



Predicting stable binding modes from simulated dimers of the D76N mutant of β 2-microglobulin



Nuno F.B. Oliveira^{a,b,1}, Filipe E.P. Rodrigues^{a,b,1}, João N.M. Vitorino^{a,b}, Rui J.S. Loureiro^a, Patrícia F.N. Faísca^{a,c,*}, Miguel Machuqueiro^{a,b,*}

^a BioISI – Biosystems & Integrative Sciences Institute, Faculty of Sciences, University of Lisboa, Campo Grande, C8 bdg, Lisboa 1749-016, Portugal

^b Department of Chemistry and Biochemistry, Faculty of Sciences, University of Lisbon, Lisboa 1749-016, Portugal

^c Department of Physics, Faculty of Sciences, University of Lisbon, Lisbon 1749-016, Portugal

ARTICLE INFO

Article history:

Received 14 July 2021

Received in revised form 2 September 2021

Accepted 2 September 2021

Available online 08 September 2021

2000 MSC:

0000

1111

Keywords:

Protein oligomerization

Folding intermediates

Molecular Dynamics

MM-PBSA

Configurational clustering

ABSTRACT

The D76N mutant of the β_2 m protein is a biologically motivated model system to study protein aggregation. There is strong experimental evidence, supported by molecular simulations, that D76N populates a highly dynamic conformation (which we originally named I_2) that exposes aggregation-prone patches as a result of the detachment of the two terminal regions. Here, we use Molecular Dynamics simulations to study the stability of an ensemble of dimers of I_2 generated via protein–protein docking. MM-PBSA calculations indicate that within the ensemble of investigated dimers the major contribution to interface stabilization at physiological pH comes from hydrophobic interactions between apolar residues. Our structural analysis also reveals that the interfacial region associated with the most stable binding modes are particularly rich in residues pertaining to both the N- and C-terminus, as well residues from the BC- and DE-loops. On the other hand, the less stable interfaces are stabilized by intermolecular interactions involving residues from the CD- and EF-loops. By focusing on the most stable binding modes, we used a simple geometric rule to propagate the corresponding dimer interfaces. We found that, in the absence of any kind of structural rearrangement occurring at an early stage of the oligomerization pathway, some interfaces drive a self-limited growth process, while others can be propagated indefinitely allowing the formation of long, polymerized chains. In particular, the interfacial region of the most stable binding mode reported here falls in the class of self-limited growth.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Beta-2 microglobulin (β_2 m) is a small protein whose native structure exhibits a typical immunoglobulin fold. The 99 residues that compose the wild-type (*wt*) form are arranged into two sheets of anti-parallel beta-strands forming a sandwich-like structure. One of the sheets comprises strands A-B-E-D, while the other is composed by strands C-F-G. The native structure is stabilized by a disulfide bridge linking residues Cys25 (strand B) and Cys80 (strand F) [1,2] (Fig. 1A–B).

The *wt* form is the causing agent of dialysis related amyloidosis (DRA), a conformational disorder associated with the formation of

amyloid fibrils in the osteoarticular system of individuals undergoing long-term (>10 yrs) hemodialysis [3]. The *ex-vivo* amyloid fibrils of patients with DRA are composed of the *wt* form, and, to a lesser extent (30%), of a truncated structural variant lacking the six N-terminal residues, which is termed Δ N6 [4]. The *wt* form does not aggregate *in vitro* under physiological conditions, while the Δ N6 does [5,6]. For this reason, it has been extensively used as a model system to study β_2 m aggregation. However, the biological significance of Δ N6 lacks consensus because it is not clear if the proteolytic cleavage of the terminal hexapeptide occurs before, or after, fibril assembly [7].

More recently, the D76N mutant, was identified as being the culprit of a rare systemic amyloidosis affecting visceral organs (liver, kidneys etc.) [8], which readily aggregates *in vitro* under physiological conditions [9]. This single point mutant thus appears to be a suitable model system to study the details of β_2 m aggregation mechanism.

* Corresponding authors.

E-mail addresses: pffaisca@fc.ul.pt (P.F.N. Faísca), machuque@ciencias.ulisboa.pt (M. Machuqueiro).

¹ Both authors contributed equally to this work.

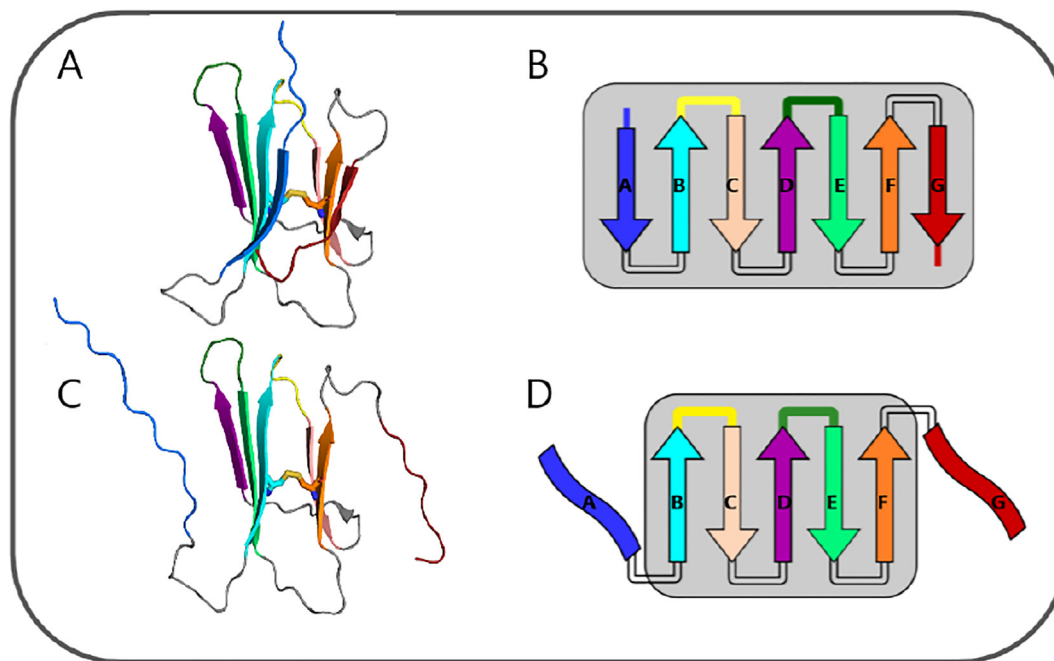


Fig. 1. Three-dimensional cartoon representation of the native structure of D76N (pdb id:2YXF[2])(A) and the schematic representation of its secondary structure (B). The I_2 intermediate state populated by D76N, which features a well-preserved core and two unstructured and decoupled termini (C, D). The N- and C-termini are coloured in blue and red, respectively. The BC- and DE-loop are colored yellow and green, respectively.

The small size of β_2m , together with its biomedical significance, renders it an ideal target for *in silico* investigations framed on molecular simulations, and a plethora of interesting results has been reported in the literature (reviewed in [10]). In particular, some of us developed a three-stage approach that integrates information obtained from complementary simulation methodologies [11–13], to provide a probabilistic, structurally-resolved picture of the initial phase of the β_2m aggregation mechanism, which starts with the formation of dimers [14]. Specifically, we use discrete Molecular Dynamics (MD) of a full atomistic structure-based model [15,16] to explore the folding space of the monomer and detect intermediate states with aggregation potential. The identified intermediate states are then used as starting conformations in constant-pH MD simulations [17–19] to create ensembles of monomers representative of different environmental conditions. In the last stage of our procedure, we deploy an in-house protein–protein Monte Carlo ensemble docking method (MC-ED), to create ensembles of dimers whose statistical analysis allows predicting which regions of the monomer are more likely to self-associate into dimers, and identifying the so-called aggregation hot-spots, i.e., the residues that establish more intermolecular interactions in the interfacial region, being therefore critical to initiate the aggregation pathway.

In the particular case of the D76N mutant, our studies predict the existence of an aggregation-prone intermediate state, termed I_2 , that is structurally characterized by having two unstructured terminal regions, which are decoupled from (a well-preserved) protein core [12] (Fig. 1C–D). We found that this intermediate is more likely to self-associate via the DE-loop and EF-loop under physiological conditions, with the termini becoming more relevant as adhesion zones under acidic pH [13]. The detachment and unfolding of the terminal regions predicted by our structure-based model is consistent with the fact that the D76N substitution breaks an important network of electrostatic interactions distributed over a large number of residues including both termini. This is in a rather good agreement with results based on ssNMR, which shows that D76N sparsely populates a highly dynamic con-

formation that exposes aggregation-prone regions as a result of a loss of β -structure at the terminal strands, i.e., an intermediate state that is topologically equivalent to I_2 [20]. It has been suggested that the shear forces present in the extracellular fluid under physiological conditions may be enough to further unfold the intermediate termini and induce amyloid formation *in vivo* [21].

Our statistical analysis of the protein–protein interfaces based on MC-ED uses a full atomistic, but, nevertheless, rigid representation of the protein. This means, in particular, that it does not take into account the conformational changes that may occur upon binding, and which can go from local rearrangements of the side-chains, to large domain motions. Additionally, it uses a cost (or scoring) function based on square-well potentials that despite considering the three driving forces of protein self-association (shape, hydrophobic, and electrostatic complementarity that includes hydrogen bonds), does not explicitly take into account desolvation effects. Therefore, it provides a qualitative assessment on the stability of the interfaces of the dimers, i.e., of their binding energies. More precise methods, combining the atomic resolution of Molecular Mechanics with an implicit continuum solvent model, have been developed to quantitatively predict the binding energies of protein–protein complexes. In contrast to the docking scoring of single structures, an ensemble of complex conformations generated using MD is evaluated in Molecular Mechanics - Poisson–Boltzmann/surface area (MM-PBSA) or MM generalized Born/surface area (MM-GBSA) approaches [22]. Both these methods have been shown to successfully predict protein–protein binding constants [23] and are particularly useful when the relative energy values between conformers are targeted since, in these cases, the elusive entropic contribution to the binding free energy mostly cancels out [24].

In the present work, we study in detail an ensemble of dimers obtained from monomers of the I_2 intermediate state using the MC-ED method. In order to do so, we designed a protocol based on MD simulations that relaxes the dimer conformation/configuration, and is followed by an evaluation of the dimer binding energy using a novel, in-house implementation of the MM-PBSA method.

Having an ensemble of predicted dimer conformations, we sought to identify binding interfaces or binding modes (BMs) that conjugate energetic stability with unlimited growth, i.e. the ability to form long, elongated polymeric chains without having to undergo large structural rearrangements, and BMs that in the absence of large structural rearrangement are limited to grow into small oligomers. The results reported in the present work may have implications regarding the understanding of the aggregation mechanism of β_2m based on the D76N mutant.

2. Methods

2.1. Molecular Dynamics simulations

We started by considering an ensemble of 221 dimers formed by monomers of the intermediate state I_2 , populated by the D76N mutant under physiological conditions (pH 7.2), which was obtained through the MC-ED method [13]. The molecular dynamics (MD) simulations of the dimers were performed using GROMACS 2018.6 [25], the GROMOS 54A7 force field [26,27], and the SPC water model [28]. The full system used in the simulations comprises a dimer of two I_2 monomers, $\sim 20,000$ water molecules, and 2 Na^+ ions to keep the overall charge neutral. All ionizable residues in the protein were kept in their most probable protonation state at neutral pH [29], including histidines which were all kept neutral. In all production runs, the Particle-Mesh Ewald (PME) electrostatics [30,31] was applied with a verlet scheme cutoff of 1.4 nm, a Fourier grid spacing of 0.12 nm, and an interpolation order of 4 (cubic). Van der Waals interactions were treated with a Lennard-Jones potential and were truncated above 1.4 nm. All bonds were constrained using the LINCS algorithm [32] for the protein while SETTLE was used for the water molecules [33]. The equations of motion were integrated every 2 fs with the neighbor lists being updated every 10 steps.

The dimer conformations were subjected to a classical protocol of energy minimization and MD initialization to ensure GROMOS 54A7 compatibility. A three-step protocol of energy minimization was performed consisting of 10,000 steps with the steepest-descent algorithm followed by 2,000 steps with the low-memory Broyden-Fletcher-Goldfarb-Shanno (l-BFGS) algorithm without constraints. Finally, the LINCS algorithm is activated on all bonds in a shorter third step (~ 100 steps) using again the steepest-descent algorithm. Any existent atomic clashes were removed in these three energy minimization steps. However, to start the temperature (NVT) and pressure (NPT) couplings in the MD initialization procedure, special attention to the position restraints is required and a three-step initialization protocol was also devised. The first step consisted on a 100 ps run in which the system temperature was raised to 310 K using the v-rescale thermostat [34] and a temperature coupling constant of 0.1 ps. Position restraints were applied to the C α atoms with a force of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. In the second step (200 ps), we turned on the Parrinello-Rahman barostat [35,36] to obtain an isotropic pressure of 1 bar with a pressure coupling constant of 0.5 ps and compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$. In this step, the position restraints force constant was reduced to $100 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. The third and final initialization step (200 ps) was focused on attenuating the position restraints, which now use a force constant of $10 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. After this protocol, we performed a visual inspection step to identify conformations in which the MC-ED artificially introduced structural entanglements between monomers (Figure S1 of Supporting Information). Such structures were found in 9 dimer systems (1, 19, 73, 102, 104, 108, 161, 180 and 194), which were excluded from the present study. The remaining 212 dimer conformations were subsequently relaxed with unrestrained MD for 100 ns. In most cases

the equilibrium properties converged relatively fast, but in a few cases the dimeric interfaces required longer equilibration times. Therefore, we performed equilibrium analyses using only the last 20 ns of each MD segment.

2.2. Structural analysis

All structural analyses of the dimers were carried out using the GROMACS package and other in-house tools. For the structural alignments and Root Mean Square Deviation (RMSD) calculations, only the C α atoms from the protein core (namely, 23–27, 36–39, 51–55, 62–66, 78–82 of each monomer) were considered. In doing so, we discard spurious structural variability due to loops, and unstructured N- and C- terminal regions. The interfacial areas of the dimers were calculated using the SAS values of both monomers in the presence and absence of the partner [37]. This procedure was also applied to each residue individually, with the final area being normalized (converted into a percentage) by the maximum SAS area of each residue type observed in all simulations. Residues with interfacial area higher than 10%, were considered to be located at the interface. All error bars indicate the standard error of the mean calculated over pseudo-replicates obtained by splitting the converged sampling into 3 segments of equal size. Images were rendered using PyMOL [38].

2.3. Structural clustering

Due to the large number of dimer configurations explored in this work, we applied a clustering protocol comprising 3 steps to reduce structural variability and try to ensure that unique dimers, representative of distinct BMs, are captured and characterized. In the first step, we started by using the cluster tool of the GROMACS package with the gromos clustering algorithm [39] to obtain a representative (average) structure of each dimer. This initial step convoluted most of the conformational variability resulting from the MD relaxation step into 212 representative conformations.

To reduce even further the structural similarity among the conformations selected in the previous step, a second clustering procedure was applied. This second step required the calculation of an RMSD matrix ($212 \times 211/2$) that accounts for the dissimilarity between dimer conformations. For the fit procedure and calculation of RMSD, we used the β -sheet core region to reduce the contribution of spurious structural variability into the RMSD. Finally, in the third step, we performed A/B to B/A monomer permutations to ensure that the best fits (lowest RMSD configurations) were chosen [40]. The final RMSD matrix was used with the clustering tool available in the HADDOCK 2.2 software [41,42] to generate structurally homogeneous clusters at a RMSD cutoff value of 4 Å and using the gromos algorithm [39].

2.4. Calculation of binding energies with MM-PBSA

All binding free energy (E_{bind}) calculations were performed using a new in-house implementation of the MM/PBSA method [43,44], written in the programming language Python (<https://github.com/mms-fcul/mmpbsa>). This method uses a single-trajectory approach to calculate 4 distinct binding energy components - 2 related to the MM energy in vacuum, namely Van der Waals (E_{vdw}) and Coulombic (E_{coul}) terms, and 2 related to the solvation energy, polar ($Solv_{polar}$) and apolar ($Solv_{apolar}$) - using the parameters of a specific force-field. In this method, E_{coul} and E_{vdw} are calculated using the equations representing the Coulomb and Lennard-Jones potentials, respectively. DelPhi4Py (<https://github.com/mms-fcul/DelPhi4Py> [29]), a DelPhi [45] python wrapper, is used as Poisson-Boltzmann (PB) solver to calculate the $Solv_{polar}$

energy. The final energy component, $Solv_{apolar}$, is calculated using a widely used Solvent Accessible Surface Area Only model (SASA-Only) of the system. The E_{bind} is, thus, comprised of the sum of the 4 aforementioned energetic terms. The program is able to read simple GROMACS coordinate files (.gro), extracted from the MD final trajectories (.xtc), and calculate energies for each given frame.

In this work, GROMOS 54A7 compatible parameters (atomic charges, atomic radii [46] and 12–6 Lennard-Jones intermolecular pair potentials) were used for the calculation of the 4 energy components. In the calculations of E_{coul} and $Solv_{polar}$, a dielectric constant of 4 was used for the protein interior [47,48]. In the P–B calculations, we used a grid scaling factor (*scale*) of 2 and a convergence criteria (*convergence*) of 0.01 kT/e [49]. We also used 500 and 50 linear (*nlit*) and non-linear (*nonit*) iterations, respectively. Additionally, the grid size (*gsz*) and the grid center (*acent*) were dynamically calculated for each file in the following manner: *gsz* was set as 2 times the maximum value of the atomic coordinates in the 3 axis, plus 1 if the final number is even; *acent* was set as the geometric center of all the atomic coordinates for a file. For each dimer, the MM-PBSA calculations were run on the full MD trajectories, every 100 ps, which resulted in 1000 frames per dimer.

3. Results

3.1. Stability analysis

An important goal of our work is to assess the stability of the 221 dimer conformations generated using the MC-ED protocol, and structurally characterise the most stable BMs. However, since the MC-ED cost function creates a few atomic clashes and artificially introduces entanglements at the dimer interface, we designed a MM/MD-based protocol to relax the conformations before evaluating their energetic stability. In particular, we performed multiple steps of energy minimization followed by several initialization MD segments where position restraints were applied to each monomer with decreasing force constants. A set of 212 dimer configurations, were further relaxed using an unrestrained MD segment (100 ns), while the remaining 9 were discarded due to the presence of artificial entanglements (Figure S1 of Supporting Information). The goal of performing this long MD segment is to relax the dimer interfaces and potentially achieve better packing through local rearrangements.

For many dimers, a few nanoseconds was enough to relax their binding interfaces, while others, which were more dynamic, took a longer time to equilibrate. Several properties were used to monitor the equilibration procedure, including RMSD (vs. the initial and final structures), radius of gyration, secondary structure content, size of the interfacial area, etc (Figure S2 of Supporting Information). Since some dimers are more structurally unstable than others, we decided to take a conservative approach and discarded the initial 80 ns of the long MD segments. Therefore, all analyses performed and presented in what follows focused exclusively on the final 20 ns segments.

We observed significantly high values of the interfacial surface area in the ensemble of dimers comprising all relaxed conformations of the 212 MD simulations (Fig. 2A), and an average value of $\sim 8 \text{ nm}^2$. High interfacial area values are usually correlated with shape-complementarity, which is one of the major drivers of protein self-association taken into account in the docking protocol. Since all dimer conformations were relaxed during long MD segments, we also evaluated how much of their equilibrated interfaces diverged from the initial ones (Fig. 2B). Most dimers ($\sim 75\%$) retained more than 50% of the original binding interface generated in the MC-ED, which suggests a high stability, and confirms the predicting ability of the MC-ED method.

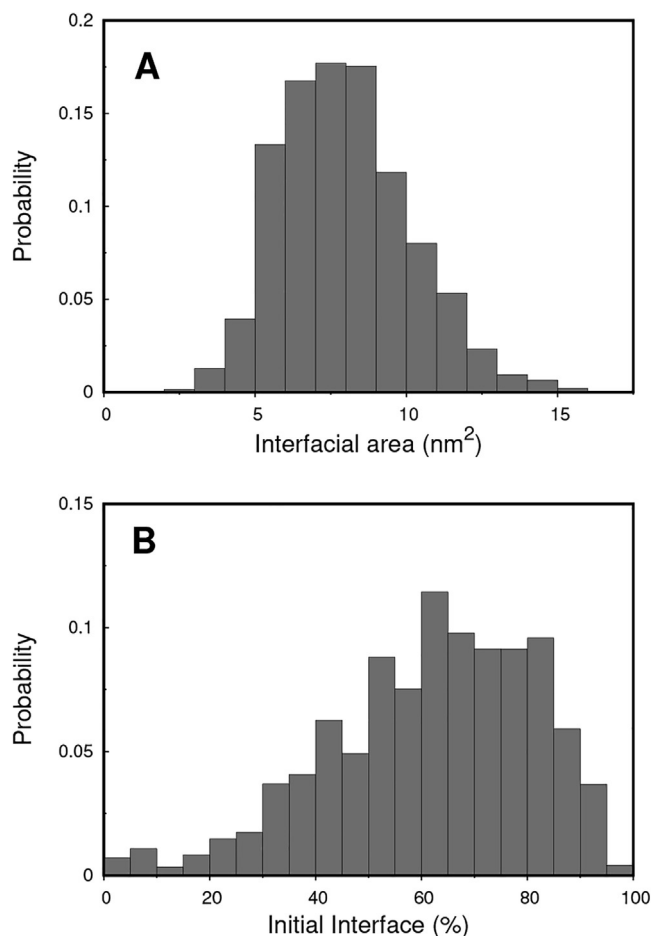


Fig. 2. Interfacial area distribution (A) and percentage of initial interface (B) of the entire relaxed ensemble composed by the last 20 ns of the 212 MD simulations. The interfacial area shown is the average of the two area values mapped on each monomer. The residues of both monomers were included in the calculation of the initial interface percentage (see Methods for further details).

To characterize the nature of the interactions governing/underlying dimer stability, we investigated the formation of cross- β structures or other less specific intermolecular hydrogen bonds (Figure S3 of Supporting Information). No significant differences were observed between the β structure content distributions of the full dimers and the sum of monomers, showing that the amount of cross- β is residual. The average number of intermolecular hydrogen bonds observed is also relatively modest with no preference between main chain and side chains.

Next, we evaluated dimer stability by means of the MM-PBSA approach, which provides binding energies. The binding energies calculated for the relaxed portion of our 212 MD simulations follow a normal distribution with an average $\sim -40 \text{ kcal mol}^{-1}$ and a total range spanning from -100 to 0 kcal mol^{-1} (Fig. 3A). These low binding energies confirm that the interfaces obtained from MC-ED and relaxed using the MD protocol are very stable, which is consistent with the high interfacial area observed for most conformations. In the MM-PBSA formalism, the binding energy is usually decomposed in vacuum and solvation terms, or in polar and apolar contributions. By separating the polar from the apolar terms to the total energy, we observe that the apolar interactions are the main contributors to dimer stability (Fig. 3B and Figure S4 of Supporting Information), suggesting that affinity and selectivity are not driven by polar contacts but rather by contact surface and shape complementarity.

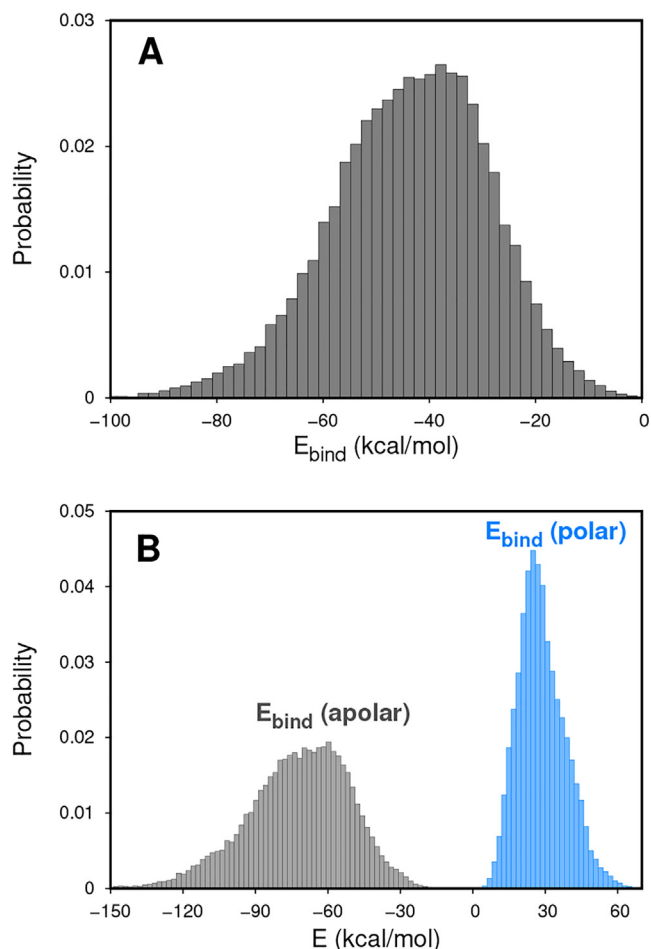


Fig. 3. Probability histograms for the binding energies (A), and their polar (blue) and apolar (black) contributions (B) calculated with the MM-PBSA method over the complete ensemble of relaxed $\beta_2\text{m}$ dimers. The polar contribution of the binding energy is the sum of coulombic and polar solvation energies, while the apolar contribution was obtained from van der Waals and apolar solvation energies.

We were intrigued by the positive polar energy contributions and tested the impact of changing the dielectric constant in the MM/PBSA methodology. This is an important parameter that despite being system-dependent, has been adjusted to fit different experimental data [48,50–52] with the value of 4 leading to the best results [47,48,50,51]. Nevertheless, we repeated the calculations of the polar energy contributions using $\epsilon = 3$ and 5 (Figure S5 of Supporting Information). As expected from the Coulomb and Poisson–Boltzmann formalisms, a higher dielectric constant leads to a larger attenuation between electrostatic interactions. Since these interactions are unfavorable, due to high energies, we observe better binding with $\epsilon = 5$. However, while artificially increasing the dielectric constant can bring these contributions closer to zero, it will not make the interactions energetically favorable.

3.2. Structural analysis

3.2.1. Identification of hot-spot residues

In protein aggregation the term “hot-spot” is often used to refer to residues (or regions) of the protein (typically non-polar groups that are buried inside the native structure) that give rise to intermolecular interactions between aggregation-prone intermediate states thus triggering the aggregation pathway. Here, we perform a fine-grained analysis of the dimer conformations in order to

determine which residues are present at the dimer interface with higher probability, being potential candidates for aggregation hot-spots. The identification of hot-spots resulting from this procedure differs from that adopted in previous studies [11–13] in which a hot-spot is a residue that establishes the most interactions within the ensemble of the most frequent interfacial contacts. An initial analysis was performed based on the relaxed conformations of all 212 dimers and the resulting probability histogram predicts several aggregation hot-spots that are located at both the N- and C-termini, as well as at the DE-loop (Figure S6 of Supporting Information). Nevertheless, this analysis does not allow one to distinguish between residues that actively stabilize the dimer interface from bystanders, i.e., residues that are highly likely to be present but do not establish any stabilizing role. With this idea in mind, we used the MM-PBSA data to sort all dimers in our ensemble according to their binding energies and constructed two data sets: one containing the best (i.e. more stable) interfaces, and another containing the worst (i.e. less stable) interfaces (10% cutoff). By subtracting from the probability of being present in the best (worst) interfaces P_B (P_W), the probability P mentioned above, we identified enriched regions of the protein (i.e. regions where the probability difference, $\Delta P = P_{B(W)} - P > 0$) and impoverished regions of the protein (i.e. regions where the probability difference, $\Delta P = P_{B(W)} - P < 0$) in the best and worst interfaces (Fig. 4). For the most stable interfaces, the enriched regions comprise the N-terminus, the BC-loop, the DE-loop and the C-terminus. Interestingly, these regions correspond to the impoverished regions for the less stable interfaces, which further supports their stabilizing role in self-association. In contrast, for the less stable dimer interfaces the enriched protein regions comprise the CD-loop and EF-loop, which are the impoverished regions for the more stable interfaces. These results suggest that the more stable interfaces may trigger an aggregation pathway, which is different from that triggered by the less stable ones.

3.2.2. Identification and structural characterization of stable binding modes

After determining potential aggregation-prone regions of the D76N mutant of $\beta_2\text{m}$, we set out to identify the most stable BMs. We started by applying a clustering protocol to ensure that only unique interfaces, corresponding to different BMs were selected. By using this approach, we obtained 161 individual BMs (clusters) from the 212 representative dimer conformations. Using the MM-PBSA binding energy values, we ranked these BMs according to their stability, and the top scorers were used in further analyses (Table 1).

In the 10 most stable BMs captured by our protocol, we noticed BM-4 as an outlier since the percentage of initial interface upon relaxation was overwhelmingly low, as a result of a large conformational rearrangement. Even though this BM was not predicted by the MC-ED, the high stability still makes it worth investigating. Overall, the top BMs obtained from our clustering analysis have low binding energies spanning from -83 to -67 kcal/mol and, as expected, they also display high interfacial areas, with the most stable BM having the highest interfacial area. However, there is no clear correlation between the size of the interface and the final binding energies, indicating that the nature of the residues involved in the intermolecular interactions is also important. We used these 10 BMs to calculate a probability histogram for the presence of residues in the interfacial region, which confirmed the results reported previously that considered the ensemble of 212 dimers (Figure S7 of Supporting Information).

The three-dimensional structural representation of the most stable BMs highlights a remarkable diversity of dimerization interfaces (Fig. 5), which suggests a complex and heterogeneous

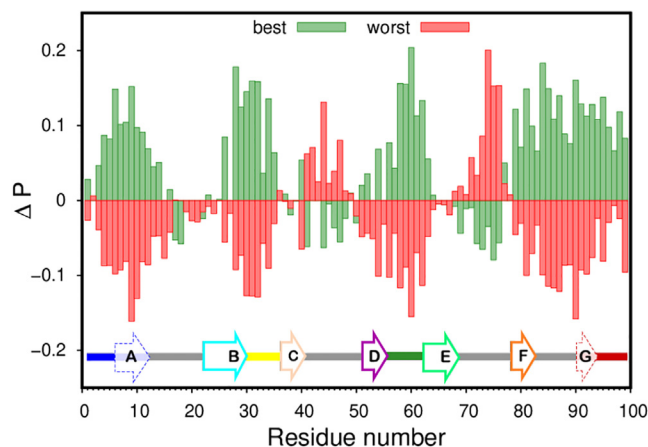


Fig. 4. The most likely residues (positive probability difference; ΔP) in the interface of the most stable (green) and less stable (red) dimer interfaces. The best and worst dimer interfaces were selected based on a 10% cutoff on their binding energy. A linear depiction of the secondary structure content (bottom) using the color code of Fig. 1. β sheets A and G have a dashed arrow and are included into their adjacent termini since these domains are significantly unstructured in the I2 β_2m intermediate [13].

oligomerization pathway for the D76N mutant of β_2m . By visual inspection, one can easily recognize the aggregation-prone regions (highlighted as colored patches) that were previously identified (Figure S7 of Supporting Information). In particular, the N- or the C-ter are present in the interfaces of all the BMs reported in Fig. 5. Since the BC- and DE-loops are usually located on the same side of the monomeric form of β_2m , binding through this interface (as in the case of BM-1, 6, and 7) should not be limited two these structural elements to allow for further oligomerization. Therefore, it is likely that the two terminal regions will play a role in the oligomerization process, by providing alternative BMs.

3.2.3. Oligomerization of stable binding modes

In order to explore how early protein–protein interactions may determine the formation of specific aggregate arrangements, we designed a simple molecular visualization protocol that generates larger oligomers by propagating the initial dimerization interface of a selected BM. In particular, we generated 6-mers by adding 4 dimeric units to a starting dimer representative of a stable BM (Fig. 6).

The oligomers thus obtained can be classified into limited and unlimited growth oligomers, depending on whether the addition of more units leads to an oligomeric chain that closes upon itself, or, in alternative, a chain that may sustain further grow. Limited growth oligomers are usually obtained when the interfaces of each monomer are oriented in a such a way that the addition of further

units leads to closed structures. The simplest case is a head-to-head arrangement, in which both monomers have the same residues in the interface. BMs 5 and 6 are examples of limited growth dimers in which both monomers interact via the same regions (BM-5: Nter + CDloop + Cter; BM-6: BCloop + DEloop + Cter). BM-1, the most stable BM reported here, leads to closed trimers with an interface involving the Nter + BCloop + DEloop from one monomer and the BCloop + DEloop + Cter from the other.

The unlimited growth oligomers can be further classified into linear (BMs 2, 3, 8, and 10) and helical oligomers (BMs 4, 7, and 9). All these BMs have contact interfaces where the monomers interact via different regions and residues. This leads to the generation of free contact surfaces that permit further oligomerization. The relative orientation of monomers at the interfaces is what determines the formation of linear or helical aggregates. For these BMs, only a small rearrangement of the interfacial/contact regions is required to accommodate another monomer, which suggests an efficient oligomerization process.

This simple analysis shows that, in the absence of major conformational rearrangements, the relative orientation of the monomers in the initial interfacial region is thus key to determine the shape and size of the final aggregates. The limitation to grow beyond a certain size can be overcome by mixing different BMs, or if the final oligomer is prone to undergo a major conformational change compatible with further growth.

4. Discussion and Conclusion

Protein aggregation is a remarkably complex process, where the formation of molecular structures of varying size and timescales is highly dependent on environmental conditions. Solving the aggregation mechanism, requires determining the size, distribution and structures of the oligomeric assemblies, filaments, protofibrils and fibrils that populate the amyloid pathway, as well as the rate constants associated with every transition [53]. In the case of β_2m it is even a bigger challenge since the *wt* form does not aggregate *in vitro* under physiological conditions. Despite these difficulties, progress has been made regarding the initial phase of the aggregation mechanism (reviewed in [10]), as well as in determining the structure of amyloid fibrils [54]. In particular, research carried out by several research groups indicate that the first phase of the β_2m aggregation mechanism is the dimerization of aggregation prone monomers [10].

Here, we studied in detail an ensemble of dimers that resulted from the association of an aggregation-prone monomeric state populated by the D76N natural mutant of β_2m . This mutant has attracted considerable attention in the last years because, contrary to the *wt* form, it aggregates readily *in vitro* in physiological conditions [8,21]. The monomeric intermediate state was previously reported in a molecular simulation study based on discrete MD

Table 1

Binding energies and dimer interface details for the top 10 best clusters/BMs. The binding energy contributions were obtained directly from the MM-PBSA calculation and the interfacial area was calculated using our SASA-based protocol. The initial interface percentage was obtained comparing the presence of the individual residues at the dimer interface after the MD relaxation step in relation to the initial structure.

Binding Mode	E_{bind} (kcal/mol)	E_{bind} (Apo) (kcal/mol)	E_{bind} (Pol) (kcal/mol)	Interf. Area (nm ²)	Initial Interf. (%)
1	-82.9 ± 7.1	-137.0 ± 10.3	54.1 ± 4.5	14.2	73.4
2	-82.1 ± 5.2	-121.5 ± 6.5	39.4 ± 2.5	12.9	60.6
3	-79.0 ± 3.2	-115.4 ± 2.8	36.5 ± 1.5	11.4	92.1
4	-77.3 ± 7.6	-104.7 ± 12.5	27.4 ± 5.0	10.7	5.8
5	-75.4 ± 1.3	-106.0 ± 1.0	30.6 ± 0.3	10.9	78.4
6	-71.8 ± 13.4	-121.9 ± 18.8	50.1 ± 5.9	13.2	82.9
7	-71.0 ± 6.9	-122.8 ± 10.3	51.8 ± 3.7	13.7	62.5
8	-68.0 ± 3.3	-110.3 ± 6.9	42.3 ± 3.9	11.4	83.6
9	-67.5 ± 3.2	-116.5 ± 5.1	49.0 ± 3.0	12.7	66.5
10	-67.2 ± 2.0	-105.9 ± 1.1	38.7 ± 1.0	10.1	66.2

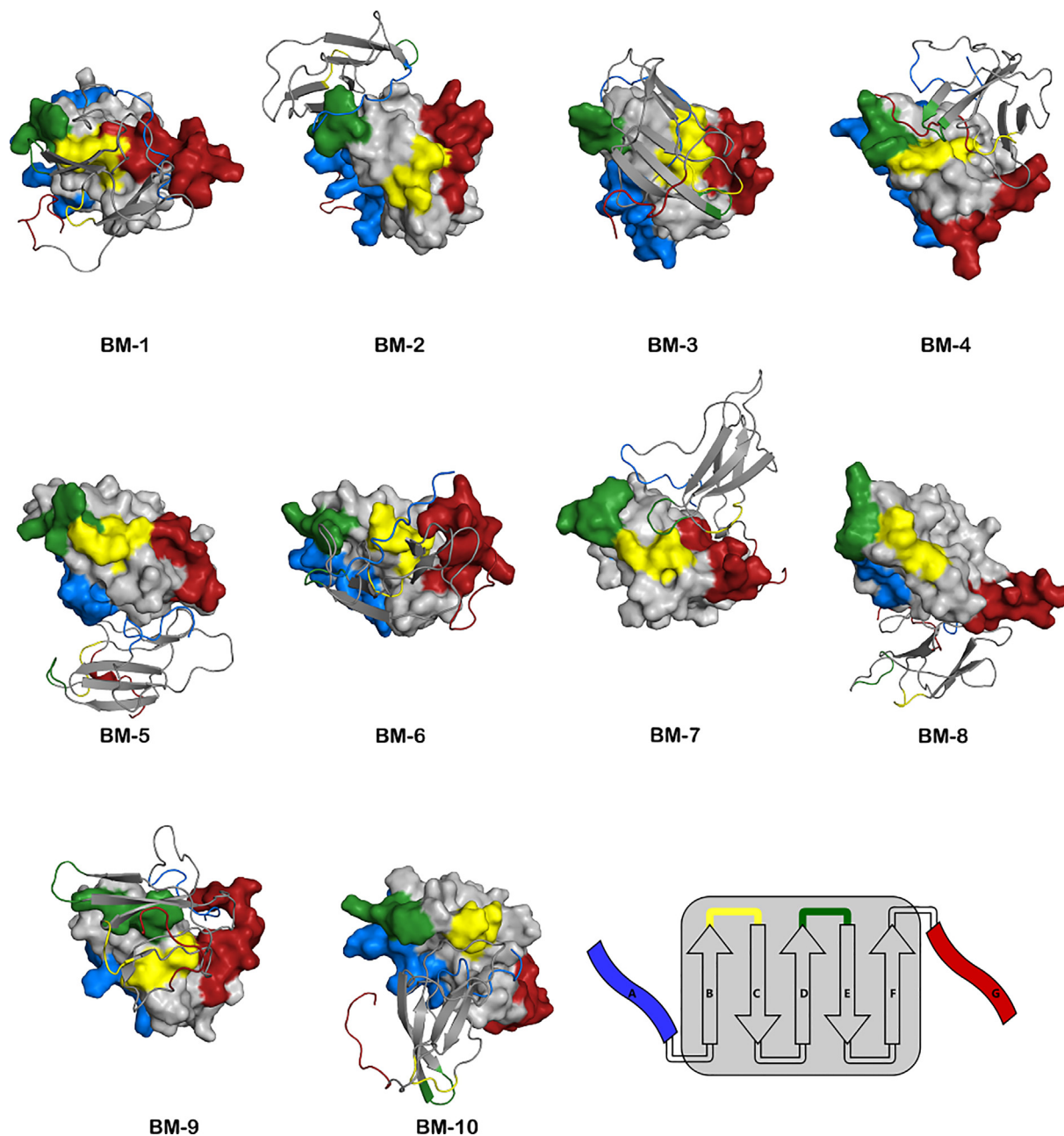


Fig. 5. Structural representation of the 10 most stable β_2m dimer binding modes. A schematic representation of the secondary structure and color code is also shown. The N- and C-terminus, and the BC- and DE-loop are colored in blue, red, yellow and green, respectively. The structures are aligned by the monomer represented with a surface, while the other monomer is represented with cartoon.

simulations of a full-atomistic structure-based model [12,13,55]. It is structurally characterized for having the two terminal regions unstructured and detached from a well-preserved core. An intermediate state sharing the same topology was later reported in a study that combined ss-NMR with MD simulations [20]. An intermediate state sharing the same topology was later reported in a study that combined ss-NMR with MD simulations [20]. More recently, the occurrence of a highly-energetic intermediate state, exhibiting a smaller loss of beta structure than that populated by the D76N mutant was reported for the *wt* form based on a similar methodological approach [56]. The observation that the murine β_2m does not populate intermediate states is in line with its considerable lower ability to oligomerize, and incapacity to form amy-

loid fibrils [57]. These findings thus reinforce the importance of intermediate states as triggers of protein self-association.

The ensemble of dimers investigated here was prepared with an in house protein-protein docking protocol that optimizes the interfacial region for shape, electrostatic and hydrophatic complementarity [13]. While the docking method provides a qualitative characterization of the dimerization interfaces, it cannot be used to assess their stability in a quantitative manner. For this reason, in the present study we designed a simulation protocol that combines classical MD with MM-PBSA to relax the dimer interfaces and accurately estimate their binding energies. We further designed a structural clustering protocol that allowed us to isolate unique dimer configurations representative of distinct BMs. Our analysis shows that the major driving force of interface stabilization are

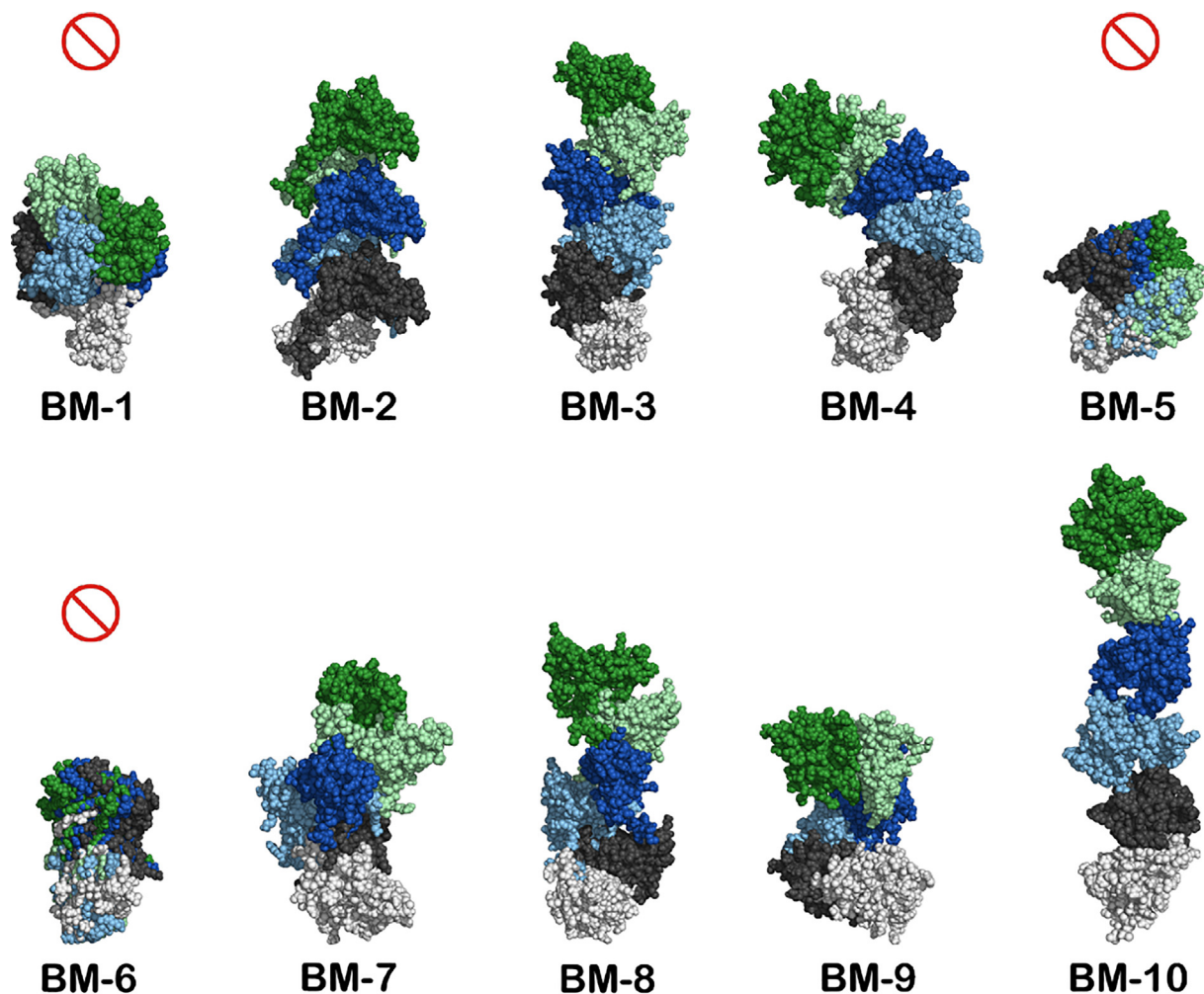


Fig. 6. Oligomerization process for the 10 best BMs. Three dimensional representation of 6-mer oligomers where the interface of the initial dimer (MonA/B in white/black spheres, respectively) was propagated by adding four new dimer units via the alignment of the new MonA to the previous MonB (using Pymol[38]). With the exception of the first dimer where MonA is colored in light grey, all MonA conformations were hidden, which results in a 6-mer with the ABBBBB monomer arrangement. The progression of the oligomerization (MonBs) follows the color code: black; light blue; blue; light green; green. A prohibition sign indicates the limited growth BMs. For visual clarity we only report 6-mers but the structure of larger oligomers can be easily inferred.

interactions between apolar residues. This is perhaps not surprising taking into account the fact that in its functional form β_2m is docked onto the β_3 domain of the major histocompatibility complex (MHC-I) through the four-stranded (A-B-E-D) beta-sheet, and that the four aromatic, bulky residues Phe56, Trp60, Phe62, and Tyr63 and the aliphatic Leu65, which are shielded from the solvent in the quaternary structure of the MHC, become highly solvent-exposed in the monomeric form, thus being able to act as sticky patches in intermolecular association. In line with this observation, we find that the interfaces of the most stable binding modes are particularly rich in residues pertaining to both the N- and C-terminus, as well residues from the BC- and DE-loops. We highlight Trp60 from the DE-loop, which is classically recognized as an aggregation hot-spot for β_2m [10]. On the other hand, the less stable interfaces are stabilized by intermolecular interactions involving residues from the CD- and EF-loops.

Since there is no clear relation between dimer stability and aggregation potential [58], in the present work, instead of focusing on the most stable dimer, we considered an ensemble of dimers. The three-dimensional structural representation of the most stable BMs revealed a remarkable variability of dimerization interfaces. In order to investigate the early protein–protein interactions that establish at the dimer level and can determine the specific aggre-

gate arrangements, we devised a simple geometric rule combined with a visualization protocol that propagates the interfaces of the 10 most stable BMs. Perhaps not surprisingly, we found that in the absence of major conformational rearrangements, the relative orientation of the monomers in the initial interfacial region is key to determine the shape and size of the final aggregates. Moreover, the dimerization interfaces in which the participating monomers contribute with the same residues (e.g. in a head-to-head arrangement) limit the growth of relatively small oligomers, while more heterogeneous dimeric interfaces in which monomers interact via different regions and residues have a potential to be propagated leading to larger oligomeric chains.

The model oligomers predicted in this work share no structural resemblance with the mature amyloid fibrils reported by Radford and co-workers [58]. It is possible that a conformational change occurring at some stage of fibril growth will lead to amyloid fibrils with a final different structural fingerprint. However, it is unlikely that the latter will share a resemblance with that reported in [58] for two major reasons. First, the amyloid structure reported therein is for the *wt* form of β_2m , while the structural models predicted here refer to the D76N mutant. Secondly, the latter were built from monomers that were relaxed with constant pH MD simulations at neutral pH 7.2, while the *wt* amyloid fibrils were obtained under

acidic conditions. Thus, it is not only the model systems that are different, but the environmental conditions of both studies also differ, with aggregation being strongly dependent on environmental conditions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge Diogo Vila Viçosa, Pedro Reis, and Tomás Silva for valuable discussions. António M. Baptista is also acknowledged for developing the fixbox tool, which was used to correct the periodic boundary conditions in the dimer trajectories. We acknowledge financial support from Fundação para a Ciência e a Tecnologia through grant CEECIND/02300/2017 and projects PTDC/FIS-OUT/28210/2017, UIDB/04046/2020, and UIDP/04046/2020.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.csbj.2021.09.003>.

References

- [1] C.H. Trinh, D.P. Smith, A.P. Kalverda, S.E.V. Phillips, S.E. Radford, Crystal structure of monomeric human β 2-microglobulin reveals clues to its amyloidogenic properties. *Proc Natl Acad Sci USA* 99 (15) (2002) 9771–9776. arXiv:<https://www.pnas.org/content/99/15/9771.full.pdf>. doi:10.1073/pnas.152337399. url:<https://www.pnas.org/content/99/15/9771>.
- [2] K. Iwata, T. Matsuura, K. Sakurai, A. Nakagawa, Y. Goto, High-resolution Crystal Structure of β 2-Microglobulin Formed at pH 7.0† *J Biochem* 142 (3) (2007) 413–419. arXiv:<https://academic.oup.com/jb/article-pdf/142/3/413/6442153/mvm148.pdf>, doi:10.1093/jb/mvm148. doi:10.1093/jb/mvm148.
- [3] Gejyo F, Yamada T, Odani S, Nakagawa Y, Arakawa M, Kunitomo T, Kataoka H, Suzuki M, Hirasawa Y, Shirahama T, Cohen A, Schmid K. A new form of amyloid protein associated with chronic hemodialysis was identified as β 2-microglobulin. *Biochem Biophys Res Commun* 1985;129(3):701–6. [https://doi.org/10.1016/0006-291X\(85\)91948-5](https://doi.org/10.1016/0006-291X(85)91948-5). url:<https://www.sciencedirect.com/science/article/pii/S0006291X85919485>.
- [4] V. Bellotti, M. Gallieni, S. Giorgetti, D. Brancaccio, Dynamic of β 2-microglobulin fibril formation and reabsorption: The role of proteolysis, *Semin Dial* 14 (2) (2001) 117–122. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1525-139x.2001.00030.x>, doi:<https://doi.org/10.1046/j.1525-139x.2001.00030.x>. url:<https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1525-139x.2001.00030.x>.
- [5] G. Esposito, R. Michelutti, G. Verdona, P. Viglino, H.H. Ández, C.V. Robinson, A. Amoresano, F.D. Piazz, M. Monti, P. Pucci, P. Mangione, M. Stoppini, G. Merlini, G. Ferri, V. Bellotti, Removal of the n-terminal hexapeptide from human β 2-microglobulin facilitates protein aggregation and fibril formation, *Protein Sci.* 9 (5) (2000) 831–845. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1110/ps.9.5.831>, doi:<https://doi.org/10.1110/ps.9.5.831>. url:<https://onlinelibrary.wiley.com/doi/abs/10.1110/ps.9.5.831>.
- [6] T. Eichner, A.P. Kalverda, G.S. Thompson, S.W. Homans, S.E. Radford, Conformational Conversion during Amyloid Formation at Atomic Resolution, *Mol Cell* 41 (2) (2011) 161–172. doi:10.1016/j.molcel.2010.11.028. doi:10.1016/j.molcel.2010.11.028.
- [7] M. Monti, S. Principe, S. Giorgetti, P. Mangione, G. Merlini, A. Clark, V. Bellotti, A. Amoresano, P. Pucci, Topological investigation of amyloid fibrils obtained from β 2-microglobulin, *Protein Sci* 11 (10) (2002) 2362–2369. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1110/ps.0206902>, doi:<https://doi.org/10.1110/ps.0206902>. url:<https://onlinelibrary.wiley.com/doi/abs/10.1110/ps.0206902>.
- [8] S. Valleix, J.D. Gillmore, F. Bridoux, P.P. Mangione, A. Dogan, B. Nedelec, M. Boimard, G. Touchard, J.-M. Goujon, C. Lacombe, P. Lozeron, D. Adams, C. Lacroix, T. Maisonneuve, V. Planté-Bordeneuve, J.A. Vrana, J.D. Theis, S. Giorgetti, R. Porcari, S. Ricagno, M. Bolognesi, M. Stoppini, M. Delpech, M.B. Pepys, P.N. Hawkins, V. Bellotti, Hereditary Systemic Amyloidosis Due to Asp76Asn Variant β 2-Microglobulin, *N. Engl. J. Med.* 366 (24) (2012) 2276–2283. doi:10.1056/NEJMoa1201356. doi:10.1056/NEJMoa1201356.
- [9] Mangione PP, Esposito G, Relini A, Raimondi S, Porcari R, Giorgetti S, Corazza A, Fogolari F, Penco A, Goto Y, Lee Y-H, Yagi H, Ceconi C, Naqvi MM, Gillmore JD, Hawkins PN, Chiti F, Rolandi R, Taylor GW, Pepys MB, Stoppini M, Bellotti V, Drug W, Unit D, Proteins AP, Sciences B, Emilia R, Proteins AP, Street RH. Structure, folding dynamics and amyloidogenesis of Asp76Asn β 2 - microglobulin: roles of shear flow, hydrophobic surfaces and α crystallin. *J Biol Chem* 2013;288(43):30917–30. <https://doi.org/10.1074/jbc.M113.498857>.
- [10] Loureiro RJS, Faisca PFN. The Early Phase of β 2-Microglobulin Aggregation: Perspectives From Molecular Simulations. *Front Mol Biosci* 2020;7:278. <https://doi.org/10.3389/fmolb.2020.578433>. url:<https://www.frontiersin.org/article/10.3389/fmolb.2020.578433>.
- [11] S.G. Estácio, H. Kroboth, D. Vila-Viçosa, M. Machuqueiro, E.I. Shakhnovich, P.F. N. Faisca, A simulated intermediate state for folding and aggregation provides insights into β 2-microglobulin amyloidogenic behavior, *PLOS Comp Biol* 10 (5) (2014) 1–17. doi:10.1371/journal.pcbi.1003606. doi:10.1371/journal.pcbi.1003606.
- [12] R.J.S. Loureiro, D. Vila-Viçosa, M. Machuqueiro, E.I. Shakhnovich, P.F.N. Faisca, A tale of two tails: The importance of unstructured termini in the aggregation pathway of β 2-microglobulin, *Proteins Struct Funct Bioinf* 85 (11) (2017) 2045–2057. doi:<https://doi.org/10.1002/prot.25358>. url:<https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.25358>.
- [13] Loureiro RJS, Vila-Viçosa D, Machuqueiro M, Shakhnovich EI, Faisca PFN. The Early Phase of β 2m Aggregation: An Integrative Computational Study Framed on the D76N Mutant and the AN6 Variant. *Biomolecules* 2019;9(8). <https://doi.org/10.3390/biom9080366>.
- [14] Rennella E, Cutuili T, Schanda P, Ayala I, Gabel F, Forge V, Corazza A, Esposito G, Brutscher B. Oligomeric states along the folding pathways of β 2-microglobulin: Kinetics, thermodynamics, and structure. *J Mol Biol* 2013;425 (15):2722–36. <https://doi.org/10.1016/j.jmb.2013.04.028>. url:<https://www.sciencedirect.com/science/article/pii/S0022283613002763>.
- [15] Kroboth H, Estácio SG, Faisca PFN, Shakhnovich EI. Identification of a Conserved Aggregation-Prone Intermediate State in the Folding Pathways of Spc-SH3 Amyloidogenic Variants. *J Mol Biol* 2012;422(5):705–22. <https://doi.org/10.1016/j.jmb.2012.06.020>. url:<http://www.sciencedirect.com/science/article/pii/S0022283612004937>.
- [16] Estácio SG, Shakhnovich EI, Faisca PFN. Assessing the Effect of Loop Mutations in the Folding Space of β 2 -Microglobulin with Molecular Dynamics Simulations. *Int J Mol Sci* 2013;14:17256–78. <https://doi.org/10.3390/ijms140917256>.
- [17] Machuqueiro M, Baptista AM. Constant-pH molecular dynamics with ionic strength effects: Protonation-conformation coupling in decalysine. *J Phys Chem B* 2006;110:2927–33.
- [18] S.R.R. Campos, M. Machuqueiro, A.M. Baptista, Constant-ph molecular dynamics simulations reveal a β rich form of the human prion protein. *J Phys Chem B* 114 (39) (2010) 12692–12700, PMID: 20843095. doi:10.1021/jp104753t.
- [19] Vila-Viçosa D, Campos SRR, Baptista AM, Machuqueiro M. Reversibility of prion misfolding: insights from constant-pH molecular dynamics simulations. *J Phys Chem B* 2012;116(30):8812–21.
- [20] T. Le Marchand, M. de Rosa, N. Salvi, B.M. Sala, L.B. Andreas, E. Barbet-Massin, P. Sormanni, A. Barbiroli, R. Porcari, C. Sousa Mota, D. de Sanctis, M. Bolognesi, L. Emsley, V. Bellotti, M. Blackledge, C. Camilloni, G. Pintacuda, S. Ricagno, Conformational dynamics in crystals reveal the molecular bases for D76N beta-2 microglobulin aggregation propensity, *Nat Commun* 9 (1) (2018) 1658. doi:10.1038/s41467-018-04078-y. doi:10.1038/s41467-018-04078-y.
- [21] M. Stoppini, V. Bellotti, Systemic amyloidosis: Lessons from β 2-microglobulin *, *J Biol Chem* 290 (16) (2015) 9951–9958. doi:10.1074/jbc.R115.639799. doi:10.1074/jbc.R115.639799.
- [22] T. Siebenmorgen, M. Zacharias, Computational prediction of protein-protein binding affinities, *WIREs Comput Mol Sci* 10 (3) (2020) e1448. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1448>, doi: 10.1002/wcms.1448.
- [23] Chen F, Liu H, Sun H, Pan P, Li Y, Li D, Hou T. Assessing the performance of the mm/pbsa and mm/gbsa methods. 6. capability to predict protein-protein binding free energies and re-rank binding poses generated by protein-protein docking. *Phys Chem Chem Phys* 2016;18:22129–39. <https://doi.org/10.1039/C6CP03670H>.
- [24] T. Hou, J. Wang, Y. Li, W. Wang, Assessing the performance of the mm/pbsa and mm/gbsa methods. 1. the accuracy of binding free energy calculations based on molecular dynamics simulations, *J Chem Inf Model* 51 (1) (2011) 69–82, PMID: 21117705. arXiv:<https://doi.org/10.1021/ci100275a>, doi:10.1021/ci100275a.
- [25] Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 2015;1–2:19–25.
- [26] W.R.P. Scott, P.H. H++nenberger, I.G. Tironi, A.E. Mark, S.R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Kr++ger, W.F. van Gunsteren, The GROMOS Biomolecular Simulation Program Package, *J Phys Chem A* 103 (19) (1999) 3596–3607.
- [27] Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark AE, van Gunsteren WF. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur Biophys J* 2011;40(7):843.
- [28] Hermans J, Berendsen HJC, Gunsteren WFV, Postma JPM. A consistent empirical potential for water-protein interactions. *Biopolymers* 1984;23 (8):1513–8.
- [29] Reis PBPS, Vila-Viçosa D, Rocchia W, Machuqueiro M, Pypka: A flexible python module for poisson-boltzmann-based pka calculations. *J Chem Inf Model* 2020;60(10):4442–8. <https://doi.org/10.1021/acs.jcim.0c00718>.
- [30] Darden T, York D, Pedersen L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J Chem Phys* 1993;98(12):10089–92.

- [31] Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *J Chem Phys* 1995;103(19):8577–93.
- [32] Hess B. P-lincs: A parallel linear constraint solver for molecular simulation. *J Chem Theory Comput* 2008;4(1):116–22.
- [33] Miyamoto S, Kollman PA. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem* 1992;13(8):952–62.
- [34] Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys* 2007;126(1):014101.
- [35] Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys* 1981;52(12):7182–90.
- [36] Nosé S, Klein ML. Constant pressure molecular dynamics for molecular systems. *Mol Phys* 1983;50(5):1055–76.
- [37] Magalhães PR, Machuqueiro M, Almeida JG, Melo A, Cordeiro MND, Cabo Verde S, Gümüş ZH, Moreira IS, Correia JD, Melo R. Dynamical rearrangement of human epidermal growth factor receptor 2 upon antibody binding: effects on the dimerization. *Biomolecules* 2019;9(11):706.
- [38] Schrödinger, LLC, The PyMOL molecular graphics system, version 2.4 (August 2020).
- [39] Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Peptide folding: When simulation meets experiment. *Angew Chem Int Ed* 1999;38(1–2):236–40.
- [40] Filipe LC, Machuqueiro M, Baptista AM. Unfolding the conformational behavior of peptide dendrimers: insights from molecular dynamics simulations. *J Am Chem Soc* 2011;133(13):5042–52.
- [41] Dominguez C, Boelens R, Bonvin AMJJ. Haddock: A protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* 2003;125(7):1731–7.
- [42] de Vries SJ, van Dijk ADJ, Krzeminski M, van Dijk M, Thureau A, Hsu V, Wassenaar T, Bonvin AMJJ. Haddock versus haddock: New features and performance of haddock2.0 on the capri targets. *Proteins Struct Funct Bioinf* 2007;69(4):726–33.
- [43] Kumari R, Consortium OSDD, Lynn A. g_mmpbsa—a gromacs tool for high-throughput mm-pbsa calculations. *J Chem Inf Model* 2014;54(7):1951–62.
- [44] Genheden S, Ryde U. The mm/pbsa and mm/gbsa methods to estimate ligand-binding affinities. *Expert Opin Drug Discov* 2015;10(5):449–61.
- [45] Rocchia W, Sridharan S, Nicholls A, Alexov E, Chiabrera A, Honig B. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. *J Comput Chem* 2002;23(1):128–37.
- [46] Teixeira VH, Cunha CA, Machuqueiro M, Oliveira ASF, Victor BL, Soares CM, Baptista AM. On the use of different dielectric constants for computing individual and pairwise terms in poisson-boltzmann studies of protein ionization equilibrium. *J Phys Chem B* 2005;109(30):14691–706.
- [47] Fogolari F, Brigo A, Molinari H. Protocol for mm/pbsa molecular dynamics simulations of proteins. *Biophys J* 2003;85(1):159–66. [https://doi.org/10.1016/S0006-3495\(03\)74462-2](https://doi.org/10.1016/S0006-3495(03)74462-2).
- [48] Wang C, Nguyen PH, Pham K, Huynh D, Le T-BN, Wang H, Ren P, Luo R. Calculating protein-ligand binding affinities with mmpbsa: Method and error analysis. *J Comput Chem* 2016;37(27):2436–46.
- [49] Teixeira VH, Vila-Vicosa D, Baptista AM, Machuqueiro M. Protonation of dmpc in a bilayer environment using a linear response approximation. *J Chem Theory Comput* 2014;10(5):2176–84.
- [50] H. Sun, Y. Li, S. Tian, L. Xu, T. Hou, Assessing the performance of mm/pbsa and mm/gbsa methods. 4. accuracies of mm/pbsa and mm/gbsa methodologies evaluated by various simulation protocols using pdbbind data set, *Phys Chem Chem Phys* 16 (2014) 16719–16729. doi:10.1039/C4CP01388C. doi:10.1039/C4CP01388C.
- [51] Sun H, Li Y, Shen M, Tian S, Xu L, Pan P, Guan Y, Hou T. Assessing the performance of mm/pbsa and mm/gbsa methods. 5. improved docking performance using high solute dielectric constant mm/gbsa and mm/pbsa rescoring. *Phys Chem Chem Phys* 2014;16:22035–45. <https://doi.org/10.1039/C4CP03179B>.
- [52] Y. Li, Y. Cong, G. Feng, S. Zhong, J.Z.H. Zhang, H. Sun, L. Duan, The impact of interior dielectric constant and entropic change on hiv-1 complex binding free energy prediction, *Struct Dyn* 5 (6) (2018) 064101. arXiv:<https://doi.org/10.1063/1.5058172>. doi:10.1063/1.5058172. doi:10.1063/1.5058172.
- [53] C.M. Gomes, P.F.N. Faisca, Protein Folding: An Introduction, Springer International Publishing, Cham, 2019, pp. 1–63. doi:10.1007/978-3-319-00882-0_1. doi:10.1007/978-3-319-00882-0_1.
- [54] Iadanza MG, Silvers R, Boardman J, Smith HI, Karamanos TK, Debelouchina GT, Su Y, Griffin RG, Ranson NA, Radford SE. The structure of a β 2-microglobulin fibril suggests a molecular basis for its amyloid polymorphism. *Nature Commun* 2018;9(1):4517.
- [55] Morand J, Nunes A, Faisca PFN. The folding space of protein beta. 2-microglobulin is modulated by a single disulfide bridge. *Physical Biology* 2021;18(056001).
- [56] Sala BM, Le Marchand T, Pintacuda G, Camilloni C, Natalello A, Ricagno S. Conformational stability and dynamics in crystals recapitulate protein behavior in solution. *Biophys J* 2020;119(5):978–88.
- [57] Achour A, Brogini L, Han X, Sun R, Santambrogio C, Buratto J, Visentin C, Barbiroli A, De Luca CMG, Sormanni P, et al. Biochemical and biophysical comparison of human and mouse beta-2 microglobulin reveals the molecular determinants of low amyloid propensity. *FEBS J* 2020;287(3):546–60.
- [58] Cawood EE, Guthertz N, Ebo JS, Karamanos TK, Radford SE, Wilson AJ. Modulation of amyloidogenic protein self-assembly using tethered small molecules. *J Am Chem Soc* 2020;142(49):20845–54.