Summary of Ph.D. Thesis

Structural investigation of peptaibols using accelerated molecular dynamics simulations

Chetna Tyagi

Supervisors:

Dr. Ferenc Ötvös
Institute of Biochemistry,
Biological Research Centre

Dr. László Kredics
Department of Microbiology,
TTIK, University of Szeged

Doctoral School of Biology

Department of Microbiology

Faculty of Science and Informatics

University of Szeged

2020
INTRODUCTION
The present study deals with bioactive peptaibols, produced by a specific group of fungi, the genus *Trichoderma*, which are known for antagonistic behavior against naturally competing fungi. Peptaibols are linear, non-ribosomally produced amphipathic polypeptides that comprise of non-standard residues like α-aminoisobutyric acid (Aib), hydroxyproline (Hyp) or D-isovaline (Div), an acetylated N-terminus and an aminoalcohol present at the C-terminus. Many peptaibols have been shown to form voltage-gated ion channels or pores within biological membranes via an aggregation step. The increasing gap between known peptaibol sequences and their three-dimensional structures can be reduced using computational modeling and molecular dynamics (MD) simulations.

MD simulations are carried out to recreate and visualize various biomolecular processes with the help of potentials formulated on the basis of physical laws that govern all chemical entities. It calculates the time evolution of properties of the macro-molecular system called as a trajectory. But classical MD simulation approaches fall short of the time scale required to obtain slow conformational transitions separated by high energy barriers. In order to enhance the sampling of conformational landscape of a biological system, many methods have been developed that follow the scheme of modifying the Hamiltonian by adding a bias potential, for example, accelerated MD (aMD) simulations. They work by adding a bias potential \( \Delta V(r) \) to the true potential and modifying the potential energy surface \( V(r) \) in such a way that the surfaces near the energy minima are raised but those
near the barriers remain unaffected. The technique promotes sampling of infrequent events of biomolecular systems without any prior knowledge of the location of energy wells or barriers. The aMD implementation includes three kinds of boost to potential, a) boost to total potential of the system, b) boost to dihedral potential of the biological system, and c) dual boost combining the previous two. We optimized the use of aMD simulations and its various parameters to accurately elucidate the complete conformational ensemble of peptaibols.

AIMS AND OBJECTIVES

The understanding of peptaibol structure and dynamic folding is important to subsequently understand and correlate with their antimicrobial mechanism of action. Various structural differences account for the variability in their ability to interact with cell membranes and therefore, their antimicrobial activity. Our main aim is to develop a reliable strategy for elucidating complete structural ensembles of peptaibol compounds by applying time- and cost-effective in silico techniques known as molecular dynamics simulations. The various steps taken in lieu of the main goal are as follows:

1) To develop an accurate representation of various non-standard amino acid residues like Aib and Div found in peptaibols, and the C-terminal amino alcohols like Pheol or Leuol in terms of their atomic partial charges, geometry and
ability to incorporate to peptide sequences for computational modeling.

2) To test the effect of various solvents and timescales on peptaibol folding using classical MD methods.

3) To optimize accelerated MD simulation parameters required to obtain complete conformational landscapes of peptaibols by comparing with a known peptaibol structure.

4) To test accelerated MD simulations for correctly modeling biological membranes and to reproduce experimental results.

MATERIALS AND METHODOLOGY

Sequence selection
We carried out classical MD simulations on two sequences: Trikoningin KA V (TRK-V) and Tripleurin XIIc (TPN XIIc). aMD simulations were carried out on Alamethicin F30/3 and later on TPN XIIc, Paracelsins B & H, Brevicelsins I & IV and TRK-V.

Force field library generation for non-standard residues
For calculation of their partial charges and creating force field libraries, the R.E.D server was used. RESP (restrained electrostatic potential) was used to calculate the charges with a HF/6-311G(d) basis set and Gaussian09 as quantum mechanical program interface. For each residue, two conformations, i.e. α-helix and β-sheet were used.
Classical molecular dynamics simulations

**In implicit water solvent:** all MD calculations were carried out with AmberTools16 with ff14SB force field using generalized born implicit solvent method. The first step is energy minimization. The maximum number of cycles was set at 10000 (maxcyc) with a convergence criterion of 0.01. The steepest descent algorithm was used for the first 100 cycles (ncyc) and then switched to conjugate-gradient algorithm for the remaining cycles. The energy minimization outputs were used for setting up the production run with 100 ns of total simulation time. All systems were maintained at 300 K using Langevin thermostat (ntt = 3, gamma_ln =1.0). The time step was set to 2 fs and no cutoff was applied for non-bonding interactions.

**In explicit water and methanol solvents:** The average structure of the last 30000 steps from the previous 100 ns long simulation was used as a starting structure for two 30 ns long simulations in explicit water and methanol solvent. The system was energy minimized for 20,000 steps and after a 50,000-long heating step, the production run was started at temperature 300 K under pressure regulation, ntp=1 condition. The periodic boundary conditions were applied, long-range effects were treated using the PME method.

**Accelerated molecular dynamics simulations**

All systems were prepared for aMD in six consecutive steps, i.e. (a) minimization of solvent, (b) water movement at 300K under isothermal and isobaric (NTP) conditions, (c) minimization of the whole system, (d) heating, (e) relax the system at 300 K for 0.5 ns, and (f) relax system at 300 K under NTP conditions for 5 ns with no
restraints. The temperature scaling was carried out using Langevin thermostat while the pressure was regulated using the default Berendsen barostat for all corresponding calculations. SHAKE bond length constraints were applied on all bonds involving hydrogen.
All aMD simulations were carried out at 300 K temperature, 2 fs time step, and energies and boost information was written at every 1000 steps. The GPU machines available through the NIIF High Performance Computing supercomputer at the University of Debrecen on the partition prod-gpu-k40-Leo nodes with 3 × Nvidia K40X CUDA8 were utilized for all aMD simulations. All simulations were carried out using pmemd.cuda implementation of Amber14, also available at the cluster. All systems were solvated using TIP3P water model for aqueous solvent except in case of TPN XIIc solvation in chloroform solvent using tleap module of AmberTools18.

**Accelerated molecular dynamics simulations applied on a bilayer membrane system: alamethicin F30/3 hexamer pore**

The hexamer pore of Alm F30/3 peptide was then embedded into a 3:1 mixture of 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE) and 1,2-dioleoyl-sn-glycero-3-[phospho-rac-(1-glycerol)] (DOPG) bilayer membranes which mimics a bacterial (*Escherichia coli*) membrane constitution. This system can be easily prepared in an Amber-ready format by using the ‘packmol-memgen’ workflow available with AmberTools18 that uses ‘Memembed’ to obtain pre- oriented protein conformation with respect to the membrane. Each aMD simulation was carried out with dual boost (iamd=3) option at 300 K temperature regulated using a Langevin thermostat. A weak
external static electric field was also applied along the z direction (across membrane) with efz values (intensity in kcal (mol × A × e)^−1) of 0.180 and 0.080 for the first and second simulation, respectively.

**RESULTS AND DISCUSSION**

**Classical molecular dynamics**

*Trikoniningin KA V in implicit solvent:* The results obtained from Trikoniningin KA V were slightly surprising as many residues like Gln6, Aib7, Aib8, Aib9, Aib12, Val14, Aib15, and Ile16 had a strong preference for the left-handed helix region. It’s not unexpected for the Aib residues but Gln and Val are associated with right-handed α-helix conformations. The remaining amino acid residues showed preference for right-handed helix region. This clash between opposing helical turns makes the resulting three-dimensional structure as an unfolded spiral. The free energy distribution of RMSD and radius of gyration values showed that a highly bent structure was the most energetically stable structure.

*Tripleurin XIIc in implicit solvent:* The presence of three D-isovalines brought a preference for left-handed helical regions while affecting the neighboring Aib residues. The Div11-Aib12-Pro13 region resulted in unwinding of the spiral due to proline while Val9-Ala10 regions show backbone reversing γ’-region geometry. The result indicated the presence of β-turns for most of the sequence with a highly bent backbone.
Tripleurin XIIc in explicit water and methanol solvent: The comparative results in both solvents showed that most residues are populated around the α- and δ-regions. Pro13 lies almost exclusively in PII conformation in both explicit water and methanol solvents. The γ’-turn conformation is not stable in explicit solvent environment especially in methanol. The loss of the Div11→Val9 γ’-turn bond seems to reduce the overall bending of the chain and bring linearity. Overall, we observed the formation of beta-bend ribbon spirals at the N-terminal segment with higher stability in methanol. For C-terminal, water solvent promotes formation of α-helix while methanol solvent promotes formation of 3_{10}-helix.

**Accelerated molecular dynamics**

Tripleurin XIIc in explicit water and chloroform solvent: A 1 µs long simulation carried out in both water and chloroform solvent with a higher dihedral boost revealed a highly folded linear conformation of TPN XIIc. The structure of TPN XIIc was found to be a continuous β-bend ribbon spiral with α/3_{10}-helix at the C-terminal. The spiral shape instead of a regular α/3_{10}-helix can be correlated with presence of three DIV residues.

**Addressing convergence:** The conformational landscape covered in the fourth 1 µs long simulation is a state which couldn’t have been achieved during shorter simulations. This shows that 500 ns are not adequate to attain the higher energy state in chloroform. The peptide seems to be stuck in a single energy state in chloroform while it is highly dynamic in water and jumps through various intermediate
states with relative ease. Irrespective of the starting configuration, all simulations must, at some point, begin to sample the same space which could not be observed in the case of chloroform. But the combined trajectory of 2.5 µs surely indicates adequate sampling.

**Optimization of aMD simulation parameters on Alm F30/3:** The simulation carried out for 1 µs using GPUs with slightly aggressive boost parameters started from the original unfolded conformation which resulted in a close-to-native conformation of AlmF30/3.

**aMD simulation on paracelsins B & H and brevicelsins I & IV:** Based on the reweighted potential of mean force (PMF) values calculated for end-to-end distance, it seems that the replacement of Aib to valine at the 17th position in paracelsin-H and brevicelsin-IV slightly lowers the energy barrier to attain linearity from a bent conformation.

**Trichoningin KA V:** aMD run shows that these residues flanking the central region show shifts from the predominant left-handed helical regions to the right-handed helical region

**Alamethicin hexamer channel all-atom enhanced sampling simulations:** The presence of peptide channel clearly introduced disorder to the membrane as shown by the lipid SCD parameter. The diffusion coefficient (DC) value is higher under a stronger electric field but is much lower than DC of water in bulk. aMD was successfully applied to model peptaibol channel and membrane interaction.
SUMMARY

1. We showed that the use of enhanced sampling methods like aMD is crucial to accurately model the folding dynamics of peptaibols in comparison to classical MD techniques.

2. We show that 1 µs long aMD simulation on a GPU with slightly aggressive boost parameters results in the complete conformational ensemble of peptaibols.

3. The study of folding dynamics also highlighted the differences that can occur with a single residue substitution.

4. The presence of helix-breaking residues like glycine and proline was highlighted to allow the hinge-like backbone bending motion of peptaibols that may have functional relevance.

5. aMD can be successfully applied to model all-atom representations of bilayer membranes and their interaction with peptaibols without introducing grave errors.
LIST OF PUBLICATIONS RELATED TO THIS THESIS


OTHER ARTICLES


**Cumulative impact factor: 87.088**

**BOOK CHAPTER:**

OTHER PUBLICATIONS


