Effect of a Single Amino Acid Substitution on the Folding of the α Subunit of Tryptophan Synthase†

C. Robert Matthews,* Mark M. Crisanti,‡ Joanna T. Manz, and Gary L. Gepner

ABSTRACT: The urea-induced unfolding of a missense mutant of the α subunit of tryptophan synthase from *Escherichia coli* involving the replacement of Gly by Glu at position 211 has been monitored by absorbance changes at 286 nm. Like the wild-type protein, the equilibrium unfolding curve demonstrates the presence of one or more stable intermediates. Comparison of these results with those from the wild-type α subunit [Matthews, C. R., & Crisanti, M. M. (1981) *Biochemistry 20*, 784] shows that the transition from the native conformation to the stable intermediates is displaced to higher urea concentration in the mutant α subunit; however, the transition from the intermediates to the unfolded form is unaffected. Kinetic studies show that the amino acid replacement slows the rate of unfolding by an order of magnitude. The effect on refolding rates is complex. One phase, previously assigned to proline isomerization [Crisanti, M. M., & Matthews, C. R. (1981) *Biochemistry 20*, 2700], is unaffected by the substitution. The rate of the second phase, which is urea dependent down to about 1 M urea, is slower than the corresponding phase in the wild-type protein by approximately a factor of 2. Below about 1 M urea, the rate of this phase becomes urea independent and identical with that of the wild-type α subunit. This change in urea dependence has been ascribed to a change in the nature of the rate-limiting step for this process from one involving folding to one involving proline isomerization. The results support the folding model for the α subunit proposed previously [Matthews, C. R., & Crisanti, M. M. (1981) *Biochemistry 20*, 784] and clarify the role of proline isomerization in limiting the rate of folding.

It is generally accepted that the amino acid sequence of a protein determines the three-dimensional structure that the protein spontaneously assumes in solution (Anfinsen, 1973). Although a great number of studies have been performed to determine the mechanism by which this rapid and efficient conformational change occurs (Baldwin, 1975; Anfinsen & Scheraga, 1975; Nemethy & Scheraga, 1977; Creighton, 1979; Privalov, 1979; Kim & Baldwin, 1982), as yet little progress has been made in obtaining a molecular level description of this complex process.

In order to obtain information on the details of the folding mechanism, we have been investigating the effects of single amino acid substitutions on the stability and folding of a globular protein. The rationale for this approach is based upon the above-cited role of the amino acid sequence in determining the structure. If this hypothesis is correct, then alterations in the sequence might be expected to influence the stability and/or folding of the protein. Identification of particular amino acids that play a significant role in folding might then lead to a better understanding of this process.

The subject of our studies has been the α subunit of tryptophan synthase. The α subunit is a single polypeptide of molecular weight 28 700 that contains no prosthetic groups (Creighton & Yanofsky, 1970); unfolding and refolding kinetics should be strictly first order. In addition to the wild-type protein, over two dozen missense mutants have been isolated and the position and nature of the amino acid replacements identified by Yanofsky and his colleagues (Yanofsky, 1967; Yanofsky & Horn, 1972; Murgola & Yanofsky, 1974).

In previous studies on this system (Matthews et al., 1980), it was observed that the substitution of glutamic acid for glycine in wild-type protein at position 211 resulted in a small increase in free energy and rather substantial increases in enthalpy and entropy for the thermal unfolding transition. In the present paper, the effect of urea, a chemical denaturant, on the stability and kinetics of folding of this Glu-211[1] α subunit are reported. Comparison of the results with those obtained previously for the wild-type α subunit (Matthews & Crisanti, 1981; Crisanti & Matthews, 1981) clearly shows that the amino acid at position 211 plays an important role in the stability and folding of this protein. The manner in which the kinetic processes are affected by the amino acid replacement substantiates the folding model proposed previously (Matthews & Crisanti, 1981; Crisanti & Matthews, 1981) and further clarifies the role of proline isomerization in limiting the rate of folding of the α subunit.

Materials and Methods

Materials

α Subunit of Tryptophan Synthase. The wild-type α subunit of tryptophan synthase from *Escherichia coli* (EC 4.2.1.20) and the Glu-211 mutant α subunit were isolated from strains B8/F'B8 and A46, respectively, by the method of Kirschner et al. (1975). For improvement of the yield, a two-step $(NH_4)_2SO_4$ precipitation procedure was inserted between the first DE-52 column and the hydroxylapatite column. The fractions from the DE-52 column that contained the α subunit were pooled, and $(NH_4)_2SO_4$ was added up to 35% at 0 °C. The precipitate was discarded, and the $(NH_4)_2SO_4$ concentration in the supernatant was increased to 70%. The precipitate was resuspended in 0.1 M imidazolium chloride, pH 7.0, 0.2 mM DTE, and 1 mM KH_2PO_4 and dialyzed against successive changes of this buffer until the residual $(NH_4)_2SO_4$ concentration was less than 1 mM. The protein was then carried through the purification described

† From the Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802. *Received August 23, 1982.* This work was supported by U.S. Public Health Service Grant GM 23303 and by Research Career Development Award 1 K04 AG 00153 (to C.R.M.).

‡ Present address: Department of Chemistry, Princeton University, Princeton, NJ 08540.

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; DTE, dithioerythritol; $NaDodSO_4$, sodium dodecyl sulfate; Glu-211, mutant protein with Glu at position 211 in the amino acid sequence; CD, circular dichroism.

0006-2960/83/0422-1445$01.50/0

by Kirschner et al. (1975). The purity of the proteins was ascertained by both $NaDodSO_4$ and native polyacrylamide gel electrophoresis, where only a single band was apparent in the electrophoretograms. The enzymes were stored in 100 mM KH_2PO_4, pH 7.8, 55% ammonium sulfate, 1 mM EDTA, and 1 mM DTE at 4 °C; the activity remained unchanged under these conditions for at least 2 months.

The activity of the α subunit was determined by measuring its ability to enhance the activity of the β_2 subunit in the condensation of indole and serine to form tryptophan (Kirschner et al., 1975); the maximum specific activity of the α subunit in this assay is 5500 units mg^{-1}. Samples of the α subunit used in the folding studies had specific activities ranging from 4500 to 5300 units mg^{-1}. Protein concentration was determined from the optical spectrum by using a specific absorption $E^{1\%}_{278} = 4.4$ for wild-type (Adachi et al., 1974) and Glu-211 mutant proteins (Matthews et al., 1980). The molar extinction coefficient was calculated to be 12 600 M^{-1} cm^{-1} by using a molecular weight of 28 700]Yanofsky et al., 1967; Li & Yanofsky, 1972).

Chemicals. Ultrapure urea was purchased from Schwarz/Mann and used without further purification. All other chemicals were reagent grade.

Methods

Spectroscopy. Ultraviolet difference spectroscopy measurements were made by using the tandem cell technique described by Herskovits (1967) on a Cary 118CX spectrophotometer. The protein concentration was held constant throughout each experiment and was generally in the range of 0.4–1.5 mg mL^{-1}. The method by which kinetic measurements of the transient response in absorbance to rapid changes in urea concentrations were obtained has been described in detail (Crisanti & Matthews, 1981). Analysis of the complex kinetic responses was done by an exponential stripping procedure described previously (Crisanti & Matthews, 1981).

Circular dichroism measurements were performed on a JASCO J-40 spectropolarimeter. The CD spectra were recorded over the wavelength range from 200 to 350 nm. Mean residue ellipticity values were calculated according to

$$[\theta]_\lambda = \frac{\theta_{obsd}\text{MRW}}{10dc}$$

where $[\theta]_\lambda$ is the mean residue ellipticity at wavelength λ, θ_{obsd} is the observed ellipticity in degrees, MRW is the mean residue weight, d is the cell path length in centimeters, and c is the protein concentration in grams per milliliter. The mean residue weight was calculated to be 107.2 g from the amino acid sequence of the α subunit. A 0.1-cm cell path length was used to keep the absorbance at 222 nm below 0.8 unit. The protein concentration was 0.3 mg mL^{-1}.

Results

The concept of determining the role of an individual amino acid in the folding of a protein by substituting other amino acids at the same site implicitly assumes that the replacement does not greatly alter the conformation of the protein. Substitutions that cause significant changes in the three-dimensional structure could also alter the folding pathway. The effect of replacing Gly with Glu at position 211 on the structure of the α subunit was determined by comparing the CD spectra for the wild-type and Glu-211 mutants in the region from 200 to 270 nm (Figure 1). The spectra are virtually coincident, demonstrating that this amino acid replacement does not perturb the secondary structure of the α subunit to any great extent. Previous studies comparing the affinities of wild-type and Glu-211 α subunits for the β_2 subunit of tryptophan synthase and the reaction of the mutant α subunit with antisera to the wild-type protein showed that changes in tertiary structure must also be rather limited (Murphy & Mills, 1968). The retention by the Glu-211 mutant of approximately 35% of the activity of the wild-type protein in the condensation of indole and glyceraldehyde 3-phosphate (Patterson et al., 1977) further substantiates the presumption that the two proteins have quite similar conformations.

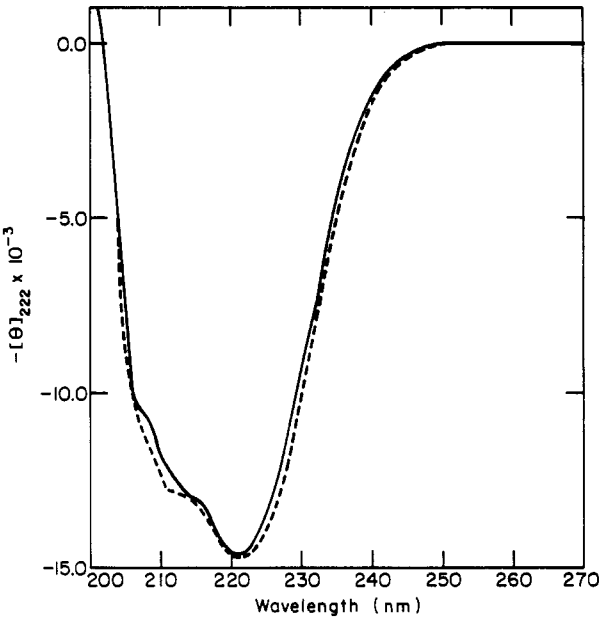


FIGURE 1: Mean residue ellipticity of wild-type (—) and Glu-211 (---) mutant α subunits in 10 mM KH_2PO_4, pH 7.8, 0.2 mM EDTA, and 0.1 mM 2-mercaptoethanol at 25 °C.

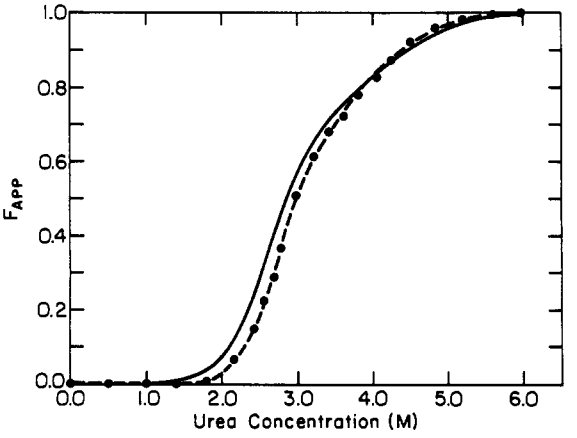


FIGURE 2: Fraction of the total change in the extinction coefficient at 286 nm, F_{app}, as a function of urea concentration for wild-type (—) and Glu-211 (---) mutant α subunits in 10 mM KH_2PO_4, pH 7.8, 0.2 mM EDTA, and 0.1 mM DTE at 25 °C.

Equilibrium Studies. Previous studies on the wild-type α subunit (Matthews & Crisanti, 1981) have shown that the urea-induced unfolding transition can be readily monitored by difference ultraviolet spectroscopy at 286 nm. The results of similar experiments on the Glu-211 mutant in a buffered solution at pH 7.8, 25 °C, are shown in Figure 2. The ordinate, F_{app}, was calculated from the equation $F_{app} = (\epsilon_{obsd} - \epsilon_N)/(\epsilon_U - \epsilon_N)$, where F_{app} is the fraction of apparent change, ϵ_{obsd} is the observed value of the extinction coefficient at 286 nm, and ϵ_N and ϵ_U are values for this parameter for the native and unfolded forms, respectively. The values of ϵ_N and ϵ_U were

found to vary linearly with urea concentration in the native and unfolded base-line regions, respectively. Values for ϵ_N and ϵ_U in the transition region were obtained by linear extrapolation from the base-line regions. For comparison, the values of F_{app} for the wild-type protein under the same conditions (Matthews & Crisanti, 1981) are also shown in Figure 2.

It has previously been shown that the urea-induced unfolding transition of the wild-type α subunit involves two stable intermediates (Crisanti & Matthews, 1981). The conversion from the native to the intermediate forms occurs in the range of 1.5 to approximately 3 M urea and the conversion from the intermediates to the unfolded forms in the range from 3 to 6 M urea at pH 7.8 and 25 °C. The appearance of these intermediates accounts for the biphasic nature of the transition curve, with an inflection at approximately 3 M urea. In Figure 2, it can be seen that the transition curve for the Glu-211 mutant is also biphasic; however, the change in extinction coefficient reflecting the conversion from the native to the intermediate forms is displaced to higher urea concentration. In contrast, the transition curve for the subsequent conversion of the intermediates to the unfolded forms coincides, within experiment error, with that of the wild-type α subunit. This behavior indicates that the replacement of Gly with Glu at position 211 serves to increase the stability of the native conformation of the mutant protein relative to the intermediates but does not alter the stability of the intermediates relative to the unfolded forms.

Kinetic Studies. The effects of the replacement of Gly by Glu at position 211 on the rates of urea-induced unfolding and refolding of the α subunit were investigated by monitoring the transient responses in the absorbance at 286 nm following rapid changes in urea concentration. As in the case of the wild-type α subunit, the unfolding kinetics are satisfactorily described by a single exponential decay that accounts for all of the change in absorbance expected from the equilibrium results (Figure 2). The refolding kinetics are more complex and are again similar to the changes observed for the wild-type protein: for jumps ending near 1.5 M urea, an initial rapid change in absorbance which accounts for approximately half of the total expected from equilibrium results is followed by two slower kinetic phases.

The dependence on the final urea concentration of the relaxation time of the single phase observed in unfolding and the two slower phases in refolding is shown in Figure 3. For comparison, relaxation times for the wild-type protein are also shown in Figure 3. No attempt was made to monitor the initial rapid phase in refolding of the Glu-211 mutant since in previous studies on the wild-type protein this phase was obscured by effects that were attributed to light scattering (Crisanti & Matthews, 1981). Unfolding jumps were inititated at 0 M urea and refolding jumps from either 5 or 6 M urea. Although the protein is not completely unfolded at 5 M urea, the relaxation times were found to depend only on the final urea concentration (data not shown), as expected (Eigen & De-Maeyer, 1963).

The relaxation time for the single phase in unfolding proceeds through a maximum near 3 M urea for the Glu-211 mutant and then decreases logarithmically at higher urea concentrations. The behavior is similar to that observed for the wild-type α subunit (Figure 3); however, the relaxation times are close to an order of magnitude longer than those of the wild-type α subunit at the same final urea concentration. The replacement of Gly with Glu at position 211 obviously has a substantial effect on the molecular events associated with this process.

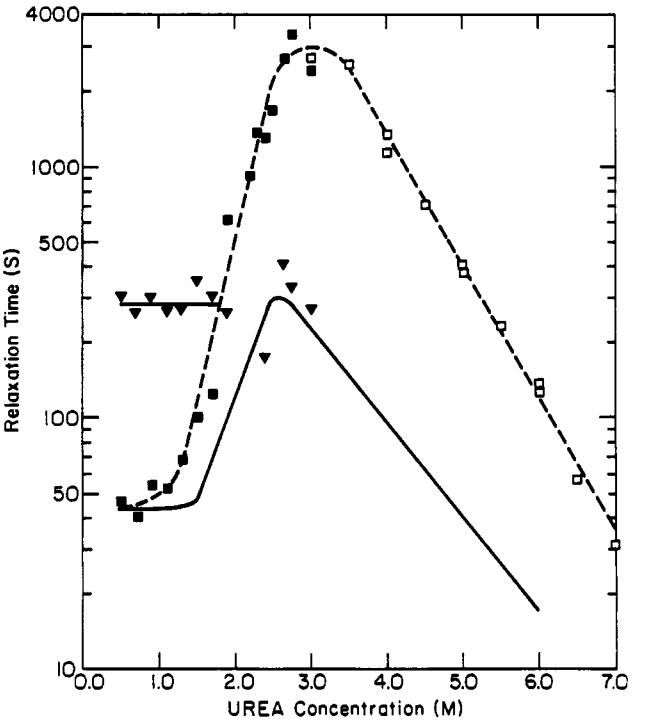


FIGURE 3: Relaxation times at the indicated final urea concentrations of the Glu-211 mutant α subunit for the single slow phase in unfolding (□) and the two slow phases in refolding (■, ▼) in 10 mM KH_2PO_4, pH 7.8, 0.2 mM EDTA, and 0.1 mM DTE at 25 °C. For comparison, solid lines indicating the values for the wild-type α subunit for unfolding and refolding (Figure 1; Crisanti & Matthews, 1981) under the same conditions are also shown. Unfolding jumps were initiated at 0 M urea and refolding jumps at 5 or 6 M urea. The lines connecting the data points are drawn to aid the eye and do not indicate fits to theory.

The dependence of the relaxation times for the two slow phases observed in refolding upon final urea concentration is more complex. For refolding jumps ending near 3 M urea, two slow kinetic phases were observed: a slower phase whose decay time is equal, within experimental error, to that of the single phase detected in unfolding and a faster phase not seen in unfolding (Figure 4). The equality of the relaxation times for the single phase in unfolding and the slower phase in refolding demonstrates that the unfolding process is reversible. The appearance of a second slow phase in refolding that is absent in unfolding can be explained in terms of a kinetic model for the folding of the α subunit (see Discussion).

As the final urea concentration in refolding is decreased below 3 M, the relaxation time of the slower phase decreases in a logarithmic fashion until it becomes equivalent to that of the previously designated faster phase near 2 M urea (Figure 3). Below this urea concentration, two kinetic phases again become apparent. The slower of the two phases below 1.8 M urea is independent of urea concentration and has an average relaxation time of 282 s, identical with that of the wild-type α subunit. The faster phase has a relaxation time that decreases as the urea concentration decreases down to approximately 1 M urea. Below 1 M urea, the relaxation time of this faster phase becomes independent of urea concentration and approaches an average value of 44 s. This value is also the same as that of the wild-type protein, within experimental error.

The abrupt change in the urea dependence of the more rapid of the two slow phases in both the wild-type and Glu-211 mutants at approximately 1 M urea and the fact that the relaxation times for both proteins become equivalent below this final urea concentration combine to suggest that the nature of the rate-limiting step for this kinetic phase changes in this

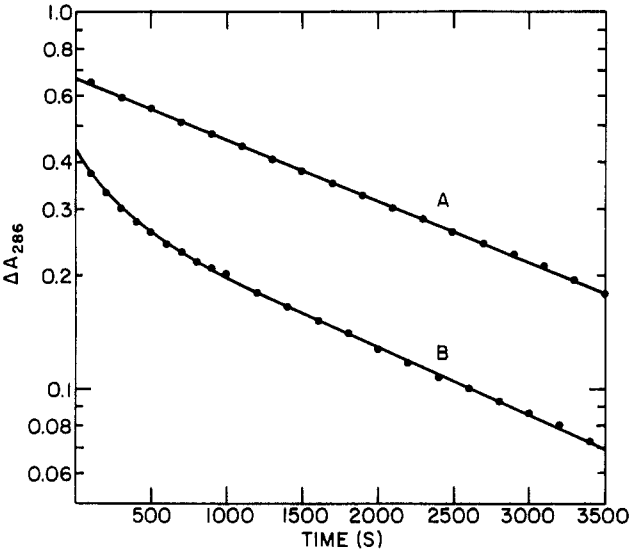
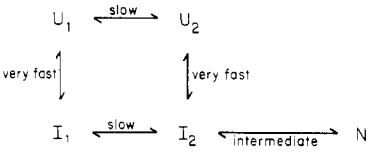


FIGURE 4: Time dependence of the change in absorbance at 286 nm for the Glu-211 mutant for (A) unfolding and (B) refolding to a final urea concentration of 3.0 M. The solid lines indicate fits of the data, shown by the filled circles, to (A) a single exponential with a relaxation time of 2680 s and (B) two exponentials with relaxation times of 273 and 2440 s. Unfolding was initiated at 0 M urea and refolding in 5 M urea in 10 mM KH_2PO_4, pH 7.8, 0.2 mM EDTA, and 0.1 mM 2-mercaptoethanol at 25 °C.

Scheme I

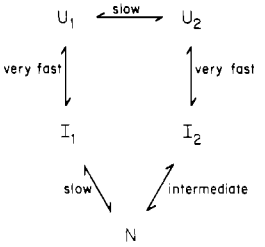


region. This possibility was tested for the wild-type protein by comparing the pH dependence and activation energy of this phase for refolding jumps ending at 1.5 and 0.5 M urea. It had previously been observed for jumps ending at 1.5 M urea that the relaxation time was pH dependent and had an activation energy of 14.5 kcal mol^{-1} (Crisanti & Matthews, 1981). In contrast, for jumps ending at 0.5 M urea, no pH dependence was observed in the relaxation time for this phase for the wild-type protein over the pH range from 6 to 8 (data not shown). Also, the activation energy was found to be 18 kcal mol^{-1} (data not shown). Thus, the change in behavior of this phase with regard to urea and pH dependence and an increase of 3.5 kcal mol^{-1} in the activation energy demonstrates that the nature of the rate-limiting step changes when the final urea concentration falls below approximately 1 M urea.

Discussion

Kinetic Model of Folding for the α Subunit. Although the dependence of the relaxation times observed in the unfolding and refolding of the Glu-211 mutant α subunit upon the final urea concentration is complex, the results can be interpreted in terms of the kinetic model for the folding of the wild-type α subunit proposed previously (Matthews & Crisanti, 1981; Crisanti & Matthews, 1981). The model (Scheme I) which is based upon the results of urea gradient gel electrophoresis and rapid mixing urea jump experiments similar to those described in this paper, is as shown. In this model (Scheme I), U_1 and U_2 are two forms of unfolded protein, I_1 and I_2 are stable intermediates, and N is the native form. The rate of the step linking U_1 and U_2 has not been observed directly but is presumed to be slow on the basis of urea gradient gel electrophoresis experiments (Matthews & Crisanti, 1981).

Scheme II



Precedence for the existence of two unfolded forms with different kinetic properties in refolding has been established (Garel & Baldwin, 1975; Nall et al., 1978). The slow step linking the unfolded forms has been attributed to cis–trans isomerization at the peptide linkage in X–Pro peptide bonds (Brandts et al., 1975). The fast steps linking U_1 to I_1 and U_2 to I_2 have relaxation times less than 10 ms and result in a decreased radius of gyration for both of the intermediates compared to the unfolded forms. The slow step linking I_1 to I_2 has been assigned to proline isomerization on the basis of its urea and pH independence and an activation energy of 17.6 kcal mol^{-1}. The step linking I_2 to N, whose relaxation time is designated intermediate in Scheme I, was not readily explained in terms of proline isomerization. The relaxation time depends upon urea concentration, at least down to ~1.5 M urea. Also, for refolding jumps ending at 1.5 M urea, the refolding rate constant is pH dependent and has an activation energy of 14.5 kcal mol^{-1}.

The results of the kinetic studies on the Glu-211 α subunit substantiate the mechanism of folding described above. An alternative kinetic model (see Scheme II) for folding that could not be ruled out on the basis of the results from the wild-type protein postulates that both intermediates are in direct equilibrium with the native conformation. The notation is the same as that used in model I (see Scheme I). For the Glu-211 mutant protein, the qualitative descriptors, slow and intermediate, that have been used to label the rates of the steps linking I_1 and I_2 to N are only appropriate at low final urea concentrations. Above ~2 M urea, the relative rates of these two steps interchange (Figure 3).

If it is assumed that the wild-type and Glu-211 mutant proteins fold by the same mechanism, the kinetic data for the Glu-211 mutant can be used to rule out model II (see Scheme II). *Refolding* jumps ending in the transition region resulted in two slow phases, one with a relaxation time of 2440 s and a second with a relaxation time of 273 s (Figure 4). *Unfolding* jumps to this same final urea concentration invariably resulted in a single exponential with a time constant within experimental error of that of the *slower* of the two phases found in refolding. These observations are not consistent with the results expected from model II (see Scheme II). If this latter model were correct, unfolding should have been dominated by the more rapid of the two slow phases observed in refolding, not the slower phase. Thus, we conclude that model II (Scheme II) is incorrect.

The interpretation of the kinetic data for the Glu-211 mutant protein requires the correct connectivity of the two slow phases detected in refolding in the transition region with the slow phases seen for jumps into the native base-line region. The ambiguity arises because the relaxation times of these two phases become identical near 2 M urea. Consideration of the refolding model discussed above and the results for the kinetics of binding of a substrate analogue (Crisanti & Matthews, 1981) provides an answer to this question. The observation of a single kinetic phase with a relaxation time of 2680 s for

unfolding jumps ending at 3 M urea shows that the process to which this corresponds is one which links the native form to subsequent forms that appear in unfolding. For refolding jumps, if this phase were to become the 282-s phase at low urea concentration, then one would expect that this phase would limit the appearance of the native form in refolding. This prediction is contrary to the results of experiments which followed the appearance of the native form by monitoring the binding of a substrate analogue (Crisanti & Matthews, 1981). These results clearly showed that the native conformation forms in two phases, approximately 75% in a phase with a 40-s time constant and 25% in a phase with a 200 ± 50 s time constant. Thus, the phase with a time constant of 2680 s at 3 M urea becomes the 44-s phase for jumps ending below 1 M urea. The ~300-s time constant phase at 3 M urea becomes the urea-independent 282-s phase below 2 M urea.

Given the kinetic model for folding and the connectivity of the relaxation times described above, it is possible to interpret the effect of a single amino acid substitution and to understand more clearly the role of proline isomerization in the folding of the α subunit. For the Glu-211 mutant, the observation of a kinetic phase with a urea-independent relaxation time for refolding jumps ending below 2 M urea supports the assignment of proline isomerization as the rate-limiting step in this process. The rate of proline isomerization is not expected to depend on urea concentration. Also, the absolute value of the relaxation time for the mutant, 282 s, is identical with that of the wild-type α subunit, 282 s. This equivalence is expected since the amino acid replacement does not involve a proline and is not at a position adjacent to proline where nearest-neighbor effects could play a role (Brandts et al., 1975; Grathwohl & Wuthrich, 1981).

In contrast, the other slow phase observed in refolding is clearly affected by the amino acid substitution. The relaxation time of this phase, like that of wild-type α subunit, is urea dependent down to final urea concentrations near 1 M. At 1.5 M urea, where both the wild-type and Glu-211 mutant proteins are in their native conformations (Figure 2), the relaxation time for the Glu-211 mutant is approximately twice as long as that of the wild-type protein. For such refolding jumps into the native base-line region, the observed relaxation time is expected to be the reciprocal of the microscopic rate constant for the refolding step linking I_2 and N in the above model. Thus, the rate of conversion of I_2 to N for the Glu-211 mutant is approximately half as fast as that for the wild-type protein.

In the case of unfolding, an order of magnitude increase in the relaxation time for the Glu-211 mutant compared to that for the wild-type protein shows that the amino acid replacement affects this process as well. According to the above model, for jumps ending in the unfolded base-line region, i.e., at and above 6 M urea (Figure 2), the observed relaxation time is the reciprocal of the microscopic rate constant for the unfolding of N to I_2. Thus, both the unfolding and refolding rate constants for the step linking I_2 to N are decreased for the Glu-211 mutant. That the decrease for the unfolding rate constant is substantially greater than that for refolding, approximately a factor of 5, is consistent with the shift of the early stages of the equilibrium unfolding transition for the Glu-211 mutant to higher urea concentrations (Figure 2). The coincidence of the equilibrium transition curves in the latter stages of unfolding, where the intermediates are being converted to unfolded forms for the wild-type and Glu-211 α subunits (Figure 2), suggests that the rate constants for these steps are unaffected by the replacement. Unfortunately, light scattering effects prevent direct detection of these faster phases.

The results for the Glu-211 protein also clarify the behavior of the intermediate phase in refolding. Two different lines of evidence suggest that the nature of the rate-limiting step for this phase changes at ~1 M urea. First, the dependence of the relaxation time on environmental parameters such as urea concentration and pH changes at this urea concentration. Second, although there is a significant difference in the absolute magnitudes of the relaxation times for wild-type and Glu-211 α-subunit proteins above 1 M urea, the values are identical below it.

One way to interpret these observations is to include an additional kinetic species in the transition from I_2 to N: $I_2 \rightarrow I_3 \rightarrow N$. One of the steps involves proline isomerization and the other protein folding; folding is rate limiting above 1 M urea, and proline isomerization is limiting below it. Other more complex mechanisms are also possible. Although we cannot conclusively assign proline isomerization and protein folding to these two steps, certain reasonable assumptions lead to a tentative assignment. We assume that the proline isomerization does not alter exposure of Tyr residues to solvent and that all of the changes in the extinction coefficient occur in the protein folding step. Also, it is assumed that for proline isomerization the rate is independent of the final urea concentration and that for protein folding the logarithm of the relaxation time decreases linearly as the final urea concentration decreases (Creighton, 1979). Given these assumptions, the two possibilities are as follows. For case A

$$I_2 \xrightarrow[\text{isomerization}]{\text{proline}} I_3 \xrightarrow[\text{folding}]{\text{protein}} N$$

and $I_2(\epsilon_{286}) = I_3(\epsilon_{286}) < N(\epsilon_{286})$.

If protein folding is rate limiting at 3 M urea, then this situation would give the observed urea dependence for the relaxation time of this kinetic phase; i.e., the relaxation time would decrease logarithmically with urea down to ~1 M urea where proline isomerization becomes rate limiting. Below this urea concentration, it would become urea independent. For case B

$$I_2 \xrightarrow[\text{folding}]{\text{protein}} I_3 \xrightarrow[\text{isomerization}]{\text{proline}} N$$

and $I_2(\epsilon_{286}) < I_3(\epsilon_{286}) = N(\epsilon_{286})$.

This situation is not consistent with the observed results. In this case, one would predict a continued logarithmic decrease in the relaxation time below ~1 M urea when the protein folding step becomes faster than proline isomerization.

Thus, if this mechanism is correct, we tentatively conclude that proline isomerization precedes protein folding in the conversion of I_2 to N. Obviously, the relaxation of the above assumptions could change the assignment. The observed activation energy at 0.5 M urea, 18 kcal mol^{-1}, suggests that the proline involved plays an essential role in limiting folding. In terms of the nomenclature of Jullien & Baldwin (1981), this is a type III proline which interferes strongly with folding.

It was previously noted that the simple proline isomerization model where the entire protein is considered as a single folding domain does not explain the observed results (Crisanti & Matthews, 1981). Since the α subunit can be cleaved by limited tryptic digestion into two stable domains (Higgins et al., 1979), it is reasonable to ask whether the extension of the simple proline isomerization model to a two-domain system accurately describes the results. The amino domain, residues 1–188, contains 13 proline residues, and the carboxyl domain, residues 189–268, contains 6 proline residues. It is assumed that the distribution of the trans:cis isomers of each X–Pro

peptide bond in the unfolded protein is 80:20 (Brandts et al., 1975) and that all X–Pro peptide bonds are in the trans form in the native conformation. Given these assumptions, one predicts that the folding of 1.4% of the population would not be limited by proline isomerization and thus should be rapid. Also, 4.1% of the population would have all of the prolines in the amino domain in the trans form and at least one in the carboxyl domain in the cis form. The relaxation time for the folding of this population would be predicted to be in the range of 20–85 s from the computer simulations of Creighton (1978). Lastly, 94.5% of the population would be expected to be limited by proline isomerization in the amino domain and have a relaxation time in the range of 40–850 s. Although the observed relaxation times for the two slow phases in refolding do indeed fall into the predicted ranges, the amplitudes do not. In a previous study (Matthews & Crisanti, 1981), the relative amplitudes of the fast, intermediate, and slow phases for refolding jumps to 1.5 M urea were found to be 46%, 34%, and 20%, respectively. Thus, we conclude that the extension of the simple proline isomerization model to a two-domain system does not describe the folding of the α subunit.

Structural Model for Folding of the α Subunit and the Role of Proline Isomerization. It is instructive to postulate a structural model for the folding of the α subunit which attempts to describe the conformations of the intermediates and the molecular nature of the kinetic phases that link these forms. On the basis of our results, proline isomerization must play an important role. Although neither the simple proline isomerization model nor its extension to a two-domain system adequately describes the results for the slow steps in the folding of the α subunit, recent studies on RNase A and bovine pancreatic trypsin inhibitor (BPTI) provide a possible explanation for the role of proline isomerization. Cook et al. (1979) have found that RNase A is capable of folding to a nativelike conformation with at least one proline residue in the incorrect isomeric form. On the basis of this result, and others on RNase A (Henkens et al., 1980) and BPTI (Jullien & Baldwin, 1981), it has been postulated that some prolines serve to limit the folding of a given protein and thus have an "essential" role in this process while others do not act in such a fashion and are "nonessential" (Schmid & Baldwin, 1978; Jullien & Baldwin, 1981).

Assuming that this hypothesis is correct, it is possible to formulate a structural model which is consistent with the optical and hydrodynamic results and which incorporates the concept of proline isomerization in limiting the folding of the α subunit. The unfolded forms of the protein, U_1 and U_2, are assumed to be structureless polypeptides that differ from each other kinetically in the isomeric form assumed by one or a few X–Pro peptide bonds in the amino-terminal domain. Refolding jumps into the native base-line region lead to the rapid collapse of U_1 and U_2 to the two intermediates I_1 and I_2, respectively.

Intermediate I_1 is more compact than U_1 and U_2 but has not adopted the native fold and has all of the seven tyrosine residues in the α subunit exposed to solvent. Intermediate I_2 has a folded amino domain and a compact but not yet folded carboxyl domain. The only tyrosine residues that are excluded from solvent are those in the amino domain that are buried in the isolated amino domain fragment (Higgins et al., 1979). The conversion of I_1 to I_2 is rate limited by the isomerization of one or a few key X–Pro peptide bonds in the amino domain; the remaining X–Pro peptide bonds in this domain are not essential. The supposition that only one or a few X–Pro peptide bonds limit the folding is based upon the observation that approximately 80% of the population of unfolded molecules is converted to I_2 and 20% to I_1 (Crisanti & Matthews, 1981). Once the amino domain has formed, the carboxyl domain can then begin to fold. The first step in this process is proline isomerization either in the carboxyl domain or perhaps in the amino domain at the interface between the two domains; the conversion of I_2 to I_3 reflects this isomerization. The conversion of I_3 to N accounts for the folding of the carboxyl domain and formation of the native conformation. The model is shown pictorially in Figure 5. Note that this model is only strictly correct for refolding; for unfolding jumps to high urea concentrations, the protein can unfold without the necessity of first executing the proline isomerization step linking I_2 and I_3.

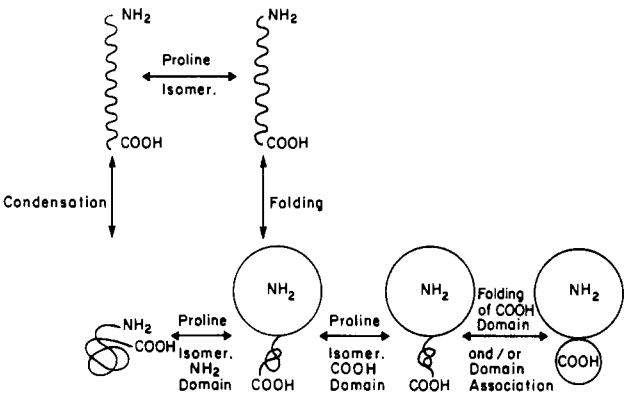


FIGURE 5: Schematic drawing of the hypothetical structural folding model of the α subunit.

The evidence supporting this model is derived from both hydrodynamic and spectroscopic properties of the intermediates. An intermediate such as I_1, with a compact but not yet folded conformation, might be expected to migrate at an intermediate position to the native and unfolded forms on urea gradient gel electrophoresis. An intermediate such as I_2 with a folded amino domain and a compact carboxyl domain might be expected to migrate at nearly the same position as the native protein at low urea concentration since the amino domain comprises 70% of the protein. These predictions are in accord with the experimental observations (Matthews & Crisanti, 1981).

From the optical spectrum of the native protein and the two isolated domains, Higgins et al. (1979) estimated that the single tyrosine in the isolated carboxyl domain is exposed to solvent and that two to three of the six tyrosine residues in the isolated amino domain are exposed to solvent. They also estimated that two to three tyrosines are exposed to solvent in the native conformation of the intact protein. Thus, association of the two domains results in the removal of one to two tyrosines from exposure to solvent. This value is in reasonable agreement with that found for the conversion of I_2 to N, 2.4 tyrosine residues (Crisanti & Matthews, 1981). That the association of the two domains should exclude Tyr residues from solvent is supported by the results of studies of second-site reversions (Yanofsky, 1967), where it was concluded that Tyr-175, in the amino domain, was in close proximity to Gly-211, in the carboxyl domain. Thus, this tyrosine may be buried when the domains associate. It is also possible that Tyr-173 is similarly affected. Also, consistent with this structural model for folding is the previous estimate (Crisanti & Matthews, 1981) that the conversion of I_2 to I_1 exposes 3.2 tyrosines to solvent. This value is in good agreement with that expected for the complete exposure of the three to four buried tyrosine residues in the unfolding of the amino domain. Thus, the results of both hydrodynamic and optical experiments are

consistent with this structural model. Obviously, further studies are required to determine the validity of this model and the structure of the intermediates; such experiments are in progress.

It is not possible to interpret the differences in the equilibrium and kinetic results between wild-type and Glu-211 proteins in terms of specific molecular interactions since an X-ray structure is not available. However, it is possible to test the consistency of these results with the structural model for folding proposed above. Second-site reversion studies place the amino acid at position 211 at the interface between the two domains (Yanofsky, 1967). If the amino acid replacement at position 211 has any effect on unfolding and refolding, one would predict that the transition from the native to the stable intermediates would be affected but that the transition from the intermediates to the unfolded forms would not. That this is indeed the case can be seen in the equilibrium unfolding curves shown in Figure 2. Thus, the folding model is consistent with the limited structural data available.

The folding of other globular proteins including penicillinase (Carrey & Pain, 1978) and the β_2 subunit of tryptophan synthase (Hogberg-Raibaud & Goldberg, 1977; Zetina & Goldberg, 1982) has also been postulated to proceed by the independent folding of subdomains.

Other Folding Studies on the α Subunit. Yutani et al. (1980) have examined the guanidine hydrochloride induced and acid-induced unfolding transitions of the wild-type α subunit of tryptophan synthase and two missense mutants at position 49. The guanidine hydrochloride induced unfolding transition was monitored by the mean residue ellipticity at 222 nm and fit to a three-state model involving one stable intermediate. On the basis of the CD data, it was postulated that the intermediate has a folded amino domain and an unfolded carboxyl domain. More recent studies by Miles et al. (1982) on the relative stabilities of the isolated amino and carboxyl domains to guanidine hydrochloride induced unfolding provide further support for this hypothesis. It was found that the amino domain is indeed more stable than the carboxyl domain.

The fact that a three-state model is sufficient to describe the equilibrium unfolding transition obtained either by CD using guanidine hydrochloride (Yutani et al., 1980) or by difference UV spectroscopy using urea (Matthews & Crisanti, 1981) may be explained by the optical properties of the intermediates I_1, I_2, and I_3. As noted above, I_1 is identical with the unfolded forms U_1 and U_2 in terms of tyrosine exposure to solvent. Thus, this species would contribute to the observed difference UV spectrum as if it were part of a manifold of unfolded states. Intermediates I_2 and I_3 are similar to each other in their exposure of tyrosines to solvent and thus appear optically as identical species. The equilibrium unfolding data would then be expected to be adequately described by a three-state model involving three optically different classes: (1) the native form; (2) the unfolded forms and I_1; and (3) the intermediates I_2 and I_3. If the secondary structure of intermediate I_1 is disrupted, then one would expect that the mean residue ellipticity data would also be well described by a three-state model. Kinetic studies of the guanidine hydrochloride induced unfolding transition are required to determine conclusively if the folding model proposed above for urea-induced unfolding also applies to the guanidine-induced unfolding.

Obviously, further experiments are required to determine the validity of the folding model proposed for the α subunit. However, such efforts are clearly justified, since the approach of studying the effects of single amino acid substitutions on the kinetics of protein folding has been extremely useful in defining a folding model for the α subunit and in providing clues to an understanding of the molecular events associated with observed kinetic phases.

Acknowledgments

We are grateful to Dr. Charles Yanofsky, Stanford University, for samples of the wild-type and Glu-211 mutant α-subunit strains.

Registry No. Tryptophan synthase, 9014-52-2; urea, 57-13-6; L-glutamic acid, 56-86-0; L-proline, 147-85-3.

References

Adachi, O., Kohn, L. D., & Miles, E. W. (1974) *J. Biol. Chem. 249*, 7756.
Anfinsen, C. B. (1973) *Science (Washington, D.C.) 181*, 223.
Anfinsen, C. B., & Scheraga, H. A. (1975) *Adv. Protein Chem. 29*, 205.
Baldwin, R. L. (1975) *Annu. Rev. Biochem. 44*, 453.
Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953.
Carrey, E. A., & Pain, R. H. (1978) *Biochim. Biophys. Acta 533*, 12.
Cook, K. H., Schmid, F. X., & Baldwin, R. L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6157.
Creighton, T. E. (1977) *Prog. Biophys. Mol. Biol. 33*, 231.
Creighton, T. E. (1978) *J. Mol. Biol. 125*, 401.
Creighton, T. E. (1979) *J. Mol. Biol. 129*, 235.
Creighton, T. E., & Yanofsky, C. (1970) *Methods Enzymol. 17*, 365.
Crisanti, M. M., & Matthews, C. R. (1981) *Biochemistry 20*, 2700.
Eigen, M., & DeMaeyer, L. (1963) *Tech. Org. Chem. 8*, 895.
Garel, J.-R., & Baldwin, R. L. (1975) *J. Mol. Biol. 94*, 611.
Grathwohl, C., & Wuthrich, K. (1981) *Biopolymers 20*, 2623.
Henkens, R. W., Gerber, A. D., Cooper, M. R., & Herzog, W. R., Jr. (1980) *J. Biol. Chem. 255*, 7075.
Herskovits, T. T. (1967) *Methods Enzymol. 11*, 748.
Higgins, W., Fairwell, T., & Miles, E. W. (1979) *Biochemistry 18*, 4827.
Hogberg-Raibaud, A., & Goldberg, M. E. (1977) *Biochemistry 16*, 4014.
Jullien, M., & Baldwin, R. L. (1981) *J. Mol. Biol. 145*, 265.
Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459.
Kirschner, K., Wiskocil, R. L., Loehn, M., & Rezeau, L. (1975) *Eur. J. Biochem. 60*, 513.
Li, S.-L., & Yanofsky, C. (1972) *J. Biol. Chem. 247*, 1031.
Matthews, C. R., & Crisanti, M. M. (1981) *Biochemistry 20*, 784.
Matthews, C. R., Crisanti, M. M., Gepner, G. L., Velicelebi, G., & Sturtevant, J. M. (1980) *Biochemistry 19*, 1290.
Miles, E. W., Yutani, K., & Ogasahara, K. (1982) *Biochemistry 21*, 2586.
Murgola, E. J., & Yanofsky, C. (1974) *J. Mol. Biol. 86*, 775.
Murphy, T. M., & Mills, S. E. (1968) *Arch. Biochem. Biophys. 127*, 7.
Nall, B. T., Garel, J.-R., & Baldwin, R. L. (1978) *J. Mol. Biol. 118*, 317.
Nemethy, G., & Scheraga, H. A. (1977) *Q. Rev. Biophys. 10*, 239.
Patterson, C. L., Jr., Hodo, H. G., III, & Hardman, J. K. (1977) *Arch. Biochem. Biophys. 181*, 428.
Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167.
Schmid, F. X., & Baldwin, R. L. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4764.

Yanofsky, C. (1967) *Harvey Lect. 61*, 145.
Yanofsky, C., & Horn, V. (1972) *J. Biol. Chem. 247*, 4494.
Yanofsky, C., Drapeau, G. R., Guest, J. R., & Carlton, B. C. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 296.
Yutani, K., Ogasahara, K., & Sugino, Y. (1980) *J. Mol. Biol. 144*, 455.
Zetina, C. R., & Goldberg, M. E. (1982) *J. Mol. Biol. 157*, 133.

Identification of the Catalytic Subunit of an Oligomeric Casein Kinase (G Type). Affinity Labeling of the Nucleotide Site Using 5′-[*p*-(Fluorosulfonyl)benzoyl]adenosine†

Jean-Jacques Feige, Claude Cochet, Fabienne Pirollet, and Edmond M. Chambaz*

ABSTRACT: Identification of the catalytic subunit of a G type [using guanosine 5′-triphosphate (GTP) as well as adenosine 5′-triphosphate (ATP) as phosphate donor], oligomeric, cyclic nucleotide independent casein kinase purified from bovine lung was carried out after reaction with 5′-[*p*-(fluorosulfonyl)-benzoyl]adenosine (FSBA) and isolation of the subunit components of the enzyme. FSBA exhibited the major characteristics of an affinity label reacting at the nucleotide (ATP, GTP) site of the casein kinase. FSBA acted as a competitive inhibitor of ATP (and GTP), led to complete inactivation of the enzyme in a reaction showing two kinetic steps, and became irreversibly bound to the protein. After being labeled with FSBA, the casein kinase (apparent molecular weight of 140 000) was separated into its two monomeric components of apparent molecular weights 38 000 (α) and 27 000 (β), respectively, after sodium dodecyl sulfate–polyacrylamide gel electrophoresis. Use of radioactive FSBA showed that specific affinity labeling was limited to the α casein kinase subunit. This result was in agreement with the fact that casein kinase activity was found associated with the α monomer after electrophoretic separation of the α and β subunits. It may thus be concluded that the largest (α) subunit contains the catalytic site of the casein kinase G. Electrophoretic analysis of purified protein kinase under denaturing conditions suggested an $\alpha_3\beta_2$ combination for an apparent molecular weight of 130 000–140 000. However, a maximum of 2 mol of FSBA could be specifically bound to the α subunit per mol of enzyme, with a concomitant complete inactivation. These data would be in agreement with an $\alpha_2\beta_2$ subunit composition for casein kinase G, as proposed by other research groups for a similar type of protein kinase of different sources. These observations suggest that the α subunits are functionally similar, each of them containing a nucleotide (ATP, GTP) binding site. The possible role of the β subunit in the enzyme activity remains to be established.

Covalent modification of proteins by phosphorylation–dephosphorylation has emerged as a widely occurring process in the regulation of several important cellular functions (Rubin & Rosen, 1975; Weller, 1979; Cohen, 1981). Various types of ATP (or GTP):protein phosphotransferases (EC 2.7.1.37; protein kinases) have been characterized in mammalian tissues, and criteria for their classification have been proposed (Krebs & Beavo, 1979; Traugh et al., 1974). Whereas messenger dependent protein kinase activities may themselves be regulated by specific intracellular effectors [adenosine cyclic 3′,5′-phosphate (cAMP),[1] cGMP, calcium calmodulin, etc.], protein kinase activities for which no specific effector is yet recognized have been characterized and thus termed messenger independent systems (Krebs & Beavo, 1979). Among this category, cyclic nucleotide independent casein kinases have been classified into two major types: the A type (casein kinase A), using only ATP as phosphate donor, and the G type (casein kinase G), using GTP as well (Cochet et al., 1980). These two types of casein kinases appear similar to the casein kinases I and II (Hathaway & Traugh, 1979) and the casein kinase S (phosphorylating only serine) and TS (phosphorylating both serine and threonine) enzymes (Meggio et al., 1977), respectively. The G-type casein kinase has been purified by several research groups from various tissue sources (Hathaway & Traugh, 1979; Dahmus & Natzle, 1977; Thornburg & Lindell, 1977; Kumon & Ozawa, 1979; Walinder, 1973; Cochet et al., 1981) and usually reported to present an oligomeric structure with an apparent molecular weight between 100 000 and 200 000 for the native enzyme. The G-type casein kinase subunit composition has been reported to be made of two (Walinder, 1973) or three (Hathaway & Traugh, 1979; Dahmus & Natzle, 1977; Kumon & Ozawa, 1979) different moieties, and the smallest subunit is usually recognized as the target of a self-phosphorylation process (Hathaway & Traugh, 1979; Dahmus & Natzle, 1977; Cochet et al., 1981).

However, nothing is known concerning the possible regulation of the cellular G-type casein kinase activity. As a prerequisite, understanding of the molecular organization of the native enzyme and knowledge of the respective roles of the different subunits in the overall catalytic activity may shed some light on the functional properties of this phosphorylation

† From the Laboratoire de Biochimie Endocrinienne, the Institut National de la Santé et de la Recherche Médicale (INSERM), and the Centre National de la Recherche Scientifique (CNRS), Université Scientifique et Médicale de Grenoble, France. *Received August 4, 1982.* This work was supported by the INSERM (U-244, ATP 77.84, and ATP 79.114), the CNRS (ERA 942), and the Fondation pour la Recherche Médicale Française.

* Address correspondence to this author at the Laboratoire de Biochimie Endocrinienne, CERMO, Université Scientifique et Médicale, Grenoble, France.

[1] Abbreviations: FSBA, 5′-[*p*-(fluorosulfonyl)benzoyl]adenosine; ATP, adenosine 5′-triphosphate; GTP, guanosine 5′-triphosphate; Tris, tris(hydroxymethyl)aminomethane; TDG buffer, tris(hydroxymethyl)-aminomethane, dithiothreitol, and glycerol buffer; $NaDodSO_4$, sodium dodecyl sulfate; cAMP, adenosine cyclic 3′,5′-phosphate; cGMP, guanosine cyclic 3′,5′-phosphate.

0006-2960/83/0422-1452$01.50/0