Physicochemical Properties of the Modeled Structure of Astacin Metalloprotease Moulting Enzyme NAS-36 and Mapping the Druggable Allosteric Space of *Heamonchus contortus*, *Brugia malayi* and *Ceanorhabditis elegans* via Molecular Dynamics Simulation

Om Prakash SHARMA1, Sonali AGRAWAL2,3, M. Suresh KUMAR1∗

1(Centre of Excellence in Bioinformatics, School of Life Sciences, Pondicherry University, Pondicherry 605014, India)
2(Centre of Excellence in Bioinformatics, School of Biotechnology, Madurai Kamaraj University, Madurai 625021, India)
3(Centre for Human Genetics, Bangalore 560100, India)

Received 7 June 2012 / Revised 6 December 2012 / Accepted 22 April 2013

Abstract: Nematodes represent the second largest phylum in the animal kingdom. It is the most abundant species (500,000) in the planet. It causes chronic, debilitating infections worldwide such as ascariasis, trichuriasis, hookworm, enterobiasis, strongyloidiasis, filariasis and trichinosis, among others. Molecular modeling tools can play an important role in the identification and structural investigation of molecular targets that can act as a vital candidate against filariasis. In this study, sequence analysis of NAS-36 from *H. contortus* (*Heamonchus contortus*), *B. malayi* (*Brugia malayi*) and *C. elegans* (*Ceanorhabditis elegans*) has been performed, in order to identify the conserved residues. Tertiary structure was developed for an insight into the molecular structure of the enzyme. Molecular Dynamics Simulation (MDS) studies have been carried out to analyze the stability and the physical properties of the proposed enzyme models in the *H. contortus*, *B. malayi* and *C. elegans*. Moreover, the drug binding sites have been mapped for inhibiting the function of NAS-36 enzyme. The molecular identity of this protease could eventually demonstrate how ex-sheathment is regulated, as well as provide a potential target of anthelmintics for the prevention of nematode infections.

Key words: *Heamonchus contortus*, *Brugia malayi*, *Ceanorhabditis elegans*, homology modeling, Molecular Dynamics Simulation, NAS-36, principal component analysis (PCA).

1 Introduction

Lymphatic filariasis (LF) is a severely debilitating, disfiguring and stigmatizing disease ascribed as neglected tropical diseases (NTDs), caused by parasitic worms. It usually causes abnormal enlargement of the limbs and the genitals (World Health Organization, 2011). As per estimated 1.34 billion people in 81 countries live in areas where filariasis is endemic and are at risk of infection (World Health Organization, WHO). In 1993, following advances in diagnosis and treatment the Carter Centre's International Task Force on Disease Eradication classified LF as a “potentially eradicable” disease. Current drugs used for MDA implementation by national elimination program only clear microfilariae temporarily without killing all adult worms (Bockarie and Deb, 2010). Thus the main problem regarding the chemotherapy of filariasis is that no safe and effective drug is available yet to combat the adult human filarial worms (Ahmad and Srivastava, 2007). Therefore there is an unequivocal call for the development of highly efficient, complementary chemotherapeutical drug with a macrofilaricidal effect.

Extra cellular matrix (ECM), called the cuticle, being tough but flexible exoskeleton maintains body shape, permits mobility via its attachments to muscles and provides a protective barrier to the worm from external environment (Page and Johnstone, 2007). The cuticle is a complex and versatile tissue made up of multiple layers, having distinct levels of structural integrity. The principal components of the cuticle are the collagens. Mutations in specific collagens and their processing enzymes result in aberrant cuticle formation that leads to distinctive body shape phenotypes such as Dpy (dumpy, short and fat), Rol (roller), or Bli (blistered) (Thein et al., 2009). By investigating the life-cycle of worms, it came into light that the microfilariae molt twice in competent insect vectors to become infective stage larvae (L3) that are infective to humans. L3 molt twice in the human host and become adult worms that are re-
productively active for years (Stepek et al., 2011). The molt L3 to L4 stage, i.e. the degradation and escape of third stage cuticle, is related to increased expression of essential metalloprotease, suggesting a crucial requirement of these enzymes in this process (Richer et al., 1992). The Nematode Astacin (NAS) metalloprotease in free-living nematodes, C. elegans, have been identified to be essential to correct development and proper shedding of the cuticle (Davis et al., 2004). Particularly NAS-36, M12 family metalloprotease, encodes functionally conserved enzyme, which are independently described as essential genes whose inactivation disrupts molting (Stepek et al., 2011). NAS-36 enzyme might also regulate the assembly of new cuticle by processing the precursors of particular extracellular matrix proteins.

Recent advances including extensive genome search, sequence homology and complementation studies have shown the characterized parasitic worms, Haemonchus contortus and Brugia malayi are homologues of C. elegans (Stepek et al., 2011) particularly with regards to the NAS-36 metalloprotease enzyme. This study has demonstrated that NAS-36 has a crucial function of ecdysis during cuticle development in free living as well as in parasitic nematodes, therefore can be a key target for developing a newer control strategy with nematocidal effect.

Knowledge of the three-dimensional structures of proteins opens the way to accelerate drug discovery and vaccine design (Sharma et al., 2011). Therefore in the absence of crystal structure of NAS-36, we modeled the tertiary structure of the NAS-36 from H. contortus, B. malayi and C. elegans. The modeled structure was verified using Molecular Dynamics Simulation (MDS) and various parameters to ensure the model structures are reliable and suitable for the further study. We have also identified the druggable pockets for the target enzyme in order to have a detailed insight into the molecular structure of NAS-36.

2 Materials and methods

2.1 Protein sequence identification and analysis of NAS-36 from Haemonchus contortus, Brugia malayi and Caenorhabditis elegans

Domain protein sequences of astacin metalloprotease from Haemonchus contortus, Brugia malayi and Caenorhabditis elegans were searched and retrieved from NCBI GenBank (Benson et al., 2000). The obtained protein sequences of H. contortus, B. malayi and C. elegans were used as input sequences for Basic Local Alignment Search Tool (BLAST) and BLASTp was performed against the Protein Data Bank (PDB) using our home made EpiBLAST tool (Sharma et al., 2012a) to find out the appropriate templates for the homology modeling. Based on BLAST result, the most identical and resembled crystal structure of proteins was downloaded from PDB for further analysis. ClustalX (Thompson et al., 2002) was used for the sequence alignment between the target and the template protein.

2.2 Homology modeling and structural analysis

The 3D structures of astacin metalloprotease from H. contortus, B. malayi and C. elegans were modeled using effective and comparative molecular modeling software named MODELLER9v9 (Sali et al., 1995). Best modeled structure was selected based on the Discrete Optimized Protein Energy (DOPE) scores (Eramian et al., 2006) defined by MODELLER. The modeled structure of the astacin metalloprotease was thoroughly analyzed with the help of PDBsum generate (Laskowski, 2001). All the selected modeled structures were viewed and analysed with the help of Chimera (Petterson et al., 2004) and PyMOL (Grell et al., 2006) molecular visualization tools. Later, the structures were subjected to the MDS using the GROMACS 4.5 (Van Der Spoel et al., 2005) package for the analysis of structural stability and flexibility.

2.3 Molecular Dynamics Simulations

The generated model of NAS-36 enzyme was used for performing MDS with the standard OPLS-AA (Optimized Potentials for Liquid Simulations) force field (Xu et al., 2007). For MDS, we have used our earlier protocol (Sharma et al., 2012b; Sharma et al., 2012c). The energy of the model protein was minimized using the steepest descent approach realized in the GROMACS 4.5 package. Then, a 500 ps position restraining simulation was carried out to restrain the protein by a 1000 kJ/mol·Å² harmonic constraint to relieve close contacts before the actual simulation. The Particle Mesh Ewald (PME) method (Wang et al., 2010) was used for computing the long-range electrostatics, a 1.4 Å cutoff for van-der Waal (vdW) interactions, a 1.2 Å cutoff for Coulomb interaction with updates every 10 steps, and the Linear Constraint Solver (LINCS) (Amiri et al., 2007) algorithm for covalent bond constraints were used. The system was simulated at 300K. Berendsen’s temperature and pressure coupling method was used to regulate the temperature and pressure of the system. After restrained dynamics, the actual MD run was set to 10 ns with the same parameters as mentioned above. All analyses were performed using the programs built within the GROMACS 4.5 and the coordinates were saved and analysed using XMGRAVE software (http://rpmfind.net/linux/RPM/opensuse/11.4/x86_64/xmgrace-5.1.22-6.2.x86_64.html). The lowest potential energy conformations were selected from the 8ns MDS trajectory and further refined by energy minimization. The refined models were validated us-