

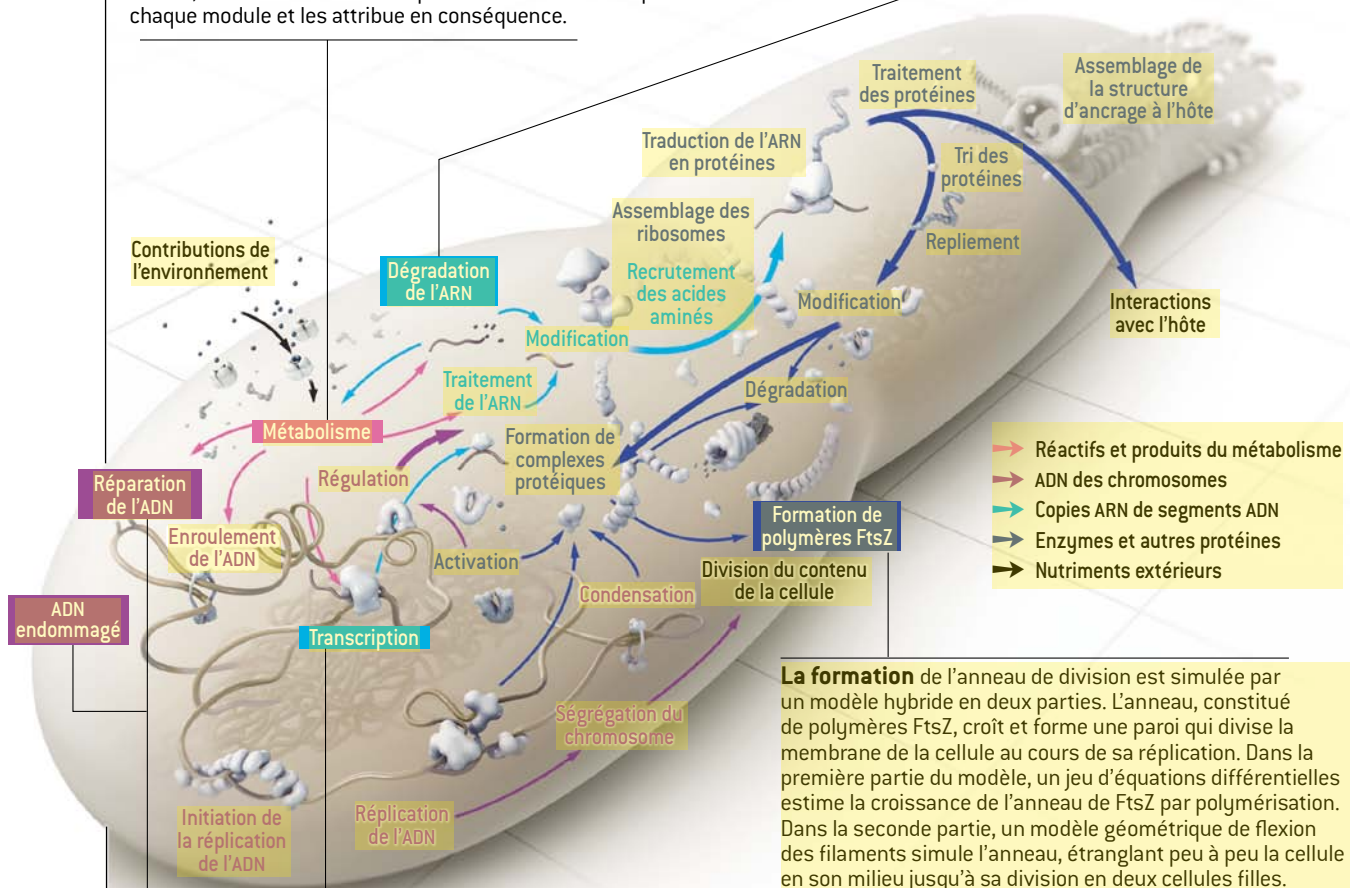
LE PRINCIPE DU SIMULATEUR

L'auteur a mis au point un modèle numérique de la bactérie infectieuse *Mycoplasma genitalium* qui représente presque tous les aspects de la vie, de la croissance et de la reproduction de ce microbe. Aucune approche mathématique ne permet de simuler toutes les fonctions biologiques de la cellule, si bien que ces fonctions ont été divisées en 28 modules distincts (*indiqués dans la cellule ci-dessous*) qui interviennent dans le traitement de l'ADN (*en violet*), de l'ARN (*en bleu clair*), des protéines (*en bleu foncé*), et de l'énergie, des nutriments et des déchets (*en rose*). Pour chaque module, les chercheurs ont sélectionné la méthode mathématique la plus appropriée (*voir ci-dessous quelques exemples de ces méthodes*).

Au début du programme, tous les modules tournent selon une séquence aléatoire, l'ensemble simulant une seconde de temps réel. Beaucoup des valeurs d'entrée sont tirées d'un grand tableau de variables représentant leurs états initiaux, d'autres sont sélectionnées à partir de fourchettes ou de calculs de probabilités. Les chercheurs peuvent simuler différents scénarios en modifiant la configuration de départ. Après ce premier pas, le programme met à jour le tableau de variables afin de refléter les sorties des modules. La séquence tourne alors pour un nouveau pas d'une seconde, met à jour le tableau, etc. La boucle continue jusqu'à ce que la cellule se divise, meure ou devienne trop vieille pour avoir un sens.

Le métabolisme des nutriments et des déchets est modélisé en utilisant l'analyse du bilan des flux. Cette méthode exploite des techniques de programmation linéaire pour calculer les vitesses de réaction qui optimisent la croissance, la production d'énergie ou toute autre caractéristique choisie par le modélisateur. Elle suppose que les réactions ont lieu assez vite pour obtenir un état stationnaire pendant la seconde correspondant à un pas de simulation. Afin d'empêcher les premiers modules d'utiliser des substances dont d'autres modules ont besoin, le simulateur estime la part de ces ressources qui revient à chaque module et les attribue en conséquence.

La dégradation et le recyclage de l'ARN et des protéines sont modélisés en utilisant des processus de comptage dits de Poisson. Ces processus utilisent un générateur de nombres aléatoires et des fonctions de probabilité pour décider si un fragment donné d'ARN ou de protéine se dégrade ou persiste jusqu'au pas suivant.



La transcription et la traduction des gènes (étapes au cours desquelles sont fabriquées de nombreuses protéines nécessaires à la croissance et la réplication des cellules) sont simulées par des algorithmes en plusieurs parties incluant des modèles de Markov (ces modèles suivent dans le temps l'état des enzymes qui copient les gènes en ARN), un algorithme probabiliste décrivant la fixation de ces enzymes sur l'ADN et une programmation linéaire pour distribuer l'énergie et autres ressources.

Les dommages subis par l'ADN et les réparations sont modélisés selon cette même méthode non déterministe.