50 mM NaCl, and 50 mM sodium cacodylate at pH 7.0 and 25°C. The dead time of the stopped-flow CD apparatus was 25 ms when a 4 mm cuvette was used. The concentration of the protein stock solution was about 1.5-2.0 mg/ml. The initial protein solutions before the concentration jump contained 1.0 M and 5.5 M GdnHCl for unfolding and refolding experiments, respectively. The diluent solution contained the same buffer (50 mM sodium cacodylate, 50 mM NaCl, and 1 mM CaCl$_2$, pH 7.0) and an appropriate concentration of GdnHCl. The two solutions were mixed with a mixing ratio of 1:10.

**X-ray crystallographic studies**

The crystal of recombinant goat $\alpha$-lactalbumin was grown by the vapor diffusion method with a hanging drop in a chamber where the temperature was controlled at 20°C. The data were collected by an automated area detector system, DIP2000, on an X-ray generator with a bent mirror system at 9.5°C. Data processing and reduction was performed using DENZO and SCALEPACK programs (Otwinowski, 1993). The crystallographic data, the diffraction intensity statistics, and the refinement statistics are listed in Table 4. The crystal structure was solved on the basis of the model structure of baboon $\alpha$-lactalbumin (Acharya _et al_, 1989) by the molecular replacement method (Brünger, 1990) and was refined by a slow-cooling molecular-simulated annealing method in the X-PLOR 3.1 program suite (Brünger, 1992).

**Theoretical estimation of $\Delta \Delta G_u$ between recombinant and authentic goat $\alpha$-lactalbumin**

The $\Delta \Delta G_u$ value was calculated by the method described by Oobatake & Ooi (1993). In this calculation, every atom was identified as belonging to one of seven atomic groups: aliphatic C, aromatic C, hydroxyl O, amide N, carbonyl C, carbonyl O, and sulphur S. In addition, the accessible surface area (ASA) of each atom in the N state (except hydrogen) was calculated by the method described by Richmond (1984) using the coordinates of the X-ray crystal structures. Because the N-terminal methionine residue is present only in the recombinant protein, the $\Delta \Delta G_u$ was assumed as a first approximation to be equal to the free energy change of unfolding ($\Delta G^U$) of the methionine residue. For the ASA of atoms in the methionine residue in the U state, the values calculated by Shrike & Rupley (1973) were used. It was also assumed that the $\Delta G^U$ and $\Delta H^U$ are proportional to the change in the ASA ($\Delta \xi$ for the ith atomic group) of the atoms on unfolding according to Oobatake & Ooi (1993). Thus:

$$\Delta G^U = \Delta G^U_0 + \Delta G^U_C$$
$$\Delta H^U = \Delta H^U_0 + T \Delta S^U_C$$
$$\Delta G^U_i = \Sigma \xi_i h_i \Delta \xi_i$$

where $\xi_i$ and $h_i$ are proportionality constants for the seven atomic groups. Although the change in the conformational entropy, $\Delta S^U_C$, was also assumed to be proportional to the $\Delta \xi_i$ values in the original Oobatake & Ooi (1993) method, this assumption may not be correct for the extra methionine residue of recombinant goat $\alpha$-lactalbumin due to the rigid nature of this residue as shown by the X-ray structure; the methyl group of the side-chain is interacting with the Gln2 side-chain, with
the distance between the Cα atom of the methyl sidechain of Met0 and the Cα atom of Gln2 side-chain being 3.5 Å. Moreover, the N-terminal amino group in the recombinant protein is hydrogen-bonded with the carboxyl oxygen atom of the Gln39 side-chain. Thus the \(-\Delta S^i_{\text{ASA}}\) value (−5.9 kcal/mol) obtained from Table 8 of Oobatake & Ooi (1993) was employed (see Discussion).

The contribution of other residues to the \(\Delta G_{\text{U}}\) value was also estimated by \(\Sigma (g_{i,j} + g_{j,i})\Delta x^i\), where \(g_{i,j}\) is a proportionality constant and \(\Delta x^i\) is the difference in the ASA value of the \(i\)th atomic group between the authentic and recombinant proteins in the N state. Here, \(-\Delta S^i_{\text{ASA}}\) was assumed to be proportional to the change in the ASA values following the original Oobatake & Ooi (1993) method. The values obtained for Mol A and Mol B of the crystal structure of the authentic protein were averaged. Since Gln1 of the authentic protein is more exposed to solvent in the unfolded state, the difference in the ASA of the atoms of Gln1 between the authentic and recombinant proteins in the unfolded state was also estimated to account for the estimation of \(\Delta G_{\text{U}}\) and \(\Delta G_{\text{S}}\). Thus the \(\Delta G_{\text{U}}\) value estimated has been found to be less than 1 kcal/mol.

Protein Data Bank accession number

The coordinates have been deposited in the Brookhaven Protein Data Bank with accession number 1HMK.

Acknowledgments

The authors thank Mr K. Maki, Mr T. Makio, Mr T. Hayashi, and Mr M. Mizuguchi for their assistance in this work. This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture and Science of Japan. T.K.C. is a postdoctoral fellow of the Japan Society for the Promotion of Science.

References


and its refolding to biologically active protein. *Protein Eng.* 3, 449-452.


Edited by P. E. Wright

(Received 22 June 1998; received in revised form 20 October 1998; accepted 22 October 1998)