Syed Lal Badshah*, Abdul Naeem Khan, Yahia Nasser Mabkhot

Molecular Dynamics Simulation of Cholera Toxin A-1 Polypeptide

1 Introduction

Cholera toxin (CT) belongs to a family of medically important bacterial toxin proteins called the AB5 toxins [1]. This toxin is responsible for the severe diarrhea, causing deaths of hundreds of thousands of people in the developing world [1]. The cholera toxin is a heterohexameric protein consisting of one A-subunit and five B-subunits (AB5) [2,3]. Each B-subunit contains 103 amino acid residues that are spread among two α-helices and ten β-strands [2,4,5]. Six of the β-strands form two sets of three-stranded antiparallel sheet [2]. The overall fold consists of six antiparallel β-strands forming a closed β-barrel, capped by an α-helix between the fourth and fifth strands [2]. The most striking architectural feature of the AB5 toxin is the presence of a central pore or channel that runs along the 5-fold axis of the B5 assembly, forming a ring-like structure [2,6]. The B–subunit is the binding domain; it binds to monosialotetrahexosyl ganglioside (GM1) of the plasma membrane of the host cell [7-9]. The A-subunit (CT-A, MW ~ 27,400) is the enzymatic domain [10,11], it contains 240 amino acid residues. The first 132 amino acids formed a compact globular unit composed of a mixture of α-helices and β-strands [10,12,13]. The catalysis is thought to occur in well define fissure on the free surface of A-subunit [10,12,14].

The CT and related AB5 toxins of other bacteria move from the plasma membrane through trans-Golgi and endoplasmic reticulum (ER) to the cytosol of host cells [6,15,16]. In the endoplasmic reticulum the A-subunit breaks into A1 and A2 chains by the action of protein disulfide isomerase (PDI) [14]. This enzyme breaks the disulfide bond between amino acid residues 187 and 199, splitting the A-subunit into A1 and A2 chains. The A1 chain of CT unfolds inside the endoplasmic reticulum and then enters the cytosol by hijacking the cellular machinery that enables misfolded proteins to cross the membrane for degradation by the proteasome, a process termed retro-translocation [14]. In the cytosol the partially unfolded CTA1 has limited in vivo activity and is prone to ubiquitin independent proteasomal degradation, so CTA1 binds to ARF6 (cytosolic protein) forming cholera toxin A1...
subunit Adenosine diphosphoribosyl factor 6 Guanosine triphosphate (CTA1-ARF6-GTP) complex for refolding and stability[11,17-19].

CTA1 is an Adenosine diphosphate (ADP)-ribosyltransferase, catalyzing the covalent transfer of an ADP-ribose moiety from nicotinamide adenine dinucleotide (NAD+) to arginine (ARG) 201 of the signalling protein, the stimulatory G protein a-subunit (Gsα) [11]. These biochemical changes generates a cascade of events that culminated into huge efflux of salts and water from the epithelial cells into the intestinal lumen, resulting in watery diarrhoea [11,20].

Currently more than 10 well resolved crystal structures of CT are available [2,11,21,22] and they offer a base for designing potent inhibitors against this toxin through a structure-based approach by using molecular modeling and dynamics simulation [22,23]. The mode of action of CT showed several important steps that can be exploited for drug design [22]. These include the blocking of the assembly of holotoxin, preventing receptor-recognition process and directly targeting the enzymatic active site in the folding, unfolding states [10-12,19]. Cholera toxin B-subunit is being used as a drug carrier and is also one of the medically important bacterial toxins and requires extra attention to explore its structural properties [16,24,25]. The aim and objective of this study is to gain an understanding into the structure of the CTA1 subunit, by performing MD simulations. The atomic structure will be clear in the presence and absence of its association with ARF6 complex at different temperatures. This study will enable one to utilize the current data for designing structural based drug inhibitors of this enzyme. Similarly, this work will further enhance the validity of the previously published experimental models [16,23] that utilized an array of biophysical and physiochemical methods and a variety of experimental conditions. This work will also be a stimulus for the beginning of MD simulation based docking, as the static structures of molecules are relatively less informative as compared to the dynamic structure of the protein - ligand system.

2 Methodology

Protein data bank (PDB) entry number 2A5F file was used for atomic coordinates file [11]. In order to obtain only the cholera toxin A1 polypeptide, the PDB file was refined by removing unnecessary molecules (water of crystallization, ions of buffer solution etc.). The CTA1 contains 185 amino acid residues in the PDB file. All MD simulations were performed with GROMACS (Groningen machine for molecular dynamics simulations) software package [26,27]. The topology file was generated from the PDB file through pdb2gmx program of GROMACS. The optimized potential liquid simulation all-atom (OPLS/AA) force field [28] was used for both the systems under study. The missing atoms were added through WHATIF server [29]. In order to provide a cell milieu to the protein, we placed apo-CTA1 in center of a cubic box using genbox program of GROMACS. The box has a dimension of 60 x 60 x 60 Å³. The distance between the wall of the box and protein molecule was 0.9 nm apart so that outside the box all types of influences were ignored. The cubic box containing CTA1 polypeptide was solvated using periodic boundary condition (PBC) with 11700 simple point charge (SPC) water molecules [30], so that CTA1 behaves as if it is inside the cell. The desired number of sodium ions were added to counter balance the system, so that the electro neutrality of the system is maintained.

Energy minimization was made to reduce the steric interactions between protein and solvent molecules, with the number of steps being performed 10,000. The Steepest decent algorithm [31] was used for the energy minimization of the MD simulation systems. The step size (emstep) was equal to 0.01 nm and tolerance (emtol) of 10.0 kJmol⁻¹nm⁻¹ was used. The frequency used to update the neighbour list (nstlist) was taken as 10, and the neighbor-searching type (ns_type) used was grid. The cutoff distance used for the short-range neighbour list (reallist) was taken as 1.0nm. The electrostatic interaction in the systems was measured with particle mesh Ewald (PME) [32]. The Fourier spacing of 0.12 nm and cubic interpolation order (pme_order) were applied. The comparative strength of the Ewald-shifted directed potential at the cut-off was given by ewald_rtol.
of 1e-5. The Van der Waal type (vdwtype) was twin range cut-off and distance for the Lennard Jones cut-off was 1.4 nm. In order to relax and avoid unnecessary distortion of protein molecules the system was equilibrated for 200 ps. The GROMACS molecular dynamics simulation engine “mdrun” program was used to carry out equilibration MD simulation. The electrostatic interactions of periodic box were determined by PME. LINEAR Constraint Solver (LINCS) algorithm was utilized to restrain all the bondings lengths [33]. The velocities were written after every 250 steps and coordinates to trajectory file after each 1000 steps. The energy of the system was written after every 10 steps under imposed periodic boundary conditions. The temperature and pressure were kept constant by separately coupling the CTA1 polypeptide and non-protein to external conditions using the Berendsen temperature-coupling scheme [30] with coupling constant, τ = 0.1 and 20 ps individually. The reference temperature was adjusted to 298 K and reference pressure to 1.0 bar. The velocities were also generated at the above mentioned temperature. After equilibration of the system, the actual molecular dynamics simulation production was performed at 283 K, 310 K, and 323 K for one nanosecond. LINCS algorithm was utilized also in actual MD simulation. The temperature and pressure of the system were kept constant by separately coupling the protein and non-protein to external temperature and pressure baths as done by others [34], having a coupling constants of τ = 1.0 and 20 ps while the reference temperature was adjusted to the temperature of respective MD simulation. The velocities were also generated at the respective simulation temperature. The velocities, coordinates and energies were written after every 1000 steps. The periodic boundary conditions were imposed on the system and long-range interactions were calculated through PME. Different structural properties were evaluated with various Gromacs programs. The simulation trajectories were examined with Visualization Molecular Dynamics (VMD) software package 1.8.6 [35]. The secondary structure stability of CTA1 was calculated through structural identification (STRIDE) of protein algorithm [36-38] of VMD.

3 Results and Discussion

3.1 Total energies of the system

The energetically favorable system represents a system that depicted the dynamic and thermodynamic behavior of a normal protein as persisted in nature. Figure 2 showed the variation in total energy with respect to simulation time for CTA1 at 283, 310 and 323 K respectively. A decrease in the total energy occurs at the beginning of the simulation and is continued up to 100 picoseconds (ps) time span, then the total energy of cholera toxin A1 polypeptide is almost constant for the rest of the simulation time. The mean total energy was found to be \(-4.47 \times 10^5\) kJ mol\(^{-1}\) at 283 K, and at physiological temperature, the mean value of total energy was \(-4.19 \times 10^5\) kJ mol\(^{-1}\). In simulation performed at physiological temperature an energy increase from \(-4.3\) to \(-4.2 \times 10^5\) kJ mol\(^{-1}\) occurred from the start, which then became constant, however a small fluctuation was observed throughout the simulation time of 1000 ps. At 323 K, a shift in total energy was observed at the beginning and the mean total energy of \(-4.075 \times 10^5\) kJ mol\(^{-1}\) was recorded, and remained almost constant with minor variations as shown in Figure 2. These results also revealed that with a rise in simulation temperature the total energy of the protein in the box is found to increase, which causes the simulated system to become unstable and provided proof that with an increase in temperature the cholera toxin A1 polypeptide became thermally unstable. For every 10-digit rise in temperature the total energy rise was about \(-0.1 \times 10^5\) kJ mol\(^{-1}\). It has been confirmed from the above mentioned figure that the total energy values remained negative and CTA1 was thermally stable in the temperature range of 283 to 323 K.

3.2 Root mean square deviation

One of the validity tests of standard MD studies is the determination of root mean square deviation (RMSD) [39].

![Figure 2: Total energy fluctuations versus simulation time for CTA1 at 283 K (black), 310 K (red) and 323 K (green).](image-url)
We also performed this obligatory analysis to determine the stability of the molecular dynamics simulation trajectories. The time variable α-Carbon (Cα) RMSD of CTA1 was calculated at the respective temperatures, taking the energetically unminimized crystal structures as a reference. Figure 3 shows a sudden jump of RMSD values at the very beginning and then making a plateau, thus convergence was observed. However, the results of Cα RMSD of CTA1 at three different temperatures have very small convergence, but the fluctuations were very prominent as shown in Figure 3. The MD simulation of CTA1 at 283 K showed Cα RMSD of 0.0267 nm at the beginning and then rose to 0.031 nm, subsequently after this a continuous fluctuation was present throughout the simulation time. The average Cα RMSD for this system was calculated to be 0.0308 nm at 283 K (Figure 3). Almost similar Cα RMSD results were obtained for MD simulation at 310 and 323 K with the same degree of fluctuations as in the case of 283 K. These high flexibilities in Cα RMSD of CTA1 might be due to the partially disordered secondary structure of cholera toxin A-1 polypeptide. These results were in agreement with the past analysis of MD simulation on various proteins [40].

3.3 Root mean square fluctuation (RMSF)

The RMSF analysis is the most productive and fruitful character of MD simulation for enzymes and other biochemical compounds that are harmful to human life in one aspect or another. It is because of these observations that we determined the fluctuations of different residues of the compounds under study. In the literature, there are two opposite views of designing the inhibitory drugs for toxins, one is by targeting either the most rigid residues and secondly those that are highly flexible ones. But both the mobile and immobile residues and regions of an enzyme have specific functional roles in different biological processes. For example molecular recognition and catalytic properties are due to the mobility possessed by definite and specific residues while the rigidity of residues produced different structures in beta-folds that are indispensable for substrate interaction and catalytic action [41]. The MD study performed by Grottesi and Sansom [42] on the venom of scorpion showed that the most immobilized residues, that have structural and functional role in the toxin, could be exploited for designing inhibitory drugs, while an opposite view has been given by Caceres et al. [43]. Based on MD investigations, they utilized the most flexible regions and loops as a substrate for designing the inhibitory drugs which act as ligands and could be readily accommodated inside those flexible pockets and regions. Similarly, the different types of interactions between the enzyme residues and drugs are possible as the mobile residues have the ability to reduce the steric clashes and could be more receptive conformationally.

In order to know the flexible and rigid behavior of residues of the toxin in a dynamic state, residue wise Cα RMSF were computed at three different temperatures. In MD simulation of CTA1 at 283 K, isoleucine 83 showed the most rigid behavior with RMSF of 0.0129 nm, as shown in Figure 4. Similarly isoleucine 14, 122; leucine 39, valine 58, aspargine 140, tyrosine 126 and tryptophan 172 were also inflexible. On the other hand, glycine 184 was the most fluctuating having Cα-RMSF of 0.0276 nm and aspargine 1 has CαRMSF of 0.025 nm, the reason was that these residues are present at the C and N-termini having no interaction with other residues. Similarly, glycine 77, alanine 101, 157 and proline 116 are also highly flexible. The average Cα RMSF of CTA1 polypeptide at 283 K was 0.0175 nm. At the physiological temperature, the CTA1 residues fluctuate differently. Glutamine 121 was the most rigid throughout the simulation having Cα RMSF of 0.0129 nm, while valine 85 and 111, tyrosine 28, 40, 126 and 148; aspartic acid 57, isoleucine 178, aspargine 140 and histidine 138 are also rigid. The high fluctuating residue is glycine 184 at the carboxyl terminus having CaRMSF of 0.028 nm. The other residues are glycine 18, 49, 71, 77, 116 and alanine 101 are also highly flexible. The average Cα RMSF at 310 K was 0.0175 nm and at 323 K aspargine 41 was highly conserved with RMSF of 0.0125 nm. Similarly valine 58, leucine 3 and 99, glutamine 121, isoleucine 122, and histidine 179 are also rigid as shown in Figure 4.
the highly flexible residues were glycine 18, 49, 71, 116 and 131 but glycine 184 has the highest Cα RMSF of 0.0259 nm. The average RMSF at 323 K was found to be 0.0175 nm. Based on these RMSF results, we can suggest to use the molecular modeling and docking methods to explore the whole enzymatic portion of CTA1 for possible binding pockets. These molecular docking studies will be helpful in designing drugs that can bind to the active sites and other pockets in CTA1. Based on the available methods of molecular docking, both rigid and flexible residues need to be targeted to discover suitable small molecule ligands that acts as inhibitors of CTA1 enzymatic activity.

3.4 Radius of gyration

In order to assess the folding and unfolding pathway of CTA1, we calculated the radius of gyration (Rg) at different temperatures. The result varies for CTA1 in the dynamic phase for 1000 ps. A continuous fluctuation in Rg values at 283 K was observed, which exhibited the folding, unfolding of protein as the time passed as shown in Figure 5. The mean Rg value was found to be 1.734 nm, with a minimum at 1.732 nm at 740 ps, while the maximum was 1.736 nm at 180 ps and at 550 ps. The MD simulation run at physiological temperature showed the mean Rg value of 1.734 nm. In the current study the highest value recorded at this temperature was 1.736 nm near 170 ps while the smallest one was 1.717 nm at 500 ps, similarly, the mean Rg value of CTA1 was observed to be 1.734 nm at 310 K (Figure 5). The radius of gyration results of CTA1 simulation at 323 K looked different from the first two-simulations of 283 and 310 K. One can observe a continuous fluctuation in the radius of gyration values but not so prominent when compared to the earlier two MD simulations at 283 and 310 K (Figure 5). Both the minimum and maximum Rg values of 1.742 at 190ps and 1.737 nm at 810ps and 990ps were observed respectively. Thus, at high temperatures the cholera toxin A1 polypeptide was expanded due to increased motion of different groups of the polypeptide, as a result it follows a stable pathway for its unfolding.

3.5 Solvent accessible surface area (SASA)

The solvent accessible surface area of CTA1 polypeptide was considered in order to find the role of different hydrophobic and hydrophilic forces and residues exposed to the solvent at different temperatures. The hydrophobic residues present at the end of the polypeptide chain are proline and tryptophan but the CTA1 is also strongly hydrophilic which has been proved from water of crystallization present in the crystal structure [7, 11]. The cholera toxin simulation at 283, 310 and 323 K showed a mean total SASA of 111 nm$^2$, which was exposed to the solvent in the system. The hydrophilic SASA was greater than the hydrophobic SASA as shown in the supplementary figures 1 A-C. The mean values of hydrophilic and hydrophobic SASA were observed to be 58.5 nm$^2$ and 52.5 nm$^2$ respectively for the CTA1; thus, there was a difference of 6 nm$^2$. The SASA calculated for CTA1 was not only equal, but nearly remained constant with a fluctuation of about 2 nm$^2$ at all the simulation temperatures during the time span. Thus, this small fluctuation can be assigned to the folding and unfolding of toxin, which has been confirmed from the radius of gyration studies. It was further confirmed from SASA studies that CTA1 exhibits nearly 50% hydrophobic and 50% hydrophilic SASA. Due to this partial hydrophobic nature, it can easily move through the plasma membrane of the cell and
membrane of the endoplasmic reticulum while the partial hydrophilic nature provides it the opportunity to make interaction in the hydrophilic cytosol.

According to Ampapathi et al. [16] cholera toxin is hydrophobic in nature. They divided the CTA1 polypeptide into three regions as CTA1α, CTA1β, and CTA1γ, where the 1st region is globular, 2nd forms an extended bridge while a 3rd is highly hydrophobic due to the hydrophobic amino acids present in these regions. Contrary to all this, Zhang et al. [2,3] and O’Neal et al. [11] obtained the crystals of cholera toxin with X-ray diffraction analysis which confirmed its hydrophilic nature because most of the crystals retained up to eighty water of crystallization molecules and depicted its hydrophilic nature. Our results are consistent with those of Zhang et al. [2,3] and O’Neal et al. [11] but are opposite to those of Ampapathi et al. [16]. From the above discussion, it is concluded that cholera toxin can spread through aerosols and it is a potent and easy to carry biological weapon. Therefore in order to prevent its misuse by rogue countries and terrorist groups as a biological weapon, inhibitory drugs and methods should be developed for its control.

3.6 Secondary structure analysis

The secondary structure of cholera toxin is of great interest because of its unusual entry mechanism of translocation inside the cell cytosol [44-46]. It utilizes the retro-translocation mechanism and deceives the proteolytic enzymes of endoplasmic reticulum and also uses the process of unfolding and refolding mechanism [15,17,20]. The secondary structure of CTA1 was calculated through the timeline plugin module of VMD having structural identification (STRIDE) of protein algorithm. The STRIDE algorithm computes the secondary structure of the given protein as a function of time based on the changes in the hydrogen bond analysis, main chain dihedral angles and hydrogen bond distances of amino acid residues. Supplementary figure 2 A-C showed strips representing the secondary structure of CTA1 at respective temperatures as a function of time. The secondary structures contain the regular conformation units of α-helices, 3-10 helices, β-sheets and β-turns while the irregular conformations are loops or coils. In the CTA1 simulation at 283 K, the secondary structure from the N-terminal, aspartic acid one made a coil continuously for 1000 ps time span. Lysine 2 showed a transition of β-sheet and coil, but the β-sheet form was dominant. The amino acid residues from leucine 3 up to aspartic acid 7 form a continuous extended bridge of β-sheet for 1000 ps simulation time. The three residues in CTA1 sequence from serine 8 to proline 10 were in irregular conformation of the coil form. Proline 11 to serine 17 was in a magenta color strip, that represented the α-helix at certain positions, transition in this secondary structure to β-turn occurred which were denoted by green patches in supplementary figure 2-A. Serine 18, which was dominant in coil form, transferred several times to β-turn form during the course of simulation time. After this, glycine 19 and leucine 20 were in β-sheet. Methionine 21 to proline 22 is in coil structure, but transition has occurred to β-turn due to methionine. Then arginine 23 to arginine 31 was in continuous α-helix. But glycine 32 to methionine 35 again was in coil conformation. The transition of tan color to white color that was representative of the transition of the isolated bridge form to the coil form was present throughout the simulation time span. Isoleucine 377 and asparagine 38 were in coil secondary structure continuously. For the third time α-helix was formed by residues from leucine 39 to arginine 44. Glycine 45 again forms a coil continuously. A continuous β-turn was formed by threonine 46 to glycine 49. A coil was formed by phenylalanine 50 and valine 51. The β-turn represented by green color was formed by arginine 52 to aspartic acid 54. The transition of β-turn into isolated bridge and vice versa, was formed by aspartic acid 55. Glycine 56 was a part of a coil. β-Sheet was then formed by tyrosine 57 and valine 58 continuously throughout the simulation. Serine 59, threonine 60 presented a transition of coil and β-sheets several times. Serine 61 formed the isolated bridge, but transition to β-sheet occurs during the simulation. Coil form was again observed due to isoleucine 62 and serine 63. The transition of isolated bridges and β-turn into one another was produced by threonine 73 to leucine 75, but leucine 64 and leucine 75 were in isolated bridge form of secondary structure. The β-turn was formed again by serine 76 to serine 79. Threonine 80 to isoleucine 86 was in the β-sheet form but threonine 80 transfers to isolated bridge form for certain time periods of simulation. Alanine 87 and threonine 88-formed coil but alanine 87 has a transition to β-sheet. The β-turn was shown by residues of alanine 89 to asparagine 91. Methionine 92 to asparagine 94 formed β-sheet while valine 95 to leucine 99 formed an α-helix. Glycine 100 to tyrosine 102 showed a transition of 3-10 helix shown by blue color band and its transition to β-turn several times during the simulation as reflected in supplemental figure 2-A. Serine 103 to histidine 5 formed β-turn. Proline 106 to aspartic acid 108 also showed transition from 3-10 helix to β-turn. Glutamine 109 produced a transition of coil and β-turn during the simulation. Aspartic acid 110 had transition of β-sheet and coil form of secondary structure several times.
during the simulation. The β-sheet was again represented by valine 111 to leucine 114. The two glycine residues of 115 and 116 formed β-turn. Then again isoleucine 117 and proline 118 formed β-sheet continuously. Another transition of secondary structure of β-turn into 3-10 helix and vice versa was shown by tyrosine 119 to glutamine 121. Isoleucine 122 to histidine 129 formed the β-sheet but transition to coil was shown by arginine 127 and to the isolated bridge by valine 128 and histidine 129 as observed in supplementary figure 2-A. An uninterrupted β-turn was produced by phenylalanine 130 and glycine 131. Valine 132 was switched to β-sheet and isolated bridge. The leucine 132 residue also showed similar behavior as above but coil form replaced the isolated bridge structure. Aspartic acid 134 to leucine 137 were in coil form but for a small time it shifted to β-sheet. The β-sheet was again formed by histidine 138 and arginine 139, and then β-turn was formed by asparagine 140 to tyrosine 143. Arginine 144 formed a coil but turned into β-turn. Another alteration of 3-10 helix into β-turn was produced by aspartic acid 145 to tyrosine 147. Next β-turn was represented by tyrosine 148 to leucine 151. Residues from aspartic acid 152 to proline 155 were in coil form. The 3-10 helix was regularly presented by alanine 156 to aspartic acid 158. Glycine 159 to glycine 164 were mostly in β-turn form, but changed to coil during the course of simulation by glycine 159 and alanine 163 and changed to 3-10 helix form and coil form by tyrosine 160 to leucine 162. A complete coil was shown by phenylalanine 165 and a full β-turn was presented by proline 166 to histidine 169. Another evolution of secondary structure of 3-10 helix to coil form was reflected in the supplementary figure 2-A by arginine 170. Alanine 171 to arginine 173 formed 3-10 helix but shifted to β-turn. Another β-turn was formed by glutamic acid 174 to proline 176. Tryptophan 177 to histidine 179 also reflects a transition of 3-10 helix into β-turn, however histidine 180 also showed coils form for several simulation times. The last conversion was produced in secondary structure of CTA1 by alanine 181 of coil form and β-turn while proline 182 to serine 185 were continuously in the coil form. From the results it was clear that the secondary structure graph produced through STRIDE algorithm showed several transitions of the secondary structure at the 283 K simulation. Further regular units were interrupted by irregular confirmation of secondary structure in this simulated system. The simulation performed at 310 K and 320 K for CTA1 (Supplementary figure 2-B, C) showed almost the same amount of secondary structure but the time and amount of transition were different. From these results we concluded that CTA1 has partially disordered secondary structure and can be placed in the intrinsically disordered (ID) protein class. The behavior of cholera toxin A1 polypeptide of regular units of helices and β-sheets interrupted by coils were in close agreement to the work of Ampapathi et al. [16] who observed the secondary structure through circular dichroism spectroscopy and 2D nuclear magnetic resonance spectroscopy and that of others [13-16].

4 Conclusions

From the molecular dynamics simulation study of cholera toxin A1 polypeptide at various temperatures, different properties were analyzed. The total energy analysis showed that CTA1 is stable thermodynamically and structurally. The Cα RMSF studies revealed a number of Cα residues, which can be used as target for designing inhibitory drugs. The CTA1 toxin has hydrophilic character and possessed a lethal threat as it can easily be dissolved in aerosols. Therefore, it is important to design and synthesize inhibitory drugs for cholera toxin. We concluded from the radius of gyration and hydrogen bond analysis that unfolding and re-folding with the passage of time is normal in its structure. The secondary structure analysis confirmed that cholera toxin A1 polypeptide has a partially disordered secondary structure. This validates the supposition of cholera toxin A1 polypeptide safe retro-translocation inside the endoplasmic reticulum. Thus it easily deceives the proteolytic enzymes of endoplasmic reticulum due to its partially disordered secondary structure and through the mechanism of unfolding.

Acknowledgement: The authors extend their sincere appreciation to the Deanship of Scientific Research at the King Saud University for its funding of this Prolific Research Group (PRG-1438-29).

References

[1] Willey J. M., Sherwood L. M., Woolverton C.J., Prescott, Harley and Klein’s Microbiology, Seventh Ed., McGraw-Hill, New York, 2008.
[2] Zhang R. G., Westbrook M. L., Westbrook E. M., Scott D. L., Otwinowski Z., Maulik P. R., Reed R. A., Shipley G. G., The 2.4 Å Crystal Structure of Cholera Toxin B subunit Pentamer: Choleragenoid., J. Mol. Biol. 1995, 251(4), 550–562, doi: 10.1006/jmbi.1995.0455.
[3] Zhang R. G., Scott D. L., Westbrook M. L., Nance S., Spangler B. D., Shipley G. G., Westbrook E. M., The Three-dimensional Crystal Structure of Cholera Toxin, J. Mol. Biol. 1995, 251(4), 563–573, doi: 10.1006/jmbi.1995.0456.
[4] Arakawa T., Harakuni T., Cholera toxin B subunit—Five-stranded α-helical coiled-coil fusion protein: “Five-to-five” molecular chimeras displays robust physicochemical stability, 2014, 32(39), 5019-5026, doi: 10.1016/j.vaccine.2014.07.016.

[5] Miyata T., Oshiro S., Harakuni T., Taira T., Matsuzaki G., Arakawa T., Physicochemically stable cholera toxin B subunit pentamer created by peripheral molecular constraints imposed by de novo-introduced intersubunit disulfide crosslinks, Vaccine 2012, 30(28), 4225–4232, doi: 10.1016/j.vaccine.2012.04.047.

[6] Lencer W. I., Tsai B., The intracellular voyage of cholera toxin: going retro, Trends Biochem. Soc., 2003, 28(12), 639–645, doi: http://dx.doi.org/10.1016/j.tibs.2003.10.002.

[7] Blessy J. I., Sharmila D. J. S., Molecular simulation of N-acetylneuraminic acid analogues and molecular dynamics studies of cholera toxin-Neu5Gc complex, J. Biomol. Struct. Dyn., 2014, 33(5), 1126–1139, doi: 10.1080/07391102.2014.931825.

[8] Blessy J. I., Sharmila D. J. S., Molecular modeling of methyl-alpha-Neu5Ac5 analogues docked against cholera toxin - a molecular dynamics study, Glycoconj. J., 2015, 32(1-2), 49–67, doi: 10.1007/s10719-014-9570-6.

[9] Sharmila D. J. S., Veluraja K., Conformations of Higher Gangliosides and Their Binding with Cholera Toxin—Investigation by Molecular Modeling, Molecular Mechanics, and Molecular Dynamics, J. Biomol. Struct. Dyn., 2006, 23(6), 641–656, doi: 10.1080/07391102.2006.10507089.

[10] Jobling M. G., Holmes R. K., Identification of motifs in cholera toxin A1 polypeptide that are required for its interaction with human ADP-ribosylation factor 6 in a bacterial two-hybrid system, Proc. Natl. Acad. Sci. U. S. A., 2000, 97(26), 14662–14667, doi: 10.1073/pnas.011442598.

[11] O’Neal C. J., Jobling M. G., Holmes R. K., Hol W. G. J., Structural basis for the activation of cholera toxin by human ARF6-GTP, Science, 2005, 309(5737), 1093–1096, doi: 10.1126/science.1113398.

[12] Jobling M. G., Gotow L. F., Yang Z., Holmes R. K., A mutational analysis of residues in cholera toxin A1 necessary for interaction with its substrate, the stimulatory G protein Gαs, Toxins (Basel), 2015, 7(3), 919–35, doi: 10.3390/toxins7030919.

[13] Massesy S., Burrell H., Taylor M., Nemec K. N.; Ray S., Haslam D. B., Teter K., Structural and functional interactions between the cholera toxin A1 subunit and ERdj3/HEDJ, a chaperone of the endoplasmic reticulum. Infect. Immun., 2011, 79(11), 6166–6171, doi: 10.1128/IAI.79.4.6166-6171.2011.

[14] Taylor M., Banerjee T., Ray S., Tatulian S. A., Teter K., Protein-disulfide isomerase displaces the cholera toxin A1 subunit from the holotoxin without unfolding the A1 subunit, J. Biol. Chem. 2011, 286(25), 22090–22100, doi: 10.1074/jbc.M111.273966.

[15] Teter K., Allyn R. L., Jobling M. G., Holmes R. K., Transfer of the cholera toxin A1 polypeptide from the endoplasmic reticulum to the cytosol is a rapid process facilitated by the endoplasmic reticulum-associated degradation pathway, Infect. Immun., 2002, 70(11), 6166–6171, doi: 10.1128/IAI.70.11.6166-6171.2002.

[16] Ampapathi R. S., Creath A. L., Lou D. I., Craft J. W., Blanke S. R., Legge G. B., Order-Disorder-Order Transitions Mediate the Activation of Cholera Toxin, J. Mol. Biol., 2008, 377(3), 748–760, doi: 10.1016/j.jmb.2007.12.075.

[17] Teter K., Jobling M. G., Sentz D., Holmes R. K., The Cholera Toxin A13 Subdomain Is Essential for Interaction with ADP-Ribosylation Factor 6 and Full Toxic Activity but Is Not Required for Translocation from the Endoplasmic Reticulum to the Cytosol, Infect. Immun., 2006, 74(4), 2259–2267, doi: 10.1128/IAI.74.4.2259-2267.2006.

[18] Morinaga N., Kaihou Y., Vitale N., Moss J., Noda, M., Involvement of ADP-ribosylation factor 1 in cholera toxin-induced morphological changes of Chinese hamster ovary cells, J. Biol. Chem., 2001, 276(25), 22838–22843, doi: 10.1074/jbc.M101184200.

[19] Banerjee T., Taylor M., Jobling M. G., Burrell H., Yang Z., Serrano A., Holmes R. K., Tatulian S. A., Teter K., ADP-ribosylation factor 6 acts as an allosteric activator for the folded but not disordered cholera toxin A1 polypeptide, Mol. Microbiol., 2014, 94(4), 898-912, doi: 10.1111/mmi.12807.

[20] Cho J. A., Chinnapen D. J. F., Aamar E., te Welscher Y. M., Lencer W. I., Massol R., Insights on the trafficking and retro-translocation of glycosphingolipid-binding bacterial toxins, Front. Cell. Infect. Microbiol., 2012, 2(51), doi: 10.3389/fcimb.2012.00051.

[21] Merritt E. A., Sarfaty S., van den Akker F., L’Hoir C., Martial J. A., Hol W. G., Crystal structure of cholera toxin B-pentamer bound to receptor GM1 pentasaccharide, Protein Sci., 1994, 3(2), 166–175, doi: 10.1002/pro.5560030202.

[22] Fan E., Merritt E. A., Zhang Z., Pickens J. C., Roach C., Ahn M., Hol W. G. J., Exploration of the GM1 receptor-binding site of heat-labile enterotoxin and cholera toxin by phenylring-containing galactose derivatives, Acta Crystallogr. Sect. D Biol. Crystallogr., 2001, 57(2), 201–212, doi: 10.1107/S0907444900016814.

[23] Sun H., Chen L., Gao L., Fang W., Nanodomains Formation of Ganglioside GM1 in Lipid Membrane: Effects of Cholera Toxin-Mediated Cross-Linking, Langmuir, 2015, 31(33), 9105–9114, doi: 10.1021/acs.langmuir.5b01866.

[24] Craft J. W., Shen T., Brier L. M., Briggs J. M., Biophysical Characteristics of Cholera Toxin and Escherichia coli Heat-Labile Enterotoxin Structure and Chemistry Lead to Differential Toxicity, J. Phys. Chem. B, 2015, 119(3), 1048–1061, doi: 10.1021/jp506509c.

[25] Basu I., Mukhopadhyay C., Insights into Binding of Cholera Toxin to GM1 Containing Membrane, Langmuir, 2014, 30(50), 15244–15252, doi: 10.1021/la5036618.

[26] Lindahl E., Hess B., van der Spoel D., GROMACS 3.0: a package for molecular simulation and trajectory analysis, Mol. Model. Annu., 2001, 7(8), 306–317, doi: 10.1007/s008940001405.

[27] Prosek S., Páll S., Schulz R., Larsson P., Bjelkmar P., Apostolov R., Shirts M. R., Smith, J. C., Kasson P. M., van der Spoel D., Hess B., Lindahl E., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics, 2013, 29(7), 845–854, doi: 10.1093/bioinformatics/btt055.

[28] Jorgensen W. L., Maxwell D. S., Tirado-Rives J., Development and Testing of the OPLS All-Atom Force Field on Conformational Energies and Properties of Organic Liquids, J. Am. Chem. Soc., 1996, 118(45), 11225–11236, doi: 10.1021/ja9621760.

[29] Vriend G., WHAT IF: A molecular modeling and drug design program, J. Mol. Graph., 1990, 8(1), 52–56, doi: 10.1016/0263-7855(90)80070-V.
[30] Berendsen H. J. C., Postma J. P. M., Gunsteren W. F., Hermans J., Intermolecular Forces, (In: B. Pullman (Ed.), Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry (13-16 April, 1981, Jerusalem, Israel), Jerusalem, 1981, 311–342.

[31] Miyamoto S., Kollman P. A., Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models, J. Comput. Chem., 1992, 13(8), 952–962, doi: 10.1002/jcc.540130805.

[32] Darden T., York D., Pedersen L., Particle mesh Ewald: An N⋅log(N) method for Ewald sums in large systems, J. Chem. Phys., 1993, 109(12), 3664-3671, doi: 10.1063/1.464397.

[33] Hess B., Bekker H., Berendsen H. J. C., Fraaije J. G. E. M., LINCS: A linear constraint solver for molecular simulations, J. Comput. Chem., 1997, 18(12), 1463–1472, doi: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

[34] Sachchidanand., Resnick-Silverman L., Yan S., Mutjaba S., Liu W., Zeng L., Manfredi J. J.; Zhou M. M., Target Structure-Based Discovery of Small Molecules that Block Human p53 and CREB Binding Protein Association, Chem. Biol., 2006, 13(1), 81–90, doi:10.1016/j.chembiol.2005.10.014.

[35] Humphrey W., Dalke A., Schulten K., VMD: Visual molecular dynamics. J. Mol. Graph., 1996, 14(1), 33–38, doi: http://dx.doi.org/10.1016/0263-7855(96)00018-5.

[36] Frishman D., Argos P., Knowledge-based protein secondary structure assignment, Proteins, 1995, 23(4), 566-579, doi: 10.1002/prot.340230412.

[37] Heinig M., Frishman D., STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins, Nucleic Acids Res., 2004, 32(2), 500–502, doi: 10.1093/nar/gkh429.

[38] Frishman D., Argos P., The future of protein secondary structure prediction accuracy,Fold Des., 1997, 2(3), 159-162, doi: 10.1016/S1359-0287(97)00022-9.

[39] Murdoch S. E., Tai K., Ng M. H., Johnston S., Wu B., Fangohr H., Laughton C. A., Essex J. W., Sansom M. S. P., Quality Assurance for Biomolecular Simulations, J. Chem. Theory Comput., 2006, 2(6), 1477–1481, doi: 10.1021/ct6001708.

[40] Ogawa H., Nakano M., Watanabe H., Starikov E. B., Rothstein S. M., Tanaka S., Molecular dynamics simulation study on the structural stabilities of polyglutamine peptides, Comput. Biol. Chem., 2008, 32(2), 102–110, doi: http://dx.doi.org/10.1016/j.compbiolchem.2007.11.001.

[41] Sun T. G., Hu J. P., Li C. H., Chen W. Z., Wang C. X., A molecular dynamics simulation study of glutamine-binding protein. J. Mol. Struct. THEOCHEM, 2005, 725(1-3), 9–16, doi: http://dx.doi.org/10.1016/j.theochem.2005.02.056.

[42] Grottesi A., Sansom M. S. P., Montal M. Molecular dynamics simulations of a K+ channel blocker: Tct toxin from Tityus cambridgei, FEBS Lett., 2003, 535(1-3), 29–33, doi: 10.1016/S0014-5793(02)03849-8.

[43] Caceres R. A., Saraiva Timmers L. F., Dias R., Basso L. A., Santos D. S., de Azavedo Jr. W. F., Molecular modeling and dynamics simulations of PNP from Streptococcus agalactiae, Bioorg. Med. Chem., 2008, 16(9), 4984–4993, doi: http://dx.doi.org/10.1016/j.bmc.2008.03.044.

[44] Pande A. H., Scaglione P., Taylor M., Nemec K. N.; Tuthill S., Moe D., Holmes R. K., Tatulian S. A., Teter K., Conformational Instability of the Cholera Toxin A1 Polypeptide, J. Mol. Biol., 2007, 374(4), 1114–1128, doi: 10.1016/j.jmb.2007.10.025.

[45] Banerjee T., Pande A., Jobling M. G., Taylor M., Massey S., Holmes R. K.; Tatulian S. A.; Teter K. Contribution of subdomain structure to the thermal stability of the cholera toxin A1 subunit, Biochemistry, 2010, 49(41), 8839–8846, doi: 10.1021/bi101201c.

[46] Massey S., Banerjee T., Pande A. H., Taylor M., Tatulian S. A., Teter K., Stabilization of the Tertiary Structure of the Cholera Toxin A1 Subunit Inhibits Toxin Dislocation and Cellular Intoxication, J. Mol. Biol., 2009, 393(5), 1083–1096, doi: 10.1016/j.jmb.2009.09.013.

[47] Supplemental Material: The online version of this article (DOI 10.1515/chem-2016-0021) offers supplementary material.