

Calculer avec l'ADN

LEONARD ADLEMAN

L'utilisation d'ADN pour résoudre des problèmes mathématiques conduit à repenser la notion de «calcul».



Quand on évoque les ordinateurs, on pense claviers, écrans, souris, mémoire vive, circuits micro-électroniques, gigaoctets et mégahertz. Les succès de l'informatique classique ont fait croire que seules des puces de silicium pouvaient recevoir des circuits capables d'effectuer des calculs, mais est-ce exact? L'ordinateur que vous utilisez inconsciemment pour lire ces mots, votre cerveau, n'a pas grand-chose à voir avec les ordinateurs classiques. Peut-être avons-nous une vision trop étroite de ce qu'est le calcul. Ne gagnerait-on pas à élargir la notion d'ordinateur? Certains peuvent-ils être des liquides, où les calculs résulteraient d'interactions de molécules? De telles machines existent : ce sont les ordinateurs à ADN.

Je m'intéresse à ces systèmes depuis un séjour dans un laboratoire de biologie moléculaire, en 1993. Mathématicien et informaticien, j'ai d'abord fait de la recherche médicale, contribuant notamment à des études du SIDA. Comment montrer aux biologistes l'intérêt des modèles mathématiques ou informatiques? Sous la houlette de Nicolas Chelyapov, j'ai commencé à apprendre les méthodes de la biologie moderne.

Ce travail est fascinant : on crée de l'ADN qui n'existe pas dans la nature et on l'introduit dans des bactéries où

il déclenche la production de protéines qui peuvent changer la nature même de l'organisme. Pendant cette période d'apprentissage, je lus la *Biologie moléculaire du gène*, un manuel classique de biologie dû à James Watson, l'un des deux découvreurs de la structure en double hélice de l'ADN, et ma conception de la biologie en fut bouleversée : la biologie moléculaire n'était plus la science des produits odorants conservés dans des réfrigérateurs ; elle gagnait en profondeur et en puissance explicative. La biologie est aujourd'hui l'étude de l'information stockée dans l'ADN par des enchaînements de quatre lettres, A, T, G et C, les quatre bases qui sont l'adénine, la thymine, la guanine et la cytosine ; la biologie étudie aussi les modifications de ces informations dans la cellule. Ces modifications se font selon des règles mathématiques.

Un soir, dans mon lit, alors que je lisais le livre de J. Watson, j'en arrivai à la description de l'ADN polymérase, la reine des enzymes, celle qui crée la vie : dans un milieu approprié, cette enzyme fabrique, à partir d'un brin d'ADN, un ADN complémentaire, où chaque C est remplacé par un G, chaque G par un C, chaque A par un T et chaque T par un A. Par exemple, en partant

d'une molécule dont la séquence est CATGTC, l'ADN polymérase engendre une nouvelle molécule de séquence GTACAG. L'ADN polymérase assure la reproduction de l'ADN, ce qui permet aux cellules de se reproduire et, indirectement, à l'espèce humaine de se développer : pour un biologiste réductionniste, la réplication de l'ADN par l'ADN polymérase est l'essence même de la vie.

L'ADN polymérase est une machine étonnante : à elle seule, cette molécule s'accroche à un brin d'ADN, se déplace dessus en «lisant» chaque base rencontrée et en «écrivant» son complément, fabriquant ainsi un nouveau brin d'ADN. Alors que je pensais à cette enzyme remarquable, je compris qu'elle fonctionnait comme une machine décrite en 1936 par le mathématicien anglais Alan Turing. Ce dernier, ainsi que Kurt Gödel, Alonzo Church et Stephen Kleene, avait analysé la notion de «calculabilité». Ces travaux théoriques, qui précédaient de plus d'une décennie les premiers ordinateurs, ont abouti à certains des résultats les plus importants des mathématiques du XX^e siècle (voir, par exemple, *Le hasard en théo-*



1. LES MOLÉCULES D'ADN stockent de l'information. Ce sont des enchaînements d'adénine, de thymine, de guanine et de cytosine, respectivement représentées par les lettres A, T, G et C. La molécule dessinée ici en couleurs, de séquence GCAGTCGGACTGGGCTATGTCCGA, code la solution du problème du chemin hamiltonien représenté page suivante.

rie des nombres, par Gregory Chaitin, *Pour la Science*, septembre 1988).

Turing avait inventé un ordinateur abstrait que l'on nomme aujourd'hui «machine de Turing». Cette machine est composée de deux bandes, ainsi que d'un système nommé contrôleur à états finis, qui se déplace le long d'une bande pour y lire des données et, simultanément, se déplace le long de la deuxième bande pour y écrire d'autres données. Le contrôleur se programme à l'aide d'instructions simples, de sorte qu'il est aisé d'écrire un programme qui lit sur la bande d'entrée une chaîne composée de A, de T, de C et de G, et qui écrit la chaîne complémentaire sur la bande de sortie. La similitude avec l'ADN polymérase est patente.

La similarité est intéressante, parce que la machine de Turing est universelle, c'est-à-dire qu'elle peut être programmée pour calculer tout ce qui est calculable. Cette idée est la «thèse de Church». Elle me fit penser que l'on pourrait programmer une machine de Turing pour produire une chaîne complémentaire, pour factoriser des nombres ou pour jouer aux échecs. Je me dressai soudain dans mon

lit et fis remarquer à mon épouse : «Bon sang! Ces choses peuvent calculer.» Je ne dormis pas de la nuit, cherchant un moyen d'utiliser l'ADN pour résoudre des problèmes.

Je voulus d'abord concevoir un ordinateur à ADN analogue à une machine de Turing, où le contrôleur à états finis serait remplacé par une enzyme. Une idée identique avait déjà été proposée dix ans plus tôt par des chercheurs de la Société IBM (voir *Les limites physiques du calcul*, par Charles Bennett et Rolf Landauer, *Pour la Science*, septembre 1985). Malheureusement, si une enzyme (l'ADN polymérase) permet de créer des ADN complémentaires, où trouverait-on les molécules qui assureraient d'autres calculs importants, telle la factorisation des nombres?

Les biologistes moléculaires ont la chance inouïe de pouvoir utiliser le contenu des cellules. Nous sommes très loin de savoir créer *ex nihilo* des machines moléculaires aussi perfectionnées que l'ADN polymérase, mais trois ou quatre milliards d'années d'évolution ont engendré des cellules pleines de petites merveilles. La biotechnologie moderne emploie certaines de ces

molécules, mais l'évolution n'a jamais produit de cellules qui jouent aux échecs. Aussi, pour construire des machines cellulaires aux capacités de calcul intéressantes, je devais utiliser les outils disponibles. Ces outils sont essentiellement les suivants :

1. L'appariement des bases. À chaque brin d'ADN correspond une molécule complémentaire. Dans une solution, lorsqu'un brin d'ADN rencontre l'ADN complémentaire, les deux brins s'apparient, c'est-à-dire qu'ils s'enroulent l'un autour de l'autre pour former une double hélice. Les brins ne réagissent pas chimiquement, mais se lient parce que des forces faibles, des «liaisons hydrogène», s'établissent entre les bases complémentaires. Quand un brin d'ADN rencontre un autre brin d'ADN qui ne lui est pas complémentaire ou qui ne possède pas de longues parties complémentaires, alors les deux brins ne se lient pas.

2. Des polymérases dupliquent l'information d'une molécule en créant une molécule complémentaire. Par exemple, l'ADN polymérase fabrique un brin d'ADN complémentaire à partir de n'importe quelle molé-