Comparing the Folding and Misfolding Energy Landscapes of Phosphoglycerate Kinase

Gergely Ágocs, Bence T. Szabó, Gottfried Köhler, and Szabolcs Osváth

Department of Biophysics and Radiation Biology, Semmelweis University, Budapest, Hungary; and Max F. Perutz Laboratories, Department of Chemistry, University of Vienna, Vienna, Austria

ABSTRACT Partitioning of polypeptides between protein folding and amyloid formation is of outstanding pathophysiological importance. Using yeast phosphoglycerate kinase as model, here we identify the features of the energy landscape that decide the fate of the protein: folding or amyloidogenesis. Structure formation was initiated from the acid-unfolded state, and monitored by fluorescence from 10 ms to 20 days. Solvent conditions were gradually shifted between folding and amyloidogenesis, and the properties of the energy landscape governing structure formation were reconstructed. A gradual transition of the energy landscape between folding and amyloid formation was observed. In the early steps of both folding and misfolding, the protein searches through a hierarchically structured energy landscape to form a molten globule in a few seconds. Depending on the conditions, this intermediate either folds to the native state in a few minutes, or forms amyloid fibers in several days. As conditions are changed from folding to misfolding, the barrier separating the molten globule and native states increases, although the barrier to the amyloid does not change. In the meantime, the native state also becomes more unstable and the amyloid more stable. We conclude that the lower region of the energy landscape determines the final protein structure.

INTRODUCTION

Understanding the sequence-specific partitioning of polypeptides between protein folding and aggregation is of great physiological and pathological importance (1). Roughly two dozen human diseases have been linked to the formation of ordered insoluble aggregates called amyloids (2). The illnesses linked to amyloidogenesis include some of the most debilitating conditions, like transmissible spongiform encephalopathy, Alzheimer’s, and Parkinson’s diseases (3,4).

Proteins that were found to form amyloids in vivo have very diverse sequences (5), and many proteins not related to diseases can form amyloid fibrils in vitro (6–8). All amyloid fibrils show a similar cross β-sheet, which is stabilized by sequence-independent hydrogen bonding within the polypeptide backbone (2,9). The role of sequence specific interactions in the formation of amyloid structures has also been shown (9–11), but it is generally believed that amyloid structure formation is a universal property of proteins (2,7,12,13).

Damaschun et al. (8) have demonstrated that yeast phosphoglycerate kinase (PGK) is able to form amyloid fibrils in vitro. Short fibrillar structures were observed few days after initiating amyloid formation. The β-sheet content of the protein in this state was already identical to that adopted in the mature amyloid. Further growth of the fibrillar structures occurs by the association of the protofibrils to form mature amyloids. Fibril growth did not reach saturation even after two weeks. The presence of amyloids was proved by electron microscopy and x-ray scattering. Analyzing the misfolding and amyloid formation of PGK, Modler et al. (14) developed an amyloid growth model based on colloid coagulation theory.

PGK is a 415-residue-large monomeric protein. The native state has two roughly equal-sized domains, linked by a helical hinge. It has a combined total of 20 α-helices, 19 β-strands, and three turns. The structure of the domains is similar, consisting of a core of a six-stranded parallel β-sheet that is surrounded by a series of helices. In addition, the C-terminal domain has three shorter β-strands. It has been used in many unfolding and refolding studies, and different denatured states of PGK have been well characterized (15–19). A hierarchical energy landscape model has been developed to predict the folding kinetics of PGK from milliseconds to the complete folding (20).

Whether the protein chooses the folding or misfolding pathway can be a matter of life and death. Understanding the mechanisms of transition between native and amyloid states and revealing the underlying physicochemical processes is of pivotal importance. The possibility of reversing different steps (21–23) or the whole process (24) of amyloid formation has been shown. Several recent articles also address the relationship between the folding and amyloid formation pathways (25–27), but a lot of research has to be done to understand how the fate of proteins is determined.

Here, using PGK as the model protein, we tune buffer conditions between folding and misfolding and record the kinetics of the structural changes from 10 ms to 20 days. The method allowed us to identify the features of the energy landscape that decide the fate of the protein: folding or amyloidogenesis.
MATERIALS AND METHODS

Expression, purification, and characterization of PGK

A histidine-tagged version of yeast (Saccharomyces cerevisiae) phosphoglycerate kinase (UniProtKB/Swiss-Prot accession number P00560) was expressed and prepared as described earlier (28). Enzyme activity of the purified PGK was verified using established methods (24,29). Samples were concentrated to 40 mM, dialyzed in 50 mM Tris, 1 mM EDTA, pH 7.0 buffer, aliquoted, and stored at −80°C.

Kinetic measurements

Structure formation of samples containing 40 μM PGK unfolded in pH 2.0 HCl solution was initiated by an 11-fold mixing with a buffer containing 10 mM Na-citrate, 10 mM Na-phosphate, and 200 mM NaCl using stopped-flow (Applied Photophysics, Leatherhead, UK) or manual mixing. The pH of the diluting buffer was varied between pH 2.0 and 7.0 in 0.5 pH steps, and structure formation of PGK was detected. All kinetic values were recorded at 4°C to avoid unspecific aggregation. Conformational changes were followed by time-resolved tryptophan fluorescence measurements (excited at 295 nm with 5-nm bandwidth and detected between 305 and 400 nm). One-hundred individual mixing experiments were averaged.

Characterization of PGK amyloid

The presence of mature amyloid structures in the sample was checked by electron microscopy (11) and Thioflavin T fluorescence using established methods (24,30).

The fluorescence emission spectrum of freshly prepared 20 μM Thioflavin T solution (in 100 mM Na-phosphate, 100 mM NaCl, pH 8.5) was measured before and after the addition of protein to a final concentration of 4 μM. Spectra were recorded on a FluoroLog-3 fluorometer (Horiba Jobin-Yvon, Edison, NJ) between 455 nm and 530 nm (λexc = 450 nm) with 1-nm bandwidth. No fluorescence change was observed when PGK in the native or unfolded state was added. On the other hand, the fluorescence of the dye increased roughly fivefold upon addition of mature amyloid fibers.

Amyloid fibrils grown in vitro were visualized with a. 7100 transmission electron microscope (Hitachi, Tokyo, Japan) after negative staining. The sample was diluted to 1 μM protein concentration and a 50 μL droplet of it applied to a Formvar-coated copper grid. An equal volume of 2% uranyl-acetate solubilized in 50% methanol was added to stain the sample for 15 min. After this time, the excess solution was removed and the sample let dry for 15 min. No fibrils were detected in samples containing the protein in the native or the unfolded state.

RESULTS

Yeast phosphoglycerate kinase (PGK) was unfolded by dialysis against pH 2 solution with low ionic strength. Structure formation of the unfolded protein was initiated by rapid mixing in buffers containing 200 mM NaCl, and followed from the millisecond to the week timescale. Solvent conditions were gradually changed from favoring amyloid formation to native folding by tuning the pH of the buffer between pH 2.0 and pH 7.0.

Early kinetics of structure formation of the acid unfolded PGK was recorded after stopped-flow mixing into various buffers with different values of pH. The pH of the solution after mixing was varied in 0.5 pH unit steps from pH 2 to pH 7. The time course of tryptophan fluorescence was monitored from 10 ms to 50 s. One-hundred individual traces were averaged, and the fluorescence was normalized to the fluorescence intensity observed at 10 ms. All curves indicate the formation of a hyperfluorescent state with a kinetics that is stretched over several orders of magnitude in time. Fig. 1 shows the kinetics recorded at pH 2.0, 4.5, and 7.0.

Earlier results demonstrated that at neutral pH, the first steps toward native structure formation of PGK proceed on a hierarchically structured energy landscape (20). Here we used this hierarchic finite level model to fit the early kinetics of PGK structure formation at all pH values.

The model considers the interconversion of the unfolded (U), intermediate (I), and folded (F) states; amyloid (A) formation is neglected because it is much slower. The measured fluorescence intensity (Fluo(t)) can be calculated from the amount of the different states and their fluorescence efficiency (ΦU, ΦI, ΦF):

\[
\text{Fluo}(t) = \Phi_U \times U + \Phi_I \times I + \Phi_F \times F.
\]

Conversion of the unfolded into the intermediate state occurs on a hierarchically structured energy landscape. A multitude of traps that can have N different depths give rise to a large number of relaxation pathways described by N different relaxation times. The amount U(t) of unfolded

![FIGURE 1 Early steps of structure formation of PGK detected by tryptophan fluorescence. Structure formation of acid unfolded PGK was initiated by stopped flow dilution and followed by tryptophan fluorescence measurements. Data were fitted with a hierarchic landscape model. The result of the fit is also shown. Residuals obtained as the difference between the measured and fitted curves demonstrate the goodness of fit. Buffer conditions after mixing: 3.6 mM PGK, 9.1 mM Na-citrate, 9.1 mM Na-phosphate, and 182 mM NaCl.

Biophysical Journal 102(12) 2828–2834
protein at time $t$ after initiating structure formation can be calculated as

$$U(t) = U(0) \times \sum_{i=1}^{N} \left( a_i / \tau_k \times \exp(-t/\tau_k) \right).$$

Here $U(0)$ is the amount of unfolded protein at time $t = 0$. Because structure formation was initiated from the acid unfolded state, $U(0)$ is the total amount of protein in our experiments. The parameters $a_i$ are the amplitudes and $\tau_k$ the time constants of the Debye contributions, which correspond to different levels of the energy landscape structure. For a landscape that consists of a multitude of hierarchically structured kinetic barriers, these components are not independent, but their amplitudes and relaxation times obey a simple scaling law:

$$a_i = a_1 \times \alpha^{i-1}, \quad \tau_k = \tau_1 \times \lambda^{i-1}.$$

The fastest relaxation pathway has a time constant of $\tau_1$. Every deeper level has $\lambda$-times longer relaxation time than the level before it. This early structure formation shows a stretched kinetics that lacks a characteristic time. The value $\tau_1$ thus is not a characteristic time of the process, but rather it is a parameter describing the entire hierarchic landscape. The $a_i/\tau_k$ ratio shows the fraction of the protein molecules that follows the pathway with the characteristic time $\tau_k$, thus the sum of the $a_i/\tau_k$ values is 1. This relationship and the scaling property of the landscape reduce the number of the free fit parameters to the constants $\tau_1$, $\alpha$, and $\lambda$ regardless of $N$.

The above model fitted well the initial steps of the structure formation at all pH values (Fig. 1). The fluorescence intensity of the intermediate structure showed small variations as the pH was changed. The number $N = 5$ of the possible trap sizes, the parameter $\alpha/\lambda = 1.0 \pm 0.15$ and $\lambda = 8.1 \pm 0.9$, did not change significantly as pH was altered. The shortest relaxation time ($\tau_1$) was the only parameter of the model that showed significant pH dependence.

**Fig. 2** shows the pH dependence of the shortest relaxation time $\tau_1$ obtained from the hierarchic finite level model fit of the fast kinetic measurements (see Fig. 1).

The folding kinetics could be fitted well with a single exponential at pH values between 4.5 and 7.0. The characteristic time ($\tau_{F,I}$) of the reaction was determined from the fit.

**Fig. 4** shows the pH dependence of the time constant of folding to the native state from the intermediate conformation (Fig. 4A) and the fluorescence change associated with it (Fig. 4B). No native structure formation was observed below pH 4.5.

**FIGURE 2** pH dependence of the shortest relaxation time $\tau_1$ obtained from the hierarchic finite level model fit of the fast kinetic measurements (see Fig. 1).

**FIGURE 3** Structural changes of PGK detected by tryptophan fluorescence. Structure formation of acid unfolded PGK recorded by tryptophan fluorescence after a quick dilution by mixing. Monoexponential fit of the observed kinetics is also indicated. Residuals obtained as the difference between the measured and fitted curves are also shown. Buffer conditions after mixing: 3.6 mM PGK, 9.1 mM Na-citrate, 9.1 mM Na-phosphate, and 182 mM NaCl.

**FIGURE 4** pH dependence of the time constant of folding to the native state from the intermediate conformation (Fig. 4A) and the fluorescence change associated with it (Fig. 4B). No native structure formation was observed below pH 4.5.
Fig. 5 represents the formation of amyloid structures monitored by Thioflavin T fluorescence from 1 min to 20 days. The gradual increase of the fluorescence intensity observed for the samples kept at low pH indicates amyloid accumulation. The observed kinetics do not change significantly between pH 2.0 and 4.5. Above pH 4.5, however, Thioflavin T fluorescence did not indicate amyloid formation.

Modler et al. (14) have shown that a good description of PGK amyloid growth can be given by the Smoluchowski coagulation theory, using size-independent association rate-constants. We used this model to fit the observed amyloid formation kinetics. The model gave a good fit to our measurements at all pH values.

If the sample contains only monomers at time \( t = 0 \), the concentrations \( n_i \) of the oligomers containing \( i \) associated protein molecules can be calculated at any time \( t \) (14):

\[
n_i(t) = n_i(0) \times \frac{(t/t_c)^{i-1}}{(1 + t/t_c)^{m+1}},
\]

Here \( t_c \) denotes the coagulation time, a constant characteristic to the process. Because there are only monomers present at \( t = 0 \), \( n_1(0) \) is the protein concentration.

It is known that Thioflavin T binds to fibrillar amyloid structures, but not to monomeric protein. We assumed that the dye binds to the forming PGK aggregate when the oligomer becomes bigger than \( m \) monomeric units. The fraction \( f_m \) of the protein molecules that are in oligomers smaller than or equal to \( m \) monomers is

\[
f_m(t) = \frac{1}{C_0} \sum_{i=1}^{m} i 
\times n_i(t).
\]

By simplifying the equation analytically, the following formula was obtained:

\[
f_m(t) = 1 - (m + 1) \left( \frac{t/t_c}{1 + t/t_c} \right)^m \left( 1 + t/t_c \right)^{m+1}.
\]

The measured Thioflavin T fluorescence intensity can be calculated as

\[
Fluo_{ThT}(t) = \Phi_0 + (\Phi_A - \Phi_0) \times (1 - f_m).
\]
Here $\Phi_0$ is the fluorescence of the dye in the absence of amyloid, and $\Phi_A$ is the Thioflavin T fluorescence if all the protein is in amyloid fibers.

The $t_c$ characteristic time of coagulation was extracted from the model fit of the kinetics. This parameter does not show a significant variation with pH.

**DISCUSSION**

Structure formation initiated from the acid unfolded state of PGK was tuned from folding to amyloid formation by adjusting pH from 7.0 to 2.0. Earlier results found that at neutral pH, PGK folded in the enzymatically active structure (15,32), whereas amyloid formation was observed at pH 2 (11,33). Conformational changes were followed at $4^\circ C$ in the 10 ms to 20 days time range by tryptophan and Thioflavin T fluorescence. Depending on the pH, either folding to the native state or amyloid formation occurred in the samples. Amyloid fibers were visualized after 20 days by electron microscopy.

Analyzing kinetics recorded at various pH indicates significant accumulation of only four states: unfolded, intermediate, native, and amyloid. In such a simple scenario, the molten globule intermediate is off pathway to the amyloid formation only if the amyloid is formed directly from the unfolded state. Earlier results, however, proved that interaction between the two folding domains of PGK directs the tertiary structure formation during amyloid growth (11). This indicates that amyloid formation happens from a state in which the domain structure of the protein is already formed, i.e., the molten globule intermediate and not the unfolded state. We also showed earlier that folding to the native state happens through the molten globule intermediate (20).

The early steps of structure formation of PGK followed stretched kinetics in the entire studied pH range (Fig. 1). Stretched kinetics was observed earlier in the folding of PGK at neutral pH (31,34), and explained by a hierarchic landscape structure with a finite number of possible relaxation pathways (20). Opposite to other landscape models, this model was developed using no a priori assumptions, giving the strongest evidence that time is important for determining the existence of complex folding landscape structures. Our landscape can be described using three parameters, one of which is a time constant. This time constant, however, is not a characteristic time of an exponential.

Changing this time constant affects the entire kinetics through the scaling relationship of the model, and a very sensitive fit of this parameter is possible. The measurements show a vibration artifact often observed in stopped-flow data. The largest modulation amplitude of the vibration artifacts, however, is $3\text{–}5\%$ of the amplitude of the stretched fluorescence increase, and the modulation decreases quickly with time. Because the fit is done to the whole process extending through three-and-a-half orders of magnitude in time (10 ms to 50 s), the artifact does not affect the fit with the hierarchic landscape model. The model fitted well to our observation at all pH values.

We found that that the parameters of the upper part of the energy landscape, which direct the early steps of structure formation, were pH-independent. The number $N = 5$ of the possible relaxation times, the parameter $a/\lambda = 1.0 \pm 0.15$, and the parameter $\lambda = 8.1 \pm 0.9$, did not change significantly as pH was altered. This means that the structure of the upper region of the energy landscape, which governs the pathway and mechanism of the early conformational changes, is the same for folding and for amyloid formation. The number $N$ and the scaling factor $\lambda$ of the relaxation times were identical to those found earlier for the refolding from the GuHCl unfolded state at neutral pH and room temperature (20). This indicates that the refolding landscapes were essentially the same in the two experiments. The $a/\lambda$ ratio of the model is the scaling factor for the number of molecules that follow different relaxation pathways. The difference between folding from acid denatured state at $4^\circ C$ ($a/\lambda = 1.0 \pm 0.15$) and refolding from GuHCl unfolded state at room temperature ($a/\lambda = 0.75 \pm 0.17$) comes from the different distribution of the unfolded ensemble in the two experiments.

The shortest relaxation time ($t_1$) was the only parameter of the model that showed significant pH dependence (Fig. 2). This pH dependence reflects the effect of acid denaturation on the upper regions of the energy landscape. As pH is decreased from neutral, early structure formation of PGK speeds up showing a flattening of the upper regions of the landscape. It was found that the quick relaxation from the unfolded to the intermediate state is fastest at pH 3.5. Decreasing the pH further results in a slowing of the early structural changes.

Based on the above analysis of the early steps of structure formation of PGK, we conclude that the folding and the misfolding preceding amyloid formation follow the same path up to the formation of a hyperfluorescent intermediate. Earlier measurements done at neutral pH and room temperature showed that this intermediate has molten globule structure (20).

From the fast kinetic measurements, we conclude that the upper region of the energy landscape does not influence the final structure of PGK. The fate of the protein is determined by the lower part of the energy landscape. To explore this landscape region, structure formation from the intermediate state was followed by tryptophan fluorescence measurements from 1 min to 1 day (Fig. 3). This part of the kinetics followed a monoeponential time dependence indicating a simple cross-barrier reaction. Folding to the native enzyme structure was observed at neutral pH in accordance with earlier studies (20). Decreasing sample temperature to $4^\circ C$ from room temperature, where earlier experiments were carried out, slowed PGK folding down roughly eightfold (20).

As the pH was decreased, the characteristic time obtained from the fit increased and the amplitude of the change...
decreased (Fig. 4). This indicates that the native structure is formed in a decreasing fraction of the molecules and more slowly as pH is decreased from the neutral. As Fig. 4 B indicates, the midpoint pH where half of the molecules are in folded, half in intermediate state in equilibrium, is roughly pH 4.7. Practically no native structure formation was observed at pH 4.0 and below it. These results suggest that the native state is destabilized compared to the intermediate state as the pH is decreased. At approximately pH 4.7 the two states become equally stable. Lowering the pH further, the intermediate becomes more stable compared to the folded conformation, and no native PGK is formed. Lengthening of the characteristic time shows the increase of the barrier separating the intermediate from the native state.

We detected amyloid formation from 1 min to 20 days using Thioflavin T fluorescence (Fig. 5). The fluorescence of the dye did not change above pH 4.5, but its intensity increased 2–3 times during the experiment at pH <4.5. This indicates that amyloid fibers are formed only below pH 4.5.

Modler et al. (14) found that the amyloid formation of PGK at pH 2.0 can be described using the Smoluchowski coagulation theory and size-independent association rate-constants. The Thioflavin T fluorescence kinetics could be well fitted with the model at all pH values (Fig. 5). For several proteins it has been found that the mechanism of amyloid formation depends on the buffer conditions, e.g., pH (7). Our results indicate that in the case of PGK, the amyloid growth mechanism is the same throughout the studied pH range. The characteristic coagulation time \( t_c \) does not show significant pH dependence, indicating that the barrier between the molten globule and amyloid states is not sensitive to pH. The lack of large change in the amyloid growth kinetics as a function of pH makes kinetic control of partitioning between amyloid and native structures unlikely.

The only rationale for kinetically controlled amyloid formation would be the appearance of a disproportionately high kinetic barrier between the intermediate and native structures as pH is decreased whereas the folded structure remains the thermodynamically most stable state on the energy landscape. The fact that the native structure of PGK unfolds quickly under acidic conditions contradicts the above rationale. The fact that no amyloid structure is accumulated above pH 4.5 even after long waiting times also indicates that the fate of the protein is not kinetically determined. Lowering the pH makes the amyloid fibers more stable than the intermediate, allowing the growth of amyloid fibers.

CONCLUSIONS

Early steps of the folding and the amyloid formation of PGK follow the same path. The protein searches through a hierarchically structured energy landscape to form a molten globule intermediate (I).

Here the pathways leading toward the native and the amyloid structures diverge. Fig. 6 illustrates the changes in the energy landscape of PGK as conditions are gradually tuned from folding to misfolding. Above pH 4.5, PGK folds to the native state (F), which is thermodynamically the most stable conformation.

As pH is decreased, the native state becomes gradually less stable and less accessible from the molten globule, whereas the amyloid state becomes more stable. The kinetic barrier separating the intermediate from the amyloid does not depend on pH. Below pH 4.5, the amyloid structure represents a global minimum of the energy landscape, and PGK forms amyloid fibers.

This work was funded by the Hungarian Scientific Research Fund, grant OTKA K-77730; the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, grant BO/00468/09; and the Social Renewal Operational Programme, grant TÁMOP-4.2.1/B-09/1/KMR-2010-0001. The research leading to these results has received funding from the European Union’s Seventh Framework Programme (FP7/2007-2013) under grant agreement HEALTH-F2-2011-278850 (INMIND).

REFERENCES


Biophysical Journal 102(12) 2828–2834