

RECHERCHE COOPÉRATIVE SUR PROGRAMME N° 25

CLAUDE WEBER

Questions de topologie en biologie moléculaire

Les rencontres physiciens-mathématiciens de Strasbourg - RCP25, 1994, tome 46
« Conférences de A. Beauville, A. Coste, J. Dubochet, A. Gramain, M. Holschneider, C. Itzykson, P. Le Tallec, V. Rivasseau, C. Weber », , exp. n° 9, p. 227-242

http://www.numdam.org/item?id=RCP25_1994__46__227_0

© Université Louis Pasteur (Strasbourg), 1994, tous droits réservés.

L'accès aux archives de la série « Recherche Coopérative sur Programme n° 25 » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

QUESTIONS DE TOPOLOGIE EN BIOLOGIE MOLÉCULAIRE.

par CLAUDE WEBER (Genève).

§1. LE MODÈLE DE CRICK ET WATSON.

En 1953 parut dans une note de trois pages publiée dans Nature le modèle célèbre de Francis Crick et James Watson. Ce modèle satisfaisait la condition nécessaire d'être compatible avec ce que l'on savait à l'époque sur l'ADN: présence de 4 bases, structure stéréochimique de ces bases, règle de Chargaff, figures de diffraction aux rayons X pour des fibres (et non pas des cristaux) d'ADN obtenues par R. Franklin et M. Wilkins. Mais son grand mérite était ailleurs: il suggérait un processus simple pour la transmission des caractères génétiques lors de la mitose. Il y a là quelque-chose d'important. A aucun moment Crick et Watson ne démontrent que l'ADN est vraiment comme ils le proposent. D'ailleurs la note commence par ces mots: "We propose a structure ...". Les confirmations ne viendront qu'au début des années '80 via diffraction aux rayons X de cristaux d'ADN (et non pas seulement de fibres). C'est l'élégance du modèle qui fut dans les années '50 un argument important pour convaincre les sceptiques.

Rappelons les grandes lignes de la structure proposée. Un brin d'ADN est constitué d'une chaîne linéaire formée alternativement d'un sucre et d'un phosphate. A chaque sucre s'attache de côté une base parmi 4 possibles: l'adénine, la cytosine, la guanine et la thymine. La structure de la liaison sucre-phosphate est telle qu'il est facile de décrire un sens de parcours (une orientation) du brin. Une molécule duplex d'ADN est obtenue à partir de deux brins que l'on dispose comme les montants d'une échelle. Les deux chaînes sucre-phosphate sont orientées en sens opposés. Les barreaux de l'échelle représentent une paire de bases; une base appartient à l'un des brins et l'autre base de la paire appartient à l'autre brin. Chaque base A est nécessairement face à une base T, tandis qu'une base C est en face d'une

base G. Ceci explique la règle de Chargaff qui avait observé expérimentalement que A et T se trouvent toujours en même proportion pour des molécules d'ADN provenant de cellules de la même espèce animale. Il en va de même pour C et G. Ce face-à-face est la règle de complémentarité. Elle implique que la connaissance du mot en les lettres A, C, G, T qu'on lit en parcourant un des brins détermine le mot qu'on lit en parcourant l'autre brin. Les liaisons entre deux bases qui se font face ne sont pas des liaisons covalentes mais des liaisons hydrogène, qui sont beaucoup plus faibles. L'énergie nécessaire pour rompre une liaison hydrogène est environ un dixième de l'énergie nécessaire pour rompre une liaison covalente.

Dans l'espace, la molécule est plongée de la façon suivante. Imaginons un cylindre ayant pour base un cercle de diamètre égal à 20A (un Angström A vaut 10^{-10} m). Sur ce cylindre traçons une hélice tournant à main droite, de pente égale à $34/20\pi$. Traçons une seconde hélice de même pente, dont les points se trouvent à distance 11A et 23A des points de l'autre hélice sur une même génératrice. Ces deux hélices représentent en gros les chaînes sucre-phosphate. Les paires de bases se trouvent à l'intérieur du cylindre. Lorsqu'une hélice fait un tour le long du cylindre, 10,5 paires de bases se sont attachées au brin correspondant.

Dans une deuxième note publiée dans Nature quelques mois après la première, Crick et Watson décrivent comment leur modèle explique la transmission des caractères génétiques lors de la mitose. Les deux brins se séparent au niveau des liaisons hydrogène entre paires de bases. Le fait qu'il s'agisse de liaisons faibles est très important. Un processus (encore à élucider à l'époque) permet sur chaque brin séparé de reconstituer le brin complémentaire. Et voilà comment l'on obtient deux molécules filles identiques à leur mère! L'ensemble de l'opération est appelé la réplication de l'ADN.

§2. QUE REPRESENTE LA DOUBLE HÉLICE?

On admet aujourd'hui que le modèle proposé par Crick et Watson (la double hélice) est un modèle local, valable pour un "petit" morceau d'ADN. Il faut aussi admettre que ce modèle est une sorte de moyenne, car le pas de l'hélice varie en fonction des paires de bases qui sont

attachées. De plus, le milieu dans lequel la molécule est plongée agit sur sa forme en fonction de la température, du degré de salinité, etc.

Ceci étant admis, qu'en est-il pour de longs morceaux d'ADN? Pendant les quelques années qui ont suivi la parution de la note à Nature, cette question ne fut pas vraiment d'actualité. On pensait que l'ADN se trouvait à l'état de petits segments dans la cellule et que, très probablement, ces segments étaient connectés entre eux par des protéines pour former des segments beaucoup plus longs. Certaines expériences de l'époque "confirmaient" d'ailleurs ce point de vue. On y voyait une foule de petits morceaux d'ADN. On sut plus tard que les procédés utilisés pour purifier l'ADN, trop brutaux, étaient responsables de la fragmentation de la molécule.

§3. LES AUTORADIOGRAPHES DE JOHN CAIRNS.

En 1963, le biologiste J. Cairns exécuta l'expérience suivante. Il marqua radioactivement (par du tritium) l'ADN d'Escherichia Coli, la bactérie universellement utilisée en biologie moléculaire. Il purifia doucement l'ADN ainsi obtenu. Il couvrit ensuite la préparation d'un film sensible aux rayons β . Après avoir laissé agir pendant deux mois, il observa au microscope les traces laissées par la radioactivité sur le film. Ces traces se nomment des autoradiographes. Le résultat fut plein d'enseignements. En effet:

- (i) Il n'y avait rien qui laissait supposer la présence de molécules de liaison entre petits segments d'ADN.
- (ii) Les composantes connexes étaient très longues. En fait l'unique chromosome d'E. Coli était visiblement constitué d'une unique molécule d'ADN, dont Cairns estima la longueur à environ $1'100\mu$. (1μ est un micron et vaut 10^{-6} m). Ceci est d'autant plus frappant que le plus grand diamètre d'E. Coli est de 2μ . La longueur estimée permet facilement de voir, en fonction des paramètres qui décrivent la double hélice, que le mot inscrit dans l'ADN d'E. Coli contient environ quatre millions de lettres. Mais le plus beau est à venir:
- (iii) La molécule d'ADN d'E. Coli n'est pas un segment mais un cercle (topologiquement).

Remarque. L'ADN d'E. Coli est à 2 brins. Par conséquent, on peut penser à cette molécule comme étant un fibré sur le cercle, dont les

fibres seraient les segments constitués par les paires de bases. Mathématiquement, il y a deux possibilités: le fibré est trivial ou homéomorphe au ruban de Möbius. Les questions d'orientation sur le bord dont nous avons parlé précédemment impliquent que le ruban de Möbius est impossible pour des raisons chimiques.

A la suite de Cairns, d'autres expériences furent faites qui montrèrent que beaucoup de virus (dont l'ADN est de taille nettement plus petite que celui d'une bactérie) ont aussi un ADN qui est circulaire à 2 brins.

L'expérience de Cairns eut plusieurs conséquences. Dans cet article je ne parlerai que d'une seule qui est la suivante. Supposons que l'on nous donne une molécule d'ADN circulaire à deux brins. Par un procédé quelconque, rompons les liaisons hydrogène entre les paires de bases. Ceci peut être obtenu, par exemple, en chauffant doucement la préparation. Nous obtenons deux molécules d'ADN circulaire à un brin. Mais leur coefficient d'enlacement est très grand. En première approximation, il est environ égal au dixième du nombre de paires de bases. Je reviendrai sur ce coefficient d'enlacement plusieurs fois dans la suite de ce texte. Pour l'ADN d'E. Coli il est donc d'environ 400'000. Mais, pour que la mitose ait lieu, il faut que les deux molécules filles se séparent. Mathématiquement, il faut qu'à la fin du processus on puisse mettre chaque fille dans une boule de R^3 et que les deux boules soient disjointes. Par conséquent le coefficient d'enlacement doit devenir nul. Comment diable la molécule mère et le milieu environnant s'y prennent-ils pour rendre ce coefficient égal à zéro? Ceci se nomme le problème topologique de la réplication. En résumé, une conséquence des expériences de Cairns est que le modèle local de Crick et Watson se heurte maintenant à une obstruction de nature globale.

§4. LES EXPÉRIENCES DE JEROME VINOGRAD.

A partir de 1963, une équipe de Caltec sous la direction de J. Vinograd a entrepris une série d'expériences qui aboutirent à la découverte du phénomène appelé "supercoiling" en anglais. Je le nommerai "contorsion". En voici un résumé.

Vinograd a constitué une colonie de virus polyoma. Il s'agit d'un virus cancérigène de la souris. Son ADN est circulaire à deux brins,

de longueur égale à environ 2μ . Ensuite, Vinograd a disposé sa population (notée I) dans une solution et placé cette dernière dans une ultracentrifugeuse (accélération de l'ordre de $50'000g$). Il a mesuré la vitesse de sédimentation s_{I} de cette population. D'autre part, il a soumis une seconde population du virus polyoma à l'action d'un enzyme appelé DNAase. Des expériences délicates ont abouti à la conclusion que cet enzyme rompt une liaison sucre-phosphate dans l'un des deux brins de l'ADN. L'ADN est encore circulaire à deux brins, mais l'un des brins est rompu en un seul endroit. On dit qu'il y a une encoche ("a nick" en anglais). La population ainsi obtenue, notée II , fut également soumise à l'ultracentrifugation. Le résultat fut que sa vitesse de sédimentation s_{II} est nettement inférieure à s_{I} . Plus précisément, le résultat obtenu fut que le rapport $(s_{\text{I}} - s_{\text{II}}) / s_{\text{I}}$ est égal à $1/5$.

Des photos obtenues par microscopie électronique ont révélé ceci. Les molécules de la population I ont une apparence très contorsionnée. La photo donne l'image par projection sur un plan de la molécule dans l'espace. Cette image a plusieurs points doubles pour les molécules de la population I tandis qu'elle n'en a en général aucun pour les molécules de la population II .

Ce fut un défi considérable pour J. Vinograd et ses collaborateurs de comprendre ce qui se passe. A cet effet, ils ébauchèrent une formule qui est capitale pour la compréhension de la structure spatiale de l'ADN circulaire à deux brins. Le prochain paragraphe est consacré à cette formule. Je reviendrai plus loin sur les expériences de Vinograd.

§5. LA FORMULE $Lk = Tw + Wr$.

Soit S^1 le cercle unité en l'origine de R^2 . Un ruban est un produit $S^1 \times J$ où $J = [-\epsilon, +\epsilon]$ pour un certain $\epsilon > 0$. L'âme du ruban est le cercle $S^1 \times \{0\}$. Un ruban R dans R^3 est l'image d'un ruban par un plongement différentiable ou PL (PL veut dire "presque linéaire" ou "linéaire par morceaux"). On oriente les deux composantes du bord d'un ruban dans R^3 parallèlement. Ceci entraîne que le coefficient d'enlacement des deux composantes est bien défini. (En effet, si l'on renverse le sens de parcours des deux composantes, le coefficient d'enlacement ne change pas). On notera ce coefficient $Lk(R)$. Il est facile de voir que la classe d'isotopie d'un ruban dans R^3 est déterminée par la classe d'isotopie du noeud constitué par l'âme et par $Lk(R) \in \mathbb{Z}$.

Soit maintenant v un vecteur de longueur 1 attaché en l'origine de R^3 . Soit E_v le plan orthogonal à v passant par l'origine de R^3 . Soit p_v la projection orthogonale de R^3 sur E_v . Si P et Q sont deux points distincts de R^3 se projetant sur le même point de E_v on dira par définition que celui qui a la plus grande composante en v via l'autre projection orthogonale est le plus haut.

Soit R un ruban dans R^3 . Soient bR_1 et bR_2 les deux composantes du bord de R . On considère R comme fixe, i.e. on s'interdit de le bouger par une isotopie. On dit que p_v est une bonne projection pour le ruban si localement en un point d'intersection de $p_v(bR_1)$ avec $p_v(bR_2)$ on voit deux courbes lisses à tangentes distinctes. Via une isotopie locale, on peut transporter cette situation de façon à ce que l'une des courbes soit envoyée sur l'axe des X au voisinage de l'origine de E_v et que l'autre courbe soit envoyée sur l'axe des Y . Ceci a un sens. En effet, on oriente R^3 une fois pour toutes. Comme la normale à E_v est orientée, le plan E_v est orienté. On choisit un système de coordonnées (X, Y) sur E_v qui représente cette orientation. Ceci étant précisé, on peut en plus demander que les orientations des brins soient conservées par l'isotopie. On attribue alors au point d'intersection le nombre $+1$ si le brin correspondant à l'axe des X passe au dessus du brin correspondant à l'axe des Y . On attribue le nombre -1 si c'est l'axe des Y qui correspond au brin supérieur.

Il est bien connu que le coefficient d'enlacement $Lk(R)$ est égal à la somme $\frac{1}{2} \sum s(B)$ où $s(B) = +1$ ou -1 est défini comme ci-dessus. La somme porte sur tous les points d'intersection de $p_v(bR_1)$ avec $p_v(bR_2)$.

Pour un tel point B , il y a un point $B_1 \in bR_1$ et un point $B_2 \in bR_2$ tels que $p_v(B_1) = p_v(B_2) = B$. Faisons tendre ϵ vers 0. On dira que B est un point double proche si B_1 et B_2 tendent vers un même point de l'âme du ruban. On dira que B est un point double éloigné sinon. On a:

$$Lk(R) = \frac{1}{2} \sum_v s(B) = \frac{1}{2} \sum'_v s(B) + \frac{1}{2} \sum''_v s(B)$$
 où la somme \sum' porte sur les points doubles proches tandis que la somme \sum'' porte sur les points doubles éloignés. Les indices v indiquent que tout ceci dépend de la projection p_v .

Faisons maintenant varier le vecteur v sur la sphère unité S^2 de R^3 . Si le ruban R est suffisamment différentiable (ou s'il est PL),

l'ensemble des vecteurs v tels que p_v est une bonne projection forme un ouvert dense. C'est un résultat classique qui remonte au moins à J. Alexander et K. Reidemeister.

Si v est un point de S^2 pour lequel p_v n'est pas une bonne projection du ruban R , on définit Σ_v , Σ'_v et Σ''_v comme étant nuls. On obtient ainsi trois fonctions sur la sphère S^2 . On peut démontrer que ces trois fonctions sont intégrables si le ruban est suffisamment différentiable (ou s'il est PL).

Comme Σ_v est égal à $Lk(R)$ pour n'importe quelle bonne projection, on a l'égalité:

$$Lk(R) = \frac{1}{4\pi} \int \frac{1}{2} \Sigma_v \, dv.$$

Notons $Tw(R)$ l'intégrale $\frac{1}{4\pi} \int \frac{1}{2} \Sigma'_v \, dv$ et $Wr(R)$ l'intégrale $\frac{1}{4\pi} \int \frac{1}{2} \Sigma''_v \, dv$.

ATTENTION: $Lk(R)$ est un nombre entier. Pour un v donné, la décomposition $Lk(R) = \frac{1}{2} \Sigma'_v + \frac{1}{2} \Sigma''_v$ est une décomposition en somme de deux nombres entiers. Mais cette décomposition varie en fonction de v , de sorte que $Tw(R)$ et $Wr(R)$ sont en général des nombres réels non entiers.

Le nombre $Tw(R)$ s'appelle en anglais le twist ou le twisting number. En français, on l'appelle la torsion du ruban. Ce terme est justifié par la proposition suivante.

PROPOSITION. $Tw(R)$ est égal à la torsion totale du ruban.

La torsion totale est définie de la façon suivante. Soit C l'âme du ruban R (orientée provisoirement). En chaque point de C , considérons la normale à C , de longueur 1, qui est tangente au ruban dans la direction de ξ positif. Notons n cette normale, t la tangente (de longueur 1) et b la binormale choisie telle que t , n et b forment un repère donnant l'orientation de R^3 usuelle. La torsion totale du ruban est alors l'intégrale

$$\frac{1}{2\pi} \int_C b \cdot dn \quad \text{où } b \cdot dn \text{ désigne le produit scalaire.}$$

Le cas de $Wr(R)$ est plus intéressant. Le terme anglais introduit par Brock Fuller est "writhe" ou "writhing number". Je propose de l'appeler en français la contorsion. On peut comprendre ce nombre de la façon suivante. Pour un v générique dans S^2 , les points doubles éloignés sont obtenus lorsque l'image par p_v de l'âme du ruban se recoupe en un point double ordinaire. Au voisinage d'un tel point dans le plan E_v on trouve deux points doubles éloignés de même signe. Ce signe est le même que le signe du point double de l'âme, obtenu pour une orientation quelconque de l'âme. (Car un changement de l'orientation de l'âme change l'orientation des deux brins locaux). Ceci suggère la définition suivante.

Définition. Soit K un noeud lisse (ou PL) dans R^3 . Soit p_v une bonne projection de K sur E_v . Le nombre de Tait $T_v(K)$ de la bonne projection p_v est le nombre entier $\sum_v s(Q)$ la somme étant prise sur tous les points doubles Q de la projection. Il est clair que l'on a la proposition suivante.

PROPOSITION. On a l'égalité $Wr(R) = \frac{1}{4\pi} \int T_v(\text{âme du ruban}) dv$.

Remarque. Le ruban étant donné, la décomposition $Lk(R) = Tw(R) + Wr(R)$ a des propriétés d'invariance intéressantes. Le coefficient d'enlacement est un entier invariant par isotopie du plongement du ruban. La torsion totale est un nombre réel invariant par isométrie du ruban. La contorsion est un nombre réel invariant par isométrie de l'âme du ruban.

§6. COMMENTAIRE HISTORIQUE.

J. Vinograd et ses collaborateurs pressentirent vers 1965 qu'il existe une décomposition de $Lk(R)$ en somme de deux nombres réels tenant compte de la géométrie de la situation et non pas seulement de la topologie. La contorsion était présentée comme de la "superhélicité". On exprimait par là qu'en plus de l'hélice locale de Crick et Watson, la forme globale est une sorte d'hélice d'ordre supérieur. Aujourd'hui encore la contorsion est la plupart du temps appelée "supercoiling" par les biologistes de langue anglaise et non pas writhe. A la demande de Vinograd, le mathématicien Brock Fuller écrivit au début des années '70 plusieurs articles pour expliciter mathématiquement la formule empirique que Vinograd employait. La contorsion y

était définie comme la différence $Lk - Tw$ où Tw apparaît comme l'intégrale de la torsion locale du ruban. L'auteur affirmait ensuite que Wr ainsi défini est la moyenne des nombres de Tait.

Sauf erreur, F. Crick contacta alors le géomètre différentiel William Pohl qui lui dit que la formule $Lk = Tw + Wr$ est une conséquence facile de résultats contenus dans la thèse de son élève James White. Pohl et White ont une définition différente de la contorsion qui est due aux travaux du mathématicien roumain Georges Calugareanu (vers 1960). En fait, comme ils le disent eux-mêmes, c'est pour simplifier et généraliser les résultats de Calugareanu que Pohl et White entreprirent leurs recherches. Voici en quelques mots de quoi il s'agit.

En 1833, Carl Friedrich Gauss donna dans une brève note (non publiée à l'époque) une formule intégrale pour exprimer le coefficient d'enlacement de deux courbes fermées disjointes dans R^3 . Concrètement, si f et g sont deux applications différentiables de S^1 dans R^3 , d'images disjointes, on construit une application $h: S^1 \times S^1 \rightarrow S^2$ par la formule $h(s,t) = \frac{f(s)-g(t)}{\|f(s)-g(t)\|}$. Le degré de cette application, exprimé sous la forme d'une intégrale, est la formule de Gauss. L'interprétation en termes d'électromagnétisme que l'on peut donner à la formule était connue de Gauss. C'est même elle qui l'y a conduit. Vers 1870, le physicien écossais Peter Tait dans ses travaux de pionnier sur la théorie des noeuds, se posa la question de la signification de l'intégrale de Gauss lorsque $f=g$ (et que $f=g$ est injective). Autrement dit, l'intégrale de Gauss pour un noeud (plutôt que pour un entrelacs à deux composantes) a-t-elle un sens et, si oui, quel est ce sens? En principe, l'intégrale n'est pas définie sur la diagonale de tore. L'interprétation physique de l'intégrale l'amena à pressentir que cette intégrale pour un noeud est liée à ce que j'ai appelé ci-dessus le nombre de Tait d'une bonne projection. C'est pour rappeler l'importance de l'oeuvre de Tait sur les noeuds que je propose ce nom. Tait ne va pas jusqu'à prendre la moyenne sur toutes les bonnes projections.

Calugareanu ne mentionne pas Tait, mais il se pose la même question que lui au sujet de l'intégrale de Gauss pour un noeud. Après coup, c'est-à-dire lorsqu'on dispose de la formule $Lk = Tw + Wr$ pour un ruban,

on s'aperçoit que le contenu des articles de Calugareanu est le suivant. Etant donné une courbe fermée simple et lisse dans R^3 , il n'existe pas nécessairement un ruban canonique dont elle serait l'âme. Mais il en existe si la courbure ne s'annule jamais. Dans ces conditions, n'importe quel vecteur du trièdre de Frenet fait l'affaire pour construire un ruban. Dans la littérature, le vecteur tangent est souvent utilisé; en ce qui me concerne, je préfère le vecteur normal. Calugareanu commence par démontrer que l'intégrale de Gauss existe lorsque $f=g$ (et que $f=g$ est injective) à condition de supposer que le plongement est de classe C^3 . Ceci fonctionne sans hypothèse sur la courbure. Ensuite, il établit et prouve la formule $Lk = Tw + Wr$ lorsque la courbe est sans point d'inflexion (ce qui est équivalent à dire que la courbure ne s'annule jamais). Le nombre entier Lk s'appelle alors le "self-linking" du noeud. Dans la formule de Calugareanu, Tw est donc la torsion totale de la courbe (au sens des formules de Serret et Frenet), tandis que Wr est défini bien sûr à l'aide de la formule de Gauss.

Le point de vue de Pohl et White, avec une très élégante démonstration, est présenté dans l'article de W. Pohl à Math. Intell. cité dans la bibliographie commentée du présent article.

Il est intéressant de noter que la formule $Lk = Tw + Wr$ est utilisée dans d'autres domaines. Pour son rôle en dynamique des fluides, voir l'article de H. Moffatt et R. Ricca.

A ma connaissance, la présentation de la formule donnée ci-dessus à l'aide des points doubles proches et éloignés ne se trouve pas dans la littérature. Mais c'est une conséquence assez immédiate de ce qui est publié sur le sujet.

§7. RETOUR AUX EXPÉRIENCES DE VINOGRAD.

La différence des deux vitesses de sédimentation fut expliquée par l'équipe de Vinograd de la façon suivante. Prenons deux populations composées de molécules ayant même poids moléculaire. Si les vitesses de sédimentation sont différentes, c'est que les molécules les plus rapides ont une forme dans l'espace plus "compacte", plus "condensée". Pour les molécules de l'ADN de polyoma utilisées par Vinograd, c'est la contorsion (ou plutôt sa valeur absolue) qui

rendrait compte du degré de "compacité", tout au moins en première approximation. C'est un problème toujours actuel que de mieux comprendre ce phénomène.

Une surprise des expériences de Vinograd est que les molécules de la population I viennent directement "de la nature" sans altération. En d'autres termes, l'ADN de polyoma est naturellement contorsionné. Vinograd entreprit alors de quantifier W_r , ou tout au moins de déterminer son signe. Après des expériences très subtiles, les résultats furent que W_r est négatif. En valeur absolue, W_r est égal à quelques pourcents de T_w . Comme T_w est positif, il en résulte que L_k est encore positif, mais un peu inférieur à ce qu'il serait si la molécule était "relaxée". On entend par là une molécule à laquelle on aurait tout d'abord fait une encoche à l'aide de DNAase, que l'on aurait ensuite laissé se décontorsionner, et à laquelle on aurait finalement rétabli la liaison manquante au niveau de l'encoche. Les biologistes savent faire tout cela.

Des expériences ultérieures montrèrent que le phénomène de "supercoiling" de l'ADN in vivo est très général. D'où la question: la contorsion sert-elle à quelque-chose? Il est légitime de penser que la contorsion favorise la présence de régions sur la molécule où la courbure est forte. Ceci est important. En effet, l'activité de l'ADN est traduite par des protéines qui viennent s'y fixer temporairement pour former ce que l'on appelle un complexe moléculaire. L'existence d'un tel complexe dépend fortement des formes (euclidiennes) des deux molécules en leurs sites actifs. Plusieurs liaisons ADN-protéines ont souvent besoin pour exister de la courbure locale de l'ADN. Certaines de ces liaisons n'ont pas lieu si l'on relaxe l'ADN. Autrement dit, on pense maintenant que la contorsion de l'ADN est une condition nécessaire pour qu'il puisse jouer son rôle.

Il est certain que la contorsion aide aussi la cellule (d'E. Coli par exemple) à résoudre le problème de stockage pour son ADN. Souvenons que le plus grand diamètre de la cellule est 2μ tandis que son chromosome circulaire a une longueur de $1'100\mu$.

§8. L'ÉLECTROPHORESE.

L'électrophorèse est un procédé utilisé depuis longtemps par les

biochimistes. Voici en gros en quoi il consiste. Supposons que l'on a un échantillon de molécules chargées électriquement (disons négativement). On se propose de les séparer en fonction de leur poids moléculaire. L'idée est de préparer un gel (il en existe de plusieurs sortes avec des pores de tailles différentes) et de faire voyager les molécules de l'échantillon le long d'une colonne de gel en plaçant convenablement des électrodes aux extrémités de la colonne. Bien sûr, c'est un procédé délicat, qui demande de la finesse de la part des expérimentateurs. Lorsque tout est bien calibré, on trouvera, "après un certain temps", les molécules contenues dans des bandes transverses; les molécules les plus petites se trouvant le plus près de l'arrivée.

En 1974, le biologiste Walter Keller eut l'idée de séparer par un procédé analogue des molécules d'ADN circulaire à deux brins ayant toutes même poids moléculaire (i.e. même longueur) mais une contorsion différente. (L'ADN étant un acide est chargé négativement en milieu aqueux). Ceci demande encore plus de finesse que dans le cas "classique". Le gel utilisé est un gel d'agarose; les colonnes ont environ 20cm de long et le voyage dure environ 10 heures. Mais ça marche! Et l'observation des bandes par microscope électronique montre que les molécules qui voyagent le plus vite sont celles ayant la plus grande contorsion. C'est un phénomène analogue à ce qui se passe par ultracentrifugation. Toutefois, il y a des différences dans les circonstances. L'ultracentrifugation est une diffusion à travers un milieu visqueux tandis que l'électrophorèse est une diffusion à travers une sorte de labyrinthe. Dans les deux cas il serait intéressant de comprendre ce qui se passe et pourquoi la contorsion est un bon indicateur de la vitesse.

Pour illustrer ce qui précède, supposons que nous ayons à disposition des échantillons des virus polyoma de type I et II. Supposons que l'on confectionne aussi une population III en cassant les deux brins. Ces derniers ADN sont donc homéomorphes à des segments. Par électrophorèse, un échantillon contenant les trois populations est séparé en trois bandes. L'échantillon III (qui n'a pas de contorsion à proprement parler) est le plus proche de l'arrivée. Ensuite on trouve la population I, puis la population II qui est la plus lente des trois.

Mais les biologistes vont plus loin dans la finesse. (Actuellement, ils utilisent pour les expériences des plasmides circulaires à deux brins plutôt que des virus. Les plasmides sont des petites molécules d'ADN, de longueur de quelques microns, disconnectées des chromosomes et dont le rôle génétique est encore mal connu.)

Pour un échantillon de plasmides préparés de la même façon on considère que, dans des situations pas trop extrêmes, les molécules ont toutes même torsion locale (donnée essentiellement par le modèle de Crick et Watson) et donc même T_w . Si l'on admet cette hypothèse, les différences dans la contorsion sont donc des différences dans le coefficient d'enlacement. Par diverses manipulations, les biologistes savent modifier le coefficient d'enlacement pour le faire varier d'une seule unité. (Dans leurs notations $\Delta Lk = 1$). Par électrophorèse, on observe que deux populations de plasmides dont le coefficient d'enlacement diffère de 1 se trouvent dans deux bandes différentes. Dit autrement, la séparation électrophorétique permet de mesurer une différence d'une unité dans les coefficients d'enlacement. C'est plus difficile de calculer exactement le coefficient. Il vaut environ 600 pour un plasmide de longueur 2μ .

§9. TOPOISOMERASES ET RECOMBINASES.

On connaît actuellement deux familles d'enzymes qui modifient la topologie des molécules d'ADN dans l'espace. La famille des topoisomérases comprend présentement deux types d'enzymes appelés Topo I et Topo II. L'enzyme Topo I agit sur l'ADN à deux brins lorsqu'il y a une encoche. Il peut alors faire passer successivement les deux brins d'une autre partie de l'ADN à travers l'encoche. Lorsque l'encoche a été refermée, il se trouve que la différence ΔLk est d'une unité. L'enzyme Topo II fait passer, sans utiliser d'encoche, une paire de brins locaux à travers une autre paire de brins locaux. Pour une telle opération manifestement $\Delta Lk = 2$. C'est ce qui a permis de distinguer l'action de ces deux enzymes par électrophorèse.

On sait maintenant que ces deux enzymes agissent lors de la réplication. En effet, si l'on empêche l'un des deux enzymes d'agir (en utilisant un produit chimique convenable) la réplication est

perturbée. Beaucoup de médicaments utilisés dans le traitement chimique des cancers agissent précisément en bloquant l'action de l'un ou l'autre des Topos. Ceci a pour effet d'arrêter la prolifération des cellules cancéreuses, dont l'une des caractéristiques est de se reproduire plus rapidement que les cellules saines. Ces médicaments ont pour effet négatif d'arrêter aussi la reproduction des cellules saines. D'où la chute des cheveux.

A l'heure actuelle, on dispose de certains scénarios qui décrivent dans les grandes lignes comment Topo I et Topo II collaborent pour résoudre le problème topologique de la réplication.

Les recombinaisons agissent en modifiant d'une unité le nombre de composantes connexes de l'ADN. Par exemple les intégrases permettent de faire se connecter deux molécules d'ADN disjointes pour former une seule molécule. Un virus qui infecte l'ADN d'une cellule a un gène qui code pour la fabrication d'une protéine (par la cellule infectée!) qui lui permettra de faire entrer son ADN comme une partie de l'ADN de la cellule. En général, ceci a des conséquences néfastes pour la cellule en question.

Il est certain que les biologistes ont un grand intérêt à comprendre comment topoisomérases et recombinaisons agissent. (Par exemple pour pouvoir intervenir chimiquement). Un schéma typique d'expériences est le suivant. On constitue un échantillon de plasmides "identiques" dont l'anneau est non noué. On ajoute à la préparation des exemplaires de l'enzyme dont on veut étudier l'action. Dans certains cas, un même plasmide peut être le siège de plusieurs actions successives des enzymes. Il se trouve que, pour certains enzymes, plus l'enzyme agit de fois, plus il complique en un certain sens la topologie plasmidique.

On sépare ensuite la préparation par électrophorèse. On examine le contenu de chaque bande au microscope électronique. Une innovation mise au point par Andrzej Stasiak consiste à faire recouvrir les molécules d'ADN qui se trouvent dans une bande par une protéine nommée RecA. Ceci a pour effet de multiplier le diamètre de la double hélice par un facteur cinq. Au microscope, on peut alors distinguer en un point double quel est le segment qui est dessus et quel est celui qui est dessous. Autrement dit, on peut reconnaître le type de

noeud représenté par l'âme du ruban.

Noter que toutes ces opérations (habillage par RecA, microscopie, etc) modifient la géométrie mais pas la topologie des molécules d'ADN.

Par exemple, en 1985 une expérience dirigée par Nicolas Cozzarelli à Berkeley a porté sur l'action de Topo I. Après électrophorèse, on a trouvé dans une même bande des noeuds ayant même crossing number (nombre de croisements) selon la table de Reidemeister. Une bande est d'autant plus rapide que le nombre de croisements des molécules qu'elle contient est grand. La différence des nombres de croisements des individus de deux bandes successives est égal à un. Cette expérience est très spectaculaire.

Le topologue De Witt Sumners a proposé de modéliser l'action de certains enzymes comme étant un "tangle move" à la John Conway. Cela veut dire que l'on se représente l'enzyme comme étant une boule dans R^3 . Cette boule intersecte l'ADN en deux segments disjoints. L'action de l'enzyme consisterait à remplacer le plongement des deux segments dans la boule par un autre plongement, puis à recoller avec ce qui se trouve à l'extérieur et qui n'a pas changé. Pour les topoisomérases, on recollerait sans changer les voisins tandis qu'il y aurait des changements dans les voisins pour les recombinaisons. En utilisant des résultats "non triviaux" de théorie des noeuds comme le "Cyclic Surgery Theorem" de Culler-Gordon-Luecke-Shalen, Sumners a réussi à démontrer à partir de l'analyse des bandes d'électrophorèse que certains recombinaisons ne peuvent agir qu'en effectuant un "tangle move" d'un certain type, qu'il est capable de calculer explicitement. C'est l'application la plus frappante que je connaisse de la théorie des noeuds à l'ADN.

REMERCIEMENTS.

Je tiens à remercier mes amis Jacques Dubochet et Andrzej Stasiak du LAU de l'Université de Lausanne des nombreuses heures qu'ils ont consacrées à mon initiation aux beautés de la biologie moléculaire. Mais je suis seul responsable des erreurs que peut contenir cet article.

BIBLIOGRAPHIE COMMENTÉE.

Comme texte de base pour s'initier à la biologie moléculaire, le livre de J. Watson est insurpassable. Il s'agit de "Molecular Biology of the Gene" édité par W. A. Benjamin. Il a été plusieurs fois remis à jour et réédité.

Comme texte de référence, je conseille le volumineux et très complet livre de B. Alberts - D. Bray - J. Lewis - M. Raff - K. Roberts - J. Watson intitulé "Molecular Biology of the Cell", édité Garland Publ. Comp.

Pour avoir une idée de la complexité des expériences en biologie, consulter le petit livre "Classical Experiments in Modern Biology" de M. Green, édité par W.H. Freeman and Comp.

J. Vinograd et son équipe ont publié de nombreux articles sur leurs travaux. Le premier où la contorsion est mentionnée est J. Vinograd - J. Lebowitz - R. Radloff - R. Watson - P. Laipis: "The twisted circular form of polyoma viral DNA" Proc. Nat. Acad. Sci. USA 49(1965) p.1104-1111.

Pour la formule $Lk = Tw + Wr$, voir l'article de W. Pohl: "DNA and differential geometry" Math. Intell. 3(1980) p.20-27. Il contient les références aux principaux articles du sujet.

Pour l'utilisation de la formule $Lk = Tw + Wr$ en dynamique des fluides, voir H. Moffatt - R. Ricca: "Helicity and the Calugareanu invariant" Proc. Roy. Soc. London A439(1992) p.411-429.

Pour l'expérience de N. Cozzarelli sur Topo I, voir l'article original F. Dean - A. Stasiak - T. Koller - N. Cozzarelli: "Duplex DNA knots produced by Escherichia Coli Topoisomerase I" Jour. Biol. Chem. 260(1985) p.4975-4983.

Les travaux de De Witt Sumners sont très bien décrits dans son survey article "Untangling DNA" Math. Intell. 12(1990) p.71-80.

Des expériences récentes concernant l'action des Topos sur l'ADN font l'objet de l'article S. Shaw - J. Wang: "DNA knot formation in aqueous solutions" Jour. Knot Theory and Appl. 3(1994) p.287-298. On y trouvera une liste de références à des résultats expérimentaux récents.

Section de mathématiques; 2-4 rue du Lièvre; case postale 240;
CH 1211 Genève 24; Suisse.

