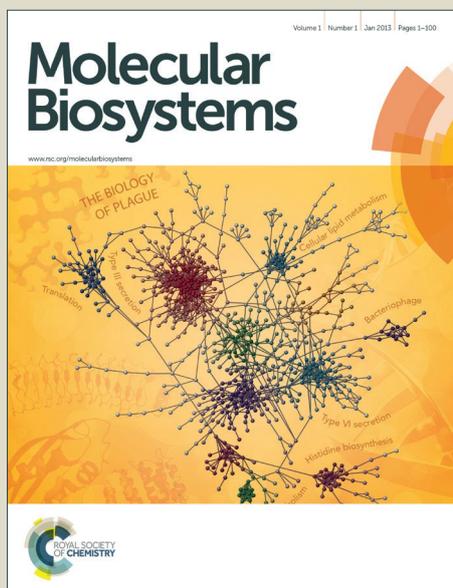


# Molecular BioSystems

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

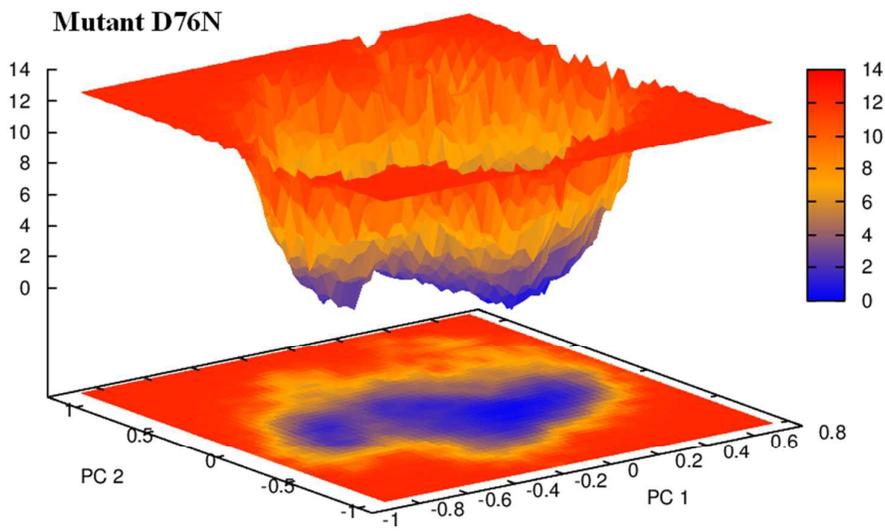
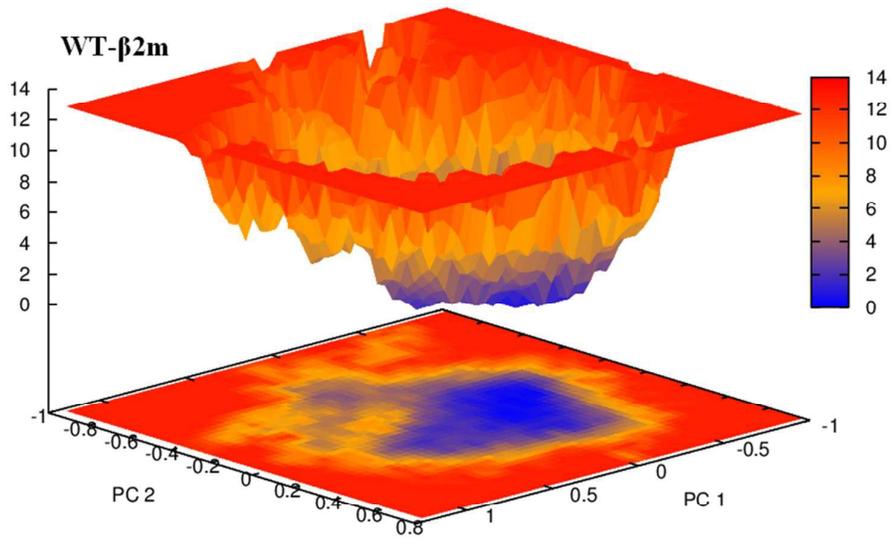
You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



[www.rsc.org/molecularbiosystems](http://www.rsc.org/molecularbiosystems)

Graphical Abstract



## Systematic molecular dynamics approach to the structural characterization of amyloid aggregation propensity of $\beta$ 2-microglobulin mutant D76N

Chandrasekaran P and Rajasekaran R\*

Bioinformatics Division, School of Biosciences and Technology, VIT University, Vellore 632 014, Tamil Nadu, India

### Abstract

Beta-2 Microglobulin ( $\beta$ 2m) is an amyloidogenic protein belongs to the immunoglobulin superfamily, responsible for the dialysis-related amyloidosis (DRA). Misfolding of  $\beta$ 2m is a prerequisite to the formation of systemic amyloidosis have an effect on the structure and function of the affected organ. The aim of our present study is to intensively explore the structural characterization of amyloid aggregation propensity of recently identified natural mutation D76N by applying the classical molecular dynamics (MD) approach. MD result revealed that mutant D76N exhibited a wide variation in the evolutionarily conserved secondary structure profile. Due to an unsatisfied position of main chain donor/acceptor atoms that unable to form essential hydrogen bonds resulted to cause misfolding of mutant D76N by disrupting the local folding of  $\beta$ -strands and turn region. Analysis of time evolution of various structural properties, especially of those of the functionally important residues: aggregation determining, initiating, and gatekeeper residues gave some possible insights into the structural characteristics of the disease mutant D76N. In a nutshell, compared to the wild-type  $\beta$ 2m, aggregation promoting propensity of mutant D76N has established a long  $\beta$ -strand D owing to an inward movement of residue, Asp<sup>53</sup>. Besides, aggregation forming characteristic of DE loop in mutant D76N showed greater flexibility along the first principal eigenvector that favored to enhance an unusual conformational dynamics may lead toward self-aggregation and amyloid fibrils.

**Key words:**  $\beta$ 2m, amyloid aggregation, misfolding, molecular dynamics.

## Introduction

Folding of proteins into their native, functional states is principally dependent on its protein sequence, the cellular environment, and the occurrence of other factors such as chaperones.<sup>1,2</sup> The balance of these conditions, if it is amended, misfolding or aggregation of proteins resulting in amyloid fibrils<sup>3,4</sup> and is associated with various diseases like Alzheimer's disease, Parkinson's disease, spongiform encephalopathies, type II diabetes, and several forms of systemic amyloidosis.<sup>5</sup> The present study, mainly focuses on the amyloid-forming protein  $\beta_2$  microglobulin ( $\beta_2$ m) whose aggregation in humans is a causative agent of the dialysis-related amyloidosis (DRA).<sup>6,7</sup>  $\beta_2$ m is a 99-residue protein with a typical immunoglobulin fold comprises of seven anti-parallel  $\beta$ -strands stabilized by a disulfide bridge between the  $\beta$ -strands, B and F by the residues 25 and 80<sup>8</sup> and it is the light chain of class I major histocompatibility complex (MHC-I).<sup>9</sup> Dissociation of human  $\beta_2$ m from the MHC-I heavy chain immediately catabolized in the kidneys. If individuals experienced with long-term hemodialysis, the elimination process is robustly impaired, and the presence of  $\beta_2$ m in the serum can increase up to 60-fold.<sup>8,10</sup> Therefore, the continuous accumulation of  $\beta_2$ m eventually leads to amyloid assembly and the onset of DRA, a pathological condition characterized by tissue erosion and destruction.<sup>12</sup> Formation of amyloid fibrils is a non-specific characteristic, perhaps all polypeptide chains driven together by main-chain/main-chain interactions, which are modulated by an array of amino acid side-chains. In the natively folded proteins, aggregation-prone residues are typically away from the solvent environment, self-assembly can take place, when the proteins undergo unfold state of the native state and exposes the aggregation-prone residues to the molecular surface. This molecular mechanism is known as the 'conformational change hypothesis' and is supported by the fact that amyloid fibril precursors can exhibit all types of secondary structural organization, whereas amyloid fibrils exhibit a characteristic, common cross- $\beta$  structure. The conformational change hypothesis is well known and supported in which a native protein undergoes a significant structural reorganization to form the amyloid precursor such as transthyretin, lysozyme, prion, and light chain amyloidoses. Consistent with this information, *in vitro* studies also have shown that denaturation or destabilization of the native state of globular proteins can increase amyloidogenic property.<sup>13</sup> It has been shown that, at physiological temperature and pH, the wild-type (WT)  $\beta_2$ m unable to form amyloid aggregation spontaneously *in vitro*, even at high protein concentration for long time periods. Approximately, 60% of the

residues in the  $\beta 2m$  are identified as highly amyloidogenic by various algorithms and also by studies of isolated peptides. The unfolding folding of the protein is capable of rapid fibril formation with acidic pH.<sup>13</sup> Recently, the single point mutant Asp76Asn (D76N), a naturally occurring variant of  $\beta 2m$  associated with late onset of a fatal hereditary systemic amyloidosis characterized by extensive primitive amyloid deposits.<sup>14</sup> Thus, the mapping of the naturally observed mutant D76N to the  $\beta 2m$  assist to explore the conformational dynamics using all atom molecular dynamics (MD) approach. Our investigations revealed that mutant D76N exemplified noteworthy level of variation in the optimal secondary structure profile resulted to cause unfolding of the structure. Furthermore, the aggregation inducing ability of mutant D76N exhibited a diverse conformational motion could be the cause of amyloid aggregation.

## Materials and method

### Set up for MD simulation

The WT- $\beta 2m$  coordinates were obtained by X-ray crystallography and stored in Protein Data Bank (PDB).<sup>15</sup> The atomic coordinates of WT- $\beta 2m$  contained in the 2YXF file showing the highest resolution of 1.13 Å was used to start the simulation. Similarly, the atomic coordinates of the mutant D76N present in the 4FXL file exhibiting the resolution of 1.40 Å was used to set up the simulation. MD simulation was carried out by the use of Gromacs 4.5.5 package.<sup>16</sup> The WT- $\beta 2m$  and its mutant D76N were parameterized with OPLS force field.<sup>17</sup> After the WT- $\beta 2m$  and its mutant have been immersed in cubic boxes separately and filled with simple point-charge (SPC) water molecules, imposing a minimal distance between the solute and the box walls of 1 nm. The counter ions were added as required to neutralize the total charge of the system by using the GENION option of Gromacs. The solvated structures were energy minimized using the steepest descent method and the Particle Mesh Ewald (PME) method<sup>18</sup> was used to treat long-range electrostatic interactions. The SHAKE algorithm<sup>19</sup> was used constrain the bond lengths involving in hydrogen atoms. The canonical ensemble (NVT) was carried out for 100 ps, during which the systems were heated from 100 to 300 K after the required energy minimization process. Harmonic restraints with force constants of 5 kcal were applied to the WT- $\beta 2m$  and its mutant D76N. Subsequently, isothermal-isobaric ensemble (NPT) MD was performed for 100 ps to adjust the solvent density. The production runs of all simulations achieved the lengths of 50 ns,

and the coordinates were saved at regular time intervals of every 2 ps. Then, a comparative analysis was carried to compute the structural variations in the WT- $\beta$ 2m and its mutant D76N.<sup>20</sup>

## 2.2 Principal component analysis

Principal component analysis (PCA) is a method commonly used for dissecting functional significance of the correlated motion of protein and their importance in biological processes, such as substrate binding or protein folding. We applied the `g_covar` module of the GROMACS package to obtain the covariance matrix of  $C_\alpha$  atomic positions from the 50 ns trajectory of the WT- $\beta$ 2m and its mutant D76N. Rotational and translational motions were removed prior to covariance matrix calculation by least-squares superposition to the averaged-structure. The covariance matrix,  $C$  for each element is represented by:

$$C_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle \quad (1)$$

Where  $x_i$  and  $x_j$  are the internal coordinates of atoms  $i$  and  $j$  and all analyzes were performed with the `g_anaeig` module of GROMACS. A set of eigenvalues and eigenvectors were identified by diagonalizing the matrix. Each eigenvector thus represented a single correlated displacement of a group of atoms in a multidimensional space and the eigenvalues were the amplitude of the motion along the eigenvector. Eigenvectors were then sorted according to their eigenvalues in decreasing order. The eigenvectors associated with the highest eigenvalues described the principal components of motion. It had been shown that the first few eigenvectors could successfully describe almost all conformational sub-states accessible to the protein. Only a few eigenvectors were needed to explain a large part of the total variance, as the motions are highly correlated, i.e., collective. To assess the width of the essential space explored by the system as a function of time, the motion described by an eigenvector could be visualized by projecting each frame of trajectory on to the eigenvector.<sup>21-23</sup>

## Results and discussion

### Misfolding of mutant D76N revealed by secondary structure profile

To understand the misfolding pathways of mutant D76N in more detail, it is an important to portray changes in the secondary structure profile along the MD trajectories, for which DSSP (Dictionary of protein secondary structure) analysis was carried out. The time evolution of

secondary structure profile of the WT- $\beta$ 2m (A) and mutant D76N (B) is shown in Fig. 1. It was found that mutant D76N displayed a slight variation in the secondary structure elements approximately up to 27 ns simulation period. After, the mutant D76N underwent a major structural deviation, particularly residues from 42 to 47. During this simulation period, mutant D76N had a tendency to form bend and  $\beta$ -bridge in those regions, where the WT- $\beta$ 2m mainly involved in forming  $\beta$ -turn and  $\beta$ -strand C'. Thus, to explore the observed changes in those regions, we have analyzed the characteristics of hydrogen bonds, because its play a prominent role in the formation of  $\beta$ -strand and turn region. In the WT- $\beta$ 2m, two hydrogen bonds were noticed between the main chain amide group (NH) of Lys<sup>41</sup> and main chain carboxyl group (C=O) of Glu<sup>44</sup> and vice versa. In addition, residue Lys<sup>41</sup> linked to Asn<sup>42</sup> with the hydrogen bond Lys<sup>41</sup>O-Asn<sup>42</sup>H <sup>$\delta$ 1</sup>. At last, residue Arg<sup>45</sup> made two hydrogen bonds with Glu<sup>47</sup> through Arg<sup>45</sup>NH1<sup>n2</sup>-Glu<sup>47</sup>O <sup>$\epsilon$ 1</sup> and Arg<sup>45</sup>NH2<sup>n2</sup>-Glu<sup>47</sup>O <sup>$\epsilon$ 2</sup>. In contrast to the WT- $\beta$ 2m, only one main chain hydrogen bond was identified between the residue Lys<sup>41</sup> (NH group) and Glu<sup>44</sup> (C=O group) in mutant D76N. In addition, residue Arg<sup>45</sup> made one hydrogen bond with Glu<sup>47</sup> through Arg<sup>45</sup>NH1<sup>n2</sup>-Glu<sup>47</sup>O <sup>$\epsilon$ 2</sup>. This result suggested that residue Lys<sup>41</sup> involved in the formation of  $\beta$ -strand C, whereas residue Lys<sup>41</sup> in mutant D76N participated in the development of turn region due to the structural destabilization<sup>24</sup> influenced by the mutation D76 $\rightarrow$ N locally. Consequently, residue Lys<sup>41</sup> in mutant D76N unable to form the optimal hydrogen bonds with the residues Asn<sup>42</sup> and Glu<sup>44</sup>. It is an important to enlighten that mutant D76N has lost one main chain-main chain, main chain-side chain and side chain-side chain hydrogen bond when compared to that of WT- $\beta$ 2m. In addition, this typical  $\beta$ -scaffold was disrupted in mutant D76N; it was found that only 50% of residues were involved in the formation of  $\beta$ -strands during 50 ns simulation period. In case of WT- $\beta$ 2m, 52% of residues were necessitated in maintaining the regular  $\beta$ -strands. Furthermore, in the WT- $\beta$ 2m, 11% and 2% of residues were required to establish the bend and  $\beta$ -bridges respectively. In contrast to the  $\beta$ 2-m, 13% and 3% of the residues had a tendency to form the bend and  $\beta$ -bridges respectively in mutant D76N. Similarly, the percentage of residues involved in coil region (27%) was less in mutant D76N, where the WT- $\beta$ 2m exhibited as 28%. It was interesting to note that 7% of residues were involved in forming turn regions in both the WT- $\beta$ 2m and mutant D76N. Finally, DSSP result proposed that mutant D76N elucidated the difference on its evolutionary conserved secondary structure pattern, particularly after 27 ns simulation period.

## Conformational dynamics of mutant D76N

Although the observations given in above provide some promising information, how changes in the regular secondary structure elements of mutant D76N propagate into the structure leads to DRA. To probe the misfolding of mutant D76N to an amyloidogenic state can be accessed by the conformational dynamics driven by its main chain and side chain atoms. Moreover, DSSP result is exemplified that mutant D76N displayed a major secondary structure transition or swapping after 27 ns simulation period when compared to the WT- $\beta$ 2m, thus it is logic to analyze the structural consequences of mutant D76N, after 27 ns simulation period. In spite of being a highly amyloidogenic state, therefore, mutant D76N displayed an unfold characteristic under the conditions engaged that, in terms of the main chain, showed prominent perturbation is shown in Fig. 2A after 27 ns simulation period. Thus, changes in the main chain atomic position directly have an effect on the folding of mutant D76N (Fig. 1B) resulted to show diverse conformational dynamics when compared to the WT- $\beta$ 2m (Fig. 2A). Furthermore, to study in more detail about the dynamic behavior of mutant D76N cause amyloid formation can be assessed by calculating the average  $C_{\alpha}$  RMSD of each residue from the ensemble of trajectories (11501). The average  $C_{\alpha}$  RMSD of each residue in the WT- $\beta$ 2m (B) and its mutant D76N (C) were illustrated in Fig. 2. Comparison of the average  $C_{\alpha}$  RMSD results revealed that of 100 residues in mutant D76N, about 80 residues showed a greater deviation in their  $C_{\alpha}$  atoms than that of the WT- $\beta$ 2m. Indeed, a  $C_{\alpha}$  atom of the AB loop (residues: 13-20) in mutant D76N exhibited a major deviation, therefore exhibit diverse conformations. Consequently, AB loop in mutant D76N may perhaps trigger a conformational flexibility of  $\beta$ -strand A (residues: 6-11) by exposing the main chain atoms to the molecular surface as represented in Fig. 2C. It was very important to enlighten that aggregation forming propensity of  $\beta$ -strand D and DE loop in mutant D76N moved well into the interior of protein with respect to the  $C_{\alpha}$  atom. Similarly, a  $C_{\alpha}$  atom of residue Pro<sup>72</sup>, end regions of  $\beta$ -strand E (residues: 66-69) and the residues from 85-88 shifted into the hydrophobic core of the mutant D76N. Except, the above-mentioned residues in mutant D76N, remaining residues were exhibited major deviations in their  $C_{\alpha}$  atoms. Further, we have calculated the side chain RMSD of the WT- $\beta$ 2m and its mutant D76N is shown in Fig. 2D. It was found that mutant D76N elucidated not much deviation in its side chain atoms approximately up to 43 ns simulation period; after that, it exhibited less atomistic deviations by comparing with the WT- $\beta$ 2m. This result indicated that side chain atoms of most of the residues

in the mutant D76N rotated into the interior of protein, thus exhibited more hydrophobicity when compared to the WT- $\beta$ 2m. Finally, our main chain RMSD result was well correlated with the experimental report; it stated that backbone atoms of  $\beta$ -strands A, B, E, and F in mutant D76N displayed more conformational flexibility.<sup>25</sup> In addition, nuclear magnetic resonance (NMR) analysis stated that unfolded characteristic of  $\beta$ 2m by acid contained non-native residue structure stabilized by a disulphide bond involves the cluster of hydrophobic residues in two regions (29-51 and 58-79).<sup>10</sup> Therefore, the likely amyloid forming potential of a few residues attained its ability by disrupting the structure, mainly stretch of residues (60-70) that determines the aggregation propensity of the entire 99 residues. The aromatic residues include Phe<sup>56</sup>, Phe<sup>62</sup>, Tyr<sup>63</sup> and the hydrophobic residue Leu<sup>65</sup> are known as aggregation determining residues. In mutant D76N, C $_{\alpha}$  atom of two aromatic residues, Phe<sup>56</sup> and Phe<sup>62</sup> pointed into the interior of protein, where the C $_{\alpha}$  atom of another two residues, Tyr<sup>63</sup> and Leu<sup>65</sup> were moved to out. In addition, residues His<sup>51</sup>, Asp<sup>53</sup>, Asp<sup>59</sup>, and Trp<sup>60</sup> were reported as aggregation initiating residues.<sup>10</sup> Among the four aggregations initiating residues, C $_{\alpha}$  atoms of Asp<sup>53</sup>, Asp<sup>59</sup>, and Trp<sup>60</sup> rotated well into the interior of the protein, where a C $_{\alpha}$  atom of residue His<sup>51</sup> slightly exposed to the molecular surface. Along with the residue Lys<sup>91</sup>, the above-mentioned residues, His<sup>51</sup> and Asp<sup>53</sup> also reported as gatekeeper residues involved in avoiding the amyloid aggregation. In contrary to the WT- $\beta$ 2m, a C $_{\alpha}$  atom of residue Lys<sup>91</sup> highly exposed to the solvent in mutant D76N as illustrated in Fig. 2C. Examination of mutant D76N by atomic level suggested that differences in the structural properties of aggregation determining, initiating, and gatekeeper residues were critical in endowing the potential to form amyloid formation.

### Hydrogen bond analysis

Further, it is an important to relate the correlation between the unfolding characteristic of mutant D76N and its hydrogen bonds. DSSP result revealed that mutant D76N displayed a major secondary structure transition after 27 ns of the simulation period. To know the cause of the misfolding of mutant D76N, we have analyzed the hydrogen bonding interactions from the ensemble of trajectories (11501) obtained from the WT- $\beta$ 2m and mutant D76N after 27 ns simulation period. The hydrogen bond profile of the WT- $\beta$ 2m (black) and mutant D76N (green) is shown in Fig. 3A. Quantitative analysis of hydrogen bonds revealed that mutant D76N accounted more numbers of intra-molecular hydrogen bonds due to the altered atomic positions

of main chain (Fig. 2A) and side chain atoms (Fig. 2D). The WT- $\beta$ 2m contained the average hydrogen bond as 61.67, where, the mutant D76N exhibited as 63.21. Furthermore, intra-molecular hydrogen bonds were classified into three types based on its atomic interactions, such as main chain-main chain (MC-MC), main chain-side chain (MC-SC) and side chain-side chain (SC-SC) hydrogen bonds. Table 1 showed a total number of hydrogen bonds in the WT- $\beta$ 2m and mutant D76N and also its three different types of hydrogen bonds. It was found that mutant D76N encompassed reduced level of MC-MC hydrogen bonds as represented in Fig. 3B and Table 1. The observed reduction in MC-MC hydrogen bonds resulted to have an effect on the local folding of  $\beta$ -sheet that was observed from DSSP plot after 27 ns simulation period (Fig. 1). Thus, disruption in the MC-MC hydrogen bonds in mutant D76N able to reduce the percentages of residues involved in  $\beta$ -sheets. Principally, local folding of proteins such as  $\alpha$ -helix,  $\beta$ -sheet and sharp turns ( $\gamma$ - and  $\pi$  turns) were formed and stabilized by the MC-MC hydrogen bonds. In contrast to the MC-MC hydrogen bonds, mutant D76N had the ability to establish more numbers of MC-SC hydrogen bonds as illustrated in Fig. 3C and Table 1. The MC-SC hydrogen bonds, mainly mediated by the local interactions between the side chain acceptor/donor ('i') and main chain donor/acceptor atom within the window of i-5 to i+5. Formation of these short range hydrogen bonds was mainly controlled by the specific combination of backbone and side chain torsion angles.<sup>26</sup> Similar to the MC-MC hydrogen bonds, it was found that mutant D76N exhibited reduced levels of SC-SC hydrogen bonds as depicted in Fig. 3D and Table 1. From this result, we propose that changes in the atomic position of main chain and side chain atoms favored increasing the MC-SC hydrogen bonds and meanwhile supported to decrease the MC-MC and SC-SC hydrogen bonds in mutant D76N. Principally, the formation and stabilization of hydrogen bonds between the residues were mainly controlled by the backbone and the side chain torsion angles. Hence, unusual rotational freedom of those angles; therefore, directly disrupt the hydrogen bonds involved in the secondary structure profile resulted to affect the overall folding of mutant D76N as observed from DSSP plot.

### Hydrogen bond network analysis in loop regions

Several experimental studies reported that mutations in  $\beta$ 2m affect the stability and conformation of loop regions lead to cause amyloid aggregation.<sup>27,28</sup> Thus, it is an essential to describe the characteristic of loop regions that are present in mutant D76N may assist to reveal

the amyloidogenic propensity of mutant D76N. Fig. 4A showed the superimposed structure of the WT- $\beta$ 2m (green) and mutant D76N (blue) obtained at the 30 ns simulation period, it was found that mutant D76N displayed more frustration in their loops. In addition, numbers of residues involved in three potential loops are shown in Fig. 4B. In the WT- $\beta$ 2m, AB loop is formed by the continuous stretches of residues from 13 to 21 (HPAENGKSN), where in mutant D76N; it is organized by the residues of 13 to 20 (HPAENGKS). It was interesting to see that the length of DE loop is extremely reduced in mutant D76N. Except two residues, Asp<sup>59</sup>, and Trp<sup>60</sup>, remaining residues were participated in  $\beta$ -strand D. Therefore, length of  $\beta$ -strand D is increased quietly, thus able to cause the conformational plasticity of  $\beta$ -strand D. Similarly, the length of EF loop also reduced in mutant D76N as compared to the WT- $\beta$ 2m. Generally, changes in the flexibility of residues favored to alter the hydrogen bonding interactions in proteins.<sup>29</sup> Therefore, to explore the reason of why the loop regions in mutant D76N displayed wide variations in their flexibility and is thus revealed by calculating hydrogen bonds involved in those regions.

**AB loop:** In the WT- $\beta$ 2m, negatively charged residue Glu<sup>16</sup> linked to Ala<sup>15</sup> with hydrogen bond Glu<sup>16</sup>O <sup>$\epsilon$ 1</sup>-Ala<sup>15</sup>H and Glu<sup>16</sup>H. In addition, residue Asn<sup>17</sup> had the ability to make two hydrogen bonds with Lys<sup>19</sup> and Ser<sup>20</sup> (Asn<sup>17</sup>O <sup>$\delta$ 1</sup>-Lys<sup>19</sup>H and Ser<sup>20</sup>H). Where, the mutant D76N lost its two hydrogen bonds between the residues of Ala<sup>15</sup> and Glu<sup>16</sup>. At the same time, residue Asn<sup>17</sup> in mutant D76N tend to form the typical hydrogen bonds with Lys<sup>19</sup> and Ser<sup>20</sup> as observed in the WT- $\beta$ 2m. Thus, loss native residue contacts mediated by hydrogen bonds could be the rationale for increased flexibility of AB loop in mutant D76N.

**DE loop:** The DE loop has been repeatedly reported to be involved in amyloid formation and it's preceded by then  $\beta$ -strand D. Several experimental reports stated that changes in the geometry of DE loop could affect the folding and stability of various mutants which substantially enhanced the aggregation propensity.<sup>27,30</sup> Our MD results elucidated that DE loop in mutant D76N exhibited more conformational flexibility as it was not observed in the WT- $\beta$ 2m. The increased frustration of DE loop in mutant D76N enhanced the mobility of  $\beta$ -strands, D and E as internal rigid body domain motion. Further, to unveil the local frustration of DE loop in mutant D76N, we have investigated the amount of hydrogen bonds involved in the DE loop region. The residue Phe<sup>56</sup> in the WT- $\beta$ 2m made contact to Ser<sup>57</sup> with hydrogen bond Phe<sup>56</sup>O-Ser<sup>57</sup>H <sup>$\delta$</sup> . In addition, two more hydrogen bonds were observed between the residues of Lys<sup>58</sup>, Trp<sup>60</sup> and Ser<sup>61</sup> (Lys<sup>58</sup>O-Trp<sup>60</sup>H and Lys<sup>58</sup>O-Ser<sup>61</sup>H). In case of mutant D76N, above-mentioned hydrogen

bonds were completely disrupted, where mutant D76N tend to form a single hydrogen bond by the residue Asp<sup>59</sup> itself (Asp<sup>59</sup>H-Asp<sup>59</sup>O<sup>δ1</sup>). As we specified above, three residues such as Phe<sup>56</sup>, Ser<sup>57</sup> and Lys<sup>58</sup> were participated in β-strand D, thus, residue Asp<sup>59</sup> moved towards to the interior. In addition, the aromatic side chain of residue Trp<sup>60</sup> completely exposed to the molecular surface, thus mutant D76N unable to form essential hydrogen bonds in the DE loop and exhibited more conformational flexibility. The information here reported furthermore induced strain in the DE loop appeared to affect not only the folding and also folding pathways of the mutant D76N.

**EF loop:** In the WT-β2m, residue Thr<sup>71</sup> made interaction with Thr<sup>73</sup> through hydrogen bond between Thr<sup>71</sup>O-Thr<sup>73</sup>H. In turn, residue Thr<sup>73</sup> linked to Asp<sup>76</sup> with hydrogen bond Thr<sup>73</sup>H<sup>γ1</sup>-Asp<sup>76</sup>O<sup>δ1</sup>. At last, residue Glu<sup>74</sup> able to formed a hydrogen bond with Lys<sup>75</sup> through Glu<sup>74</sup>O<sup>ε1</sup>-Lys<sup>75</sup>H. On the other hand, mutant D76N made a single hydrogen bond between the residue of Thr<sup>73</sup> and Asn<sup>76</sup> (Thr<sup>73</sup>O<sup>γ1</sup>-Asn<sup>76</sup>H<sup>δ2</sup>) and it lost hydrogen bond between Thr<sup>71</sup>-Thr<sup>73</sup> and Glu<sup>74</sup>-Lys<sup>75</sup>. Thus, mutation D76N in EF loop exhibited more frustration due to the loss of molecular contact between the residues mediated by the hydrogen bonds.

### Hydrogen bonding relationship of residue 76 in the WT-β2m and mutant D76N

Further, it is an essential to explore the local structural changes influenced by the substituted residue asparagine (D76→N) in mutant D76N. In the WT-β2m, residue Asn<sup>42</sup> was strongly allied to Asp<sup>76</sup> with the hydrogen bond Asn<sup>42</sup>H<sup>δ2</sup>-Asp<sup>76</sup>O<sup>δ2</sup>. In turn, residue Asn<sup>42</sup> was hydrogen bonded with Glu<sup>77</sup> via two almost conserved hydrogen bonds (Asn<sup>42</sup>H-Glu<sup>77</sup>O and Asn<sup>42</sup>O<sup>δ1</sup>-Glu<sup>77</sup>H). In addition, residue Asp<sup>76</sup> interacted to Thr<sup>73</sup> with hydrogen bond between Thr<sup>73</sup>H<sup>γ1</sup>-Asp<sup>76</sup>O<sup>δ1</sup>, all these hydrogen bonds were found at 30 ns simulation, and these results were well correlated with the experimental report<sup>14</sup> as illustrated in Fig. 4C. In case of mutant D76N, only one hydrogen bond was noticed between the residues of Asn<sup>76</sup> and Thr<sup>73</sup> (Thr<sup>73</sup>O<sup>γ1</sup>-Asn<sup>76</sup>H<sup>δ2</sup>). Except, the above-stated hydrogen bond, mutant D76N totally lost its essential hydrogen bonds between the residues Asn<sup>42</sup>-Asp<sup>76</sup> and Asn<sup>42</sup>-Asp<sup>77</sup>; it suggested that substituted residue Asn<sup>76</sup> exemplified more variations in its structure locally.

## Amyloid propensity of $\beta$ -strand D

How could the structural rearrangement of mutant D76N facilitate self-assembly? Several studies reported that amyloid formation mainly achieved by the disruptions of  $\beta$ -strand D.<sup>5,9,31</sup> Hence, to understand the mechanism behind the amyloid forming ability of  $\beta$ -strand D can be analyzed at the atomistic level intensely. In contrast to the WT- $\beta$ 2m, the length of  $\beta$ -strand D observed in the mutant D76N was high in the majority of trajectories after 27 ns simulation and also displayed a short  $\beta$ -strand C in that simulation period. The length of  $\beta$ -strand was primarily controlled by the presence of hydrogen bonds between the main chain atoms of two  $\beta$ -strands resulted to form the local folding of  $\beta$ -sheet. In case of WT- $\beta$ 2m, a total of five hydrogen bonds were observed between the  $\beta$ -strand D and E mediated by the main chain atoms of them. In the WT- $\beta$ 2m, residue Glu<sup>50</sup> connected to Tyr<sup>67</sup> by a hydrogen bond between Glu<sup>50</sup>H-Tyr<sup>67</sup>O. In addition, residue His<sup>51</sup> mediated an interaction to Tyr<sup>67</sup> by forming two hydrogen bonds between His<sup>51</sup>H-Tyr<sup>67</sup>O and vice versa. At last, residue Ser<sup>55</sup> linked to Tyr<sup>63</sup> with hydrogen bond Ser<sup>55</sup>H-Tyr<sup>63</sup>O and vice versa. On the other hand, mutant D76N had the ability to show more numbers of hydrogen bonds between the  $\beta$ -strand D and E, because the numbers of residues involved in  $\beta$ -strand D were considerably higher. In the mutant D76N, residue Glu<sup>50</sup> had the tendency to form a hydrogen bond with Tyr<sup>67</sup> through Glu<sup>50</sup> H-Tyr<sup>67</sup>O. In addition, residue His<sup>51</sup> able to associate with Tyr<sup>67</sup> by establishing hydrogen bond between His<sup>51</sup>O-Tyr<sup>67</sup>H. More importantly, the gatekeeper residue Asp<sup>53</sup> linked to Leu<sup>65</sup> through two hydrogen bonds Asp<sup>53</sup>O-Leu<sup>65</sup>H and vice versa. Besides, residue Ser<sup>55</sup> had the propensity to form two hydrogen bonds with Tyr<sup>63</sup> through Ser<sup>55</sup>H-Tyr<sup>63</sup>O and vice versa. Finally, the residue Lys<sup>58</sup> connected to Ser<sup>61</sup> with hydrogen bond Lys<sup>58</sup>O-Ser<sup>61</sup>H. On contrary to the WT- $\beta$ 2m, the gate-keeper residue Asp<sup>53</sup> rotated inward, thus able to form hydrogen bonds with the residue Leu<sup>65</sup> resulting in loss of  $\beta$ -bulge in mutant D76N, forming a more aggregation-competent surface.

Several experimental studies reported that  $\beta$ -strand D and DE loop were promising regions for mediating interactions between the two monomers resulting in amyloid aggregation.<sup>5,9,30,31</sup> Thus, it is necessary to characterize the residues that cooperate in the amyloid formation. The four residues such as Phe<sup>56</sup>, Phe<sup>62</sup>, Tyr<sup>63</sup>, and Leu<sup>65</sup> identified to be an aggregation determining residues<sup>10</sup> as mentioned above showed substantial structural rearrangement in mutant D76N. Indeed, apart from those residues, residues His<sup>51</sup>, Asp<sup>53</sup>, Asp<sup>59</sup>

and Trp<sup>60</sup> play a potential role in initiating amyloid formation. In particular, mutations in DE loop have been involved in amyloid formation by affecting the geometry and the stability of DE loop, principally the replacement of Asp<sup>59</sup> and Trp<sup>60</sup>. In the mutant D76N, only two residues, Asp<sup>59</sup>, and Trp<sup>60</sup> were involved in establishing of DE loop. Thus, to emphasize the possible mechanism behind the amyloid formation mediated by the interactions between the two monomers can be obtained by analyzing the relative solvent accessibility. It was interesting to portray that out of seven residues (His<sup>51</sup>, Phe<sup>56</sup>, Asp<sup>59</sup>, Trp<sup>60</sup>, Phe<sup>62</sup>, Tyr<sup>63</sup>, and Leu<sup>65</sup>), backbone and side chain atom of residue His<sup>51</sup> highly exposed to the solvent. This was because of an inward characteristic of side chain atoms of residue Glu<sup>50</sup> which facilitated to move the residue His<sup>51</sup> to out, especially the side chain of His<sup>51</sup> more protrude to the solvent. Where, the other six residues exhibited relatively less accessible to the solvent when compared to the residues in WT- $\beta$ 2m. In particular, residue Asp<sup>59</sup> moved well into the interior, and it lost its ideal hydrophilic characteristic as represented in Table 2. This analysis explored that both the aggregations determining and initiating residues in mutant D76N exhibited more hydrophobicity and probably renders it more prone to edge-to-edge  $\beta$ -aggregation by forming the inter-subunit  $\beta$ -hydrogen bonds.

Regular  $\beta$ -sheets are intrinsically aggregation-prone because hydrogen bonding of one  $\beta$ -strand with any other  $\beta$ -strand is available. Principally,  $\beta$ -sheet proteins can be designed to evade this interaction by the use of various blocking features to avoid edge-to-edge  $\beta$ -aggregation. For  $\beta$ -scaffold proteins, to avoid aggregation is the presence of a charged side chain of Asp, Arg, Glu, Lys, and His. In the  $\beta$ -scaffold proteins, such charged residues can form a hydrophobic core through packing interactions, while exposing their charged side chains to the solvent helps to prevent edge-to-edge association by electrostatic repulsion between these exposed charges. A typical  $\beta$ 2m comprised of four edge  $\beta$ -strands:  $\beta$ -strands A and G form one pair at one side of the protein and  $\beta$ -strands C and D form a second pair at the opposite side. Focus on these susceptible  $\beta$ -strands,  $\beta$ 2m contained several gatekeeper residues including His<sup>51</sup> to protect  $\beta$ -strand D and Lys<sup>91</sup> to protect  $\beta$ -strand G, thus avoiding edge-to-edge aggregation. Again, the presence of  $\beta$ -bulge in the WT- $\beta$ 2m was able to bend the  $\beta$ -strands are another approach to avoiding edge-to-edge aggregation by altering the geometry and brought out the twist of the associating  $\beta$ -strands. Remarkably, the residue Asp<sup>53</sup> is a key residue that has usually forming a  $\beta$ -bulge by pointing outwards and altering the geometry of the  $\beta$ -strand D. The accessible surface area of three

gatekeeper residues in the WT- $\beta$ 2m and its mutant D76N were represented in Table 3. In the mutant D76N, the backbone of gatekeeper residue Lys<sup>91</sup> was highly exposed to the solvent, where side chain atoms moved well into the interior of protein than that of residue Lys<sup>91</sup> in the WT- $\beta$ 2m. Moreover, residue His<sup>51</sup> in mutant D76N pointed outward to the molecular surface, and it was totally opposite to the WT- $\beta$ 2m resulting in loss of gatekeeper function, thus it may initiate edge-to-edge aggregation of  $\beta$ -strand D with the  $\beta$ -strand of another monomer. In case of mutant D76N, the gatekeeper residue Asp<sup>53</sup> rotated into the hydrophobic core of the protein consequently made a hydrogen bond with residue Leu<sup>65</sup> and gets lost its edge-to-edge aggregation preventing aspect. Subsequently, the mutant D76N exhibited a long  $\beta$ -strand D that was probably more prone to an inter-molecular pairing of  $\beta$ -strands. In particular, the backbone atoms of residue Asp<sup>53</sup> remained in the same atomic position, while the side chain atoms were completely buried in the interior of mutant D76N by comparing with the WT- $\beta$ 2m. This conspicuous rearrangement of mutant D76N causes a significant increase in the surface hydrophobicity of the protein, apparently reducing its intrinsic solubility. The solvent accessibility profile of the WT- $\beta$ 2m (A) and its mutant D76N (B) as illustrated in Supplementary Fig. 1. It should be noted that  $\beta$ -strand A (residues: 6 to 11) exhibited more hydrophobic characteristic than that of WT- $\beta$ 2m. Concomitantly, the inward rotation of residues Asp<sup>53</sup> and Asp<sup>59</sup> can also contribute to increasing of the surface hydrophobicity of mutant D76N. This result was well correlated with the side chain RMSD of mutant D76N, most of the side chain atom of residues was considerably moved well into the interior of the protein and exhibited significant difference on its surface hydrophobicity.

### **Conformational drift by principal component analysis (PCA)**

Whatever cause made by the mutant D76N on its structure, it is clear that further conformational changes are required to promote assembly of adjacent monomers into amyloid fibrils, for which PCA analysis was performed. Proteins are frequently carried out their specific function by collective atomic motions, if changes occurred in its inherent motion directly to cause various diseases. Therefore, it is requisite to know the concerted motion of mutant D76N to describe the possible mechanism underlying the amyloid formation. PCA is a method often used to get insights into the correlated motion of protein into a few principal motions, each of which is characterized by an eigenvector and eigenvalues. Applying, MD trajectories obtained

after 27 ns simulation period to the PCA aid to extract the collective atomic motion of the WT- $\beta$ 2m (black) and its mutant D76N (green) in an essential subspace as illustrated in Fig. 5. It showed the conformational sampling of WT- $\beta$ 2m (Fig. 5A) in the essential subspace by projecting the  $C_{\alpha}$  atom, showing a cluster representative of explored tertiary conformations along the eigenvector 1 and 2. Such a collective motion represented by these two eigenvector was mainly due to the energetic contribution of its corresponding eigenvalues, thus describing its global motion. Fig. 5B and C described the residue displacements of WT- $\beta$ 2m (black) described by the eigenvector 1 and 2 respectively. It was found that the bend and coil region exposed more frustration occurred between the  $\beta$ -strands A and B along the eigenvector 1. Moreover, these regions exhibited a trivial mobility along the eigenvector 2. In contrast to the WT- $\beta$ 2m, a clear separation of conformational motion was observed in mutant D76N (green) along the represented eigenvector as depicted in Fig. 5D. Contrary to the WT- $\beta$ 2m, it can be seen that mutant D76N elucidated the difference in its concerted motion, alteration in conformational dynamics of mutant D76N could be of an unusual flexibility of each residue as illustrated in Fig. 5E and F along the eigenvector 1 and 2 respectively. Consequently, residues in the bend and coil regions in mutant D76N showed less frustration that were occurred between the  $\beta$ -strand of A and B when compared to the WT- $\beta$ 2m. Also, it was interesting to see that the residue Lys<sup>58</sup> occurred at the end of  $\beta$ -strand D followed by the residues Asp<sup>59</sup> and Trp<sup>60</sup> from DE loop displayed an immense level of flexibility, and it could not be noticed in the WT- $\beta$ 2m. The increased local frustration of those residues in mutant D76N might be the loss of MC-MC hydrogen bonds between the residues Lys<sup>58</sup> and Trp<sup>60</sup>. In addition, the bend and loop regions positioned between the  $\beta$ -strands A and B that portrayed more residue flexibility along the eigenvector 2, which could not be observed in the WT- $\beta$ 2m. Thus, local fluctuations of one or more regions in the mutant D76N aid to generate the precursors that are prone to spontaneous self-assembly are also at the origin of DRA.<sup>7</sup> Hence, this unusual conformational dynamics of mutant D76N could be the starting precursor for the development of amyloid aggregation.

To distinguish the conformational states of WT- $\beta$ 2m and its mutant D76N associated with energy, the Gibbs free energy landscape (FEL) was calculated using the first two principal components as reaction coordinates. The FEL plot was not accurate to explain the energy barriers and to portray the location of metastable and barriers in detail. Nevertheless, it could provide the valuable information to describe the conformational landscape accessible to the

molecule in the ensemble of trajectories. Fig. 6 showed the FEL of WT- $\beta$ 2m (A) and its mutant D76N (B) along the two principal components. The global energy minima of conformational states were designated by the blue color; while the orange color signified the meta-stable states. The WT- $\beta$ 2m displayed a well-defined large single global energy minima basin associated with its conformational state. Whereas, misfolding of mutant D76N exhibited a large and also a small global energy minima basin associated with its conformational state. Misfolding of mutant D76N may cause by distortion in the ideal turn region between the  $\beta$ -strand C and C' resulted to adopt thermodynamically unfavorable conformational states (Fig. 6B). Because, turn region has been considered to be a nucleation site for folding of proteins, it may play a passive role in structure formation, in contrast to an active role in which the turn initiates and promotes folding. The unfolded states of mutant D76N was not randomly sampled in all the areas of  $\phi$  (phi) and  $\psi$  (psi) angles space, but those conformational states were best described as populating an interconverting set of native-like local conformations, with an occasional long range native-like interactions or hydrophobic clusters.

## Conclusion

In summary, the evidences reported here suggest that mutation in EF loop (D76N) that cause misfolding of mutant D76N by disrupting the regular secondary structure elements. The repacking of mutant D76N could be of wide variation in the atomistic position of main chain and side chain that showed the combined effect on the molecular contacts mediated by hydrogen bonds and the protein misfolding. Thus, both the amyloid determining and initiating residues are disordered notably in mutant D76N that may favor to produce a precursor for amyloid aggregation. More importantly, the gatekeeper residues, His<sup>51</sup>, Asp<sup>53</sup>, and Lys<sup>91</sup> showed more structural consequences when compared to the residues in the WT- $\beta$ 2m. In addition, amyloidogenic state inducing ability of  $\beta$ -strand D in mutant D76N showed a long  $\beta$ -strand D thus provide, the more aggregation-competent surface due to the inward rotation of residue, Asp<sup>53</sup>. Besides, residues in DE loop exhibited more frustration along the first eigenvector could be the cause of change in the conformational dynamics of mutant D76N lead to produce more specific aggregation-competent species which facilitate self-assembly. Finally, these results unlock the way to get insights into the molecular mechanism of amyloid aggregation propensity of mutant D76N and were good agreement with the experimental report.

## Acknowledgement

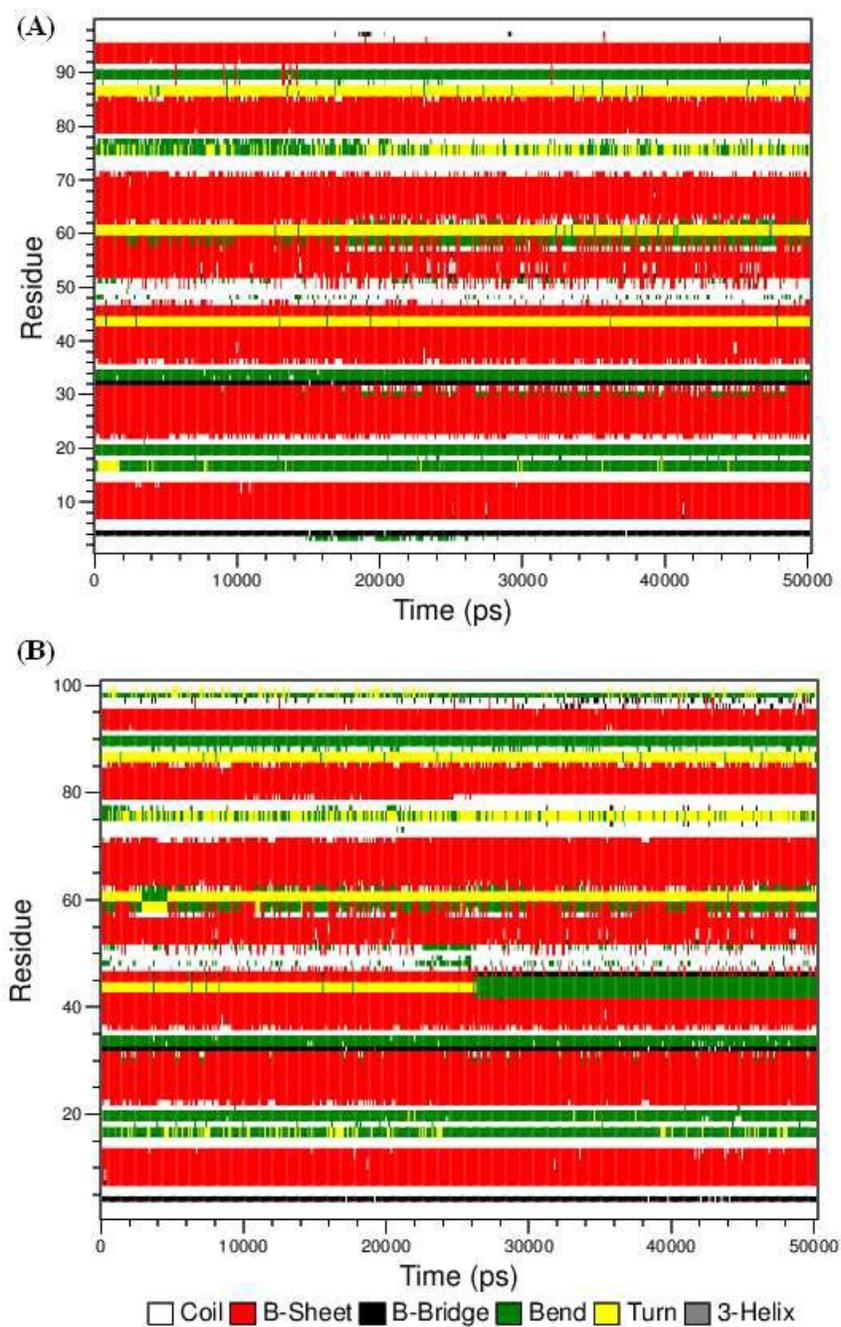
We thank the management of VIT University for providing facilities to carry out this work.

## References

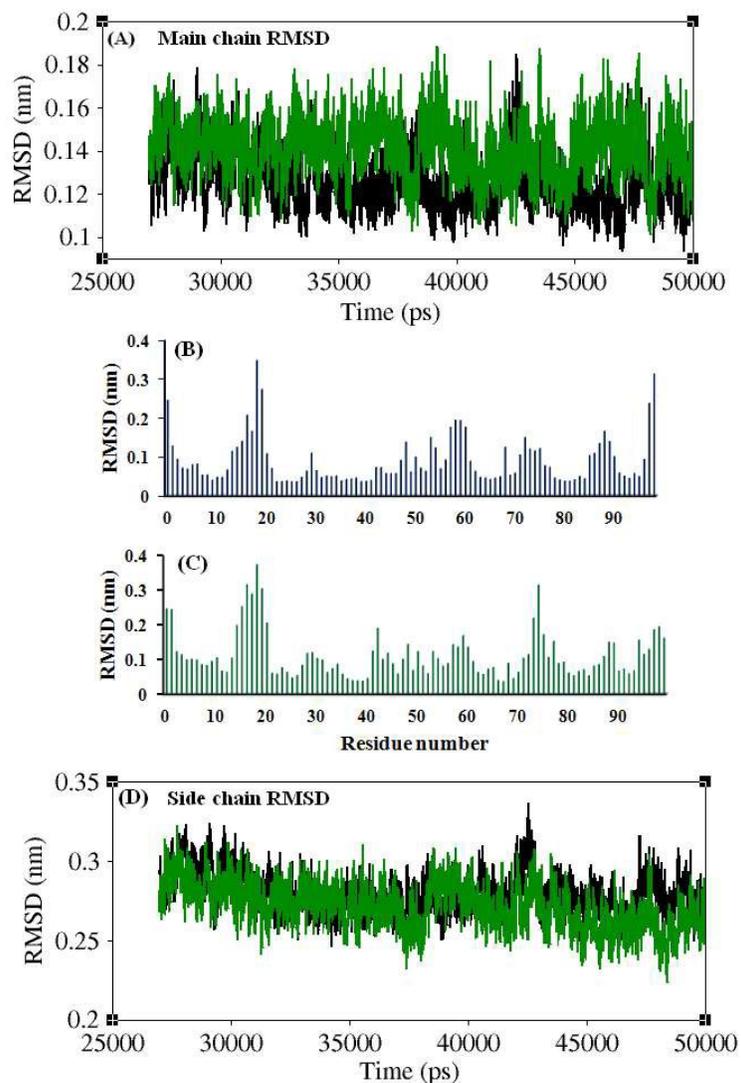
1. C. M. Dobson, *Nature*, 2003, **426**, 884-890.
2. F. U. Hartl, *Nature*, 1996, **381**, 571-579.
3. T. Eichner and S.E. Radford, *Mol. Cell*, 2011, **43**, 8-18.
4. F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.*, 2006, **75**, 333-366.
5. S. Park and J. G. Saven, *Protein Sci.*, 2006, **15**, 200-207.
6. F. Gejyo, T. Yamada, S. Odani, Y. Nakagawa, M. Arakawa, T. Kunitomo, H. Kataoka, M. Suzuki, Y. Hirasawa, T. Shirahama, A. S. Cohen and K. Schmid, *Biochem. Biophys. Res. Commun.*, 1985, **129**, 701-706.
7. T. Eichner and S. E. Radford, *FEBS J.*, 2011, **278**, 3868-3883.
8. S. G. Estácio, H. Krobath, D. Vila-Viçosa, M. Machuqueiro, E. I. Shakhnovich and P. F. N. Faisca, *PLoS Computational Biology*, 2014, **10**, e1003606 (2014).
9. M. Colombo, M. de Rosa, V. Bellotti, S. Ricagno and M. Bolognesi, *FEBS J.*, 2012, **279**, 1131-1143.
10. G. W. Platt and S. E. Radford, *FEBS Lett.*, 2009, **583**, 2623-9.
11. S. Giorgetti, A. Rossi, P. Mangione, S. Raimondi and S. Marini, *Protein Sci.*, 2005, **14**, 696-702.
12. N. H. Heegaard, *Amyloid*, 2009, **16**, 151-73.
13. J. P. Hodkinson, S. E. Radford and A. E. Ashcroft, *Rapid Commun. Mass Spectrom.*, 2012, **26**, 1783-92.
14. 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. Maisonobe, 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 and V. Bellotti, *N. Engl. J. Med.*, 2012, **366**, 2276-83.
15. H. M. Berman, J. Westbrook, Z. Feng, G. G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
16. K. Berendsen, H. J. C. Van der Spoel and R. Van Drunen, *Comput. Phys. Commun.*, 1995, **91**, 43-56.
17. M. Kaminski, G. A. Friesner, R. A. Tirado-Rives and L. W. Jorgensen, *J. Phys. Chem. B*, 2001, **105**, 6474-6487.
18. N. Darden, T. York, and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
19. O. Hess, B. Kutzner, D. Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435-447.
20. P. Chandrasekaran and R. Rajasekaran, *Mol. Biosyst.*, 2014, **10**, 1869-1880.
21. S. Haider, G. N. Parkinson and S. Neidle, *Biophys. J.*, 2008, **95**, 296-311.
22. M. Laberge and T. Yonetani, *Biophys. J.*, 2008, **94**, 2737-2751.
23. J. S. Hub and B. L. de Groot, *PLoS Comput. Biol.*, 2009, **5**, doi:10.1371/journal.pcbi.1000480.
24. L. Halabelian, S. Ricagno, S. Giorgetti, C. Santambrogio, A. Barbiroli, S. Pellegrino, A. Achour, R. Grandori, L. Marchese, S. Raimondi, P. P. Mangione, G. Esposito, R. Al-Shawi, J. P. Simons, I. Speck, M. Stoppini, M. Bolognesi and V. Bellotti, *J. Biol. Chem.*, 2014, **289**, 3318-3327.
25. P. P. Mangione, G. Esposito, A. Relini, S. Raimondi, R. Porcari, S. Giorgetti, A. Corazza, F. Fogolari, A. Penco, Y. Goto, Y. H. Lee, H. Yagi, C. Cecconi, M. M. Naqvi, J. D. Gillmore, P.

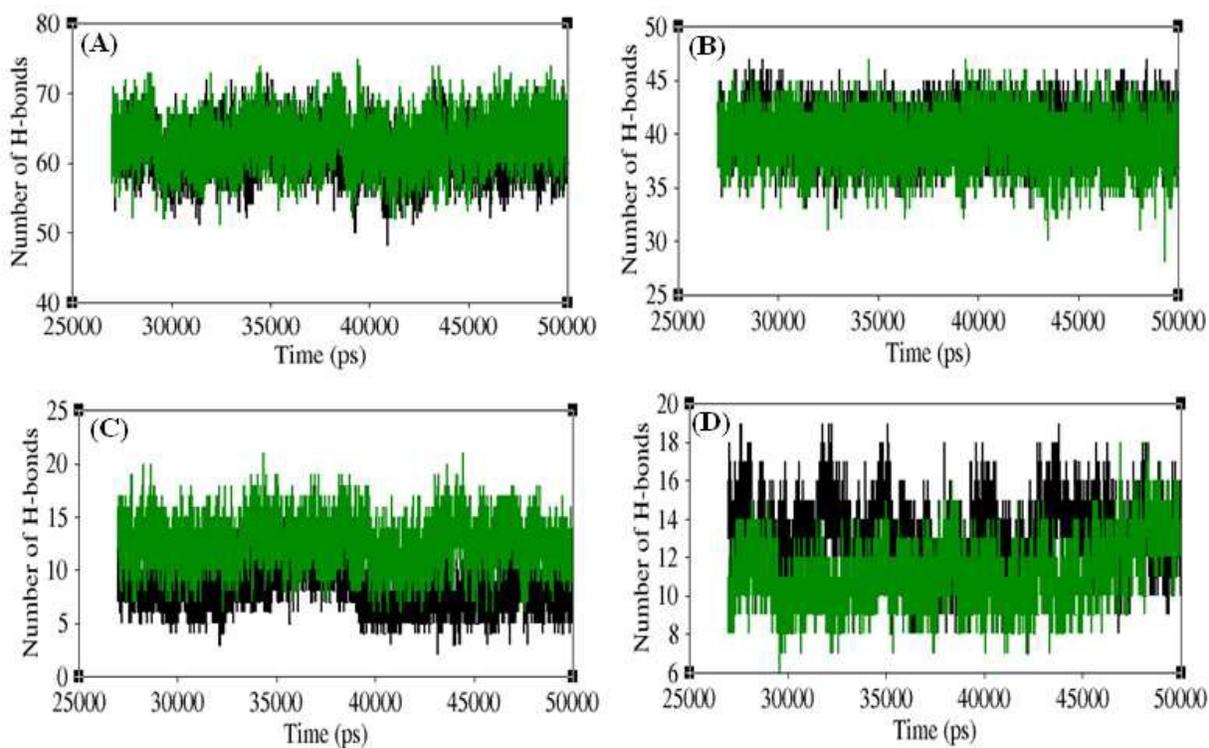
- N. Hawkins, F. Chiti, R. Rolandi, G. W. Taylor, M. B. Pepys, M. Stoppini and V. Bellotti, *J. Biol. Chem.*, 2013, **288**, 30917-30.
26. N. Eswar and C. Ramakrishnan, *Protein Eng.*, 2000, **13**, 227-38.
27. A. Natalello, A. Relini, A. Penco, L. Halabelian, M. Bolognesi, S. M. Doglia, S. Ricagno S, *PLoS One*, 2015, **10**, e0122449.
28. C. Santambrogio, S. Ricagno, M. Colombo, A. Barbiroli, F. Bonomi, V. Bellotti, M. Bolognesi and R. Grandori, *Protein Sci.*, 2010, **19**, 1386-94.
29. H. Gohlke, L. A. Kuhn and D. A. Case, *Proteins*, 2004, **56**, 322-37.
30. S. Ricagno, M. Colombo, M. de Rosaz, E. Sangiovanni, S. Giorgetti, S. Raimondi, V. Bellotti and M. Bolognesi, *Biochem. Biophys. Res. Commun.*, 2008, **377**, 146-150
31. A. C. Leney, C. L. Pashley, C. A. Scarff, S. E. Radford and A. E. Ashcroft, *Mol. Biosyst.*, 2014 **10**, 412-420.



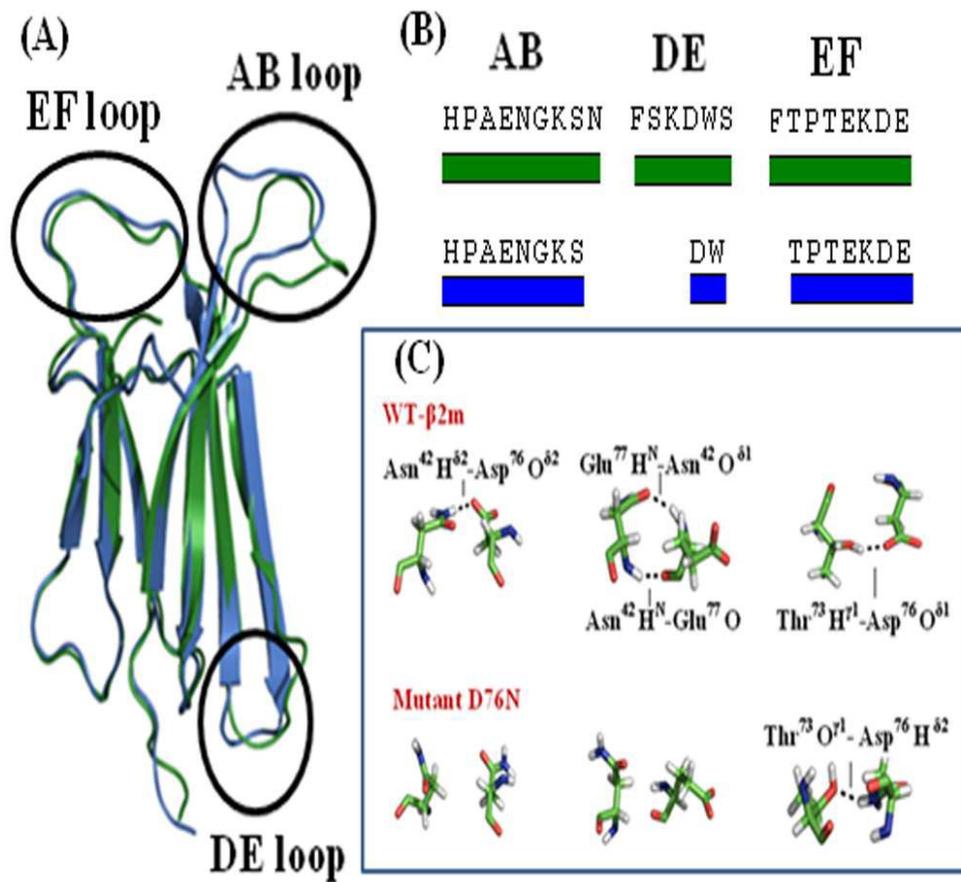
**Fig. 1** Time evolution of the secondary structure profile observed in the WT-β2m (A) and its mutant D76N (B) during 50 ns simulation period.



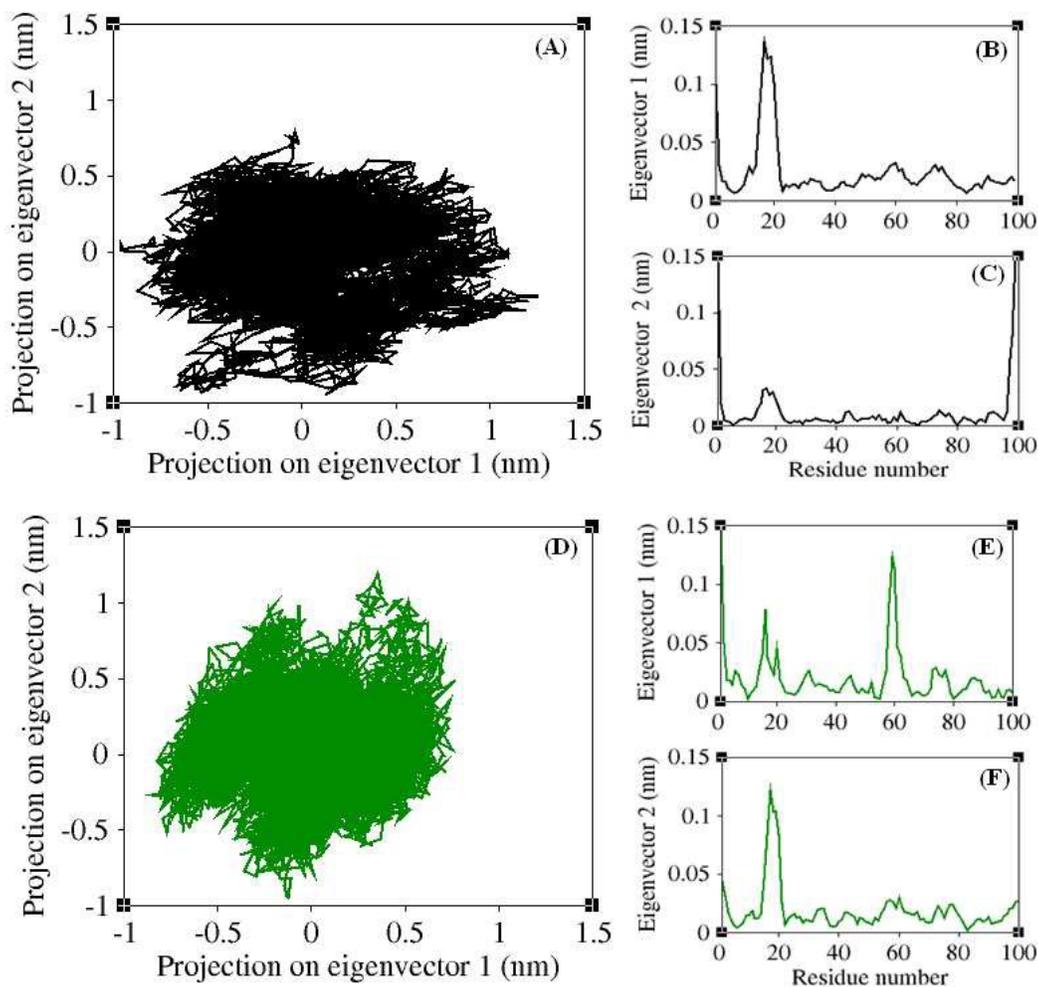
**Fig. 2** (A) Main chain RMSD of the WT-β2m (black) and its mutant D76N (green) after 27 ns of simulation period. The C<sub>α</sub> RMSD profile of the WT-β2m (B) and its mutant D76N (C). (D) Side chain RMSD of the WT-β2m (black) and its mutant D76N (green)



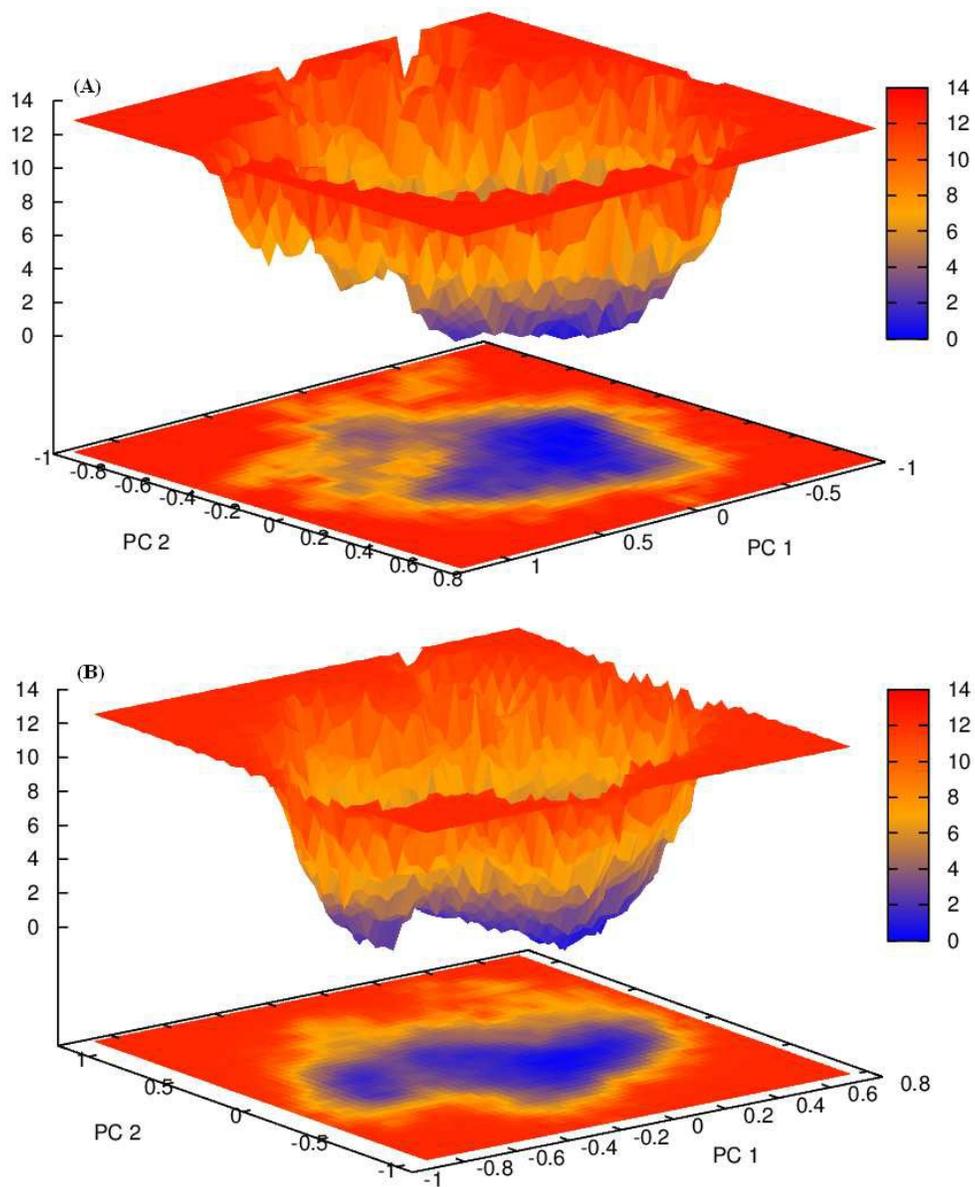
**Fig. 3** (A) Intra-molecular hydrogen bonds involved in the WT-β2m (black) and its mutant D76N (green). (B) Main chain-Main chain, (C) Main chain-Side chain and, (D) Side chain-Side chain hydrogen bonds of the WT- β2m domain and its mutant D76N



**Fig. 4** (A) Superimposed structures of the WT-β2m (green) and its mutant D76N (blue). (B) Residues involved in the loop regions of the WT-β2m (green) and its mutant D76N (blue). (C) Hydrogen bonding relationship of residue 76 in the WT-β2m and its mutant D76N



**Fig. 5** Projection of the motion of the protein in the essential subspace along the eigenvector 1 and 2 for the WT- $\beta$ 2m (A) domain and its mutants D76N (D). B and C Represented the residue displacements of the WT- $\beta$ 2m along the eigenvector 1 and 2 respectively. E and F Illustrated the residue displacements of the mutant D76N along the eigenvector 1 and 2 respectively.



**Fig. 6** FEL of the WT-β2m (A) and its mutant D76N (B) along the two principal components

**Table 1** Numbers of hydrogen bonds observed in the WT- $\beta$ 2m its mutant D76N

	<b>Total H</b>	<b>M-M H</b>	<b>M-S H</b>	<b>S-S H</b>
WT- $\beta$ 2m	709271	463532	96720	149019
Mutant D76N	72,7074	455029	143669	128376

**Table 2** Relative accessible surface area of aggregation determining residues in the WT- $\beta$ 2m and its mutant D76N

WT- $\beta$ 2m				Mutant D76N		
Residues	Tot-Ass	B-Ass	S-Ass	Tot-Ass	B-Ass	S-Ass
Phe <sup>56</sup>	36.35	4.37	31.97	34.84	7.93	26.91
Asp <sup>59</sup>	34.49	8.40	26.08	18.64	0.21	18.43
Trp <sup>60</sup>	38.70	2.63	36.07	31.18	0.00	31.18
Phe <sup>62</sup>	11.39	0.21	11.18	7.33	0.00	7.33
Tyr <sup>63</sup>	19.26	0.00	19.26	18.50	0.52	17.98
Leu <sup>65</sup>	13.28	0.00	13.28	10.07	0.46	9.61

**Table 3** Relative accessible surface area of the gatekeeper residues in the WT- $\beta$ 2m and its mutant D76N

WT- $\beta$ 2m				Mutant D76N		
Gatekeeper residues	Tot-Ass	B-Ass	S-Ass	Tot-Ass	B-Ass	S-Ass
His <sup>51</sup>	21.20	0.69	20.50	29.22	1.74	27.47
Asp <sup>53</sup>	23.85	3.00	20.85	17.84	3.14	14.70
Lys <sup>91</sup>	33.28	0.69	32.59	23.45	4.75	18.69