

Modélisation in silico des réseaux métaboliques

M. Beurton-Aimar, S. Pérès, N. Parisey, C. Nazaret, J.P. Mazat.

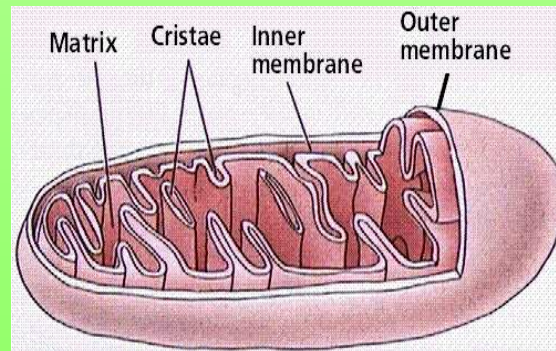
LaBRI - Université Bordeaux, Laboratoire de physiologie mitochondriale - Université Bordeaux

- Objectif du projet MitoScoP:
Construire un modèle informatique pour la représentation de l'activité mitochondriale.
- Application :
Étudier et simuler le comportement du **“système”** mitochondrial, normal ou pathologique.

La mitochondrie

Un organelle intra-cellulaire qui possède ses propres caractéristiques métaboliques :

- Le cycle de Krebs
- La chaîne respiratoire
- ...



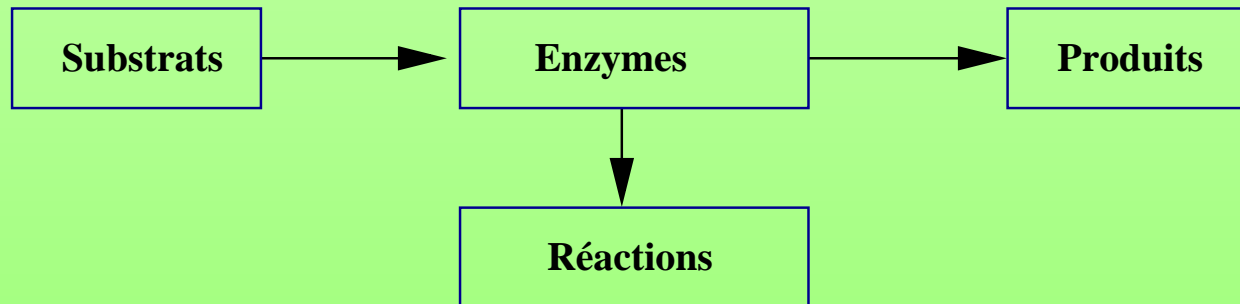
Modéliser les données biologiques

- Un **modèle** pour l'informatique.
- Modéliser séparément les **données** et les **traitements**.
- Capacités à **partager** les données entre plusieurs modules de traitements.

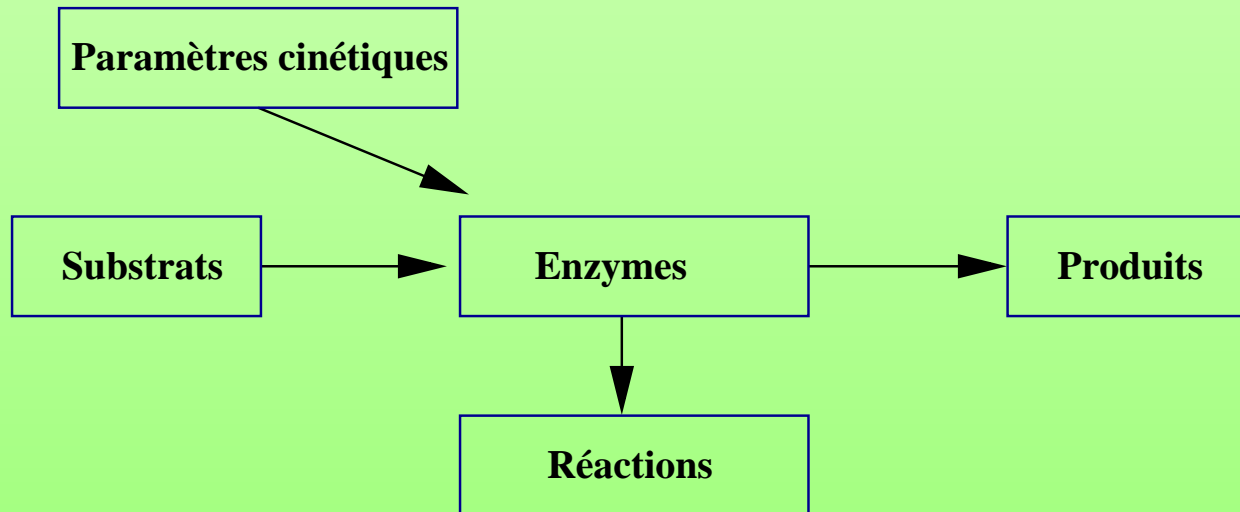
Les différents outils existants

- La capitalisation de connaissances : les banques de données, généralement un ***environnement statique***.
 - Les banques pour la mitochondrie : Mitodat, Mitomap, Brenda, KEGG ...
- Les logiciels de simulation ***"programmables"***, souple mais qui exigent une connaissance approfondie de leur fonctionnement.
- Les projets de modélisation ***"in silico"*** des systèmes biologiques : E-Cell, Virtual Cell ...

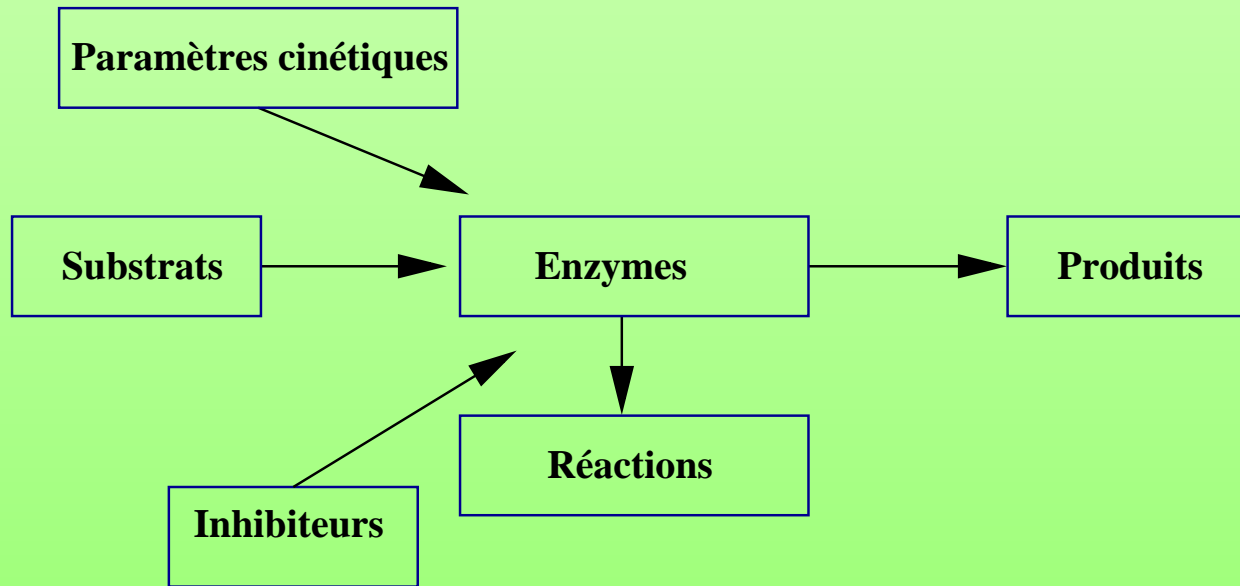
L'équation de base d'une réaction métabolique



L'équation de base d'une réaction métabolique



L'équation de base d'une réaction métabolique



Modèle de données E-Cell

L'ensemble des paramètres et des opérations sont organisées autour de 3 ***concepts*** :

1. ***Substances*** : toutes les molécules constituant la cellule ou le milieu de culture.
2. ***Reactors*** : description des inter-actions entre les différents composants de la cellule dans une base de règles.
3. ***System*** : organisation spatiale et fonctionnelle de la cellule.

Modèle orienté objet

Les principes :

- Introduire la **sémantique** des données.
- Modéliser indépendamment d'un outil de traitement.
- Mais en prenant en compte un **objectif**, un **environnement**, un **domaine** d'application.
- Concevoir une architecture modulaire et décrire le système à un **“bon”** niveau d'abstraction..

Objectif :

- Construire un environnement intégré de **simulation** .

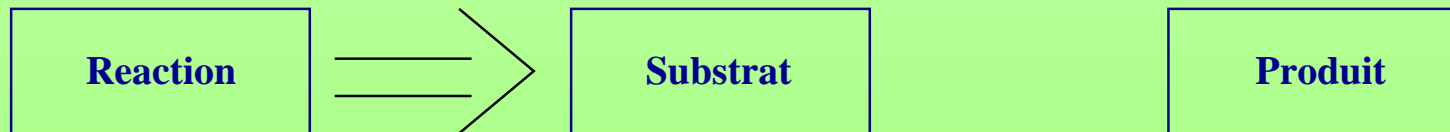
Organisation des données

Reaction



```
graph LR; Reaction[Reaction] --> Arrow[>]
```

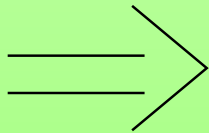
Organisation des données



Organisation des données

Objet Biologique

Reaction



Substrat

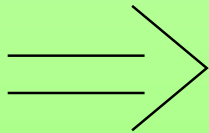
Produit

Organisation des données

Objet Biologique

- _ **Nom**
- _ **Sequence**
- _ **Masse molaire**
- _ **Reference Bibliographique**

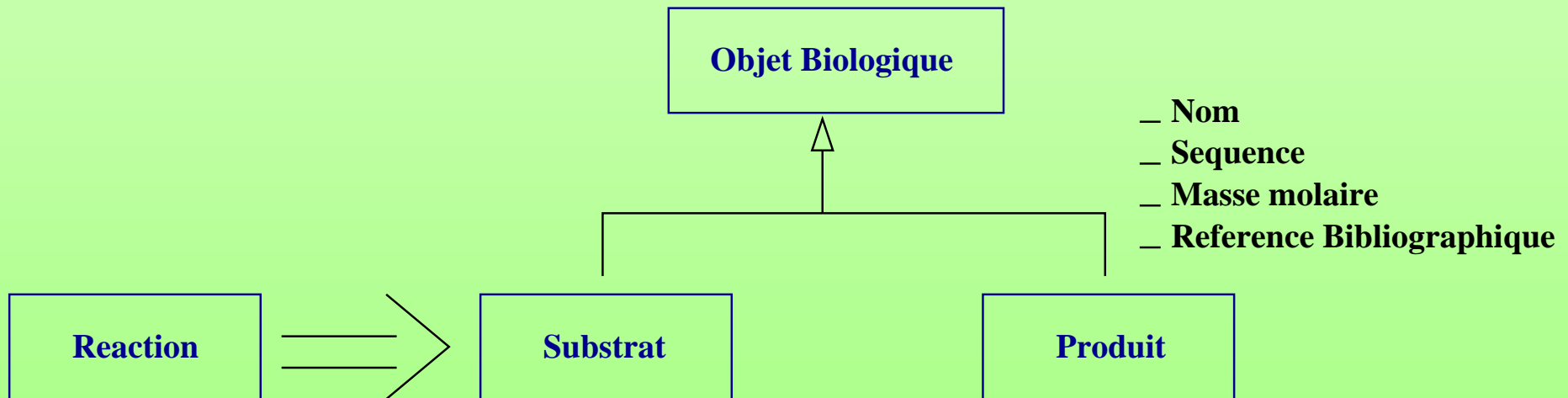
Reaction



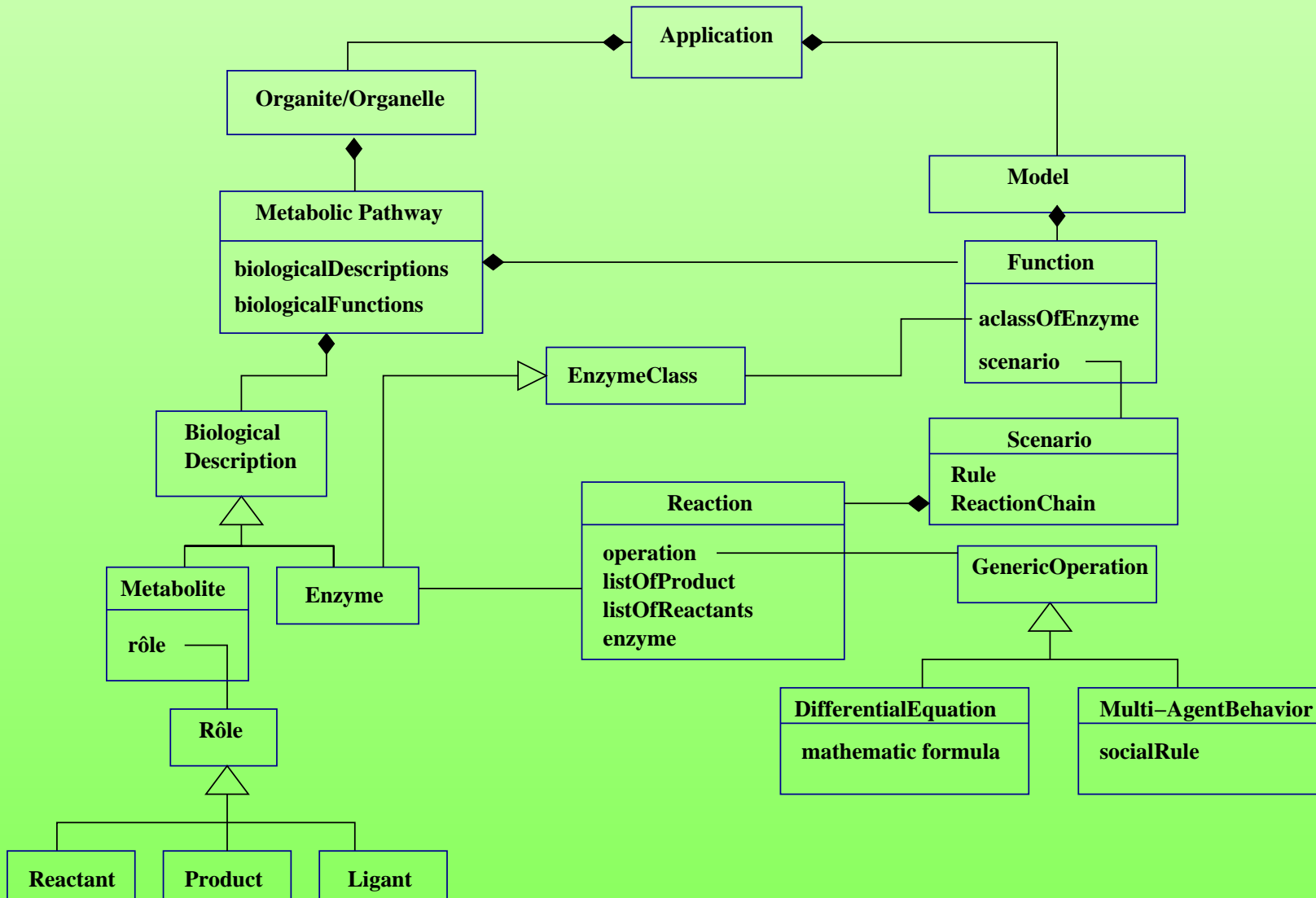
Substrat

Produit

Organisation des données



Une architecture orientée objet



Utiliser un langage de description de données

- Indépendant des langages de programmation .
- **XML** : eXtended Markup Language
 1. Langage à base de **TAGS**.
 2. Création de fichiers de texte pouvant être relus (**parser**) ou écrits automatiquement.
 3. Communications aisées avec des logiciels existants ou des programmes “*maison*” interfacés.
- **SBML** : une version adaptée à la biologie par la création de **TAGS** spécifiques.

Schéma SBML

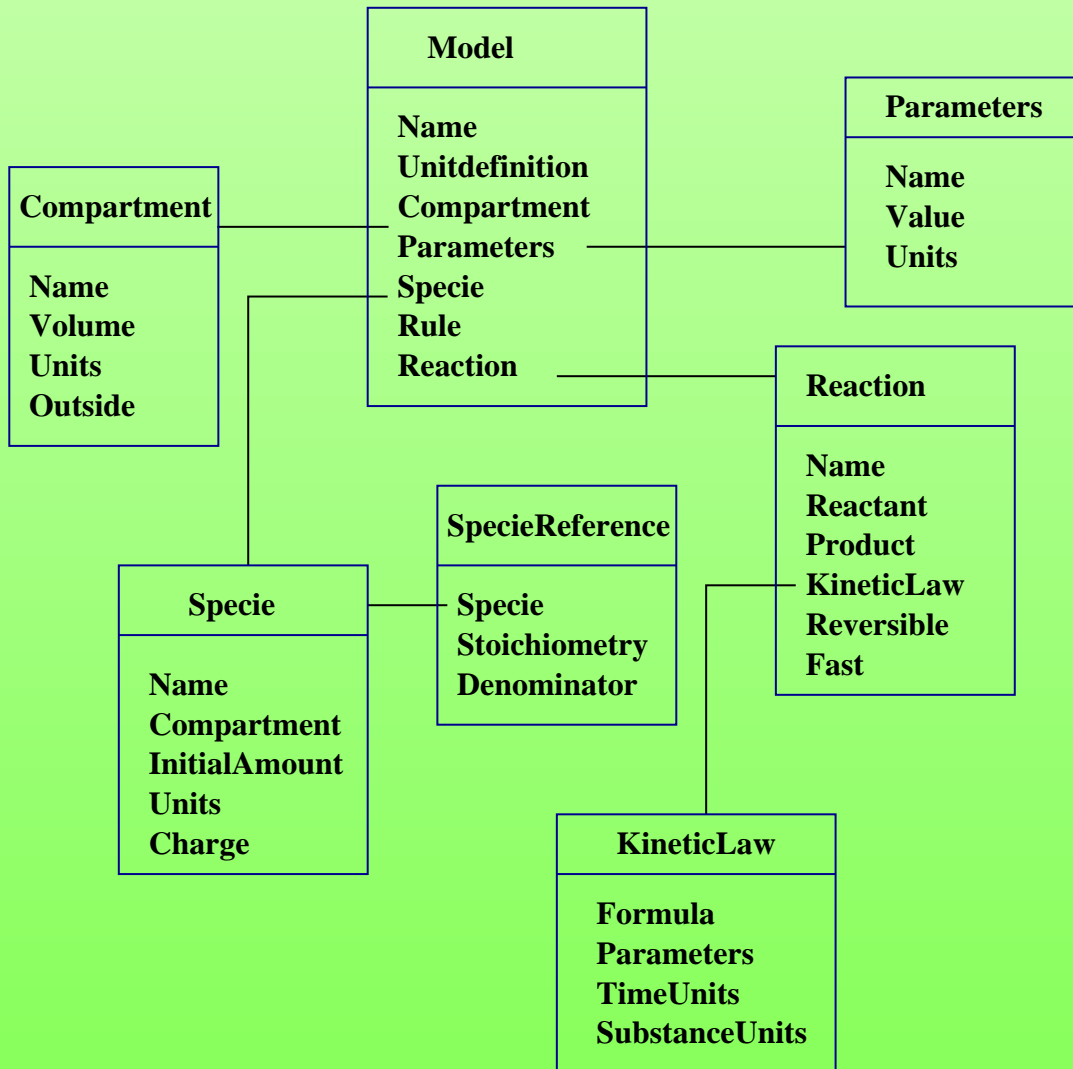
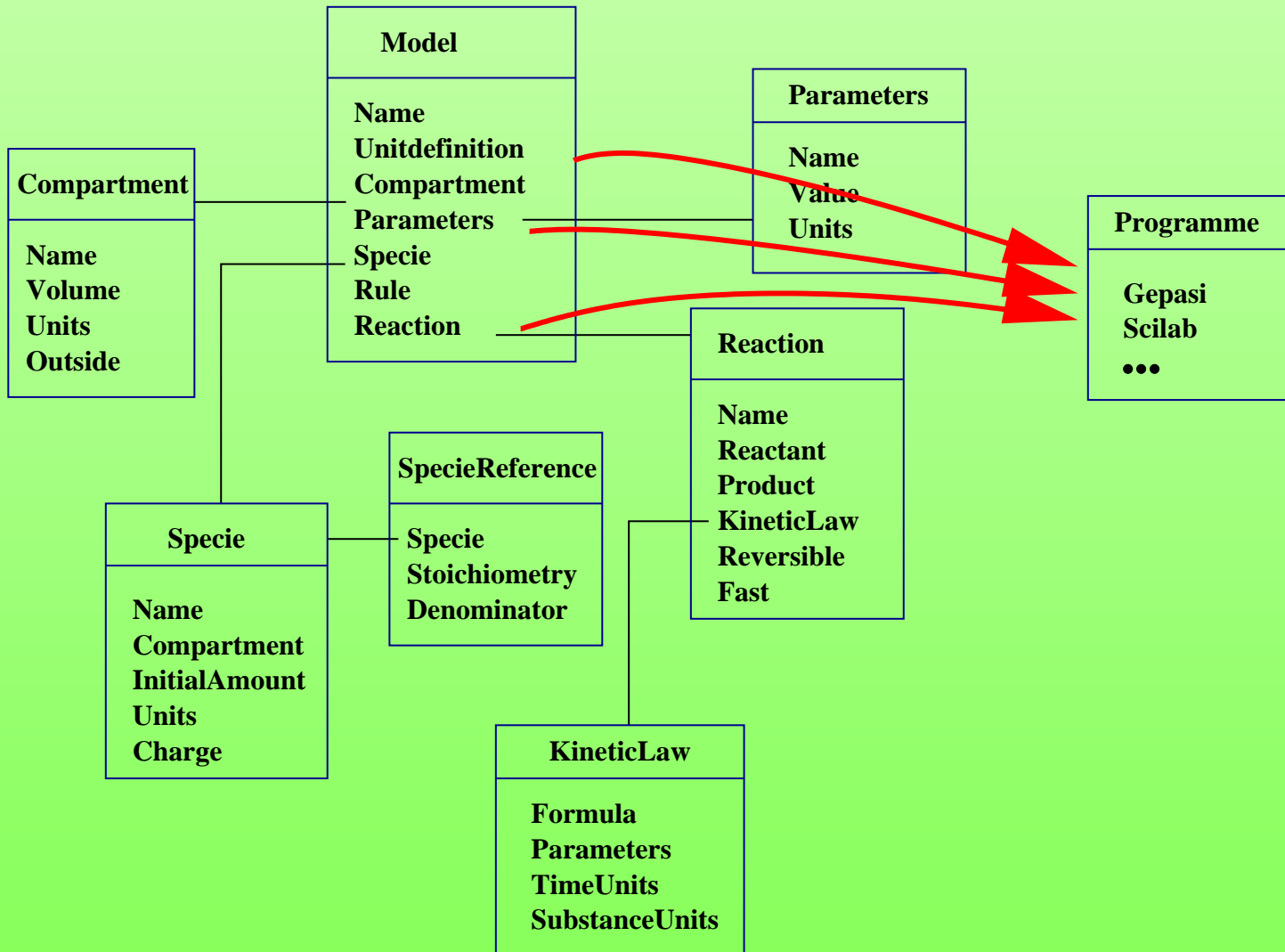


Schéma SBML



Le cycle de Krebs en SBML

```
<model name="cycle-de-krebs">
  <listOfCompartments>
    <Compartment name="Mitochondria"/>
    <Compartment name="Matrice" outside="Mitochondria"/>
  </listOfCompartments>

  <!-- METABOLITES -->
  <Species name="acetyl-CoA"
    Compartment="Matrice"
    initialAmount="0.0"/>
  <Species name="citrate"
    Compartment="Matrice"
    initialAmount="0.0" />

  <!-- ENZYMES -->
  <Species name="citrate synthase"
    Compartment="Matrice"
    initialAmount="0.0?" />
```

Le cycle de Krebs en SBML

```
<!--REACTIONS DE BARBARA WRIGHT-->
<Reaction name="W1" reversible="false">
  <listOfReactants>
    <specieReference specie="OAA2" />
    <specieReference specie="acetyl-CoA" />
  </listOfReactants>
  <listOfProducts>
    <specieReference specie="citrate" />
    <specieReference specie="CoA" />
    <specieReference specie="H+" />
  </listOfProducts>
  <listOfEnzymes>
    <specieReference specie="citrate synthase" />
  </listOfEnzymes>
```

Le cycle de Krebs en SBML

```
<kineticLaw formula =  
''vcs * aco * oaa2 / (k11oaa2 * aco + k11aco * (1.0 + CoA / k11icoa) * oaa2 + aco * oaa2)''>  
  <listOfParameters>  
    <Parameter name='''vcs''' value= ''8.23 '' />  
    <Parameter name= ''k11aco'' value= ''0.01 '' />  
    <Parameter name= ''k11oaa2'' value= ''0.007 '' />  
    <Parameter name= ''k11icoa'' value= ''0.11 '' />  
  </listOfParameters>  
</ kineticLaw>  
</Reaction>
```

Conception du système

Modifications structurelles de SBML :

- Amélioration du niveau d'abstraction, ajout de nouveaux concepts, explicitation de toutes les entités.
- La racine de la description n'est plus un traitement.

Ajouts de fonctionnalités :

- Élaboration d'un langage de description de scénario (ou de graphes de processus).
- Prise en compte des informations concernant la localisation des composants et les unités temporelles.

Organiser les simulations

- Décrire un système comme un *assemblage* de fonctions.
 - Instancier le modèle avec un processus de calcul d'équations linéaires ou différentielles.
 - Implémenter des algorithmes de traitements des graphes - réseaux - orientés.
 - Instancier le modèle avec le système multi-agents, ajout de la notion d'espace, de structure, d'hypersructure.

La chaine respiratoire - Modèle de Bohnensack

Respiratory chain-Hydrogen supply $SH_2 + A_{ox} \xrightarrow{v_h} S + A_{red}$ $v_h = V_h a_{ox}$	Energy transformation-H^+ gradient generation $A_{red} + 2D_{ox} + n_r H_I^+ \xrightleftharpoons{v_r} A_{ox} + 2D_{red} + n_r H_E^+$ $v_r = V_r \left(1 - \frac{a_{ox} d_{red}^2}{a_{red} d_{ox}^2} 10^{-n_r(\Delta\mu - \Delta\mu_r)/Z} \right)$
Terminal oxidation $2D_{red} + 1/2O_2 \xrightarrow{v_o} 2D_{ox} + H_2O$ $v_o = V_o d_{red}$	Phosphate translocation $P_{i,E}^{2-} + 2H_E^+ \rightleftharpoons P_{i,I}^{2-} + 2H_I^+$ $[P_i]_I = [P_i]_E \varphi_p 10^{\Delta pH}$
ATPase $ADP_I + P_{i,I} + n_a H_E^+ \xrightleftharpoons{v_a} ATP_I + n_a H_I^+$ $v_a = V_a \left(1 - \frac{[ATP]_I}{[ADP]_I [P_i]_I} \varphi_p 10^{n_a(\Delta\mu - \Delta\mu_a)/Z} \right)$ $\varphi_p = f_p + (1 - f_p) 10^{\Delta pH}$ $\Delta pH = -f_\mu \Delta\mu / Z$	Adenine nucleotides translocation $ADP_E + ATP_I \xrightleftharpoons{v_t} ADP_I + ATP_E$ $v_t = V_t \frac{1 - \frac{[ATP]_E [ADP]_I}{[ADP]_E [ATP]_I} 10^{\Delta\psi/Z}}{\left(1 + \frac{[ATP]_E}{[ADP]_E} 10^{f_\psi \Delta\psi/Z} \right) \left(1 + \frac{[ADP]_I}{[ATP]_I} 10^{(1-f_\psi)\Delta\psi/Z} \right)}$ $\Delta\psi = \Delta\mu + Z\Delta pH$
Proton leak $H_E^+ \xrightarrow{v_l} H_I^+$ $v_l = V_l 10^{-\Delta\mu/Z}$	Steady-state relations $v_r = v_h = v_o = v_{resp}$ $n_r v_r = n_a v_a + v_t + v_l$ $v_a = v_t + v_i$ $v_t = v_e + v_{ex}$
$v_i = k_i \frac{[ATP]_I}{[ADP]_I}$	$v_e = V_e \frac{[ATP]_E}{[ATP]_E + [ADP]_E}$

Modélisation d'un réseau métabolique

Le cas du cycle de Krebs :

- La base de données KEGG cite 23 enzymes comme appartenant au TCA toutes espèces confondues.
 - 16 enzymes chez Homo sapiens.
 - 14 enzymes chez S. cerevisiae.
 - 13 enzymes chez Schizosaccharomyces pombe.
 - S. Schuster fournit le programme Metatool qui calcule la matrice de stochiométrie, les modes élémentaires et les modes fondamentaux du réseau (méthode des cones convexes).
- A General Definition of Metabolic Pathways Useful for Systematic Organization and Analysis of Complex Metabolic Networks. Nature Biotechnol. 18 (2000) 326-332.
- METATOOL: for studying metabolic networks. Bioinformatics 15(3): 251-257 (1999)

Recherche des chemins dans un réseau

- Un **graphe** est un ensemble de **noeuds** (sommets) auquel est associé un ensemble de **transitions**, orientées ou non.
- Définition extrêmement proche du réseau métabolique.
- On dispose d'un ensemble d'algorithmes bien connus qui fournissent tous les chemins possibles, les chemins minimaux, l'ensemble minimal de sommets qui garantissent la connexité ...
- Stockage plus économique si la matrice est creuse, coût inférieur pour les ajouts ou les délétions de sommets.

Intérêts d'un modèle orienté objet

- **Comparaison** entre les tissus ou les espèces facilitée par le niveau d'**abstraction**.
- Effet de **zoom** ou d'échelle automatiquement intégré grâce à la **composition**.
- **Collaborations** possibles entre différents types de traitements.
 - Recherche des configurations possibles pour des réseaux métaboliques complexes.
- **Évolution** du système, **réutilisation** des modules.
- Construction d'un **système intégré** et **communicant**.

Les problèmes difficiles qui vont se poser

- Structuration des données :
 - Trouver le bon niveau de traduction des fonctions biologiques.
 - Gestion de la revision de connaissances, d'un historique.
- Organisation des traitements :
 - Interpréter les résultats de simulations complexes *“qui marchent”*.
 - Gestion de résultats de formats hétérogènes.

Un réseau de mitochondries ou une mitochondrie en réseau ?

