Systems biology modelling of lipid metabolism in the malaria parasite Plasmodium / Modélisation intégrée du métabolisme des lipides chez Plasmodium, parasite causal du paludisme

Résumé. **Le lipidome de Plasmodium** est d'un intérêt certain car il paraît essentiel à la survie de ce parasite et est l'objet d'une nouvelle approche pharmacologique pour le traitement de la maladie dont il est responsable. Plasmodium se prête particulièrement bien à **une étude du type biologie des systèmes** car il comprend une diversité rare de voies métaboliques que le parasite met en œuvre lors de sa réplication. La reconstitution des voies de synthèse dans le métabolisme des lipides chez Plasmodium est un problème crucial pour la compréhension de la biologie de cet agent infectieux. L'identification des intermédiaires et des produits finaux de ces voies, des co-facteurs et des partenaires, ainsi que leurs changements en fonction des stades de développement ou sous l'effet d'effecteurs pharmacologiques dans une approche transcriptomique et maintenant métabolomique développées par l'équipe du Dr. Henri Vial (en collaboration avec Dr.Karine Leroch, University of California Riverside, Dr. Isabelle Lefebvre-Tournier, Prof. C Périgaud, UMR CNRS-UM2 **5247**, département de chimie du CNRS) servira de source d'informations pour la construction de modèles intégrés.

Les méthodes de la biologie des systèmes permettent d'intégrer, dans le cadre du même modèle mathématique, des phénomènes biologiques très différents. L'objectif de cette thèse est de développer et de mettre en pratique une méthodologie pour construire un modèle dynamique du type réseau mixte (métabolique couplé à un réseau génétique) en utilisant l'information contenue dans des données hétérogènes, transcriptomiques et métabolomiques. Cela conduira à élucider les interactions métaboliques, déterminer les régulations et les nœuds de régulations et comprendre les relations métaboliques/ transcriptomiques. Ce modèle servira par la suite pour tester in silico l'effet de différentes perturbations sur les étapes importantes du cycle de vie du parasite en différents milieux.

Ce projet est couplé au réseau d'Excellence EVImalAR (European Virtual Institute dedicated to Malaria Research, October 2009 - September 2014, A Waters, University of Glasgow, UK, Coordinator) qui propose comme un objectif important de la recherche dans ce domaine, le développement des approches de la biologie des systèmes pour l'étude de la malaria. Les travaux de cette thèse conduiront à une interaction forte avec les membres du consortium ainsi qu'avec des membres du Centre de Biologie Intégrative des Systèmes (CISBIC) et du Service de Bionformatique de Imperial College à Londres (Prof. Stephen Muggleton, Dr. Sarah Butcher). L'Imperial College de Londres a acquis une position mondiale dans le domaine.

Summary. Malaria research has entered the post-genomic era, the complete genome sequence of several *Plasmodium* species, strains and field isolates has been determined. However the complexity of this deadly parasite resides not simply in the composition of its structures, but more importantly in the wide-ranging networks of regulatory interactions within the parasite, and between the parasite and it's hosts. This concept, termed 'Systems Biology', opens up new ways of thinking and working, integrating data from various high throughput disciplines to understand structural and regulatory networks of the cell. In an integrated approach, to understand how parasite lipid metabolism is controlled we will combine current, and newly generated knowledge from the parasites' transcriptome, proteome, and metabolome into quantitative mathematical models.

Fluxomics, Metabolomic and transcriptomic approaches developed by the team of Dr. Henri Vial (in collaboration with Dr. Karine Leroch, from University of California Riverside, and with Dr. Isabelle Lefebvre-Tournier, and Prof. C Périgaud, from the department of Chemistry of CNRS) will provide information for constructing integrated models.

The objective of this thesis is to develop a methodology for constructing dynamic mixed metabolic and genetic networks by using heterogeneous data. This methodology will be applied to fluxomic transcriptomic and metabolomic data and enzyme kinetics in order to build a model for lipid metabolism in Plasmodium. This model will be used for in silico studies of the life cycle and of the defences against perturbations of the parasite.

The project responds to the needs of the Network of Excellence EVImalAR, (European Virtual Institute dedicated to Malaria Research, October 2009 - September 2014, A Waters, University of Glasgow, UK, Coordinator) whose cluster 4 proposes the development of systems biology approaches of malaria as an work package. The theoretical work will benefit from the strong interaction with members of the consortium and members of the Centre for Integrative Systems Biology (CISBIC) and the Bioinformatics Support Service in Imperial College London (Prof. Stephen Muggleton, Dr. Sarah Butcher).

Introduction, state of the art

Metabolite levels and fluxes represent the downstream amplification of changes occurring in the transcriptome and proteome and are the final mediators of biological function. This level of analysis, that has so far not been exploited to any significant extent in *Plasmodium*, is that of the 'metabolome', the complement of "all" the low molecular weight molecules measured in the parasite in a particular physiological state. Understanding and modelling metabolic networks is an important goal, not only for a fundamental understanding of parasite cellular physiology, but also for providing new weapons against this deadly parasite.

We will examine the role of individual reactions and pathways in the context of an entire metabolic network. We will prioritize lipid metabolism in *P. falciparum*, because this lower eukaryotes used a panoply of metabolic pathways rarely reunified in a single organism. We will study also stress- or drug-induced changes in this metabolism, using mutated parasites as appropriate. The use of the most advanced techniques including high resolution mass spectrometry (MS), will allow the detection of low-level metabolites knowledge of which will permit us to build a comprehensive network of metabolic pathways operating in an organism at a given time.

The Malaria Lipidome: Lipids are among the most critical components involved in cell physiology as crucial structural, functional and regulatory components. Phospholipids (PL) are the major malaria lipids and *P. falciparum* utilizes a bewildering variety of metabolic pathways to synthesize considerable amounts of membrane. One of the purposes of our project is to understand why *Plasmodium* uses so many pathways. Final and intermediate metabolites of these metabolic pathways will be identified and quantified. We will study the respective contributions of the pathways to the production of individuals PL and investigate eventual inter-regulations and key regulations. Modulations will be induced by treating the parasites with drugs (including a novel class of potent antimalarial compounds that were discovered in the team of Dr. Henri Vial). Alternatively, alteration of these metabolisms will be obtained from *P.falciparum* cell lines transgenic for enzymes involved in the lipid biosynthesis pathway (over-expressing cell lines or null mutants).

Systems biology models: Most mathematical methods for modelling regulatory mechanisms are based on formal methods of chemical kinetics developed for studying chemical or biochemical systems. Modelling

of large biochemical networks, based on standard mathematical approaches, faces obstacles such as incompleteness of network description (structural and parametric) and lack of exact knowledge of kinetic parameters.

Constraint based approaches (such as flux balance analysis FBA) circumvent these obstacles by using optimality principles and replacing the network by a set of stoichiometric constraints. FBA is well suited for global studies of perturbations of metabolism. Thermodynamics imposes constraints that can be dealt with within the same approach. FBA has been successfully applied to the global study of the metabolism of *Leishmania* to identify lethal and growth reducing gene knock-outs in various media, as well as for defining the concept of minimal supporting growth media [1]. In spite of these successes FBA has two major drawbacks. It can not deal with time dynamics. Moreover it can not predict concentrations of metabolites (the predicted variables are the fluxes), that is a major defect when dealing with metabolomic studies. Dynamic effects are particularly important in parmacokinetik when, depending on the dose time-scenario, the application of a drug could trigger or not compensatory mechanisms or cytostasis.

Dynamic modeling approaches using differential equations have been employed to test in silico the hypothesis that compartmentalisation of the first steps of trypanosome glycolysis in a ATP buffered glycosome, allows to avoid turboexplosion of metabolic intermediates as the effect of an otherwise autocatalytic feed-back loop [2].

When used for genetic networks, dynamic modeling allows reconstructing dynamic genetic regulatory networks from transcriptome time series. Such time series containing about 50 microarrays (taken every hour) have been obtained very early for Plasmodium falciparum blood stage [3]. More recent studies use similar timeseries to reveal perturbation-specific compensatory mechanisms [4].

In spite of some relevant progress in modelling trypanosome apicomplexa, nothing has been done on systems biology of *Plasmodium*. Our project will be the first systems biology project on malaria. In particular, combining metabolomic and transcriptomic data will provide a very interesting framework for developing models containing the two layers of the regulatory picture (metabolic and genetic) and will allow a step forward towards understanding of the life cycle and of the defences against perturbations of these parasites.

Main objectives

In order to answer the above questions, the thesis will be directed towards the following two main objectives:
1) Construct medium and large scale dynamic models of lipid metabolism in Plasmodium together with its main regulations.
2) Develop a methodology to extract information from metabolomic, fluxomic and transcriptomic data.
The objectives are presented in detail below.

Model building. Building systems biology models is a multi-step process, each step depending on specific data and on the level of complexity we target.
In a first approach we will build a medium scale model using existing information on the lipid synthetic pathways also kinetic information on the enzymatic reactions obtained in the group of Dr. Henri Vial. This model will be challenged to predict known behaviour under basic perturbations and in various media.

A more complex model will be built starting from the main metabolites that were identified in the metabolomic studies. Annotated genome from PlasmoDB and pathway databases (Hagai Ginsburg biochemistry database, PlasmoCyc, Ecocyc, etc) will be used for reconstructing the reaction pathways

between metabolites. Machine learning could be used to fill in gaps in the model (missing reactions). Kinetic information will surely be incomplete for the larger model. To solve the problem we will use original model reduction methods previously developed to cope with incomplete information [5,7]. These methods can employ information about the order of the reaction constants instead of precise numerical values. The orders of the constants will be obtained from metabolomic data (see next objective). The remaining critical values (in much smaller number) that need to be known with more precision could be reverse engineered using optimisation techniques [5,6].

Methodology for extracting information from heterogeneous data. Metabolomics, fluxomics and transcriptomic data contain complementary information. Thus, it is well known in metabolic control that fluxes are more stable (more buffered) than metabolites. Small localized perturbations (for instance partial inhibition of an enzyme) can produce significant variations of the metabolites but only small variations of the fluxes. Larger perturbations can redistribute metabolic fluxes such as in metabolic switches and this is also accompanied by significant metabolite variations.
Furthermore, the distribution of the metabolites along pathways depends on the kinetic parameters of the reactions. Thus, in a chain, metabolites are accumulated before and depleted after a limiting step. More general rules were developed in our recent theory of limitation in linear networks with separated constants [5,7] and will be extended to the non-linear and partially separated cases. Thus, metabolite concentrations can be used to estimate orders of parameters and time-scales of reactions which can be used in our dynamical model reconstruction.
Finally, metabolomic, fluxomic and transcriptomic variations are correlated. There are two sources for metabolite variations, the first one being the variation of fluxes and the second the variations of enzyme activities. Using metabolic control theory and general limitation theory we will be able to related fluxes and metabolite variations to enzyme activity variations. These can be compared to transcriptome variations. The conflicts will indicate unknown regulations at the post-transcriptional and post-translational level, that are quite frequent in apicomplexa (these inherit many of the regulation mechanisms of plants).

References:

[1] AK Chavali, JD Whittemore, JA Eddy, KT Williams, and JA Papin. Systems analysis of metabolism in the pathogenic trypanosomatid *Leishmania major*. Mol Syst Biol. 4: 177 (2008).
[2] JR Haanstra, A van Tuijl, P Kessler, W Reijnders, P Michels, HV Westerhoff, M Parsons, BM Bakker. Compartmentation prevents a lethal turbo-explosion of glycolysis in trypanosomes. PNAS,105:17718-23 (2008).
[3] Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, DeRisi JL., The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. PLoS Biol., 1: E5 (2003).
[4] Anna C van Brummelen, Kellen L Olszewski, Daniel Wilinski, Manuel Llinás, Abraham I Louw, and Lyn-Marie Birkholtz. Co-inhibition of *Plasmodium falciparum S*-Adenosylmethionine Decarboxylase/Ornithine Decarboxylase Reveals Perturbation-specific Compensatory Mechanisms by Transcriptome, Proteome, and Metabolome Analyses. *The Journal of Biological Chemistry, 284: 4635-4646 (2009).*
[5]O.Radulescu, A.N.Gorban, A.Zinovyev, A.Lilienbaum, "Robust simplifications of multiscale biochemical networks", BMC Systems Biology, 2008, 2:86.
[6] P. Blavy, F. Gondret, H. Guillou, S. Lagarrigue, P.G.P., Martin, J. van Milgen, O. Radulescu and Siegel, A. A minimal model for hepatic fatty acid balance during fasting: Application to PPAR alpha-deficient mice, Journal of Theoretical Biology, 261 (2009), pp. 266-278.
[7]A.N. Gorban, O.Radulescu, A. Zinovyev, Asymptotology of Chemical Reaction Networks,Chemical Engineering Science, available online Sept.2009.