Simulating Botulinum Neurotoxin with Constant pH Molecular Dynamics in Generalized Born Implicit Solvent

Yongzhi Chen [a], Xin Chen [b], Yuefan Deng [a,*]

[a] *Department of Applied Mathematics and Statistics*

[b] *Department of Physics*

Stony Brook University

Abstract

A new method was proposed by Mongan et al for constant pH molecular dynamics simulation and was implemented in AMBER 8 package. Protonation states are modeled with different charge sets, and titrating residues are sampled from a Boltzmann distribution of protonation states generated by Monte Carlo sampling based on Generalized Born (GB) derived energies. However, when this approach was applied to a bio-toxin, Botulinum Neurotoxin Type A (BoNT/A) at pH 4.4, 4.7, 5.0, 6.8 and 7.2, the pK_a predictions yielded by the method were inconsistent with the experimental values. The systems being simulated were divergent. Furthermore, the system behaviors in a very weak acidic solution (pH6.8) and in a very weak basic solution (pH7.2) were significantly different from the neutral case (pH7.0). Hence, we speculate this method may require further study for modeling large biomolecule.

Key words: Constant pH molecular dynamics; Botulinum Neurotoxin; Generalized Born Method

1. Introduction

Botulinum neurotoxins (BoNTs) are among the most potent toxins to human beings. There are seven different serotypes (A-G) of BoNT [1]. They were believed causative agents of botulism, a potentially fatal condition of neuromuscular paralysis and researched by many groups [2–6]. Currently no effective chemical antidote is available against botulism principally because of the lack of knowledge of the molecular structures and the mechanism of toxicity [1].

In the whole three-step toxic action, translocation domain across membranes attracts great attention since there are interesting phenomena in channel forming and enormous size transport. An acidic environment is believed to induce conformational changes in the translocation domain. Unfortunately, there was no experimental approach available for the range pH 4.0 to 5.0 [3]. For this reason, the property of BoNT/A in this range is of greater interests and an efficient constant pH MD modeling technique is highly desirable.

2. Constant pH MD

It has been long and well known that protein structures are strongly dependent on solvent pH [7,8]. On the other hand, it is known that a structural change also affects the pK_a values of titratable residues. Therefore, there is a tight coupling between protein conformation and protonation state [8]. Modeling such correlation requires a reliable constant pH molecular dynamics method and such a method was indeed proposed and implemented in

* Corresponding author.

Email address: Yuefan.Deng@StonyBrook.edu (Yuefan Deng).

AMBER 8 package by Mongan et al [9]. This method considers the above coupling factor and is in contrast to the methods that set a constant protonation state for each titratable group. A user is allowed to directly specify the value of solvent pH as an input parameter. Before the constant pH MD could be executed, the user needs to prepare relative files and define the types of residues for titrating. AMBER 8 suggests titrating AS4 (Aspartate), GL4 (Glutamate) and HIP (Histidine) for the acidic case and titrating LYS (Lysine), TYR (Tyrosine) and HIP for the basic case [9]. This version is the first software package that implement constant pH molecular dynamic with a Generalized Born (GB) model.

The basic idea is to implement GB MD and periodically sampling protonation states with Monte Carlo method. At each scheduled Monte Carlo step, a titratable site and a new protonation state for that site are randomly chosen. The total transition energy ΔG is calculated and used as the Metropolis criterion [8].

The total transition energy is calculated according to the following formula [8]:

$$\Delta G = ln10 k_B T(pH - pK_{a,r}) + \Delta G_e - \Delta G_{e,r}$$

where k_B is the Boltzmann constant, T is temperature, pH is the specified solvent pH, $pK_{a,r}$ is the pK_a of the reference compound, ΔG_e is the electrostatic component of the transition energy for the titratable group, and $\Delta G_{e,r}$ is the electrostatic component of the transition energy for the reference compound [8]. If the transition is accepted, MD will continue with the titratable group in the new protonation state. Otherwise MD will continue without any protonation state change [8].

This approach was applied to BoNT/A at different pH values, especially pH 4.7, to seek the special conformational information that could not be obtained by experimental approach. Since there is very few constant pH MD applications for a biomolecule as large as BoNT/A, the detail of this system should be greatly helpful to the developers of model and package and other researchers.

3. Numerical Experiment

Our MD experiments were performed by using AMBER 8. The ff99 force field was employed. A refined GB model (igb=5) was used for solvation. Salt concentration was set at 0.2 M. The cutoff for nonbonded interactions and computation of effective Born radii was 30 Å. The time step was 2 fs. Sander [9] was used as MD engine and ptraj [9] was used to analyze the trajectories and calculate the RMSD values in each case.

4. Results

A valid constant pH MD simulation method should yield pK_a predictions consistent with experimental values. It can rapidly converge to those predictions and maintain a stable trajectory. Furthermore, the method should be computationally efficient [8].

4.1. *pK_a prediction*

A method for constant pH MD simulations is expected to reproduce accurate titration curves. But this method did not work well in BoNT/A system. Table 1 summarizes the pK_a prediction of all 12 HIP residues in BoNT/A. Where Comp.1 was calculated from all predictions and Comp.2 was calculated only from the predictions with absolute offset less than 2.0.

We chose HIP residue to evaluate since protonation states of Histidine change dominantly in this pH range. Results showed that the titration curves were quite far from experimental curves. For example, figure 1 depicted the best fit titration curve for HIP-542 (in the translocation domain) and HIP-868 (in the binding domain). The curves have pK_a of 6.81 and 7.16 respectively, but the theoretical value is 6.0.

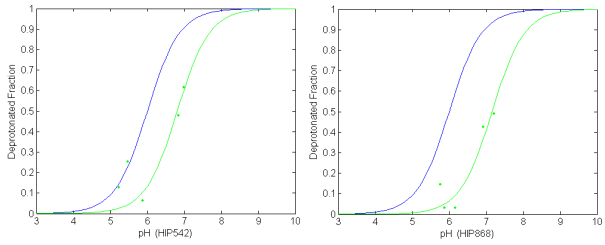

Fig. 1. Deprotonated fraction for HIP-542(left) and HIP-868(Right) from a 10-ns simulation. Green line represents best fit titration curve.

4.2. *Convergence*

The approach was firstly used to simulate BoNT/A at constant pH4.7 solvent. The system was gradually divergent. The RMSD was selected as

	pH4.4	pH4.7	pH5.0	pH6.8	pH7.2	Comp. 1	Comp. 2	Exp.
HIP-38	-0.387	4.476	4.389	4.870	4.879	-9.96	3.96	6.0
HIP-169	4.252	4.915	4.672	6.472	7.326	4.51	4.51	6.0
HIP-222	3.220	3.589	3.911	3.784	5.668	2.18	2.22	6.0
HIP-226	2.624	1.852	2.751	4.057	4.528	-0.50	0.75	6.0
HIP-229	3.027	2.393	3.321	4.238	6.241	0.61	1.65	6.0
HIP-268	4.852	5.515	5.243	6.389	6.030	5.67	5.67	6.0
HIP-533	3.348	-	4.125	5.512	-	2.40	2.40	6.0
HIP-542	5.240	5.875	5.470	6.840	6.997	6.81	6.81	6.0
HIP-868	5.877	6.176	5.773	6.929	7.219	7.16	7.16	6.0
HIP-1029	4.199	6.012	5.198	6.444	6.500	6.20	6.20	6.0
HIP-1045	5.742	5.992	5.819	6.835	7.394	7.26	7.26	6.0
HIP-1234	5.150	5.133	5.027	6.431	7.339	5.61	5.61	6.0

Table 1. pK_a Predictions for HIP residues

the major metric since the conformational change was of great interest in this system.

To find the exact location of BoNT/A which cause the divergence of RMSD, we did a full scan at the RMSD peak value of 14.34 Å at 12.69 ns to find the residues which cause the big RMSD for the translocation domain. In the scanning, we changed the size of the residue group we covered from one residue to 30 residues, and also move the starting point of these residue group in the full range of BoNT/A.

The results showed that there were two peak of RMSD, one was near the residue 440, another was near residue 830. Both are in the translocation domain. The translocation domain were greatly distorted and the whole protein tent to be torn apart and denatured.

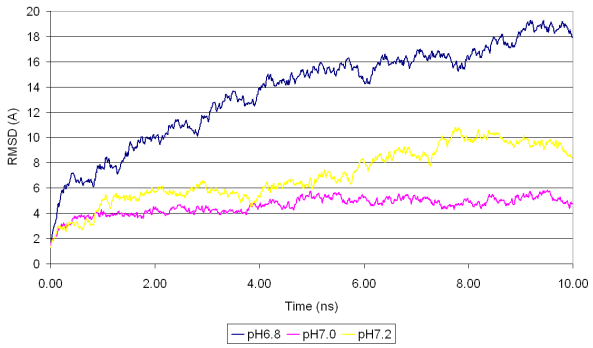

Fig. 2. Constant pH MD of BoNT/A at different pH.

To understand these results, the verification of the method on this system was necessary. The cases at other pH values could be helpful. If it was valid for this system, it supposed to work well in the whole pH range. The neutral cases was selected as the reference since its data was easily obtained by other method and could be easily and accurately crosschecked. Therefore a very weak acidic solution (pH6.8) case and a very weak basic solution (pH7.2) case were set up. The simulation results (depicted as Figure 2) showed that they were also gradually divergent and significantly different from the neutral case in which the system only evolved according to GB solvated MD.

5. Analysis and Conclusions

To explain these results, the following reasons are considered.

(1) Sufficiently conformational sampling is infeasible for the system of this size.

The instantaneous pK_a is strongly dependent on conformation, so it is reasonable that they would have different protonation state population if two simulations sample conformation space differently. The random error due to incomplete conformational sampling may produce larger effects than those caused by a small change in pH [8]. Of all 1277 residues in BoNT/A, the residues need to titrate are 165 in acidic case and 184 in basic case. Note that the default number of residue could be titrated in the package is 50. Even the parameter that determines the period for Monte Carlo steps was already specified as the most often, that is, sampling a titratable site in each step. It still could not meet the request of conformational sampling since there

are a lot of residues need sampling and titrating. Therefore the incomplete conformational sampling is the major reason that accounts for the inaccuracy of pK_a predictions and the system divergence.

(2) GB model is not accurate enough for a macromolecule.

This constant pH approach employed GB electrostatics [8]. The GB electrostatics is adopted because the potentials used for dynamics could be consistent with those used to choose protonation states in the Monte Carlo steps and the calculation of transition energies using GB is fast since no need for solvent equilibration [8]. These two problems must be seriously considered and solved if an explicit solvent model is adopted in constant pH MD. It is known that GB model may not work well in some cases, especially for a macromolecule. In fact, GB model was originally developed for small compounds where it was found to work quite well. However, its performance on larger molecules was worse than the expectations [10,11]. Note that Figure 2 showed the GB model itself could not account for the system divergence since the neutral system based on GB did not diverge in a 10-ns simulation. But the inaccuracy resulted from GB model still may work with other factors to result in the inaccuracy of pK_a predictions and the system divergence.

(3) This constant pH MD method may not work well in the whole pH range.

The system mentioned in [8] was only considered and verified in HEWL with the pH ranges from 2.0 to 6.5 and from 9.0 to 12.0. Its performance in pH range from 6.5 to 9.0 need the further validation. Besides the accuracy of model need more verifications of successful applications.

Therefore the current code may not work well for a macromolecular system. But the situation can be ameliorated if the method could be further improved on more efficient conformational sampling aspect such as incorporating multiple titratable site Monte Carlo moves into the protonation state sampling. Definitely this method will also get benefit from continuing improvement in GB models. Another suggestion is to find a way to implement constant pH MD in explicit solvent.

This computational expensive simulation was a pioneer job in this field. Although the project did not produce the expected conformational information, it should be helpful for developers to refine their model and code.

6. Acknowledgement

We thank Brookhaven National Laboratory(BNL) for their support and assistance.

References

[1] B.R. Singh, Intimate details of the most poisonous poison. Nat Struct Bio 7(8) (2000) Aug 617-619.

[2] H. Sugiyama, Clostridium botulinum neurotoxin. Micro-biol. Rev. 44 (1980) 419-448.

[3] D.B. Lacy, R.C. Stevens, Recombinant Expression and Purification of the Botulinum Neurotoxin Type A Translocation Domain. Protein expression and purification 11 (1997) 195-200.

[4] L. Li, B.R. Singh, Spectroscopic Analysis of pH-Induced Changes in the Molecular Features of Type A Botulinum Neurotoxin Light Chain. Biochemistry 39 (2000) 6466-6474.

[5] B.R. DasGupta, H. Sugiyama, A common subunit structure in Clostridium botulinum type A, B and E toxins. Biochem. Biophys. Res. Commun. 48 (1972) 108-112.

[6] L.L. Simpson, Kinetic studies on the interaction between botulinum toxin type A and the cholinergic neuromuscular junction. J. Pharmacol. Exp. Ther. 212 (1980) 16-21.

[7] R. Burgi, P.A. Kollman, W.F. Gunsteren, Simulating Proteins at Constant pH: An Approach Combining Molecular Dynamics and Monte Carlo Simulation. PROTEINS: Structure, Function, and Genetics. 47 (2002) 469-80.

[8] J.Mongan, D.A. Case, J.A. McCammon, Constant pH Molecular Dynamics in Generalized Born Implicit Solvent. J Comput Chem. 25 (2004) 2038-2048.

[9] D.A. Case et al, Amber 8 Users' Manual. 2004.

[10] A. Onufriev, D.A. Case, D. Bashford, Effective Born Radii in the Generalized Born Approximation: The Importance of Being Perfect. J Comput Chem. 23 (2002) 1297-1304.

[11] W.C. Still, A. Tempczyk, R.C. Hawley, T. Hendrickson, A General Treatment of Solvation for Molecular Mechanics. J. Am. Chem. Soc. 112 (1990) 6127.