

problème était «NP-complet» : la plupart pensent qu'on ne trouvera jamais d'algorithme rapide pour résoudre ce problème, bien que la preuve de cette supposition reste à fournir (voir *Les lois du tout ou rien*, par Jean-Paul Delahaye, *Pour la Science*, juillet 1995). Le problème du chemin hamiltonien n'est pas insoluble, mais le temps de calcul risque d'être excessivement long.

Considérons par exemple l'algorithme suivant :

Soit un graphe possédant n sommets.

1. Engendrer un ensemble de chemins aléatoires.
2. Pour chaque chemin de l'ensemble :
 - a. Vérifier si le chemin part du sommet de départ et se termine au sommet d'arrivée. Sinon, le retirer de l'ensemble.
 - b. Vérifier si ce chemin passe exactement par n sommets. Sinon, le retirer de l'ensemble.
 - c. Pour chaque sommet, vérifier que ce chemin y passe. Sinon, retirer ce chemin de l'ensemble.
3. Si l'ensemble n'est pas vide, alors il existe un chemin hamiltonien. Si l'ensemble est vide, alors il n'existe pas de chemin hamiltonien.

Cet algorithme est imparfait, mais si l'ensemble des chemins est engendré de manière suffisamment aléatoire et si l'ensemble résultant est suffisamment grand, alors cet algorithme donnera probablement la solution. C'est cet algorithme que

j'ai mis en œuvre dans la première expérience de calcul avec l'ADN.

Sept jours dans un laboratoire

Pour cette première expérience, j'ai cherché un problème suffisamment simple pour qu'un prototype d'ordinateur à ADN le résolve, mais suffisamment important pour constituer une preuve indubitable de l'intérêt de ces nouvelles machines. J'ai choisi un graphe comportant sept villes et 14 vols, représenté page 58. À la main, on met en moyenne une minute pour trouver l'unique chemin hamiltonien de ce graphe (vous pouvez commencer maintenant...).

Pour examiner la technique utilisée, considérons le graphe qui contient quatre villes seulement : Toulouse, Grenoble, Rennes et Lille, reliées par six vols. On cherche un chemin hamiltonien qui part de Toulouse et qui finit à Lille.

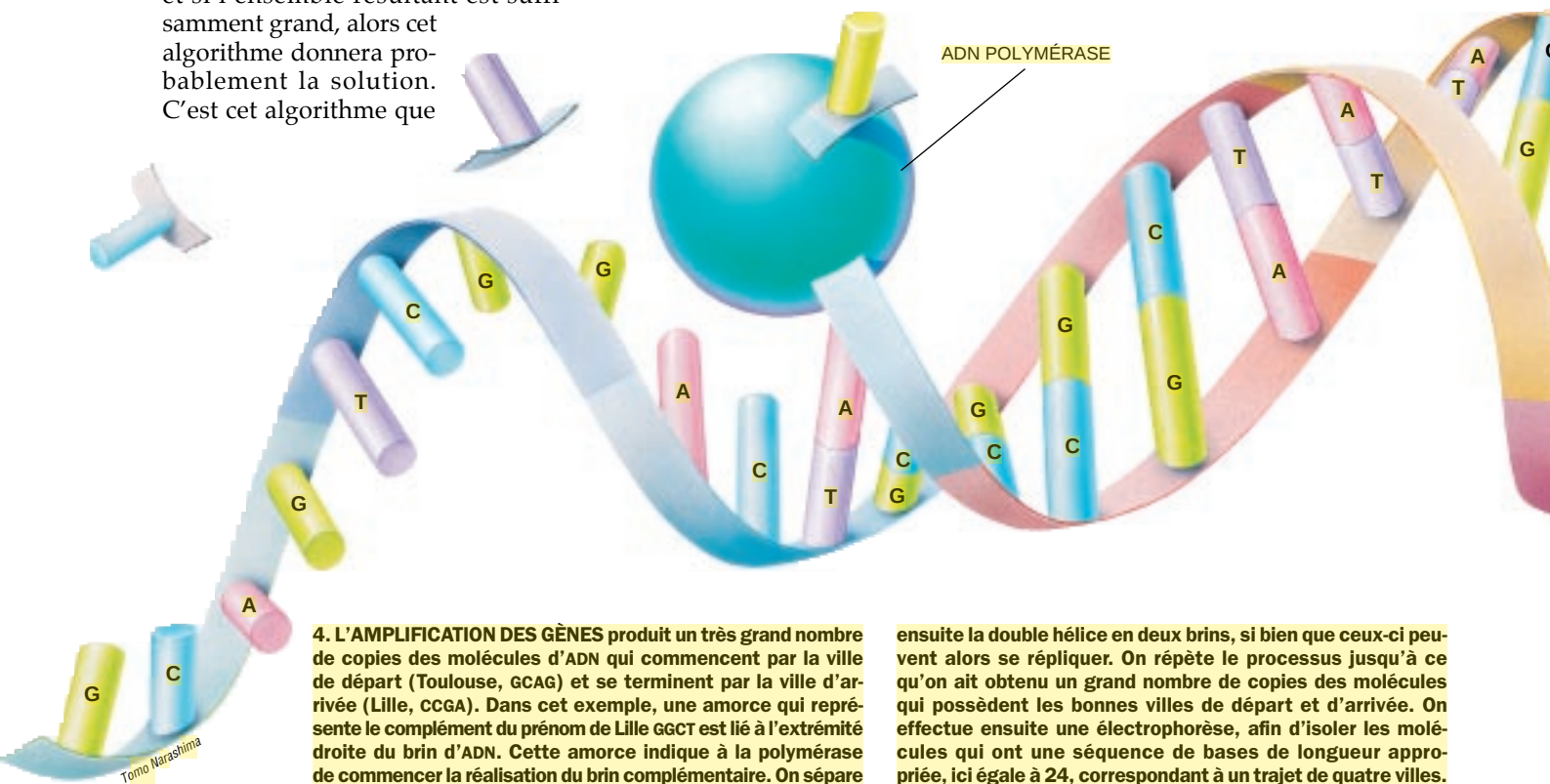
J'ai commencé par attribuer une séquence aléatoire d'ADN à chacune des villes : à Toulouse, j'ai associé ACTT GCAG, et à Grenoble TCGGACTG. J'ai alors imaginé que la première moitié des séquences était le prénom des villes et la seconde moitié leur nom propre. Ainsi le nom propre de Toulouse est GCAG, et le prénom de Grenoble TCGG.

J'ai ensuite codé chaque vol en mettant bout à bout le nom propre de la ville d'origine avec le prénom de la ville d'arrivée. Dans notre exemple, le code du vol reliant Toulouse à Grenoble est GCAGTCGG.

Comme on peut former un ADN complémentaire de tout brin d'ADN, chaque ville possède un nom complémentaire. Le nom complémentaire de Toulouse est TGAACGTC.

Après avoir effectué ces codages, j'ai synthétisé les séquences d'ADN complémentaires des noms de ville et les séquences des codes des vols. J'ai ensuite pris une pincée de chaque séquence, environ 10^{14} molécules que j'ai placées dans un tube à essai. Pour que le calcul commence, j'ai simplement ajouté de l'eau, de la ligase, du sel et quelques autres composants qui reproduisent les conditions qui règnent à l'intérieur d'une cellule. En tout, j'ai utilisé un cinquième de cuillère à café de solution et, une seconde plus tard, j'avais la solution du problème du chemin hamiltonien.

Afin de comprendre le déroulement du calcul, examinons les réactions dans le tube. Le numéro du vol reliant Toulouse à Grenoble (GCAGTCGG), par exemple, et le nom complémentaire de Grenoble (AGCCTGAC) peuvent se rencontrer par hasard. Or, par construction, la première séquence se termine

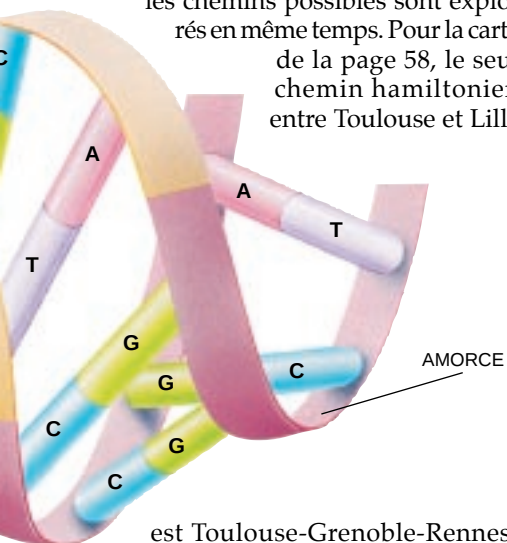


par TCGG, et la deuxième commence par AGCC. Ces séquences étant complémentaires, elles s'associent. La molécule complexe ainsi formée peut ensuite rencontrer la molécule associée au vol Grenoble-Rennes (ACTGGGCT), qui se liera à son tour au complexe, parce que la fin de celui-ci (TGAC) est complémentaire du début du code du vol (ACTG). Les molécules s'allongent ainsi à mesure que les codes des vols s'associent aux noms complémentaires des villes. Les ligases du mélange relient en permanence les chaînes d'ADN codant les vols. Dans le tube, les molécules restantes représentent des chemins aléatoires passant par les différentes villes. La première étape de notre algorithme est terminée.

Comme on introduit initialement un très grand nombre de molécules et que le problème ne fait intervenir qu'un petit nombre de villes, on obtient très vraisemblablement au moins une molécule qui code le chemin hamiltonien cherché. Ainsi la solution d'un problème mathématique est stockée dans une simple molécule!

Remarquons aussi que tous les chemins sont créés en même temps grâce aux interactions simultanées de centaines de milliers de milliards de molécules. Cette réaction biochimique est un gigantesque calcul parallèle : tous les chemins possibles sont explorés en même temps. Pour la carte

de la page 58, le seul chemin hamiltonien entre Toulouse et Lille



est Toulouse-Grenoble-Rennes-Lille. La molécule qui code cette solution possède donc la séquence GCAGTCGGACTGGGCTATGTCCGA. Cette solution est présente dans le tube sous la forme double brin, c'est-à-dire avec la séquence complémentaire appariée.

Toutefois, cette molécule se trouvait parmi les centaines de milliards

Cristal d'ADN

La science est très loin de pouvoir créer des ordinateurs moléculaires et, sans l'exemple donné par la vie elle-même, il ne serait pas évident qu'une telle organisation est possible. Toutefois, l'expérience décrite dans cet article a montré comment les bases complémentaires assurent l'auto-assemblage des molécules d'ADN en une structure linéaire. Des structures autres que linéaires peuvent-elles être construites de la sorte?

Les travaux de Nadrian Seeman, de l'Université de New York, sur les structures d'ADN branchées apportent un élément de réponse. Les structures d'ADN branchées, des molécules d'ADN avec plusieurs extrémités, sont présentes dans la cellule lors de la réplication de l'ADN et pendant la recombinaison. Avec ses collègues, N. Seeman a appris à façonner des molécules d'ADN en des structures branchées. Sa construction d'un cube dont les

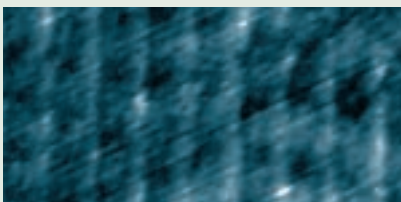
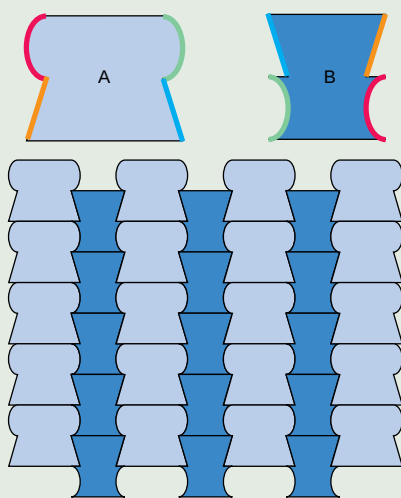
arêtes sont des molécules d'ADN marque le début de ce que l'on nomme «les nanotechnologies de l'ADN».

En 1997, je fis le lien entre les travaux de N. Seeman et ceux de L. Adleman grâce à la théorie des pavages (voir les pavages de Penrose dans l'article de Martin Gardner, page 86). La théorie des pavages est intéressante non seulement par sa nature géométrique, mais aussi parce qu'elle est liée à la logique et au calcul. Dans les pavages, la forme des bords des pavés détermine quels pavés peuvent être assemblés. Dans un brin d'ADN, la séquence des bases détermine quels brins s'associent. Le passage aux «pavés moléculaires» constitués de brins d'ADN était alors naturel.

Avec le groupe de N. Seeman, nous avons construit le premier pavage moléculaire en laboratoire. Deux pavés moléculaires A et B (deux structures branchées), s'assemblent d'une façon unique pour produire un pavage rayé. Chaque pavé moléculaire comporte quatre extrémités avec des séquences uniques ; chaque extrémité du pavé A est complémentaire d'une extrémité du pavé B, de sorte que, lorsque l'on mélange de nombreux pavés dans une solution, ils s'assemblent spontanément en un cristal bidimensionnel d'ADN (voir la figure).

À première vue, ce simple motif rayé ne ressemble en rien à du calcul. Toutefois, les pavés moléculaires sont «programmables» : on peut modifier facilement les extrémités d'assemblages de chaque pavé pour déterminer leurs interactions. En fait, on peut construire ainsi une infinité de pavés, dont des pavés qui, comme dans l'expérience d'Adleman, simuleraient une machine de Turing par la «logique» de leur assemblage. L'auto-assemblage des pavés d'ADN peut fonctionner comme un ordinateur universel.

Erik WINFREE,
Institut de technologie de Californie



Le premier pavage bidimensionnel d'ADN est obtenu en mélangeant deux sortes de pavés d'ADN, des structures d'ADN qui comportent des extrémités complémentaires (segment colorés, en haut). Ces pavés s'assemblent selon une structure rayée unique (au centre), révélée par microscopie à force atomique (en bas).