Research Article
Open Access

Dynamics of the Large Progenitor Toxin Complex of *Clostridium botulinum*

Yosuke Kondo and Satoru Miyazaki*

Department of Medicinal and Life Science, Faculty of Pharmaceutical Sciences, Tokyo University of Science, Yamazaki, Noda-shi, Chiba, Japan

Abstract

Botulinum neurotoxins (BoNTs) are one of the potent toxins in nature but the toxicity is immediately eliminated by the harsh environment in the digestive tract because BoNTs are just a protein. However, BoNTs can get to synaptic junctions thanks to support of other proteins by aggregating with each other, combining with BoNTs and forming a large progenitor toxin complex (L-PTC). In order to explain how the complex formation enables the BoNTs to intrude into our body, we found that the three-dimensional structure of the L-PTC consists of an ovoid body with three legs and speculated important roles in the body and the legs. In the legs part, it is helpful for promoting absorption especially from the small intestine. Because experimental results showed that the legs are flexible and have specific binding sites of saccharides, the flexibilities may help the L-PTC to easily access the binding sites to the saccharides on the intestinal surface. However, such flexibilities have been only investigated by experimental methods. This means that we still have not objectively discussed what motions are generated from the shape of the L-PTC and how the structure is changed gradually. Therefore, we developed a new method integrating an analysis based on anisotropic network model and principal component analyses in order to measure the dynamics of the large-sized protein consisting of several subunits. The results showed that the L-PTC had characteristic motions which have large movements of the three legs. In addition, the flexible motions were appeared regardless of the theoretical models. Our application to measure the dynamics of the L- PTC suggested the importance of the flexibility which enables the L-PTC to break the epithelial barrier. We hope that the activity of the L-PTC is applied for developing a new oral drug delivery system.

Keywords: Botulinum neurotoxin; Three-dimensional structure; Anisotropic network model; Principal component analysis; Drug delivery system

Introduction

At the neuromuscular junction between the presynaptic terminal of a motor neuron and the postsynaptic membrane of a skeletal muscle, acethylcholine (ACh), 2-acetoxy-N,N,N-trimethylethanaminium, works as a neurotransmitter. The reason why AChs synthesized in the presynaptic cell and taken into a vesicle are released from the presynaptic cell is a membrane fusion, which is triggered by binding synaptosomal- associated protein of 25 kDa (SNAP-25) and vesicle-associated membrane protein (VAMP) on the vesicle surface with syntaxin on the inner cellular membrane [1]. As a result, ACh can bind to an ACh receptor on the postsynaptic membrane. However, if AChs are not released from the presynaptic cell, we cannot transmit signals from the motor neuron to the skeletal muscle. Botulinum neurotoxin (BoNT) inhibits releasing AChs from the presynaptic cell by cleaving SNAP-25, VAMP or syntaxin. Before this cleavage, it should be required to eventually get BoNTs to reach the presynaptic terminal.

One of the intruding roots of BoNTs is oral. In the digestive tract, however, BoNTs are inactivated by the acidic condition or enzymes. Consequently, if BoNTs are taken orally, we can inhibit BoNTs to get to the presynaptic terminal. However, several other proteins called non-toxin components, which are simultaneously produced with BoNTs by Clostridium botulinum, help BoNTs to delivery into the bloodstream. The non-toxic components, which consist of a non-toxic non-hemagglutinin (NTNHA) and three hemagglutinins (HA70, HA17 and HA33) [2], can form various stoichiometric complexes with BoNTs. Such complexes are called as progenitor toxin complexes (PTCs) and include the minimally functional PTC (M-PTC) and the large PTC (L-PTC) [3-5]. Recent studies showed that the stoichimetries of the M-PTC and the L-PTC are BoNT:NTNHA=1:1 [6] and BoNT:NTNHA:HA70:HA17:HA33=1:1:3:3:6 [7], respectively.

Some researchers suggested that the structures of the M-PTC and the L-PTC play important roles in the delivery. There are two effects. The NTNHA of the M-PTC is effective for protecting against inactivation of the BoNT by directly binding with the BoNT [6]. The HA parts of the L-PTC are effective for promoting absorption from the intestinal surface [7]. In the latter, for two reasons that the HA has specific binding sites of saccharides such as sialic acid, galactose, lactose, luctulose or isopropyl β- D-1-thiogalactopyranoside [8-11] and the HA parts are flexible [12], the flexibilities may help the L-PTC to easily access the binding sites to the saccharides on the intestinal surface. However, such flexibilities have not been investigated by computer analysis or simulation. This has caused two problems. The first one is that we cannot define what motions are generated from the structure of the L-PTC. The second one is that we cannot know how the structure of the L-PTC is changed gradually.

In this paper, we construct the three-dimensional structure of the L-PTC from crystal structures and then consider three theoretical models: anisotropic network model (ANM), coarse grained (CG) model and united atom model. Based on each model, the dynamics of the L-PTC are estimated by normal mode analysis (NMA) or molecular dynamics (MD) simulation. Finally, the results of the NMA and MD simulations are compared.

*Corresponding author: Satoru Miyazaki, Department of Medicinal and Life Science, Faculty of Pharmaceutical Sciences, Tokyo University of Science, 2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan, Tel: 81-471213630; E-mail: smiyazak@rs.noda.tus.ac.jp

Materials and Methods

Modeling of the L-PTC

From the Protein Data Bank (PDB) [13], we downloaded PDB ID:3V0A (BoNT-NTNHA) [6], PDB ID:4LO4 (a trimer of HA70s) [7] and PDB ID : 4LO7 (HA70-HA17-HA33) [7]. In order to model an HA complex, HA70:HA17:HA33=3:3:6, the C-terminal region of HA70 in 4LO7 was superimposed to each C-terminal region of HA70 in 4LO4 by the Align and Superimpose Proteins protocol of Discovery Studio 3.1 [14]. In order to model an L-PTC, we conducted two things: constructing N-loop of NTNHA in 3V0A by the Loop Refinement MODELER protocol and superimposing the HA complex and BoNT-NTNHA to L-PTC (EMDB ID:EMD2417) [7] by the gmconvert and the gmfit programs by setting 10 as a number of Gaussian distribution functions and 10,000 as a number of initial configurations by random generation [15] and then some structures of the L-PTC were modeled.

ANM analysis

We used the ProDy Python package for ANM analysis [16]. In one of the modeled structures, 6,535 Cα atoms were extracted and a cutoff and a γ parameter were set to 1.5 nm and 1.0, respectively.

MD simulations

We used the GROMACS software [17] for MD simulation. Although various force fields and parameters were tried, we eventually selected two computational conditions. In the first one, we specified the GRO-MOS53A6 force field [18]. Then, we constructed a rectangular box whose size is $\approx 28 \times 28 \times 24$ nm^3 and put $\approx 600,000$ TIP-3P water molecules [19] and 107 sodium ions into the box. We conducted energy minimizations of 10,000 steps by the steepest descent method and 10,000 steps by the conjugate gradient method. 0.1 ns MD simulations were conducted by increasing the reference temperature to 310 K by the velocity rescale method. The time step was set to 2 fs by using the LINCS algorithm for hydrogen atoms [20]. The switch method was used and pair list, long range and short range cutoffs were set to 1.2 nm, 1.1 nm and 1.0 nm, respectively. The Coulomb energy was computed by the particle mesh Ewald (PME) algorithm [21]. Next, 10 ns MD simulations were conducted with the Parrinello-Rahman method [22] and the reference pressure was set to 1.0 bar. Independent MD simulations were conducted 10 times and the MD trajectories were analyzed by Cartesian principal component analysis (PCA) [23] by extracting 100 snapshots from each independent MD simulation.

In the second one, we specified the MARTINI21 force field [24] with an elastic network which is 500 kJ mol^{-1} nm^{-2} of the spring constant, 0.5 and 0.9 nm of the lower and the upper cutoffs, respectively [25]. Then, we constructed a rectangular box whose size is $\approx 29 \times 28 \times 25$ nm^3 and put $\approx 160,000$ CG water molecules [24] into the box. We conducted energy minimizations of 10,000 steps by the steepest descent method and $\leq 10,000$ steps by the conjugate gradient method. 1.0 ns MD simulations were conducted by the Berendsen method [26] and the reference temperature was set to 310 K. The time step was set to 20 fs by using the LINCS algorithm [20]. The shift method was used and the pair list cutoff, the long or short range cutoff of VdW energy and the short range cutoff of Coulomb energy were set to 1.4 nm, 1.2 nm, 0.9 nm and 0.0 nm, respectively. Next, 1,000 ns MD simulations were conducted by the Berendsen method and the reference pressure was set to 1.0 bar. Independent MD simulations were conducted 10 times and the MD trajectories were analyzed by Cartesian PCA [23] by extracting 1,000 snapshots from each independent MD simulation.

In order to estimate similarities between the ANM analysis and/or PCAs [11], two eigenvectors were represented as A=(a$_1$, a$_2$, . . . , a$_3$n) and B=(b$_1$, b$_2$, . . . , b$_3$n) and cosine similarity S was computed as

$$S = \frac{A.B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{3n} a_i b_i}{\sqrt{\sum_{i=1}^{3n} a_i^2}\sqrt{\sum_{i=1}^{3n} b_i^2}} \qquad (1)$$

where n is a number of Cα atoms.

Visualization

Eigenvalues, projections and heat maps were visualized by the matplotlib Python package [27]. Three-dimensional structures were visualized by the VMD program [28].

Results

We conducted ANM analysis or PCAs of 10 or 1,000 ns MD simulations of the L-PTC. Figure 1A shows that 20 collective motions which involve to the whole structure of the L-PTC were obtained by the ANM analysis. Figure 1B shows that, in all eigenvalues, rates of eigenvalues of PCA axes 1 and 2 are 27.8% and 16.4%, respectively. Figure 1C shows that, in all eigenvalues, rates of eigenvalues of PCA axes 1 and 2 are 36.7% and 23.1%, respectively.
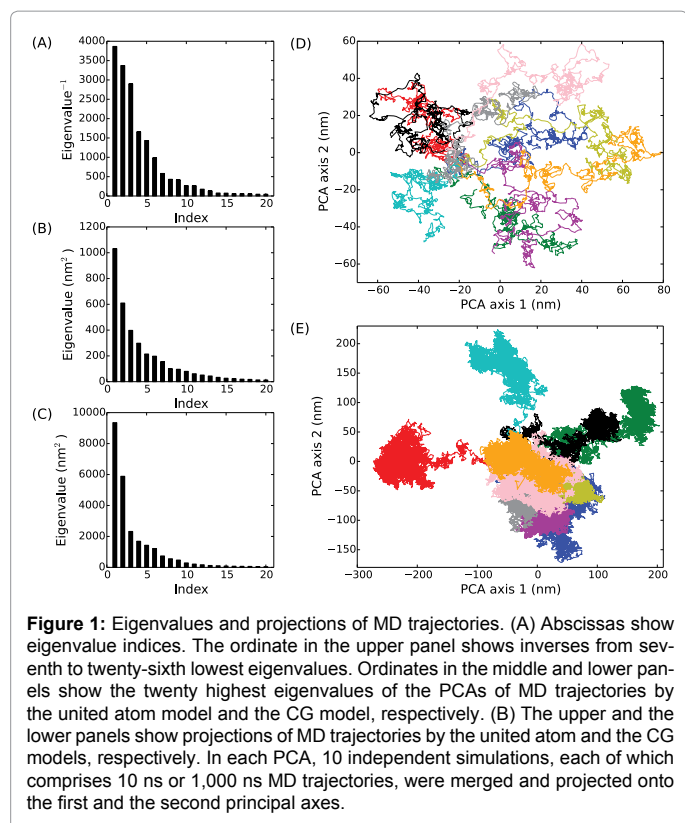
Figure 1D shows that the trajectories move a narrower space than the trajectories shown in Figure 1E. Figure 1E shows that some trajectories move well to the positive or the negative direction on PCA axis 1 or 2. For example, the red trajectories move to the negative direction on PCA axis 1, the green trajectories move to the positive direction on PCA axis 1 and the cyan trajectories move to the positive direction on PCA axis 2.

We computed cosine similarities from the ANM analysis and PCAs. Figure 2 shows that the highest value is 0.750 which is calculated from ANM mode 1 and PCA axis 2 by the CG model. The second highest value is 0.705, which is calculated from ANM mode 2 and PCA axis 1 by the CG model. The third highest value is 0.608, which is calculated from PCA axis 1 by the united atom model and PCA axis 1 by the CG model.

The followings are details of the movements of the L-PTC. Animation S1 (Supplementary material) shows that when the left leg moves up and the body moves left, the right leg moves down and the middle leg moves right and vice versa. Therefore, when the left leg moves down and the body moves right, the right leg moves up and the middle leg moves left. Animation S2 (Supplementary material) shows that the middle leg and the body move up and down and the left and the right legs move down and up, respectively. Animation S3 (Supplementary material) shows that the middle, the left and the right legs move up and down.

Discussion

Dynamics of the L-PTC were estimated by three computational models. The first one only considers Cα atoms and the forces between atoms are represented as a harmonic potential. The second one is the CG model whose amino acid residues are represented as 1-5 particles. The forces between particles are also represented as a harmonic potential. As other forces, bond, angle and dihedral angle and Coulomb and VdW interactions and water effects are considered. The third one is the united atom model whose carbon atom and hydrogen atoms connecting to the carbon atom are represented as 1 particle. As forces, bond, angle and dihedral angle and Coulomb and VdW interactions and water effects

Figure 1: Eigenvalues and projections of MD trajectories. (A) Abscissas show eigenvalue indices. The ordinate in the upper panel shows inverses from seventh to twenty-sixth lowest eigenvalues. Ordinates in the middle and lower panels show the twenty highest eigenvalues of the PCAs of MD trajectories by the united atom model and the CG model, respectively. (B) The upper and the lower panels show projections of MD trajectories by the united atom and the CG models, respectively. In each PCA, 10 independent simulations, each of which comprises 10 ns or 1,000 ns MD trajectories, were merged and projected onto the first and the second principal axes.

are also considered. On the basis of the computational models, the motion of the L-PTC was investigated. ANM analysis enables us to obtain sets of direction vectors of Cα atoms. The motion defined by the direction vectors is called as a collective motion and we can select the motion which involves all parts of the structure. MD simulation enables us to obtain trajectories of the atoms or the particles and PCA enables us to obtain direction vectors involving the large variance of the trajectories.

The ANM analysis and the PCAs of MD trajectories were compared by a square of the cosine similarity. The range of the square of the cosine similarity is from 0 to 1. The larger the value is, the more similar the vector direction is. In Figure 2, the highest value is 0.750 and the second highest value is 0.705. This means that the MD trajectories by the CG model include the motion resembling to ANM modes 1 and 2 as a large movement. As described above, the ANM analysis and the MD simulation by the CG model also consider harmonic potentials. Therefore, such a harmonic potential may be important for the motions which have flexible three legs. Additionally, the results also show that such motions are included even if the model considers waters or not. This implies that the shape of the L-PTC is important for the motions.
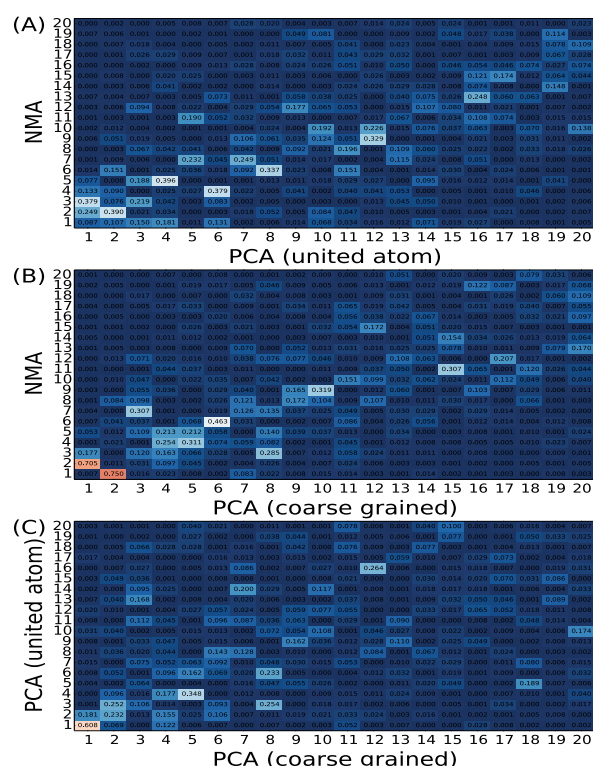
Meanwhile, the third highest value in Figure 2 is 0.608. This means that the MD trajectories by the CG model include the motion resembling to MD trajectories by the united atom model as a large movement. Therefore, the motion like Animation S3 (Supplementary material) shown by the red trajectories in Figure 1E is included in MD trajectories by the united atom model because the red trajectories move well to the negative direction on PCA axis 1. This indicates that the movements such as Animation S3 (Supplementary material) could be also reproduced by the model which does not consider the harmonic potential. Therefore, if the simulation time of the L-PTC by the

united atom model increased, we might be able to obtain the motion like Animation S3 (Supplementary material). For these reasons, such flexible motions could be obtained regardless of the three theoretical models.

Recent studies showed that the legs also interact with glycoprotein 2 [29] or E-cadherin [30,31]. This implies that for binding with these molecules, the flexibilities of the three legs are also important. Therefore, we should take account of such flexibility for understanding how the L-PTC facilitates translocation and disruption of the intestinal epithelial barrier [31,32].

Conclusions

In this study, a complex structure, which consists of 14 subunits, was investigated by computer analysis and simulation. The results showed that the L-PTC had flexible movements of the three legs and the flexible motions were included in different types of the theoretical models. This indicates that the structure which is an ovoid body with three legs is important for large movements of the three legs. Based on the flexibility, we suggest a story that the flexible structure enables the L-PTC to easily access to saccharides on the intestinal surface and then the L-PTC can stay in the surface for a long time and make an opportunity to be absorbed from the surface. Therefore, we supposed that the flexibility is effective for making an initial step to translocate the epithelial cell and should consider relationships between such flexibility and interactions with saccharides, glycoprotein 2 and E-cadherin. In addition, such flexibility is generated by formation of the quaternary structure. This implies that the complex formation is important for obtaining an ability which does not appear in each subunit only.

Figure 2: Similarities between the NMA and PCAs. 20 × 20 combinations of squares of cosine similarities were computed from (A) ANM analysis and PCA of the united atom model, (B) ANM analysis and PCA of the CG model and (C) PCA of the united atom model and PCA of the CG model.

References

1. Schiavo G, Matteoli M, Montecucco C (2000) Neurotoxins affecting neuroexocytosis. Physiol Rev 80: 717-766.

2. Dineen S, Bradshaw M, Johnson E (2003) Neurotoxin gene clusters in Clostridium botulinum type A strains: Sequence comparison and evolutionary implications. Curr Microbiol 46: 345-352.

3. Ohishi I, Sugii S, Sakaguchi G (1977) Oral toxicitees of Clostridium-botulinum toxins in response to molecular-size. Infect Immun 16: 107-109.

4. Ohishi I, Sakaguchi G (1980) Oral toxicities of Clostridium-botulinum type-C and type-D toxins of different molecular sizes. Infect Immun 28: 303-309.

5. Inoue K, Fujinaga Y, Watanabe T, Ohyama T, Takeshi K, et al. (1996) Molecular composition of Clostridium botulinum type A progenitor toxins. Infect Immun 64: 1589-1594.

6. Gu S, Rumpel S, Zhou J, Strotmeier J, Bigalke H, et al. (2012) Botulinum neurotoxin is shielded by NTNHA in an interlocked complex. Science 335: 977-981.

7. Lee K, Gu S, Jin L, Le TTN, Cheng LW, et al. (2013) Structure of a bimodular botulinum neuro- toxin complex provides insights into its oral toxicity. PLoS Pathogens 9: e1003690.

8. Nakamura T, Tonozuka T, Ide A, Yuzawa T, Oguma K, et al. (2008) Sugar-binding sites of the HA1 subcomponent of Clostridium botulinum type C progenitor toxin. J Mol Biol 376: 854-867.

9. Yao G, Lee K, Gu S, Lam KH, Jin R (2014) Botulinum neurotoxin A complex recognizes host carbohydrates through its hemagglutinin component. Toxins 6: 624-635.

10. Lee K, Lam WH, Kruel AM, Perry K, Rummel A, et al. (2014) High-resolution crystal structure of HA33 of botulinum neurotoxin type B progenitor toxin complex. Biochem Biophys Res Commun 446: 568-573.

11. Lee K, Lam KH, Kruel AM, Mahrhold S, Perry K, et al. (2015) Inhibiting oral intoxication of botulinum neurotoxin a complex by carbohydrate receptor mimics. Toxicon 107: 43-49.

12. Benefield DA, Dessain SK, Shine N, Ohi MD, Lacy DB (2013) Molecular assembly of botulinum neurotoxin progenitor complexes. Proc Natl Acad Sci USA 110: 5630-5635.

13. Berman H, Westbrook J, Feng Z, Gilliland G, Bhat T, et al. (2000) The protein data bank. Nucleic Acids Res 28: 235-242.

14. Accerlys Inc. (2011) Discovery Studio, Version 3.1. San Diego, CA, USA.

15. Kawabata T (2008) Multiple subunit fitting into a low-resolution density map of a macromolecular complex using a gaussian mixture model. Biophys J 95: 4643-4658.

16. Bakan A, Meireles LM, Bahar I (2011) ProDy: Protein dynamics inferred from theory and exper- iments. Bioinformatics 27: 1575-1577.

17. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, et al. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 29: 845-854.

18. Oostenbrink C, Villa A, Mark A, Van Gunsteren W (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem 25: 1656-1676.

19. Jorgensen W, Chandrasekhar J, Madura J, Impey R, Kkein M (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79: 926-935.

20. Hess B, Bekker H, Berendsen H, Fraaije J (1997) LINCS: A linear constraint solver for molecular simulations. J Comput Chem 18: 1463-1472.

21. Darden T, York D, Pedersen L (1993) Particle mesh Ewald - an N.log(N) method for Ewald sums in large systems. J Chem Phys 98: 10089-10092.

22. Parrinello M, Rahman A (1981) Polymorphic transitions in single-crystals - a new molecular- dynamics method. J Appl Phys 52: 7182-7190.

23. Amadei A, Linssen A, Berendsen H (1993) Essential dynamics of proteins. Proteins-structure function and genetics 17: 412-425.

24. Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, et al. (2008) The MARTINI coarse-grained force field: Extension to proteins. J Chem Theory Comput 4: 819-834.

25. Periole X, Cavalli M, Marrink SJ, Ceruso MA (2009) Combining an elastic network with a coarse- grained molecular force field: structure, dynamics, and intermolecular recognition. J Chem Theory Comput 5: 2531-2543.

26. Berendsen H, Postma J, Vangunsteren W, Dinola A, Haak J (1984) Molecular-dynamics with coupling to an external bath. J Chem Phys 81: 3684-3690.

27. Hunter JD (2007) Matplotlib: A 2D graphics environment. Comput Sci Eng 9: 90-95.

28. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. J Mol Graphics Modell 14: 33-38.

29. Matsumura T, Sugawara Y, Yutani M, Amatsu S, Yagita H, et al. (2015) Botulinum toxin A complex exploits intestinal M cells to enter the host and exert neurotoxicity. Nat Commun 6: 6255.

30. Sugawara Y, Yutani M, Amatsu S, Matsumura T, Fujinaga Y (2014) Functional dissection of the Clostridium botulinum Type B hemagglutinin complex: identification of the carbohydrate and E-Cadherin binding sites. PLoS ONE 9: e111170.

31. Lee K, Zhong X, Gu S, Kruel AM, Dorner MB, et al. (2014) Molecular basis for disruption of E-cadherin adhesion by botulinum neurotoxin A complex. Science 344: 1405-1410.

32. Lam TI, Stanker LH, Lee K, Jin R, Cheng LW (2015) Translocation of botulinum neurotoxin serotype A and associated proteins across the intestinal epithelia. Cell Microbiol 17: 1133-1143.