Characterization of an Intermediate in the Folding Pathway of Phosphoglycerate Kinase: Chemical Reactivity of Genetically Introduced Cysteinyl Residues during the Folding Process†

Nathalie Ballery, Michel Desmadril, Philippe Minard, and Jeannine M. Yon*

Laboratoire d'Enzymologie Physicochimique et Moléculaire, Unité de Recherche du Centre National de la Recherche Scientifique, Université de Paris-Sud, 91405 F Orsay, France

Received June 23, 1992; Revised Manuscript Received October 7, 1992

ABSTRACT: The unfolding–refolding kinetics of yeast phosphoglycerate kinase were studied using the chemical reactivity of genetically introduced cysteinyl residues as conformational probes and far-ultraviolet circular dichroism. A unique internal cysteinyl residue was introduced in several mutants at selected positions in the N- and C-domains. The cysteinyl residues were at positions 97 (the unique cysteinyl residue of the wild-type enzyme), 183 in the N-domain, 285 and 324 in the C-domain. A similar strategy has been used to study the unfolding–refolding transition under equilibrium conditions [Ballery et al. (1990) *Protein Eng.* 3, 199–204]. Except for the mutant C97A,A183C, whose cysteinyl residue is located at the domain interface, three labeling phases were observed during the refolding process, indicating the presence of three species, the unfolded, intermediate, and folded proteins. The comparison of the data obtained following the accessibility of the thiol group to 5,5′-dithiobis(2-nitrobenzoate) and ellipticity at 218 nm indicated that all mutants have the same folding pathway and allowed us to characterize the intermediate. In this species, each domain appeared to have a high content of secondary structure but a flexible tertiary structure; this intermediate, which had the characteristics of a molten globule, remained in fluctuating equilibrium with a widely unfolded form. The same folding intermediate was detected for mutant C97A,A183C; however, the cysteinyl residue being totally accessible to the reagent, it is likely that in this intermediate the interdomain interactions are not established. Domain pairing and formation of the native tertiary structure occur simultaneously in the slow phase of refolding. The validity and limitations of the methodology are discussed.

The molecular mechanisms by which a polypeptide chain folds into a definite three-dimensional structure is an essential but not clearly understood aspect of structural biology. For instance, the early events which occur in several parts of the polypeptide chain are not elucidated and the nature of the intermediates on the folding path are unknown. A classical strategy to approach this problem consists of the characterization of stable or transient intermediate states observed during the refolding process of denatured proteins.

In the past few years, two-dimensional nuclear magnetic resonance in conjunction with rapid hydrogen–deuterium exchange methods have been used to describe, with a high spatial and temporal resolution, the folding process of a few proteins (Baldwin & Roder, 1991). A more classical approach uses either spectroscopic probes reflecting the global secondary structure such as far-ultraviolet circular dichroism or local probes of the tertiary structure such as the fluorescence of a given tryptophan residue. This latter approach, previously limited by the natural distribution of these residues within the structure, can now be extended to proteins in which such probes (e.g., tryptophan) have been inserted by site-directed mutagenesis (Smith et al., 1991).

We have previously introduced a similar strategy (Ballery et al., 1990). The chemical reactivity, toward a thiol reagent, of a protein containing a unique cysteine is very dependent on the accessibility of the thiol group. The introduction, by site-directed mutagenesis, of a unique internal cysteinyl residue in a selected position can therefore provide a local probe of the conformational variations around this residue.

We have used this approach to study the folding process of yeast phosphoglycerate kinase. This monomeric protein is made up of a single polypeptide chain folded into two domains of approximately equal size (Watson et al., 1982). Phosphoglycerate kinase has been chosen as a convenient model for folding studies and more specifically to investigate the consequences of its two-domain structure on the folding process (Yon et al., 1988, 1990). A first mutant in which the single wild-type cysteine is replaced by an alanine (C97A) was constructed. Several other mutants, each possessing a unique cysteine buried in the native molecule, have been constructed from the original mutant. The positions have been selected to be buried, not highly conserved, and distributed all over the structure, i.e., 183 (at the end of βF) in the N-domain and 285 (on βJ) and 324 (on helix X) in the C-domain. Previous studies have shown that the introduction of a buried cysteine in these positions did not cause any perturbation in the structural and functional properties of the molecule. The unfolding–refolding transition of these yeast phosphoglycerate kinase mutants have been previously studied under equilibrium conditions (Ballery et al., 1990).

To further elucidate the rate of folding in different regions of the molecule, we present in this paper a kinetic study of the unfolding and refolding of wild-type PGK and these mutants, using both circular dichroism and the reactivity of the thiol group as conformational probes. This method shows that phosphoglycerate kinase refolds through a metastable intermediate and gives some information about the properties of this intermediate. The possibilities, requirements, and limitations of the methodology are discussed.

† This work was supported by the Centre National de la Recherche Scientifique, by a contract with the Centre National de Transfusion Sanguine who provides circular dichroism facilities, and by a grant from the Fondation pour la Recherche Médicale Française.

* To whom correspondence should be addressed.

0006-2960/93/0432-708$04.00/0

MATERIALS AND METHODS

Preparation of the Different Mutants. Different mutations were introduced into the yeast PGK[1] gene by site-directed mutagenesis according to Minard et al. (1990a). The following mutants were obtained from a first mutant (C97A) whose cysteine was replaced by an alanine: C97A,A183C, C97A,I285C, C97A,T324C. The mutant proteins were prepared as already described by Minard et al. (1990b). All the following experiments were done in a 50 mM Tris buffer, pH 7, containing 0.5 mM EDTA at 20 °C, unless otherwise indicated.

Purified proteins were precipitated and stored in ammonium sulfate. Before use, each sample was centrifuged, and the protein pellet was dissolved in a 50 mM Tris buffer, pH 8, containing 25 mM DTT. The fully reduced protein was then transferred to a suitable buffer by desalting on a Sephadex G25 column and used within 24 h.

Kinetics of Unfolding and Refolding Followed by the Reactivity of the Thiol Group. Kinetics of unfolding were monitored by recording the reaction of the thiol group with NbS_2. The protein was added at time zero in a Tris buffer containing 1 M GdnHCl and 0.25 mM NbS_2. The final concentration of protein was 3 μM.

For the kinetics of refolding, the protein was previously incubated overnight in a Tris buffer containing 6 M GdnHCl. Then, aliquots were refolded by dilution in a regenerating buffer (without GdnHCl), the final concentration of the denaturant being 0.12 M. At different times, the thiol group was titrated by the addition of NbS_2 at a final concentration of 0.25 mM. The reaction rate of the protein with NbS_2 was monitored by variations in absorbance at 412 nm. The measured absorbances were corrected taking into account the basal absorbance level due to NbS_2. The concentration of thiol groups was calculated using ϵ_{412nm} = 14 150 M^{-1} cm^{-1} for NbS^-.

Kinetics of Unfolding and Refolding Followed by Circular Dichroism. The unfolding reaction was initiated by rapid mixing of the native wild-type or mutant protein in a solution containing the denaturant and the variation in ellipticity at 218 nm was recorded as a function of time. The final concentration of denaturant was varied between 0.5 and 1.3 M GdnHCl. For the kinetics of refolding, the protein was previously denatured overnight in 6 M GdnHCl and then mixed in a regeneration buffer containing a final concentration of denaturant varying between 0.12 and 1 M GdnHCl. The final concentration of protein was 2 μM for both kinds of experiments.

Data Analysis. The kinetics of unfolding were analyzed according to one exponential term as follows:

$$\frac{y(t) - y(eq)}{y(N) - y(U)} = Ae^{-\lambda t} \quad (1)$$

$y(t)$ being the value of the observable at time t, $y(N)$ and $y(U)$ this value for the native and unfolded protein, respectively, and $y(eq)$ this value when the equilibrium is reached. A is the amplitude of the exponential and λ is the macroscopic rate constant.

[1] Abbreviations: ANS, anilinonaphthalenesulfonate; EDTA, ethylenediaminetetraacetic acid; GdnHCl, guanidine hydrochloride; NbS_2, 5,5′-dithiobis(2-nitrobenzoate); PGK, phosphoglycerate kinase (EC 2.7.2.3).

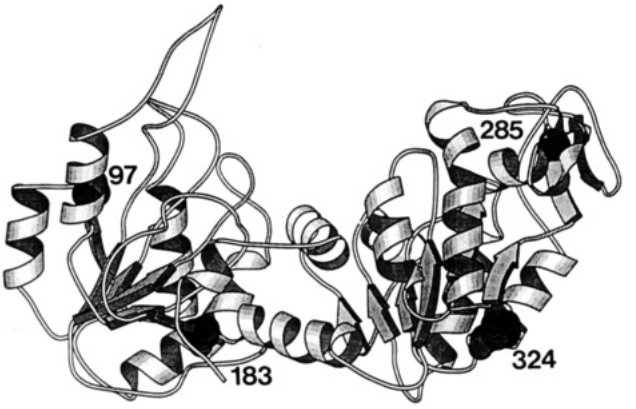

FIGURE 1: Three-dimensional structure of yeast PGK showing the positions of the different cysteinyl residues in the wild-type and mutant proteins.

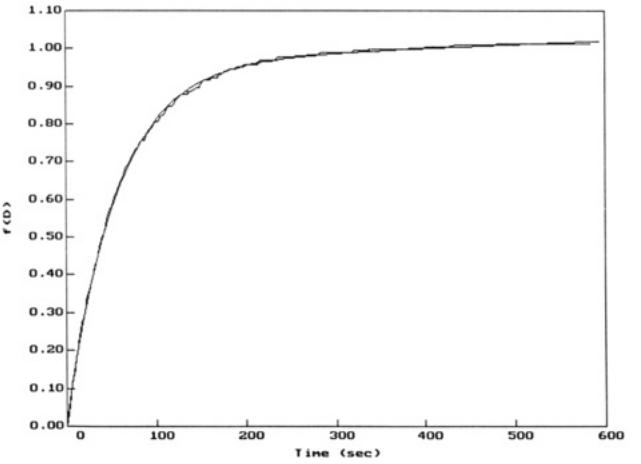

FIGURE 2: Kinetics of unfolding in 1 M GdnHCl as assessed by the titration of the thiol group (C97) by NbS2 at pH 7.0 and 20 °C for a 3 μM final concentration of wild-type PGK.

The kinetics of refolding were analyzed as the sum of two exponential terms according to the following:

$$\frac{y(t) - y(eq)}{y(N) - y(U)} = A_1e^{-\lambda_1 t} + A_2e^{-\lambda_2 t} \quad (2)$$

The experimental curves were fitted according to eq 1 or 2 using a nonlinear regression procedure.

RESULTS

Properties of the Mutant Enzymes. The four thiol groups used as local probes in this work are located at positions 97 (cysteinyl residue of the wild-type enzyme) and 183 in the N-domain and 285 and 324 in the C-domain (Figure 1).

As previously reported (Ballery et al., 1990), all the constructed mutants had the same conformational properties as the wild-type protein as assessed by circular dichroism. The enzyme specific activity of the mutants was also found to be identical to that of the wild-type enzyme. Furthermore, it has been shown that the unfolding–refolding transition induced by GdnHCl is totally reversible for these mutants.

Unfolding–Refolding of Yeast Phosphoglycerate Kinase and Mutants as Assessed by the Reactivity of the Thiol Groups. Kinetics of unfolding were recorded in 1 M GdnHCl final concentration by titration of the thiol group by NbS_2. The kinetics were analyzed according to a single exponential term as shown in Figure 2 for the wild-type protein, the amplitude at the end of the reaction corresponding to the titration of one thiol group per molecule. The macroscopic rate constant was evaluated to be 0.017 ± 0.002 s^{-1}. This

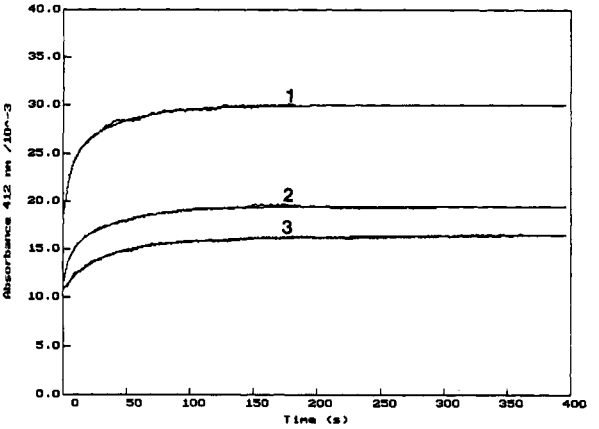

FIGURE 3: Kinetics of labeling of the thiol group (C97) by NbS2 after 1 min (1), 2 min (2), and 3 min (3) of refolding of wild-type PGK in 0.12 M GdnHCl, pH 7.0, 20 °C, and 3 μM final protein concentration.

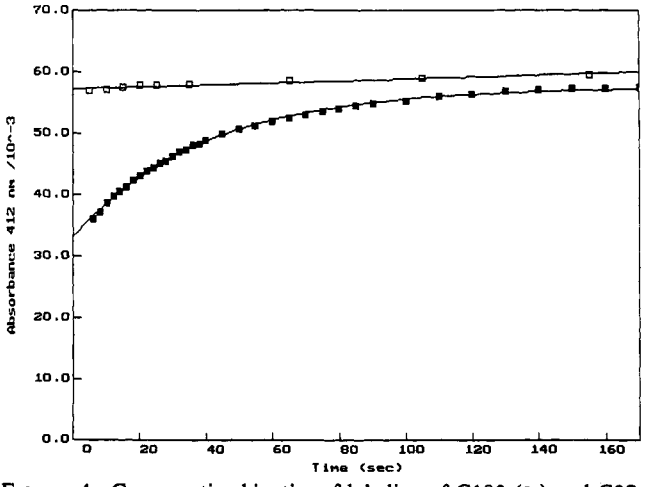

FIGURE 4: Comparative kinetics of labeling of C183 (□) and C97 (■) after 30 s of renaturation. Experimental conditions: protein final concentration 5 μM, 0.2 M GdnHCl, and 0.25 mM NbS2. The curves were fitted according to a single exponential term.

value is similar to the unfolding rate constant measured by circular dichroism under identical conditions. The same behavior was observed for wild-type PGK and all mutants.

For the refolding process, typical NbS_2 titration recordings after different times of refolding are shown in Figure 3. The kinetics of labeling by NbS_2 after various times of refolding exhibited three phases widely separated on the time scale. A first phase represented by an initial burst corresponded to the labeling of a completely accessible cysteinyl residue, this phase being a second-order reaction. When the protein was renatured in a buffer containing NbS_2, the kinetics of labeling included only this phase followed by a plateau with the amplitude expected for the complete titration of one thiol group per molecule. When the labeling by NbS_2 was initiated after different refolding times, a second first-order phase with a rate constant of 0.025 ± 0.001 s^{-1} was observed for all proteins with the exception of mutant C97A,A183C (Figure 4). It should be noted that this phase was observed for cysteines located in the N- (C97) and in the C-domain (C285 and C324). A third very slow first-order phase appeared with an amplitude increasing with the refolding time, the rate constants being equal for wild-type PGK and each mutant to those reported for the native proteins (Ballery et al., 1990). This result contrasts with the reactivity of these thiol groups under equilibrium conditions which did not show the intermediate labeling phase described in the present work. This fact and the absence of the intermediate labeling phase for mutant C97A,A183C (Figure 4) clearly show that it is not a mixing artifact but actually reflect an intrinsic property of the molecule population during the folding process.

These results indicate the presence of at least three species during the refolding process, each being characterized by a different reactivity of the thiol group: the unfolded species with a totally accessible cysteine, an intermediate form, and the native protein. According to this criterion, only two species were detected for mutant C97A,A183C.

The fraction of protein with a fast-reacting thiol obtained from the amplitude of the initial burst at different refolding times allowed us to follow the kinetics of disappearance of the "unfolded form", unfolded according to the criterion of thiol accessibility. These kinetics are shown in Figure 5a–d for the wild-type and mutant enzymes. In the inset, the semilog plot clearly indicates that the disappearance of the species with a fast-reacting thiol is biphasic for all proteins expected for mutant C97A,A183C; in this case, only the slow phase of disappearance of the unfolded form was detected. The macroscopic rate constants of the two phases are reported in Table I. For all proteins with biphasic refolding kinetics, the amplitude of the fast phase represents about 30% of the total amplitude. The slow refolding phase was determined at three final GdnHCl concentrations and, as expected, the rate constant increased with decreasing denaturant concentrations.

The kinetics of formation and transformation of the intermediate species from wild-type PGK were also deduced from the amplitude of the different phases of labeling at a given time of refolding. The fraction of the intermediate at different times was deduced from the amplitude of the intermediate labeling phase. The fraction of native protein was obtained by the difference between the value corresponding to a complete titration of one thiol group and the value of the apparent plateau of the labeling kinetics at a given time of refolding. Figure 6 shows the kinetics of disappearance of the unfolded form, variations of the intermediate, and formation of the native wild-type protein. This procedure gives a correct value for the concentration of the unfolded form, which was measured in a very short time. The competition between folding and labeling of the intermediate did not affect the determination of the folding macroscopic rate constant. However, the concentration of the native form was overestimated since the labeling constant (0.025 ± 0.001 s^{-1}) of the intermediate was only three times higher than the rate of transformation of the intermediate into the native form (Table I). Similarly, the intermediate population was underestimated by about 7.5%. Thus, the kinetics suggest that the intermediate species was formed from the unfolded form within a few seconds.

Kinetics of Unfolding and Refolding of Wild-Type Phosphoglycerate Kinase and Mutants As Assessed by Circular Dichroism. Kinetics of unfolding for PGK and mutants were recorded by circular dichroism at 218 nm in a range of final concentrations of denaturant from 0.5 to 1.3 M. All were fitted by a single exponential term.

Kinetics of refolding of proteins previously denaturated in 6 M GdnHCl were monitored in a GdnHCl concentration range varying from 0.1 to 1 M. All were biphasic; the very fast phase was completed in the dead time of the experiment (Figure 7). The variations of the macroscopic rate constants of unfolding and those of the slow phase of refolding with GdnHCl were found to be identical for all mutants (Figure 8) including the mutant devoid of cysteine (C97A). Compared with the wild-type enzyme, the variations of the macroscopic

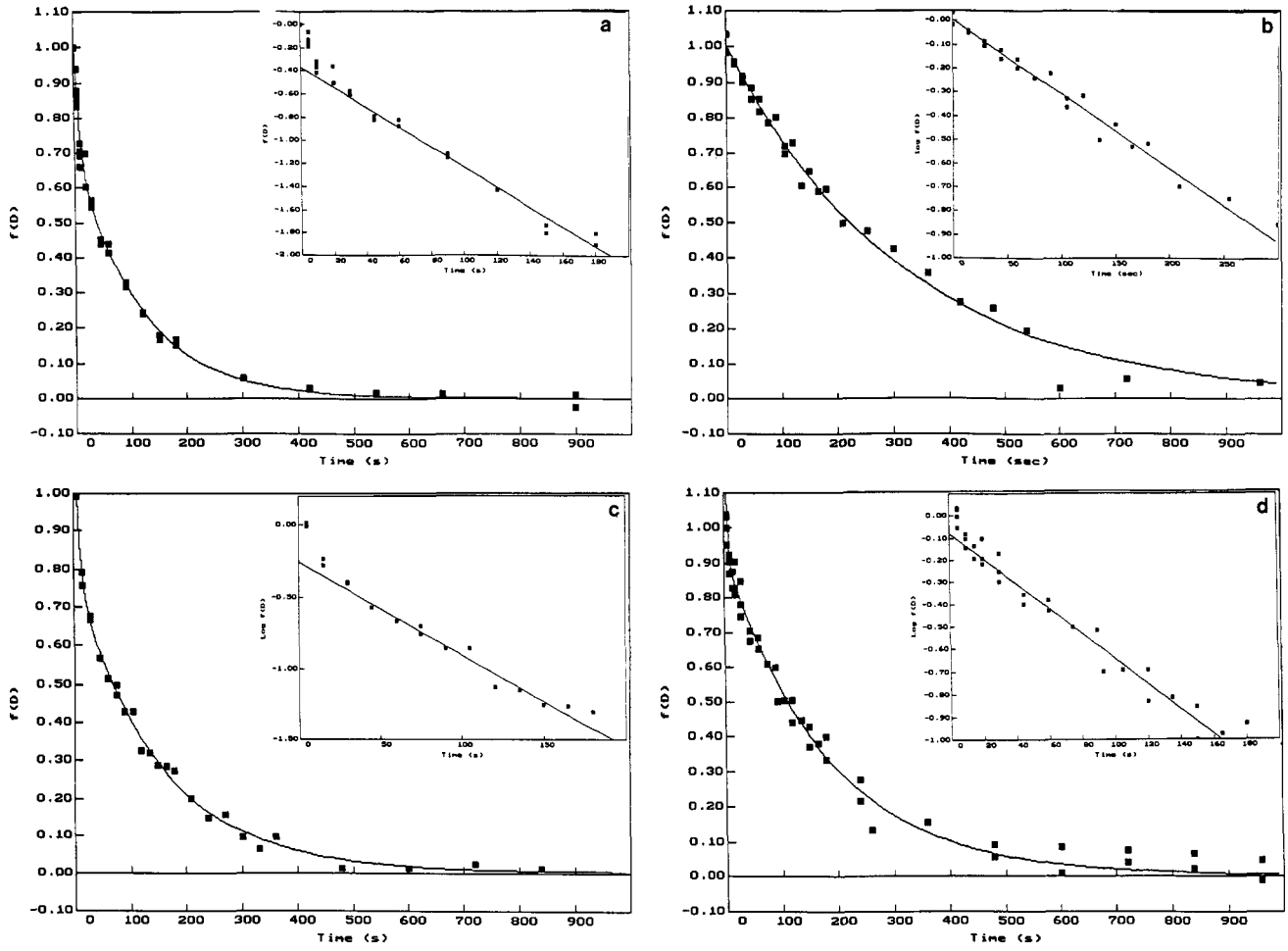

FIGURE 5: Kinetics of refolding as assessed by the titration of the thiol group by NbS2, for (a) wild-type PGK, (b) mutant C97A,A183C, (c) mutant C97A,I285C, (d) mutant C97A,T324C. Experimental conditions: 0.12 M final denaturant concentration, pH 7.0, 20 °C, and 3 μM final protein concentration. In the inset, the log plots versus time indicate the biphasicity of the kinetics at the exception of those for mutant C97A,A183C.

Table I: Rate Constants of the Different Refolding Steps as Assessed (a) by the Variations in Accessibility of the Thiol Group and (b) by Circular Dichroism in 0.12 M GdnHCl, pH 7.0, 20 °C, 3 μM Protein Final Concentration

protein	rate constant of the fast step (a) (s^{-1})	rate constant of the slow step (a) ($\times 10^{-3}$ s^{-1})	rate constant of the slow step (b) ($\times 10^{-3}$ s^{-1})
wild-type PGK	>0.15	7.9	4.7
C97A,A183C		3.0	5.5
C97A,I285C	>0.10	6.5	6.0
C97A,T324C	>0.10	6.8	6.0

rate constant, λ, with GdnHCl concentration are shifted toward lower concentrations of denaturant.

The profiles of λ versus GdnHCl concentration being similar, it is likely that the mutants and wild-type PGK follow the same folding pathway. The shift in the λ profiles reflected the lower c_m values observed in equilibrium studies for all mutants compared with the wild-type enzyme.

The kinetics of refolding obtained from the variations of the two observables, i.e., the reactivity of the thiol group and the ellipticity at 218 nm, were found to be biphasic with the same relative amplitudes between the fast and the slow phases, with the exception of mutant C97A,A183C, whose thiol group became buried only during the slow phase. Furthermore, the variations of the macroscopic rate constant, λ, with GdnHCl were found to be identical for the two observables. This suggests a minimum scheme of folding including three species,

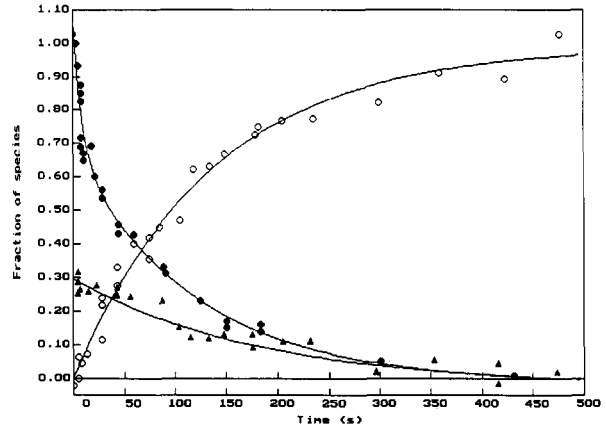

FIGURE 6: Progress curves corresponding to the unfolded species (●), the native protein (○), and the intermediate (▲), as deduced from the kinetics of labeling of the cysteinyl residue from wild-type PGK.

the unfolded protein U, an intermediate X, and the folded protein N:

$$U \Leftrightarrow X \Leftrightarrow N$$

Characterization of the Intermediate. An alternative interpretation of multiphasic refolding kinetics is the possibility of parallel refolding pathways of different unfolded species. Indeed, fast and slow refolding species produced by proline isomerization are, for several proteins, responsible for the complexity of the folding process (Brandts et al., 1975; Kim & Baldwin, 1990). A first argument indicating that this slow

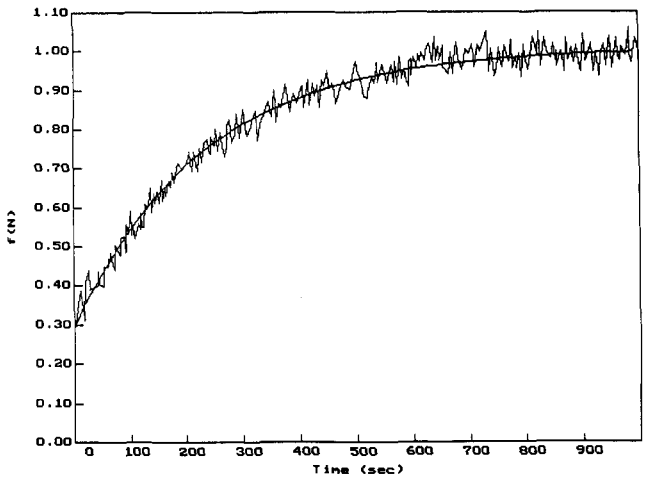

FIGURE 7: Kinetics of refolding of wild-type PGK as assessed by circular dichroism in 0.12 M final concentration of denaturant, pH 7.0, 20 °C, and 3 μM final protein concentration.

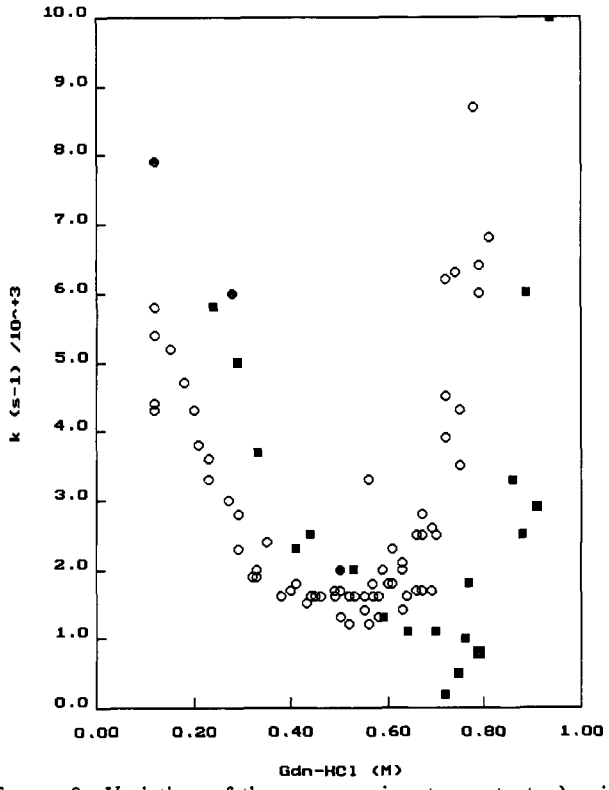

FIGURE 8: Variations of the macroscopic rate constants, λ, with GdnHCl concentration for the unfolding and refolding as assessed by circular dichroism at 218 nm for the following PGK and mutants: (■) wild-type enzyme; (○) all mutants, i.e., C97A, C97A,A183C, C97A,I285C, and C97A,T324C mutants; (●) λ values of the slow refolding phase evaluated from the accessibility of the thiol group in wild-type PGK.

step was not reflecting proline isomerization was that the macroscopic rate constant varies with GdnHCl concentration. Furthermore, a double-jump experiment was carried out. Refolding was monitored using the reactivity of the cysteinyl residue as a conformational probe, the protein being previously incubated only 30 s at 4 °C in 6 M GdnHCl, which is a time too short to allow a complete isomerization of the prolines (Brandts & Lin, 1986). Under these conditions, the same kinetic behavior as that observed after denaturing the protein overnight was observed. The kinetics of disappearance of the fast reacting thiol were found to be biphasic. The amplitudes of the two refolding phases were found to be identical whatever the previous incubation time in the denaturant, allowing us

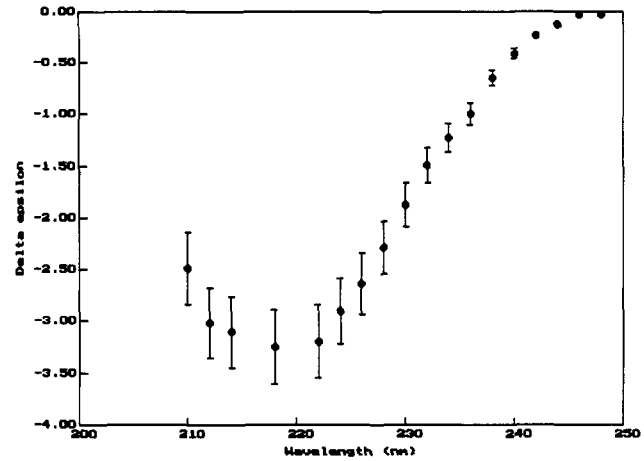

FIGURE 9: Circular dichroism spectrum of the intermediate (see text). Error bars were obtained from the standard deviations of multiple measurements.

to discard proline isomerization as responsible for the slow phase.

The intermediate was characterized from the kinetic studies monitored by both circular dichroism and reactivity of the thiol group. Under the same conditions of complete refolding (0.12 M GdnHCl), the amplitude of the fast refolding phase for both signals, ellipticity at 218 nm, and thiol titration were found to be quite identical and close to 30% of the total amplitude. The stoichiometry of the labeling suggested that an intermediate species represented about a third of the total population. Taking into account the large difference in reactivity of the cysteinyl residue in the different species, it was deduced that the amount of native protein at the end of the fast phase was negligible. Since the thiol group located at different positions in the molecule reacted as in the unfolded PGK, it was reasonable to consider that the fast-reacting species are devoid of structure. Therefore, the amplitude of the fast phase observed by circular dichroism could result from a linear combination of a 1:2 intermediate to denatured species ratio. The amplitude of the fast phase was determined at different wavelengths. Taking into account the relative amount of the intermediate, a putative circular dichroism spectrum of the intermediate was obtained (Figure 9). This spectrum, characteristic of a high content in helices, was found to be about the same as that of native PGK. However, if the species reacting with NbS_2 as the unfolded protein had a partial content of secondary structure, $\Delta\epsilon$ values of the intermediate were overestimated.

DISCUSSION

The main goal of the experiments described above was to refine our description of the folding process of PGK, a classical example of a monomeric two-domain protein. The kinetics of the refolding process were monitored using, as local probes, the reactivity of a unique cysteinyl residue distributed all over the structure in different mutants and, as a global probe, far-ultraviolet circular dichroism. A previous study under equilibrium conditions (Ballery et al., 1990) has shown that all the constructed mutants having a unique internal cysteine behave identically.

In the three mutants and wild-type PGK, the reactivity of the unique thiol group was increased by several orders of magnitude upon denaturation of the protein. Thus, at a given time of refolding, it was possible to titrate the fractions of folded and unfolded molecules. An important point is that partially folded molecules with an exposed thiol group or with a thiol group buried within rapidly interconverting structures,

might appear as "unfolded" according to this unique criterion. In the thiol titration, three labeling phases were observed for all mutants except C97A,A183C for which only two labeling phases were detected. In a rapid folding phase corresponding to less than 15 s, only 30% of the dichroism signal was recovered. The four thiol groups were then reacting as in the unfolded protein for about 70% of the molecule population (100% in C97A,A183C mutant). In these species, the thiol group reacted as in the denatured protein, whatever its position, suggesting that, in this state, the molecule had a low degree of structure. At the end of the fast refolding phase, the concentration of the thiol group reacting as in the native protein was negligible, suggesting that the folded form only appeared in the last folding phase.

An intermediate labeling phase which was not foreseen using such a methodology, was detected revealing the existence of a third state of the protein distinct from the native and unfolded species. This intermediate was observed in all PGKs except the C97A,A183C mutant. In these three PGKs, for short renaturation times, a third of the molecules reacts with NbS_2 with a rate constant very different from, and between those of, both the native and denatured species. In this intermediate species, positions 97 in the N-domain and 285 and 324 in the C-domain were not freely accessible to the reagent. The rate constant of the "intermediate labeling phase" being independent on NbS_2 concentration, it is likely that the rate-determining step of the reaction with NbS_2 in this state is the partial or total unfolding of the protein structure. This suggests that, in this intermediate form, the thiol group is buried but the protein is much more flexible that in the native state.

Properties of the Intermediate. The above description of the observed intermediate is obviously reminiscent of the molten globule model. This compact but fluctuating state has been proposed by Ptitsyn et al. (1990) to be a kinetic intermediate in the refolding of several proteins, including yeast PGK, and is suggested to be quite general in protein folding. A property attributed to the molten globule is its ability to bind nonspecifically a fluorescent hydrophobic probe, anilinonaphthalene sulfonate (ANS). The rate of formation and disappearance of the molten globule can be measured by the variation of ANS fluorescence (Ptitsyn et al., 1990). In order to compare quantitatively the rate constant followed by several methods, we have reproduced the experiments described by Ptitsyn et al. (1990) in the refolding conditions used in the present study. Results similar to those described by Ptitsyn et al. (1990) were obtained showing that the rate constant of ANS expulsion is very close to that of the slow refolding phase as assessed by thiol titration experiments (data not shown). Therefore, it seems possible that the ANS binding species, "the molten globule", and the species with a thiol group buried in a flexible environment represent the same intermediate since these two signals show similar half-lives under identical conditions.

In this intermediate, each domain appears to have some degree of structure since positions 97, 285, and 324 are buried. This observation supports the idea that the complexity of the refolding process of yeast PGK is not due to a sequential refolding of the two domains (Missiakas, 1990; Ballery et al., 1990).

Two possibilities could explain the absence of the intermediate labeling phase with the C97A,A183C mutant. Either the intermediate does not exist for this mutant or the intermediate does exist but with the side chain of residue 183 freely accessible to the reagent. The following data suggest that the second hypothesis is correct. Indeed, no qualitative differences in the folding process of this mutant protein were detected when the refolding was monitored by circular dichroism and by ANS fluorescence; these two methods clearly show that the same kinetic intermediate exists for this mutant protein as for wild-type PGK.

In the native structure, residues 185–188 in the first turn of helix V and residues 6–8 and residue 411, respectively, on the N-terminal and C-terminal extremities are in contact with A183. The two ends of the sequence are close together in the three-dimensional structure. These stretches of structure are in the interdomain region. The high reactivity of C183 in the intermediate form suggests that in this species the interdomain interactions have not yet formed.

The evident similarities in the folding process of the four mutant proteins (see Figure 8) strongly suggest that the introduction of a cysteinyl residue is essentially neutral with respect to the folding process. In other words, the different mutants can be considered as reflecting the same protein. The difference between the wild-type enzyme and all the mutants is related to a destabilization caused by the C97A mutation (Ballery et al., 1990). The amplitudes of the fast refolding phase determined by circular dichroism and thiol titration experiments are not different for the mutants and wild-type PGK. This suggests that the effect of the C97A mutation appears late in the folding process and that the tertiary structure around C97 is not fully formed in the intermediate X. Thiol titration experiments show that C97 is indeed in a flexible environment in this species.

Events Occurring during the Slow Refolding Phase. Although the kinetics of disappearance of the fast-reacting thiol reflect a local effect in different parts of the molecule, the regain of the native dichroism signal occurs in a slow refolding phase, with a rate constant in the same order of magnitude as the rate constant deduced from thiol titration, for all mutants. The largest difference was observed for the wild-type enzyme; it is likely that the region around C97 becomes inaccessible slightly faster than the regain of secondary structure. Semisotnov et al. (1991) have pointed out that a specific feature in the folding process of both yeast PGK and pig muscle PGK is the presence of a quite slow refolding phase with a rate constant equal to 0.3 min^{-1} in 0.25 M GdnHCl; this slow step has been followed by several methods including far-ultraviolet circular dichroism, fluorescence of the tryptophan residues, enzyme activity, and ANS explusion. The same slow phase has recently been observed for yeast PGK, using pulse proteolysis as a conformational probe (Betton et al., 1992). A very similar behavior has been described for horse muscle PGK (Betton et al., 1985). These different reports lead to a conclusion that in these different proteins the slow phase apparently does not result from proline isomerization.

Can this slow refolding phase be explained by the pairing of folded but unpaired domains? The large amplitude of the slow refolding phase as observed by circular dichroism results from the fact that the intermediate form is incompletely populated. The present study shows that, during this slow phase, there are at least two different nonnative species significantly populated. A population, referred to as the "unfolded form", in which no tertiary structure is apparent, is in equilibrium with an intermediate form. The intermediate species could have the characteristics of a molten globule in each of the two domains. In the intermediate, the pairing and the stabilization of the two domains occurs simultaneously. The rate constant determined for the transformation of this intermediate to the native protein is about the same as that determined for the slow refolding phase by different methods, circular dichroism, ANS explusion, reactivation, fluorescence,

and pulse proteolysis, suggesting that the same process is observed.

Recently, it has been shown that a slow phase was observed by both circular dichroism and fluorescence for the isolated C-domain (Missiakas et al., 1992). So, for this domain the slow phase is not caused by interactions with the N-domain, but rather by intradomain interactions. For the isolated N-domain, the situation is more complex. This domain, when isolated, regains its native properties, as assessed by circular dichroism and NMR, but remains more flexible as shown by a higher isotope exchange rate and the high reactivity of its thiol group (Fairbrother et al., 1990). Therefore, interdomain interactions have probably an important role in decreasing the flexibility of the N-domain. However, the refolding kinetics of the isolated N-domain are also biphasic, as observed by the variations in ellipticity, suggesting that some slow events also exist in the folding process of this isolated domain.

Potentialities and Limitations of the Methodology. The method described in this paper can only be used for proteins in which cysteine residues are either absent or neutral for the protein structure. The accuracy of the description is then limited by the number of suitable mutants which can be constructed. The mutation should not drastically modify the folding properties of the protein, and the reaction with NbS_2 must be sufficiently different in the native and the denatured forms of the protein. The results previously reported (Ballery et al., 1990) and the present work suggest that such mutations can be predicted from simple criteria with a reasonable chance of success. It is important to note that the effect of the label on the properties of the labeled protein do not have to be taken into account. Since our precedent report on the use of single cysteine mutants for folding studies, the same experimental strategy has been used independently by Freksgård et al. (1991) to study the refolding of carbonic anhydrase II. Interestingly, this study shows that, in this protein, the local structures around the C-terminus are formed very early on during the folding process.

An interesting feature of this methodology is that as long as their rate constants of reaction with NbS_2 are sufficiently different, the different states of the protein appear as discrete subpopulations. When the labeling rate constants are higher than the folding rate constants, the concentration of the different species can be directly calculated from the amplitude of the phases. With classical spectroscopic probes such as fluorescence or circular dichroism, the signals of the different subpopulations are averaged and it is not possible to describe, without an assumption of their specific properties, the distribution of the different species. An essential result of the present work is precisely that the intermediate observed is in equilibrium with an "unfolded" form (according to the thiol accessibility criterion) which remains highly populated for minutes even under native conditions. A possible interpretation is that the kinetic intermediate in PGK is marginally stable and the structure is stabilized only as the fully folded form appears. A similar observation has been reported by Baldwin and co-workers (1990) for the refolding of RNase A followed by pulse hydrogen exchange. Indeed, in this protein, I_1, an H-bonded early intermediate on the direct pathway of folding is not fully populated and remains in solution with the unfolded species. Early intermediates in the refolding of α-lactalbumin and staphylococcal nuclease detected by stopped-flow circular dichroism were also shown to be only marginally stable (Kuwajima, 1989).

This study does not give information about possible early intermediates between the unfolded and intermediate species, since the fast rate constant cannot be accurately determined by manual methods. Preliminary circular dichroism stopped-flow data would suggest that the fast circular dichroism phase observed by manual mixing includes several different folding phases.

ACKNOWLEDGMENT

We acknowledge Dr. M. Hodges for carefully reading the manuscript and revising the English.

REFERENCES

Adams, B., Burgess, R. J., & Pain, R. H. (1985) *Eur. J. Biochem. 152*, 715–720.

Baldwin, R. L., & Roder, H. (1991) *Curr. Biol. 1*, 218–220.

Banks, R. D., Blake, C. C. F., Evans, P. R., Haser, R., Rice, D. W., Hardy, G. H., Merrett, M., & Phillips, D. (1979) *Nature 279*, 773–777.

Ballery, N., Minard, P., Desmadril, M., Betton, J. M., Perahia, D., Mouawad, L., Hall, L., & Yon, J. M. (1990) *Protein Eng. 3*, 199–204.

Betton, J. M., Desmadril, M., Mitraki, A., & Yon, J. M. (1984) *Biochemistry 23*, 6654–6661.

Betton, J. M., Desmadril, M., Mitraki, A., & Yon, J. M. (1985) *Biochemistry 24*, 4570–4577.

Betton, J. M., Desmadril, M., & Yon, J. M. (1990) *Biochemistry 28*, 5421–5428.

Betton, J. M., Missiakas, D., & Yon, J. M. (1992) *Arch. Biochem. Biophys. 296*, 95–101.

Brandts, J. F., & Lin, L. N. (1986) *Methods Enzymol. 131*, 107–126.

Brandts, J. F., Halvorson, H. R., Brennan, M. (1975) *Biochemistry 14*, 4953–4963.

Fairbrother, W. J., Minard, P., Hall, L., Betton, J. M., Missiakas, D., Yon, J. M., & Williams, R. J. P. (1990) *Protein Eng. 3*, 5–11.

Freskård, P. O., Carlsson, U., Mårtensson, L. G., & Jonsson, B. H. (1991) *Eur. J. Biochem. 289*, 117–122.

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631–660.

Kuwajima, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 87–103.

Minard, P., Bowen, D. J., Hall, L., Littlechild, J. A., & Watson, H. (1990a) *Protein Eng. 3*, 515–521.

Minard, P., Hall, L., Betton, J. M., Missiakas, D., & Yon, J. M. (1990b) *Protein Eng. 3*, 55–60.

Missiakas, D., Betton, J. M., Chaffotte, A., Minard, P., & Yon, J. M. (1992) *Protein Sci.* (in press).

Mitraki, A., Betton, J. M., Desmadril, M., & Yon, J. M. (1987) *Eur. J. Biochem. 163*, 29–34.

Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., & Razgulyaev, O. I. (1990) *FEBS Lett. 262*, 20–24.

Semisotnov, G. V., Vas, M., Chemerov, V. V., Kashparova, J., Kotova, N. V., Razgulyaev, O. I., & Sinev, M. A. (1991) *Eur. J. Biochem. 202*, 1083–1089.

Smith, C. J., Clarke, A. R., Chia, W. N., Irons, L. I., Atkinson, T., & Holbrook, J. J. (1991) *Biochemistry 30*, 1028–1036.

Udgaonkar, J. B., & Baldwin, R. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8197–8201.

Watson, H. C., Walker, N. P. C., Shaw, P. J., Bryant, T. N., Wandell, P. L., Fothergill, L. A., Perkins, R. E., Conroy, S. C., Dobson, M. J., Tuite, M. F., Kingsman, A. J., & Kingsman, S. M. (1982) *EMBO J. 1*, 1635–1640.

Yang, J. T., Wu, C. S. C., & Martinez, H. M. (1986) *Methods Enzymol. 130*, 208–269.

Yon, J. M., Betton, J. M., Desmadril, M., Mitraki, A., Minard, P., Gaillard, S., Ballery, N., & Missiakas, D. (1988) *J. Chromatogr. 440*, 421–437.

Yon, J. M., Desmadril, Betton, J. M., Minard, P., Ballery, N., Missiakas, D., Gaillard-Miran, S., Perahia, D., & Mouawad, L. (1990) *Biochimie 72*, 417–429.