Molecular Dynamics (MD) for Cancer control Protocol

CURRENT STATUS: POSTED

Claudio Nicolini
Claudio Ando Nicolini's Lab, NanoWorld, USA; HC Professor Nanobiotechnology, Lomonosov Moscow State University, Russia; Foreign Member Russian Academy of Sciences; President NWI Fondazione EL.B.A. Nicolini; Editor-in-Chief NWJ, Santa Clara, USA

✉ clannicolini@gmail.com Corresponding Author

Marine Bozdaganyan
Claudio Ando Nicolini's Lab, NanoWorld, USA; Lomonosov Moscow State University (MSU), Moscow, Russia; Fondazione EL.B.A. Nicolini

Nicola Luigi Bragazzi
Claudio Ando Nicolini's Lab, NanoWorld, USA

Philippe Orekhov
Claudio Ando Nicolini's Lab, NanoWorld, USA

Eugenia Pechkova
Claudio Ando Nicolini's Lab, NanoWorld, USA

Abstract

In the present protocol, we describe how molecular dynamics (MD) can be applied for studying Langmuir-Blodgett (LB)-based crystal proteins. MD can therefore play a major role in nanomedicine, as well as in personalized medicine.

Introduction

The computer simulation of dynamics in molecular systems is widely used in molecular physics, biotechnology, medicine, chemistry and material sciences to predict physical and mechanical properties of new samples of matter.

In the molecular dynamics method there is a polyatomic molecular system in which all atoms are interacting like material points, and behavior of the atoms is described by the equations of classical mechanics. This method allows doing simulations of the system of the order of 106 atoms in the time range up to 1 microsecond. Despite of some limitations such approximation may describe the dynamics of macromolecules at the atomic level. However, this method does not consider the chemical reactions and the formation or breaking of chemical bonds. For these purposes there are approaches which combine classical and quantum mechanics simulations.

The behavior of each atom can be described as Newton's equation of motion.

Resultant force is the gradient of the potential, which includes the sum of all interactions of atoms in the system. For pairs of atoms there are different types of interactions: covalent and hydrogen bonds, Coulomb and van der Waals interactions. For triplets of atoms - the bond angles interactions, and for fours - torsion angles interactions.

The fundamental point of integration procedure is the choice of the value of the integration step Δt. Step should be less than the period of the fastest motions in the system. For organic molecules the characteristic time of the valence oscillations of the C-H bond is 10 fs. Thus, for full-atom systems we usually choose value of the integration step 1-2fs. Initial velocities usually are randomly selected according to Maxwell distribution for a given temperature.

In the context of molecular modeling, a force field refers to the form and parameters of mathematical functions used to describe the potential energy of a system of particles (typically molecules and

atoms). Force field functions and parameter sets are derived from both experimental work and high-level quantum mechanical calculations. "All-atom" force fields provide parameters for every type of atom in a system, including hydrogen, while "united-atom" force fields treat the hydrogen and carbon atoms in each terminal methyl and each methylene bridge as a single interaction center. "Coarse-grained" force fields, which are frequently used in long-time simulations of proteins and membranes, provide even more crude representations for increased computational efficiency.

The potential energy in the different force fields is common; however they may differ considerably by the calculation of individual contributions and parameterization.

The energy of the covalent bonds is approximated by a harmonic potential.

Spring constant and the equilibrium bond length are chosen to reproduce the geometry and vibrational frequencies of simple model compounds. The main drawback of such representation of valence bonds is that at a great distance for two atoms the bond does not break, and the energy of the system increases dramatically which can lead to artifacts. Most software packages allow to fix valence bonds in the calculation using algorithms SHAKE [1] and LINKS [2]. This approach allows to increase the integration step to 4-5 fs.

The potential energy of the bond angles is also represented as harmonic potential.

Spring constant and the equilibrium value of the angle are selected the same like for bonds to reproduce the equilibrium geometry and vibrational frequency of simple molecules.

The torsion angle for the atoms i, j, k, and l is the dihedral angle between the planes which are defined by atoms ijk and jkl. The angle is 0° when molecule is in cis-conformation.

There are two common energy potentials for dihedral angles: periodic function and potential Ryckaerta-Bellemans.

Both potentials can give an error in the case of aliphatic hydrocarbons where the transition between the cis and trans conformations are sterically hindered. Therefore for a more precise description of system behavior there the non-bonded interactions (Coulomb and van der Waals forces) between the outermost atoms of the torsion angle, so called "1-4 interactions".

In MD atoms are material points and the existence of electrons is not taken into account. However,

the electron density distribution in the molecule is unevenly and partial charge of each atom is used to describe the distribution of electron density in the force fields.

For non-bonded interactions the potential energy is not calculated for each pair of atoms in the system but only for atoms located at a distance of less than a certain value - the cutoff radius.

In structured environments, such as lipid bilayers, using cut-off radius can lead to artifacts so for these systems special approximations are applied to take into account long range Coulomb interactions.

The most common approach is calculating forces using the method of Ewald sums. In this method the sum of all pair interactions is done in the reverse Fourier space. A modified method of Ewald sums is PME (Particle Mesh Ewald) is based on fast Fourier transforms in the summation in reciprocal space.

During MD simulation atoms may be at too small distance to each other (in the forbidden region of the potential field) and at the next step the atom can receive the excess kinetic energy, which subsequently lead to the warming of the system. To avoid this situation, there are special algorithms - thermostats. Here we discuss only two thermostats.

For keeping the temperature of the system constant it is heated externally with fixed temperature Velocity is scaled at each step.

This thermostat has some drawbacks. It does not satisfy canonical Gibbs distribution and does not take into account the law of equipartition of kinetic energy of the molecule degrees of freedom. In the computational experiments with small molecular systems it was observed pumping of energy from high to low frequency vibration modes.

The thermostat becomes an integral part of the system by adding an artificial variable s, coupled with "mass" Q> 0.

Ensembles with constant pressure are often used in molecular dynamics simulations. In particular this can be applied to systems where there are different phases (lipid-water sysrems as example). The most common barostats are Berendsen and Parrinello-Raman.

The pressure tensor can be calculated using the virial theorem of Clausius.

All program software for molecular modeling can be divided by two groups:

1) Molecule/topology builders and visualization. In this type of programs there is graphical interface which allows visualizing different molecules and their editing, building of initial systems. Some examples: Automated Topology Builder (ATB) - automated molecular topology building service for small molecules; UCSF Chimera - 3D molecular viewer, amino acid rotamers and other building functions; Visual Molecular Dynamics (VMD) - 3D molecular viewer, molecule building, analysis of MD trajectories; PyMOL - Python based molecular viewer and editor, has a lot of additional plugins; RasMol - fast viewer of molecular systems;

2) Programs for running MD simulations. LAMMPS - has potentials for soft and solid-state materials and coarse-grain systems, uses spatial-decomposition techniques to partition the simulation domain into small 3d sub-domains; GROMACS - has potentials for all types of biological systems, extremely fast due to algorithmic and processor-specific optimization, typically running 3-10 times faster than many simulation programs; NAMD - has potentials for biological molecules, written using the Charm++ parallel programming model, noted for its parallel efficiency and often used to simulate large systems (millions of atoms); SCIGRESS – is molecular modelling, computational chemistry and materials science software suite, has MM2 and MM3 force fields, using DFT, semi-empirical methods, parallel MD, docking.

First is setting the main aim of simulation: which properties and characteristics are needed to get from the MD trajectories of certain object. The second is choosing the software and finally after all computational procedures – analysis of the results and conclusions.

Parameters for MD simulation are: temperature, pressure, time step, duration of simulation, ensemble type and force field which fits the best for the chosen type of system.

The microcanonical ensemble maintains constant number of particles (N), constant volume (V) and constant total energy (E), so it is also known as the NVE ensemble. The canonical ensemble keeps a constant number of particles, constant volume and constant temperature (T), so it is also called the NVT ensemble. The isothermal-isobaric ensemble keeps constant number of particles, constant pressure and constant temperature, so it is also known as the NPT ensemble. The NPT ensemble is the most important in chemistry, because many chemical processes are performed under constant

pressure and temperature.

Periodic boundary conditions are a set of boundary conditions that are often used to simulate a large system by modeling a small part that is far from its edge.

Molecular dynamics is typically applied to small systems containing thousands of particles. Unless the behaviors near the walls (surface effects) are of interest, they can be eliminated and periodic boundary conditions are employed to simulate the bulk material. Particles are generated in a volume V which is called the primary cell.

The bulk is assumed to be composed of the primary cell surrounded by its exact replicas to model a macroscopic sample. The image cells not only have the same size and shape as the primary one but also contain particles that are images of the particles in the primary cell. Cells are separated by open boundaries so particles can freely enter or leave any cell. When particles leave the cell, their images simultaneously enter the cell through the opposite face. Therefore the shape of the cells must be space filling. The number of image cells needed depends on the range of intermolecular forces. When the forces are sufficiently short ranged (e.g. in truncated Lennard-Jones model), only image cells that adjoin the primary cell are needed (minimum image convention). For squares in two dimensions, there are eight adjacent image cells whereas for cubes in three dimensions, the number of adjacent images is 26.

Atomic structure is an important information about a protein activity and function, since structure and function are known to be interconnected. This information can be obtained by X-ray, nuclear magnetic resonance (NMR), and other less used techniques in the field of crystallography. However, knowledge of all-atom structure represents only the first step towards a full complete understanding of proteins functionality. Bioinformatics, MD and other computational procedures, coupling static structure with its dynamics and fluctuations study, could overcome the limitations of alone crystallography and provide useful structural and functional insights.

X-ray usually provides a protein modeling a static single conformation [3], even though indirect information about potential conformations, flexibility and rigidity can be inferred from B-factor (isotropic and anisotropic), thermal motion parameters, and theoretical estimates. NMR exploits the

6

relaxation parameters [4]. Some advanced techniques – multi-start simulated annealing refinement [5], multi-copy refinement [6], multi-path simulated annealing refinement [7], time-averaging with multiple refinements [8], dynamic-ensemble refinement [9] – have been proposed to study conformational back-bone and side-chain mobility, but MD remains fundamental in order to gain direct insights on protein motion and behavior. Protocols of simultaneous determination of protein native structure and its dynamics have been advocated as vital approaches to capture structural and functional variability [10].

MD has been used to better elucidate about potential thermodynamics (estimating enthalpy and entropy) [11], stability, folding and unfolding disorders [12, 13], structural variants of selected motifs and domains as well as mutated proteins [14-16], in order to foster advancement in the field of protein engineering, protein behavior in different environments, protein-protein interactions [17], from molecular docking [18] to the formation of macromolecular complexes [19], enzyme dynamics, binding [20] and kinetics. MD has promising applications for new drugs design and discovery [21].

For these purposes, MD has been coupled with X-ray, NMR, Atomic Force Microscopy (AFM), Electron Microscopy (EM), spectroscopy, and other experimental imaging procedures, or even integrating all of them in an ambitious multi-scale macromolecular simulation [22-28]. So far little efforts have been made to investigate the advantages and the drawbacks of the crystallization techniques via MD [29].

Different crystallization techniques have been proposed during the decades, like classical vapor drop hanging and its variant sitting drop, dialysis, cryo-temperature, batch, and even in space (using techniques like free interface diffusion or FID, counter-ion diffusion, CID).

Recently a novel crystallization technique based on LB (Langmuir-Blodgett) has been proposed in order to improve crystal growth [30, 31]. Using MD analysis, we show how LB-based crystals and space ones can be compared with classically grown crystals.

Further, MD can be used for nanomedicine. Nanomedicine is the medical application of nanotechnology. Nanomedicine ranges from the medical applications of nanomaterials, to nanoelectronic biosensors and possible future applications of molecular nanotechnology. Current problems for nanomedicine involve understanding the issues related to toxicity and environmental

impact of nanoscale materials.

MD simulation method can be used in different fields of nanomedicine. Here we just overview nanotoxicology part: prediction of interaction of the fullerene and its derivative with biomembrane. Nowadays with increasing production of nanoparticles the problem of its toxicity is becoming more and more important. Fullerenes and its derivatives are being tested continuously to be used in different fields of nanomedicine [33]. As an object of our research we have chosen fullerene C60 and its tris-malonic derivative – radical scavenges – to test its interaction with biomembrane. These radical scavengers have shown to protect cell growth from various toxins that can induce apoptotic injuries in vitro [34, 35] in different cell types such as neuronal cells [36], hepatoma cells [37] and epithelial cells [38].

MD simulations allow to study the process of penetration of nanoparticles through biomembranes in all-atom computational models. For understanding the toxic properties of our molecules we calculated potential of mean force for the systems: fullerene and tris-malonic fullerene with biomembrane. This method allows to analyze the interaction energy between molecules and membrane as function of the distance.

The analysis of the trajectories of the equilibrium molecular dynamics shows that during the first nanoseconds of the simulation fullerene comes to be adsorbed onto the area of lipid heads. However, on the third nanosecond it spontaneously into the tail region (Figure 2.2.1) what is consistent with [39]. The average speed of this jump is approximately 0.35 m/s.

From the Fig. 2.2.1 one can see that fullerene in the beginning overcomes the energy barrier of 10 kJ/mol to pass the lipid heads. The energy minimum (-34 kJ/mol) is located at the distance of 0.75 nm from the center of the membrane. This means that being close to the membrane fullerene molecules got inside the membrane and stays in the region of hydrophobic tails.

Fullerenes, being hydrophobic molecules, in water solution aggregate in clusters. To study the permeability of the membrane for a fullerene cluster, we assembled a system, consisting of 10 fullerene placed above the surface of the membrane (Figure 2.2.2.). At the initial time the fullerenes did not contact with each other or membrane, but after a few nanoseconds, they aggregate and

absorb in the head area. After 4 ns one of the eight fullerenes in the aggregate penetrates the membrane, other seven latter follow it. As a result, after 100 ns 9 of 10 fullerenes penetrate into the membrane.

The presence of fullerene in the membrane does not significantly affect its basic physical properties. However, the membrane is deformed by the addition of a large number of fullerenes: it curves (Figure 4), the area per lipid head is reduced from 69 to 56 Å2, while thickness is increased of 0.5 nm. In general such changes upon the penetration of a large number of fullerenes into the membrane could preface further formation of a pore, a micelle or a membrane rupture, which were not observed due to the limited simulation time and/or the restrained size of the system.

In the simulation with C3, it was adsorbed in the area of hydrophilic lipid heads on the first nanosecond and further preferably remained in this area, rather than in the bulk (Figure 2.2.3). The calculations showed that the molecule of C3 has a stable and well-structured hydration shell of 40 water molecules.

As can be seen from the plot of potential of mean force, the penetration of membrane appears to be energetically unfavorable for C3 derivative – with decrease of the distance between C3 and the membrane the repulsive force driving the fullerene away increases. Such a hindered penetration may occur because of a negative charge carried by C3 as well as due to a high energetical cost of the water shell dissociation which is essential for C3 to pass through the membrane.

We conclude that MD simulations can be used for predictions of toxicity of different molecules to biological objects. The non-modified fullerene (a single molecule and a cluster of ten molecules) spontaneously penetrate the membrane and remain there throughout the entire simulation time. After entry into the eukaryotic membrane, fullerenes remain at a distance of about 1 nm from the center of the membrane. Cluster of fullerenes deform membrane causing its curvature and changing of thickness. This fact means that C60 is potentially toxic molecule for eukaryotic cells. C3 does not penetrate eukaryotic membranes and causes no changes in their structure – the toxicity of this molecule is lower than C60.

Equipment

One of the previously mentioned software for MD design and visualization.

Procedure

MD for proteomics

1. Take some proteins (for example, lysozyme, proteinase K and thaumatin) crystallized by different methods;

2. Put all proteins inside the box;

3. Solvate all proteins;

4. Perform minimization;

5. Perform equilibration of the system running it for 10 ns at different temperatures (for example, 300 and 500 K), in order to understand the stability of proteins (it is expected that under higher temperature protein can have much more conformations then at low temperatures, so we can study better for short time the potential energy surface of the protein);

6. characterize the protein status for example, investigating its radius of gyration, which is indicative of the level of compaction in the structure, i.e. how folded or unfolded the polypeptide chain is, or the root mean square deviation (RMSD); for comparison purpose, normalize it (for example, using a scaling law of Rg ∝ N2/5 [32] for the radius of gyration or computing the normalized root-mean-square deviation or error, NRMSD or NRMSE).

References

1. S. Miyamoto, K.K.K., An analytical version of the shake and rattle algorithm for rigid water models. J. Comput. Chem, 1992. 13: p. 952–962.

2. B. Hess, H.B., H.J. Berendsen, J.G.E.M. Fraaije, A linear constrant solver for molecular simulation. J. Comput. Chem, 1997. 18: p. 1463–1472.

3. DePristo, M.A., P.I. de Bakker, and T.L. Blundell, Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography. Structure, 2004. 12(5): p. 831-8.

4.   Kay, L.E., Protein dynamics from NMR. Nat Struct Biol, 1998. 5 Suppl: p. 513-7.

5.   Rice, L.M., Y. Shamoo, and A.T. Brunger, Phase improvement by multi-start simulated

       annealing refinement and structure-factor averaging. Journal of Applied

       Crystallography, 1998. 31: p. 798-805.

6.   Burling, F.T. and A.T. Brunger, Thermal Motion and Conformational Disorder in

       Protein Crystal-Structures - Comparison of Multi-Conformer and Time-Averaging

       Models. Israel Journal of Chemistry, 1994. 34(2): p. 165-175.

7.   Liu, X.G., et al., Averaging tens to hundreds of icosahedral particle images to resolve

       protein secondary structure elements using a Multi-path Simulated Annealing

       optimization algorithm. Journal of Structural Biology, 2007. 160(1): p. 11-27.

8.   Levin, E.J., et al., Ensemble refinement of protein crystal structures: validation and

       application. Structure, 2007. 15(9): p. 1040-52.

9.   Klepeis, J.L., et al., Long-timescale molecular dynamics simulations of protein

       structure and function. Current Opinion in Structural Biology, 2009. 19(2): p. 120-

       127.

10.   Karplus, M. and J. Kuriyan, Molecular dynamics and protein function. Proceedings of

       the National Academy of Sciences of the United States of America, 2005. 102(19): p.

       6679-6685.

11.   Levitt, M., et al., Potential-Energy Function and Parameters for Simulations of the

       Molecular-Dynamics of Proteins and Nucleic-Acids in Solution. Computer Physics

       Communications, 1995. 91(1-3): p. 215-231.

12.   Beck, D.A. and V. Daggett, Methods for molecular dynamics simulations of protein

       folding/unfolding in solution. Methods, 2004. 34(1): p. 112-20.

13.   Day, R. and V. Daggett, All-atom simulations of protein folding and unfolding. Adv

       Protein Chem, 2003. 66: p. 373-403.

14. Friedland, G.D., et al., A Correspondence Between Solution-State Dynamics of an Individual Protein and the Sequence and Conformational Diversity of its Family. Plos Computational Biology, 2009. 5(5).

15. Kondrashov, D.A., et al., Protein structural variation in computational models and crystallographic data. Structure, 2007. 15(2): p. 169-77.

16. Morrow, J.K. and S.X. Zhang, Computational Prediction of Protein Hot Spot Residues. Current Pharmaceutical Design, 2012. 18(9): p. 1255-1265.

17. Wass, M.N., A. David, and M.J.E. Sternberg, Challenges for the prediction of macromolecular interactions. Current Opinion in Structural Biology, 2011. 21(3): p. 382-390.

18. Ritchie, D.W., Recent progress and future directions in protein-protein docking. Current Protein & Peptide Science, 2008. 9(1): p. 1-15.

19. Terentiev, A.A., N.T. Moldogazieva, and K.V. Shaitan, Dynamic proteomics in modeling of the living cell. Protein-protein interactions. Biochemistry (Mosc), 2009. 74(13): p. 1586-607.

20. Aqvist, J., V.B. Luzhkov, and B.O. Brandsdal, Ligand binding affinities from MD simulations. Accounts of Chemical Research, 2002. 35(6): p. 358-365.

21. Bienstock, R.J., Computational Drug Design Targeting Protein-Protein Interactions. Current Pharmaceutical Design, 2012. 18(9): p. 1240-1254.

22. Kohn, J.E., et al., Evidence of Functional Protein Dynamics from X-Ray Crystallographic Ensembles. Plos Computational Biology, 2010. 6(8).

23. Connelly, L., et al., Effects of point substitutions on the structure of toxic Alzheimer's beta-amyloid channels: atomic force microscopy and molecular dynamics simulations. Biochemistry, 2012. 51(14): p. 3031-8.

24. Chan, K.Y., et al., Cryo-Electron Microscopy Modeling by the Molecular Dynamics

Flexible Fitting Method. Biopolymers, 2012. 97(9): p. 678-686.

25.   Evans, J.E. and N.D. Browning, Enabling direct nanoscale observations of biological reactions with dynamic TEM. Journal of Electron Microscopy, 2013. 62(1): p. 147-156.

26.   Jana, S., et al., Study of microheterogeneous environment of protein Human Serum Albumin by an extrinsic fluorescent reporter: A spectroscopic study in combination with Molecular Docking and Molecular Dynamics Simulation. Journal of Photochemistry and Photobiology B-Biology, 2012. 112: p. 48-58.

27.   Lv, C.M., et al., Low Folding Cooperativity of Hp35 Revealed by Single-Molecule Force Spectroscopy and Molecular Dynamics Simulation. Biophysical Journal, 2012. 102(8): p. 1944-1951.

28.   Frauenfelder, H., et al., A unified model of protein dynamics. Proceedings of the National Academy of Sciences of the United States of America, 2009. 106(13): p. 5129-5134.

29.   Burnley, B.T., et al., Modelling dynamics in protein crystal structures by ensemble refinement. Elife, 2012. 1: p. e00311.

30.   Pechkova, E. and C. Nicolini, Protein nucleation and crystallization by homologous protein thin film template. Journal of Cellular Biochemistry, 2002. 85(2): p. 243-251.

31.   Pechkova, E. and C. Nicolini, Protein nanocrystallography: a new approach to structural proteomics. Trends in Biotechnology, 2004. 22(3): p. 117-122.

32.   Hong, L. and J.Z. Lei, Scaling Law for the Radius of Gyration of Proteins and Its Dependence on Hydrophobicity. Journal of Polymer Science Part B-Polymer Physics, 2009. 47(2): p. 207-214.

33.   Bosi, S., et al., Fullerene derivatives: an attractive tool for biological applications. Eur J Med Chem, 2003. 38(11-12): p. 913-23.

34.   Lin, A.M.Y., et al., Carboxyfullerene prevents iron-induced oxidative stress in rat

brain. Journal of Neurochemistry, 1999. 72(4): p. 1634-1640.

35.   Lin, A.M.Y., et al., Local carboxyfullerene protects cortical infarction in rat brain.
        Neuroscience Research, 2002. 43(4): p. 317-321.

36.   Dugan, L.L., et al., Carboxyfullerenes as neuroprotective agents. Proceedings of the
        National Academy of Sciences of the United States of America, 1997. 94(17): p.
        9434-9439.

37.   Huang, Y.L., et al., Blockage of apoptotic signaling of transforming growth factor-
        beta in human hepatoma cells by carboxyfullerene. European Journal of Biochemistry,
        1998. 254(1): p. 38-43.

38.   Straface, E., et al., C3-fullero-tris-methanodicarboxylic acid protects epithelial cells
        from radiation-induced anoikia by influencing cell adhesion ability. FEBS Lett, 1999.
        454(3): p. 335-40.

39.   Qiao, R., et al., Translocation of C60 and its derivatives across a lipid bilayer. Nano
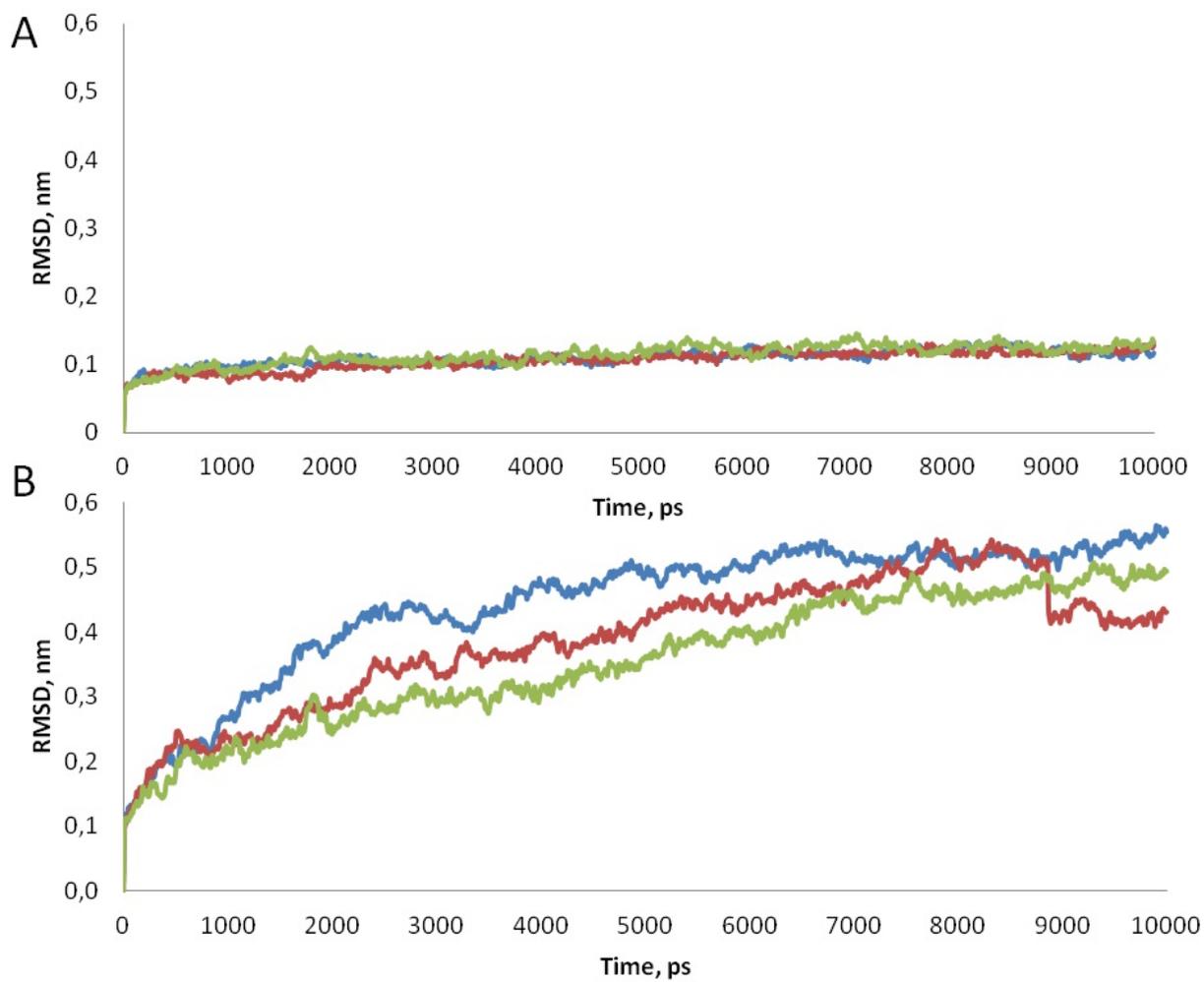   Lett, 2007. 7(3): p. 614-9.

Figures

Figure 1

Averaged RMSD for all proteins at 300 K (A) and 500 K (B). Blue line is classically (HD) crystallized proteins, red is LB and green is space grown crystals.
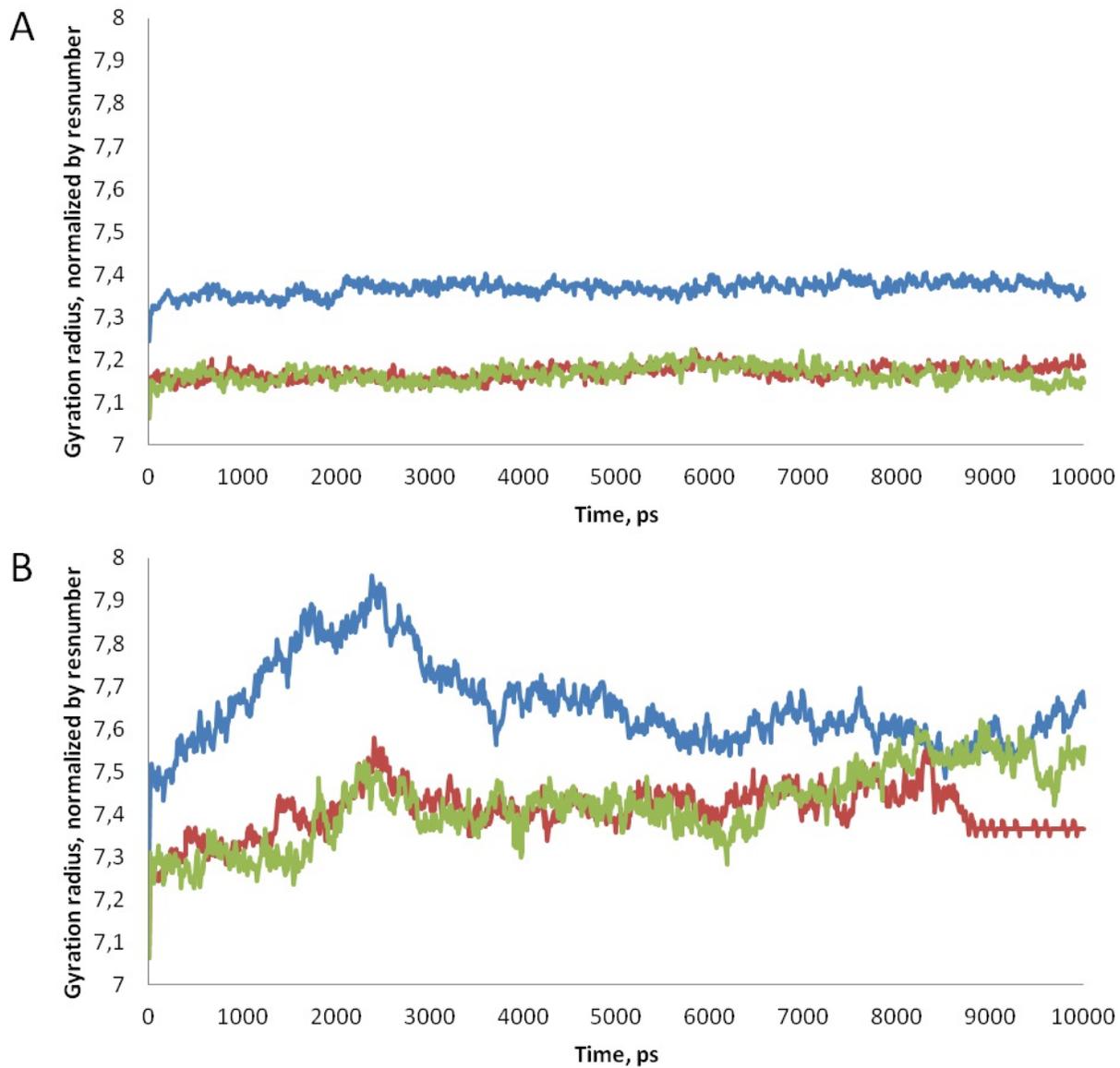
Figure 2

Normalized gyration radius of all proteins prepared by different techniques at 300 K (A) and 500 K (B). Blue line is classically (HD) crystallized proteins, red is LB and green is space grown crystals.
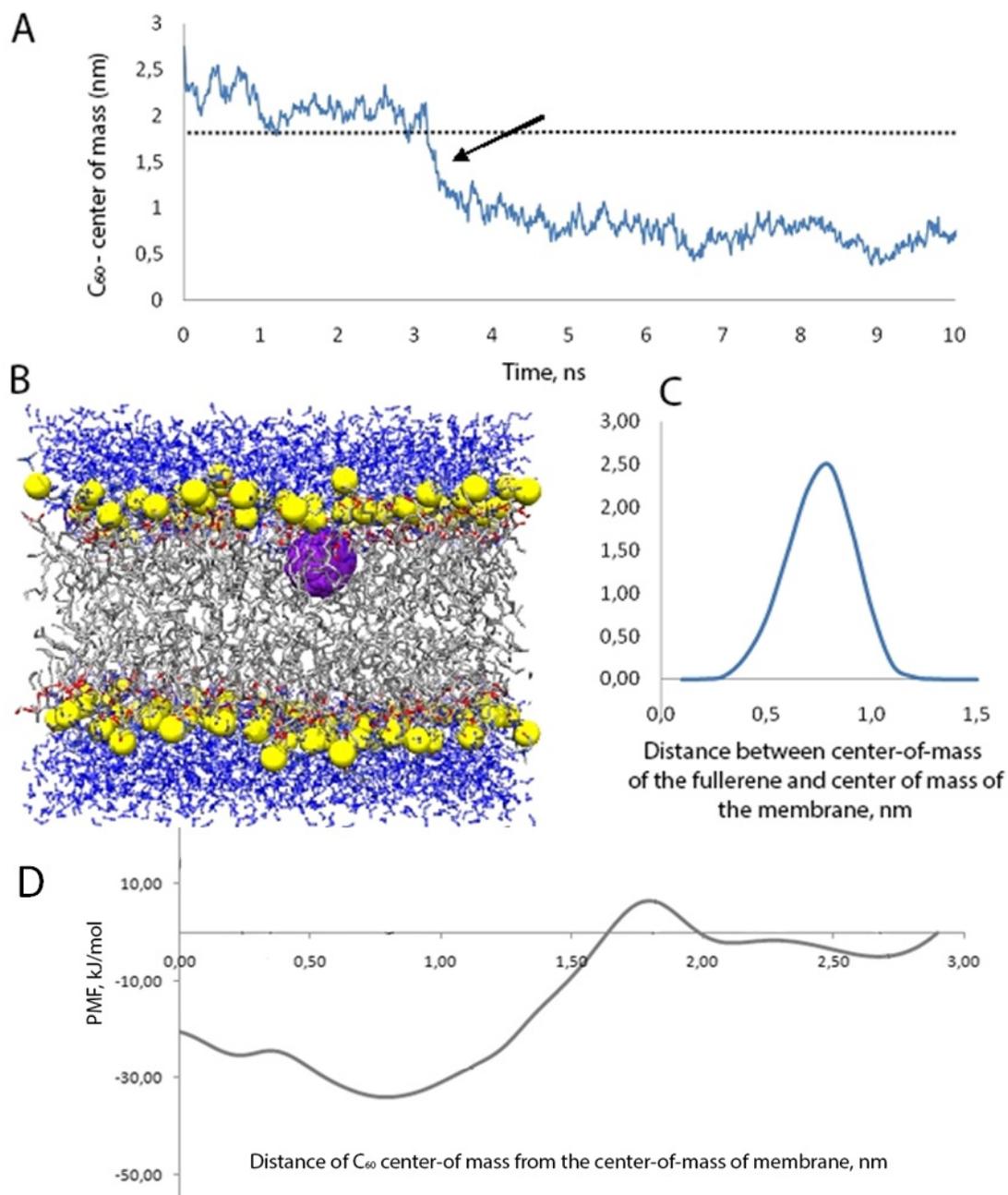
Figure 3

A. Distance between center of mass of the fullerene and center of mass of the membrane during the simulation. On the third nanosecond fullerene spontaneously "jump" into the membrane (shown with an arrow). B. Snapshot of the system with the fullerene insi
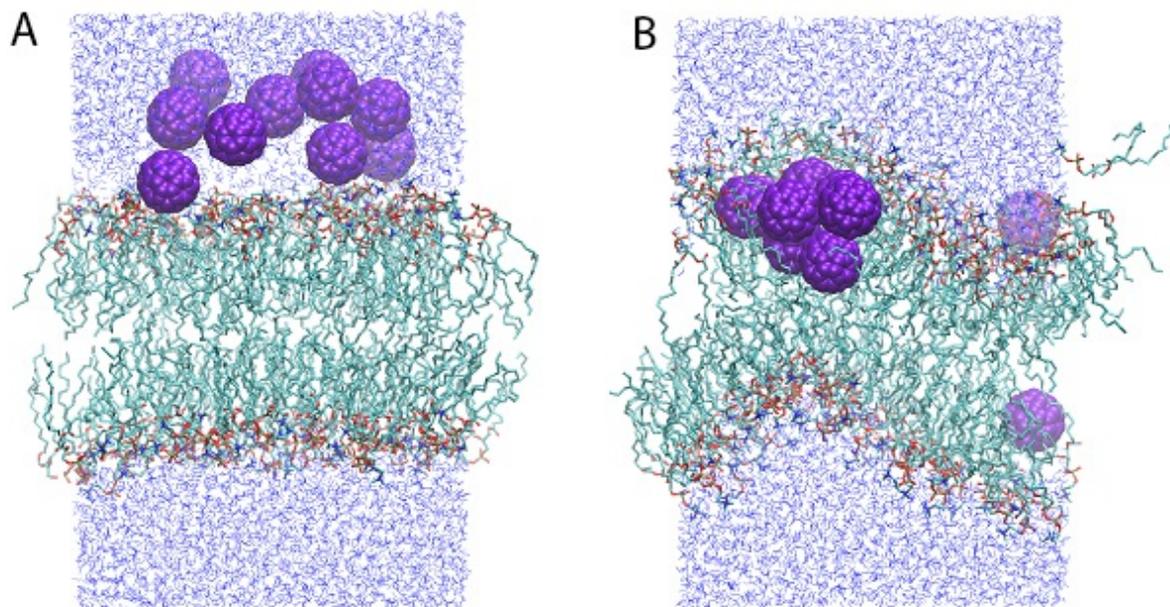
Figure 4

Snapshots of the membrane and ten fullerenes in the initial moment of time and after 100
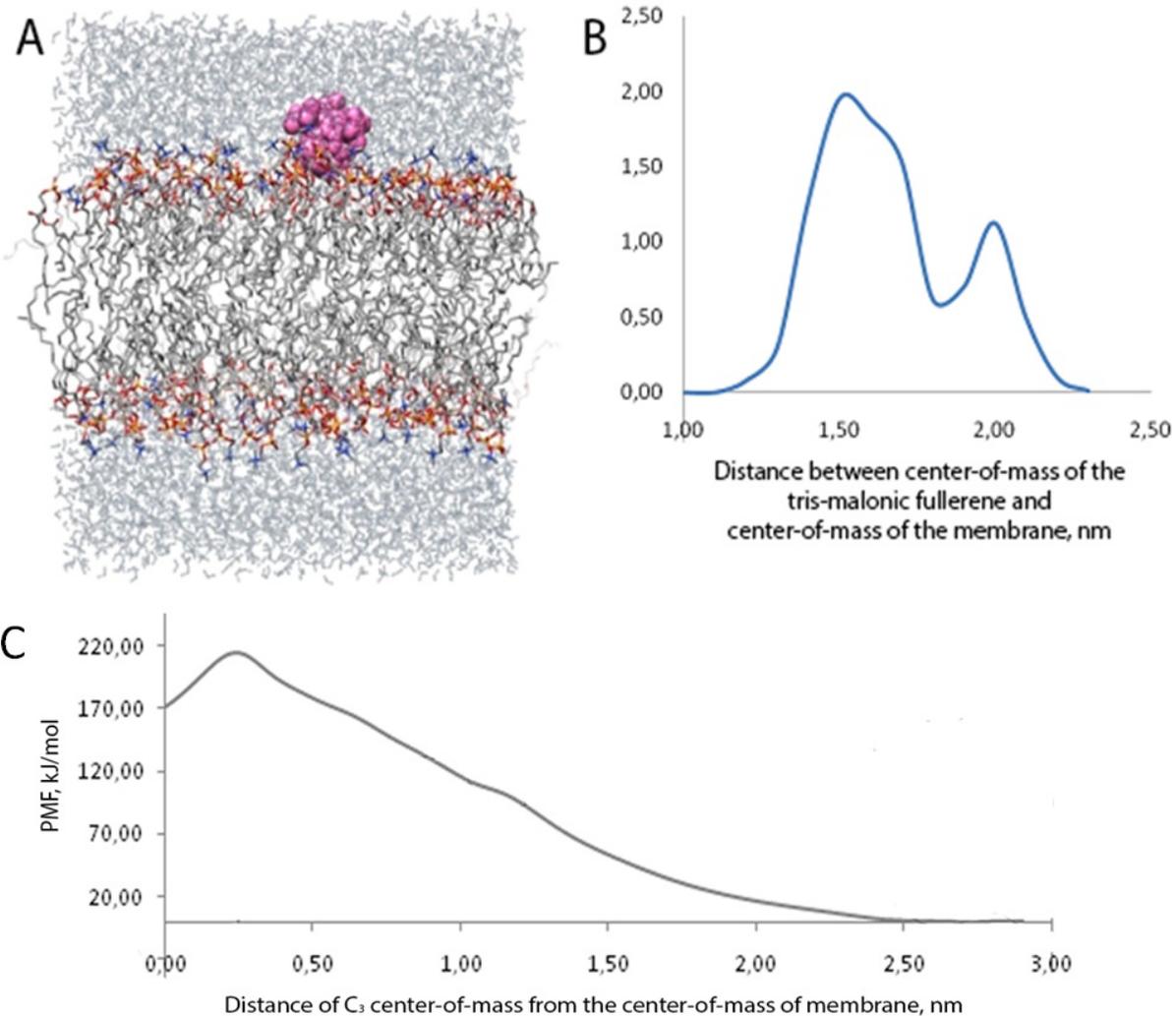
ns of simulation.

Figure 5

A. Tris-malonic fullerene on the surface of the membrane B. The probability distribution of

location of the tris-malonic fullerene at different distances from the center-of-mass of

membrane. C. Potential of mean force profile for tris-malonic fullerene

Comparative Computational Study of Interaction of C60-Fullerene and Tris-Malonyl-C60-Fullerene Isomers with Lipid Bilayer: Relation to Their Antioxidant Effect