

Modélisation de phénomènes biologiques : état de l'art et bilan du projet microMéga

Carole Bernon, Davy Capera, Pierre Glize, Jean-Pierre Mano, Sylvain Videau

Mai 2009

Rapport interne IRIT / RR--2009-12--FR

Résumé :

En raison du manque de théories satisfaisantes pour expliquer les phénomènes biologiques, les biologistes s'appuient habituellement sur des outils de modélisation ainsi que sur la simulation pour essayer de comprendre le comportement de ces systèmes. L'objectif du projet ANR microMéga (BLAN-0202-05, 2005-2008) a été de fournir une plate-forme de modélisation et de simulation reposant sur des systèmes multi-agents auto-organiseurs. Ce rapport présente, dans le cadre de ce projet, des outils et systèmes capables de modéliser des phénomènes et systèmes biologiques en se focalisant plus particulièrement sur les phénomènes intracellulaires. Dans un second temps, il fait un bilan du projet en en présentant les grandes lignes.



Contenu

| | |
|---------------------------------------------------------------------------------------------------|-----------|
| 1. Introduction | 5 |
| 1.1. Contexte du projet microMéga | 5 |
| 1.2. Objectif pour l'équipe SMAC | 5 |
| 1.3. Contenu du rapport..... | 6 |
| 2. Modélisation de phénomènes biologiques | 6 |
| 2.1. Simulations biologiques et modélisations fonctionnelles de cellules | 6 |
| 2.1.1. Modélisation cellulaire générique par réseaux d'interaction..... | 7 |
| 2.1.2. Modélisation spatiale de la signalisation, transduction et diffusion intracellulaires..... | 11 |
| 2.2. Agents, systèmes multi-agents et modélisation biologique..... | 12 |
| 2.2.1. Modélisation de réactions biologiques | 12 |
| 2.2.2. Modélisation de cellules | 13 |
| 2.2.3. Modélisation de population de cellules | 15 |
| 3. La plate-forme de modélisation de microMéga | 16 |
| 3.1. Approche adoptée | 16 |
| 3.2. Composants de base d'un modèle | 17 |
| 3.3. Comportement des agents | 18 |
| 3.3.1. Comportement nominal..... | 18 |
| 3.3.2. Comportement coopératif..... | 19 |
| 3.4. Exemples de modèles | 20 |
| 4. Conclusion | 22 |
| 5. Références bibliographiques | 23 |
| 5.1. Travaux en relation avec les technologies autres qu'agents | 23 |
| 5.2. Travaux en relation avec les technologies agents | 25 |



1. Introduction

1.1. Contexte du projet microMéga

Ce document a été rédigé dans le cadre du projet ANR microMéga¹ (www.irit.fr/microMega) dont l'objectif à long terme est de permettre aux biologistes de réaliser des expérimentations virtuelles.

En effet, en raison du manque de théories satisfaisantes pour expliquer les systèmes biologiques, les biologistes s'appuient habituellement sur des outils de modélisation et la simulation pour essayer de comprendre le comportement de ces systèmes. [Ren08] considère les systèmes biologiques comme « des systèmes complexes caractérisés par des non-linéarités, i.e. l'émergence, ce qui signifie aussi "que le tout n'est pas égal à la somme de ses parties". Le comportement global d'un système biologique est un comportement émergent des interactions entre de nombreux composants locaux (tels que l'ADN, les protéines, les cellules)². [...] Chaque composant possède ses règles comportementales simples, tandis qu'un groupe de composants peut exhiber un comportement complexe émergent (i.e. adaptabilité, évolution, sécurité, auto-maintenance). [...] Le système biologique est distribué de manière inhérente sans élément central le contrôlant ». Ceci résumé extrêmement bien les écueils auxquels doivent faire face les outils de modélisation et simulation des phénomènes et systèmes biologiques : les non-linéarités, l'ouverture, la compréhension des phénomènes d'auto-organisation et émergents [Holland98, Maturana73], etc.

Comme le montre l'état de l'art des paragraphes suivants, la plupart des modèles utilisés pour simuler des phénomènes biologiques, qu'ils reposent sur des outils mathématiques (systèmes d'équations différentielles, par exemple), des réseaux de neurones, des réseaux de Petri ou autres, agissent souvent comme des boîtes noires en empêchant de comprendre la structure réelle du système simulé et sont essentiellement statiques. Par exemple, il est généralement impossible d'ajouter de nouveaux éléments au modèle en cours de fonctionnement ou de le faire évoluer dans un environnement dynamique. Toutefois, depuis quelques temps, un lien commence à se former entre le domaine biologique et celui des technologies agents. Ainsi, un groupe de travail européen s'est intéressé à ces relations [Luck05] et d'autres chercheurs ont évalué jusqu'à quel point utiliser des systèmes multi-agents pour simuler des systèmes biologiques peut être fructueux [Amigoni07]. Il est vrai que les caractéristiques des systèmes multi-agents et de leurs composants cadrent parfaitement à celles énoncées ci-dessus (système distribué, sans contrôle central, ouvert, entités en interaction, émergence d'une fonction collective, adaptation, etc.).

1.2. Objectif pour l'équipe SMAC

Dans ce contexte, le projet microMéga a été l'occasion pour l'équipe SMAC de confronter la théorie des AMAS (Adaptive Multi-Agent Systems [Camps98]) à un nouveau domaine, celui de la biologie computationnelle. L'objectif étant alors de fournir aux biologistes les moyens de construire un modèle intracellulaire assez flexible pour leur permettre de comprendre des phénomènes cellulaires simulés et de découvrir de nouvelles expérimentations. Le micro-organisme cible est *Saccharomyces Cerevisiae*.

L'idée du travail entrepris est donc de faire en sorte que le modèle reflète à tout instant les données expérimentales ainsi que les souhaits de l'utilisateur i.e. qu'il reste fonctionnel quels que soient les changements apportés à son environnement d'exécution. On doit alors le doter de capacités à s'adapter à ces changements et donc à s'auto-construire en fonction notamment des données expérimentales sur lesquelles il va « apprendre » à devenir le modèle adéquat. Une fois ce modèle considéré comme stable et donnant les bons résultats pour ces données expérimentales, il doit être possible de

¹ N° ANR-BLAN-0202-05, 2005-2008, partenaires : équipes « Génie Microbiologique » et « Physiologie Microbienne des Eucaryotes » du LISBP/INSA, équipe SMAC de l'IRIT, groupe DISCO du LAAS, LSP/UPS.

² Constatations corroborées par [O'Malley05, Hazen05]

déconnecter cette capacité d'auto-construction. On peut alors utiliser le modèle pour réaliser des simulations à partir de données différentes de celles qui lui ont permis « d'apprendre » sa structure.

1.3. Contenu du rapport

L'objectif de ce document est de positionner microMéga par rapport aux travaux similaires existants puis de faire un rapide bilan du projet pour l'équipe de l'IRIT qui y a participé, à savoir l'équipe SMAC³ (Systèmes Multi-Agents Coopératifs).

Ainsi, l'état de l'art présenté ci-après s'intéresse d'abord aux approches « classiques » de modélisation de systèmes biologiques avant de se focaliser plus particulièrement sur l'apport des technologies agent et multi-agent à ce domaine.

La dernière partie décrit rapidement la plate-forme de modélisation développée pour le projet ainsi qu'un exemple de résultats obtenus grâce à elle.

2. Modélisation de phénomènes biologiques

2.1. Simulations biologiques et modélisations fonctionnelles de cellules

Il est utile de préciser la distinction entre la biologie systémique et la bioinformatique. Cette dernière s'applique, en relation avec le domaine d'étude qui concerne microMéga, à manipuler les énormes et disparates quantités de données issues des recherches en biologie [Castrillo04], que ce soit le traitement, la normalisation et l'analyse statistique des relevés [Fang03, Tilstone03], ou encore des analyses multivariées comparatives [Kell00, Bryant01], du « clustering » sur les réseaux d'interaction géniques couplés ou non aux réseaux métaboliques [de la Fuente02, Mendes02, Wu02, Fiehn03], voire à un niveau plus intégré, la gestion globale des différents niveaux d'études de l'information génomique comme dans les plates-formes GIMS [Cornell03], MAPPFinder ou encore geneMAPP [Doniger03].

La biologie systémique, en revanche, utilise l'outil informatique pour modéliser et simuler les systèmes biologiques. Cela implique forcément une première étape de « data mining » afin de travailler sur des données souvent pré-analysées par les outils statistiques de la bioinformatique, pour nourrir les modèles ou simplement les valider. La deuxième étape est les constructions du modèle (le plus souvent mathématique) qui devront satisfaire, au moins partiellement, une partie du corpus des données. À partir du modèle, une étape de simulation permettra d'en valider ou non la valeur prédictive.

Aujourd'hui, plusieurs méthodes de modélisation en biochimie, métabolomique et génomique sont déployées pour couvrir la modélisation des mécanismes cellulaires à l'œuvre dans les cellules ou, plus généralement, lors des interactions biochimiques.

On distingue, en premier lieu, les réseaux booléens sous forme de réseaux de Petri ou encore d'automates cellulaires [Li07, Huang99, Shmulevich02], bien adaptés à l'analyse qualitative des systèmes étendus où l'état d'un nœud sera déterminé par l'état des nœuds qui lui sont directement connectés, et en second lieu, les équations différentielles, qui dans des conditions particulières, permettent d'approcher très finement le détail des cinétiques réactionnelles. Les modèles à équations différentielles, le plus souvent ODE (pour « Ordinary Differential Equations ») décrivent les relations non linéaires entretenues par les variables et/ou les paramètres du système afin d'en décrire les propriétés quantitatives.

Lorsque le modèle décrit les lois d'action de masse d'une réaction biochimique, les équations différentielles prennent une forme stoechiométrique qui décrit l'équilibre entre les substrats et produits de la réaction. À partir de là, deux types principaux de résolution sont à la disposition des modélisateurs, une résolution probabiliste (stochastique) ou une résolution déterministe. Ces deux techniques se différencient en fonction du nombre de molécules impliquées dans le processus étudié. Lorsque l'on s'adresse à un petit nombre

³ www.irit.fr/SMAC



d'éléments en interaction, où des fluctuations locales de densité vont avoir une répercussion sur les vitesses de fonctionnement, une simulation stochastique sera la plus appropriée. Cette technique très gourmande en termes de calcul (chaque action élémentaire des entités en interaction est prise en compte sur une échelle de temps souvent très petite) n'est pas applicable lorsque l'effectif des éléments en interaction explose. La résolution déterministe est alors plus rapide à fournir des résultats.

On recense plusieurs algorithmes déterministes pour calculer les variables ; à chaque pas de la simulation, la valeur d'une variable est calculée à partir de sa valeur au pas précédent (algorithme explicite : méthode d'Euler de premier ordre) ou à partir d'une estimation de la valeur au pas suivant (méthode implicite : méthode d'Euler de deuxième ordre, méthode de Runge-Kutta d'ordre 4, méthode adaptative de Runge-Kutta-Fehlberg, méthodes Gear, CVODE, LSODES, Crank-Nicolson...). Ces algorithmes sont quasi ubiquitaires dans les plates-formes de simulation dont nous allons parler par la suite. Des approches probabilistes plus confidentielles existent cependant, que ce soit à l'aide de réseaux Bayésiens [Friedman00] ou de réseaux de neurones artificiels [Weaver99].

2.1.1. Modélisation cellulaire générique par réseaux d'interaction

Nous nous intéressons ici exclusivement aux modèles fonctionnels des cellules, des modèles algorithmiques qui vont simuler certains des comportements cellulaires à partir des interactions biochimiques dans un environnement prédéterminé. Il existe des modèles tissulaires électriques et ou mécaniques d'organes complets (cœur, os...) dont nous ne parlerons pas car ils se situent à un niveau d'intégration trop macroscopique.

- *CellWare*

www.cellware.org

CellWare est une plate-forme « open source » qui intègre des outils de modélisation textuels via un accès aux bases de données du web (KEGG...) ou graphiques, permettant de travailler sur des réseaux volumineux grâce à des outils de compression du graphe. Le moteur de simulation utilise à la demande de l'utilisateur soit différents types d'algorithmes stochastiques, soit prédictifs (ODE). La particularité de CellWare réside dans sa capacité à conduire des calculs sur grilles pour un traitement plus rapide.

- *Cytoscape*

www.cytoscape.org

Cytoscape permet la construction de modèles à partir des bases de données SBML, et la visualisation des modèles construits. Les outils graphiques sont d'une grande qualité et permettent l'exploration de graphes très complexes et étendus. Des outils mathématiques et statistiques permettent d'unifier les données en provenance de différentes sources. Cytoscape permet ensuite de visualiser les résultats des simulations que l'on aura exécutées sur d'autres plates-formes.

- *NeuralWorks Predict*

NeuralWorks Predict combine plusieurs approches : réseaux de neurones, cartes auto-organisatrices (SOM), algorithmes génétiques et logiques floues, afin de rechercher des associations de données et de construire des applications de prédiction et de classification. Les différentes techniques automatisées d'apprentissage et de sélection travaillent sur des flux de données brutes et permettent de fournir à l'utilisateur un réseau de neurones « clé en main ».

- *JWS*

jjj.biochem.sun.ac.za

JWS est une plateforme en ligne de modélisation en ligne de voies métabolique développée par Triple-J. Cette plate-forme est basée sur une résolution conventionnelle de système d'équations différentielles. L'approche utilisée cherche à mettre en avant, dans le modèle, les relations de contrôle tant dans les états d'équilibre que dans les processus dynamiques. Cette plate-forme permet une présentation très didactique des

modèles d'équations différentielles à partir d'une matrice des coefficients stoechiométriques et d'un vecteur des cinétiques réactionnelles. De plus, de très nombreux modèles sont disponibles sur le site ce qui favorise grandement les échanges « didactiques » dans ce domaine.

- *PySCeS*

pysces.sourceforge.net

PySCeS dans la même catégorie, mais nettement moins ergonomique, est un simulateur de systèmes cellulaires écrit en Python, il a l'avantage d'être très paramétrable par l'utilisateur.

- *Gepasi*

www.gepasi.org

Gepasi est un ensemble de logiciels de modélisation des systèmes biochimiques à partir de données réelles. Les vitesses des réactions sont associées à plusieurs outils algorithmiques d'optimisation qui opèrent sur les données et/ou le modèle afin d'ajuster l'adéquation entre les deux. Ainsi, Gepasi fournit une assistance à la construction du modèle pour réaliser la transposition du modèle « chimique » d'un enchaînement de réactions vers un modèle mathématique de description de variables. Les outils se veulent suffisamment précis pour une réelle modélisation exploratrice et les interfaces sont suffisamment simples pour une utilisation pédagogique. Au niveau observatoire, Gepasi intègre une analyse du contrôle métabolique (dépendances fonctionnelles), de même qu'une analyse des phases de stabilité linéaires [Mendes93].

- *Kinsolver*

lstdis.cs.uga.edu/~aleman/kinsolver

Kinsolver est un simulateur déterministe assez proche dans ses fonctionnalités du précédent (*Gepasi*) ; il semble plus spécifiquement conçu pour l'étude des réseaux de régulation géniques.

- *StochSim*

www.ebi.ac.uk/~lenov/stochsim.htm

Il est intéressant également de présenter ici StochSim, pour sa conception « objets » autonomes qui s'approche d'une technologie multi-agent. C'est une plate-forme qui exécute un environnement 2D dans lequel les molécules entretiennent des relations probabilistes de voisinage, en fonction de leurs concentrations et de taux constants.

- *Simulateurs interprétant le SBML*

Il existe de très nombreux simulateurs qui interprètent le langage SBML, un « Mark-up Language » dédié à la description des modèles biologiques :

- SBMLSim (www.dim.uchile.cl/~dremenik/SBMLSim) qui utilise le moteur mathématique de MATLAB et permet de simuler des modèles SBML et de visualiser les résultats.
 - MathSBML (sbml.org/software/mathsbml) qui a la même fonction que le précédent, mais s'interface avec Mathematica.
 - libSBML (www.sbml.org/libsbml.html) est une bibliothèque de programmation « open source » pour lire, écrire, manipuler, traduire et valider des fichiers SBML.
 - BIOCHAM (contraintes.inria.fr/BIOCHAM) est un environnement de programmation qui permet de récupérer des modèles SBML, mais également de coder le modèle dans un langage spécifique qui utilise une logique temporelle pour définir de propriétés (booléennes ou numériques) aux entités manipulées. Une visualisation graphique sommaire du modèle est disponible ainsi que analyse fonctionnelle complète de celui-ci.
- *Systems Biology Workbench*
www.sys-bio.org, sbw.sourceforge.net



Systems Biology Workbench se veut un environnement complet de biologie computationnelle. Il permet de créer manipuler et analyser des modèles biologiques au niveau moléculaire cellulaire et multicellulaire. Différents aspects des réseaux d'interaction biochimiques sont pris en compte : voies métaboliques, lois d'action de masse, comportements stochastiques, transduction du signal, régulation et expression des gènes. Il faut noter que l'ensemble de logiciels s'articulant autour de ce projet est « open source ».

Jarnac (simulateur) et JDesigner (module graphique de création de réseaux) participent au Systems Biology Workbench (SBW) et constituent un ensemble complet pour la modélisation et la simulation (équations différentielles ou stochastique) des systèmes métaboliques sans pour autant intégrer toutes les interactions possibles entre les gènes dans les modèles. Par exemple, *PNK2e* (page.mi.fu-berlin.de/trieglaf/PNK2e/index.html) utilise les réseaux de Pétri afin d'importer des modèles SBML et de les simuler sous forme de processus Markovien stochastique et donc déterministe. Des outils et de mesures des propriétés statiques et dynamiques du modèle sont disponibles.

Cette plate-forme soutenue par de nombreuses institutions (DARPA, NSF, DOE Genomics) offre d'excellentes possibilités au modélisateur confirmé qui sait ce que doit être son modèle.

- *CellDesigner*

www.celldesigner.org

Dans la continuité de SBW, CellDesigner étend les outils précédents avec une interface plus performante qui offre à l'utilisateur des outils graphique de modélisation et d'analyse des réseaux métaboliques et génétiques. CellDesigner permet avant tout de créer des modèles, puis de les simuler et de les analyser via les autres outils du SBW [Funahashi08].

- *Bio-SPICE*

biospice.sourceforge.net

Bio-SPICE est un autre ensemble de logiciels « open source » dont le but est d'aider à l'étude des systèmes cellulaires, plus particulièrement du point de vue de l'analyse spatio-temporelle des processus. Cette plate-forme polymorphe contient des outils de fouille de données, de développement de modèles mathématiques des interactions gène-protéines, de simulation... jusqu'aux outils de validation expérimentale. Elle offre la possibilité à de nombreux logiciels de s'interconnecter afin d'utiliser au mieux les compétences de chacun. Par exemple, Simpatica pour la simulation et l'analyse des voies métaboliques (bioinformatics.nyu.edu/Projects/Simpatica), JigCell pour une fonctionnalité similaire à laquelle s'ajoute une capacité d'estimation de paramètres (jigcell.biol.vt.edu), JDesigner pour la conception graphique des réseaux métaboliques et Charon, un langage objet dédié à la modélisation des systèmes métaboliques... Ainsi, plus de 80 outils fédérés en une même plate-forme offre une redondance de service qu'il est parfois difficile à démêler, mais en contrepartie ils rendent disponibles des plugins extrêmement spécifiques, les uns dédiés à l'analyse des puces ADN, les autres à l'identification morphologique des cellules...

- *CADLIVE*

www.cadlive.jp

CADLIVE est un système de construction et de simulation des réseaux d'interaction gène-protéomique. Il utilise une GUI pour définir le modèle à partir de ses réactions et des éléments acquis les régulent. Le modèle est stocké sous une forme XML propre à la plate-forme. Le simulateur est à base de résolution de systèmes d'équations différentielles (S-system), et le moteur de l'exploration paramétrique pour l'ajustement de modèle est un algorithme génétique.

- *BioUML*

www.biouml.org

BioUML permet une intégration graphique de modèles formels (structurels et fonctionnels) de biologie systémique, avec accès aux trois principales bases de données ouvertes. Cette plate-forme qui comporte des outils de simulation et de visualisation, permet également la gestion des formes polymères des protéines.

- *Moleculizer*

www.molsci.org/~lok/moleculizer/index.html

Moleculizer est un simulateur stochastique (Monte Carlo) de réseaux réactionnels cellulaires. Sa conception non spatiale permet cependant de simuler le comportement temporel des voies de signalisation intercellulaire en réaction à des perturbations. Le modèle fourni reste ouvert et les espèces moléculaires ainsi que les réactions peuvent être générées à la volée. Le résultat est exportable au format SBML [Lok05].

- *Cellerator*

www.cellerator.org

Cellerator est une ancienne extension de Mathematica pour simuler des voies de signalisation cellulaires et par composition pour simuler la communication intercellulaire. Basé sur la résolution d'équations différentielles, une nouvelle version serait en développement : xCellerator (xcellerator.info).

- *DynaFit*

www.biokin.com/dynafit/index.html

DynaFit est un outil de simulation de réactions métaboliques afin d'en estimer leur cinétique. L'analyse utilise une régression non linéaire (méthode des moindres carrés).

- *NetBuilder*

strc.herts.ac.uk/bio/maria/NetBuilder

NetBuilder est un environnement de simulation dédié à l'étude des réseaux de régulation génétique qui intègre des outils mathématiques limités permettant de définir les cinétiques réactionnelles. Cette plate-forme permet la conception graphique par l'utilisateur de modèles d'organismes cellulaires et multicellulaires.

- *Copasi (Complex Pathway Simulator)*

www.copasi.org

Copasi est un logiciel renommé de simulation et d'analyse des réseaux métaboliques même très étendus. Plus facile d'accès qu'E-Cell (voir paragraphe suivant), ce logiciel est associé à un projet de collaboration avec CellDesigner du Systems Biology Workbench afin d'apporter les outils de construction de modèles.

- *Virtual Cell*

Virtual Cell est une plate-forme en ligne de résolution de problèmes développée par NRCAM, utilisant une base de données centralisée. Virtual Cell permet de définir une topologie de la cellule, d'y exécuter des simulations et d'analyser les résultats. Cette plate-forme a donné lieu à un grand nombre de publications (www.nrcam.uchc.edu/applications/published%20models.html). À partir d'un modèle biologique associé à une micro-imagerie pour valider la localisation des objets du modèle, Virtual Cell permet la traduction du modèle au niveau mathématique par des approximations automatiques d'état d'équilibre et des règles de conservation de la masse. Le modèle mathématique fourni tend vers un maximum de simplicité, Virtual Cell est doté d'une interface Web et d'une interface adressée au biologiste qui donne à l'utilisateur la possibilité, via un éditeur mathématique, de corriger les propositions du système. Cette plate-forme intègre également des outils bien adaptés pour modéliser les phénomènes de réaction-diffusion en systèmes compartimentés [Moraru08].



- *JSIM*

bioeng.washington.edu/jsim

JSIM est le dernier environnement de simulation cellulaire dont nous parlerons ici, il fait partie du projet international « physiome project ». Il permet d'importer des modèles aux formats SBML ou CellML, mais également la construction de novo d'un modèle s'appuyant sur un module graphique (Jmodel). Il met en œuvre soit une conception processus-interaction (transition de phase), soit une conception événementielle, toutes deux pouvant mélanger des équations différentielles ordinaires, des équations aux dérivées partielles, des équations implicites, intégrales, sommes, événements discrets et fonctions de transition d'états... JSIM basé sur un formalisme assez proche du DEVS (Discrete Event System Specification) vérifie la cohérence des dimensions employées et l'équilibrage des équations.

2.1.2. Modélisation spatiale de la signalisation, transduction et diffusion intracellulaires

- *MCell*

www.mcell.psc.edu

MCell permet d'intégrer une topologie de très haute résolution dans des modèles de diffusion Brownienne des ligands et des voies de signalisation. Ainsi, le niveau de l'analyse se situe au-dessus de la dynamique moléculaire et ne peut intégrer la cellule dans son ensemble. Le moteur de simulation reste probabiliste (toutes les données des modèles sont transposées en probabilité de Monte Carlo), en revanche, c'est, avec MesoRD, l'une des rares plates-formes qui intègre la notion cardinale des molécules et leurs interactions dans un environnement dynamique et complexe.

- *MesoRD*

mesord.sourceforge.net

MesoRD offre des outils pour étudier plus spécifiquement les phénomènes de réaction/diffusion dans des environnements en deux ou trois dimensions afin d'explorer, par exemple, les mécanismes de ségrégation moléculaire qui interviennent lors de la spécialisation locale des membranes cellulaires.

- *E-Cell Simulation Environment*

ecell.sourceforge.net

E-Cell est un environnement de simulation orienté-objet dédié à l'étude des systèmes complexes. Issu d'un projet international dont les intentions couvrent tous les moyens nécessaires à reconstruction *in silico* des phénomènes biologiques, E-Cell vise une modélisation complète et précise de la cellule en utilisant une résolution composée où plusieurs algorithmes avec leurs échelles de temps propres coexistent. Ainsi, cette plate-forme peut potentiellement mettre en œuvre tous les mécanismes présents dans la cellule, incluant la réplication de l'ADN, et la régulation génétique de phénotypes cellulaires d'un ordre supérieur. À ce projet durable, et toujours en activité, est associé un grand nombre de publications (www.e-cell.org/publications). C'est une excellente plate-forme pour l'analyse et la simulation qui manque cependant d'outils de conception des modèles [Tomita99].

- *SigTran*

www.cellsystems.org/teams/modeling/projects/sigtran

SigTran a été conçu spécifiquement pour supporter la simulation de réseaux vastes et complexes de transduction du signal. Il intègre plusieurs algorithmes déterministes (ODE) ou stochastiques (Firth-Bray, Gillespie, Gillespie-Gibson...) dont un propre qui permet une gestion uniforme du temps sur l'ensemble du réseau. Le logiciel BALSÀ, développé dans la même université, s'associe à Sigtran pour modéliser et visualiser plus spécifiquement les phénomènes membranaires. Il est à noter que SigTran et BALSÀ constituent l'une des rares plates-formes écrites en Java.

- *Dizzy*

magnet.systemsbioology.net/software/Dizzy

Dizzy est un simulateur de réactions biochimiques sur une base stochastique (Gillespie, PGibson-Bruck, Tau-Leap) ou déterministe (ODE). Écrit en Java, il permet d'importer et exporter des modèles en SBML de gros réseaux d'interaction géno-métaboliques. L'affichage des graphes est délégué à Cytoscape. Cet outil offre une estimation du bruit à l'état d'équilibre et une spatialisation des réactions via la compartimentalisation des réactions.

- *STOCKS*

www.sysbio.pl/stocks

STOCKS est un logiciel de simulation stochastique des réseaux de réactions biochimiques [Kierzek02].

- *Smoldyn*

www.smoldyn.org

Smoldyn est un autre simulateur stochastique qui l'avantage d'être très récent et qui donc met en œuvre des méthodes de calcul lourdes mais précises. Il calcule les positions des particules (molécules d'intérêt) dans un espace continu (de 1 à 3 dimensions) à partir des équations Einstein-Smoluchowski de description du mouvement brownien.

2.2. Agents, systèmes multi-agents et modélisation biologique

Comme il a été vu précédemment, de nombreux outils existent pour simuler ou modéliser des phénomènes biologiques, toutefois l'utilisation d'agents ou de systèmes multi-agents dans ce genre de domaine est relativement récente et peu de travaux existaient au démarrage du projet microMéga [Keele05]. Pourtant, dans [Paton93], la cellule avait été décrite comme un objet biologique complexe pouvant être vu comme un ordinateur parallèle distribué aux propriétés émergentes. Cette définition cadre parfaitement avec les propriétés des systèmes multi-agents qui peuvent, eux-mêmes, être vus comme des systèmes distribués composés d'entités fonctionnant en parallèles. Et en se focalisant sur les systèmes multi-agents auto-organiseurs, objets d'étude du projet microMéga, l'aspect « émergent » est aussi présent puisque le comportement collectif des entités composant ces systèmes produit une fonctionnalité émergente.

Toutefois depuis quelques années, le domaine de la biologie apparaît être un candidat potentiel d'application pour les technologies agents. Ainsi, les agents sont principalement utilisés en bioinformatique pour traiter et analyser les gigantesques quantités de données biologiques collectées [Luck06, Merelli06]. Les agents sont aussi utilisés afin de simuler des systèmes biologiques pour comprendre leur comportement ou réaliser des expérimentations *in silico/in virtuo* qui ne sont pas toujours faisables dans la réalité. Dans ce contexte, les travaux s'orientent autour de trois axes : la simulation ou la modélisation de réactions biologiques, la modélisation d'une partie du comportement de cellules ou bien la simulation du comportement de cellules en interaction au sein d'une population. D'autres travaux [Cannata05, Montagna06, Kawazu06, Ren08] s'orientent vers l'ingénierie de simulation de systèmes biologiques complexes ou la production de plates-formes. Même si les choses commencent à bouger, à l'heure actuelle, aucun système multi-agent ne peut se targuer d'avoir résolu le problème posé par microMéga, à savoir apprendre à modéliser le comportement fonctionnel d'un organisme unicellulaire.

2.2.1. Modélisation de réactions biologiques

Les premiers travaux à utiliser des agents, voire des systèmes multi-agents, ont été consacrés à la simulation de voies métaboliques.

Ainsi [Querrec03] s'intéresse à la simulation de la voie des MAP kinase (MAPK) qui constitue l'une des voies de signalisation les plus importantes au niveau cellulaire. Le principal apport du système multi-agent sous-jacent et du langage utilisé (oRis) est la possibilité de pouvoir en cours de simulation, perturber les cascades de la voie, en



simulant l'ajout d'un produit, et d'observer la réaction induite sur le comportement de la cellule. Ceci permet une expérimentation *in virtuo* et la compréhension d'un phénomène sans devoir redémarrer la simulation à nouveau.

Cellulat [Gonzalez00, Lagunez-Otero02, Gonzalez03] permet de construire un modèle de voies métaboliques de signalisation intracellulaire. Une cellule et ses composants (récepteur, protéine, etc.) sont vus comme des agents adaptatifs autonomes possédant chacun des capacités cognitives particulières et interagissant par l'intermédiaire d'un tableau noir. La cellule perçoit son environnement externe grâce à des récepteurs de surface et agit dessus en générant de nouveaux signaux internes qui pourront affecter le comportement de l'organisme à laquelle elle appartient. La cellule décide des relations entre signaux externes (molécules de signalisation) et signaux internes (signaux d'activation/inactivation) de manière à atteindre ses buts (réalisation de ses fonctions cellulaires). Une méthode est proposée pour construire un modèle de voies métaboliques (du style MAPK), toutefois le modèle ne se construit pas de lui-même.

[Khan03] utilise une plate-forme multi-agent (DECAF) pour simuler des réseaux biologiques (de transduction de signal). Contrairement aux systèmes d'équations différentielles, les systèmes multi-agents permettent le passage à l'échelle et la modélisation de grands réseaux. Cet article présente une stratégie hybride qui prend en compte des mesures du niveau macro (constantes de taux de réaction) pour calculer la cinétique stochastique au niveau de la molécule. Certaines propriétés physiques (diffusion, compartimentation) sont modélisées ce qui est difficile avec les équations différentielles. Un type de molécule est modélisée par un agent avec des structures de réseau de tâches hiérarchiques pour représenter les réactions initiées par lui ou l'extérieur. L'identité d'un agent est déterminée par un fichier de règles spécifiant les réactions auxquelles il peut participer et sa concentration initiale. Les réactions au sein du système sont déclenchées par la communication entre agents. Les résultats sont obtenus pour l'étude de la voie de facteur de croissance épidermique (EGF).

[Webb04, Webb06] décrivent CellAK un outil de modélisation d'une cellule qui a été validé en comparant les résultats obtenus pour la voie métabolique glycolytique avec ceux produits par le simulateur biochimique Gepasi [Mendes93]. Un diagramme UML exprime les relations d'héritage entre les différents composants d'une cellule eucaryote (membrane, cytoplasme, noyau) qui sont décomposés à leur tour (ainsi, une membrane contient une bicouche lipidique et des protéines de transport). Lorsqu'ils sont voisins dans la hiérarchie ainsi définie, les agents, instances de ces classes, sont capables d'échanger des messages afin de faire évoluer leur état. C'est la manière dont sont définis ces échanges de messages qui permet de définir les comportements des composants et celui du système modélisé. L'avantage souligné de cette approche est qu'une telle décomposition en agents est plus compréhensible que les systèmes d'équations différentielles utilisés dans Gepasi.

D'autres travaux se sont intéressés au problème de l'assemblage [Troisi04], du repliement [Armano05, Besse05] ou de l'appariement de molécules [Bortolussi05] mais leurs préoccupations sont un peu trop éloignées de celles de microMéga pour être détaillées ici.

Parmi les autres travaux en relation avec la simulation ou la modélisation de réactions ou de réseaux biologiques on trouvera aussi ceux de [Pogson06, Pogson08] avec la construction d'un modèle basé sur les agents pour les voies de signalisation NFkB, de [Amigoni07] qui étudie la capacité d'un système multi-agent à simuler une voie de transduction du signal ou une validation mathématique d'un système multi-agent appliqué au mécanisme de coagulation [Redou05].

2.2.2. Modélisation de cellules

L'objet de [Lints05] est de montrer la pertinence de l'utilisation des systèmes multi-agents afin de modéliser des organismes, y compris unicellulaires. Un modèle multi-agent d'une cellule bactérienne est développé, il est basé sur le modèle biologique du titrage de DnaA et traite de la réplication de l'Adn et de la division cellulaire. L'article part du

postulat que les systèmes d'équations différentielles sont peu compréhensibles au modélisateur moyen et que les caractéristiques des agents (autonomie, réactivité, proactivité et sociabilité) dotent les systèmes multi-agents d'avantages dans le contexte de la modélisation cellulaire. Il est ainsi possible d'avoir une correspondance directe entre le système biologique et les parties du modèle, ce qui est une approche plus intuitive qu'un modèle mathématique. Le modèle est plus flexible car les agents sont autonomes donc faciles à ajouter, enlever, modifier, distribuer en cours d'exécution de la simulation. De plus, différentes parties d'une même simulation peuvent être représentées avec différents niveaux d'abstraction pour ajouter de la flexibilité au modèle. Enfin, il est facile de produire un comportement émergent dans le modèle car les systèmes à agents s'appuient sur les concepts d'interaction et de parallélisme.

D'autres travaux impliquant la technologie agent visent à modéliser plus particulièrement le comportement de cellules. C'est le cas d'AgentCell [Emonet05] qui est un modèle basé sur la technologie agent destiné à l'étude des relations entre les procédés stochastiques intracellulaires et le comportement individuel des cellules. Le cas d'étude est le chimiotactisme d'une bactérie (E. Coli). Chaque bactérie est représentée par un agent équipé de son propre réseau de chimiotactisme, de moteurs et d'un flagelle. Les cellules qui nagent sont libres de se déplacer dans un environnement 3D. L'objectif d'AgentCell est d'étudier la manière dont le bruit moléculaire influence le comportement de cellules nageant dans un tel environnement 3D. Des essais reproduisent les données expérimentales obtenues pour des cellules individuelles et des populations bactériennes. L'implantation de la simulation a été effectuée sous Repast. Bien que ce travail utilise un agent pour représenter une cellule, les composants de cette cellule (flagelle, moteur, etc.) restent de simples objets dépourvus d'autonomie ou de capacités d'apprentissage, ils suivent un algorithme établi pour reproduire les données expérimentales et n'ont pas pour objectif d'expliquer les phénomènes qui se produisent au sein de la cellule.

L'objectif décrit dans [Bosse06, Jonker07, Jonker08] est similaire, à savoir utiliser la simulation agent afin de modéliser le fonctionnement d'une bactérie et s'abstraire d'intégrations coûteuses de systèmes d'équations différentielles pour lesquelles certains paramètres ne sont pas toujours obtenus expérimentalement, voire connus. La structure de la cellule étudiée (Escherichia Coli) est figée, l'objectif principal étant d'étudier la manière dont ses propriétés dynamiques se comportent en fonction de propriétés de niveau inférieur. Une architecture Agent-Groupe-Rôle (AGR) [Ferber98] est utilisée dans [Bosse06] pour décrire l'organisation structurelle interne d'un organisme unicellulaire, architecture ensuite étendue afin de formuler des propriétés dynamiques au sein de cette organisation. [Jonker07] complète ce travail en utilisant des concepts agents afin de modéliser la dynamique à deux niveaux : au sein d'un agent (BDI) et au niveau de son comportement observable extérieurement. Dans [Jonker08], des relations temporelles entre propriétés cellulaires sont définies en se basant sur une connaissance biochimique et des simulations sont effectuées dans lesquelles il est montré que la cellule est capable d'atteindre un état interne stable comparable à la réalité pour un environnement stable et qu'elle est capable de s'adapter à des changements relativement espacés dans le temps si son environnement est dynamique. Dans les autres cas, elle peut osciller entre des états stables, ce qui est conforme aussi à la réalité, ou osciller sans atteindre d'état stable. Des hypothèses ont été faites quant aux paramètres temporels utilisés dans les relations temporelles entre propriétés d'état intentionnel de la cellule et aux concentrations des substances intervenant dans la voie biochimique étudiée. L'approche semble trop complexe pour simuler la totalité du comportement de la cellule ; du moins, pour décrire toutes les relations temporelles des différentes voies au sein de la cellule. De plus, ces relations temporelles sont paramétrées par des durées ou des concentrations, paramètres qui ne sont pas forcément parfaitement connus ou que l'on ne peut pas obtenir de manière expérimentale. L'approche proposée ne semble pas répondre à cet inconvénient précis qu'il souligne pour d'autres approches de modélisation. Si la cellule est capable d'adapter son comportement à son environnement, rien ne peut laisser espérer que de nouvelles interactions non exprimées dans le modèle temporel apparaissent lors de cette adaptation. Ce modèle est figé, parfaitement défini. Il n'y a donc aucune forme d'apprentissage ou d'émergence. Ces travaux se focalisent plus sur les aspects



relationnels entre dynamiques de niveau différents que sur la volonté de produire un modèle fonctionnel d'une cellule. L'approche employée bien que pouvant aider à comprendre les relations entre agent et multi-agent ou micro et macro niveaux, ne cherche pas particulièrement à aider les biologistes en leur fournissant des outils leur permettant de découvrir de nouveaux phénomènes au sein d'une cellule.

Dans le cadre du projet MitoScop, dont l'objet est de modéliser le métabolisme d'une organelle intracellulaire, les travaux présentés dans [Lales05, Lales07] cherchent à simuler une membrane cellulaire. À ces fins, des agents réactifs ont été utilisés afin de simuler, entre autres, les différents lipides que l'on peut trouver dans une membrane. Il est ainsi possible de réaliser des bicouches avec deux types de lipides aux propriétés différentes. L'apport des systèmes multi-agents est, ici encore, la capacité à décrire des situations qui ne peuvent être décrites par des systèmes d'équations différentielles (cas moyen inconnu, association transitoire de molécules, influence de leur conformation, perturbation du réseau, trop grand nombre d'équations, etc.), le fait que le réseau à modéliser peut se percevoir comme un système composé d'entités en interaction et donc peut correspondre à la définition même d'un système multi-agent.

Parmi les autres travaux référencés, on trouve celui décrit dans [Shafaei08] qui utilise un système multi-agent holonique pour simuler une cellule bien que la présentation soit, pour l'instant, limitée à la modélisation d'un réseau biologique. Le fait qu'une cellule puisse être vue comme un système ouvert et complexe et la structure hiérarchique d'un organisme, d'un organe, d'une cellule, d'une réaction ont justifié le choix de la structure du système multi-agent établissant le modèle biologique puisqu'un holon peut-être vu comme un agent à structure récursive. Cette dernière approche est sans doute celle qui se a le plus de points communs avec microMéga, du moins dans la gestion des réactions (transformation, transport et liaison) qui y sont vues comme des interactions entre deux holons et un ensemble de règles. Chaque holon gère sa vitesse et sa stoechiométrie par inférence sur le moteur de règles. Pour se rapprocher de situations réelles, l'utilisateur peut, en cours de simulation, ajouter ou enlever des substrats ou des produits, changer les règles ou réguler la meilleure vitesse pour une réaction ; cette dernière pouvant aussi être apprise à l'aide de réseaux de neurones ou d'algorithmes génétiques. Toutefois, les détails donnés sur la manière dont l'apprentissage est effectué au niveau des holons sont insuffisants pour en conclure que le modèle s'auto-construit.

2.2.3. Modélisation de population de cellules

Le comportement des cellules souches a été simulé et étudié dans [Inverno05a, Inverno05b, Gatti08]. L'objectif des travaux décrits dans [Inverno05a, Inverno05b] est de modéliser le comportement de telles cellules dans une niche en utilisant un système multi-agent afin d'étudier l'organisation globale émergente de cette population. La problématique dans ce cas n'est pas de modéliser les processus internes d'une cellule mais de modéliser pourquoi et comment elle se divise et/ou se différencie ou non dans un environnement donné dans lequel d'autres cellules similaires évoluent. Cependant ce travail vise aussi à prouver que, dans un tel contexte, les agents sont des outils de modélisation plus adéquats que les automates cellulaires ou les fonctions mathématiques. En effet, les systèmes multi-agents sont un moyen efficace de modéliser des systèmes complexes dont le comportement global émerge au niveau macro des interactions du niveau micro. Une telle approche ascendante peut alors favoriser la connaissance des mécanismes d'auto-organisation sous-jacents et des mécanismes biologiques mis en jeu. [Conceicao08, Gatti08, Gatti07] proposent une structure s'appuyant sur des agents pour simuler différents types de cellules, en prenant pour exemple particulier, des cellules souches. Les conclusions de cet article sont similaires à celles précédemment exposées : les systèmes multi-agents sont un outil qu'il est plus facile à aborder que les systèmes d'équations différentielles pour des médecins, en particulier, qui souhaiteraient expérimenter virtuellement pour des raisons de coût, de temps ou d'éthique. De plus, ces systèmes multi-agents représentent sans doute une solution à l'explosion du nombre d'équations différentielles qui devient inévitable lorsque la modélisation porte sur un système biologique complexe. Enfin, les phénomènes émergents que l'on veut étudier

peuvent être modélisés à partir de règles simples gouvernant le comportement de chaque agent du système.

L'étude de la formation de tissus est aussi un point abordé dans les travaux utilisant la technologie agent. Ainsi [Santos04], un tissu est vu comme un système multi-agent, une agrégation de cellules où chaque cellule est représentée par un agent. Les échanges chimiques au sein de ce tissu sont modélisés par une communication entre agents. Un contexte environnemental maintient l'état du système et fournit les perceptions aux agents. Cet environnement reflète aussi les actions (se déplacer, mourir, se diviser, réaliser un échange chimique) que les agents décident de réaliser en fonction de leurs perceptions et des buts (multiples et éventuellement conflictuels) qu'ils doivent réaliser (correspondant aux fonctions des cellules). Cette modélisation avait pour objectif d'étudier un mécanisme d'auto-organisation dans un tissu : la formation d'un groupe de cellules ayant un but commun. En effet, une cellule saine qui détecte une cellule tumorale convertit d'autres cellules saines voisines pour qu'elles vérifient cette détection et forment ainsi un groupe.

[Guo07] présente un cadre de modélisation et de simulation de systèmes adaptatifs complexes en fournissant un ensemble de modèles individuels et un modèle d'interactions. Ce modèle, plus précisément orienté vers une capacité d'adaptation à court terme, est validé par l'étude de phénomènes d'auto-organisation se produisant pendant le cycle de vie d'un protozoaire mycétozoaire, *Dictyostelium*. Son cycle de vie se caractérise par une alternance de formes unicellulaires et de phases pluricellulaires ; l'agrégation, qui est le processus de transformation entre les deux formes, peut servir à étudier les relations entre une auto-organisation multi-cellulaire et un comportement unicellulaire. La population de *Dictyostelium* est vue comme un système complexe adaptatif auquel correspond un système multi-agent ; chaque cellule correspondant alors à un agent. L'argument mis en avant pour utiliser un modèle s'appuyant sur un système multi-agent par rapport à un modèle purement conceptuel (qui peut donner des indications permettant de comprendre des mécanismes sous-jacents) est sa capacité de pouvoir être relié à des phénomènes ou des faits biologiques réels.

Si la plupart des travaux énumérés ici sont d'accord pour considérer les technologies agents ou systèmes multi-agents comme des outils pertinents pour construire des systèmes aussi complexes et flexibles que les systèmes biologiques, [Amigoni07] souligne aussi un inconvénient : la plupart du temps, les résultats obtenus ne sont pas validés en les comparant avec des données expérimentales réelles. Ces données sont en effet, peu souvent disponibles en pratique et, lorsqu'elles le sont, ne sont pas forcément facilement interprétables ; c'est pourquoi de nouvelles méthodes de validation sont nécessaires.

3. La plate-forme de modélisation de microMéga

3.1. Approche adoptée

Un système multi-agent peut-être simplement défini comme un système formé d'entités autonomes (nommés agents) qui interagissent au sein d'un environnement commun afin de réaliser une fonction collective.

L'équipe SMAC s'intéresse à la mise au point de fonctions complexes pour lesquelles l'algorithme permettant de les réaliser n'est pas connu *a priori*. En ce sens, le projet microMéga qui vise à fournir un modèle du fonctionnement d'un micro-organisme, tel que *Saccharomyces Cerevisiae*, à partir de données expérimentales constitue un exemple d'application idéal puisque les éléments constitutifs de la cellule cible sont nombreux, leurs interactions le sont aussi, la fonction du micro-organisme cible est complexe et n'est pas connue dans sa totalité. Le fait de devoir prendre en compte des données expérimentales quantitativement et qualitativement hétérogènes (données sur les transcrits et sur les éléments macroscopiques) et parfois entachées de bruit de mesure, et de considérer les milliers d'éléments mis en jeu au sein de la cellule constitue un défi sur le plan informatique.



Comme il a été souligné dans l'introduction, le modèle recherché doit s'auto-construire pour être assez dynamique et flexible. L'approche adoptée est donc une approche ascendante qui consiste à identifier les composants de base du modèle et à les laisser interagir de façon à trouver une bonne organisation, celle qui conduit au modèle fonctionnel attendu. Ce modèle est implanté sous la forme d'un système multi-agent auto-organisateur et adaptatif dans lequel les composants de base précédemment identifiés ont été « agentifiés » [Bernon09].

Dans un tel système, selon la théorie des AMAS [Camps98], la fonction globale du système est une fonction émergeant des interactions entre agents du système. De plus, la théorie assure qu'en dotant les agents de comportements dits coopératifs, cette fonction globale sera celle effectivement souhaitée. Ce comportement peut se dissocier en un comportement nominal et un comportement coopératif (ou adaptatif) :

- Le comportement nominal d'un agent correspond à sa fonctionnalité de base.
- Le comportement coopératif de l'agent lui permet de faire face à des situations qui sont néfastes pour l'attitude coopérative qu'il doit exhiber envers son environnement. Ce comportement coopératif consiste à :
 - Tenter de résorber ces situations néfastes en essayant d'abord d'ajuster ses paramètres de fonctionnement (comportement de réglage) ;
 - Essayer de reconsidérer ses liens avec les autres, si le comportement de réglage ne suffit pas à résoudre le problème (comportement de réorganisation) ;
 - Provoquer la création d'autres agents ou bien à reconsidérer sa place au sein du système en disparaissant si un agent se juge complètement inutile (comportement d'évolution).

Cet état coopératif est jugé d'un point de vue purement local à l'agent. Cela signifie que globalement, le système multi-agent ne saura pas déterminer si son comportement collectif correspond à la fonctionnalité attendue par l'utilisateur. C'est à ce dernier que revient la charge de décider de l'adéquation du modèle obtenu. Lorsque le comportement coopératif des agents est celui qui permet d'atteindre la fonction collective souhaitée, il doit donc être possible de « figer » les comportements et par là même, de figer la structure du modèle obtenu.

La théorie des AMAS a été, au travers de différentes applications, validée empiriquement dans différents domaines tels que l'aéronautique [Welcomme08, Capera05], la prévision de crues [Georgé03], le commerce électronique [Gleizes02], etc. et microMéga est l'occasion de s'intéresser au domaine de la biologie computationnelle.

3.2. Composants de base d'un modèle

Les différents composants nécessaires à la modélisation fonctionnelle du micro-organisme cible ont été identifiés en interaction avec nos partenaires biologistes.

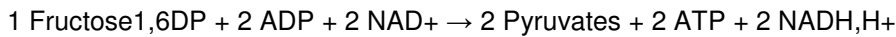
En considérant le fait que les données expérimentales sont incomplètes et hétérogènes et que les réactions intracellulaires sont complexes et imparfaitement connues, l'architecture de la plate-forme se devait d'être stable et assez souple pour supporter des remises en cause éventuelles. C'est pourquoi, les composants ont été identifiés au sein des éléments de base du réseau d'interactions locales de la cellule, à savoir les gènes, les protéines, les métabolites intracellulaires, les sucres de réserve et la respiration. Au final, différents types d'agents ont été identifiés :

- Des agents représentant les éléments chimiques de base de la cellule, une quantité leur est associée ;
- Des agents représentant les réactions chimiques pouvant se produire au sein de la cellule. Ces agents sont de quatre types :
 - Les transporteurs qui déplacent une quantité d'un élément d'un compartiment à l'autre, ce transport peut être actif (consommant de l'ATP) ou passif ;

- Les catalyseurs qui transforment une quantité d'un élément en deux autres, cette catalyse peut être éventuellement régulée ;
- Les synthétiseurs qui effectuent la transformation inverse d'un catalyseur en assemblant deux éléments pour en produire un, une régulation est aussi possible ;
- Les gènes qui doivent corroborer les données sur les transcrits et qui sont vus comme des réactions particulières.

Les « agents-éléments » et les « agents-réactions » sont appelés « agents fonctionnels ».

Combiner ces différents agents dans un réseau d'interactions permet de construire des réactions plus complexes. Par exemple, une réaction telle que :



se traduira par le réseau de la Figure 1. Les carrés y sont des agents éléments, les étoiles sont des réactions de synthèse et les octogones sont des réactions de catalyse qui consomment ou produisent les éléments (relations exprimées par les liens pleins). Un lien pointillé traduit la régulation d'une réaction par un élément (ici, c'est une protéine PE3 qui joue le rôle de régulateur de la réaction transformant du fructose en pyruvate).

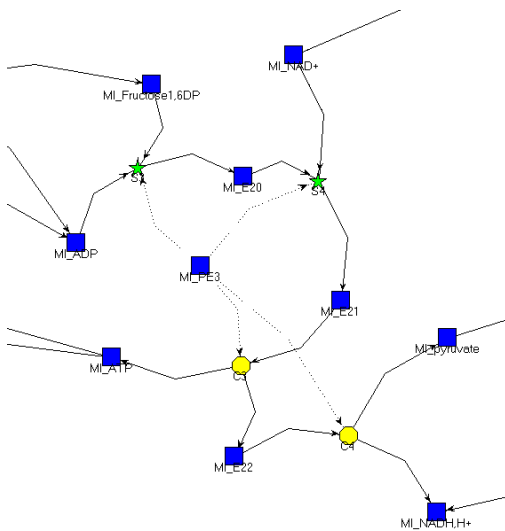


Figure 1. Exemple de réaction.

Enfin, il est aussi nécessaire de pouvoir interagir avec la plate-forme dans le but d'observer ou d'interpréter le modèle en fonctionnement, ou bien d'intégrer les données expérimentales sur lesquelles le modèle qui se construit va apprendre. Enfin, il est intéressant d'avoir un moyen d'influencer la structure du modèle en cours de construction, de le contraindre sur certains aspects qui sont biologiquement connus et qu'il est inutile de redécouvrir. Par exemple, le nombre de

compartiments cellulaires est *a priori* connu et agir sur le modèle pour lui faire respecter cette contrainte est utile. Un type d'agents supplémentaire, agents qui n'existent que pour exercer une pression sur le modèle ou bien l'observer, a donc été identifié, il s'agit des « agents-vues ».

compartiments cellulaires est *a priori* connu et agir sur le modèle pour lui faire respecter cette contrainte est utile. Un type d'agents supplémentaire, agents qui n'existent que pour exercer une pression sur le modèle ou bien l'observer, a donc été identifié, il s'agit des « agents-vues ».

3.3. Comportement des agents

Une fois les agents identifiés, leurs comportements nominal et coopératif ont été définis.

3.3.1. Comportement nominal

Le comportement nominal d'un agent est fondé sur un cycle de vie classique « percevoir, décider et agir » et correspond à ce qu'un agent fait lorsqu'aucune action d'apprentissage n'est effectuée.

Le comportement nominal d'un agent-réaction consiste à calculer la quantité consommée et/ou produite pour chaque agent-élément auquel il est lié et à la faire connaître à l'élément en question. Le calcul de la quantité dépend des valeurs stœchiométriques pour l'élément considéré et de la vitesse de réaction. Cette vitesse de réaction est fonction des valeurs des régulateurs i.e. qu'on associe une vitesse à chaque intervalle de valeurs possibles pour les régulateurs. La vitesse choisie pour une réaction à un instant donné dépend donc du contexte dans lequel s'effectue cette réaction.

Le comportement nominal d'un agent-élément consiste simplement à mettre à jour sa quantité en fonction des activités de consommation ou de production des agents-réactions auxquels il est lié.



Le comportement nominal d'un agent-vue consiste soit à injecter des données expérimentales dans le système, soit à récupérer les données concernant les agents fonctionnels et à les stocker. Certains agents-vues, qui connaissent les données expérimentales, peuvent ainsi permettre aux agents-éléments d'avoir ont un retour (« feedback ») sur ce qu'ils font.

3.3.2. Comportement coopératif

Au stade actuel de développement de la plate-forme, seuls les comportements de réglage et de réorganisation ont été définis, le comportement d'évolution étant encore à l'étude.

- *Phase de réglage*

La première composante du comportement coopératif consiste donc pour un agent à analyser son comportement nominal pour déterminer si certaines situations sont potentiellement néfastes pour son état coopératif.

Un agent-réaction peut rencontrer deux types de situations : il se trouve dans un contexte de fonctionnement qui n'a pas encore été rencontré jusque là et ne peut donc trouver une vitesse associée, ou bien, un agent-élément auquel il est lié semble ne pas accepter la manière dont il est consommé ou produit par cette réaction. Dans le premier cas (situation d'improductivité selon la théorie des AMAS), l'agent-réaction essaie de résoudre cette situation en créant un nouveau contexte qui correspond à la nouvelle situation. Dans le second (situation d'incompétence), il va essayer de modifier sa stœchiométrie ou sa vitesse de réaction pour satisfaire l'agent-élément concerné. Modifier sa stœchiométrie n'est pas encore réalisable à ce stade du développement, l'agent va donc essayer d'ajuster sa vitesse en changeant la vitesse du contexte qu'il est en train d'utiliser (contexte courant) ou bien en essayant de changer de contexte. Ce changement de contexte peut se réaliser en faisant évoluer les intervalles des régulateurs du contexte courant ou en propageant cette demande de modification vers les agents-éléments qui régulent la réaction (ceci dans le but d'influencer le choix d'un meilleur contexte de fonctionnement).

Un agent-vue qui sert d'interface avec les données expérimentales calcule, lors de son comportement nominal, la différence entre la quantité théorique et la quantité réelle d'un agent élément. Si une différence est détectée, il considère ceci comme une situation de conflit lors de sa phase de réglage et essaie de résoudre ce problème en avertissant l'agent-élément concerné. Au contraire, s'il juge que cet agent-élément possède la bonne quantité, l'agent-vue lui enverra un message de renforcement pour le conforter dans son comportement actuel.

Lors de son comportement nominal, un agent-élément peut faire face à une situation d'incompétence endogène lorsque sa quantité devient négative ce qui n'a pas de sens, ou bien exogène dans le cas où sa quantité courante ne correspond pas à la quantité demandée par un agent-vue (qui reflète les données expérimentales) ou un agent-réaction que cet élément régule (pour trouver un meilleur contexte de fonctionnement). Dans tous les cas, l'agent-élément réagit en fixant sa quantité à celle demandée (ou à 0 si cette quantité était négative) et en propageant son problème aux agents-réactions en leur demandant d'augmenter ou de diminuer leur vitesse de réaction en conséquence.

- *Phase de réorganisation*

Le comportement de réglage d'un agent peut ne pas entièrement résoudre les situations jugées néfastes qui ont été détectées lors du comportement nominal ou en faire apparaître de nouvelles. C'est le comportement de réorganisation d'un agent qui va donc essayer de résorber ces situations. De manière générale, ces situations peuvent, selon la classification donnée par la théorie des AMAS, être de trois types : inutilité totale, inutilité partielle ou incompétence.

Dans le premier cas, un agent est inutile s'il n'a pas de lien avec d'autres agents. Théoriquement, pour résoudre ce problème, l'agent recherche des agents avec lesquels interagir. C'est le cas, par exemple, lorsqu'un agent-élément n'est relié à aucun agent-

réaction, il va diffuser un message pour rechercher un agent-réaction susceptible d'être intéressé par une consommation ou une production de cet élément. Il en est de même avec un agent-réaction qui ne serait lié à aucun agent-élément ; être isolé n'a aucun sens pour réaliser son comportement nominal et il devra trouver un (ou des) agent(s)-élément(s) avec lesquels interagir.

L'inutilité devient partielle lorsqu'un agent ne peut réaliser son comportement nominal car des données lui manquent. De même que précédemment, il doit trouver de nouveaux partenaires pour acquérir ces informations. Par exemple, un agent-élément est produit par un agent-réaction mais il n'est lié à aucun autre agent-réaction. Ou bien, un agent-réaction consomme bien un (ou des) agent(s)-élément(s) mais ne produit rien.

Enfin, une situation d'incompétence apparaît lorsqu'un agent est incapable d'atteindre un comportement satisfaisant les autres agents (ces derniers continuent à demander des changements à cet agent) et que le comportement de réglage ne suffit pas, bien entendu, à résoudre le problème. Cet agent doit donc remettre en question les agents avec lesquels il interagit, soit en choisissant de ne plus interagir avec eux, soit en essayant de modifier ses liens actuels. C'est le cas lorsqu'un agent-réaction est confronté à des demandes contradictoires de modification de la vitesse associée à un certain contexte de fonctionnement et qu'il ne peut donc pas ajuster cette vitesse. Pour l'aider à distinguer les différentes situations, il peut chercher de nouveaux régulateurs. Un agent-élément qui n'est pas arrivé à modifier sa quantité de manière satisfaisante jusque-là peut remettre en question les liens de régulation qu'il possède en considérant que sa confiance envers un agent-réaction n'est plus fondée et qu'il aurait plutôt avantage à réguler une (d')autre(s) réaction(s).

3.4. Exemples de modèles

La plate-forme microMéga est opérationnelle et permet, au stade actuel de son développement, de construire manuellement, à partir de l'expertise d'un biologiste, des modèles biologiques dynamiques fonctionnels. Les outils fournis permettent d'effectuer des simulations grâce à un modèle afin d'obtenir et visualiser des résultats *in silico*, de modifier un modèle et de le sauvegarder.

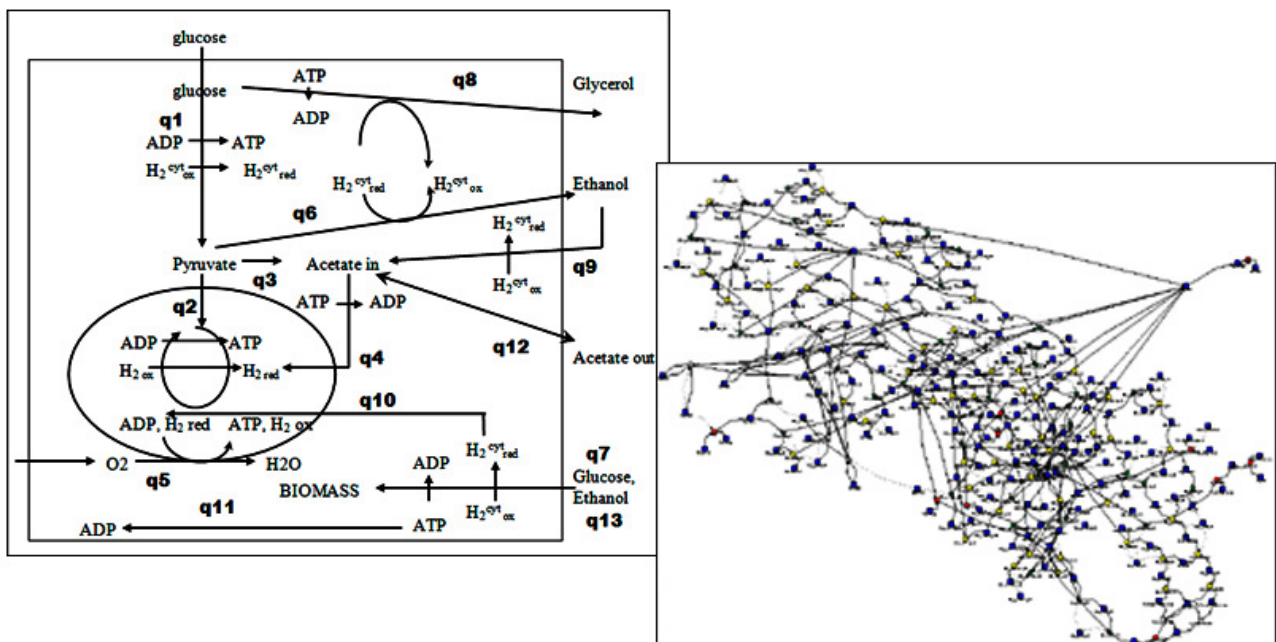


Figure 2. Modèle simplifié de la glycolyse chez *Saccharomyces Cerevisiae* (à gauche) et réseau généré sous la plate-forme microMéga (à droite).

Cela permet de construire des modèles semblables à celui visible sur la Figure 2. Il s'agit d'un modèle simplifié de la glycolyse chez *Saccharomyces Cerevisiae*, en culture selon plusieurs modes opératoires. Les résultats associés à des simulations effectuées avec ce modèle sont présentés dans [Bernon09].

Pour illustrer notre propos, nous présenterons un réseau plus simple visualisé en Figure 3. Ce réseau représente un modèle simulant les mécanismes de respiration et de photosynthèse des végétaux. Il est important de noter qu'il ne s'agit pas d'une modélisation biologique fidèle de ce phénomène complexe, mais plutôt d'une image pertinente de sa dynamique. Les réactions présentes dans cette simulation, la photosynthèse, d'une part, et la respiration, d'autre part, sont elles-mêmes décomposées en deux « sous-réactions » formant l'ensemble décrit en Table 1.

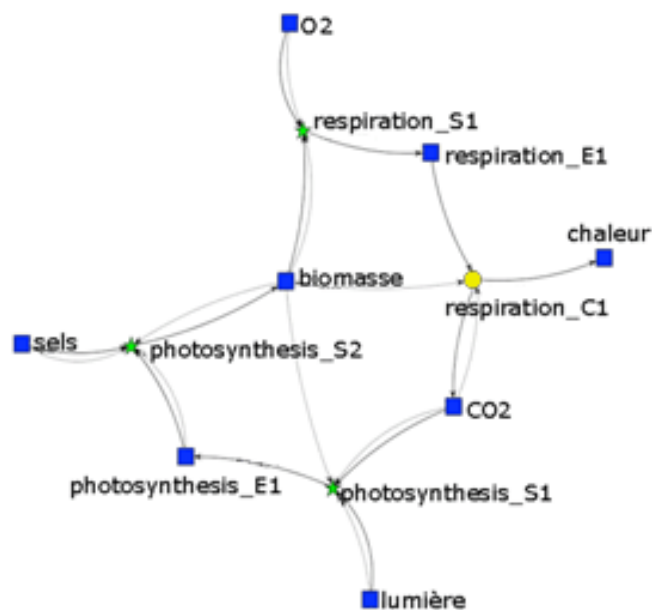


Figure 3. Réseau microMega correspondant à un modèle de photosynthèse/respiration.

La photosynthèse consomme de la *Lumière*, du CO_2 et des *Sels* afin de produire de la *Biomasse*, tandis que la respiration consomme cette *Biomasse* et de l' O_2 afin de produire du CO_2 et de la *Chaleur*. À cela s'ajoutent de nombreux liens de régulation, représentés en pointillés sur la Figure 3, qui décrivent l'impact des différents éléments, n'étant ni produits ni consommés, sur les réactions. Par exemple, la quantité de *Biomasse*, bien que n'apparaissant pas dans la réaction *Photosynthesis_S1* comme un élément produit ou consommé, régule la quantité de l'élément *Photosynthesis_E1* créé.

| Réactions | Consommés | Produits |
|-------------------|--------------------------|-------------------|
| Photosynthesis_S1 | CO_2 + Lumière | Photosynthesis_E1 |
| Photosynthesis_S2 | Photosynthesis_E1 + Sels | Biomasse |
| Respiration_S1 | Biomasse + O_2 | Respiration_E1 |
| Respiration_C1 | Respiration_E1 | Chaleur + CO_2 |

Table 1. Liste des réactions du modèle de la photosynthèse.

La *Lumière* est l'élément central de cette simulation, sa présence permettant la transition entre les réactions de respiration et de photosynthèse, réactions ne pouvant avoir lieu simultanément. De ce fait, la simulation considère initialement la lumière allumée (modélisée par une valeur fixe arbitrairement haute), et l'éteint (valeur nulle) à $t=90$ afin d'observer la réaction du système. Notons que les quantités initiales des différents éléments apparaissant dans cette simulation ont été déterminées empiriquement.

La Figure 4 fournit les résultats du déroulement de l'intégralité de la simulation. Lorsque la lumière est présente, nous observons une forte augmentation de la quantité de *biomasse*, ainsi qu'une diminution de la quantité de CO_2 . Ces indicateurs mettent en évidence le déroulement des réactions de photosynthèse.

A partir de $t=90$, la quantité de lumière est réduite jusqu'à son extinction. Les conséquences de cette action sont d'une part, la stabilisation, puis la diminution de la biomasse, et d'autre part, l'augmentation de la quantité de CO_2 du système. Ainsi, nous

observons les signes de la présence des réactions de respiration. Les effets correspondant au déroulement de la photosynthèse n'apparaissant plus dans les résultats, nous constatons que le système s'est adapté aux différents éléments présents afin de produire la réaction appropriée.

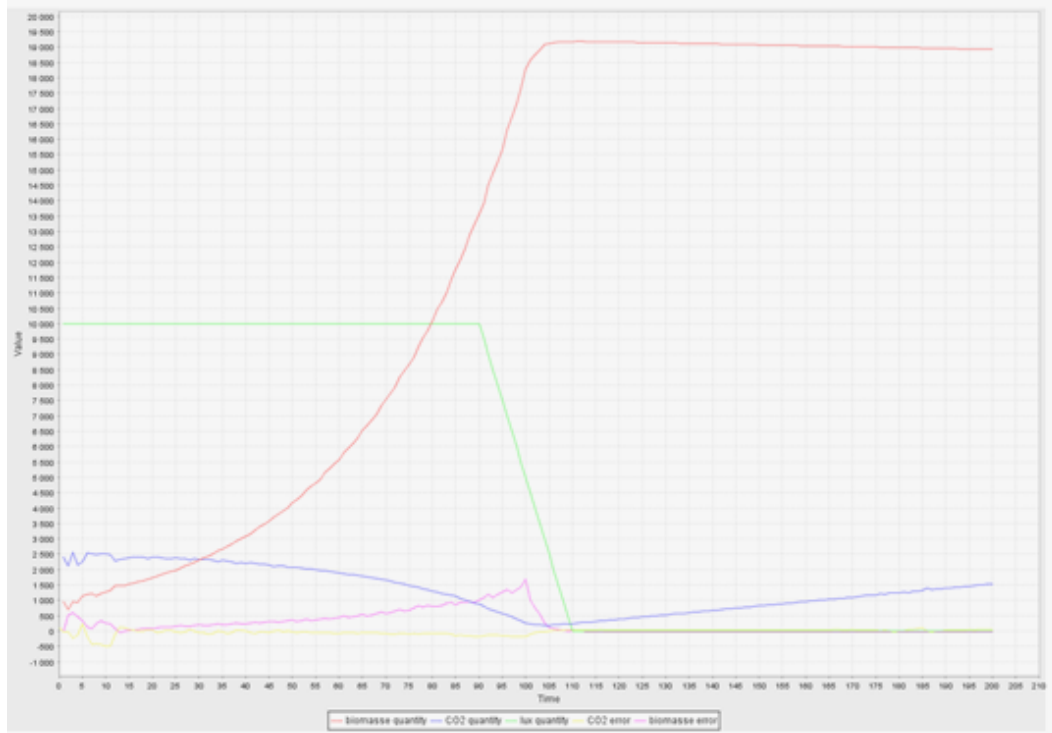


Figure 4. Résultats de la simulation de la photosynthèse/respiration.

4. Conclusion

Un modèle construit à l'aide de microMéga peut être présenté comme des centaines de modèles simplifiés s'exécutant en parallèle. Pour chaque situation, la dynamique du modèle global s'exprime par l'expression d'une cinétique de transformation de certaines quantités d'éléments en d'autres quantités d'éléments. La variation des dynamiques au cours d'une simulation est prise en compte par l'existence simultanée au sein des agents fonctionnels de toutes les cinétiques potentielles activables en fonction du contexte. En cela, on s'approche d'une formalisation à événements discrets (DEVS) où les phases de transition ne sont pas gérées par les mêmes modèles que les états d'équilibre. microMéga offre l'avantage d'un même moteur de simulation quelle que soit la phase du système à décrire, la complexité du système biologique modélisé se traduisant par un accroissement du nombre d'agents et du nombre de contextes « appris » par les agents.

L'apprentissage d'un modèle dynamique soulève en soi la problématique de l'ajustement de paramètres. En effet, le système doit apprendre pour chacun de ses paramètres la valeur que celui-ci doit avoir à un moment donné. Toutefois, cette valeur change au cours du temps et le système doit en permanence ajuster ses valeurs. Pour pallier cette dépendance, nous avons mis en place un apprentissage contextuel. Ainsi dans un contexte de fonctionnement donné, les éléments et fonctions du système apprennent quelles valeurs affecter à leurs paramètres. Cette approche nous permet de construire des modèles dynamiques associant la combinaison de plusieurs états stables, pour voir l'émergence de dynamiques spécifiques dues à l'enchaînement temporel de contextes de fonctionnement différents (une cinétique moyenne apparaît de l'enchaînement d'au moins deux valeurs de vitesses différentes). Cependant notre première approche s'est révélée improductive du fait que les agents fonctionnels devaient, pour apprendre un contexte, simultanément identifier les plages de valeurs définissant le contexte et y associer une valeur de vitesse. Or, cette valeur de vitesse



ayant elle-même un impact plus ou moins direct sur les valeurs définissant le contexte, il s'ensuit des phénomènes d'amplification des erreurs et du bruit liés à ce couplage qui peuvent faire diverger le système de sa fonction. Étant donné l'intérêt fondamental de la résolution de ce type de problème, nous poursuivons le développement d'un nouveau type d'apprentissage des contextes, basé cette fois, sur une discrétisation préalable, établie en fonction du bruit supposé des valeurs des paramètres. Les contextes deviennent des entités moins labiles qui auront simplement à associer une vitesse à une plage de valeurs. Pendant l'apprentissage, le couplage amenant à un glissement des plages de valeurs des paramètres n'affectera pas un contexte devenu caduc, mais générera l'affectation de la bonne vitesse à un nouveau contexte de fonctionnement.

Parvenir à un modèle qui s'auto-construit réellement semble possible, lorsque le comportement coopératif des agents et la mise au point des contextes de fonctionnement des réactions seront validés.

Comme le souligne [Amigoni07], la plupart du temps, les modèles biologiques basés sur des systèmes à agents ou multi-agent ne sont pas validés en comparant les résultats obtenus à des données expérimentales. Ceci s'explique par le fait que ces données ne sont pas forcément disponibles en pratiques ou que les interpréter n'est pas chose facile. L'un des points forts de l'approche mise en jeu dans microMéga est que la validation des modèles issus de la plate-forme de modélisation devrait s'effectuer sur des données expérimentales non utilisées lors de leur phase d'auto-construction, cette validation pouvant être réalisée par des informaticiens en collaboration avec des experts du domaine biologique.

5. Références bibliographiques⁴

[Hazen05] Robert M. Hazen (2005), *Genesis. The Scientific Quest for Life's Origin*, Washington: Joseph Henry Press.

[Holland98] John Holland (1998), *Emergence: From Chaos to Order*, Cambridge, MA: Perseus Books.

[Morange05] Michel Morange (2005), *Les secrets du vivant. Contre la pensée unique en biologie*, Paris: La Découverte.

[O'Malley05] Maureen O'Malley, John Dupré (2005), *Fundamental Issues in Systems Biology, BioEssays*, 27, 1270-1276.

[Maturana73] Humberto Maturana, Francisco Varela (1973) *Autopoiesis: The Organization of the Living*, Dordrecht.

5.1. Travaux en relation avec les technologies autres qu'agents

[Alon06*] U. Alon (2006) *An Introduction to Systems Biology: Design Principles of Biological Circuits*, Chapman & Hall.

[Bryant01] C. H. Bryant, S. H. Muggleton, S. G. Oliver, D. B. Kell, P. Reiser, R. D. King (2001) Combining Active Inductive Programming, Active Learning and Robotics to Discover the Function of Genes, *Electron. Trans. Artif. Intell.*, 5, 1-36.

[Castrillo04] J. I. Castrillo, S. G. Oliver (2004) Yeast as a Touchstone in Post-genomic Research: Strategies for Integrative Analysis in Functional Genomics, *Journal of Biochemistry and Molecular Biology*, 37(1), January 2004, 93-106.

[Cornell03] M. Cornell, N. W. Paton, C. Hedeler, P. Kirby, D. Delneri, A. Hayes, S. G. Oliver (2003) GIMS: An Integrated Data Storage and Analysis Environment for Genomic and Functional Data, *Yeast* 20, 1291-1306.

[DeJong02*] H. De Jong (2002) Modelling and Regulation of Genetic Regulatory Systems: a Literature Review, *Journal of Computational Biology*, 9, 67-103.

⁴ Certaines références listées ici ne sont pas citées dans le texte ci-dessus (leur nom se termine par un astérisque). Toutefois, elles ont été utilisées dans les publications liées à la plate-forme microMéga et, à ce titre, ont été jugées d'intérêt pour le lecteur.

- [Delneri01*] D. Delneri, F. L. Brancia, S. G. Oliver (2001) Towards a Truly Integrative Biology through the Functional Genomics of Yeast, *Current Opinion in Biotechnology*, 12(1), February 2001, 87-91.
- [Doniger03] S. Doniger, N. Salomonis, K. D. Dahlquist, K. Vranizan, S. C. Lawlor, B. R. Conklin (2003) MAPPFinder: Using Gene Ontology and genMAPP to Create a Global Gene-expression Profile from Microarray Data, *Genome Biol.*, 4, R7. On-line publication (<http://genomebiology.com/2003/4/1/R7>).
- [Efroni03*] S. Efroni, D. Harel, I.R. Cohen (2003) Toward Rigorous Comprehension of Biological Complexity: Modeling, Execution, and Visualization of Thymic T-Cell Maturation, *Genome Research*, 13, 2485-2497.
- [Fang03] Y. Fang, A. Brass, D. C. Hoyle, A. Hayes, A. Bashein, S. G. Oliver, D. Waddington M. Rattray (2003) A Model-based Analysis of Microarray Experimental Error and Normalisation, *Nucleic Acids Res.*, 31, e96.
- [Fiehn03] O. Fiehn, W. Weckwerth (2003) Deciphering Metabolic Networks, *Eur. J. Biochem.*, 270, 579-588.
- [Fisher07*] J. Fisher, T. Henzinger (2007) Executable Cell Biology, *Nature Biotechnology*, 25(11), November, 1239-1249.
- [de la Fuente02] A. de la Fuente, P. Brazhnik, P. Mendes (2002) Linking the Genes: Inferring Quantitative Gene Networks from Microarray Data, *Trends Genet.*, 18, 395-398.
- [Friedman00] N. Friedman, M. Linial, I. Nachman, D. Pe'er (2000) Using Bayesian Networks to Analyze Expression Data, *Journal of Computational Biology*, August 2000, 7(3-4), 601-620.
- [Funahashi08] A. Funahashi, Y. Matsuoka, A. Jouraku, M. Morohashi, N. Kikuchi, H. Kitano (2008) CellDesigner 3.5: A Versatile Modeling Tool for Biochemical Networks, *Proceedings of the IEEE*, 96(8), August 2008, 1254-1265.
- [Heinrich96*] R. Heinrich, S. Schuster (1996) *The Regulation of Cellular Systems*, Chapman & Hall.
- [Huang99] S. Huang (1999) Gene Expression Profiling, Genetic Networks and Cellular States: an Integrating Concept for Tumorigenesis and Drug Discovery, *J. Mol. Med.*, 77, 469-480.
- [Hynne01*] F. Hynne, S. Dano, P. G. Sorensen (2001) Full-scale Model of Glycolysis in *Saccharomyces Cerevisiae*, *Biophys. Chem.*, 94, 121-163.
- [Kell00] D. B. Kell, R. D. King (2000) On the Optimization of Classes for the Assignment of Unidentified Reading Frames in Functional Genomics Programmes: the Need for Machine Learning, *Trends Biotechnol.*, 18, 93-98.
- [Kephart03*] J. Kephart, D. Chess (2003) The Vision of Autonomic Computing, *Computer*, 36(1), 41-50.
- [Kierzek02] A. M. Kierzek (2002) STOCKS: STOChastic Kinetic Simulations of Biochemical Systems with Gillespie Algorithm, *Bioinformatics*, 18, 470-481.
- [Kitano02*] H. Kitano (2002) Systems Biology: A Brief Overview, *Science*, 295, 1662-1664.
- [Li07] C. Li, Q. W. Ge, M. Nakata, H. Matsuno, S. Miyano (2007) Modelling and Simulation of Signal Transductions in an Apoptosis Pathway by using Timed Petri Nets, *J. Biosci.*, 32, 113-127.
- [Lok05] L. Lok, R. Brent (2005) Automatic Generation of Cellular Reaction Networks with Molecuizer 1.0., *Nature Biotechnology*, 1, 131-136.
- [Ma03*] H. Ma, A-P. Zeng (2003) Reconstruction of Metabolic Networks from Genome Data and Analysis of their Global Structure for Various Organisms, *Bioinformatics*, 19(2), 270-277.
- [Mendes93] P. Mendes (1993) GEPASI: A Software Package for Modelling the Dynamics, Steady States and Control of Biochemical and other Systems, *Computer Applications in the Biosciences*, 9(5), 563-571.
- [Mendes02] P. Mendes (2002) Emerging Bioinformatics for the Metabolome, *Briefings in Bioinformatics*, 3, 134-145.
- [Moraru08] I. I. Moraru, J. C. Schaff, B. M. Slepchenko, M. L. Blinov, F. Morgan, A. Lakshminarayana, F. Gao, Y. Li, L. M. Loew (2008) Virtual Cell Modelling and Simulation Software Environment, *Systems Biology*, IET, September 2008, 2(5), 352-362.
- [Paton00*] N.W. Paton, S.A. Khan, A. Hayes, F. Moussouni, A. Brass, K. Eilbeck, C.A. Goble, S.J. Hubbard, S.G. Oliver (2000) Conceptual Modelling of Genomic Information, *Bioinformatics*, 16, 548-557.
- [Raamsdonk01*] L.M. Raamsdonk, B; Teusink, D. Broadhurst, N. Zhang, A. Hayes, M.C. Walsh, J.A. Berden, K.M. Brindle, D.B. Kell, J.J. Rowland et al. (2001) Functional Genomics via the Metabolome: a FANCY Approach to Characterising Mutations with a Silent Phenotype, *Nat Biotechnol.*, 19, 45-50.



- [Ripoll98*] C. Ripoll, J. Guespin-Michel, V. Norris, M. Thellier (1998) Defining Integrative Biology, *Complexity*, 4 (2), 19-20.
- [Shmulevich02] I. Shmulevich, E. R. Dougherty, S. Kim, W. Zhang (2002) Probabilistic Boolean networks: A Rule-based Uncertainty Model for Gene Regulatory Networks, *Bioinformatics*, 18, 261-274.
- [Tilstone03] C. Tilstone (2003) Vital Statistics, *Nature*, 424, 610-613.
- [Tomita99] M. Tomita, K. Hashimoto, K. Takahasi, T. S. Shimizu, Y. Matsuzaki, F. Miyoshi, K. Saito, S. Tanida, K. Yugi, J. C. Venter, C. A. Hutchison (1999) E-CELL: Software Environment for Whole-cell Simulation, *Bioinformatics*, 15, 72-84.
- [Weaver99] D. C. Weaver, C. T. Workman, G. D. Stormo (1999) Modeling Regulatory Networks with Weight Matrices, *Pac. Symp. Biocomput.*, 4, 112-123.
- [Wu02] L. F. Wu, T. R. Hughes, A. P. Davierwala, M. D. Robinson, R. Stoughton S. J. Altschuler (2002) Large-scale Prediction of *Saccharomyces Cerevisiae* Gene Function using Overlapping Transcriptional Clusters, *Nat. Genet.*, 31, 255-265.

5.2. Travaux en relation avec les technologies agents

- [Alves06] Rui Alves, Fernando Antunes, Armindo Salvador (2006) Tools for Kinetic Modeling of Biochemical Networks, *Nature Biotechnology*, 24, 667-672.
- [Amigoni07] F. Amigoni, V. Schiaffonati (2007) Multiagent-Based Simulation in Biology: A Critical Analysis, *Model-Based Reasoning in Science, Technology, and Medicine*, Springer, *Studies in Computational Biology*, 64, Lorenzo Magnani and Ping Li (Eds), 179-191.
- [Armano05] G. Armano, G. Mancosu, A. Orro, E. Vargiu (2005) A Multi-agent System for Protein Secondary Structure Prediction, *Transactions on Computational Systems Biology III*, LNCS 3737, Springer, 14-32.
- [Bernon09] C. Bernon, D. Capera, J-P. Mano (2009) Engineering Self-Modeling Systems: Application to Biology, *International Workshop on Engineering Societies in the Agents World (ESAW'08)*, Saint-Etienne, 24-26/09/2008, A. Artikis, G. Picard, L. Vercoeur (Eds.), Springer-Verlag, LNCS, 2009 (à paraître).
- [Besse05] C. Besse (2005) *Recherche de conformation de molécules et apprentissage du potentiel de Lennard-Jones par systèmes multi-agents adaptatifs*, Rapport de Master Recherche IARCL, Université Paul Sabatier, juin 2005.
- [Bortolussi05] L. Bortolussi, A. Dovier, F. Fogolari, Multi-Agent Simulation of Protein Folding (2005) First Workshop on *Multi-Agent Systems for Medicine, Computational Biology, and Bioinformatics (MAS*BIOMED'05@AAMAS'05)*, 91-106, 2005.
- [Bosse06] T. Bosse, C. Jonker, J. Treur (2006) Simulation and Analysis of Complex Biological Processes: an Organisation Modelling Perspective, *39th Annual Simulation Symposium*.
- [Camps98] V. Camps (1998) *Vers une théorie de l'auto-organisation dans les systèmes multi-agents basée sur la coopération : application à la recherche d'information dans un système d'information répartie*, Thèse de doctorat, Université Paul Sabatier, janvier 1998.
- [Cannata05] Nicola Cannata, Flavio Corradini, Emanuela Merelli, Andrea Omicini, Alessandro Ricci (2005) An Agent-Oriented Conceptual Framework for Systems Biology, *Transactions on Computational Systems Biology III*, Corrado Priami, Emanuela Merelli, Pedro Pablo Gonzalez, Andrea Omicini (Eds), LNCS 3737, Springer Verlag, 102-122.
- [Capera05] Davy Capera (2005) *Systèmes multi-agents adaptatifs pour la résolution de problèmes : Application à la conception de mécanismes*, Thèse de doctorat, Université Paul Sabatier, juin 2005.
- [Conceicao08] D. Conceição, M. Gatti, C. de Lucena (2008) An Agent-based Framework for Stem Cell Behavior Modeling and Simulation, *Research report 17/08*, Department of Computer Sciences, Pontificia Universidade Catolico do Rio de Janeiro.
- [D'Inverno05a] Mark d'Inverno, Jane Prophet (2005) Multidisciplinary Investigation into Adult Stem Cell Behaviour, *Biology, Computer Science and Bioinformatics: Multidisciplinary Models, Metaphors and Tools*, *Transactions on Computational System Biology III*, Springer-Verlag, LNCS 3737, Emanuela Merelli, Pablo Gonzalez and Andrea Omicini (Eds), 49-64.
- [D'Inverno05b] M. d'Inverno, R. Saunders (2005) Agent-based Modelling of Stem Cell Organisation in a Niche, *Engineering Self-Organising Systems: Methodologies and Applications*, Springer, Brueckner S., Di Marzo Serugendo G., Karageorgos A., Nagpal R. (Eds), LNCS 3464, Springer, 52-68.

- [Emonet05] T. Emonet, C. Macal, M. North, C. Wickersham, P. Cluzel (2003) AgentCell: a Digital Single-cell Assay for Bacterial Chemotaxis, *Bioinformatics Advance Access*, *Bioinformatics*, 21, 2714-2721, 2005.
- [Gonzalez 03] P. González, M. Cárdenas, D. Camacho, A. Franyuti, O. Rosas, J. Lagúnez-Otero, Cellulat: an Agent-based Intracellular Signalling Model, *Biosystems*, 68(2-3), 171-185.
- [Ferber98] Jacques Ferber, Olivier Gutknecht (1998) A Meta-Model for the Analysis and Design of Organizations in Multi-Agent Systems, *Proceedings of the Third International Conference on Multi-Agent Systems (ICMAS98)*, 128-135, Paris, France.
- [Gatti08] Maíra A. de C. Gatti, Carlos J. P. de Lucena (2008) Cell Simulation: An Agent-based Software Engineering Approach, *Monografias em Ciência da Computação*, No. 18/08 ISSN: 0103-9741 *Research report 18/08*, Department of Computer Sciences, Pontificia Universidade Catolico do Rio de Janeiro, April 2008.
- [Gatti07] Maíra Gatti, José Eurico, Renato Raposo, Guilherme Pate (2007) Modeling and Simulating Self-Organizing Stem Cells using Multi-Agents Systems, Position paper, *Third Workshop on Software Engineering for Agent-oriented Systems (SEAS'07)*, João Pessoa.
- [Georgé03] Jean-Pierre Georgé, Marie-Pierre Gleizes, Pierre Glize, Christine Régis (2003) Real-time Simulation for Flood Forecast: an Adaptive Multi-Agent System STAFF, *Symposium on Adaptive Agents and Multi-Agent Systems*, University of Wales, Aberystwyth, Society for the Study of Artificial Intelligence and the Simulation of Behaviour, 109-114.
- [Gleizes02] Marie-Pierre Gleizes, Pierre Glize (2002) ABROSE: Multi Agent Systems for Adaptive Brokerage, *International Bi-Conference Workshop on Agent-Oriented Information Systems*, Toronto, Canada, Vol. 57, P. Giorgini, Y. Lespérance, G. Wagner, E. Yu (Eds.), *CEUR Workshop Proceedings*.
- [Guo07] D. Guo, E. Santos, A. Singhal, E. Santos, Q. Zhao (2007) Adaptivity Modeling for Complex Adaptive Systems with Application to Biology, *IEEE International Conference on Systems, Man and Cybernetics*, 272-277.
- [Jonker08] C. Jonker, J. Snoep, J. Treur, H. Westerhoff, W. Wijngaards (2008) BDI-modelling of Complex Intracellular Dynamics, *Journal of Theoretical Biology*, 251, 1-23.
- [Jonker07] Catholijn Jonker, Jan Treur (2007) Agent-oriented Modeling of the Dynamics of Biological Organisms, *Applied Intelligence*, 27(1), 1-20, August 2007.
- [Kawazu06] Toshihiro Kawazu, Shinsuke Odai, Noriko Shibuya, and Taishin Nomura (2006) Development of a Support Tool for Multi-agent Based Biological Modeling, *Proc. of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '06)*, NY, 4263-4166
- [Khan03] S. Khan, R. Makkena, W. Gillis, C. Schmidt (2003) A Multi-agent System for the Quantitative Simulation of Biological Networks, *Second International Joint Conference on Autonomous Agents & Multiagent Systems (AAMAS'03)*, Melbourne, ACM, 385-392, 2003.
- [Lales 07] C. Lales (2007) *Modélisation gros grain et simulation multi-agents. Application à la membrane mitochondriale*, Thèse de Doctorat de l'Université de Bordeaux I, décembre 2007.
- [Lales05] C. Lales, N. Parisey, J-P. Mazat, M. Beurton-Aimar (2005) Simulation of Mitochondrial Metabolism using Multi-agents System, *First Workshop on Multi-Agent Systems for Medecine, Computational Biology, and Bioinformatics (MAS*BIOMED'05 at AAMAS'05)*, 137-145, 2005.
- [Lints05] T. Lints (2005) Multiagent Modelling of a Bacterial Cell, a DnaA Titration Model Based Agent Model as an Example, *Ninth Symposium on Programming Languages and Software Tools*, Tartu, Estonia, Vene V., Meriste M. (Eds.), 82-96.
- [Luck 06] M. Luck, E. Merelli (2005) TFG on Agents in Bioinformatics, *Knowledge Engineering Review*, 20(2), 117-125.
- [Mano06] J.P. Mano (2006) *Étude de l'émergence fonctionnelle au sein d'un réseau de neuro-agents coopératifs*, Thèse de Doctorat de l'Université Paul Sabatier, IRIT, Toulouse III, 30 mai 2006.
- [Mendes93] P. Mendes (1993) GEPASI: A Software Package for Modelling the Dynamics, Steady States and Control of Biochemical and other Systems, *Computer Applications in the Biosciences*, 9(5), 563-571, 1993.
- [Merelli06] E. Merelli, G. Armano, N. Cannata, F. Corradini, M. d'Inverno, A. Doms, P. Lord, A. Martin, L. Milanesi, S. Moller, M. Schroeder, M. Luck (2006) Agents in Bioinformatics, *Computational and Systems Biology, Briefings in Bioinformatics*, 8(1), 45-59.
- [Paton93] R. Paton (1993) Some Computational Models at the Cellular Level, *BioSystems*, 29, 63-75.
- [Pogson06] M. Pogson, R. Smallwood, E. Qvarnstrom, M. Holcomb (2006) Formal Agent-based Modelling of Intracellular Chemical Interactions, *BioSystems*, 85, 37-45.



- [Querrec 03] G. Querrec, V. Rodin, J.F. Abgrall, S. Kerdelo, J. Tisseau (2003) Uses of Multiagents Systems for Simulation of MAPK Pathway, *Third IEEE Symposium on Bioinformatics and Bioengineering (BIBE'03)*, 421-425.
- [Ren08] L-H. Ren, Y-S. Ding, Y-Z. Shen, X-F. Zhang (2008) Multi-agent-based Bio-network for Systems Biology: Protein-protein Interaction Network as an Example, *Amino Acids*, 35(3), 565-572.
- [Troisi05] A. Troisi, V. Wong, M. Ratner (2005) An Agent-based Approach for Modeling Molecular Self-organization, *Proceedings of the National Academy of Sciences of the USA (PNAS)*, 102(2), 255-260.
- [Santos04] E. Santos, D. Guo, E. Santos Jr., W. Onesty (2004) A Multi-Agent System Environment for Modelling Cell and Tissue Biology, *International Conference on Parallel and Distributed Processing Techniques and Applications*, Las Vegas, USA, CSREA Press, Arabnia H. R. (Eds), 3-9.
- [Shafaei08] S. Shafaei, N. Aghaee (2008) Biological Network Simulation Using Holonic Multiagent Systems, *Tenth International Conference on Computer Modeling and Simulation (UKSIM'08)*, 1-3 April, 617-622.
- [Webb04] Kenneth Webb, Tony White (2004) Cell Modeling using Agent-based Formalisms, 3rd International Joint *Conference on Autonomous Agents and Multiagent Systems (AAMAS'04)*, 19-23 August 2004, New York, NY, IEEE Computer Society, 1190-1196.
- [Webb06] K. Webb, T. White (2006) Cell Modeling with Reusable Agent-based Formalisms, *Applied Intelligence*, 24(2), 169-181.
- [Welcomme08] Jean-Baptiste Welcomme (2008) *Simulation adaptative par systèmes multi-agents auto-organiseurs, Application à l'agencement de composants logiciels pour la conception d'avions*, Thèse de doctorat, Université Paul Sabatier, mars 2008.