THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Transient Helical Structure during PI3K and Fyn SH3 Domain Folding

Yoshitaka Matsumura,[†] Masaji Shinjo,*[,†] Seung Joong Kim,[‡] Nobuyuki Okishio,[§] Martin Gruebele,[‡,⊥] and Hiroshi Kihara[†,∥,#]

[†]Department of Physics, Kansai Medical University, 2-5-1 Shinmachi, Hirakata, 573-1010, Japan

[‡]Department of Physics and [⊥]Department of Chemistry, and Center for Biophysics and Computational Biology, University of Illinois, Urbana, Illinois, 61801, United States

[§]Japanese Red Cross Hokkaido Blood Center, 2-2 Yamanote, Nishi-ku, Sapporo 063-0002, Japan

[∥]SR center, Ritsumeikan University, 1-1-1 Noji-Higashi Kusatsu, Shiga, 525-8577, Japan

[#]Synchrotron Radiation Research Center, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan

ABSTRACT: A growing list of proteins, including the β-sheet-rich SH3 domain, is known to transiently populate a compact α-helical intermediate before settling into the native structure. Examples have been discovered in cryogenic solvent as well as by pressure jumps. Earlier studies of λ repressor mutants showed that transient states with excess helix are robust in an all-α protein. Here we extend a previous study of src SH3 domain to two new SH3 sequences, phosphatidylinositol 3-kinase (PI3K) and a Fyn mutant, to see how robust such helix-rich transients are to sequence variations in this β-sheet fold. We quantify helical structure by circular dichroism (CD), protein compactness by small-angle X-ray scattering (SAXS), and transient helical populations by cryo-stopped-flow CD. Our results show that transient compact helix-rich intermediates are easily accessible on the folding landscape of different SH3 domains. In molecular dynamics simulations, force field errors are often blamed for transient non-native structure. We suggest that experimental examples of very fast α-rich transient misfolding could become a more subtle test for further force field improvements than observation of the native state alone.

INTRODUCTION

The formation of α-helical structure from a polypeptide chain begins with a localized nucleation event between residues i and $i + 4$.[1] Thus, it is plausible that non-native helical structure could form transiently during the earliest stages of protein folding, only to dissolve when longer-range native contacts of even lower free energy are made. Residual helical structure exists in denatured states,[2] fragments α-helix and β-sheet proteins,[3,4] and "equilibrium intermediates" stabilized in acidic solution.[5] It is also seen in simulations, even with force fields that have relatively low helix propensity.[6]

Excess α-helix content has been reported during the earliest stages of protein folding. In particular, cryogenic solvent conditions have tuned the free energy landscape to promote transient helix formation in β-lactoglobulin,[7−9] src SH3 domain,[10] the FHA domains of Rad53 and Ki67,[11] two sequences of ubiquitin,[12,13] and two λ_{6-85} sequences.[14,15] In the case of the five helix bundle λ_{6-85}, we were able to tune the amount of transient helix by introducing helix-destabilizing residues, but the helix-rich transients were not eliminated.

SH3 domains are β-sheet proteins containing two antiparallel β-sheets connected by various-length loops (Figure 1).[16] The src SH3 domain has been studied in particular detail. The structure formed in its transition state in room temperature solvent is known (central β-sheet),[17,18] but the transition state structure can be altered substantially by circular permutation.[19] Dasgupta and Udgaonkar detected a rapidly formed compact folding intermediate, showing that the early stages of assembly are more complex than previously thought.[20] Under cryogenic solvent conditions, src SH3 domain formed a transient α-helix-rich intermediate before folding to the native state.[10] The helix-rich intermediate is a compact globular state: its radius of gyration ($R_g \approx 1.9$ nm) is closer to the native state (1.5 nm) than to the denatured state (2.7 nm).

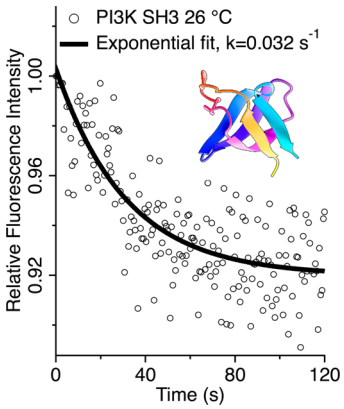

Figure 1. Generic SH3 fold ribbon structure and intrinsic tyrosine fluorescence of PI3K SH3 domain refolding by GuHCl-jump at 26 °C (aqueous buffer without ethylene glycol). The solid curve is a single exponential fit. The rate constant was 0.032 ± 0.004 s^{-1}.

Received: January 6, 2013
Revised: March 25, 2013
Published: March 28, 2013

Table 1. Comparison of Protein Sequences and AGADIR (Helix Propensity) Scores[a]

```
Src   ---------------------------------TFVALYDYESRTETDLSFKKGERLQI—
Fyn   ----------------------------VQIVTLFVALYDYEARTEDDLSFHKGEKFQI--
Fyn mutant   --------------------VQISTLFEALYDYEARTEDDLSFHKGEKFQI--
PI3K  ------------------------GSMSAEGYQYRALYDYKKEREEDIDLHLGDILTVNK

Src     -----VNNTEG--------DWWLAHSLTTGQTGYIPSNYVAPSD----------------
Fyn     -----LNSSEG--------DWFEARSLTTGETGYIPSNYVAPVDRLDYKDDDDKHHHHHH
Fyn mutant -LNSSEG--------DWWEVRSLTTGETGYIPSNYFAPVDRLDYKDDDDKHHHHHH
PI3K    GSLVALGFSDGQEARPEEIGWLNGYNETTGERGDFPGTYVEYIGRKKISPP---------
```

AGADIR helix % score (pH 6, 278 K, 0.1 M ionic strength)
Src SH3 domain: 0.77; Fyn SH3 domain: 0.61; Fyn SH3 domain (mutant): 0.69
PI3K SH3 domain: 1.67

[a]Fyn mutations are shown in red, highly conserved residues in yellow.

Here we study whether two sequence variants of the SH3 fold also show a helical burst phase under cryogenic stopped flow conditions. In an all-helical protein like λ_{6-85}, or an α/β protein such as ubiquitin, helical propensity turned out to be robust, but is it for a β-sheet protein? We look in detail at the PI3K SH3 domain, and at a new mutant of Fyn SH3 domain. The two new sequences are compared to SH3 domains studied in the past in Table 1. We chose the Fyn SH3 mutant to study the effect of a small sequence difference relative to SH3 sequences previously studied in cryosolvent. We chose PI3K because its sequence is very different from other SH3 sequences previously studied in cryogenic solvent, but folding studies exist. Guijarro et al. initiated the folding kinetics of the PI3K SH3 domain by guanidine hydrochloride (GuHCl) jump and detected it by real-time one-dimensional ^{1}H nuclear magnetic resonance (NMR) and several optical spectroscopies at 6 and 20 °C.[21] They showed that all probes monitored the same slow phase. Dasgupta and Udgaonkar reported that the initial collapse of the PI3K SH3 domain polypeptide chain is completed within submilliseconds.[20] They initiated the reaction by stopped flow and continuous flow and detected it by fluorescence resonance energy transfer (FRET) experiments at 25 °C.[20] However, these studies were not designed to detect fast formation of secondary structure.

If both PI3K and mutant Fyn SH3 domains also form α-helix-rich compact folding intermediates, then the formation of transient helix structure would appear to be quite robust for this β-sheet-rich fold. SH3 domain then joins the ranks of proteins whose sequence variants have compact non-native states with excess helical structure that lie low enough on the free energy landscape to be stabilized by cryogenic folding conditions.

Molecular dynamics simulations have also shown that helical transients form easily during the early stages of folding.[22−24] This is frequently blamed on force field errors, and indeed there is evidence for such "superhelicity" of force fields used to fold a protein to the *in silico* minimum free energy.[23] However, the experiments described here, as well as the earlier studies referenced above, provide evidence that non-native helical traps can be formed very rapidly in experiments. Further evidence for this is offered by recent pressure jump experiments on λ_{6-85}, also interpreted in terms of transient helical overshoots.[25]

These results could be useful for testing force fields at a finer level of mechanistic detail than the current paradigm of native structure and average folding/unfolding rates.[26] While many small proteins are apparent two-state folders, their folding barriers are quite small (typically 2−10 k_BT), and may just barely "hide" free energy landscape roughness, which takes the form of transiently populated non-native structural ensembles. The detection of such transient structures by experiment and molecular dynamics simulations could lead to the next level of refinement and validation of force fields used for protein folding calculations.

METHODS

Proteins. The PI3K SH3 domain was expressed in *E. coli* and purified as described.[27] The Fyn SH3 domain mutant (Table 1) was expressed in *E. coli* and purified as described.[28] Hereafter, the mutant is denoted as just "Fyn mutant".

Equilibrium CD. The samples of PI3K and Fyn mutant were prepared in 50 mM sodium phosphate buffer at pH 6 with various concentrations of GuHCl in the presence of 45% ethylene glycol unless otherwise noted. GuHCl is of ultrapure reagent grade from ICN Biomedicals Inc. (lot no. 3814J). The concentration of GuHCl was calibrated by refractive index measurements.[29,30] The temperature was controlled within ±1 °C by a recirculating refrigerator. The concentration of both SH3 domains was less than 130 μM. CD measurements were performed at various temperatures with a spectropolarimeter specially designed by Unisoku Inc. for cryogenic operation. Cuvettes of 1 mm path-length were used for all measurements.

Kinetic Refolding: Instrument and Dead Time. The stopped flow instrument was constructed for use with high viscosity solutions at low temperature in collaboration with Unisoku Inc.[31] Its dead time was estimated to be *ca.* 6 ms by the reduction reaction of 2,6-dichloroindophenol by ascorbic acid[32] at 4 °C and −28 °C. For SAXS measurements, dead time was limited by the data sampling rate of 200 ms.

Kinetic Refolding: Tyrosine and Tryptophan Fluorescence. Refolding was measured by diluting protein in GuHCl buffer with excess GuHCl-free buffer. The cryogenic buffer was made with 50 mM sodium phosphate in the absence (T > 4 °C) or presence ($T \leq 4$ °C) of 45% ethylene glycol. One volume of protein in 5 M GuHCl buffer was mixed with 6 volumes of buffer, thus giving 0.7 M GuHCl after mixing. A 50 mM sodium phosphate buffer at pH 6 was used to prepare all solutions. The final concentration of PI3K SH3 domain was <40 μM, for Fyn mutant <70 μM.

Fluorescence was monitored with a specially designed optical fiber of less than 1 mm thickness, which was inserted into the stopped flow observation cell. Tyrosine fluorescence was excited at 268 nm and integrated at wavelengths above 305 nm with a cutoff filter. Tryptophan fluorescence was excited at

dx.doi.org/10.1021/jp400167s | J. Phys. Chem. B 2013, 117, 4836−4843

295 nm and integrated at wavelengths above 325 nm with a cutoff filter.

Kinetic Refolding: CD. The same conditions were used as for fluorescence, with final concentration of PI3K SH3 domain 10−30 μM and Fyn mutant 30−50 μM. The refolding process was monitored by CD at 222 ± 2 nm ($[\theta]_{222}$) at several temperatures. The averaged data was normalized to give molar ellipticity.

At each condition, in addition to the refolding experiment, two more experiments were always done: (1) Mixing the unfolded protein in the GuHCl buffer with the same GuHCl buffer gives us the initial CD level. (2) Letting the solution relax for 5 to 20 min gives us the final CD level. In many cases, the first data point of the refolding experiments already differs from the initial level, indicating the existence of a burst phase. Measured ellipticities were reliable to ±200 deg cm^2 dmol^{-1} during the ≤10 s time scale of kinetics measurements, although drifts of 500 to 1000 deg cm^2 dmol^{-1} were observed over 5−20 min. However, all burst phases have amplitudes significantly larger than even this slow drift.

SAXS. The SAXS experiment monitors the decreasing size of the protein coil during the refolding process.[33] SAXS experiments were performed at beamline 15A of the Photon Factory at KEK (Tsukuba, Japan).[34,35] The X-ray wavelength was λ = 1.504 Å. A CCD detector was used for both kinetic and static measurements.[36,37] Time-resolved SAXS measurements of PI3K SH3 domain were performed at −5 °C as described elsewhere.[38,39]

SAXS measurements were also performed at the Biophysics Collaborative Access Team Beamline of the Advanced Photon Source at Argonne National Laboratory (Argonne, IL).[40] The X-ray wavelength was λ = 1.0 Å. An Aviex CCD detector was used for kinetic measurements as described previously.[28] Time-resolved SAXS measurements of Fyn mutant were performed at −28 °C. The dead time was 0.1 to 0.2 s.

Helium gas was flowed in and around the observation cell in order to prevent frost from forming on the window. X-ray scattering data were collected for both protein solution and the corresponding buffer to allow background subtraction. The scattering data of the buffer were subtracted from those of the protein solution. X-ray scattering data were analyzed by the Guinier approximation, assuming that the logarithm of the scattering intenisty depends linearly on q^2, where $q = 4\pi \sin\theta/\lambda$ and θ is half of the scattering angle.[41] The linear Guinier region was fitted to obtain R_g, where the value of $R_g q$ is ≲1.3 to satisfy the Guinier condition.

Equilibrium experiments were performed using a static-flow cell[42] at ≤4 °C to prevent radiation damage of the proteins. Concentrations of PI3K and Fyn SH3 domain were less than 300 μM.

RESULTS

Comparison of PI3K Kinetics with Previous Measurements. We observed the slow apparent two-state refolding of PI3K SH3 domain at room temperature to confirm the consistency of our data with previous results.[20,21] Figure 1 shows kinetic refolding signals by intrinsic tyrosine fluorescence of PI3K SH3 domain at 26 °C. The signal fits a single-exponential decay with rate coefficient k_{obs} = 0.032 ± 0.004 s^{-1}. In the kinetic refolding experiments of intrinsic tyrosine fluorescence by Guijarro and co-workers,[21] two rate coefficients were reported: 0.0585 ± 0.0016 s^{-1} and 0.0169 ± 0.0003 s^{-1}, for two-state folding of different proline isomeric forms. Their average rate coefficient matches well with our single value. (Our signal-to-noise ratio did not require a double-exponential fit.)

GuHCl-Induced Equilibrium Unfolding of PI3K SH3 Domain. We next performed equilibrium unfolding experiments at subzero temperatures to determine optimal conditions for refolding kinetics. Experiments were done at pH 6 in the presence of ethylene glycol as antifreeze (see Methods).[9] We checked the effect of ethylene glycol on the conformation of PI3K SH3 domain by circular dichroism (CD) spectroscopy. CD ellipticities at 222 nm ($[\theta]_{222}$) were within 5% from 0% to 90% ethylene glycol. This result for these solvent conditions is consistent with our previous work.[9−13,15]

Figure 2A shows far-UV CD spectra of PI3K SH3 domain at two temperatures with and without GuHCl. The two 0 M GuHCl spectra are nearly identical, so the secondary structure of PI3K SH3 domain remains stable from −20 to 4 °C. They are also very similar to spectra of the native state at pH 7.2 at 25 °C reported by Okishio et al.[27] The two 5 M GuHCl spectra show slightly different random coil shapes, so residual secondary structure is temperature-dependent. PI3K SH3 domain is extensively denatured by 5 M GuHCl in 45% ethylene glycol regardless of temperature. (Guanidine absorption precludes spectral collection below 210 nm.)

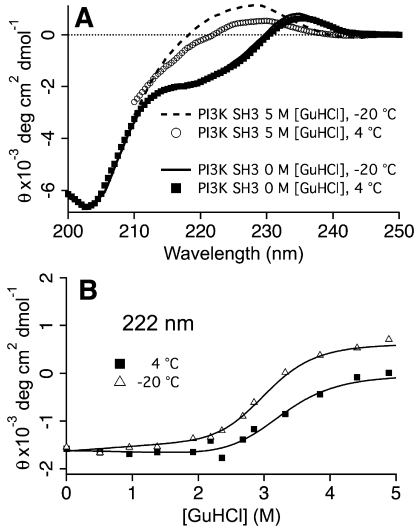

Figure 2. GuHCl-induced unfolding of PI3K SH3 domain monitored by CD at 4 and −20 °C in 45% ethylene glycol aqueous buffer. (A) Far-UV CD spectra of PI3K SH3 domain with and without GuHCl denaturant. In the short wavelength region with 5 M GuHCl, the data are not shown due to the strong absorption by GuHCl. (B) Unfolding transitions of PI3K SH3 domain on GuHCl concentrations monitored by $[\theta]_{222}$. The solid lines are fits to the cooperative model in eq 1: c_M = 3.26 ± 0.06 M and m = 7.80 ± 0.93 M^{-1} at 4 °C, and c_M = 3.08 ± 0.04 M and m = 8.28 ± 0.66 M^{-1} at −20 °C.

The guanidine denaturation curves at 222 nm ($[\theta]_{222}$) are shown in Figure 2B. The midpoints lie at c_M = 3.08 ± 0.04 M (−20 °C) and 3.26 ± 0.06 M, so denaturant jumps from 5 to 0.7 M GuHCl should allow for complete refolding of the PI3K SH3 domain.

The guanidine denaturation curve was also detected by SAXS at 4 °C. R_g values were calculated from Guinier plots of each scattering curve, and plotted in Figure 3A. The R_g values of PI3K are in good agreement with those for src SH3 domain from ref 10: 14.5 ± 0.6 Å (folded PI3K) vs 14.6 Å (folded src) and 27.5 ± 0.7 Å (denatured PI3K) vs 26.7 Å (denatured src).

dx.doi.org/10.1021/jp400167s | J. Phys. Chem. B 2013, 117, 4836−4843

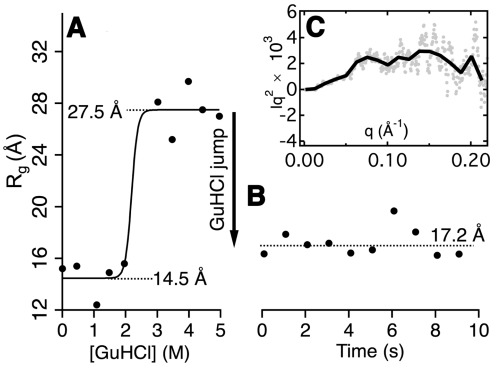

Figure 3. GuHCl-induced unfolding and refolding of PI3K SH3 domain monitored by X-ray solution scattering in 45% ethylene glycol aqueous buffer. (A) Dependence of R_g on GuHCl concentration at 4 °C; the solid line is a fit to the cooperative model in eq 1. The fitted thermodynamic parameters are $m = 17 \pm 4$ M^{-1} and $c_M = 2.7 \pm 0.5$ M. (B) GuHCl concentration jump by the cryo-stopped flow method at −5 °C, on the same *y*-scale as (A). R_g changed from the unfolded value of 27.5 Å to 17.2 Å within the dead time of the apparatus. (C) Kratky plot for data from (B) averaged from 0 to 1 s shows a broad peak (the solid line is a smoothed version of the same data).

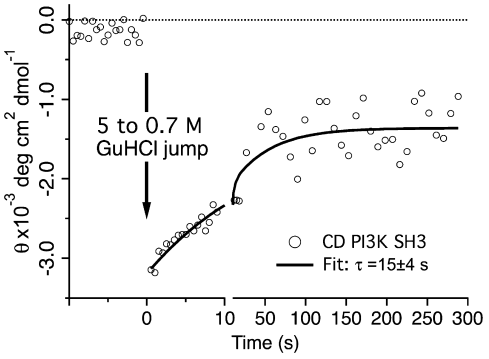

Figure 4. Kinetic refolding experiments of PI3K SH3 domain at −5 °C in aqueous ethylene glycol buffer. The reaction was initiated by cryo-stopped flow from 5 to 0.7 M GuHCl, and monitored by $[\theta]_{222}$, the circular dichroism signal at 222 nm. The black curve is the single exponential fit, also shown on the split time axis (see Table 2 for all kinetic rate coefficients).

The SAXS data yielded $c_M = 2.7 \pm 0.5$ M GuHCl where the R_g increased cooperatively. The large uncertainty of c_M arises because the SAXS data was sampled at 0.5 M increments of denaturant.

The CD and SAXS fits were obtained with a sigmoid model (two-state, native = N and unfolded = U). The signal *S* is represented by[7]

$$S(c) = \frac{S_N + S_U\, e^{-m(c-c_M)}}{1 + e^{-m(c-c_M)}} \quad (1)$$

Here *m* represents the cooperativity of the transition, and S_N and S_U are the signal values of the N and U states, represented as linear baselines. Note that the smaller midpoint $c_M = 2.7 \pm 0.5$ M obtained at 4 °C by SAXS agrees within experimental uncertainty with the more accurate value of 3.26 ± 0.06 M obtained by CD. The midpoints for both CD and SAXS lie at about 3 M, so denaturant jumps from 5 to 0.7 M GuHCl between −20 and 4 °C should allow for complete refolding of the PI3K SH3 domain.

Our c_M in ethylene glycol buffer lies 1.3 M higher than the result reported by Guijarro et al. and detected by intrinsic tyrosine fluorescence (c_M is 1.4 M).[21] As they measured at room temperature without ethylene glycol, this suggests PI3K SH3 domain is more stable at lower temperatures in the presence of 45% ethylene glycol. We have reported that mutants of ubiquitin, FHA1 domain of Rad53, and λ repressor proteins were also slightly more stable in the presence of cryosolvent.[11,13,15]

Refolding Kinetics of PI3K SH3 in Cryosolvent. We carried out time-resolved refolding experiments by GuHCl-jump from 5 to 0.7 M with our cryo-stopped flow apparatus. Secondary structure was detected by CD and protein compactness was detected by SAXS at subzero temperatures (Figures 3B and 4).

Time-resolved R_g values after the GuHCl-jump are shown in Figure 3B. R_g dropped from the unfolded value of 27.5 Å (Figure 4A) to 17.2 ± 1.0 Å within our dead time of 200 ms. This radius is 19% larger than that of the native state (14.5 ± 0.6 Å), but much smaller than that of the denatured state (27.5 ± 0.7 Å). Therefore any transient intermediate of PI3K SH3 domain is nearly as compact as the native state. We averaged Kratky plots over the 0 to 1 s time window. Albeit noisy, the plot has a peak (Figure 3C), further suggesting that the transient intermediate is a compact globular state.

In Figure 4, secondary structure is monitored during refolding of PI3K SH3 domain. The refolding trace at −5 °C shows a large burst phase in the CD spectrum at 222 nm (Table 2), followed by a single exponential decay toward the native value. The transient ellipticity is about twice the magnitude of the native CD signal (*ca.* −1600 deg cm^2 dmol^{-1}). This suggests that PI3K SH3 domain forms a transient α-helix-rich intermediate in <6 ms (our dead time), followed by slower interconversion to the native state (directly or *via* less compact conformations). We also performed kinetic refolding CD experiments at −20 °C. The α-helical burst was even larger at the lower temperature (Table 2).

Together, the kinetic refolding experiments in cryogenic solvent detected by SAXS and CD show that the PI3K SH3 domain at low temperature populates a compact α-helix-rich state before folding to the native state on a much longer time scale. The radius of the transient intermediate is only 19% larger than that of the native state.

Kinetic Refolding of Fyn Mutant in Cryosolvent. We also studied a new Fyn mutant, which is much more similar to previous SH3 domains studied under cryogenic solvent conditions.[10] All experiments at low temperatures were done at pH 6 in the presence of 45% ethylene glycol as antifreeze (see Methods).

Figure 5A shows circular dichroism spectra of Fyn mutant in 0 and 5 M GuHCl buffer. Its CD spectra in 0 or 0.7 M GuHCl at 4 °C are almost identical to that of the native wild-type Fyn SH3 domain reported by Di Nardo et al.[43] This indicates that Fyn SH3 domain remains folded in 0.7 M GuHCl at cryogenic temperatures, just like PI3K SH3 domain. In contrast, the CD spectrum in 5 M GuHCl is nearly zero at 222 nm.

We measured refolding kinetics by jumping the Fyn mutant from 5 to 0.7 M GuHCl at −28 °C (Figure 5B). The signal jumps from a molar ellipticity ≈0 to −5600 deg cm^2 dmol^{-1} in <6 ms. Following the burst phase, the native state CD signal is recovered in ≈1 s. The resolvable refolding kinetics was not a single-exponential process. Two rate coefficients of 17 ± 3 s^{-1} and 2.5 ± 0.5 s^{-1} fitted the data within measurement uncertainty. The two-step process is consistent with the

dx.doi.org/10.1021/jp400167s | J. Phys. Chem. B 2013, 117, 4836−4843

Table 2. Burst Phase Amplitudes and Refolding Rate Coefficients of SH3 Domains at pH 6

protein	CD value at 222 nm of the burst phase ($deg.cm^2.dmol^{-1}$)	rate constant of the fast phase (s^{-1})	rate constant of the slow phase (s^{-1})
src SH3 domain (ref 10)	−6700 (at 4 °C)	5.5	-
Fyn SH3 pseudowild type (ref 10)	−5700 (at −20 °C)	32	1.1
Fyn mutant	−5200 ± 400 (at −24 °C)	20 ± 3[a]	1.5 ± 0.3[a]
(this work)	−5600 ± 400 (at −28 °C)	17 ± 3[a]	2.5 ± 0.5[a]
PI3K SH3 domain	−3200 ± 200 (at −5 °C)	0.07 ± 0.03	-
(this work)	−4100 ± 200 (at −20 °C)	0.06 ± 0.04	-

[a]Global fit of CD- and fluorescence-detected kinetics.

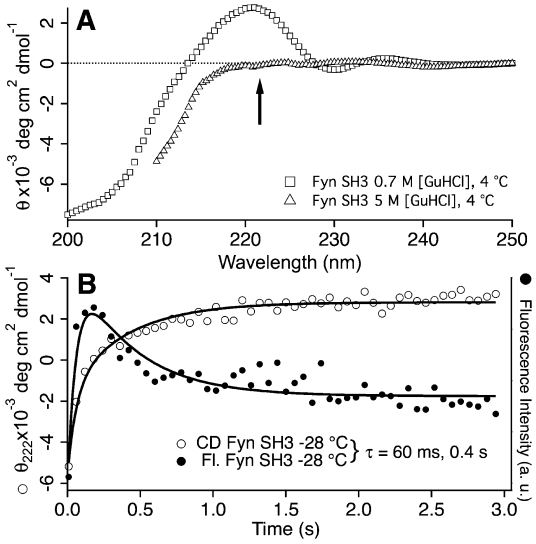

Figure 5. GuHCl-induced unfolding and refolding of Fyn mutant in 45% ethylene glycol aqueous buffer. (A) Far-UV CD spectra of the native (0 M GuHCl) and denatured (5 M GuHCl) states of Fyn mutant at 4 °C. In the short wavelength region of the 5 M GuHCl CD spectrum, the data are not shown due to the strong absorption of GuHCl. The arrow at 222 nm indicates the kinetic probe wavelength. (B) Refolding of Fyn mutant monitored by CD ($[\theta]_{222}$) and tryptophan fluorescence at −28 °C. The GuHCl concentration was jumped from 5 to 0.7 M. Double-exponential functions with the same rate coefficients but different amplitudes were fitted globally to the CD and fluorescence data (solid lines, see Table 2 for values of the rate coefficients).

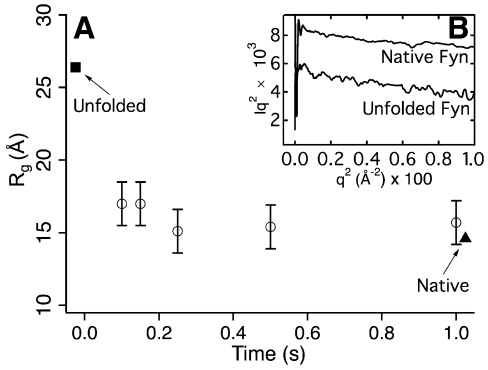

Figure 6. X-ray solution scattering experiments of Fyn SH3 domain in aqueous ethylene glycol buffer. (A) Refolding of Fyn mutant by cryo-stopped flow SAXS at −28 °C. R_g decreases to the native value within 0.1 to 0.2 s. Native and unfolded R_g values were obtained by equilibrium experiments. (B) Guinier plots of the native (0 M GuHCl) and denatured (4 M GuHCl) state of Fyn SH3 domain at 4 °C at equilibrium. The R_g values of the native and denatured state were 14.6 ± 0.8 Å and 27.3 ± 3.9 Å, respectively.

previously published kinetics of Fyn SH3 domain pseudo wild type W37F.[10]

To further characterize the double-exponential refolding subsequent to the burst phase, we also monitored tryptophan fluorescence under the same conditions as in the CD experiments (Figure 5B). The resolvable refolding signal for both CD and fluorescence detection could be fitted globally to the same double exponential (Table 2). Unlike the CD signal, fluorescence is non-monotonic, clearly showing the need for a double-exponential fit. Before reaching the native state the Fyn mutant proceeds through a hyperfluorescent stage, which has been interpreted as a signature of a highly flexible native-like conformation in other proteins.[44]

We also measured SAXS of Fyn SH3 domain (Figure 6). Guinier plots in 0 and 4 M GuHCl at 4 °C are shown in Figure 6B. R_g was 14.6 ± 0.8 Å and 27.3 ± 3.9 Å, respectively, confirming that the Fyn mutant is denatured in >4 M GuHCl. The 14.6 Å value is in good agreement with the native value of the protein reported previously.[28] The 27.3 Å is also in good agreement with the denatured values of src and PI3K SH3 domain as described above.[10]

Kinetic refolding at −28 °C and monitored by SAXS was initiated by cryo-stopped flow (Figure 6A). Within 0.1 to 0.2 s, the R_g approaches the native value.

The CD, fluorescence, and SAXS data suggest a similar picture for the Fyn mutant as for the PI3K SH3 domain. First, the protein rapidly collapses and forms excess helical structure (<6 ms). Then the CD signal at 222 nm approaches the smaller native magnitude. Unlike PI3K SH3 domain, the Fyn mutant recovers from the compact excess helical state in two steps. Thus, even neglecting the rapid collapse and secondary structure formation at <6 ms, the Fyn mutant, like the pseudo wild type, is not an apparent two-state folder.

■ DISCUSSION

Quite an array of data has accumulated to suggest that the formation of compact α-helical intermediates during protein folding is commonplace.[7,9−15,45,46] Even though such intermediates are not populated at physiological temperature, they are easily elicited in cryogenic solvent, proving that they are not much higher in free energy than the native state. The cryogenic conditions do not necessarily elicit helical intermediates by destabilizing the native state relative to the unfolded state: in Figure 2A the PI3K SH3 domain denaturation midpoints are very similar at 4 and −20 °C.

The transient population of compact helix-rich states early on during folding should not be surprising: Rapid protein collapse is driven by hydrophobicity (albeit less so at low temperatures[47]), and α-helical formation requires only local (i to i + 4)

dx.doi.org/10.1021/jp400167s | J. Phys. Chem. B 2013, 117, 4836−4843

contacts, which can be explored by the unfolded chain on very fast time scales.

In this context, it is worth noting that a β-sheet protein has been reported that forms two different structures from the same sequence in equilibrium (one with only β-sheet, the other with some helix).[48] The more α-helix-rich structure has a different biological function than the more β-rich structure. Such proteins may well have evolved starting from transient helix-rich states that lie at low free energy relative to the native state.

SH3 domain is the latest entry in the series of proteins that can form α-helix rich compact intermediates: three different sequences or mutants of this domain (Fyn, src, and PI3K) transiently populate states with excess helical structure (Table 2). The 222 nm ellipticities of the burst phase of src SH3 domain,[10] Fyn mutant, and PI3K SH3 domain all had values much greater in magnitude than the native state. The transient helix-rich intermediate "I_1" is not specific to any variant.

This is not to say that all three sequences fold in the same way: src SH3 domain and Fyn SH3 domain mutants have very similar sequences (Table 1). Their helical intermediates have similar compactness and excess helical content (Table 2). The PI3K sequence is quite different. Even though its AGADIR helical score is slightly higher (Table 1),[49] the burst phase of PI3K SH3 domain is significantly smaller than that of the other two SH3 domains. The Fyn mutant and PI3K SH3 domain also differ in other ways. The latter folds in a single exponential step to the native state after the helical burst that populates I_1. The Fyn mutant populates an additional intermediate I_2 with more native-like helix content, and folds more rapidly overall. All observations are consistent with the following kinetic scheme for SH3 domain folding: From the GuHCl denatured state U, I_1 is a compact helix-rich intermediate rapidly populated in cryosolvent. From I_1, the native state N forms slowly, either in one step by losing the non-native helical structure of I_1, or through a second intermediate I_2. Whether I_1 is a trap that has to unfold completely before acquiring native structure, or just partly off-pathway because excess helix can be lost without increasing the R_g too much, it is easily accessible on the free energy landscape

Evolutionary selection should filter out states like I_1 if they reduce protein function or lead to aggregation. Clearly, evolution of SH3 domain proteins has not selected against such states strongly: they are populated easily during folding by adding a carbohydrate-like cosolute (ethylene glycol), and/or $a \leq 0.1\ k_BT$ decrease of the available thermal energy. Indeed, the interior of a cell is not an aqueous buffer, but filled with carbohydrates, nucleic acids, and other crowders that influence protein stability and kinetics;[50] a 45% ethylene glycol/water mixture is likely no further removed from the cytoplasmic environment than the aqueous phosphate buffer commonly used for *in vitro* protein folding studies.

In silico, metastable helical states are also commonly observed. For example, the β-sheet WW domain was shown to fold to a helical state by the CHARMm force field, and the force field was modified to better represent β-sheet hydrogen bonding to correct this.[23,51] Ref 24 reports further recent results and reviews some of this literature. In simulations, helical bursts are generally considered a nuisance, but our results here, and those for other proteins, indicate just how subtle the nuisance is: Helical states really exist that lie low on the free energy landscape, and force fields will have to be delicately balanced to place them at the appropriate free energy. More data on metastable states containing non-native helix could provide benchmarks for a further level of force field refinement, which takes into account not just native and denatured ensembles, but also metastable ensembles populated during folding.

Recently, several computational papers have suggested functional properties of kinetic helical intermediates. Chikenji and Kikuchi[52] concluded from a simple polar−nonpolar model that non-native α helix restricts the conformational search and guides folding of β-lactoglobulin. Mohanty et al.[53] reported that the N-terminus of Top7-CFr initially forms a helical extension to hide a native β-sheet that could lead to harmful aggregation during folding. Such reports point out that non-native helix could assist folding instead of simply delaying it.

While our cryogenic conditions foster rapid formation of helix-rich intermediates, folding to the native state is generally very slow (seconds) in cryogenic experiments. Microsecond pressure jumps hold promise for showing that formation of excess helical structure can occur on a microsecond time scale, where comparison with simulation is currently feasible.[25]

CONCLUSION

Combined with literature results, four different sequences of SH3 domain, including the PI3K sequence that is very different from src- and Fyn-based sequences, all show transient helical populations in cryogenic ethylene glycol/water solvent. At the same time, this solvent only slightly stabilizes the proteins overall. Subtle tuning of these proteins' free energy landscapes (by a few k_BT) sufficiently stabilizes helix over sheet structure for the former to become transiently populated: the fast kinetics of local contacts briefly dominates over the more favorable free energy of native structure. Non-native "overshoots" will be useful for calibrating molecular dynamics force fields' helix/sheet propensities, if such "overshoots", known from our data to occur in <6 ms, are resolved on the submillisecond time scale.

AUTHOR INFORMATION

Corresponding Author

*E-mail: shinjom@hirakata.kmu.ac.jp. Tel.: +81 72 804 2358. Fax: +81 72 850 0733.

Present Address

S.J.K.: Department of Bioengineering and Therapeutic Sciences, University of California, 1700 Fourth Street, San Francisco, CA 94158, USA. Y.M.: Research & Development Division, Kobe Laboratory, TAO Health Life Pharma Co., Ltd. Institute of Biomedical Research and Innovation, Kobe Hybrid Business Center, 6-7-6, Minamimachi, Chuo-ku, Kobe, 650-0047, Japan.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are grateful to Professor Stuart L. Schreiber of Harvard University for kindly providing the expression vector containing PI3K SH3 domain, and to Dr. Masayuki Morita of Kansai Medical University for kindly helping with *E. coli* expression. We thank Dr. Liang Guo at Argonne National Laboratory helpful assistance in setting up the beamline for experiments, and Charles Dumont for help with data collection. This study was supported by a Grant-in-Aid for Scientific Research (no. 20540400) from the Ministry of Education, Culture, Sports, Science, and Technology. M.G. and S.J.K. were supported by National Institutes of Health grant R01-

GM093318. Time on BioCAT-18, managed by Prof. Thomas Irving, was supported by proposal GUP-6360 of the Advanced Photon Source at the Argonne National Laboratory. SAXS measurements were performed under proposal number 2006G407 and 2008G684 of the Photon Factory.

REFERENCES

(1) Poland, D.; Scheraga, H. A. *Theory of Helix-Coil Transitions in Biopolymers*; Academic Press: New York, 1970.
(2) Yao, J.; Dyson, H. J.; Wright, P. E. Chemical Shift Dispersion and Secondary Structure Prediction in Unfolded and Partly Folded Proteins. *FEBS Lett.* **1997**, *419*, 285−289.
(3) Kuroda, Y.; Hamada, D.; Tanaka, T.; Goto, Y. High Helicity of Peptide Fragments Corresponding to Beta-Strand Regions of Beta-Lactoglobulin Observed by 2d-Nmr Spectroscopy. *Folding Des.* **1996**, *1*, 255−263.
(4) Dyson, H. J.; Merutka, G.; Waltho, J. P.; Lerner, R. A.; Wright, P. E. Folding of Peptide Fragments Comprising the Complete Sequence of Proteins Models for Initiation of Protein Folding I. Myohemerythrin. *J. Mol. Biol.* **1991**, *226*, 795−817.
(5) Ikeguchi, M.; Kato, S.; Shimizu, A.; Sugai, S. Molten Globule State of Equine Beta-Lactoglobulin. *Proteins* **1997**, *27*, 567−575.
(6) Voelz, V.; Bowman, G.; Beauchamp, K.; Pande, V. Molecular Simulation of Ab Initio Protein Folding for a Millisecond Folder Ntl9(1−39). *J. Am. Chem. Soc.* **2010**, *132*, 1526−1528.
(7) Kuwajima, K.; Yamaya, H.; Sugai, S. The Burst-Phase Intermediate in the Refolding of Beta-Lactoglobulin Studied by Stopped-Flow Circular Dichroism and Absorption Spectroscopy. *J. Mol. Biol.* **1996**, *264*, 806−822.
(8) Arai, M.; Ikura, T.; Semisotnov, G. V.; Kihara, H.; Amemiya, V.; Kuwajima, K. Kinetic Refolding of Beta-Lactoglobulin - Studies by Synchrotron X-Ray Scattering, and Circular Dichroism, Absorption and Fluorescence Spectroscopy. *J. Mol. Biol.* **1998**, *275*, 149−162.
(9) Qin, Z.; Hu, D.; Shimada, L.; Nakagawa, T.; Arai, M.; Zhou, J.; Kihara, H. Refolding of Beta-Lactoglobulin Studied by Stopped-Flow Circular Dichroism at Subzero Temperatures. *FEBS Lett.* **2001**, *507*, 299−302.
(10) Li, J. S.; Shinjo, M.; Matsumura, Y.; Morita, M.; Baker, D.; Ikeguchi, M.; Kihara, H. An Alpha-Helical Burst in the Src SH3 Folding Pathway. *Biochemistry* **2007**, *46*, 5072−5082.
(11) Matsumura, Y.; Shinjo, M.; Mahajan, A.; Tsai, M. D.; Kihara, H. Alpha-Helical Burst on the Folding Pathway of Fha Domains from Rad53 and Ki67. *Biochimie* **2010**, *92*, 1031−1039.
(12) Larios, E.; Li, J. S.; Schulten, K.; Kihara, H.; Gruebele, M. Multiple Probes Reveal a Native-Like Intermediate During Low-Temperature Refolding of Ubiquitin. *J. Mol. Biol.* **2004**, *340*, 115−125.
(13) Qin, Z.; Ervin, J.; Larios, E.; Gruebele, M.; Kihara, H. Formation of a Compact Structured Ensemble without Fluorescence Signature Early During Ubiquitin Folding. *J. Phys. Chem. B* **2002**, *106*, 13040−13046.
(14) Dumont, C.; Matsumura, Y.; Kim, S. J.; Li, J.; Kondrashkina, E.; Kihara, H.; Gruebele, M. Solvent-Tuning Collapse and Helix Formation Time Scales of Lambda6−85. *Protein Sci.* **2006**, *15*, 2596−2604.
(15) Kim, S. J.; Matsumura, Y.; Dumont, C.; Kihara, H.; Gruebele, M. Slowing Down Downhill Folding: A Three-Probe Study. *Biophys. J.* **2009**, *97*, 295−302.
(16) Noble, M. E. M.; Musacchio, A.; Saraste, M.; Courtneidge, S. A.; Wierenga, R. K. Crystal Structure of the SH3 Domain in Human Fyn; Comparison of the Three-Dimensional Structures of SH3 Domains in Tyrosine Kinases and Spectrin. *EMBO J.* **1993**, *12*, 2617−2624.
(17) Grantcharova, V. P.; Riddle, D. S.; Santiago, J. V.; Baker, D. Important Role of Hydrogen Bonds in the Structurally Polarized Transition State for Folding of the Src SH3 Domain. *Nat. Struct. Biol.* **1998**, *5*, 714−720.
(18) Riddle, D. S.; Grantcharova, V. P.; Santiago, J. V.; Alm, E.; Ruczinski, I.; Baker, D. Experiment and Theory Highlight Role of Native State Topology in SH3 Folding. *Nat. Struct. Biol.* **1999**, *6*, 1016−1024.
(19) Grantcharova, V. P.; Baker, D. Circularization Changes the Folding Transition State of the Src SH3 Domain. *J. Mol. Biol.* **2001**, *306*, 555−563.
(20) Dasgupta, A.; Udgaonkar, J. B. Evidence for Initial Non-Specific Polypeptide Chain Collapse During the Refolding of the SH3 Domain of PI3 Kinase. *J. Mol. Biol.* **2010**, *403*, 430−445.
(21) Guijarro, J. I.; Morton, C. J.; Plaxco, K. W.; Campbell, I. D.; Dobson, C. M. Folding Kinetics of the SH3 Domain of PI3 Kinase by Real-Time NMR Combined with Optical Spectroscopy. *J. Mol. Biol.* **1998**, *276*, 657−667.
(22) Shell, M. S.; Ritterson, R.; Dill, K. A. A Test on Peptide Stability of Amber Force Fields with Implicit Solvation. *J. Phys. Chem. B* **2008**, *112*, 6878−6886.
(23) Freddolino, P. L.; Park, S.; Roux, B.; Schulten, K. Force Field Bias in Protein Folding Simulations. *Biophys. J.* **2009**, *96*, 3772−3780.
(24) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Systematic Validation of Protein Force Fields against Experimental Data. *Plos One* **2012**, *7*, e32131.
(25) Prigozhin, M.; Liu, Y.; Kapoor, S.; Winter, R.; Schulten, K.; Gruebele, M. Misplaced Helix Slows Down Ultrafast Pressure-Jump Refolding of Lambda6−85. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, www.pnas.org/cgi/doi/10.1073/pnas.1219163110.
(26) Piana, S.; Sarkar, K.; Lindorff-Larsen, K.; Guo, M.; Gruebele, M.; Shaw, D. E. Computational Design and Experimental Testing of the Fastest-Folding Beta-Sheet Protein. *J. Mol. Biol.* **2011**, *405*, 43−48.
(27) Okishio, N.; Nagai, M.; Fukuda, R.; Nagatomo, S.; Kitagawa, T. Interactions of Phosphatidylinositol 3-Kinase Src Homology 3 Domain with Its Ligand Peptide Studied by Absorption, Circular Dichroism, and Uv Resonance Raman Spectroscopies. *Bioploymers* **2000**, *57*, 208−217.
(28) Kim, S. J.; Dumont, C.; Gruebele, M. Simulation-Based Fitting of Protein-Protein Interaction Potentials to Saxs Experiments. *Biophys. J.* **2008**, *94*, 4924−4931.
(29) Pace, C. N. Determination and Analysis of Urea and Guanidine Hydrochloride Denaturation Curves. *Methods Enzymol.* **1986**, *131*, 266−280.
(30) Nozaki, Y. The Preparation of Guanidine Hydrochloride. *Methods Enzymol.* **1970**, *26*, 43−50.
(31) Kihara, H. Stopped-Flow Apparatus for X-Ray Scattering and Xafs. *J. Synchrotron Radiat.* **1994**, *1*, 74−77.
(32) Matsumura, K.; Enoki, Y.; Kohzuki, H.; Sakata, S. A Simple Procedure for Determination of the Dead Time of a Stopped-Flow Instrument. *Jpn. J. Physiol.* **1990**, *40*, 567−571.
(33) Semisotnov, G. V.; Kihara, H.; Kotova, N. V.; Kimura, K.; Amemiya, Y.; Wakabayashi, K.; Serdyuk, I. N.; Timchenko, A. A.; Chiba, K.; Nikaido, K.; et al. Protein Globularization During Folding - a Study by Synchrotron Small-Angle X-Ray Scattering. *J. Mol. Biol.* **1996**, *262*, 559−574.
(34) Amemiya, Y.; Ito, K.; Yagi, N.; Asano, Y.; Wakabayashi, K.; Ueki, T.; Endo, T. Large-Aperture TV Detector with a Beryllium-Windowed Image Intensifier for X-Ray-Diffraction. *Rev. Sci. Instrum.* **1995**, *66*, 2290−2294.
(35) Amemiya, Y.; Wakabayashi, K.; Hamanaka, T.; Wakabayashi, T.; Matsushita, T.; Hashizume, H. *Nucl. Instrum. Methods* **1983**, *208*, 471−477.
(36) Ito, K.; Kamikubo, H.; Yagi, N.; Amemiya, Y. Correction Method and Software for Image Distortion and Nonuniform Response in Charge-Coupled Device-Based X-Ray Detectors Utilizing X-Ray Image Intensifier. *Jpn. J. Appl. Phys.* **2005**, *44*, 8684−8691.
(37) Ito, K.; Kamikubo, H.; Arai, M.; Kuwajima, K.; Amemiya, Y.; Endo, T. Calibration Method for the Contrast Reduction Problem in X-Ray Image Intensifiers. *Photon Factory Activity Rep.* **2001**, *18*, 275−281.
(38) Eliezer, D.; Chiba, K.; Tsuruta, H.; Doniach, S.; Hodgson, K. O.; Kihara, H. Evidence of an Associative Intermediate on the Myoglobin Refolding Pathway. *Biophys. J.* **1993**, 912−917.

(39) Arai, M.; Ito, K.; Inobe, T.; Nakao, M.; Maki, K.; Kamagata, K.; Kihara, H.; Amemiya, Y.; Kuwajima, K. Fast Compaction of Alpha-Lactalbumin During Folding Studied by Stopped-Flow X-Ray Scattering. *J. Mol. Biol.* **2002**, *321*, 121−132.
(40) Fischetti, R.; Stepanov, S.; Rosenbaum, G.; Barrea, R.; Black, E.; Gore, D.; Heurich, R.; Kondrashkina, E.; Kropf, A. J.; Wang, S.; et al. The Biocat Undulator Beamline 18id: A Facility for Biological Non-Crystalline Diffraction and X-Ray Absorption Spectroscopy at the Advanced Photon Source. *J. Synchrotron Rad.* **2004**, *11*, 399−405.
(41) Guinier, A.; Fournet, G. *Small-Angle Scattering of X-Rays*; John Wiley & Sons, Inc.: New York, 1955.
(42) Kihara, H.; Nagamura, T. Development of a Static Flow Cell for the Measurement of SAXS. *Photon Factory Activity Report* **1999**, *16*, 246−249.
(43) Di Nardo, A. A.; Larson, S. M.; Davidson, A. R. The Relationship between Conservation, Thermodynamic Stability, and Function in the SH3 Domain Hydrophobic Core. *J. Mol. Biol.* **2003**, *333*, 641−655.
(44) Ervin, J.; Larios, E.; Osvath, S.; Schulten, K.; Gruebele, M. What Causes Hyperfluorescence: Folding Intermediates or Conformationally Flexible Native States? *Biophys. J.* **2002**, *83*, 473−483.
(45) Qin, Z. J.; Vyas, S.; Fink, A. L.; Li, J. S.; Kihara, H. Transient Alpha-Helical Structure During Folding of Src SH3 Domain at Subzero Temperatures. *J. Kansai Med. Univ.* **2006**, *58*, 163−169.
(46) Matsumura, Y.; Li, J.; Ikeguchi, M.; Kihara, H. Helix-Rich Transient and Equilibrium Intermediates of Equine Beta-Lactoglobulin in Alkaline Buffer. *Biophys. Chem.* **2008**, *134*, 84−92.
(47) Southall, N. T.; Dill, K. A.; Haymet, A. D. J. A View of the Hydrophobic Effect. *J. Phys. Chem. B* **2002**, *106*, 521−533.
(48) Tuinstra, R. L.; Peterson, F. C.; Kutlesa, S.; Elgin, E. S.; Kron, M. A.; Volkman, B. F. Interconversion between Two Unrelated Protein Folds in the Lymphotactin Native State. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 5057−5062.
(49) Muñoz, V.; Serrano, L. Development of the Multiple Sequence Approximation within the Agadir Model of Alpha-Helix Formation. Comparison with Zimm-Bragg and Lifson-Roig Formalisms. *Biopolymers* **1997**, *41*, 495−509.
(50) Ebbinghaus, S.; Dhar, A.; McDonald, J. D.; Gruebele, M. Protein Folding Stability and Dynamics Imaged in a Living Cell. *Nat. Methods* **2010**, *7*, 319−323.
(51) Freddolino, P. L.; Liu, F.; Gruebele, M.; Schulten, K. Ten-Microsecond Molecular Dynamics Simulation of a Fast-Folding WW Domain. *Biophys. J.* **2008**, *94*, L75−L77.
(52) Chikenji, G.; Kikuchi, M. What is the Role of Non-Native Intermediates of Beta-Lactoglobulin in Protein Folding. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 14273−14277.
(53) Mohanty, S.; Meinke, J. H.; Zimmermann, O.; Hansmann, U. H. E. Simulation of Top7-Cfr: A Transient Helix Extension Guides Folding. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 8004−8007.