Metadynamics update

Maria Celeste Maschio[1] and stefano.corni[1]

[1]Affiliation not available

September 28, 2018

Abstract

In this work we want to investigate with Metadynamics how the presence of the mutations affects the free energy barriers that characterize the cis to trans transition barrier of the Proline32 isomerisation.

Introduction

Among all the amino acids that are the basics of protein composition, Proline is the most unusual one: it has a cyclic side chain that is bound to the amide nitrogen of the backbone

[1]

The isomerisation of Pro32 from its native cis to the non-native trans configuration is indicated as a trigger factor for b2m misfolding and the subsequent amyloidogenical mechanism.

Cis-configuration for the Proline 32 (the other 4 proline are in trans configuration) is required in order to maintain the soluble native state of B2M wild type, while trans-configuration is required for amyloid elongation at neutral pH. From 10 years ago work [2] , we know that assembly is shown to involve the transient formation of a non-native monomer containing a trans backbone conformation. subsequently there is the formation of dimeric species and higher aggregates that accumulate before the development of amyloid fibrils. (This is the milestone for Radford and co to say that Pro32 isomerisation)

In this work, fluorinated proline derivatives are used as probe to investigate the structure-function relationships in b2m. Such substitution alters the equilbrium population of trans and cis isomers via stereoelectronics effects and also lowers the barrier for the isomerisation.

Previous studies showed that trans to cis prolyl bond isomerisation of Pro32 is the rate-limiting step in the protein folding mechanism.

Experimental works suggested that fibril formation occurs via metastable, partially unfolded protein conformers. These species are obtained when Pro32, starting from the normal cis-conformation, slowly adapts to trans geometry, leading to destabilizing effects on the structure of the protein: in particular there is an exposition of the hydrophobic structure and also a rearrangement of the D strand that lead to intermolecular aggregation.

Beta-2 microglobulin (β2m) is part of the Major Histocompatibility Complex Class I (MHC I) and when monomeric becomes an aggregation prone protein that is responsible for a human disorder known as dialysis-related amyloidosis. In 2012 Valleix et al. described a new familial systemic amyloidosis: an unreported β2m mutant (D76N) is the etiological agent of such disease. Main symptoms were chronic diarrhea, loss of weight and polyneuropathy: large amyloid deposits were found in internal organs. From the biophysical point of view, the D76N β2m is much less stable and more amyloidogenic than wt β2m; however, its crystal structure reveals very minor conformational changes compared with the wt protein [3]

The six Asp in β2m sequence have been mutated to Asn (D-to-N mutants): their thermal stability and aggregation propensity closely resemble those of wt β2m; from the structural point of view only the D38N mutation triggers non-negligible conformational changes, however the structural rearrangements observed in the C-D region do not correlate with an increased aggregation propensity . In summary, these data indicate that the loss of a negative charge is not sufficient to explain the amyloidogenic properties of the D76N mutant. The second set of mutants focuses on the position 76: four mutants have been prepared: D76A, D76H, D76E, D76K. Although data are preliminary and their analysis is still in progress, it seems that any substitution in the 76 position affects to great extent both stability and amyloidogenicity of β2m in solution without major impact on its structure.

Position 76 is very crucial to determine β2m thermal stability and aggregation propensity and only an extensive interaction such the one within the MHC I is able to compensate the loss in stability due to the D76N mutation.

Proline dipeptide in water

Molteni: Ca-Cd-O-Ca

Celeste: Cd-Ca-O-Ca

wild type

w60g

d76n

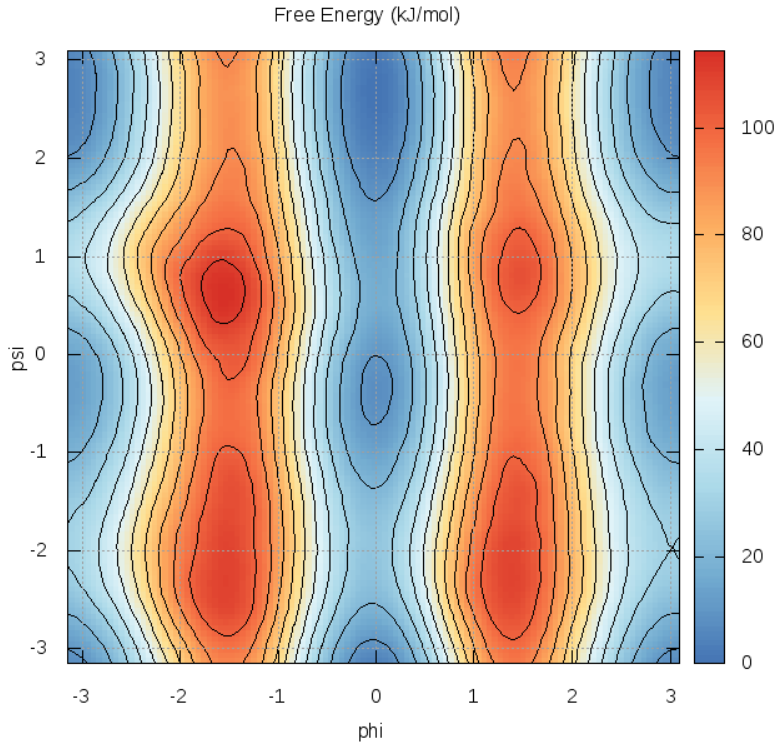

Figure 1: Free energy surface, proline dipetide in SPC explicit water, OPLS-AA ff

The behavior of the proteins during the classical MD trajectories

OBSERVING TRANS-PRO32 MD

The wild type

In this dynamics, the most relevant portion is the AB loop, showing high mobility in water. The apical part involving BC, DE e FG loop maintains an almost constant conformation during time.

Evaluating the RDF for the Oxygen of the dihedral angle defining the Proline isomerisation, it can be said that neither from the outside nor the inside with water molecules inside the structure there is a relevant interaction with the solvent for the hotspot Pro32.

The Proline environment interaction with water molecules has been evaluated with the Radial Distribution Function, looking at Calpha 484 and O498 of HIS31, and O512 of Pro32.

Furthermore, it can be observed that during the dynamics His31, Ser33 and Trp60 are kind of covering Pro32.

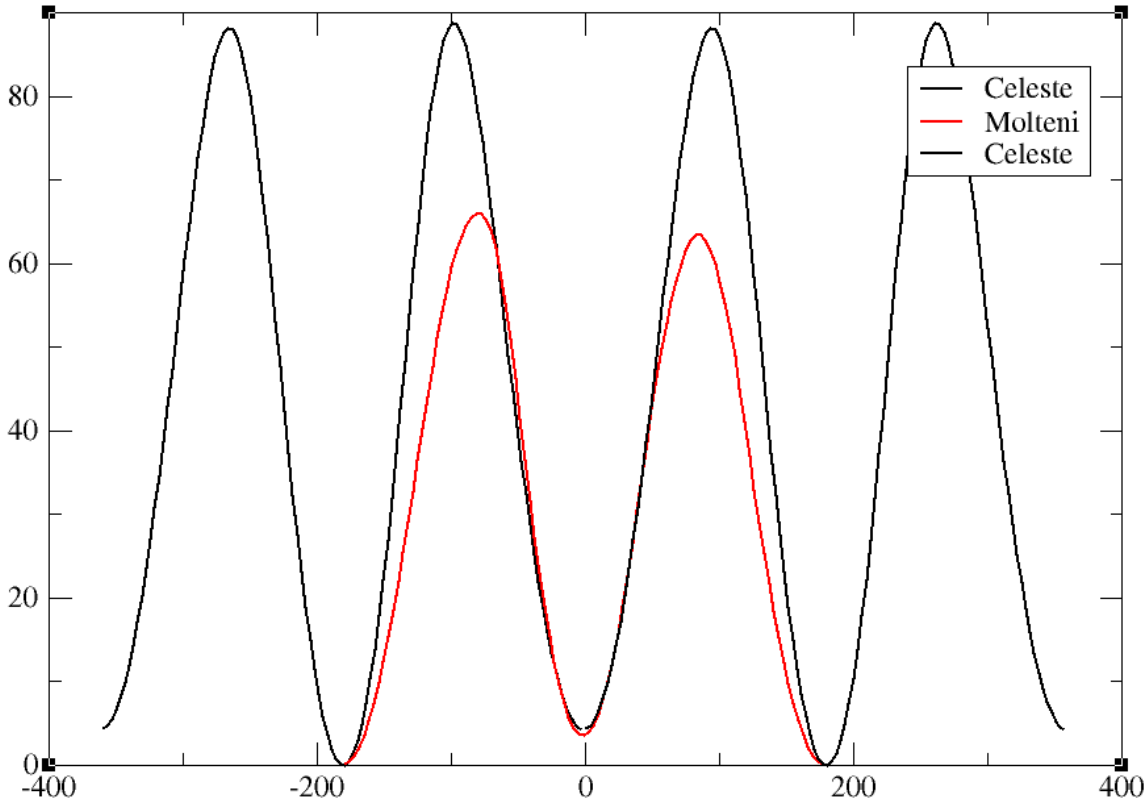

Figure 2: Comparison between different Zeta angle definition

Another very relevant fact is that in trans configuration there an almost constant H-bond between Ser33 and His31 while proline is in trans. Distance on average is 1.85 A.

The anti-aggregation variant W60G

As expected from the experiments [4] and other simulations [5], the most important change in the structure of the protein is given by DE loop.

While the other portions of the protein are vibrating interacting with the water molecules, DE loop rearranges itself starting from a very open configuration towards the solvent and getting closer to the BC loop where the trans Pro32 is located. This can be observed from the total RMSD of the system, that lowers during time suggesting a more compact structure, and the distances between Gly60 and Pro32 (DE and BC loop) and Thr86 and Pro32 (FG and BC loop). While the last bond does not vary in time, the first describes with an abrupt change how the DE loop closes.

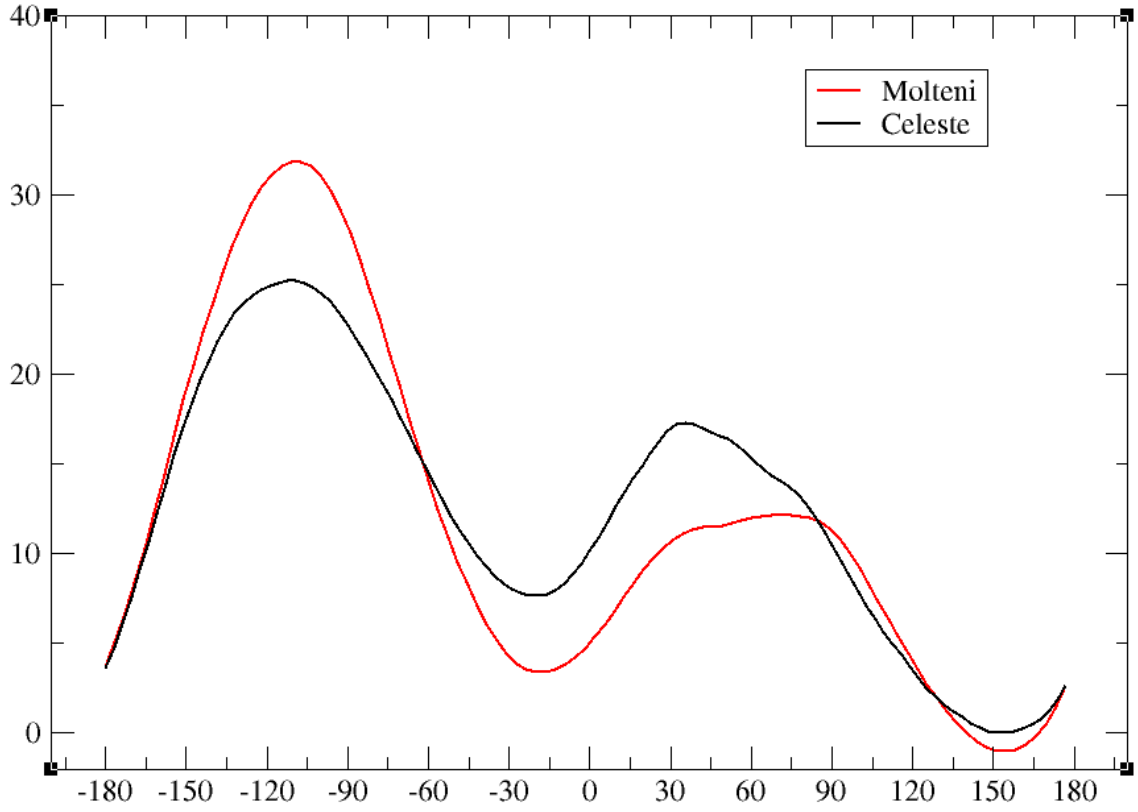

Figure 3: Comparison between psi angle

The calculated radial distribution function suggests that the Pro32 remains buried inside the other side chain. In fact, it is exposed to water when DE loop is open but when it closes, a Salt Bridge forms between Asp59 and Arg3.

OBSERVING CIS-PRO32 MD

The wild type

In this dynamics, the monomer does not experience heavy rearrangement but the most mobile portion of the protein are AB loop and DE loop. The high mobility of AB loop is strongly related to the interaction with water more than a formation of salt bridge of hbond with EF loop.

The Proline environment interaction with water molecules has been evaluated with the Radial Distribution Function, looking at Calpha 484 and O498 of HIS31, and O512 of Pro32.

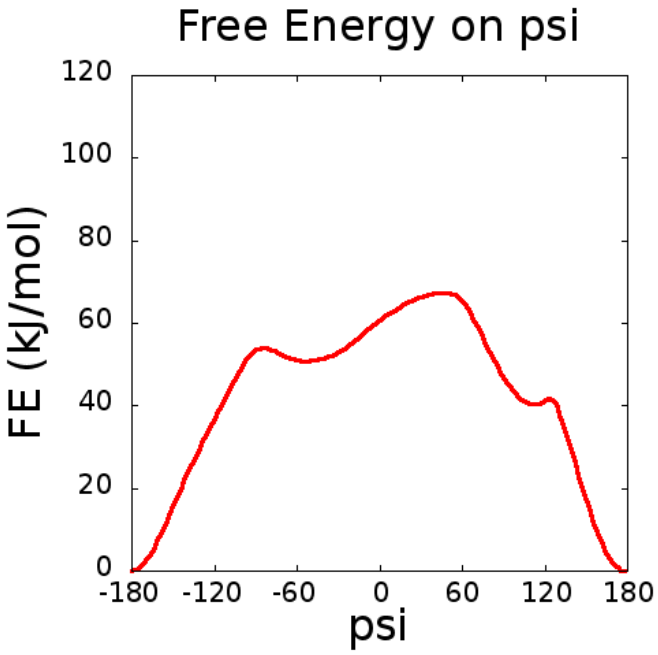

Figure 4: This is a caption

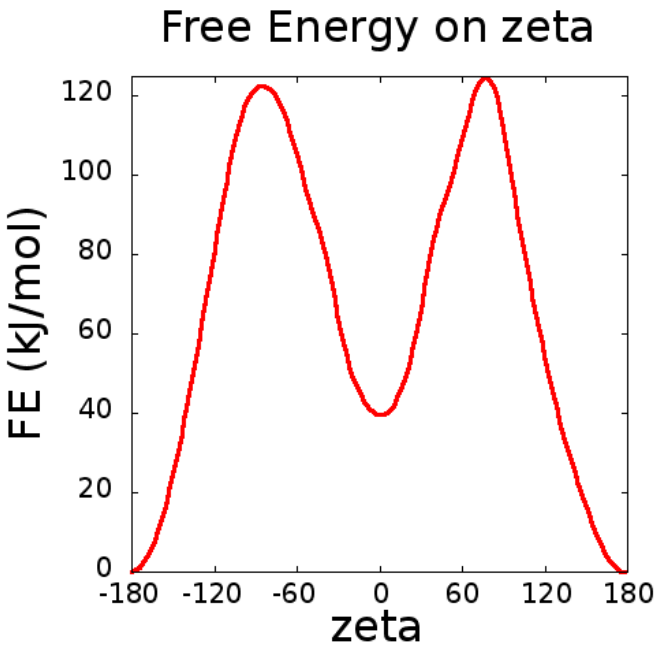

Figure 5: This is a caption

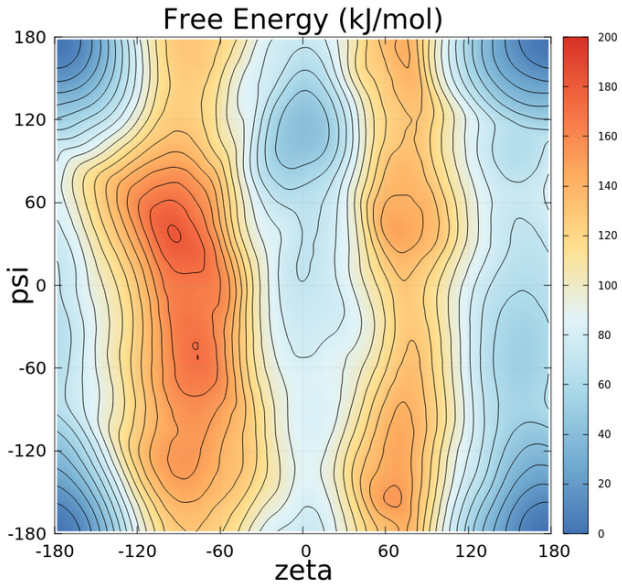

Figure 6: This is a caption

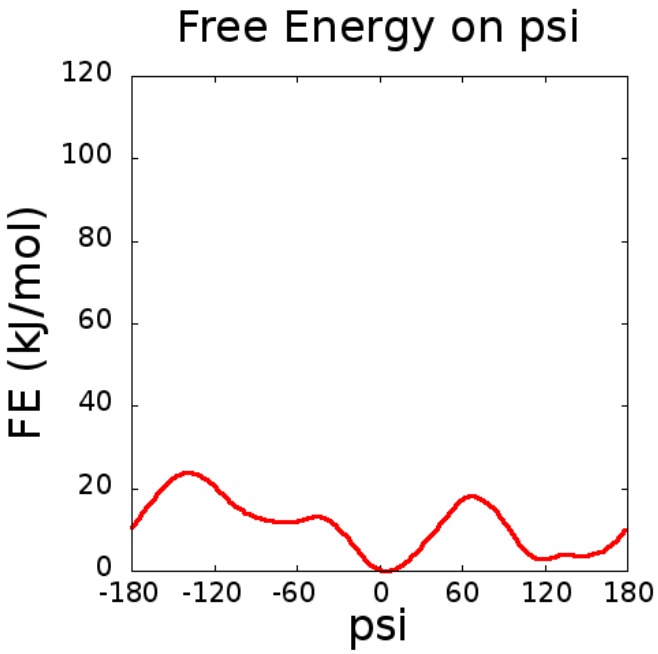

Figure 7: This is a caption

During the classical dynamics where the proline32 is in cis configuration, we find, contrarily to the trans dynamics a high distance between serine33 and histidine31, with a distance about 8 A.

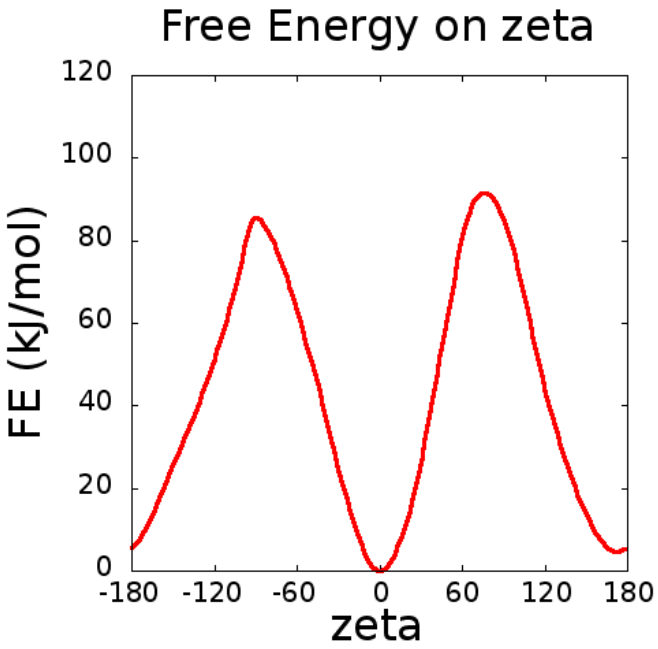

Figure 8: This is a caption

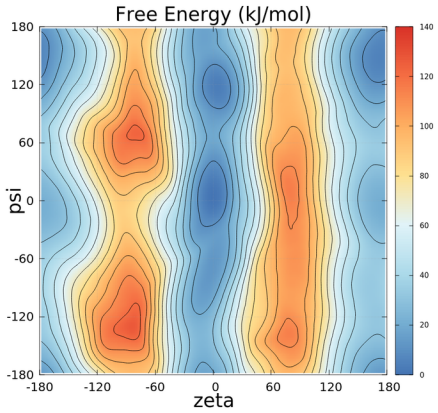

Figure 9: This is a caption

The amyloidogenical variant D76N

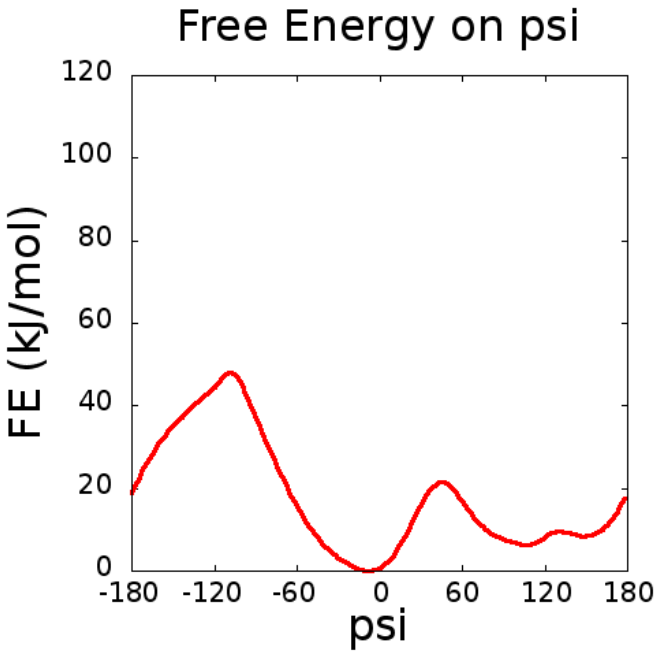

Figure 10: D76N psi (not averaged)

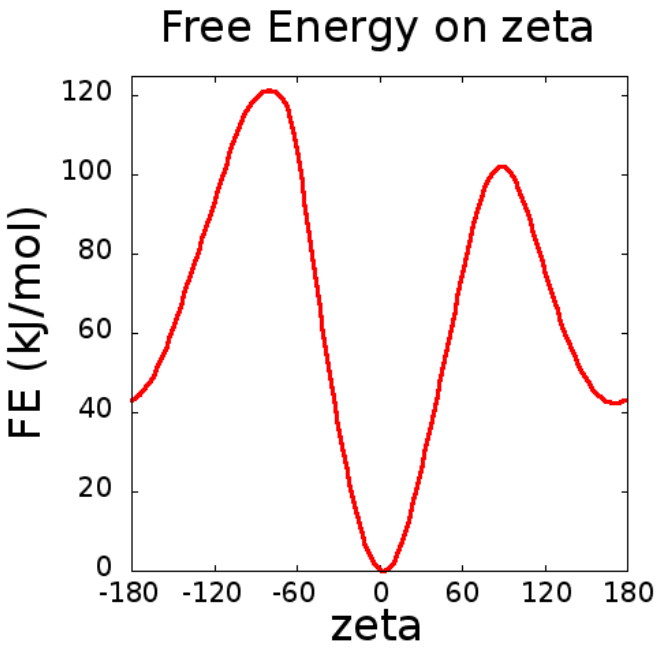

Figure 11: D76N zeta (not averaged)

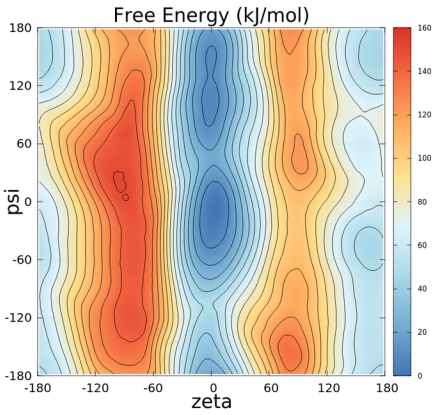

Figure 12: This is a caption

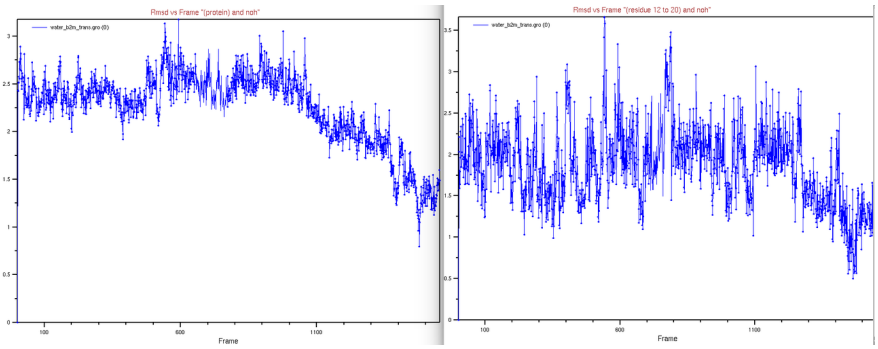

Figure 13: B2M: RMSD for the entire protein in water, and for the AB loop

The Metadynamics convergence

Here it is reported how the CVs behaves during the simulations. We can say, observing the gaussian deposition and the variation of the collective variables in time that the more the simulation goes on, the less the isomerisation happens. The hypothesis we can do is that, considering singularly each case, it is more easy to remain stuck in a certain basin and now, as the height of the gaussian is very low, it takes a very long time to overcome the barrier and visit the other conformation.

Then, observing the behavior in the last part of the simulation, after 1microsecond, could be misleading as we are maybe seeing conformation led by the water more than the CVs we selected.

Comparing the evolution of the Zeta dihedral, the most relevant one, we decided to perform analysis in the interval between 800ns and 1000ns, as for each variant both of the geometry are well explored (even if not with a diffusive behavior).

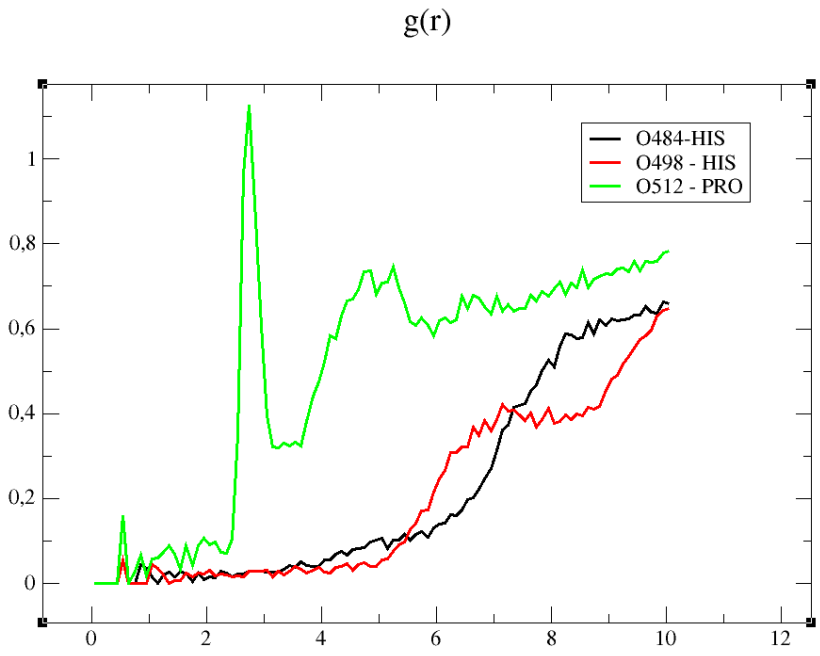

Figure 14: This is a caption

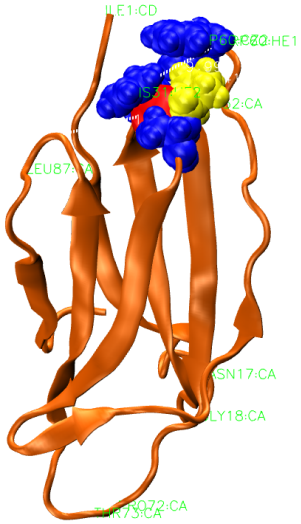

Figure 15: This rearrangement is almost constant during the dynamics

The wild type

We will evaluate separately *cis* and *trans* geometry.

Analysing the **CIS configuration**, we discovered a good stability in the protein structure as the RMSD of the protein shows, although there is a deep shift of the AB loop toward the strand D (see the plot). Such

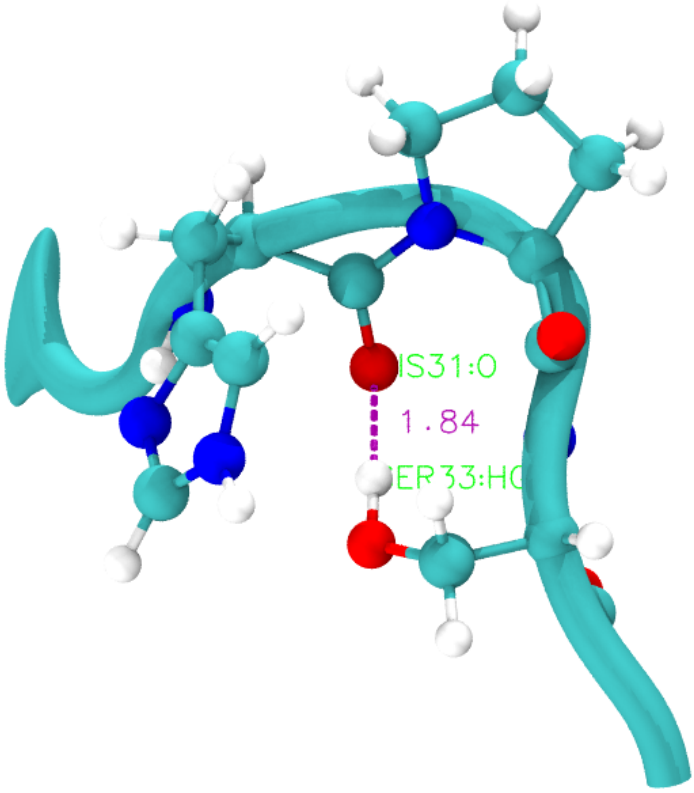

Figure 16: Insight of bond between serine and histidine during classical MD with proline in trans configuration

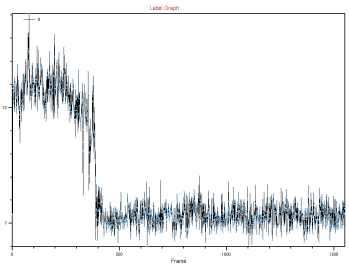

Figure 17: BC-DE loop distance (A) in time

rearrangement, that does not vary during the metadynamics, is stabilized by the formation of a salt bridge between Glu16 and Arg12.

Furthermore, very few molecules of water are able to enter in the protein structure.

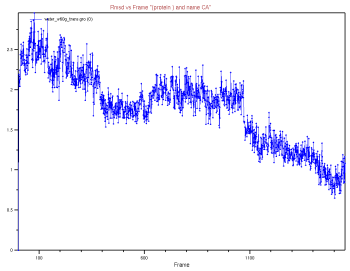

Figure 18: Protein RMSD (A) in time

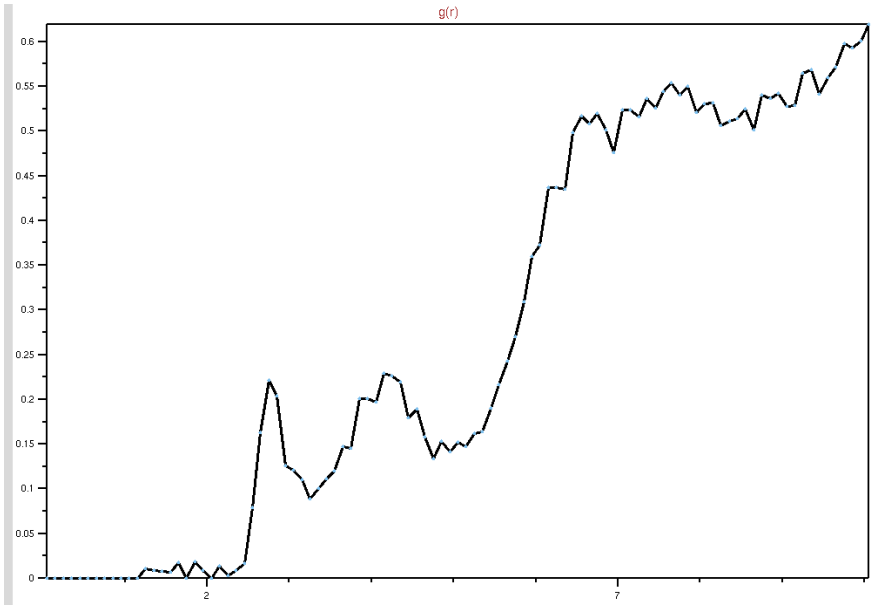

Figure 19: RDF of O515 in Zeta dihedral, wrt water oxygens.

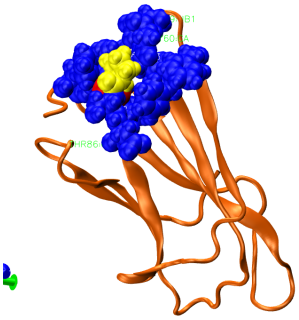

Figure 20: W60G configuration showing O515 in red, buried during MDtransPro. In the Back a SB forms a stable pocket.

Looking at Pro32 and the surroundings, it can be seen from the calculation of RDF for O498 that no water

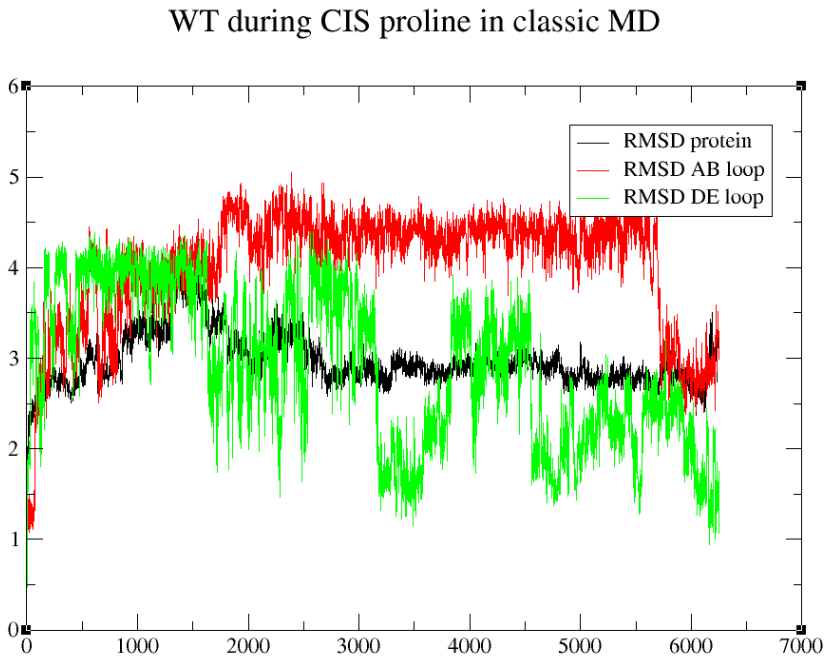

Figure 21: RMSD for the wild type

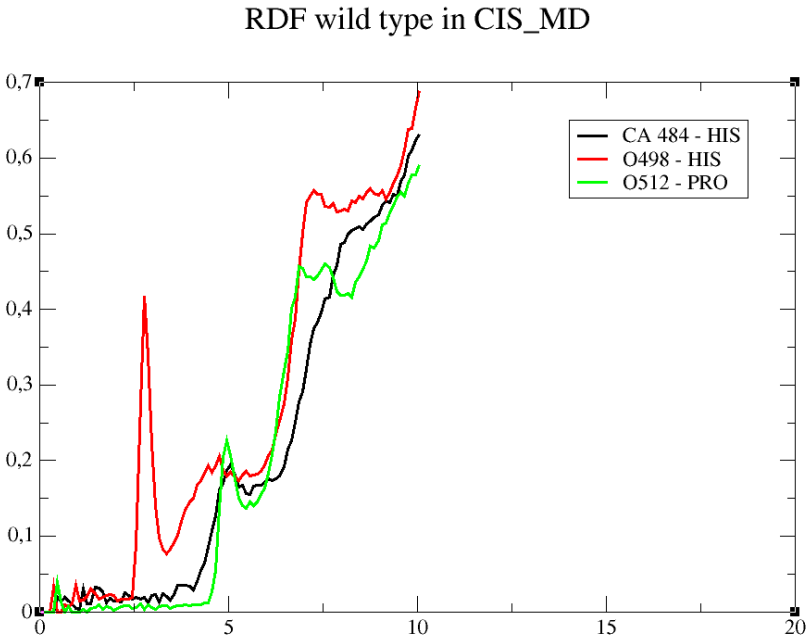

Figure 22: This is a caption

molecules are in contact with such atoms. In particular, we observe a formation of a coverage by Trp60 and Ile1 for the residue 31 32 33.

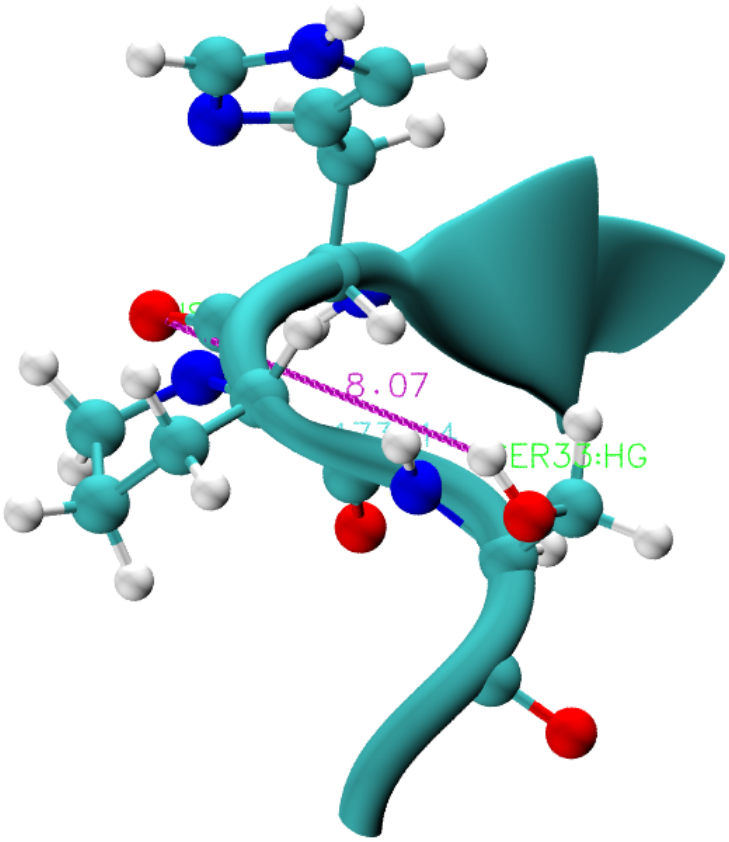

Figure 23: His31-Ser33 distance in cis configuration

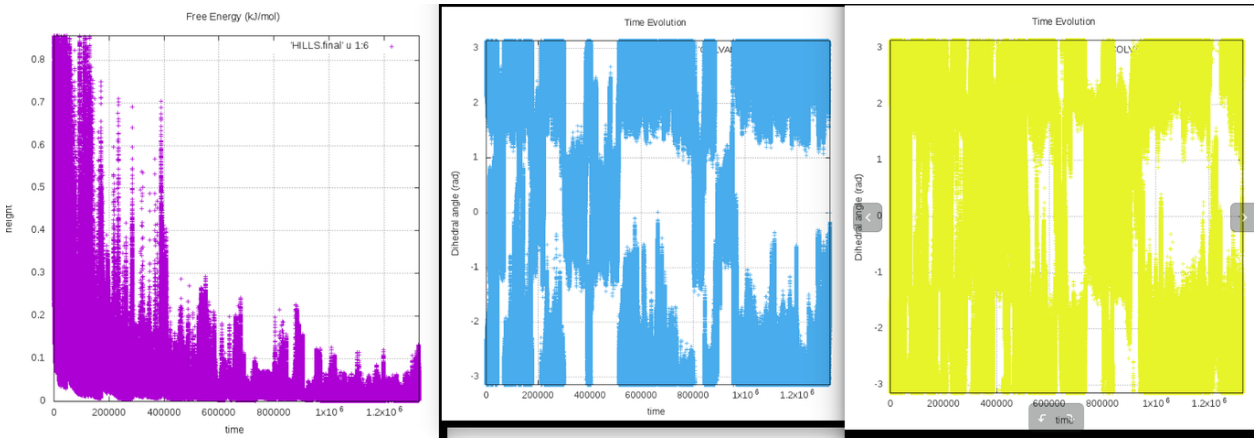

Figure 24: B2M wild type: l-r, gaussian curves deposition, Zeta in time, Psi in time

Analysing the **TRANS configuration**, we find the same behavior for the AB loop in terms of rearrange-

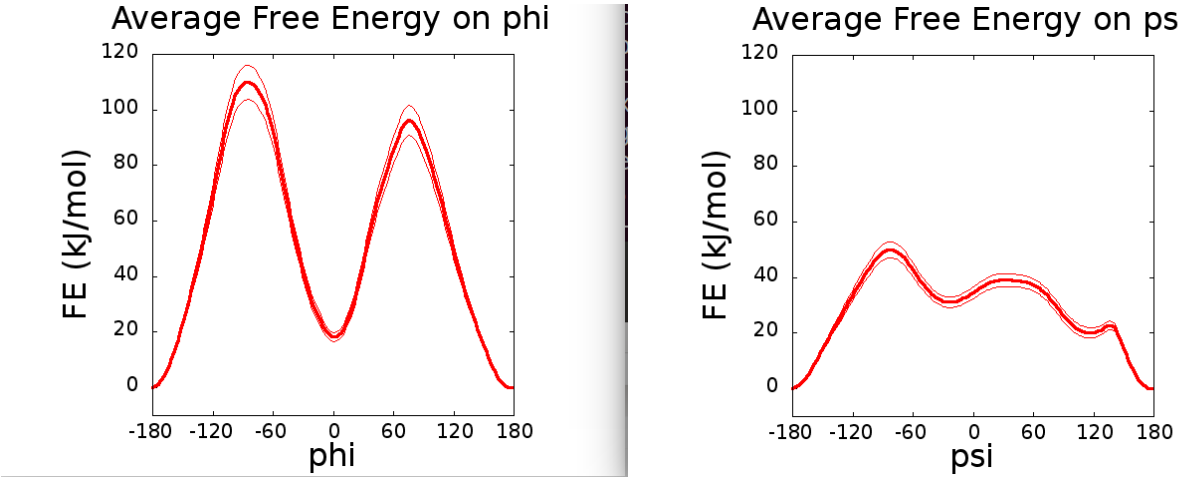

Figure 25: B2M wild type: zeta(phi) and psi dihedral profile calculated in the interval from 800ns to 1000ns with block analysis (FE average on 20 blocks, FE error is standard deviation /sqrt(Nblocks))

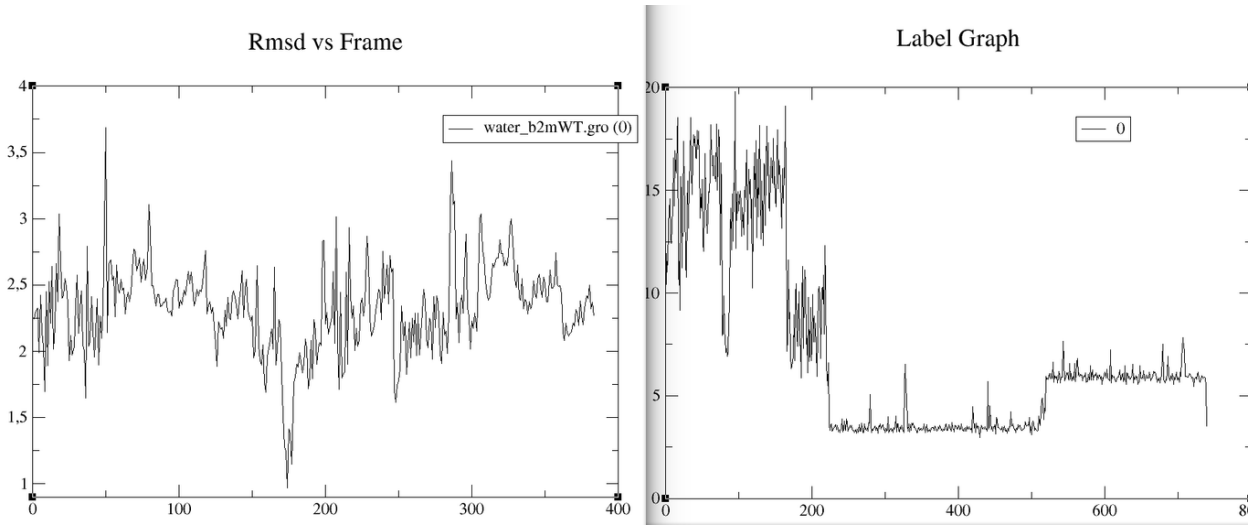

Figure 26: Left: B2M wild type CIS RMSD. Right: Distance between AB loop and D strand.

ment, whereas in the classical MD for the TransProline is not present. In such configuration, some water molecule is able to enter the structure, in particular around Arg45.

DIstances and hydrogen bonds

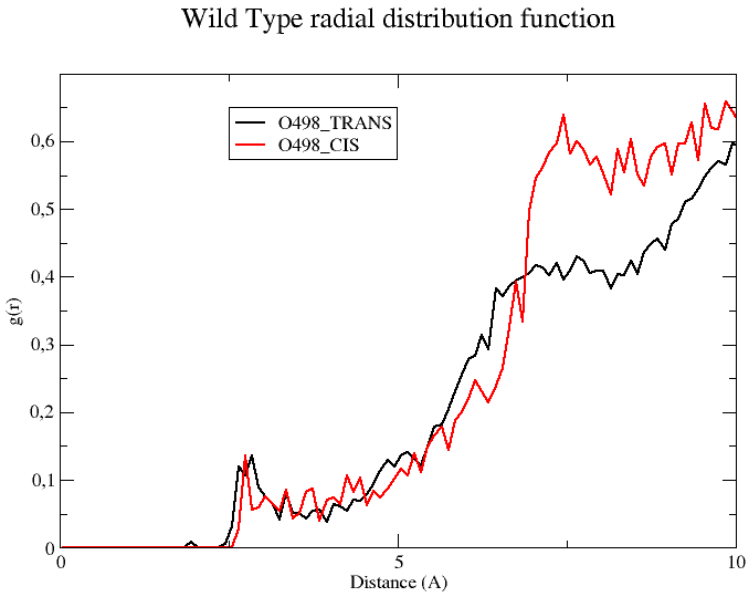

Figure 27: RDF for O498

The amyloidogenical variant D76N

Proceeding with simulation, the Pro32 is stuck for a very long time in trans configuration, that is expected for the amyloidogenical variant.

Analyzing the **TRANS** configuration, we observe that the Pro32 is completely buried by His31 and Ser33. Ser33 is openend toward the solvent in the first 200ns of the simulation but then there is the pocket formation. In fact, there is a profound difference between the RDF for O529 in trans (black curve) wrt to CIS.

Analyzing **CIS** configuration, we can see that here the O529 is largely exposed to water.

Furthermore, it is found that there is a high number of molecule inside the apical part of the protein opposite to Pro32, namely CD-EF-AB loop.

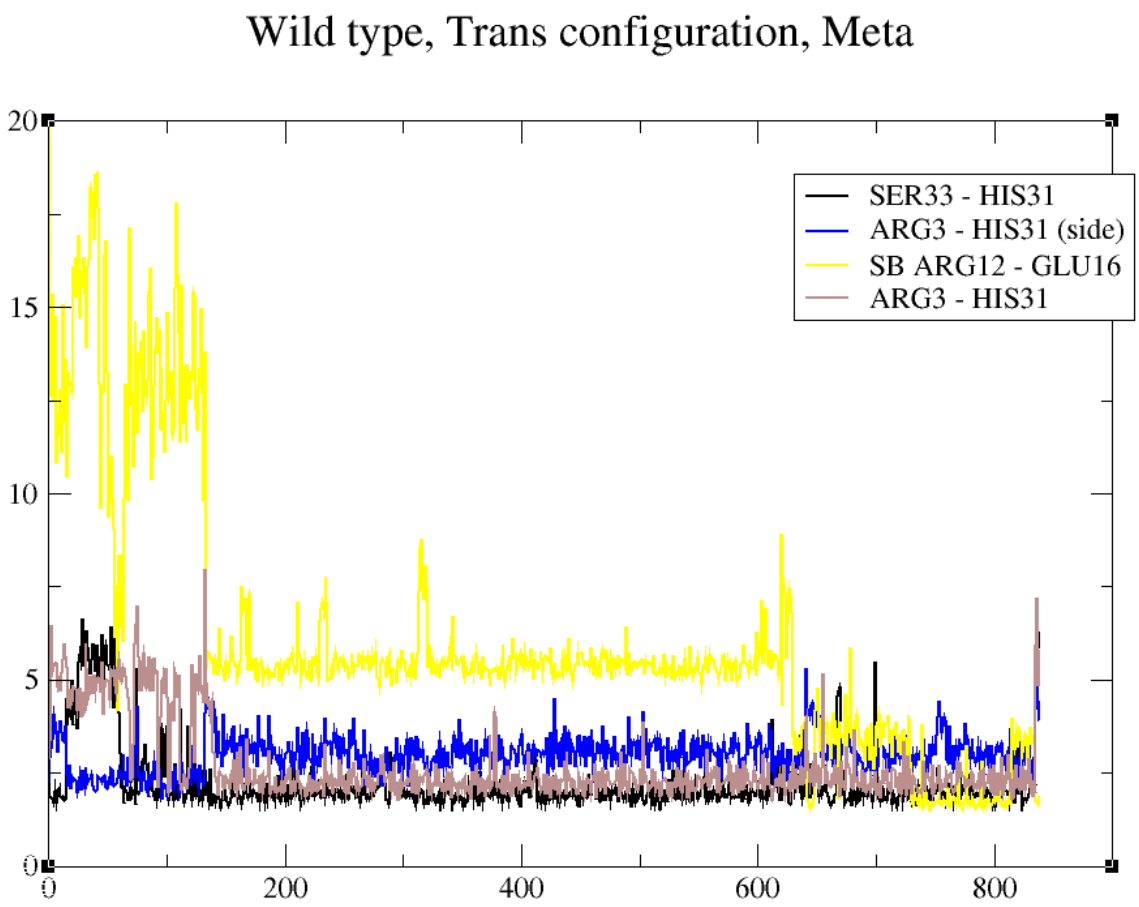

Figure 28: Hbonds around Pro32 and correlation with SB formation, TRANS population

The anti-aggregation variant W60G

For such variant, the cis and trans configuration are visited for very long time.

tipo, la CV zeta della d76n rimane incastrata in trans (da me è scambiato, è attorno allo zero) per tantissimo tempo e quello che succede alla prolina nel frattempo è che rimane attaccata con il suo anello alla thr86 e poi alla gly60 che se ne stanno nei due loop lì intorno. nella seconda configurazione, tra i foglietti beta c'è spazio per delle molecole d'acqua dentro la struttura della proteina ma nella seconda invece no, le varie side chain vanno a chiudere questo buco

Analyzing the **TRANS configuration**, Pro32 switch from being very close to Thr86 to being close to Gly60. In the first rearrangement, there is no space for water molecules inside the protein structure as the side chain of the underlying residues occupy the space. When BC loop goes toward DE loop, some solvent

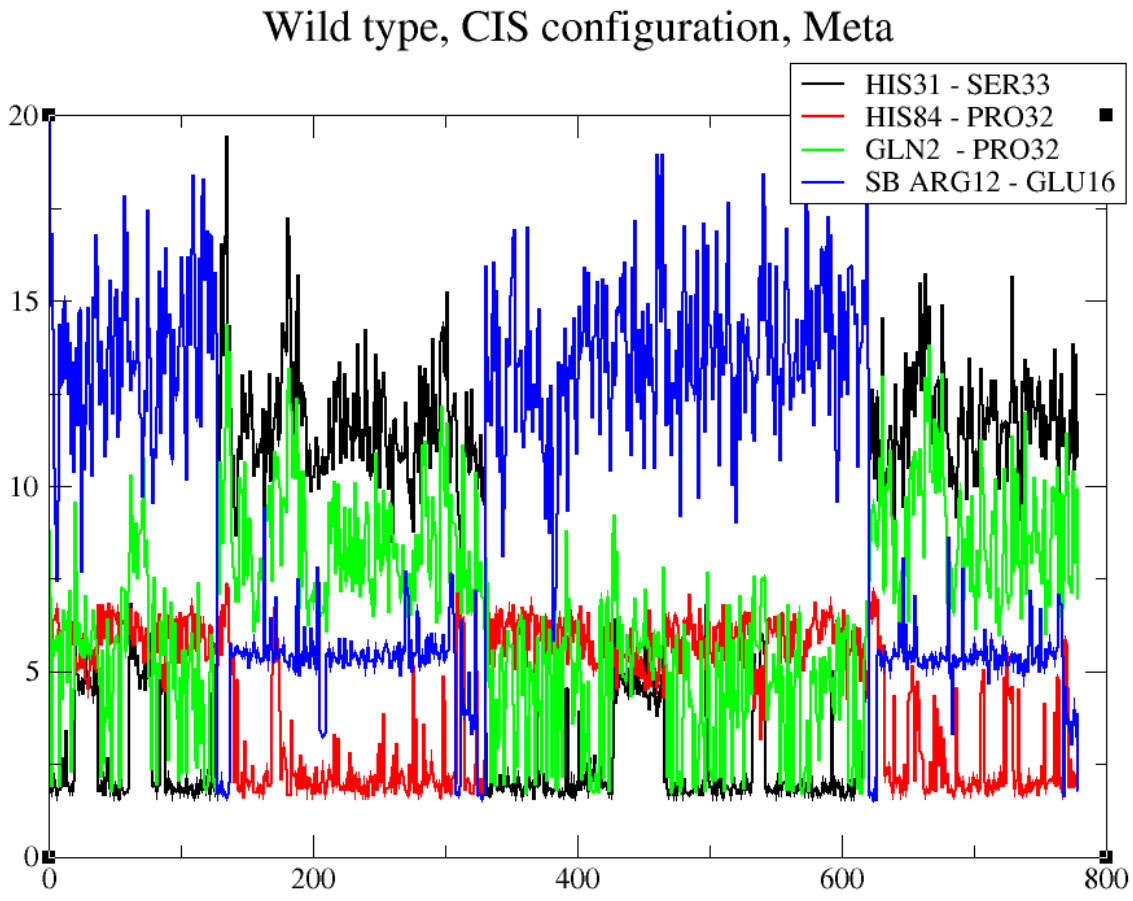

Figure 29: Hbonds around Pro32 and correlation with SB formation, CIS population

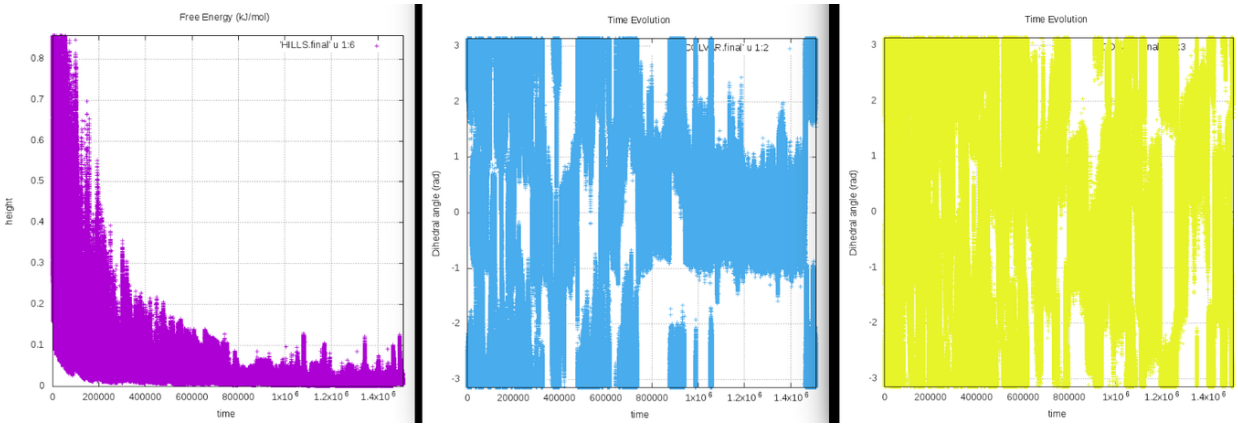

Figure 30: D76N: l-r, gaussian curves deposition, Zeta in time, Psi in time

can actually enters (more or less 5 in the same empty space).

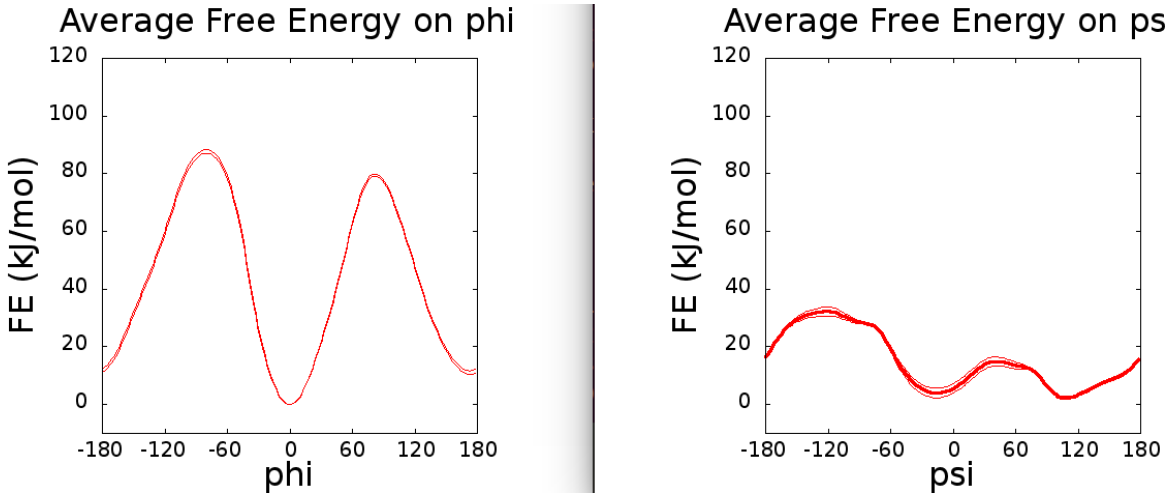

Figure 31: D76N: zeta(phi) and psi dihedral profile calculated in the interval from 800ns to 1000ns with block analysis (FE average on 20 blocks, FE error is standard deviation /sqrt(Nblocks))

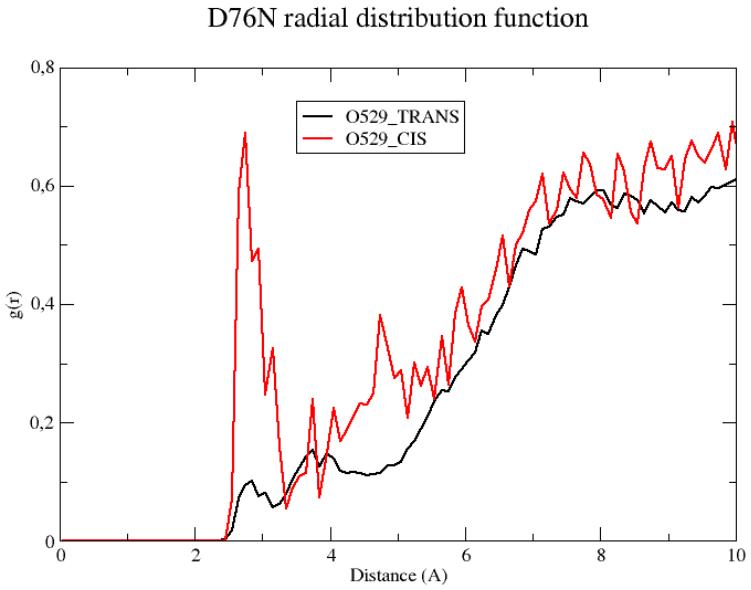

Figure 32: RDF for D76N

Analyzing the **CIS configuration**, here Pro32 has a constant configuration and same distance from DE and FG loops for the entire dynamics.

As RDF shows, in both conformations, His31-Pro32 peptide bond is exposed and interacting with water and we always have few molecules trapped between BC and FG loops.

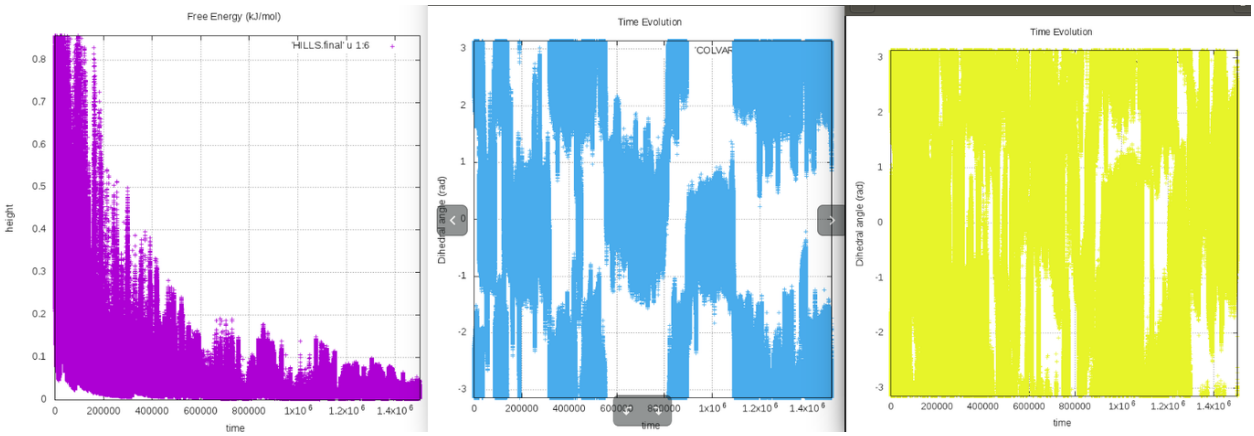

Figure 33: W60G: l-r, gaussian curves deposition, Zeta in time, Psi in time

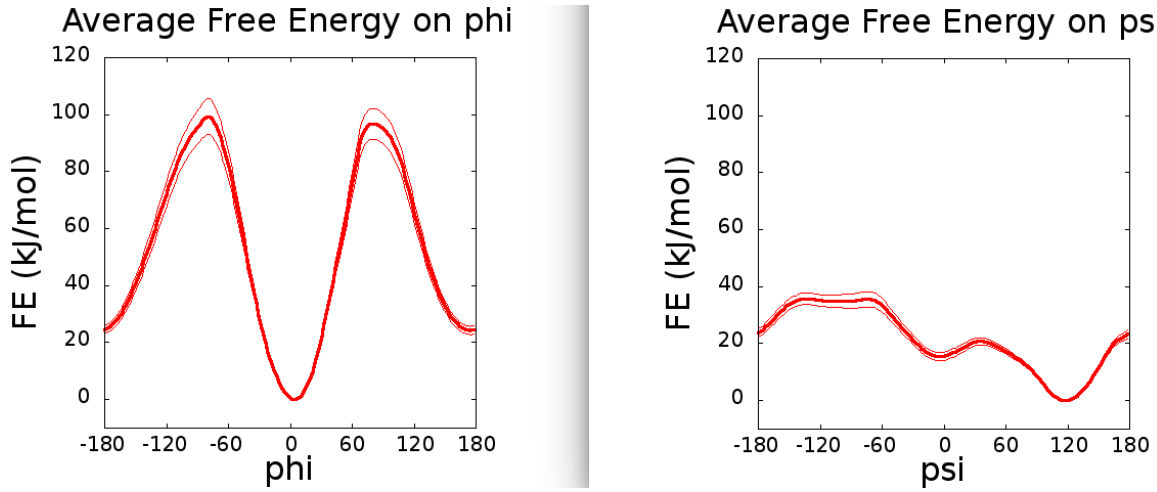

Figure 34: W60G: zeta(phi) and psi dihedral profile calculated in the interval from 800ns to 1000ns with block analysis (FE average on 20 blocks, FE error is standard deviation /sqrt(Nblocks))

Comparison between WT and the other variants

Analyzing the wild type during metadyamics simulation and observing separately the sampling in CIS and in TRANS conformation (in particular referring to the basins where the zeta angle is isomerized), we were able to observe which kind of environment is stabilizing the two conformers.

THE WILD TYPE

The classical MD simulation in CIS configuration for the peptidil-prolil bond is analyzed after 500ns.

It is identified a stabilization of the CO group in HIS31, the one responsible for the isomerization of PRO32, in an hydrogen bond with ARG3. In this case, PRO32 is bonded with HIS84.

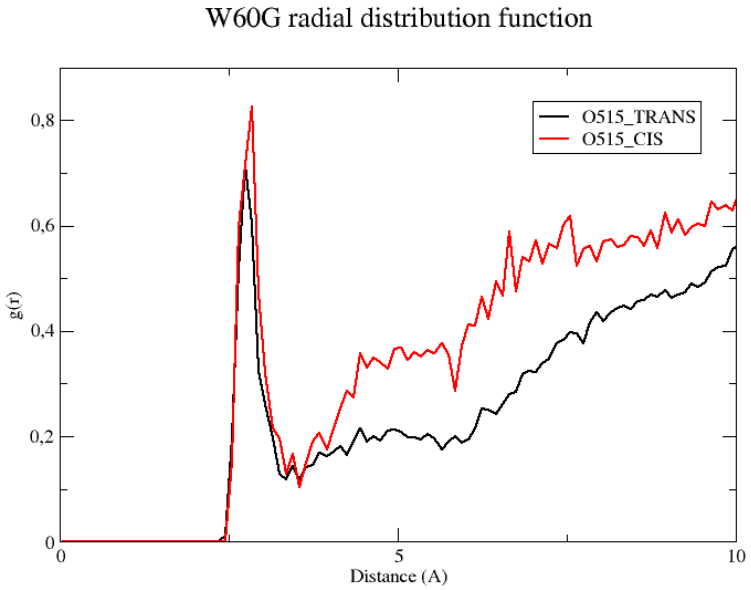

Figure 35: RDF for W60G, considering O515

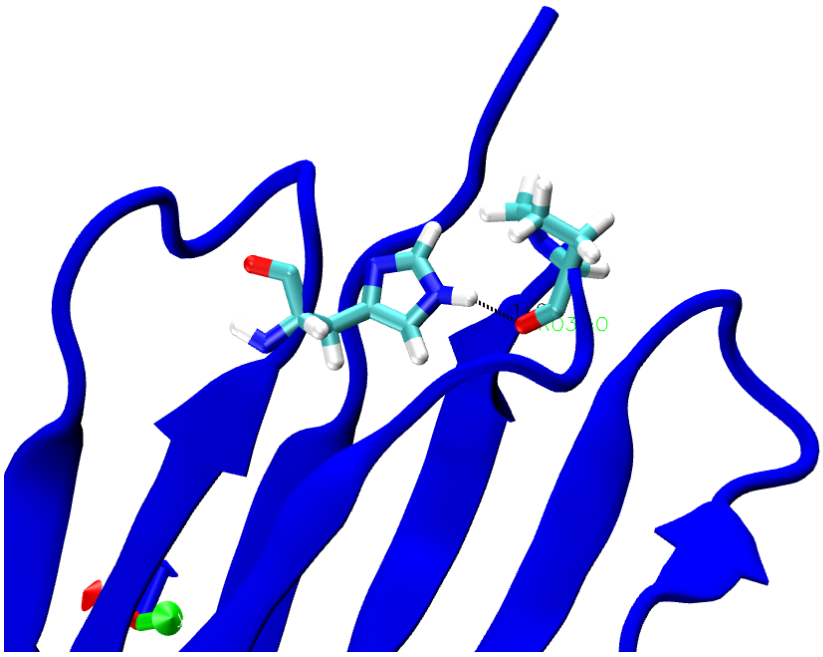

Figure 36: WT cis cluster, hbond between PRO32 and HIS84. When not isomerized the proline can easily interact with the histidine and being stuck in such configuration for most of the dynamics.

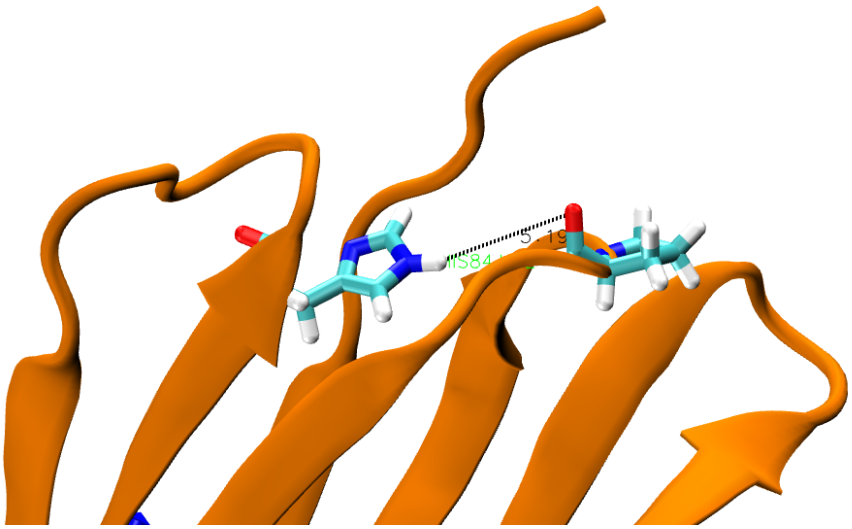

Figure 37: When in trans, the PRO32 inside WTb2m has the O too far from HIS84 to interact. QUESTION: does it interact with water then?

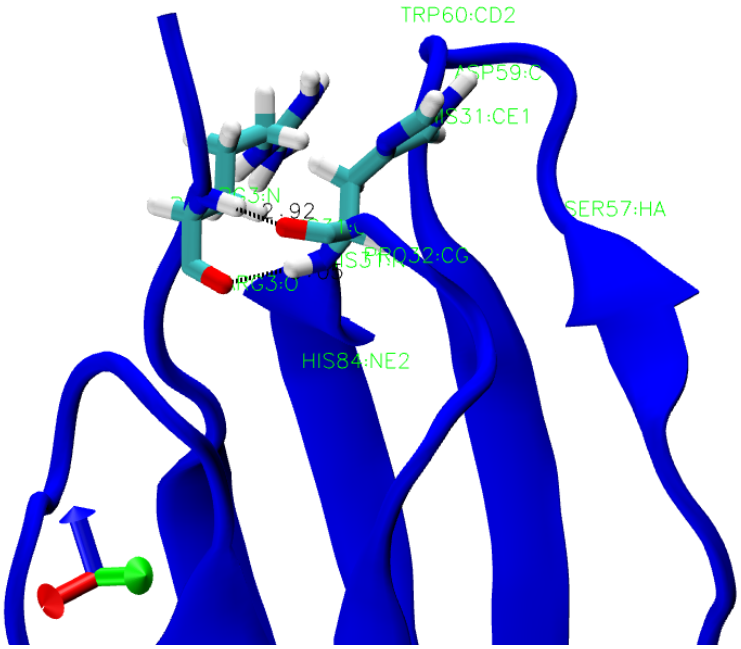

Figure 38: in CIS configuration, HIS31 is very often bonded with the CO group controlling the isomerization to the ARG3

The TRP60 is involved in hydrogen bonds with SER57 and ASP59 in both the cis and the trans configuration: both the cluster of the configuration

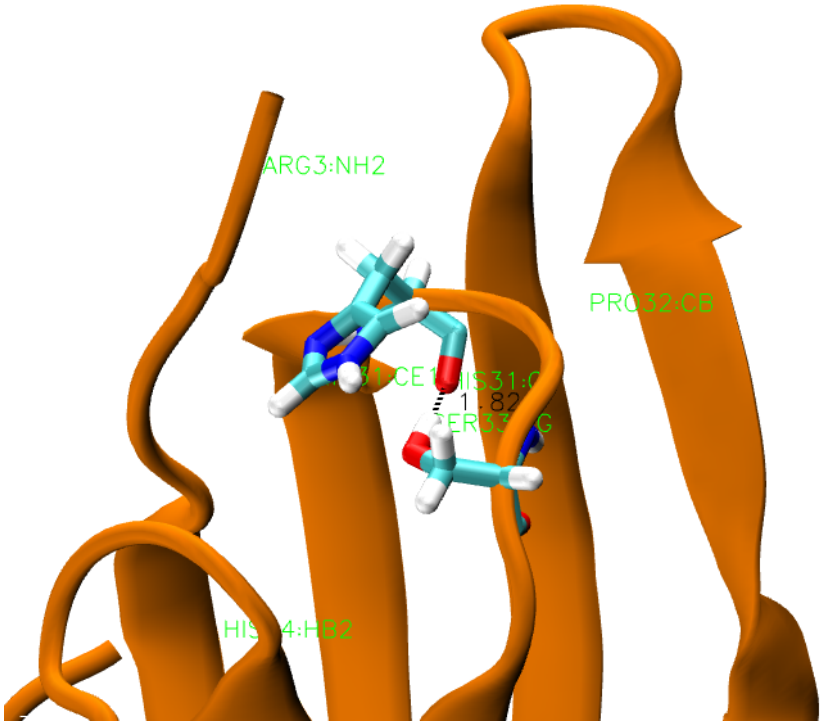

Figure 39: When the isomerization happens, the rotation of the CO group bring the HIS31 in being bonded to SER33

(using VMD Hbond tools, residue X wrt protein, 3.5A, 30°)

CIS classical MD

donor	acceptor	occupancy
HIS31-Main-N	ARG3-Main-O	92.42%
ARG3-Main-N	HIS31-Main-O	23.56%
HIS84-Side-NE2	PRO32-Main-O	36.76%
PHE30-Side-CD1	PRO32-Main-O	13.75%
TRP60-Main-N	SER57-Main-O	37.95%
ARG97-Side-NH2	ASP76-Side-OD2	19.69%
ARG97-Side-NH1	ASP76-Side-CG	17.39%
ARG97-Side-NH1	ASP76-Side-OD1	18.57%
ARG97-Side-NH2	ASP76-Side-CG	14.83%

ARG97-Side-NH1	ASP76-Side-OD2	17.81%
ARG97-Side-NH2	ASP76-Side-OD1	19.92%

TRANS classical MD

donor	acceptor	occupancy
SER33-Side-OG	HIS31-Main-O	38.89%

NO relevant hydrogen bonds for PRO32

SER33-Side-OG	HIS31-Main-O	123.50%
HIS84-Side-NE2	SER33-Side-OG	32.00%
TRP60-Main-N	SER57-Main-O	64.91%

During Metadynamics These selections are made for specific area of the free energy map. We selected the interval of the isomerization for zeta but all the possible values for the psi angle. Is psi value during the MD affecting the geometry in a way the bond is more favoured?

CIS WellTempered

donor	acceptor	occupancy
HIS31-Main-N	ARG3-Main-O	36.53%
ARG3-Main-N	HIS31-Main-O	58.81%
HIS84-Side-NE2	PRO32-Main-O	22.54%
SER33-Side-OG	GLN2-Side-OE1	47.15%
HIS84-Side-NE2	SER33-Side-OG	13.47%
TRP60-Main-N	SER57-Main-O	29.53%
ARG97-Side-NH2	ASP76-Side-CG	20.21%
ARG97-Side-NH2	ASP76-Side-OD2	12.95%
ARG97-Side-NH2	ASP76-Side-OD1	17.36%
ARG97-Side-NE	ASP76-Side-OD2	16.06%

TRANS WellTempered

donor	acceptor	occupancy
SER33-Side-OG	HIS31-Main-O	16.17%
HIS31-Main-N	ARG3-Main-O	23.05%

no relevant Hbond for PRO32

SER33-Side-OG	HIS31-Main-O	63.07%
HIS84-Side-NE2	SER33-Side-OG	27.90%

TRP60-Main-N	SER57-Main-O	85.18%
ARG97-Side-NH2	ASP76-Side-CG	25.61%
ARG97-Side-NE	ASP76-Side-OD2	20.62%
ARG97-Side-NH2	ASP76-Side-OD1	21.43%
ARG97-Side-NE	ASP76-Side-CG	23.05%
ARG97-Side-NH2	ASP76-Side-OD2	21.97

THE AMYLOIDOGENICAL VARIANT D76N

let's first compare the residues that seems to be extremely relevant for the wild type in order to characterize the variants with respect to the benchmark system of the wild type

CIS classic MD

donor	acceptor	occupancy

HIS 31 is poor in hbonds with the other protein residues that are relevant in the cis md: CHECK WHAT IS HAPPENING WITH WATER

PRO32 is poor as well, and SER33 too

TRP60-Main-N	PHE56-Main-O	73.88%

ASN76-Side-ND2	GLU77-Main-O	51.84%
ASN76-Side-ND2	ASN42-Side-OD1	64.44%
ASN76-Side-ND2	ASN42-Side-CG	19.42%

TRANS classic MD

donor	acceptor	occupancy
HIS31-Main-N	GLN2-Side-OE1	34.17%

PRO32 does not have any relevant Hbond

SER33-Main-N	PHE30-Main-O	32.63%
SER33-Side-OG	PHE30-Main-O	56.72%
TRP60-Main-N	PHE56-Main-O	33.33%
ASN76-Side-ND2	GLU77-Main-O	63.59%
ASN76-Side-ND2	LYS41-Main-O	109.80%

CIS WellTempered

donor	acceptor	occupancy
HIS31-Main-N	HISH84-Main-O	84.24%

(as we do not observe it in the classic MD, may be this hbond is influenced by the psi angle)

nor PRO32 nor SER33 are characterized by relevant Hbond

TRP60-Side-NE1	PHE30-Main-O	41.82%
ASN76-Side-ND2	GLU74-Main-O	16.97%
ASN42-Side-ND2	ASN76-Side-OD1	1.21%
ASN42-Side-CB	ASN76-Side-OD1	1.21%
ASN76-Side-ND2	GLU77-Main-O	1.21%

TRANS WellTempered

donor	acceptor	occupancy
ASP34-Main-N	HIS31-Main-O	32.79%
SER33-Main-N	HIS31-Side-ND1	13.66%
SER33-Side-OG	HIS31-Side-ND1	14.75%

no relevant interaction for PRO32

SER33-Side-OG	HIS31-Side-ND1	35.52%

SER33-Side-OG HIS31-Side-CE1 3.83%
SER33-Side-OG HIS31-Side-CG 2.19%

TRP60 in trans config seems to loose the hbond of the cis config

ASN76-Side-ND2 ASN42-Side-OD1 31.69%

THE ANTI-AGGREGATION W60G

CIS classic MD

donor acceptor occupancy
HIS31-Main-N ARG3-Main-O 25.51%
ARG3-Main-N HIS31-Main-O 53.84%

no relevant interactions for PRO32 nor SER33

GLY60-Main-N LYS58-Main-O 25.09%
GLY60-Main-N ASP59-Side-OD1 26.43%

ASP76-Main-N THR73-Main-O 80.34%
LYS41-Side-NZ ASP76-Side-OD2 16.14%
LYS41-Side-NZ ASP76-Side-CG 15.93%
LYS41-Side-NZ ASP76-Side-OD1 11.28%

TRANS classic MD

donor acceptor occupancy
SER33-Side-OG HIS31-Main-O 16.20%

SER33-Side-OG HIS31-Main-O 36.69%

GLY60-Main-N SER57-Main-O 73.73%
GLY60-Main-N SER57-Side-OG 9.37%

ASP76-Main-N THR73-Main-O 108.45%

CIS WellTempered

donor acceptor occupancy
HIS31-Main-N GLY60-Main-O 75.46%

nothing relevant for PRO32 and SER33

GLY60-Main-N	SER57-Main-O	112.27%
HIS31-Main-N	GLY60-Main-O	37.73%
ASP76-Main-N	THR73-Main-O	46.01%
ASN42-Side-ND2	ASP76-Side-OD2	57.36%

TRANS WellTempered

donor	acceptor	occupancy
HIS31-Main-N	GLY60-Main-O	108.29%
SER33-Main-N	HIS31-Side-ND1	18.42%
SER33-Side-OG	PRO32-Main-O	20.73%
SER33-Side-OG	HIS31-Main-O	15.16%
GLY60-Main-N	SER57-Side-OG	15.50%
GLY60-Main-N	SER57-Main-O	102.11%
ASN42-Side-ND2	ASP76-Side-OD1	22.23%
ASP76-Main-N	THR73-Main-O	40.38%
ASN42-Side-ND2	ASP76-Side-OD2	58.80%

Acknowledgements

References

[1]V. Y. Torbeev and D. Hilvert, "Both the cis-trans equilibrium and isomerization dynamics of a single proline amide modulate 2-microglobulin amyloid assembly", *Proceedings of the National Academy of Sciences*, vol. 110, no. 50, pp. 20051–20056, Nov. 2013.

[2]T. Eichner and S. E. Radford, "A Generic Mechanism of 2-Microglobulin Amyloid Assembly at Neutral pH Involving a Specific Proline Switch", *Journal of Molecular Biology*, vol. 386, no. 5, pp. 1312–1326, Mar. 2009.

[3]M. de Rosa, L. Halabelian, A. Barbiroli, M. Bolognesi, V. Bellotti, and S. Ricagno, "An Asp to Asn mutation is a toxic trigger in beta-2 microglobulin: structure and biophysics", *Amyloid*, vol. 24, no. sup1, pp. 15–16, Mar. 2017.

[4]A. Natalello, A. Relini, A. Penco, L. Halabelian, M. Bolognesi, S. M. Doglia, and S. Ricagno, "Wild Type Beta-2 Microglobulin and DE Loop Mutants Display a Common Fibrillar Architecture", *PLOS ONE*, vol. 10, no. 3, p. e0122449, Mar. 2015.

[5]C. Camilloni, B. M. Sala, P. Sormanni, R. Porcari, A. Corazza, M. D. Rosa, S. Zanini, A. Barbiroli, G. Esposito, M. Bolognesi, V. Bellotti, M. Vendruscolo, and S. Ricagno, "Rational design of mutations that change the aggregation rate of a protein while maintaining its native structure and stability", *Scientific Reports*, vol. 6, no. 1, May 2016.

[6]J. Zhou, Z. Xue, Z. Du, T. Melese, and P. D. Boyer, "Relationship of tightly bound ADP and ATP to control and catalysis by chloroplast ATP synthase", *Biochemistry*, vol. 27, no. 14, pp. 5129–5135, Jul. 1988.

[7]P. D. Boyer, "Energy Life, and ATP (Nobel Lecture)", *Angewandte Chemie International Edition*, vol. 37, no. 17, pp. 2296–2307, Sep. 1998.

[8]K. Zinszer, K. Morrison, A. Verma, and J. S. Brownstein, "Spatial Determinants of Ebola Virus Disease Risk for the West African Epidemic.", *PLoS Curr*, vol. 9, 2017.

[9]"Both the cis-trans equilibrium and isomerization dynamics of a single proline amide modulate 2-microglobulin amyloid assembly".

[10]"Redirecting".

[11]"A Generic Mechanism of 2-Microglobulin Amyloid Assembly at Neutral pH Involving a Specific Proline Switch - ScienceDirect".