

Nos travaux sur *M. genitalium* ne sont que la première des nombreuses étapes à franchir pour modéliser les cellules ou tissus humains à l'échelle des gènes et des molécules. Notre modèle est loin d'être parfait, et les mycoplasmes sont l'expression la plus simple de la vie autosuffisante. Nos simulations, le code source, les données expérimentales, le code de visualisation, la base de connaissances, tout est disponible gratuitement sur Internet. Avec d'autres chercheurs, nous travaillons déjà à améliorer le simulateur et à l'étendre à divers organismes, tels que la bactérie *E. coli* et la levure *Saccharomyces cerevisiae*, toutes deux des organismes utilisés fréquemment tant dans les laboratoires que dans l'industrie.

Chez ces espèces, la régulation des gènes est beaucoup plus complexe, et la localisation des événements dans la cellule joue un rôle bien plus important (la levure, par exemple, présente un noyau). Autant de problèmes à résoudre avant de s'attaquer à des cellules plus complexes encore, tels des macrophages de souris ou

humains. Faciles à cultiver, ces cellules du système immunitaire seront sans doute les premières utilisées pour ajuster et valider le modèle.

Il est trop tôt pour estimer le temps que ces améliorations prendront. Comparées aux bactéries, les cellules humaines comportent beaucoup plus de compartiments. De plus, leur contrôle génétique, plus complexe que celui des mycoplasmes, demeure en grande partie mystérieux. Enfin, de par leurs rôles complémentaires dans les tissus multicellulaires, les cellules humaines interagissent plus avec d'autres types cellulaires que les bactéries.

Début 2008, modéliser la plus simple des cellules ne semblait pas pouvoir se concrétiser avant une décennie, et modéliser une forme de vie plus complexe paraissait inenvisageable. Aujourd'hui, l'idée de simuler une cellule humaine ne paraît plus si saugrenue. Et si le logiciel se révèle inapproprié, au moins cet échec mettra-t-il en lumière les nombreuses inconnues à expliciter encore sur nos propres cellules. ■

■ BIBLIOGRAPHIE

E. Gonçalves *et al.*, **Bridging the layers: toward the integration of signal transduction, regulation and metabolism into mathematical models**, *Molecular Biosystems*, vol. 9, pp. 1576-1583, 2013.

J. R. Karr *et al.*, **A whole-cell computational model predicts phenotype from genotype**, *Cell*, vol. 150, pp. 389-401, 2012.

P. L. Freddolino et S. Tavazoie, **The dawn of virtual cell biology**, *Cell*, vol. 150, pp. 248-250, 2012.

W. Gibbs, **La biologie virtuelle**, *Pour la Science* n° 288, pp. 38-44, 2001.

france culture

LA MARCHÉ DES SCIENCES

Découvertes, inventions, aventures savantes au fil de l'Histoire

Aurélie Luneau
14h-15h / chaque jeudi

franceculture.fr

POUR LA SCIENCE



La biologie de vers une ingénierie

François KÉPÈS

Concevoir un génome *ab initio*, inventer une cellule minimale, produire de nouvelles fonctions biologiques... Telles sont quelques-unes des ambitions de la biologie de synthèse. Son objectif : ouvrir de nouveaux horizons en biologie en s'inspirant des méthodes de l'ingénierie.

Au début des années 1960, l'invention des circuits intégrés a métamorphosé l'ingénierie électronique : un schéma imprimé sur un masque suffisait à déposer sélectivement des couches métalliques et chimiques sur un support, et à produire ainsi le circuit imaginé. Bientôt, les ingénieurs ont puisé dans les circuits déjà conçus pour les combiner en des montages toujours plus complexes. La modélisation informatique aidant, les progrès furent considérables, conduisant aux nombreux circuits intégrés que comportent les téléphones portables, ordinateurs et autres dispositifs électroniques.

Et si une telle combinaison de technique et de méthode était transposable en biologie ? S'il était possible, en biologie aussi, de disposer d'une banque d'outils et de

règles rationnelles qui, combinés selon les besoins, aideraient à comprendre le vivant en le fabriquant, ou à concevoir et produire à la demande de nouvelles fonctions biologiques n'existant pas dans la nature ? Depuis quelques années, cette idée fait son chemin parmi les biologistes. Rassemblées sous le nom de biologie de synthèse, leurs approches se construisent aux interfaces de la biologie et de l'ingénierie, avec des apports de la chimie, de la physique, des mathématiques et de l'informatique. Elles ont pour objectif d'injecter en biologie les principes fondant toute ingénierie. De quoi s'agit-il ? Quelles sont les ambitions de la biologie de synthèse ? Que permet-elle aujourd'hui ? Quelles sont ses perspectives et limites ? Telles sont les questions que nous examinerons ici.

La biologie de synthèse est à la biologie des systèmes – l'analyse des interactions des biomolécules et des comportements des systèmes biologiques – ce que la chimie de synthèse est à la chimie analytique (l'analyse des substances chimiques). Après le triomphe de la chimie analytique dans les années 1850, les chimistes se sont tournés vers la synthèse pour obtenir plus de produits naturels à analyser, puis des composés n'existant pas dans la nature. Cette recherche a donné naissance à l'industrie chimique et à la pétrochimie du XX^e siècle. De même, dans les années 1990, les progrès de la biologie analytique ont été considérables grâce au développement des techniques dites « -omiques » : le séquençage du génome (génomique), la cartographie des ARN (transcriptomique) ou des protéines (protéomique) d'un système biologique, ou des modifications non génétiques du génome (épigénomique)... Stimulée par un volume croissant de données, la biologie des systèmes s'est développée, attirant physiciens et informaticiens. Quelques années plus tard est apparue la biologie de synthèse, qui s'appuie sur la compréhension ainsi acquise.

La biologie de synthèse est l'ingénierie rationnelle de la biologie. En d'autres termes,



synthèse

du vivant

elle vise à la conception rationnelle et à l'ingénierie de systèmes complexes fondés sur le vivant ou inspirés par le vivant, et dotés de fonctions absentes dans la nature. Un exemple de cette démarche est la synthèse par voie biologique de l'artémisinine, médicament antipaludéen de référence dont l'industrialisation a débuté en 2013.

Un médicament contre le paludisme

L'artémisinine est naturellement produite par une plante, l'armoise annuelle. Son usage dans le traitement du paludisme remonte à plus de deux millénaires en Chine et, depuis 2002, l'Organisation mondiale de la santé recommande son utilisation (combinée à d'autres médicaments pour prévenir le développement de résistances du parasite responsable de la maladie). L'artémisinine fait partie d'une classe de molécules naturelles produites par des plantes, les terpénoïdes, utilisées dans des secteurs aussi variés que l'alimentation (arômes), les cosmétiques (fragrances) ou certaines thérapies anticancéreuses (outre la lutte contre le paludisme). Le menthol et le camphre en sont deux exemples.

L'ESSENTIEL

- Apparue dans les années 2000, la biologie de synthèse vise à rationaliser les pratiques en biotechnologie.
- Une fonction biologique peut être vue comme un circuit génétique constitué de plusieurs modules – des gènes produisant des fonctions de base.
- En modélisant et combinant ces modules, les biologistes espèrent concevoir et fabriquer des fonctions biologiques à la demande.
- Cette approche aiderait à mieux comprendre le vivant. Ses applications potentielles sont multiples.

© Kenn Brown/Mandolitic Studios

Au début des années 2000, Jay Keasling, de l'Université de Californie à Berkeley aux États-Unis, et ses collègues ont cherché des méthodes alternatives de fabrication de ces molécules, dont la production naturelle était trop faible et fluctuante en qualité et en quantité. Ils se sont notamment concentrés sur l'artémisinine. La synthèse chimique demandant un grand nombre d'étapes et n'offrant elle aussi qu'un faible rendement, ils se sont tournés vers la synthèse par voie biologique, ou bioproduction. En 2013, après une dizaine d'années d'efforts, leurs travaux ont abouti à la bioproduction, par la société pharmaceutique *Sanofi*, d'environ 60 tonnes annuelles d'artémisinine de synthèse aux caractéristiques stables.

Pour parvenir à cette bioproduction, ils ont combiné plusieurs éléments : un micro-organisme connu, à croissance rapide et aisé à modifier génétiquement, des gènes issus de l'armoise annuelle et des techniques de biologie moléculaire pour optimiser l'expression des gènes. Après quelques tâtonnements, le choix s'est porté sur la levure du boulanger. Cette cellule a l'intérêt, comme d'autres cellules eucaryotes (c'est-à-dire à noyau) non végétales, de produire un précurseur de l'artémisinine. En lui ajoutant les gènes de plusieurs enzymes de l'armoise annuelle impliquées dans la production de l'artémisinine, les biologistes ont réussi à construire, dans la levure, une voie de synthèse complète d'un précurseur de l'artémisinine, l'acide artémisinique. Voie qu'ils ont ensuite améliorée en flanquant chaque gène d'une séquence (un promoteur) favorisant son expression et en effectuant quelques autres modifications raisonnées de la levure. Le rendement est ainsi passé de 1,6 à 28 grammes par litre de fermenteur.

Cette concentration élevée entraîne une précipitation du produit, une caractéristique qui permet d'extraire l'acide artémisinique durant la fermentation. Celui-ci est ensuite transformé en artémisinine par voie chimique (chimie de synthèse optimisée par photochimie).

Cet exemple illustre le principe de la biologie de synthèse : concevoir de nouveaux systèmes biologiques en combinant des biomolécules, des circuits biochimiques servant un métabolisme ou la régulation d'une fonction biologique, et des cellules ou des particules qui abriteront l'ensemble. Trois démarches la caractérisent. La première est une approche constructiviste et hiérarchique :

on part de biobriques, des segments d'ADN remplissant des fonctions définies, et on les assemble dans des dispositifs – des circuits de régulation ou métaboliques –, eux-mêmes insérés dans des cellules qui produiront les fonctions. La bioproduction d'artémisinine en est un exemple.

La deuxième démarche est, à l'inverse, déconstructiviste: on part d'un micro-organisme fonctionnel et on remplace une partie de son génome par un bout de génome de synthèse qui produit la molécule ou la fonction souhaitée. Les biologistes espèrent, à terme, réussir l'opération avec le génome entier, nous y reviendrons. La troisième démarche, enfin, consiste à concevoir et construire des systèmes vivants simplifiés, tels que des protocellules – des vésicules aqueuses fermées par une enveloppe lipidique. Des équipes cherchent aussi à synthétiser de nouveaux nucléotides (les constituants de l'ADN) et acides aminés (les constituants des protéines), jamais observés dans la nature.

Bien que ces approches soient très différentes, elles ont en commun une même volonté: fonder ces recherches sur les prin-

cipes de l'ingénierie, afin de les optimiser, avec l'idée que cela facilitera l'émergence de nouvelles idées et la compréhension du vivant. Qu'est-ce que cela signifie? Les ingénieries de systèmes électroniques ou mécaniques nécessitent des cadres bien établis pour gérer la complexité, des outils fiables pour manipuler les états du système, et des plateformes de tests. La biotechnologie, en revanche, est encore dépourvue de tels cadres, outils et plateformes. Pour mesurer la distance à parcourir, reprenons une à une les propriétés qui caractérisent toute ingénierie. Le projet immédiat de la biologie de synthèse est de raccourcir cette distance.

Une propriété fondamentale de l'ingénierie est le découplage de la conception et de la fabrication: on commence par concevoir les plans d'une voiture ou d'un circuit intégré avant de les fabriquer. Actuellement, pour développer des composants ou circuits biochimiques opérationnels, il faut suivre des protocoles compliqués de construction de fragments d'ADN, user de modèles simples pour guider la conception, et procéder par approxima-

■ L'AUTEUR



François KÉPÈS est directeur de recherche à l'institut de Biologie des Systèmes et de Synthèse (ISSB, Genopole, UEVE, CNRS) et directeur du Programme d'épigénomique à Genopole, à Évry. Il est professeur invité permanent au Collège impérial de Londres.

LA BOÎTE À OUTILS DE LA BIOLOGIE DE SYNTHÈSE

Si, depuis quelques années, il est possible de développer l'ingénierie rationnelle des systèmes biologiques, c'est grâce aux techniques visant à modifier les génomes. Dans le passé, ces méthodes étaient limitées à l'insertion ou à la mutation de gènes soit en des sites aléatoires, soit en des points précis du génome. Les quelques méthodes disponibles pour

diriger les modifications de génomes avaient plusieurs défauts, en particulier une faible efficacité et un coût élevé en main-d'œuvre. L'essor de la biologie de synthèse s'appuie sur plusieurs progrès récents qui ont considérablement accru les capacités d'intervention dirigée et de combinaison des modifications génétiques. Voici les principaux outils utilisés aujourd'hui.



LES RECOMBINASES sont des enzymes qui permettent d'insérer de longs fragments d'ADN dans le génome en des sites spécifiques.



CERTAINS INTRONS DE GROUPE II sont des éléments génétiques « égoïstes » qui s'incorporent eux-mêmes dans l'ADN hôte *via* une molécule intermédiaire d'ARN.



LES NUCLÉASES à doigts de zinc ou à domaines TAL sont des protéines dont la séquence peut être manipulée afin de cibler une séquence particulière d'ADN. Cette activité de fixation spécifique peut être associée à une autre, telle que le clivage de l'ADN ou l'interaction avec une autre protéine.



L'INGÉNIERIE RECOMBINOGÈNE fait appel aux propriétés des virus bactériens (les phages) pour effectuer des recombinaisons ciblées du génome, c'est-à-dire y insérer des gènes en des endroits précis.



LES NUCLÉASES GUIDÉES par ARN du type CRISPR offrent plus de flexibilité que les précédentes. Les CRISPR naturels sont des systèmes immunitaires bactériens dégradant l'ADN étranger de virus. On les utilise dans nombre d'organismes pour effectuer des modifications dirigées par des molécules d'ARN conçues sur mesure.



LE GÉNIE GÉNÉTIQUE EN MULTIPLEX permet d'incorporer simultanément dans la même cellule plusieurs fragments d'ADN en différents points ciblés de son chromosome. Dans la méthode MAGE, le ciblage est effectué grâce à des fragments d'ADN que la cellule assimile à des intermédiaires de la réplication du génome lors de la division cellulaire. Ces fragments se retrouvent ainsi incorporés dans le génome.

tions successives suivies d'interminables réglages fins. En conséquence, le cycle de conception, fabrication et dépannage de composants et circuits biochimiques reste lent et sujet à erreurs.

Il s'agirait, dans le futur, de concevoir l'objet virtuel et de tester ses propriétés avant de le construire. Durant l'époque intermédiaire que nous vivons actuellement, ces deux phases de conception puis de fabrication, bien que distinctes, doivent en général être réitérées en raison de l'imperfection des modèles utilisés. En ce cas, les résultats des tests sur l'objet fabriqué servent à recalibrer le modèle plus près de la cible. En somme, une spirale d'amélioration passant alternativement par la conception et la fabrication doit être instaurée pour parvenir à un résultat satisfaisant.

La conception comprend la mise en place d'un modèle mathématique; sa simulation numérique permet de modifier les paramètres jusqu'à obtenir le résultat souhaité. La fabrication comprend la construction du circuit de synthèse respectant les paramètres ainsi obtenus. La construction fait appel au génie génétique de l'ADN codant les biomolécules nécessaires (*voir l'encadré page ci-contre*); puis les molécules et circuits ainsi construits sont caractérisés par des tests adéquats.

Des modules biologiques indépendants

Une deuxième propriété fondatrice de l'ingénierie est l'indépendance du contexte ou l'orthogonalité: chaque composant électronique fonctionne de la même façon quel que soit le circuit auquel il est intégré. Dans la cellule vivante, les interactions des composants de synthèse et cellulaires sont difficiles à prédire ou à caractériser. Un circuit bien caractérisé dans une bactérie fonctionnera différemment dans une autre souche, même si elle est proche. Cela ne respecte pas le principe d'indépendance d'un circuit de synthèse par rapport à son contexte (ici cellulaire). De nombreux travaux visent à diminuer cette dépendance vis-à-vis du contexte. Un exemple en est l'introduction répétée d'isolateurs transcriptionnels de part et d'autre des gènes introduits dans la cellule. Les isolateurs sont de courtes séquences d'ADN ayant pour effet de terminer la transcription une fois que celle des gènes d'intérêt a été effectuée, afin de ne pas «contaminer» l'expression d'un gène voisin.

Au-delà de l'indépendance par rapport au contexte cellulaire se pose la question de la modularité, qui mesure l'indépendance entre les divers composants du circuit de synthèse introduit dans la cellule. La biologie de synthèse se propose de constituer des systèmes (nanoparticule, cellule) par assemblage de dispositifs (oscillateur, interrupteur, etc. *voir l'encadré page 33*), eux-mêmes réalisés à partir de briques (fragments d'ADN, protéines, ARN). La fabrication d'un objet fondé sur la biologie, ou inspiré par elle, devient donc un processus hiérarchique d'assemblage de modules. Idéalement, les propriétés de chaque module ne devraient pas dépendre du circuit de synthèse dans lequel il est immergé. Tout praticien de la biologie moléculaire sait que cet objectif est aujourd'hui lointain. Néanmoins, chaque progrès sur ce chemin sera porteur d'améliorations pour le biologiste.

Enfin, dans la dimension hiérarchique et modulaire décrite ci-dessus, la biologie de synthèse exige une normalisation des composants, nécessaire pour permettre leur assemblage en dispositifs, puis l'assemblage de ces dispositifs en systèmes. On est cependant encore loin d'atteindre une norme unique stabilisée: récemment encore, de nouveaux standards ont été proposés. De plus, il s'agit de normaliser non seulement les composants, mais aussi les dispositifs et la production biologique, sans oublier les systèmes hybrides qui combinent nanobiologie et nanoélectronique.

La normalisation concerne aussi la dimension mathématico-informatique de la biologie de synthèse. En général, deux familles de formalismes sont utilisées pour exprimer les questions scientifiques: les équations différentielles et les approches booléennes. Dans le premier cas, chaque interaction entre deux biomolécules peut être représentée par une équation différentielle qui définit la dépendance entre leurs concentrations, c'est-à-dire l'évolution après un délai court des concentrations de ces molécules dans la cellule. Souvent, la même molécule intervient dans plusieurs interactions, et sa concentration se retrouve donc dans plusieurs équations. On dit alors que les équations sont couplées. Le jeu consiste à intégrer ces équations couplées pour calculer au prochain pas de temps les nouvelles concentrations de toutes les biomolécules. On trace ainsi l'évolution temporelle de ces concentrations, c'est-à-dire le «comportement» du système.

■ BIBLIOGRAPHIE

N. Annaluru *et al.*, **Total synthesis of a functional designer eukaryotic chromosome**, *Science*, vol. 344, pp. 55-58, 2014.

K. M. Esvelt et H. H. Wang, **Genome-scale engineering for systems and synthetic biology**, *Mol. Syst. Biol.*, vol. 9, pp. 641-657, 2013.

C. J. Paddon *et al.*, **High-level semi-synthetic production of the potent antimalarial artemisinin**, *Nature*, vol. 496, pp. 528-532, 2013.

F. Képès, **La biologie de synthèse plus forte que la nature ?** Le Pommier, 2011.

W. Weber et M. Fussenegger, **Emerging biomedical applications of synthetic biology**, *Nature Reviews Genetics*, vol. 13, pp. 21-35, 2011.

D. G. Gibson *et al.*, **Creation of a bacterial cell controlled by a chemically synthesized genome**, *Science*, vol. 329, pp. 52-56, 2010.

■ SUR LE WEB

Informations générales sur la biologie de synthèse : www.biologie-de-synthese.fr

Observatoire français de la biologie de synthèse : biologie-synthese.cnam.fr

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

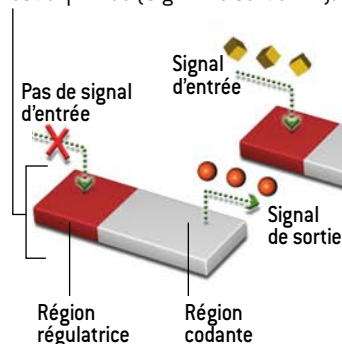
Des assemblages de gènes et de séquences régulatrices d'ADN peuvent être conçus pour fonctionner comme des composants électroniques : ils effectuent une opération booléenne, c'est-à-dire fournissent une réponse binaire [0 ou 1] à un signal d'entrée. Chaque assemblage ainsi construit est une biobrique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

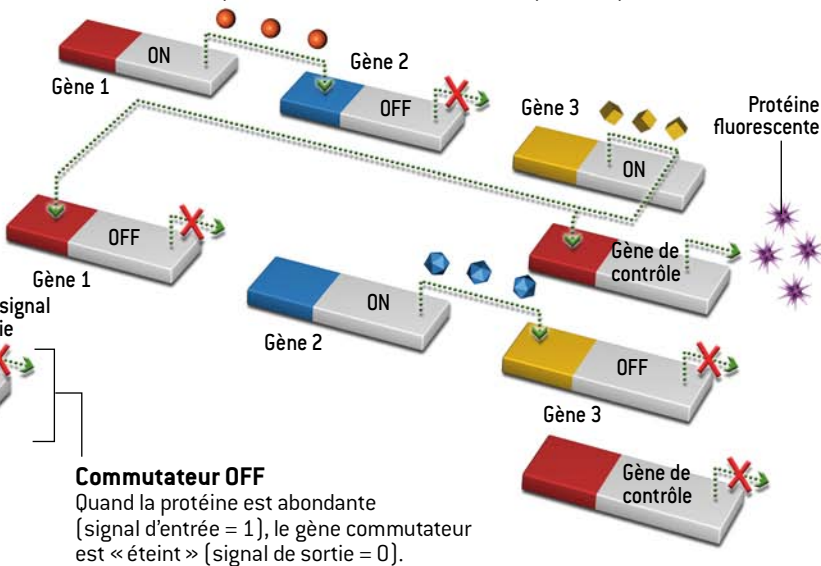
Commutateur ON

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].

UN EXEMPLE DE CIRCUIT GÉNÉTIQUE

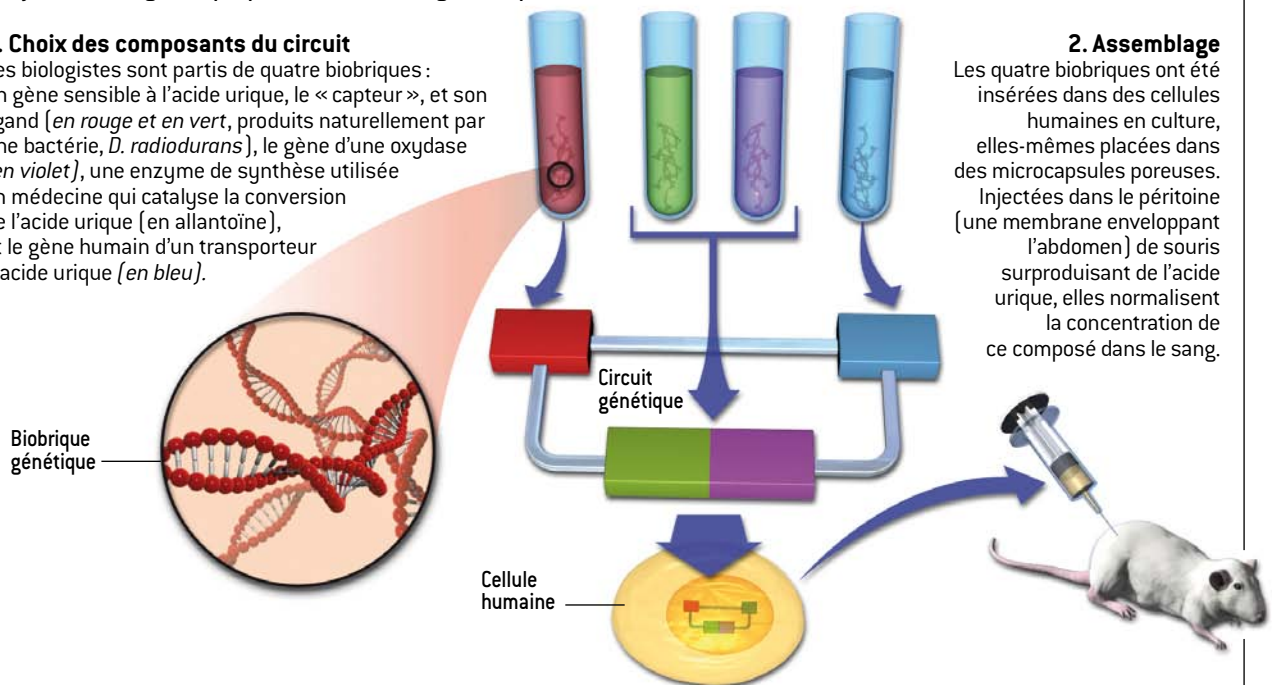
Une concentration élevée d'acide urique dans le sang peut entraîner une maladie, la goutte [des dépôts d'acide urique se forment dans les articulations, la peau, les reins...]. Des biologistes de l'École polytechnique fédérale de Zurich, en Suisse, ont conçu un circuit génétique pour contrôler et réguler en permanence cette concentration. Celui-ci fonctionne chez la souris.

1. Choix des composants du circuit

Les biologistes sont partis de quatre biobriques : un gène sensible à l'acide urique, le « capteur », et son ligand (en rouge et en vert, produits naturellement par une bactérie, *D. radiodurans*), le gène d'une oxydase (en violet), une enzyme de synthèse utilisée en médecine qui catalyse la conversion de l'acide urique (en allantoine), et le gène humain d'un transporteur d'acide urique (en bleu).

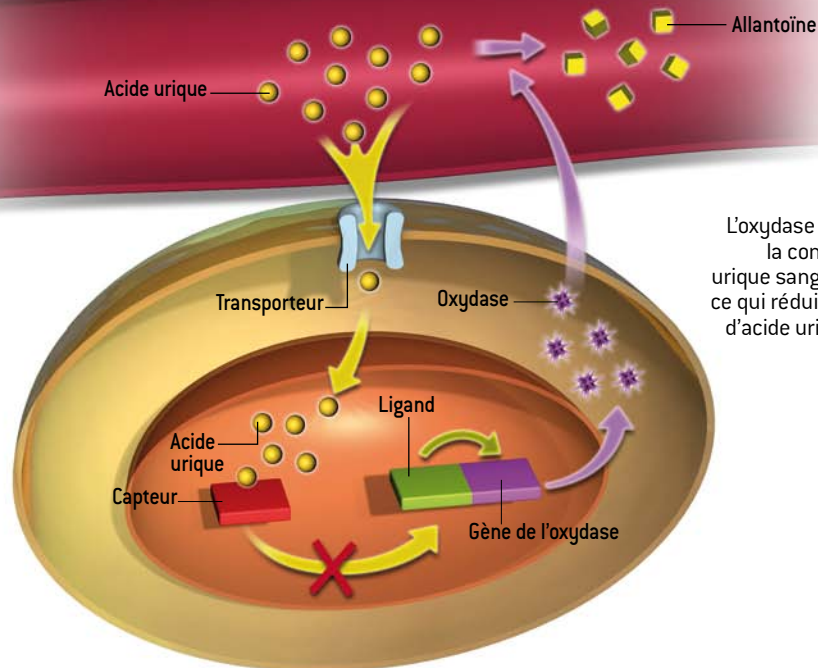
2. Assemblage

Les quatre biobriques ont été insérées dans des cellules humaines en culture, elles-mêmes placées dans des microcapsules poreuses. Injectées dans le péritoine (une membrane enveloppant l'abdomen) de souris surproduisant de l'acide urique, elles normalisent la concentration de ce composé dans le sang.



3. Transport

Les cellules modifiées expriment le transporteur d'acide urique (en bleu). Une partie de l'acide urique (en jaune) en excès dans le sang pénètre ainsi dans la cellule. L'acide urique entre dans le noyau de la cellule, où il se lie au « capteur » (en rouge).



5. Catalyse

L'oxydase produite catalyse la conversion de l'acide urique sanguin en allantoin, ce qui réduit la concentration d'acide urique dans le sang.

4. Régulation

En se liant à son « capteur » (en rouge), l'acide urique l'empêche de se fixer à son ligand (en vert). Or le ligand, lorsqu'il n'est pas masqué par le capteur, agit comme un déclencheur (un promoteur) de l'expression de l'enzyme oxydase (en violet). En d'autres termes, le capteur régule l'expression de l'enzyme oxydase : quand il se fixe à son ligand, l'expression de l'enzyme est bloquée, et inversement.

En quête d'un génome minimal

Un des grands projets de la biologie de synthèse est la création d'un génome minimal – un génome qui ne contiendrait que les gènes nécessaires au fonctionnement de la cellule qui l'héberge, et ceux qui produiront les fonctions biologiques souhaitées.

L'intérêt d'utiliser un génome réduit est multiple. D'une part, cela diminuerait le coût énergétique de la réplication du génome lors de la division cellulaire. D'autre part, l'ADN codant les fonctions synthétiques pourrait être inséré à la place libérée.

Enfin, certaines séquences du génome, dites recombinogènes, sont connues pour favoriser les recombinaisons génétiques. En les retirant, on diminue le risque que le génome subisse des

modifications spontanées. Plusieurs équipes sont impliquées dans la conception et la fabrication de tels génomes réduits. Ainsi, la société américaine *Scarab Genomics* commercialise des souches de colibacille dont l'ADN a été réduit de 15 pour cent, notamment par élimination des séquences recombinogènes. Cette société a été fondée en 2002 par Fred Blattner, dont l'équipe séquençait le premier génome de colibacille.

À l'Université Johns Hopkins, aux États-Unis, Jef Boeke (aujourd'hui à l'Université de New York) et Joel Bader ont lancé le projet Sc2.0 d'un génome synthétique d'une cellule eucaryote, la levure de bière *Saccharomyces cerevisiae*. À ce jour, un chromosome a été synthétisé entièrement. Là aussi, les séquences recombinogènes ont été supprimées. En revanche, des séquences ont été introduites pour induire à volonté de nombreux réarrangements aléatoires de morceaux de chromosome. Il est donc possible de disposer d'un génome stable, mais, par un simple stimulus, d'induire son évolution accélérée. Ensuite

les variants issus de cette évolution peuvent être sélectionnés sur un critère quelconque.

En 2010, l'Institut américain J. Craig Venter a publié la synthèse chimique du génome d'une bactérie, un mycoplasme (environ un million de paires de nucléotides). Introduit dans une souche cousine, ce génome était toujours fonctionnel : les cellules obtenues étaient capables de se diviser. Cette prouesse technique ouvre la voie à des génomes totalement synthétiques. En revanche, la séquence synthétisée était une copie quasi identique de la version naturelle. À l'évidence, le chemin sera encore long avant de pouvoir concevoir *de novo* un génome...

d'une dizaine de réactions. C'est le cas du circuit de bioproduction de l'artémisinine. Comment briser ce plafond pour construire des circuits biochimiques plus complexes et augmenter d'autant les possibilités de la biologie de synthèse ?

Pour une part, ce plafond tient à notre incapacité à concevoir un génome complet. Pour étendre fortement la taille et la complexité des circuits, il faudrait pouvoir introduire un nombre arbitraire de gènes qui coopéreraient de façon optimale selon un cahier des charges fixé par avance. Cela sous-entend de connaître l'ensemble des contraintes qui pèsent sur les génomes naturels des hôtes envisagés – une levure ou une bactérie, par exemple. La version modifiée de ce génome devrait non seulement tenir compte de ces contraintes, mais en tirer parti pour optimiser les circuits de synthèse.

Concevoir un génome *ab initio*

Plusieurs équipes se sont ainsi lancées dans des recherches visant à synthétiser ou à réduire des génomes complets d'organismes unicellulaires. C'est le cas, par exemple, de l'équipe de l'Institut J. Craig Venter, à Rockville aux États-Unis, et de celle de Fred Blattner, de l'Université du Wisconsin, à Madison, pour les bactéries,

ou encore de celle de Jef Boeke et Joel Bader, de l'Université Johns Hopkins, à Baltimore, pour la levure (voir l'encadré ci-dessus).

Ces travaux représentent des tours de force techniques. Paradoxalement, ils soulignent aussi les carences dans notre capacité à concevoir *ab initio* un génome. Par exemple, le génome de mycobactérie synthétisé par l'équipe de J. Craig Venter est quasi copié du génome naturel. C'est pourtant bien la conception à l'échelle du génome et non la copie qui permettra d'agrandir la taille et la complexité des circuits biochimiques, la synthèse d'ADN permettant ensuite de réaliser ce génome à façon.

Dans mon équipe, nous cherchons précisément à exploiter les contraintes connues des génomes de micro-organismes pour la conception *ab initio* de génomes porteurs de grands circuits de synthèse. En particulier, nous avons observé que les gènes qui fonctionnent ensemble (par exemple parce qu'ils sont régulés par les mêmes protéines) ne sont pas disposés au hasard le long des génomes. Au contraire, on les retrouve côte à côte ou à intervalles périodiques. Ces schémas semblent optimiser l'expression des gènes. Ce type d'observation conduit à proposer de respecter ces schémas « naturels » lors de l'insertion de gènes co-fonctionnels dans des chromosomes microbiens – une donnée

qui devrait aider à augmenter la complexité des circuits.

On le voit, la biologie de synthèse n'est pas simplement un nouveau domaine de la biologie. Elle apporte un vaste socle méthodologique et conceptuel, ainsi qu'un état d'esprit nouveau en biotechnologie. Construire un système biologique qui fonctionne comme prévu est une façon de s'assurer que l'on a compris les phénomènes sous-jacents. En ce sens, la biologie de synthèse permet de faire progresser les connaissances sur le monde vivant. L'approche de l'ingénieur étant dominante, les applications industrielles ne sont cependant jamais loin.

Le potentiel économique de ces applications est considérable et concerne de nombreux domaines : santé (diagnostic, médicaments), agroalimentaire, environnement (biocapteurs de pollution), énergie (biocarburants), matériaux (biomatériaux « intelligents »), etc. En ouvrant toutes ces possibilités, la biologie de synthèse renouvelle les questions éthiques concernant la responsabilité des hommes à « artificialiser » le vivant. Jusqu'à quel point voulons-nous modifier ou recréer le vivant ? Quelle gouvernance de la biologie de synthèse adopter pour qu'elle corresponde bien à nos attentes ? Il est indispensable que ces questions soient dès maintenant continuellement débattues. ■

Biologie de synthèse et sciences sociales, un dialogue difficile

Pierre-Benoît Joly et Benjamin Raimbault

Quel sera l'impact de la biologie de synthèse sur la société ? Comment éviter les dérives ? Les sciences sociales apporteraient un éclairage décisif si elles n'étaient pas cantonnées à un rôle inapproprié de médiateur.

Domaine de recherche en plein essor depuis le début des années 2000, la biologie de synthèse nourrit d'importantes ambitions théoriques puisqu'il s'agit d'appliquer les principes de l'ingénierie à la biologie. Conjointement, en s'annonçant comme la nouvelle révolution industrielle, elle revendique un grand potentiel d'applications et, par conséquent, de profonds changements économiques et sociaux, ses applications devant toucher des domaines aussi variés que la santé, l'agroalimentaire, l'environnement ou la production d'énergie. En pratique, si certains discours laissent entendre que la biologie de synthèse sera la corne d'abondance du début du XXI^e siècle, très peu d'applications sont aujourd'hui disponibles, à l'exception de certains cas emblématiques tels que la semi-synthèse de l'artémisinine, un médicament antipaludéen produit désormais en grande quantité par *Sanofi*.

Néanmoins, la biologie de synthèse alimente déjà plusieurs débats au sein de la société. Par exemple, à la suite de la recommandation de l'Organisation mondiale de la santé d'utiliser des combinaisons thérapeutiques à base d'artémisinine, des milliers de petits producteurs asiatiques

L'ESSENTIEL

- Pour anticiper les conséquences sociales des futures applications de la biologie de synthèse, les biologistes ont développé des collaborations avec les chercheurs en sciences sociales.
- Plusieurs approches ont été utilisées, mais des difficultés sont apparues.
- Les sciences sociales ont un rôle à jouer, mais celui-ci n'est ni de vaincre la contestation ni d'être le porte-parole de la société.
- Il est nécessaire de clarifier l'agenda de recherche et les questions de fond sur la biologie de synthèse et sa gouvernance.

© Seamartin Graphics/shutterstock

et africains ont développé la culture de la plante qui la produit naturellement, l'armoïse annuelle. Après une période alternant surproduction et sous-production d'artémisinine naturelle, une initiative de planification à l'échelle mondiale de médicaments à base d'artémisinine a vu le jour en 2009, permettant de consolider la filière et de répondre aux besoins. Produite au même prix à partir de 2013, l'artémisinine de synthèse vient ainsi concurrencer les petits producteurs et perturber un marché encore fragile sans pour autant constituer un réel progrès médical. Dès lors, comment s'assurer que les applications de la biologie de synthèse seront bien en adéquation avec les besoins sociétaux et environnementaux, et limiter leurs effets négatifs ?

Une des principales originalités de la biologie de synthèse a été de vouloir anticiper les conséquences sociales de ses futures applications en développant de façon systématique une collaboration avec les sciences sociales. Depuis les débuts de la biologie de synthèse, plusieurs approches ont été menées dans ce sens, mais celles-ci ont leurs limites. Quelles sont-elles et comment les dépasser ? C'est ce que nous allons examiner.