Properties of Some Variants of Human $\beta_2$-Microglobulin and Amyloidogenesis*^[S]

Alessandra Corazza‡, Fabio Pettirossi‡, Paolo Viglino‡, Giuliana Verdone‡, Julian Garcia‡, Pascal Dumy§, Sofia Giorgetti¶, Palma Mangione∥, Sara Raimondi∥∥, Monica Stoppini∥∥∥, Vittorio Bellotti∥∥∥∥, and Gennaro Esposito‡∗

From the ‡Dipartimento di Scienze e Tecnologie Biomediche and Microgravity, Aging, Training, and Mobility, Centre of Excellence, Università di Udine, Piazzale Kolbe 4, 33100 Udine, Italy, §Laboratoire d’Études Dynamiques et Structurelles de la Sélectivité, University Joseph Fourier of Grenoble, 38041 Grenoble Cedex 9, France, the ¶Dipartimento di Biochimica via Taramelli 3/b Università di Pavia, 27100 Pavia, Italy, and the ∥Laboratorio di Biotecnologie Istituto di Ricerca e Cura a Carattere Scientifico, Policlinico S. Matteo, 27100 Pavia, Italy

Three variants of human $\beta_2$-microglobulin ($\beta_2$-m) were compared with wild-type protein. For two variants, namely the mutant R3A$\beta_2$-m and the form devoid of the N-terminal tripeptide ($\Delta N3\beta_2$-m), a reduced unfolding free energy was measured compared with wild-type $\beta_2$-m, whereas an increased stability was observed for the mutant H31Y$\beta_2$-m. The solution structure could be determined by $^1$H NMR spectroscopy and restrained modeling only for R3A$\beta_2$-m that showed the same conformation as the parent species, except for deviations at the interstrand loops. Analogous conclusions were reached for H31Y$\beta_2$-m and $\Delta N3\beta_2$-m. Precipitation and unfolding were observed over time periods shorter than 4–6 weeks with all the variants and, sometimes, with wild-type protein. The rate of structured protein loss from solution as a result of precipitation and unfolding always showed pseudo-zeroth order kinetics. This and the failure to observe an unfolded species without precipitation suggest that a nucleated conformational conversion scheme should apply for $\beta_2$-m fibrillogenesis. The mechanism is consistent with the previous and present results on $\beta_2$-m amyloidogenesis, provided a nucleated oligomeric species be considered the stable intermediate of fibrillogenesis, the monomeric intermediate being the necessary transition step along the pathway from the native protein to the nucleated oligomer.

Over the last several years, an overwhelming number of reports have addressed the phenomenon of amyloidogenesis. The interest in the subject stems not only from the social relevance of amyloidopathies such as Alzheimer’s disease, or spondyloarthropathy, or the various systemic amyloidoses, but also from the general implications in the issue of protein folding. Amyloidoses have been recognized as conformational diseases that arise from the conversion of globular proteins into insoluble fibrillar aggregates (1). Despite the diversity of the involved proteins, amyloid fibrils exhibit a common structure known as cross-$\beta$ structure, which appears to be a particularly stable, generic protein fold, accessible to many polypeptide chains under specific conditions in vitro and in vivo (2). The amyloid deposition of $\beta_2$-microglobulin ($\beta_2$-m), the nonpolymeric light chain of the class I major histocompatibility complex (MHC-I), is associated to dialysis-related amyloidosis (3). The disease is the result of long term hemodialysis in individuals with chronic renal failure, a widespread pathology with high social costs that are further increased by the inevitable dialysis-related amyloidosis complication. Recently ankylosing spondylitis has also been proposed to originate from $\beta_2$-m deposition (4). We determined the solution structure of isolated $\beta_2$-m by NMR spectroscopy (5) and showed that the most important rearrangements of the protein, with respect to its structure in MHC-I, were observed for strands D and E, interstrand loop D-E, and strand A, including the N-terminal segment. We stated that these modifications can be considered as the prodromes of the amyloid transition that starts at sheet 1 with the rupture of strand A pairing, and leads to polymerization, through intermolecular pairing at strand D and probably strand C, and precipitation into fibrils, i.e. according to the scheme we had proposed earlier from a comparative investigation on the full-length protein and the form devoid of the six N-terminal residues ($\Delta N6\beta_2$-m) (6). This latter species, which accounts for $\sim$30% of $\beta_2$-m content in ex vivo fibrils, had been shown to possess a higher tendency to self-aggregate than the full-length protein and to fail reaching a fully folded native state at the end of the refolding procedure (7). Successive works reporting $^1$H,$^15$N NMR studies of pH and urea-driven denaturation experiments (8) and deuterium exchange of backbone amides (9) confirmed the amyloid pairing pattern of $\beta_2$-m that was inferred from our former results (6). Moreover, in analogy with the solution structure, even the crystal structure of isolated human $\beta_2$-m has been claimed to reveal the clues to its amyloidogenic properties (10). However, the revelations that

---

*S This work was supported by grants from the Italian Ministry of Education (Fondo per gli Investimenti della Ricerca di Base, Cofin), the Italian Ministry of Health (Progetto Finalizzato Alzheimmer 2000), and the Dipartimento di Scienze e Tecnologie Biomediche - University of Udine. 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.

[S] The on-line version of this article (available at http://www.jbc.org) contains Figures S1 and S2.

1 The abbreviations used are: $\beta_2$-m, $\beta_2$-microglobulin; $\Delta N3\beta_2$-m, $\beta_2$-microglobulin 4–9 fragment; $\Delta N6\beta_2$-m, $\beta_2$-microglobulin 7–99 fragment; GdmCl, guanidinium chloride; MD, molecular dynamics; MHC-I, class I major histocompatibility complex; NCC, nucleated conformational conversion; NOESY, nuclear Overhauser spectroscopy; ppm, parts per million; r.m.s.d., root mean square deviation; TOCSY, total correlation spectroscopy; UHBD, University of Houston Brownian Dynamics; HLA, human leukocyte antigen.
are disclosed by the crystal structure appear quite different from the elements of the solution structure that were recognized to forerun the amyloid transformation of \( \beta \)-m and hence considered as prodromes of the process (5). To confirm and reinforce our model, we present here the results of the characterization by classical biochemical methods and \(^1\)H NMR of three variants of human \( \beta \)-m, namely the truncated form \( \Delta N 3 \beta \)-m and the mutants R3\( \alpha \)A\( \beta \)-m and H31Y\( \beta \)-m. These species were designed to test the role of some structural elements in the stability of the wild-type protein in solution (5, 6). The truncation from corresponds to the wild-type protein devoid of the three N-terminal residues. In the mutants, the wild-type sequence residues in position 3 and 31, i.e., an arginyl and a histidyl residue, respectively, are replaced with an alanyl and a tyrosyl residue, respectively.

**EXPERIMENTAL PROCEDURES**

**Preparation of Recombinant Protein Samples**—Expression and purification of recombinant proteins was carried out as previously reported for \( \beta \)-m and \( \Delta N 3 \beta \)-m (6). CDNA corresponding to the truncated species of \( \beta \)-m lacking the first three (\( \Delta N 3 \beta \))-m residues was accomplished by PCR amplification of the full length CDNA with the primers 5\'-GCACA CATACATCCAAAGATTCCAGG and 3\'-GCTATGTGCTGGGTTTTCTCCATCCGACAT TGAAG and 3\'-CCATGTGCTGGGTTTTCTCCATCCGACAT TGAAG and 3\'-CCCATGTGCTGGGTTTTCTCCATCCGACATAGC for R3\( \alpha \), 5\'-GCACAATGATCCAGGCTACTCCAAAGATTCAGG and 3\'-CGCACATATGACTCCAAAGATTCCAGG for H31S, 5\'-CTCTGAATGTGTCTGGGTTTTCTCCATCCGACAT TGAAG and 3\'-CCCTGAAATTTTTGAGTACCTGTTGAGT for R3\( \alpha \)A\( \beta \)-m. Mutagenesis of His-31 into Tyr and Ser, and of Arg-3 into Ala were performed by using the QuikChange™ site-directed mutagenesis kit supplied by Stratagene (La Jolla, CA). The following primers were used: for H31Y, 5\'-GCATGGTGTGGTGGTTTTCTCCATCCGACAT TGAAG and 3\'-CTCTGAATGTGTCTGGGTTTTCTCCATCCGACATAGC for R3\( \alpha \), 5\'-GCACAATGATCCAGGCTACTCCAAAGATTCAGG and 3\'-CGCACATATGACTCCAAAGATTCCAGG for H31S, 5\'-CTCTGAATGTGTCTGGGTTTTCTCCATCCGACAT TGAAG and 3\'-CCCTGAAATTTTTGAGTACCTGTTGAGT. The constructs were always introduced in the same BL21 Escherichia coli strain. A methionine residue was present at the N-terminal position of all recombinant products that will be removed to Met-0. All the recombinant proteins gave a single species when assayed by electrospray-ionization mass spectrometry with 20 μM samples. The concentration of the protein samples was determined spectrophotometrically at 280 nm: \( A \)\( _{280} \) = 16.17 (\( \beta \)-m), 16.29 (R3\( \alpha \)A\( \beta \)-m), 16.24 (H31\( \beta \)-m), 17.21 (H31Y\( \beta \)-m), 16.73 (\( \Delta N 3 \beta \)-m), and 17.22 (\( \Delta N 6 \beta \)-m).

**Determination of Folding Stability**—The thermodynamic stability was determined by the measurement of intrinsic tryptophan fluorescence on guanidinium chloride (GdmCl) concentration. Excitation was at 295 nm and emission monitored in the range 300–500 nm (11) using slit widths of 5 nm. The protein concentration of all samples was 0.5 mg/ml sodium phosphate buffer (70 mM phosphate, 100 mM NaCl, pH 6.7 (pH* is the pH meter reading without isotope effect correction)). When necessary, samples were centrifuged and/or filtered using filters with a threshold of 0.22 μm (Millipore, Bedford, MA). The studies were carried out at 310 K, with only a few experiments obtained at 298, 303, or 307 K with H31Y\( \beta \)-m for assignment controls. A number of two-dimensional NOESY spectra (17) and NOESY (18) spectra were acquired. The adopted experimental schemes included solvent suppression by presaturation, or WATER-GATE (19) and excitation sculpting (20), 1–5-s steady state recovery time, mixing times (\( t _{m} \)) of 15–30 ms for TOCSY and 50–150 ms for NOEY, \( t _{1} \), quadrature detection by time proportional phase incrementation (21), or States method (22), or gradient-assisted coherence selection (echo-antiecho) (23). The spin-lock mixing of the TOCSY experiment was obtained with MLEV17 (24) or DIPSI-2 (25) pulse trains at \( \gamma B / 2 \pi = 7–10 \) kHz. The acquisitions were performed over a spectral width of 7002.801 Hz in both dimensions (or, sometimes, 8012.820 Hz), with matrix size of 1024 × 1024 points each. FIDs, FTs, and water suppression were performed. For the observation of very rapidly relaxing spin systems, the acquisition time was reduced in both dimensions by collecting data matrices of smaller size (512–1024 points in \( t _{1} \) and 128–256 points in \( t _{2} \)). Precipitation was followed by collecting series of one-dimensional spectra, and sometimes also two-dimensional TOCSY, at regular time intervals after periods of 7–20 days. Acquisition conditions for two-dimensional data were the same as above reported, whereas for three-dimensional time domain were set at 4096–8192 points, with relaxation delays of 1.1 or 10 s, depending on whether a steady or an equilibrium state integral was needed. Refolding of the wild-type protein was also followed by one-dimensional NMR measurements. After incubating at ambient temperature, for approximately 1 h, 1 mg of protein in 100 μl of 3 M GdmCl was transferred into the NMR tube and diluted 20-fold with the renaturing buffer (70 μM protein, 100 mM NaCl, in D<sub>2</sub>O, pH* = 7.5). The sample was then put into the magnet and, after quick equilibration and spectrometer adjustments, one-dimensional spectra were collected at regular intervals (1.5 min). At 280 K, with excitation sculpting (20) to suppress the residual HOD peak, 2048 time-domain data points, 4 scans/experiment, and 8 dummy scans. The first acquisition was completed at 6.5 min from dilution, and, overall, the refolding kinetics was monitored by this real-time NMR approach for approximately 90 min. The achievement of a steady state was checked by recording spectra at higher intervals (>24 h) on the sample left in the magnet at the same instrumental conditions. pH Determination using Felix (Accelrys Inc., San Diego, CA) software with shifted (60–90°) square sinebell apodization, zero filling (up to 2048 × 2048 real points) and polynomial base-line correction for NOESY data. All \( \Delta N 3 \beta \)-m and R3\( \alpha \)A\( \beta \)-m spectra were referenced on the Leu-23 \( \text{CH}_3 \) resonance at 0.58 ppm as previously reported (6). For H31Y\( \beta \)-m spectra were referenced using the Leu-23 \( \text{CH}_3 \) resonance peak (0.568 ppm at 310 K), whose chemical shift was determined at varying temperature with respect to internal dioxygen (always at 3.750 ppm). Interaxial distances were quantified only for R3\( \alpha \)A\( \beta \)-m from the NOEY cross peak volumes of Trp-95, as previously reported (5). Tethered molecular dynamics (MD) simulations were performed by using the Discover software (Accelrys Inc., San Diego, CA) and following protocol for simulated annealing (8.6-ps MD, 50 conjugate-gradient minimization steps) (26). Molecular structures were generated with InsightII (Accelrys Inc., San Diego, CA), MOLMOL (27), and Swiss-PdbViewer (www.expasy.ch/spdbv/mainpage.html).

**pK<sub>a</sub> and Free Energy Calculations**—The electrostatic contributions to the energy were computed using the UHBD (University of Houston Brownian Dynamics) program, which is capable of solving the linearized Poisson-Boltzmann equation using a finite-difference method (28) in a continuum solvent model. The \( pK \) values for the ionizable residues of the protein and the related electrostatic free energy contributions (\( \Delta G_e \)) were calculated according the methodology of Antosiewicz (29–32) and referenced to the free energy change (\( \Delta G_m \)) calculated according to Eisenberg et al. (31) with a routine implemented in the program Tinker (J. W. Ponder, Washington University School of Medicine, Saint Louis, MO). The method is semipirical and calculates, for every atom i, the free energy of transfer, \( \Delta G_i \).
from the environment of the protein interior to solution, according to $\Delta G_i = \Delta G_{1/2}$, where $\Delta G_i$ is the atomic solvent parameter (31) and $A_i$ is the accessible surface area of the atom. The total $\Delta G_{1/2}$ for the protein is obtained by summing over all protein atoms.

RESULTS

GdmCl Denaturation—The denaturation of wild-type and variant $\beta_2$-m species was measured from the changes in fluorescence upon addition of GdmCl, and the results are presented in Table I and Fig. 1. $\Delta N6\beta_2$-m appears to be the least stable species followed by $\Delta N3\beta_2$-m and R3A$\beta_2$-m, which show intermediate $\Delta G_{1/2}$ and $C_m$ values between wild type and $\Delta N6\beta_2$-m. The alanyl substitution for arginyl in position 3 destabilizes the variant by 0.4 kcal mol$^{-1}$, whereas the truncation at the same location destabilizes by 1.9 kcal mol$^{-1}$. On the other hand, the presence of a tyrosyl residue in position 31 stabilizes $\beta_2$-m by 1.5 kcal mol$^{-1}$, in qualitative agreement with the result obtained for the $\beta_2$-m His-31 → Phe mutant (2.6 kcal mol$^{-1}$ stabilization under different conditions (Ref. 32)), whereas the introduction of a serine in the same position results in a reduced stabilization (6.9 kcal mol$^{-1}$).

Unfolding/Refolding Kinetics—In wild-type $\beta_2$-m, the refolding pathway is characterized by three phases, the slowest of which has a rate constant of $0.0032 \pm 0.0005 \text{ s}^{-1}$ (14). This slow refolding phase of $\beta_2$-m was also investigated by real-time NMR. The spectra collected within the first 30 min from dilution with phosphate buffer of the GdmCl solution exhibited progressive changes toward the typical resonance pattern of the folded protein. A few resolved peaks appear to evolve in the upfield and aromatic regions (Fig. 2). In particular, the Ile-35 δ-methyl signal and the upfield-shifted H$^\alpha$ resonance of Leu-23 slowly grow up to the final intensity, whereas the γ-methyls of Val-37 seem to increase at the expense of another pair of signals that, most likely, arise from the methyls of the same residue in a different geometry within a transient conformer of the molecule. Simultaneous changes observed in the aromatic region involve the ring resonances of Tyr-66, Phe-30 and Phe-70, and Trp-95, i.e. the aromatic groups that are largely responsible for the upfield shifts of Val-37, Ile-35, and Leu-23, respectively. Overall, the NMR pattern suggests that a rearrangement of the hydrophobic core takes place throughout the slow refolding phase of $\beta_2$-m, consistently with the corresponding increase of 8-anilino-1-naphthalene sulfonic acid binding and changes in fluorescence and circular dichroism pattern, previously observed during the same refolding phase (14). Based on the NMR data, the rearrangement process involves a transient conformer that slowly converts into the stable, native form of the protein, with no evidence in favor of an equilibrium between the two species (15). The persistence of some 14 ± 8% of total protein in the refolding intermediate conformation (L2) (15) could not be confirmed, although the occurrence of smaller amounts of the same species, i.e. below 2–3%, cannot be ruled out in principle because such small quantities would be hardly detectable by NMR. Likewise, we cannot exclude possible equilibria involving spectroscopically silent species (14), which for NMR would be either species with unresolved signals buried under the main envelope, or large aggregates with very broad linewidths, in any case other than the slowly converting form observed during refolding.

As for the mutants, the increased folding stability of H31Y$\beta_2$-m and H31S$\beta_2$-m prompted us to investigate the effect of the mutation on unfolding/refolding kinetics. Although the H31Y$\beta_2$-m very fast and fast folding kinetics phases are comparable with those of wild-type $\beta_2$-m, a remarkable difference was appreciated for the slow phase that displays a rate constant of 0.031 s$^{-1}$ for H31Y$\beta_2$-m (Fig. 3), i.e. 10 times faster than the value reported for the wild-type species (14). Preliminary results show that the slow folding phase of the H31S mutant exhibits a rate constant of 0.0043 ± 0.0006 very close to that of the wild type protein. A lower unfolding rate constant for H31Y$\beta_2$-m ($6.9 \pm 1.2 \times 10^{-5} \text{ s}^{-1}$) with respect to the wild-type protein ($4.9 \pm 1.8 \times 10^{-4} \text{ s}^{-1}$) is in perfect agreement with the increased stability of the mutant at equilibrium.

NMR Characterization—The detailed interpretation of the $^1$H NMR spectra of the mutants and fragments of $\beta_2$-m, based on the parent spectrum of the wild-type species, is far from being trivial. As already observed with $\Delta N6\beta_2$-m (6), the task requires checking and redetermining most of the resonance assignments of the variant under investigation according to the standard methodology, i.e. going through scalar and dipolar connectivity pattern for each amino acid residue (33). This work was performed entirely for the truncated fragment $\Delta N3\beta_2$-m and the two mutants R3A$\beta_2$-m and H31Y$\beta_2$-m. The corresponding assignment lists are available at the BioMagRes data bank (entry points 5782, 5783, and 5784). The relevant differences of H$^\alpha$ chemical shifts ($\Delta \delta$H$^\alpha$) with respect to the corresponding values of wild-type protein are shown in Fig. 4. For the sake of comparison, the $\Delta N3\beta_2$-m and R3A$\beta_2$-m data (Fig. 4A) are not plotted in the same panel as those of H31Y$\beta_2$-m (Fig. 4B) because of the gaps in the assignment list of the latter mutant. The backbone resonances within fragments 29–34, 57–58, and 60–62 of H31Y$\beta_2$-m could not be observed in the NMR spectrum (either totally or partially for residues 58 and 62) because of exchange broadening. For the other two variants instead, the missing resonances (as a matter
of fact only a few HN peaks) were sparsely distributed along the sequence and the trends of their deviations appeared similar. Indeed, Fig. 4 shows that the deviations from wild-type HN-N3-2-m and R3A-2-m. However, the deviations are remarkably larger in the loop region between strands D and E and particularly for Trp-60. Similar results were obtained also for the HN chemical shift deviations of the same molecules (data not shown). The overall deviation extent appears significantly reduced with the corresponding data of H31Y-2-m, i.e. less than ±0.05 ppm for most of the resonances, but larger differences could conceivably occur in segments 29–34 (loop B-C) and 57–62 (loop D-E), which were not detected. The failure to observe most of the backbone resonances in these segments was ascribed to chemical exchange. Difficult detectability of some NMR signals in the same regions had been experienced with H31Y-2-m and R3A-2-m and, previously, also with wild-type HN-2-m, in particular for a few residues of the interstrand loops B-C and D-E. For this latter loop, the problems arose from a conformational equilibrium leading to exchange broadening of the backbone resonances in segment 56–62, as first recognized by Okon and colleagues (34). In the interstrand loop B-C, instead, the resonances of Pro-32, which adopts a cis peptide bond geometry also in solution, were hardly observable probably as a result of the critical density of the local dipolar network, in the absence of any evidence in favor of an additional, local exchange process. Under proper conditions, the adverse consequences on the detection sensitivity of either exchange and dipolar broadening could be overcome with wild-type HN-2-m, as well as for H31Y-2-m and R3A-2-m (except for a few HN resonances). Analogue experimental conditions, however, failed to restore detectability with H31Y-2-m, leaving unassigned the B-C and D-E loop segments. As a matter of fact, the latter regions and the previously mentioned N-terminal and F-G loop fragments are all spatially positioned at one edge of the HN-2-m molecule and therefore relatively close in space. Therefore, besides the dipolar crowding effects, the broadening that hinders observation of some backbone resonances in loop B-C could be the additional result of the exchange occurring at the adjacent D-E loop. The question arises as to why the exchange rate should be so unfavorable as to bleach out several backbone resonances in H31Y-2-m, contrary to wild-type protein and the other studied variants. Such an effect could be related to the increased fold
stability of H31Yβ2-m, resulting in unfavorably slow conformational dynamics of the D-E loop for NMR detection. Besides the missing resonances in H31Yβ2-m, the same area exhibits also the largest δH values measured for the other two molecules (Fig. 4A). Such a result highlights the fine sensitivity of 1H chemical shifts in reporting local conformational deviations. 1H NMR chemical shifts show an established correlation with secondary structure only for Hβ1 values (35). None of the δH values of Fig. 4 alter the secondary structure profile inferred from the NMR chemical shift deviations of wild-type H92522-m with respect to the limiting values of statistically disordered conformations. In other words, the addressed variants adopt the same secondary structure as the parent sequence. This conclusion is felt to hold true also for the unassigned segments of H31Yβ2-m, after considering the substantial persistence of the tertiary structure in the whole molecule, as inferred from the very small variations of the backbone hydrogen chemical shifts with respect to wild-type H92522-m in all the remaining (assigned) regions.

NMR-based Modeling—NMR-restrained modeling was not possible for either the truncated form, ΔN3β2-m, or the H31Y mutant. A peculiarity of the truncated form was the occurrence in solution of a second minor species in equilibrium with the main component (ratio ≈ 1:4), as suggested by the presence of a second resonance for 28 residues: 7–10, 22, 25, 26, 37, 38, 49, 50, 55, 60, 66–69, 78, 79, 81–85, 88, 89, 91, 93, and 95–97. The vast majority of the residues that revealed resonance splitting belong to β-strands and are located in the central part of the molecule (Fig. 5). The only exception was Trp-60 that is in the loop D-E. Unfortunately, the extent of the splittings (often limited, see Fig. S1 in the Supplemental Material) and the degeneracy of most resonances made impossible defining a precise structural distinction between the two forms that were responsible for signal doubling pattern. Because of the presence of two different conformations, a proper quantitative analysis of the NOESY data was not viable, and therefore only qualitative inference could be drawn from the evidence on ΔN3β2-m.

No NMR-based restrained modeling was attempted for H31Yβ2-m either, in this case because of the lack of any data from the B-C and D-E loop regions. The NOESY pattern of the mutant was largely superimposable to the corresponding data of wild-type β2-m structure for the assigned resonances, which would have led to close conformational restraining, whereas no assessment at all was possible for the two mentioned loop regions where deviations could conceivably have emerged. Hence there was no point in running restrained modeling without any information from those loop regions. NMR restraints were instead collected and employed with R3Aβ2-m. The structural inference from the substantial similarity between wild-type and mutant chemical shifts for both HN (data not shown) and Hα (Fig. 4A) led us to use a tethered modeling approach in

![Fig. 4](image-url) The assigned backbone Hα chemical shifts of the three investigated β2-m variants are compared with the corresponding values of wild-type species obtained under similar conditions as those adopted in the present study (5). The ΔHα values (ppm) of R3Aβ2-m (A, black bars), ΔN3β2-m (A, red bars), and H31Yβ2-m (B) are reported as (δvariant − δwt).

![Fig. 5](image-url) Ribbon model of ΔN3β2-m backbone, based on wild-type β2-m structure in water (5) according to the corresponding limited δHα values (Fig. 4A). The yellow tracts highlight the residues of the truncated variant for which resonance doubling was observed in the NMR spectrum. The plot was drawn using MOLMOL (27).
which all atoms within fragments 8–26, 38–52, 65–81, and 92–99 were kept fixed while restraining the remaining ones. With a starting point chosen from the 20 best conformers that were determined for wild-type β²-m, the tethered MD run used 280 experimental distance restraints (129 intraresidue and 151 inter-residue), for residues 0–7, 27–37, 53–64, and 82–91. The restraints came from data that had been collected within 36 h from sample preparation. It is worth noting that, for the same region of the wild-type protein, 485 restraints could be employed. Despite this restraint number reduction, the resulting new family of R3Aβ²-m conformers showed a pairwise global root mean square deviation (r.m.s.d.) for the backbone atoms of 1.50 ± 0.23 Å. The analysis of the spreading within the family revealed that, besides the N-terminal region (0–7), the loop between strands D and E was the most dispersed region, followed by the loop between strands F and G (Fig. 6).

The spreading extents, however, were distinctly larger than the corresponding ones observed within the structure family of wild-type β²-m in solution (Fig. 7). A comparison between the mean structures from the 20 best-fit conformer families of wild type and R3A mutant showed a displacement maximum again in the loop D–E (r.m.s.d. 2.41 Å), with substantial deviations in the N-terminal region (r.m.s.d. 1.19 Å), loop F–G (r.m.s.d. 1.19 Å), and loop B–C (r.m.s.d. 0.49 Å). The smaller number of experimental restraints with respect to wild-type β²-m was certainly determining for the latter outcomes (conformational dispersion and displacements), but reflected inherent definition limits of R3Aβ²-m conformational distribution rather than defective restraining, as suggested by comparison with the wild-type protein nuclear Overhauser effect data set.

Precipitation Studies—A general feature of β²-m solutions is the limited stability leading to precipitation of amorphous or fibrillar material over variable time periods. Protein precipitation depends on several factors, including pH, chaotropes, temperature, organic cosolvents, etc., that would deserve each one a specific analysis in the particular and unique context of amyloidogenic proteins. For β²-m precipitation, sets of data obtained under different experimental conditions are available (36–38). Further systematic exploration is still necessary, however, because no comprehensive rationale has yet emerged. Moreover, a general caveat concerns the careful control of the initial protein preparation that may affect the experimental reproducibility of the precipitation studies. In our hands, for instance, sometimes different batches of wild-type β²-m showed, under similar conditions, a remarkable diversity of precipitation pattern. Precipitation started within a few days when the initial solution was not clear, as a result of poorly soluble particles that catalyzed the process acting as seeds. By repeating the refolding procedure, the protein could be recovered and gave clear solutions without any early precipitation problem. In the present study, the typical conditions that support stability of β²-m solutions for 1 month at least (T = 310 K, pH = 6.6, 100 mM NaCl, 70 mM phosphate buffer, 0.7–1 mM protein concentration) were tested with the three variants under investigation. For the truncated species, a clear tendency to precipitate was observed only with the freshly prepared, unfiltered solution. Centrifugation and filtering gave a stable sample, i.e. without any visible precipitate over a long time period (~1 month). An analogous filtration-dependent precipitation pattern was observed with H31Yβ²-m. Even after initial centrifugation/filtration, however, a precipitate was formed, and the spectrum of H31Yβ²-m soluble species appeared to deviate, over a 2-week time period, from the initial pattern, i.e. that typical of the structured protein. This time span was approximately halved without any initial treatment of the sample. Because the stability of any β²-m variant solution is crucially dependent on the presence of insoluble protein material, the earlier precipitation observed with H31Yβ²-m solutions compared with wild-type protein, despite the increased folding stability of the mutant, suggests that the preparation of H31Yβ²-m is quite critical. Probably formation of poorly soluble adducts becomes more prominent with the slight hydrophobicity increase of the His-to-Tyr mutation.

Unlike the truncated variant and H31Y mutant, the solutions of R3Aβ²-m were observed to last more briefly than those of wild-type protein at similar concentration. In particular, R3Aβ²-m dissolved very quickly to give clear solutions, but...
after 24–36 h (lag phase) started to show floating particles. This behavior was qualitatively independent of concentration and filtration, performed even though the starting solution appeared clear. A 5–6-fold reduction of R3Aβ2-m concentration (0.15–0.2 mM) led, in fact, only to a corresponding lengthening of the lag phase and subsequent precipitation duration. At 1 mM, precipitation was monitored to occur continuously and reproducibly over periods of ~7–8 days after the initial 24–36-h lag phase. NMR spectra taken at regular intervals showed, however, that, besides precipitation, the decrease of the structured protein signal was also the result of the onset of new resonances that clustered in the disordered peptide regions (Fig. 8). Thus, the mutant R3Aβ2-m underwent also an unfolding process leading to a soluble species. The detection of soluble species coming from unfolding of native structure after observing precipitation seemed to be the rule among the addressed β2-m derivatives, although the unfolding extent and the observed signal intensity thereof could vary. The loss of the structured species resonances, as well as the onset of new resonances from the soluble unfolded species, were followed by either one-dimensional and two-dimensional NMR measurements with R3Aβ2-m, initially in an attempt to identify specific regions of the molecule with different rates of unfolding. No such region was found by treating the data on either an absolute or a normalized scale, i.e. using the integral (volume) of an entire representative spectral moiety as normalization factor. In addition to the resolved upfield-shifted resonances (the signature of any structured protein) that could be followed also in one-dimensional spectra, a number of resolved two-dimensional cross-peaks originating from the folded structure were analyzed and, after the initial lag phase, observed to decrease with time and enable monitoring the loss of folded protein caused by precipitation and unfolding. The latter can be appreciated from the increase of resonance clustering in the disordered peptide regions. The starred peak identifies an impurity.

This, most conceivably, should reflect the variable overlap extents and stoichiometries of the measurable cross-peak volumes from the unfolded species. Because no assignment was obtained for any of the new cross-peaks, the onset-rate data were not considered further.

pK_a Values—A calculation of pK_a values of ionizable residues was conducted for β2-m, R3Aβ2-m, and H31Yβ2-m. The solution structure of β2-m exhibits a number of deviations with respect to the x-ray conformation (Fig. 7) that may be related to the solvation of areas which are buried in the HLA complex (6). We investigated therefore the electrostatic properties of β2-m using both the solution structure ensemble (5) (referred to as β2-m_nmr), and the crystal structure of HLA complex (19), either with or without the heavy chain (referred to as β2-m_HLA or β2-m_X-ray, respectively). Hence the solvation effects, computed by a continuum model, were estimated for either solid-state structures, corresponding to an in silico solvent transfer, and for the solution ensemble. On considering the 20 best-fitting members of the solution structure family, the individually determined parameters related to the electrostatic properties were averaged.

In the case of R3Aβ2-m, the input structures for UHBD calculations were based on the family of conformers determined by restrained dynamics. In the absence of a NMR-restrained structure of H31Yβ2-m, the relative UHBD calculations were performed on a model that was built based on the wild-type protein solution ensemble, in consideration of the persistence of secondary and tertiary structure inferred from the backbone chemical shifts (Fig. 4B).

Table II summarizes the UHBD pK_a values, and Fig. 10 provides their comparison in the region that appears always relevant for stability. A number of residues in β2-m_HLA showed unusual pK_a values that could be ascribed to the effect of inter- or intramolecular hydrogen bonds or salt bridges. The same holds for β2-m_X-ray and β2-m_nmr because of intramolecular H-bonds or salt bridges. Our attention was focused particularly on the pK_a variation of the histidines, the only residues whose acid-base equilibrium could be affected at experimental pH values around neutrality. Upon elimination of the MHC-I heavy chain, the computed solvation effects extend over the whole β2-m surface. In response to such a change, the pK_a of H13 remained around a value close to the reference one.
Fig. 9. Amplitude time course of the characteristically upfield-shifted methyl resonances of Leu-23 and Ile-35 from the $^1$H NMR spectrum of 2-m$_{R3A}$. Because the corresponding signals overlap partially, a single cumulative integral was measured. Following the initial lag phase, the signal loss exhibits a linear trend, i.e. a concentration-independent behavior that corresponds to a pseudo-zeroth order kinetics.

**Table II**

Calculated $pK_a$ values of ionizable residues of wild-type and mutant 2-m$_{2}$ species at 310 K

| Residue | $\beta_{2-m}_{\text{HLA}}$ | $\beta_{2-m}_{\text{X-ray}}$ | $\beta_{2-m}_{\text{nrm}}$ | $\beta_{3A\beta_{2-m}}$ | $\beta_{311Y\beta_{2-m}}$ |
|---------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Met-0   | 7.68            | 7.28            | 7.36 ± 0.19     | 7.56 ± 0.24     | 7.4 ± 0.2       |
| Arg-3   | 14.64           | 13.49           | 12.79 ± 0.33    | 12.6 ± 0.2      | 12.6 ± 0.2      |
| Lys-6   | 12.29           | 10.73           | 10.90 ± 0.29    | 11.30 ± 0.40*   | 10.8 ± 0.1      |
| Tyr-10  | 13.11           | 10.16           | 10.17 ± 0.34    | 10.22 ± 0.37    | 10.2 ± 0.5      |
| Arg-12  | 14.17           | 13.30           | 13.25 ± 0.55    | 13.12 ± 0.40    | 12.9 ± 0.3      |
| His-13  | 6.05            | 6.25            | 5.72 ± 0.54     | 5.75 ± 0.49     | 5.9 ± 0.37      |
| Glu-16  | 3.09            | 2.87            | 3.76 ± 0.45     | 3.75 ± 0.48     | 4.0 ± 0.3      |
| Lys-19  | 11.31           | 11.90           | 11.46 ± 0.47    | 11.38 ± 0.48    | 11.1 ± 0.3      |
| Tyr-26  | 12.26           | 10.05           | 10.29 ± 0.31    | 10.48 ± 0.58    | 9.9 ± 0.2      |
| His-31/Tyr-31 | 4.17 | 5.26 | 5.41 ± 0.47 | 5.60 ± 0.46 | 9.8 ± 0.2 |
| Asp-34  | 2.41            | 3.33            | 3.59 ± 0.13     | 2.84 ± 0.57*    | 3.7 ± 0.2      |
| Glu-36  | 3.72            | 3.93            | 3.97 ± 0.39     | 3.90 ± 0.24     | 4.3 ± 0.3      |
| Lys-38  | 1.25            | 2.80            | 1.96 ± 0.47     | 2.20 ± 0.59*    | 2.4 ± 0.5*     |
| Glu-44  | 4.32            | 4.26            | 3.39 ± 0.44     | 3.24 ± 0.54     | 3.6 ± 0.4      |
| Arg-45  | 14.41           | 13.49           | 13.46 ± 0.48    | 13.51 ± 0.45    | 13.4 ± 0.5     |
| Glu-47  | 3.87            | 3.96            | 3.83 ± 0.30     | 3.89 ± 0.28     | 3.8 ± 0.3      |
| Lys-48  | 11.07           | 11.00           | 11.42 ± 0.35    | 11.35 ± 0.34    | 11.1 ± 0.1     |
| Glu-50  | 4.16            | 4.19            | 3.92 ± 0.26     | 3.94 ± 0.24     | 3.9 ± 0.1      |
| His-51  | 7.17            | 7.49            | 6.64 ± 0.16     | 6.72 ± 0.17     | 6.7 ± 0.2      |
| Asp-53  | 1.66            | 3.68            | 3.41 ± 0.26     | 3.65 ± 0.19     | 3.7 ± 0.1      |
| Lys-58  | 11.22           | 10.51           | 10.62 ± 0.09    | 11.19 ± 0.36*   | 10.7 ± 0.1     |
| Asp-59  | 1.94            | 3.06            | 3.43 ± 0.27     | 2.77 ± 0.46*    | 3.7 ± 0.3      |
| Tyr-63  | 12.79           | 10.15           | 10.62 ± 0.31    | 10.15 ± 0.17*   | 10.5 ± 0.2     |
| Tyr-66  | 10.56           | 11.03           | 11.60 ± 0.38    | 11.59 ± 0.41    | 11.5 ± 0.2*    |
| Tyr-67  | 11.97           | 11.02           | 10.63 ± 0.35    | 10.56 ± 0.34    | 10.7 ± 0.3     |
| Glu-69  | 3.97            | 3.92            | 3.68 ± 0.36     | 3.78 ± 0.36     | 4.0 ± 0.1      |
| Glu-74  | 3.85            | 3.86            | 4.03 ± 0.30     | 4.09 ± 0.28     | 4.2 ± 0.2      |
| Lys-75  | 10.76           | 10.79           | 10.98 ± 0.32    | 10.94 ± 0.32    | 10.8 ± 0.1     |
| Asp-76  | 2.52            | 2.34            | 2.96 ± 0.52     | 2.05 ± 0.56     | 3.0 ± 0.3      |
| Glu-77  | 2.66            | 3.12            | 3.23 ± 0.55     | 3.31 ± 0.54     | 3.7 ± 0.5*     |
| Tyr-78  | 10.25           | 9.95            | 9.56 ± 0.45     | 9.65 ± 0.46     | 8.7 ± 0.6*     |
| Arg-81  | 14.03           | 13.31           | 13.60 ± 0.32    | 13.61 ± 0.33    | 13.3 ± 0.3     |
| His-84  | 3.80            | 3.96            | 4.14 ± 0.39     | 4.49 ± 0.65     | 4.8 ± 0.2      |
| Lys-91  | 10.83           | 10.54           | 10.50 ± 0.13    | 10.55 ± 0.26    | 10.5 ± 0.1     |
| Lys-94  | 12.26           | 12.11           | 11.53 ± 0.50    | 11.44 ± 0.53    | 11.0 ± 0.3     |
| Asp-96  | 3.32            | 3.62            | 3.62 ± 0.37     | 3.71 ± 0.34     | 3.7 ± 0.4      |
| Arg-97  | 13.76           | 13.56           | 13.62 ± 0.57    | 13.58 ± 0.63    | 13.2 ± 0.5     |
| Asp-98  | 1.95            | 4.00            | 4.00 ± 0.26     | 3.97 ± 0.24     | 4.1 ± 0.2      |
| Met-99  | 0.55            | 3.81            | 3.41 ± 0.52     | 3.48 ± 0.44     | 3.5 ± 0.3      |
(see Table II), wherever the considered structures came from, i.e. X-ray data or NMR determinations. His-31, the side chain of which in the complex experiences a salt bridge with the heavy chain counterpart, shifted its pK$_a$ value from 4.17 to 5.26 in β$_{2-m}$, X-ray, i.e. within the standard deviation of the average value of 5.41 obtained with β$_{2-m}$, nmr (Table II), where the original orientation also proves changed toward a more outward-pointing geometry (5). His-51 showed minor variations of pK$_a$, always close to the reference value, whereas His-84, which maintains the imidazole ring always embedded in the core of the protein, conserved an unusually low pK$_a$ value in any of the considered structures.

The pK$_a$ values computed for R3Aβ$_{2-m}$ were largely similar to those obtained for β$_{2-m}$, nmr, including the general trend of the histidine residues. Residues such as Lys-6, Asp-34, Lys-58, Asp-59, and Tyr-63 were predicted to display conspicuous deviations (more than 0.4 pK$_a$ units, Table II) without any relevant effect, however, on the total molecular charge at the experimental pH value of 6.6. An interesting feature was the small pK$_a$ increase estimated for His-31 with respect to the corresponding value of wild-type protein in solution, i.e. 5.60 versus 5.41. Despite the limited extent, this change is significant and in line with the expected trend. By mere substitution of alanine for arginine in position 3 of wild-type protein solution ensemble, an average pK$_a$ increase of 0.7 units could be calculated for His-31 that should represent exclusively the variation resulting from the drop of the positive charge at the mutation site. The conformational contributions conveyed by the solution ensemble of R3Aβ$_{2-m}$ balance the 0.7 units pK$_a$ increase of His-31, but only in part. As a result, the residual small pK$_a$ increase of His-31 should affect the molecular charge under the experimental pH conditions.

The effect of the mutation of histidine in position 31 with a tyrosine were estimated to cause remarkable shifts on pK$_a$ values of Asp-38, Tyr-66, Asp-76, Glu-77, Tyr-78, and Lys-94 with respect to the corresponding values in wild-type solution structure. Besides the slight decrease of positive charge as a result of the direct consequence of the mutation, these shifts should contribute again no additional major molecular charge variation near neutrality conditions.

Unfolding free energy can be split into an electrostatic, pH-dependent term, and a non-electrostatic term. The electrostatic contribution to the folded form free energy is directly related to pK$_a$ values according to the equation described by Yang and Honig (40). Analogously, for the unfolded form, a similar computation can be performed assuming that the pK$_a$ values of ionizable groups in denatured state are identical to those obtained for isolated amino acids.

In Table III, the calculated ΔG$_{el}$ values are listed for wild-type β$_{2-m}$ (either crystal and solution structure) and the two investigated mutants, along with the experimental global unfolding ΔG values. A direct comparison of calculated and experimental free energy differences is not feasible because the experimental ΔG values of unfolding include the non-electrostatic contributions. A more sensible parameter is ΔG$_{el}$ with respect to wild-type species, when compared with the corresponding experimental ΔG. The computed data show a higher electrostatic contribution to stability for both mutants with respect to wild-type protein and a reduced electrostatic contribution for β$_{2-m}$, X-ray. The latter cannot be directly compared with experimental values, but could be considered to account for the rearrangements the protein undergoes in solution leading to an increased electrostatic stabilization. A similar rationale cannot be applied directly to the other results of Table III because of the starting charge differences between wild-type and mutants. Hence the ΔG$_{el}$ values obtained for both mutants can no longer be interpreted only in terms of local conformational differences with respect to the natural sequence.

For His-31Yβ$_{2-m}$, the increased thermodynamic stability expressed by the experimental ΔG of unfolding is in qualitative agreement with the computed increase of electrostatic contribution to free energy. On the other hand, the computed ΔG$_{el}$ values for the mutant R3Aβ$_{2-m}$ indicates that the electrostatic contribution can hardly account for the experimental ΔG of unfolding of the species.

To take into account hydrophobic interactions between the protein and the solvent, which contribute to the non-electrostatic terms of free energy, a calculation of the solvation component (ΔG$_{sol}$) in the factorization of global unfolding ΔG was undertaken. The procedure to calculate ΔG$_{sol}$ values entails defining a reference state, typically the unfolded state, and hence introduc-
The experimental data obtained for solutions of β2-m, R3Aβ2-m, and H31Yβ2-m are compared with the calculated values for the NMR structural families, averaged over the ensemble of 20 best-fitting conformations, or the crystal structure of β2-m in MHC-I. Only for H31Yβ2-m were calculations performed on a model that was built based on the wild-type protein solution ensemble. All values are in kcal/mol.

The differences are determined with respect to the experimental value obtained for wild-type β2-m.

The differences in folding are calculated for (β2-m_nmr), which was shown to occur primarily at His-31 and involve partially His-13 as well, the next available His residue along the transition pathway (5).

The solvation sensitivity of β2-m surface at the interface with the heavy chain of MHC-I was first recognized after realizing that the extended hydrophobic patch embedding the four positive charges of Arg-3, Lys-6, Lys-58, and Lys-91 (spread over the surface shown in Fig. 7) had to release the MHC-I packing (39) and accommodate some mobility of the charged side chains to supply a favorable contribution to free energy (6). The lack of such a contribution and the ensuing instability (Table I) could explain the properties and behavior of fragment ΔN6β2-m observed in vitro (7, 6).

In vivo the same fragment accounts for ~30% of the protein material in β2-m fibrils and could originate from proteolysis (41), once strand A detachment from sheet 1 has made accessible and mobile the N-terminal segment, in agreement with the amyloid transition pathway already described. Although the exact timing of ΔN6β2-m formation, i.e., whether prefibrillar or postfibrillar, is not yet established, the available data on structure and the mild conditions used for fibrillogenesis, along with the results from limited proteolysis (6, 41), strongly indicate that the species can be considered the closest representation of the amyloidogenic conformation of β2-m.

To check further the solvation sensitivity-driven mechanism outlined above, the variants ΔN3β2-m and R3Aβ2-m were designed and characterized. The unfolding free energies were intermediate between the values of wild-type and fragment ΔN6β2-m (Table I), confirming that the positive charges on the N-terminal strand are a stability factor for the isolated protein. Compared with ΔN6β2-m, the fragment ΔN3β2-m does not exhibit quick precipitation and, based on the ΔΔH" values (Fig. 4A), its structure in solution should remain quite similar to that of the parent protein over relatively long periods (weeks), without any hint of increased aggregation as observed with the shorter fragment (6). The occurrence of two different conformers, in slow exchange on the NMR time scale (Fig. 5 and Supplemental Material, available in the on-line version of this article), does not seem to affect this conclusion. In response to loss of the N-terminal tripeptide, it is likely that the fragment ΔN3β2-m is capable of arranging in solution a vicarious role of the N-terminal amino group of Met-0, which is located close to the former location of the missing Arg-3 residue. The importance of the positive charge at position 3 to complement the

### Table III

| Experimental | Calculateda | Calculatedb |
|-------------|-------------|-------------|
| β2-m X-ray  | 7.5         | 34 (±13)    |
| (β2-m_nmr)  | -1.8        | -1 ± 26     |
| (R3Aβ2-m)   | 6.3 ± 0.2   | 9.3 ± 1.3   |
| (H31Yβ2-m)  | 5.9 ± 0.2   | 10.1 ± 2.1  |
| (H31Yβ2-m)  | 7.8 ± 0.4   | 13.1 ± 2.1  |

* α Unfolding free energies were determined at pH = 7.3 and T = 293K (see Table I).
* β ΔG_{sol}^{un} represents the electrostatic contribution to unfolding calculated with UHBD program at 293 K and pH = 7.5.

### DISCUSSION

#### Stability and Structure of β2-m: Variants Versus Wild Type—

The solvation structure of human β2-m, determined by 1H NMR spectroscopy and restrained modeling calculations (5), exhibits several deviations compared with the crystal structure of the protein associated to MHC-I heavy chain component (Fig. 7). These structural differences can be considered as the prodromes of the amyloid transition. Solvation of the protected regions in MHC-I leaves β2-m molecule in a poorly stabilized state, where even minor charge and solvation changes, in response to pH or ionic strength variations, can easily compromise the hydrophobic/hydrophilic balance and drive a transition to relieve the hydration stress. This transition should imply unpairing of strand A and should lead to polymerization into fibrils or amorphous aggregates and, eventually, precipitation (5, 6). The same mechanism can account for the partial unfolding and fiber formation subsequent to Cu2+ binding, which was shown to occur primarily at His-31 and involve partially His-13 as well, the next available His residue along the transition pathway (5).

The experimental data obtained for solutions of β2-m, R3Aβ2-m, and H31Yβ2-m are compared with the calculated values for the NMR structural ensemble, averaged over the ensemble of 20 best-fitting conformations, or the crystal structure of β2-m in MHC-I.
number of positive charges inserted in a surface hydrophobic patch is highlighted by the results obtained with R3Aβ2-m. Apparently its NMR solution structure exhibits limited variations with respect to the wild-type protein, mostly clustered at the D-E loop, but the samples do not last more than a couple of weeks as a result of precipitation and unfolding. The thermodynamic stability of R3Aβ2-m mutant is higher than that of the examined truncated variant, ΔN3β2-m, because of the additional N-terminal tripeptide of the mutant and the conceivable free energy contributions thereof. On the other hand, the absence of the positive side chain at position 3 in R3Aβ2-m brings about a small but significant destabilization with respect to wild type that should be ascribed to the hydrophobic solvation contribution, according to the earlier interpretation (6). The calculated values from theoretical dissection of the unfolding free energy confirm the expectations. From Table III it is clear that the destabilization of R3Aβ2-m with respect to wild-type protein is not because of the electrostatic term (that, instead, appears to stabilize slightly), but to the solvation term. The result, albeit only qualitative, is valuable because experimentally reliable if one considers that it was obtained from the NMR-based conformational family of R3Aβ2-m. The effects of the altered solvation term, which can nicely account for the decreased stability of R3Aβ2-m, should not be confined only to the mentioned hydrophobic surface. Other sensible surface areas are the loops B-C and D-E, the end of strand D and the start of strand E, all located on the side of the N terminus of the protein, above the extended hydrophobic patch region (Fig. 7).

In particular, exposed surface changes must be certainly involved at His-31 side chain, orientation change (5) and pK_a increase (Table II) of which can be directly related to an increase of the local positive charge at neutral or slightly acidic pH, and loss of packing with the N-terminal residues of strand A. This mechanism, although contributing to destabilization of R3Aβ2-m with respect to wild-type protein, as inferred from the corresponding His-31 pK_a value (Table II), should be considered the basic one to accelerate strand A detachment from sheet 1 upon further increase of His-31 positive charge by either pH decrease or Cu^{2+} binding (5). Because the link between Cu^{2+} binding and partial unfolding of β2-m is the local charge increase on His-31, equivalent to acidification of a neutral β2-m solution, to design a more stable protein we reasoned that a mutant His-31→Tyr might prevent the destabilizing imidazole titration, while preserving the local aromatic character. The expectation was successful as the designed mutant has been shown to display increased folding stability relative to the wild-type protein (Table I), along with reduced Cu^{2+} binding (42). Although the His-31-mediated mechanism for Cu^{2+} specificity was judged reasonable, an additional, alternative view was advanced involving a copper binding site in non-native (unfolded) β2-m states (32). The issue would deserve further consideration, out of the scopes of the current work, but it is worth noting here that the crucial role of His-31 for the stability of isolated β2-m is confirmed by the intermediate stability of mutant His-31→Ser, more stable than wild-type β2-m as a result of absence of positive charge at Ser-31, but less stable than H31Yβ2-m because of the decreased hydrophobic character of the same residue. That the main contribution to H31Yβ2-m stabilization comes from the loss of positive charge at the mutation site follows also from theoretical results of Table III. The electrostatic contribution to free energy accounts for the whole stability gain of the mutant, with no resulting solvation contribution. Although the results for H31Yβ2-m may be biased by the specific conformation chosen for calculations (modeled from the wild-type ensemble in solution), recent x-ray structural evidence are in fairly good agreement with the conformational features imposed to H31Yβ2-m solution structure (43). Thus, no incorrect conformational bias should affect the computed free energy components of Table III.

The increased stability of H31Yβ2-m compared with wild type and R3Aβ2-m has also been related to the different rates of conformational exchange occurring at the D-E loop. Indeed the dynamics at this loop, which can only be inferred indirectly from NMR data, are an additional facet of the solvation arrangement that β2-m and variants thereof reach in aqueous solution. In MHC-I, in fact, loop D-E and the contiguous strand segments of β2-m are also part of the interface between the two proteins of the complex (39). The analysis of unfolding and refolding kinetics of H31Yβ2-m reveals features similar to those reported for wild-type β2-m (14). Clearly the slower rate of unfolding measured for the mutant with respect to wild type correlates with the increased stability of the former. On the other hand, the increased rate of the slow refolding phase measured for both H31Y and H31S mutants can be tentatively interpreted, based on the rearrangement of the hydrophobic core observed to occur in wild-type protein during the same phase (Fig. 2). Both mutants bear a sequence change at His-31, an extremely critical residue for the consequences of its pK_a variation. This type of variation is intimately linked with local conformation changes. Besides considering the related effects on the solvation and packing of the N-terminal moiety of strand A, it is worth remembering that the orientation of the imidazole group in the solution structure of wild-type β2-m (5) is distinctly different from the corresponding orientation in MHC-I crystal (39). This should affect the neighboring Phe-30 and, in turn, the packing of the latter with Ile-35, i.e. residues involved in the complex, and probably cooperative, rearrangement of the hydrophobic core undergone by the protein during refolding (Fig. 2). Hence, it should not be surprising that mutations at His-31 may influence the rearrangement process and its kinetics. In particular, because both H31Y and H31S mutants are more stable than wild type and exhibit a faster slow phase-refolding rate (Fig. 3), it should be concluded that their I_2 intermediates are closer to the final structure (or the transition state toward it) and thus reach it more easily than wild-type protein, because of the effects of the mutations in position 31. Although this may be the result of different scenarios, at least for H31Yβ2-m the available structural evidence (43) could suggest that an orientation of Tyr-31 side chain similar to that of β2-m in MHC-I (39) may greatly facilitate the hydrophobic core rearrangement and therefore speed up the final step of refolding.

To summarize, a thorough analysis of the structure and stability of two β2-m mutants, R3A and H31Y, along with additional data from a truncated variant, ΔN3β2-m, and another mutant, H31S, fully support the plausibility of an amyloidogenic pathway of β2-m based on recognition of the transition prodomes from the solution structure of the isolated protein (5, 6). The results of a systematic investigation on 13 amyloidogenic β2-m mutants have recently been published (44, 45). Although no correlation was proposed between the location of the mutations and the extent of destabilization with respect to wild-type sequence, a unique role in amyloid formation was envisaged for the N and C-terminal β-strands of β2-m (λ and G, respectively) from the increased fibrillogenesis rate at acidic and mildly acidic pH of the variants with mutations in strands A and G. The underlying rationale is based on the loss of local hydrophobic packing by β-strand unpairing, leading to increased population of conformers with exposed assembly-competent surfaces. By different arguments, from previous (5, 6, 41) and present evidence, we have proposed that detachment of strand A is the crucial event leading to intermolecular pairing
of β2-m. Because the terminal β-strand unfolding and the subsequent local mobility increase should enhance proteolytic cleavage at the N-terminal, as already mentioned (6, 41), as well as at the C-terminal β-strand of wild-type β2-m sequence, we note that the terminal strand unfolding appears more likely to occur at the N terminus because of the much higher abundance of N-terminally truncated species (>30%) than C-terminally truncated ones (<1%) in natural fibrils (7).

**On the Amyloid Transition**—The experimental and theoretical results obtained with the examined variants of β2-m confirm, overall, the important role that charge and solvation changes assume in determining the stability of the protein. Besides the extended hydrophobic patch interrupted by the positive charges of Arg-3 and Lys-6 (N-terminal region), Lys-99 (D-E loop), and Lys-91 (G strand), another potentially solvation-sensitive location is His-31 in the B-C loop. The inherently stressed structure of isolated β2-m reacts to any charge and solvation modification at these hot spots by accelerating a transition, the amyloid one, which starts with unfolding and disordering of strand A and leads to intermolecular pairing through strands D and C and, eventually, to precipitation. Several lines of evidence have been accumulated in favor of this mechanism (5, 6, 8, 9, 41, 43, 46–48), which prompts some considerations concerning the amyloid transition of β2-m. It is widely accepted, in fact, that amyloid formation proceeds through partial unfolding transformations of the parent protein, which reaches a crucial intermediate conformation that can either undergo further unfolding or become trapped into an intermolecular aggregation process (2). This alternative pathway leads eventually to precipitation: the fibrillar, amorphous, or mixed nature of the precipitate depending on the actual experimental conditions. Most generally, the mechanism does not necessarily imply the intermediate to be a metastable species, although it is much more than intuitive to postulate a metastable intermediate at any bifurcation point. It is clear that the more stable the intermediate, the easier its observation. Metastable intermediates may even be considered addressable targets, not necessarily for blocking the amyloid transition, but, rather, also for studying their properties to prevent the process. On the other hand, partial unfolding intermediates may also be very transient species. This indeed appears to be the case with β2-m amyloid transition. No trace of stable intermediate species has ever been observed in the NMR spectra of either the wild-type or variant proteins here examined, whereas for the least stable and fastest fibril-forming variant of Table I, ∆N6β2-m, only evidence of intermolecular association was obtained from gel filtration and NMR data (6). Extensively unfolded species were detected in conjunction with the observation of precipitates in R3Aβ2-m and H31Yβ2-m solutions, and similar results had been obtained occasionally also with some preparations of wild-type β2-m, in particular when precipitation was observed to occur shortly after solution preparation because of the presence of insoluble particles. Quantitative measurements were limited to R3Aβ2-m because of the particularly convenient time frame of the process in this mutant. The observation of soluble unfolded species always and only in the presence of precipitates, and the pattern of the structured protein resonance loss, which exhibits a clear lag phase and a pseudo-zeroth order kinetics (Fig. 9), suggest the occurrence of a pathway where the intermediate formation is a necessary, but most likely transient step, along the more complex and involved route of nucleated conformational conversion (NCC) (49–51). According to this mechanism, soluble, conformationally dynamic oligomers should undergo a nucleation transition, to form the assembly-competent species that grow into fibrillar aggregates. The nucleation event is the rate-de-

**FIG. 11. Proposed NCC-type mechanism (49–52) for β2-m fibrillogenesis.** The rate-determining step is the conversion of the dynamic oligomer (OliD) into the nucleated oligomer (OliN) and entails a concerted transition by which Nn closely spaced monomers convert simultaneously into a nucleus of intermediates. This is highlighted by the pictorial representation of the transition state above the arrows of the rate-determining step. In general, OliD could be a series of aggregates with various stoichiometries or a series of adducts that are formed by association equilibria. OliD should be composed of either folded species (Mf) and conformationally activated or partially unfolded species (Mf). In the general scheme above outlined, Mf monomers are considered in equilibrium with Mf, i.e., as thermodynamic intermediates. However, Mf monomers, whether partially unfolded species or activated folded states, could also be transient species, i.e., kinetic intermediates. This latter nature of Mf could be the case with β2-m based on the NMR evidence (Figs. 2 and 9), despite the former possibility (thermodynamic intermediate) cannot be definitely ruled out, as also suggested by previous results (15, 42). Once formed, OliN elongates by addition of Mf, Mf, or OliD and subsequent conformational rearrangement of the incorporated monomers. Elongated nuclei (OliNk) keep growing to form eventually a fibrillar, amorphous, or mixed precipitation, but may also release partially unfolded monomers, in a conformational state other than Mf. These released species undergo irreversible further unfolding to reach a soluble unfolded state (Mf). Soluble unfolded species may be obtained, of course, also directly from partially unfolded or destabilized monomers, a route that is not necessarily alternative to the release from nucleated species, but may coexist. Note the different colors and arrangement of the monomers in the Mf nucleus and in OliN. The symbols indicate that different intra- and intermolecular conformational arrangements are adopted within the Mf nucleus and the fibril nucleus. The rate-determining step that should explain a lag phase because of the simultaneous and precise partial unfolding transition that a minimum number of monomers (the nucleation number, Nn) must undergo to establish a nucleus, whereas the lack of concentration dependence of the structured protein signal decrease (pseudo-zeroth order kinetics) should be the consequence of the direct dependence of the precipitation and unfolding rates on the actual concentration of the nucleated protein reservoir, a group of species that are largely unobservable by NMR as a result of their dimensions. As checked with R3Aβ2-m, the increase of lag phase duration upon decreasing the protein concentration, and the accompanying slope reduction for the linearly decreasing part of the time-course function corresponding to the trend shown in Fig. 9, are in line with the expected pattern from a NCC-type mechanism (51). A nucleation-dependent polymerization model has already been proposed for β2-m fibril formation *in vitro* and *in vivo*, based on kinetic data analysis (36). The present qualitative kinetic observations, while suggesting the extension of the NCC mechanism to β2-m fibrillogenesis, also provide some clues that can be of general interest for amyloid transition of globular proteins. In particular, the link between unfolded species and precipitate, which was invariably seen in our experiments, suggests that extensive protein unfolding leading to soluble species should occur via release of properly unfolded monomers from the nucleated species, and irreversible, further unfolding of the same mono-
mers to remain in solution, according to the scheme of Fig. 11. The rate-limiting step (the dynamic-to-nucleated oligomer conversion) entails achieving the transient intermediate structure in $N_n$ closely-spaced monomers within the conformationally dynamic oligomer. Rather than an aggregate with definite stoichiometry, the dynamic oligomer of $\beta_2$-m could most likely be one of the adducts along a series of ongoing association equilibria (6). The NCC mechanism is a general, flexible scheme, which can account for different specific aspects of a broad class of nucleation-based processes, as recently shown by Bitan and colleagues (52). No matter whether a proper oligomerization or a nucleation event. This means considering an ensemble of molecules thereof. Similar considerations extend also to the formation of a stabilized protein form undergoing an easier amyloidogenic transformation.

In the context of a NCC mechanism, the actual intermediate of $\beta_2$-m along the amyloid transition should be the nucleated oligomer. However, the possibility of transient intermediate(s) formation before the nucleation step should not be ruled out a priori. Within the NCC framework (49–51), the features of the dynamic oligomer are so loose that postulating monomers, already in partially unfolded conformation(s) or in a sort of activated state like the $I_2$ species described from NMR evidence, that occur either sporadically over the dynamic oligomer population, or even promote the same dynamic oligomerization being thus responsible for it, could fit the mechanism anyway (Fig. 11). The really qualifying aspect of the NCC scheme is the requirement of a concerted transition involving a precise number of closely spaced monomers, as opposed to the single molecule event. This means considering an ensemble of molecules before and after the transformation, and the thermodynamics thereof. Similar considerations extend also to the formation of unfolded soluble species. First, an oligomeric nature for the unfolded soluble species cannot be excluded. The above-mentioned notation $M_2$ stems from the NMR line sharpening, i.e. a qualitative argument, observed in solutions of R3A$\beta_2$-m, an ideally suited mutant to follow, within reasonably short periods, precipitation, and unfolding. Second, obtaining an extensively unfolded species in solution, whether monomeric or not, not only from the dissociation equilibrium of an elongated nucleus, could well be related to the presence of a partially unfolded intermediate prior to nucleation.

CONCLUSIONS

The NCC mechanism, which was demonstrated for Alzheimer Aβ42 peptide (49, 52) and Sup35 NM domain (50) and proposed as the most general and reliable scheme for amyloidogenesis (51), seems to apply also to $\beta_2$-m fibril formation. The NCC scheme is consistent with the experimental findings from the present study on structure, precipitation, and unfolding of $\beta_2$-m variants, along with the role of the previously identified amyloid transition prodomes of wild-type $\beta_2$-m (5) for the onset of a transient intermediate. The consistency extends also to the Goto group proposal concerning the dependence of $\beta_2$-m amyloidogenic potential on the mobility restrictions imposed by the single disulfide bridge of the molecule (47, 48). By reducing this disulfide bridge, the helical propensity of strand E takes over and spoils the quality of the intermolecular $\beta$-strand pairing that should occur at the adjacent D strand. On the other hand, by keeping the disulfide, a rigidity constraint is introduced at strand E, which prevents spoiling of the intermolecular pairing. The NCC scheme entails oligomer nucleation, which can take place only via intermolecular interactions. Polypeptide chains can maximize intermolecular interactions by establishing pairings between $\beta$ strands of individual chains, to form sheets where the acid-base potential of backbone amide and carboxyl groups are satisfied intermolecularly. Conversely, polypeptides can maximize intramolecular interactions by adopting $\alpha$-helical conformations where the same acid-base potentials are internally balanced. The result of the competition between inter- and intramolecular interactions, that, in turn, are competed anyway by solvent-polypeptide interactions, is determined by the specific amino acid sequence. Therefore, the amyloidogenic potential of any polypeptide chain could be ultimately related to its intramolecular interaction potential in water, i.e. the same potential that determines the overall folding in water. Exceptions to this scheme, such as intermolecular interactions established by individual helices, may well be possible, but the general character of the conclusions appears consistent with the unifying soundness of the NCC mechanism for amyloidogenesis, as suggested by recent results on the recognition of soluble oligomers of different amyloidogenic proteins by a single antibody, which also inhibits the corresponding in vitro toxicity (54).

Acknowledgments—G. E. thanks Dr. F. Al Piaz for helpful discussions and Dr. A. Makek for assistance.

REFERENCES

1. Merlino, G., and Bellotti, V. (2003) Bioessays 25, 849–856
2. Dobson, C. M. (1999) Trends Biochem. Sci. 24, 329–332
3. Gejyo, F., Yamada, T., Odani, S., Nakagawa, Y., Arakawa, M., Kunitomo, T., Kataoka, H., Suzuki, M., Hirawasa, Y., Shishara, T., Cohen, A. S., and Schmid, K. (1985) Biochim. Biophys. Res. Commun. 128, 701–706
4. Uchanska-Ziegler, B., and Ziegler, A. (2003) Trends Immunol. 24, 73–76
5. Verdine, G., Corazza, A., Vignolo, P., Petrossi, F., Giorgietti, S., Mangione, P., Andreola, A., Stoppani, M., Bellotti, V., and Esposito, G. (2002) Protein Sci. 11, 487–499
6. Esposito, G., Michelutti, R., Verdine, G., Vignolo, P., Hernández, H., Robinson, C. V., Amoresano, A., Diaz, F., Mantzi, M., Pucci, P., Mangione, P., Asti, L., Stoppani, M., Merlino, G., Ferri, G., and Bellotti, V. (2000) Protein Sci. 9, 831–845
7. Bellotti, V., Stoppani, M., Mangione, P., Sunde, M., Robinson, C. V., Asti, L., Brancaccio, D., and Ferri, G. (1998) Eur. J. Biochem. 256, 61–67
8. McParland, V. J., Kalverda, A. P., Homans, S. W., and Radford, S. E. (2002) Nat. Struct. Biol. 9, 526–531
9. Hoshino, M., Kato, H., Hagiwara, Y., Hasegawa, K., Naihi, K., and Goto, Y. (2002) Nat. Struct. Biol. 9, 332–336
10. Trinh, C. H., Smith, D. P., Kalverda, A. P., Phillips, S. E., and Radford, S. E. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 9771–9776
11. Tsunenaga, M., Goto, Y., Kawata, Y., and Hamaguchi, K. (1987) Biochemistry 26, 6044–6051
12. Santoro, M. M., and Bolen, D. W. (1988) Biochemistry 27, 8063–8068
13. Pace, C. N. (1986) Methods Enzymol. 131, 266–280
14. Chiti, F., Mangione, P., Andreola, A., Giorgietti, S., Stefani, M., Dobson, C. M., Bellotti, V., and Taddei, N. (2001) J. Mol. Biol. 307, 379–391
15. Chiti, F., De Lorenzo, E., Grossi, S., Mangione, P., Giorgetti, S., Cacciabue, G., Dobson, C. M., Merlino, G., Ramponi, G., and Bellotti, V. (2001) J. Biol. Chem. 276, 46714–46721
β2-M Variants and Amyloidogenesis

18. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) J. Chem. Phys. 71, 296–299
19. Piotto, M., Saudek, V., and Sklenar, V. (1992) J. Magn. Reson. 98, 97–104
20. Hwang, T. L., and Shaka, A. J. (1995) J. Magn. Reson. 112, 275–279
21. Marion, D., and Wu (1995) J. Magn. Reson. 113, 201–205
22. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) J. Magn. Reson. 48, 296–299
23. Keeler, J., Clowes, R. T., Davis, A. L., and Laue, E. D. (1994) Methods Enzymol. 239, 145–207
24. Bax, A., and Davis, D. G. (1985) J. Magn. Reson. 65, 355–360
25. Shaka, A. J., Lee, C. J., and Pines, A. (1988) J. Magn. Reson. 77, 274–293
26. Dariyider, H. E., Tromontano, A., Sprang, S. R., and Fletterick, R. J. (1986) J. Mol. Graphics 6, 82–87
27. Korsak, R., Billette, M., and Wuthrich, K. (1996) J. Mol. Graphics 14, 51–55
28. Madura, J. D., Briggs, J. W., Wade, R. C., Davis, M. E., Luty, B. A., Ilin, A., Antosiewicz, J., Gilson, M. K., Bagheri, B., Scott, L. R., and McCammon, J. A. (1995) Computer Phys. Commun. 91, 57–95
29. Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1994) J. Mol. Biol. 238, 415–436
30. Fogelari, F., Ragona, L., Lizziardi, S., Romagnoli, S., Michelutti, R., Ugolini, R., and Molinari, H. (2000) Proteins 39, 317–330
31. Eisenberg, D., and McLachlan, A. D. (1986) Nature 319, 199–203
32. Eakin, C. M., Knight, J. D., Morgan, C. J., Gelfand, M. A., Miranker, A. D. (2002) Biochemistry 41, 10646–10655
33. Wuthrich, K. (1986) NMR Spectroscopy of Proteins and Nucleic Acids, Wiley & Sons, New York
34. Okon, M., Bray, P., and Vucelic, D. (1992) Biochemistry 31, 8906–8915
35. Wishart, D. S., and Sykes, B. D. (1994) Methods Enzymol. 239, 363–392
36. Nuki, H., Hashimoto, N., Suzuki, S., Kimura, H., Nakakuki, K., and Geyo, F. (1997) Amyloid Int. J. Exp. Clin. Investig. 4, 225–232
37. McParland, V. J., Kad, N. M., Kalverda, A. P., Brown, A., Kirwin-Jones, P., Hunter, M. G., Sunde, M., and Radford, S. E. (2000) Biochemistry 39, 8735–8746
38. Morgan, C. J., Gelfand, M., Attrey, C., and Miranker, A. D. (2001) J. Mol. Biol. 309, 339–345
39. Saper, M. A., Bjorkman, P. J., and Wiley, D. C. (1991) J. Mol. Biol. 219, 277–289
40. Yang, A., and Honig, B. (1993) J. Mol. Biol. 231, 459–474
41. Monti, M., Prince, S., Giorgetti, S., Mangione, P., Merlini, G., Clark, A., Bellotti, V., Amoresano, A., and Pucci, P. (2002) Protein Sci. 11, 2362–2369
42. De Lorenzi, E., Grossi, S., Massolini, G., Giorgetti, S., Mangione, P., Andreola, A., Chiti, F., Bellotti, V., and Caccialanza, G. (2002) Electrophoresis 23, 918–925
43. Rosano, C., Zuccotti, S., Mangione, P., Giorgetti, S., Bellotti, V., Pettirussi, F., Corazza, A., Viglino, P., Esposito, G., and Bolognesi, M. (2003) J. Mol. Biol. 335, 1051–1064
44. Jones, S., Smith, D. P., and Radford, S. E. (2003) J. Mol. Biol. 330, 935–941
45. Smith, D. P., Jones, S., Serpell, L. C., Sunde, M., and Radford, S. E. (2003) J. Mol. Biol. 330, 943–954
46. Kanahiro, G. V., Hagiwara, Y., Kawakami, T., Hasegawa, K., Naka, H., and Goto, Y. (2002) J. Biol. Chem. 277, 1310–1315
47. Kato, H., Kanno, T., Hoshino, M., Hagiwara, Y., Tanaka, H., Kai, T., Hasegawa, K., Naka, H., and Goto, Y. (2002) Protein Sci. 11, 2218–2229
48. Ohhashi Y., Hagiwara, Y., Koshizuka G. V., Hasegawa, K., Yamaguchi, I., Naka, H., and Goto, Y. (2002) J. Biochem. (Tokyo) 131, 45–52
49. Lomakin, A., Teplow, D. B., Kirschner, D. A., and Benedek, G. B. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 7942–7947
50. Serino, T. R., Cashikar, A. G., Koval A. S., Sawicki, G. J., Serpell, L., Arnsdorf, M. F., and Lindquist, S. L. (2000) Science 289, 1317–1321
51. Kelly, J. W. (2000) Nat. Struct. Biol. 7, 824–826
52. Bittar, G., Kirkikidze, B. D., Lomakin, A., Vlakov, S. S., Benedict, G. B., and Teplow, D. B. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 330–335
53. Zhang, S., Iwata, K., Luchemann, M. J., Peng J. W., Li, S., Stimson, E. R., Lu, Y.-a, Felix, A. M., Maggio, J. E., and Lee, J. P. (2000) J. Struct. Biol. 130, 130–141
54. Kayed, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W., and Glabe, C. G. (2003) Science 300, 486–489