

Dans le second cas, les approches booléennes permettent de décrire le système étudié en termes de relations logiques du type : « si A est vrai et si B est faux, alors C est vrai. » Il est possible d'ajouter des notions temporelles, par exemple : « si A se produit avant B, alors C sera toujours vrai. » Ces approches booléennes sont particulièrement efficaces lorsque les informations quantitatives complètes ne sont pas disponibles, comme souvent en biologie. Elles permettent alors des prédictions sûres, à rebours des approches différentielles qui sont plutôt utilisées pour apporter des explications scientifiques aux phénomènes observés.

En résumé, il n'y a pas de standards admis par tous pour ces formalismes ni pour les méthodes numériques associées. En revanche, il existe des standards *de facto*, c'est-à-dire quelques formalismes et méthodes répandus dans la communauté scientifique concernée. Mais il n'y a pas d'unicité de ces standards *de facto*, parce que leur choix rationnel repose sur les questions posées, les jeux de données disponibles, les observables biologiques et les objectifs de l'étude.

Quant à la normalisation de l'interface de la conception informatique et de la fabrication biomoléculaire, presque rien n'a été fait, malgré d'intéressantes initiatives de la jeune pousse américaine *Ginkgo Bioworks*. Par exemple, pour caractériser ses micro-organismes, *Ginkgo Bioworks* a automatisé le processus complet – une gageure, car le processus fait intervenir une succession de robots incompatibles. En pratique, des écrans décrivant le statut des échantillons vis-à-vis du processus en cours trônent sur les étagères au-dessus des paillasses entre les bouteilles de réactifs. Le résultat en est un nombre réduit d'erreurs et un meilleur contrôle de la qualité.

Pourtant, d'importants gains sont prévisibles si était développé un environnement de conception qui non seulement fournirait une interface de haut niveau commandant les robots biomoléculaires, mais aussi analyserait automatiquement les résultats des expériences et leur reproductibilité. Cette analyse servirait à améliorer le prochain plan expérimental, lui aussi numérisé. Aujourd'hui, c'est encore l'expert humain qui établit cette boucle entre résultats expérimentaux et conception de l'expérience suivante.

Pour comprendre ce que nous réserve la biologie de synthèse dans le futur, il est

important de garder en mémoire son état d'esprit. En faisant des propriétés discutées ci-dessus autant d'objectifs de recherche, cet état d'esprit accélère leurs progrès au long cours. Prenons un exemple. Il y a 15 ans, nul n'aurait pensé la biologie en termes d'« indépendance vis-à-vis du contexte » ; par conséquent, aucun effort n'était réalisé en ce sens. À partir du moment où cette propriété devient une cible, des efforts sont accomplis qui, même si cet objectif reste difficile à atteindre, permettent de s'en approcher.

Malgré quelques avancées, la conception et fabrication de circuits biochimiques de synthèse réalisant les fonctionnalités souhaitées reste un artisanat spécialisé. Comme pour chaque révolution industrielle, la donne ne sera changée que lorsque ce socle conceptuel et méthodologique quittera la tour d'ivoire de quelques chercheurs très spécialisés pour devenir accessible à de nombreux ingénieurs. Cette ouverture requiert de nouveaux « outils », terme pris ici dans sa plus large acception.

L'étape limitante : la phase de conception

Quels types d'outils – et de formations ou filières éducatives correspondantes – manquent le plus pour ouvrir à de nombreux ingénieurs ce qui est aujourd'hui du ressort de quelques équipes de recherche ? Par exemple pour que, si elle devait être refaite, une mise au point comme celle de la biosynthèse d'artémisinine ne requière que le travail durant un ou deux ans d'une équipe réduite, qualifiée certes, mais dont l'expertise ne soit pas quasi unique à l'échelle mondiale ?

La biologie de synthèse fait appel à des méthodes provenant de deux catégories : des méthodes biomoléculaires et d'autres, mathématico-informatiques. La phase de fabrication relève des méthodes biomoléculaires, qui font l'objet d'efforts intenses depuis plusieurs décennies. L'expertise paraît s'être ouverte dans ce domaine : plus d'un million de personnes de par le monde connaissent et utilisent les méthodes avancées de la biologie moléculaire ; des kits et protocoles sont disponibles à tous pour appliquer ces méthodes ; enfin, même les méthodes biomoléculaires moins avancées se sont récemment démocratisées avec le mouvement dit de « biologie de garage », dont l'association *La Paillasse* est le principal représentant en France.

1995 : Le premier génome bactérien est séquencé ; le premier long gène est synthétisé.

1999 : Le premier génome viral est synthétisé.

2000 : Il est possible de concevoir des circuits de régulation et de les implanter dans des bactéries ; le génome humain est séquencé.

2003 : Un repli protéique non naturel est conçu et réalisé.

2008 : Un génome bactérien entier est synthétisé *de novo*.

2010 : Ce génome est introduit avec succès dans une bactérie hôte capable de se reproduire.

2014 : Un chromosome réduit fonctionnel de la levure est synthétisé.

La marge d'amélioration dans le domaine biomoléculaire relève donc plus de l'automatisation des phases de clonage et de caractérisation que d'autres aspects. Par exemple, Amyris, la société de biologie de synthèse qui a développé le processus de bioproduction d'artémisinine, a automatisé ses méthodes de clonage à l'échelle de volumes usuels, puis à des microvolumes. Cette automatisation a considérablement augmenté les chances de succès de ses expériences.

En revanche, la phase de conception, qui relève de modélisations mathématiques et de simulations informatiques, est soumise à plusieurs verrous scientifiques et méthodologiques – une situation qui reflète sans doute un investissement bien moindre. Voici quelques-uns de ces verrous.

D'une part, chaque logiciel dans ce domaine a été conçu par une équipe différente, qui a suivi la logique propre à sa discipline d'origine. Par exemple, pour concevoir une voie métabolique, on fera coexister des modèles métaboliques et d'autres sur la structure et le repliement

des protéines si une enzyme de la voie métabolique doit être modifiée. De plus, toute modification des biomolécules de la voie se fera par génie génétique. Ce dernier s'appuiera sur des robots de biologie moléculaire provenant de divers fabricants. Or les logiciels de l'ingénierie métabolique, de la biologie structurale des protéines, du génie génétique, et des robots de biologie moléculaire « parlent » des langages incompatibles, que ce soit au plan conceptuel ou dans leur implémentation informatique. Il serait donc utile de développer des programmes faisant le lien entre ces différents langages.

Par ailleurs, les outils informatiques disponibles sont encore trop spécialisés et complexes. Il manque des outils qui permettraient d'abaisser le niveau d'expertise nécessaire à un projet. Du point de vue industriel, la biologie de synthèse est encore en enfance. L'expertise est répartie entre une centaine d'équipes à travers le monde. Encore bien plus rares sont les équipes qui maîtrisent l'ensemble de la chaîne requise pour un

grand projet tel celui de l'artémisinine. La biologie de synthèse ne deviendra accessible à une large population d'ingénieurs biotechnologues que lorsque des outils génériques normalisés et aux interfaces simples auront été mis en place.

Un autre saut qualitatif sera effectué quand on utilisera mieux les capacités sensorielles et cognitives des experts. Actuellement, les outils de modélisation ne font appel qu'à la vue d'un écran et au toucher d'un clavier. Il s'agirait de développer les interfaces avec les autres sens : les objets virtuels pourraient être façonnés par le toucher avec retour de force ou par l'ouïe, ce qui permettrait de régler simultanément de multiples paramètres. L'intelligence collaborative devrait aussi être mieux exploitée : ensemble, plusieurs individus résolvent des énigmes plus complexes qu'un individu seul.

Une barrière, enfin, est la taille limite des circuits biochimiques que l'on réussit à construire. Aujourd'hui, la taille des circuits biochimiques viables introduits dans un organisme hôte plafonne autour

DE LA BIOBRIQUE AU CIRCUIT GÉNÉTIQUE

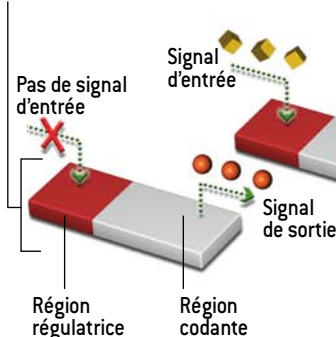
Des assemblages de gènes et de séquences régulatrices d'ADN peuvent être conçus pour fonctionner comme des composants électroniques : ils effectuent une opération booléenne, c'est-à-dire fournissent une réponse binaire [0 ou 1] à un signal d'entrée. Chaque assemblage ainsi construit est une biobrique qui peut être insérée dans un circuit génétique plus vaste.

UNE BIOBRIQUE

Un exemple simple de biobrique est le commutateur biochimique : un gène effectue l'opération booléenne SAUF en réponse à un signal d'entrée – une protéine codée par un autre gène.

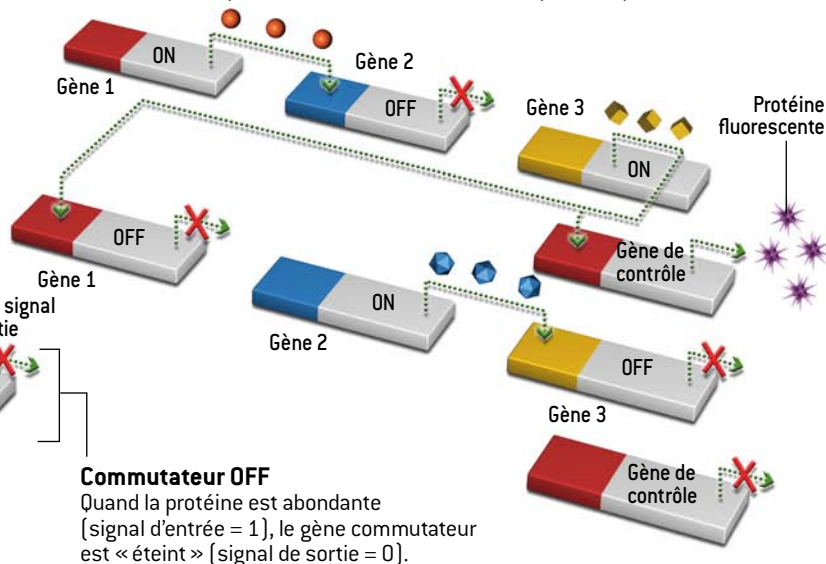
Commutateur ON

Quand la protéine est absente [signal d'entrée = 0], le gène commutateur est « allumé » : la protéine qu'il code est exprimée [signal de sortie = 1].



UN CIRCUIT GÉNÉTIQUE

On construit un circuit génétique simple en connectant trois commutateurs, chacun constitué d'un gène différent (gènes 1, 2 et 3). Les gènes oscillent entre les états ON et OFF selon le signal propagé le long du circuit. Le comportement peut être vérifié par un gène de contrôle qui intercepte la protéine produite par un des commutateurs (ici le gène 3) et produit alors une protéine détectable (fluorescente, par exemple).



Commutateur OFF

Quand la protéine est abondante [signal d'entrée = 1], le gène commutateur est « éteint » [signal de sortie = 0].