Native-unlike Long-lived Intermediates along the Folding Pathway of the Amyloidogenic Protein \(\beta_2\)-Microglobulin Revealed by Real-time Two-dimensional NMR*†

Received for publication, September 17, 2009, and in revised form, December 4, 2009. Published, JBC Papers in Press, December 22, 2009, DOI 10.1074/jbc.M109.061168

Alessandra Corazza 1, Enrico Rennella 1, Paul Schanda 2, Maria Chiara Mimmi 3, Thomas Cutuil 4, Maria Chiara Mimmi 3, Thomas Cutuil 4, Sofia Giorgetti 4, Federico Fogolari 4, Paolo Viglino 4, Lucio Frydman 4, Maayan Gal 4, Vittorio Bellotti 1, Bernhard Brutscher 6, and Gennaro Esposito 5

From the 4Department of Biomedical Science and Technology, University of Udine, Piazzale Kolbe 4, 33100 Udine, Italy, the 5Institut de Biologie Structurale Jean-Pierre Ebel, 41 Rue Jules Horowitz, UMR5075 CNRS-Université Joseph Fourier-Centre d’Etude Atomique, 41, rue Jules Horowitz, 38027 Grenoble Cedex 9, France, the 6Department of Biochemistry, University of Pavia, Via Taramelli 3b, 27100 Pavia, Italy, the 1Biotechnology Laboratory, Fondazione IRCCS, PoliclinicoS, Matteo, 27100 Pavia, Italy, and the 2Department of Chemical Physics, Weizmann Institute of Science, Rehovot 76100, Israel

\(\beta_2\)-microglobulin (\(\beta_2m\)), the light chain of class I major histocompatibility complex, is responsible for the dialysis-related amyloidosis and, in patients undergoing long term dialysis, the full-length and chemically unmodified \(\beta_2m\) converts into amyloid fibrils. The protein, belonging to the immunoglobulin superfamily, in common to other members of this family, experiences during its folding a long-lived intermediate associated to the \(\text{trans}\)-\(\text{cis}\) isomerization of Pro-32 that has been addressed as the precursor of the amyloid fibril formation. In this respect, previous studies on the W60G \(\beta_2m\) mutant, showing that the lack of Trp-60 prevents fibril formation in mild aggregating condition, prompted us to reinvestigate the refolding kinetics of wild type and W60G \(\beta_2m\) at atomic resolution by real-time NMR. The analysis, conducted at ambient temperature by the band selective flip angle short transient real-time two-dimensional NMR techniques and probing the \(\beta_2m\) states every 15 s, revealed a more complex folding energy landscape than previously reported for wild type \(\beta_2m\), involving more than a single intermediate species, and shedding new light into the fibrolytic pathway. Moreover, a significant difference in the kinetic scheme previously characterized by optical spectroscopic methods was discovered for the W60G \(\beta_2m\) mutant.

Among the amyloidogenic proteins that have been most frequently studied over the last years, there is \(\beta_2\)-microglobulin (\(\beta_2m\)) 4 that is responsible for dialysis-related amyloidosis.

**This work was supported by grants from Ministero dell’Istruzione, dell’Università e della Ricerca (FIRB Grant RBNE03PX83 and RBRNO7BMCT and PRIN 2007_XY59ZJ_002), the European Union (EU) (EURAMY Project LSHM-CT-2005-037525; EU-NMR Contract 026145), and the French Agence Nationale pour la Recherche (Grant ANR-08-BLAN-0210).**

† The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables 1 and 2 and Figs. 1 and 2.

‡ To whom correspondence may be addressed. Tel.: 39-0432-494322; Fax: 39-0432-494301; E-mail: alessandra.corazza@uniud.it.

§ To whom correspondence may be addressed. Tel.: 39-0432-494322; Fax: 39-0432-494301; E-mail: gennaro.esposito@uniud.it.

1 The abbreviations used are: \(\beta_2m\), \(\beta_2\)-microglobulin; SOFAST, selective flip angle short transient; HMQC, heteronuclear multiple quantum coherence; FTA, fluid turbulence-adapted; I, intermediate; N, native; U, unfolded state.

‡‡1 This work was supported by grants from Ministero dell’Istruzione, dell’Università e della Ricerca (FIRB Grant RBNE03PX83 and RBRNO7BMCT and PRIN 2007_XY59ZJ_002), the European Union (EU) (EURAMY Project LSHM-CT-2005-037525; EU-NMR Contract 026145), and the French Agence Nationale pour la Recherche (Grant ANR-08-BLAN-0210).

‡‡‡1 The abbreviations used are: \(\beta_2m\), \(\beta_2\)-microglobulin; SOFAST, selective flip angle short transient; HMQC, heteronuclear multiple quantum coherence; FTA, fluid turbulence-adapted; I, intermediate; N, native; U, unfolded state.
that, in contrast to P32G, the mutant P32V does not undergo any amyloid transition. They also pointed out the necessity of introducing a coupling between the trans-cis isomerization and the adjacent kinetic step and suggested a minimal folding model (7). The apparently monoeponential unfolding pattern of the spectroscopic traces was thus rationalized as the consequence of the occurrence of two processes with similar timing: unfolding and trans-cis peptidyl-prolyl isomerization between positions 31 and 32. Recently, further evidence of a substantially native-like structure of the refolding intermediate for the constant domain of an antibody light chain (C\textsubscript{l}) was provided by NMR experiments, performed at very low temperature, and molecular dynamics simulations (8). It is worth noting that \(\beta2m\) and C\textsubscript{l} belong to the same immunoglobulin superfamily, but despite the presence of a slow refolding intermediate, C\textsubscript{l} does not convert into amyloid fibrils.

Traditionally, protein folding studies have been performed by following the reaction by optical spectroscopies such as fluorescence or circular dichroism coupled with efficient mixing devices to obtain a good temporal resolution. Although these methods are characterized by high sensitivity and good time resolution, they do not provide detailed residue-resolved information on the folding process because of their reliance on only very few detectable probes. Multidimensional NMR spectroscopy overcomes some of these limitations as kinetic and structural information can be obtained in real-time during the refolding process for a large number of sites within the protein. A major breakthrough for the application of real-time NMR to atom-resolved studies of protein folding has been the introduction of fast and ultrafast multidimensional NMR techniques (9, 10). In particular, the SOFAST-HMQC (11–13) and ultraSOFAST-HMQC (14, 15) experiments allow the recording of two-dimensional \(^1\text{H}-\text{\textsuperscript{15}}\text{N}\) correlation spectra of proteins at \(-0.1–1\ \text{s}^{-1}\) repetition rates, thus providing the required time resolution for the study of kinetic molecular processes occurring on the seconds-to-minutes time scale.

A number of mutations of \(\beta2m\) have been exploited to reveal possible determinants in the fibrillogenesis phenomenon; in particular, Trp-60 was identified by molecular dynamics simulations as one of the essential residues for the aggregation process (16). In this work, SOFAST real-time NMR techniques are employed to gain more insight into the folding process of wild type and W60G \(\beta2m\) at ambient temperature. The results, obtained by probing the folding every 15 s at atomic level, show that the process is more complex than previously reported and seems to involve more than a single long-lived intermediate. Moreover, our NMR data reveal a significant difference in the folding pathway of W60G \(\beta2m\) when compared with wild type \(\beta2m\), which was not detectable by optical spectroscopic methods (17).

**EXPERIMENTAL PROCEDURES**

**Proteins**—Expression and purification of wild type and W60G \(\beta2m\) was carried out as reported previously (17, 18) with additional \textsuperscript{15}N uniform labeling. A methionine residue was always present at the N-terminal position of all recombinant products. All the recombinant proteins gave a single species when assayed by electrospray-ionization mass spectrometry.

**Refolding Protocol**—The pH-jump protocol here adopted has already been reported by Kameda et al. (3). The protein was unfolded in the denaturing solution (23 mM HCl, 1.5 M urea, pH 2.2), and the pH was determined to range between 2.4 and 3.0. The refolding was started by injecting 100 \(\mu\text{L}\) of refolding buffer (300 mM phosphate, 1.5 M urea, pH 7.4) into 360 \(\mu\text{L}\) of unfolded \(\beta2m\) solution using a fast injection device for rapidly mixing the two solutions inside the NMR magnet (13). At the end of the folding process and of the measurements, the final pH was determined to range between 6.6 and 6.8. The temperature of all refolding measurements was 23.9 °C.

**NMR Data Acquisition and Processing**—All NMR experiments were performed with a Varian INOVA spectrometer operating at 800 MHz (\(^1\text{H}\) frequency), equipped with a cryogenic probe. The refolding process was started inside the spectrometer, using an injection device, as already described elsewhere (13). The reactions were then followed through a series of FTA-SOFAST-HMQC spectra (13) of about 15-s duration. Short interscan delays could be used because of the efficient \(T_1\) relaxation, resulting in overall single-scan times of \(-100\ \text{ms}(11)\). Nitrogen decoupling was achieved by the use of a WURST-40 train (19), whereas the States et al. (20) scheme was employed for sign discrimination with respect to the nitrogen carrier frequency. Each experiment was performed with only one scan for \(T_1\) increment, and phase cycling was performed in subsequent spectra to eliminate artifacts, as described before (13). The experiments were processed with NMRPipe (21) in the indirect dimension F1, the original data set of 60 real points was extended by linear prediction (30 points) and zero-filled to 128 points, and a squared sine-bell function was employed, with a shift of \(\pi/4\). In the direct dimension F2, a sine-bell function with a shift of \(2\pi/5\) was employed. Following two-dimensional Fourier transform, a fourth order polynomial baseline correction was applied.

**NMR Data Analysis**—Spectra inspection, peak assignment, and volume calculation were performed using NMRView (22). Different classes of peaks, detailed under “Results,” were analyzed. Here, it is worth noting that for native (N) and intermediate (I) species peaks, the boxes, used in volume measurements to confine the peak areas, comprise only one peak, whereas for I+N peaks, the boxes are bigger and contain two peaks belonging to I and N species. As a consequence, the error on the measured volumes is likely to be higher for the peaks of the latter class than for those of the first two classes. NMR assignments were taken from Biological Magnetic Resonance Bank (BMRB) entries 3078 (23) and 15480 (18), with slight adjustments arising from thermal variation of the chemical shifts. Fits of the time course of the peak volumes were done using Mathematica 6 (Wolfram Research). Statistical analysis of the fit goodness was performed using Student’s \(t\) test for the fitting parameters, using \(F\)-tests on the sum of the squares of the residuals, and comparing the adjusted \(R^2\) square values.

**RESULTS**

A Single Slow Phase Does Not Account for \(\beta2m\) Folding—In a recent study (17), and in agreement with published data (1, 2), we confirmed the change in tryptophan fluorescence emission during the folding process of wild type \(\beta2m\), consistent with the
Two-dimensional NMR Reveals the Folding Pathway of \( \beta_{2m} \)

presence of a slow phase ascribed to Pro-32 isomerization (3, 4). Analogous experiments were performed on the W60G mutant, for which a reduced association and aggregation propensity had been revealed (17). Unlike wild type \( \beta_{2m} \), no folding slow phase could be detected through fluorescence spectra for this mutant, suggesting that W60G \( \beta_{2m} \) reaches a native-like conformation in a couple of minutes, at least as sensed by the fluorescence probes. It is worth noting that the fluorescence changes upon folding of \( \beta_{2m} \) are essentially due to Trp-95 that becomes buried in the native conformation, whereas Trp-60 contributes only marginally because it remains solvent-exposed also in the folded state, as shown by Kihara et al. (24).

To gain further insight into the folding process of \( \beta_{2m} \) species, we used real-time two-dimensional SOFAST NMR spectroscopy. The refolding of wild type and W60G \( \beta_{2m} \) was studied at 23.9 °C starting from protein samples in 1.5 M urea at pH = 2.2. To increase the timing efficiency, the refolding reaction was initiated inside the magnet (13) by rapidly injecting the refolding phosphate buffer at pH = 7.4. The pH that was finally achieved ranged from 6.6 to 6.8 for the different samples. The refolding process was followed by acquisition of a series of SOFAST-HMQC spectra recorded at a 0.07 s\(^{-1}\) rate using a sequence optimized for the use of a fast mixing device (FTA-SOFAST-HMQC) (13). The measurements were repeated twice for each protein (wild type and W60G mutant) to ensure reproducibility of the results and to evaluate their accuracy. Fig. 1 shows NMR spectra recorded before injection, immediately after injection, and once the refolding process was completed.

Surprisingly, the NMR data obtained for the wild type and W60G proteins showed similar features, especially the presence of two protein states characterized by distinct sets of peaks that can be ascribed to the native state (N) and to an intermediate state (I). Thus, in contrast to fluorescence data, our NMR data reveal the presence of a long-lived folding intermediate for both wild type and W60G \( \beta_{2m} \). The absence of unfolded state (U) peaks in the very first NMR spectra acquired after initiating the refolding process indicates that the conformational transition from U to I was completed within the dead time (~10 s) of the kinetic experiment.

For a more detailed quantitative analysis of the NMR data, we classified the observed NMR correlation peaks in three different classes. 1) The first class was pure native state peaks (N) for which the chemical shift values match those of native \( \beta_{2m} \) under equilibrium conditions. In this class, we included only the well isolated peaks with a volume ratio between the first and the last spectrum below 15%. These peaks report on the growth of the native state population over time. 2) The second class was pure intermediate state peaks (I) for which the chemical shift values differ from both the unfolded and the folded states. These peaks correspond to amide sites with a local chemical environment in the intermediate state that is different from the one in the native state. The decay of these peaks is related to the decrease of the intermediate species (I) over time. 3) The third class was peaks arising from overlap of intermediate and native species resonances (I+N) due to similar chemical environments of the corresponding amide groups in the I-state and the N-state. Intensity changes over time observed for this class of peaks provide useful information on the presence of additional folding intermediates and differences in local dynamics between the two states.

This classification enabled the identification of 27/27 (class N), 9/18 (class I), and 18/18 (class I+N) peaks in the spectra of wild type/W60G \( \beta_{2m} \), respectively. Examples of residue-specific kinetic folding curves obtained from the volume changes of class N, I, and I+N peaks over time are shown in supplemental Fig. 1. As expected, within each peak family, the same kinetic behavior is observed within the experimental error. Therefore, to increase the precision of the quantitative kinetic analysis, the volumes for all peaks of a given class, N, I, and (I+N), were added. The resulting kinetic traces are shown in Fig. 2 for wild type and W60G \( \beta_{2m} \). In the case of a simple two-state transition from I to N, one would expect a monoexponential behavior for both, the decrease of I-state, and the increase of N-state, and as a consequence, no change in the (N+I) peak volumes over time. Fitting reveals that a monoexponential curve does not account for the observed folding kinetics of the N-state, particularly in the case of the wild type protein. This observation is supported by the statistical Student’s t test and by the analysis of the residuals (supplemental Tables 1 and 2), according to which a biexponential model leads to a significant improvement in the fit quality and reliability. The statistical tests reject very high probability the redundancy of the biexponential with respect to the monoexponential model.

According to all the examined data sets for both wild type and W60G \( \beta_{2m} \), two slow folding phases are required to account for the formation of the native state (Fig. 2, a and b, and Table 1) with a first rate constant on the order of 10\(^{-3}\) s\(^{-1}\), \( k_2 \), and a second, slower on the order of 10\(^{-4}\) s\(^{-1}\), \( k_3 \). An important difference between the wild type and mutant proteins is observed with respect to the amplitudes of the two kinetic phases. For the W60G mutant, the slower phase only accounts for about 5%, whereas it is more pronounced in the wild type protein with values ranging from 16 to 25%, depending on the experimental data set used (Table 1). Instead, the decay of the observed intermediate species follows a monoexponential kinetics for both \( \beta_{2m} \) variants (Fig. 2, c and d, and Table 1). This decay is characterized, in the case of wild type \( \beta_{2m} \), by a slightly slower rate constant (\( k_i = (1.56 \pm 0.14) \times 10^{-3} s^{-1} \)) than for the mutant (\( k_i = (2.35 \pm 0.07) \times 10^{-3} s^{-1} \)). Interestingly, the decay rate of the intermediate species for wild type \( \beta_{2m} \) does not match any of the rates obtained from the biexponential fitting of the native species buildup, whereas for W60G \( \beta_{2m} \), the rates \( k_1 \) and \( k_2 \) are equal, within the statistical uncertainty.

The presence of two different kinetic constants suggests the existence of an additional intermediate species that does not give rise to detectable NMR signals. This conclusion is further supported by the kinetic traces obtained from the third class of peaks (I+N), shown in Fig. 2, e and f, for wild type and W60G \( \beta_{2m} \), respectively. For both proteins, the signal intensity of class I+N peaks, and consistently the sum of I-state and N-state populations, are not constant throughout the experimental observation interval. Again, this observation points toward the presence of some “invisible” species that converts into the native state together with the distinctly observed slow refolding intermediate I.
The data from class I+N peaks were fitted for the wild type protein to a biexponential equation, yielding rate constants of $(4.6 \pm 0.1) \cdot 10^{-3} \text{s}^{-1}$ and $(6.25 \pm 0.95) \cdot 10^{-4} \text{s}^{-1}$. These values are of the same order of magnitude as those extracted from the native state growth curve (class N peaks). For the W60G mutant protein, a monoexponential function satisfactorily accounts for the class I+N peak data, with a kinetic rate constant of $(1.6 \pm 0.2) \cdot 10^{-3} \text{s}^{-1}$. The contribution of the additional NMR-invis-
aggregation or internal mobility of the protein in the I-states and the N-states, resulting in different spin relaxation properties and thus in unequal signal intensities in the NMR spectra even for equal populations of the two states. To investigate whether differences in amide $^1$H and $^{15}$N relaxation between the N-state and the I-state are at the origin of the missing peak intensities, we evaluated line widths of typical cross peaks assigned to the N-state and the I-state. $^1$H and $^{15}$N line widths are identical within experimental error, thus weakening the differential relaxation property argument as a possible explanation for the signal gain observed for class I+N peaks as folding proceeds. In addition, refolding measurements performed with longer recycling delay (d1) give the same results as obtained with shorter d1 (see supplemental Fig. 2), which further rules out the differential relaxation rate argument.

The real-time two-dimensional NMR data provide clear evidence that the slow folding phase observed for $\beta2m$ is not a simple two-state process from an intermediate state accumulated during the first seconds of the folding process to the native state but that additional long-lived intermediate states are required to explain the observed folding kinetics. The burst phase amplitude, corresponding to the amount of protein that folds directly to the native state via a pathway that does not require any slow phase or long-lived intermediate, has been calculated by extrapolating to zero the fitting curve of the N-state volumes. This burst phase amplitude is about (5.6 ± 1.3)% for wild type $\beta2m$ and slightly higher (10.6 ± 3.4)% for the W60G mutant. These values are significantly lower than the 20% figure previously obtained by Kameda et al. (3) for wild type $\beta2m$ also from real-time NMR measurements, although under different experimental conditions.

Possible Kinetic Schemes for $\beta2m$ Folding—Different models have been proposed for $\beta2m$ refolding, ranging from simple linear two-step schemes (1, 3) to parallel models with two or three intermediate species (4, 7). As already pointed out above,
Kinetic parameters of β2m folding

For each class of species, the sum of the measured peak intensities as a function of time is fitted to the following set of equations, for wild type and W60G 2m respectively. Wild type, I(t) = k1 e−k2t; N(t) = −k2 e−k2t; [1 + N(t)] e−k3t. W60G, I(t) = k1 e−k2t; Nmono(t) = −k2 e−k2t; Nexp(t) = −k2 e−k2t; [1 + N(t)] e−k3t. The burst phase (bp) is calculated by extrapolating the value of N(t) and I(t) to t = 0.

| Species        | Wild type — first set | Wild type — second set | W60G — first set | W60G — second set |
|----------------|-----------------------|------------------------|------------------|--------------------|
|                | k1                    | k2                     | k3               | k4                 |
| I             | (1.42 ± 0.02) × 10^−3 | (1.85 ± 0.04) × 10^−3  | (2.4 ± 0.5) × 10^−3 | (4.5 ± 0.4) × 10^−3 |
| N             | (4.5 ± 0.4) × 10^−3   | (7.2 ± 1.6) × 10^−4    | 80.4 ± 2.0       | 15.3 ± 0.7         |
| I+N           | 25 ± 1.6              | 11 ± 1.3               | 64 ± 12          |

| Species        | A3                    | A4                   | A5               | bp     |
|----------------|-----------------------|----------------------|------------------|--------|
| Wild type — first set | 43 ± 0.3              | 84 ± 0.3             | 85 ± 2.4         |
| Wild type — second set | 6.8 ± 0.6             | 14 ± 0.6             | 66 ± 6           |

| Species        | A3                    | A4                   | A5               | bp     |
|----------------|-----------------------|----------------------|------------------|--------|
| W60G — first set | 84 ± 0.3              | 7.2 ± 2.3            | 85 ± 2.4         |
| W60G — second set | 15.7 ± 0.1            | 14.1 ± 2.1           | 87.4 ± 4.2       |

A simple parallel model with two intermediate species (Fig. 3a) was not viable. Models with a single intermediate species are not sufficient to explain the time-sequence of the kinetic parameters. A more complex model (Fig. 3b) involving two intermediates was evaluated successfully, leading to similar functional forms and spectra. The presented scheme is consistent with wild type and W60G data. A simpler parallel model with two intermediate species (Fig. 3a) was not viable. However, the presented scheme is consistent with wild type and W60G data. A more complex model (Fig. 3b) involving two intermediates was evaluated successfully, leading to similar functional forms and spectra. The presented scheme is consistent with wild type and W60G data.
rate constant of $2.35 \times 10^{-3} \text{s}^{-1}$, whereas $I^*$ represents the least populated (14 ± 5%) and NMR-silent intermediate and converts to N at a rate of $1.6 \times 10^{-3} \text{s}^{-1}$. A further (11 ± 3)% of unfolded W60G β2m transforms directly into N without passing through any of these intermediate species, conceivably because of an already attained cis configuration at the His-31–Pro-32 ω angle. Note that the five-state model A reduces to the simpler model B assumed for the mutant W60G if the step from $I^{**}$ to I is very fast.

**Structural Properties of the Intermediate Species**—The real-time NMR study reported here allowed us to identify at least two distinct long-lived intermediates of β2m folding, both for the wild type species and for the W60G mutant. These two intermediate states have completely different spectral signatures. Although the I-state is directly observed in the NMR spectra as a set of $^{1}H-^{15}N$ correlation peaks characteristic of a well defined globular conformation, the $I^*$ together with $I^{**}$ states do not give rise to any detectable NMR signal.

Even without the sequential resonance assignment of the I-state, some structural information is obtained from the analysis of its spectral pattern, observed at the beginning of the slow folding phase. In both β2m variants, some peaks of this intermediate species were found to overlap those of the native state (Fig. 4a), indicating that for those amides, the local chemical environment in the I-state is close to that in the N-state. The distribution of amide groups with similar chemical shifts in the I- and N-states clusters in precise locations of the molecule (Fig. 4b), at the opposite side of Pro-32, which has a cis peptide configuration in the native state but may have not yet adopted this configuration in the intermediate species. Indeed, the magnitude of the decay rate of the native-like intermediate is typical of proline peptide bond isomerization (25). The native-like intermediate is thus likely to possess a trans configuration at Pro-32, a structure similar to that of mutant P32A (a crystallographic dimer, PDB entry 2FBO (6)). Further work is in progress to determine the solution structure of the intermediate through fast multidimensional NMR techniques. All these observations are consistent with the previously reported results (3, 4) that stressed the key role of trans-cis isomerization as the rate-limiting step for the folding of β2m. It is interesting to note that the peak integrals of the native-like intermediate are lower in wild type species and for the W60G mutant. These two intermediate states do not give rise to any detectable NMR signal.

The structural nature of the additional, presumably native-unlike, intermediates is unknown. By NMR, we were only able to observe the native-like intermediate but not the other species. Two mechanisms can reasonably account for the substantial line broadening determining the NMR invisibility of the additional intermediate species: either a conformational equilibrium at unfavorable exchange rates ($\mu$s to ms range), plausibly within an ensemble of unfolded forms, or an extensive association/aggregation process of the same forms. Finally, because all the slow folding rates found in this work are compatible, as mentioned, with a proline peptide bond isomerization, it is reasonable to assume that a trans configuration for the ω angle preceding Pro-32 is also present in the additional intermediate(s).

**DISCUSSION**

The present work takes advantage of the recent advances in fast multidimensional NMR techniques to gain a more detailed view into the folding process of β2m at ambient temperature. A similar approach was recently applied to bovine α-lactalbumin (13), where a good agreement in the measured folding kinetics was found between NMR and fluorescence methods.

Instead, in this work, a surprising mismatch between NMR and fluorescence was found; in contrast to CD and fluorescence data (17), our NMR results reveal the presence of a slow phase for W60G β2m. Moreover, the presence of at least two folding slow phases for wild type β2m had neither been recognized before nor considered in any of the previously proposed folding kinetic schemes (1, 3, 4).

Kameda et al. (3) performed a similar analysis on wild type β2m using conventional HSQC NMR experiments at 2.8 °C, but they were unable to detect the multiexponential kinetics of native state formation we observed at ambient temperature. Apart from differences in the experimental temperature that could profoundly affect the folding mechanism, the lower time resolution of conventional NMR experiments in comparison with the SOFAST counterparts could also justify the different ability to go deeper into the fine details of the kinetic processes. A further difference between the present analysis and the results reported by Kameda et al. (3) regards the values of the burst phase amplitude. We found burst phase amplitudes around 5%, i.e. significantly below the previously reported value of 20% (3). Again, the differences in data collection and modeling could well be responsible for this discrepancy.

Close inspection of the NMR pattern observed for the I intermediate enabled recognizing a native-like structure in the region opposite to Pro-32, in agreement with previous findings (3, 4). On the contrary, the residues close to Pro-32, the facing N-terminal end, and the DE loop of the observable intermediate species could not be identified, mainly because this is the region where most structural differences are expected in comparison with the native state. An additional reason could be the extensive broadening of these resonances because of the unfavorable rates of conformational exchange, whose remnants are also present in the fully folded wild type protein, as revealed by $R_2/R_1$ relaxation data (17).

As detailed above, the biexponential growth of the native species and the apparent monoexponential decay of the intermediate species, characterized by three different apparent rates, rule out a model involving only two intermediate species and prompt us to propose a more complex mechanism. The proposed five-state kinetic scheme can account for the discrepancy between optical and NMR results observed upon comparing wild type and W60G β2m. In this model, three intermediate species are included, one, I, observable by NMR and two others, $I^*$ and $I^{**}$, corresponding to NMR-silent species. No data are available for $I^*$ and $I^{**}$, but the failure to observe them in the NMR spectra could be due to unfavorable conformational dynamics and/or oligomerization.
As experienced by W60G mutant, the parallel folding scheme (model B) predicts a preferential pathway, with I accumulating more than 5-fold with respect to I*. This is consistent with a virtually undetectable folding slow phase for W60G β2m in fluorescence spectra and demonstrates that the I intermediate, highly native-like, is fluorometrically indistinguishable from the native state. In other words, fluorescence does not discriminate between natively folded species and native-like intermediates of β2m. Hence, the slow phase observed by fluorescence must be a native-unlike, NMR-silent species. In wild type β2m, a slow folding phase is detected by fluorescence because the additional long-lived I* and I** intermediates that accumulate to a higher extent must be structurally different from the N-state in the vicinity of residue Trp-95.

Therefore, although the NMR undetectability of I* and I** gives some indications about their oligomeric and dynamic properties, the fluorescence response suggests that I* and I** are intermediate states that have not yet reached a native-like packing of the hydrophobic core. The higher population of the native and native-like conformations, N and I, for W60G, in comparison with wild type β2m, correlates with the higher thermodynamic stability of the mutant protein (17).

The presence of folding intermediates and the fibrillogenesis propensity of a protein is regarded as an important correlation when studying amyloidosis (26, 27). Along this view, Chiti et al. (2) proposed that I2, a folding intermediate associated with the slow process, is involved in the amyloid deposition, on the basis of the enhanced seed elongation during the folding; later, this idea was recovered by other groups (3, 4, 24). Controversial cases are represented, however, by the different amyloidogenic propensity of P32G and P32V β2m (4, 7) or by the inability of the CL antibody to form fibrils (8).

In this work, we show that the correlation between folding and fibrillogenesis for β2m is more complex than previously
believed. In fact, we demonstrate that more than a single intermediate species is transiently populated during the folding slow process of β2m and that not all of them correlate with fibrillogenic capability. In particular, the inability of W60G β2m to elongate fibrils in 20% trifluoroethanol (17) appears coupled with the occurrence of a predominant native-like intermediate species and with a very reduced occurrence of native-unlike intermediate(s). The latter species transiently populates during the folding slow phase of wild type β2m that, in fact, is capable of elongating fibrils in vitro. Because both native-like and native-unlike species seem related to trans-cis isomerization of the peptide bond preceding Pro-32, our findings, although confirming the crucial role of that peptide bond conversion in the β2m folding, cast some doubts on previous claims for amyloidogenic capability. In particular, the inability of W60G β2m to elongate fibrils in vitro of elongating fibrils 2m that, in fact, is capable for the onset of the pathology. On the basis of the data in our hands, however, it is not possible to unequivocally identify the NMR-silent species as the conformer(s) responsible in vivo for the onset of the pathology. Future work will focus on this problem but also on a more detailed characterization of all the species involved in the folding slow phase, also from a structural point of view.

REFERENCES

1. Chiti, F., Mangione, P., Andreola, A., Giorgetti, S., Stefani, M., Dobson, C. M., Bellotti, V., and Taddei, N. (2001) J. Mol. Biol. 307, 379–391
2. Chiti, F., De Lorenzi, E., Grossi, S., Mangione, P., Giorgetti, S., Caccialanza, G., Dobson, C. M., Merlino, G., Ramponi, G., and Bellotti, V. (2001) J. Biol. Chem. 276, 46714–46721
3. Kameda, A., Hoshino, M., Higurashi, T., Takahashi, S., Naiki, H., and Goto, Y. (2004) J. Biol. Chem. 279, 9176–9189
4. Eakin, C. M., Berman, A. J., and Miranker, A. D. (2006) Nat. Struct. Mol. Biol. 13, 202–208
5. Sakata, M., Chatani, E., Kameda, A., Sakurai, K., Naiki, H., and Goto, Y. (2008) J. Mol. Biol. 382, 1242–1255
6. Feige, M. J., Grosse, S., Marcinowski, M., Yew, Z. T., Truffault, V., Paci, E., Kessler, H., and Buchner, J. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 13373–13378
7. Freeman, R., and Kuce, E. (2003) J. Biomol. NMR 27, 101–113
8. Malmodin, D., and Billeter, M. (2005) Prog. Nucl. Magn. Reson. Spectrosc. 46, 109–129
9. Schanda, P., and Brutscher, B. (2005) J. Am. Chem. Soc. 127, 8014–8015
10. Schanda, P., Van Meleckebeke, H., and Brutscher, B. (2006) J. Am. Chem. Soc. 128, 9042–9043
11. Schanda, P., Forge, V., and Brutscher, B. (2007) Proc. Natl. Acad. Sci. U.S.A. 104, 11257–11262
12. Gal, M., Schanda, P., Brutscher, B., and Friedman, L. (2007) J. Am. Chem. Soc. 129, 1372–1377
13. Gal, M., Kern, T., Schanda, P., Friedman, L., and Brutscher, B. (2009) J. Biomol. NMR 43, 1–10
14. Fogolari, F., Corazza, A., Viglino, P., Zuccato, P., Pieri, L., Faccioli, P., Bellotti, V., and Esposito, G. (2007) Biophys. J. 92, 1673–1681
15. Esposito, G., Ricagno, S., Corazza, A., Rennella, E., Gümr, D., Mimmi, M. C., Betto, E., Pucillo, C. E., Fogolari, F., Viglino, P., Raimondi, S., Giorgetti, S., Bolognesi, B., Merlini, G., Stoppani, M., Bolognesi, M., and Bellotti, V. (2008) J. Mol. Biol. 378, 887–897
16. Esposito, G., Michelutti, R., Verdone, G., Viglino, P., Hernández, H., Robinson, C. V., Amoresano, A., Dal Piaz, F., Monti, M., Pucci, P., Mangione, P., Stoppani, M., Merlini, G., Ferri, G., and Bellotti, V. (2000) Protein Sci. 9, 831–845
17. Kupce, E., and Freeman, R. (1995) J. Magn. Reson. A 115, 273–276
18. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) J. Magn. Reson. 48, 286–292
19. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
20. Blevins, R. A., and Johnson, B. A. (1994) J. Biomol. NMR 4, 603–614
21. Okon, M., Bray, P., and Vucelić, D. (1992) Biochemistry 31, 8906–8915
22. Kihara, M., Chatani, E., Iwata, K., Yamamoto, K., Matsuura, T., Nakagawa, A., Naiki, H., and Goto, Y. (2006) J. Magn. Reson. A 273–276
23. Okon, M., Bray, P., and Vucelić, D. (1992) Biochemistry 31, 8906–8915
24. Kihara, M., Chatani, E., Iwata, K., Yamamoto, K., Matsuura, T., Nakagawa, A., Naiki, H., and Goto, Y. (2006) J. Biol. Chem. 281, 31061–31069
25. Cheng, H. N., and Bovey, F. A. (1977) Biopolymers 16, 1465–1472
26. Chiti, F., and Dobson, C. M. (2006) Annu. Rev. Biochem. 75, 333–366
27. Chiti, F., and Dobson, C. M. (2009) Nat. Chem. Biol. 5, 15–22

Two-dimensional NMR Reveals the Folding Pathway of β2m