Thermodynamic Equilibrium of Biological Macromolecules Under Mechanical Constraints

A Preprint

Ashkan Shekaari  
Department of Physics  
K. N. Toosi University of Technology  
Tehran, Iran  
shekaari@email.kntu.ac.ir

Mahmoud Jafari *  
Department of Physics  
K. N. Toosi University of Technology  
Tehran, Iran  
jafari@kntu.ac.ir

February 11, 2021

Abstract

Equilibrating proteins and other biomacromolecules is cardinal for molecular dynamics simulation of such biological systems in which they perform free dynamics without any externally-applied mechanical constraint, until thermodynamic equilibrium with the surrounding is attained. However, in some important cases, we have to equilibrate the system of interest in the constant presence of certain constraints, being referred to as constrained equilibration in the present work. A clear illustration of this type is a single amyloid β-strand or RNA, when the reaction coordinate is defined as the distance between the two ends of the strand and we are interested in carrying out replica-exchange umbrella sampling to map the associated free energy profile as the dependent quantity of interest. In such cases, each sample has to be equilibrated with the two ends fixed. Here, we introduced a simulation trick to perform this so-called constrained equilibration using steered molecular dynamics. We then applied this method to equilibrate a single, stretched β-strand of an amyloid beta dodecamer fibril with fixed ends. Examining the associated curves of the total energy and the force exerted on the practically-fixed SMD atom over the total timespan broadly supported the validity of this kind of equilibration.

Keywords  Constrained equilibration · Biomacromolecules · Replica-exchange umbrella sampling

1 Introduction

Examining the underlying free energy profile (FEP) is still of foremost importance in fully understanding a large number of chemical processes such as protein-ligand binding [1], or drug transportation across the cell membrane [2], which cannot be accurately predicted, for instance in the field of computer-aided rational drug design [3], without having the associated free energy behaviors [4]. A rich set of numerical tools for reliable determination of free energy changes based on the fundamental principles of statistical mechanics [5] has so far been developed and is now within reach. This so-called methodological set, in conjunction with recent advancements in computational power, has unprecedentedly enhanced the fields of free energy calculations, and therefore computational structural biology [6], widening their areas of application as well.

The theoretical framework for free energy calculations was established a long time ago and several approximations have so far been accordingly developed [7, 8, 9, 10, 11, 12]. However, central to accurate determination of free energy difference between initial and final states of a system is to sufficiently explore the configurational space of the initial state so that relevant, low-energy configurations of the final state could adequately be sampled. Conventional methods, particularly molecular dynamics (MD) [13] and Monte Carlo (MC) [14], are not indeed successful in this respect [15] based on the fact that the system of interest could be trapped in a few conformational states (i.e., local

*Corresponding author
minima) during the simulation and the resulting potential of mean force (PMF) would accordingly be dependent on the starting conformation of the system due to the so-called unrepresentative sampling.

To circumvent the problem, since the late 1960s, a number of sophisticated techniques and strategies have so far been proposed and advanced based on performing random walks using non-Boltzmann probability weight factors [16]. Replica-exchange umbrella sampling (REUS) [17] is a nearly recent of them, which combines replica-exchange (RE) [18] and umbrella sampling (US) [19] techniques, and therefore, has both merits of RE and US. The former belongs to the so-called generalized-ensemble algorithms [20], while the latter is indeed a particular application of the more general importance sampling [21] used in statistics. In contrast to RE, all replicas in REUS have the same temperature value and this is the potential energy that is exchanged.

Based on the fact that US is a quasi-equilibrium method—in contrast to non-equilibrium approaches such as steered molecular dynamics (SMD) [22], or targeted molecular dynamics (TMD) [23, 24]—each replica in REUS has to be simulated and therefore equilibrated in the canonical ensemble independently. The problem arises when the reaction coordinate is defined as the distance between the two ends of the system, such as a single RNA or amyloid beta-strand, and the associated PMF is to be calculated along that collective variable. In such a case, we may have several replicas, each of which has to be equilibrated independently in a way that the related end-to-end distance (or equivalently the two ends of the strand) must be kept fixed during the equilibration simulation. To overcome this problem, we carried out such a kind of constrained equilibration on a single, stretched beta-strand of an amyloid beta dodecamer fibril using the constant-velocity protocol of SMD in that a differential value of the order of \(10^{-20}\) \(\text{Å}/\text{fs}\) was set to the pulling velocity of the SMD atom, which, in turn, kept this atom practically fixed over the total timespan. Examining the time-dependent total energies and forces acting on the SMD atom for all the replicas eventually verified the accuracy of the constrained equilibration. The computational setup was also described in Sec. 2 in detail.

## 2 Computational details

The initial atomic positions of a single amyloid beta-strand was taken from the RCSB [25] PDB file of amyloid beta A\(\beta_{1-42}\) dodecamer fibril with the entry code 2MXU [26]. The simulations were carried out on Debian style [27] Linux [28] systems supported by MPICH (version 2-1.4) [29, 30] using NAMD (version 2.14b2) [31] computer software with the July 2018 update of CHARMM36 force fields [32]. The VMD program (version 1.9.4a9) [33] was also used for post-processing. The fibril was solvated in a box of water under periodic boundary conditions with a unit-cell padding of about 2 nm to decouple the periodic interactions. Minimization was then carried out for 50000 conjugate gradient steps (100 ps) followed by a 2-ns free-dynamics equilibration at 310 K and 1.01325 bar in order to reach the equilibrium conformation of the fibril. Langevin forces with a Langevin damping constant of 2.5\(\text{ps}\) along with the Langevin piston pressure control were applied for the NPT equilibration to keep fixed the temperature and pressure of the system. The integration timestep was also 1 fs.

After equilibrating the whole fibril in water, a single beta-strand was taken and extended by SMD simulations in vacuum in a way that the distance between the first atom (Glu\(^{11}\)\(C_n\); where \(C_n\) is carbon, 11 is the residue number, and Glu is the associated amino acid) and the last one (Ala\(^{42}\)\(C_n\)), defined as the reaction coordinate (Fig. 1), increased from 31.4 Å (the initial value) to about 42.4 Å according to the function \(31.4 + i\) \((i = 0, ..., 11)\). Therefore, 12 independent replicas numbered consecutively from 0 to 11 were generated. Each replica equilibrated for 1 ns at 310 K in vacuum during an SMD simulation in that Glu\(^{11}\)\(C_n\) and Ala\(^{42}\)\(C_n\) were respectively defined as the fixed and SMD atoms, and the latter was pulled along the reaction coordinate (+z; one-dimensional pulling) with an infinitesimal, constant velocity of \(v_p = 10^{-20}\) \(\text{Å}/\text{fs}\), which is practically zero over the total timespan. The force acting on the SMD atom was described as

\[
f(r, t) = -\nabla U(r, t) = -\nabla \left( \frac{k}{2} [v_p (r - r_0) \hat{n}]^2 \right),
\]

where \(U\) is the harmonic potential (spring), \(k\) is the spring constant, \(t\) is time, \(r\) and \(r_0\) are respectively the instantaneous and initial position vectors of the SMD atom, and \(\hat{n} = (0, 0, 1)\) is the normalized pulling vector (along \(+z\)). The harmonic biasing potential of the form \(U(\xi, i) = k(\xi - \xi_i)^2/2\) was added to the Hamiltonian of the system, where \(k = 2\text{ kcal/(mol.Å}^2\) corresponding to a thermal fluctuation of the SMD atom of \(\sqrt{k_B T/\xi^2} = 0.555\) Å at 310 K, with \(k_B\) the Boltzmann constant—\(\xi\) is the collective variable, and \(\xi_i = (31.4 + i)\) Å defines the center of the \(i\)th bias potential.
3 Results and discussion

Fig. 2 illustrates the initial and equilibrated conformations of each replica shown in cyan (at $t = 0$) and magenta (at $t = 1$ ns), respectively. As is seen, from replica #0 to the last one, the end-to-end extension increases 1 Å at a time, and each replica independently equilibrates at 310 K. Moreover, the right ends of the two strands for each replica exactly coincide, approving that the SMD atom, in practice, is fixed during the simulation, as expected. To verify whether the equilibrations have been achieved, the associated time-dependent total energy curves were accordingly provided, as shown in Fig. 3. All the curves take similar, initially-decreasing trends and then converge to values around 700 kcal/mol with long tails, demonstrating that the corresponding equilibrium states have been attained. It could also be observed that the larger the end-to-end distance is stretched, the higher the value to which the total energy converges. Further examining the energy–time diagrams revealed that the specific pattern of the total energy was uniquely determined by that of the electrostatic energy based on the observation that the other energy terms in the CHARMM potential energy function—including bond, angle, dihedral, improper dihedral, and van der Waals (vdW)—fluctuated about constant values over the total timespan. This could be inferred from Fig. 4(a) which depicts the time dependence of these various energy contributions associated to the replica #11. Fig. 4(b) also illustrates the time dependence of the instantaneous force exerted on the SMD atom during the constrained equilibration of the replica #11. The relatively-large fluctuations over the first half of the trajectory are mainly due to the fact that the strand is initially out of equilibrium, and therefore considerable forces act on the SMD atom over this period, in contrast to the second half over which equilibration is achieved and the fluctuating pattern then becomes narrower. That these fluctuations mostly take negative values over the first 0.5 ns further reveals that the force is exerted by the strand itself, not by the external pulling factor which is practically vanishing during the simulation according to Eq. 1 and $v_p = 10^{-20}$ Å/fs. The average value of the force over the second half of the trajectory is also clearly zero due to equilibration as well as the infinitesimal pulling velocity.

4 Conclusions

We proposed a simulation trick to carry out equilibration simulations on biological macromolecules, which are constantly under the action of externally-applied, mechanical constraints such as stretching or contraction. Such cases could frequently be encountered in replica-exchange umbrella sampling simulations particularly where the reaction coordinate is explicitly related to the mechanical deformation of the system and the potential of mean force (PMF) is the dependent quantity of interest. We accordingly applied our approach to equilibrate a single, stretched β-strand, taken from an amyloid beta dodecamer fibril, where the reaction coordinate was defined as the distance between the two ends of the strand, and was accordingly extended by 11 Å in a way that every 1 Å of extension generated the initial conformation of one replica, leaving 12 replicas as well. We then simulated each replica independently in the canonical ensemble at 310 K using constant-velocity SMD in a way that the SMD atom was being pulled with a differential velocity of magnitude $10^{-20}$ Å/fs along the reaction coordinate, making this atom to be practically fixed during simulation.
Figure 2: Conformational transition of each replica from the associated initial state at $t = 0$ (cyan) to the equilibrated conformation at $t = 1$ ns (magenta). The right ends of the two strands for each replica exactly coincide, showing that the SMD atom is practically fixed during the simulations.

As a result, the SMD simulation turned into an equilibrating process and each replica consequently equilibrated using free dynamics plus two mechanical constraints corresponding with the two fixed ends. The validity of this kind of constrained equilibration was strongly approved by examining the time-dependent variations of the total energy and the force exerted on the SMD atom associated to each replica.

References

[1] Woo H -J, Roux B 2005 PNAS 102 6825–6830.
[2] Yang N J, Hinner M J 2015 Methods Mol. Biol. 1266 29–53.
[3] Chao W -R, Yean D, Amin K, Green C, Jong L 2007 J. Med. Chem. 50 3412–3415.
[4] Chipot C, Pohorille A 2007 Free Energy Calculations: Theory and Applications in Chemistry and Biology Springer-Verlag: Berlin, Heidelberg.
[5] Pathria R K, Beale P D 2011 Statistical Mechanics Third ed. Butterworth-Heinemann: Oxford.
[6] Nussinov R, Tsai C J, Shehu A, Jang H 2019 Molecules 24 637.
[7] Kirkwood J G 1935 J. Chem. Phys. 3 300–313.
[8] Zwanzig R W 1954 J. Chem. Phys. 22 1420–1426.
Figure 3: The time-dependent total energy curves of the replicas during constrained equilibration—rendered in Gnuplot (version 5.2) [35]. The long-tailed converging patterns are evident, demonstrating that all the replicas are finally in equilibrium.

[9] Widom B 1963 *J. Chem. Phys.* 39 2808–2812.
[10] Valleau J P, Card D N 1972 *J. Chem. Phys.* 57 5457–5462.
[11] Bennett C H 1976 *J. Comput. Phys.* 22 245–268.
[12] McDonald I R, Singer K 1967 *J. Chem. Phys.* 47 4766–4772.
[13] Haile J M 1997 *Molecular Dynamics Simulation: Elementary Methods* First ed. Wiley-Interscience.
[14] Landau D P, Binder K 2000 *A Guide to Monte Carlo Methods in Statistical Physics* Cambridge University Press: Cambridge.
[15] Owicki J C, Scheraga H A 1977 *J. Am. Chem. Soc.* 99 7403–7412.
[16] Abreu C R A, Escobedo F A 2006 *J. Chem. Phys.* 124 054116.
[17] Sugita Y, Kitao A, Okamoto Y 2000 *J. Chem. Phys.* 113 6042.
[18] Hukushima K, Nemoto K 1996 *J. Phys. Soc. Japan* 65 1604.
[19] Torrie G M, Valleau J P 1977 *J. Comput. Phys.* 23 187–199
[20] Okamoto Y 2004 *J. Mol. Graph. Model.* 22 425–439.
[21] Liu J S 2008 *Monte Carlo Strategies in Scientific Computing* Springer Science & Business Media.
[22] Gao M, Wilmanns M, Schulten K 2002 *Biophys. J.* 83 3435–3445.
[23] Schlitter J, Engels M, Krüger P 1994 *J. Mol. Graphics* 12 84–89.
[24] Schlitter J, Engels M, Krüger P, Jacoby E, Wollmer A 1993 *Mol. Simul.* 10 291–308.
[25] https://www.rcsb.org/
Thermodynamic equilibrium of biological macromolecules under mechanical constraints

Figure 4: (a) The various energy contributions of the CHARMM potential energy function including bond, angle, dihedral, improper dihedral, electrostatic, van der Waals (vdW), and total energy, and (b) the associated instantaneous force exerted on the SMD atom, as functions of time over the total timespan of simulating the replica #11. The average force is clearly zero over the last 0.5 ns due to achieving thermodynamic equilibrium.

[26] Xiao Y, Ma B, McElheny D, Parthasarathy S, Long F, Hoshi M, Nussinov R, Ishii Y 2015 *Nat. Struct. Mol. Biol.* 22 499–505.

[27] https://www.debian.org.

[28] Torvalds L 1999 *Commun. ACM* 42 38–39.

[29] https://www.mpich.org.

[30] Gropp W, Lusk E, Skjellum A 1999 *Using MPI: Portable Parallel Programming with the Message-Passing Interface* MIT Press.

[31] Phillips J C, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel R D, Kale L, Schulten K 2005 *J. Comput. Chem.* 26 1781–1802.

[32] MacKerell Jr A D, Bashford D, Bellott M, Dunbrack Jr R L, Evanseck J D, Field M J, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau F T K, Mattos C, Michnick S, Ngo T, Nguyen D T, Prodhom B, Reiher W E, Roux B, Schlenkrich M, Smith J C, Stote R, Straub J, Watanabe M, Workiewicz-Kuczera J, Yin D, Karplus M 1998 *J. Phys. Chem. B* 102 3586–3616.

[33] Humphrey W, Dalke A, Schulten K 1996 *J. Mol. Graph.* 14 33–38.

[34] http://jedi.ks.uiuc.edu/~johns/raytracer/

[35] http://www.gnuplot.info