Karolinska Institutet

Institutionen för Cell och Molekylärbiologi

Comparative Genomics and Molecular Characterization of N-alpha Acetyltransferase in Trypanosomes for Drug Target Identification

AKADEMISK AVHANDLING
som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i Föreläsningssal CMB, Berzelius väg 21.

Fredag den 15 Mars, 2013, kl 09.30

av
Stephen O Ochaya
Herr

Huvudhandledare:
Professor Björn Andersson
Karolinska Institutet
Institutionen för Cell och Molekylärbiologi

Bihandledare:
Universitetslektor Lena Åslund
Uppsala Universitet
Institutionen för immunologi, genetik och patologi

Professor Martin Rottenberg
Karolinska Institutet
Institutionen för mikrobiologi, tumör- och cellbiologi

Fakultetsopponent:
Dr. Anders Hofer
Umeå Universitet
Medicinsk kemi och biofysik

Betygsnämnd:
Professor Marita Troye-Blomberg
Stockholms Universitet
Wenner-Grens institut
Avdelningen för immunologi

Docent Piergiorgio Percipalle
Karolinska Institutet
Institutionen för Cell och Molekylärbiologi

Docent Kutty Selva Nandakumar
Karolinska Institutet
Institutionen för medicinsk biokemi och biofysik

Stockholm 2013

ABSTRACT

Trypanosoma cruzi and *Trypanosoma brucei* are flagellated *protozoan* parasites that cause Chagas disease and African trypanosomiasis in Latin American and African countries, respectively. Currently, over 8 million people are infected with *T. cruzi* and about 25 million more are at risk. About half a million people are affected by *T. brucei*. Trypanosome species share many peculiar biological and biochemical features, such as RNA editing. In contrast, they exhibit profound differences at the level of host-parasite interaction and disease pathology. Both parasites are transmitted to their host via different insect vectors. There are no available vaccines, and the current treatments have severe adverse effects. We were involved in sequencing the *T. cruzi* genome, an initiative launched by the WHO to increase our knowledge of the molecular basis of the parasite.

The aim of this thesis was to participate in the sequencing and analysis of the *T. cruzi* genome, and use the data to investigate acetyltransferase enzymes, presumably linked to important metabolic pathways, as possible drug targets. In Papers I and II, we describe genome sequencing and analysis of two distinct *T. cruzi* strains. One of the selected strains, CL Brener, was found to be a genetic hybrid of two divergent strains; and it contains about 22 000 genes, encoded on 700 scaffolds with a total genome size of 110 Mb. About 50% of the genes are of unknown function, and lack homology to other sequenced eukaryotes. Large numbers of members of surface molecule gene families, such as trans-sialidase, mucin, mucin-associated protein, and GP63 were found. Comparative analyses revealed that TcI had a smaller genome by up to about 11 Mb. The genome size difference was linked to genes encoding surface molecules and to other repeats and repeated genes. Additionally, six reading frames present in TcVI were not detected in TcI. Genetic polymorphisms such as, indels, microsatellites and SNPs were identified and analyzed. Many genes were found to be under different selective pressures in *T. cruzi*, indicating differential evolutionary rates, signifying their importance to parasite biology. Within syntenic regions, the two genomes have the same gene complement. Identified features warrant sequencing of further *T. cruzi* strains, and findings from our studies offer opportunities for more targeted functional studies as well as tools for epidemiology.

In the second part of this thesis, Papers III to V, a *Trypanosoma cruzi* acetyltransferase gene family, identified in the genome project, was chosen for functional characterization as a first step to evaluate its potential as drug target.

Acetyltransferases are responsible for protein acetylation, where an acetyl molecule is transferred from acetyl-Coenzyme A to lysine residues in a protein sequence, N-epsilon acetylation, and to N-termini of proteins or peptides, protein N-alpha acetylation. N-alpha acetylation is linked to many metabolic pathways, influences protein stability, protein-protein interaction, localization to organelles and acts as degradation signals. The impact of this post-translational modification in parasite is not known.

We have identified *T. cruzi* NatC and A, and show that they are expressed in the three life cycle stages (epimastigote, trypomastigote, and amastigote). The catalytic and auxiliary subunits form a complex *in vivo*. Additionally, they partially co-sediment with the ribosome and may have both co-translational and post-translational protein acetylation functions. In epimastigote, the catalytic subunit of *T. cruzi* NatA was localized both in the nuclear periphery and in cytoplasm, whereas NatC was predominantly assigned to the cytoplasm. The auxiliary subunit of NatA was mainly confined to the cytoplasm with cytoskeletal-like labelling, whereas NatC showed a punctate profile. Interestingly, the staining patterns of the different subunits analysed for NatA and NatC differ between the life cycle stages, which suggests differential regulation and expression.

The native substrates for NatC and predicted NatA, are similar to those described in yeast and humans, suggesting evolutionary conserved functions. The proteins appear to acetylate a large number of proteins N-terminally, suggesting that manipulation of the enzymes may simultaneously affect many cellular functions and thereby could interfere with or abolish infection. Additionally, our data indicate that NatC and A, may have both N-alpha and N-epsilon acetylation potential.

Collectively, the genome analyses presented here have provided more molecular insights into the parasite's biology, and have narrowed the gaps between scientific communities working on parasite research. The identification of Nats and native substrates has hopefully laid a solid foundation for future study of Nats, which could provide chemotherapeutic targets for parasitic diseases.