

Structural Events during the Refolding of an All β -Sheet Protein*

Received for publication, July 6, 2000, and in revised form, October 18, 2000
Published, JBC Papers in Press, October 18, 2000, DOI 10.1074/jbc.M005921200

Dharmaraj Samuel‡, Thallampuranam Krishnaswamy Suresh Kumar‡,
Krishnaswamy Balamurugan§, Wann-Yin Lin¶, Der-Hang Chin§, and Chin Yu‡||

From the ‡Department of Chemistry, National Tsing Hua University, Hsinchu 300, the §Department of Chemistry, National Changhua University of Education, Changhua 500, and the ¶Department of Chemistry, National Taiwan University, Taipei 106, Taiwan

The refolding kinetics of the 140-residue, all β -sheet, human fibroblast growth factor (hFGF-1) is studied using a variety of biophysical techniques such as stopped-flow fluorescence, stopped-flow circular dichroism, and quenched-flow hydrogen exchange in conjunction with multidimensional NMR spectroscopy. Urea-induced unfolding of hFGF-1 under equilibrium conditions reveals that the protein folds via a two-state (native $\leftrightarrow$ unfolded) mechanism without the accumulation of stable intermediates. However, measurement of the unfolding and refolding rates in various concentrations of urea shows that the refolding of hFGF-1 proceeds through accumulation of kinetic intermediates. Results of the quenched-flow hydrogen exchange experiments reveal that the hydrogen bonds linking the N- and C-terminal ends are the first to form during the refolding of hFGF-1. The basic β -trefoil framework is provided by the simultaneous formation of β -strands I, IV, IX, and X. The other β -strands comprising the β -barrel structure of hFGF-1 are formed relatively slowly with time constants ranging from 4 to 13 s.

It is still unknown as to how a polypeptide chain folds into a unique structure out of an astronomical number of energetically possible conformations in a short period of time (1–3). Characterization of the conformational properties of the early stages sampled during refolding are crucial to our understanding the rules determining the folding process (4–8). Much of our knowledge on the kinetics of refolding/unfolding is derived employing the stopped-flow optical spectroscopic methods or the quenched-flow hydrogen/deuterium exchange techniques (9–14). By using these techniques, several proteins have been shown to fold via the formation of transient kinetic intermediate(s) (15–19). Understanding the structural properties of the intermediates that occur in the folding/unfolding pathway(s) would explain how the vast conformational space available could be narrowed down early during folding to yield the native state of the protein (20–22).

The refolding kinetics of a number of all-helical proteins have been investigated (23, 24). It is found that helix formation in proteins occurs within the low millisecond scale of refolding (24). This aspect is not surprising because formation of a helix

involves intrastrand, local hydrogen bonding interactions. By contrast, very little information exists on the modes of refolding of all β -sheet proteins. As the β -sheet formation involves inter-strand interactions between distant parts of the polypeptide chain, proteins belonging to this structural class (all β -sheet proteins) are predicted to refold slowly (25). However, investigation of folding kinetics of several $\alpha + \beta$ proteins using quenched-flow hydrogen/deuterium (H/D) exchange revealed that stable β -sheet, and α -helix formation occurs on a similar time scale (18, 19, 26). Interestingly, folding studies on interleukin-1 β , an all β -sheet protein, showed that stable β -sheet formation only occurs on a time scale greater than 50 s (25, 27). Thus, it appears that the data base on the refolding kinetics of all β -sheet proteins needs to be significantly expanded to draw generalized conclusions on the rates of formation of β -sheets in proteins (28, 29).

Human acidic fibroblast growth factor (hFGF-1)¹ is a single chain, heparin-binding protein involved in a variety of important cellular processes like the proliferation and differentiation of cells (30, 31). hFGF-1 is a potent mitogen, and due to its involvement in the wound repair process, it is a potential therapeutic agent (31). hFGF-1 is a 15-kDa, all β -sheet protein with no disulfide bonds (32, 33). It has a single tryptophan residue, whose emission properties are known to describe effectively the conformational changes occurring during the folding $\leftrightarrow$ unfolding transition of the protein (33, 34). Furthermore, high resolution crystal (31) and NMR structures (32) of hFGF-1 are available (Fig. 1). These characteristics render hFGF-1 as an useful model to understand the folding/unfolding pathways of all β -sheet proteins. In the present study, we investigate the events in the refolding kinetics of hFGF-1 using a variety of biophysical techniques such as stopped-flow fluorescence, stopped-flow CD, and quenched-flow hydrogen exchange. The results obtained in this study reveal that refolding of the protein occurs very slowly via the formation of transient intermediate(s). In addition, the β -strands constituting the heparin binding domain appear to form significantly faster than the other β -strands constituting the β -barrel structure of hFGF-1.

MATERIALS AND METHODS

1-Anilino-8-naphthalenesulfonic acid-NH₄⁺ salt was purchased from Sigma. Ultrapure urea was procured from Merck, and heparin-Sepharose was from Amersham Pharmacia Biotech. Labeled ¹⁵NH₄Cl and urea-d₄ were obtained from Cambridge Isotope Laboratories. All other chemicals used were of high quality analytical grade. Unless otherwise mentioned, all solutions were made in 100 mM phosphate buffer (pH 7.0) containing 100 mM ammonium sulfate. All experiments were performed at 20 °C.

Protein Expression and Purification—Expression vector for the truncated form of the human FGF-1 (hFGF-1, residues 15–154) was con-

* This work was supported by research grants from National Science Council, Taiwan, Dr. C. S. Tsong Memorial Medical Research Foundation (to C. Y.), and the National Health Research Institute, Taiwan (to D.-H. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

|| To whom correspondence should be addressed. Tel.: 886 35 715131 (ext. 5605); Fax: 886 35 711082; E-mail: cyu@mx.nthu.edu.tw.

¹ The abbreviations used are: hFGF, human fibroblast growth factor; ANS, 1-anilino-8-naphthalene sulfonate; H/D, hydrogen/deuterium.

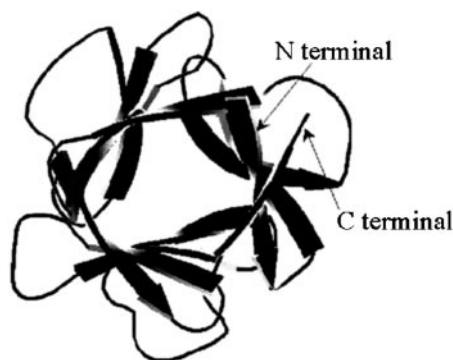


FIG. 1. MOLSCRIPT (50) representation of the structure of hFGF-1. The protein has 12 β -strands folded into a trigonal pyramid structure.

structed and inserted between the *Nde*I and *Bam*HI restriction sites in pET20b. The plasmid containing the hFGF-1 insert was transformed into *Escherichia coli* BL21 (DE3) pLysS. The expressed protein was purified on heparin-Sepharose column using a NaCl gradient (0–1.5 M). The protein was desalted by ultrafiltration using an Amicon set up. The homogeneity of the protein was assessed using SDS-polyacrylamide gel electrophoresis. The authenticity of the sample was further verified by electron spray-mass analysis. The concentration of the protein was estimated from the extinction coefficient value at 280 nm.

Preparation of Isotope-enriched hFGF-1—Uniform ^{15}N isotope labeling was achieved using M9 minimal medium containing $^{15}\text{NH}_4\text{Cl}$. To realize maximal expression yields, the composition of the M9 medium was modified by the addition of a mixture mixture of vitamins. The expression host strain *E. coli* BL21(DE3)pLysS is a vitamin B_1 -deficient host, and hence the medium was supplemented with thiamine (vitamin B_1). Protein expression yields were in the range of 25–30 mg/liter of the isotope-enriched medium. The extent of ^{15}N labeling was verified by electron spray-mass analysis.

Non-fluorescence Measurements—All fluorescence experiments were carried out on a Hitachi F-4500 spectrofluorimeter using a 1-cm quartz cell. The final concentration of the protein used was 100 $\mu\text{g}/\text{ml}$. Intrinsic tryptophan fluorescence was measured using an excitation wavelength of 280 nm. The tryptophan emission was monitored between 300 and 450 nm. The excitation and emission slit widths were set at 2.5 and 10 nm, respectively. The fraction of unfolded species formed at various concentrations of urea were estimated based on the 308 to 350 nm ratio.

Far-UV CD Measurements—All far-UV CD experiments were performed on a Jasco J-720 spectropolarimeter using a quartz cell of 1-cm path length. Each spectra was an average of 5 scans. The concentration of the protein used was 0.5 mg/ml. Necessary background corrections were made in all spectra.

Stopped-flow Fluorescence—The kinetics of unfolding and refolding were monitored by fluorescence using an SF-61 stopped-flow spectrofluorimeter (Hi-Tech Scientific Co., UK). Intrinsic tryptophan fluorescence measurements were made by exciting the protein at 280 nm (2.5 nm band pass) via a 0.2-cm path length cell and monitoring the emission intensity at 350 nm (at 20 $^\circ\text{C}$). For 1:10 mixing ratios in both unfolding or refolding experiments gave a final protein concentration of 1 μM hFGF-1 in 100 mM phosphate buffer (pH 7.0), at 20 $^\circ\text{C}$ (containing 100 mM ammonium sulfate), in different urea concentrations. From 0.2–5 μM protein, no deviations longer than 5% were detected in the rates or relative amplitudes. The data were fit to a two exponential equation and analyzed using the RKBIN software provided by Hi-tech Scientific Co., UK.

For the stopped-flow ANS binding experiments, the stock solution of ANS (300 μM) was appropriately mixed (in the concentration range between 50 and 250 μM) into the native buffer (100 mM phosphate buffer containing 100 mM ammonium sulfate (pH 7.0)). 10 μM of equilibrium-unfolded hFGF-1 (in 4 M urea) was diluted 10-fold into 270 μl of the native buffer containing ANS. This procedure ensured that the concentration of the protein (1 μM) and urea (0.4 M) remained unchanged during the refolding reaction in the presence of different concentrations of ANS. The binding of ANS was monitored by setting the excitation and emission wavelengths at 418 and 517 nm, respectively. The bandwidth used in all the measurements was 2 nm.

The protein concentration-dependent ANS binding experiments were carried out under similar conditions by varying the protein concentration (in the range of 0.5 to 5 μM).

Stopped-flow CD Measurements—Stopped-flow CD experiments were performed using SFM-3 Biologic stopped-flow apparatus attached to a Jasco J-715 spectropolarimeter. The cell path length was 1 mm, and the dead time of the instrument was estimated to be about 5 ms. Refolding was initiated by a 10-fold dilution of the denatured hFGF-1 (in 4 M urea) sample with the refolding buffer (100 mM phosphate buffer containing 100 mM ammonium sulfate (pH 7.0), 20 $^\circ\text{C}$). The data presented are an average of 30 scans. The reaction between sodium salt of 2,6-dichlorophenol indophenol and L-ascorbic acid (using a mixing ratio of 1:1) was used to determine the performance limits of the instrument. The disappearance of the color of 2,6-dichlorophenol indophenol was followed at 524 nm. All solutions contained 0.2 M NaCl to minimize the ionic strength effects.

The ellipticity changes at 228 nm were analyzed employing an in-built Biologic software.

Quenched-flow NMR Experiments—All experiments were carried out at 20 ± 0.1 $^\circ\text{C}$ using a RQF-63 rapid mixing quenched-flow apparatus (Hi-Tech Scientific Co., UK). Complete denaturation and exchange of the backbone amide protons with deuterium was achieved by dissolving 4 M urea into a uniformly ^{15}N labeled hFGF-1 solution, previously equilibrated for 24 h in D_2O at pD 6.6 ± 0.2 . Refolding of the denatured protein was initiated by a 10-fold dilution with 100 mM phosphate buffer (pH 5.0, containing 100 mM ammonium sulfate) prepared in H_2O . At this pH, negligible labeling occurred. After variable refolding times ranging from 8 ms to 100 s, the solution was diluted again to 10 times the initial protein volume with 100 mM phosphate buffer (pH 9.0) (containing 100 mM ammonium sulfate), to initiate labeling of the deuterated amides in hFGF-1 with protons. After a lapse of 10 ms, the labeling phase was stopped by a further 3.4 times dilution of the initial protein volume with 1 M phosphoric acid in water. It should be mentioned that there were no signs of aggregation during any of the refolding steps in the quenched-flow experiments.

^1H - ^{15}N HSQC spectrum of each sample was recorded on a Bruker DMX 600-MHz NMR spectrometer at 20 $^\circ\text{C}$. A 5-mm inverse probe with a self-shielded z -gradient was used to obtain all gradient-enhanced ^1H - ^{15}N HSQC spectra. ^{15}N decoupling during acquisition were accomplished using the GARP sequence. 2048 complex data points were collected in the ^1H dimension of the ^1H - ^{15}N HSQC experiments. In the ^{15}N dimension of the spectra, 512 complex data points were obtained. ^{15}N chemical shifts were referenced using the consensus ratio of 0.0101329118. All spectra were processed on a Silicon Graphics Work station using the UX NMR and Aurelia softwares.

The intensities of the cross-peaks in the HSQC spectra collected at various refolding times were normalized based on the peak height of the Ile-70 γ -methyl protons (-0.2 ppm) in the one-dimensional ^1H NMR spectra collected each time prior to recording the HSQC spectra. The time courses of change in proton occupancies were fitted to a single exponential decay ($y = A \exp^{-kt} + C$, where A is the amplitude of the phase, k is the apparent rate constant, and C is the final amplitude) by the Levenberg-Marquardt nonlinear least squares method, yielding rate constants and phase amplitudes in the kinetic refolding experiment. All data analysis was performed using Kaleidagraph software (Synergy software).

RESULTS AND DISCUSSION

Two-state Equilibrium Unfolding—The fluorescence spectrum of hFGF-1 shows an emission maxima at around 308 nm (Fig. 2, inset A). The fluorescence of the lone tryptophan residue located at position 121 of the sequence is completely quenched in the native state of the protein (33, 34). The quenching effect is attributed to the presence of proximal imidazole and pyrrole groups in the three-dimensional structure of hFGF-1. However, unfolding of the protein relieves the quenching effect, and the fluorescence spectrum of the protein in the unfolded state shows an emission maxima at around 350 nm (Fig. 2, inset A). Hence, the refolding/unfolding kinetics of hFGF-1 was monitored by the changes in the emission intensity at 350 nm.

Fig. 2 shows the urea-induced unfolding curve of hFGF-1 monitored by fluorescence spectroscopy. The protein is completely unfolded at concentrations of urea greater than 3.2 M. The urea-induced structural transitions are completely reversible. The concentration of denaturant (C_m) at which the protein (hFGF-1) is half unfolded ($\Delta G_u = 0$) is estimated to be 2.3 M.

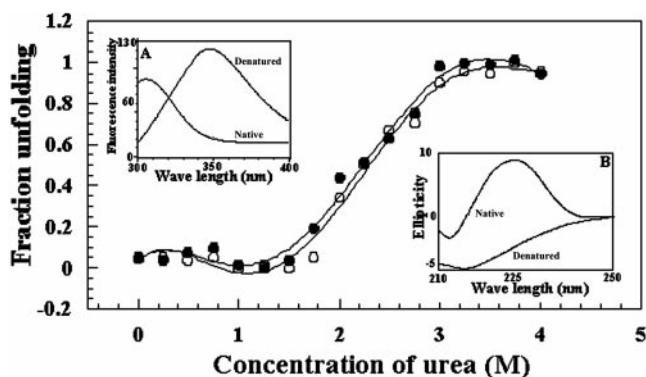


FIG. 2. Urea-induced unfolding profiles of hFGF-1, monitored by fluorescence (●) (308/350 nm) and far-UV CD (○) (228 nm). It could be noticed that the unfolding profiles monitored by the two spectroscopic techniques are nearly superimposable, implying that the urea-induced unfolding process of hFGF-1 is cooperative without the accumulation of stable equilibrium intermediates. Inset A depicts the intrinsic fluorescence spectra of hFGF-1 in the native and 4 M urea-unfolded states. Inset B shows the far-UV CD spectra of hFGF-1 in the native and 4 M urea unfolded states.

The m value, which is a measure of the cooperativity of the unfolding transition, is found to be $1.53 \pm 0.4 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. The free energy of unfolding in the absence of the denaturant ($\Delta G_u(\text{H}_2\text{O})$) obtained by extrapolation of ΔG_u to zero denaturant concentration is calculated to be $3.46 \pm 0.2 \text{ kcal} \cdot \text{mol}^{-1}$. To test if the unfolding transition monitored by fluorescence reflected a total disruption of the overall structure of the protein, or just a local unfolding, we analyzed the chemical denaturation of hFGF-1 induced by urea using far-UV CD spectroscopy. The far-UV CD spectrum of hFGF-1 shows a prominent positive ellipticity band centered at 228 nm. This CD band (at 228 nm) is a composite signal and is representative of the secondary and tertiary structural interactions in the protein (Fig. 2, inset B). Upon complete unfolding of the protein (at urea concentrations greater than 3.2 M), the 228 nm far-UV CD band shows negative ellipticity values (Fig. 2, inset B). The urea-induced equilibrium unfolding curves followed by fluorescence and CD are nearly superimposable. Within experimental error(s), the $\Delta G_u(\text{H}_2\text{O})$ ($3.40 \pm 0.4 \text{ kcal} \cdot \text{mol}^{-1}$) estimated from the urea-induced far-UV ellipticity is identical to that calculated based on fluorescence spectroscopy. These results strongly suggest that the urea-induced equilibrium denaturation of hFGF-1 follows a two-state (native $\leftrightarrow$ unfolded state) mechanism, without the accumulation of stable intermediate(s).

Slow Refolding hFGF-1—Fig. 3 shows the stopped-flow time traces of the refolding of urea-denatured hFGF-1 (at 25 °C) monitored by the changes in the tryptophan fluorescence (at 350 nm). Complete signal evolution (at 350 nm) occurs within a refolding time of 50 s. The fluorescence trace could not be satisfactorily fit to a single exponential function. A sum of two exponentials best fits the refolding curve yielding refolding rate constant values of 0.131 ± 0.02 and $0.023 \pm 0.07 \text{ s}^{-1}$ for the faster and slower phases, respectively (Table I). The first phase associated with the faster rate accounts for 80–85% of the total refolding amplitude. The slower phase has an amplitude of about 10–20%. This phase of refolding is probably associated with the *cis-trans*-proline isomerization (discussed below). No burst phase is observed to occur as is evident from the close agreement between the fluorescence signals at time 0 and that extrapolated from the equilibrium unfolded hFGF-1 base line. However, we cannot completely rule out the possibility of occurrence of the burst phase intermediate(s) because the burst phase changes cannot be traced under conditions where the intermediate(s) has similar fluorescence as the unfolded state.

hFGF-1 like many other β -barrel proteins shows a far-UV

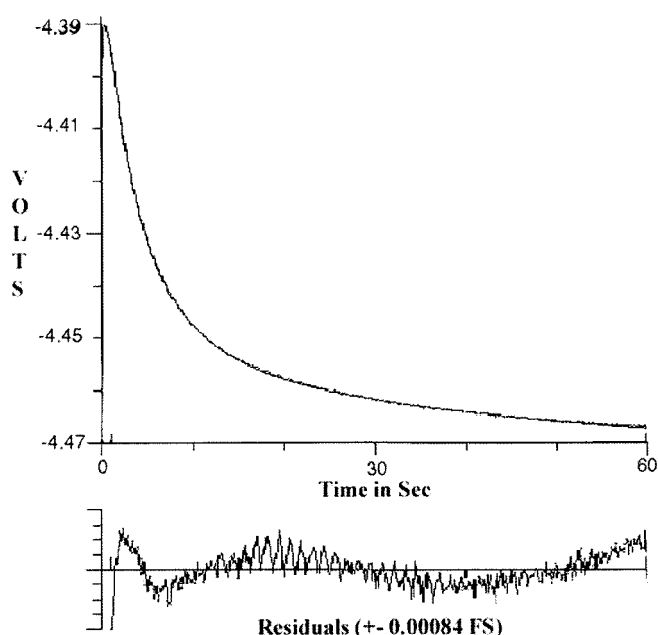


FIG. 3. Refolding kinetics of hFGF-1 monitored by stopped-flow fluorescence (at 350 nm). The refolding trace is found to best fit with a sum of two exponentials. The refolding rates (k) and the amplitude (A) of the two phases are as follows: $k_1 = 0.131 \pm 0.02 \text{ s}^{-1}$; $A_1 = 81.73 \pm 0.34\%$; $k_2 = 0.023 \pm 0.07 \text{ s}^{-1}$; $A_2 = 15.2 \pm 0.3\%$.

positive ellipticity band at 228 nm, signifying the secondary and, to a lesser extent, the tertiary structural interactions in the protein. The far-UV ellipticity changes (at 228 nm) are complete within a time span of 3 s (Fig. 4). The ellipticity changes could be best fit to a single exponential equation yielding a refolding time constant of $1.24 \pm 0.65 \text{ s}^{-1}$. The slow minor phase (observed in the stopped-flow fluorescence experiment) possibly representing the *cis-trans*-proline isomerization is not observed in the stopped-flow CD experiments. In addition, no burst phase ellipticity change(s) could be detected during the refolding of hFGF-1. Within experimental error, the ellipticity (at 228 nm) at time 0 is similar to that obtained from the equilibrium unfolded state base line. It is interesting to note that the refolding rate constant value of the major phase estimated using the stopped-flow fluorescence ($0.131 \pm 0.02 \text{ s}^{-1}$) and stopped-flow CD do not match (Table I). The fast recovery of the 228 nm ellipticity signal presumably reflects the rapid acquisition of some portion(s) of the native structure prior to the burial of the tryptophan in the interior of the protein. Thus, the results of the stopped-flow fluorescence and stopped-flow CD experiments strongly suggest the accumulation of transient intermediate(s) in the kinetic refolding pathway of hFGF-1.

Occurrence of Kinetic Intermediate(s)—It has been observed that in many proteins where an intermediate in equilibrium conditions is not found, the analysis of the refolding kinetics shows the unambiguous presence of an intermediate. The kinetics of refolding were examined at 20 °C over a range of urea concentrations (up to 6 M) by monitoring the changes in the tryptophan fluorescence using the stopped-flow fluorescence technique. Fig. 5 shows the urea dependence of the natural logarithm of the observed rate constant of folding/unfolding. Interestingly, the slow phase (minor phase) shows no denaturant dependence and is consistent with it being a rate-limiting proline isomerization event. The unfolding rate decreases linearly with the increase in urea concentration (Fig. 5). However, a prominent curvature could be observed at low concentrations of the denaturant (Fig. 5). In general, a deviation from linearity ("roll over") is believed to reflect the accumulation of intermediate species at low denaturant concentrations (36). For a pro-

TABLE I
Rate and time constant of the various phases of refolding of hFGF-1

	First phase		Second phase	
	Rate constant s^{-1}	Refolding time constant s	Rate constant	Refolding time constant
Fluorescence	0.131 ± 0.02	7.63 ± 1.0	0.023 ± 0.07	43.47 ± 14.32
CD	1.24 ± 0.65	0.806 ± 0.27	NO ^a	NO
ANS	0.21 ± 0.012	2.76 ± 0.21	0.018 ± 0.008	53.3 ± 17

^a NO, not observed.

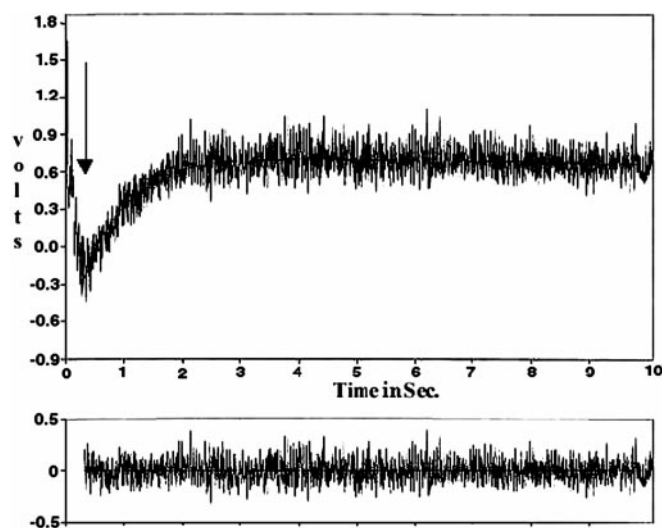


FIG. 4. Stopped-flow kinetics of the refolding of hFGF-1 monitored by far-UV circular dichroism (at 228 nm). The refolding trace fits well to a single exponential equation yielding a time constant of 0.8075 s. The arrow in the figure indicates the start of the refolding trace.

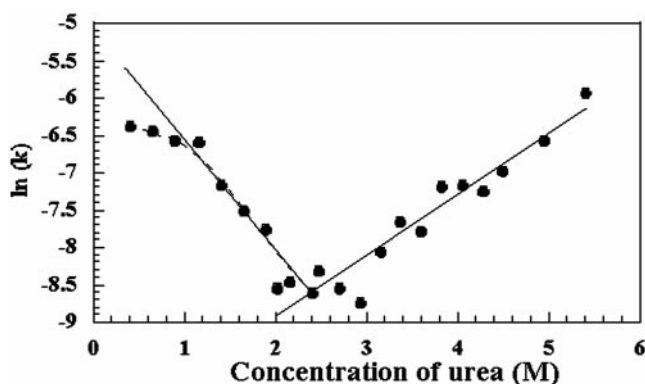


FIG. 5. Urea-dependent folding kinetics of hFGF-1. The solid circles represent the observed rate constants for refolding and unfolding obtained from the stopped-flow fluorescence experiments. The solid line shows the predicted fit of the kinetic data for a two-state kinetic model. The small curvature observed at low urea concentrations is ascribed to the formation of transient kinetic intermediate(s) during the refolding of hFGF-1.

tein folding typically by a two-state mechanism, the Chevron plot ($\ln k$ versus concentration of the denaturant) is expected to be V-shaped, with no roll over. In summary, these results clearly suggest that the refolding of hFGF-1 occurs via the formation of transient intermediate(s).

There is a lot of controversy on the interpretation of the roll over in the Chevron plot (35–38). Curvatures in Chevron plots, in addition to accumulation of kinetic intermediates, could also be ascribed to the movements of the transition state ensemble (36, 39–41). According to the broad barrier model of protein folding, the folding process is proposed to occur isoenergetically

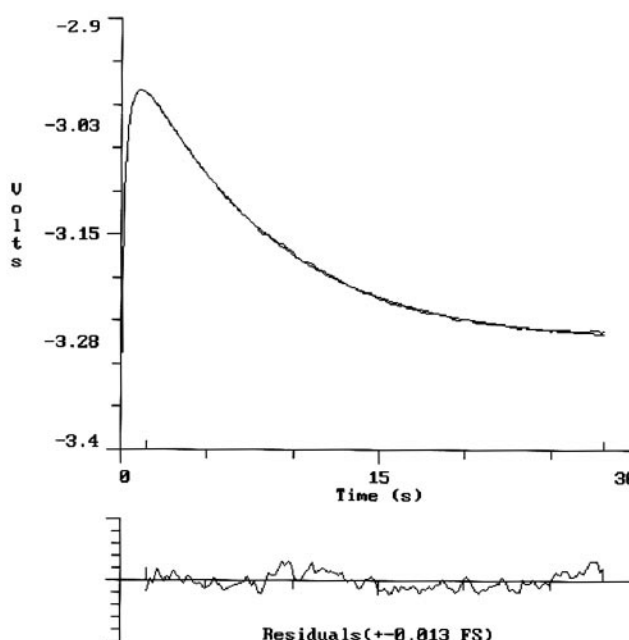


FIG. 6. ANS binding affinity during the refolding of hFGF-1 monitored by stopped-flow fluorescence at 520 nm. The steep rise in the ANS fluorescence intensity is attributed to the binding of the fluorescent dye to the heparin-binding domain formed in the time scale of about 2.2 s.

at the transition state level, and hence the position of the transition state ensemble is expected to be sensitive to denaturant conditions (36). Recent studies on the human spliceosomal protein have demonstrated that the roll over in the Chevron plot could also be due to the movements of the transition state along the top of a very broad and flat activation barrier (Hammond effect, see Refs. 41–44). In general, Hammond effects (during protein folding/unfolding) are manifested as curvatures in both the limbs of the Chevron plot (35). In this context, as the roll over is observed in only one of the limbs of the Chevron plot of hFGF-1, it could be possibly attributed to the accumulation transient intermediate(s) rather than to the movement of the transition state ensemble.

Absence of Hydrophobic Collapse—ANS is a popular hydrophobic dye used to probe early events in protein folding (45–47). ANS is known to bind to transiently exposed hydrophobic patches during folding of many proteins, leading to a prominent increase in the fluorescence intensity of the dye. Refolding of hFGF-1 was monitored in the presence of ANS at 520 nm. No prominent intensity changes were observed within the dead time of the instrument. Thus, it appears that the refolding hFGF-1 does not involve early burst phase events such as the hydrophobic collapse. Interestingly, after 1 s of initiation of refolding, the ANS fluorescence intensity shows a dramatic increase and reaches a maximum value at around 2.2 s (Fig. 6). Beyond this refolding time (>2.2 s), the emission intensity exponentially decreases to reach a steady state value that is much higher than that of the free ANS (Fig. 6). The changes in

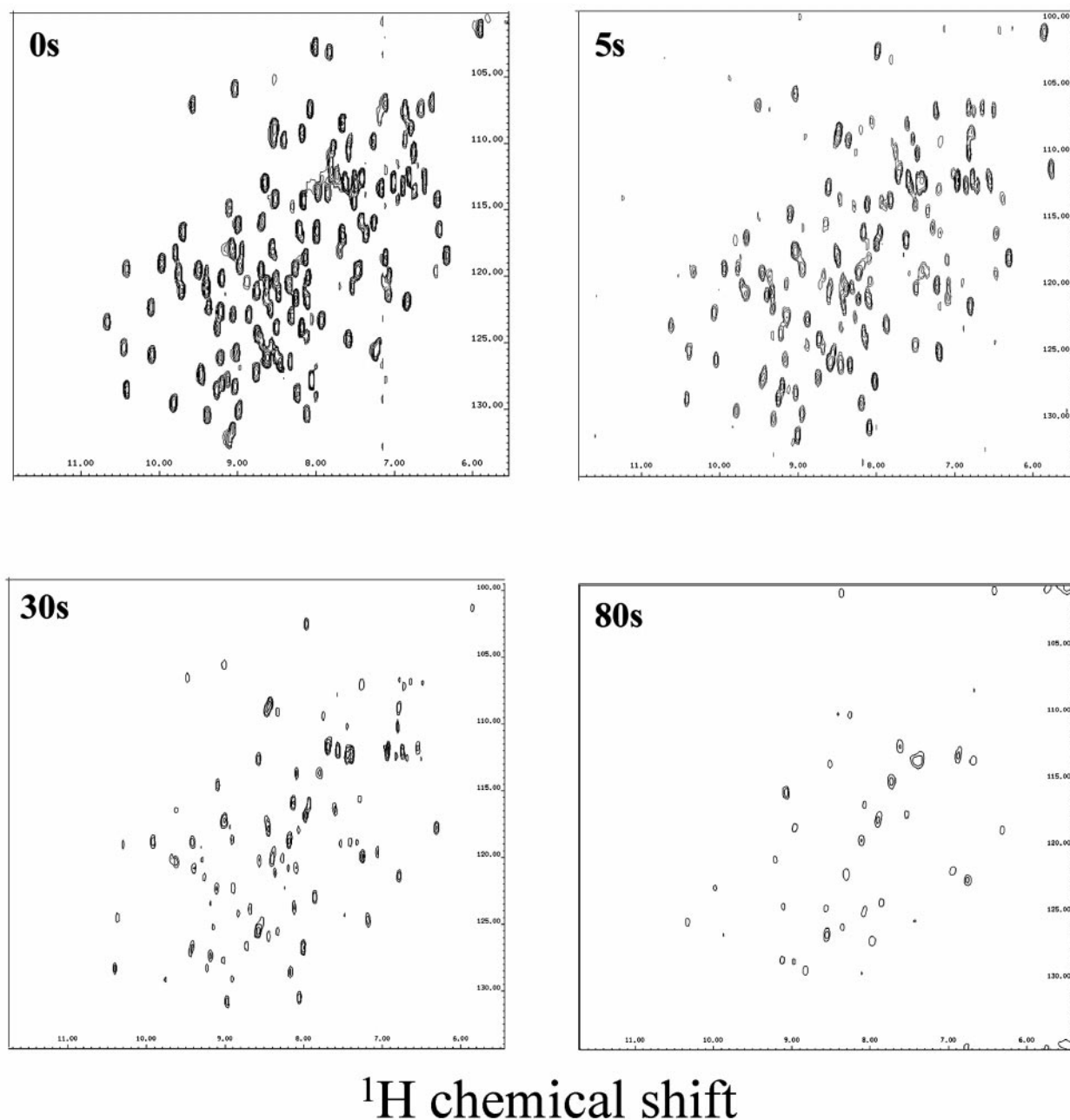
 ^1H chemical shift

FIG. 7. ^1H - ^{15}N HSQC spectra of hFGF-1 samples prepared by quenched-flow hydrogen exchange experiments at various refolding time periods.

the ANS fluorescence intensity (beyond 2.2 s) could be best fit to a two exponential term yielding folding rate constant values of 0.21 ± 0.012 and $0.018 \pm 0.008 \text{ s}^{-1}$, respectively (Table I). The amplitudes of the fast and slow phase have been estimated to be 90 ± 0.1 and $10 \pm 0.2\%$, respectively.

The use of an extrinsic probe such as ANS to monitor the folding reactions raises the possibility that the binding of the probe to the protein may perturb the folding reaction (46, 47). We examined this possibility by studying the refolding of hFGF-1 at varying concentrations (50–250 μM) of ANS in the refolding buffer. It was found that the hydrophobic dye did not have significant effect(s) on the refolding time constant of the two phases of folding (data not shown). In addition, the kinetics of refolding of hFGF-1 monitored by changes in the intrinsic tryptophan fluorescence (at 350 nm) was also unaffected by the presence of ANS (up to a concentration of 250 μM) in the refolding buffer (data not shown). Thus, these results suggest

that ANS does not significantly perturb the refolding kinetics of hFGF-1.

It is important to understand the dramatic increase in the ANS fluorescence (at ~ 2.2 s) observed during the refolding of the protein. hFGF-1 is a heparin-binding protein, and its cell regulatory properties are known to be strongly dependent on binding to the proteoglycan. Available crystal and solution structures of hFGF-1 reveal that the heparin-binding site is in the segment (consisting of residues, 110–130) of the protein spanning β -strands IX and X (31–33). Recent studies indicate that several polysulfonated compounds could nonspecifically bind to the positively charged heparin-binding site (33). In this background, the drastic increase in the ANS fluorescence intensity observed at around 2 s could be attributed to the transient binding of ANS to the solvent-exposed, heparin binding domain (residues spanning β -strands IX and X) formed in this time scale (~ 2.2 s). This explanation is consistent with the

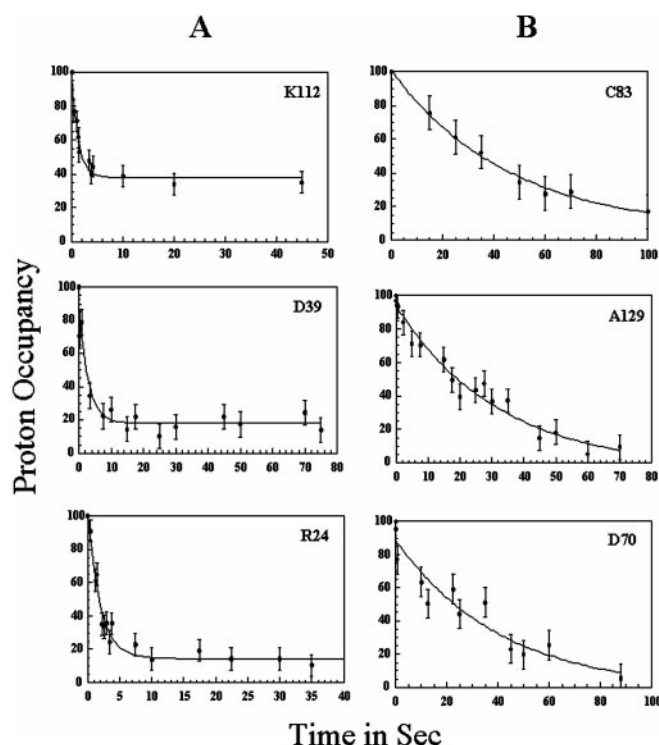


FIG. 8. Comparison of the observed time courses of selected residues as a function of length of the refolding time. All residues were fit to a first order exponential decay. A depicts the percentage of proton occupancy of residues located in the various β -strands. The percentage of proton occupancy of residues located in the unstructured region is represented in B.

average time constant of refolding of the β -strands IX and X (constituting the polyanion binding site), estimated from the quenched-flow H/D exchange experiments (see below). The heparin-binding site is densely populated by cationic residues and is interspersed with many nonpolar residues such as Asn-109, Tyr-111, Ala-111, Trp-121, and Val-123. It appears that the strong ANS binding affinity of the folding species realized at ~ 2.2 s is mediated by both the charge and hydrophobic interactions with the fluorescent dye. However, in the subsequent stages of folding of the protein (>2.2 s), the nonpolar groups in the heparin-binding site appear to be sequestered into the interior of the protein, leading to a decrease in the binding affinity of ANS to the protein. This aspect is exemplified by the exponential decrease in the ANS emission intensity beyond 2.2 s (Fig. 6).

Recently, transient aggregates have been shown to form during the *in vitro* refolding of proteins from their random coil states (39–41). In this context, it could be argued that the increase in ANS intensity observed during the refolding (at ~ 2.2 s) of hFGF-1 could be due to the higher binding affinity of the hydrophobic dye to the transiently formed aggregates (formed in the time scale of ~ 2.2 s). In the event of formation of transient aggregates (during refolding), the refolding rate constants of the various phases of folding are expected to show a linear decrease with the increase in the concentration of the refolding protein. However, it is found that the rate constants of both the phases (major and minor) of refolding (observed by monitoring the changes in the tryptophan fluorescence) do not significantly change as a function of the protein concentration (data not shown). These results clearly suggest that refolding of hFGF-1 does not proceed via the formation of transient protein aggregates. The observed increase in the ANS emission signal (at ~ 2.2 s) is probably due to binding of the dye to the heparin binding domain formed in this time scale (~ 2.2 s).

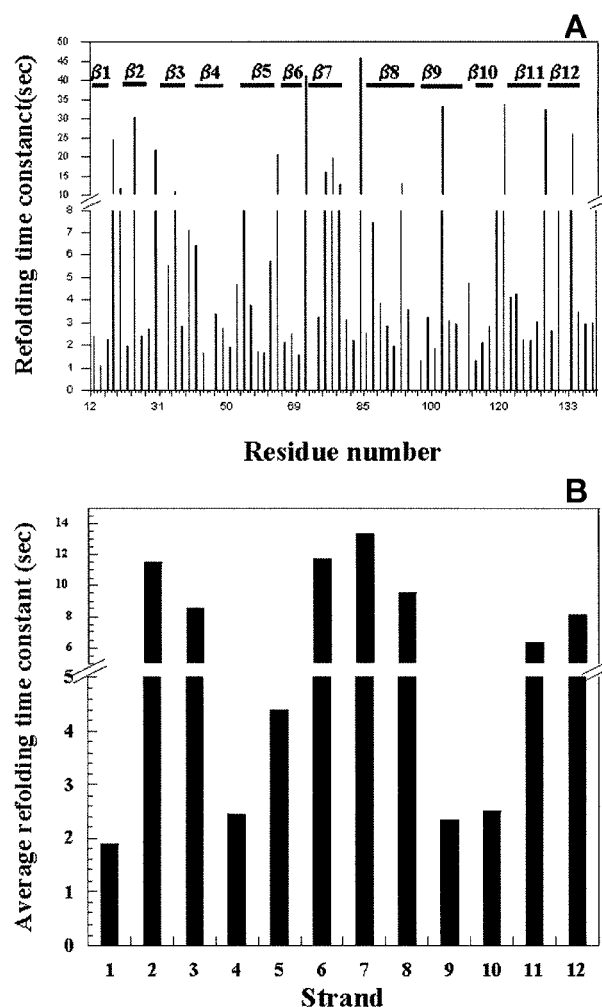


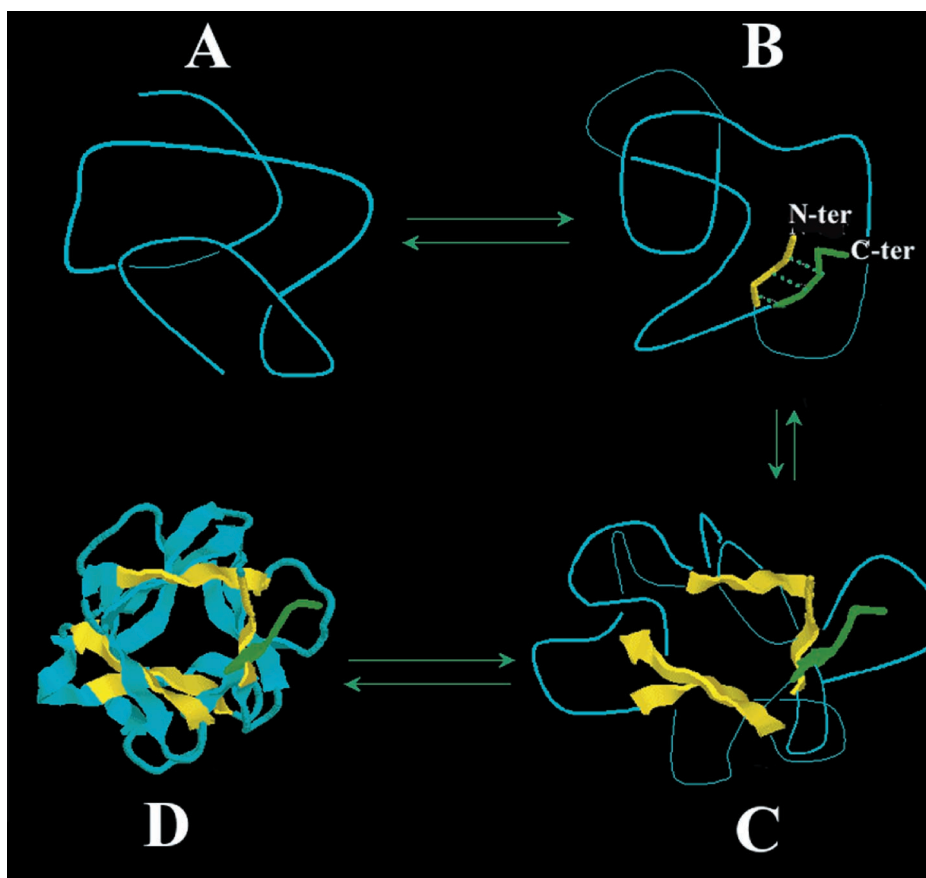
FIG. 9. A plot of the time constant of refolding versus the amino acid residue number (A). It could be observed that most of the residues located in β -strands I, IV, IX, and X have low refolding time constant values. The y axis scale is broken to depict clearly the differences in the refolding time constant values of the fast folding residues. B, average time constant of various β -strands in hFGF-1. The y axis scale is broken to indicate clearly the trends in the average time constant values of various β -strands in the protein.

Refolding Events Detected by Quenched-flow Hydrogen Exchange—Stopped-flow optical techniques only report gross conformational changes that occur during the refolding of proteins. They do not provide any information on the structural changes that occur at a residue level. However, the quenched-flow H/D exchange measurements allow the determination of the time scales of formation of various hydrogen bonds involved in secondary structure during refolding of the protein (23). It is in this context that we used the quenched-flow H/D exchange to monitor the chronology of events in the refolding pathway of hFGF-1.

The ^1H - ^{15}N HSQC spectra of hFGF-1 has been completely assigned (32). Hence, it is possible to follow unambiguously the folding kinetics of 75 well separated, slowly exchanging amide residues involved in secondary and tertiary structural interactions in the protein (Fig. 7). These amide protons are distributed throughout the protein molecule. The exchange kinetics of all residues are adequately described by a single exponential (Fig. 8).

Representative ^1H - ^{15}N -HSQC spectra of hFGF-1 samples collected after various refolding times are depicted in Fig. 7. Only a few stable hydrogen bonds were observed to form during the dead time (~ 8 ms) of the quenched-flow apparatus. This

FIG. 10. Schematic representation of the initial events in the refolding of hFGF-1. A, hFGF-1 in the 4 M urea-unfolded state. B, the bridging of the N- (indicated in yellow) and C-terminal ends (indicated in green) of the hFGF-1 molecule through backbone hydrogen bonds (indicated by green dotted lines) between Leu-13-NH and Leu-135-CO, Tyr-15-NH and Leu-133-CO, and Val-137-NH and Pro-11-CO appears to be the first detectable event in the refolding of hFGF-1. The N- and C-terminally linked topology probably provides the structural mold for subsequent events. C among the various β -strands in hFGF-1, β -strands, I, IV, IX, and X form (indicated in yellow) in similar time scales (~ 2 – 2.5 s) and appear to provide the basic β -trefoil framework. Subsequently, D, the trigonal pyramid structure (as seen in the native state) is built by progressive layering of the various other β -strands (indicated in blue) (in the order of strand V, strand XI ($\tau = 5$ – 7 s) > strand XII, strand VIII, strand III, strand II, strand VI, and strand VII ($\tau = 10$ – 13 s)).



aspect is consistent with the conclusions drawn from the stopped-flow fluorescence experiments. After 3 s, there were only about 20 amide protons whose intensities decrease by 50–60%. This observation is in marked contrast to the stopped-flow CD data (at 228 nm) where the ellipticity changes (primarily representing the secondary structural interactions) are found to be almost complete ($\sim 95\%$) within 3 s of initiation of folding (Fig. 4). The backbone hydrogen bonds that are formed and detected by the far-UV CD signal could be unstable due to the local breathing or sliding of one strand relative to the other. The net consequence of this structural motion(s) is the rapid formation and disruption of hydrogen bonds. The short lifetime of the hydrogen bonds probably accounts for their weak protection (against exchange) observed in the ^1H - ^{15}N HSQC spectra collected at refolding times lesser than 5 s (Fig. 7). Interestingly, 50–60% resonance intensity decrease for half of the amide protons involved in the secondary structure formation occurred only after 5 s of refolding. The secondary structure formation in another β -barrel protein interleukin-1 β , monitored by quenched-flow H/D exchange, also revealed a similar trend (26).

If we arbitrarily subdivide the time constants of refolding (τ) of various residues estimated by the quenched-flow H/D exchange technique into three classes, as fast ($1 < \tau < 3$ s), medium ($5 < \tau < 10$ s), and slow ($\tau > 20$ s) folding, most of the residues in β -strands I, IV, IX, and X fit into the fast folding category (Fig. 9). The slow folding residues are mostly confined to the unstructured loop regions inter-spread between the various β -strands. There are 11 residues whose refolding time constant fits into the medium folding set. These residues are mostly concentrated in β -strands III, V, and VIII (Fig. 9).

It is interesting to note that the amide protons of some residues, which are highly protected from deuterium exchange in the native state, could not be observed even in the ^1H - ^{15}N

HSQC spectra collected after 8 ms of folding. These include the amide protons of Leu-13, Tyr-15, and Val-137. This feature could be rationalized on the basis of the three-dimensional structure of hFGF-1. hFGF-1 is a β -barrel protein and the N and C termini of the molecule are strongly bridged by three hydrogen bonds between, Leu-13-NH and Leu-135-CO, Tyr-15-NH and Leu-133-CO, Val-137-NH and Pro-11-CO (31, 32). The absence of the ^1H - ^{15}N cross-peaks representing the amide protons of Leu-13, Tyr-15, and Val-137 in the first ^1H - ^{15}N HSQC spectrum obtained following the dead-time of the quenched-flow apparatus (< 8 ms) strongly suggests that the formation of hydrogen bonds linking the extreme ends (N and C termini) of the β -barrel is probably the first event in the refolding pathway of hFGF-1 (Fig. 10).

The three-dimensional structure of hFGF-1 could be best described as a trigonal pyramid with three topological units consisting each of four antiparallel β -sheets (31, 32). The β -strands comprising the structural units A are β -strands I–III and XII. The structural units B and C consist of β -strands, IV–VII and β -strands VIII–XI, respectively.

Comparison of the time constant of folding of various β -strands comprising the β -barrel structure of hFGF-1 shows that there is a clear pattern in the rates of the formation of the 12 β -strands. The average time constants of the β -strands in the protein are as follows (Fig. 9): β -strand I, $\tau = 1.89 \pm 0.24$ s; β -strand II, $\tau = 11.52 \pm 0.25$ s; β -strand III, $\tau = 8.50 \pm 0.31$ s; β -strand IV, $\tau = 2.44 \pm 0.1$ s; β -strand V, $\tau = 4.39 \pm 0.25$ s; β -strand VI, $\tau = 11.78 \pm 0.41$ s; β -strand VII, $\tau = 13.33 \pm 0.15$ s; β -strand VIII, $\tau = 9.50 \pm 0.30$ s; β -strand IX, $\tau = 2.33 \pm 0.34$ s; β -strand X, $\tau = 2.5 \pm 0.21$ s; β -strand XI, $\tau = 6.32 \pm 0.34$ s; and β -strand XII, $\tau = 11.69 \pm 0.41$ s. The hFGF-1 molecule with the N- and C-terminal ends bridged together appears to constitute the basic mold for subsequent folding events (Fig. 10). The β -strands in the three topological (A, B,

and C) units comprising the trigonal pyramid structure (of hFGF-1) appear to form in a phased manner. β -Strands I, IV, IX, and X appear to form simultaneously to establish the first layer of the trigonal pyramid structure (Figs. 9 and 10). The drastic increase in the ANS fluorescence intensity observed after 2.2 s could be attributed to the transient binding of the polysulfonated dye to the heparin binding domain (comprising of the residues in β -strands IX and X) formed in this time scale ($\tau \sim 2.2$ ms). The simultaneous formation of β -strands I, IV, IX, and X could be ascribed to the hydrogen bonds linking these β -strands (strands I, IV, IX, and X) with each other in the native state of the protein (Figs. 9 and 10).

In the second phase of folding (average $\tau = 4$ to 6.5 s), β -strands V and XI are formed (Fig. 9B). The formation of β -strands, II, III, VI–VIII, and XII occurs simultaneously in the final phase (average $\tau = 9$ –12 s) of refolding of hFGF-1. Thus, it appears that the protein regains most of its native interactions within a time span of about 20 s. The slow phase of folding ($\tau > 20$ s) probably pertains to conformational reorganization events arising out of *cis-trans*-proline isomerization.

Comparison with the Refolding Kinetics of Interleukin-1 β —The three-dimensional structures of hFGF-1 and interleukin-1 β exhibit strong structural homology. Although both proteins share only 10–15% sequence similarity, their backbone in 9 of the 12 β -sheet strands could be superimposed with a root mean square deviation of 0.5 Å (48). The folding pathway of interleukin-1 β has been investigated using a variety of spectroscopic techniques including quenched-flow hydrogen exchange experiments (26, 49). It is informative to compare the events in the refolding kinetics of hFGF-1 and interleukin-1 β . Striking similarities could be observed in the refolding kinetics of hFGF-1 and interleukin-1 β . Both the proteins fold very slowly, and complete folding of these proteins occurs on a time scale greater than 50 s (26). As observed in hFGF-1, the formation of unstable β -sheets in interleukin-1 β is shown to occur rapidly, but stabilization of the three-dimensional structure with the establishment of native hydrogen bonds begins only after 1 s of initiation of refolding. Additionally, the refolding of both hFGF-1 and interleukin-1 β involves the formation of kinetic intermediate(s).

The refolding kinetics of hFGF-1 and interleukin-1 β also show significant difference(s). Unlike hFGF-1, where no burst phase hydrophobic collapse has been observed, the refolding pathway of interleukin-1 β involves clustering of the hydrophobic residues as evidenced by the strong ANS binding in the burst phase of refolding. In hFGF-1 the formation of hydrogen bonds between the N- and C-terminal ends is the first detectable event. Thus, comparison of the folding kinetics of these two structurally homologous proteins reveals that although the time scale for refolding of the proteins is dependent on their overall structural architecture, the individual events in the refolding pathway appear to be governed by the unique local and long range interactions present in the protein in context.

The results of the present study support the general notion that folding/unfolding of large proteins (>100 amino acids) is complex and essentially involves the formation of intermediate(s). Detailed folding studies using appropriate mutants of hFGF-1 are in progress to understand the relationship between the conformational stability and rate of refolding of the protein.

Acknowledgments—We thank Prof. Ing-Ming Chiu, Ohio State University, for providing us the hFGF-1 clone. We also thank the anonymous referee for valuable comments on an earlier version of the manuscript.

REFERENCES

- Grantchacova, V. P., and Baker, D. C. (1997) *Biochemistry* **36**, 15685–15692
- Socci, N. D., Onuchic, J. N., and Wolynes, P. G. (1998) *Proteins Struct. Funct. Genet.* **32**, 136–158
- Mirny, L. A., Abkevich, V. I., and Shakhnovich, E. I. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 4976–4974
- Ropson, I. J., and Dalessio, P. M. (1997) *Biochemistry* **36**, 8594–8601
- Matagne, A., Jamin, M., Chung, E. W., Robinson, C. V., Radford, S. E., and Dobson, C. M. (2000) *J. Mol. Biol.* **297**, 193–210
- Plaxco, K. W., Guijaro, J. J., Morton, C. J., Pitkeathly, M., Campbell, I., and Dobson, C. M. (1998) *Biochemistry* **37**, 2529–2537
- Jonsson, T., Waldburger, C. P., and Saur, R. T. (1996) *Biochemistry* **35**, 4795–4802
- Staniforth, A. R., Dean, J. L. E., Zhong, Q., Zerovnik, E., Clarke, A. R., and Waltho, J. P. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5790–5795
- Jackson, S. E., and Fersht, A. R. (1994) *Biochemistry* **30**, 10436–10443
- Kuszewski, J., Clore, G. M., and Gronenborn, A. M. (1994) *Protein Sci.* **3**, 1945–1952
- Schonbrunner, N., Koller, K. P., and Kiefhaber, T. (1997) *J. Mol. Biol.* **268**, 526–538
- Viguera, A. R., Martinez, J. C., Filimonov, V. V., Matco, P. L., and Serrano, L. (1994) *Biochemistry* **33**, 2142–2150
- Chiti, F., Van Nuland, N. A. J., Taddei, N., Magherini, F., Stefani, M., Ramponi, G., and Dobson, C. M. (1998) *Biochemistry* **37**, 1447–1455
- Ionescu, R. M., Smith, V. J., O'Neil, J. C., and Matthews, C. R. (2000) *Biochemistry* **39**, 9540–9550
- Matthews, C. R. (1993) *Annu. Rev. Biochem.* **62**, 652–683
- Pittsytyn, O. B. (1992) in *Protein Folding* (Creighton, T. E., eds) pp. 243–300, W. H. Freeman & Co., New York
- Mann, C. J., and Matthews, C. R. (1993) *Biochemistry* **32**, 5282–5290
- Briggs, M. S., and Roder, H. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 2017–2024
- Lu, J., and Dahlquist, F. W. (1992) *Biochemistry* **31**, 4749–4756
- Hogan, S. J., and Eaton, W. A. (2000) *J. Mol. Biol.* **297**, 781–789
- Heidary, D. K., O'Neill, J. C., Roy, M., and Jennings, P. A. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5866–5870
- Clementi, C., Jennings, P. A., and Onuchic, J. N. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5871–5876
- Van Nuland, N. A. J., Chiti, F., Taddei, N., Raugei, G., Ramponi, G., and Dobson, C. M. (1998) *J. Mol. Biol.* **283**, 883–891
- Roder, H., Elove, S. W., and Englander, S. W. (1988) *Nature* **335**, 700–704
- Shakhnovich, E., Abkevich, V., and Pittsytyn, O. B. (1996) *Nature* **379**, 96–98
- Varley, P., Gronenborn, A. M., Christensen, H., Wingfield, P. T., Pain, R. H., and Clore, G. M. (1993) *Science* **260**, 1110–1113
- Radford, S. E., Dobson, C. M., and Evans, P. A. (1992) *Nature* **358**, 302–306
- Heidary, D. K., Gross, L. A., Roy, M., and Jennings, P. A. (1997) *Nat. Struct. Biol.* **4**, 725–731
- Sivaraman, T., Kumar, T. K. S., Lin, W. Y., Chang, D. K., and Yu, C. (1998) *J. Biol. Chem.* **273**, 10181–10189
- Burgess, W. H., and Maciag, T. (1989) *Annu. Rev. Biochem.* **58**, 575–608
- Zhu, X., Komiya, H., Chirino, A., Faham, S., Fox, G. M., Arakawa, T., Hsu, B. T., and Rees, D. C. (1991) *Science* **251**, 90–93
- Ogura, K., Nagata, K., Hatanaka, H., Habuchi, H., Kimata, K., Tati, S. I., Ravera, M. W., Jaye, M., Schlessinger, J., and Inagaki, F. (1999) *J. Biomol. NMR* **13**, 11–24
- Lozano, R. M., Pineda-Lucena, A., Gonzalez, C., Jimenez, M. A., Cuevas, P., Redono-Horcajo, M., Sanz, J. M., Rico, M., and Gimenez, G. (2000) *Biochemistry* **39**, 4982–4993
- Burke, C. J., Volk, D. B., Mach, H., and Middaugh, C. R. (1993) *Biochemistry* **32**, 6419–6428
- Matouschek, A., Matthews, J. M., Johnson, C. M., and Fersht, A. R. (1994) *Protein Eng.* **7**, 1089–1095
- Jackson, S. E. (1998) *Folding Design* **3**, 81–91
- Silow, M., and Oliveberg, M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 6084–6086
- Parker, M. J., and Marqusee, S. (1999) *J. Mol. Biol.* **293**, 1195–1210
- Otzen, D. E., Kritensen, O., Proctor, M., and Oliveberg, M. (1999) *Biochemistry* **38**, 6499–6511
- Silow, M., Tan, Y. I., Fersht, A. R., and Oliveberg, M. (1999) *Biochemistry* **38**, 13006–13012
- Oliveberg, M., Tan, Y. J., Silow, M., and Fersht, A. R. (1998) *J. Mol. Biol.* **277**, 933–943
- Cota, E., and Clarke, J. (2000) *Protein Sci.* **9**, 112–120
- Hammond, G. S. (1955) *J. Am. Chem. Soc.* **77**, 334–338
- Main, E. R. G., Fulton, K. F., and Jackson, S. E. (1999) *J. Mol. Biol.* **291**, 429–444
- Jones, B. E., Jennings, P. A., Pierre, R. A., and Matthews, C. R. (1994) *Biochemistry* **33**, 15250–15258
- Shasthy, M. C. R., and Udgaonkar, J. B. (1995) *J. Mol. Biol.* **247**, 1013–1027
- Engelhard, M., and Evans, P. A. (1995) *Protein Sci.* **4**, 1553–1562
- Eriksson, A. E., Cousens, L. S., Weaver, L. H., and Mathews, B. W. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 3441–3445
- Finke, J. M., Roy, M., Zimm, B. H., and Jennings, P. A. (2000) *Biochemistry* **39**, 575–583
- Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950