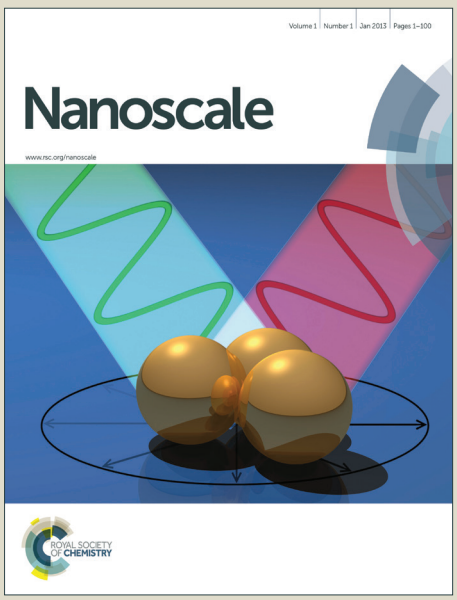
Nanoscale

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/nanoscale

Can small hydrophobic gold nanoparticles inhibit β_2-microglobulin fibrillation?

Giorgia Brancolini,[a‡] Dimitrios Toroz,[b] and Stefano Corni[c‡]

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
First published on the web Xth XXXXXXXXXX 200X
DOI: 10.1039/b000000x

Inorganic nanoparticles stabilized by a shell of organic ligands can enhance or suppress the natural propensity of proteins to form fibrils. Functionalization facilitates targeted delivery of the nanoparticles to various cell types, bioimaging, drug delivery and other therapeutic and diagnostic applications. In this study, we provide a computational modelization of the effect of a prototypical thiol-protected gold nanoparticle, $Au_{25}L_{18}^-$(L=$S(CH_2)_2Ph$) on the β_2-microglobulin natural fibrillation propensity. To reveal the molecular basis of the protein-nanoparticle association process, we performed various simulations at multiple levels (Classical Molecular Dynamics and Brownian Dynamics) that cover multiple length- and timescales. The results provide a model of the ensemble of structures constituting the protein-gold nanoparticles complexes, and insights into the driving forces for the binding of β_2-microglobulin to hydrofobic small size gold nanoparticles. We have found that the small nanoparticles can bind the protein to form long-life complexes. This binding of nanoparticles is able to block the active sites of domains from binding to another protein, thus leading to a potential inhibition of the fibrillation activity. A comparison with the binding patches identified for the interaction of the protein with a known inhibitor of fibrillation, supports our conclusion.

1 Introduction

Nanoparticles (NPs) are highly promising candidates for various important biological applications, such as gene delivery[1], cellular imaging[2], tumor therapy[3], and experimental biotechnology[4]. Meanwhile, the interaction between gold nanoparticles (GNPs) and the biological systems has received great attention since this may bring biosafety concerns[5]. In particular recent experimental works investigated the effect of NPs on protein fibrillation, *in vitro*[6]. Protein fibrillation is involved in many human diseases, including Alzheimer, Creutzfeld-Jacob's and dialysis-related amyloidosis[10]. The NPs have been found either to enhance or to inhibit the rate of formation of fibrils[11,12], therefore they can potentially lead to novel mechanisms for amyloid diseases as well as to therapeutic opportunities for their treatment[6]. Most experimental studies indicate that the composition of the NPs and their surface characteristics, rather than surface curvature or surface concentration, is determining their impact on fibrillogenesis.[7–9] Despite current efforts are being devoted to design NPs which could be used as drugs to inhibit fibril formation or disrupt fibrillar aggregates[13], *in vitro* experiments indicate that some NPs actually accelerate fibril assembly kinetics, leading to concerns about the risks of increasing environmental exposure to NPs[6,14]. Therefore, more information is needed to evaluate the possible impact of NPs on fibrillogenesis.

We consider β_2-Microglobulin (β_2m), one of the best characterized amyloidogenic protein, belonging to the superfamily of the constant domains of Immunoglobulins (Ig). Its homology among different species is quite high ($\sim$ 50 %)[15]. It is normally part of the MHC-I complex, whose function is to display antigenic proteins to the Tcells (and the MCH-β_2m complex has structural analogies with the homologous association in Igs). After performing its function the protein un-binds from the MHC complex and it is lost from the cell surface into the plasma and reaches the kidneys to be eventually degraded. However, when the renal function is compromised, its concentration can increase 25-35 times[16]; if the high concentrations is maintained for years, the proteins tends to aggregate in amyloid fibrils. This is a problem especially in haemodialized patients, causing amyloidose and arthropathies. β_2m represents an ideal prototypical system to study the interaction with engineered NPs which could pro-

† Electronic Supplementary Information (ESI) available: [Details on the molecular dynamics simulation results. Table S1 reports results of the MD trajectories with a single NP at different inital velocities (d1,d2,d3,d4) (three-dimensional structures and contact residues). Table S2 reports results of the MD trajectories with a couple of NPs at different inital velocities (initial orientations, three-dimensional structures, contact residues and root-mean-square deviations). Table S3 reports root-mean-square fluctuations and divergence of protein structure with respect to the NMR model. Tables S4 describes the average energy of the final complexes. See DOI: 10.1039/b000000x/

[a] *CNR Institute of Nanoscience, S3 Center, Via Campi 213/A, 41125 Modena, Italy. Fax: 39 059 2055651; Tel: 39 059 2055333; E-mail: giorgia.brancolini@nano.cnr.it*
[b] *University of Leeds, Faculty of Engineering, Leeds LS2 9JT, UK*
[c] *CNR Institute of Nanoscience, S3 Center, Via Campi 213/A, 41125 Modena, Italy Fax: 39 059 2055651; Tel: 39 059 2055205; E-mail: stefano.corni@nano.cnr.it*

vide new routes to inhibit fibril formation and for any application which depends on the physicochemical characteristics of the NPs.

Generally, NPs are composed of two parts: the core material, and a surface modifier that may be employed to change the physicochemical properties of the core material [17]. The core materials may be biological materials like peptides, phospholipids, lipids, lactic acid, dextran or chitosan, or may be formed of a chemical polymer, carbon, silica, or metals [17–19]. The physicochemical properties of NPs, such as charge and hydrophobicity, can be altered by attaching specific chemical compounds, peptides or proteins to the surface [17,18,20,21].

Here we consider one of the most studied and stable form of thiol-protected gold nanoparticle (GNP) for which X-Ray structure is available. The nanocluster $Au_{25}L_{18}^{-}$(L=S$(CH_2)_2$Ph) with Ph=C_6H_5, is known to display molecule-like behavior and for its distinct electrochemical and optical features. [22] More specifically, $Au_{25}L_{18}^{-}$ has a core composed by an Au_{13} centered icosahedron surrounded by six Au_2L_3 units in which the 12 Au atoms are stellated on 12 faces of the core (see Fig. 1). The 18 ligands thus split into a group of 12 inner ligands, with sulfur bonded to one stellated Au atom and one core Au atom, and a second group of 6 outer ligands, with sulfur clamped by two stellated Au atoms. In its native form, as prepared, the NP is negatively charged i.e. $Au_{25}L_{18}^{-}$ with net charge equal to -1.00 *e*, diamagnetic [23–26] and like the majority of gold nanoclusters $Au_{25}L_{18}^{-}$ is scarcely water soluble, which limits the biological application requiring water-solubility. Although the presence of such important experimental limitation, we believe that the proposed thiol-protected gold NP represents a clear test case to rationalize at a molecular level, the ability of GNPs to bind proteins and their effects on the protein itself. While the investigated GNP cannot be used as is in experiments, our results can inspire the design of metal NPs that keep similar size and general apolar character but displaying higher solubility. Many well defined thiolate-protected gold nanoclusters have been reported with precise control at the atomic level [27–30]. However the syntesis of water soluble metal nanoclusters capable of combining the stability of alkythiol-capped gold nanocluster with water solubility, has been largely limited to the gold-glutathione system. Recently, among water soluble thiol ligands for protecting Au_{25} nanoclusters and suitable for biological applications, the pyrrolidine-2-carboxylic acid (commercial name: captopril) has been shown to combine thermal stability and excellent water solubility. We believe that captopril or its possible derivatives, could be an interesting candidate based not only on its water solubility but also on the presence of its five-memebered pyrrolidone ring which is expected to provide the structural rigidity and stability required for biological applications. [31]

The focus is on the capability of hydrophobic functional groups attached to the surface of GNPs to alter the aggregation properties of amyloid proteins. Since the interaction of GNPs with proteins can cause perturbation of both protein structure and function, it is not surprising that interaction of GNPs with amyloid proteins or peptides might inhibit or facilitate amyloid formation. The pursued computational modeling is able to provide detailed information about the protein-GNPs interactions which could help to elucidate the mechanism of fibril formation at the molecular level, which remains elusive so far. To our knowledge, no atomistic simulation of the problem has been performed so far due to the significant challenge to modeling such interactions including lateral application of novel force fileds parameters. We have generated our own force field parameters for the functionalized GNP based on *ab initio* calculations that take into account the quantum nature of such a small GNP in a consistent and compatible way with the former GolP force field [32]. We believe that our findings may provide novel concepts to engineer small GNPs for the interaction with amyloid proteins and peptides.

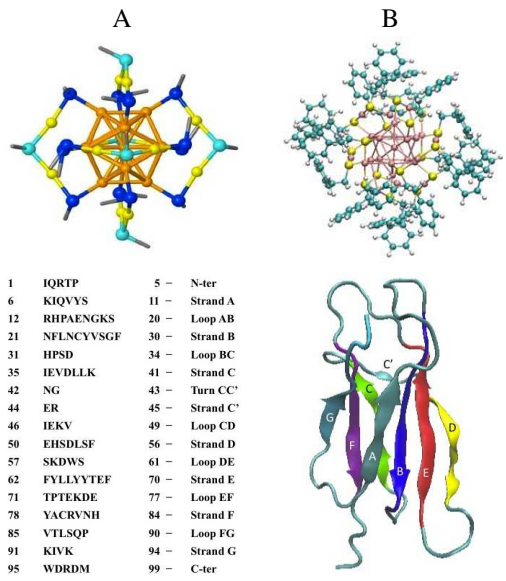


Fig. 1 Panel A (top): Structure of $Au_{25}L_{18}^{-}$: core Au atoms (dark yellow), stellated Au atoms (light yellow), inner-ligand S atoms (blue), outer-ligand S atoms (light blue). For clarity, the -$(CH_2)_2$Ph groups have not been shown. Panel B (top): Structure of the gold nanoparticle with -$(CH_2)_2$Ph groups shown. Panel A (bottom): Protein Sequence. Panel B (bottom): Secondary structure and Topology of β_2-Microglobulin.

2 Results and discussion

3 Docking of β_2m to Au nanoparticle

In this section, we focus on the binding of β_2m to a single $Au_{25}L_{18}^{-}$(L=S$(CH_2)_2$Ph) nanoparticle. To identify possible adsorption orientations of β_2m on the GNP faces (and the corresponding driving forces), we use the rigid-body docking method which was implemented in SDA 6.0[33], to predict the structure of protein-GNP encounter complexes. The structures of protein-GNP encounter complexes are generated by running Brownian Dynamics (BD) simulations during which the protein and the NP are modeled as rigid bodies. Around 5000 BD simulations were run. The adsorption free energies of the protein-nanoparticle encounter complexes obtained during the BD simulation are computed and the trajectories are clustered to identify genuinely different protein orientations. For each of the most populated complexes, which are ranked by size, a representative structure is selected.

When this docking procedure was applied to the β_2m system with the GNP ($Au_{25}L_{18}^{net}_{chg}$ = -1.00 *e*), it yielded six different orientations accounting for more than 98 per cent of the encounter complexes obtained, see Tab. 1. The representative structure of each computed complex is shown in Fig. 2. The protein residues contacting the nanoparticle differ in the various complexes, and are listed in Tab. 2.

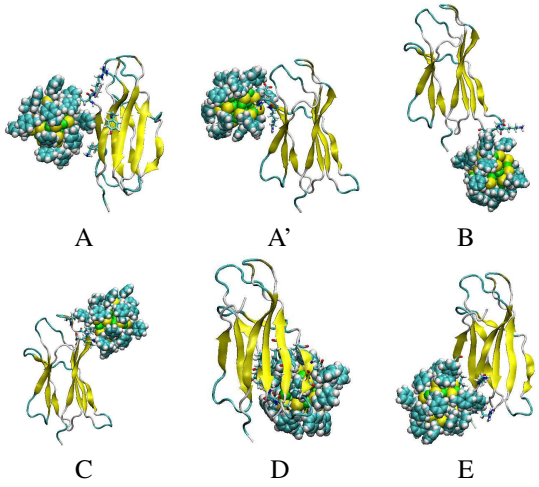


Fig. 2 Most populated encounter complexes of β_2m on gold nanocluster obtained by BD simulation. The structures of representative complexes for each of the six clusters are shown, ordered by decreasing cluster size. The protein backbone is shown in cartoon representation. The residues contacting the gold surface are shown in stick representation while all other atoms are shown in van der Waals representation.

During docking, the interaction energy of the protein with the nanoparticle is described by three main terms:[34] van der Waals energy described by site-site Lennard-Jones, E_{LJ}, interactions, adsorbate-nanoparticle electrostatic interaction energy, U_{EP}, (dominated by the contribution from the charged side chains), and the desolvation energy of the complex, U_{ds} (see Tab. 1).

Table 1 Resultant protein-nanoparticle encounter complexes from rigid-body BD docking. A hierarchical clustering algorithm (based on a minimum distance linkage function) was applied to the diffusional encounter complexes after docking to a $Au_{25}L_{18}^{net}_{chg}$ = −1.00 *e* nanoparticle. The reported complexes represent 98 per cent of the encounter complexes obtained by BD simulation.

Label	RelPop % [a]	U_{Repr} [b]	$E_{LJ} + U_{ds}^{p}$ [c]	U_{EP} [d]	spread [e]
A	28	-47.980	-27.417	-20.560	2.283
A'	28	-44.540	-27.49	-17.05	15.23
B	18	-41.890	-31.674	-10.220	2.67
C	16	-42.400	-48.304	5.911	7.926
D	4	-47.310	-44.478	-2.829	1.830
E	6	-46.480	-49.36	2.871	1.393

[a] Relative population of this cluster
[b] U_{Repr}: total interaction energy of the representative of the given cluster in kT with T= 300 K
[c] E_{LJ}: Lennard-Jones energy term for the representative complex, U_{ds}^{p}: non-polar (hydrophobic) desolvation energy of the representative complex, in kT
[d] U_{EP}: total electrostatic energy of the representative complex, in kT
[e] RMSD of the structures within the cluster with respect to the representative complex

Table 2 Contact Residues of each patch

No	Contact Residues
A	TYR10,LYS91,ASP96,ARG97
A'	GLY43,GLU44,ARG45
B	LYS58,ASP59,TRP60
C	MET99,HIS13,PRO14,GLU16,LYS19
D	SER33,ASP34,ILE35,LEU54,ASP53,LEU64
	GLU36,VAL37,HIS51,TYR66
E	GLN2,ARG3,THR4,LYS6
	THR86,LEU87,SER88,GLN89,LYS91

In all cases the binding is driven by E_{LJ} interactions but in complexes A, A' and B, electrostatic is also relevant. On the contrary, complexes C, D, E are driven only by E_{LJ} interactions since the electrostatic term is unfavorable. The difference between the binding energies in complexes A and B of $\sim$ 6 kT comes from a more favorable electrostatic energy for complex A. The difference between complexes B respect to D, E of $\sim$ 5 kT mainly comes from the favourable electrostatic energy for complex B. The difference between complex E and

C of ~ 4 kT is mainly due to a more favorable hydrophobic desolvation energy for complex E. The strongest binding seems to be associated with the total amount of charged residues contacting the surface of the nanoparticle (see complex A in Tab. 1), with a small preference for LYS with respect to other aliphatic and negatively charged residues. Moreover, except for A, A', the contact patches seem to be composed predominantly of hydrophobic amino acids e.g. TRP60, with high affinity for ligand hydrophobic patches on the surface of the NP.[35]

From the present docking results, we may conclude that β_2m on this small sized GNP makes two types of bound complexes: A, A', B (the most populated) in which a small contact area is compensated by electrostatic interactions of charged residues that simultaneously anchor the protein to the surface of the nanoparticle and C, D, E in which many contacts are used in order to optimize the binding energy. Docking positions D and E appear much less frequently than the others.

4 Refinement of docked encounter complexes by atomistic MD simulations

To assess the stability of the docked encounter complexes and to include the effect of structural relaxation, a series of eight indipendent MD simulations of β_2m interacting with a single $Au_{25}L_{18}^-$ in solution were performed starting with the two most representatives and populated complexes obtained from rigid-body BD docking, i.e. complexes A and B. The stability of the selected docked complexes is examined by using 50 ns of standard MD simulations at 300 K (results are reported in Tab. S1 of Supporting Information). The simulation time was effective in sampling different orientations and positions of the NP with respect to the protein. Simulations are repeated four times using a different seed for the initial velocity distribution (d1, d2, d3, d4) to improve the statistics of the search of the energy minima on the Potential Energy Surface, PES. Simulations are based on the OPLS/AA force field with the SPC/E water model as implemented in the GROMACS package[36–38]. Force field parameters for the GNP are developed from *ab initio* and implemented in the GolP[32] (for details, see Methods). This force field is compatible by construction with the implicit solvent model[34] used for rigid-body docking.

The results are summarized in Tab. S1 of the S.I. and only the final most stable and most populated orientations, complexes B and C, are reported in Fig.3. The potential energy components for each run are analyzed and reported in Tab. S4 of S.I. A detailed discussion of the MD refinement starting from the most representative docking complexes is presented below.

Complex A: The four independent MD runs revealed the global instability of the complex which can easily be converted into other stable docking complexes, either E, B or C (see Tab. S1 of S.I.). This result reflects the fact that the population and stability of this orientation partially relies on the initial conformation of the C-terminal tail which can be affected by conformational relaxation of the C-terminus during MD refinement. In two cases over four (simulations with initial velocities d3 and d4) the protein undergo a global rearrangement of C-terminus and its neighboring apical part (loops AB and EF) which drives the NP to bind through different patches. The final complexes are very similar to complex C, involving contact residues at the C-terminal domain. Residues PRO14 ALA15 GLU16 ASN17 (loop AB), GLU74, LYS75 (loop EF) and MET99 ASP98 ARG97 (C-terminal tail) bind to the surface of the NP at the end of the simulations. Even when extending the length of the simulation (at T = 300 K) to 80 ns, the interacting residues were conserved. On the contrary, initial velocities seeds d2 induced a global rotation of the NP with respect to the protein resulting to a final complex which is very similar to complex B, testifing the easiness of this initial complex to be converted into very different patches. The final complex is involving the contact residues: GLN2, ARG3, HIS31, ASP59, TRP60 and SER61 (loop BC, DE and N-terminal tail). Finally, the complex with initial velocity d1 is also converted to a quite different encounter complex

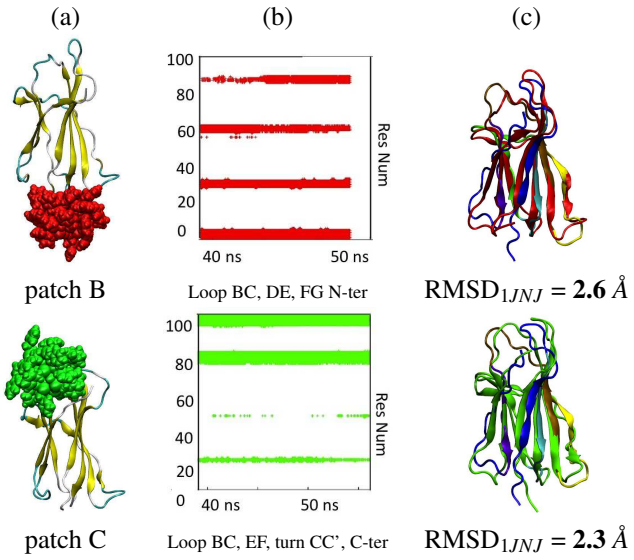


Fig. 3 (a) Final most stable orientations after 50 ns MD at T=300 K starting from complexes A and B resulted from the docking, (b) the time evolutions of contact residues of the protein respect to the surface of the nanoparticle along the last 10 ns dynamics (c) the divergence of each refined structure (excluding the N-terminal and C-terminal tails) with respect to the starting NMR structure (native state)

with contact patch made of residues ASN83, HIS84, VAL85, LEU87, SER88, GLN89 and PRO90 (loop FG) in analogy with complex E.

Complex B: The system was initially binding with N-terminal tail region and remained stable in three MD runs (d1, d2, d3) and β_2m changed orientation in one case (d4). (see tab. S1 in S.I.) In the simulations with initial velocities d1, d2 and d3, the protein is able to accomodate the NP through an extended contact patch, involving loop BC, DE and FG. Residues at 3 *Å* distance from the GNPs in the final most representative patch are: PHE30, HIS31, PRO32, SER33, ASP34, ILE35 (BC loop), LEU54, SER55, PHE56, PHE62 (DE loop) and HIS84 (loop FG). For the initial conditions d4, the β_2m reorient to a new orientation with the strand G and loop CC' interacting with the surface, labelled complex A', and shown in Fig. 1. Binding of β_2m to the NP in this orientation is stabilized by a single contact *via* LYS94 (strand G) and strengthened by ARG81 (strand F), PRO90 and GLN89 (loop FG). The residues of the loop CC' which are included in the binding are LEU40, ASN42 and GLY43. Energetically, the complex is more stable respect to complexes binding through the C-TER region (see Tab. S4 in S.I.)

In Fig. 3 we report: (a) the representative structure of the most stable orientation from MD; (b) the time evolutions of contact residues of the protein respect to the surface of the nanoparticle along the last 10 ns dynamics; (c) the RMSD of each refined structure (excluding the N-terminal and C-terminal tails) with respect to the starting NMR structure. It is shown that the lifetime of each contact is well conserved during the last 20 ns of MD. Moreover, RMSD is always between 2.3 and 2.6 Å (1.96 Å for the protein in water) in the last part of dynamics (i.e. the last 20 ns MD) which points to modest internal rearrangements of the protein.

To summarize, with a number of independent MD refining runs, we are able to discern only two stable complexes, B and C, starting from A, B out of six global orientations predicted by rigid-body BD docking. In all cases of the identified complexes, the protein is contacting the nanoparticle through the apical region representing the edges of the A, B, D β-strands.

5 Comparison with the effect of a molecule with anti-fibrillation activity

The present findings on the protein-GNP interactions are compared with previous studies of the identical protein interacting with an anti-fibrillation drug i.e. doxycycline. Doxycycline is a drug which is known to be able to suppress fibril formation *in vitro*[41]. Recent molecular dynamics simulation of highly concentrated doxycycline (six molecules) in the presence of β2m have provided binding modes for the drug and a rationale for its effect[40]. The authors were able to discern striking differences between their simulation and the identical simulation performed in the absence of doxycycline, in particular referring to the N-terminal tail and at the AB loop. Also BC, EF, FG loops are reported to be stabilized by the presence of doxycycline[40]. In order to strenghten the comparison with available theoretical and experimental data on doxycycline and to discern the nature of protein-GNP interactions towards GNP-GNP association propensity, we perform a series of ten indipendent MD simulations of β_2m interacting with two identical $Au_{25}L_{18}^-$ GNPs in solution. Simulations are started by the complex A (able to interconvert to other complexes, as discussed below) to which the second nanoparticle is added in four different geometrical configurations respect to the initial binding complex (see previous Section). In each simulation the second nanoparticle is located near one of the four vertices of the simulation box by keeping the initial binding complex at the center of the box. In order to properly sample the extended number of contacts made available to the system upon addition of the second GNP, longer simulation times are used. The stability of the final complexes is examined by using 100 ns of standard MD simulations at 300 K (results are reported in tab. S2 and S3 of S.I.).

Results with two nanoparticle interacting with the protein are discussed below, referring to Fig. 4 and S.I.

Simulations are started with a different initial velocity for each arrangement of the system. For geometrical reasons, in one case over four the two GNPs are located at the same vertex of the simulation box at a closer distance between each other. Starting the simulation at this conformation lead the protein to be contacted by two GNPs joined in a dimer (complexes BE and A'C in Fig. 4) despite the repulsive electrostatic interaction between them. In this case, we observe that the binding of the GNPs joined in a dimer is mediated by the presence of positively charged residues of the protein contacting the GNPs. However, in order to improve the statistics related to this initial state, we decided to run four additionally indipendent simulations with different initial velocities of the two NPs starting at the same vertex of the box and to compare them with the same series of simulations starting with the NPs at two opposite vertex of the box. In the end the total number of indipendent simulations was equal to ten. Interestingly, starting the simulations with the two GNPs at the same vertex of the simulation box, always provided the protein to be finally contacted by the two GNPs joined in a dimer. On the contrary, the formation of a dimer was never observed when the simulations were started with the two GNPs positioned at different vertices of the box. More specifically, we have identified two stable binding patches for the NPs contacting the protein as two separated units i.e. complexes BC and A'D in Fig. 4 and two binding patches for the GNPs joined in dimer i.e. complexes BE and A'C in Fig. 4, among the most populated. Below, we classify the complexes obtained in the presence of two

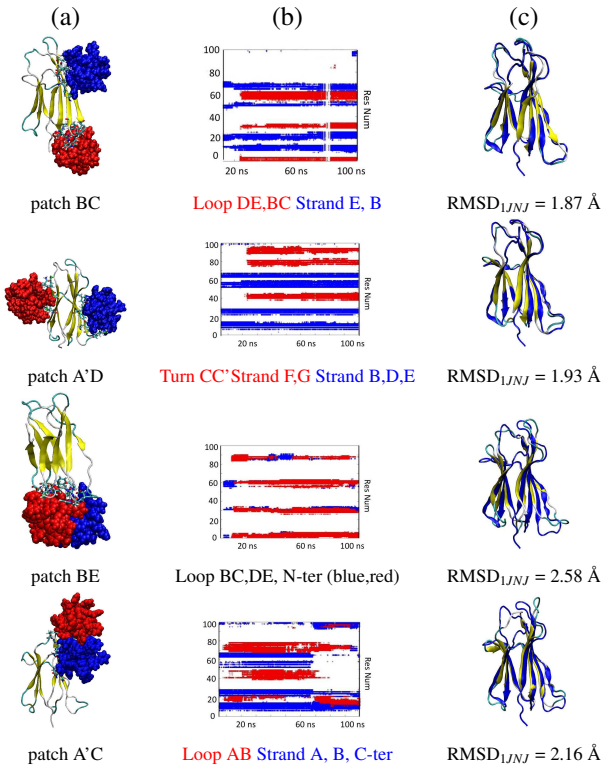


Fig. 4 (a) Most stable orientations of the couple of NPs after 100 ns MD at T=300 K (b) time evolutions of contact residues respect to the surface of the two NPs along the entire 100 ns dynamics. The binding patches of each NP are differentiated by color (blue and red) (c) the divergence of each protein refined structure (excluding the N-terminal and C-terminal tails) with respect to the starting NMR structure (native state)

NPs on the basis of the labels used for the single NP results. In Fig. 4(a) we report the representative structure of the most stable orientation from MD, in Fig. 4(b) the time evolutions of contact residues of the protein respect to the surface of the two GNPs along the entire 100 ns dynamics (each NP is differentiated from the other by color) and in Fig. 4(c) we report the divergence of each refined structure (excluding the N-terminal and C-terminal tails) with respect to the starting NMR structure. The long lifetime of each contact is clearly shown in Fig. 4(b) since all contacts are well conserved during the entire lenght of MD except for patch A'C in which some contacts are lost in the last part of the dynamics. We observe a lowered value of the RMSD among the most populated patches BC and A'D respect to the single NP cases. RMSD values of patches BC and A'D are between 1.8 and 1.9 Å (1.96 Å for the protein in water) which points to a stabilization of the native state of the protein which can be associated to the binding at those specific sites of the protein. In the other cases RMSD is a bit larger but this is mainly due to the presence of two distinct conformation of the DE loop in water which is emphasized in the case of complexes BE.

Complex BC: This binding patch is the most populated, it appeared in four cases over ten, and the binding with the protein is occuring through two single GNPs separated from each other. Once the number of contacts are formed they tend to be stable during the dynamics. Two different contact patches are identified involving residues TYR10, SER11, ARG12, HIS13 (loop AB) PHE22, ASN24 (strand B) LEU65, TYR67 (strand E) for the binding with the first nanoparticle (blue) and residues HIS31, PRO32, SER33 (loop BC) PHE56 TRP60 PHE62 (loop DE) for the binding with the second nanoparticle (red). Both binding patches are involving a large number of hydrophobic anchoring groups (TYR67, LEU65, PHE22, TYR10 and PHE56, TPR60, PHE62) and positively charged residues (HIS13, ARG12 and HIS31). The interactions between the protein and the nanoparticle are involving not only the hydrophobic ligands but also the core gold atoms.

Complex A'D: This binding complex is also well populated, it is seen in two cases over ten. The two GNPs bind to the protein, separately. Each of the nanoparticle is kept apart from the other through the protein β_2m barrell since the binding here is involving the lateral β sheets of the protein instead of the apical parts. Two patches are made of residues TYR10 ARG13 (loop AB) TYR26 SER55 SER57 LYS58 TYR63 LEU65 (loop DE, strand E) in the binding with the first nanoparticle (blue) and LEU40 LYS41 ASN42 GLY43 GLU44 ARG45 (turn CC') ALA79 ARG81 ILE92 (strand F) in the binding with the second nanoparticle (red). Interestingly this latter binding patch is not very often seen in the presence of a single nanoparticle suggesting a concerted binding mechanism between the two GNPs which are able to optimize in this way their repulsive interaction. The present binding pose

is associated to the strongest reduction of the local fluctuations of loops AB, BC, DE, EF, FG driving the protein to a stabilization of its native state, as supported also by the smallest RMSD respect to NMR strucuture, reported in Fig. 4.

Complex BE: The binding involves the two GNPs joined in a dimer. The complex occurs in two cases over ten. The patch is defined by residues HIS31 PRO32 (loop BC) ASP34 PHE56 PHE62 (loop DE) THR86 (loop FG) contacting the first nanoparticle (blue) and GLN2 ARG3 THR4 PRO5 (N-terminal tail) SER28 GLY29 PHE30 HIS31 PRO32 (loop BC) PHE56 TRP60 SER61 PHE62 (loop DE) SER88 (loop FG) contacting the second nanoparticle (red). The number of contacts tends to be stable during the simulation as shown in Fig. 4. Here the interaction with one of the nanoparticle (blue) occurs through a small contact area which is compensated by electrostatic interactions by two charged residues (HIS31, ASP34) that simultaneously anchor the nanoparticle and two hydrophobic anchoring groups (PHE56, PHE62) which optimize stacking interactions with the hydrophobic ligand of the nanoparticle. The interaction with the second nanoparticle (red) occurs by a large number of contacts which are used in order to optimize the binding and which are including four hydrophobic ancoring groups (PHE30, PHE56, PHE62, TRP60) and two charged residues (ARG3, HIS31). In this complex the binding of the protein with the dimeric nanoparticle is occurring mainly by the hydrophobic ligands of the GNPs but gold atoms of the core are also directly involved in the binding due to the shorth lenght of the ligands.

Complex A'C: In this last complex, the binding still occur through the two GNPs joined in a dimer. The patch is involving the apical part of the protein at the C-terminal tail. The number of identified contacts is not globally stable during the time evolution of the dynamics (as can be seen in Fig. 4) . Especially the contacts with one of the two nanoparticle (red) tends to be estinguished at the end of the dynamics. The result reflects the fact that the stability of the binding rely on the structural fluctuations of the C-terminal tail which can undergo strong conformational rearrangements during MD. The most extended contact patch includes residues GLN8 VAL9 TYR10 SER11 HIS13 PRO14 ALA 15 ASN21 (loop AB) TRP95 ASP96 ARG97 ASP98 MET99 (C-terminal tail) (blue) whereas the small contact patch is made of HIS13 PRO14 and ALA15 (loop AB) (red).

Summarizing this section, we find that in most cases, the addition of a second GNP to the protein-GNP system yields a binding patch, BC, which is exibithing identical binding patches of complexes B and C of the single GNP case. The results provide a reliable characterization of the binding patches for the benchmark system and strenghten the comparison with available experimental and theoretical data on our benchmark molecule, doxycycline. Akin to doxycycline, the interaction of the protein with a couple of $Au_{25}L_{18}^{-}$ NPs, has the effect of reducing local fluctuations of the protein[40]. In Figure 4(c) we show that the distortion of the protein with respect to the native NMR structure points to reduced internal rearrangements as a function of the increased number of contacting GNPs, like doxycycline. (This is supported by the RMSF plots reported in Tab. S3 of SI). The protein contacts with the couple of GNPs have been analyzed by plotting the number of contacts versus residue number in both cases[40] (in Panel B of Fig. 5(top) direct contacts with a couple of doxycicline molecules, in Panel B of Fig. 5(bottom) with a couple of $Au_{25}L_{18}^{-}$). The reasons for this similar behavior could be associated with the presence of terminal phenyl rings on the hydrophobic NPs which may confer similar physico-chemical properties to that of the doxycycline drug which also made of a plurality of aromatic rings (see Panel B of Fig. 5).

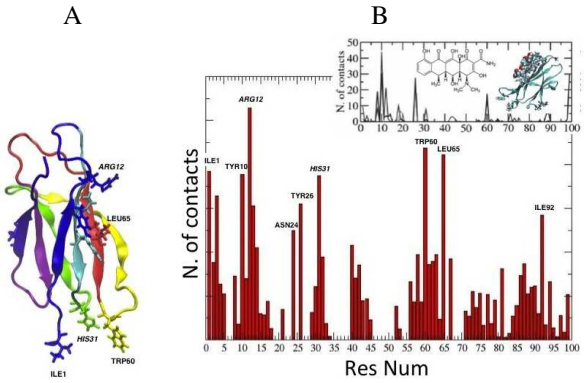


Fig. 5 A panel: Secondary structure of β_2m protein. The most contacted residues are displayed as solid bonds. B panel: Number of protein contacts with the two NPs are compared with doxycycline contacts versus protein residue number. The number of contacts are avergaged over 100 ns at simulation time in both cases.

As clearly shown in Fig. 5, the most contacted residues of the GNPs belong to the same binding domain of the doxycicline drug. In both cases, the preferred binding is occurring through stacking interactions of doxyxycline/NPs with TYR10, ARG12 in strand A and with TRP60 in loop DE. The complexes formation do not disturb the inner core of the protein, but rather aids a general conformational stability favoring the native-like protein structure as already indicated by the reduced RMSD respect to native structure. The simulations suggest that the effect of the hydrophobic GNPs may be that of inhibiting protein fibrillation propensity. This hypotesis is even strengthen by the recent identification for the β_2m protein of sites associated with an early amyloidogenic intermediated and related to the protein self-association[42]. From experimental findings, binding patches B, E and D (as labelled in the present work) are supposed to be directly involved in

protein self-association process. Our results show that binding of the hydrobic GNPs frequently occurs at those active sites thus being able of blocking the protein from binding to another protein and leading to a further inhibition of the fibrillation activity.

6 Methodology

In order to investigate the microscopic mechanisms regulating the interaction of β_2m with GNPs, we have performed electronic structure calculations on the available experimental structure[23] of the GNP, to develop force fields parameters to be used in rigid docking and classical molecular dynamics simulations, which allow us to take into account the effect of thermal fluctuations associated to the inherent flexibility of the protein and the NP as well as the mediation of the water solvent and counterions.

6.1 Materials

6.2 Development of force field parameters (GolP) for $Au_{25}L_{18}^{-}$(L=S(CH_2)$_2$Ph) nanoparticle

The OPLS/AA force field that we adopted in the simulations does not contain parameters for the GNP which is interacting with the protein and that we wish to simulate. Therefore, we generated force field parameters for $Au_{25}L_{18}^{-}$(L=S(CH_2)$_2$Ph) consistent with the OPLS/AA force field. To this aim, we performed structural optimization and electronic structure calculations starting from the available crystal structure of the nanoparticle at different level of DFT theory and localized basis set, namely B3LYP/lanl2dz, BHH/lanl2dz, BP86/lanl2dz, BP86/TZP using the Gaussian09 and ADF computer packages, respectively. Finally the best fitting to the experimental crystal structure was obtained using the BP86/lanl2dz using the Gaussian09 package. The missing bond and angle parameters were derived on selected fragments of the NPs, performing extensive *ab initio* PES scanning followed by validation of the newly derived force field parameter to reproduce the *ab initio* results (details are reported in Tab. S5 of S.I.). Charges were derived by fitting the *ab initio* ESP grid points following the Merz-Kollman scheme[43], and from these a set of RESP (Restrained ElectroStatic Potential) atomic charges was derived[44]. The interactions of the NPs with biomolecules are assumed to be dominated by dispersive forces, which are described with a Lennard-Jones potential. To model these interactions we have created a model of the NP interacting with a methane molecule via MP2 calculations, which was used to derive atom-atom 12-6 LJ parameters. Short test MD simulations of the target $Au_{25}L_{18}^{-}$(L=S(CH_2)$_2$Ph) over 20 ns gave satisfactory results, in terms of stability of the system as inspected by the analysis of partial and total root mean square deviations of the structural parameters (not reported here).

6.3 Brownian Dynamics

Rigid-body docking simulations were carried out using Brownian dynamics (BD) techniques with the ProMetCS continuum solvent model augmented with protein-GNP interactions parameters[34]. The calculations were performed using the SDA version 6 software[45,46]. As a model for the $Au_{25}L_{18}^{-}$(L=S(CH_2)$_2$Ph) NPs, the representative structure extracted from 20 ns of MD in water was taken.[32] As a model for the protein we took the NMR structure of the native state (PDB: 1JNJ). Human β_2m is a 99-residue-long, 11.9 kDa protein, with a single disulphide bridge between the two Cys residues of the sequence at positions 25 and 80. The protein folds into the classical β-sandwich motif of the immunoglobulin superfamily, i.e. seven antiparallel β-strands (A, B,..., G) forming two facing sheets (ABED and CFG) (as in Panel B (bottom) of Fig. 1). This fold is essentially conserved also when the protein is isolated, both in the crystal[50] and the solution states[51]. Experimental salt concentration of 30 mM was included as a non-specific screening effects on the electrostatic potential of the protein which was calculated using the APBS program.[47] All titratable protein side chains, were assigned their standard protonation state at pH 6.8 with H++,[48]. 5000 BD trajectories were computed starting with the protein positioned uniformly random orientations at a center-to-center separation distance b, and finish when the protein centers reach a separation distance $c > b$ or the two docking partners meet. Parameters b was set to 100 Å and c to 150 Å. The specified number of docked complexes was extracted directly from the runs and clustered with a clustering algorithm. The relative translational diffusion coefficient was 0.0123 Å^2/ps and the rotational diffusion coefficient for the protein was 1.36 x 10^{-4} in radian^2/ps. The simulation timestep was set to 0.50 ps. Parameters for the calculation of hydrophobic desolvation energy/forces was set to -0.019 kcal/mole/Å^2 and for the electrostatic desolvation energy/forces to 1.67 according to ref.[52]. BD trajectories were generated in a sphere (ibox = 0). At each BD step, the protein-nanoparticle interaction energy and forces acting on the protein were computed using the implicit-solvent ProMetCS force field[34], parametrized and implemented for protein-GNP interactions. Two clustering algorithms were tested and evaluated for this system. These were top-down splitting (hierarchical based on a reference structure) and bottom-up aggregating (single-linkage based on RMSD). We selected the algorithm and parameters providing the smallest number of physically distinct orientations of β_2m on GNP. Thus, we applied a single-linkage clustering method (based on CA atoms, with RMSD= 3.0 Å) for all the results given in the manuscript.

6.4 Molecular Dynamics

The same protein and GNPs as for the BD simulations were used as the initial coordinates for the MD simulations. A cubic simulation box of dimension including SPC/E water molecules, the protein and the GNP was built. In the case of a system with a single NP, the dimensions of the box including NP, protein, water and ions, were (70 x 70 x 70)Å^3, whereas in the case of a couple of NPs the dimensions were (90 x 90 x 90)Å^3. The protein was placed at the center of the simulation box and the NPs were kept as in the representatives complexes of the docked clusters obtained from the BD docking simulations. Before the addition of the water molecules, the center of mass of the NP was translated by 5 Å from the center of the protein, retaining the original docked orientation with respect to the protein. The choice of this distance was motivated by various tests that we performed in the past showing that if the simulations were started with the protein in direct contact to the NPs, it results in a kinetically trapped state where only minor relaxation could take place on the time-scale of tens of ns. During equilibration dynamics, all systems contacted the NPs surface within the first 1 ns of MD, in several cases quickly re-orienting with respect to the protein.

All simulations were performed with the Gromacs 4.5.4 package[36]. GolP[32] and newly developed OPLS/AA parameters[37] were used for the NP (as described in Section "Development of force field parameters (GolP) for $Au_{25}L_{18}^-$(L=S$(CH_2)_2$Ph) nanoparticle") and the protein and the SPC/E water model[38] was applied. The lengths of bonds were constrained with the LINCS algorithm. Core gold atoms of the NP were frozen during all simulations. Classical MD simulations were performed at constant volume and temperature (T = 300 K). Periodic boundary conditions and the Particle-Mesh-Ewald algorithm were used. A 2 fs integration time step was used.

In the single NP cases, four independent runs starting from each of the most representative protein-NP docking conformations from docking were performed, each starting with different initial velocities (d1, d2, d3, d4). Simulations performed upon addition of a second NPs to the most representative complex from single NP docking were started with four different initial velocities in two cases and with a single initial velocites other two. The gold NP(s) was(were) coupled to 300 K with a *Nosé-Hoover* thermostat[39]. For the systems with one NP and two NPs, standard MD simulations were conducted for 50 ns and 100 ns, respectively at a constant temperature of 300 K in the NVT ensemble. Trajectories were analyzed in terms of density, temperature, potential energy and other macroscopic properties with the Gromacs tools (*e.g.* g_traj, g_rms, g_clusters *etc.*).

6.4.1 Acknowledgements This work was funded by the Italian Institute of Technology through Platform Computations and Seed project "MOPROSURF-MOdeling PROtein SURFace interactions". Funding from MIUR through PRIN 2012A7LMS3_003 is gratefully acknowledged. The ISCRA staff at CINECA (Bologna, Italy) is acknowledged for computational facilities and technical support. Oak Ridge National Laboratory by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy is acknowledged for the supercomputing project CNMS2013-064. Facilities of the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, are also acknowledged. G.B. and S.C. acknowledge F. Fogolari, A. Corazza and G. Esposito for fruitful discussions.

7 Conclusions

The results provided by our simulations on human β_2-microglobulin consistently suggest that the functionalization of small gold GNPs by hydrophobic ligands can allow the formation of long-life complexes with the protein. This binding of GNPs is able to block active sites of protein domains for the binding to another protein, thus leading to a potential inhibition of the fibrillation activity. Previous simulation with a known inhibitor of fibrillation, doxycycline, provides a matching picture with an overall reduction in protein flexibility as a consequence of stacking interactions of both doxyxycline/NPs with TYR10, ARG12 in strand A and with TRP60 in loop DE. The complexes formation do not disturb the inner core of the protein, but rather aids a general conformational stability strongly favoring the native-like protein structure.

References

1 N. L. Rosi, D. A. Giljohann, C. S. Thaxton, A. K. R. Lytton-Jean, M. S. Han and C. A. Mirkin, *Science*, 2006, **312**, 1027-1030.
2 X. Michalet, F. F. Pinaud, L. A. Bentolila, J. M. Tsay, S. Doose, J. J. Li, G. Sundaresan, A. M. Wu, S. S. Gambhir and S. Weiss, *Science*, 2005, **307**, 538-544.
3 X. Wang, L. L. Yang, Z. Chen and D. M. Shin, *CA Cancer J. Clin.*, 2008, **58**, 97-110.
4 H. K. Li, J. H. Huang, J. H. Lv, H. J. An, X. D. Zhang, Z. Z. Zhang, C. H. Fan and J. Hu, *Angew. Chem., Int. Ed.*, 2005, **44**, 5100-5103.
5 N. Gilbert, *Nature*, 2009, **460**, 937-937.
6 S. Linse, C. Cabaleiro-Lago, W.-F. Xue, I. Lynch, S. Lindman, E. Thulin, S. E. Radford and K. A. Dawson, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 8691-8696.
7 A. S. Pai, I. Rubinstein and H. Onyuksel, *Peptides*, 2006, **27**, 2858-2866.
8 J. E. Kim and M. Lee, *Biochem. Biophys. Res. Commun.*, 2003, **303**, 576-579.
9 S. Rocha, A. F. Thunemann, C. Pereira Mdo, M. Coelho, H. Mohwald and G. Brezesinski, *Biophys. Chem.*, 2008, **137**, 35-42.
10 F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.*, 2006, **75**, 333-366.
11 V. L. Colvin and K. M. Kulinowski, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 8679-8680.

12 M. Mahmoudi, I. Lynch, M. R. Ejtehadi, M. P. Monopoli, F. Baldelli Bombelli and S. Laurent, *Chem Rev.*, 2011, **111**, 5610-5637.

13 R. C. Triulzi, Q. Dai, J. Zou, R. M. Leblanc, Q. Gu, J. Orbulescu and Q. Huo, *Colloids Surf. B Biointerfaces*, 2008, **63**, 200-208.

14 W. H. Wu, X. Sun, Y. P. Yu, J. Hu, L. Zhao, Q. Liu, Y. F. Zhao and Y. M. Li, *Biochem. Biophys. Res. Commun.*, 2008, **373**, 315-318.

15 P. J. Bjorkman, M. A. Saper, B. Samaroui, W. S. Bennet, J. L. Strominger and D. C. Wiley, *Nature*, 1987, **329**, 506-512.

16 J. Floege and M. Ketteler, *Kidney Int.*, 2001, **59**, 164-171.

17 W. H. De Jong and P. J. Borm, *Int. J. Nanomedicine*, 2008, **3**, 133-149.

18 P. M. Heegaard, U. Boas and D. E. Otzen, *Macromol. Biosci.*, 2007, **7**, 1047-1059.

19 L. Crombez, M. C. Morris, S. Deshayes, F. Heitz and G. Divita, *Curr. Pharm. Des.*, 2008, **14**, 3656-3665.

20 D. Aili, K. Enander, J. Rydberg, I. Nesterenko, F. Bjorefors, L. Baltzer and B. Liedberg, *J. Am. Chem. Soc.*, 2008, **130**, 5780-5788.

21 L. Gao, J. Zhuang, L. Nie, J. Zhang, Y. Zhang, N. Gu, T. Wang, J. Feng, D. Yang, S. Perrett and X. Yan, *Nat. Nanotechnol.*, 2007, **2**, 577-583.

22 S. Antonello, N. V. Perera, M. Ruzzi, J. A. Gascoń and F. Maran, *J. Am. Chem. Soc.*, 2013, **135**, 15585-15594.

23 M. W. Heaven, A. Dass, P. S. White, K. M. Holt, and R. W. Murray, *J. Am. Chem. Soc.*, 2008, **130**, 3754-3755.

24 M. Zhu, C. M. Aikens, F. J. Hollander, G. C. Schatz and R. Jin, *J. Am. Chem. Soc.*, 2008, **130**, 5883-5885.

25 Y. Negishi, K. Nobusada and T. Tsukuda, *J. Am. Chem. Soc.*, 2005, **127**, 5261-5270.

26 J. B. Tracy, G. Kalyuzhny, M. C. Crowe, R. Balasubramanian, J.-P. Choi and R. W. Murray, *J. Am. Chem. Soc.*, 2007, **129**, 6706-6707.

27 A. G. Kanaras, F. S. Kamounah, K. Schaumburg, C. J. Kiely and M. Brust, *Chem. Commun.*, 2002, 2294-2295.

28 H. Häkkinen, *Nature Chemistry*, 2012, **4**, 443-455.

29 P. D. Jadzinsky, G. Calero, C. J. Ackerson, D. A. Bushnell and R. D. Kornberg, *Science*, 2007, **318**, 430-433.

30 Y. Pei and X. C. Zeng, *Nanoscale*, 2012, **4**, 4054-4072.

31 S. Kumar and J. Rongchao, *Nanoscale*, 2012, **4**, 4222-4227.

32 F. Iori, R. Di Felice, E. Molinari and S. Corni, *J. Comp. Chem.*, 2009, **30**, 1465-1476.

33 http://projects.villa-bosch.de/mcmsoft/sda/6.00/

34 D. B. Kokh, S. Corni, P. J. Winn, M. Hoefling, K. E. Gottschalk and R. C. Wade, *J. Chem. Theory Comput.*, 2010, **6**, 1753-1768.

35 M. Hoefling, F. Iori, S. Corni and K. E. Gottschalk, *Langmuir*, 2010, **26**, 8347-8351.

36 D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comp. Chem.*, 2005, **26**, 1701-1718.

37 W. L. Jorgensen, D. S. Maxwell and J. TiradoRives, *J. Am. Chem. Soc.*, 1996, **118**, 11225-11236.

38 B. Hess and N. F. van der Vegt, *J. Phys. Chem. B*, 2006, **110**, 17616-17626.

39 A. Cheng and K. M. Jr. Merz, *J. Phys. Chem.*, 1996, **100**, 1927-1937.

40 F. Fogolari, A. Corazza, N. Varini, M. Rotter, D. Gumral, L. Codutti, E. Rennella, P. Viglino, V. Bellotti and G. Esposito, *Proteins*, 2011, **79**, 986-1001.

41 J. E. Ward, R. Ren, G. Toraldo, P. Soohoo, J. Guan, C. O'Hara, R. Jasuja, V. Trinkaus-Randall, R. Liao, L. H. Connors and D. C. Seldin, *Blood*, 2011, **118**, 6610-6617.

42 K. Domanskaa, S. Vanderhaegena, V. Srinivasana, E. Pardona, F. Dupeuxc, J. A. Marquezc, S. Giorgetti, M. Stoppini, L. Wynsa, V. Bellotti and J. Steyaerta, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 1314-1319.

43 N. Spackova, I. Berger, M. Egli and J. Sponer, *J. Am. Chem. Soc.*, 1998, **120**, 6147-6151.

44 E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak and F.-Y. Dupradeau, *Nucl. Acids Res.*, 2011, **39**, W511-W517.

45 R. R. Gabdoulline and R. C. Wade, *Biophys. J.*, 1997, **72**, 1917-1929.

46 www.h-its.org/mcm

47 N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 10037-10041.

48 http://biophysics.cs.vt.edu/H++

49 G. Esposito, A. Corazza and V. Bellotti, *Subcell Biochem.*, 2012, **65**, 1917-1929.

50 C. H. Trinh, D. P. Smith, A. P. Kalverda, S. E. Phillips and S. E. Radford, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 9771-9776.

51 G. Verdone, A. Corazza, P. Viglino, F. Pettirossi, S. Giorgetti, P. Mangione, A. Andreola, M. Stoppini, V. Bellotti and G. Esposito, *Protein Sci.*, 2002, **11**, 487-499.

52 A. H. Elcock, R. R. Gabdoulline, R. C. Wade and J. A. McCammon, *J. Mol. Biol.*, 1999, **291**, 149-162.