Eur Biophys J (1997) 25: 211–219

ARTICLE

Klaus Gast · Angela Nöppert · Marlies Müller-Frohne
Dietrich Zirwer · Gregor Damaschun

Stopped-flow dynamic light scattering as a method to monitor compaction during protein folding

Received: 8 August 1996 / Accepted: 18 October 1996

Abstract Kinetic dynamic light scattering is a useful tool to follow compaction during protein folding. In contrast to measurements of the formation of secondary structure and side chain ordering, kinetic measurements of compactness are not well established up to now. This work describes the adaptation of a stopped-flow system (SFM-3) to a dynamic light scattering apparatus, which allows one to monitor the compaction of protein molecules by measuring the hydrodynamic Stokes radius R. The feasibility of such investigations was demonstrated by measuring R and the integrated scattered intensity I during refolding of ribonuclease A and phosphoglycerate kinase from yeast. Refolding was initiated by rapid mixing of protein solutions containing high concentrations of guanidine hydrochloride with buffer. Between 20 and 50 mixing events were performed in these experiments. Measuring both R and I in one and the same experiment is important to distinguish between true folding of individual molecules and cases where folding is accompanied by the appearance of transient oligomers or associated misfolded structures. On refolding of ribonuclease A (0.6 M GuHCl, 25 °C), after a fast phase the Stokes radius decreased from 2.26 nm to 1.95 nm with a time constant of 27 s without detectable aggregates. By contrast, transient and stable oligomers have been observed during the more complex folding of phosphoglycerate kinase. In general, the time-resolution of the method is of the order of 1 s. It can be extended to the subsecond time-range if the number of shots is not limited by the amount of protein available.

K. Gast (✉) · A. Nöppert · M. Müller-Frohne
Max-Delbrück-Center for Molecular Medicine Berlin-Buch,
Robert-Rössle-Strasse 10, D-13122 Berlin, Germany
(Fax: +49 30 9406 2548; e-mail: gast@orion.rz.mdc-berlin.de)

D. Zirwer · G. Damaschun
Humboldt University at Berlin, Institute of Biology,
c/o Max-Delbrück-Center for Molecular Medicine Berlin-Buch,
Robert-Rössle-Strasse 10, D-13122 Berlin, Germany

Key words Protein folding · Dynamic light scattering · Stopped-flow kinetics · Ribonuclease A · Phosphoglycerate kinase

Abbreviations DLS, dynamic light scattering; CD, circular dichroism; GuHCl, guanidine hydrochloride; RNase A, ribonuclease A; PGK, phosphoglycerate kinase

Introduction

The mechanisms and the dynamics of how a protein folds into its native, biologically active structure are the subjects of a still increasing number of theoretical and experimental studies. For reviews of basic principles, see for example (Bryngelson et al. 1995; Dill et al. 1995; Karplus and Weaver 1994). Particularly, the characterisation of both equilibrium and kinetic intermediates is considered to be the key for understanding the folding process (Kuwajima 1989). For a long time the existence of transient folding intermediates was assumed to be essential for fast and efficient folding (Kim and Baldwin 1990). However, recent theoretical and experimental studies have questioned this useful role and suggest that populated intermediate states may represent trapped non-productive conformations (Wolynes et al. 1995; Sosnick et al. 1994; Kiefhaber 1995). Evidence for fast folding in the absence of detectable intermediates was provided by Schindler et al. (1995). The problem of on-pathway and off-pathway intermediates was discussed by Baldwin (1996). Therefore, investigations of the characteristics and the role of folding intermediates have reached a new stage.

Important parameters describing the folding process are: 1. the degree of collapse of the polypeptide chain, 2. the formation of secondary structure and 3. side chain ordering. In the case of kinetic intermediates secondary structure formation and side chain ordering were followed down to the millisecond time range using spectroscopic techniques, mainly circular dichroism (CD) spectroscopy (Kuwajima et al. 1987; Elöve et al. 1992), and NMR com-

bined with hydrogen exchange labelling and quenched flow methods (Udgaonkar and Baldwin 1988; Roder et al. 1988). By contrast, kinetic studies of the dimensions of proteins describing compaction upon folding have not yet reached a comparable level. X-ray or neutron scattering experiments give the best and most direct access to geometric quantities in terms of the radius of gyration or the distance distribution function of the atomic structure. These techniques have been successfully applied to study unfolded states and partially folded states under equilibrium conditions (Damaschun et al. 1991; Flanagan et al. 1992; Damaschun et al. 1993; Gast et al. 1994). By contrast, kinetic X-ray studies on transient folding intermediates are extremely difficult and have been reported in only a few cases (Phillips et al. 1988; Eliezer et al. 1993; Eliezer et al. 1995). The most impressive experiment was performed by Eliezer et al. (1995), who obtained the radius of gyration of apomyoglobin 100 ms after initiation of folding by rapid dilution of a solution containing 5.6 M urea.

Alternatively, quantities less directly related to the degree of compaction can be derived from fluorescence (Buckler et al. 1995) and hydrodynamic data. Suitable techniques for a rapid estimation of the hydrodynamic Stokes radius of proteins in different conformational states are dynamic light scattering (DLS) and size-exclusion chromatography (Uversky 1993). Though size-exclusion chromatography has been used to monitor the kinetics of protein folding (Zerovnik and Pain 1987; Shalongo et al. 1993), its application is restricted to the time range of minutes.

There is an increasing number of applications of DLS to study the time-dependence of particle size distributions in macromolecular systems. Frequently, these investigations have been performed to monitor the kinetics of micellation or aggregation. A particularly interesting experiment was reported by Chu et al. (1995), who studied the kinetics of the coil globule collapse for polystyrene in cyclohexane. Such a macromolecular collapse is also predicted theoretically for the initial stages of protein folding (Dill 1985; Dill et al. 1995). Whether this collapse precedes or is coupled to secondary structure formation is a central question in the protein folding problem.

Recently, we have demonstrated that DLS is a useful method to study the kinetics of unfolding and refolding of proteins (Gast et al. 1992, 1993). In these experiments, folding and refolding were initiated by temperature jumps. However, temperature induced folding can be observed only in particular cases, because many proteins aggregate on heating and only a few proteins undergo cold denaturation at easily attainable temperatures. Thus, most of the folding reactions were initiated by changing the solvent using rapid mixing techniques. DLS investigations monitoring the compaction of hen egg white lysozyme when refolding is initiated by 10-fold dilution from 5 M guanidinium hydrochloride (GuHCl) have been reported (Feng and Widom 1994). These authors have coupled a DLS apparatus to a continuous-flow system (Feng et al. 1993). In such an experiment, the refolding time is determined by the propagation of the mixed solution from the mixer to the observation cell. This method is only useful for measuring the Stokes radius at a few delay times within a time interval, which is defined by appropriate flow conditions and dead volumes. In this work we report on the adaptation of a stopped-flow system SFM-3 (Biologic, France) to a DLS apparatus. The time resolution of a stopped-flow DLS experiment and the advantages and limitations for protein folding investigations are discussed. To demonstrate this we have performed experiments with two proteins, which reveal folding processes consistent with the attainable time resolution of about 1 s. The folding of bovine ribonuclease A (RNase A) and phosphoglycerate kinase (PGK) from yeast was studied after a GuHCl-dilution jump. For both proteins, particular barriers slow down the folding process. In such cases of slow folding it is very important to differentiate between on-pathway intermediates and misfolded side products of folding. DLS in combination with static light scattering not only allows one to monitor compaction but also can distinguish between changes in molecular size and variation in the state of association.

Materials and methods

Materials

Ultra pure GuHCl was obtained from ICN Biomedicals, Inc. The molarity of GuHCl was controlled refractometrically (Pace et al. 1989). Bovine pancreatic RNase A was purchased from Sigma (type XII-A) and was used without further purification. RNase A was dissolved in 10 mM sodium acetate buffer, pH 4, containing 6 M GuHCl. Protein concentrations were determined photometrically using A (0.1%, 1 cm) = 0.68 at 277 nm. Yeast PGK was obtained from Boehringer Mannheim GmbH (FRG). The precipitated protein of the commerical ammonium sulfate suspension was dissolved in 20 mM sodium phosphate, pH 6.5, 1 mM EDTA, 1 mM DTT. The sample was applied to a FPLC-Superose 12 column (Pharmacia LKB Biotechnology) equilibrated with the same buffer in order to remove ammonium sulfate and protein aggregates. Collected peak fractions were then dialysed against buffer containing 2 M GuHCl. PGK concentrations were determined using A (0.1%, 1 cm) = 0.495 at 277 nm.

Experimental set-up, general considerations

The laboratory built DLS apparatus was described previously (Gast et al. 1992). The adaptation of the DLS set-up to the SFM-3 stopped-flow system (Bio-logic, France) is shown in Fig. 1. It consists mainly of an argon laser Lexel 3500 (Lexel Laser, Inc., USA) operating at $\lambda = 514.5$ nm and mostly at 1 W output power and a 90-channel multibit multiple-τ correlator that calculates the homodyne autocorrelation function $G^2(\tau)$. A Turbo Pascal program, which controls the entire experiment is run on one of two microcomputers (PC) incorporated into the system. It performs data acquisition, data storage on the

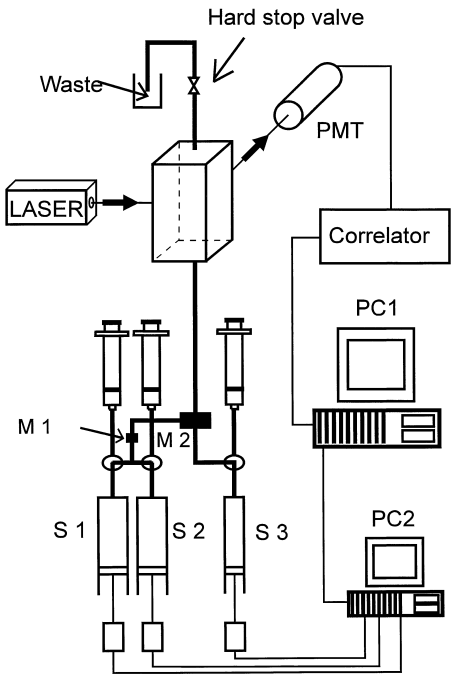

Fig. 1 Schematic diagram of the stopped-flow DLS apparatus. *S1–S3*: drive syringes, *M1, M2*: mixers, *PMT*: photomultiplier tube. Buffer and protein solutions are loaded via 100 nm pore size filters into syringes *S1*, *S2* and *S3*, respectively. *PC 1* controls the entire experiment and sends a trigger signal to *PC 2* before each shot. *PC 2* controls the stepping motors, which move the plungers of the syringes and activates the hard stop valve. In the present configuration, buffer and protein solutions are mixed by *M2*. The exit of *M2* leads to the observation cell. A photomultiplier tube detects the scattered light at 90 degree scattering angle

hard disc and provides the trigger signals for the stopped flow system. The second microcomputer is exclusively used to control the stopped-flow system. The photon detection system, which is mounted on a modified X-ray goniometer (Praezisionsmechanik Freiberg GmbH, FRG) in the case of routine applications, was linked to the optical-mechanical part of a SFM-3 stopped-flow system. It will be shown below that the demands concerning the dead-time of the instrument are not very high in the case of DLS investigations. Thus, we could test several configurations of a stopped-flow optical design for use in connection with dynamic and static light scattering. The SFM-3 system was found to fit the requirements rather well. The main problem in performing stopped-flow DLS, as compared to other experimental techniques, is that the mixture in the observation cell must be extremely free of dust and bubbles. Attempts to include a filter behind the mixer have been unsuccessful, because rapid mixing is inconsistent with subsequent filtration through small pore-size filters. As a consequence, the entire stopped-flow system had to be flushed with filtered water and solvent prior to the experiment. We have used 100 nm pore-size filters (Protein solutions, UK).

The SFM-3 stopped-flow module contains 3 syringes and 2 mixers. The pistons of the syringes are driven independently by stepping motors allowing suitable mixing ratios over a wide range. For our denaturant dilution experiments, we have used two 18 mL syringes (S1, S2) for the dilution buffer and a 4 mL syringe (S3) for the solution containing the protein and the denaturing agent. The contents of S1 and S2, merged together by mixer 1, were finally mixed by mixer 2 with the contents of S3. Mixer 2 was a specially designed mixer (HDS) for high density solutions. This type of mixer is absolutely necessary, if one is working with solutions containing high concentrations of GuHCl or urea. Syringes 1 and 2 containing the dilution buffer were manufactured of stainless steel. For Syringe 3, which contained the denaturant, we have used the Kel-F version. Stainless steel was found to be inappropriate for solutions containing high concentrations of GuHCl. We have used rectangular cells with 1.5 mm pathlength (FC-15) for the DLS measurements. With these cells the dead volume is about 30 μL (from the mixer to the centre of the cell) allowing dead times of the order of a few milliseconds. The outlet of the cuvette was controlled by a so-called hard-stop valve provided by the manufacturer of the SFM-3 module. This valve should avoid pressure-artifacts and therefore the appearance of bubbles, which is crucial for light scattering experiments. The syringes and the observation head, which surrounds the optical cell, can be thermostated by two independent temperature regulating circuits. They were held at the same temperature in the experiments described below.

The incident laser beam was focused into the scattering cell with a lens of 90 mm focal length. This lens is part of an adaptor kit for the application of SFM-3 in connection with a J-720 CD spectrometer (Jasco, Japan). This has the advantage that the SFM-3 module can be applied for both DLS and CD measurements with only minor modifications. The scattered light was detected at 90 degree scattering angle. The photon detection system was mounted on an optical rail which could be adjusted relative to the observation head of the SFM-3 module.

Data acquisition scheme and time-resolution

In previous work (Gast et al. 1992) we estimated the minimum data acquisition time $T_{A,min}$, which is needed to obtain the diffusion coefficient D or the corresponding Stokes radius

$$R = kT/6\pi\eta D \qquad (1)$$

with an accuracy of about ±1%. k is Boltzmann's constant, T the temperature in K and η the solvent viscosity. $T_{A,min}$ is of the order of a few seconds for a correlation time $\tau_c = (q^2D)^{-1} \cong 2 \cdot 10^{-5}$ s and a mean photon count rate per correlation time $r > 1$. q is the scattering vector. This is fulfilled in DLS experiments with proteins of molecular masses greater than 10,000 g/mol and at concentrations of the order of 1 mg/mL. It is noteworthy that the signal-to-noise ratio for a given T_A cannot be increased further by increasing the laser power once r is greater than 1, i.e. after the transition from the shot noise limit to the high pho-

ton flux limit. Such a short acquisition time is indeed sufficient, if the system under study is essentially monodisperse. In this case, simple procedures such as the cumulant method (Koppel 1972) can be used to calculate the diffusion coefficient from the autocorrelation function. A somewhat longer acquisition time is necessary if the solution contains an additional slower diffusing species, for example protein aggregates. An inverse Laplace transformation yielding a distribution function of diffusion coefficients must then be applied to obtain the diffusion coefficient of protein monomers. We have used the program CONTIN (Provencher 1982a, b) for this purpose. To get stable results, a typical acquisition time of 20 s is needed.

In the following we will take $T_{A,min} = 20$ s as a reasonable total acquisition time. If the kinetic process under study is slow, a time resolution $\Delta t = T_{A,min}$ is sufficient to measure the kinetics of the change in R after one mixing event (shot). In order to increase the time resolution, hence extending stopped-flow DLS to fast kinetics, N_S shots have to be performed to get an entire acquisition time $N_S \cdot \Delta t = T_{A,min}$. This suggests that the kinetic sample time $\Delta t = T_{A,min}/N_S$ could be shortened down to the millisecond time range if the number of shots (N_S) can be made sufficiently large. Limitations due to the available amount of protein will be discussed below. Another point to consider is the large amount of data to be handled in one kinetic experiment. Each kinetic data point, R(t), is primarily represented by a correlation function of 90 data points, which is stored as a data file on the hard disc. N_t correlation functions calculated during time-intervals Δt are needed to follow changes in R after each shot. Therefore, a total of $N_S \cdot N_t$ correlation functions has to be handled during one kinetic experiment. Typically, 10–50 shots with 50–100 kinetic data points have been performed in the experiments described below. In order to reduce the amount of data, all correlation functions belonging to the i-th kinetic data point could be averaged during the running experiment. However, this was not useful, because some of the individual correlation functions were distorted by unavoidable bubbles or by traces of dust or large protein aggregates. Thus, the entire set of correlation functions and the scattering intensities averaged over each acquisition time, $I(\Delta t)$, had to be stored on the hard disk. Afterwards, correlation functions distored by spurious scattering events were discarded. Typically, no more than 30% of the original data was lost in this way. The "lifetime" of bubbles within the scattering volume was mostly of the order of 1 s. Thus, it was useful to acquire data with relatively small Δt, to reject distorted intervals and to integrate afterwards over some Δt to get the desired time resolution. During the experiment the measurement control program displays an estimate of the diffusion coefficient calculated by the cumulant method and the average scattered intensity for each acquisition time after an individual shot. After the experiment, undistorted correlation functions of corresponding data acquisition intervals were averaged yielding a set of N_t correlation functions, from which the time dependence of D or R_S could be obtained. In the absence of large amounts of protein aggregates, this could be done by the cumulant method, but mostly the CONTIN program had to be applied to separate the Stokes radius of the protein molecules from that of aggregates.

Estimation of the Stokes radius of protein monomers in the presence of small oligomers

Evidence is accumulating that during refolding of some proteins small oligomers, mostly dimers or trimers, are formed transiently (Gast et al. 1992; Pecorari et al. 1996). Such behaviour is indicated by a decreasing integrated light scattering intensity in the course of refolding. One could expect that this effect may prevent any estimation of the radius of monomers. The radius of small oligomers is too close to that of monomers to be distinguished correctly by a DLS experiment and only an average value of monomers and oligomers is obtained as a stable result. In this section, we derive a correction formula, which can be applied to calculate an approximate Stokes radius of monomers even in this case if the following conditions are satisfied: the weight fraction of oligomers is small compared to that of the monomers and only monomers are present at the end of the experiment. It turns out that the particular type of the oligomers is of less importance than one would expect.

The following assumptions are made for these calculations. The total weight concentration c_0 is constant during the observation time and only the weight fraction of oligomers f is changing. The mass of the oligomers is $n \cdot M_0$, where M_0 is the molecular mass of monomers and $n = 2, 3, 4$ because we consider only dimers, trimers and tetramers. The protein molecules and oligomers are small compared to the wavelength of light resulting in a scattering form factor ≈ 1. This is fulfilled for proteins with $M_0 < 10^5$ g/mol. The diffusion coefficient of oligomers D_n is smaller by a factor $r_n = D_0/D_n$ than that of the monomers D_0.

The total relative excess scattered intensity $I = \text{const} \cdot (c_0 \cdot (1-f) \cdot M_0 + c_0 \cdot f \cdot n \cdot M_0)$ will be normalized to that of a solution containing only monomers $I_0 = \text{const} \cdot c_0 \cdot M_0$ yielding:

$$I_r = I/I_0 = 1 + f \cdot (n-1) \qquad (2)$$

const is a proportional constant regarding all experimental parameters and the optical properties of the solutions. Particularly it is assumed that the refraction index increment is the same for both monomers and oligomers.

If the solution contains different diffusing species, the apparent diffusion coefficient measured by DLS, D_{app}, is a z-average. For a solution containing monomers and one type of oligomer we get $D_{app} = (c_0 \cdot (1-f) \cdot M_0 \cdot D_0 + c_0 \cdot f \cdot n \cdot M_0 \cdot D_n)/(c_0 \cdot (1-f) \cdot M_0 + c_0 \cdot f \cdot n \cdot M_0)$. The ratio of D_{app} to D_0 can then be expressed in the form

$$D_{app}/D_0 = (1 + f \cdot (n/r_n - 1))/(1 + f \cdot (n-1)) \qquad (3)$$

The problem is to choose appropriate estimates of r_n for particular oligomeric structures because experimental values of r_n are not attainable. As an approximation we will take the values calculated for oligomers consisting of

spherical monomeric subunits (Garcia Bernal and Garcia de la Torre 1981), particularly $r_2 = 1.392$ for dimers, $r_3 = 1.67$ for trimers and $r_4 = 1.90$ for tetramers. The values for trimers and tetramers are averages over different configurations. Using the Stokes-Einstein equation (1) and (3), R_{app}/R_0 can be expressed as a function of the weight fraction of different types of oligomers (Fig. 2). Defined relations between f, I_r and R_{app}/R_0 exist for each kind of oligomers. Accordingly, one can substitute f in (3) by I_r from (2) yielding

$$R_0/R_{app} = (1 + (I_r - 1) \cdot (n/r_n - 1)/(n-1))/I_r \quad (4)$$

R_{app}/R_0 versus I_r is shown in Fig. 3 for dimers, trimers and tetramers. If the scattered intensity is changing during the experiment indicating the presence of transient oligomers (see Fig. 5), $R_0(t)$ can be estimated from the experimentally determined $I_r(t)$ and $R_{app}(t)$ using a correction $R_0(t) = R_{app}(t) \cdot (a + b/I_r)$. $R_0(t)$ is the approximate Stokes radius of monomers during refolding and the values of a and b can be calculated for particular oligomers. An interesting result is that this correction does not depend strongly on the particular type of the oligomer. This becomes evident from the relatively small differences between the curves for dimers, trimers and tetramers in Fig. 3. We have used a mean correction function which agrees with that for trimers. The quality of the correction depends mostly on how well the following assumptions are fulfilled: a) the size and shape of a monomer unit are the same when isolated and within the oligomeric structure, b) the theoretically derived r_n can be successfully applied to oligomers of globular proteins.

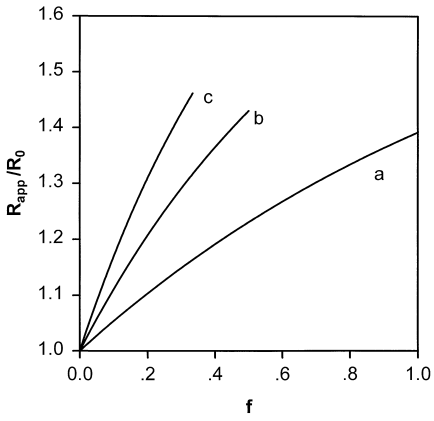

Fig. 2 Dependence of the relative apparent radii, R_{app}/R_0, on the weight fraction of oligomers, *a* dimers, *b* trimers, and *c* tetramers

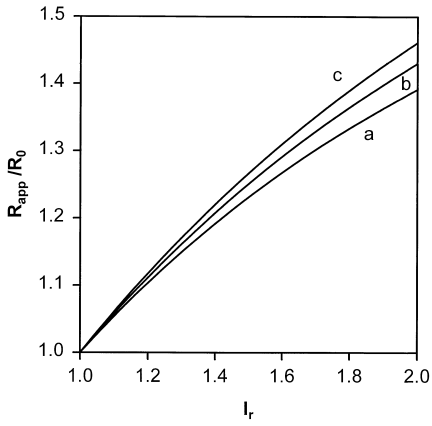

Fig. 3 Relative apparent radii, R_{app}/R_0, versus I_r for a mixture of monomers and different kind of oligomers, *a* dimers, *b* trimers, and *c* tetramers. $I_r = 2$ corresponds to 100% dimers, 50% trimers + 50% monomers, and 33% tetramers + 67% monomers, respectively

Results and discussion

Refolding investigations of RNase A

All folding investigations have been performed on RNase A with the four disulfide bridges intact. Prior to the experiments the protein was unfolded by the presence of 6 M GuHCl. The kinetics of refolding of RNase A as monitored by DLS after ten-fold dilution of a solution containing 6 M GuHCl with 10 mM sodium acetate buffer, pH 4 at 25 °C is shown in Fig. 4. The final protein concentration was 1.1 g/L. 50 shots have been carried out to accumulate the correlation functions, from which the radii in Fig. 4 have been obtained by using the program CONTIN. 40 data acquisitions were performed after each shot. The acquisition time defining the time resolution Δt was 4.2 s.

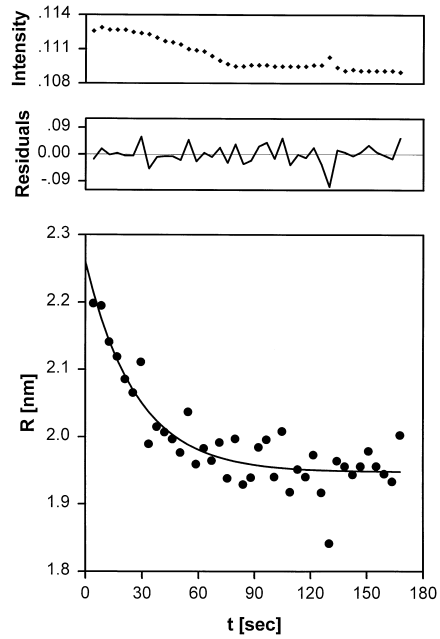

Fig. 4 Refolding kinetics of RNase A monitored by DLS at a time resolution of $\Delta t = 4.19$ s. RNase A was denatured by 6 M GuHCl. Refolding was carried out at T = 25 °C by 1:10 dilution in sodium acetate buffer, pH 4, with a final protein concentration of 1.1 g/L. The plotted curve is an average of 48 runs. The continuous line corresponds to a monoexponential fit with $R(0) = 2.26 \pm 0.04$ nm, $R(\infty) = 1.95 \pm 0.04$ nm, and $\tau = 27 \pm 5$ s

For the 50 shots accomplished during this experiment, the total acquisition time for each data point was 3.5 min. This is sufficiently long to obtain reliable results using the program CONTIN. Though the actual dead volume was only about 50 μL, a total volume of 600 μL per shot was used to achieve perfect washing of the cuvette between the shots. The integrated scattered intensity versus refolding time, which is also shown in Fig. 3, exhibits only a weak decrease of about 3%. This points to the absence of a significant association process during refolding. Nevertheless, we have applied the correction for oligomers to avoid any overestimation of the observed changes in R. The time-dependence of the Stokes radius was fitted by a single exponential function yielding a macroscopic decay time of 27 ± 5 s, a final Stokes radius of 1.95 ± 0.05 nm and a Stokes radius of 2.26 ± 0.05 nm at $t = 0$. The radius at $t = 0$ without correction for oligomers is only 2% larger. These values have to be compared with those measured under equilibrium conditions. In previous work (Nöppert et al. 1996), we measured $R = 1.90 \pm 0.04$ nm for the native protein and $R = 2.60 \pm 0.05$ nm for RNase a unfolded by 6 M GuHCl. In the present experiments the refolding buffer still contained a residual amount of GuHCl. Therefore, we have performed additional equilibrium investigations in the presence of 0.6 M and 1.2 M GuHCl in order to check whether there is a perceptible change in R with increasing GuHCl concentration before the unfolding transition, which proceeds between 2 and 4 M GuHCl. Extrapolating to zero protein concentration we found $R = 1.95$ nm in 0.6 M GuHCl and 1.99 nm in 1.2 M GuHCl, respectively. Only a weak concentration dependence was observed in both cases. Thus, the radii measured at 1.1 g/L in the kinetic experiment should be close to that at zero protein concentration.

Refolding experiments at 1.2 M final GuHCl concentration yielded $R = 2.29 \pm 0.05$ nm immediately after initiation of refolding, a final radius of 1.92 ± 0.05 nm and a remarkably larger decay time of 59 ± 9 s.

Refolding of chemically denatured, disulfide-intact RNase A has been studied by various physical and biochemical probes, mostly by tyrosine absorption and fluorescence (Schmid 1981, 1983; Lin and Brandts 1983; Houry et al. 1994; Dodge and Scheraga 1996). RNase A shows complex folding kinetics, which arises from a mixture of conformational isomers in the unfolded state. As a consequence, fast and slow folding molecules are observed (Garel and Baldwin 1973). Brandts et al. (1975) first proposed that non-native proline isomers give rise to the slow folding species. In the native state, Pro93 and Pro114 form cis X-Pro peptide bonds, while Pro42 and Pro117 have trans X-Pro bonds. The generally accepted model of folding comprises three different parallel pathways all leading to the native protein (Schmid 1983):

$U_F \rightarrow N$ (20%, fast),
$U_S^{II} \rightarrow N$ (60–70%, slow),
$U_S^{I} \rightarrow N$ (10–20%, very slow).

The rate constants depend on the particular folding conditions. In general the fast species folds within milliseconds, while slow folding takes seconds or minutes. We suggest that the observed slow changes in R_S (Fig. 1) can be related to the major slow folding species U_S^{II}. Folding is not blocked by incorrect proline isomers, but nativelike folded molecules may be formed before proline isomerization takes place (Schmid 1983). For the folding of U_S^{II} molecules a minimal sequential folding scheme has been proposed (Kim and Baldwin 1982, Schmid 1983): $U_S^{II} \rightarrow I_1 \rightarrow I_N \rightarrow N$. I_N is a nativelike intermediate which still contains a wrong proline isomer. Pulsed H/D-exchange studies (Udgaonkar and Baldwin 1990, 1995) have been performed to follow the protection of amide protons during folding. According to the measured protection factors, even the early intermediate I_1 should have a pronounced secondary structure and some fixed side chain structure. Protection is similar in I_1 and I_N, but I_1 and I_N differ from each other in the extent of buried tyrosine side chains. I_N is formed in less than 14 s under strong refolding conditions at 10 °C used in these studies.

In our refolding studies the Stokes radius drops rapidly from 2.6 nm down to 2.26 nm within the dead time of our experiment. This is partially due to the complete folding of U_F molecules and obviously mostly a consequence of a collapse of slow folding molecules. We have to take into account that the measured radii are averages resulting from unfolded molecules, partially folded molecules, but also from about 20% molecules, which are already in the native state. How can these data be related to the folding scheme described above? The time scales of folding cannot be compared directly owing to the differences in the folding conditions used in the various folding experiments. We suggest that the observed changes in R are due to a transition from unfolded molecules to essentially native molecules. Although the transition proceeds via structural intermediates, they do not accumulate significantly under the conditions used in our studies, i.e. T = 25 °C, in the presence of GuHCl and in the absence of stabilising salts (Schmid 1983). Therefore, the rate limiting step in the entire process $U_S^{II} \rightarrow I_1 \rightarrow I_N \rightarrow N$ is not the isomerization step $I_N \rightarrow N$, but rather the folding step $U_S^{II} \rightarrow I_1, I_N$.

This is supported by the concentration dependence of the time constant. 27 s and 59 s were measured at 0.6 and 1.2 M GuHCl, respectively. Proline isomerization was found to be independent of GuHCl concentration (Schmid 1981). R = 2.26 nm immediately after initiation of folding differs from an average radius of 2.47 nm, calculated for a mixture of 80% unfolded molecules (R = 2.6 mm) and 20% native, folded molecules (R = 1.95 nm). It should not be influenced by the radius of I_1 or I_N, because these intermediates are not well populated in our experiments. Thus, the question is whether there is a collapse of unfolded molecules preceding the formation of the specific intermediates (I_1, I_N) leading to a radius of 2.26 nm. This and other questions cannot be answered unequivocally by the presented data, which does, however, demonstrate the potentials of kinetic DLS investigations for protein folding studies. Particularly, further experiments at different folding conditions combined with kinetic CD measurements in the near- and far-UV regions are needed to give a more de-

tailed picture of conformational changes during folding of RNase A. Our results further demonstrate that RNase A folds to the native state without being trapped in misfolded, aggregated structures.

Refolding investigations of yeast PGK

The main purpose of these investigations is to compare folding of PGK starting either from the chemically denatured state or from the cold-denatured state as described previously (Gast et al. 1993). PGK was denatured by 2 M GuHCl where it is in a highly unfolded state (Missiakas et al. 1990, Damaschun et al. 1993). Refolding was initiated at 22.5 °C by diluting the solution to a final GuHCl concentration of 0.2 M with 20 mM sodium phosphate buffer, pH 6.5. 20 shots were performed. Primarily, 100 kinetic data points in time-intervals of 8.4 s were recorded. The time-dependences of the scattered intensity and of the Stokes radii after correction for transient association are shown in Fig. 5. The number of data points in the long-time region has been reduced by integrating over 2 or 3 data acquisition intervals. The continuous curve represents a single-exponential fit yielding a time-constant of 3.7 ± 0.2 min. The change in intensity is much larger than that in the case of RNase A and amounts of about 30%. Therefore, the correction for oligomers is essential. The apparent Stokes radius immediately after initiation of folding without correction is 4.73 nm, while the corrected value is 4.04 nm. The error in the experimentally determined (uncorrected) radius is about ±2%. For the corrected radius an additional systematic error cannot be excluded. The total effect of the correction is 17%. The final radius of 3.13 nm is only slightly larger than that obtained under equilibrium conditions, 3.01–3.09 nm (Gast et al. 1992, 1993). Probably, besides the transiently formed oligomers, trace amounts of stable oligomers are still present after folding is essentially completed. The appearance of transient oligomers is typical of folding of yeast PGK and was observed starting either from the GuHCl unfolded state (Pecorari et al. 1996) or from the cold denatured state (Gast et al. 1992). The correction for oligomers is of importance for the interpretation of the results. Without correction, the dimensions of the protein immediately after initiation of folding are quite similar to those of PGK in the cold-denatured state (Gast et al. 1993). With correction, a considerable collapse within the dead-time of the experiments has to be taken into consideration.

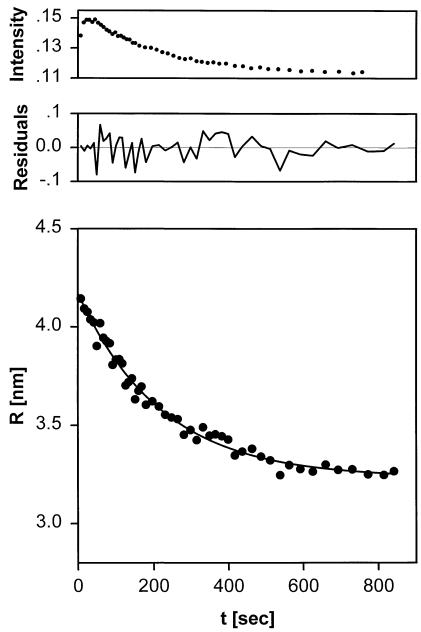

Fig. 5 Refolding kinetics of PGK monitored by DLS at a time resolution of $\Delta t = 8.4$ s at 22.5 °C. PGK was denatured by 2 M GuHCl. Refolding was initiated by a 1:10 dilution in sodium phosphate buffer, pH 6.5. The final protein concentration was 0.36 g/L. The plotted curve is an average of 20 runs. The continuous line corresponds to a monoexponential fit with $R(0) = 4.04 \pm 0.08$ nm, $R(\infty) = 3.13 \pm 0.06$ nm, and $\tau = 3.7 \pm 0.2$ min

To compare refolding either from the chemically or from the cold-denatured state directly, the refolding conditions should be identical. The optimum solvent conditions for cold denaturation are found for yeast PGK in the presence of 0.7 M GuHCl, where PGK has practically the native structure at room temperature (Damaschun et al. 1993). Therefore, refolding from the GuHCl denatured state was also investigated with 0.7 M final GuHCl concentration. Under these conditions, a considerable number of the protein molecules were trapped in irreversibly associated oligomeric structures. The average hydrodynamic radius was 3.9 nm. This means, that the pathways of folding are not the same when folding is started either from the cold denatured or the GuHCl denatured state. Accordingly, the results of spectroscopic refolding investigations (Missiakas et al. 1992) at GuHCl concentrations above 0.5 M must be considered with some care.

The origin of the slow folding phase, which is also observed in kinetic CD experiments (Missiakas et al. 1992; Gast et al. 1993) is still unclear. According to the results of double jump experiments (Missiakas et al. 1992) proline isomerization should not be involved in the folding process.

Conclusions

Our results have shown that it is possible to study protein folding reactions in terms of the Stokes radius with stopped-flow dynamic light scattering. By increasing the number of shots it should be possible to extend the time resolution into the subsecond region. In practice, the number of shots attainable will be mostly limited by the amount of protein available. For example, about 35 mg of protein was needed for the experiment with RNase A. Therefore, a time resolution of the order of one second is realistic without concern about the availability of a particular protein. Therefore, stopped-flow DLS is well suited to investigate late folding events or folding reactions, which are slowed down

by particular folding conditions. It is very important to measure both the time-autocorrelation function from which the Stokes radius is derived, and the time-integrated scattered intensity, which yields an independent estimate of the apparent molecular mass in one and the same experiment. This enables one to differentiate between true folding of individual molecules and cases where molecular association accompanies protein folding. Association is mostly favoured by partly folded or misfolded protein conformations and complicates the estimation of molecular dimensions. On the other hand the technique is a sensitive tool for the detection of the transient association of intermediates, e.g. the molten globule or the appearance of misfolded structures. Though the Stokes radius is not a geometric parameter, time-resolved measurements of its changes yield a good description of how compaction progresses when a protein folds. In addition, stopped-flow DLS can be used to check the feasibility of kinetic size measurements before expensive kinetic X-ray scattering experiments are performed.

Acknowledgements This work was supported by grants from the Deutsche Forschungsgemeinschaft Da292/1-3, and Ga530/1-1, by a grant from the Bundesministerium for Bildung, Wissenschaft, Forschung und Technologie 0310188A, and by a grant from the Fonds the Chemischen Industrie to G.D. We thank D. Otto for skillful technical assistance.

References

Baldwin RL (1996) On-pathway versus off-pathway intermediates. Folding and Design 1:R1–R8

Brandts JF, Halvorson HR, Brennan M (1975) Consideration that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. Biochemistry 14:4953–4963

Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. Proteins Struct Funct Genet 21:167–195

Buckler DR, Haas E, Scheraga HA (1995) Analysis of the structure of ribonuclease A in native and partially denatured states by time-resolved nonradiative dynamic excitation energy transfer between site-specific extrinsic probes. Biochemistry 34:15965–15978

Chu B, Ying QC, Grosberg AY (1995) Two stages kinetics of single chain collapse: polysterene in cyclohexane. Macromolecules 28:180–189

Damaschun G, Damaschun H, Gast K, Gernat C, Zirwer D (1991) Acid denatured apocytochrome c is a random coil: evidence from small-angle X-ray scattering and dynamic light scattering. Biochim Biophys Acta 1978:289–295

Damaschun G, Damaschun H, Gast K, Misselwitz R, Müller JJ, Pfeil W, Zirwer D (1993) Cold denaturation-induced conformational changes in phosphoglycerate kinase from yeast. Biochemistry 32:7739–7746

Dill KA (1985) Theory for the folding and stability of globular proteins. Biochemistry 24:1501–1509

Dill KA, Bromberg S, Yue K, Fiebig KM, Yee DP, Thomas PD, Chan HS (1995) Principles of protein folding – A perspective from simple exact models. Protein Sci 4:561–602

Dodge RW, Scheraga HA (1996) Folding and unfolding kinetics of the proline-to-alanine mutants of bovine pancreatic ribonuclease A. Biochemistry 35:1548–1559

Eliezer D, Chiba K, Tsuruta H, Doniach S, Hodgson KO, Kihara H (1993) Evidence of an associative intermediate on the myoglobin refolding pathway. Biophys J 65:912–917

Eliezer D, Jennings PA, Wright PE, Doniach S, Hodgson KO, Tsuruta H (1995) The radius of gyration of an apomyoglobin folding intermediate. Science 270:487–488

Elöve GA, Chaffotte AF, Roder H, Goldberg ME (1992) Early steps in cytochrome c folding probed by time-resolved circular dichroism and fluorescence spectroscopy. Biochemistry 31:6876–6883

Feng HP, Scherl DS, Widom J (1993) Lifetime of the histone octamer studied by continuous-flow quasielastic light scattering: test of a model for nucleosome transcription. Biochemistry 32:7824–7831

Feng HP, Widom J (1994) Kinetics of compaction during lysozyme refolding studies by continuous-flow quasielastic light scattering. Biochemistry 33:13382–13390

Flanagan JM, Kataoka M, Shortle D, Engelman DM (1992) Truncated staphylococcal nuclease is compact but disordered. Proc Natl Acad Sci USA 89:748–752

Garcia Bernal JM, Garcia de la Torre J (1981) Transport properties of oligomeric subunit structures. Biopolymers 20:129–139

Garel JR, Baldwin RL (1973) Both the fast and the slow refolding reactions of ribonuclease A yield native enzyme. Proc Nat Acad Sci USA 70:3347–3351

Gast K, Damaschun G, Misselwitz R, Zirwer D (1992) Application of dynamic light scattering to studies of protein folding kinetics. Eur Biophys J 21:357–362

Gast K, Damaschun G, Damaschun H, Misselwitz R, Zirwer D (1993) Cold denaturation of yeast phosphoglycerate kinase: Kinetics of changes in secondary structure and compactness on unfolding and refolding. Biochemistry 32:7747–7752

Gast K, Damaschun H, Misselwitz R, Müller-Frohne M, Zirwer D, Damaschun G (1994) Compactness of protein molten globules: temperature-induced structural changes of the apomyoglobin folding intermediate. Eur Biophys J 23:297–305

Houry WA, Rothwarf DM, Scheraga HA (1994) A very fast phase in the refolding of disulfide-intact ribonuclease A: implication for the refolding and unfolding pathways. Biochemistry 33:2516–2530

Karplus M, Weaver DL (1994) Protein folding dynamics: the diffusion-collision model and experimental data. Protein Sci 3:650–668

Kiefhaber T (1995) Kinetic traps in lysozyme folding. Proc Natl Acad Sci USA 92:9029–9033

Kim PS, Baldwin RL (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. Annu Rev Biochem 51:459–489

Kim PS, Baldwin RL (1990) Intermediates in the folding reactions of small proteins. Annu Rev Biochem 59:631–660

Koppel DE (1972) Analysis of macromolecular polydispersity in intensity correlation spectroscopy: the method of cumulants. J Chem Phys 57:4814–4820

Kuwajima K, Yamaya H, Miwa S, Sugai S, Nagamura T (1987) Rapid formation of secondary structure framework in protein folding studied by stopped-flow CD. FEBS Lett 221:115–118

Kuwajima K (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular protein structure. Proteins Struct Funct Genet 6:87–103

Lin LN, Brandts JF (1983) Mechanisms for the unfolding and refolding of ribonuclease A. Kinetic studies utilizing spectroscopic methods. Biochemistry 22:564–573

Missiakas D, Betton JM, Minard P, Yon J (1990) Unfolding-refolding of the domains in yeast phosphoglycerate kinase: comparison with the isolated domains. Biochemistry 29:8683–8689

Missiakas D, Betton JM, Chaffotte A, Minard P, Yon JM (1992) Kinetic studies of the refolding of yeast phosphoglycerate kinase: comparison with the isolated engineered domains. Protein Sci 1:1485–1493

Nöppert A, Gast K, Müller-Frohne M, Zirwer D, Damaschun G (1996) Reduced-denatured ribonuclease A is not in a compact state. FEBS Lett 380:179–182

Pace CN, Shirley BA, Thomson JA (1989) Measuring the conformational stability of a protein. In: Creighton TE (ed) Protein structure. A practical approach. IRL Press, Oxford, pp 311–330

Pecorari F, Minard P, Desmadril M, Yon JM (1996) Occurrence of transient multimeric species during refolding of a monomeric protein. J Biol Chem 271:5270–5276

Phillips JC, LeGrand AD, Lehnert WF (1988) Protein folding observd by time-resolved synchrotron X-ray scattering. A feasibility study. Biophys J 53:461–464

Provencher SW (1982a) A constained regularization method for inverting data presented by linear algebraic or integral equations. Comp Phys Commun 27:213–227

Provencher SW (1982b) CONTIN: A general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. Comp Phys Commun 27:229–242

Roder H, Elöve GA, Englander SW (1988) Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. Nature 335:700–704

Schindler T, Herrler M, Marahiel MA, Schmid FX (1995) Extremely rapid folding in the absence of intermediates. Nature Struct Biol 2:663–673

Schmid FX (1981) A native-like intermediate on the ribonuclease A folding pathway. Eur J Biochem 114:105–109

Schmid FX (1983) Mechanism of folding of ribonuclease A. Slow folding is a sequential reaction via structural intermediates. Biochemistry 22:4690–4696

Shalongo W, Jagannadham M, Stellwagen E (1993) Kinetic analysis of the hydrodynamic transition accompanying protein folding using size exclusion chromatography. 2. Comparison of spectral and chromatographic kinetic analysis. Biopolymers 33:135–145

Sosnick TR, Mayne L, Hiller R, Englander SW (1994) The barriers in protein folding. Nature Struct Biol 1:149–156

Udgaonkar JB, Baldwin RL (1990) Early folding intermediates of ribonuclease A. Proc Natl Acad Sci USA 87:8197–8201

Udgaonkar JB, Baldwin RL (1995) Nature of early folding intermediates of ribonuclease A. Biochemistry 34:4088–4096

Uversky VN (1993) Use of fast size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. Biochemistry 32:13288–13298

Wolynes PG, Onuchic JN, Thirumalai D (1995) Navigating the folding routes. Science 267:1619–1620

Zerovnik E, Pain RH (1987) Refolding of β-lactamase followed by f.p.l.c. Protein Engineering 1:248