Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle x-ray scattering

LOIS POLLACK*, MARK W. TATE*, NICHOLAS C. DARNTON‡, JAMES B. KNIGHT‡, SOL M. GRUNER*, WILIAM A. EATON§, AND ROBERT H. AUSTIN‡

*Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, NY 14853; ‡Physics Department, Princeton University, Princeton, NJ 08544; and §Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Communicated by Hans Frauenfelder, Los Alamos National Laboratory, Los Alamos, NM, June 24, 1999 (received for review March 27, 1999)

ABSTRACT Time-resolved small-angle x-ray scattering was used to measure the radius of gyration of cytochrome c after initiation of folding by a pH jump. Submillisecond resolution was obtained with a microfabricated diffusional mixer and synchrotron radiation. The results show that the protein first collapses to compact denatured structures before folding very fast to the native state.

The stability and folding speed of a protein depend on the structures of the denatured as well as the native state, raising the question: do proteins fold more rapidly from a denatured state of expanded structures or from one of compact structures? Lattice simulations of simplified representations of proteins suggest that slow folding amino acid sequences collapse to compact structures with non-native topologies before folding, while fast folders collapse and fold simultaneously (1–4). We have begun to address this question experimentally with submillisecond small-angle x-ray scattering (SAXS) using synchrotron radiation and a microfabricated diffusional mixer to rapidly initiate folding. SAXS yields the radius of gyration, the most unambiguous measure of compactness. Here we show that, in contrast to the simulations, one of the fastest-folding proteins (cytochrome c: \( t_{\text{folding}} = 400 \mu\text{s} \)) first collapses to compact structures before forming the final native state.

X-ray scattering by proteins in solution is sensitive to spatial variations in electron density. Scattering at the smallest angles yields the radius of gyration, \( R_g \), which in conjunction with the protein molecular weight provides a measure of the compactness of globular proteins. Additional structural information can be obtained from scattering at larger angles, which reflects electron density correlations on length scales shorter than \( R_g \). For a compact polymer, such as the native protein or compact denatured structures, \( I(q)q^2 \) increases at low \( q \), goes through a maximum, and decreases at large \( q \) where \( I(q) \) is the scattered intensity; \( q = \frac{4\pi\sin \theta}{\lambda} \), is the momentum transfer; \( \lambda \) is the x-ray wavelength; 1.54 Å; and \( 2\theta \) is the scattering angle] (5). In contrast, for a polymer chain undergoing a random walk in space, as can occur for an unfolded protein under strongly denaturing conditions, \( I(q)q^2 \) first increases, then plateaus (6), and, at large \( q \), where \( I(q)q^2 \) increases linearly (7). In this way, using Kratky plots, \( I(q)q^2 \) vs. \( q \), one can distinguish easily between random coil and compact conformations, making time-resolved SAXS a powerful tool for investigating structure in protein folding experiments (8–10).

Because of their simpler kinetic behavior, small single-domain proteins have been the focus of both experimental and theoretical investigations of the mechanism of protein folding. Folding times of these proteins in the absence of denaturant are often less than \( \sim 10 \) ms (11), which is too fast to be studied with the mixing technologies used in previous time-resolved SAXS studies (8–10). We have increased the time resolution by almost 2 orders of magnitude by using a lithographically microfabricated mixing device. Its design is based on the principle that diffusional mixing times scale as the square of the linear dimension of the mixing volume. Short mixing times are achieved by hydrodynamic focusing, which creates a micron or submicron stream of protein solution in contact with a surrounding flowing reservoir. We fabricated an x-ray compatible device at the Cornell Nanofabrication Facility (Fig. 1). Synchrotron radiation was necessary for these experiments because of the combination of micro sample volumes and the normally weak scattering of protein solutions. Time-resolved scattering was obtained by passing a micro-collimated x-ray beam through the sample at fixed locations in the flowing stream of protein solution. Relative time was determined from the velocity of the protein stream and the position of the x-ray beam.

MATERIALS AND METHODS

Fig. 1 shows a cross-sectional view of the central portion of the diffusive mixer. The channels are 100 μm wide and 390 μm deep. In operation, 2.5 mM of cytochrome c in HCl at pH 2 is injected into the inlet port (Fig. 1 Left). We chose cytochrome c (104 aa, Sigma catalog no. C7752) for our initial study using this method because its folding kinetics to times as short as 50 μs already had been characterized (13–18). A 0.1 M phosphate buffer at pH 7 containing 0.2 M imidazole, introduced into the orthogonal channels at a higher pressure, hydrodynamically focuses the low pH, protein-containing solution into a thin stream (12). This focused protein stream, sheathed by the pH 7 buffer, shoots out the fourth exit port of the device (Fig. 1 Right). Rapid increases in pH occur by diffusion of protons out of and buffer into the focused protein solution. Although the stream eventually reaches pH 7, equilibrium studies indicate that cytochrome c is almost fully native above pH 3 (19). Once the pH of the focused protein solution is above 3, the protein folds as it flows. Scattering data at different times can be obtained by moving the device relative to the fixed position of the x-ray beam.

The x-ray experiments were conducted at the Cornell High Energy Synchrotron Source’s D-line, using a multilayer monochromator to enhance the bending magnet beam. The beam at the sample had a flux of \( 10^8 \) x-rays/s into an area that was 40 μm along the width of the protein stream and 120 μm along the direction of flow. A charge-coupled device area detector (20) was used to record the scattering patterns.

As a result of the large dimensions of these channels, low driving pressures for the fluids are involved and both the width and speed of the focused protein solution vary significantly as a function of depth in the channel. At the surface of the device, the protein stream was close to 40 μm wide; its width rapidly decreased as the stream entered the mixing region. The beam size was not critical because the data were collected with low-energy x-rays, where scattering is dominated by the form factor of the molecule, not structure factor contributions from interatomic interactions.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviation: SAXS, small-angle x-ray scattering.

†To whom reprint requests should be addressed.
tryptophan fluorescence, which results from increased

During this period, however, there is contribution to the

this paper was 30 cm

speed of the focused protein solution for the work described in

the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

of the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,

the average protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

RESULTS AND DISCUSSION

Kratky plots of the data from three positions in the device are

shown in Fig. 2. For the initial state, measured at a position

before mixing, the product $Iq^2$ increases nearly linearly at large

$q$, indicating that the average structure is expanded and

random-coil like. It is, however, not possible to extract an $R_g$

from the data at low $q$ because of interference caused by the

high concentration of a highly charged protein (21) (2.5 mM

per mL). Imidazole may not completely bind before non-native

histidines produce misfolded structures that fold more slowly than the

protons. At the exit of the outlet channel, the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

RESULTS AND DISCUSSION

Kratky plots of the data from three positions in the device are

shown in Fig. 2. For the initial state, measured at a position

before mixing, the product $Iq^2$ increases nearly linearly at large

$q$, indicating that the average structure is expanded and

random-coil like. It is, however, not possible to extract an $R_g$

from the data at low $q$ because of interference caused by the

high concentration of a highly charged protein (21) (2.5 mM

per mL). Imidazole may not completely bind before non-native

histidines produce misfolded structures that fold more slowly than the

protons. At the exit of the outlet channel, the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.

The result from our experiments is the x-ray scattering

pattern at the intermediate time interval of 150–500 μm

as determined by confocal microscopy with a pH-
sensitive fluorescent dye (the corresponding position in the
device is defined to be $t = 0$ for the analysis). This technique

provides a direct measure of proton diffusion. The protein

stream also broadens as a result of diffusion; however, because of
the larger mass of the macromolecules, they diffuse much

more slowly than the protons. At the exit of the outlet channel,
the average final protein concentration is expected to be a

factor of two less than the initial concentration. The mean

speed of the focused protein solution for the work described in

this paper was 30 cm/s.
sponding to a mean heme-tryptophan distance <2.0 nm (com-
pared with ~1 nm in the native structure and ~4 nm in the acid
denatured structures). According to this model, then, our ob-
served Kratky plot in the intermediate time range, at
150–500 μs, is primarily a linear combination of Kratky plots
from two states: compact denatured and native. Because of in-
complete mixing, there is also a small contribution from
denatured protein at pH 2. Using the populations from the
fluorescence kinetics we have calculated an upper limit for the
R_g of the compact denatured state, 18.1 ± 0.9 Å. The R_g of
the compact denatured structures is therefore no more than 30%
larger than that of the native structure. This is the same ratio
observed for the denatured state formed by the addition of salt
at pH 2, which collapses the structure because of counterion
shielding of the electrostatic repulsion among the 24 positively
charged residues (23). Thus, in the absence of denaturant, one
of the fastest denaturing proteins known first collapses to form a
relatively compact denatured state.

It is interesting to compare this result with lattice simul-
ations (1–4) of simplified representations of proteins, which
suggest that proteins fold faster if they do not first collapse to
form compact structures. In the simulations these collapsed
denatured structures are misfolded. The subsequent step to
form the native topology is slowed because it involves breaking
many non-native contacts. If, on the other hand, folding and
collapse occur simultaneously, fewer non-native contacts for,
and folding in the simulations is faster. An interesting ex-
planation for the apparent disagreement between our exper-
imental result and the simulations is suggested by the rate ex-
pression for folding, considered as diffusion of the polypep-
tide on a partially rough, free energy landscape (24, 25). In
this description the folding time, τ_f, is given by (24):

\[ \tau_f \approx (\Delta Q^2) D_{\text{pol}}^{-1} \exp \left( \frac{\Delta E^2}{k_B T} \right) \exp \left( \frac{\Delta F}{k_B T} \right) \]  

where \( \Delta Q^2 \) is the mean square fluctuation in the configura-
tional coordinate in the denatured state, \( \Delta E^2 \) is the roughness
of the energy landscape, \( D_{\text{pol}} \) is the diffusion constant in the
absence of roughness, and \( \Delta F \) is the free energy barrier
separating the denatured and native states. If the compact
denatured state has a native-like rather than misfolded topol-
omy, Eq. 1 predicts that folding is faster because all three terms
in this expression for τ_f are expected to be smaller. These
considerations suggest that one contribution to fast folding in
cytochrome c is a native-like topology of the denatured state
(30). There is experimental evidence to support this hypothesis
from amide exchange experiments, which show that the com-
 pact denatured state contains segments of native helical
structure (26). It would, of course, be quite important if a
native-like topology were a general property of compact
denatured states (27). The “folding problem” at the level of
overall topology then would be solved for this class of proteins
in the initial rapid collapse of the polypeptide chain (\( \tau_{\text{collapse}} \approx 50 \mu s \) for cytochrome c; ref. 18). Rapid collapse, moreover, to
native-like topologies with burial of hydrophobic residues may
have a biological advantage by reducing aggregation and
proteolysis (28).

The current work suggests that it will be important to obtain
systematic data on folding kinetics and time-resolved SAXS for
a number of proteins. Such studies will be greatly aided by the
brilliant third-generation synchrotron sources that have mi-
cro-sized microfocused beams 100-fold more intense than
those used here. Together with faster flows and thinner
streams time resolution an order of magnitude faster will be
possible. Acquisition of higher signal-to-noise scattering data
also should permit the determination of additional structural
information on the compact denatured state (29).

We thank E. Fontes, M. Novak, M. Skvarla, P. Infante, W. Zipfel,
and R. Williams for valuable assistance, and J. Hofrichter, V. Muñoz,
A. Szabo, S. J. Hagen, J. Trewella, and D. Shalloway for helpful
discussions. This work used Cornell High Energy Synchrotron Source,
a National User Facility supported by the National Science Founda-
tion through the Cornell Nanofabrication Facility, supported by the National
Science Foundation, Cornell University, and industrial affiliates; and
the Developmental Resource for Biophysical Imaging and Opto-
electronics at Cornell, a National Institutes of Health-National Center
for Research Resources facility. This work was supported by grants
from the National Science Foundation and the Department of Energy.

4732–4743.

Design 1, 103–116.

4070–4073.


387–432.


9. Semisotnov, G. V., Kihara, H., Kotova, N. V., Kimura, K.,
Ameniya, Y., Wakabayashi, K., Serduyuk, Y. N., Timchenko,
A. A., Chiba, K., Nikaido, K. et al. (1996) J. Mol. Biol. 262,
559–574.

10. Segel, D. J., Bachmann, A., Hofrichter, J., Hodgson, K. O.,
489–499.


33, 6925–6935.

1779–1784.

16. Takahashi, S., Yeh, S.-R., Das, T. K., Chan, C.-K., Gottfried,

17. Sosnick, T. R., Shitlereman, M. D., Mayne, L. & Englander, S. W.


20. Tate, M. W., Eikenberry, E. F., Barna, S. L., Wall, M. E.,
28, 196–205.


Biotechnology 37, 12443–12451.

Biol. 229, 591–596.


254, 960–967.