

Séquençage de l'ADN par ouverture mécanique de la double hélice: une évaluation théorique

Ultrafast sequencing of DNA by mechanical opening of the double helix: a theoretical investigation

JEAN-LOUIS VIOVY ⁽¹⁾, CHRISTOPH HELLER ⁽¹⁾, FRANÇOIS CARON ⁽²⁾, PHILIPPE CLUZEL ⁽³⁾, DIDIER CHATENAY ⁽³⁾

⁽¹⁾ Laboratoire de Physicochimie Théorique (Unité Associée au CNRS 1382), ESPCI, 10, rue Vauquelin, 75231 Paris Cedex 05, France.

⁽²⁾ Laboratoire de Génétique Moléculaire, UA CNRS 1302, École Normale Supérieure, rue d'Ulm, 75005 Paris, France.

⁽³⁾ Laboratoire PSI, Institut Curie, UA CNRS 1379, 11, rue Pierre-et-Marie-Curie, 75005 Paris, France.

RÉSUMÉ

Nous proposons un modèle mécanique pour l'ouverture contrôlée d'une double hélice d'ADN, envisagée en tant que méthode de séquençage. En utilisant diverses données moléculaires, nous proposons des évaluations numériques qualitatives des forces mises en jeu et des temps caractéristiques importants du système. Nous montrons que le principal obstacle à la réalisation pratique de cette approche réside dans l'élasticité moléculaire des portions d'ADN déjà séquencées, qui « lissent » le profil de force en fonction du déplacement, et favorisent l'ouverture thermique spontanée des paires de bases. ▲

Mots clés : DNA, séquençage, nucléotide, dynamométrie.

ABSTRACT

We propose and evaluate a model experiment, in which the sequence of a DNA fragment is determined by mechanically opening the double helix in a controlled manner (e.g. pulling on the 3' end of one strand), and measuring the variation of the force exerted by the base pairs on a nanodynamometer (e.g. on a bead in an optical trap or a glass microneedle attached to the 5' end of the other strand). We show that the major limitation of the approach is the longitudinal elasticity of the already sequenced single strand sections, which soften the displacement-force function, and facilitate spontaneous thermal opening of the base pairs. ▲

Key words : DNA, sequencing, nucleotide, dynamometry.

Abridged version (see p. 799)

Les dernières années ont vu des progrès considérables dans les méthodes de cartographie des génomes [1]. Par opposition, les techniques de séquençage (lecture de l'enchaînement des nucléotides ou « bases » Adénine, Thymine, Guanine et Cytosine, A, T, G, C) progressent lentement, et constitueront pendant de longues années encore le principal goulot d'étranglement des projets de recherche sur le génome [2]. Actuellement, le séquençage s'effectue par électrophorèse, et cette technique reste lente malgré de gros efforts d'automatisation. Une nouvelle méthode, imaginée récemment par notre équipe et par d'autres [3, 4], consisterait à « ouvrir »

mécaniquement la double hélice comme une « fermeture Éclair » tout en mesurant les fluctuations de la force mise en jeu pour cela. Étant donné la différence de nombre de liaisons hydrogène et d'enthalpie de liaison entre les paires AT et GC, respectivement, on pense que leur force de cohésion est également différente. Une mesure en continu de la force d'attraction entre brins devrait donc permettre de remonter à la séquence, avec une rapidité de lecture beaucoup plus grande qu'en électrophorèse. Cet article présente à notre connaissance la première évaluation théorique de la faisabilité de cette méthode.

Principe

Le principe de la mesure est simple à énoncer.

Dans un premier temps, on prépare une molécule d'ADN en « Y », comme représentée sur la Figure 1. Le « tronc »

Note présentée par Pierre-Gilles de Gennes.

Note remise le 22 juin 1994, acceptée après révision le 27 juillet 1994.

Correspondance : J.-L. Viovy.

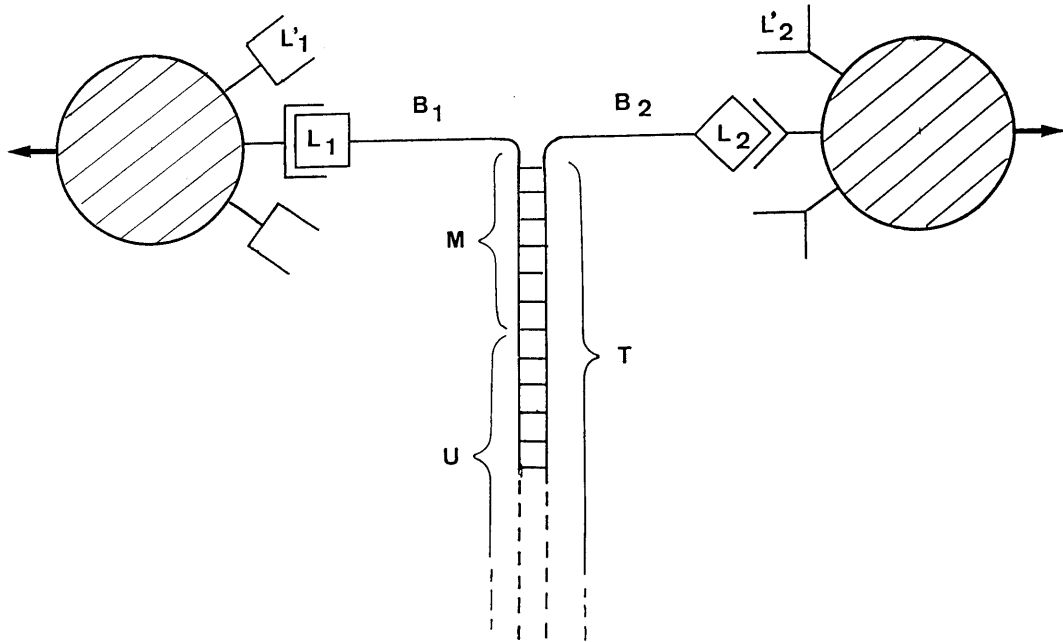


Figure 1. **Représentation schématique d'une double hélice d'ADN en cours d'ouverture mécanique par un picodynamomètre.** U: séquence inconnue; M: marqueur de repérage; B: « bras » déjà séquencés; L₁, L₂: ligands; L'₁, L'₂: récepteurs permettant l'accrochage sur le dynamomètre (ici deux billes de latex).

T du Y est une paire double-brin contenant la séquence inconnue (U), commençant par un marqueur de séquence connue (M). Les deux « bras » B₁ et B₂, qui peuvent être constitués d'ADN simple-brin ou partiellement double-brin, sont accrochés sur des billes ou une autre surface faisant partie d'une micro « machine de force ». Le greffage d'ADN sur des surfaces peut être réalisé par une variété de méthodes connues en chimie et en biologie moléculaire. Une possibilité consiste à ajouter deux ligands L₁ et L₂ (par exemple, de la biotine (bio-) ou de la digoxygénine (dig-) à l'extrémité de l'ADN par des réactions enzymatiques ou chimiques [5-9].

Subséquentement, les ligands servent à lier chacun des bras à des récepteurs complémentaires de L₁ et L₂ (L'₁ et L'₂ sur la Fig. 1) portés par les surfaces (dans l'exemple cité plus haut, L'₁ représenterait la streptavidine et L'₂ l'antidigoxygénine, mais d'autres couples antigène/anticorps pourraient évidemment être utilisés). Ces différentes étapes peuvent en principe être réalisées avec des protocoles connus de biologie moléculaire (cependant, il est honnête de signaler que leur mise en œuvre coordonnée peut poser des problèmes insoupçonnés, et que la faisabilité pratique de l'ensemble reste à démontrer).

On écarte ensuite de façon contrôlée les deux bras du Y, en mesurant la force exercée entre eux. Celle-ci doit passer par un maximum quand on « ouvre » une liaison entre deux bases, et on s'attend que ce maximum dépende de la nature de la paire de bases puisque l'énergie des paires AT et GC est différente. La micromachine de force pourrait utiliser un principe déjà décrit dans la littérature. Smith *et al.* [10], par exemple, ont greffé un côté d'une molécule linéaire d'ADN à une surface de verre, et l'autre à une bille magnétique soumise à un champ magnétique,

pour mesurer l'élasticité de l'ADN [10]. On pourrait aussi utiliser, pour mesurer la force, la déflexion d'une fine aiguille de verre, comme dans [11], le déplacement d'une bille de latex dans un « piège optique », comme dans [12], ou un levier réfléchissant, comme dans la microscopie à force atomique [13]. En pratique, cependant, on aurait besoin pour le séquençage d'une résolution spatiale de l'ordre de l'angström avec un taux d'erreur très faible. Les performances des dispositifs les plus performants actuellement décrits dans la littérature sont loin de satisfaire ces exigences : dans, par exemple, la résolution spatiale est de l'ordre du micromètre. Dans [11, 12], des déplacements de quelques nanomètres ont été mis en évidence, mais seulement à travers une analyse statistique. Enfin, la microscopie à force atomique [13] atteint bien des résolutions de l'ordre de l'angström, mais seulement sur des molécules adsorbées sur une surface, ce qui dans le cas qui nous occupe empêcherait la libre rotation du « tronc » nécessaire à l'ouverture de la double hélice. On voit donc que, bien que simple en principe, le séquençage par ouverture de la double hélice requerra encore un ou plusieurs sauts technologiques avant d'être praticable.

Notons, finalement, qu'il faudrait également être capable de lever l'indétermination qui subsiste, avec la méthode proposée plus haut, entre une paire AT et une paire TA (resp. GC et CG), qui ont la même énergie. On peut espérer que de subtiles interactions de proches voisins modifient suffisamment les interactions entre bases pour cela, mais cela requerrait des mesures encore plus précises. Une autre possibilité serait de substituer certains nucléotides par des nucléotides modifiés présentant une énergie d'appariement différente (par exemple l'inosine ou le 7-déaza-guanosine remplaçant la guanosine [14-17]). En résumé, le séquençage par dynamométrie est un projet séduisant, mais sa faisabilité avec les

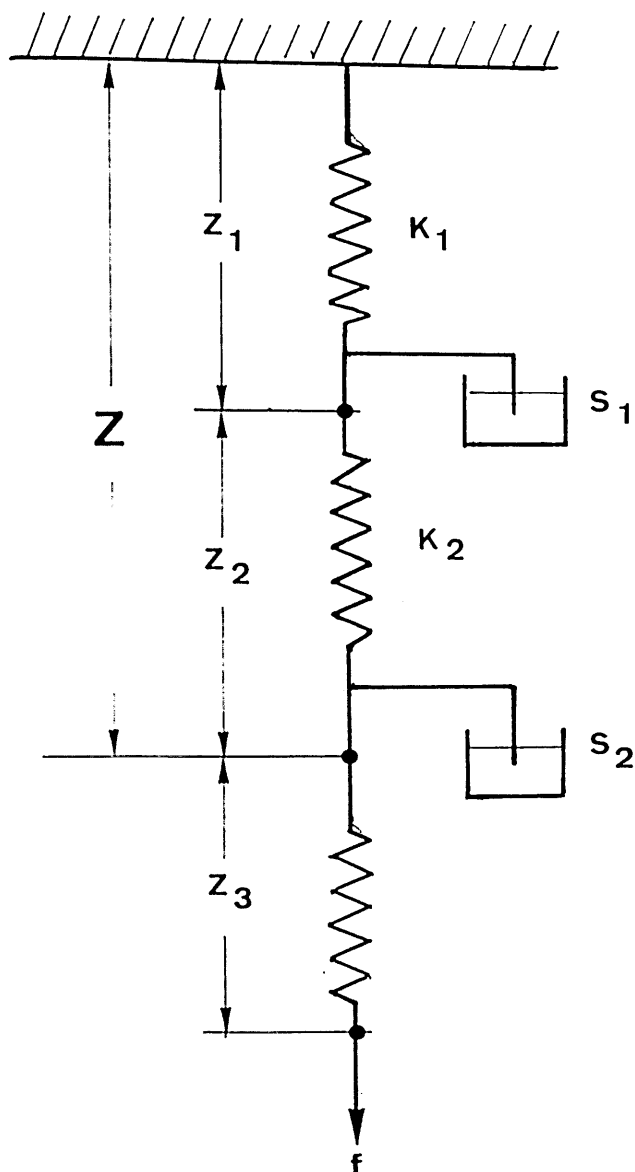


Figure 2. *Modèle mécanique en « ressorts/amortisseurs ».* k_1 , z_1 : raideur et allongement du dynamomètre; k_2 , z_2 : raideur et allongement de la paire de « bras » (une fois déduite la longueur étirée des bras sans allongement ou flexion de liaisons); k_3 , z_3 : raideur et allongement de la paire de bases en cours d'ouverture.

technologies existantes est tout sauf évidente. Il semble donc raisonnable, pour guider les progrès dans ce programme, d'évaluer plus précisément la véritable nature des sauts technologiques nécessaires en utilisant les connaissances disponibles sur la physique de l'ADN.

Pour le physicien, le dispositif de mesure décrit dans l'introduction peut être représenté schématiquement par un assemblage de couples « ressorts/amortisseurs » comme celui représenté dans la Figure 2, où on trouve le « dynamomètre » proprement dit, de raideur k_1 et de friction ζ_1 (microlevier ou piège optique), les deux « bras » représentés par un système unique de raideur k_2 et de friction ζ_2 , et finalement la paire de bases en cours de traction représentée par k_3 (on peut raisonnablement

négliger sa friction devant celle des bras). En général, ces systèmes ne sont pas harmoniques. Considérons tout d'abord l'équilibre statique.

Caractéristique force/déplacement des paires de bases

L'énergie de la paire double-brin doit présenter, en fonction de z_3 , la forme d'un escalier aux marches plus ou moins arrondies et de hauteur inégale (de l'ordre de (40/N) kJ pour une paire AT et (60/N) kJ pour une paire GC, où N est le nombre d'Avogadro) [18]. Par la suite, on raisonne sur une séquence $(GC)_n$, pour simplifier. La force correspondante doit avoir l'allure représentée dans la Figure 3, où a est la distance totale sur laquelle se fait sentir l'attraction entre bases, et b la longueur libérée par l'ouverture d'une paire. Seuls des calculs de modélisation moléculaire assez poussés permettraient d'évaluer la courbe $f(z_3)$ de façon précise. En l'absence de ceux-ci, on représente approximativement la force par une sinusoïde :

$$f \cong f_M (1 + \cos(2\pi z/a)) \quad (1).$$

Il est probable que les liaisons hydrogène ne « résistent » pas de façon coordonnée, et on évalue a à 0,4 nm environ. Pour b , on choisit la longueur de deux bases entièrement étirées en configuration trans, $b \cong 1,5$ nm, car la tension doit empêcher l'appariement des bases du « bras » en simple hélice (voir paragraphe suivant). On évalue alors aisément :

$$f_M \cong (2/a) \int_{-a/2}^{a/2} f dz_3 \cong 500 \text{ pN} \quad (2).$$

Raideur des « bras »

Un point essentiel de la présente modélisation est que les chaînes d'ADN sont en permanence soumises à une tension très supérieure à la force d'attraction entre cycles π (forces de « base stacking ») et à celles mises en jeu par les forces thermiques. Ainsi chaque base d'une paire désappariée, une fois « extraite » de la double hélice, doit prendre une configuration « trans » par rapport à toutes les liaisons à rotation libre. Une fois ce réarrangement configurationnel effectué, l'essentiel de l'élasticité de la chaîne provient de l'extension des liaisons elles-mêmes, et de l'ouverture des angles de liaison. On représente donc les « bras » par l'association en série d'un segment rigide, de longueur approximative 7,5 Å/base, et d'un ressort harmonique représentant la raideur des liaisons. Typiquement, la raideur associée à l'extension de liaisons simples, k_{λ} , est de l'ordre de 5×10^{-8} N/Å [19]. Cela conduit, pour une base dont le squelette comporte 6 liaisons, et après projection sur la direction de traction, à une valeur approximative :

$$k_{2,\lambda} \cong 1 \text{ N/Å/base} \quad (3).$$

La raideur d'ouverture d'angles de liaison est typiquement de l'ordre de 5×10^{-9} N/Å [19]. L'énergie associée à un petit déplacement Δz est donnée par :

$$\Delta E = k_{\theta} \Delta z^2 / l^2 \sin^2 \theta \quad (4),$$

8. Zischler H., Nabda I., Schäfer R., Schmid M., Epplen J. T. 1989. Digoxigenated oligonucleotide probes specific for simple repeats in DNA fingerprinting and hybridization *in situ*. *Hum. Genet.* 82: 227-33.
9. Urdea M. S., Warner B. D., Running J. A., Stempien M., Clyne J., Horn T. 1988. A comparison of non-radioisotopic hybridization assay methods using fluorescent, chemiluminescent and enzyme labeled synthetic oligodeoxyribonucleotide probes. *Nucleic Acids Res.* 16: 4937-56.
10. Smith S. B., Finzi L., Bustamante C. 1992. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science* 258: 1122-6.
11. Ishijima A., Doi T., Sakurada K., Yanagida T. 1991. Sub-piconewton force fluctuations of actomyosin *in vitro*. *Nature* 352: 301-6.
12. Finer J. T., Simmons R. M., Spudich J. A. 1994. Single myosin molecule mechanics : piconewton forces and nanometre steps. *Nature* 368: 113-9.
13. Lyubchenko Y., Shlyakhtenko L., Harrington R., Oden P., Lindsay S. 1993. Atomic force microscopy of long DNA : imaging in air and under water. *Proc. Natl. Acad. Sci. USA* 90: 2137-40.
14. Imman R. B., Baldwin R. L. 1964. Helix-random coil transition in DNA homopolymer pairs. *J. Mol. Biol.* 8: 452-69.
15. Mills D. R., Kramer F. R. 1979. Structure independent nucleotide sequence analysis. *Proc. Natl. Acad. Sci. USA* 76: 2232-5.
16. Mizusawa S., Nishimura S., Seela F. 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. *Nucleic Acids Res.* 14: 1319-24.
17. Barr J. P., Thayer R. M., Layborn P., Najarian R. C., Seela F., Tolan D. R. 1986. 7 deaza-2'-deoxyguanosine-5'-triphosphate : enhanced resolution in M13 dideoxy sequencing. *BioTechniques* 4: 428-32.
18. Stryer L. 1988. *Biochemistry*, 3rd ed. New York : W. H. Freeman Publ.
19. Hopfinger A. J. 1973. *Conformational properties of macromolecules*. New York : Academic Press.