THE JOURNAL OF PHYSICAL CHEMISTRY B

Anatolian University Libraries Consortium a·n·k·o·s

Subscriber access provided by HACETTEPE UNIVERSITESI KUTUPHANESI

Article

Molecular Dynamics Simulation of a RNA Aptasensor

Min Ruan, Mahamadou Seydou, Vincent Noel, Benoit Piro, François Maurel, and Florent Barbault

J. Phys. Chem. B, **Just Accepted Manuscript** • DOI: 10.1021/acs.jpcb.6b12544 • Publication Date (Web): 31 Mar 2017

Downloaded from http://pubs.acs.org on April 2, 2017

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.

ACS Publications

The Journal of Physical Chemistry B is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

Molecular Dynamics Simulation of a RNA Aptasensor

Min Ruan[1,2], Mahamadou Seydou[1], Vincent Noel[1], Benoit Piro[1], François Maurel[1] and Florent Barbault[1*]

1 Université Paris Diderot, Sorbonne Paris Cité, ITODYS, UMR 7086, CNRS 15 rue J-A de Baïf, 75013, Paris, France

2 School of Materials and Metallurgy, Hubei Polytechnic University, Huangshi, Hubei, China

* Corresponding author: florent.barbault@univ-paris-diderot.fr

phone: (+33) 157 276 861

Abstract

Single-stranded RNA aptamers have emerged as novel biosensor tools. However, the immobilization procedure of the aptamer onto a surface generally induces a loss of affinity. To understand this molecular process, we conducted a complete simulation study for the Flavin mononucleotide aptamer for which experimental data are available. Several molecular dynamics simulations (MD) of the Flavin in complex with its RNA aptamer were conducted in solution, linked with six thymidines (T6) and, finally, immobilized on an hexanol-thiol-functionalized gold surface. Firstly, we demonstrated that our MD computations were able to reproduce the experimental solution structure and to provide a meaningful estimation of the Flavin free energy of binding. We also demonstrate that the T6 linkage, by itself, does not generate a perturbation of the Flavin recognition process. From the simulation of the complete biosensor system, we observe that the aptamer stays oriented parallel to the surface at a distance around 36 Å avoiding, this way, interaction

with the surface. We evidenced a structural reorganization of the Flavin aptamer binding mode related to the loss of affinity and induced by an anisotropic distribution of sodium cationic densities. This means that ionic diffusion is different between the surface and the aptamer than above this last one. We suggest that these findings might be extrapolated to other nucleic acids systems for the future design of biosensors with higher efficiency and selectivity.

Introduction

Aptamers are highly structured single-stranded RNA or DNA oligonucleotides with approximately 20-100 bases in length. These can be selected to recognize a wide range of biomedical relevant proteins or small-molecules with a great affinity and selectivity, with dissociation constants (K_d) ranging from picomolar to nanomolar. Besides, aptamers present a lowest immunogenicity than antibodies, can be reversibly denatured and refolded without loss of activity, may be chemically synthesized in large quantities at lower price and can be site-specifically modified [1]. Therefore, aptamers are widely regarded as ideal recognition elements for biosensing applications [2-5].

The immobilization procedure of the aptamer onto a surface is a keystone in the design of a biosensor but this step usually induces a decrease of the aptamer affinity and, *in fine*, a reduction of the biosensor device sensitivity [6]. Several molecular effects are hypothesized for explaining this loss of affinity such as: interaction with the surface, binding site occlusion, steric constraint and/or electrostatic repulsion [7-9]. The underlying molecular mechanism is hard to be obtained experimentally, so that molecular simulations represent an affordable alternative to understand the structural factors responsible of the decrease in the binding properties of the aptamer when this one is grafted on a surface.

Recently, a remarkable comparative study of the heterogeneous binding of 12 different immobilized

aptamers was performed using surface plasma resonance [10]. The comparison of the obtained binding affinity constants to the homogeneous situation demonstrates the loss of affinity when the aptamers are chemically immobilized onto gold surfaces. Among the different aptamer/target systems, the Flavin mononucleotide (hereafter called Flavin) RNA aptamer can be used as an archetypal target for the following reasons [11]:

- Both its heterogeneous and homogenous affinity constants are known. The Kd of Flavin with its aptamer is of 500 nM in solution [12] whereas it increases to 710 nM when grafted on an hexanol-thiol self-assembled monolayer on gold surface [10]. The immobilization process clearly induces a loss of affinity even if the recognition ability is still retained.
- The 3D structure of the RNA aptamer/Flavin complex has been experimentally elucidated [13] (pdb code 1FMN) and this structure represents a good starting point for a computational study.
- On the reference work, binding assays were realized with surface plasmon resonance (SPR) so that the detection process cannot be incriminated in the loss of affinity [10]. The loss of affinity should be explained only by the linkage of the aptamer onto the surface.
- The T6 sequence which links the anti-Flavin aptamer to the gold-SAM surface is commonly used in nucleic-acids biosensors [14]. This one is not connected at the vicinity of the Flavin ligand so that we cannot expect that the linkage perturbs the RNA recognition. Besides, SAM Au surfaces modified with alkyl thiol are widely used [15-17] so that our findings would be easier to extrapolate to other systems.

The simulation of a biosensor system at the atomistic level of description requires the consideration of a high number of atoms which can be efficiently handled with MD based on classical force-field. MD technique have been widely used to study the flexibility and to decipher the behavior of biological macromolecules such as proteins, carbohydrates and nucleic acids uses specific potential energy functions with empirical parameters (called a force-field) able to calculate the all-atom energy and forces of large

molecules such as protein, carbohydrates and nucleic-acids [18-19] in their water and ionic environment. Despite the potency of MD techniques and its wide application in biological context, only few works report studies of nucleic-acids at the interface of an inorganic surface. For example, Schatz and collaborators simulated DNA-functionalized gold nanoparticles [20] and the DNA A to B conformational transition at the gold surface interface [21]. To our knowledge, there is yet no molecular modeling study of a complete RNA aptamer based biosensor. Investigation of biomolecule-surface interaction considering both the molecular aspect of solvent and a realistic model of the surface is an important step for the understanding of chemical properties at the interfaces. The role of solvent should be especially taken into account since the surface hydrophobicity can drastically change the water organization around the biomolecule at the interface, leading to a change in its affinity.

Within this context, we report here a complete molecular simulations study of a RNA-Flavin pair immobilized on Au(1,1,1) surface through a flexible spacer. By comparison of computations made in homogeneous and heterogeneous conditions, we demonstrated that the decrease of affinity is due to a RNA helix distortion induced by an anisotropic distribution of cations around the nucleic acids. We then proposed some modifications rules which might help to rationalize the design of more efficient biosensor devices.

Simulation details

1. *Description of simulation systems*

a. RNA/Flavin

Coordinates of the RNA aptamer were extracted from the PDB under the accession code 1FMN [13]. This

entry represents a solution structure of the RNA aptamer in complex with Flavin mononucleotide, called Flavin hereafter (see figure 1). The sequence of the single strand RNA aptamer (5'-GGCGU GUAGG AUAUG CUUCG GCAGA AGGAC ACGCC-3') is characterized by a stable hairpin structure. MD simulations were performed with the amber 12 software [22]. A force-field composed by the ff99SBildn [23] and gaff [24] set of parameters was used to describe the RNA system in interaction with Flavin. The AM1-BCC method [25-26] was used to determine the Flavin partial atomic charges. These conditions, employed for all systems, have already been demonstrated to be consistent for the description of RNA/ligand interactions [27-28]. To take into account the solvation, the RNA/Flavin system was immersed in a rectangular box of 6475 TIP3P water molecules extending to a distance of 10 Å from any solute atom, leading this way to a system of 20644 atoms. These typical values have been demonstrated able to mimic the enthalpic and entropic water solvent behavior for a RNA [29].

b. T6-RNA/Flavin

In the T6-RNA/Flavin system, a chemical linker, composed by 6 Thymine nucleotides (T6) was added on the 5' position of the RNA aptamer, thus on the G1 residue. The 3' position of the T6 tail was connected to an hexanol-thiol (HS-$(CH_2)_6$-O) build with the maestro software [30]. Similarly to the Flavin ligand, the partial atomic charges of the hexanol-thiol moiety were determined with the AM1-BCC method. The T6-RNA/Flavin system was immersed in a rectangular box of 9435 TIP3P water molecules extending to a distance of 10 Å from any solute atom, leading this way to a system of 29747 atoms.

c. Hexanol-thiol SAM Gold Surface

The gold surface model used in this work is the $(2\sqrt{3} \times \sqrt{3})$R30° Au(111) plane with 2 slab layers (six Au atoms for each layer) with saturation coverage c(2 × 2) of hexanol-thiol SAM that means one thiol

molecule per three gold surface atoms. The structure was optimized in the framework of density-functional theory (DFT) at PBE level of calculation [31] under periodic boundary conditions as implemented in VASP 5.4.1 software [32]. The generalized gradient approximation (GGA) was used with the Perdew-Burke-Ernzerhof functional (PBE) [33-34]. The electron-ion interactions were described by the projector augmented wave (PAW) [35-36] method, representing the valence electrons, as provided in the code libraries. The convergence of the plane-wave expansion was obtained using a cut off of 500 eV. The sampling in the Brillouin zone was performed on a grid of k-points of 6×6×1 for the geometry optimizations. The convergence tolerance in the electronic self-consistency loop was 10^{-4} eV/cell, while convergence in the relaxation of the ionic positions was defined by all force components falling below 1 meV/Å. An optimized interatomic distance of 5.10 Å, was found between two sulfur atoms. The hexanol-thiol molecules take an asymmetric bridge position on the surface giving the S-Au bonds length to 2.39 and 2.15 Å. The molecules are oriented relative to the surface with tilt angle (angle between the main axis of the molecule and the surface normal vector of 49.6°. The $(2\sqrt{3} \times \sqrt{3})$R30° supercell describing the surface, displayed on figure 1, was then employed to generate a surface large enough to contain the T6-RNA/Flavin system.

Insert figure 1 here.

d. Gold-SAM-T6-RNA/Flavin

A representative structure of T6-RNA/Flavin was grafted onto the Gold-SAM surface (see figure 1). A water box of TIP3P molecules was built with the solvate module of VMD software [37], leading to a global systems of 72531 atoms. This module was used in order to solvate the system only on the z direction and also to add a 15 Å layer of vacuum to avoid interactions between layers in the PBC box. Position of Au, S and its first covalent carbon atom were kept frozen during the simulation so that distance and tilt angle,

determined previously at the QM level, were respected on the MD simulations. Interaction with the gold atom of the surface is described with the Van der Waals term of the force-field according to the following equation:

$$E_{ligand/surface} = \sum_{i=1}^{Gold} \sum_{j=i+1}^{NB} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] (1)$$

On equation (1) empirical parameters are needed to compute interactions with the gold atoms on the surface. We used the combined rules of Lorentz-Berthelot [38-39] to consider all interactions with the Au and X other atoms according to equations 2 and 3:

$$\sigma_{Au-X} = \frac{\sigma_{Au} + \sigma_X}{2} (2) \epsilon_{Au-X} = \sqrt{\epsilon_{Au} \times \epsilon_X} (3)$$

Values of σ_{Au} and ϵ_{Au} were taken from the literature [40] and respectively set to 2.569 Å and 0.458 eV with the help of the parmed module included in the Amber package software.

e. Gold-SAM/water

For comparative purpose, the Gold-SAM/water system was constructed with the same dimensions as the Gold-SAM-T6-RNA but without the nucleic acid grafted. 45 Na^+ and Cl^- ions were added to this system in order to get the same ionic strength of the Gold-SAM-T6-RNA system. Other parameters are identical to the previous system.

2. MD simulations

All systems were neutralized by adding Na^+ counter ions so that the Particle Mesh Ewald method [41] can be employed to compute electrostatic interactions under periodic boundary conditions. The simulations started with a first minimization for the solvent whereas all solute atoms were restrained with a harmonic potential with a force-field constant of 50 kcal.mol^{-1}.Å^{-1}. This step was followed by an identical energy minimization process without restraint on the solute. The heating phase consisted in 6 steps of 10 ps of MD in the NVT ensemble with temperature jumps of 50 K (from 0 to 300 K) where a weak harmonic potential (5 kcal.mol^{-1}.Å^{-1} force-field constant) are set for the solute atoms to avoid meaningless structural distortion. The SHAKE algorithm [42] was used for all covalent bonds containing a hydrogen atom allowing to use a time increment of 2 fs.

For the RNA/Flavin and T6-RNA/Flavin systems production trajectories of 100 ns were obtained in the NTP ensemble. Temperature regulation at 300 K was ensured through Langevin dynamics [43-44] with a collision frequency of 2 ps^{-1}. For the Gold-SAM-T6-RNA/Flavin and Gold-SAM/water systems, NTP simulations are inconsistent since pressure cannot be isotropic because of the surface. Therefore, NVT simulations were conducted for 100 ns with the Nosé-Hoover method [45] with a thermostat time constant of 0.1 ps to control the temperature at 300K.

3. Analyses

a. Structural analyses

VMD was used to observe the MD trajectories of systems and make the figures. Root Mean Square Deviations (RMSD) and other structural data were computed with the cpptraj module of AmberTools [22]. Representative structures of MD simulations were extracted from their trajectories by selecting the

conformation which displayed the lowest RMSD of their average coordinates.

b. Free energy of binding

Free energies of binding of the RNA/Flavin association were determined with the MMGBSA technique [46]. Other more precise techniques such as Steered Molecular Dynamics or Umbrella Sampling are here difficult to undertake because the dissociation of Flavin could trigger the unfolding of the RNA aptamer. Indeed these structures exist only in the presence of their respective ligands and are folded by these ones. In the case of MMGBSA, only the interaction is computed and this method has proven its efficiency to provide meaningful values of binding free energies for protein/ligand systems [47-49], DNA/ligand [50], RNA/ligand [27] and even carbohydrates systems [51-52]. Besides, the purpose of this paper is not a quantitative description of RNA aptamer biosensor but a rough evaluation of the interaction energy of ligand using MM/GBSA. Therefore, perhaps the trend of binding energy could be accessible by this method, but the absolute values could include significant inaccuracy. The method uses the following set of equations (4-9) to determine the free energy:

RNA + Flavin → complex

$$\langle \Delta G_{binding} \rangle = \langle G_{complexe} \rangle - (\langle G_{RNA} \rangle + \langle G_{FMN} \rangle) \quad (4)$$

$$\langle G_i \rangle = \langle G_i^{gas} \rangle + \langle G_i^{solvent} \rangle \quad (5)$$

$$\langle G_i^{solvent} \rangle = \langle G_i^{polar} \rangle + \langle G_i^{non-polar} \rangle \quad (6)$$

$$\langle G_i^{gas} \rangle = \langle E_i^{gas} \rangle - T\langle S_i^{gas} \rangle \quad (7)$$

$$\langle S_i^{gas} \rangle = \langle S_{translational}^{gas} \rangle + \langle S_{rotational}^{gas} \rangle + \langle S_{vibrational}^{gas} \rangle \quad (8)$$

$$\langle \Delta G_{binding} \rangle = \langle \Delta E_{binding}^{gas} \rangle - T\langle \Delta S_{binding}^{gas} \rangle + \langle \Delta G_{binding}^{solvent} \rangle \quad (9)$$

Binding free energies analyses were made on the last 50 ns of each 100 ns trajectories ensuring this way that equilibrations are reached for all systems. Values in brackets denote the average obtained from calculations of snapshots coordinates, extracted every 20ps. Subscripts i represent the species used for the calculation (complex, nucleic acid or Flavin). Ei, Si and Gi are, respectively, the energy, entropy and free energy of i. The solvent free energy is shared into non-polar and polar contributions (equation 6). For the gas phase entropic contributions (equation 8), the vibrational contribution were estimated with a normal mode analysis, for a subset of 10% of the total snapshots, according to the method developed by Kottalam and Case [53]. Although MMGBSA give a rather crude estimation with a relative weak accuracy of the binding free energy, this method is useful for understanding observed affinities and trends [54].

c. Conformational free energy

The free energy landscape of a macromolecule can be deduced from a MD trajectory by using statistical mechanics. 2D maps were computed from the MD trajectories using two arbitrary structural coordinates, Φ_1 and Φ_2. The free energy is computed according to the formula:

$$\Delta G(\Phi_1, \Phi_2) = -k_B T \cdot \ln\left[\frac{P\left(\rho_{\Phi_1,\Phi_2}\right)}{P_{max}\left(\rho_{\Phi_1,\Phi_2}\right)}\right] (10)$$

In equation 10, k_B and T are the Boltzmann constant and the temperature (300 K) of simulations. P(ρ) is

the probability density function of the reaction coordinates defined by the Φ_1 and Φ_2 variables. $P_{max}(\rho)$ is the most probable state probability. The script of this software is available in supplementary data.

Results and Discussion

For all MD simulations we monitored the kinetics, potential, and overall energies along the trajectory, as well as the densities, pressures or volumes, and temperatures. All these data demonstrate the consistency of the studied systems. For all systems, RMSD curves of selected residues were produced, along the 100 ns trajectories, to observe molecular movements and fluctuations. Figure 2 summarizes the RMSD curves obtained for the RNA aptamer, the Flavin ligand, and the T6 extension of all systems.

Insert figure 2 here.

1. MD of the RNA/Flavin system

The 100ns MD simulation of the RNA/Flavin system shows that there are few evolutions from the experimental structure. A look at figure 2 illustrates this behavior since only small variations are observed along the MD. Indeed, the RNA aptamer presents a RMSD evolution below 5 Å from the experimental starting structure (at t=0). Then, the curve quickly reached a plateau and the variations recorded are weak. A similar conclusion can be raised for the Flavin moiety with an evolution below 2 Å which stays stable along the trajectory. The initial growth of RMSD is due to the fact that a nucleic acid experimental structure is not the most stable one, mainly because in the solid state the RNA is influenced by crystal packing forces[55]. The global flexibility of the RNA aptamer is in accordance with our knowledge on nucleic acids structural studies [27, 50, 56-58]. Besides, this low value of RMSD clearly indicates that we should not observe a difference of binding interaction mode for the analyte. Figure 3 illustrates this fact with a

superimposition of the representative structure of the MD trajectory and the experimental structure (in grey color). We observe that the G10-U12-A25 triple helix, as described on experimental studies[13], remains stable during the MD simulation. This fact emphasizes the stability of this RNA motif required for the Flavin interaction.

Insert figure 3 here

The free energy of binding for the Flavin analyte was determined with the MMGBSA technique and its various contributions are compiled in table 1. A negative value of -24.7 kcal.mol^{-1} signifies that the binding process is spontaneous. We also pay attention on the various contributions of the free energy. The enthalpic gas phase electrostatic contribution is highly repulsive since the aptamer carries numerous negative charges which are close together. However, this repulsive contribution is compensated by the solvent free energy which incorporates the electrostatic shield effect of water molecules and cations. This behavior is in accordance with our knowledge on MMGBSA computations for nucleic acids [27, 50]. The free energy values elucidated here will be further used as references for the other systems.

Insert table 1 here

2. MD of the T6-RNA/Flavin system

A nucleic acids biosensor contains three different elements: a surface, a nucleic acid strand and a linker which is grafted on the surface and connected to the oligonucleotide strand. Only few investigations have been made on the role of this linker. Nevertheless, recently, Martin and coworkers produced an interesting methodological study on the optimal length and nature of nucleic-acids linker on microarrays [6]. They, experimentally, demonstrated that oligo-thymidine linkers do not perturb the aptamer behavior, in contrast

with guanine linkers, for example. They also established that the optimal sequence length should be between 4 and 10. Therefore, the T6 linker fits these requirements and this is why it is widely used for grafting nucleic-acids onto surfaces [14].

We conducted a MD simulation of the T6-RNA/Flavin system in order to investigate the role of this T6 tail on the recognition process. The RMSD curves displayed on figure 2 nicely illustrate the behavior of the new molecular system. Since we started from the experimental solution structure of the RNA/Flavin aptamer, weak RMSD values of the aptamer and Flavin ligand indicate that there are no large structural variations for these parts. Besides, a comparison of values previously obtained for the RNA/Flavin system suggests that the comportment of these molecular fragments is identical. The visualization of the molecular trajectory confirms this assumption: the structure organization of the aptamer and its interaction with the Flavin is identical to the RNA/Flavin reference system. Similarly, the binding free energy of Flavin, as shown on table 1, is identical to the previous system with a value of -26.2 kcal.mol^{-1}. Therefore, we can conclude that the T6 tail do not interferes with the aptamer and its flexibility is high as observed on its RMSD curve (figure 2)..

The second objective of the T6-RNA/Flavin MD exploration was to elucidate the preferential orientation of the T6 fragment toward the aptamer. Indeed, we expected that the simulation of the Gold-SAM-T6-RNA/Flavin system would require a long equilibration time if the system starts far from an optimal orientation of the T6 linker. To determine the best orientation of the T6 tail we defined two distances to describe its flexibility. The first distance, $d_{S\text{-}G1}$, is located between the sulfur atom of the hexanol-thiol moiety, thus the extremity of the T6 tail, and the O5' atom of G1 which represents the first residue of the aptamer where the T6 is linked. The second distance, $d_{S\text{-}G10}$, is located between the same sulfur atom and the O2' of the G10 residue, which is roughly close to the center of gravity of the aptamer. These two distances were chosen in order to describe the elongation of the T6 tail (d1) and the folding to the center

of the aptamer (d2). The 3D free energy landscape map obtained by employing equation 10 is displayed on figure 4. The analysis of this map indicates that only one favorable zone is present in the 3D free energy map. This means that transitions between structures, according to these reaction coordinates, do not require crossing around high energy barriers. The favorable zone is broad which denotes that the flexibility of the T6 linker is substantial. Even if the favorable energy valley is broad, only one optimal conformation exists and this one has a d1 and d2 equals to, respectively, 18.7 Å and 29 Å. This conformation, also displayed on figure 4, is selected for being grafted on the Gold-SAM surface.

Insert figure 4 here

3. MD of the Gold-SAM-T6-RNA/Flavin system

MD of the Gold-SAM-T6-RNA/Flavin system was conducted in the canonical NVT ensemble and with a restrained surface. These conditions are classical when a rigid surface (organic or inorganic) is included on a MD simulation [59-64]. The visualization of the MD trajectory was very informative: the aptamer stays oriented in parallel to the surface at a distance of around 36 Å. We did not observe an event where the aptamer lies down onto the surface - which was our first hypothesis to explain the Flavin affinity decrease. Besides, we did not perceive any interaction of the aptamer with the hydroxyl groups of the hexanol-thiol moieties through hydrogen bond. Therefore, if grafting has a role on the recognition process of the Flavin this one is subtler.

The related RMSD curves of figure 2, advise about the flexibility of the system. Globally, a plateau is reached after some nanoseconds of MD indicating the ability of the system to be equilibrated. The RMSD of the Flavin ligand show only small variations, around 2 Å, for all systems. This is due to the aromatic nature of this ligand which is thus poorly flexible. The RMSD curve of the aptamer is higher than for the

T6/RNA-Flavin and RNA-Flavin systems suggesting a structural change due to the surface vicinity. On the other hand, the RMSD of the T6 tail is lower than for the previous system indicating that the T6 fragment is less flexible when it is grafted on the gold surface, which is expected. From this molecular trajectory, we computed the free energy landscape as function of the coordinates d1 and d2 defined previously. This map, displayed as supplementary data (figure S1) and compared to that obtained for the T6-RNA/Flavin system, shows a translation of the global minimum. The new surface is narrower than for the previous isotropic conditions (i.e. not grafted on a surface) and suggests that the grafting on the functionalized gold surface limits the flexibility of the T6 tail.

Like for the previous systems, we determined the free energy of binding of Flavin through MMGBSA analysis (table 1). The computed global free energy of binding has a value of -11.6 kcal.mol^{-1} which corresponds to a drastic decrease of affinity when compared to the systems computed in isotropic conditions. This loss of affinity is in accordance with experimental results for Flavin's aptamer where the dissociation constant evolves from 500 nM in solution to 710 nM when the aptamer is grafted on the functionalized gold surface [10, 12]. Besides, even if the free energy of binding has been increased when the aptamer was immobilized on the gold surface, this value is significantly negative (-11.6 kcal.mol^{-1}) which means that there is still a molecular recognition of Flavin by the biosensor. All these results are in accordance with experimental data and emphasize the reliability of the simulation. A detailed analyze on the various energy contributions indicates that the loss of affinity may be explained by a decrease of dispersive effects. These effects encompass Van der Waals, hydrogen bonds or stacking interactions and thus suggest a structural reorganization.

Insert figure 5 here

As suspected from the RMSD analyzes, the binding mode of the Flavin evolves drastically when the

aptamer is grafted onto the surface. The G10-U12-A25 triple helix configuration is disturbed and forms a five planes paralleling with Flavin from two sides in the RNA-Flavin-T-S-surface system. Indeed, the aromatic fragment of Flavin is encompassed in a stacking interaction involving residues U12, A13, G10, the Flavin and A25, as shown on figure 5. We also record three weak hydrogen bonds between the phospho-hydroxyl moiety of Flavin and the bases of A11 and G27 residues (figure 5).The association of Flavin with RNA aptamer is thus thermodynamically favorable. However, after the structural evolution of the binding mode, the base-pair association with residue A26 is missing. This base pair is known as a strong stabilizing element of the Flavin association [13]. The evolution of the binding mode is thus in accordance with a higher free energy of binding. To elucidate this phenomenon, we carefully visualized the MD trajectory of the Gold-SAM-T6-RNA/Flavin with a specific focus on the Flavin's binding pocket. After 7,898 ns we observe quick rotations of phosphodiester dihedral angles which trigger the disruption of the A13-G24 and U12-A25 base pairs. Released, the U12 and A13 bases generate a stacking with the G10 residue. This provokes a 90° global orientation of the Flavin molecule which then stacks up with these three bases.

The remaining question is then: why the Flavin's binding mode evolves when the aptamer is grafted on the surface? Our observations demonstrate that we cannot directly incriminate the gold atoms or the hexanol-thiol fragments since they are located too far, around 32 Å, for creating interactions with the ligand. Therefore, we analyzed the ionic distribution around the binding pocket of Flavin since ionic concentration is known to have a drastic influence on nucleic acid structure [18, 65-67]. Moreover, structural rotations around phosphodiester torsion angles triggered by Na^+ concentration evoke the well-known B-Z transition on DNA [68-69]. Therefore, we focused our investigation on the role of Na^+ cations. Firstly, we computed the radial distribution function (RDF) of water and Na^+ around Flavin for the T6-RNA/Flavin and Gold-SAM-T6-RNA/Flavin systems. These graphs are available on supplementary data (figure S2). The comparison of these curves show that the densities of water molecules around Flavin are almost the same

in the different two systems, when that of Na^+ is 25% more for the surface system than the T6-RNA/Flavin system at the direct proximity of the binding pocket (radius of 2.8 Å). Besides, the ionic distribution in the T6-RNA/Flavin system highlights a periodicity which is classical in isotropic system whereas this periodicity is not observed when the aptamer is immobilized on the surface (figure S2, right). This fact was also perceived when we determined the mean square displacements (MSD, equation 11) of the Na^+ cations along the x, y and z axis during the trajectory time for the T6-RNA/Flavin and Gold-SAM-T6-RNA/Flavin systems (see figure S3). From these computations we can determine the diffusion coefficient, D, according to the Einstein equation:

$$D = \frac{1}{2n} \times \lim_{t\to\infty} \left\{ \frac{1}{t} \times \frac{1}{T} \sum_{t=1}^{T} (x(t) - x_0)^2 + (y(t) - y_0)^2 + (z(t) - z_0)^2 \right\} = \frac{1}{2n} \times \lim_{t\to\infty} \frac{MSD}{t} \quad (11)$$

In this last equation, n is the number of dimensions and is classically set to 3 which signifies that the ions can move on the three axes (3D). According to equation 11, we found a similar diffusion coefficient for the T6-RNA/Flavin and Gold-SAM-T6-RNA systems with respective values of 1.5 10^{-5} and 1.2 10^{-5} cm^2/s. Nevertheless, on figure S3, we see that the MSD curves along the x, y and z axis are almost the same for the T6-RNA/Flavin system whereas, for the Gold-SAM-T6-RNA/Flavin system, the MSD curve of the z axis is flat and never rise which is quite different than the curves along the x and y. This means that the Na^+ cations cannot move freely and are confined along the z axis because of the functionalized gold surface.

Insert figure 6 here

Therefore, we deciphered the ionic density along the z-axis of the Gold-SAM-T6-RNA/Flavin system. To do so, we divided the whole space in 22 layers of 4 Å each, along the z-axis (see illustration on figure 6).

Then we counted, for each i layer (L_i), the number of ions (p_i) along the trajectory (t to T). Finally, we normalized the values by dividing by the total number of trajectory frames (T). The equation 12 summarizes this calculation:

$$L_i = \frac{1}{T}\sum_{t=1}^{T} p_i(t) \quad (12)$$

This equation was computed for the Gold-SAM-T6-RNA/Flavin and the Gold-SAM/water systems and the histograms are displayed on figure 6. The script of this software is available on supplementary data. These histograms are then a good witness of the sodium cationic densities along the z-axis. For both systems, we observe low cationic densities on the first layer (from 0 to 4 Å) because the first layer is mainly occupied by the hexanol-thiol moieties where there is no room for Na^+. The cationic histograms densities of the Gold-SAM/water system show that the Na^+ are distributed uniformly along the z-axis when there is no nucleic acid grafted on its surface (figure 6). In the Gold-SAM-T6-RNA/Flavin system, the distribution is not uniform (figure 6) and densities raised around 36 Å, which is the position of the aptamer organized parallel to the surface. The increase of cationic densities is due to the presence of the negative charges of the RNA phosphodiester groups which attract the sodium ions. Besides this electrostatic effect, we observed an anisotropic distribution around the RNA aptamer of cationic densities. Indeed, this density is 25 % higher between the surface and the RNA than upon RNA. We thus assert that this anisotropic distribution of sodium ions around the Flavin is responsible for the structural evolution to the aptamer binding mode. This conclusion echoes the previous results found on DNA chips [21, 70] where a similar increase of ionic density (around 20%) was observed and associated to an increase of DNA melting temperature [70].

In addition, it has been demonstrated that nucleic acid structures are stabilized by an isotropic distribution

of cations [67], this fact being even amplified for RNA [65]. Such structural perturbations triggered by local variations cationic distributions, should be observed in the range of the nanosecond or even sub-nanosecond [18] which is in accordance with our simulations.

This way, we propose that the global parallel orientation of the aptamer, when it is grafted on a functionalized gold surface, induces a perturbation of the cationic distribution around the aptamer which triggers a structural variation of the binding mode. We think that this type of event might be extrapolated to other RNA aptamers grafted on gold surfaces. Moreover, we believe that a chemical modification of the T6 linker could re-orient the RNA aptamer perpendicularly to the surface avoiding, this way, the anisotropic distribution of cations around the ligand binding pocket. We are now working in this direction.

Conclusions

The presented study represents a complete simulation of the Flavin mononucleotide RNA aptamer biosensor system. This type of investigation, even if not totally new, is scarce and only few references are available since these simulations require a deep control of the surface/biomolecule interface. Therefore, by itself, this study represents a significant improvement on the atomic description of a nucleic acid behavior when this one is immobilized on a surface. This work was divided into several parts. Firstly, the results show that MD computations were able to reproduce the experimental solution structure and to provide an estimation of the Flavin free energy of binding to its RNA aptamer. Besides, the influence of the T6 linkage was investigated and shows that, by itself, the T6 tail does not generate a perturbation of the Flavin recognition process.

The modeling of the complete biosensor system was then considered. Visualization of the trajectory shows that the aptamer stays oriented parallel to the surface at a distance around 36 Å which thus avoids any

direct interaction. However, the Flavin's free energy of binding increase significantly when the system is grafted on the surface which is in accordance with the experimental data. A structural reorganization of the Flavin RNA aptamer explained this fact and is associated with the jump of binding free energy. Indeed, the evolution of the Flavin aptamer binding mode is triggered by an anisotropic distribution of sodium cationic densities perpendicularly to the surface. This signifies that ionic diffusion is different between the surface and the aptamer than above this last one. These findings are in accordance with literature data found for other systems and nicely explain the discrepancies of free energies encountered for nucleic acids biosensors. Moreover, this study identifies that the parallel/perpendicular orientation of the aptamer toward the surface is a keystone for the recognition and must be mastered. This result can then be extrapolated to other nucleic acids biosensors and our prospect is now to design and/or modify the linkage so that the nucleic acids will be perpendicularly oriented to the surface and the cation densities would remain isotropic.

Acknowledgments

ANR (Agence Nationale de la Recherche) and CGI (Commissariat Général à l'Investissement) are gratefully acknowledged for their financial support and the post-doctoral fellowship of Miss Min Ruan through Labex SEAM (Science and Engineering for Advanced Materials and devices) ANR 11 LABX 086, ANR 11 IDEX 05-02. This work benefited of the access to the supercomputing facilities of the GENCI (Grand Equipement National pour le Calcul Informatique).

References

(1) Klussman, S. *The Aptamer Handbook: Functional Oligonucleotides and Their Applications*. 2006.
(2) Han, K.; Liang, Z.; Zhou, N. Design Strategies for Aptamer-based Biosensors. *Sensors Basel* **2010,** *10*, 4541-4557.
(3) Kong, R. M.; Zhang, X. B.; Chen, Z.; Tan, W. Aptamer-assembled Nanomaterials for Biosensing and Biomedical Applications. *Small* **2011,** *7*, 2428-2436.

(4) O'Sullivan, C. K. Aptasensors: the Future of Biosensing? *Anal. Bioanal. Chem.* **2002,** *372*, 44-48.
(5) Song, S.; Wang, L.; Li, J.; Fan, C.; Zhao, J. Aptamer-Based Biosensors. *TrAC Trend. Anal. Chem.* **2008,** *27*, 108-117.
(6) Martin, J. A.; Chushak, Y.; Chavez, J. L.; Hagen, J. A.; Kelley-Loughnane, N. Microarrays as Model Biosensor Platforms to Investigate the Structure and Affinity of Aptamers. *J. Nucleic Acids* **2016,** *2016*, 9718612.
(7) Cho, E. J.; Collett, J. R.; Szafranska, A. E.; Ellington, A. D. Optimization of Aptamer Microarray Technology for Multiple Protein Targets. *Anal. Chim. Acta* **2006,** *564*, 82-90.
(8) Collett, J. R.; Eun, J. C.; Ellington, A. D. Production and Processing of Aptamer Microarrays. *Methods* **2005,** *37*, 4-15.
(9) Kautilius, E.; Flores, C.; Woodbury, N. W. Exploring the Sequence Space of a DNA Aptamer using Microarrays. *Nucleic Acids Res.* **2007,** *35*, 7626-7635.
(10) Chang, A. L.; McKeague, M.; Liang, J. C.; Smolke, C. D. Kinetic and Equilibrium Binding Characterization of Aptamers to Small Molecules using a Label-free, Sensitive, and Scalable Platform. *Anal. Chem.* **2014,** *86*, 3273-3278.
(11) Kao, Y. T.; Tan, C.; Song, S. H.; Ozturk, N.; Li, J.; Wang, L.; Sancar, A.; Zhong, D. Ultrafast Dynamics and Anionic Active States of the Flavin Cofactor in Cryptochrome and Photolyase. *J. Am. Chem. Soc.* **2008,** *130*, 7695-7701.
(12) Burgstaller, P.; Famulok, M. Isolation of RNA Aptamers for Biological Cofactors by In Vitro Selection. *Angew. Chem. Int. Edit.* **1994,** *33*, 1084-1087.
(13) Fan, P.; Suri, A. K.; Fiala, R.; Live, D.; Patel, D. J. Molecular Recognition in the FMN-RNA Aptamer Complex. *J. Mol. Biol.* **1996,** *258*, 480-500.
(14) Wu, Y.; Lai, R. Y. Effects of DNA Probe and Target Flexibility on the Performance of a "Signal-on" Electrochemical DNA Sensor. *Anal. Chem.* **2014,** *86*, 8888-8895.
(15) Brown, K. R.; Fox, A. P.; Natan, M. J. Morphology-Dependent Electrochemistry of Cytochrome c at Au Colloid-Modified SnO2 Electrodes. *J. Am. Chem. Soc.* **1996,** *118*, 1154-1157.
(16) Xu, C.; Cai, H.; He, P.; Fang, Y. Electrochemical Detection of Sequence-Specific DNA using a DNA Probe Labeled with Aminoferrocene and Chitosan modified Electrode Immobilized with ssDNA. *Analyst* **2001,** *126*, 62-65.
(17) Lei, C. X.; Hu, S. Q.; Gao, N.; Shen, G. L.; Yu, R. Q. An Amperometric Hydrogen Peroxide Biosensor Based on Immobilizing Horseradish Peroxidase to a Nano-Au Monolayer Supported by Sol-Gel Derived Ccarbon Ceramic Electrode. *Bioelectrochemistry* **2004,** *65*, 33-39.
(18) Cheatham, T. E. Simulation and Modeling of Nucleic Acid Structure, Dynamics and Interactions. *Curr. Opin. Struct. Biol.* **2004,** *14*, 360-367.
(19) Leach, A. R. *Molecular modelling: principles and applications*. 2 ed.; Harlow, 2001.
(20) Lee, O.-S.; Schatz, G. C. Molecular Dynamics Simulation of DNA-Functionalized Gold Nanoparticles. *J. Phy. Chem. C* **2009,** *113*, 2316-2321.
(21) Lee, O.-S.; Cho, V. Y.; Schatz, G. C. A- to B-Form Transition in DNA Between Gold Surfaces. *J. Phys. Chem. B* **2012,** *116*, 7000-7005.
(22) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R.; Zhang, W.; Merz, K. M., et al. *AMBER*, 12.0; University of California: San Francisco, 2012.
(23) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins* **2010,** *78*, 1950-1958.
(24) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004,** *25*, 1157-1174.

(25) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC model: I. Method. *J. Comput. Chem.* **2000,** *21*, 132-146.
(26) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC model: II. Parameterization and Validation. *J. Comput. Chem.* **2002,** *23*, 1623-1641.
(27) Barbault, F.; Ren, B.; Rebehmed, J.; Teixeira, C.; Luo, Y.; Smila-Castro, O.; Maurel, F.; Fan, B.; Zhang, L.; Zhang, L. Flexible Computational Docking Studies of New Aminoglycosides Targeting RNA 16S Bacterial Ribosome Site. *Eur. J. Med. Chem.* **2008,** *43*, 1648-1656.
(28) Duca, M.; Malnuit, V.; Barbault, F.; Benhida, R. Design of Novel RNA Ligands that Bind Stem-Bulge HIV-1 TAR RNA. *Chem. Commun.* **2010,** *46*, 6162-6164.
(29) Auffinger, P.; Westhof, E. Melting of the Solvent Structure Around a RNA Duplex: a Molecular Dynamics Simulation Study. *Biophys. Chem.* **2002,** *95*, 203-210.
(30) Schrödinger: Maestro suite, S. M., *LLC New York, NY* **2011**.
(31) Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. Atoms, Molecules, Solids, and Surfaces: Applications of the Generalized Gradient Approximation for Exchange and Correlation. *Phys. Rev. B* **1992,** *46*, 6671-6687.
(32) Kresse, G.; Hafner, J. Ab initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993,** *47*, 558-561.
(33) Hammer, B.; Hansen, L. B.; Norskov, J. K. Improved Adsorption Energetics within Density-Functional Theory using Revised Perdew-Burke-Ernzerhof Functionals. *Phys. Rev. B* **1999,** *59*, 7413-7421.
(34) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple (vol 77, pg 3865, 1996). *Phys. Rev. Lett.* **1997,** *78*, 1396-1396.
(35) Blochl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994,** *50*, 17953-17979.
(36) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999,** *59*, 1758-1775.
(37) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996,** *14*, 33-38.
(38) Berthelot, D. Sur le Mélange des Gaz. *CR. Acad. Sci.* **1898,** *1898*, 1703-1706.
(39) Lorentz, H. A. Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. *Annalen der Physik* **1881,** *248* (1), 127-136.
(40) Agrawal, P. M.; Rice, B. M.; Thompson, D. L. Predicting Trends in Rate Parameters for Self-Diffusion on FCC Metal Surfaces. *Surf. Sci.* **2002,** *515*, 21-35.
(41) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993,** *98*, 10089-10092.
(42) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J Comput. Phys.* **1977,** *23*, 327-341.
(43) Pastor, R. W.; Brooks, B. R.; Szabo, A. An Analysis of the Accuracy of Langevin and Molecular Dynamics Algorithms. *Mol. Phys.* **1988,** *65*, 1409-1419.
(44) Uberuaga, B. P.; Anghel, M.; Voter, A. F. Synchronization of Trajectories in Canonical Molecular-Dynamics Simulations: Observation, Explanation, and Exploitation. *J. Chem. Phys.* **2004,** *120*, 6363-6374.
(45) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé–Hoover chains: The Canonical Ensemble via Continuous Dynamics. *J. Chem. Phys.* **1992,** *97*, 2635-2643.
(46) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W., et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Accounts Chem. Res.* **2000,** *33*, 889-897.

(47) Li, A.; Barbault, F.; Maurel, F.; Delamar, M.; Wang, B. Interaction Mode and Selectivity of the 2PU Inhibitor with the CDK4 and CDK2 Cyclin-Dependant Kinases: A Molecular Dynamics Study. *J.Mol. Struc. THEOCHEM* **2008,** *849*, 62-75.

(48) Barbault, F.; Maurel, F. Is Inhibition Process Better Described with MD(QM/MM) Simulations? The Case of Urokinase Type Plasminogen Activator Inhibitors. *J. Comput. Chem.* **2012,** *33*, 607-616.

(49) Li, Y.; Delamar, M.; Busca, P.; Prestat, G.; Le Corre, L.; Legeai-Mallet, L.; Hu, R.; Zhang, R.; Barbault, F. Molecular Modeling Study of the Induced-Fit Effect on Kinase Inhibition: the Case of Fibroblast Growth Factor Receptor 3 (FGFR3). *J. Comput. Aid. Mol. Des.* **2015,** *29*, 619-641.

(50) Li, A.; Maurel, F.; Barbault, F.; Delamar, M.; Wang, B.; Zhou, X.; Wang, P. Molecular Modeling Study of Binding Site Selectivity of TQMP to G-quadruplex DNA. *Eur. J. Med. Chem.* **2010,** *45*, 983-991.

(51) Gourmala, C.; Luo, Y.; Barbault, F.; Zhang, Y.; Ghalem, S.; Maurel, F.; Fan, B. Elucidation of the LewisX–LewisX Carbohydrate Interaction with Molecular Dynamics Simulations: A Glycosynapse Model. *J.Mol. Struc. THEOCHEM* **2007,** *821*, 22-29.

(52) Luo, Y.; Barbault, F.; Gourmala, C.; Zhang, Y.; Maurel, F.; Hu, Y.; Fan, B. T. Cellular Interaction through LewisX Cluster: Theoretical Studies. *J. Mol. Model.* **2008,** *14*, 901-910.

(53) Kottalam, J.; Case, D. A. Langevin Modes of Macromolecules: Applications to Crambin and DNA Hexamers. *Biopolymers* **1990,** *29*, 1409-1421.

(54) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Dis.* **2015,** *10*, 449-461.

(55) Neidle, S.; Nunn, C. M. Crystal Structures of Nucleic Acids and their Drug Complexes. *Nat. Prod. Rep.* **1998,** *15*, 1-15.

(56) Barbault, F.; Huynh-Dinh, T.; Paoletti, J.; Lancelot, G. A New Peculiar DNA Structure: NMR Solution Structure of a DNA Kissing Complex. *J. Biomol. Struct. Dyn.* **2002,** *19* (4), 649-58.

(57) Girard, F.; Barbault, F.; Gouyette, C.; Huynh-Dinh, T.; Paoletti, J.; Lancelot, G. Dimer Initiation Sequence of HIV-1Lai Genomic RNA: NMR Solution Structure of the Extended Duplex. *J. Biomol. Struct. Dyn.* **1999,** *16*, 1145-1157.

(58) Kieken, F.; Arnoult, E.; Barbault, F.; Paquet, F.; Huynh-Dinh, T.; Paoletti, J.; Genest, D.; Lancelot, G. HIV-1(Lai) Genomic RNA: Combined used of NMR and Molecular Dynamics Simulation for Studying the Structure and Internal Dynamics of a Mutated SL1 Hairpin. *Eur. Biophys. J. Biophy.* **2002,** *31*, 521-531.

(59) Kubiak-Ossowska, K.; Burley, G.; Patwardhan, S. V.; Mulheran, P. A. Spontaneous Membrane-Translocating Peptide Adsorption at Silica Surfaces: a Molecular Dynamics Study. *J. Phys. Chem. B* **2013,** *117*, 14666-14675.

(60) Kubiak-Ossowska, K.; Mulheran, P. A. Multiprotein Interactions During Surface Adsorption: a Molecular Dynamics Study of Lysozyme Aggregation at a Charged Solid Surface. *J. Phys. Chem. B* **2011,** *115*, 8891-8900.

(61) Kubiak-Ossowska, K.; Mulheran, P. A.; Nowak, W. Fibronectin Module FN(III)9 Adsorption at Contrasting Solid Model Surfaces Studied by Atomistic Molecular Dynamics. *J. Phys. Chem. B* **2014,** *118*, 9900-9908.

(62) Snyder, J. A.; Abramyan, T.; Yancey, J. A.; Thyparambil, A. A.; Wei, Y.; Stuart, S. J.; Latour, R. A. Development of a Tuned Interfacial Force Field Parameter Set for the Simulation of Protein Adsorption to Silica Glass. *Biointerphases* **2012,** *7*, 56.

(63) Wei, T.; Carignano, M. A.; Szleifer, I. Molecular Dynamics Simulation of Lysozyme Adsorption/Desorption on Hydrophobic Surfaces. *J. Phys. Chem. B* **2012,** *116*, 10189-10194.

(64) Yancey, J. A.; Vellore, N. A.; Collier, G.; Stuart, S. J.; Latour, R. A. Development of Molecular Simulation Methods to Accurately Represent Protein-Surface Interactions: The Effect of Pressure and its Determination for a System with Constrained Atoms. *Biointerphases* **2010,** *5*, 85-95.

(65) Chen, S.-J. RNA Folding: Conformational Statistics, Folding Kinetics, and Ion Electrostatics. *Annu. Rev. Biophys.* **2008,** *37*, 197-214.
(66) Dolinnaya, N. G.; Fresco, J. R. Single-Stranded Nucleic Acid Helical Secondary Structure Stabilized by Ionic Bonds: d(A(+)-G)10. *Proc. Natl. Acad. Sci. U. S. A.* **1992,** *89*, 9242-9246.
(67) Tan, Z.-J.; Chen, S.-J. Nucleic Acid Helix Stability: Effects of Salt Concentration, Cation Valence and Size, and Chain Length. *Biophys. J.* **2006,** *90*, 1175-1190.
(68) Cherstvy, A. G. Effect of DNA charge helicity on B-Z DNA transition. *J. Chem. Phys.* **2005,** *123*, 116101.
(69) Gueron, M.; Demaret, J.; Filoche, M. A Unified Theory of the B-Z transition of DNA in High and Low Concentrations of Multivalent Ions. *Biophys. J.* **2000,** *78*, 1070-1083.
(70) Lee, O.-S.; Schatz, G. C. Interaction Between DNAs on a Gold Surface. *J. Phys. Chem. C* **2009,** *113*, 15941-15947.

Tables and figures

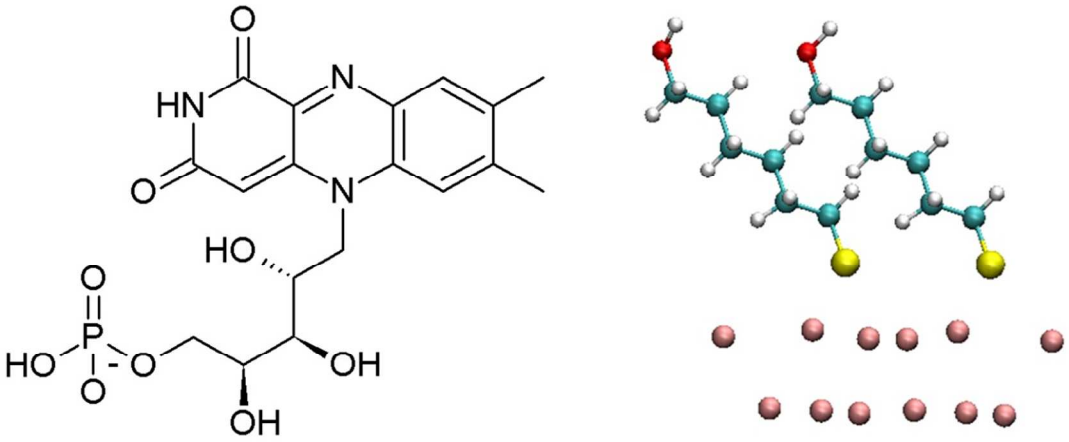


Figure 1. Left, molecular scheme Flavin mononucleotide. Right, The$2\sqrt{3} \times \sqrt{3}$Au(111)SAM-surface unit system used for Gold-SAM QM optimization.

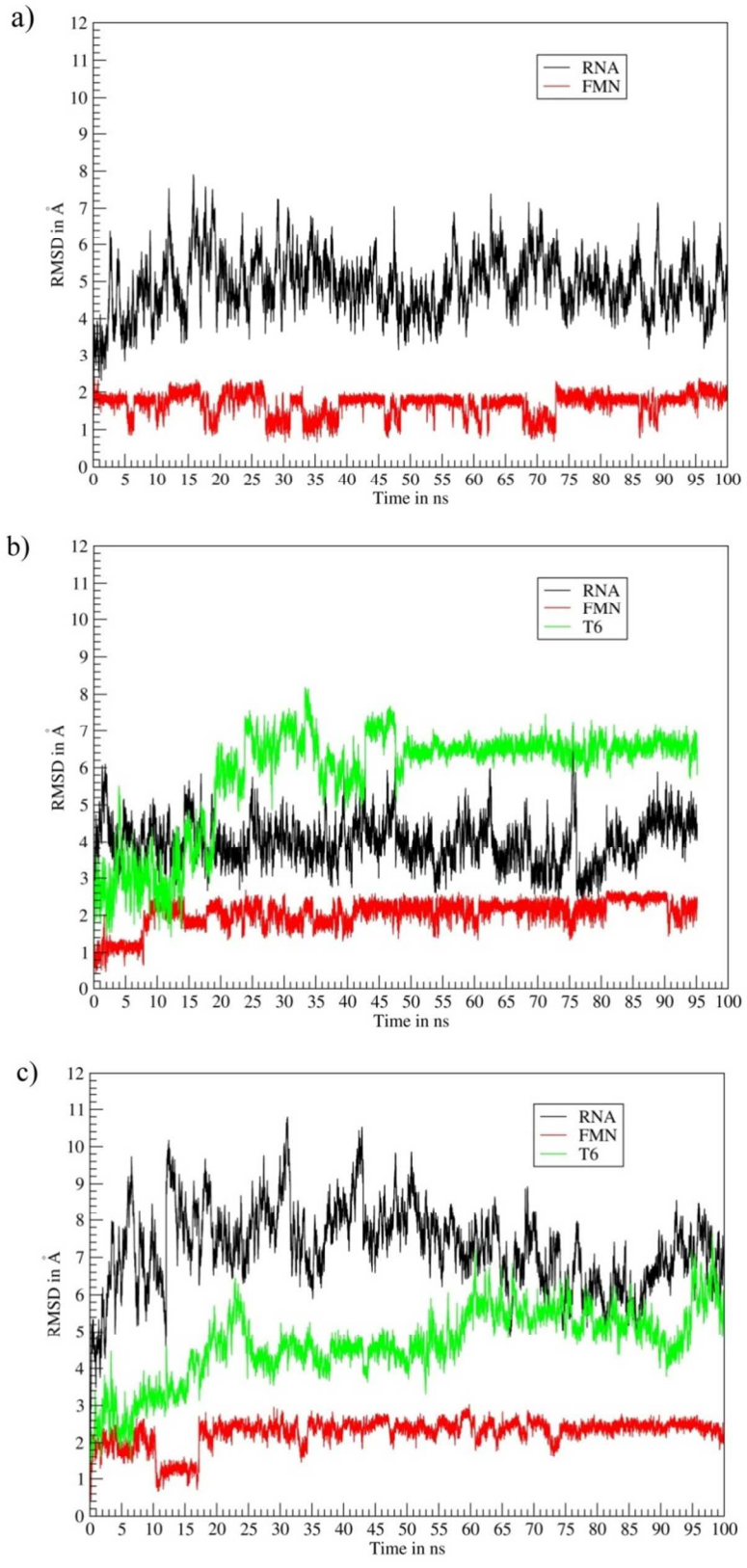


Figure 2: RMSD curves recorded along the 100 ns molecular dynamics of the RNA/Flavin (a) T6-RNA/Flavin (b) and Gold-SAM-T6-RNA/Flavin (c) systems. For each curves, when available, RNA,

Flavin (FMN) and T6 fragments are respectively displayed with black, red and green curves.

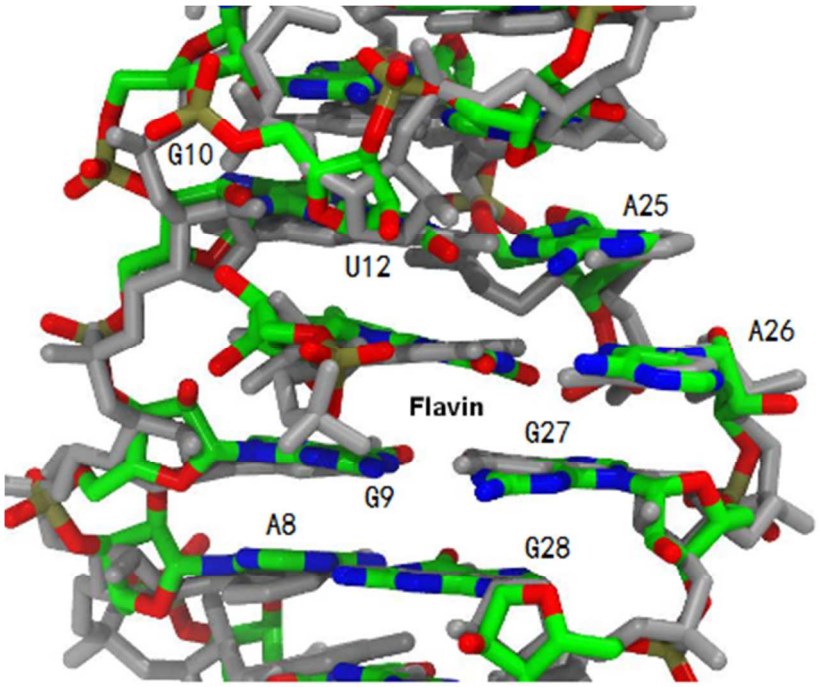


Figure 3: superimposition of the experimental solution structure of the Flavin aptamer (pdb code 1FMN in grey color) and the representative structure extracted from the molecular dynamics simulation of the RNA/Flavin system (with CPK color, carbon in green).

Table 1: The FMN ligand binding free energies on the different studied systems. All values are in kcal.mol^{-1} and were determined with the MMGBSA technique (see material and method).

System name	RNA/FLavin	T6-RNA/Flavin	Gold-SAM-T6-RNA/Flavin
Gas contribution			
ΔE_{elec}(1)	576.8	682.2	571.7
ΔE_{vdw} (2)	-58.2	-66.4	-44.6
ΔH_{gas} (1+2)	518.6	615.8	527.1
$T\Delta S_{gas}$ (3)	-10.6	-12.4	-11.8
ΔG_{gas}(4=1+2-3)	529.2	628,2	538.9
Solvent contribution			
ΔE_{surf} (5)	-3.8	-4.5	-3.8
ΔE_{gb} (6)	-550.1	-650.0	-546.8
$\Delta G_{solvent}$ (7=5+6)	-553.9	-654.4	-550.5
Total			
ΔG_{elec} (1+6)	26.7	32.2	24.9
ΔG_{disper}(2+5)	-62.0	-70.9	-48.4
$\Delta G_{binding}$ (4+7)	-24.7	-26.2	-11.6

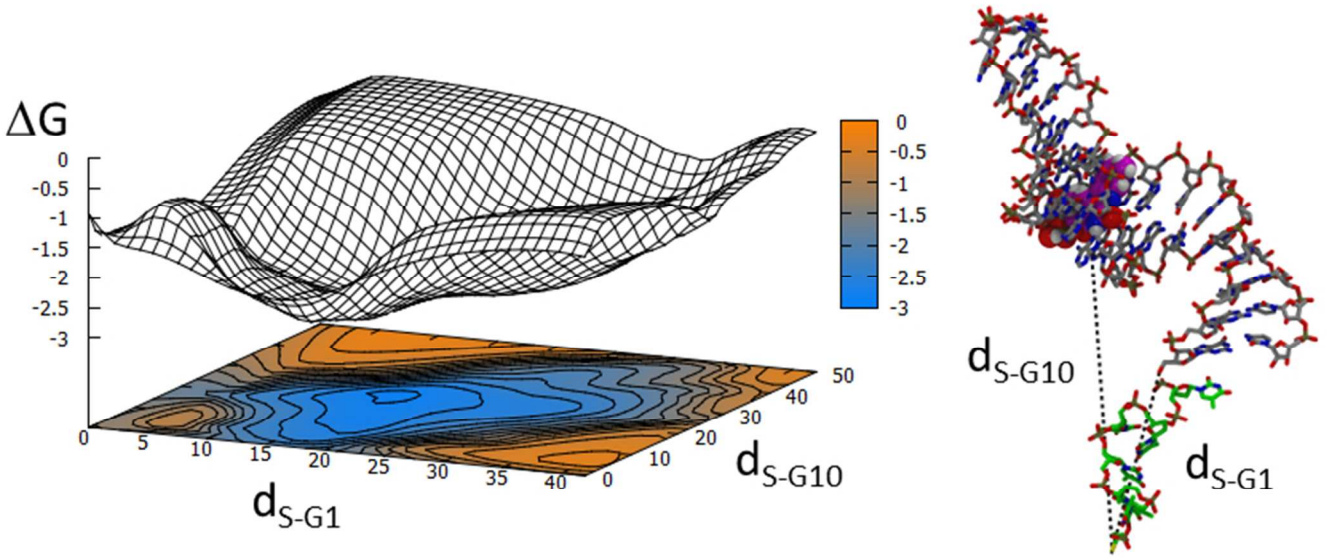


Figure 4: Left : 3D free energy landscape of the T6-RNA/Flavin system. The free energy ΔG, on the z axis, is given in kcal/mol. The x and y axes represent d_{S-G1} and d_{S-G10} in Å. The surface is presented with mesh lines on the xyz coordinates whereas, on the xy plane, a map is presented as isocontours, and with a color ranking from blue (-3.0 kcal mol^{-1}) to orange (0.0 kcal mol^{-1}). Right : the low energy conformation of the T6-RNA/Flavin system.

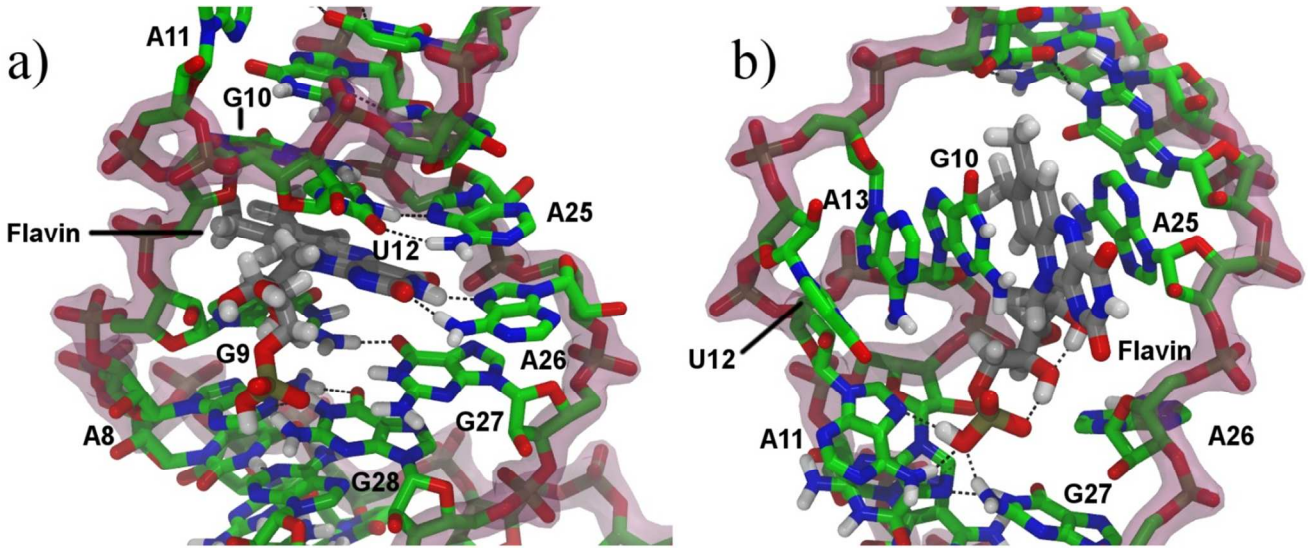


Figure 5: Structural organization of the Flavin inside the aptamer binding pocket before (a) and after (b) its grafting on the gold functionalized surface.

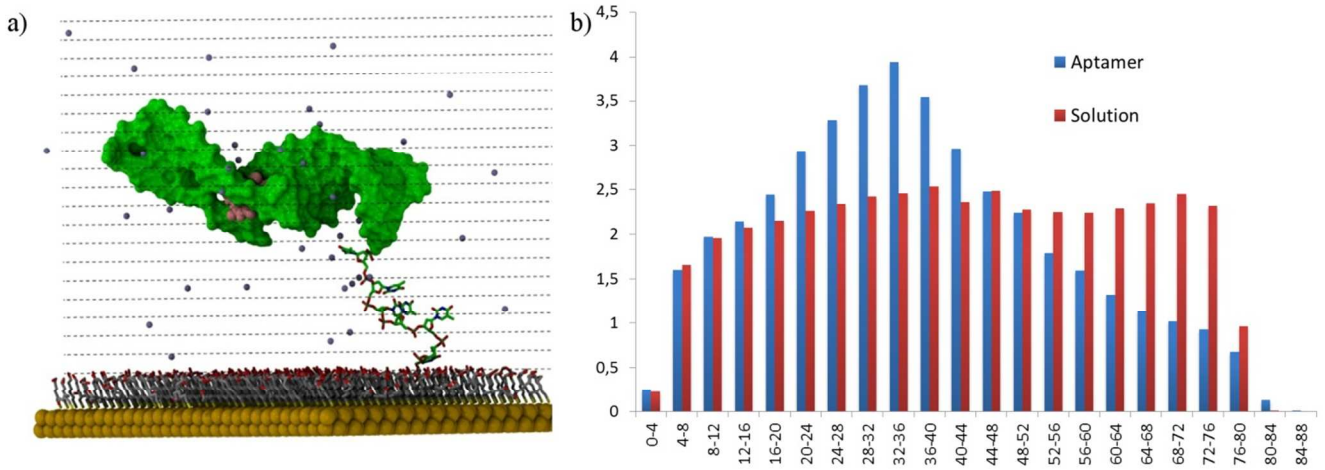


Figure 6: a) The structural organization of the RNA aptamer grafted on the functionalized gold surface. The aptamer and flavin are displayed with molecular surface colored, respectively, in green and pink. The gold surface atoms are presented in yellow with Van der Waals balls whereas the hexanol-thiol self-assembled monolayer moieties are exposed with sticks and CPK atomic colors (for the sake of clarity, the hydrogen atoms are not displayed on this figure). The T6 tail is also shown with sticks but the carbon atom color is green to distinguish this part from the hexanol-thiol fragments. Sodium ions are exhibited with blue Van der Waals balls whereas water molecules are removed from the visualization. Dashed lines, parallel to the gold surface and spaced by 4 Å, illustrate the accounting of cations along the molecular dynamics (see text). b) histograms of the cations distributions (see equation 12) along the molecular dynamics trajectories of the Gold-SAM-T6-Aptamer/Flavin (blue) and the Gold-SAM/water systems (red).The horizontal axis represents the distances to the gold surface whereas the vertical axis is the normalized computation of cations.

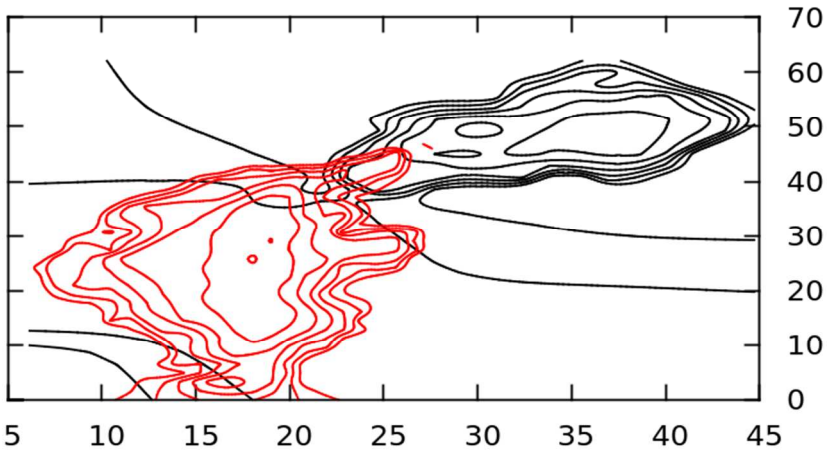


Figure S1: Comparison of the free energy landscape maps obtain for the T6-RNA/Flavin system in isotropic conditions (in red) and when this system is grafted onto the gold functionalized surface (in black). The x and y axis represent, respectively, the distances d1 and d2 as defined in the text.

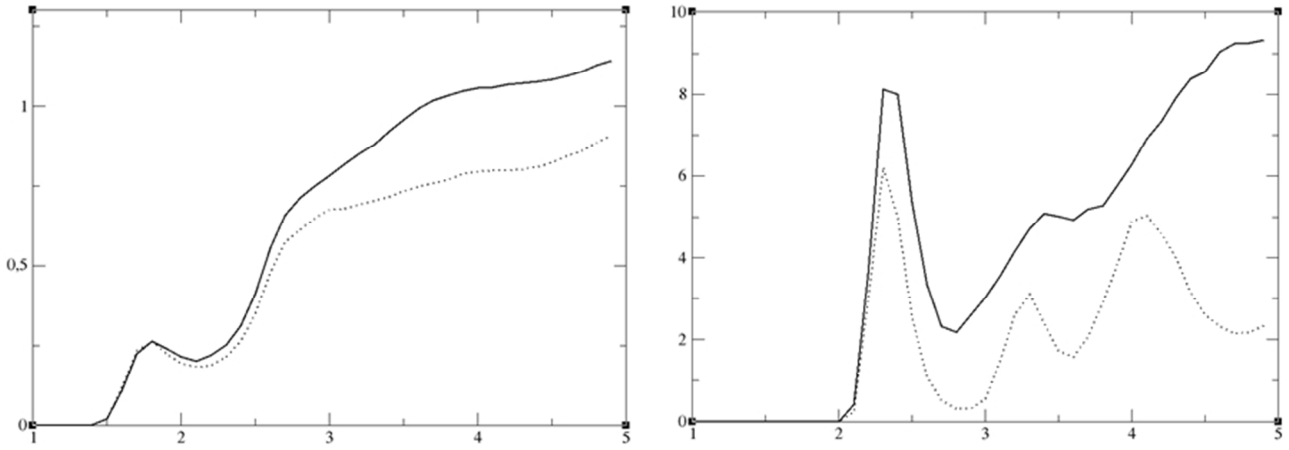


Figure S2: Radial Distribution Function (RDF) of water (left) and Na+ (right) in the Gold-SAM-T6-RNA/Flavin (plain line) and T6-RNA/Flavin (dash line) systems. See text for the interpretation of these curves.

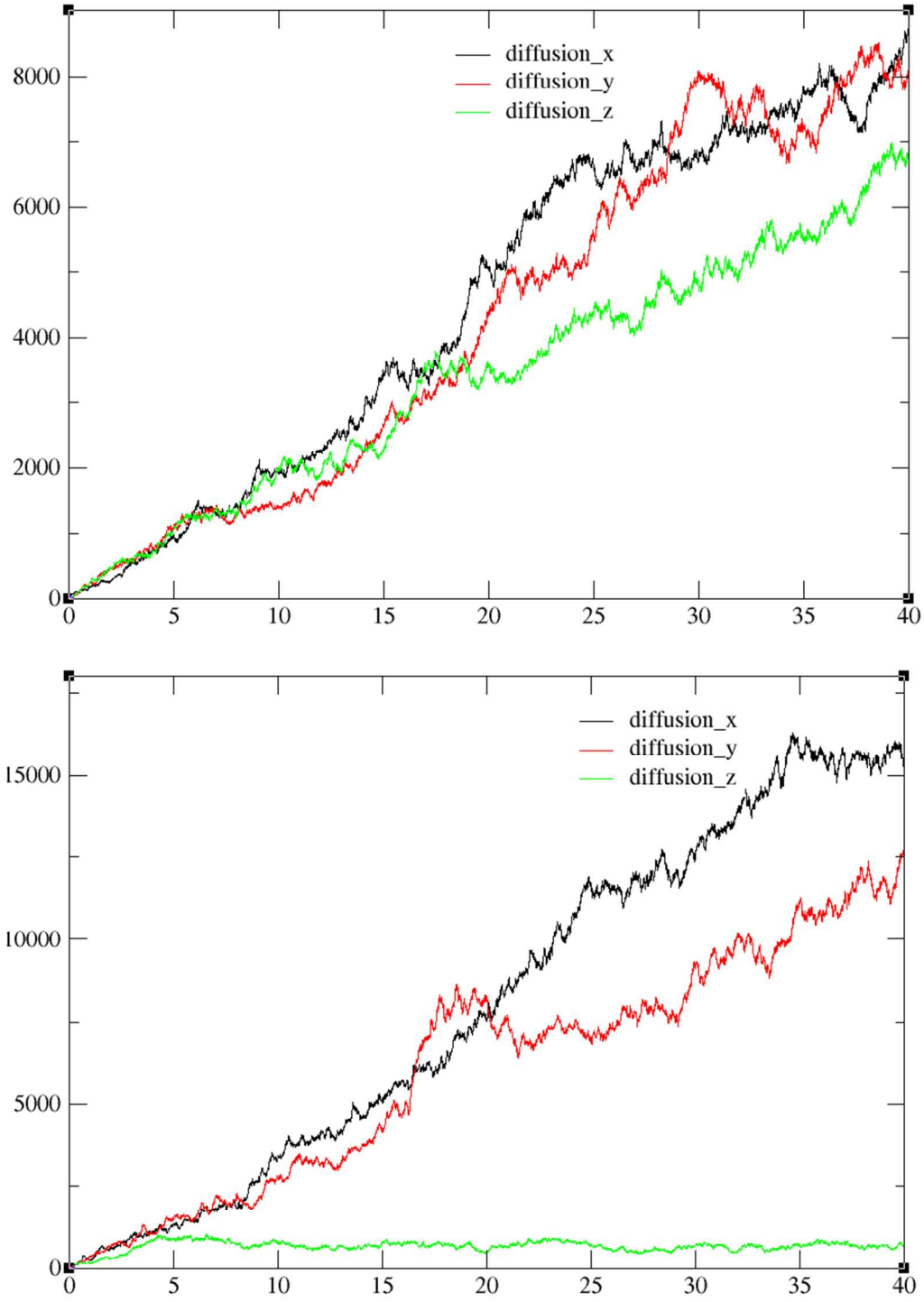


Figure S3: Molecular Square Displacement (MSD) in function of time (in ns) of Na^+ cations along the x, y and z directions. Top and bottom graphs respectively represent the T6-RNA/Flavin and Gold-SAM-T6-RNA/FMN systems.

Supplementary data: additional scripts for structural analyses.

Script 1: map_DG.pl

This perl script use statistical mechanics to compute a conformational free energy landscape with two arbitrary structural coordinates (equation 10 on the manuscript). Temperature of simulation is 300 K but might be change (variable $T). DG unit is kcal/mol but may be change to kJ/mol (variable $kcal). This software needs, as input, a file with three columns: an index, X and Y values. Such file may be easily obtained, for example, with the VMD software.

```
#!/usr/bin/perl
# interface
print "\n\n";
print "******************\n";
print "* Free energy map *\n";
print "******************\n";
print "\nHi, this soft determine the DG map for two calculated metrics.";
print "usage map_DG.pl ^ file\n";
print "\n\nThe file must contains three columns:\n";
print "\tFirst  : an index (1,2,3,...)\n";
print "\tSecond : X values\n";
print "\tThird  : Y values\n";
print "\tThe first row must be removed (no label)\n";
print "\n\nEnjoy.        Florent Barbault\n";
#
# Variable declaration
#
# input variables
#
$file=$ARGV[0];
#
# reading the data
#
$i=0;
open (FILE, "< $file");
 while (<FILE>)
 {
  @tab = split(' ',$_);
  $X[$i]=$tab[1];
  $Y[$i]=$tab[2];
  $i=$i+1;
 }
$tot=$i;
print "You have $tot data\n";
#
# checking the lowest and highest values of x and y
#
$xmin=$X[0];
```

```
$xmax=$X[0];
$ymin=$Y[0];
$ymax=$Y[0];
for ($i=0;$i<$tot;$i++)
{
 if ($X[$i] < $xmin){ $xmin=$X[$i];}
 if ($X[$i] > $xmax){ $xmax=$X[$i];}
 if ($Y[$i] < $ymin){ $ymin=$Y[$i];}
 if ($Y[$i] > $ymax){ $ymax=$Y[$i];}
}
print "\n your values are ranked between";
print "\n $xmin < X < $xmax";
print "\n $ymin < Y < $ymax";
#
# questions: x,y min and max and number of bin
#
print "\nDo you want to keep these extremum values as axis limit?";
print "\nType Y for yes, all other values will lead to no\t";
$choix=<STDIN>;
chop($choix);
if ($choix eq 'Y')
{
 $xm=$xmin;
 $xM=$xmax;
 $ym=$ymin;
 $yM=$ymax;
}       else {
print "\nX min value:\t";
$xm=<STDIN>;
chop($xm);
print "X max value:\t";
$xM=<STDIN>;
chop($xM);
print "Y min value:\t";
$ym=<STDIN>;
chop($ym);
print "Y max value:\t";
$yM=<STDIN>;
chop($yM);
}
print "\nHow many bin you want in the X and Y directions:\t";
$bin_tot=<STDIN>;
chop($bin_tot);
$dx=($xM-$xm)/$bin_tot;
$dy=($yM-$ym)/$bin_tot;
print "\nName of the DG map file:\t";
$file_out=<STDIN>;
chop($file_out);
#
# counting occurence
#
open (FOUT, "> map.dat");
$xcoord=0;
$ycoord=0;
$xbas=0;
$xhaut=0;
$ybas=0;
$yhaut=0;
for ($i=1;$i<=$bin_tot;$i++)
```

```
{
 for ($j=1;$j<=$bin_tot;$j++)
 {
  $xcoord=$xm+$i*$dx-$dx/2;
  $ycoord=$ym+$j*$dy-$dy/2;
  $occur=0;
#
  for ($ind=0;$ind<$tot;$ind++)
  {
  $xbas=$xcoord - $dx/2;
  $xhaut=$xcoord + $dx/2;
  $ybas=$ycoord - $dy/2;
  $yhaut=$ycoord + $dy/2;
  if ($X[$ind] >= $xbas)
  {
   if ($X[$ind] < $xhaut)
   {
    if ($Y[$ind] >= $ybas)
    {
     if ($Y[$ind] < $yhaut)
     {
      $occur=$occur+1;
     }
    }
   }
  }
 }
  printf FOUT ("\n %6.2f   %6.2f   %6.2f",$xcoord,$ycoord,$occur);
}
}
#
# calculate DG
#
$R=8.314;
$T=300;
$kcal=1000*4.18;
$factor=$R*$T/$kcal;
close(FOUT);
open (FG, "> map_DG.dat");
$i=0;
$total=0;
$OCCUR=0;
$OCCUR_max=0;
$max=0;
open (FILE, "< map.dat");
 while (<FILE>)
 {
  @table = split(' ',$_);
  $xx[$i]=$table[0];
  $yy[$i]=$table[1];
  $oc[$i]=$table[2];
  $OCCUR=$OCCUR+$oc[$i];
  if ($oc[$i] > $OCCUR_max){$OCCUR_max=$oc[$i];}
  $i=$i+1;
 }
$total=$i;
#print "$OCCUR $factor";
$DG=0;
for ($ind=0;$ind<$total;$ind++)
```

```
{
if ($oc[$ind] != 0.0){ $DG=-$factor*log($oc[$ind]/$OCCUR_max); if ($DG >
$max){$max=$DG;}}
if ($oc[$ind] == 0.0){ $DG=0;}
 printf FG ("\n %6.2f   %6.2f   %6.2f",$xx[$ind],$yy[$ind],$DG);
}
close(FG);
close(FILE);
$i=0;
open (FSOR, "> $file_out");
open (FILE, "< map_DG.dat");
while (<FILE>)
 {
  @table = split(' ',$_);
  if ($table[2] != 0.0){ $table[2]=$table[2]-$max; }
  printf FSOR ("\n %6.2f   %6.2f   %6.2f",$table[0],$table[1],$table[2]);
  $i=$i+1;
 }
$total=$i;
close (FSOR);
system "rm map.dat";
system "rm map_DG.dat";
```

Script 2: sodium count.

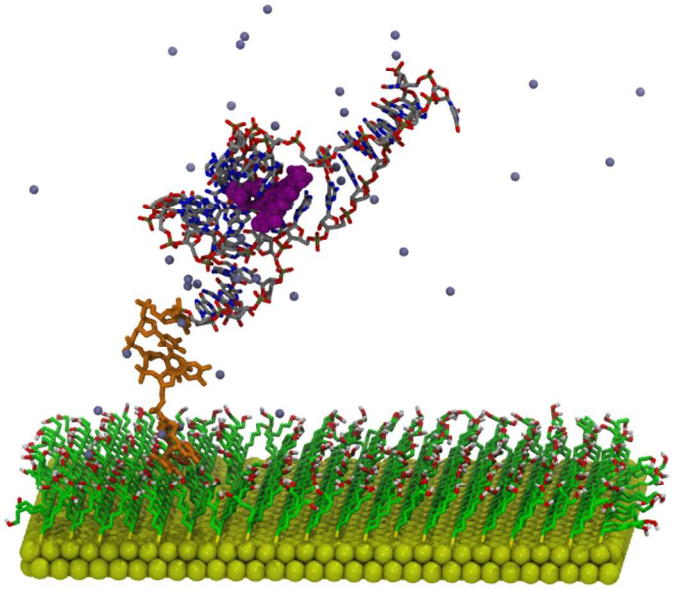
This software count the number of sodium ions between user-defined layers from a series of pdb. These pdb should be obtained from a cpptraj script and a file listing their names must be given as input. The output file is a “comma separated variable” file (csv) and could be open in any excel-like software. Other ions or moieties may be computed with this script (change line 57 Na+ in other ions). The average distance of the SAM is also computed.

```
#!/usr/bin/perl
# interface
print "\n\n";
print "****************\n";
print "* Sodium count *\n";
print "****************\n";
print "\nHi, this soft count the number of sodium in layers. The layers and number
of layers are defined by the user.";
print "\nYou have to first create a file with the list of pdb files inside";
print "\nFor example, typing in a shell : ls -1 distance.* > file_input.flo will
give you this required file";
print "\n\nusage :\t prog_distance_sodium.pl file_input.flo";
print "\n\nEnjoy.        Florent Barbault, 23 June 2015\n";
#
# input variables
#
$file_input=$ARGV[0];
```

```
print "Height of each layer \t :";
$height =<STDIN>;
chop($height);
print "Number of layer \t :";
$num_layer=<STDIN>;
chop($num_layer);
open (FILE, "< $file_input");
open (FSOR, "> dist.csv");
while (<FILE>)
{
$name=$_;
$i=0;
$dist_SAM=0;
$j_moy=0;
print "\nAnalysis of file $name";
open (FICH, "< $name");
 while (<FICH>)
 {
  @tab = split(' ',$_);
  $atom[$i]=$tab[2];
  $res[$i]=$tab[3];
  $Z[$i]=$tab[7];
  if ($res[$i] eq 'SAM')
   {
     if ($atom[$i] eq 'O1')
      {
       $dist_SAM=$dist_SAM+$Z[$i];
       $j_moy=$j_moy+1;
      }
    }
  $i=$i+1;
 }
close (FICH);
$tot=$i;
$dist_SAM=($dist_SAM/$j_moy)-12.314;
print "File $name height of SAM $dist_SAM";
print FSOR "File $name height of SAM $dist_SAM";
$zcor=$dist_SAM+12.314; #12.314 height of the gold
for ($i=0;$i<$tot;$i++)
 {
 if ($res[$i] eq 'Na+')
  {
   $zna=$Z[$i]-$zcor;
   for($inc=0;$inc<$num_layer;$inc=$inc+1)
    {
      $inci=$inc+1;
      if ($zna >= $inc*$height)
       {
         if ($zna < $inci*$height)
           {
            $layer[$inc]=$layer[$inc]+1;
           }
       }
    }
  }
 }
}

for($inc=0;$inc<$num_layer;$inc=$inc+1)
```

```
{
 $inci=$inc+1;
 print FSOR "\n layer $inc between $inc*$height and $inci+$height count for
$layer[$inc]";
 $bas=$inc*$height;
 $haut=$inci*$height;
 print "\n layer $inc between $bas and $haut count for $layer[$inc]";
}
print "\n";
close(FILE);
```



TOC Graphic