

ARTICLE

DOI: 10.1038/s41467-018-04078-y

OPEN

# Conformational dynamics in crystals reveal the molecular bases for D76N beta-2 microglobulin aggregation propensity

Tanguy Le Marchand<sup>1</sup>, Matteo de Rosa<sup>2</sup>, Nicola Salvi<sup>3</sup>, Benedetta Maria Sala<sup>2</sup>, Loren B. Andreas<sup>1</sup>, Emeline Barbet-Massin<sup>1</sup>, Pietro Sormanni<sup>4</sup>, Alberto Barbiroli<sup>5</sup>, Riccardo Porcari<sup>6</sup>, Cristiano Sousa Mota<sup>7</sup>, Daniele de Sanctis<sup>7</sup>, Martino Bolognesi<sup>2,8</sup>, Lyndon Emsley<sup>1</sup>, Vittorio Bellotti<sup>6</sup>, Martin Blackledge<sup>3</sup>, Carlo Camilloni<sup>2</sup>, Guido Pintacuda<sup>1</sup> & Stefano Ricagno<sup>2</sup>

Spontaneous aggregation of folded and soluble native proteins *in vivo* is still a poorly understood process. A prototypic example is the D76N mutant of beta-2 microglobulin ( $\beta$ 2m) that displays an aggressive aggregation propensity. Here we investigate the dynamics of  $\beta$ 2m by X-ray crystallography, solid-state NMR, and molecular dynamics simulations to unveil the effects of the D76N mutation. Taken together, our data highlight the presence of minor disordered substates in crystalline  $\beta$ 2m. The destabilization of the outer strands of D76N  $\beta$ 2m accounts for the increased aggregation propensity. Furthermore, the computational modeling reveals a network of interactions with residue D76 as a keystone: this model allows predicting the stability of several point mutants. Overall, our study shows how the study of intrinsic dynamics in crystallo can provide crucial answers on protein stability and aggregation propensity. The comprehensive approach here presented may well be suited for the study of other folded amyloidogenic proteins.

<sup>1</sup>Centre de RMN à Très Hauts Champs, Institut des Sciences Analytiques (UMR 5280 CNRS/UCB Lyon 1/ENS Lyon), Université de Lyon, 69100 Villeurbanne, France. <sup>2</sup>Dipartimento di Bioscienze, Università degli Studi di Milano, 20133 Milano, Italy. <sup>3</sup>Institut de Biologie Structurale, CNRS, CEA, UGA, 30044 Grenoble, France. <sup>4</sup>Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK. <sup>5</sup>Dipartimento di Scienze per gli Alimenti, la Nutrizione e l'Ambiente, Università degli Studi di Milano, 20133 Milano, Italy. <sup>6</sup>Wolfson Drug Discovery Unit, Centre for Amyloidosis and Acute Phase Proteins, University College London, London NW3 2PF, UK. <sup>7</sup>ESRF - The European Synchrotron, 38043 Grenoble, France. <sup>8</sup>Centro di Ricerca Pediatrica Romeo ed Enrica Invernizzi, Università degli Studi di Milano, 20133 Milano, Italy. Correspondence and requests for materials should be addressed to C.C. (email: [carlo.camilloni@unimi.it](mailto:carlo.camilloni@unimi.it)) or to G.P. (email: [guido.pintacuda@ens-lyon.fr](mailto:guido.pintacuda@ens-lyon.fr)) or to S.R. (email: [stefano.ricagno@unimi.it](mailto:stefano.ricagno@unimi.it))

Numerous pathologies are related to the transformation of functional well-soluble proteins into amyloid fibrils. Many involve disordered peptides or proteins (e.g., A $\beta$ ,  $\alpha$ -synuclein). More surprisingly, globular proteins (e.g., beta-2 microglobulin, superoxide dismutase, and lysozyme) can also be associated to amyloidosis diseases<sup>1</sup>. Yet, evolution has highly protected their native state from aggregation, and for these proteins special conditions, such as low pH or high metal ion concentration, are required for their aggregation *in vitro*. In contrast, specific mutations are able to dramatically increase the potential of these proteins to form fibrils. A recent case of systemic amyloidosis involves the highly unstable D76N mutation of beta-2 microglobulin ( $\beta$ 2m)<sup>2</sup>. Patients carrying this mutation exhibit large amyloid deposits in internal organs, leading to progressive gastrointestinal symptoms and autonomic neuropathy. *In vitro*, the native fold of the D76N mutant displays a strikingly low thermodynamic stability compared to the wild type (wt) protein<sup>2,3</sup>. The D76N mutant is strongly amyloidogenic and, in contrast to the wt protein, aggregates efficiently and abundantly without seeding and under native conditions<sup>4</sup>.

Common effects of point mutations are structural perturbations of the backbone, the impairment of the hydrophobic core or the introduction of hydrophobic patches on the protein surface, with potentially very drastic outcomes involving a marked change in size or in the chemical properties<sup>1</sup>. None of these effects can nevertheless explain the enhanced reactivity of D76N  $\beta$ 2m. Indeed, the 76 position is located in a solvent exposed loop between  $\beta$ -strands E and F (Fig. 1a). The D76N mutation does not introduce a hydrophobic residue, which could lower protein solubility and Asp and Asn are isosteric. From the structural point of view the D76N mutation does not trigger major effects neither in solution nor in the crystalline state<sup>2,4</sup>. In particular, the crystal structures of wt  $\beta$ 2m and of D76N mutant match very closely and the residue 76 can be well superposed in the two structures<sup>2</sup>.

However, the crystal structure does not exhaustively define the native ensemble of a protein. Not only the position of each atom rapidly (picosecond to nanosecond) fluctuates around its equilibrium position, but also entire parts of proteins may undergo coordinated slow motions (from microseconds to longer than seconds). Such local and long-range conformational dynamics are intimately connected to biomolecular function and malfunction. Sparsely populated conformations, interconverting with the major state can be more reactive and aggregation prone, becoming a driving force in amyloid formation<sup>1,5</sup>. In order to better understand the amyloidogenic properties of D76N, it is therefore necessary to investigate the free energy landscape of both the wt and D76N  $\beta$ 2m variants.

Nuclear magnetic resonance (NMR) has the unique ability to reach both spatial (atomic) and temporal (from tens of picoseconds to milliseconds) resolution. Although D76N  $\beta$ 2m rapidly converts into non-soluble oligomers and amyloid fibrils in solution, the intrinsic dynamics of both wt and mutant forms of the protein can be studied in the crystals using solid-state NMR (ssNMR). Additionally, the absence of molecular tumbling and the capacity of the probes to achieve higher spin-lock powers in ssNMR experiments allow probing dynamics in the biologically relevant  $10^{-8}$ – $10^{-5}$  s timescale. The advent of fast magic-angle spinning (MAS) probes and high field magnets has recently<sup>6,7</sup> enabled the investigation of dynamics in samples ranging from microcrystals<sup>8,9,11</sup> to membrane proteins<sup>12–14</sup> and amyloid fibrils<sup>15</sup>. The wealth of crystallographic data available on  $\beta$ 2m and the inherent protection against aggregation makes MAS NMR a technique of choice for a comparative study of the dynamics of wt and D76N  $\beta$ 2m.

Here by combining ssNMR and ensemble modeling<sup>16</sup>, we shed light on the determinants of the aggregation propensity of D76N

$\beta$ 2m. Our results show the presence of distinct minor conformers exchanging with a major form, on sub-millisecond timescales, within the crystal lattice of both wt and D76N. The minor states are characterized by a local loss of  $\beta$ -strand structure that is enhanced in D76N. A detailed analysis of the interaction energies of the ensembles reveals that D76 is at the center of a network of stabilizing electrostatic interactions that are lost in the D76N mutant. These observations provide a molecular explanation for the pathological aggregation propensity of D76N, together with a model that accurately predicts the impact of further mutations on  $\beta$ 2m stability.

## Results

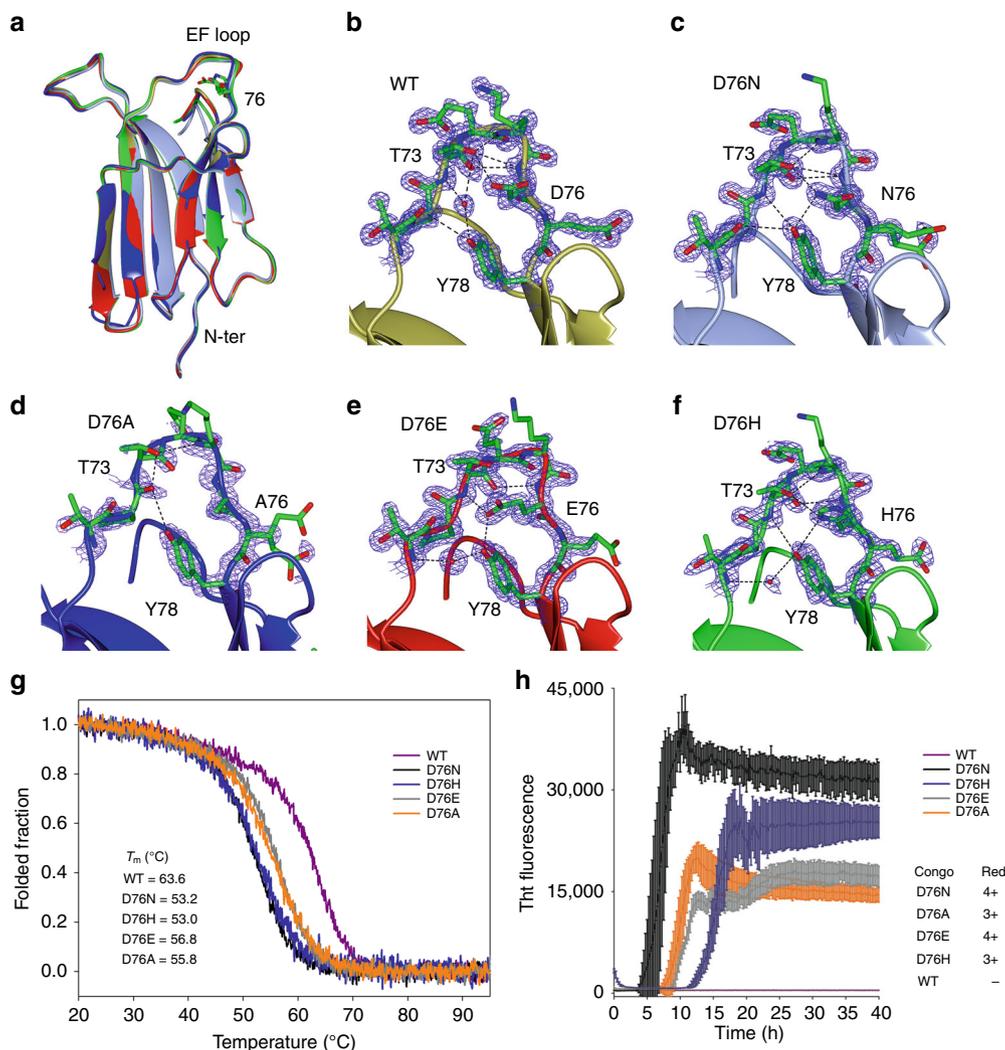
### Structural and biophysical studies on position 76 mutants.

Previous work suggests that D76 residue is key in determining the thermal stability and aggregation properties of  $\beta$ 2m. Indeed, among all D to N mutations along the  $\beta$ 2m sequence only D76N affects significantly  $\beta$ 2m stability and aggregation propensity<sup>17</sup>. To strengthen this hypothesis, four mutations were inserted at position 76: D76A, D76E, D76H, and D76K. While D76K resulted too unstable to be purified, the other three variants were structurally and biophysically characterized.

The crystal structures of D76A, D76E, and D76H were determined to high resolution (Fig. 1 and Supplementary Table 1). Analogously to the structure of D76N<sup>2</sup>, these structures do not show any major conformational change compared to wt  $\beta$ 2m (Fig. 1 and Supplementary Table 2). However, although the backbone of the EF loop is well superimposable between the structures of wt  $\beta$ 2m and the mutants (Supplementary Fig. 1), changing the residue in position 76 alters the H-bond network within the loop (Fig. 1), in particular in D76A and D76E. In keeping with this observation, the D76A structure where Ala76 side chain cannot establish any H-bond with neighboring residues, displays an electron density for the EF loop poorer than those of all other structures indicating an increased disorder (Fig. 1d). The last three C-terminal residues (residues 97–99) are the only stretch undergoing some rearrangement upon mutation at position 76. The side chain of Arg97 presents different conformations, and the strength of its interactions with the EF loop and to Trp95 varies in different mutants; residue Met99 interacts with Ser11 or with Arg12 (Supplementary Fig. 1). Such observations suggest a role for residue 76 in determining the interactions within the EF loop and with  $\beta$ 2m C-terminal region.

The thermal stability and aggregation propensity of these variants were evaluated (Fig. 1g, h). Similarly to the D76N mutant, all three mutants display a decreased stability and increased aggregation propensity, confirming the key role played by position 76 in determining  $\beta$ 2m biophysical properties. Nevertheless, despite the high resolution structural characterization of these different D76 mutants, the static crystal structures do not allow to provide a clear explanation for their distinct aggregation propensities.

**Fast protein dynamics.** Figure 2a, b show the  $^{15}\text{N}$ - $^1\text{H}$  correlation spectra of wt and D76N  $\beta$ 2m, corresponding to around 70% of the amide groups (see Supplementary Data 1 for the resonance assignments). The absent signals are associated with residues located in loops, which are subject to either large conformational disorder or extreme mobility, preventing their detection in CP-based experiments. Interestingly, signals from all residues in the BC-, EF- (encompassing position 76) and FG-loops are present in the spectra, showing that these portions of the molecule are sufficiently rigid and ordered, probably as a result of a network of strong interactions within the loops and/or with the neighboring side chains. Signals from residues 51–53 are missing



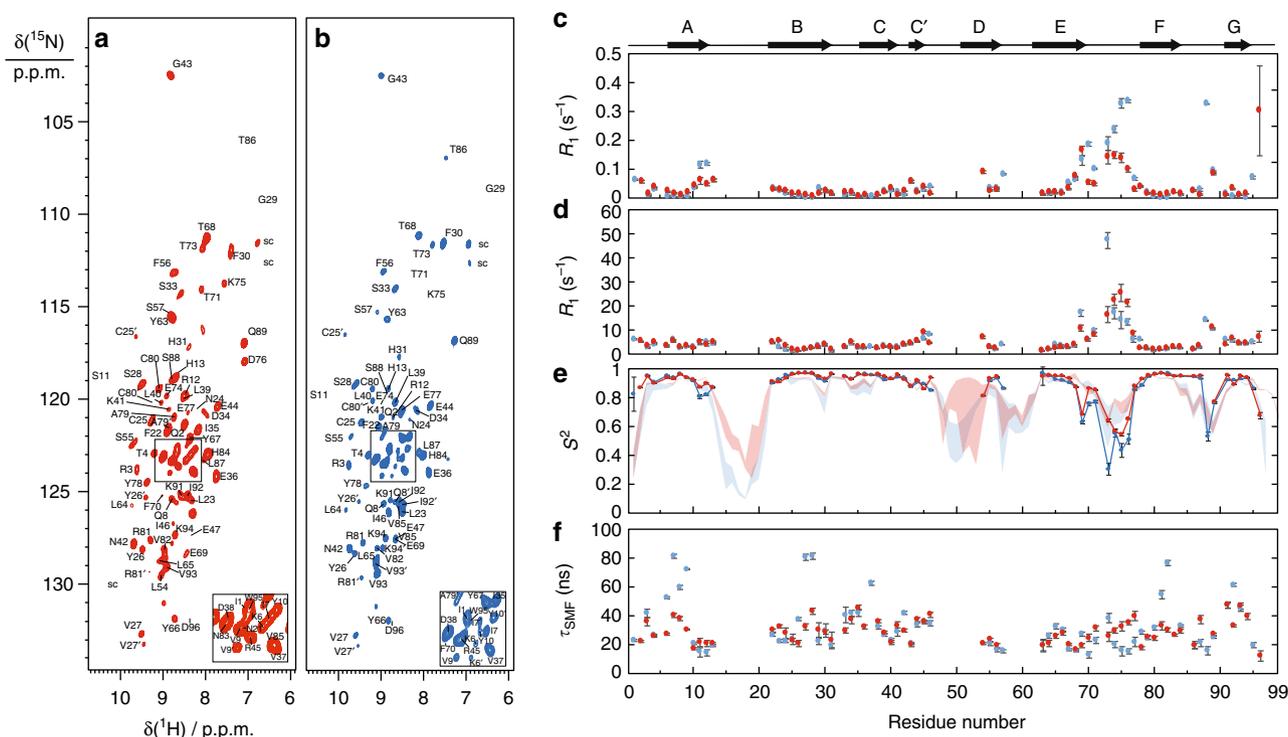
**Fig. 1** Structural and biophysical characterization of  $\beta 2m$  mutants at position 76. **a** Superposition of the structures of wt  $\beta 2m$  (2YXF, gold) and of the D76N (4FXL, light blue), D76A (blue), D76E (red), and D76H (green) variants. The residues in position 76 are modeled as cylinders. **b–f** Zoom-in of the structures shown in **a**. H-bonds are displayed as dash lines and electron density ( $\sigma = 1.2$ ) is clipped around the EF loop. **g** Normalized temperature ramps monitored by far-UV circular dichroism at 202 nm for wt  $\beta 2m$  and of all four 76 mutants. **h** Aggregation kinetics monitored by thioflavin T fluorescence and a Congo Red evaluation of the resulting aggregates (number are indicating the fluorescence intensity on a qualitative scale)

from the spectra, consistent with the fact that this portion of the molecule has been shown to be particularly dynamic in solution, and to display different secondary structures in different PDB structures<sup>18</sup>.

Site-specific  $^{15}\text{N}$  spin–lattice relaxation rates in the laboratory ( $R_1$ ) and rotating frame ( $R_{1\rho}$ )<sup>9</sup> were measured for the two molecules (pulse sequences and typical decay curves are shown in Supplementary Fig. 3), and are depicted in Fig. 1c, d (see also Supplementary Data 2). These rates are mainly determined by the fluctuation of the  $^{15}\text{N}$  chemical shift anisotropy (CSA) and of the  $^{15}\text{N}$ – $^1\text{H}$  dipole coupling. These data were interpreted using the simple model-free approach<sup>19</sup>, determining order parameters ( $S^2$ ) and correlation times ( $\tau$ ) for all visible amide groups, which respectively report on the amplitude and on the timescale of local motions (Fig. 2e, f and Supplementary Data 3). For most of the residues, the obtained order parameters are close to 0.9, and extremely similar for the two variants. These results indicate that the two proteins are overall very rigid, as expected for a beta-sandwich fold and that the rigidity of the beta-sheet core of the protein is not affected by the mutation. The residues in the EF-loop (containing residue 76) display lower order parameters. While this is in principle not surprising, as loops may feature

significant mobility also in the presence of crystal packing, it is intriguing to observe that the fitted order parameters are consistently lower in D76N than in wt  $\beta 2m$ , reflecting an increase in flexibility upon mutation. Small differences in the order parameters were also observed for residues 11 and 12 at the end of the A-strand. To better compare these data with crystallographic data, the crystal structures of wt and D76N were determined from diffraction data collected at room temperature (RT) rather than at 100 K. The RT structures are highly comparable in terms of conformation and B factors with structures from frozen crystals (Supplementary Fig. 2 and Supplementary Table 2).

Molecular dynamics (MD) simulation offers a mechanistic description of fast protein dynamics<sup>20</sup>. We compare the experimental order parameters with those calculated from two 50 ns MD runs (Fig. 2e) of both wt and D76N  $\beta 2m$  in a water box. Overall, the simulation reproduced the experimental sequence dependence with good accuracy. In the simulation, the AB loop and the C-terminal region are predicted to be highly flexible, more so in the D76N mutant than in the wt, in agreement with the fact that these regions are not visible in CP-based experiments. The high-quality reproduction of dynamics



**Fig. 2** Detection of fast local motions by  $^{15}\text{N}$   $R_1$  and  $R_{1\rho}$  relaxation experiments.  $^{15}\text{N}$ - $^1\text{H}$  CP-based HSQC of wt- (a) and D76N- (b) [ $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ]  $\beta$ -2 microglobulin with 100% reprotonation at the exchangeable sites recorded on an 800 MHz spectrometer at 60 kHz MAS rate (283 K, pH 5.5). All resonances were assigned using a set of experiments detailed in ref. 26. The color code (red for wt and blue for D76N  $\beta$ 2m) is used for all the figures presented here. Site-specific  $^{15}\text{N}$   $R_1$  (c) and  $R_{1\rho}$  (d) relaxation rates of the two proteins were used to calculate the order parameters  $S^2$  (e) and the correlation time (f) with the model-free approach. The results of a 50 ns MD trajectory are plotted together with the experimental results in E (red and blue bundles).  $\beta$ 2m secondary structures are plotted on top of the panels. Experimental error bars (s.d.) were obtained from 1000 Monte Carlo simulations

observed throughout the two proteins experimentally and via simulation suggests that the observed motions are intrinsic to the protein and are essentially independent of immediate environment (solution and crystalline forms). It is noteworthy that the timescale of the motions detected here are in the tens of nanoseconds, which falls in the time range that is especially difficult to detect by spin relaxation in solution due to protein rotational diffusion. Reduced flexibility is found in the EF loop in both D76N and wt, possibly due to additional slower motions that are not captured by the 50 ns simulations.

**Slow protein dynamics.** The contribution of conformational exchange on microsecond to millisecond timescales to transverse spin relaxation can be characterized by measuring  $^{15}\text{N}$   $R_{1\rho}$  rates as a function of the amplitude of the spin-lock RF field in so called relaxation dispersion experiments.  $R_{1\rho}$  was measured for all residues of wt and D76N  $\beta$ 2m, using  $^{15}\text{N}$  spin-lock RF amplitudes ranging from 3 to 20 kHz. For several residues, mainly located within the  $\beta$ -sheet core, the relaxation rates are independent of the applied RF field. Interestingly, however, the rates of a non-negligible fraction of the signals feature ample dispersion profiles as the RF field is varied (Supplementary Fig. 4). Figure 3a shows selected dispersion profiles for residues located at the end of the A-strand and on the BC- and EF-loops. These data are indicative of conformational exchange in the microsecond to millisecond range<sup>21</sup>.

Figure 3b reports the amplitude of the dispersion of the  $R_{1\rho}$  rates ( $\Delta R_{\text{ex}}$ ). For the wt, the residues showing the highest dispersion are located in the BC-loop, comprising Pro32, as well as residues located in the nearby N-terminus of the protein. These residues display a very similar dispersion also in the D76N

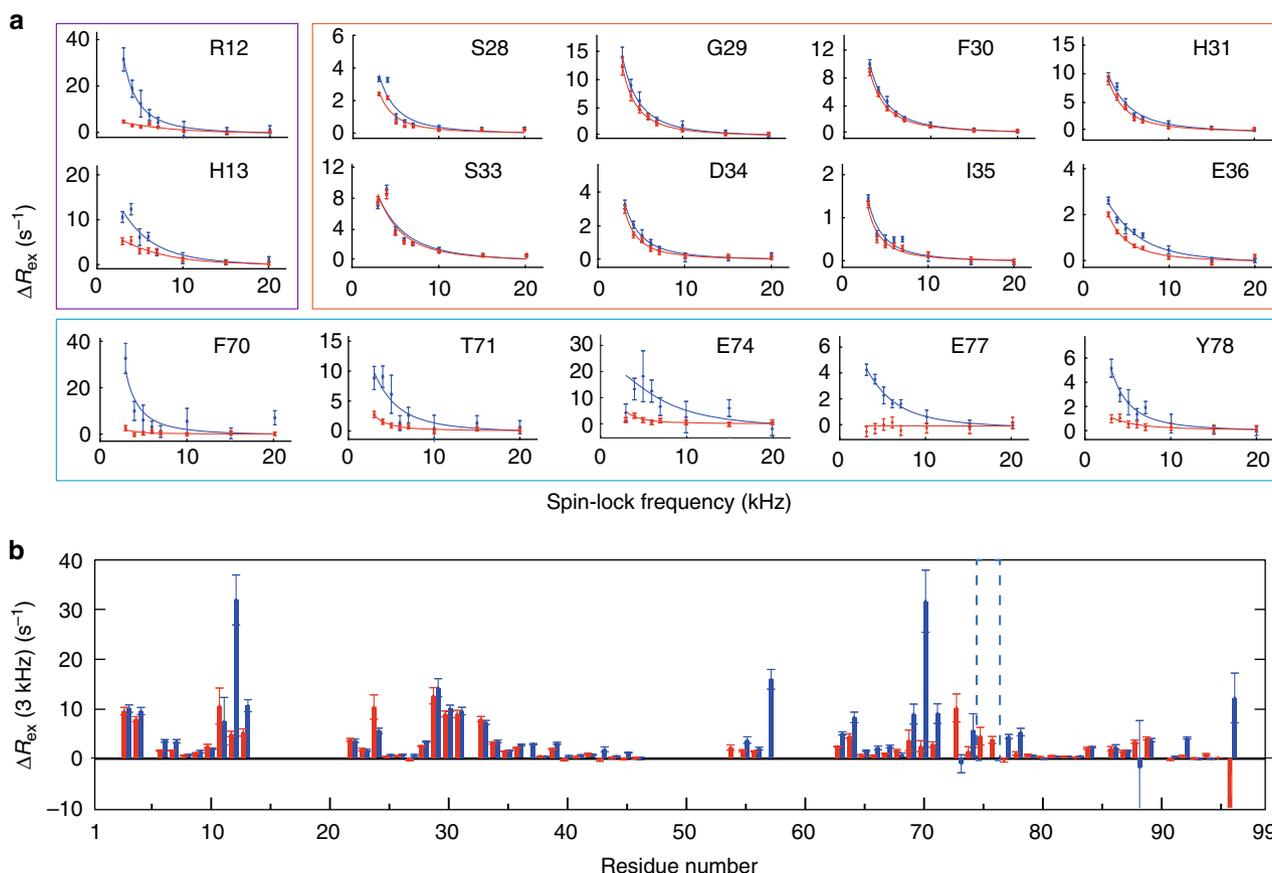
mutant, as it could be expected given the large distance with respect to the mutation site. However, in D76N, major changes are clearly observed in the EF-loop, as well as for the residues at the end of the A-strand. In the case of residues 75–76 in D76N,  $^{15}\text{N}$   $R_{1\rho}$  rates were so large for low spin-lock fields that the signals disappeared within the shortest relaxation delay, indicating extensive exchange contribution for these two residues.

These results reveal the presence of significant conformational exchange, consistent with two global processes. A first mode, common to wt and to D76N variant, involves the residues in spatial proximity to Pro32. A second mode affects the EF-loop and remote residues at the end of the A-strand, and produces a clear experimental signature only for D76N  $\beta$ 2m.

#### Replica-averaged metadynamics (RAM) ensemble simulations.

In order to gain further insights into the conformational exchange process and its consequences for the amyloid fibril formation, RAM simulations were performed<sup>22,23</sup>. RAM is used to integrate experimental data, in this case ssNMR chemical shifts, in molecular force fields to accurately explore the conformational space over timescales longer than those accessible during conventional MD runs.

On the basis of  $^1\text{H}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$ , and  $^{13}\text{C}'$  chemical shifts, converged (Supplementary Figs. 5 and 6) ensembles of structures were calculated for both the wt and the D76N mutant. Free energy surfaces were generated (Fig. 4a, b) as a function of the side-chain rotamer distribution (AlphaBeta collective variable (CV), AB, see Methods) and the antiparallel  $\beta$ -structure content (AntiBetaRMSD CV, anti- $\beta$ , see Methods). In both cases, the calculated structures clustered around two bundles, describing a major populated substate close to the crystal structure (less than



**Fig. 3** Conformational exchange of  $\beta 2m$  probed by relaxation dispersion experiments. Selected examples of  $^{15}N$   $R_{1\rho}$  relaxation dispersion curves for wt (red) and D76N (blue)  $\beta 2m$ . **a** The purple, orange, and blue rectangles show respectively residues located at the end of the A-strand, in the CD-loop and EF-loop. Best Lorentzian fits are depicted with solid lines. The difference between  $^{15}N$   $R_{1\rho}$  rates at 3 kHz and the fitted values at 20 kHz spin-lock field are depicted versus the sequence in **b**. Error bars (s.d.) were obtained from 1000 Monte Carlo simulations

2 Å full backbone average RMSD) and a minor populated substate, which features significant differences in secondary structural content (Fig. 4c, e) and chemical shifts differences qualitatively in agreement with those measured (Supplementary Fig. 5). The major substate (85 and 95% populated for the wt and D76N, respectively) is characterized by a high  $\beta$ -sheet content ( $\sim 48\%$  for both ensembles) while the minor substates are characterized by a significantly lower  $\beta$ -sheet content ( $\sim 38$  and  $30\%$  for wt and D76N, respectively).

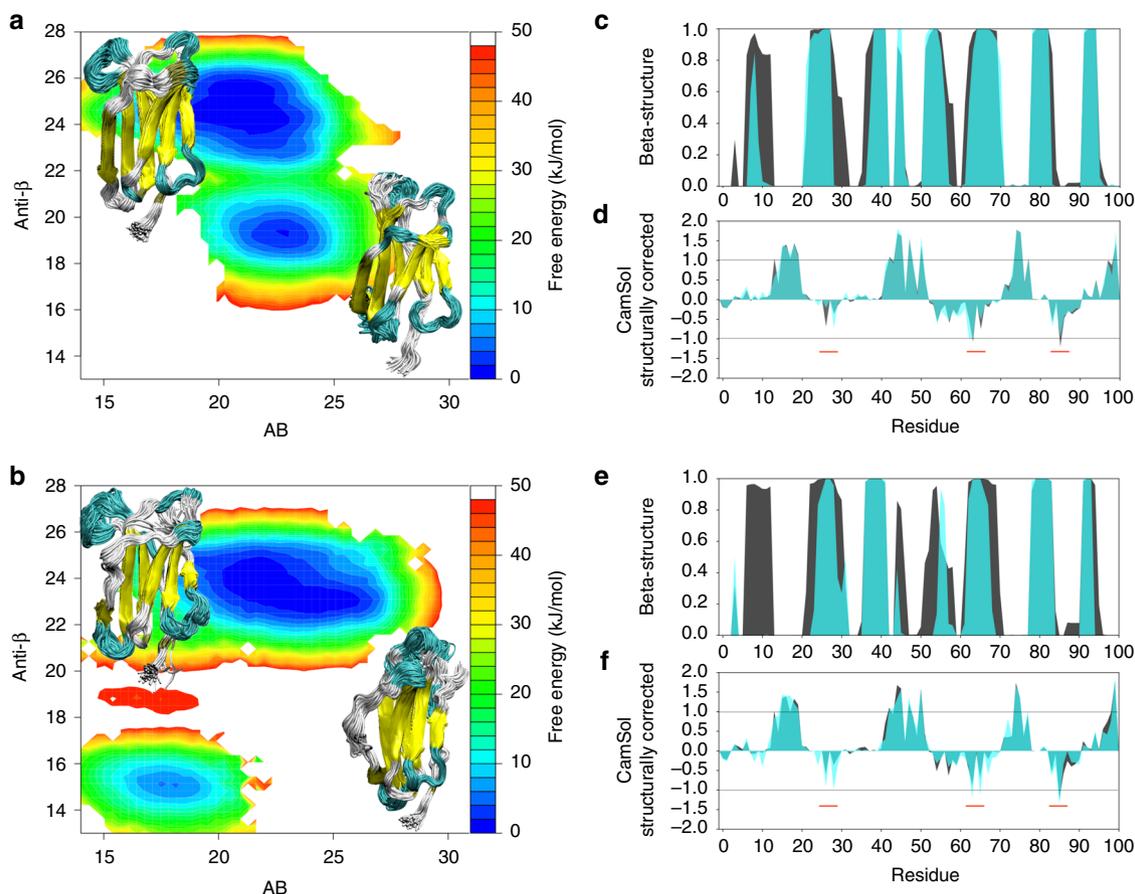
An analysis of the predicted solubility of the two ensembles performed with CamSol<sup>24</sup> shows that the D76N mutation in the sequence causes a small reduction of the solubility for residues 74–77 (Supplementary Fig. 7). The solubility profiles corrected for the structure and averaged over the ensembles show more distributed differences (Fig. 4d, f, black shades), with a general trend of reduced solubility for the D76N mutation but with no significant changes. The same analysis performed instead only on the minor substates (Fig. 4d, f, light blue shades), reveals a loss of solubility for the D76N variant. In particular, the number of residues below the aggregation prone threshold increases from one to four for the D76N minor substate and all the aggregation-prone sites belong to regions (around residues 65 and 85) that have been recently shown to be crucial for modulating the aggregation properties of the protein<sup>25</sup>. Comparison of the secondary structure content and the solubility profiles in the minor substates shows how the larger loss of beta structure in the A-strand for the D76N variant with respect to the wt corresponds to a decrease of solubility for B-strand region that is left more exposed to the solvent.

**Interaction network analysis.** The RAM results provide a molecular framework for understanding the dramatic increase in the aggregation propensity of the D76N mutant, but do not explain why the mutation is also reducing the overall stability of the protein.

In order to understand the energetic contribution of the D76 mutation, we calculated the nonbonded interaction energies for each pair of residues, averaged over the ensemble, thus resulting in two interaction matrices. By subtracting the wt matrix from the D76N matrix, we identified the largest energy gap in the distribution of the elements. We then performed a network analysis based only on the residue pairs for which the energy difference was larger than the gap, i.e., on residue pairs experiencing large differences in their interaction energy between the two ensembles. Figure 5a shows the result of this network analysis in the form of a graph.

Not surprisingly, it is evident that the largest differences in nonbonded interaction energies between wt and D76N are related to the interactions of residue 76 with the neighboring residues (Fig. 5a). Mutation of D76 to N weakens a number of interactions with charged residues. In particular a salt-bridge with K41 is lost and the interaction with the R97 side chain that mediates an interaction between the C- and the N-termini of the protein disappears. Overall this explains the higher loss of  $\beta$  structure at the N- and C-termini in D76N minor substate, thus linking the lower fold stability with the increased aggregation propensity in such minor substate.

The analysis of the ensembles suggests that D76 is in a key position with respect to the electrostatic interactions among



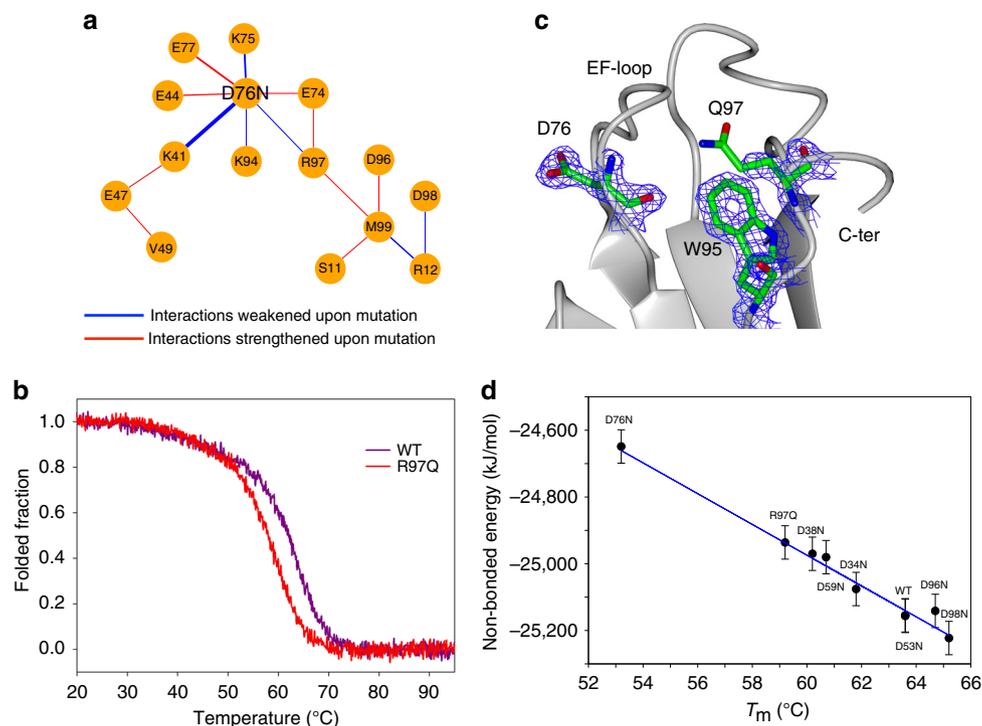
**Fig. 4** Conformational ensembles for native  $\beta 2m$  variants. Free energies surfaces (kJ/mol), secondary structure content and solubility for the wt (upper panels) and D76N (lower panels) ensembles. **a, b** Free energy surfaces for the wt and D76N variant, respectively. The surfaces are shown as a function of the side chains rotameric state, AB, and the antiparallel  $\beta$ -structure content (anti- $\beta$ ) with a set of representative conformations for the major and minor substates. **c, e** Beta structure content per residue (within each ensemble) for the wt and D76N variants, respectively. The major state is shown in black, while the minor in cyan. **d, f** CamSol solubility scores for wt and D76N, respectively. Scores greater than 1 suggest a highly soluble site while a score lower than  $-1$  suggest an aggregation-prone site. The major state is shown in black while the minor in cyan

surface residues. In order to confirm the role of the long-range interaction between R97 and D76, the R97Q mutant was prepared: in keeping with the above considerations, the mutant protein presents a significantly lower  $T_m$  value and intriguingly Q97 side chain is disordered in the structure (Fig. 5b, c). Finally, to further strengthen the hypothesis about the importance of surface electrostatic interactions centered around D76 in determining the stability of  $\beta 2m$ , the nonbonded energy was recalculated from the RAM simulation of the wt by modifying one-by-one the charge state for all the D-N mutants previously reported<sup>17</sup>, as well as for R97Q. The resulting energies were then compared with the corresponding melting temperatures (Fig. 5d) showing a remarkable correlation. Particularly relevant is the agreement between the prediction of the increased stability for the D98N mutant and the effect of the R97 mutation: this shows an energy increase higher than any of the D-N mutations but lower than the N76 mutant in agreement with the hypothesis that the interaction between the EF loop and the C-terminal residues have an important role for the stability of the protein.

**Structural investigation of D76N  $\beta 2m$  fibrils by ssNMR.** In order to link our observations to the mechanism of aggregation, we investigated amyloid fibrils grown *in vitro* by ssNMR. Resonance assignment was performed using the same  $^1H$ -detected correlation experiments as for the resonance assignment of the

microcrystalline form ref. <sup>26</sup> (Fig. 6a). Sequential resonance assignment was obtained for 53 residues, suggesting that this portion of the molecule corresponds to the most rigid and ordered core of the fibrils. The secondary structure of  $\beta 2m$  in the fibrils was determined with TALOS-N software<sup>27</sup> on the basis of  $^1H^N$ ,  $^{13}CA$ ,  $^{13}CB$ ,  $^{13}CO$ , and  $^{15}N$  chemical shifts (Fig. 6b, c). Five  $\beta$ -strands are predicted to be present in the fibrils. Interestingly, the native B-strand is shortened at the N-terminal side upon fibril formation. Residues corresponding to the C'-strand are not observed, indicating disorder or substantial backbone flexibility. The D-strand is extended toward the N-terminus, the E-strand is subject to an important rearrangement and the EF-loop is involved in  $\beta$  structure. The  $^{13}CA$  and  $^{13}CB$  chemical shifts of C25 (55.6 and 39.3 p.p.m., respectively) are indicative of an oxidized state as in the native monomer<sup>28</sup>. This is in agreement with previous data showing that the disulfide bond is necessary for the growth of wt  $\beta 2m$  fibrils<sup>29</sup>.

We next aimed at characterizing the tertiary/quaternary structure, defined as the arrangement of the intermolecular  $\beta$ -strands in fibrils<sup>30</sup>. We thus performed  $^1H$ - $^1H$  through-space correlations<sup>31</sup>, using radio frequency-driven recoupling mixing<sup>32</sup>. These experiments were acquired as 3D or 4D datasets, where the contacts between amide protons were encoded with the chemical shifts of the  $^1H$  and/or  $^{15}N$  atoms of starting and/or landing amide groups. These experiments revealed contacts between residues spaced in sequence by less than five residues. This



**Fig. 5** Perturbation of the energy network upon D76N mutation. **a** The network shows the interactions between couples of amino acids that are strengthened (red) or weakened (blue) upon mutation as calculated over the two ensembles of structures. **b** Temperature ramp of wt and R97Q variants monitored by far-UV ( $T_m$ : WT = 63.6 °C, R97Q = 59.2 °C). **c** Crystal structure of the R97Q mutant. Electron density (2FoFc) at 1.2  $\sigma$  is clipped around residue D76, W95, and Q97 shown as cylinder models. The lack of electron density for Q97 side chain indicates its high flexibility. **d** Correlation between thermal stabilities for multiple  $\beta 2m$  variants and the predicted change in energy obtained from the wild type ensembles by removing (or inverting) the side chain charges. Error bars (s.e.m.) were obtained through a block-averaging procedure

pattern is strongly indicative of a parallel, in register  $\beta$ -sheet intermolecular arrangement, where each residue faces its homologous residue in a neighboring molecule.

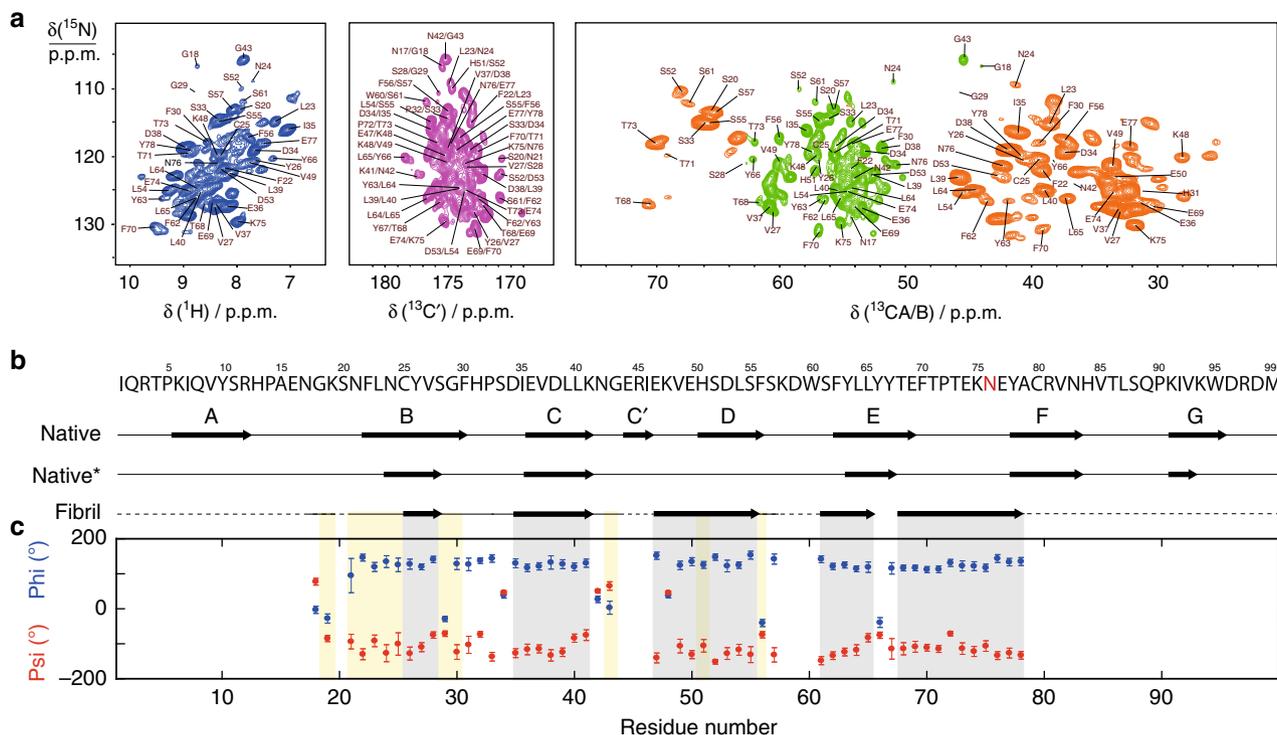
## Discussion

In 2012 the  $\beta 2m$  D76N mutant was reported to be responsible for an aggressive form of systemic amyloidosis with large fibrillar deposits in internal organs. Exceptional D76N aggregation propensity was also verified in vitro: this variant abundantly aggregates under native conditions in the absence of seeds<sup>2</sup>. While in principle it is unexpected that a D–N mutation on an exposed loop should trigger such a dramatic effect, the D76N mutant also displays a dramatic decrease in thermodynamic stability<sup>4</sup>. In contrast, the mutation of one Asp to Asn—with the consequent loss of a negative charge—in different parts of the protein did not result in any marked effects in terms of fold stability or aggregation propensity<sup>2,17</sup>.

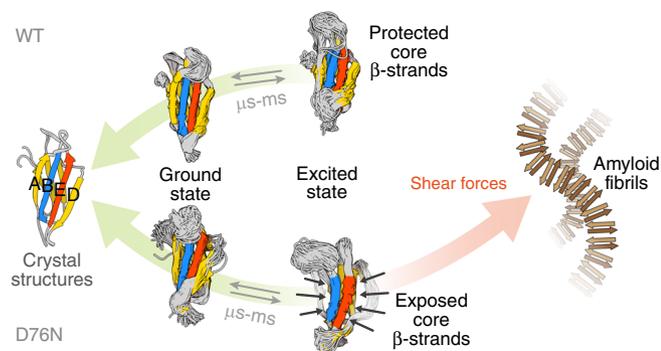
A saturation site mutagenesis (D76 to E, H, and A) allowed us to show here how position 76 is indeed crucial in the  $\beta 2m$  fold, any mutation at this position results in a notable decrease in thermal stability and a remarkable increase in aggregation propensity as observed for D76N. Intriguingly, the crystal structures suggest that D76 is an important node in a network of H-bond within the EF loop. We then used ssNMR to detect the change in dynamics upon mutation. One of the advantages of ssNMR, specific to the particular case of aggregation-prone proteins, consists in studying protein dynamics in crystallo, where the crystal lattice plays a “shielding role” preserving an overall native conformation. We are therefore able to exclusively probe intrinsic dynamics of individual protein molecules and decouple them from intermolecular interactions.

A comparison of fast (sub-microsecond) timescale dynamics measured using relaxation reveals similar motional characteristics in the wt and D76N mutant, with slightly higher levels of flexibility in the mutant. The position and amplitude of the dynamic modes observed in the crystalline lattice are well reproduced by fully solvated MD simulation, suggesting that the observed dynamics are intrinsic to the protein. Remarkably, ssNMR <sup>15</sup>N  $R_{1\rho}$  relaxation dispersion experiments in crystalline  $\beta 2m$ <sup>11,33–35</sup>, reveal the existence of excited states that exchange with the ground state on the micro to millisecond timescale for both the wt and the D76N variant. The D76N excited state displays more extensive exchange phenomena than wt  $\beta 2m$ , with seven sites exhibiting markedly high relaxation dispersion effects. Such sites are almost all located in the protein region neighboring the mutation: in the EF-loop, in the E-strand and at the end of the A-strand.

In keeping with these observations, ensembles resulting from RAM simulations showed that  $\beta 2m$  displays two distinct conformational states in both mutant and wt. As sketched in Fig. 7, while globally wt  $\beta 2m$  excited state resembles very closely the ground state, D76N populates a conformational ensemble, which is less structured. In the excited state the D and the N-terminal A-strands lose their beta structure and the C-terminus is partially detached from protein core. As a result, the aggregation-prone core strands (B, E, and F) lose the protection of the aggregation resistant edge strands (A, C, D, and G). Remarkably all the data indicate an increased level of dynamics in D76N N-terminal region that is reminiscent of the  $\Delta N6$   $\beta 2m$  mutant, which is known for a strong aggregation propensity<sup>36,37</sup>. While our data clarify the differences in aggregation propensity between D76N and wt, which likely yield to different aggregation behaviors both in vitro and in vivo,



**Fig. 6** Structural characterization of D76N fibrils. **a** Assigned  $^1\text{H}$ - $^{15}\text{N}$  2D projection of (H)CANH (dark blue) correlation spectrum and  $^{13}\text{C}$ - $^{15}\text{N}$  2D projections of (H)CONH (purple), (HCA)CB(CA)NH (orange) and (H)CANH (green) correlation spectra of [2H,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ] D76N  $\beta_2\text{m}$  fibrils with 100% reprotonation at the exchangeable sites (283 K, pH 7). Spectra were recorded on a 1 GHz spectrometer at 60 kHz MAS rate. **b** Secondary structure prediction of D76N  $\beta_2\text{m}$  fibrils by TALOS-N 39 (black arrows correspond to  $\beta$ -strands, straight lines to loops and dashed lines to an absence of prediction). The  $\beta$ -strands of the native fold are depicted above the primary sequence. The secondary structures of fibrillar wt  $\beta_2\text{m}$ <sup>40</sup> are also shown for comparison. **c** Backbone dihedral angles phi and psi predicted by TALOS-N. The gray and yellow boxes show TALOS-N for  $\beta$ -strands and ambiguous predictions, respectively. Error bars (s.d.) are based on the 10 best TALOS-N database matches



**Fig. 7** Proposed mechanism for the first steps of D76N aggregation. Under native conditions both wt and D76N are in equilibrium between low and high-energy conformations. While wt high-energy conformation has a contained aggregation propensity, the D76N excited state is less structured, highly dynamic, and exposes highly aggregation-prone regions shown in orange. Thus, modest energies such as shear forces present in the circulation system are sufficient to further unfold D76N  $\beta_2\text{m}$  termini (model) destabilizing completely  $\beta_2\text{m}$  fold and leading to amyloid formation

however, they do not provide specific insights on the aggregation pathways for the two  $\beta_2\text{m}$  variants.

Furthermore, the conformational ensembles suggest a link between the differences in the excited states and the differences in the thermal stabilities where the mutation in position 76 breaks a large network of electrostatic interactions distributed over a large number of residues including the C- and the N-termini. RAM simulations indicate that the hydrophobic core around the

evolutionary conserved W95 is not optimally shielded from the solvent. A specifically designed R97Q mutant next to the C-terminus allowed us to further test this hypothesis showing that the weakening of the interaction between positions 76 and 97 significantly decreases  $\beta_2\text{m}$  thermal stability.

Moving perspective, we complemented the analysis of the native state of D76N  $\beta_2\text{m}$  with ssNMR characterization of amyloid fibrils grown in vitro. Interestingly, as shown in Fig. 6b, we find a progressive evolution of structural elements between the starting, native conformation derived from the crystal structure, the ensemble representing the excited state derived by RAM simulations, and the final fibrillar state. Notably, the A and G edge strands of  $\beta_2\text{m}$ , which are destabilized by the point mutation, are not part of the cross  $\beta$ -structure. Additionally, the N-terminal portion of the B-strand loses secondary structure both in the excited state (native\*) and in the fibrils. While these data do not represent an evidence for a particular aggregation pathway, we note that the rearrangement of the secondary structure elements in the core of the fibrils may be facilitated by the loss of structure of the corresponding native  $\beta$ -elements. This is particularly evident for the region encompassing residues 44–56 where the C' and D strands are absent in the excited state while they appear to form a unique strand in the fibrillar state. A similar observation can be made for the E-strand and EF-loop region. The loss of interaction between EF- and C'D-loops, as evidenced by the network analysis, may therefore be related to their reorganization in  $\beta$ -structures within the fibrils.

The secondary structure of wt  $\beta_2\text{m}$  fibrils has previously been characterized by  $^{13}\text{C}$ -detected ssNMR<sup>38–40</sup>. We note common structural and dynamical features in both wt and D76N fibrils, namely the position of the rigid core, a parallel in register beta

arrangement, the preservation of the disulfide bridge and flexible termini, but also differences can be observed in the fine details of the secondary structures, namely the positions of the  $\beta$ -strands. This may indicate comparable fibril architecture even in the presence of likely different aggregation pathways.

In summary, in the present work, by using a combination of X-ray crystallography, ssNMR and MD simulations, we have been able to determine the impact of the highly pathogenic D76N mutation on the native state of  $\beta$ 2m. Thermodynamic stability of the native fold is lowered due to the weakening of a set of electrostatic interactions and to the diminished interaction between the C-terminal residues and the  $\beta$ 2m hydrophobic core. Enhanced aggregation propensity in D76N appears to be correlated with the existence of a high-energy state characterized by a loss of structure in the protective edge  $\beta$ -strands. Previous observations that low energetic contributions, e.g., the friction in extracellular space in vivo or the shear forces at the interface air/liquid in vitro are sufficient to trigger amyloid aggregation<sup>2,4</sup>, are well rationalized by the presence of an amyloid-competent state under native conditions. In conclusion, our observations provide a molecular explanation for the biophysical properties leading to D76N pathogenicity, while showing the effectiveness of our integrated approach to investigate the molecular bases of amyloid aggregation propensity, a method that may well be applied to other amyloidogenic folded proteins.

## Methods

**Protein production.**  $\beta$ 2m mutants (D76E, D76A, D76H, D76K, and R97Q) were produced using the Phusion Site-Directed Mutagenesis Kit (lifetechnologies), following the manufacturer protocol. Forward primers used for mutagenesis are available in Supplementary Table 3. Any attempt to obtain amount of the D76K mutant suitable for structural and biophysical characterization failed. All other variants together with wt and D76N proteins were instead expressed and purified using standard protocols<sup>36</sup>.

**Crystallization structure determination and analysis.** All the proteins were crystallized under the standard conditions as used for most of the  $\beta$ 2m monomeric variants: protein concentration 8–10 mg/ml, 21–27% w/v PEG 4000, 15% v/v glycerol, 0.2 M ammonium acetate, 0.1 M sodium acetate, pH 5.0–5.5.

Crystals were flash-frozen to 100 K in liquid nitrogen without supplementation of additional cryoprotectants. Diffraction data were collected at the European Synchrotron Radiation Facility (Grenoble France) at the ID29, ID23-2, and ID23-1 beamlines<sup>41–43</sup>.

Data collection at RT was performed as follows. Beta2-microglobulin crystals were mounted on a microRT Room Temperature kit (MiTeGen), sealed with a MicroRT polyester capillary, filled at the extremity with 5  $\mu$ l mother liquor. RT X-ray diffraction data were collected at beamline ID29 at ESRF using Pilatus 3 M detector in a shutterless mode at 0.9840 Å wavelength. Data collection strategy were calculated using BEST<sup>44</sup> from two images collected with a 50  $\times$  30  $\mu$ m<sup>2</sup> Gaussian beam and a flux limited to 5.58  $\times$  10<sup>10</sup> photons/s (corresponding to 5% of the full beam) to limit radiation damage. Strategy was calculated using a susceptibility factor of 40 to take into account the faster decay rate that occurs at RT. For the wt crystal, 880 images were collected with 0.02 s exposure time and 0.15° oscillation angle, with a flux of 1.12  $\times$  10<sup>10</sup> photons/s (corresponding to 1.35% of full beam). For the D76N mutant, 1800 images were collected with 0.02 s exposure time and 0.1° oscillation angle, with a flux of 8.30  $\times$  10<sup>9</sup> photons/s (1% of full beam).

Data were processed with MOSFLM<sup>45</sup> and either SCALA or AIMLESS<sup>46</sup> and structures solved by molecular replacement using the program PHASER<sup>47</sup> and wt  $\beta$ 2m as search model (pdb ID 2YXF). All the structures were refined with Phenix refine<sup>48</sup> and manual model building and structure analysis were performed with COOT<sup>49</sup>. CCP4MG was used for the preparation of the figures<sup>50</sup>.

**Thermal denaturation.** Thermal unfolding was monitored at 202 nm in the 20–95 °C range using a JASCO J-810 spectropolarimeter equipped with a Peltier device and a fluorescence detector (JASCO corporation, Tokyo, Japan). For all  $\beta$ 2m variants involved in this study ramps were carried out in 50 mM sodium phosphate, pH 7.4. Protein concentration was set to 0.1 mg/mL (cell path 0.1 cm), and the temperature increase was 50 °C/h.  $T_m$  values were determined as the minima of the first derivative of the unfolding profiles. Thermal unfolding experiments display high reproducibility.

**Aggregation propensity.** A volume of 100  $\mu$ l of recombinant variants D76N, D76A, D76H, D76E, and wt  $\beta$ 2m at 40  $\mu$ M (PBS pH 7.4) containing 10  $\mu$ M Thioflavin T (ThT) (SIGMA), were incubated at 37 °C in Costar 96-well black-wall plates closed with clear sealing film (4TITUDE) and were subjected to 900 rpm double-orbital shaking. ThT fluorescence was recorded every 15-min (BMG LABTECH FLUOstar Omega). Experiments have been performed in triplicate.

**Isotopically labeled sample preparation.** Isotopically labeled D76N and wt  $\beta$ 2m were expressed in triple labeled (<sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N) M9 medium (D2O, d<sub>5</sub>-<sup>13</sup>C-glucose, and <sup>15</sup>N-ammonium chloride). The purification and refolding step were performed as described above in non-deuterated water allowing proton back-exchange at every amide site. Crystals were grown using the sitting drop vapor diffusion technique (10 mg/mL protein concentration, crystallization solution: 22% w/v PEG, 20% v/v glycerol, 0.1 mol/L sodium acetate, pH 5.5). After 2 weeks, microcrystals of D76N and wt  $\beta$ 2m were filled into 1.3 mm NMR rotor by centrifugation (14,000 r.p.m.).

**NMR spectroscopy.** NMR experiments were conducted on 18.8 or 23.5 T narrow-bore Bruker Avance III spectrometers, corresponding respectively to <sup>1</sup>H Larmor frequency of 800 MHz and 1 GHz at 60 kHz MAS. Sample temperature was monitored by bulk-water chemical shift using the PEG line as a reference and set to 283 K.

Resonance assignment of D76N  $\beta$ 2m was performed with the same procedure that described<sup>51</sup> using the whole set of experiments ((H)CANH, (H)(CO)CA(CO)NH, (H)CONH, (H)CO(CA)NH, (H)(CA)CB(CA)NH, and (H)(CA)CB-(CA)(CO)NH).

The pulse sequences used for <sup>15</sup>N  $R_1$  and  $R_{1\rho}$  relaxation rates measurement are as described<sup>9</sup> (Supplementary Fig. 1). Relaxation delays were set to 0, 0.5, 1, 2.5, 5, 10, 20, and 40 s for <sup>15</sup>N  $R_1$  measurements and 1, 3, 8, 20, 45, 100, 180, and 300 ms for <sup>15</sup>N  $R_{1\rho}$  measurements. The <sup>15</sup>N  $R_{1\rho}$  relaxation dispersion curves were obtained with <sup>15</sup>N spin-lock frequency of 3, 4, 5, 6, 7, 10, 15, and 20 kHz. Examples of decay curves are shown in Supplementary Fig. 3.

The NMR data processing was done with Topspin, the resonance assignment with CARA and the analysis of the relaxation with Sparky.

**Relaxation data fitting.** All relaxation data fitting has been done using Matlab. Amide <sup>15</sup>N relaxation was assumed to be dominated by the dipolar coupling with its bonded <sup>1</sup>H and by its CSA. The relaxation rates can therefore be expressed as:

$$R_1 = \frac{d^2}{4} [J_0(\omega_H - \omega_N) + 3J_1(\omega_N) + 6J_2(\omega_H + \omega_N)] + \frac{\omega_N^2 \Delta\sigma^2}{3} J_1(\omega_N), \quad (1)$$

$$R_{1\rho} = \frac{d^2}{8} [4J_0(\omega_{SL}) + J_0(\omega_H - \omega_N) + 3J_1(\omega_N) + 6J_1(\omega_H) + 6J_2(\omega_H + \omega_N)] + \frac{\omega_N^2 \Delta\sigma^2}{18} [4J_0(\omega_{SL}) + 3J_1(\omega_N)], \quad (2)$$

where  $d = \frac{\mu_0 \gamma_H \gamma_N}{4\pi r_{NH}^3}$ ,  $\mu_0$ , the vacuum permittivity,  $\gamma_H$  and  $\gamma_N$ , respectively, the <sup>1</sup>H and <sup>15</sup>N gyromagnetic ratios,  $r_{NH}$  the NH bond distance (1.02 Å),  $\omega_H$  and  $\omega_N$ , respectively, the <sup>1</sup>H and <sup>15</sup>N Larmor-precession frequency,  $\Delta\sigma$  the <sup>15</sup>N CSA (170 p.p.m.), and  $\omega_{SL}$  the rf spin-lock frequency.

We applied the Lipari–Szabo model-free analysis. In this framework, the spectral density function  $J_m(\omega)$  is independent of the value of  $m$  and is given by:

$$J_{SMF}(\omega) = \frac{2}{5} \left[ \frac{(1 - S^2)\tau_{eff}}{1 + (\omega\tau_{eff})^2} \right], \quad (3)$$

where  $S^2$  is the order parameter and  $\tau_{eff}$  the effective model-free timescale. This approach does not account for the anisotropy. Intensity decay was therefore treated as a mono-exponential process.

The relaxation decay curves were fitted by minimizing the chi-square function:

$$\chi^2 = \frac{1}{N} \sum_{k=1}^N \frac{(I_{X,exp}(t_k) - I_{X,calc}(t_k))^2}{\sigma_{X,exp}^2}, \quad (4)$$

$X$  stands for <sup>15</sup>N  $R_1$  or <sup>15</sup>N  $R_{1\rho}$ ,  $N$  is the number of points of the curve, and  $\sigma$  is the experimental error calculated as the spectrum RMSD.

Error estimate was done with 1000 Monte Carlo simulations. In these simulations, synthetic datasets are produced by adding Gaussian random noise to the back-calculated decay curves. The error was then defined as the standard deviation of the ensemble of dynamical parameters obtained by fitting the synthetic datasets.

**MD simulations.** The simulation system consists of either wt or D76N  $\beta$ 2m in a dodecahedron box. Na<sup>+</sup> and Cl<sup>-</sup> ions were added to neutralize charges. Two 50-ns independent simulations were seeded from both the wt (PDB code 2YXF) and mutant (4FXL) crystal structures. GROMACS<sup>52</sup> was used to calculate the

trajectories. Prior to the production runs of 50 ns with an integration step of 2 fs in the NPT ensemble, energy minimization with the steepest-descent minimization algorithm is performed, followed by 2-ns equilibration steps in the NPT ensemble. Amber ff99SB-ILDN<sup>53</sup> is used as force field with TIP4P/2005<sup>54</sup> water model. Holonomic constraints are applied on all hydrogen-heavy atom bond terms to remove fast modes of oscillation. A cutoff of 10 Å is used for the Lennard-Jones and electrostatic interactions. Particle-mesh Ewald summation with a grid spacing of 16 Å is used to calculate long-range electrostatic interactions. A modified Berendsen thermostat with a damping constant of 0.1 ps is used to keep the temperature of the system at 298 K. A Berendsen barostat with a relaxation time of 2 ps is used to control the pressure at the target value of 1.01325 bar.

**RAM simulations.** The simulations were carried out using GROMACS<sup>52</sup> and PLUMED<sup>55</sup>. The system was described using the Amber03W<sup>56</sup> force field in explicit TIP4P05 water at 278 K. The starting conformations were taken from the 2YXF and 4FXL X-ray structures for the wt and D76N mutant, respectively. The structures were protonated and solvated with ~8200 water molecules in a dodecahedron box of ~260 nm<sup>3</sup> of volume. The RAM protocol was applied using chemical shifts as replica-averaged restraints<sup>54</sup> and bias-exchange metadynamics<sup>23,57,58</sup>. Four replicas of the system were simulated in parallel with a restraint applied on the average value of the back-calculated NMR chemical shifts with a force constant of 24 kJ/(mol p.p.m.<sup>2</sup>).

Each of the four replica is biased along one of the following four CVs: the antiparallel  $\beta$  content (the “anti- $\beta$ ” CV), the parallel  $\beta$ -sheet content (the “Para- $\beta$ ” CV), the AlphaBeta CV defined over all the  $\chi_1$  angles for the hydrophobic side-chains (the “AB” CV), and the AlphaBeta CV defined over all the  $\phi$  and  $\psi$  backbone dihedral angles of the protein (the “bbAB” CV). Definition of the CVs are available in the PLUMED manual and publication. Gaussians deposition was performed with  $\sigma$  values set to 0.1, 0.04, 0.16, and 0.25, for  $\beta$ , P $\beta$ , AB, and bbAB, respectively; an initial energy deposition rate of 0.125 kJ/mol/ps and a bias-factor of 8. Each replica has been run for a nominal time of 700 ns, with exchange trials every 50 ps.

The cumulative sampling of the four replicas was used to generate a free energy landscape as a function of the before mentioned CVs. A set of microstates is identified by dividing the four-dimensional CV-space into a homogeneous grid of small dimensional hypercubes and their free energy is obtained using a standard weighted histogram analysis<sup>59</sup>. Lower-dimensional FESes are obtained integrating out the CVs not showed.

**Data availability.** Coordinates and structure factors for D76E, D76A, D76H, R97Q, wt RT, and D76N RT have been deposited in the Protein Data Bank under accession codes 4RMU, 4RMW, 4RMV, 5CSG, 5CS7, and 5CSB, respectively. The datasets generated and/or analyzed during the current study are available from the corresponding authors upon reasonable request.

Received: 2 March 2018 Accepted: 29 March 2018

Published online: 25 April 2018

## References

- Chiti, F. & Dobson, C. M. Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **5**, 15–22 (2009).
- Valleix, S. et al. Hereditary systemic amyloidosis due to Asp76Asn variant beta2-microglobulin. *N. Engl. J. Med.* **366**, 2276–2283 (2012).
- Halabelian, L. et al. Class I major histocompatibility complex, the trojan horse for secretion of amyloidogenic beta2-microglobulin. *J. Biol. Chem.* **289**, 3318–3327 (2014).
- Mangione, P. P. et al. Structure, folding dynamics, and amyloidogenesis of D76N beta2-microglobulin: roles of shear flow, hydrophobic surfaces, and alpha-crystallin. *J. Biol. Chem.* **288**, 30917–30930 (2013).
- Jahn, T. R. & Radford, S. E. The Yin and Yang of protein folding. *FEBS J.* **272**, 5962–5970 (2005).
- Andreas, L. B., Le Marchand, T., Jaudzems, K. & Pintacuda, G. High-resolution proton-detected NMR of proteins at very fast MAS. *J. Magn. Reson.* **253**, 36–49 (2015).
- Bockmann, A., Ernst, M. & Meier, B. H. Spinning proteins, the faster, the better? *J. Magn. Reson.* **253**, 71–79 (2015).
- Chevelkov, V., Fink, U. & Reif, B. Quantitative analysis of backbone motion in proteins using MAS solid-state NMR spectroscopy. *J. Biomol. NMR* **45**, 197–206 (2009).
- Knight, M. J. et al. Structure and backbone dynamics of a microcrystalline metalloprotein by solid-state NMR. *Proc. Natl Acad. Sci. USA* **109**, 11095–11100 (2012).
- Lewandowski, J. R. et al. Measurement of site-specific C-13 spin-lattice relaxation in a crystalline Protein. *J. Am. Chem. Soc.* **132**, 8252 (2010).
- Tollinger, M., Sivertsen, A. C., Meier, B. H., Ernst, M. & Schanda, P. Site-resolved measurement of microsecond-to-millisecond conformational-exchange processes in proteins by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **134**, 14800–14807 (2012).
- McDermott, A. Structure and dynamics of membrane proteins by magic angle spinning solid-State NMR. *Annu. Rev. Biophys.* **38**, 385–403 (2009).
- Good, D. B. et al. Conformational dynamics of a seven transmembrane helical protein Anabaena Sensory Rhodopsin probed by solid-state NMR. *J. Am. Chem. Soc.* **136**, 2833–2842 (2014).
- Saurel, O. et al. Local and global dynamics in klebsiella pneumoniae outer membrane protein a in lipid bilayers probed at atomic resolution. *J. Am. Chem. Soc.* **139**, 1590–1597 (2017).
- Smith, A. A., Testori, E., Cadalbert, R., Meier, B. H. & Ernst, M. Characterization of fibril dynamics on three timescales by solid-state NMR. *J. Biomol. NMR* **65**, 171–191 (2016).
- Bonomi, M., Heller, G. T., Camilloni, C. & Vendruscolo, M. Principles of protein structural ensemble determination. *Curr. Opin. Struct. Biol.* **42**, 106–116 (2017).
- de Rosa, M. et al. Decoding the structural bases of D76N ss2-microglobulin high amyloidogenicity through crystallography and Asn-Scan mutagenesis. *PLoS ONE* **10**, e0144061 (2015).
- Esposito, G. et al. The controlling roles of Trp60 and Trp95 in beta2-microglobulin function, folding and amyloid aggregation properties. *J. Mol. Biol.* **378**, 887–897 (2008).
- Lipari, G. & Szabo, A. A model-free approach to the interpretation of Nmr relaxation in macromolecules. *Biophys. J.* **33**, A307–A307 (1981).
- Mollica, L. et al. Atomic-resolution structural dynamics in crystalline proteins from NMR and molecular simulation. *J. Phys. Chem. Lett.* **3**, 3657–3662 (2012).
- Palmer, A. G. 3rd & Massi, F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chem. Rev.* **106**, 1700–1719 (2006).
- Camilloni, C., Cavalli, A. & Vendruscolo, M. Replica-averaged metadynamics. *J. Chem. Theory Comput.* **9**, 5610–5617 (2013).
- Camilloni, C., Robustelli, P., De Simone, A., Cavalli, A. & Vendruscolo, M. Characterization of the conformational equilibrium between the two major substates of RNase A using NMR chemical shifts. *J. Am. Chem. Soc.* **134**, 3968–3971 (2012).
- Sormanni, P., Aprile, F. A. & Vendruscolo, M. The CamSol method of rational design of protein mutants with enhanced solubility. *J. Mol. Biol.* **427**, 478–490 (2015).
- Camilloni, C. et al. Rational design of mutations that change the aggregation rate of a protein while maintaining its native structure and stability. *Sci. Rep.* **6**, 25559 (2016).
- Barbet-Massin, E. et al. Rapid proton-detected NMR assignment for proteins with fast magic angle spinning. *J. Am. Chem. Soc.* **136**, 12489–12497 (2014).
- Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
- Sharma, D. & Rajarathnam, K. <sup>13</sup>C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* **18**, 165–171 (2000).
- Sarell, C. J. et al. Expanding the repertoire of amyloid polymorphs by co-polymerization of related protein precursors. *J. Biol. Chem.* **288**, 7327–7337 (2013).
- Petkova, A. T., Yau, W. M. & Tycko, R. Experimental constraints on quaternary structure in Alzheimer’s beta-amyloid fibrils. *Biochemistry* **45**, 498–512 (2006).
- Knight, M. J. et al. Fast resonance assignment and fold determination of human superoxide dismutase by high-resolution proton-detected solid-state MAS NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **50**, 11697–11701 (2011).
- Bennett, A. E., Ok, J. H., Griffin, R. G. & Vega, S. Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* **96**, 8624–8627 (1992).
- Krushelnitsky, A., Zinkevich, T., Reichert, D., Chevelkov, V. & Reif, B. Microsecond time scale mobility in a solid protein as studied by the <sup>15</sup>N R (rho) site-specific NMR relaxation rates. *J. Am. Chem. Soc.* **132**, 11850–11853 (2010).
- Lewandowski, J. R., Sass, H. J., Grzesiek, S., Blackledge, M. & Emsley, L. Site-specific measurement of slow motions in proteins. *J. Am. Chem. Soc.* **133**, 16762–16765 (2011).
- Kurauskas, V. et al. Slow conformational exchange and overall rocking motion in ubiquitin protein crystals. *Nat. Commun.* **8**, 145 (2017).
- Esposito, G. et al. Removal of the N-terminal hexapeptide from human beta2-microglobulin facilitates protein aggregation and fibril formation. *Protein Sci.* **9**, 831–845 (2000).

37. Karamanos, T. K., Kalverda, A. P., Thompson, G. S. & Radford, S. E. Visualization of transient protein-protein interactions that promote or inhibit amyloid assembly. *Mol. Cell* **55**, 214–226 (2014).
38. Barbet-Massin, E. et al. Fibrillar vs crystalline full-length beta-2-microglobulin studied by high-resolution solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **132**, 5556–5557 (2010).
39. Debelouchina, G. T., Platt, G. W., Bayro, M. J., Radford, S. E. & Griffin, R. G. Magic angle spinning NMR analysis of beta(2)-microglobulin amyloid fibrils in two distinct morphologies. *J. Am. Chem. Soc.* **132**, 10414–10423 (2010).
40. Su, Y. C. et al. Secondary structure in the core of amyloid fibrils formed from human beta(2)m and its truncated variant delta N6. *J. Am. Chem. Soc.* **136**, 6313–6325 (2014).
41. de Sanctis, D. et al. ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *J. Synchrotron Radiat.* **19**, 455–461 (2012).
42. Flot, D. et al. The ID23-2 structural biology microfocus beamline at the ESRF. *J. Synchrotron Radiat.* **17**, 107–118 (2010).
43. Nurizzo, D. et al. The ID23-1 structural biology beamline at the ESRF. *J. Synchrotron Radiat.* **13**, 227–238 (2006).
44. Bourenkov, G. P. & Popov, A. N. A quantitative approach to data-collection strategies. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 58–64 (2006).
45. Leslie, A. G. W. in *Joint CCP4+ESF-EACMB Newsletter on Protein Crystallography Number 27* (eds Wolf, W. M. & Wilson, K. S.) 30–31 (SERC, 1992).
46. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
47. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
48. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
49. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
50. McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 386–394 (2011).
51. Barbet-Massin, E. et al. Rapid proton-detected NMR assignment for proteins with fast magic angle spinning. *J. Am. Chem. Soc.* **136**, 12489–12497 (2014).
52. Abraham, M. J. et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
53. Lindorff-Larsen, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).
54. Abascal, J. L. & Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **123**, 234505 (2005).
55. Tribello, G. A., Bonomi, F., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* **182**, 604–613 (2014).
56. Best, R. B. & Mittal, J. Protein simulations with an optimized water model: cooperative helix formation and temperature-induced unfolded state collapse. *J. Phys. Chem. B* **114**, 14916–14923 (2010).
57. Camilloni, C. & Vendruscolo, M. Statistical mechanics of the denatured state of a protein using replica-averaged metadynamics. *J. Am. Chem. Soc.* **136**, 8982–8991 (2014).
58. Piana, S. & Laio, A. A bias-exchange approach to protein folding. *J. Phys. Chem. B* **111**, 4553–4559 (2007).
59. Marinelli, F., Pietrucci, F., Laio, A. & Piana, S. A kinetic model of trp-cage folding from multiple biased molecular dynamics simulations. *PLoS Comput. Biol.* **5**, e1000452 (2009).

## Acknowledgements

C.C. acknowledges the CINECA award under the ISCRA initiative, for the availability of high performance computing resources and support. We thank the members of the technical staff of the Institut des Sciences Analytiques for assistance with the NMR spectrometers. Financial support from the CNRS (IR-RMN FR3050) and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC-2015-CoG GA no. 648974) are gratefully acknowledged. L.B.A. was supported by a MC incoming fellowship (REA Grant agreement no. 624918 "MEM-MAS"). S.R. was funded by the Italian Ministry of University (Project FIRB RBFIR109EOS) and by Fondazione Cariplo (Cariplo Giovani Project 2016-0489).

## Author contributions

T.L.M., M.D.R., N.S., B.M.S., L.B.A., P.S., E.B.-M., A.B., R.P., C.S.M., D.D.S., and C.C. performed the experiments; M.Bol., L.E., V.B., M.Bla., C.C., G.P., and S.R. designed the study; T.L.M., C.C., G.P., and S.R. wrote the paper with contributions from all other authors.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-018-04078-y>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018