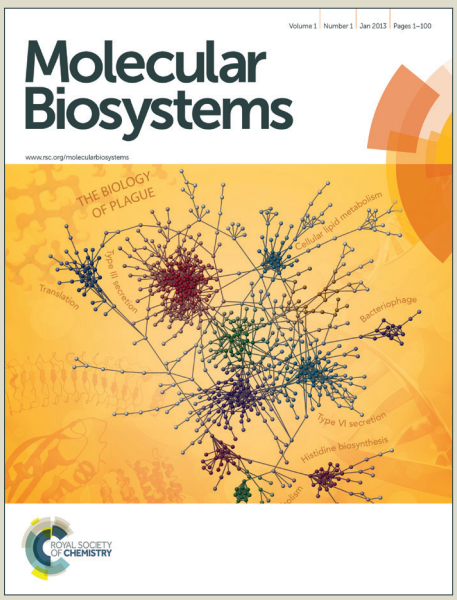
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

Journal Name

RSCPublishing

PAPER

Molecular dynamics simulations of wild type and mutants of botulinum neurotoxin A complexed with synaptic vesicle protein 2C

Cite this: DOI: 10.1039/xoxxoooooox

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/xoxxoooooox

www.rsc.org/

Feng Wang,[a] Hua Wan, [b] Jian-ping Hu [c] and Shan Chang [*b],

Botulinum neurotoxins (BoNTs) are known as the most poisonous biological substances, and they are also used to treat a wide range of medical conditions as well as the cosmetic applications. Recently, the complex structure of BoNT/A receptor-binding domain (BoNT/A-RBD) and the synaptic vesicle protein 2C luminal domain (SV2C-LD) was determined by X-ray crystallography. In this article, the wild type (WT) and four mutants of the new structure are studied by molecular dynamics (MD) simulations. The differently decreased structural stabilities of the mutants relative to WT are shown to be consistent with the experimental data of binding affinities. The conformational changes of the five systems are explored by using principal component analysis (PCA) and free energy landscape (FEL) methods. Based on the calculation of interactions at the binding interface, we divide the interface between BoNT/A-RBD and SV2C-LD into two crucial binding regions. Through the comparison of WT and four mutants, we further propose the relationship between the conformational changes of BoNT/A-RBD:SV2C-LD and the interfacial interactions. This study would provide some new insights into the understanding of the dynamics and interaction mechanism of BoNT/A-RBD:SV2C-LD.

Introduction

Botulinum neurotoxins (BoNTs) are proteins produced by the bacteria *Clostridium botulinum* and they have been classified into seven serotypes (BoNT/A-BoNT/G).[1] They are known as the most poisonous biological substances.[2] Despite their lethal toxicity, BoNTs have emerged as the valuable therapeutic tools for the treatment of a wide range of medical conditions, including migraines, spastic muscle disorders and pain management.[3-5] In the seven serotype of BoNTs, the BoNT/A is the first one utilized to treat human neurological diseases, and it is also widely used as an antiwrinkle agent in cosmetic applications.[6,7]

The X-ray structures of BoNTs promote our understanding for the properties and the mechanism of these toxins. The structures of BoNTs are composed of a light chain (LC, ~50kDa) and a heavy chain (HC, ~100kDa), which are linked by a disulfide bond.[8] The light chain (residues 1~448) is a Zn^{2+}-dependent metalloprotease. The heavy chain contains an N-terminal translocation domain (residues 449~872) and a C-terminal receptor-binding domain (residues 873~1295). In 1998, the crystal structure of the entire 1285 amino acid BoNT/A was determined at 3.3 Å resolution.[9] In 2007, Allen et al. reported the crystal structures of BoNT/A-LC in complexes with two potent small-molecule inhibitors, and explored the active site and conformational flexibility of BoNT/A-LC.[10] In 2008, Stevens et al. solved the crystal structure of the BoNT/A binding domain alone and in complex with the ganglioside GT1b at 1.7 Å and 1.6 Å, respectively.[11] These important structure data provided some new clues to the functional study of BoNTs.

Theoretical studies, such as the coarse-grained models[12-14] and the all-atom molecular dynamics (MD) simulation[15-17], have become important tools in exploring the dynamics and interaction mechanisms of proteins. These simulation methods also have been performed to investigate the interactions and conformational flexibility of BoNTs. Some groups have applied MD methods to explore the conformational flexibility of BoNT/A-LC and the inhibitor binding modes.[18,19] The entire BoNT/A structure also has been studied by the long-time MD simulations and the properties of BoNT/A at various temperatures and pH values were analyzed to understand the toxicity and structural variations.[20] The all-atom MD simulations have been used to investigate the conformational property for the oligosaccharide structure of GD1A and GD1B in aqueous environment.[21,22] In addition, both the urea denaturation experiments and theoretical modeling were performed to understand the folding mechanism of BoNT/A.[23]

In the previous studies,[24] the synaptic vesicle protein 2 (SV2) has been proposed to be a protein receptor for BoNT/A, but the structural details and theoretical studies of the interactions between BoNT/A and SV2 are relatively sparse. Recently, Kammerer et al. determined the high-resolution crystal structure of the BoNT/A receptor-binding domain (BoNT/A-RBD) in complex with the SV2C luminal domain (SV2C-LD).[25] This study provides a strong structural basis for the

interactions of BoNT/A and SV2. However, several questions still remain unclear. What specific interactions are formed at the interface between BoNT/A-RBD and SV2C-LD? How do the different mutants reduce the binding affinities? Do the conformational changes occur when the interface residues of BoNT/A-RBD:SV2C-LD are mutated? What is the relationship between the conformational changes and the interfacial interactions?

In this article, in order to probe the above issues, the new crystal structure of BoNT/A-RBD:SV2C-LD and four relevant mutant systems were analyzed by MD simulations. The principal component analysis (PCA) and free energy landscape (FEL) methods were applied to explore the functional dynamics and conformational changes of the five systems. We investigated the crucial interfacial interactions of BoNT/A-RBD:SV2C-LD and explained the loss of binding affinities of four mutants relative to the wild type. Finally, we further suggested the relationship between the conformational changes of BoNT/A-RBD:SV2C-LD and the interfacial interactions.

Systems and methods

Protein systems

The crystal structure of BoNT/A-RBD:SV2C-LD complex (PDB code: 4JRA)[25] was obtained from the Protein Data Bank. The unresolved residues (N1169, K1170, D1228, Q1229, G1230) in the BoNT/A-RBD were modeled based on the chain A in the BoNT/A-RBD:GT1b complex structure (PDB code: 2VU9).[11] In the BoNT/A-RBD:SV2C-LD complex, BoNT/A-RBD consists of the C-terminal and N-terminal subdomains. SV2C-LD forms a right-handed, quadrilateral β-helix. In Fig. 1, the important interface residues are highlighted in stick model. Besides the wild type, we also simulated four relevant mutants for comparison with the experimental data. The mutations of residues were performed with VMD1.9.[26] The first three mutations were R1294A, F563A and R1156E, respectively. In the fourth mutant system, both two BoNT/A-RBD threonines T1145 and T1146 were mutated to the alanines. For convenience, the five systems were referred to as WT, R1294A, F563A, R1156E and TT1145/6AA, respectively.

(Fig. 1)

Molecular dynamics simulations

Five independent molecular dynamics simulations were performed with the NAMD 2.9 program[27] by using the CHARMM 27 all-atom force field.[28] In each system, the protein was solvated using explicit TIP3P water models and the water molecules extended about 12 Å from the protein atoms in a cubic periodic box. Sodium and chloride ions were added to the systems to get a final ion concentration of 0.15 mol/L. Each system was energetically minimized with 20000 steps and then slowly heated up from 0 to 310 K over a period of 1.0 ns with a harmonic constraint of 0.1 kcal·mol^{-1}·Å^{-2} placed on all backbone atoms. Next, the non-constrained MD simulation was performed at constant pressure (1 atm) and constant temperature (310 K) for 42 ns. The Particle Mesh Ewald (PME)[29] summation algorithm was used to calculate the long-range electrostatic interactions. A cut-off of 10 Å was used for the construction of non-bonded list. The covalent bonds involving hydrogen atoms were constrained with the SHAKE algorithm.[30] For each system, the trajectory snapshots were saved every 2.0 ps, and then total 21000 conformations were collected for further analyses.

Principal component analysis and free energy landscape

Principal component analysis (PCA) is a widely used approach to extract the slow and functional motions of bio-molecules from the MD trajectories by applying the dimensional reduction method.[17,31] This analysis is based on the calculation and diagonalization of the covariance matrix. The elements C_{ij} in the matrix are defined as[32]:

$$C_{ij} = \frac{\left\langle \Delta r_i \cdot \Delta r_j \right\rangle}{(\left\langle \Delta r_i^2 \right\rangle \cdot \left\langle \Delta r_j^2 \right\rangle)^{1/2}} \quad (1)$$

where Δr_i (Δr_j) is the displacement vector corresponding to i th(j th) atom of the systems, and $\left\langle \cdots \right\rangle$ indicates an ensemble average. The eigenvectors of the matrix are also called the principal components (PCs), which represent the directions of the concerted motions. Usually, the first few PCs will describe the slow motion modes of the system, which are related to the functional motions of a bio-molecular system.[32,33] The eigenvalues of the matrix indicate the magnitude of the motions along the direction. In this article, PCA was performed with Gromacs 4.5 package[34] in order to investigate and compare the motion modes of the five systems.

Free energy landscape (FEL) can promote an understanding for the dynamic process occurring in a biological system.[33,35] In the FEL, the free energy minima usually represent the conformational ensemble in the stable states while the free energy barriers denote the transient states.[36] The FEL was constructed on the basis of the above PCA data.[32] The corresponding expression is:

$$\Delta G(X) = -K_B T \ln P(X) \quad (2)$$

where the reaction coordinate X is the PCs, K_B is the Boltzmann constant, T is the absolute temperature, and $P(X)$ is the probability distribution of system along the PCs. In this study, we calculated the FEL to identify the dominant conformational states of the five systems.

Analysis of the interactions

The interactions between BoNT/A-RBD and SV2C-LD are dominated by the hydrogen bonds and cation-π interaction.[25] In this article, the hydrogen bonds were calculated by VMD 1.9[26] with a distance cut-off value of 3.5 Å and an angle cut-off value of 35°.[37] The cation-π interaction was analyzed by using the energy-based and geometry-based methods.[38,39] The program CaPTURE (Cation-π Trends Using Realistic Electrostatics)[38] was used to identify cation-π interactions between the cationic group residues (i.e., lysine or arginine) and the aromatic rings residues (i.e., phenylalanine, tyrosine and tryptophan). This program provides an energetic evaluation of potential cation-π interactions using a variant of the optimized potentials for liquid simulations (OPLS) force field.[40] Besides the energy calculation, we also measured the distance and angle of the cation-π interaction. The distance (R) was calculated from the cation to the centroid of the aromatic ring. The angle (θ) was defined between the cation-centroid vector and the aromatic ring plane. These geometric parameters are shown in Fig.2.

(Fig. 2)

In order to analyze the relative binding energy between BoNT/A-RBD and SV2C-LD, the program g_mmpbsa[41] was

applied to the five systems. It implements the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) approach to estimate interaction free energies. This program can calculate molecular mechanics potential energy (include both electrostatic and van der Waals interactions) and free energy of solvation (include polar and nonpolar solvation energies). Although the entropy contribution was not included in the calculations, the binding energy calculated using the g_mmpbsa showed an apparent correlation of 0.80~0.85 with the experimental binding free energy in previous study.[41]

Results and discussion

Comparative analysis of the MD trajectories

The protein stability during MD simulation is monitored through the time evolution of the root mean square deviation (RMSD). Five 42 ns MD simulations were carried out for the WT, R1294A, F563A, R1156E and TT1145/6AA systems, respectively. Fig. 3A shows the RMSD values of backbone atoms for the five simulated structures. For the modeled structures, the systems usually need longer time (around 10 ns) to reach equilibrium states than that starting from the native one.[42,43] The modeled WT system remains stable after 12 ns, so we choose timescales 12~42 ns for further analyses. Fig. 3B displays the distributional probability of RMSD from 12~42 ns trajectories. The mean RMSD values of the five systems (WT, R1294A, F563A, R1156E and TT1145/6AA) are 2.1, 2.6, 3.1, 3.4 and 3.2 Å, respectively. The fluorescence anisotropy experiments[25] showed that the dissociation constant K_d for the five systems are 0.26±0.02, 1±0.2, 1.7±0.4, ≥7 and ≥25μM, respectively. The RMSD analysis shows that the structural stabilities in MD simulations are consistent with the experimental binding affinity data. Notably, the RMSD value of R1156E system is higher than that of TT1145/6AA. In the R1156E system, the positively charged residue R1156 is mutated to negatively charged Glutamic acid, whereas the opposite binding surface of SV2C-LD is slightly negatively charged. Thus, this electrostatic repulsion will increase the fluctuation of the structure remarkably.

(Fig. 3)

The flexibility of each residue is assessed by its root mean square fluctuation (RMSF). Fig. 3C and 3D shows the RMSF values of BoNT/A-RBD and SV2C-LD calculated from 12~42 ns trajectories, respectively. For the five systems, the middle part of the C-terminal BoNT/A-RB subdomain (residues K1170~I1250) and the N-terminal of SV2C-LD (residues V473~N480) exhibit large fluctuation values. In contrast, the binding site residues (see Fig. 1) of BoNT/A-RBD:SV2C-LD shows relatively lower fluctuation values. For the WT system, the RMSF values of BoNT/A-RBD and SV2C-LD are lower than those of the mutant systems. For the R1294A, R1156E and TT1145/6AA systems, the mutations lie in the BoNT/A-RBD, so these mutations mainly lead to the high fluctuation of BoNT/A-RBD. Similarly, the F563A mutation increases the fluctuation of SV2C-LD. Notably, the R1156E mutation also causes the RMSF increase of SV2C-LD. This result is consistent with the above RMSD analysis. The mutation of R1156E induces the electrostatic repulsion, which will affect the stability of both BoNT/A-RBD and SV2C-LD.

Motion modes and conformational changes of BoNT/A-RBD:SV2C-LD

In order to further inspect the direction of the fluctuation in the five systems, both PCA and FEL were performed for all Cα atoms of BoNT/A-RBD:SV2C-LD complex structure from the last 30 ns trajectory. The first and second motion modes (PC1 and PC2) of WT system are shown in Fig. 4A and 4B. Fig. 4C displays the corresponding free energy contour map for the WT system at 310 K with deeper color indicating the lower energy. Fig. 4D shows the representative structures in the energy basins compared with the crystal structure. As shown in Fig. 4A, the SV2C-LD have slightly swing motions in the first slowest motion mode. The C-terminal BoNT/A-RB subdomain (residues K1170~I1250) shows some irregular motions. Similarly, the N-terminal part and some loop region of BoNT/A-RBD also exhibit the irregular motions in the second slowest motion mode. Compared with other crystal structures of BoNT/A,[9,11] these irregular motion regions are interacted with the gangliosides and the translocation domain of BoNT/A. Thus, the irregular motions may contribute to adjust the local conformation of BoNT/A-RBD for binding. As shown in Fig. 4C, the free energy landscape has two local basins. The PC2 values of these two basins are around 0, which shows the magnitude of PC2 motion is not remarkable. In the PC1 direction, the two basins are around -1 and 1, respectively. Compared with Fig. 4A, it implies that the SV2C-LD structure will undergo the reciprocating swing in the PC1 direction. Similarly, the representative structures of the two basins in the WT system show remarkable conformational changes at the C-terminal BoNT/A-RB subdomain. In contrast, the binding interface between BoNT/A-RBD and SV2C-LD is relatively stable. Since the PC1 value of the second representative structure is -1, the SV2C-LD exhibits the slight swing in the negative direction of the first slowest motion mode.

(Fig. 4)

In the mutant systems, the motion modes are different from those of the WT system. As shown in Fig. S1A (ESI†), the N-terminal BoNT/A-RB subdomain in the R1294A system has larger movement than that of WT system, so the RMSF values of some parts in the N-terminal BoNT/A-RB subdomain are increased remarkably in Fig. 3C. In PC1 and PC2 modes, the SV2C-LD and the N-terminal BoNT/A-RB subdomain have the opposite motions. The free energy landscape of the R1294A system also has two local basins, but the magnitudes of the PC1 and PC2 motions are larger than those of the WT system.

As shown in Fig. S1B (ESI†), since the mutant residue F563 lies in the SV2C-LD, the average structure of SV2C-LD changed remarkably in the F563A system. In the PC1 mode, the C-terminal BoNT/A-RB subdomain has the opposite movement with SV2C-LD. The free energy landscape of the F563A system also has two local basins, but most of the conformations are located in the negative direction of PC1 and the positive direction of PC2.

In the R1156E system, the average structure has large changes at the binding site (see Fig. S1C, ESI†). The left side of SV2C-LD moves away from the BoNT/A-RBD. In the PC1 and PC2 modes, the SV2C-LD has the opposite movements against the C-terminal and the N-terminal subdomains of BoNT/A-RB, respectively. The free energy landscape of the R1156E system has some basins in the negative direction of PC1 mode, so SV2C-LD moves anticlockwise and the left side will be far away from BoNT/A-RBD.

Similar to R1156E, the average structure of TT1145/6AA also has large changes at the binding site, but the position is on the right side (see Fig. S1D, ESI†). In the PC1 mode, the BoNT/A-RBD rotates anticlockwise and SV2C-LD moves

clockwise. The free energy landscape of the TT1145/6AA system has one basin in the positive direction of PC1 mode, so the SV2C-LD moves away from BoNT/A-RBD on the right side.

(Fig. S1)

Similar to WT system, Fig. 5 shows the representative structures of mutant systems in different energy basins compared with the crystal structure. As shown in Fig. 5A, the two representative conformations of R1294A have large conformational changes in N-terminal BoNT/A-RB subdomain, but most of the interface residues are relatively stable. Compared to the crystal structure, the mutation site A1294 exhibits some fluctuations in the representative conformations. In the F563A system, besides BoNT/A-RB, the SV2C-LD in the representative conformations also changes remarkably (see Fig. 5B). The interface between BoNT/A-RBD and SV2C-LD becomes unstable, especially at the mutation site A563. In the R1156E system, the representative structure has large changes at the binding site (see Fig. 5C). The left side of SV2C-LD moves away from the BoNT/A-RBD. Similar to R1156E, the representative structure of TT1145/6AA also has conformational changes at the binding interface, but the position is on the right side (see Fig. 5D).

(Fig. 5)

In order to take a consistent view of the conformational changes of WT and mutants, the mutant trajectories are also projected on to the first and second motion modes of WT. As shown in Fig. S2 (ESI†), each mutant system has only one basin in this free energy landscape. The local basins of the R1294A and F563A systems located in the negative direction of WT's PC2 mode, but the PC1 values are relatively small (see Fig. S2A and S2B, ESI†). Therefore, although the conformational changes of R1294A and F563A are different to that of WT, the fluctuations of SV2C-LD in these two systems are not remarkable. The basin of the R1294A seems to be nearer to the local minimum of WT, so its average structure is more similar to the WT system. The local basin of the R1156E system located in the negative direction of WT's PC1 mode, so the SV2C-LD moves away from BoNT/A-RBD on the left side (see Fig. S2C, ESI†). In contrast, the local basin of TT1145/6AA located in the positive direction of PC1 mode (see Fig. S2D, ESI†). Then, the SV2C-LD rotates clockwise and the right side is far away from BoNT/A-RBD.

(Fig. S2)

From the above motion analysis, it is found that the SV2C-LD has swing motions in all the five systems. In the WT system, this motion is reciprocating swing and the magnitude is small, so the structure remains stable. In the R1294A and F563A, the mutation may break some interactions in the interface, and the motion magnitude increases. Fortunately, these systems also have two basins in the free energy landscape, so the average structures are similar to the WT system. In contrast, the R1156E and TT1145/6AA systems have only one basin or some basins with same motion directions in the free energy landscape. The mutations in these two systems may break the important interactions in the interface, so the SV2C-LD only swings in the fixed direction. Hence, the average structures and the representative structures are different from those of the WT system, and then the binding affinities are reduced remarkably. It can be seen that these conformational changes in the four mutant systems are correlated with the mutation positions and the interactions in the interface. Then, the detailed interactions in the interface will be analyzed below.

Interactions between BoNT/A-RBD and SV2C-LD

The previous study has implied that the interactions of BoNT/A-RBD:SV2C-LD complex are mainly the hydrogen bonds and the cation-π interaction.[25] The cation-π interaction is between R1156 of BoNT/A-RBD and F563 of SV2C-LD. These interactions were calculated and shown in Tables 1~2, Table S1 (ESI†) and Fig. 6. Table 1 shows the hydrogen bonds with occupancy over 10% in the WT system. These hydrogen bonds mainly concentrate on residues S1142~T1146 in BoNT/A-RBD and residues E556~F562 in SV2C-LD. The S1142 and M1144 mainly form the backbone hydrogen bonds with F562 and C560. As shown in Fig. 6A, these backbone hydrogen bonds locate at the middle part of interface. In contrast, the T1145 and T1146 form the side-chain hydrogen bonds with E556, F557 and N559. These side-chain hydrogen bonds lie in the right side of interface. In addition, K951 and R1294 form the low occupancy hydrogen bonds with N559 and D539, respectively (see Fig. 6B). The R1294-D539 hydrogen bond locates at the left side of interface and is close to the cation-π interaction.

(Table 1)

(Fig. 6)

Table S1 (ESI†) shows the hydrogen bonds with occupancy over 10% in the four mutant systems. Interestingly, the mutant systems possess more hydrogen bonds compared to WT. The R1294A system only loses some low occupancy hydrogen bonds, and most of the high occupancy hydrogen bonds are maintained in the interface. The F563A and R1156E mutations mainly break the cation-π interaction, and have little influence on the high occupancy hydrogen bonds formed by S1142~T1146. Notably, the residue R1294 forms much more hydrogen bonds in F563A and R1156E systems. It implies that after the cation-π interaction is broken, the R1294 contributes much more interactions to stabilize the systems. The TT1145/6AA mutation has remarkable influence on the high occupancy hydrogen bonds. In this system, the side-chain hydrogen bonds of T1145 and T1146 are completely broken. Moreover, the residue M1144 lose a backbone hydrogen bond with C560, whereas R1156 forms many side-chain hydrogen bonds with N565 and D543.

(Table S1)

Another important interaction in the interface is the cation-π interaction (see Table 2). The cation-π interaction between BoNT/A-RBD and SV2C-LD is mainly formed by residues R1156 and F563, and located at the left side of the interface. The F563A and R1156E mutations mainly break the cation-π interaction. In the F563A system, the cation-π interaction is completely lost in the interface. Nevertheless, in the R1156E system, R1294 forms the cation-π interactions with F563 or Y478, but both the energy and occupancy are very low. The WT, R1294A and TT1145/6AA systems still maintain the cation-π interactions. Although the occupancies are more than 96% in the three systems, the binding energies are definitely different. The WT and R1294A systems possess the similar binding energies, where E_{elec} < -4.0 kcal/mol and E_{vdW} < -3.0 kcal/mol. It implies that the cation-π interactions are strong in these two systems. In the R1294A system, the binding energy is even lower than that of the WT system. The position of R1294 is close to the cation-π interaction, so the cation-π interaction will also help to stabilize the system after the hydrogen bond of R1294 lost. Compared with other systems, the cation-π interaction has much higher occupancy in the TT1145/6AA system, but the energy is not favourable. As shown in Fig. 7, the angle in the TT1145/6AA system decreases to around 55°,

and the distance also increases to more than 4.5 Å. The motion modes in Fig. S1D (ESI†) can give the explanations for these changes. The SV2C-LD moves away from BoNT/A-RBD on the right side, which causes the changes of angle and distance. In addition, R1156 forms some hydrogen bonds with N565 and D543 in the TT1145/6AA system (see Table S1, ESI†), so these wrong hydrogen bonds may also attract the side-chain of R1156 and then weak the cation-π interaction.

(Table 2)

(Fig. 7)

The binding energies of the five systems are calculated by g_mmpbsa and shown in Table 3. The van der Waals energies of WT, R1294A and F563A are lower than those of R1156E and TT1145/6AA. In the R1156E and TT1145/6AA systems, the SV2C-LD moves away from BoNT/A-RBD at the interface, so the van der Waals energies between the two proteins are increased. In the F563A and TT1145/6AA systems, the electrostatic energies are much lower than other systems. These two systems formed much more new hydrogen bonds in the interface, so electrostatic energies become more favourable. Notably, the R1156E structure also formed the new hydrogen bonds, but the electrostatic energy is higher than other systems. In the R1156E system, the positively charged residue R1156 is mutated to negatively charged Glutamic acid. Then, this mutation will cause the electrostatic repulsion in the interface, so the electrostatic energy is increased in R1156E system. Because the polar residues in surface forms much more new hydrogen bonds in the F563A, R1156E and TT1145/6AA systems, the polar interactions between protein and waters may be decreased. Thus, the polar solvation energies are increased in these three systems. The values of nonpolar solvation energies seem to be similar, except the TT1145/6AA system. The movement of SV2C-LD in TT1145/6AA breaks the interfacial interactions and decreases the interface area between BoNT/A-RBD and SV2C-LD. Therefore, the van der Waals energy and nonpolar solvation energy are increased in the TT1145/6AA system. Although the total binding energy of WT is very low, it is not the lowest one in the five systems. The entropy contribution was not included in the g_mmpbsa,[41] so this binding energy will have some deviations compared with the real binding free energy.

(Table 3)

From the above analyses, we can divide the interface between BoNT/A-RBD and SV2C-LD into two binding regions. One is the hydrogen bonds between S1142~T1146 in BoNT/A-RBD and residues E556~F562 in SV2C-LD, which locates at the middle and right sides of the interface. The other one consists of the cation-π interaction and the R1294-D539 hydrogen bond, which lies in the left side of the interface. The TT1145/6AA mutation mainly breaks the first binding region. The high occupancy hydrogen bonds are lost on the middle and right sides, so the SV2C-LD moves away from BoNT/A-RBD on the right side. This movement also changes the angle and the distance of cation-π interaction. Thus, this mutation reduces the binding affinity remarkably. The R1156E mutation breaks the second binding region. The correct cation-π interaction is lost, but R1294 make some wrong cation-π interactions with F563 or Y478. These wrong interactions also affect the correct hydrogen bond between R1294 and D539. However, the hydrogen bonds in the first region are kept, so SV2C-LD moves away from BoNT/A-RBD on the left side. The F563A mutation also breaks the cation-π interaction. However, the hydrogen bonds between R1294 and D539 increase the occupancy significantly, which helps to stabilize the left side of the interface. In the R1294A system, the mutation only breaks the low occupancy hydrogen bond between R1294 and D539. Meanwhile, the energy of cation-π interaction decrease to even lower than that of the WT system, which also helps to stabilize the left side of the interface. Therefore, the F563A and R1294A systems have the relatively lower binding energies and remain some binding affinities.

Conclusions

In this study, MD simulation was utilized to investigate the crucial interactions and dynamics of the new crystal structure of BoNT/A-RBD:SV2C-LD complex. The stability comparison of the five systems (WT, R1294A, F563A, R1156E and TT1145/6AA) is consistent with the experimental data. The conformational changes of the five systems are further revealed by PCA and FEL methods. It is found that the SV2C-LD has swing motions in all the five systems. If the mutations break the important interactions in the interface, the SV2C-LD will only swing in the fixed direction and the systems will have large conformational changes. Thus, we further calculated the interactions at the interface between BoNT/A-RBD and SV2C-LD. The interface is proposed to be divided into two binding regions. The first region is the high occupancy hydrogen bonds between S1142~T1146 in BoNT/A-RBD and residues E556~F562 in SV2C-LD. The second one consists of the strong cation-π interaction and the low occupancy hydrogen bond of R1294-D539. If the mutations completely break one region in the interface, the SV2C-LD will move far away from BoNT/A-RBD at that region and the binding affinities decrease remarkably (such as TT1145/6AA and R1156E). In contrast, if the mutations only breaks partial interactions and simultaneously other critical interactions can help to stabilize the two regions in the interface, the systems will also remain the binding affinities to some extent (such as F563A and R1294A). This study provides the dynamics and atomic level information for the BoNTs/A:SV2C interactions, which will be helpful to the rational design of BoNT/A variants and the development of novel antitoxin agents.

Acknowledgements

This work was funded by the National Natural Science Foundation of China (31200990, 11247018, 31372116), the Foundation for Outstanding Young Teachers in Higher Education of Guangdong (Yq2013027), the Science and Technology Bureau Foundation of Sichuan Province (2011JYZ007), the Key Project of Education Department of Sichuan Province (12ZA066) and the National Supercomputer Center in Guangzhou.

Notes and references

[a] School of information science & engineering, Changzhou University, Changzhou, China.

[b] College of Informatics, South China Agricultural University, Guangzhou, China. E-mail: schang@scau.edu.cn

[c] Faculty of Biotechnology industry, Chengdu University, Chengdu, China

† Electronic Supplementary Information (ESI) available: Fig. S1-S2 and Table S1. See DOI: 10.1039/b000000x/

1 G. Schiavo, M. Matteoli and C. Montecucco, *Physiol. Rev.*, 2000, **80**, 717-766.
2 S. S. Arnon, R. Schechter, T. V. Inglesby, D. A. Henderson, J. G. Bartlett, M. S. Ascher, E. Eitzen, A. D. Fine, J. Hauer, M. Layton, S. Lillibridge, M. T. Osterholm, T. O'Toole, G. Parker, T. M. Perl, P. K. Russell, D. L. Swerdlow and K. Tonat, *JAMA-J. Am. Med. Assoc.*, 2001, **285**, 1059-1070.
3 R. Hackett and P. C. Kam, *Med. Chem.*, 2007, **3**, 333-345.
4 R. Bhidayasiri and D. D. Truong, *J. Neurol. Sci.*, 2005, **235**, 1-9.
5 A. Wheeler and H. S. Smith, *Toxicology*, 2013, **306**, 124-146.
6 A. C. Markey, *Clin. Exp. Dermatol.*, 2000, **25**, 173-175.
7 C. Colasante, O. Rossetto, L. Morbiato, M. Pirazzini, J. Molgo and C. Montecucco, *Mol. Neurobiol.*, 2013, **48**, 120-127.
8 C. Montecucco and G. Schiavo, Q. Rev. Biophys., 1995, 28, 423-472.
9 D. B. Lacy, W. Tepp, A. C. Cohen, B. R. DasGupta and R. C. Stevens, *Nat. Struct. Mol. Biol.*, 1998, **5**, 898-902.
10 N. R. Silvaggi, G. E. Boldt, M. S. Hixon, J. P. Kennedy, S. Tzipori, K. D. Janda and K. N. Allen, *Chem. Biol.*, 2007, **14**, 533-542.
11 P. Stenmark, J. Dupuy, A. Imamura, M. Kiso and R. C. Stevens, *PLoS Pathog.*, 2008, **4**, e1000129.
12 I. Bahar, T. R. Lezon, L. W. Yang and E. Eyal, *Annu. Rev. Biophys.*, 2010, **39**, 23-42.
13 M. G. Saunders and G. A. Voth, *Annu. Rev. Biophys.*, 2013, **42**, 73-93.
14 S. Chang, X. Jiao, J. P. Hu, Y. Chen and X. H. Tian, *Int. J. Mol. Sci.*, 2010, **11**, 4014-4034.
15 J. Weng and W. Wang, *Adv. Exp. Med. Biol.*, 2014, **805**, 305-329.
16 P. Z. Jiao, W. W. Xue, Y. L. Shen, N. Z. Jin and H. X. Liu, *Mol. Biosyst.*, 2014, **10**, 767-777.
17 H. Wan, J. P. Hu, X. H. Tian and S. Chang, *Phys. Chem. Chem. Phys.*, 2013, **15**, 1241-1251.
18 J. C. Burnett, J. J. Schmidt, C. F. McGrath, T. L. Nguyen, A. R. Hermone, R. G. Panchal, J. L. Vennerstrom, K. Kodukula, D. W. Zaharevitz, R. Gussio and S. Bavari, *Bioorg. Med. Chem.*, 2005, **13**, 333-341.
19 P. Šilhár, M. A. Lardy, M. S. Hixon, C. B. Shoemaker, J. T. Barbieri, A. K. Struss, J. M. Lively, S. Javor and K. D. Janda, *ACS Med. Chem. Lett.*, 2013, **4**, 283-287.
20 X. Chen and Y. Deng, *J. Mol. Model.*, 2007, **13**, 559-572.
21 S. Venkateshwari and K. Veluraja, *J. Struct. Biol.*, 2012, **180**, 497-508.
22 S. Venkateshwari and K. Veluraja, *J. Biomol. Struct. Dyn.*, 2012, **30**, 255-268.
23 R. Kumar, R. V. Kukreja, L. Li, A. Zhmurov, O. Kononova, S. Cai, S. A. Ahmed, V. Barsegov and B. R. Singh, *J. Biomol. Struct. Dyn.*, 2014, **32**, 804-815.
24 M. Dong, F. Yeh, W. H. Tepp, C. Dean, E. A. Johnson, R. Janz and E. R. Chapman, *Science*, 2006, **312**, 592-596.
25 R. M. Benoit, D. Frey, M. Hilbert, J. T. Kevenaar, M. M. Wieser, C. U. Stirnimann, D. McMillan, T. Ceska, F. Lebon, R. Jaussi, M. O. Steinmetz, G. F. X. Schertler, C. C. Hoogenraad, G. Capitani and R. A. Kammerer, *Nature*, 2014, **505**, 108-111.
26 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33-38, 27-38.
27 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781-1802.
28 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. MacKerell, *J. Comput. Chem.*, 2010, **31**, 671-690.
29 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
30 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
31 S. Chang, J. P. Hu, P. Y. Lin, X. Jiao and X. H. Tian, *Mol. Biosyst.*, 2010, **6**, 2430-2438.
32 G. G. Maisuradze, A. Liwo and H. A. Scheraga, *J. Chem. Theory Comput.*, 2010, **6**, 583-595.
33 H. Wan, J. P. Hu, K. S. Li, X. H. Tian and S. Chang, *PLoS One*, 2013, **8**, e76045.
34 D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701-1718.
35 Y. Yang, H. X. Liu and X. J. Yao, Mol. Biosyst., 2012, 8, 2106-2118.
36 J. P. Hu, H. Q. He, X. Jiao and S. Chang, *Mol. Simul.*, 2013, **39**, 828-836.
37 M. M. Babu, S. K. Singh and P. Balaram, *J. Mol. Biol.*, 2002, **322**, 871-880.
38 J. P. Gallivan and D. A. Dougherty, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 9459-9464.
39 B. P. Dimitrijevic, S. Z. Borozan and S. D. Stojanovic, *RSC Adv.*, 2012, **2**, 12963-12972.
40 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225-11236.
41 R. Kumari, R. Kumar, A. Lynn and C. Open Source Drug Discovery, *J. Chem Inf. Model.*, 2014, **54**, 1951-1962.
42 A. Grossfield, *Biochim. Biophys. Acta-Biomembr.*, 2011, **1808**, 1868-1878.
43 H. Fan and A. E. Mark, *Protein Sci.*, 2004, **13**, 211-220.

Table 1 Hydrogen bonds between BoNT/A-RBD and SV2C-LD in WT system

WT		
BONT/A-RBD	SV2C-LD	Occupancy
M1144-N	C560-O	93.0%
S1142-O	F562-N	90.8%
M1144-O	C560-N	79.2%
T1146-OG1	F557-N	78.8%
S1142-N	F562-O	71.6%
T1146-N	F557-O	54.8%
T1145-OG1	N559-N	47.2%
T1146-OG1	E556-OE2	29.0%
T1146-OG1	E556-OE1	23.2%
K951-NZ	N559-OD1	19.1%
R1294-NH1	D539-O	16.4%

Table 2 Cation-π interactions between BoNT/A-RBD and SV2C-LD in WT and mutate systems

TYPE	BONT/A-RBD	SV2C-LD	E_{elec} [a]	E_{vdW}	Occupancy
WT	R1156	F563	-4.4±1.2	-3.2±0.5	96.3%
R1294A	R1156	F563	-4.7±1.2	-3.2±0.6	96.6%
F563A	--	--	--	--	--
R1156E	R1294	Y478	-2.0±0.6	-1.5±0.8	0.1%
	R1294	F563	-2.2±0.9	-1.6±0.5	3.0%
TT1145/6AA	R1156	F563	-2.9±0.8	-2.3±0.4	98.8%

[a] E_{elec} and E_{vdW} are the electrostatic and van der Waals energies, respectively. The unit of energy is kcal/mol.

Table 3 Comparison of binding energy components between BoNT/A-RBD and SV2C-LD in WT and mutate systems

TYPE	ΔE_{vdW} [a]	ΔE_{elec}	ΔG_{polar}	$\Delta G_{nonpolar}$	$\Delta G_{binding}$
WT	-60.4±3.4	-234.1±6.3	100.3±8.9	-7.1±0.3	-201.4±8.3
R1294A	-59.4±5.6	-232.7±9.8	100.7±9.4	-6.7±0.4	-197.6±9.2
F563A	-61.8±6.3	-297.0±9.3	156.6±10.0	-7.4±0.5	-209.5±9.8
R1156E	-47.0±6.1	-214.2±8.1	133.6±11.1	-6.8±0.5	-134.7±8.5
TT1145/6AA	-34.9±3.4	-292.9±9.7	188.3±9.9	-5.8±0.3	-145.2±9.2

[a] ΔE_{vdW}, ΔE_{elec}, ΔG_{polar}, $\Delta G_{nonpolar}$ are binding energy components of van der Waals, electrostatic, polar and nonpolar solvation energies, respectively. $\Delta G_{binding}$ is the total binding energy. The unit of energy is kcal/mol.

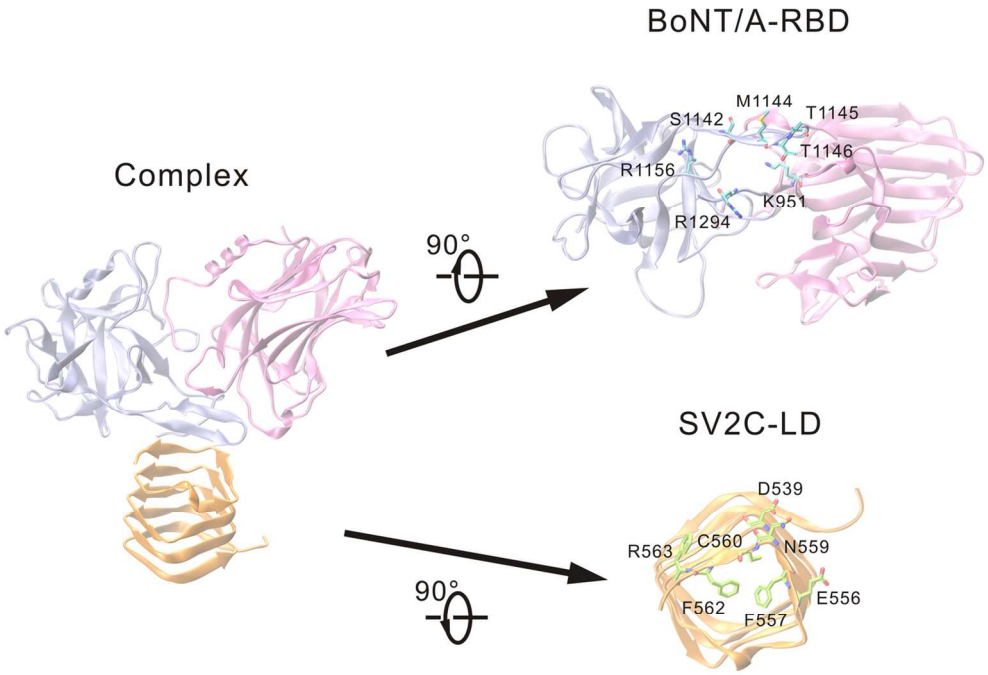


Fig. 1 Cartoon and open book representation of the BoNT/A-RBD: SV2C-LD complex. BoNT/A-RBD consists of the C-terminal (ice blue) and N-terminal (mauve) subdomains. SV2C-LD forms a right-handed, quadrilateral β-helix (orange). The important interface residues are highlighted in stick model.
127x90mm (300 x 300 DPI)

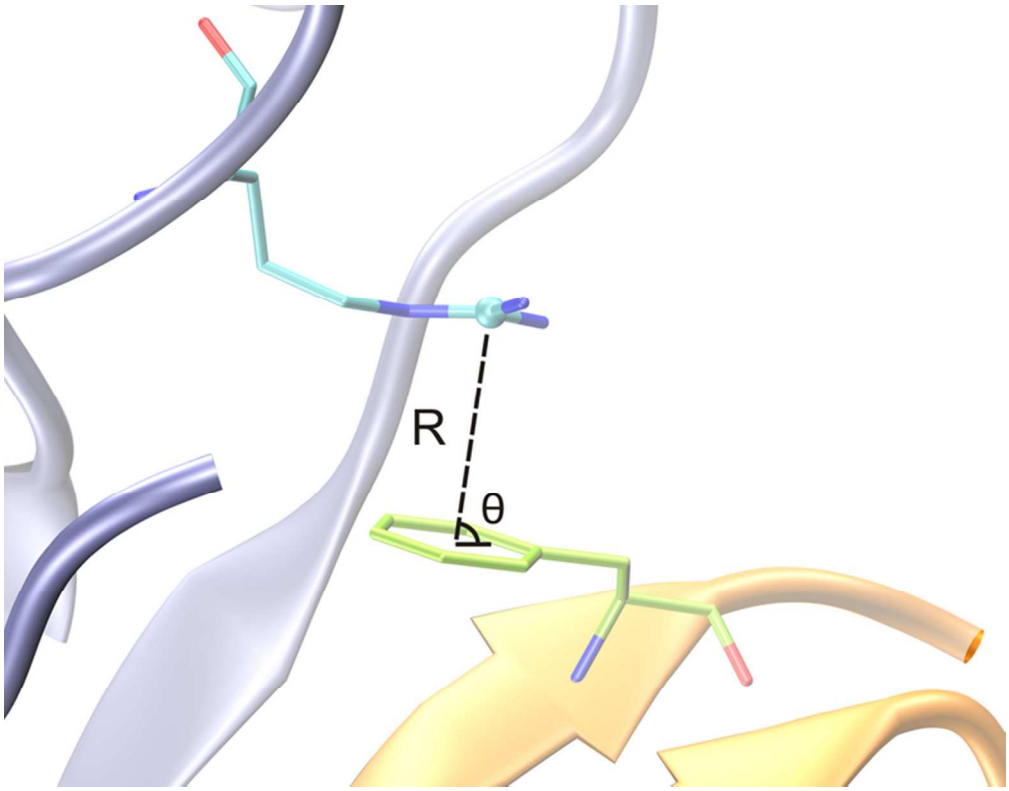


Fig. 2 The geometric parameters for cation-π interactions. The distance (R) was calculated from the cation to the centroid of the aromatic ring. The angle (Θ) was defined between the cation-centroid vector and the aromatic ring plane.
67x52mm (300 x 300 DPI)

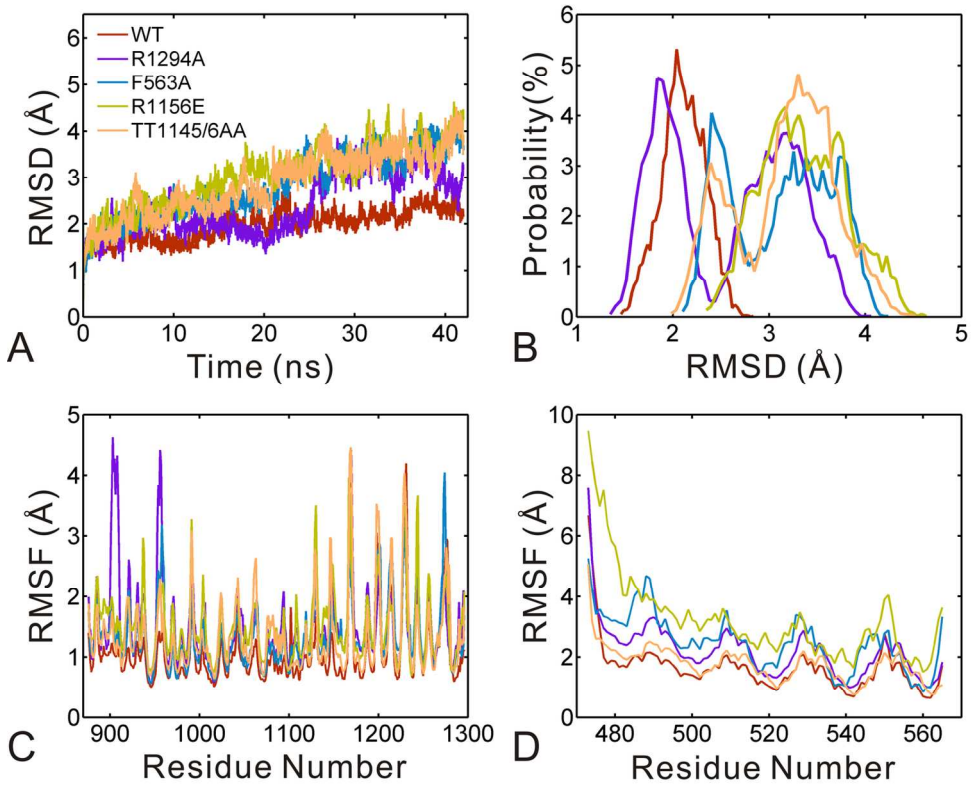


Fig. 3 Comparison of RMSD and RMSF for WT, R1294A, F563A, R1156E and TT1145/6AA systems. (A) The RMSD values of backbone atoms versus simulation time. (B) The probability distribution of RMSD calculated from 12~42 ns trajectories. (C) The RMSF values of BoNT/A-RBD calculated from 12~42 ns trajectories. (D) The RMSF values of SV2C-LD calculated from 12~42 ns trajectories.
141x116mm (300 x 300 DPI)

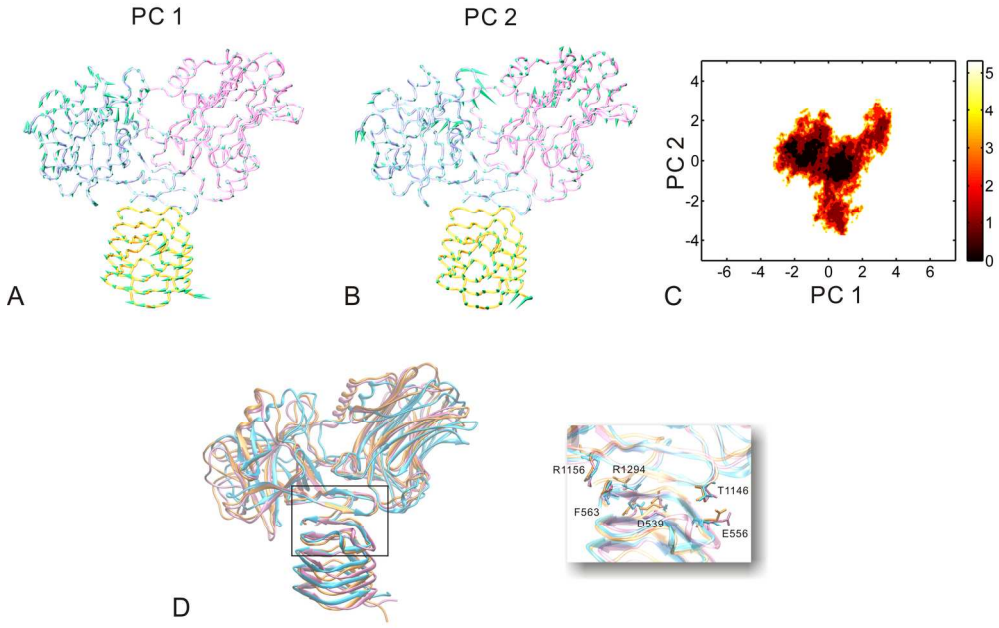


Fig. 4 The first and second motion modes (PC1 and PC2) of WT, the corresponding free energy contour map and the representative structures in the energy basins. The average structure of BoNTs/A-RBD:SV2C-LD complex is based on the last 30 ns trajectory and depicted with the tube model. The first (A) and second (B) motion modes are shown with the cone model. The length of cone is positively-correlated with the motion magnitude, and the orientation of cone indicates the motion direction. (C) Free energy contour map versus the motion modes at 310 K, with deeper color indicating the lower energy. (D) The representative structures in the energy basins compared with the crystal structure. The interface structure is highlighted in the inset figure. The crystal structure is colored blue. The first and second representative structures are colored orange and mauve, respectively. The important interface residues are represented by the stick model.
170x106mm (300 x 300 DPI)

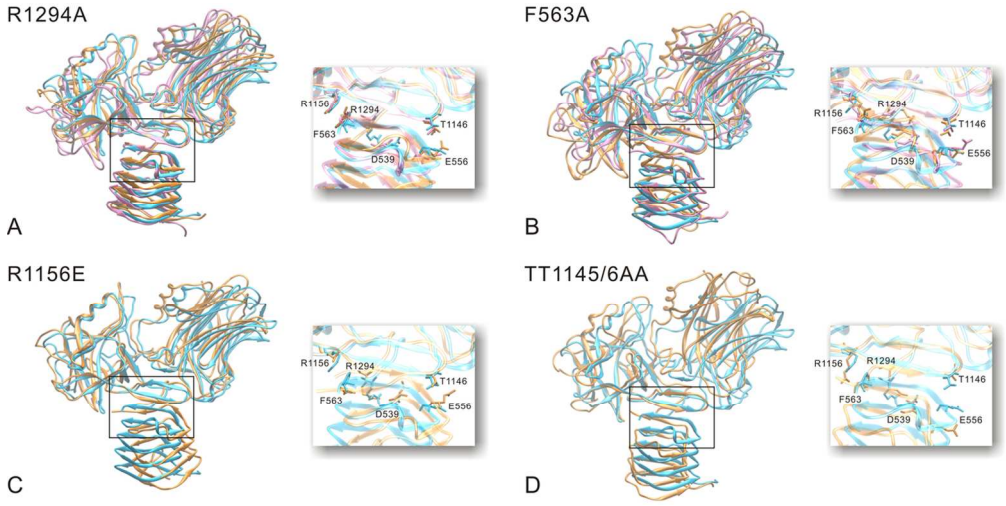


Fig. 5 The representative structures of mutant systems in the energy basins compared with the crystal structure. The interface structure is highlighted in the inset figure. The crystal structure is colored blue. The important interface residues are represented by the stick model. For the R1294A (A) and F563A (B) systems, the first and second representative structures are colored orange and mauve, respectively. For the 1156E (C) and TT1145/6AA (D) systems, the only one representative structure is colored orange.
120x60mm (300 x 300 DPI)

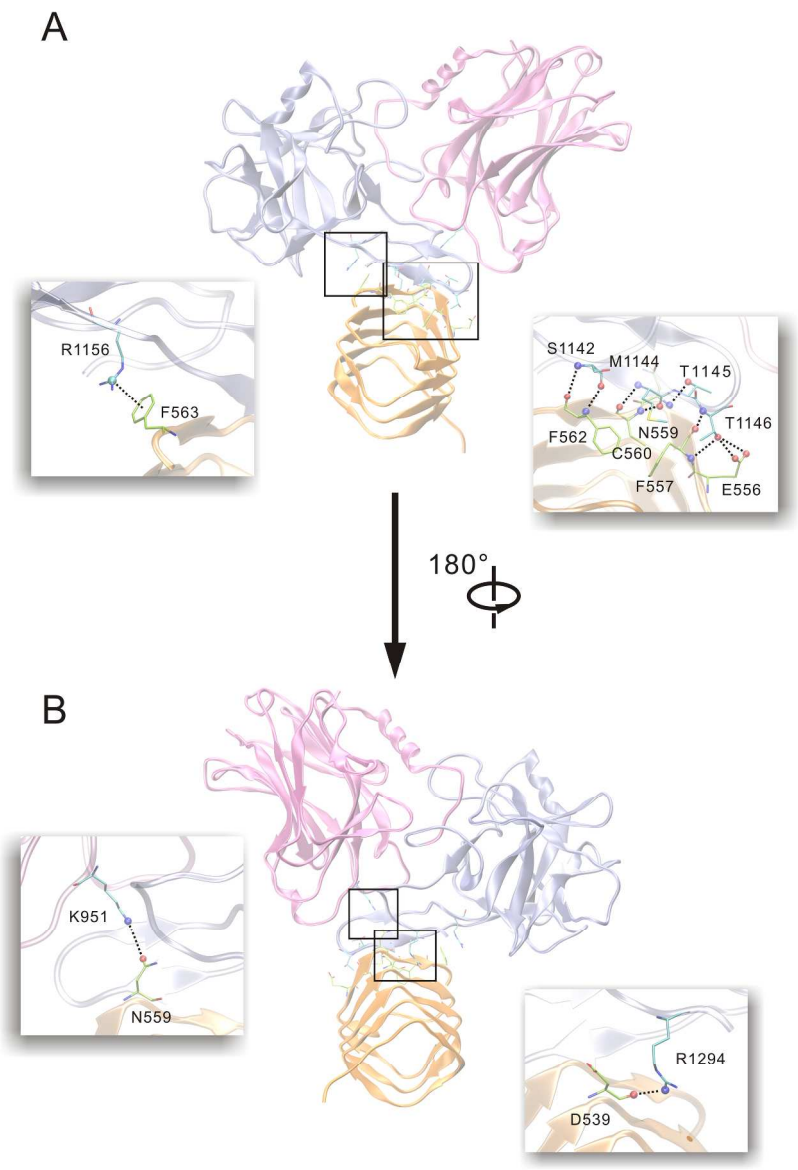


Fig. 6 The crucial interactions between BoNT/A-RBD and SV2C-LD. (A) Cartoon representation of the complex structure of BoNT/A-RBD: SV2C-LD. (B) The complex structure from a different view. The close-up views of crucial interactions are shown in the four inset figures. The interaction residues are represented by the stick model, and the crucial atoms are depicted with CPK model.
215x322mm (300 x 300 DPI)

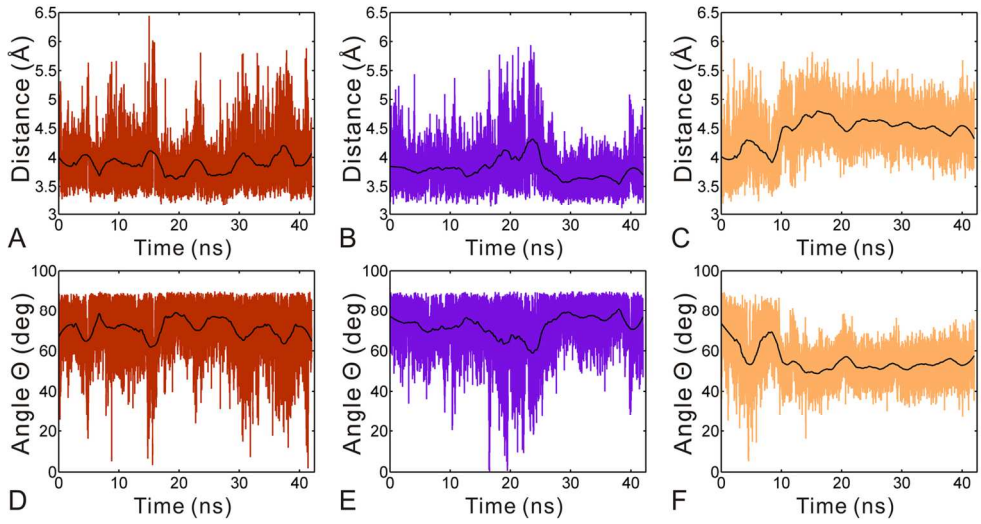


Fig. 7 Time series of the distance and angle for the cation-π interactions in the WT (red), R1294A (purple) and TT1145/6AA (orange) systems. The smooth values are represented by the black curves.
137x72mm (300 x 300 DPI)