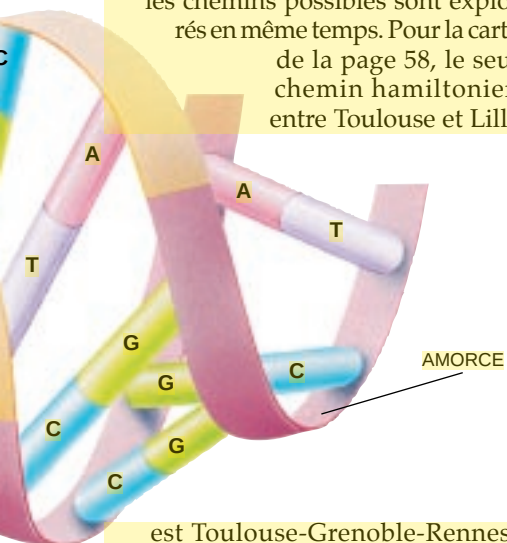


par TCGG, et la deuxième commence par AGCC. Ces séquences étant complémentaires, elles s'associent. La molécule complexe ainsi formée peut ensuite rencontrer la molécule associée au vol Grenoble-Rennes (ACTGGGCT), qui se liera à son tour au complexe, parce que la fin de celui-ci (TGAC) est complémentaire 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

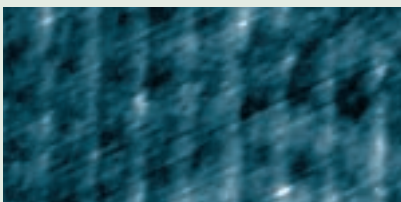
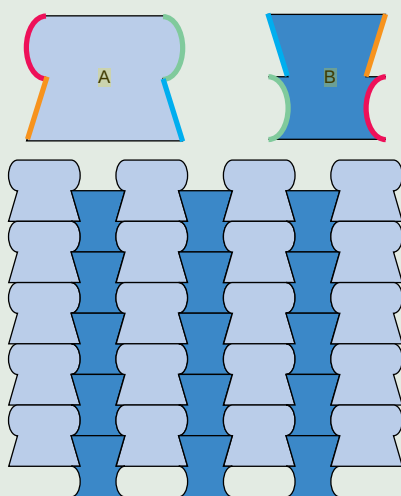
arêtes sont des molécules d'ADN marque le début de ce que l'on nomme «les nanotechnologies de l'ADN».

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.

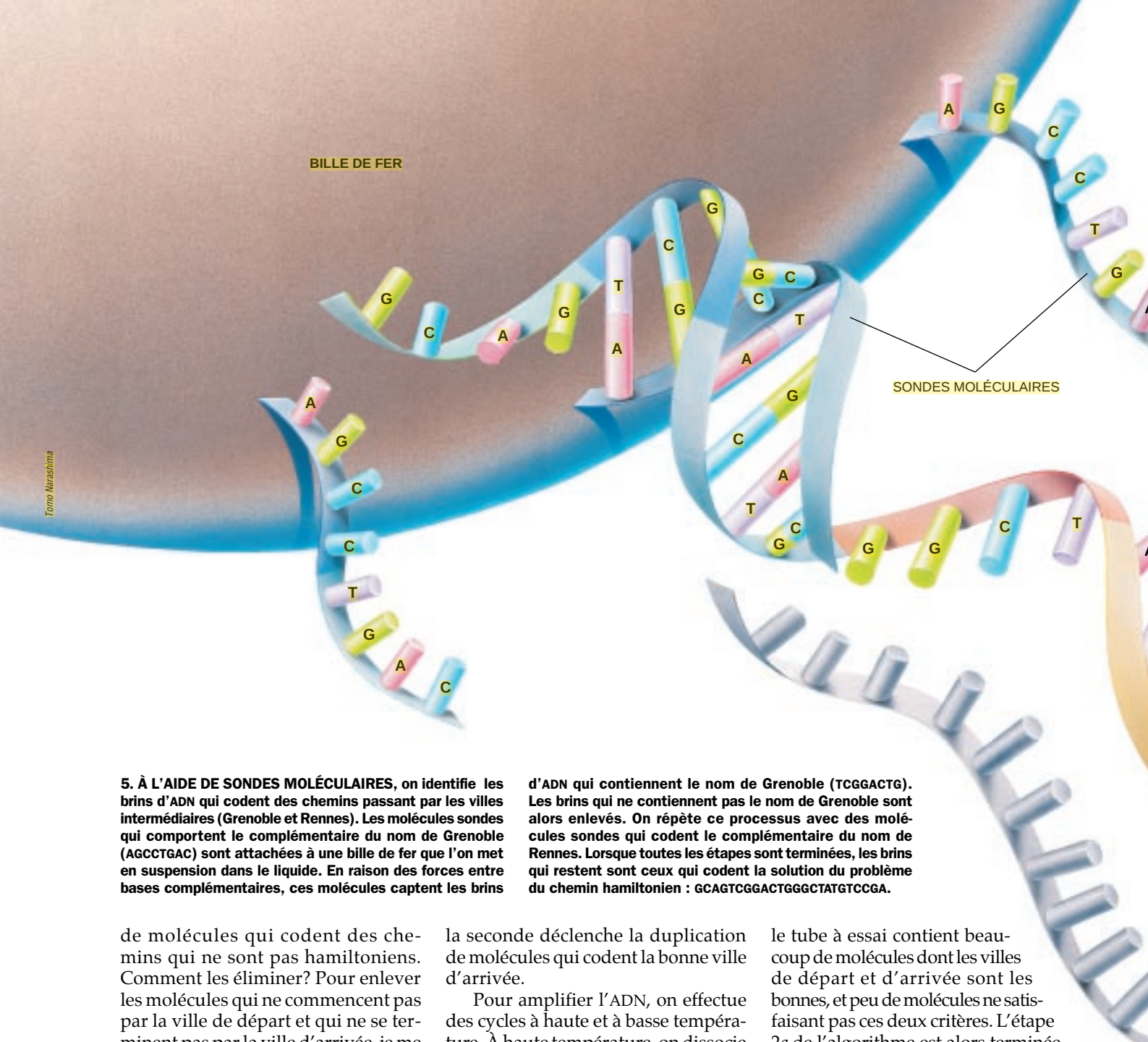
Avec le groupe de N. Seeman, nous avons construit le premier pavage moléculaire en laboratoire. Deux pavés moléculaires A et B (deux structures branchées), s'assemblent d'une façon unique pour produire un pavage rayé. Chaque pavé moléculaire comporte quatre extrémités avec des séquences uniques ; chaque 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).



5. À L'AIDE DE SONDAS MOLÉCULAIRES, on identifie les brins d'ADN qui codent des chemins passant par les villes intermédiaires (Grenoble et Rennes). Les molécules sondes qui comportent le complémentaire du nom de Grenoble (AGCCTGAC) sont attachées à une bille de fer que l'on met en suspension dans le liquide. En raison des forces entre bases complémentaires, ces molécules captent les brins

d'ADN qui contiennent le nom de Grenoble (TCGGACTG). Les brins qui ne contiennent pas le nom de Grenoble sont alors enlevés. On répète ce processus avec des molécules sondes qui codent le complémentaire du nom de Rennes. Lorsque toutes les étapes sont terminées, les brins qui restent sont ceux qui codent la solution du problème du chemin hamiltonien : GCAGTCGGACTGGGCTATGTCCGA.

de molécules qui codent des chemins qui ne sont pas hamiltoniens. Comment les éliminer? Pour enlever les molécules qui ne commencent pas par la ville de départ et qui ne se terminent pas par la ville d'arrivée, je me suis servi de la technique d'amplification des gènes (PCR). Cette technique utilise de nombreux exemplaires de deux séquences «amorces» complémentaires des deux extrémités de la séquence d'ADN que l'on désire amplifier. Ainsi, l'ADN polymérase effectue plusieurs cycles de réplication, alternativement sur chacun des brins. Les amorces utilisées sont, d'une part, le nom propre de la ville de départ (GCAG pour Toulouse) et, d'autre part, le complémentaire du prénom de la ville d'arrivée (GGCT pour Lille). Ces deux amorces agissent simultanément : la première prévient l'ADN polymérase de copier les séquences qui possèdent la bonne ville de départ,

la seconde déclenche la duplication de molécules qui codent la bonne ville d'arrivée.

Pour amplifier l'ADN, on effectue des cycles à haute et à basse température. À haute température, on dissocie les complexes double brin pour libérer les séquences complémentaires des amorces. À basse température, on apparie les amorces, et l'ADN polymérase effectue la réplication, recréant une molécule double brin.

Ainsi, les molécules qui possèdent les bonnes villes de départ et d'arrivée se multiplient à une vitesse très élevée (exponentielle). En revanche, les molécules qui possèdent la bonne ville de départ, mais pas la bonne ville d'arrivée, ou réciproquement, se multiplient à une vitesse inférieure (linéaire). Les séquences d'ADN qui ne possèdent ni la bonne ville de départ, ni la bonne ville d'arrivée ne se reproduisent pas du tout. Après la réaction,

le tube à essai contient beaucoup de molécules dont les villes de départ et d'arrivée sont les bonnes, et peu de molécules ne satisfaisant pas ces deux critères. L'étape 2a de l'algorithme est alors terminée.

J'utilise ensuite un gel d'électrophorèse pour identifier les molécules qui possèdent la bonne longueur, 24 dans notre exemple ; toutes les autres molécules sont éliminées. C'est l'étape 2b de notre algorithme.

Afin de vérifier que les séquences qui restent codent bien des chemins qui passent par les villes intermédiaires, j'utilise la complémentarité des séquences d'ADN dans une procédure nommée séparation par affinité. On utilise des sondes d'ADN qui codent le complémentaire du nom d'une ville particulière, par exemple Grenoble. Ces sondes sont fixées à de microscopiques sphères de fer d'un micromètre de diamètre environ.