

Les mécanismes moléculaires de la synthèse des protéines cellulaires

Autor(en): **Scherrer, Klaus**

Objektyp: **Article**

Zeitschrift: **Verhandlungen der Schweizerischen Naturforschenden Gesellschaft. Wissenschaftlicher und administrativer Teil = Actes de la Société Helvétique des Sciences Naturelles. Partie scientifique et administrative = Atti della Società Elvetica di Scienze Naturali**

Band (Jahr): **146 (1966)**

PDF erstellt am: **13.09.2024**

Persistenter Link: <https://doi.org/10.5169/seals-90662>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Les mécanismes moléculaires de la synthèse des protéines cellulaires

KLAUS SCHERRER

(Institut de Biologie Physico-Chimique, Paris-5^e)

Dans cet exposé nous tenterons de retracer les connaissances actuelles concernant la synthèse des protéines telles qu'elles se dégagent des progrès réalisés au cours des dernières années dans le domaine de la Biologie moléculaire. Du fait même que cette science nouvellement née trouve l'origine de son développement dans les domaines les plus divers de la science fondamentale (je veux nommer la Chimie, la Physique, la Génétique et la Biologie classique), il nous faut emprunter, afin de décrire et comprendre les phénomènes qui sont à la base même de la vie, des termes, modèles et abstractions caractéristiques de toutes ces sciences. En effet, les racines de la véritable révolution qui a ébranlé les concepts jusqu'à présent affirmés en biologie prennent naissance dans l'ensemble de ces sciences classiques.

En un laps de temps trop court pour permettre d'autre développement qu'un rapide survol, j'aimerais évoquer les démarches qui ont conduit à élucider le fonctionnement chimique des gènes. Après un bref aperçu concernant le *milieu cellulaire* dans lequel se déroule l'ensemble des mécanismes biochimiques et biophysiques conduisant à la genèse des macromolécules, nous aborderons ce que l'on est accoutumé d'appeler le *dogme central de la biologie moléculaire* en retraçant la démonstration du conditionnement de la séquence des acides aminés dans les protéines par la séquence des bases dans le DNA. Nous parlerons ensuite de l'élément-clé du transfert de l'information génétique du chromosome au peptide naissant : c'est *l'adaptateur* qui permet la transformation du système informatif à base «4» – le code génétique –, en un autre à base «20» – le code des 20 acides aminés constituant les chaînes polypeptidiques et les protéines fonctionnelles. Ensuite, nous évoquerons la prédiction théorique et la découverte expérimentale du *RNA messenger* avant de passer à une description des mécanismes impliqués dans le *transfert de l'information génétique du DNA vers les protéines*, lesquels sont à la base de l'expression phénotypique du génome. Enfin, après avoir rappelé les hypothèses concernant *la régulation* des mécanismes décrits, nous tâcherons de dégager les conclusions générales.

Le milieu cellulaire

Panta rhei! Tout s'écoule! Telle est la meilleure description du milieu dans lequel se déroulent les mécanismes chimiques et physiques dont nous allons parler par la suite. La cellule n'est ni un vase alchimique dans lequel

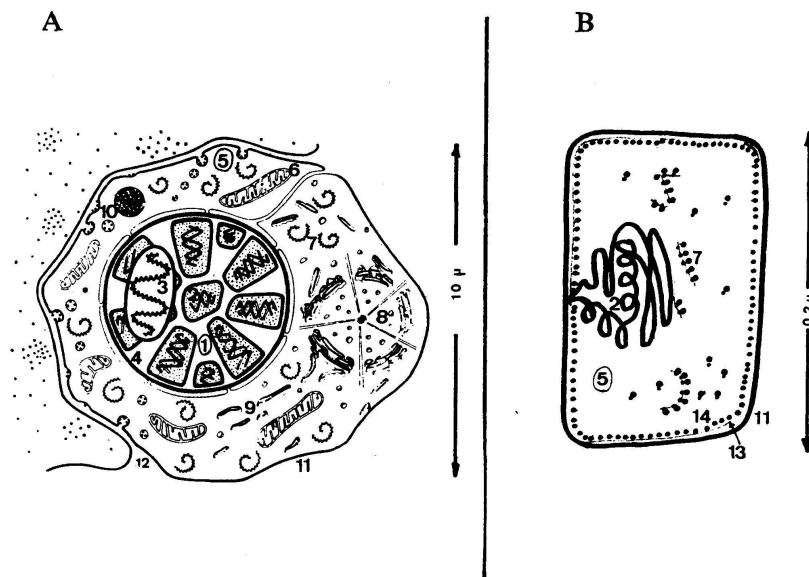


Figure 1

La cellule animale (A) et bactérienne (B)

1 : noyau; 2 : chromosomes comprenant le DNA en double hélice (et les histones); 3 : nucléole; 4 : membrane nucléaire; 5 : cytoplasme; 6 : mitochondries; 7 : polysomes constitués de ribosomes; 8 : appareil de Golgi; 9 : reticulum endoplasmique; 10 : vacuoles formées par le processus de pinocytose; 11 : paroi cellulaire; 12 : espace intercellulaire; 13 : membrane cellulaire; 14 : membrane cellulaire intérieure renfermant des enzymes, etc.

les réactions chimiques sont déterminées rigoureusement par la simple présence de constituants définis et leurs concentrations respectives. Elle n'est pas non plus organisée et conditionnée par une architecture rigide telle qu'un instrument physique ou électronique. L'organisation structurale de la cellule est bien au contraire dans un état de dynamisme incessant. Toutefois, au sein de ces structures en renouvellement permanent qui participent elles-mêmes aux réactions chimiques et physiques existe un ordre rigoureux qui assure l'aboutissement des mécanismes de synthèse et de dégradation qui sont à la base de la vie cellulaire. Il ne faut pas perdre de vue ce «*protodynamisme*» si l'on veut comprendre les réactions décrites plus loin. Ce phénomène est unique et caractéristique de la cellule vivante, dont le schéma simplifié de la *figure 1* ne donne qu'une représentation fugace, sorte de prise de vue instantanée. Le fonctionnement des structures représentées et des macromolécules qui les constituent est conditionné à la fois par leurs réactivités chimiques, leur localisation ou leur transfert, à l'intérieur de la cellule, d'un compartiment subcellulaire à un autre.

Le dogme central

La séquence des 4 bases nucléiques dans l'acide desoxyribonucléique (DNA) – qui constitue le matériel génétique – détermine la séquence des 20 acides aminés, laquelle représente la structure primaire de la chaîne pep-

tidique. Ainsi peut être formulé le dogme central de la biologie moléculaire. En termes plus classiques, nous pouvons dire que le gène dirige l'expression phénotypique d'une cellule donnée. Avant de passer aux mécanismes qui assurent cette expression, nous devons définir les unités matérielles qui conditionnent le phénotype.

Pendant longtemps, la génétique classique a principalement considéré les aspects morphologiques de l'expression phénotypique des gènes, telle que la forme d'une aile d'insecte, leur couleur, etc. Mais le postulat de toute la biochimie moderne est que ces différences spécifiques ne font qu'exprimer en dernière analyse des équipements cellulaires distincts en protéines, lesquelles, vous le savez, sont les éléments de soutien et les catalyseurs biologiques essentiels. Il faut donc admettre que le DNA doit contrôler la spécificité des protéines afin d'assurer la pérennité des caractères héréditaires. Des résultats génétiques aussi bien que nos connaissances actuelles concernant les structures protéiques ont amené à prouver ce postulat. Comme nous allons le voir, en effet, les produits de l'expression des gènes ne sont pas tant les protéines, transmises avec leurs caractéristiques distinctives spécifiques, que les sous-unités polypeptidiques qui les constituent. La phase essentielle de l'expression phénotypique du gène repose donc dans le passage de l'information génétique aboutissant à la structure primaire de la chaîne peptidique.

Il est peut-être utile, avant d'explicitier ce dogme central, de rappeler certaines notions chimiques fondamentales concernant les acides nu-

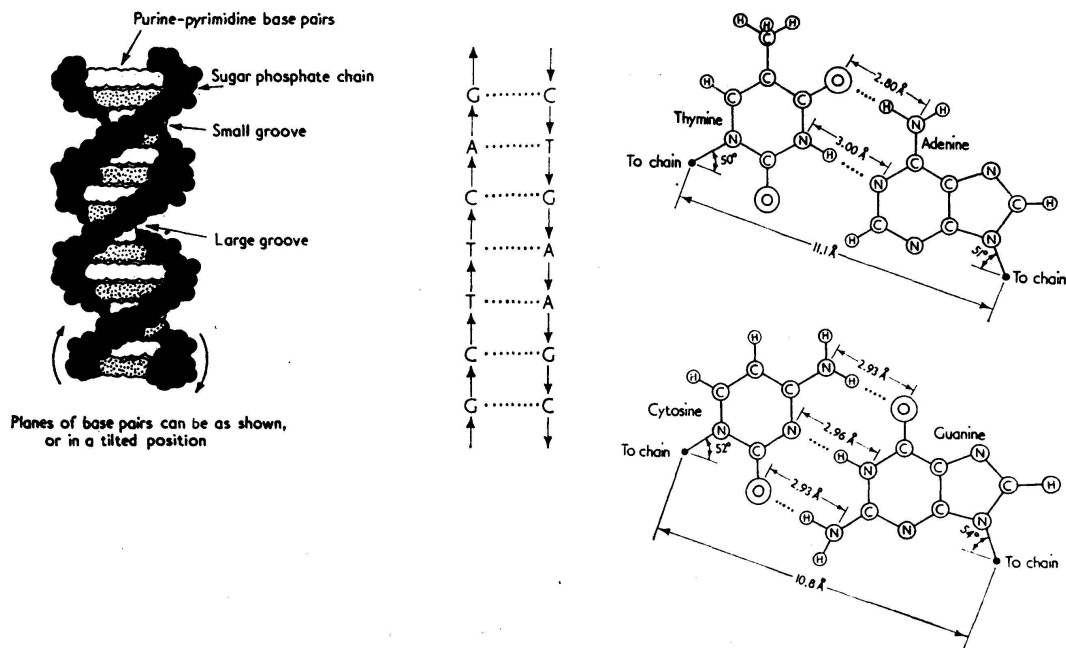


Figure 2

La structure du DNA

A. Appariement des bases; B. structure hélicoïdale et séquence des paires de bases complémentaires (d'après MICHELSON, «The chemistry of Nucleosides», Acad. Press, 1963).

cléiques et les protéines. Les premiers sont constitués, vous le savez, par l'enchaînement de résidus de sucres en C₅ reliés par des groupes phosphates en liaisons « diester ». Ce squelette est constitué dans le cas du DNA par le desoxyribose, chez le RNA par le ribose. A cette chaînette linéaire sont fixées latéralement, par le groupe réducteur du sucre, les bases nucléiques telles que l'adénine (A), la thymine (remplacée par l'uracile [U] dans le RNA), la guanine (G) et la cytosine (C). La règle de l'appariement des bases découverte par WATSON et CRICK postule alors que deux de ces chaînettes peuvent s'apparier base par base de telle sorte qu'une adénine soit toujours placée en regard d'une thymine ou d'une uracile et une guanine en regard de la cytosine. Les paires de bases ainsi formées peuvent se partager des électrons grâce aux liaisons dites « hydrogène ». Les exigences stériques de ces liaisons conduisent les deux brins à s'entrelacer en double hélice, comme le démontre la *figure 2*. Cette structure permet à la fois le maintien de l'information génétique et sa transmission de génération en génération par les mécanismes de replication sur lesquels nous ne voulons pas nous étendre ici.

La double hélice du DNA forme souvent des structures circulaires continues. Ainsi la « molécule génétique » d'un virus forme un anneau comprenant environ 6000 paires de bases, celle d'une bactérie une structure circulaire d'une circonférence d'un millimètre renfermant environ 10 millions de paires de bases nucléiques. Le RNA ne se trouve qu'exceptionnellement en forme de double hélice. Son rôle principal est lié au transfert de l'information intracellulaire et ne sert que rarement comme dans les virus à RNA à la perpétuation des caractères génétiques.

Passons aux molécules protéiques qui vous sont certainement assez familières. Rappelons que les 20 acides aminés se condensent grâce à leurs fonctions COOH et NH² pour former des macromolécules linéaires. Leur *séquence* est appelée *structure primaire*, celle-ci conditionnant toutes les transformations ultérieures que ces molécules sont susceptibles de subir. Par *structure secondaire*, nous entendons la formation de régions hélicoïdales et ouvertes sous l'influence de facteurs stériques et de transferts des charges dont sont porteuses les chaînes latérales. La *structure tertiaire* représente la structure tridimensionnelle de ces chaînes; elle est le résultat d'un équilibre entre charges, régions hélicoïdales et emplacement des groupements sulfhydriles qui se prêtent à la formation de ponts -S-S-. Toutes les propriétés biologiques des protéines résultent de cette structure tertiaire qui peut ou bien telle quelle être douée d'activité biologique, ou bien n'acquérir celle-ci qu'après formation d'une *structure quaternaire* résultant de la condensation de plusieurs chaînes. La gamme des protéines s'étend en conséquence d'un poids moléculaire de quelque mille Daltons (dans le cas de polypeptides à chaîne unique) jusqu'à plusieurs millions pour les protéines hautement associées. Malgré la complexité de ces assemblages, on est en droit de dire que toutes les propriétés des protéines sont inscrites dans la séquence des acides aminés de la structure primaire.

Nous voici amenés à nous demander quelle fut la preuve du déterminisme reliant la structure protéique primaire à la séquence des bases dans le DNA. Cette preuve a pu être apportée grâce à l'analyse d'un phénomène physiologique que l'on appelle les *maladies moléculaires*. Il existe en effet de nombreuses variétés d'anémies transmissibles selon les lois classiques de l'hérédité. Or, il a été démontré notamment par INGRAM, ITANO et PAULING que la cause de ces maladies réside dans un défaut de l'hémoglobine. Chaque type de maladie est associé à une mutation de l'hémoglobine correspondante entraînant le remplacement dans la chaîne peptidique d'un acide aminé par un autre. Le DNA constituant le matériel génétique, il faut donc conclure que la cause primaire de ces remplacements d'acides aminés réside dans une modification de la séquence des bases nucléiques.

Une mutation génétique consiste donc en une altération chimique dans une paire de nucléotides dans le DNA. Dès lors, si l'on détermine par analyse génétique la distribution relative des mutations à l'intérieur d'un cistron, et si l'on procède par ailleurs à l'analyse biochimique des changements correspondants dans la séquence polypeptidique, on peut établir si les deux types de modification sont colinéaires ou non. Des chercheurs américains du groupe de YANOFKY ont ainsi réussi à démontrer dans le cas de la tryptophane synthétase que les distances relatives de ces mutations dans le DNA sont exactement les mêmes que les distances relatives des acides aminés modifiés dans la chaîne polypeptidique.

A l'échelle moléculaire, cette *colinéarité* est parfaite si nous admettons que la séquence de trois bases nucléiques détermine la sélection d'un acide

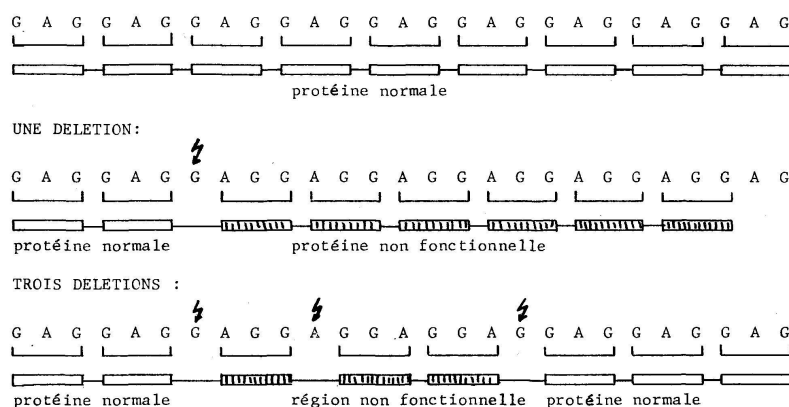


Figure 3

La démonstration du code à triplet

La délétion (ou l'insertion) d'une base à l'intérieur d'un cistron contrôlant la formation d'une chaîne polypeptidique rend celle-ci inactive parce que, à partir de l'emplacement de la base disparue, le cadre de lecture est changé. Le codon GAG est devenu AGG. Mais deux délétions (ou insertions) supplémentaires restituent la lecture correcte et redonnent de l'activité à la protéine correspondante, à condition que la région restant erronée ne soit pas trop large et ne comprenne pas d'endroits vitaux.

aminé. La démonstration que l'information génétique est transmise par des codons à trois unités, ou si l'on préfère que le code génétique est un code constitué d'éléments à triplets, a pu être acquise il y a trois ans par des preuves génétiques et a été confirmée récemment par les brillantes expériences de KHORANA. Je me bornerai pourtant ici à évoquer dans la *figure 3* les expériences antérieures de BRENNER et CRICK, lesquelles ont apporté la première suggestion d'après laquelle la traduction de l'information génétique a lieu par groupe de triplets successifs.

L'adaptateur

Si nous considérons comme acquise la démonstration que les gènes ont pour fonction d'ordonner l'agencement séquentiel des acides aminés dans la chaîne polypeptidique, il faut considérer le problème de la traduction d'un code à 4 lettres – celui du DNA – en un autre à 20 lettres, celui des polypeptides. La comparaison couramment utilisée d'un tel processus avec le déchiffrement d'un texte écrit en un langage particulier ne me paraît pas très précise. Mieux vaut rapprocher ce problème de celui de l'expression d'une information par des systèmes numériques différents. De même que les grands ordinateurs transforment le système décimal que nous utilisons couramment en un système binaire pour manipuler plus facilement l'information, ainsi la nature utilise un système quaternaire pour l'emmagasinage de l'information génétique, tandis qu'elle applique un système à base de vingt pour son expression phénotypique. Il nous faut donc retrouver dans la cellule l'ordinateur et son programme pour comprendre cette transformation.

Du fait qu'il n'existe aucune relation chimique évidente entre la structure du DNA et les acides aminés permettant une interaction directe au niveau d'une séquence prédéterminée, il fallait postuler un *adaptateur* qui

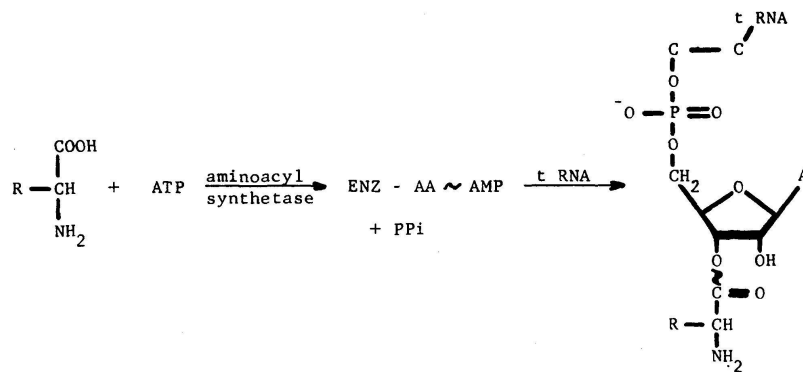


Figure 4

L'activation de l'acide aminé

L'acide aminé forme avec l'enzyme aminoacyl synthétase un complexe grâce à l'énergie fournie par l'ATP. De ce complexe riche en énergie, l'acide aminé est transféré à l'hydroxyle en 3' de l'adénosine terminale du tRNA. L'adénosine suivi de deux cytosines forme toujours la terminaison acceptrice des tRNA.

L'adaptateur porte une séquence de trois nucléotides que nous appelons *anticodon*, mise en évidence dans la *figure 5*. Dans chaque tRNA ce triplet est complémentaire – suivant la règle de WATSON et CRICK élucidée antérieurement – du codon correspondant à l'acide aminé véhiculé par ce tRNA. Ainsi que nous allons l'envisager plus loin, les tRNA acylés sont choisis consécutivement grâce à leur anticodon, par appariement aux bases du codon complémentaire avant que la liaison peptidique ne se forme.

Le RNA messenger

A ce stade des connaissances, il nous faut relier les différents faits pour expliquer d'une manière simple le contrôle exercé par les gènes sur la synthèse des protéines spécifiques. On savait depuis un certain temps que la synthèse des protéines est effectuée au contact des ribosomes, les particules cytoplasmiques où figurent en association protéines et RNA. Ces particules nucléoprotéiques sont composées de deux sous-unités dont chacune renferme une molécule de RNA caractéristique. Il serait facile d'imaginer que le DNA transmette l'information correspondante d'un gène à un RNA ribosomal. Le ribosome ainsi formé serait transféré hors du chromosome et dirigerait en contact avec les adaptateurs activés la formation de la protéine spécifique.

Deux raisons essentielles pouvaient jeter un doute sur le rôle probable du RNA des ribosomes comme véhicule direct de l'information génétique :

1° Les RNA ribosomiques affichent toujours, à quelques variations près, la même composition globale en nucléotides, et la longueur de chaîne des molécules qui les constituent n'est toujours que de deux types distincts. Une telle monotonie ne devait pas manquer de surprendre puisque les structures génétiques elles-mêmes varient en fonction du spectre des protéines à la fois par leur composition et leur taille.

2° Il s'est avéré que les RNA ribosomiques sont *métaboliquement stables*. Or, le véhicule d'information hypothétique doit être – du moins chez les bactéries – une substance de très faible durée de vie.

Ce dernier postulat découle d'un ensemble d'observations biologiques dont je ne veux citer que *deux*. L'une porte sur le fait assez général que les bactéries sont capables de *s'adapter rapidement* à un changement de leur milieu nutritif en changeant leur équipement enzymatique. Si le ribosome portait l'information, la période d'adaptation devrait être caractérisée par une forte néosynthèse du RNA ribosomique et par une dégradation des ribosomes superflus. Or, l'expérience montre le contraire, pendant cette phase, il n'y a ni synthèse ni dégradation de ribosomes. La *deuxième* démonstration qui est plus rigoureuse, découle de la cinétique de synthèse d'un enzyme après l'apparition du gène correspondant dans le cytoplasme des bactéries. Pour cette expérience, PARDEE, JACOB et MONOD se sont servis du croisement bactérien. Dans certaines conditions, des bactéries dites « mâles » peuvent injecter leur chromosome dans les cellules d'une souche femelle déficiente en un enzyme donné. On constate alors que cet enzyme est synthétisé d'emblée au taux maximum sitôt que

le gène fonctionnel est injecté. En d'autres termes, si un RNA est impliqué en tant qu'intermédiaire dans l'expression des gènes, il ne peut s'agir du RNA des ribosomes, lequel, étant stable, eût conduit à une cinétique d'expression autocatalytique.

JACOB et MONOD ont donc supposé que le vecteur d'information, ou messenger, devait être un RNA instable et ont proposé que chaque gène élabore un RNA messenger spécifique, copie conforme du gène, ce messenger s'associant aux ribosomes, pour être traduit en protéines. Grâce aux travaux de GROS et WATSON, cette hypothèse a rapidement trouvé sa confirmation expérimentale. Une fraction ne constituant que 5% du RNA bactérien fut identifiée en tirant parti précisément de sa rapidité de synthèse et de son instabilité métabolique.

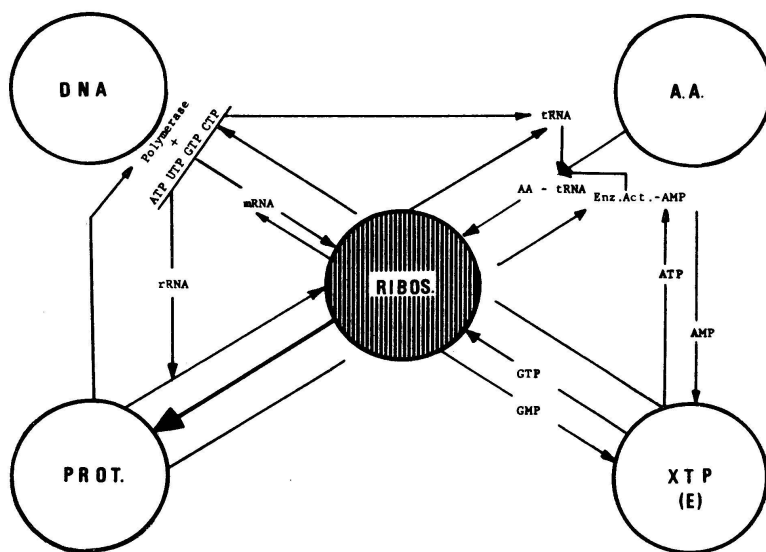


Figure 6

Représentation diagrammatique de la synthèse protéique

Le ribosome est central dans l'ensemble des mécanismes liés à la synthèse protéique : là, les protéines se forment. Il est entouré par les 4 facteurs essentiels : le DNA contrôle la réaction par l'intermédiaire du mRNA. Les ribosides triphosphates (XTP) fournissent l'énergie. Les acides aminés (AA) servent de substrat. Le pool des protéines produites fournit les enzymes nécessaires à la synthèse et les éléments structuraux des ribosomes eux-mêmes.

Nous venons ainsi de décrire les mécanismes généraux et les structures impliqués dans la formation des protéines. La figure 6 résume schématiquement l'ensemble de ces données. Il devient dorénavant plus facile de comprendre les expériences qui ont permis de briser le secret du code génétique. Les biochimistes ont appris à faire fonctionner les ribosomes dans leur tube à essai. S'inspirant des expériences de TISSIÈRES, véritable pionnier dans ce domaine, les chercheurs NIRENBERG et MATTHAI ont eu

l'idée d'introduire des messagers artificiels sous la forme de polynucléotides constitués exclusivement par un enchaînement de résidus d'uracile (poly U) dans un mélange de ribosomes et d'extraits solubles de bactéries. En analysant le produit formé à partir d'acides aminés divers, ils ont identifié un polypeptide unique, formé exclusivement par l'acide aminé phénylalanine. Ainsi ils ont pu conclure que le codon UUU correspondait à la phénylalanine dans le langage génétique. En procédant de façon identique avec une multiplicité de messagers artificiels, la correspondance entre chaque acide aminé connu et son codon spécifique a pu être établie. Un raffinement de cette technique a permis de déterminer la séquence des bases à l'intérieur du triplet et ainsi d'établir la carte totale du code génétique donné dans la *figure 7*.

2e 1er	U	C	A	G	3e
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Ochre	?	A
	Leu	Ser	Amber	Tryp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Glu N	Arg	A
	Leu	Pro	Glu N	Arg	G
A	I leu	Thr	Asp N	Ser	U
	I leu	Thr	Asp N	Ser	C
	I leu	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Figure 7
Le code génétique

Des expériences de même nature ont été conduites en parallèle avec des messagers naturels. Le grand et ancien rêve des biochimistes de pouvoir former des protéines *in vitro* a donc trouvé sa réalité. Il faut pourtant admettre que, de tous les résultats publiés jusqu'à présent à ce sujet, seule la synthèse *in vitro* d'une protéine virale sous la direction du RNA extrait du virus paraît être incontestable.

Le transfert de l'information génétique

Nous distinguons deux phases dans le transfert de l'information aboutissant à la synthèse des protéines: la *transcription* du DNA en RNA messenger et la *traduction* du RNA en polypeptide.

Obéissant à la même loi d'appariement des bases dont nous avons vu les caractéristiques lors de la replication du DNA, le RNA messenger est synthétisé par un enzyme, la RNA polymérase. Cet enzyme forme des liaisons ester entre le RNA amorcé et le nucléoside triphosphate sélectionnée par la base opposée dans le DNA. C'est ainsi que s'imprime dans le mRNA la séquence des bases du DNA. La structure hélicoïdale du DNA n'est dissociée que sur une région relativement courte à l'intérieur de laquelle le RNA néoformé prend la place de la chaîne opposée. Comme le montre le schéma de la *figure 8*, cette polymérisation admet une polarité déterminée: le nucléoside -5'-triphosphate entrant en jeu se condense avec le groupe 3'-hydroxyle de la chaîne RNA amorcée. La polarité du RNA naissant mène donc de la terminaison 5'-hydroxyle vers la terminaison 3'. Les RNA «naissants» peuvent comprendre un nombre de nucléotides qui varie de 80 pour les plus petits jusqu'à plus de 20000 pour les RNA polycistroniques d'origine animale, ainsi que nous avons pu le mettre en évidence récemment.

Au cours de la transcription, une seule des deux chaînes du DNA est copiée. Cette règle établie par de nombreuses expériences souffre pourtant quelques exceptions démontrant que les deux chaînes du DNA peuvent, en des endroits différents, être porteuses d'information génétique.

En vous rappelant la continuité physique circulaire du DNA, il faut postuler que la polymérase doit reconnaître des signaux inscrits dans le DNA délimitant le gène à transcrire. Ce problème reste encore à éclaircir.

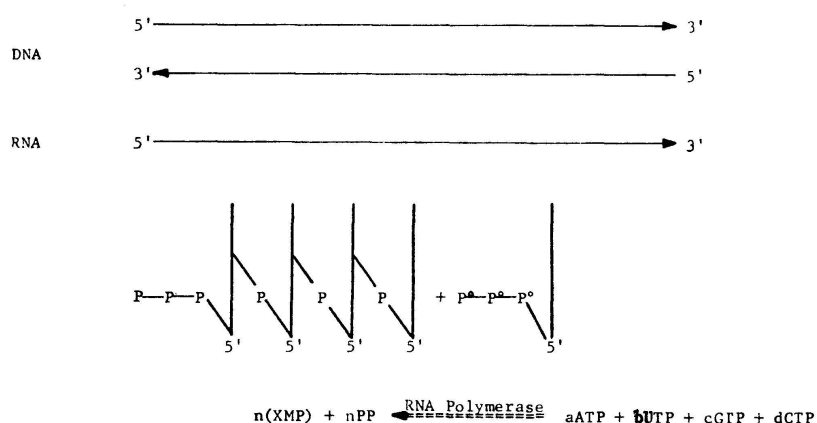


Figure 8

La direction de la transcription du DNA

La synthèse du RNA messenger s'effectue suivant une direction définie : des ribosides-5'-triphosphates sont ajoutés aux hydroxyle-3' à l'extrémité du polynucléotide en formation. La synthèse de la chaîne suit donc une polarité qui mène du 5' au 3'.

Cependant, l'analyse génétique met en évidence des régions «promotrices» supposées être les sites d'attachement de la RNA polymérase. On a d'ailleurs démontré par microscopie électronique que l'enzyme s'attache bien à un nombre de points limités sur le DNA et non pas de façon arbitraire.

Il est d'un intérêt capital de déterminer les unités de transcription pour une meilleure compréhension de la régulation de la synthèse protéique. Un fait peut être considéré comme acquis: la transcription concerne presque toujours plusieurs cistrons en bloc et presque jamais des unités d'information génétique correspondant à des chaînes polypeptidiques isolées. Nous verrons plus loin que ce fait est une des bases physiques de l'existence des opérons, théorie énoncée par JACOB et MONOD. Cette théorie spécifie que l'activation de ces opérons aboutit à la synthèse synchrone de plusieurs enzymes à la fois. Les RNA messagers correspondants ont pu être mis en évidence expérimentalement. Le seul type de RNA connu actuellement pour être synthétisé comme produit d'un cistron unique est le tRNA. On admet le même principe de transcription unitaire pour le RNA ribosomal des bactéries. Par contre, nous avons pu mettre en évidence que le RNA ribosomal de cellules animales est synthétisé par bloc de plusieurs unités, et ceci est vrai également pour une fraction considérable des RNA messagers. Les RNA ribosomiques animaux synthétisés sous la forme d'un produit multicistronique sont métabolisés pour donner naissance à des molécules plus petites et fonctionnelles. Or, nous avons des raisons de penser que les messagers subissent la même altération, bien que ceci ne puisse encore être considéré comme un fait vraiment acquis.

Les modifications chimiques subies par les RNA après leur synthèse sont de nature variée: les bases nucléiques peuvent être méthyles dans leur groupe aminé et des résidus sulfhydriles ajoutés. L'uracile peut être transformé en pseudouracile *in situ*. Toutes ces altérations chimiques influencent la structure secondaire des molécules de RNA. Or, la structure secondaire et, par voie de conséquence, la structure tertiaire d'un RNA joue très probablement un rôle important dans l'activité biologique de ce type de molécule.

Les RNA ayant acquis ainsi leur structure et leur conformation finales assument leurs fonctions respectives: le tRNA est condensé avec l'acide aminé et le rRNA s'unit aux protéines ribosomales pour former la particule sur laquelle la synthèse protéique a lieu. Quant au RNA messenger, il n'existe guère à l'état libre. Dès sa formation, il est capté par les ribosomes qui, dans les bactéries, ont accès direct au chromosome. Chez les cellules animales, une particule spéciale s'attache au messenger nouvellement formé dans le noyau et le transfère au cytoplasme où la synthèse protéique a lieu.

Ainsi nous sommes enfin arrivés à une des phases centrales de notre discussion: la formation de la chaîne polypeptidique elle-même. Le RNA messenger est attaché à la sous-unité 30 S du ribosome, le tRNA portant la chaîne polypeptidique naissante grâce à sa liaison avec le dernier acide aminé incorporé dans la chaîne est fixé sur la grande sous-unité 50 S. Un

nouvel adaptateur porteur de l'acide aminé prochain est «choisi» par les trois bases du codon suivant sur le mRNA, lesquelles s'apparient à l'anticodon. Le tRNA porteur de la chaîne peptidique et le nouveau venu étant tous deux appariés avec le mRNA et fixés côte à côte sur le ribosome dans des sites bien définis, il y a formation d'une nouvelle liaison peptidique, comme cela est illustré dans la *figure 9*. Une fois l'acide aminé incorporé, le nouveau tRNA portant maintenant la chaîne peptidique se déplace vers le site occupé au début par le tRNA précédent qui a été libéré par la réaction. D'après des données publiées ces dernières semaines, il faut penser que deux facteurs enzymatiques participent à la formation de la liaison peptidique. L'énergie pour cette réaction est entièrement fournie par la liaison riche en énergie d'acyle-tRNA. Par contre, on est amené à penser que le déplacement du messenger par rapport au ribosome dépend de l'énergie fournie par le GTP grâce à la participation du deuxième de ces facteurs.

Le messenger peut servir à la synthèse de plusieurs chaînes peptidiques à la fois: en effet, les ribosomes se suivent à un rythme rapide sur le mRNA. Cette situation donne naissance à la structure unique appelée *polysome* (poly-ribosome) représentée sur la *figure 10*. Le ribosome se déplaçant de codon à codon sur le mRNA, la chaîne polypeptidique peut ainsi s'allonger.

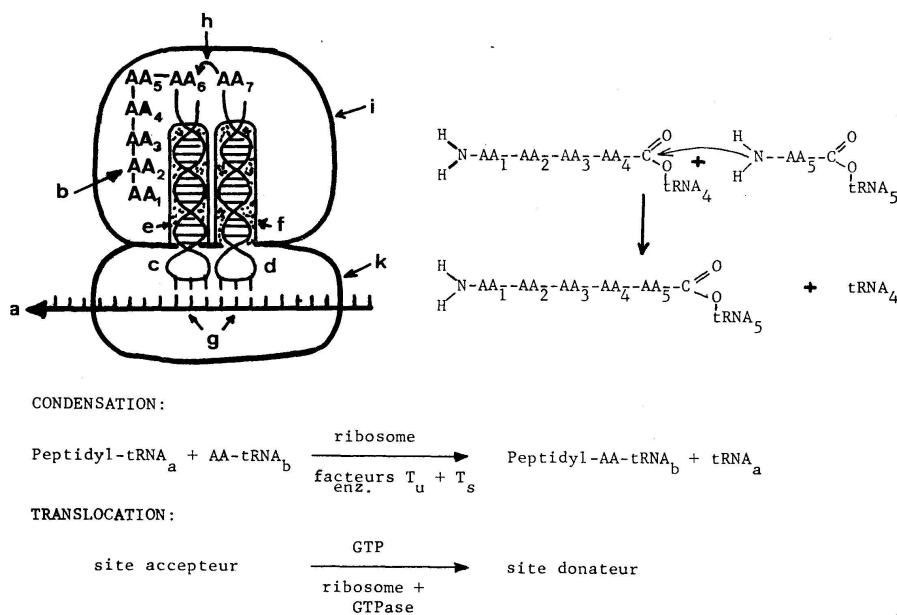


Figure 9

La formation du polypeptide grâce à la traduction du RNA messenger sur le polysome

a: RNA messenger; b: chaîne polypeptidique en formation fixée au tRNA₆; c: tRNA₆; d: tRNA₇; e: site accepteur, f: site donateur du tRNA; g: codons du mRNA appariés aux anticodons des tRNA; h: site de condensation des acides aminés; i: particule 50 S du ribosome; k: partie 30 S du ribosome.

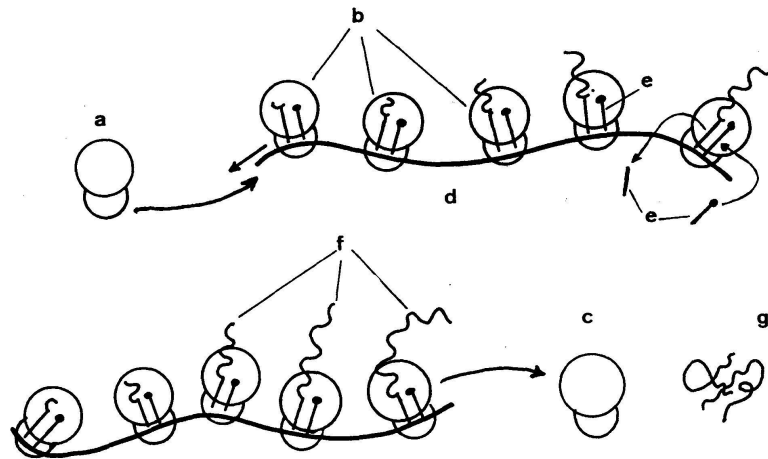


Figure 10

Les polysomes

a : ribosome libre en voie d'attachement; b : ribosomes engagés dans le polysome; c : ribosome détaché du polysome après passage; d : mRNA; e : tRNA; f : chaînes polypeptidiques en formation; g : chaînes polypeptidiques terminées après détachement revêtant leurs structures tertiaire et quaternaire.

Le message est lu dans le polysome dans la même direction que celle de sa synthèse. La polarité 5'-hydroxyle vers 3'-hydroxyle du mRNA est donc parallèle à celle du polypeptide qui commence par un groupe α aminé libre et se termine par une fonction acide.

Reste à envisager le problème de l'initiation et de la terminaison de la synthèse d'un polypeptide. Bien que l'on puisse envisager dans le cas du polysome formant de l'hémoglobine que le début et la terminaison du message conditionnent la longueur de la chaîne polypeptidique formée, le problème se pose d'une façon plus compliquée si nous envisageons des messages polycistroniques tels qu'un RNA viral. Or, il découle de résultats récents que des codons spéciaux existent pour le commencement aussi bien que pour la terminaison de la lecture. Le début paraît être signifié par le codon AUG reconnu par l'adaptateur de la formyle-méthionine. Pour les signaux de terminaison on discute de la participation des triplets UAA et UAG. Du fait que ces codons (appelés *ambre* et *ocre*) ne correspondent à aucun acide aminé activé par son adaptateur, la chaîne polypeptidique serait libérée instantanément du ribosome. Ces codons *non sens* peuvent être produits à partir de codons normaux par des mutations que l'on appelle *ocre* et *ambre*. A la suite de leur formation des chaînes polypeptidiques incomplètes apparaissent dans les cellules mutées.

Bien que la transcription du génome en RNA *polycistronique* puisse être considérée comme un fait établi, la traduction de ce message en une seule unité est encore sujet de discussion. Il faut aussi signaler que le RNA message fonctionnel dans le polysome peut être de stabilité variable. Chez les bactéries la plupart des mRNA ont une durée de vie qui ne dépasse pas trois minutes. Il existe toutefois aussi des messages bactériens plus

stables. La stabilité des messagers des cellules animales varie entre 3 heures et plus que 24 heures. Les raisons de cette variation de stabilité de molécules chimiquement identiques restent l'un des mystères à résoudre.

La dernière phase de la formation des protéines comprend la libération des chaînes polypeptidiques naissantes du ribosome. Ceci peut être un processus spontané dans certains cas, parfois une réaction provoquée par d'autres chaînes déjà en solution. La chaîne libérée prend spontanément sa structure tertiaire, conditionnée par le milieu ionique et de petites molécules, les *effecteurs* dits *allostériques*. La forme biologiquement active de la protéine est donc acquise. Pour des protéines de grande taille, elle s'accompagne d'une condensation de plusieurs chaînes polypeptidiques formant la structure quaternaire. Ainsi sont engendrées les protéines, éléments fondamentaux de l'expression phénotypique des gènes.

La régulation

La régulation cellulaire représente un problème d'une telle ampleur que nous ne pouvons pas évoquer dans leur totalité les faits, connaissances et hypothèses principales. Ce qui nous intéresse ici, c'est la régulation primaire de la synthèse protéique, celle concernant la formation de la structure primaire de la chaîne polypeptidique. Nous laisserons ainsi de côté les facteurs influençant l'élaboration des protéines dans leur conformation finale et leur activité biologique.

Il faut admettre que règne actuellement encore le plus grand mystère en ce qui concerne la régulation de la synthèse protéique. Le grand mérite de JACOB et MONOD a été de jeter pour la première fois de la lumière dans ces ténèbres. Il faut attirer l'attention sur le fait que l'acquisition la plus importante de leurs travaux est d'avoir élaboré là où régnait la plus grande confusion d'hypothèses un mécanisme moléculaire de la régulation de l'expression du génome dans le cas d'une cellule spécifique. Bien que les schémas, abstractions et définitions nouvellement introduites représentent une véritable révolution dans un des domaines les plus importants de la biologie, il importe de préciser cependant que leur démonstration repose sur l'emploi d'un matériel cellulaire particulier, la cellule bactérienne et plus spécifiquement celle d'*Escherichia coli*. Il importe donc de rechercher si ce schéma général peut être étendu au cas d'autres bactéries (ce qui paraît déjà assuré pour certaines d'entre elles) et à celui des organismes supérieurs. Le contenu informatif des systèmes biologiques varie de 10^3 codons chez certains virus à 10^9 dans le cas de l'homme (en passant par les bactéries avec 10^6). Il paraît peu vraisemblable que la nature puisse avoir recours à un seul et unique système de régulation pour gouverner l'activité du matériel héréditaire, quelles que soient ses dimensions et sa complexité. Il nous faudra donc dans un avenir immédiat essayer de tirer profit de l'impulsion donnée par le modèle de JACOB et MONOD pour élucider les systèmes de régulation chez les eucaryotes mêmes.

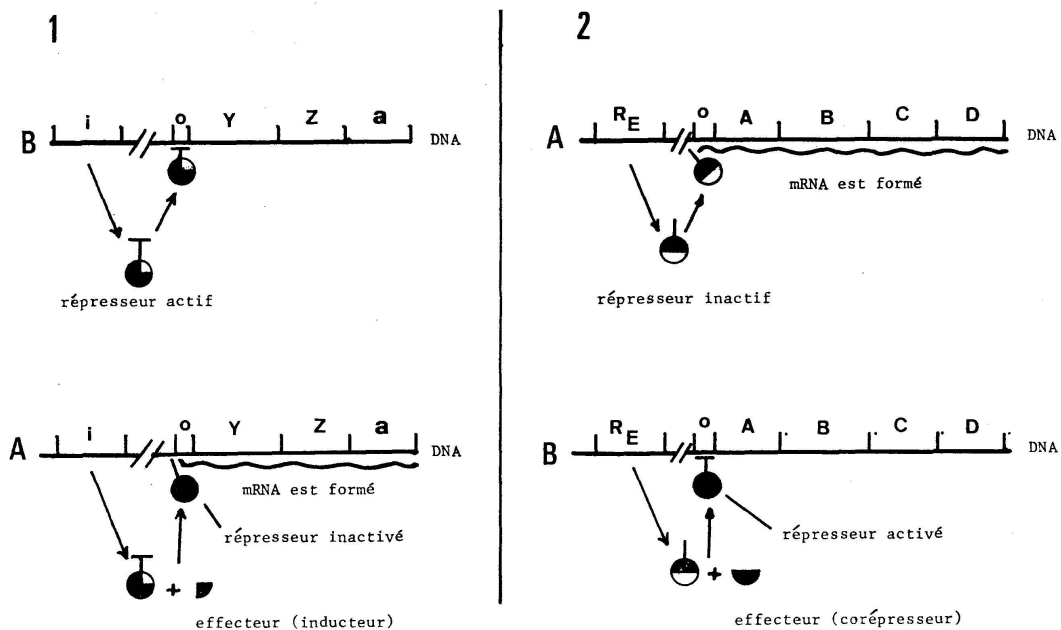


Figure 11

Le modèle de régulation d'après Jacob et Monod

1° le répresseur est actif en soi. 2° le répresseur est inactif en soi (aporépresseur).
 A: induction; B: répression; i, R_E gènes de régulation, A, B, C, D et y, z, a, gènes de structure.

Une représentation simplifiée de ce modèle est donnée dans la figure 11. Son idée centrale est la distinction entre *gènes de structure* définissant la structure primaire des protéines et *gènes de régulation*. Le fait qu'un gène de régulation puisse gouverner l'expression de l'information pour plusieurs chaînes protéiques – ou de plusieurs *cistrons*, en langage génétique – a conduit à la notion de *l'opéron*. Un opéron comprend donc un ensemble de cistrons gouvernés par le même gène de régulation. L'activité de ce dernier donne naissance à un *répresseur* consistant en une macromolécule (fort probablement en une protéine) qui peut être pourvue d'activité inhibitrice telle quelle ou après interaction avec une petite molécule – du type des cofacteurs bien connus en enzymologie – que l'on nomme *effecteur*. La présence de cet effecteur induit un changement structural – dit *allostérique* – du répresseur, ce qui peut activer ce dernier ou, dans le cas où il est déjà pourvu d'activité en soi, peut neutraliser celle-ci. En conséquence, il doit exister deux types d'effecteurs: les uns *induisant* et les autres *réprimant* l'activité d'un opéron. Le répresseur sous sa forme active se combine avec *l'opérateur* qui marque le début de l'opéron sur le DNA, ou sur son produit, le RNA messenger, empêchant ainsi son expression. La présence ou absence de ces effecteurs qui peuvent être des substrats ou produits des enzymes à synthétiser ou encore, éventuellement, des molécules régulatrices spécialisées telles que des hormones, peut donc gouverner l'expression de fractions bien définies de l'information génétique.

L'unité de régulation est donc l'opéron qui donne naissance à un messager unique. L'unité de traduction en protéines peut être ce *messenger polycistronique* lui-même ou les cistrons qui le constituent individuellement. L'unité d'expression phénotypique comme nous l'avons définie en termes de biologie moléculaire est donc le cistron lui-même. Il faut toutefois attirer l'attention sur le fait que le produit protéique de cette unité est rarement pourvu d'activité biologique en soi. Souvent il faut en effet l'activité de plusieurs cistrons pour que se forme une protéine pourvue d'activité biologique, grâce à l'association de chaînes polypeptidiques réunies en structure quaternaire.

En passant aux organismes supérieurs, la question de la régulation de la synthèse protéique devient extrêmement compliquée. Nous devons donc nous limiter à quelques remarques sommaires. Il découle de travaux parmi les plus récents que chez les eukaryotes la régulation de l'expression des gènes a lieu vraisemblablement à plusieurs niveaux et non pas au niveau de la transcription du DNA exclusivement. En effet, si une régulation paraît bien se manifester à ce niveau, comme ceci a été mis surtout en évidence par le système des chromosomes géants chez lesquels on peut observer avec un simple microscope comment l'action d'une hormone règle l'activité de segments précis du génome, d'autres expériences tendent à démontrer qu'un contrôle existe aussi après la formation du mRNA, ainsi qu'au niveau des polysomes. Le fait paraît déjà acquis qu'une fraction plus large du génome est en activité productrice de messagers que celle finalement traduite en protéines et exprimée phénotypiquement. Vous me permettez de citer ici un résultat de mon propre laboratoire qui tend à démontrer que dans la cellule sanguine produisant presque exclusivement de l'hémoglobine, environ 5% du génome est en activité. Ces 5% doivent porter l'information pour au moins 10^5 cistrons. Nous sommes donc amenés à penser qu'une « cascade d'événements régulateurs » élimine à des niveaux successifs l'excès de mRNA transcrit. Je rappelle à ce sujet le chemin parcouru par le RNA du chromosome aux polysomes et les nombreuses transformations qu'il subit, dont la dernière serait la traduction elle-même dans le polysome. Mais nous touchons ici aux limites de la pure spéculation. J'ai simplement essayé de vous donner brièvement une appréciation très superficielle de ce qui est acquis dans ce domaine et de ce qui nous reste à élucider.

Conclusion

J'ai essayé de décrire quelques aspects des mécanismes moléculaires de la synthèse des protéines dans les limites permises par l'immensité du sujet, mes faibles connaissances et la limitation du temps. En terminant, j'aimerais essayer d'attirer votre attention sur l'universalité qui est caractéristique de la plupart des réactions et des mécanismes dont nous venons de parler. En effet, sans se soucier désormais de la formidable complexité des espèces, la biologie moléculaire regarde surtout les processus communs au niveau des molécules. Elle constate alors que les molécules d'un

éléphant et d'une bactérie, celles d'un arbre et d'un homme ont, sinon certes la même architecture apparente, du moins les mêmes éléments formateurs. La base matérielle de l'existence de l'homme, le DNA renfermant l'information génétique, obéit aux mêmes lois chez lui ou chez la plante. Les macromolécules essentielles, charpentes de toutes les cellules, sont engendrées par des mécanismes remarquablement identiques à travers toute l'échelle des êtres vivants.

C'est ainsi que nous comprenons désormais le langage qu'autorise l'alphabet nucléique. Grâce à quoi nous commençons à comprendre comment un gène fonctionne. La prochaine étape sera précisément d'expliquer comment, à partir de ces mécanismes moléculaires communs, est possible cette diversité éblouissante de la nature. Le problème de la différenciation cellulaire est considérable. Mais attention! La biologie moléculaire n'a fait qu'un premier pas.