Evidence of Discrete Substates and Unfolding Pathways in Green Fluorescent Protein

Giancarlo Baldini,* Fabio Cannone,* Giuseppe Chirico,* Maddalena Collini,* Barbara Campanini,† Stefano Bettati,† and Andrea Mozzarelli†

*Department of Physics, University of Milano-Bicocca, Milan, Italy; and †Department of Biochemistry and Molecular Biology, University of Parma, Parma, Italy

ABSTRACT We present evidence of conformational substates of a green fluorescent protein mutant, GFPmut2, and of their relationship with the protein behavior during chemical unfolding. The fluorescence of single molecules, excited by two infrared photons from a pulsed laser, was detected in two separate channels that simultaneously collected the blue or the green emission from the protein chromophore chemical states (anionic or neutral, respectively). Time recording of the fluorescence signals from molecules in the native state shows that the chromophore, an intrinsic probe sensitive to conformational changes, switches between the two states with average rates that are found to assume distinct values, thereby suggesting a multiplicity of protein substates. Furthermore, under denaturing conditions, the chromophore switching rate displays different and reproducible time evolutions that are characterized by discrete unfolding times. The correlation that is found between native molecules' switching rate values and unfolding times appears as direct evidence that GFPmut2 can unfold only along distinct paths that are determined by the initial folded substate of the protein.

INTRODUCTION

In the last decades, the existence of a multiplicity of conformational substates in folded proteins has been foreseen (1,2), and the experimental evidence of their presence and functional properties has been inferred (3–7) from the nonexponential behavior of reaction kinetics (2,8,9) in ensemble measurements and from the analysis of the heterogeneous response in single protein approaches (10–12). Indeed, as suggested in the literature (8,12,13), a suitable single molecule approach should be pursued to ascertain the existence of different substates (14) and to detect the expected multiplicity of unfolding routes that occur when perturbing native proteins (15–18). Here we show clear evidence of the composite nature of the native state of a green fluorescent protein (GFP) (19,20) mutant, GFPmut2 (21), and of its unfolding pathways by exploiting single molecule fluorescence spectroscopy.

GFPmut2, a mutant engineered to enhance emission and stability (22–24), appears as an ideal candidate for investigating substates and folding-unfolding (25–27) pathways since the anionic (A) and neutral (N) chemical states of the chromophore, associated to different conformations of the protein matrix (28–31), can be monitored by the fluorescence occurring at two separate wavelengths, green (508 nm) and blue (460 nm), respectively (24,26). Moreover, it has been recently found (27) that the switching rates between the A and N states, although usually random, become, unexpectedly, very regular when in proximity to the unfolding event and display a discrete number of allowed frequencies (27). These findings suggest that the A-N fluorescence switching rate, detected in single molecules, could be employed to explore the possible multiplicity of native conformations, which are spectrally undistinguishable in bulk experiments. Therefore, a search aimed at uncovering both the presence and role of substates in protein unfolding was undertaken.

Despite little experimental evidence, expectations in line with these assumptions have been put forward by recent theoretical studies on the existence of a basin of substates in folded proteins and on the feasibility of multiple folding pathways (17,32–36). A few unfolding studies have been performed on single protein molecules in solution, where the observation time can last only a few milliseconds (12,37–40) due to the fast molecular diffusion across the focused laser beam and, consequently, data collection along the entire unfolding pathway cannot be achieved (12,38,40). On the contrary, in this work the unfolding dynamics of the same single protein molecules could be followed for long times (hours to days) (26) since GFPmut2 was entrapped in wet nanoporous silica gels (24,26). The adopted experimental procedure (proper excitation intensity, very sensitive detection, nitrogen flux, etc.) allowed good noise rejection and negligible bleaching effects (24,26). By a detailed analysis of the A-N switching rate, we find here that native GFPmut2 proteins exist in a few conformational substates which lead to a set of unfolding pathways under the action of denaturants.

MATERIALS AND METHODS

GFPmut2 protein

GFPmut2 is a GFP mutant containing a triple substitution (S65A, V68L, S72A) responsible for an enhanced fluorescence emission (21–24). The GFPmut2 gene, cloned in a pKEN1 vector, was kindly provided by Dr. Brendan P. Cormack (21). Protein expression and purification was carried out as previously described (22). GFPmut2 stock solutions were dialyzed against 50 mM Tris buffer, pH 8.0, and kept at −80°C.

Submitted July 17, 2006, and accepted for publication November 10, 2006.

Address reprint requests to Giancarlo Baldini, E-mail: baldini@mib.infn.it.

0006-3495/07/03/1724/08 $2.00

doi: 10.1529/biophysj.106.093567

The two-photon excitation (TPE) spectrum of single GFPmut2 molecules encapsulated in silica gels exhibits two main components at 820 ± 2 and 885 ± 3 nm (24). The excitation at 885 nm is largely favored at high pH with respect to the band at 820 nm, suggesting that the 885-nm excitation band is due to the two-photon absorption of the anionic state of the GFPmut2 chromophore, whereas the excitation at 820 nm corresponds to the absorption of the neutral chromophore. The single molecule TPE emission spectrum, upon excitation at 820 nm, shows two bands at 450 and 510 nm (24), in good agreement with the one-photon emission spectra (23). At alkaline pH, the component at 510 nm is largely favored and it is ascribed to the anionic state of the GFPmut2 chromophore. The emission at ≅460 nm is due to the neutral state of the chromophore (24).

The optical absorption and fluorescence properties of GFP are commonly described in terms of three chemical states of the chromophore (20,24): a neutral N (protonated), an anionic A (deprotonated), and a zwitterionic dark state Z (41). It is likely that A, N, and Z are taxonomic states (1,2). Excitation at 860 nm gives rise to fluorescence emission switching between the anionic and neutral form of the chromophore (27) that is due to a proton transfer on the GFP chromophore (23). The hydroxyl group of the GFP chromophore, which is part of an intricate network of hydrogen bonds that favors the anionic form (23), exchanges a proton in a time shorter than a microsecond. The A-N switching is assumed to be a sequential process involving two steps: a), proton transfer from solvent to the chromophore, and b), internal structural rearrangements to stabilize a protonated chromophore (29–31,42). In particular, when the chromophore is in the anionic state, a negative charge is placed on Tyr-66 due to H-bond formation with Thr-203 and His-148 (43). When the chromophore switches to the neutral form, the Thr-203 rotates by ~100° and His-148 slightly moves away from the chromophore and breaks the connecting H-bonds (42,44). Although the proton transfer involved in the anionic-neutral switching occurs in a short time, the slower rearrangement of the side chains near the chromophore and the cooperative movements of the β-sheet backbone (29,30) slow down the process to hundreds of microseconds (27,29–31).

Protein encapsulation in wet silica gels

Observation of single proteins for an extended period of time has been previously achieved by their immobilization either on a solid surface or in gels (40). However, owing to the usually very low conformational stability of proteins, their interaction with a surface is known to disturb the folding reaction, whereas for trapped proteins in sugar matrices (45) one cannot modulate their environment. A good strategy to circumvent these difficulties was the encapsulation of proteins in wet silica gels (46). Encapsulation of GFPmut2 in silica gels was carried out according to the procedure of Bettati and Mozzarelli (47,48), as previously described (24,26,27). The protein's functional properties were not significantly perturbed by the gel as shown for hemoglobin and several enzymes (49). Repeated unfolding-refolding experiments could be performed on the same molecule (26,27).

Chemicals and buffers

All chemicals, purchased from Sigma-Aldrich (St. Louis, MO), except guanidinium chloride (GdnHCl) (Fluka, Buchs, Switzerland), were used without further purification. For denaturation experiments a stock solution of 5.3 M guanidinium hydrochloride (GuHCl) was prepared in 600 mM NaCl, 50 mM Tris at pH 6.8. NaCl was added to the buffer to screen the charges of the gel matrix. Because silica gels bear a net negative charge at pH around neutrality, sodium chloride was added to the buffer to shield the gel matrix charges and avoid partitioning of the denaturant molecules between the pores of the gel and the surrounding medium (16,46).

Experimental setup

GFPmut2 samples embedded in wet silica gels were mounted on the scanning stage (Physik Instrumente, Karlsruhe/Palmbach, Germany) (50) of an inverted microscope (TE200, Nikon, Tokyo, Japan). Circularly polarized 860-nm light from a mode-locked femtosecond pulsed infrared Ti:Sapphire laser (0.1-ps pulse width and 80-MHz repetition rate, Tsunami 3960, Spectra Physics, Mountain View, CA) was fed into the microscope and focused on the sample by means of an oil immersion objective (numerical aperture 1.4, 100×, Nikon). Under TPE (51), the excitation intensity on the sample was 10 kW cm^{-2}. Mechanical shutters (response time: 16–20 ms) were inserted in the optical path to stop excitation when needed: to avoid bleaching each single molecule underwent cycles of excitation, 55 s, followed by absence of excitation, 5 s (see Data Analysis). The molecule fluorescence collected by the objective, with 100-μs time resolution, was separated from the excitation by a dichroic mirror (495DCLP, Chroma Technology, Rockingham, VT). Blue and green emissions of GFPmut2 were selected with two long-pass filters (460/30LP, 515/30LP, Chroma Technology) and detected simultaneously by two single photon avalanche diodes (SPCM 15, EG & G, Salem, MA). All the experiments were performed at room temperature.

Unfolding-refolding experiments

Drops of the denaturant solution (5.3 M GuHCl) were added to the gel, and the acquisition of the fluorescence signal from a single molecule was started immediately. The denaturant reached the protein in a few seconds as clearly seen from the occurrence of a slight quenching in the fluorescence signal (26,27). GFPmut2 refolding was obtained by rapidly rinsing the silica gel with a buffer solution at pH 6.8 (26,27). The focal plane was maintained by microadjustment of the piezo control during the long-lasting denaturation experiments.

Data analysis

Each segment of the fluorescence trace (55 s at 100-μs resolution, i.e., 550,000 points) was processed through a Mathlab code which detects the occurrence of zero signals in the A channel correlated with a steady signal in the N channel. The extracted information consists of the number of switchings per trace, their duration, and the time of their disappearance. The A-N switching process was easily separated from the blinking process (27) by its duration (at most a few milliseconds with respect to hundreds of milliseconds or seconds for blinking) and by the lack of anticorrelation with the neutral channel. A-N switchings (27) as short as 100 μs (setup resolution) have not been included in the analysis. By following each protein molecule from the native to the denatured/unfolded state (the latter occurring when the fluorescence vanishes in both channels), one can record the K ''trajectories'', i.e., the rate of A-N switchings, evaluated for each 55-s timespan, versus time.

RESULTS

Spectroscopic characterization of GFPmut2 in wet silica gels

In single molecule experiments fluorescence bleaching limits the observation period (22,24,26,52). Two procedures have been adopted here to minimize bleaching effects. First, to reduce oxygen quenching, all measurements have been performed in a nitrogen atmosphere. Nitrogen does not affect protein unfolding, as verified on 50 single GFPmut2 molecules, it increases molecular brightness and it decreases bleaching rates. In particular, when exciting single proteins with $I = 10$ kW cm^{-2}, the adopted excitation intensity, the bleaching time is 82 ± 8 s.

Second, to make the observation time as long as possible, excitation has been periodically interrupted (55 s of steady

illumination followed by 5 s of absence of excitation) since under TPE (51) the fluorescence bleaching is related to thermal effects (52). In this way the same GFPmut2 molecule can be followed for long periods (from hours to days) (26). At ~10 kW cm^{-2} excitation intensity, the signal/noise ratio of the fluorescence emitted by single GFPmut2 molecules both in anionic and neutral channels is sufficiently high to resolve the switching events (Fig. 1 *a*).

A-N switching rate in native GFP

A section of the fluorescence trace of a single GFPmut2 molecule, simultaneously recorded in the A and N channels with 100-μs time resolution, is reported in Fig. 1 *a* where two A-N switching events are observed. In the following the symbol K^0 refers to A-N switching rates (number of switching events per second, averaged over a period of 55 s) of folded molecules, whereas K refers to the switching of unfolding proteins. It has been found that each molecule displays a particular value of K^0, shown in the histogram of Fig. 1 *b* for some 150 folded molecules. The histogram does not appear to obey a random distribution but, rather, some bunching of the K^0 values seems to occur at ~0.04 (0.03–0.06 range), 0.075 (0.06–0.08 range), 0.13 (0.10–0.15 range), 0.19 (0.18–0.21 range), and 0.23 (0.22–0.26 range) s^{-1}. The K^0 values are found to be independent of excitation intensity and stable in time for each GFPmut2 native molecule (observations extending to 24 h). On the contrary when the same molecule is unfolded and then refolded under 5.3 M GuHCl (26), its K^0 value may change to one that is different but that still belongs to the observed distribution, as shown for four molecules that have undergone a few unfolding-refolding cycles (Fig. 1 *b*, *inset*). Moreover, the observed absence of room temperature interconversions among the native substates, except upon unfolding-refolding cycling, indicates that the protein energy barriers are larger than the thermal energy, unlike other proteins such as myoglobin (1,5,53).

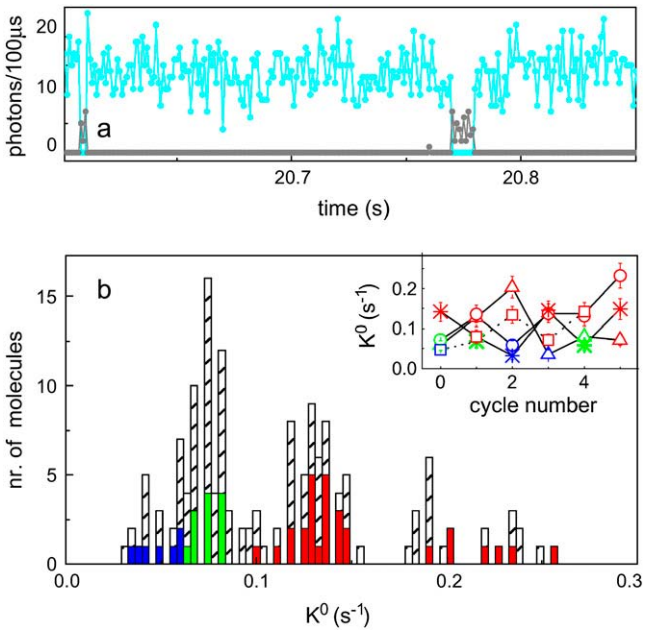

FIGURE 1 Fluorescence switching statistics of native GFPmut2 molecules. (*a*) Typical section of the fluorescence trace, simultaneously recorded in the cyan (anionic) and gray (neutral) channels, that shows two A-N switchings of a native molecule and displays the utmost anticorrelation of the signals. (*b*) Histogram of 150 native molecules versus K^0, the A-N switching rate, average of five runs. The colored bars represent ~50 molecules that have been followed during denaturation. Colors are coded according to the values of the preunfolding oscillations (17), K_{FIN} ~ 440 (*red*), ~ 720 (*green*), and ~ 930 Hz (blue). (*Inset*) K^0 values of four native molecules in the native state measured after each of six unfolding-refolding cycles. Each line connects the values of K^0, taken in consecutive cycles for the same molecule (same symbol). Colors are coded as indicated above.

A-N switching rate during GFP unfolding

The switching rate between the A and N states, K, of nearly 100 single GFPmut2 molecules has been recorded in the presence of 5.3 M GuHCl during the entire unfolding process. The completion of unfolding is indicated by the loss of fluorescence from both the A and N channels (26,27) at T_{UN}, the time elapsed after denaturant addition. Fig. 2 illustrates the time evolution of K for molecules grouped according to their unfolding time. Some unfolding times are found to occur with a larger probability than others, as can be seen in the histogram of Fig. 3 *a*. Fig. 2 reports, as examples, the molecules that unfolded in ~8, ~28, ~36, and ~86 min. In the presence of denaturant, the A-N switching rate, which is found to increase versus time more than 10-fold from the initial K^0 value, displays steep rises when unfolding times T_{UN} are short (e.g., 8–9 min, Fig. 2 *a*) and much more structured trends in the case of longer unfolding times (e.g., 86 min, Fig. 2 *d*) are seen.

Furthermore, just before (≅40 ms) the unfolding event, K increases even more and displays a very regular behavior on a ≅20-ms time stretch (Fig. 2, *insets*). By computing K_{FIN}, the frequency of the switching events occurring in the regular part of the trace (≅20 ms), we find that all the GFPmut2 molecules display only three distinct values: K_{FIN} = 440 (*red*), 720 (*green*), and 930 ± 30 s^{-1} (*blue*) (27), (Fig. 2, *insets*). It is remarkable that each molecule displays one and only one of the above K_{FIN} values after each refolding-unfolding cycle (27). Thus, K_{FIN} (as coded by the colors in the figures) can be assumed as a good indicator of the unfolding behavior of a single molecule, as shown in the following.

The K trajectories of some molecules are remarkably similar (well within the signal/noise ratio) as can be seen in Fig. 2: those molecules that display highly reproducible ''humps'' and ''jumps'' details in the ''K traces'' happen to show unfolding times T_{UN} and preunfolding frequencies

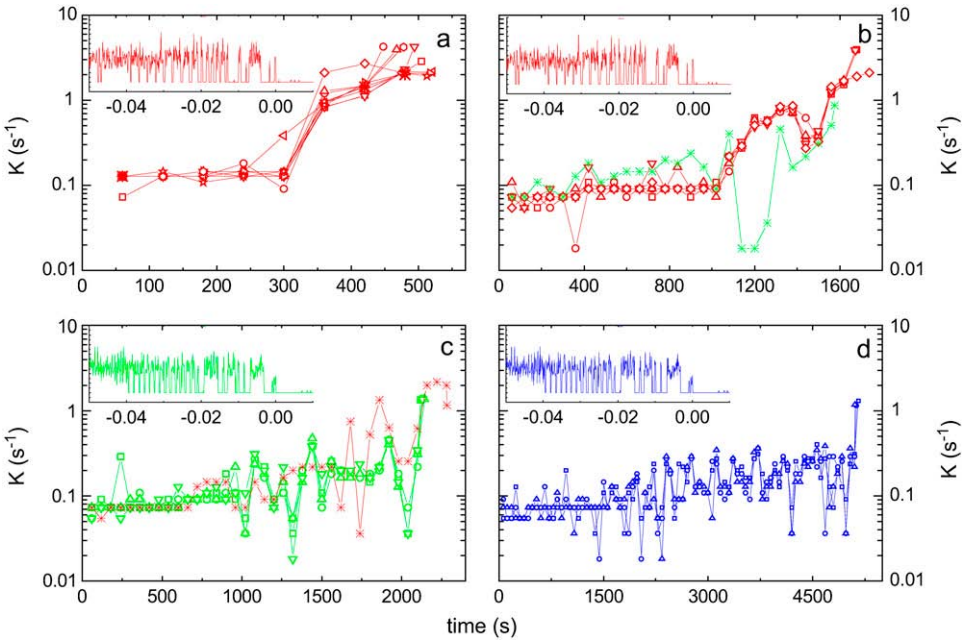

FIGURE 2 A-N switching rate K versus time for molecules unfolding in the presence of 5.3 M GuHCl. (*a*) The K versus time patterns for the molecules displaying T_{UN} = 8–9 min (eight molecules). (*b*) K patterns for the molecules displaying T_{UN} = 27–29 min (five molecules). (*c*) K patterns for the molecules displaying T_{UN} = 36–39 min (five molecules). (*d*) K patterns for the molecules displaying T_{UN} = 86 min (three molecules). Different symbols correspond to different molecules. Color code as in Fig. 1 *b*. (*Insets*) Final 50 ms of fluorescence emission of three GFPmut2 molecules displaying the characteristic A-N fluorescence oscillations at K_{FIN} = 440 (*red*, *panels a* and *b*), 720 (*green*, *panel c*), and 930 s^{-1} (*blue*, *panel d*). Ordinate values range from −5, offset value, to 30 photons/100 μs. Unfolding events have been translated to $t = 0$ for the three molecules shown here.

K_{FIN} that are very close. On the other hand, those molecules that display the same unfolding time but different K_{FIN} values are characterized by K trajectories that are markedly different from each other. For example, the four red trajectories (K_{FIN} = 440 s^{-1}, $T_{UN} \sim$ 28 min) appearing in Fig. 2 *b* share the same K trajectory, whereas the green one (K_{FIN} = 720 s^{-1}, $T_{UN} \sim$ 27 min) displays a different trend. Results very similar to those just described are found when the proteins are unfolded at lower GuHCl concentrations. The validity of the above observations is supported by the result that shows the K trend to be independent of excitation intensity, whether single or double photon, and of the excitation duty cycle (15–55 s excitation, 5–300 s no excitation).

DISCUSSION

Native conformational substates

Since the A-N transition reflects a GFP chemical/conformational change (proton transfer coupled to the β-barrel dynamics (23,30,31)), the observed structured distribution of the K^0 values suggests a multiplicity of conformational substates in the folded protein ensemble (1,36,54) contrary

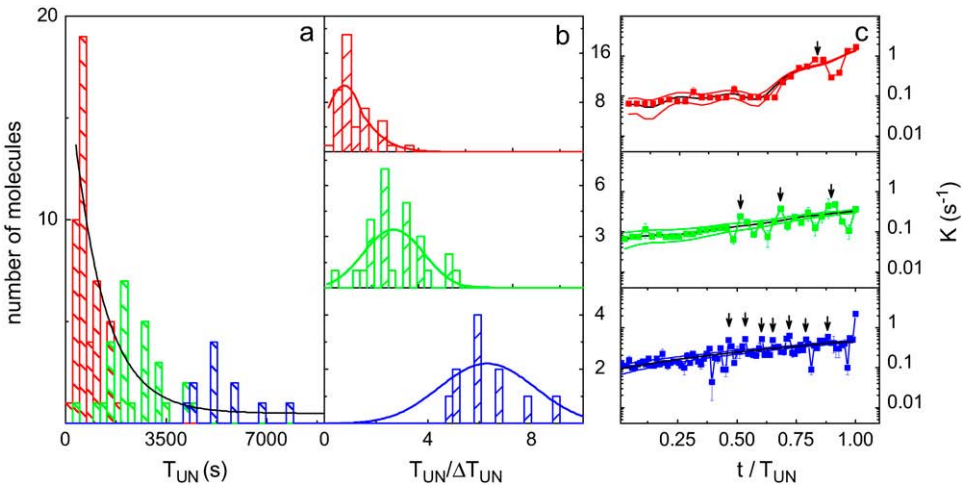

FIGURE 3 Unfolding statistics of single molecule GFPmut2. (*a*) Histogram reporting the number of unfolding GFPmut2 molecules versus T_{UN}, the time required to unfold (*red*, *green*, and *blue bars*). The black line is an exponential fit of the histogram (1200 ± 120 s time constant). Color code for the K_{FIN} values as in Fig. 1 *b*. (*b*) Histograms of the number of GFPmut2 molecules versus their unfolding time in units of $\langle \Delta T_{UN} \rangle$, the average time spacing between adjacent clusters of unfolding times. The average values are $\Delta T_{UN} \cong$ 630 ± 100, 790 ± 140, and 870 ± 110 s for the proteins that correspond to $K_{FIN} \cong$ 440, 720, and 930 s^{-1}, respectively. The solid lines are binomial distribution fits of the data. (*c*) Fluctuations of the switching rate K versus time in the presence of 5.3 M GuHCl for molecules displaying K_{FIN} values according to the color code of panel *a*. The solid lines in each panel are the average K trend (Fourier averaging with 1/300 s^{-1} cutoff frequency), and the dashed lines correspond to $\pm 3\sigma$ in the K values. The panels correspond to: top $T_{UN}/\langle \Delta T_{UN} \rangle = 1$, $K_{FIN} \cong 440$ s^{-1}; middle $T_{UN}/\langle \Delta T_{UN} \rangle = 3$, $K_{FIN} \cong 720$ s^{-1}; and bottom $T_{UN}/\langle \Delta T_{UN} \rangle = 6$, $K_{FIN} \cong 930$ s^{-1}.

to other parameters, such as brightness or lifetime, whose values do not reveal the existence of components. The assumption of protein substates characterized by K^0 gains further support when labeling the native molecules according to the K_{FIN} values displayed at the unfolding instant. Fig. 1 *b* shows that molecules with the largest K^0 values display the smallest K_{FIN} values and vice versa, thereby suggesting a close relationship between native states and unfolding pathways (38).

Unfolding analysis

The 86 single molecules' distribution versus unfolding time T_{UN} is compatible with an exponential decay, when binned over times of the order of 10 min or larger (not shown), in agreement with previously reported bulk experiments (25). When binning times ≅4–6 min are assumed for the histogram analysis, some structure becomes apparent, as shown in Fig. 3 *a*, with a peak at $T_{UN} \sim$ 7–8 min, and significant fluctuations in the number of molecules are seen at longer unfolding times, suggesting a heterogeneous response of the protein ensemble. In fact, when classifying the molecules according to the specific values of their final oscillations, K_{FIN} (color code in Fig. 2) (27), then the T_{UN} histogram appears characterized by three distinct distributions (see *colored bars*, Fig. 3, *a–b*). From the association of T_{UN} with K_{FIN}, it becomes evident that the fastest unfolding molecules, $T_{UN} =$ 6–30 min, display the lowest preunfolding frequency, $K_{FIN} = 440\ s^{-1}$, the slowest unfolding ones, $T_{UN} =$ 60–100 min, oscillate at $930\ s^{-1}$, whereas those in between exhibit $K_{FIN} = 720\ s^{-1}$. The faster the unfolding, the lower the associated preunfolding frequency.

It should be noted that the vertical bars in the three bell-shaped distributions are almost equally spaced by a time ΔT_{UN} as shown in Fig. 3 *b*. For each subpopulation of molecules displaying the same K_{FIN} value, we can estimate an average value of ΔT_{UN}: $\langle \Delta T_{UN} \rangle_{440} = 630 \pm 100$, $\langle \Delta T_{UN} \rangle_{720} = 790 \pm 140$, $\langle \Delta T_{UN} \rangle_{930} = 870 \pm 110$ s. When the timescale of the unfolding time histograms is given in units of $\langle \Delta T_{UN} \rangle$, then, to a good approximation, the three distributions fall at integer values of $T_{UN}/\langle \Delta T_{UN} \rangle$ (Fig. 3 *b*), suggesting a binomial description according to

$$P(x,p,N) = \binom{N}{x} p^x (1-p)^{N-x},$$

where $x = t/\langle \Delta T_{UN} \rangle$ is the step number obeying $0 < x < N$, p is the probability of each step, and $\binom{N}{x}$ is given by

$$\binom{N}{x} = \frac{f(N)}{f(x)f(N-x)},$$

where the factorial f in the Stirling approximation is $f(x) = x^x e^{-x}\sqrt{2\pi x}[1+1/(12x)]$.

In particular, the population of molecules unfolding with $K_{FIN} = 440\ s^{-1}$ is fitted by an average number of steps $\langle N \rangle = 1.6 \pm 0.1$; the molecules with $K_{FIN} = 720\ s^{-1}$ yield $\langle N \rangle = 3.4 \pm 0.2$; the molecules with $K_{FIN} = 930\ s^{-1}$ give $\langle N \rangle = 6.2 \pm 0.6$, when p is assumed to be 0.5 (Fig. 3 *b*). The binomial data fit shows that the recorded unfolding processes are well described by a finite number of steps N that grows with the unfolding time. Surprisingly, the average number of steps, $\langle N \rangle$, is found to be close to the number of humps in the unfolding K traces that are visible in Fig. 2.

The time course of the switching rate, K, of single GFPmut2 molecules in the presence of 5.3 M GuHCl (Fig. 2) displays humps and bumps whose amplitude clearly exceeds the standard error as estimated from the number of events collected per sampling time (~55 s). The K fluctuations are more clearly seen by comparing the data with their average trend (Fourier averaging with $1/300\ s^{-1}$ cutoff frequency) as shown in Fig. 3 *c*. To this purpose we have considered the average K time trajectory of the molecules displaying $T_{UN} =$ 28, 36, and 86 min, respectively, and reported in Fig. 2, *b–d*. For each trajectory a discrete number of peaks is evident, their number depending on the cutoff adopted for separating fluctuations from noise ($\pm 3\sigma$ lines in Fig. 3 *c*). The number of peaks indicated by the arrows in Fig. 3 *c* corresponds fairly well to the best fit value $\langle N \rangle$ obtained from the binomial fit of the T_{UN} distributions, Fig. 3 *b*. For the molecules with $K_{FIN} \cong 440$, 720, and $930\ s^{-1}$, the number of peaks emerging from the background is 1, 3, and 7, respectively, to be compared to the values of $\langle N \rangle = 1.6$, 3.4, and 6.2. The observed correlation suggests that a relation exists between the occurrence of discrete unfolding times and the number of fluctuations in the switching rate, K, observed versus time under the action of denaturant.

Native-unfolded correlation

A strong support to the previous analysis is found when comparing all the results: unfolding times (T_{UN}), preunfolding oscillation frequencies (K_{FIN}, color coded), and native switching rates (K^0) (Fig. 4). Native substates, characterized by higher values of K^0, correspond to molecules that unfold faster and vice versa, a finding that allows us to make straightforward predictions of the unfolding time from the initial measurement of K^0 for any single molecule. As an example, molecules with $K^0 \sim 0.23\ s^{-1}$ unfold displaying $K_{FIN} = 440\ s^{-1}$ and $T_{UN} < 300$ s, whereas those with $K^0 \sim 0.04\ s^{-1}$ display $K_{FIN} = 930\ s^{-1}$ and $T_{UN} > 5000$ s. Fig. 4 shows the K^0-T_{UN} correlation, with four distinct regimes as indicated by the solid line.

The inescapable consequence of these observations is that the unfolding process can only occur along discrete pathways in the conformational space of the protein and that a given pathway (or a set of them) is dictated by a particular initial substate (or set). Moreover, the highly reproducible unfolding K patterns suggest that a deterministic kinetics, probably due to the denaturant driving the protein to a new equilibrium, dominates over the stochastic components. It is

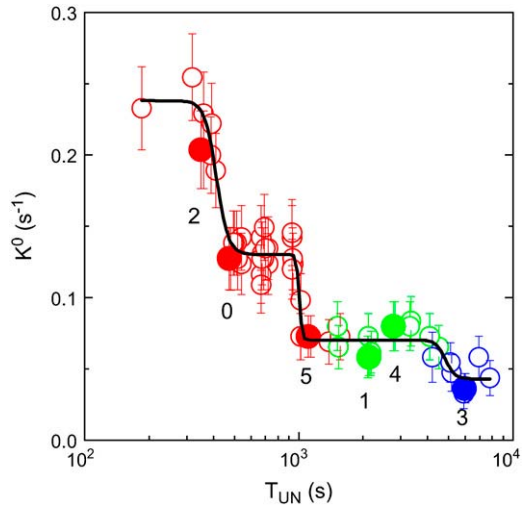

FIGURE 4 Average switching rate K^0 of 56 GFPmut2 native molecules versus unfolding time, T_{UN}. Colors code according to the K_{FIN} values, as in Fig. 2. The bars represent the standard error. The solid circles, labeled 0–5, correspond to the same molecule that has undergone five consecutive unfolding-refolding cycles.

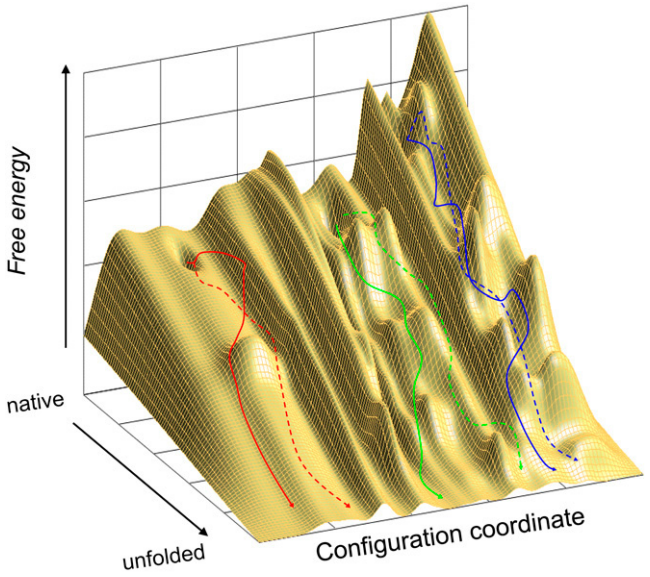

FIGURE 5 Pictorial representation of the free energy landscape explored by unfolding GFPmut2. The native substates, starting points of the colored unfolding paths, are placed at different reaction coordinate values that correspond to different free energy values (*top* of picture). Continuous and broken lines are free representations of different unfolding pathways. The routes end at low free energy values where the molecules are expected to assume a variety of unfolded conformations. The rugged free energy landscape versus reaction coordinate in the picture is suggested by the bumpy K traces shown in Fig. 2.

also interesting that the same molecule, upon each unfolding-refolding cycle, displays a value of the unfolding time that, although appearing to be chosen at random, obeys the K^0-T_{UN} -K_{FIN} correlation (Fig. 4, *solid circles*). This suggests that one could explore extended regions of the GFPmut2 energy landscape by repeated unfolding-refolding cycles performed on just one molecule. At room temperature a molecule cannot cross the barriers that trap the native substate, whereas after a strong albeit reversible perturbation such as unfolding-refolding, one of the other native substates becomes accessible.

Energy landscape

A visual summary of the results of this investigation is attempted in the pictorial representation of Fig. 5 where an imaginary, albeit plausible, energy landscape versus a reaction coordinate is the scenario for the unfolding kinetics of GFPmut2. The three major valleys host the unfolding pathways and correspond to the three binomial components in the T_{UN} histogram of Fig. 3 *b*. The bumps in the energy landscape stand for transition states encountered by the protein during unfolding (55,56). The unfolding paths are drawn by connecting a native state, or a group of them, to an unfolded one along discrete downhill winding routes. The pathway colors code for the characteristic preunfolding oscillations (27).

CONCLUSIONS

According to the energy landscape theory description (1,17,36), folding and unfolding occur through multiple paths winding around local energy barriers in the energy landscape. These discrete paths and the heterogeneous protein kinetics emerge when the landscape becomes rough and local traps become important, such as below the folding temperature. We have shown here that the fluorescence signal induced by the chemical/conformational states' dynamics (30,31) is the key to access the heterogeneous response of GFPmut2 proteins in the native state and during unfolding. The observed different values of K^0, the A-N switching rate, lead to clear experimental evidence of a discrete multiplicity of native substates in GFPmut2. The data in Fig. 1 suggest a fine structured population of the protein states, and a kind of hierarchy is apparent.

These native substates are found to evolve, under the action of denaturants, along specific K paths toward unfolding. The K trajectories of molecules taking the same time to unfold show reproducible details, thereby suggesting the existence of a few selected routes that are those allowed to accomplish the process. These observations are strengthened by the striking correlation between the native state properties, expressed by K^0, and those of the unfolded state, summarized by T_{UN} and K_{FIN}, together with the observed nonexponential distribution of the unfolding times. This finding is in agreement with the analysis of the protein dynamics along an energy landscape whose roughness becomes relevant below the glassy temperature (8,13,36,57).

The detection of a multiplicity of native substates and unfolding pathways in a protein, accomplished with single molecule techniques as shown here, should allow a more quantitative description of the heterogeneities hidden in biological processes. The coupling of single molecule physical investigations to biological methodologies, such as in situ specific mutagenesis, would deepen the comprehension of the molecular basis of relevant biological processes, such as folding/unfolding and enzymatic reactions.

This work was partially supported by INFM-CNR Fondo Integrativo per la Ricerca di Base, Cofin Ministero dell'Istruzione Università e Ricerca Scientifica, Progetto RBNE03PX83_006, and Fondazione Cariplo 2005-1073 and PRIN 2006-025255.

REFERENCES

1. Frauenfelder, H., F. Parak, and R. D. Young. 1988. Conformational substates in proteins. *Annu. Rev. Biophys. Biophys. Chem.* 17:451–479.
2. Frauenfelder, H., S. Sliga, and P. Wolynes. 1991. The energy landscapes and motions of proteins. *Science.* 254:1598–1603.
3. Eisenmesser, E. Z., O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, D. A. Bosco, J. J. Skalicky, L. E. Kay, and D. Kern. 2005. Intrinsic dynamics of an enzyme underlies catalysis. *Nature.* 438: 117–121.
4. Huang, Y. P. J., and G. T. Montelione. 2005. Structural biology—-proteins flex to function. *Nature.* 438:36–37.
5. Austin, R. H., K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus. 1975. Dynamics of ligand binding to myoglobin. *Biochemistry.* 14:5355–5373.
6. Koehler, C. W., and J. Friedrich. 1989. Probing of conformational relaxation processes of proteins by frequency labeling of optical states. *J. Chem. Phys.* 90:1270–1273.
7. Friedrich, J., J. Gafert, J. Zollfrank, J. Vanderkooit, and J. Fidy. 1994. Spectral hole burning and selection of conformational substates in chromoproteins. *Proc. Nati. Acad. Sci. USA.* 91:1029–1033.
8. Zhou, Y. Q., C. Zhang, G. Stell, and J. Wang. 2003. Temperature dependence of the distribution of the first passage time: results from discontinuous molecular dynamics simulations of an all-atom model of the second beta-hairpin fragment of protein G. *J. Am. Chem. Soc.* 125:6300–6305.
9. Leeson, D. T., D. A. Wiersma, K. Fritsch, and J. Friedrich. 1997. The energy landscape of myoglobin: an optical study. *J. Phys. Chem. B.* 101:6331–6340.
10. Zhuang, X., L. E. Bartley, H. P. Babcock, R. Russell, T. Ha, D. Herschlag, and S. Chu. 2000. A single-molecule study of RNA. Catalysis and folding. *Science.* 288:2048–2051.
11. Jia, Y., D. S. Talaga, W. L. Lau, H. S. M. Lu, W. F. DeGrado, and R. M. Hochstrasser. 1999. Folding dynamics of single GCN-4 peptides by fluorescence resonant energy transfer confocal microscopy. *Chem. Phys.* 247:69–83.
12. Deniz, A. A., T. A. Laurence, G. S. Beligere, M. Dahan, A. B. Martin, D. S. Chemla, P. E. Dawson, P. G. Schultz, and S. Weiss. 2000. Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. Proc. Natl. Acad. Sci. USA. 97:5179–5184.
13. Leite, V. B. P., L. C. P. Alonso, M. Newton, and J. Wang. 2005. Single molecule electron transfer dynamics in complex environments. *Phys. Rev. Lett.* 95:118301.
14. Peleg, G., P. Ghanouni, B. K. Kobika, and R. N. Zare. 2001. Single-molecule spectroscopy of the beta(2) adrenergic receptor: observation of conformational substates in a membrane protein. *Proc. Natl. Acad. Sci. USA.* 98:8469–8474.
15. Finke, J. M., and J. N. Onuchic. 2005. Equilibrium and kinetic folding pathways of a TIM barrel with a funneled energy landscape. *Biophys. J.* 98:488–505.
16. Leite, V. B., J. N. Onuchic, G. Stell, and J. Wang. 2004. Probing the kinetics of single molecule protein folding. *Biophys. J.* 87:3633–3641.
17. Onuchic, J. N., and P. G. Wolynes. 2004. Theory of protein folding. *Curr. Opin. Struct. Biol.* 14:70–75.
18. Nymeyer, H., A. E. Garcia, and J. N. Onuchic. 1998. A landscape analysis of friction effects in protein folding. *Proc. Natl. Acad. Sci. USA.* 95:5921–5928.
19. Tsien, R. Y. 1998. The green fluorescent protein. *Annu. Rev. Biochem.* 67:509–544.
20. Zimmer, M. 2002. Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. *Chem. Rev.* 102:759–781.
21. Cormack, B. C., R. H. Valdivia, and S. Falkow. 1996. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene.* 173:33–38.
22. Chirico, G., F. Cannone, S. Beretta, A. Diaspro, B. Campanini, S. Bettati, R. Ruotolo, and A. Mozzarelli. 2002. Dynamics of green fluorescent protein mutant2 in solution, on spin-coated glasses, and encapsulated in wet silica gels. *Protein Sci.* 11:1152–1161.
23. Abbruzzetti, S., E. Grandi, C. Viappiani, S. Bologna, B. Campanini, S. Raboni, S. Bettati, and A. Mozzarelli. 2005. Kinetics of acid-induced spectral changes in the GFPmut2 chromophore. *J. Am. Chem. Soc.* 127:626–635.
24. Cannone, F., M. Caccia, S. Bologna, A. Diaspro, and G. Chirico. 2004. Single molecule spectroscopic characterization of GFP-MUT2 mutant for two-photon microscopy applications. *Microsc. Res. Tech.* 65:186–193.
25. Campanini, B., S. Bologna, S. Bettati, F. Cannone, G. Chirico, and A. Mozzarelli. 2005. Unfolding of green fluorescent protein mut2 in wet nanoporous silica gels. *Protein Sci.* 14:1125–1133.
26. Cannone, F., S. Bologna, B. Campanini, A. Diaspro, S. Bettati, A. Mozzarelli, and G. Chirico. 2005. Tracking unfolding and refolding of single GFPmut2 molecules. *Biophys. J.* 89:2033–2045.
27. Baldini, G., F. Cannone, and G. Chirico. 2005. Pre-unfolding resonant oscillations of single green fluorescent protein molecules. *Science.* 309:1096–1100.
28. Dickson, R. M., A. B. Cubitt, R. Y. Tsien, and W. E. Moerner. 1997. On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature.* 338:355–358.
29. Haupts, U., S. Maiti, R. Schwille, and W. W. Webb. 1998. Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. *Proc. Natl. Acad. Sci. USA.* 95:13573–13578.
30. Scharnagl, C., R. Raupp-Kossmann, and S. F. Fischer. 1999. Molecular basis for pH sensitivity and proton transfer in green fluorescent protein: protonation and conformational substates from electrostatic calculations. *Biophys. J.* 77:1839–1857.
31. Saxena, A. M., J. B. Udgaonkar, and G. Krishnamoorthy. 2005. Protein dynamics control proton transfer from bulk solvent to protein interior: a case study with a green fluorescent protein. *Protein Sci.* 14:1787–1799.
32. Onuchic, J. N., Z. Luthey-Schulten, and P. G. Wolynes. 1999. Analyzing single molecule trajectories on complex energy landscapes using replica correlation functions. *Chem. Phys.* 271:175–184.
33. Onuchic, J. N., H. Nymeyer, A. E. Garcia, J. Chahine, and N. D. Socci. 2000. The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. *Adv. Protein Chem.* 53:87–152.
34. Laurents, D. V., and R. L. Baldwin. 1998. Protein folding: matching theory and experiment. *Biophys. J.* 75:428–434.
35. Wang, J., and G. M. Verkhivker. 2003. Energy landscape theory, funnels, specificity, and optimal criterion of biomolecular binding. *Phys. Rev. Lett.* 90:188101–1.
36. Wang, J. 2003. Statistics, pathways and dynamics of single molecule protein folding. *J. Chem. Phys.* 118:952–958.

37. Eaton, W. A., V. Munoz, A. J. Hagen, G. S. Jas, L. J. Lapidus, E. R. Henry, and J. Hofrichter. 2000. Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 29:327–359.

38. Schuler, B., E. A. Lipman, and W. A. Eaton. 2002. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature.* 419:743–747.

39. Lipman, E. A., B. Schuler, O. Bakajin, and W. A. Eaton. 2003. Single-molecule measurement of protein folding kinetics. *Science.* 301: 1233–1235.

40. Schuler, B. 2005. Single-molecule fluorescence spectroscopy of protein folding. *ChemPhysChem.* 6:1206–1220.

41. Weber, W., V. Helms, J. A. McCammon, and P. W. Langhoff. 1999. Shedding light on the dark and weakly fluorescent states of green fluorescent proteins. *Proc. Natl. Acad. Sci. USA.* 96:6177–6182.

42. Mallik, R., J. B. Udgaonkar, and G. Krishnamoorthy. 2003. Kinetics of proton transfer in a green fluorescent protein: a laser-induced pH jump study. *Proc. Indian Natl. Sci. Acad. B Biol. Sci.* 115:307–317.

43. Brejc, K., T. K. Sixma, P. A. Kitts, S. R. Kain, R. Y. Tsien, M. Ormö, and S. J. Remington. 1997. Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proc. Natl. Acad. Sci. USA.* 94:2306–2311.

44. Elsliger, M. A., R. M. Wachter, G. T. Hanson, K. Kallio, and S. J. Remington. 1999. Structural and spectral response of green fluorescent protein variants to changes in pH. *Biochemistry.* 38:5296–5301.

45. Collini, M., L. D'Alfonso, and G. Baldini. 2003. Trehalose induced changes of the ethidium hydration shell detected by time resolved fluorescence. *Photochem. Photobiol.* 77:376–382.

46. Badjic, J. D., and N. M. Kostic. 1999. Effects of encapsulation in sol-gel silica glass on esterase activity, conformational stability, and unfolding of bovine carbonic anhydrase II. *Chem. Mat.* 11:3671–3679.

47. Bettati, S., B. Pioselli, B. Campanini, C. Viappiani, and A. Mozzarelli. 2004. Protein-doped nanoporous silica gels. *In* Encyclopedia of Nanoscience and Nanotechnology, Vol. 9. H. S. Nalwa, editor. American Scientific Publishers, Stevenson Ranch, CA. 81–103.

48. Bettati, S., and A. Mozzarelli. 1997. T state hemoglobin binds oxygen noncooperatively with allosteric effects of protons, inositol hexaphosphate, and chloride. *J. Biol. Chem.* 272:32050–32055.

49. Viappiani, C., S. Bettati, S. Bruno, L. Ronda, S. Abbruzzetti, A. Mozzarelli, and W. A. Eaton. 2004. New insights into allosteric mechanisms from trapping unstable protein conformations in silica gels. *Proc. Natl. Acad. Sci. USA.* 101:14414–14419.

50. Malengo, G., R. Milani, F. Cannone, S. Krol, A. Diaspro, and G. Chirico. 2004. High sensitivity optical microscope for single molecule spectroscopy studies. *Rev. Sci. Instrum.* 75:2746–2751.

51. Diaspro, A., G. Chirico, and M. Collini. 2005. Two-photon fluorescence excitation in biological microscopy and related techniques. *Q. Rev. Biophys.* 38:97–166.

52. Chirico, G., F. Cannone, G. Baldini, and A. Diaspro. 2003. Two-photon thermal bleaching of single fluorescent molecules. *Biophys. J.* 84:588–598.

53. Frauenfelder, H., B. H. McMahon, R. H. Austin, K. Chu, and J. T. Groves. 2001. The role of structure, energy landscape, dynamics, and allostery in the enzymatic function of myoglobin. *Proc. Natl. Acad. Sci. USA.* 98:2370–2374.

54. Frauenfelder, H. 2002. Proteins: paradigms of complexity. *Proc. Natl. Acad. Sci. USA.* 99:2479–2480.

55. Itzhaki, L. S., D. E. Otzen, and A. R. Fersht. 1995. The structure of the transition-state for folding of chymotrypsin inhibitor-2 analyzed by protein engineering methods—evidence for a nucleation-condensation mechanism for protein-folding. *J. Mol. Biol.* 254:260–288.

56. Dill, K. A., and H. S. Chan. 1997. From Levinthal to pathways to funnels. *Nature.* 4:10–19.

57. Wang, J. 2004. The complex kinetics of protein folding in wide temperature ranges. *Biophys. J.* 87:2164–2171.