Journal of Structural Biology 214 (2022) 107876

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Research Article

Elucidation of critical pH-dependent structural changes in Botulinum Neurotoxin E

Christophe J. Lalaurie [a], Andrew Splevins [b,d], Teresa S. Barata [c,d], Karen A. Bunting [d], Daniel R. Higazi [e], Mire Zloh [f], Valentina A. Spiteri [g], Stephen J. Perkins [g], Paul A. Dalby [a,*]

[a] Department of Biochemical Engineering, Bernard Katz Building, University College London, Gordon Street, London WC1H 0AH, UK
[b] Evox Therapeutics Ltd, Oxford Science Park, Medwar Center, Oxford, England OX4 4HG, UK
[c] FairJourney Biologics, 823 Rua do Campo Alegre, Porto, Porto 4150-180, Portugal
[d] Ipsen Bioinnovation, 102 Park Drive, Milton Park, Abingdon, Oxfordshire OX14 4RY UK
[e] Ipsen Biopharm Ltd., Wrexham Industrial Estate, 9 Ash Road, LL13 9UF, UK
[f] UCL School of Pharmacy, University College London, 29-39 Brunswick Square, London WC1N 1AX, UK
[g] Department of Structural and Molecular Biology, Division of Biosciences, Darwin Building, University College London, Gower Street, London WC1E 6BT, UK

ARTICLE INFO

Keywords:
Botulinum Neurotoxin
Molecular dynamics
Small-angle X-ray scattering

ABSTRACT

Botulinum Neurotoxins (BoNT) are the most potent toxins currently known. However, they also have therapeutic applications for an increasing number of motor related conditions due to their specificity, and low diffusion into the system. Although the start- and end- points for the BoNT mechanism of action are well-studied, a critical step remains poorly understood. It is theorised that BoNTs undergo a pH-triggered conformational shift, activating the neurotoxin by priming it to form a transmembrane (TM) channel. To test this hypothesis, we combined molecular dynamics (MD) simulations and small-angle x-ray scattering (SAXS), revealing a new conformation of serotype E (BoNT/E). This conformation was exclusively observed in simulations below pH 5.5, as determined by principal component analysis (PCA), and its theoretical SAXS profile matched an experimental SAXS profile obtained at pH 4. Additionally, a localised secondary structural change was observed in MD simulations below pH 5.5, in a region previously identified as instrumental for membrane insertion for serotype A (BoNT/A). These changes were found at a critical pH value for BoNTs *in vivo*, and may be relevant for their therapeutic use.

1. Introduction

The Botulinum NeuroToxin (BoNT) family of proteins, produced by *Clostridium botulinum*, are currently the most potent toxins known. With an LD_{50} reaching as low as 1 ng/kg (Pirazzini et al., 2017) depending on the serotype, they are considered a potential bio-terror threat (Rossetto et al., 2015) due to their common presence in soils around the world. There are seven widely accepted serotypes, termed A through to G (Kukreja et al., 2010); which can be divided into further subtypes (termed A1, A2 etc). A new X serotype has also been reported (Masuyer et al., 2015). More subtle variations of the known sero- and sub-types are likely to be discovered over time, with advances in deep-sequencing techniques. BoNTs are responsible for Botulism, a flaccid paralysis condition which can be lethal if not detected in time. The most common method of infection is through ingestion of contaminated meat, through the gastro-intestinal tract. BoNTs are produced with a progenitor complex, which protects them from the highly acidic conditions until they enter the bloodstream (Rummel, 2015). These proteins are highly specific to nerve terminals, and when injected locally have low diffusion into the surroundings. Their extreme potency and specificity is also of significant interest to the pharmaceutical industry, with many new medical applications discovered since the first reported use for the treatment of blepharospasm (Cooper, 2007; Fonfria et al., 2018; Jankovic, 2004), a list that is likely to expand as our understanding of these complex proteins grows.

BoNTs and their overall mechanism of action have been well studied, from entry into the system through to the target proteins and their resulting effects. The 150 kDa protein is initially produced as a single

Abbreviations: MD, Molecular Dynamics; SAXS, Small Angle X-ray Scattering; BoNT/E, Botulinum Neurotoxin - serotype E; LD50, Lethal dose 50%; RMSD, Root mean square deviation; SASA, Solvent Accessible Surface Area; BoNT/A, Botulinum Neurotoxin - serotype A.

* Corresponding author.
E-mail address: p.dalby@ucl.ac.uk (P.A. Dalby).

https://doi.org/10.1016/j.jsb.2022.107876
Received 22 March 2022; Received in revised form 13 June 2022; Accepted 16 June 2022
Available online 20 June 2022

chain and is then split into a heavy (100 kDa, HC) and light (50 kDa, LC) chain. The HC contains a binding domain (BD) and a translocation domain (TD), while the LC forms a single domain and is the catalytic part of the toxin. A disulphide bond formed between the TD and the LC, maintains the protein as a single entity after a small region is cleaved between the TD and the LC, and is crucial for *in vivo* function. Two structural conformations of BoNTs have been identified through crystallography. The open conformation has the LC and the BD on opposite sides of the TD (Fig. 1a), while the compact conformation brings the LC and the BD into proximity, folded around the TD (Fig. 1b). Most BoNT serotypes have been identified as being in the open conformation under physiological pH conditions. However, serotype E (BoNT/E) has been observed in the closed conformation, which is closer to the known crystal structure of Tetanus Neurotoxin (TeNT), a close homologue with similar structure and activity (Fig. 1c).

Mechanistically, the BD enables entry into the nerve cell by binding to the surface of the neuron. Upon neuro-transmitter release, the BD latches onto luminal vesicle protein domains exposed on the cell surface and the toxin is internalised during re-entry into the emitting neuron. The protein is inside the vesicle as it is prepared for a new cycle of neurotransmission, during which the pH drops to ≈ 5 (Fischer and Montal, 2007). At this point, the TD forms a transmembrane channel in the vesicle, and the LC enters the cytoplasm where the disulphide bond is reduced. After reduction, the LC acts as an endopeptidase, and targets proteins of the SNARE family. These proteins form the SNARE complex whose function is to fuse the vesicle, which contains the neurotransmitter, to the pre-synaptic membrane and facilitate neurotransmitter release into the synaptic cleft. By targeting this family of proteins, BoNTs prevent formation of the SNARE complex and hence silence neuro-transmission (Montal, 2010). The modular design of BoNTs, and the functionally distinct domains, increases their potential for therapeutics, by leaving open the possibility to combine different domains for a particular purpose (Keith and John, 2010); such as Miyashita et al (Miyashita et al., 2021) achieved to treat botulism in mice by fusing antibodies to the BD and TD of a BoNT.

While this paints a clear picture of the process from start to finish, the physical effect of the pH drop on the protein structure remains poorly understood, though a conformational change is likely to take place prior to formation of the transmembrane channels. A study on TeNT (Masuyer et al., 2017), a very close homologue to BoNTs (see sequence alignment in Supplementary Fig S1a,b), revealed a more extended structure at lower pH conditions, but a more compact structure, close to that of BoNT/E, at higher pH values. This suggests that a similar structural change may occur in BoNT/E. However, previous computational studies of BoNTs have focused on BoNT/A, and were also limited to short simulation times and pH values (Chen and Deng, 2007; Chen et al., 2007; Wang et al., 2015). Despite shorter simulation times (between 60 and 200 ns), Chen et al were able to identify some small localised conformational changes in BoNT/A. In order to detect larger scale global changes such as inter-domain movement, longer simulations would be required. A number of other proteins are also known to require acidic pH environments to release their payload, such as influenza virus hemagglutinin or Ebola virus GP. These proteins rely on a smaller scale secondary structure change in which a hydrophobic loop is buried at neutral pH but exposed in acidic environments (Gregory et al., 2011; Lam et al., 2018). A similar mechanism was identified in BoNT/A, and is conserved in BoNT/E and TeNT, by Lam et al (2018), and warrants analysis of this region in BoNT/E. Finally, pH is known to affect the stability of proteins in ways specific to each protein (O'Brien et al., 2012). In some cases, proteins display increased flexibility at lower pH values, such as chymotrypsin inhibitor 2, where lower forces are required to achieve unfolding. This effect is in part due to the protonation of specific residues (Tollinger et al., 2003). Identifying a similar behaviour in BoNT/E may shed light on the pH dependent activity of this protein.

To explore these hypotheses for BoNT/E and expand on the MD

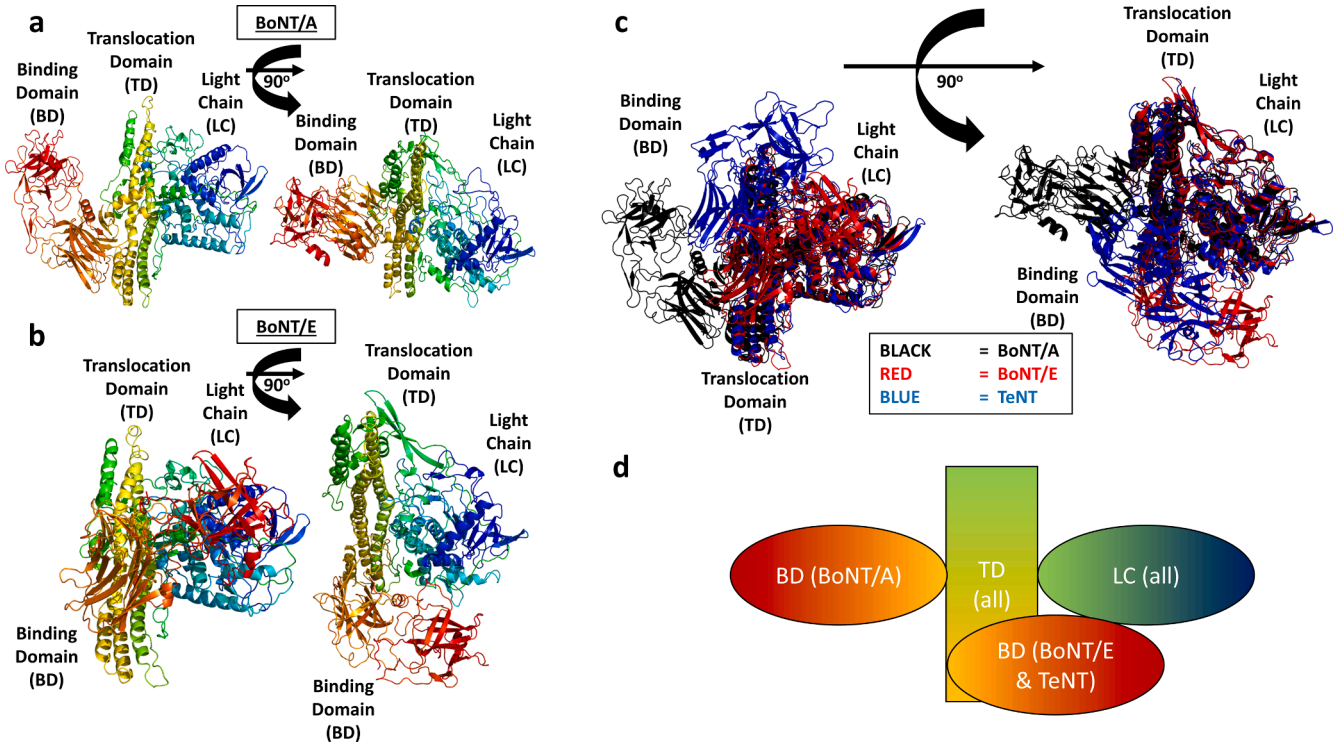

Fig. 1. Cartoon representation of BoNT/A and BoNT/Eserotypesand aligned structures of BoNT/A, BoNT/E and TeNT. (a) represents BoNT/A, with the BD and LC on either side of the TD. **(b)** represents BoNT/E, with the BD near the LC. **(c)** shows the alignment of the three structures of BoNT/A (PDB ID: 3BTA), BoNT/E (PDB ID: 3FFZ) and TeNT (PDB ID: 5N0B). This displays the likeness of BoNT/E and TeNT, and how they differ from the currently known structure of BoNT/A. **(d)** Schematic representation of the domain organization in BoNT/A, BoNT/E and TeNT.

landscape of BoNTs, a set of 400 ns molecular dynamic (MD) simulations were conducted at several pH conditions in the range from 4 to 7. By increasing the final runtime of these MD simulations, we aimed to observe similar local structural changes but also expected larger scale (eg. inter-domain) conformational changes to occur. In particular, these could identify the role of the BD / LC interface in BoNT/E conformational changes, and how this may affect the activity. Global root mean square deviation (*RMSD*, a measure of the deviation of a structure with respect to a reference structure) analyses of the trajectories, using the crystal structure of BoNT/E (PDB ID: 3FFZ) (Kumaran et al., 2009) as a point of reference, explored protein flexibility; while the radius of gyration (R_G, a measure of protein extension) and solvent accessible surface area (*SASA*) tracked any global conformational changes over time at each pH. Structural changes within each of the individual domains were also analysed in more detail. Clustering of the full trajectories by principal component analysis (PCA) determined that the protein occupied a small number of discrete conformations across the full pH range studied, and did not simply undergo a broad range of random motions. Finally, the conformations explored by the MD simulations were used to generate theoretical small angle x-ray scattering (SAXS) curves, and compared to experimental curves obtained at instrument B21 (Cowieson et al., 2020) (Diamond Light Source).

2. Results

2.1. MD simulations reveal pH dependent behaviour

MD simulations allow for the study of solution behaviour of proteins with atomic level resolution. Despite relying on a number of user-defined conditions, they have been shown to accurately mimic experimental behaviour consistently (Hollingsworth and Dror, 2018). With the current advances in computing power, they are increasingly used as a screening ahead of experimental tests (Suan Li and Khanh Mai, 2013). The trajectories presented here were first analysed using *RMSD* and R_G calculations within Gromacs (Van Der Spoel et al., 2005), in view of identifying any local and/or global conformational changes. The R_G values (Supplementary Fig S2a) showed that for pH 5 and below, most repeats explored different conformations in which an increase in R_G values was observed. The two biggest changes in R_G values occurred in the first pH 4 and the third pH 5 repeats. In both of these, the protein transitioned from an R_G value starting at 3.60 nm and extended gradually towards a final value of ≈ 3.85 nm; an increase of 7%. The R_G value for the third pH 5 repeat had not fully stabilised by the end of the simulation and so may have increased further if the simulation time were to be extended. By contrast, at pH 5.5 and above, the R_G values remained at a close to constant 3.60 nm, which matches that of the crystal structure of the closed form, obtained at pH 7 (PDB ID: 3FFZ) (Kumaran et al., 2009).

RMSD analysis can potentially detect conformational changes which do not affect the R_G value. This was also calculated as a function of trajectory time (Supplementary Fig S2b). At pH 5 and below, the increase in *RMSD* values corresponded broadly to the increase in R_G values, indicating that the overall changes in structure were due to movements which also affected the R_G of the protein. At pH 5, 5.5 and 6 however, some trajectories had a large change in *RMSD* values while the R_G values remained unchanged. The first repeat at pH 6 settled at a constant *RMSD* value, however the second pH 5.5 repeat ended on an upward trend and may have increased further if the simulation had been extended beyond 400 ns. This pointed to a different conformational change in which the corresponding R_G value was unaffected. These three trajectories are addressed further below, when clustering the data through PCA. By contrast, at pH 7 both the *RMSD* and the R_G values remained largely constant throughout the four simulations.

Fig. 2 shows the *RMSD* (Fig. 2a) and R_G (Fig. 2b) value distributions at each pH, taken from a total of 140,008 frames combined from the complete set of four MD simulation repeats, and excluding the first 50 ns of equilibration. At pH 5.5 and above, the R_G value never exceeded 3.70 nm; while at pH 5 and below, the majority of counts came from R_G values >3.70 nm. The proportion of frames beyond 3.65 nm at pH 5.5 to 7 were almost negligible, accounting for less than 1% of the total. The *RMSD* value bin distribution mostly matched the R_G value distribution, except for the three trajectories at pH 5, 5.5 and 6. The increased distribution of R_G and *RMSD* values at pH 5 and below indicated an increase in flexibility, and the ability to undergo structural changes into new and potentially distinct conformations. Conformational changes that increased R_G value were only observed at pH 5 and below, whereas different changes were detected at pH 5, 5.5 and 6 that increased *RMSD*, but did not affect the R_G of BoNT/E.

The *RMSD* values above were calculated by aligning the entire protein at every time point to the structure at t = 0 which does not deconvolute inter- and intra-domain variances. To deconvolute these contributions to the global *RMSD* value, two selections were aligned individually (BD as one entity, and LC + TD as another). The intra-domain *RMSD* values were then calculated within each domain alone,

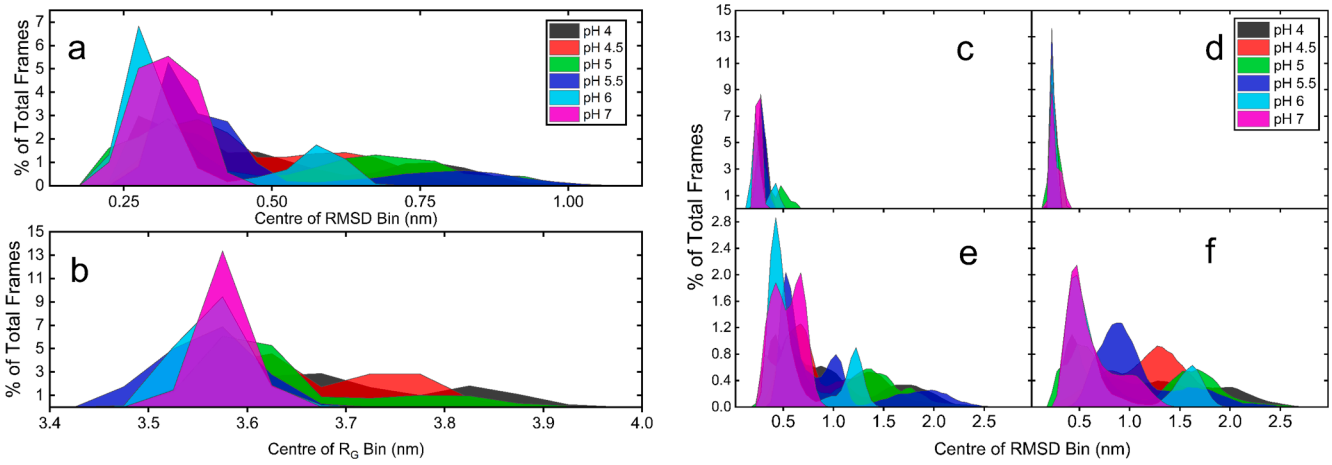

Fig. 2. *RMSD* analysis of MD simulations, with exclusion of the first 50 ns. (a) represents the binned data from *RMSD* analysis, with bin width 0.05 nm. **(b)** represents the binned data from the R_G analysis, with bin width 0.05 nm. **(c)** represents the *RMSD* bin counts of the BD after aligning it to the same selection of residues in the BoNT/E (PDB ID: 3FFZ). **(d)** represents the *RMSD* bin counts of the TD + LC after aligning it to the same selection of residues in BoNT/E (PDB ID: 3FFZ). **(e)** represents the *RMSD* bin counts of the BD after fixing the TD + LC selection, i.e. the *RMSD* of the BD with respect to the starting position of this selection in the PDB ID 3FFZ. **(f)** represents the *RMSD* bin count of the TD + LC after fixing the BD selection, i.e. the *RMSD* of the TD + LC with respect to the starting position of this selection in BoNT/E (PDB ID: 3FFZ). The sum of all the bins presented per quarter is 100%. Data collected from 840,048 backbone-only frames.

and separately the inter-domain *RMSD* values were calculated for each domain with respect to its starting position after aligning only the other domain. The distributions for these *RMSD* values at each pH are presented in Fig. 2c-f.

The intra-domain *RMSD* values for each single domain (BD or TD + LC) remained very low throughout. This indicates that the changes in global *RMSD* values were not due to protein unfolding or conformational changes within any of the individual domains, and so must have instead derived from translational movements of one domain relative to the other (inter domain). An exception to this was in the first pH 6 simulation, which had a slightly higher intra-domain *RMSD* value for the BD, which explains the observation of an overall increase in *RMSD* value but with no change in R_G value. The translational motion of the domains was confirmed from the two inter-domain *RMSD* value distributions. However, even in this more detailed analysis, the inter-domain *RMSD* values at pH 6 and 7 remained below 1 nm in most frames. This indicated that the protein was very stable at pH 7, while only a small change occurred in just one trajectory at pH 6. By contrast, at the lower pHs the inter-domain *RMSD* values were widely distributed up to 2.5 nm, again indicating a higher flexibility of the protein and an ability to change the relative position of its domains.

Given that the TD-membrane interactions may be driven by solvent exposure at lower pH values, for otherwise buried structure, any changes in solvent exposure resulting from the observed inter-domain conformational changes, were located and quantified. The overall *SASA* of the full protein was calculated for every frame. The average *SASA* from the last 200 ns of four repeats per pH showed a net increase when the pH was decreased from 6.0 to 4.5 (Fig. 3a), with the transition centred at between pH 5.0 and 5.5. The largest overall change came from the BD with a large portion shielded by the LC at pH 6.0, however this domain became increasingly and steadily solvent exposed as the pH decreased (Fig. 3b).

While the *SASA* of the LC and BD showed a clear transition at pH 5.0–5.5, the TD was found to have an increased *SASA* at pH 5.5 and 6 compared to pH 7. At pH 5 the *SASA* decreased coming closer to that of the reference structure at pH 7, however from pH 4.5 to 4 the SASA increased gradually. This decrease in SASA between pH 5.5 and 5, averaged over four trajectories, reflected an increase in flexibility and conformational freedom at this pH; with some repeats leaving the TD more exposed, while others resulted in either unchanged SASA or a more shielded domain. Fig. 4a presents the localisation of the change in *SASA* on the TD relative to that at pH 7, and highlights the residues most affected by the conformational change. A small section near the "lower" half (residues 808 – 823) of the TD was consistently affected, and is near the hinge region that links the TD and the BD together. Another area of importance was a short helix (residues 614 – 620) that protruded from the TD in the "upper" half (Fig. 4b,c). A homologous region was previously identified, through crystallography and functional studies, as playing a crucial role for membrane insertion in BoNT/A (Lam et al., 2018). This region saw an increase in *SASA* at pH 5 and below, and yet remained largely unchanged at pH 5.5 and 6, with respect to pH 7. This region forms several α-helices which maintained their secondary structure at pH 5.5 to 7 throughout the full trajectories. By contrast, one of the helices briefly unfolded and increased the solvent exposure of a lysine residue (K618) exclusively in pH 4 to 5 simulations.

In order to explain the observed increase in flexibility below pH 5, the charge distribution was analysed at the starting point of each pH condition. At the transition from pH 5.5 to pH 4.5 the charge on the TD + LC domain increased at the steepest rate, while the net charge switched from negative to positive at just below pH 6 (Fig. 5). This resulted in a net positive charge on both domains that increased rapidly below pH 6, and the resulting inter-domain repulsions could explain the sudden surge in flexibility and domain movement relative to each other. The number of salt bridges also decreased rapidly at below pH 5.5 which would further facilitate the loss of stability at the domain interface (Fig. 5). Of particular interest were two areas identified as interacting regions at the BD / LC interface, in which electrostatic interactions were lost between pH 5.5 and pH 5. Key salt bridges included Asp338/Lys834, where the Asp changed from negatively charged to neutral, and Lys323/Glu1058 where the Glu shifted from negatively charged to neutral. In both cases, the Lys residues remained positively charged. Additionally, a short helix motif conserved from diphteria toxin (Ratts et al., 2005) underwent a strong electrostatic shift between pH 5.5 and pH 5 with the net loss of two negatively charged residues, Glu810 and Asp817 which became neutral upon protonation.

2.2. PCA clustering separates pH ≤ 5 and pH > 5

While the data presented above indicated a shift in the protein behaviour at below pH 5, it was not enough to confirm whether the changes were due to a concerted conformational shift between discrete populations, or whether each peak in *RMSD* or R_G values contained broad distributions of conformations sampled through random motions. To investigate this, the trajectories were combined using only the backbone atoms (C, CA, N) and subjected to PCA to evaluate the conformational clusters present in the combined trajectories.

All 24 MD trajectories were saved as backbone-only files (C, CA, N), and then concatenated together into a single file of 70,004 frames per pH, that included only the last 350 ns of the trajectories, with a timestep of 40 ps. Most of the variance could be reasonably accounted for using

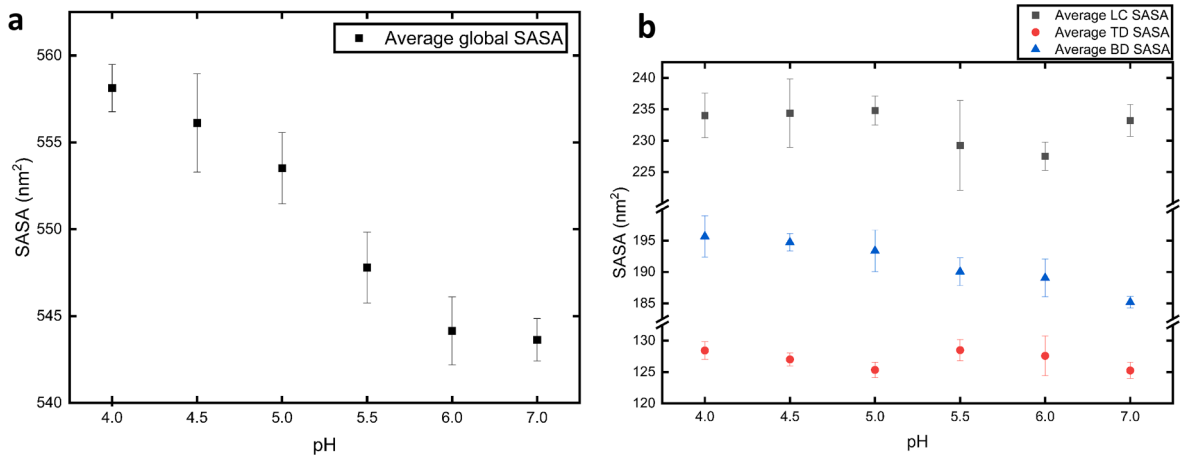

Fig. 3. Average Solvent Accessible Surface Area of BoNT/E. (a) represents the average *SASA* of the full BoNT/E at each pH, with one standard deviation error bar. **(b)** represents the *SASA* broken down by domain by summing the *SASA* of all residues within the domain, with one standard deviation error bar. *SASA* values in both cases are the averaged *SASA* from the last 200 ns of four repeats per pH.

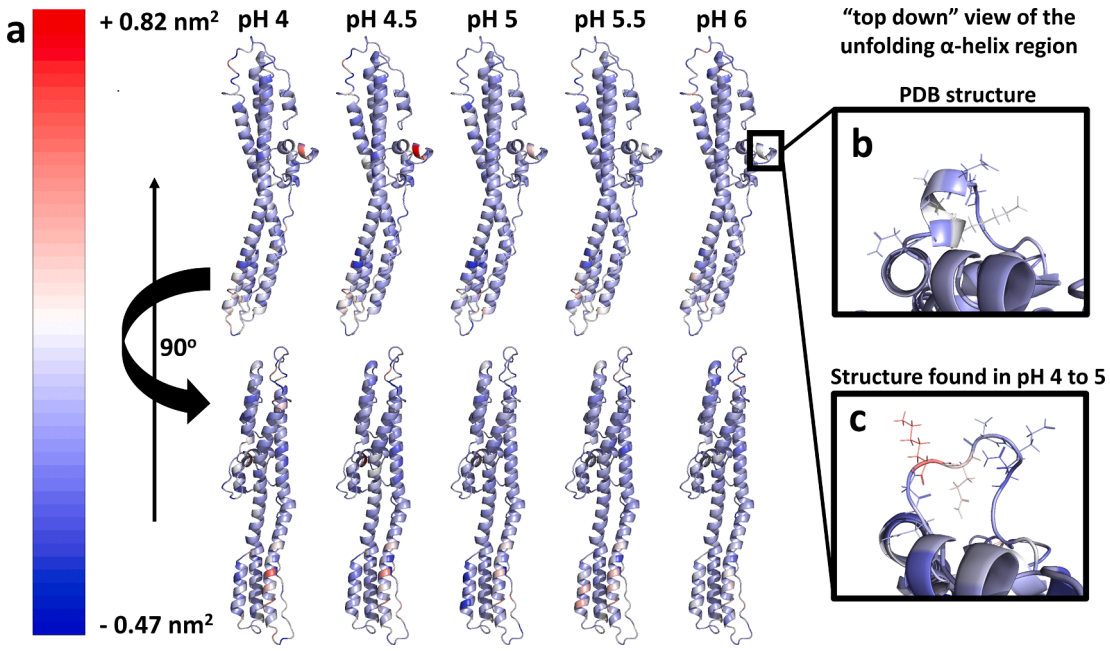

Fig. 4. TD of BoNT/E, coloured by *SASA* difference with respect to pH 7. (a) represents the location of the highest *SASA* change with respect to pH 7 (blue = largest negative change, red = largest positive change). **(b)** and **(c)** are zoomed visuals on the short helix region which had the largest increase in SASA at pH 4 to 5 **(b)** and a structure where the secondary structure is lost **(c)**, exposing a lysine residue (white **(b)**/red **(c)**) K618. This conformation was exclusively found in pH 4 to 5 simulations. The *SASA* difference with respect to pH 7 was taken from the average of four repeats per pH.

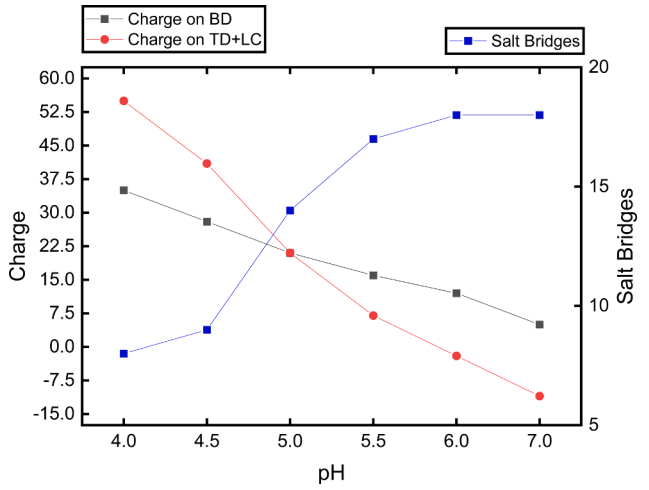

Fig. 5. Charge and salt-bridge content of BoNT/E per pH. Charge content per domain and salt-bridge content in the starting structure for each pH. Charge calculated in PyMOL; salt-bridges counted in VMD 1.9.3.

only 4 clusters (Supplementary Fig S3). The PCA biplots and scree plot are presented in Supplementary Fig S4. Supplementary Fig S5 presents the contribution of each pH trajectories to the total frames per cluster (Supplementary Fig S5a) and the contribution of each cluster to the total frames per pH (Supplementary Fig S5b).

We found that at pH 5.5, 6 and 7, the BoNT MD conformations remained primarily in structural cluster 1 for the entire 350 ns. Only a single repeat at pH 5.5 explored cluster 4, and a single repeat at pH 6 explored cluster 3, each with no effect on the R_G of the protein. The structures from cluster 1 were observed at every pH, but their contribution decreased from 97% at pH 7, to 75% at pH 5.5 and pH 6.0, and then again to 50–54% at pH 5.0 and below. By contrast, structures from cluster 2 contributed increasingly to the total frames per pH as it decreased from pH 5 (25%), to pH 4.5 (40%), and then pH 4.0 (50%), thus indicating overall a gradual switch from cluster 1 to cluster 2 at below pH 5.5. To expand on this, it is important to understand what each cluster represents and how robust the clustering is. To determine this, the average *RMSD* with respect to the cluster centroid was calculated as shown in Table 1, where the centroid was the structure that is closest to all the other structures within the cluster. Cluster 2 had the highest *RMSD* with respect to its centroid, meaning that there was more variation within that cluster than in the others. However, the average *RMSD* of this cluster remained below 0.6 nm, demonstrating that the structures placed into it were all closely related. The other clusters were also reliable, with an average *RMSD* lower than 0.4 nm with respect to their centroids. These clusters are presented as a dendrogram in Supplementary Fig S6, and as surface representations in Fig. 6b. In this representation, the LC + TD selection is aligned and shows the major movements of the BD. Clusters 1 and 2 dominate the distribution and may form two ends of a pH dependent shift required for activity. Clusters 3 and 4 are less populated and may be transient states.

Combining the clustering data with the R_G and *RMSD* data per frame, the average R_G and *RMSD* values per cluster were calculated (Table 1). This, in combination with the mid-point *RMSD* values, confirms the robustness of the clustering, with low standard deviations with respect

Table 1
Summary of clustering data.

		Cluster 1	Cluster 2	Cluster 3	Cluster 4
% total frames per pH	4	50.03	49.92	0.05	0.00
	4.5	54.80	41.10	0.86	3.24
	5	52.13	23.57	23.44	0.86
	5.5	75.36	2.00	0.64	22.00
	6	75.14	0.27	24.59	0.00
	7	97.30	0.00	2.70	0.00
***RMSD* from centroid**	**nm**	0.37 ± 0.08	0.53 ± 0.14	0.34 ± 0.16	0.30 ± 0.09
***RMSD* from reference**		0.33 ± 0.06	0.63 ± 0.13	0.63 ± 0.13	0.75 ± 0.14
R_G		3.58 ± 0.04	3.76 ± 0.07	3.59 ± 0.02	3.61 ± 0.03

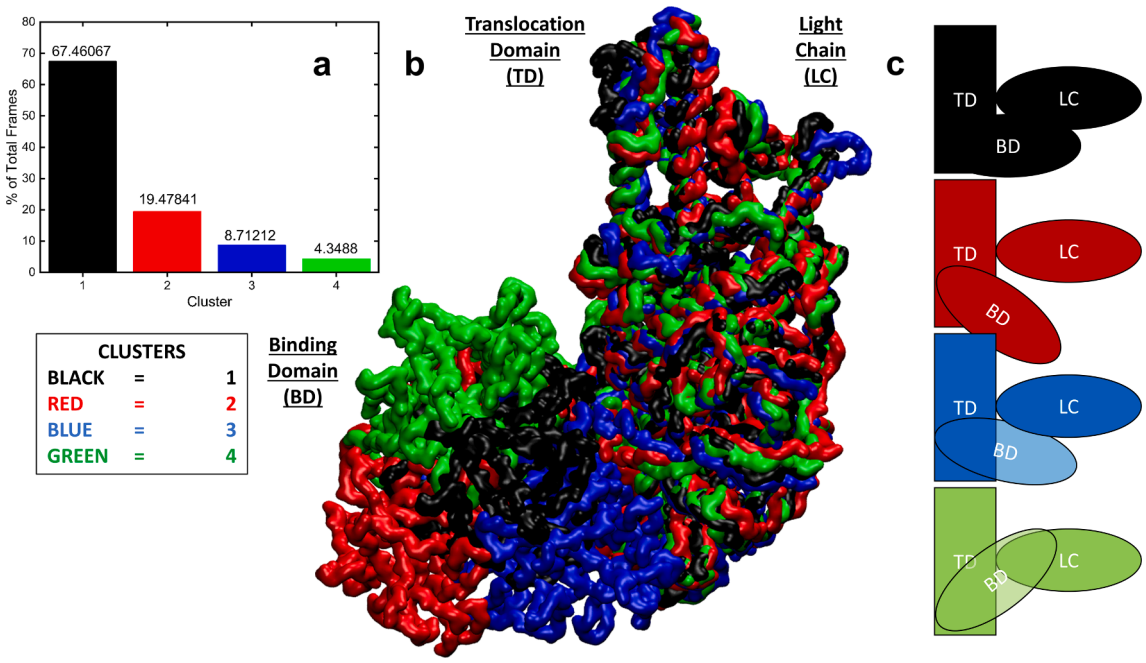

Fig. 6. Surface representation of the cluster centroids, identified by PCA. (a) shows the percentage each cluster contributes to the total frames used for the analysis (420,024 frames). **(b)** represents the surface of the four mid-point structures, i.e. the structure which has the lowest *RMSD* with respect to the rest of the clusters' frames. **(c)** Schematic representations of the domain organization in the clusters as seen from a 90° counter-clockwise angle with respect to **(b)**.

to the average. Supplementary Fig S7 shows the population weighted average *RMSD* and R_G values per pH, and confirms the more extended state of the protein starting from pH 5 down to pH 4.

2.3. SAXS validates the structural change seen in the MD

The MD simulation analyses were validated using experimental SAXS measurements, which provide insight into the solution structure of proteins and can detect conformational differences. Samples were dialysed into six different pH buffers, concentrated, and immediately loaded onto the HPLC-SEC system which feeds directly into the path of the beam. The resulting SAXS data is presented in Fig. 7a,b. The experimental I(Q) curves for most of the samples represented a compact globular structure, with a single peak in the Kratky plot and flat tail at the higher Q-values.

$$Q = \frac{4.\pi.\sin(\theta)}{\lambda}$$

With 2θ the scattering angle, and λ the wavelength of the beam.

At pH 4 however, two elution peaks were observed (Fig. 7c), which produced two different scattering curves. The first peak matched the data from all the other conditions and implied a similarly compact structure. The second peak had a steeper low-Q region in the I(Q) vs Q plot, indicating an increased R_G of the protein, and a tail in the Kratky plot which indicated that a second domain was now "visible" by the beam, with its own scattering profile.

Using the SASSIE web server (Curtis et al., 2012), theoretical SAXS curves were generated from each trajectory and matched against

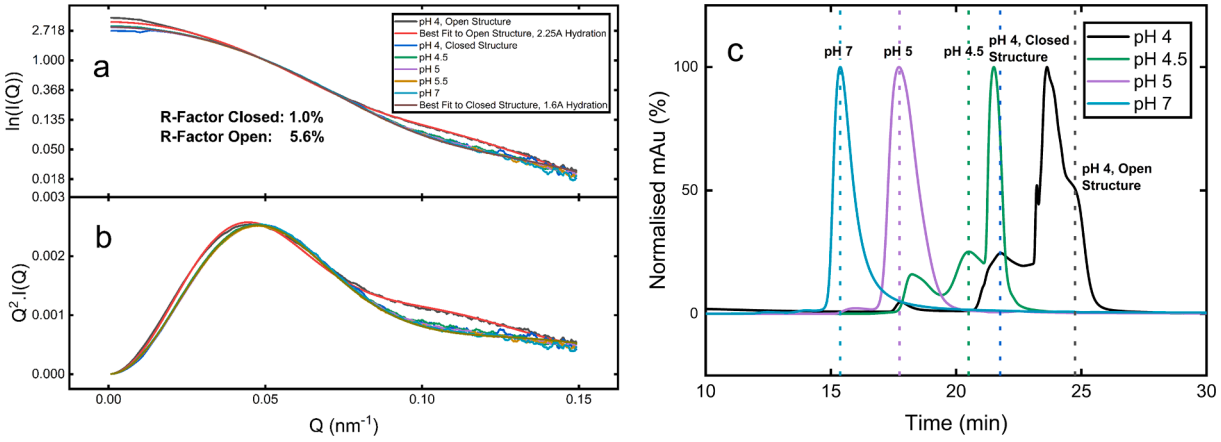

Fig. 7. ln(I(Q)) vs Q and Kratky plots. (a) ln(I(Q)) vs Q of BoNT/E at different pH conditions, with the best-fit curves to both structures. pH 4 showed two distinct populations which eluted at different times, and resulted in two different scattering profiles (shown in Fig. 7c). **(b)** Kratky plot displaying the difference between the two populations, with the higher tail in the red & black curves indicating a separation of two domains, compared to the lower tail in the other curves (blue, green, purple, yellow, cyan & brown). **(c)** HPLC trace of the elution prior to SAXS analysis. Most conditions eluted in a single peak, while pH 4.5 and pH 4 showed several. The highest peak in pH 4 & pH 4.5 was found to be aggregated protein; however in pH 4 a smaller population distinct from this peak was identified and matched a distinct conformation.

experimental curves obtained at the Diamond Light Source B21 beamline using SEC-SAXS. The best-fit curves were obtained by loading all 160,004 frames per pH from the MD simulations onto the SASSIE server, and using an R-factor fit to the experimental curve (Supplementary Figs S8a,b).

$$R = \frac{\sum_Q \left| B_f(I(Q)) - Exp(I(Q)) \right|}{\sum_Q |Exp(I(Q))|}$$

With $B_f(I(Q))$ the theoretical best fit intensity from the model, and $Exp(I(Q))$ the experimental intensity at the same Q value (Curtis et al., 2012; Press et al., 2007).

An initial screen of the backbone-only trajectory identified a best-fit structure that was then extracted with full atoms and a hydration shell included. This decreased further the R-factor between the theoretical curve and the experimental data. It was found that a 2.25 Å hydration shell best fit the open configuration, while a 1.6 Å shell best suited the closed configuration, to give final R-factors of 5.6% and 1% respectively. Fig. 8 shows a superposition of the 100 best fit structures for both conformations. The 100 best-fit conformations of the closed conformation corresponded closely to cluster 1 as identified by PCA, while the partially open conformations corresponded closely to the cluster 2 structures found predominantly within trajectories at pH 4 and pH 4.5. Both the experimentally observed structures, closed and partially open, had distinct pair distance distribution functions (P(r)) with the best fit structures generating good fits in real space (Supplementary Fig S9).

3. Discussion

This joint MD and scattering study has notably clarified the conformational effect of pH on the structure of BoNT/E. Recent advances in computing power and increasingly accurate tools for MD have made simulations a key tool for major discoveries (Bai et al., 2021; Wang, 2021). In the work presented here, MD simulations have been used to extensively study the conformational space available to BoNT/E, with a total of 840,048 individual conformations generated. An overall trend was identified whereby at pH 5 and below, the protein exhibited much higher flexibility with a number of trajectories in the pH range 4 to 5 exhibiting higher R_G values, and transitioned from one conformation (cluster 1) to another (cluster 2) over time. By contrast, at pH 5.5, 6 and 7 the R_G of BoNT/E remained constant throughout all 4 repeats, and just three trajectories displayed a conformational change detected by *RMSD* analysis. By separating the protein into two distinct selections (BD as one and TD + LC as another), it was shown that the conformational change was not within any of the domains, but derived from a translational movement of one domain with respect to the other, at the lower pH range. Cases of higher *RMSD* values that did not simultaneously affect the R_G appeared at pH 5, 5.5 and 6, but were rarely populated as determined by PCA, and primarily derived from only one or two trajectories in each case. Given the extensive changes observed in the lower pH range, and the absence of any notable changes in any of the trajectories at pH 7, the simulations were terminated at a total of 400 ns, with analysis excluding the first 50 ns for equilibration. The typical timescale of domain movements in protein MD is $10^{-11} - 10^{-8}$ s (Skjærven et al., 2011), and falls within the timeframe presented here. The previous MD studies identified were closer to the lower limit of this range (Chen and Deng, 2007; Chen et al., 2007; Wang et al., 2015), and only focused on BoNT/A or single domains of BoNT/A (Evander Emeltan Tjoa et al., 2019). No similar computational studies were identified focusing on BoNT/E. By performing more repeats for a longer duration, and on the whole protein alone in solution, we have presented here a deeper look into the conformational landscape of BoNT/E; with some results in good agreement with previous studies of BoNT/A (see below).

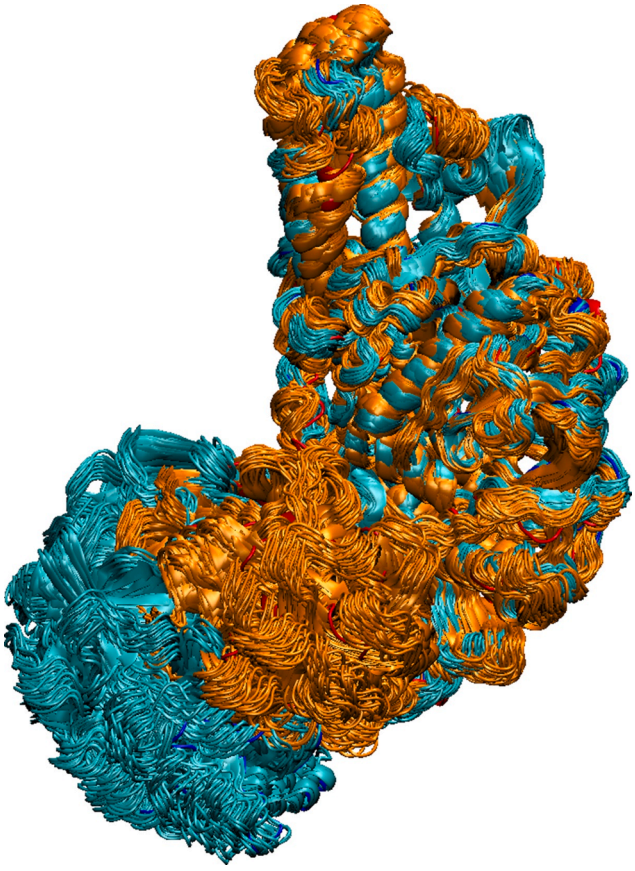

Fig. 8. Superposition of the 100 best fits to the SAXS closed (orange) and open (cyan) structures from the MD trajectories. In blue and red are the top best fit structures to the open and closed structures, respectively.

In order to explain the changes in behaviour at the switch to pH 5.0 and below, a transition in the charge distribution was observed, with a large drop in salt-bridge content and a large increase in positive charge on both domains at the lower pH range. It was therefore likely that the greater number of salt-bridges, combined with lower but opposing net charges on the two domains, were important for stability of the compact globular state at pH 6 and 7. Starting at pH 5.5, the salt-bridge content decreased to a point where some inter-domain movement was possible, and at pH 5 the sudden increase in positive charge on both domains increased the repulsion between them. This resulted in a much broader space explored in those simulations, with most trajectories moving in and out of more extended conformations. This transition at pH 5.5 and below matched the known activation pH required for this protein *in vivo*, with many studies demonstrating the absence of catalytic activity at above pH 5 (Keller et al., 2004; Pirazzini et al., 2011). It is therefore likely that a major conformational change involving the separation of the BD from the LC domain is required for activity. This change also increased the *SASA* of the whole protein; and specifically the TD, which transitions into a transmembrane channel through which the LC is released into the cytosol. Notably, a small region was detected with a large increase in the *SASA* at and below pH 5 which has homologues in all serotypes, as well as in TeNT. This small region was found to be crucial for membrane insertion in BoNT/A (Lam et al., 2018). Further analysis or mutagenesis of this region may reveal deeper insight into the structural change enabling membrane insertion in BoNT/E. Furthermore, due to the importance of this region for this protein's toxicity, it could become a key area to target with antibodies in order to prevent the membrane interaction.

To validate the population statistics and dynamics observed from the MD simulations, the structures associated with each cluster were compared to experimental SAXS data obtained for each species separated by SEC at each pH (Fig. 7c). The high-throughput server SASSIE was used to generate theoretical SAXS curves from every structure obtained (Perkins et al., 2016; Sarachan et al., 2013). These were then

compared to experimental curves obtained at Diamond Light Source, using R-factor to quantify the goodness of fit. Two scattering profiles were obtained experimentally at pH 4, separated as two peaks by SEC, indicating two distinct conformations present in solution. The two conformations matched the scattering profiles of a compact globular protein in one case, and in the other a more extended double-domain structure. These two experimental curves were compared to the theoretical curves from the MD, and good fits were found, with R-factors of 1% and 5.6%, respectively for the closed and open conformations. The increased depth of the hydration shell around the open conformation is in good agreement with the increased surface charge associated with it. The best fit to the closed conformation came from a structure very close to the crystal structure of BoNT/E (PDB ID: 3FFZ) (Kumaran et al., 2009), while the best fit for the open conformation came from cluster 2 identified by PCA.

Taken together we conclude that the BoNT/E begins its conformational journey in the compact form (i.e. within cluster 1) at physiological pH, and ends it in the open form (i.e. within cluster 2) at lower pH, some of the other structures identified in the MD are likely to be intermediate states through which the protein passes (clusters 3 and 4), before reaching a final state for creating the transmembrane channel. The separation of the two domains adjacent to the TD indicates that it is left more exposed, which may facilitate the formation of the transmembrane channel necessary for the activity of this toxin; while localised structural changes within the TD favour membrane interaction by exposing hydrophobic residues.

4. Materials and methods

4.1. Molecular dynamics simulations

The MD simulations were all started from the PDB structures (PDB ID: 3FFZ) (Kumaran et al., 2009), from which a single copy of BoNT/E was extracted through text editing of the PDB. Due to the nature of X-Ray crystallography, some of the more flexible regions in the protein structure were unresolved. These missing residues were rebuilt using SwissPDB Viewer (Guex and Peitsch, 1997), which infers the positions of the missing residues using residues in the same positions in homologous structures as a reference. The complete structure was uploaded to the PDB2PQR web server (Dolinsky et al., 2004) to simulate the protonation states of each ionisable group at each required input pH, and add hydrogen atoms to titrable residues using the PropKa (Olsson et al., 2011) algorithm. The file generated (.pqr) was converted back to a PDB file, and prepared for a standard MD simulation in Gromacs (Abraham et al., 2014; Van Der Spoel et al., 2005). The Amber03 (Duan et al., 2003) force field was used, as it included a topology and parameters for the Zinc ion present in the LC and has been reported to produce results in good agreement with experimental results on well-studied proteins (Hornak et al., 2006; Hu and Jiang, 2010). The protein was centred in a cubic box such that there was 2 nm between any periodic images of itself, solvated with water using the SPC/E model (Berendsen et al., 1987), and enough solvent molecules replaced with counter-ions to neutralise charge within the system. After setting up the system, it was energy-minimized using a steepest descent algorithm, temperature equilibrated at 300 K using a modified Berendsen thermostat, and pressure equilibrated at 1 bar using Parrinello-Rahman coupling. Hydrogen bonds were constrained during the temperature and pressure equilibration steps. The minimized, temperature and pressure equilibrated system was then left unrestrained and simulation trajectories were generated up to 400 ns, with a snapshot saved every 10 ps. Four repeats were obtained for each condition of pH 4, 4.5, 5, 5.5, 6 and 7.

The trajectories were generated using Gromacs 5.0.4 on an external High Power Computer (HPC) and were corrected for periodic boundary condition artefacts using "gmx trjconv" on Gromacs 2018.4, after transferring the data to a local computer. Once all the trajectories were complete, a file with backbone atoms only (C, CA, N) was generated from the 50 ns mark; and every MD trajectory (24 in total) was concatenated into a single file using VMD 1.9.3 (Humphrey et al., 1996), comprising 840,048 frames. A tool command language (Tcl) script, compiled from existing commands in VMD 1.9.3 and looped over every frame of the merged trajectory, was used to align the protein backbone to that of the BoNT/E (PDB ID: 3FFZ) structure, measure the *RMSD* and R_G of this; then align the BD alone and measure the *RMSD* of two selections (BD, and TD + LC) separately; then align the TD + LC, and measure the *RMSD* of the two selections separately. All images of the protein were generated in PyMOL (Schrödinger, LLC, 2015) or VMD 1.9.3.

4.2. Protein production and purification

A plasmid containing an endopeptidase-negative (i.e. non-toxic) BoNT/E DNA sequence was provided by Ipsen. The plasmid was transformed into E. coli BL21(DE3) (New England Biolabs) cells, and grown in modified terrific broth (mTB), using Kanamycin at 30 μg/mL. An overnight starter culture was used to inoculate larger cultures, which were grown at 37 oC until OD_{600} reached 0.6. The temperature was lowered to 16 °C and once cooled, 1 mM IPTG was added for induction. The induced culture was incubated for a further 21 h, after which the cells were harvested by centrifugation at 3,196 RCF. The pellets were re-suspended with 3 mL lysis buffer (35 mM NaCl, 50 mM Tris, pH 8.0) per gram of cell paste. The cells were lysed by sonication with 10 cycles of 30 s bursts, 30 s rest. The lysed mixture was centrifuged at 19,802 RCF for 30 min. The supernatant was then diluted with an equal volume of 2 M ammonium sulfate, 50 mM Tris, pH 8.0; and centrifuged again at 19,802 RCF for 30 min. A Butyl column (hydrophobic interaction) equilibrated with 1 M ammonium sulfate, 50 mM Tris, pH 8.0 was used to capture the protein, using a 50 mM Tris, pH 8.0 elution buffer. The relevant elution fractions were pooled and desalted using a 53 mL desalting column, exchanging into 10 mM Tris, pH 8.0. The desalted sample was loaded onto a Q column (ion exchange), and eluted using 500 mM NaCl, 10 mM Tris, pH 8.0. A high purity fraction of this elution was isolated and concentrated/diluted to 0.5 mg/mL. It was then pooled with 5 μg Trypsin per mg protein at 37 °C for 40 mins, to cleave the link between LC and HC ("activation" step). The activated sample was then mixed with 1.6X its volume of 5 M NaCl, 10 mM Tris, pH 8.0 and loaded onto a Butyl column equilibrated with 3 M NaCl, 10 mM Tris, pH 8.0. The final product was highly pure, activated (but still safe, endopeptidase-negative) BoNT/E and was eluted using 10 mM Tris, pH 8.0. The purity was assessed by SDS-PAGE, using 4–12% Bis-tris gels (Thermo-Fisher) and the activation was assessed by adding β-Mercapto-Ethanol in one of the lanes, which separated the band at 150 kDa into two bands at 100 and 50 kDa, respectively.

4.3. SAXS

The purified protein sample was buffer exchanged into six different buffers, at the six different pH values, using 10 mM sodium acetate for pH 4, 4.5, 5, and 5.5, and 10 mM sodium phosphate for pH 6 and 7. All the buffers were made at 20 mM ionic strength using NaCl. The final purified sample was diluted or concentrated to 0.5 mg/mL and separated into 0.5 mL aliquots. Each aliquot was then buffer exchanged against 500 mL of the appropriate buffer in a coldroom overnight, using 0.5 mL Slide-A-Lyser dialysis cassettes (Thermo-Fisher) with a 20 K MWCO.

Samples were concentrated to ≈ 4 mg/mL using spin concentrators, loaded onto a 96-well plate for autosampling into a SEC-HPLC (KW 403, SHODEX) with buffer-matched mobile phase to each sample, then directly into a temperature-controlled quartz cell capillary with a diameter of 1.5 mm of X-ray scattering instrument B21 at the Diamond Light Source (Harwell Science and Innovation Campus, Didcot, UK), operating with a ring energy of 3 GeV, and an operational energy of 12.4 keV. This recorded 3 s snapshots for 32 min, while the sample was

loaded and eluted off the SEC. The start of the elution, corresponding to buffer, was subtracted from the sample signal. Buffer subtraction was performed on ScÅtter 3 (BIOISIS) (Rambo, n.d.); and the generated scattering curves saved for analysis. Theoretical scattering curves were generated from the MD structures using the SASCALC module on the SASSIE web-server (Curtis et al., 2012; Watson and Curtis, 2013) in order to find a structure which fit the experimental curve best. The raw trajectories from Gromacs were converted to .dcd format using VMD, and loaded onto the SASSIE server with a reference PDB at the relevant pH. For analysis, the best fit structure was assessed by R-factor. A hydration shell around the protein was included using the explicit solvent of the trajectories for the best-fit frame obtained by the full-atom trajectory analysis.

To facilitate comparison between all the experimental curves and the theoretical curves generated using SASSIE, the experimental data was interpolated to a fixed number of Q points using a MATLAB script. To ensure a good fit to the true data, 800 Q points were used as this was close to the actual number of Q values collected experimentally. All the intensities of every sample were then normalised by dividing by their intensity at $Q = 0.05\ nm^{-1}$. This maintained the shape of the curve, while allowing for comparison between each condition. Analysis of ln(I(Q)) vs Q curves and Kratky plots, provided information about the approximate size and shape of the protein.

4.4. Principal component analysis

Principal component analysis was used to analyse the MD results in increased detail. For this, the raw trajectories (in .xtc format) from Gromacs were converted to the .dcd format using VMD, and only backbone atoms were kept in the trajectory, so as to remove all the solvent molecules. A stride of 2 frames was used, starting from 50 ns, leaving 17,501 frames per MD trajectory, with a frame every 20 ps. Once this reduced trajectory was generated for every pH repeat, a final file with all 24 MD trajectories was compiled by concatenating all the backbone-only files together. The PCA was performed using the Bio3D package (Grant et al., 2006) with the "fastcluster" library (Müllner, 2013) to expand the power of the "hclust" function. A first script helped determine a sensible number of clusters to use, based on how much of the variance was covered in using varying numbers of clusters. Using the determined appropriate number of clusters, a second script performed the PCA and clustered the data accordingly. This data was sorted into bins for in-depth analysis. Six clustering algorithms were trialled (single, median, centroid, mcquitty, complete and average) (Müllner, 2013); but only the "average" algorithm was taken forward as the best method (see Supplementary information), using 4 clusters.

Significance Statement

With the growing therapeutic potential of BoNTs, it is paramount to understand their complete mechanism of action, and their structural behaviour in formulations. Specifically, the importance of the pH drop during the vesicle acidification step, and its effect on the conformation of the toxin, remains elusive. Here we have identified a major conformational state of BoNT/E when the pH falls below 5.5, using a combination of MD simulations and SAXS. This conformational shift could be critical in facilitating the formation of the BoNT transmembrane channel, essential for function.

Funding information

This research was jointly funded by Ipsen, and the Engineering and Physical Sciences Research Council (EPSRC) Centre for Doctoral Training in Emergent Macromolecular Therapies (EP/L015218/1) for financial support of CJL.

Data Availability Statement

Upload of SAS data to SASBDB: https://www.sasbdb.org/data/SASDNV5/

MD data available on request.

VMD and Bio3D scripts used available at https://www.github.com/ucbecla/BoNT-E_PCA_VMD.git.

CRediT authorship contribution statement

Christophe J. Lalaurie: Investigation, Data curation, Formal analysis, Writing – original draft, Editing. **Andrew Splevins:** Conceptualization, Funding acquisition, Study coordination, Supervision. **Teresa S. Barata:** Conceptualization, Funding acquisition, Study coordination, Supervision. **Karen A. Bunting:** Conceptualization, Supervision, Validation, Writing – review & editing. **Daniel R. Higazi:** Conceptualization, Supervision, Validation, Writing – review & editing. **Mire Zloh:** Conceptualization, Supervision, Validation, Writing – review & editing. **Valentina A. Spiteri:** Resources, Data curation, Formal analysis for the SAXS data, Writing – review & editing. **Stephen J. Perkins:** Resources, Data curation, Formal analysis for the SAXS data, Writing – review & editing. **Paul A. Dalby:** Project administration, Conceptualization, Funding acquisition, Supervision, Validation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. K. A. Bunting and D. Higazi are employed by IPSEN.

Acknowledgments

C.J.L. was supported by the EPSRC Centre for Doctoral Training in Emergent Macromolecular Therapies (000033549) and IPSEN Bioinnovation. S.J.P. was supported by a joint EPSRC (EP/K039121/1) and NSF (CHE01265821) grant for CCP-SAS. We thank Dr. Akash Pandya for help in setting up the MD work; Dr Katsuaki Inoue for setting up the HPLC SEC-SAXS measurements on Instrument B21 at Diamond Light Source; and Dr Kevin Moore for his input on optimizing the protein purification procedures.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsb.2022.107876.

References

Abraham, M., Van Der Spoel, D., Lindahl, E., Hess, B., Spoel, D. van der, Lindahl, E., 2014. GROMACS User Manual version 5.0.4. www.Gromacs.org. https://doi.org/10.1007/SpringerReference_28001.

Bai, Q., Liu, S., Tian, Y., Xu, T., Banegas-Luna, A.J., Pérez-Sánchez, H., Huang, J., Liu, H., Yao, X., 2021. Application advances of deep learning methods for de novo drug design and molecular dynamics simulation. WIREs Comput. Mol. Sci. 12 (3) https://doi.org/10.1002/wcms.1581.

Berendsen, H.J.C., Grigera, J.R., Straatsma, T.P., 1987. The missing term in effective pair potentials. J. Phys. Chem. 91 (24), 6269–6271. https://doi.org/10.1021/j100308a038.

Chen, X., Deng, Y., 2007. Long-time molecular dynamics simulations of botulinum biotoxin type-A at different pH values and temperatures. J. Mol. Model. 13 (5), 559–572. https://doi.org/10.1007/s00894-007-0178-7.

Chen, Y., Chen, X., Deng, Y., 2007. Simulating botulinum neurotoxin with constant pH molecular dynamics in Generalized Born implicit solvent. Comput. Phys. Commun. 177 (1-2), 210–213. https://doi.org/10.1016/j.cpc.2007.02.095.

Cooper, G., 2007. Therapeutic uses of botulinum toxin. Ther. Uses Botulinum Toxin 5, 1–238. https://doi.org/10.1007/978-1-59745-247-2.

Cowieson, N.P., Edwards-Gayle, C.J.C., Inoue, K., Khunti, N.S., Doutch, J., Williams, E., Daniels, S., Preece, G., Krumpa, N.A., Sutter, J.P., Tully, M.D., Terrill, N.J., Rambo, R.P., 2020. Beamline B21: High-throughput small-angle X-ray scattering at

Diamond Light Source. J. Synchrotron Radiat. 27 (5), 1438–1446. https://doi.org/10.1107/S1600577520009960.
Curtis, J.E., Raghunandan, S., Nanda, H., Krueger, S., 2012. SASSIE: A program to study intrinsically disordered biological molecules and macromolecular ensembles using experimental scattering restraints. Comput. Phys. Commun. 183 (2), 382–389. https://doi.org/10.1016/j.cpc.2011.09.010.
Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., Baker, N.A., 2004. PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res. 32 (Web Server), W665–W667. https://doi.org/10.1093/nar/gkh381.
Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., Kollman, P., 2003. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. J. Comput. Chem. 24 (16), 1999–2012. https://doi.org/10.1002/jcc.10349.
Evander Emeltan Tjoa, S., Maria Vianney, Y., Emantoko Dwi Putra, S., 2019. In silico mutagenesis: decreasing the immunogenicity of botulinum toxin type A. J. Biomol. Struct. Dyn. 37 (18), 4767–4778. https://doi.org/10.1080/07391102.2018.1559100.
Fischer, A., Montal, M., 2007. Single molecule detection of intermediates during botulinum neurotoxin translocation across membranes. Proc. Natl. Acad. Sci. U. S. A. 104 (25), 10447–10452. https://doi.org/10.1073/pnas.0700046104.
Fonfria, E., Maignel, J., Lezmi, S., Martin, V., Splevins, A., Shubber, S., Kalinichev, M., Foster, K., Picaut, P., Krupp, J., 2018. The expanding therapeutic utility of botulinum neurotoxins. Toxins (Basel). 10, 1–27. https://doi.org/10.3390/toxins10050208.
Grant, B.J., Rodrigues, A.P.C., ElSawy, K.M., McCammon, J.A., Caves, L.S.D., 2006. Bio3d: An R package for the comparative analysis of protein structures. Bioinformatics 22 (21), 2695–2696. https://doi.org/10.1093/bioinformatics/btl461.
Gregory, S.M., Harada, E., Liang, B., Delos, S.E., White, J.M., Tamm, L.K., 2011. Structure and function of the complete internal fusion loop from Ebolavirus glycoprotein 2. Proc. Natl. Acad. Sci. U. S. A. 108 (27), 11211–11216. https://doi.org/10.1073/pnas.1104760108.
Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 18 (15), 2714–2723. https://doi.org/10.1002/elps.1150181505.
Hollingsworth, S.A., Dror, R.O., 2018. Molecular Dynamics Simulation for All. Neuron 99 (6), 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011.
Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C., 2006. Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins Struct. Funct. Genet. 65 (3), 712–725. https://doi.org/10.1002/prot.21123.
Hu, Z., Jiang, J., 2010. Assessment of biomolecular force fields for molecular dynamics simulations in a protein crystal. J. Comput. Chem. https://doi.org/10.1002/jcc.21330.
Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: Visual molecular dynamics. J. Mol. Graph. 14 (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.
Jankovic, J., 2004. Botulinum toxin in clinical practice. J. Neurol. Neurosurg. Psychiatry 75 (7), 951–957. https://doi.org/10.1136/jnnp.2003.034702.
Keith, F., John, C., 2010. Targeted secretion inhibitors-innovative protein therapeutics. Toxins (Basel) 2 (12), 2795–2815. https://doi.org/10.3390/toxins2122795.
Keller, J.E., Cai, F., Neale, E.A., 2004. Uptake of Botulinum Neurotoxin into Cultured Neurons. Biochemistry 43 (2), 526–532. https://doi.org/10.1021/bi0356698.
Kukreja, R.V., Sharma, S.K., Singh, B.R., 2010. Molecular basis of activation of endopeptidase activity of botulinum neurotoxin type E. Biochemistry 49 (11), 2510–2519. https://doi.org/10.1021/bi902096r.
Kumaran, D., Eswaramoorthy, S., Furey, W., Navaza, J., Sax, M., Swaminathan, S., 2009. Domain Organization in Clostridium botulinum Neurotoxin Type E Is Unique: Its Implication in Faster Translocation. J. Mol. Biol. 386 (1), 233–245. https://doi.org/10.1016/j.jmb.2008.12.027.
Lam, K.-h., Guo, Z., Krez, N., Matsui, T., Perry, K., Weisemann, J., Rummel, A., Bowen, M.E., Jin, R., 2018. A viral-fusion-peptide-like molecular switch drives membrane insertion of botulinum neurotoxin A1. Nat. Commun. 9 (1) https://doi.org/10.1038/s41467-018-07789-4.
Masuyer, G., Conrad, J., Stenmark, P., 2017. The structure of the tetanus toxin reveals pH?mediated domain dynamics. EMBO Rep. 18, e201744198 https://doi.org/10.15252/embr.201744198.
Masuyer, G., Davies, J.R., Moore, K., Chaddock, J.A., Ravi Acharya, K., 2015. Structural analysis of Clostridium botulinum neurotoxin type D as a platform for the development of targeted secretion inhibitors. Sci. Rep. 5 (1) https://doi.org/10.1038/srep13397.
Miyashita, S.-I., Zhang, J., Zhang, S., Shoemaker, C.B., Dong, M., 2021. Delivery of single-domain antibodies into neurons using a chimeric toxin–based platform is therapeutic in mouse models of botulism. Sci. Transl. Med. 13 (575).
Montal, M., 2010. Botulinum Neurotoxin: A Marvel of Protein Design. Annu. Rev. Biochem. 79 (1), 591–617. https://doi.org/10.1146/annurev.biochem.051908.125345.
Müllner, D., 2013. Fastcluster: Fast hierarchical, agglomerative clustering routines for R and Python. J. Stat. Softw. 53, 1–18. https://doi.org/10.18637/jss.v053.i09.
O'Brien, E.P., Brooks, B.R., Thirumalai, D., 2012. Effects of pH on proteins: Predictions for ensemble and single-molecule pulling experiments. J. Am. Chem. Soc. 134 (2), 979–987. https://doi.org/10.1021/ja206557y.
Olsson, M.H.M., Søndergaard, C.R., Rostkowski, M., Jensen, J.H., 2011. PROPKA3: Consistent treatment of internal and surface residues in empirical p K a predictions. J. Chem. Theory Comput. 7 (2), 525–537. https://doi.org/10.1021/ct100578z.
Perkins, S.J., Wright, D.W., Zhang, H., Brookes, E.H., Chen, J., Irving, T.C., Krueger, S., Barlow, D.J., Edler, K.J., Scott, D.J., Terrill, N.J., King, S.M., Butler, P.D., Curtis, J.E., 2016. Atomistic modelling of scattering data in the ollaborative Computational Project for Small Angle Scattering (CCP-SAS). J. Appl. Crystallogr. 49 https://doi.org/10.1107/S160057671601517X.
Pirazzini, M., Rossetto, O., Bolognese, P., Shone, C.C., Montecucco, C., 2011. Double anchorage to the membrane and intact inter-chain disulfide bond are required for the low pH induced entry of tetanus and botulinum neurotoxins into neurons. Cell. Microbiol. 13 (11), 1731–1743. https://doi.org/10.1111/j.1462-5822.2011.01654.x.
Pirazzini, M., Rossetto, O., Eleopra, R., Montecucco, C., Witkin, J.M., 2017. Botulinum Neurotoxins: Biology, Pharmacology, and Toxicology. Pharmacol. Rev. 69 (2), 200–235. https://doi.org/10.1124/pr.116.012658.
Press, W.H., Teukolsky, S.A., Vettering, W.T., Flannery, B.P., 2007. NUMERICAL RECIPES The Art of Scientific Computing Third Edition, CAMBRIDGE UNIVERSITY PRESS. https://doi.org/10.1017/CBO9781107415324.004.
Rambo, R.P., n.d. BIOISIS [WWW Document]. URL https://www.bioisis.net/ (accessed 1.5.22).
Ratts, R., Trujillo, C., Bharti, A., vanderSpek, J., Harrison, R., Murphy, J.R., 2005. A conserved motif in transmembrane helix 1 of diphtheria toxin mediates catalytic domain delivery to the cytosol. Proc. Natl. Acad. Sci. U. S. A. 102 (43), 15635–15640. https://doi.org/10.1073/pnas.0504937102.
Rossetto, O., Pirazzini, M., Montecucco, C., 2015. Current gaps in basic science knowledge of botulinum neurotoxin biological actions. Toxicon 107, 59–63. https://doi.org/10.1016/j.toxicon.2015.07.002.
Rummel, A., 2015. The long journey of botulinum neurotoxins into the synapse. Toxicon 107, 9–24. https://doi.org/10.1016/j.toxicon.2015.09.009.
Sarachan, K.L., Curtis, J.E., Krueger, S., 2013. Small-angle scattering contrast calculator for protein and nucleic acid complexes in solution. J. Appl. Crystallogr. 46 (6), 1889–1893. https://doi.org/10.1107/S0021889813025727?10.1107/S0021889813025727/fs5057sup1.pdf.
Schrödinger, LLC, 2015. The {PyMOL} Molecular Graphics System, Version~1.8.
Skjærven, L., Reuter, N., Martinez, A., 2011. Dynamics, flexibility and ligand-induced conformational changes in biological macromolecules: A computational approach. Future Med. Chem. 3 (16), 2079–2100. https://doi.org/10.4155/fmc.11.159.
Suan Li, M., Khanh Mai, B., 2013. Steered Molecular Dynamics-A Promising Tool for Drug Design. Curr. Bioinform. 7 https://doi.org/10.2174/157489312803901009.
Tollinger, M., Crowhurst, K.A., Kay, L.E., Forman-Kay, J.D., 2003. Site-specific contributions to the pH dependence of protein stability. Proc. Natl. Acad. Sci. U. S. A. 100 (8), 4545–4550. https://doi.org/10.1073/pnas.0736600100.
Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H.J.C., 2005. GROMACS: Fast, flexible, and free. J. Comput. Chem. 26 (16), 1701–1718. https://doi.org/10.1002/jcc.20291.
Wang, F., Wan, H., Hu, J.-P., Chang, S., 2015. Molecular dynamics simulations of wild type and mutants of botulinum neurotoxin A complexed with synaptic vesicle protein 2C. Mol. BioSyst. 11 (1), 223–231. https://doi.org/10.1039/C4MB00383G.
Wang, W., 2021. Recent advances in atomic molecular dynamics simulation of intrinsically disordered proteins. PCCP 23 (2), 777–784.
Watson, M.C., Curtis, J.E., 2013. Rapid and accurate calculation of small-angle scattering profiles using the golden ratio. J. Appl. Crystallogr. 46 (4), 1171–1177. https://doi.org/10.1107/S002188981301666X.