PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Studying submicrosecond protein folding kinetics using a photolabile caging strategy and time-resolved photoacoustic calorimetry

Hsin-Liang Chen,[1,2] Jack C.-C. Hsu,[1,3] Man Hoang Viet,[4] Mai Suan Li,[4] Chin-Kun Hu,[5,6] Chia-Hsun Liu,[1] Frederick Y. Luh,[1] Silvia S.-W. Chen,[1] Evan S.-H. Chang,[1] Andrew H.-J. Wang,[1,3] Min-Feng Hsu,[1,7] Wunshain Fann,[2] and Rita P.-Y. Chen[1,3]*

[1] Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan
[2] Institute of Atomic and Molecular Sciences, Academia Sinica, Taipei 10617, Taiwan
[3] Institute of Biochemical Sciences, National Taiwan University, Taipei 10617, Taiwan
[4] Institute of Physics, Polish Academy of Sciences, 02-668 Warsaw, Poland
[5] Institute of Physics, Academia Sinica, Taipei 11529, Taiwan
[6] Center for Nonlinear and Complex Systems and Department of Physics, Chung-Yuan Christian University, Chungli 32023, Taiwan
[7] Core Facility for Protein Production and X-Ray Structural Analysis, Academia Sinica, Taipei 11529, Taiwan

ABSTRACT

Kinetic measurement of protein folding is limited by the method used to trigger folding. Traditional methods, such as stopped flow, have a long mixing dead time and cannot be used to monitor fast folding processes. Here, we report a compound, 4-(bromomethyl)-6,7-dimethoxycoumarin, that can be used as a "photolabile cage" to study the early stages of protein folding. The folding process of a protein, RD1, including kinetics, enthalpy, and volume change, was studied by the combined use of a phototriggered caging strategy and time-resolved photoacoustic calorimetry. The cage caused unfolding of the photolabile protein, and then a pulse UV laser ($\sim 10^{-9}$ s) was used to break the cage, leaving the protein free to refold and allowing the resolving of two folding events on a nanosecond time scale. This strategy is especially good for monitoring fast folding proteins that cannot be studied by traditional methods.

Proteins 2010; 00:000–000.

Key words: protein folding; RD1; cage; photolabile; photoacoustic calorimetry.

INTRODUCTION

To understand the intrinsic principles of protein folding, the events in the folding process have to be systematically explored from small to large time scales. Traditional methods for triggering protein folding using stopped-flow or quenched-flow instruments are often too slow to measure the earliest kinetic processes of refolding. In the last 2 decades, new experimental methods permitting the observation of protein folding and unfolding on the previously inaccessible nanosecond–microsecond time scale have been developed.[1–4] The capillary mixing method (continuous flow) allows observation periods of 50 μs to about 1 ms, depending on the length of the flow cell and the flow rate.[5–9] To further improve the time resolution to the nanosecond range, initiation of protein folding by light was developed.[10] Jones *et al.*[11] used laser-triggered nanosecond photodissociation of the heme–CO complex to study the refolding of reduced cytochrome *c*, whereas Pascher *et al.*[12] described the use of a photoinduced electron transfer reaction to initiate the folding of reduced cytochrome *c*, taking advantage of the difference in stability of the oxidized and reduced protein. Laser temperature jump is currently receiving more attention and has been used to observe the formation of α-helices[13,14] and β-hairpins.[15,16]

Recently, a method combining a photolabile caging strategy and laser flash photolysis was developed to study fast peptide folding. A photolabile linker, bromoacetyl-carboxymethoxybenzoin (BrAcCMB), is used to cyclize a short peptide and con-

Additional Supporting Information may be found in the online version of this article.
Abbreviations: BrAcCMB, bromoacetyl-carboxymethoxybenzoin; BrDMC, 4-(bromomethyl)-6,7-dimethoxycoumarin; DMC, 6,7-dimethoxy-coumarin-4-ylmethyl group; DMNB, 4,5-dimethoxy-2-nitrobenzyl group; DMNBB, 4,5-dimethoxy-2-nitrobenzyl bromide; P8, an eight-residue control peptide (sequence DAEFRHDC); P8-DMC, P8 with a DMC group on the side chain of Cys; RD1-A7C, protein RD1 with an Ala-7 → Cys mutation; RD1-A7C-DMC, RD1-A7C with a DMC group on the side chain of Cys-7.

Grant sponsor: National Science Council, Taiwan, R.O.C.; Grant numbers: NSC-95-2113-M-001-049, NSC-97-2113-M-001-014-MY2, NSC-96-2911-M-001-003-MY3, NSC-98-2911-I-001-028; Grant sponsor: Academia Sinica; Grant numbers: AS-94-TP-A01, AS-95-TP-A07; Grant sponsor: Ministry of Science and Informatics in Poland; Grant number: 202-204-234; Grant sponsor: National Center for Theoretical Sciences in Taiwan.
Hsin-Liang Chen and Jack C.-C. Hsu contributed equally to this work.
*Correspondence to: Rita P.-Y. Chen, Institute of Biological Chemistry, Academia Sinica, No. 128, Sec. 2, Academia Rd, Nankang, Taipei 11529, Taiwan. E-mail: pyc@gate.sinica.edu.tw.
Received 7 May 2010; Revised 24 June 2010; Accepted 2 July 2010
Published online 23 July 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.22823

strains the structure of the "caged" peptide in a well-defined non-native conformational state.[17,18] To initiate peptide refolding, the photolabile linker is cleaved by irradiation with pulses of UV light from an Nd:YAG laser. The linker is cleaved within 10^{-9} s, giving a time resolution adequate for observing the rapid refolding of the photocleaved, linearized peptide back toward its native structure.

This method is excellent and has been used to study various small structural motifs, such as α-helix, helix-turn-helix, β-hairpin, and three-stranded β-sheet.[17–19] However, the photolabile linker BrAcCMB has the drawback of low chemical stability. It is easily hydrolyzed during the coupling reaction with a polypeptide chain under basic conditions. Moreover, it is difficult to cyclize a longer polypeptide chain, which limits the application of this method to proteins. To use the photolabile caging strategy to study the fast folding of proteins rather than small peptides, a more stable cage has to be used. Hirota *et al.*[20] used 4,5-dimethoxy-2-nitrobenzyl bromide (DMNBB) as a new cage compound to unfold apoplastocyanin (apoPC) and analyzed the folding kinetics after cage photolysis using transient grating as the detection method.

The next challenge for the photolabile caging strategy is the selection of the caged protein and the cage-labeling site. To study protein refolding, we aimed to find a caging site at which interference by the bulky photolabile cage coupled to a side chain of the protein would lead to unfolding of the protein. Because the cage reacts with the sulfhydryl group of cysteine, the target protein must have only one reactive cysteine. In the case of proteins without any cysteines in the native sequence, one can easily be introduced by site-directed mutagenesis; however, the introduced cysteine must not interfere with the folding of the protein after the cage is released. In this study, our target protein was a small protein, RD1, a Type III antifreeze protein from the Antarctic eel pout, *Lycodichthys dearborni*.[21] RD1 contains only 64 amino acids and has no cysteine. Its crystal structure has been determined at a resolution of 0.62 Å.[21] Its structure is compact and consists of many short pieces of secondary structures [Fig. 1(A)]. It is composed of two internal tandem motifs arranged about a pseudo-dyad symmetry. Each motif consists of four short β-strands and a one-turn 3_{10} helix, and the two motifs are connected by a 15-residue loop, which contains a two-turn 3_{10} helix. Sixty-nine percent of the amino acid residues are involved in the secondary structure. The protein fold of RD1 has been found in other proteins, such as the C-terminal domain of human and bacterial sialic acid synthase.[22] This "antifreeze" domain can have other functions in proteins. For example, it is involved in the formation of the quaternary structure of *E. coli* sialic acid synthase, and deletion of this domain results in loss of enzyme activity.[23] As RD1 lacks an intrinsic fluorescence probe and any clear spectroscopic differences between the folded and unfolded states, its folding cannot be monitored by general detection methods.

In this article, we used 4-(bromomethyl)-6,7-dimethoxycoumarin (BrDMC) as the photolabile cage

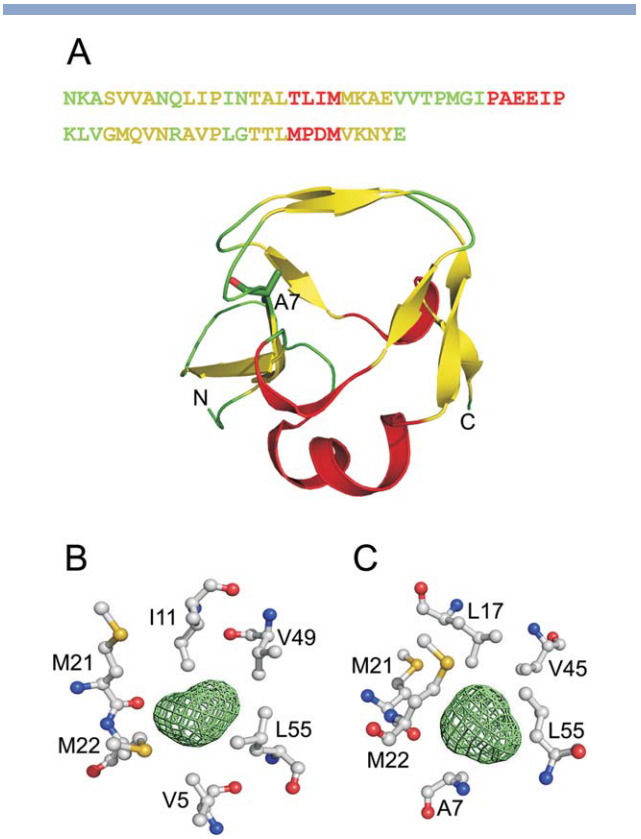


Figure 1

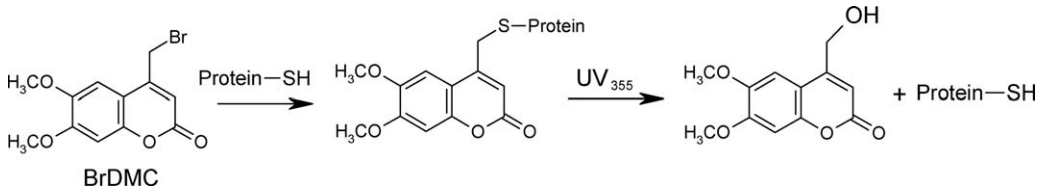
(**A**) Sequence and overall structure of the RD1 molecule (PDB ID: 1UCS). The structure is shown in ribbon format. The colors indicate different secondary structures, with helix, β-strand, and loop being colored red, yellow, or green, respectively. (**B**, **C**) The cavity (shown as green mesh) in the interior of the RD1 molecule. (B) and (C) are orthogonal to each other. The nine residues surrounding the cavity are labeled.

(Scheme 1). Its chemical stability is better than that of BrAcCMB, and its photolysis rate is faster than that of the nitrobenzyl derivative.[24] Photoacoustic calorimetry (PAC) was used to observe the refolding process on a nanosecond time scale.[25] This strategy is especially good for monitoring the folding of fast folding proteins that cannot be explored by traditional methods, but does not allow the detailed monitoring of folding events at the atomic level. This problem can be solved by using all-atom simulations with explicit solvent. Using Gromos force field 43a1, the simulation result was in reasonable agreement with the experimental result and helped us explain the experimental data.[26]

MATERIALS AND METHODS

Protein expression and purification

The RD1 gene was cloned into the pET11a plasmid. Ala-7 was mutated to Cys by site-directed mutagenesis. The mutant protein RD1-A7C was expressed in *E. coli*



Scheme 1

Reaction scheme for photolabile caging.

BL21 (DE3) cells after induction with 0.1 m*M* isopropyl-β-D-thiogalactopyranoside. After induction, the cells in LB broth containing 100 μg/mL of ampicillin were incubated at 16°C for 3 days with vigorous shaking and then were centrifuged at 7000*g* at 4°C for 20 min. The cells collected from 4 L of broth were resuspended in 150 mL of freeze-thaw buffer (100 m*M* Tris-HCl, 1 m*M* EDTA, 5 m*M* DTT, and 1 m*M* benzamidine HCl, pH 8) and the suspension frozen in liquid nitrogen, then thawed in a 37°C water bath for 15 min, and this freeze-thaw cycle being repeated a further twice. Soluble proteins were separated from cell debris by centrifugation at 18,000*g* at 4°C for 25 min and then were precipitated as the 20–80% ammonium sulfate cut. After dialysis against 20 m*M* Tris-HCl, 1 m*M* EDTA, 1 m*M* DTT, and 1 m*M* benzamidine HCl, pH 8, RD1-A7C was purified by anion exchange chromatography on a HiPrep™ 16/10 Q FF ion-exchange column at a flow rate of 3 mL/min and elution with a linear 0–1*M* NaCl gradient in 20 m*M* Tris-HCl, 1 m*M* EDTA, 1 m*M* DTT, and 1 m*M* benzamidine HCl, pH 8.5. The eluted fractions were monitored with a UV detector and analyzed by SDS-PAGE, and fractions containing RD-A7C were pooled and concentrated by ultrafiltration at 2100*g* at 4°C on an Amicon YM-3 Centriplus® filter membrane to a final volume of 3 mL, which was then applied to a Superdex-75 gel filtration column and eluted with 50 m*M* ammonium bicarbonate, 2 m*M* DTT, pH 7.6, at a flow rate of 1 mL/min. The purity of the protein was assessed using SDS-PAGE and MALDI (matrix-assisted laser desorption ionization) mass spectroscopy, and its concentration was determined by amino acid analysis.

Peptide synthesis

The octapeptide DAEFRHDC, denoted as P8, was synthesized by the batchwise Fmoc (fluorenylmethoxycarbonyl)-polyamide method on a PS3 peptide synthesizer (Rainin). Preloaded Fmoc-Cys-Wang resin was purchased from Anaspec and used as the solid support in the synthesis. Fmoc-amino acid derivatives (0.4 mmol) were coupled to 0.1 mmol of resin using 0.4 mmol of benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate in dimethylformamide (DMF) containing 4.45% (v/v) of *N*-methyl morpholine. Fmoc cleavage was performed using 20% (v/v) piperidine in DMF. The peptide was cleaved from the resin by stirring at room temperature for 1–2 h with a mixture of 9.4 mL of trifluoroacetic acid, 0.1 mL of triisopropylsilane, 0.25 mL of water, and 0.25 mL of ethanedithiol and precipitated with three volumes of ice-cold methyl *t*-butyl ether by centrifugation at 2000*g* for 10 min at 4°C, and then the pellet was washed twice with methyl *t*-butyl ether and dried under vacuum. The resulting white powder was purified by reverse-phase HPLC using a Vydac C18 column (10 mm × 250 mm) and acetonitrile–water mixtures containing 0.1% trifluoroacetic acid (v/v). The final products were analyzed by MALDI mass spectroscopy. Fractions containing the desired product were lyophilized and stored at −20°C.

Synthesis of the caged protein and peptide

Purified RD1-A7C (0.46 m*M*) or peptide P8 (0.8 m*M*) was dissolved in 2*M* GdnHCl, 100 m*M* Tris-HCl, and 1 m*M* tris(2-carboxyethyl)phosphine at pH 9, and BrDMC was dissolved in DMF to a final concentration of 2 m*M*. Equal volumes of the protein (or peptide) solution and the BrDMC stock solution were mixed with gentle shaking at 37°C for 1 h in the dark, and then the caged protein or peptide was purified by reverse-phase HPLC and identified by mass spectroscopy.

Photolysis of the caged protein and peptide

To carry out photolysis of the caged protein or peptide, samples were placed in a quartz cuvette and irradiated in a "merry-go-round" photoreactor PR-2000 (Pan-Chum, Taiwan) equipped with 16 UV lamps at 352 nm.

Mass spectroscopy

The molecular masses of all peptides and proteins were determined by matrix-assisted laser desorption/ionization (MALDI) mass spectroscopy (micro MX™, Waters, USA). The samples were prepared in 1% acetic acid/50% acetonitrile.

Nuclear magnetic resonance spectroscopy

All nuclear magnetic resonance (NMR) spectra were recorded on a Bruker AVANCE 600 AV NMR spectrometer. RD1-A7C (1 m*M*) or RD1-A7C-DMC (0.5 m*M*) was dissolved in water (H_2O/D_2O = 9/1; pH 3.3) alone or containing 1.2*M* GdnHCl. The quoted pH values were those read on the pH meter, with no correction for the D/H isotope effect. Two-dimensional total correlated spectroscopy (2D-TOCSY) spectra were collected using standard phase-cycling sequences at 298 K; a 60 ms mixing time was used in this experiment. Two-dimensional spectra were acquired with 2K data points in the direct dimension and 512 increments in the indirect dimension. Data were processed using XWINNMR software (Bruker). The shifted square sine bell window functions in both dimensions were applied to all spectra.

Photoacoustic calorimetry

The light source, operated at 355 nm, was the third harmonic of a Q-switch Nd-YAG laser (New Wave Research, Fremont, USA). The laser pulse width was ~5 ns, and the repetition rate was 2 Hz. The typical power used in the experiment was a few microjoules. PAC signals were measured between 15 and 35°C, and the temperature variation during an experiment was less than 0.5°C. The data were analyzed by comparison with those for the calorimetric reference compound, potassium dichromate ($K_2Cr_2O_7$), which releases all the energy absorbed upon photoexcitation as heat with a quantum efficiency of 1.0.[27] The optical density of the reference solution was adjusted to match that of the sample solution (range 0.2–0.3). The photoacoustic pressure waves were detected using a 1 MHz bandwidth microphone (GE Panametrics V-103, USA), and the PAC signals from 300 shots were averaged and recorded using a digital oscilloscope. The instrument response time was 20 ns.

The PAC signal is a convolution of the instrumental response function ($R(t)$) and a function representing the time evolution ($S(t)$) of the heat upon laser excitation. The form describing the signal can be written as $O(t) = R(t)^{*}S(t)$. $S(t)$ is often written as the summation of the single exponential terms [see Eq. (1)].

$$S(t) = \sum_i \frac{\phi_i}{\tau_i} e^{-\frac{t}{\tau_i}}, \qquad (1)$$

where ϕ_i and τ_i are the respective amplitude and decay time for the ith component in the sum of the exponentials. Because the heat release of the reference compound is much faster than the instrument response time, the signal for the reference compound was used as the instrument response time function $R(t)$. $O(t)$, the total signal from the sample, was recorded using a digital oscilloscope. The software program "Sound analysis 32" (Quantum Northwest, USA) was used to fit the data. The program is based on the least squares iterative reconvolution method.

Simulation method

All-atom simulations with explicit water were performed using the GROMACS program (version 4.0.3), and the interactions were described using the GROMOS 43a1 force field.[28] The SPC water model was used to describe the solvent. The protein was placed in an octahedral box with edges of 6.2, 6.2, and 4.4 nm with ≈5000 water molecules. The equations of motion were integrated using a leap-frog algorithm with a time step of 2 fs. The Linear Constraint Solver program was used to constrain bond lengths with a relative geometric tolerance of 10^{-4}.[29] Long-range electrostatic interactions were calculated using the particle-mesh Ewald algorithm.[30] The nonbonded interaction pair list was updated every 10 fs, using a cutoff of 1.2 nm. The starting structures were generated from the crystal structure of the protein (PDB ID: 1UCS), adjusted for an Ala-7 to Cys substitution. The native conformer was gradually heated from 300 to 450 K. At 450 K, a molecular dynamics (MD) calculation was performed until the total number of native contacts in the protein backbone became zero. The last extended conformation was saved as the starting configuration for the collapse study. To study the collapse process, MD simulation was carried out using values of T = 300 K for 200 ns, starting from the randomly selected denatured conformer, in four independent runs. For analysis, conformations were recorded every 10 ps. A hydrogen bond (HB) was considered as formed if the distance between the donor D (atom N) and acceptor A (atom O) was ≤3.5 Å, and the angle D-H-A was ≥145°.

RESULTS

Design of the cage-labeling site

In the RD1 molecule, there is an interior cavity that is not occupied by protein or solvent atoms. It has a volume of 45 Å^3 and is surrounded by nine conserved nonpolar residues, Val-5, Ala-7, Ile-11, Leu-17, Met-21, Met-22, Val-45, Val-49, and Leu-55, which are situated on different β-strands [Fig. 1(B,C)]. Because of this hydrophobic cavity in the interior of RD1, our experimental design was to replace one of the residues facing the interior cavity with a cysteine residue by site-directed mutagenesis for cage labeling. Ala-7 was the prime choice because its size is closest to that of Cys, and the substitution might not affect the refolded structure after the cage was released. A second reason was that Ala is the smallest residue facing the cavity, and increasing the size at this position is more likely to unfold the protein. The mutant

protein, denoted RD1-A7C, was expressed, purified, and reacted with BrDMC. It was expected that the DMC group would be bulky enough to push away the neighboring residues by hydrophobic repulsion and cause unfolding of RD1-A7C.

Synthesis and photolysis of RD1-A7C-DMC

BrDMC was efficiently coupled to Cys-7 of RD1-A7C in the presence of 2*M* GdnHCl. The caged product, denoted as RD1-A7C-DMC, was purified by HPLC and identified by mass spectroscopy [Fig. 2(A)]. After addition of the cage, RD1-A7C-DMC became less soluble and tended to aggregate, which implied that the protein was unfolded and that the hydrophobic interior became exposed and formed intermolecular associations. Because of this, RD1-A7C-DMC was dissolved in 1.2*M* GdnHCl. In the absorption spectrum of RD1-A7C-DMC [Fig. 2(B)], the absorption band at 350 nm comes from the DMC moiety. This band is distant from the protein absorption band and thus irradiation at 355 nm can photocleave the cage without affecting the protein moiety. When RD1-A7C-DMC was irradiated in a "merry-go-round" photoreactor, an absorbance decrease at 355 nm and absorbance increase at 280 nm after photolysis were clearly seen by comparing the absorption spectrum of the black line (before photolysis) and the red line (after photolysis) [Fig. 2(B)]. The mass result showed that RD1-A7C was reformed after irradiation.

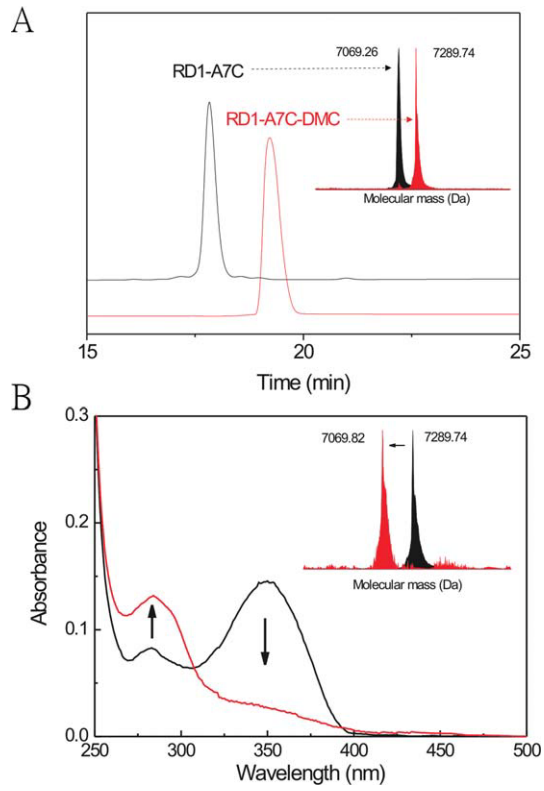


Figure 2

(**A**) HPLC chromatographs of RD1-A7C (black) and RD1-A7C-DMC (red) using a Vydac C18 column. Elution was performed by linearly increasing the percentage of solvent B from 30 to 60% in 30 min (solvent A: 5% acetonitrile and 0.1% TFA in water; solvent B: 0.1% TFA in acetonitrile). Inset: Mass spectra of RD1-A7C (black) and RD1-A7C-DMC (red). The theoretical mass of RD1-A7C is 7069.6 Da. The theoretical mass of RD1-A7C-DMC was calculated as 7069.6 + (299.1 − 81) = 7287.7. (**B**) Absorption spectra of RD1-A7C-DMC before (black) and after (red) 10-min irradiation at 352 nm in a photoreactor. RD1-A7C-DMC was dissolved in 20 m*M* HEPES buffer (pH 8) and 1.2*M* GdnHCl. Inset: Mass spectra of RD1-A7C-DMC before (black) and after (red) photolysis.

For the caging strategy to be used successfully in studying protein folding, addition of the cage should unfold the protein. The structures of RD1-A7C and the cage-modified RD1-A7C-DMC were compared using NMR spectroscopy (Fig. 3). The first and second traces in Figure 3(A) show the 1D-NMR spectrum of RD1-A7C and RD1-A7C-DMC in water, respectively. When a protein is unfolded, the upfield shifts of the methyl and methylene protons of the hydrophobic side chains are no longer seen, and comparison of the spectra clearly showed that addition of a DMC group to Cys-7 unfolded the protein. In Figure 3(B), the fingerprint regions (composed of H^{α} and H^{N} crosspeaks) of the 2D TOCSY spectra for RD1-A7C-DMC (red) and RD1-A7C (black) are superimposed. The H^{α} and H^{N} crosspeaks of RD1A7C-DMC were poorly dispersed, as in a random coil conformation. The signal was very weak, probably because aggregation increases the T_1 relaxation. Based on these results, it was concluded that the addition of BrDMC substantially unfolded the protein. To avoid aggregation, 1.2*M* GdnHCl was used to dissolve RD1-A7C-DMC. The 1D-NMR spectrum of RD1-A7C in 1.2*M* GdnHCl [third spectrum in Fig. 3(A)] was close to that in water [first spectrum in Fig. 3(A)], suggesting that the protein structure remained native like in the presence of 1.2*M* GdnHCl. This spectrum was also similar to that of RD1-A7C-DMC after photolysis [fourth spectrum in Fig. 3(A)], suggesting that the caged protein folds back to its native structure after the cage is released from the protein.

Refolding kinetics monitored by PAC

The folding of RD1-A7C-DMC in 20 m*M* HEPES buffer, pH 8, containing 1.2*M* GdnHCl after laser flash photolysis was monitored by PAC. The cage was broken by photoirradiation and released into the solvent; at the same time, the protein started to fold from the unfolded state to the native state and the process was monitored by PAC. A synthetic eight amino acid random-coil peptide (sequence DAEFRHDC), denoted P8, was also modified with BrDMC for comparison, and the modified peptide was named P8-DMC.

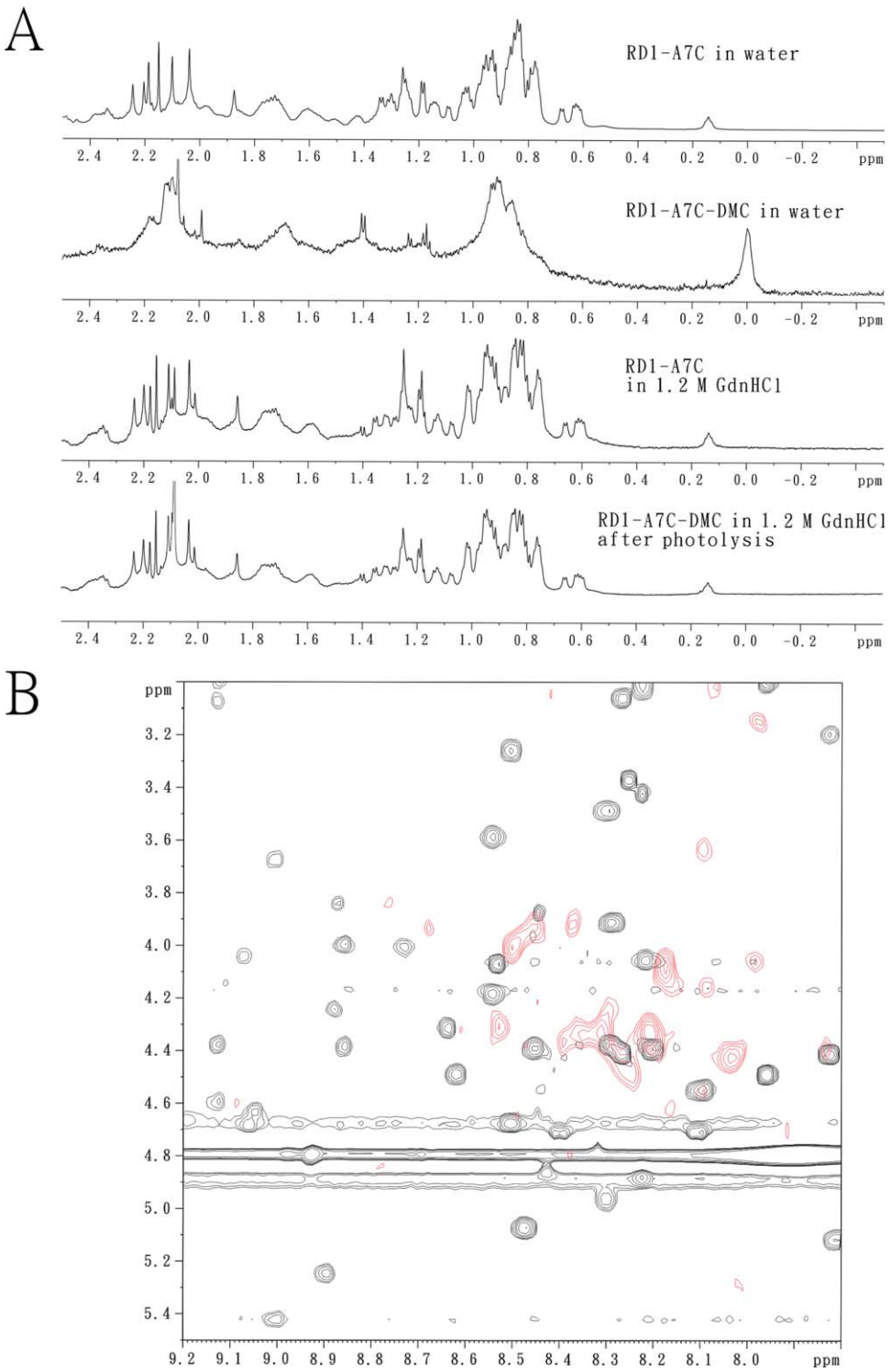


Figure 3

(**A**) Comparison of the 1D-NMR spectra of RD1-A7C in water, RD1-A7C-DMC in water, RD1-A7C in 1.2*M* GdnHCl, and RD1-A7C-DMC in 1.2*M* GdnHCl after photolysis. (**B**) Fingerprint regions of the 2D-TOCSY spectra of RD1-A7C (black) and RD1-A7C-DMC (red) in water. All samples contained 10% D_2O. The spectra were recorded at 298 K.

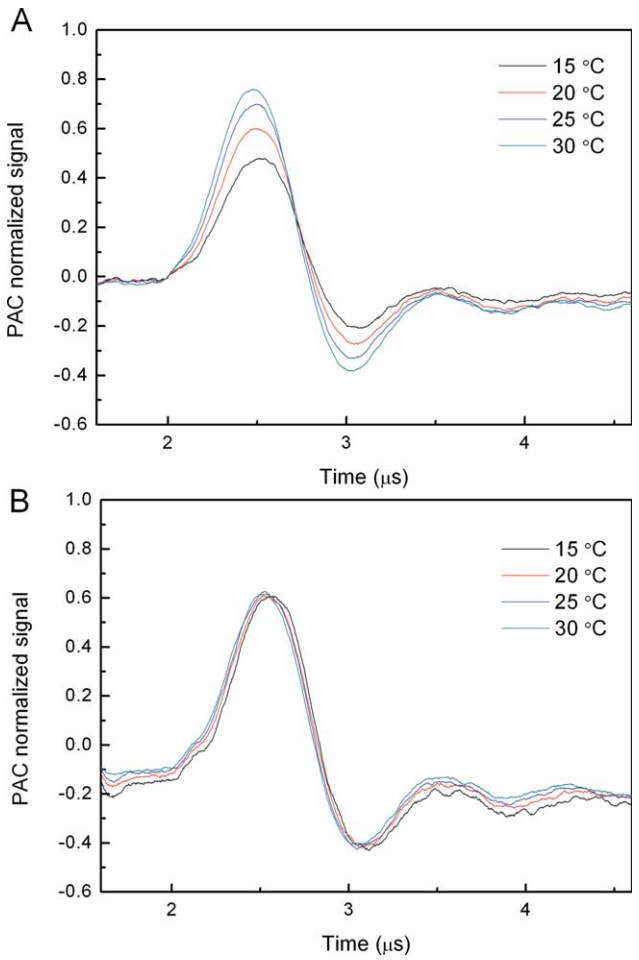


Figure 4

The photoacoustic waves generated after laser irradiation of RD1-A7C-DMC (**A**) or P8-DMC (**B**) at 15, 20, 25, and 30°C. The amplitude was normalized to the reference amplitude at the same temperature.

Figure 4 shows normalized PAC traces for RD1-A7C-DMC [Fig. 4(A)] and P8-DMC [Fig. 4(B)] recorded at different temperatures. The PAC amplitude, which was normalized with respect to the reference compound $K_2Cr_2O_7$, for RD1-A7C-DMC varied depending on the temperature, whereas that for P8-DMC remained constant at different temperatures. Because P8-DMC contained only eight amino acids and the cage was conjugated to the C-terminus of the peptide, cage release should only affect heat and there should be no volume change in the peptide that would contribute to the observed PAC signal. The difference in amplitude between P8-DMC and RD1-A7C-DMC implies that protein folding occurs after cage release from RD1-A7C-DMC.

The PAC signals were resolved using Sound Analysis (Version 1.50D). The PAC signal for RD1-A7C-DMC was well fitted by a function with two exponential decays. The two components derived from the fitting represented two events in the refolding of RD1-A7C. At 15°C, the PAC signal was fitted with amplitude $\phi_1 = \sim 0.48$ and $\tau_1 < 20$ ns for the fast component and $\phi_2 = \sim 0.05$ and $\tau_2 = 520 \pm 100$ ns for the slow component (see Supporting Information). In contrast, the PAC signal for P8-DMC at 15°C was well fitted by a function with one exponential decay, with ϕ_1 of ~0.60 and $\tau_1 < 20$ ns. The fitting results for RD1-A7C-DMC and P8-DMC suggested that photolysis of the cage occurred in less than 20 ns, a result consistent with the reported photolysis rate of the cage.[24]

In a PAC experiment, the contribution of heat release and volume change to the pressure wave for the transient species (ϕ_i) can be described by Eq. (2)[31]

$$E_{h\nu}\phi_i = q_i + \Delta_R V_i \left(\frac{C_p \rho}{\beta}\right)_T, \qquad (2)$$

where $E_{h\nu}$ is the excitation energy (80.6 kcal/mol for 355 nm); q_i is the heat released in the system after excitation by 1 mol of photons; $\Delta_R V_i$ is the observed volume change in the ith process, $\Delta_R V_i = \Phi_c \Delta V_i$, where Φ_c is the photocleavage quantum yield and ΔV_i is the molar volume change of the solute in the ith process; β is the volumetric thermal expansion coefficient of the solvent; and ρ and C_P are the density and heat capacity of the solvent, respectively. The $(C_p\rho/\beta)$ ratio of the solvent was measured using the method described by Oertling *et al.*[32] The PAC signals for RD1-A7C-DMC and P8-DMC after laser irradiation were measured in the temperature range 15–35°C, and data analysis was performed as described above.

According to Eq. (2), the slope of the line corresponds to the observed volume change ($\Delta_R V_i$) of the sample, and the heat release value (q_i) can be obtained from the intercept on the y axis. As shown in Figure 5, in the case

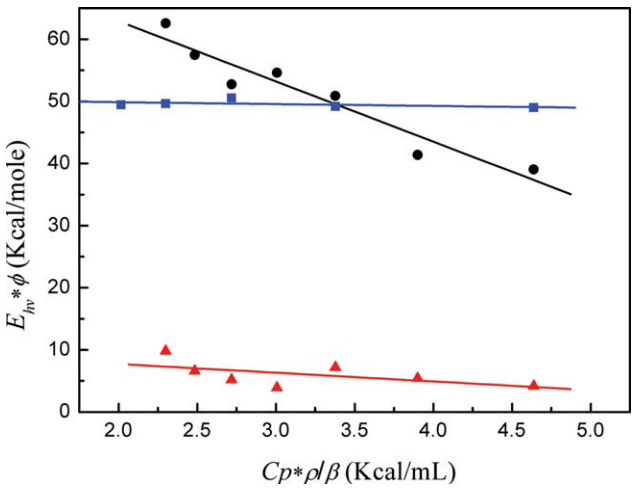


Figure 5

Volume and enthalpy changes for RD1-A7C-DMC and P8-DMC. The plots are of $E_{h\nu}{}^{*}\phi$ versus $C_p{}^{*}\rho/\beta$ for the fast phase (black circles) and slow phase (red triangles) for RD1-A7C-DMC and for P8-DMC (blue squares).

of P8-DMC (blue squares), the observed volume change was −0.3 mL/mol, and the heat released to the solution was 50 kcal/mol. Considering the structure of P8-DMC, it was reasonably expected that there would be no substantial volume change during photolysis, and the experimental result supported this assumption, implying that the cysteine-cage breaking process did not induce a volume change in the peptide. In the case of RD1-A7C-DMC, the volume change and heat release for the fast and slow components were calculated. The fast component (black circles) showed a volume change of −9.7 mL/mol ($\Delta_R V_1$) and a heat release (q_1) of 83 kcal/mol, whereas the slow component (red triangles) showed a small volume change of −1.4 mL/mol ($\Delta_R V_2$) and a heat release (q_2) of 10 kcal/mol.

Simulation of the folding process of RD1-A7C

The folding process of RD1-A7C was simulated using an all-atom model with Gromos force field 43a1[28] and explicit water. Four independent runs of 200 ns were generated. The time dependence of the gyration radius and the native HB formation of the protein are shown in Figure 6, the results being averaged over four trajectories. In the case of the gyration radius [Fig. 6(A)], the fit (smooth red curve) was given by the equation $R_g = R_{g0} + R_{g1} \exp(-t/\tau_{Rg})$, where R_{g0} = 11.6 Å, the pre-exponential factor R_{g1} = 1.4 Å, and the collapse time τ_{Rg} = 3.8 ± 0.1 ns. The gyration radius data clearly showed that the protein underwent a rapid decrease in volume. It should be noted that a previous all-atom simulation study performed on a number of short proteins[33] gave collapse times compatible with our estimate.

To fit the data for HB formation [Fig. 6(B)], a biexponential function was needed, namely, $H = H_0 - H_1\exp(-t/\tau_{HB1}) - H_2\exp(-t/\tau_{HB2})$, where H is the number of native HBs formed in the recorded conformers during simulations, H_0 was 32 ± 5, the pre-exponential factors H_1 and H_2 were 5.9 ± 0.2 and 11 ± 5, respectively, and the time scales τ_{HB1} and τ_{HB2} were 4.2 ± 0.2 ns and 663.3 ± 322.5 ns, respectively. Obviously, τ_{HB1} is close to τ_{Rg}, as it corresponds to the burst collapse stage. The longer time τ_{HB2} describes the slow folding event in which most HBs were formed. Interestingly, the theoretical value of τ_{HB2} is not far from the experimental one. Thus, folding is a two-stage process. As is evident from Figure 6(B), about 20% of the native HBs were formed during the fast volume contraction, following which more and more native HBs were formed, but the gyration radius [Fig. 6(A)] of the collapsed globular structure only showed small fluctuations. This suggests that the protein undergoes structural rearrangement after fast structural collapse.

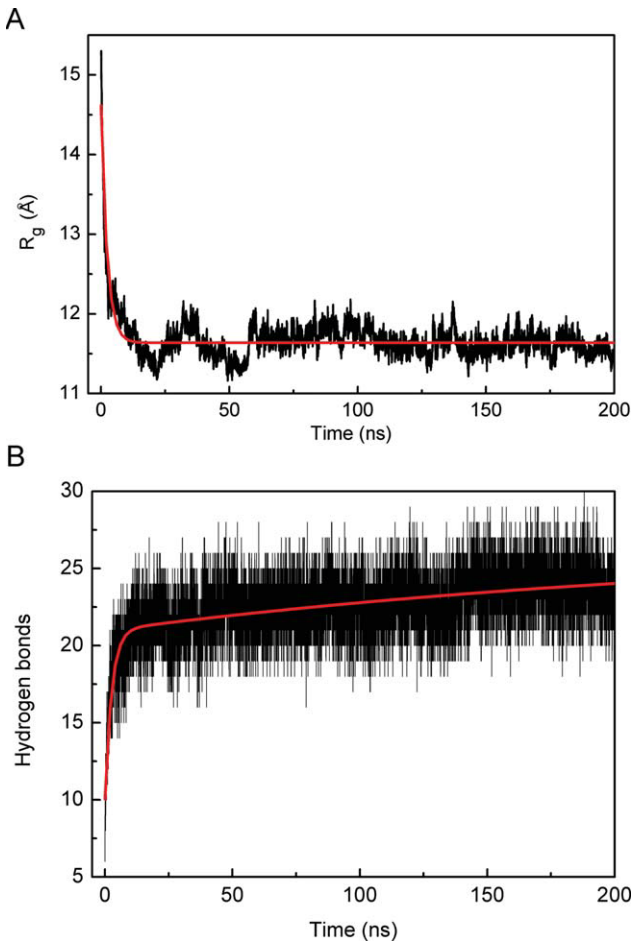


Figure 6

Changes in the gyration radius (**A**) and number of native HB formed (**B**) for RD1-A7C with time during the simulation. The values shown (black lines) are the averaged results from four independent simulations, while the fitting result is shown in red.

DISCUSSION

The NMR results indicated that the presence of the cage resulted in unfolding of the RD1 protein structure. However, the caged protein was prone to aggregate, so photolysis and refolding were performed in the presence of a low concentration of denaturant (GdnHCl). Aggregation was not surprising, as unfolding must result in the exposure of hydrophobic residues, favoring association due to hydrophobic interaction. In a previous study, Hirota *et al.*[20] used DMNBB as the cage to unfold apoPC. Interestingly, the caged-labeled apoPC seemed to be unfolded (although their NMR data suggested there was still some folded protein), but remained soluble in buffer. Comparing the amino acid composition of RD1 and apoPC, apoPC has more charged residues (24 charged residues, 16.2% of the protein) than RD1 (10 charged residues, 7.7% of the protein), whereas hydrophobic plus aromatic residues account for 45 and 58% of the total

residues in apoPC and RD1, respectively. As RD1 has many more hydrophobic residues and fewer charged residues, it is prone to aggregate when unfolded. In addition, it took 400 ns to cleave the DMNB group from apoPC in the study of Hirota *et al.* In our study, RD1 (<20 ns for hydrophobic collapse and ~520 ns for structural rearrangement) was found to fold faster than apoPC (270 μs for hydrophobic collapse, 23 ms for intermediate conversion, and 30–1000 s for *trans*-to-*cis* isomerization of two proline residues), so we were unable to obtain a PAC signal when we tried to use the DMNBB cage in our initial studies. Comparing BrDMC and DMNBB, BrDMC has a faster photolytic rate ($\sim 10^9$ s^{-1})[24] and a reasonable photocleavage quantum yield (~0.2; see Supporting Information), suggesting that it is a good choice for a cage for fast protein folding studies. However, BrDMC itself emits fluorescence when excited, so it is not suitable when the reaction has to be monitored by fluorescence spectroscopy.

However, although we used a caged peptide, P8-DMC, as a control, we cannot rule out the possibility that the existence of the cage after photocleavage might affect protein folding kinetics, because the environments near the caging site for the protein and the peptide are different. Even so, this system can be used to study how protein sequence affects the early stages of folding kinetics on a comparative base by studying different mutant proteins with the cage at the same position.

The folding of small proteins (residue number < 100, including subdomain and secondary structure motifs) has been reported in different types of experiments and simulation calculations.[34] The rate of initial hydrophobic collapse and water depletion of a subdomain protein (or peptide) ranges from a few nanoseconds to tens of microseconds in both simulations and experimental studies, depending on its hydrophobic properties.[33,35,36] It can be very fast; for example, using the T-jump and FRET methods, Sadqi *et al.*[37] measured the temperature-triggered hydrophobic compaction of an acid-denatured protein BBL and found that the relaxation time for hydrophobic collapse is only ~60 ns at 305 K, faster than secondary structure formation. Using the PAC results for RD1-A7C-DMC, two folding events were identified in our study.

The fast event involving volume contraction was assumed to be the initial collapse of the protein. RD1 is a small protein (64 residues) without any disulfide bonds or bound metal, and its structural stability is mainly contributed by its hydrophobic residues, which make up 58% of the total residues. The collapse is probably driven by hydrophobic interactions. Moreover, the observed volume change was −9.7 mL/mol, and the photocleavage quantum yield (Φ_c) of the cage was estimated to be around 0.2. Thus, the ΔV_1 (volume change of the protein during the hydrophobic collapse) corresponded to −80.5 $Å^3$ for one protein molecule. According to the X-ray structure, the surface of RD1 is 3783 $Å^2$, and the total volume can be estimated as ~11,946 $Å^3$ for one protein molecule (the volume and accessible area were calculated using the ARVO method).[38,39] Thus, the percentage volume change during the collapse step was ~0.67%. This suggests that the cage-induced unfolding does not lead to a large size difference in the protein. Finally, the corresponding enthalpy change during the fast event can be calculated using Eq. (3)[40]

$$E_{hv} = q_1 + E_f \times \Phi_f + \Phi_c \times \Delta H_1, \quad (3)$$

where E_f is the energy of the lowest excited state (66.28 kcal/mol) of RD1-A7C-DMC; Φ_f is the fluorescence quantum yield of RD1-A7C-DMC (see Supporting Information); and ΔH_1 is the enthalpy change during the fast event. The fast event had a ΔH_1 of −101 kcal/mol, suggesting that it is an exothermic process and that the system releases heat to its surroundings.

The second event showed a small volume change (ΔV_2 = −1.4 mL/mol), corresponding to a volume change of ~−11 $Å^3$ per protein molecule and a percentage volume change of ~0.09%. These data are consistent with the simulation results showing that the gyration radius only shows a small fluctuation after the initial collapse. We therefore surmised that structural rearrangement toward the global minimum of energy occurred during the second event. The enthalpy change (ΔH_2) for the structural rearrangement, calculated from the released heat during the slow event using the equation: $\Delta H_2 = -q_2/\Phi_c$, was −50 kcal/mol.

RD1 has an internal cavity of 45 $Å^3$ surrounded by nine nonpolar residues [Fig. 1(B,C)]. Mizuno *et al.*[41] proposed that a hydrophobic cavity plays an important role in stabilizing enzyme structure. This is clearly the case for RD1, as, when we mutated V49 to Ala (or Trp) and L55 to Ala, the expressed proteins formed inclusion bodies and could not be refolded, whereas soluble proteins were obtained when V49 or L55 was replaced with a residue of similar size and hydrophobicity, for example, mutation of V49 to Leu and L55 to Val (or Ile) (our unpublished data). These results show that the cavity plays a significant role in maintaining the native structure. Unfortunately, whether this cavity plays a role in the folding kinetics of RD1 cannot be examined experimentally because the protein is not folded in the absence of this cavity. However, the simulation results suggested that this cavity is not formed during the early stage of folding. We analyzed all the simulated conformers that had a hydrophobic cavity (the volume of the cavity was estimated by the enveloping triangulation method)[42] and estimated the probability that residues V5, C7, I11, L17, M21, M22, V45, V49, and L55, which are involved in the cavity in the native form, were involved in the simulated cavities

(see Supporting Information) and found that, in 20% of the simulated conformers, none of these residues were involved in the formed cavity. The maximum number of these residues involved in cavity formation during simulation was six. Thus, we surmised that the native cavity might form after structural rearrangement.

In addition, temperature-induced unfolding of RD1 using molecular dynamic simulation has been reported recently.[43] The authors found that the loss of the three helices of the protein happened earlier than the disruption of the β-sheets. The snapshots of our folding simulation (see Supporting Information) showed that the β-strands formed earlier than the helices. Thus, the loss of interactions during unfolding occurs in the opposite order to the gaining of contact during the folding process, demonstrating that folding and unfolding of RD1 are microscopically reversible.[44,45]

CONCLUSION

An important goal of protein folding studies is to quantitatively predict the contributions of amino acid sequence, pH, temperature, and salt concentration to the kinetics and thermodynamics of the protein folding process. To achieve this goal, detailed knowledge of the structural and thermodynamic properties of the different states populated during folding and of the kinetics involved in going from one state to another is required. Although experimental studies on protein folding have become a popular field in recent years, the bulk of the work has been directed toward the late stage of the folding process. Despite almost 40 years of effort, because of the limitations of kinetics methods, there is still a paucity of data on the early kinetic events in the refolding process. For any kinetic study to be meaningful, the event that perturbs the equilibrium must occur on a time scale much shorter than the process of interest. Unfortunately, traditional methods for triggering a protein folding event in a stopped-flow or quenched-flow instrument are often too slow to measure the early folding kinetics because these instruments inevitably have a dead time of several milliseconds, depending on the method used to change solvent conditions. The folding kinetics of RD1 have been difficult to study not only because its folding is too fast to be detected by traditional methods but also because it does not have a characteristic circular dichroism spectrum and does not contain tryptophan to serve as an intrinsic fluorescence probe for monitoring the folding process. Here, we, for the first time, report that BrDMC can be used as a photolabile cage to study fast protein folding. Combining a photolabile caging strategy and the PAC method, the fast folding kinetics of RD1 were monitored on the nanosecond time scale. The same strategy can be used to study the fast folding kinetics of other proteins that could not be previously studied.

ACKNOWLEDGMENTS

The NMR spectra were obtained at the High-Field Nuclear Magnetic Resonance Center (HF-NMRC) supported by the National Research Program for Genomic Medicine.

REFERENCES

1. Eaton WA, Thompson PA, Chan CK, Hage SJ, Hofrichter J. Fast events in protein folding. Structure 1996;4:1133–1139.
2. Eaton WA, Munoz V, Thompson PA, Chan CK, Hofrichter J. Submillisecond kinetics of protein folding. Curr Opin Struct Biol 1997;7:10–14.
3. Dumont C, Emilsson T, Gruebele M. Reaching the protein folding speed limit with large, sub-microsecond pressure jumps. Nat Methods 2009;6:515–519.
4. Bartlett AI, Radford SE. An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. Nat Struct Mol Biol 2009;16:582–588.
5. Shastry MCR, Luck SD, Roder H. A continuous-flow capillary mixing method to monitor reactions on the microsecond time scale. Biophys J 1998;74:2714–2721.
6. Shastry MCR, Roder H. Evidence for barrier-limited protein folding kinetics on the microsecond time scale. Nat Struct Biol 1998;5:385–392.
7. Park SH, Shastry MC, Roder H. Folding dynamics of the B1 domain of protein G explored by ultrarapid mixing. Nat Struct Biol 1999;6:943–947.
8. Chan CK, Hu Y, Takahashi S, Rousseau DL, Eaton WA, Hofrichter J. Submillisecond protein folding kinetics studied by ultrarapid mixing. Proc Natl Acad Sci USA 1997;94:1779–1784.
9. Akiyama S, Takahashi S, Ishimori K, Morishima I. Stepwise formation of alpha-helices during cytochrome c folding. Nat Struct Biol 2000;7:514–520.
10. Chan CK, Hofrichter J, Eaton WA. Optical triggers of protein folding. Science 1996;274:628–629.
11. Jones CM, Henry ER, Hu Y, Chan CK, Luck SD, Bhuyan A, Roder H, Hofrichter J, Eaton WA. Fast events in protein folding initiated by nanosecond laser photolysis. Proc Natl Acad Sci USA 1993;90:11860–11864.
12. Pascher T, Chesick JP, Winkler JR, Gray HB. Protein folding triggered by electron transfer. Science 1996;271:1558–1560.
13. Thompson PA, Eaton WA, Hofrichter J. Laser temperature jump study of the helix<==>coil kinetics of an alanine peptide interpreted with a 'kinetic zipper' model. Biochemistry 1997;36:9200–9210.
14. Williams S, Causgrove TP, Gilmanshin R, Fang KS, Callender RH, Woodruff WH, Dyer RB. Fast events in protein folding: helix melting and formation in a small peptide. Biochemistry 1996;35:691–697.
15. Xu Y, Purkayastha P, Gai F. Nanosecond folding dynamics of a three-stranded beta-sheet. J Am Chem Soc 2006;128:15836–15842.
16. Munoz V, Thompson PA, Hofrichter J, Eaton WA. Folding dynamics and mechanism of beta-hairpin formation. Nature 1997;390:196–199.
17. Chen RPY, Huang JJT, Chen HL, Jan H, Velusamy M, Lee CT, Fann WS, Larsen RW, Chan SI. Measuring the refolding of beta-sheets with different turn sequences on a nanosecond time scale. Proc Natl Acad Sci USA 2004;101:7305–7310.
18. Kuo NNW, Huang JJT, Miksovska J, Chen RPY, Larsen RW, Chan SI. Effects of turn stability on the kinetics of refolding of a hairpin in beta-sheet. J Am Chem Soc 2005;127:16945–16954.

19. Rock RS, Hansen KC, Larsen RW, Chan SI. Rapid photochemical triggering of protein unfolding in a nondenaturing environment. Chem Phys 2004;307:201–208.
20. Hirota S, Fujimoto Y, Choi J, Baden N, Katagiri N, Akiyama M, Hulsker R, Ubbink M, Okajima T, Takabe T, Funasaki N, Watanabe Y, Terazima M. Conformational changes during apoplastocyanin folding observed by photocleavable modification and transient grating. J Am Chem Soc 2006;128:7551–7558.
21. Ko TP, Robinson H, Gao YG, Cheng CHC, DeVries AL, Wang AHJ. The refined crystal structure of an eel pout type III antifreeze protein RD1 at 0.62 Å resolution reveals structural microheterogeneity of protein and solvation. Biophys J 2003;84:1228–1237.
22. Hamada T, Ito Y, Abe T, Hayashi F, Guntert P, Inoue M, Kigawa T, Terada T, Shirouzu M, Yoshida M, Tanaka A, Sugano S, Yokoyama S, Hirota H. Solution structure of the antifreeze-like domain of human sialic acid synthase. Protein Sci 2006;15:1010–1016.
23. Hwang TS, Hung CH, Teo CF, Chen GT, Chang LS, Chen SF, Chen YJ, Lin CH. Structural characterization of *Escherichia coli* sialic acid synthase. Biochem Biophys Res Commun 2002;295:167–173.
24. Furuta T, Noguchi K. Controlling cellular systems with Bhc-caged compounds. Trends Anal Chem 2004;23:511–519.
25. Larsen RW, Miksovska J. Time resolved thermodynamics of ligand binding to heme proteins. Coord Chem Rev 2007;251:1101–1127.
26. Schaeffer RD, Fersht A, Daggett V. Combining experiment and simulation in protein folding: closing the gap for small model systems. Curr Opin Struct Biol 2008;18:4–9.
27. Braslavsky SE, Heibel GE. Time-resolved photothermal and photoacoustic methods applied to photoinduced processes in solution. Chem Rev 1992;92:1381–1410.
28. Berendsen HJC, Vanderspoel D, Vandrunen R. Gromacs—a message-passing parallel molecular-dynamics implementation. Comput Phys Commun 1995;91:43–56.
29. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. J Comput Chem 1997;18:1463–1472.
30. Darden T, York D, Pedersen L. Particle mesh ewald: An N.Log(N) method for ewald sums in large systems. J Chem Phys 1993;98:10089–10092.
31. Gensch T, Viappiani C. Time-resolved photothermal methods: accessing time-resolved thermodynamics of photoinduced processes in chemistry and biology. Photochem Photobiol Sci 2003;2:699–721.
32. Oertling WA, Cornellison CD, Treff NR, Watanabe J, Pressler MA, Small JR. Photoacoustic characterization of protein dynamics following CO photodetachment from fully reduced bovine cytochrome c oxidase. J Inorg Biochem 2007;101:635–643.
33. Camilloni C, Sutto L, Provasi D, Tiana G, Broglia RA. Early events in protein folding: Is there something more than hydrophobic burst? Protein Sci 2008;17:1424–1433.
34. Kubelka J, Hofrichter J, Eaton WA. The protein folding 'speed limit'. Curr Opin Struct Biol 2004;14:76–88.
35. Qiu LL, Zachariah C, Hagen SJ. Fast chain contraction during protein folding: "foldability" and collapse dynamics. Phys Rev Lett 2003;90:1681031–1681034.
36. Zhou RH, Huang XH, Margulis CJ, Berne BJ. Hydrophobic collapse in multidomain protein folding. Science 2004;305:1605–1609.
37. Sadqi M, Lapidus LJ, Munoz V. How fast is protein hydrophobic collapse? Proc Natl Acad Sci USA 2003;100:12117–12122.
38. Busa J, Dzurina J, Hayryan E, Hayryan S, Hu CK, Plavka J, Pokorny I, Skrivanek J, Wu MC. ARVO: a Fortran package for computing the solvent accessible surface area and the excluded volume of overlapping spheres via analytic equations. Comput Phys Commun 2005;165:59–96.
39. Hayryan S, Hu CK, Skrivanek J, Hayryan E, Pokorny I. A new analytical method for computing solvent-accessible surface area of macromolecules and its gradients. J Comput Chem 2005;26:334–343.
40. dos Santos RMB, Lagoa ALC, Simoes JAM. Photoacoustic calorimetry. An examination of a non-classical thermochemistry tool. J Chem Thermodyn 1999;31:1483–1510.
41. Mizuno T, Hasegawa C, Tanabe Y, Hamajima K, Muto T, Nishi Y, Oda M, Kobayashi Y, Tanaka T. Organic ligand binding by a hydrophobic cavity in a designed tetrameric coiled-coil protein. Chem Eur J 2009;15:1491–1498.
42. Busa J, Hayryan S, Hu CK, Skrivanek J, Wu MC. Enveloping triangulation method for detecting internal cavities in proteins and algorithm for computing their surface areas and volumes. J Comput Chem 2009;30:346–357.
43. Kunda S, Roy D. Temperature-induced unfolding pathway of a type III antifreeze protein: insight from molecular dynamics simulation. J Mol Graph Model 2008;27:88–94.
44. Daggett V. Protein folding-simulation. Chem Rev 2006;106:1898–1916.
45. Li MS, Kouza M, Hu CK. Refolding upon force quench and pathways of mechanical and thermal unfolding of ubiquitin. Biophys J 2007;92:547–561.