

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/279202392>

Ondes d'échelle et régulation épigénétique de la biosynthèse des protéines (1991)

Data · December 1991

DOI: 10.13140/RG.2.1.3917.4247

CITATIONS

0

READS

1,033

1 author:



[Joël Sternheimer](#)

Genodics

31 PUBLICATIONS 18 CITATIONS

SEE PROFILE

ONDES D'ECHELLE ET REGULATION EPIGENETIQUE DE LA BIOSYNTHESE DES PROTEINES

Joël Sternheimer

(1991)

Lorsqu'un individu ressent une émotion, l'ébranlement qui parcourt son être met en correspondance l'échelle, microscopique, de l'expression des gènes avec l'échelle, macroscopique, de l'individu lui-même. Cette correspondance présente les caractéristiques physiques d'une onde d'échelle, c'est-à-dire d'un phénomène vibratoire qui se propage dans l'échelle en ayant pour fonction de préserver l'invariance d'échelle à partir de la singularité introduite à une échelle donnée par un phénomène individuel, ici moléculaire. Cette onde constitue une généralisation directe de l'onde de Broglie, qui préserve l'invariance relativiste à partir de la singularité spatio-temporelle constituée par une particule individuelle. Si cette interprétation est correcte, elle entraîne une double conséquence: il doit exister une correspondance structurelle précise entre l'expression d'un gène et un phénomène sensible situé à notre échelle; et, l'onde pouvant parcourir l'échelle dans les deux sens, il doit être possible d'agir de façon spécifique, par l'intermédiaire des sens, sur l'expression des gènes.

Pour ce dernier point, les exemples sont en fait nombreux: regarder le ciel nous fait faire du pigment bleu, et éprouver la sensation colorée correspondante; un soleil trop vif nous fait faire des protéines "heat shock", etc... On connaît également des gènes associés à l'odorat et au goût. Mais y a-t-il un rapport précis, dont la présence serait la "signature" d'une onde d'échelle, entre le pigment bleu et la couleur du même nom? Cette vieille question du "pourquoi le bleu nous paraît-il bleu" est soeur jumelle de "pourquoi tel morceau de musique est-il gai et tel autre triste"; la fin même de la musique, selon Descartes (*Musicae Compendium*, 1618) étant d'"émouvoir en nous des passions diverses", en agissant, selon le traité chinois ancien *Li Ki*, sur le "tsing chen", ce qui signifie "substances vitales sécrétées par l'organisme" (on traduirait aujourd'hui par "synthèse des protéines"). Si l'émotion suscitée en nous par une musique est bien l'expression d'une onde d'échelle provenant de l'expression d'un gène et de la synthèse d'une protéine, il doit exister une correspondance précise entre cette musique et ce gène.

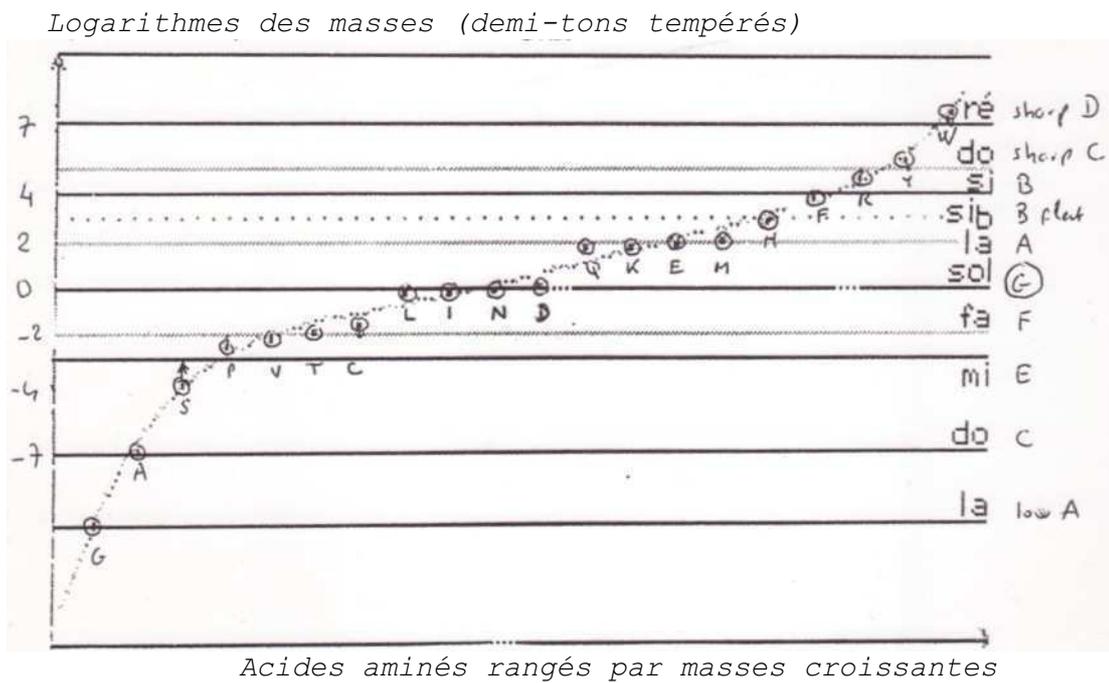
Le présent travail procède de l'observation qu'il en est bien ainsi, avec les conséquences que cela entraîne de pouvoir agir de façon précise et prédictive, par la musique, sur l'expression des gènes -- ainsi que, par les couleurs, sur la stabilité de la protéine obtenue. Nous montrerons plus précisément que les suites

d'ondes brogliennes associées à la synthèse des protéines constituent bien, une fois transposées en vibrations audibles, des mélodies à la "musicalité" d'autant plus remarquable que les protéines correspondantes sont typiquement plus sensibles à des facteurs environnementaux, et au "style" même généralement bien caractéristique de musiques accompagnant une situation environnementale donnée (par exemple, le kininogène dont la synthèse est provoquée par une nourriture pimentée comme dans la cuisine arabe est une musique de style arabe; la chaîne légère de la myosine, protéine musculaire, évoque de façon frappante la musique militaire du réveil ou de la corvée; la lipotropine beta, précurseur de la beta-endorphine, qui induit l'endormissement, est une berceuse, etc...); et que les transpositions en question ont effectivement, lorsqu'on les écoute profondément, une action qui a toutes les caractéristiques de l'expression du gène correspondant.

Toutefois, comme il ne saurait être question, pour nous, et pour des raisons tenant tant à l'éthique qu'au sujet même de ce travail, de prouver ces faits -- que nous avons nous-même observés à maintes reprises -- par une expérimentation (humaine ou animale) qui ne tiendrait pas compte de la nécessité manifeste de faire ici intervenir, pour toute application, l'individu dans son intégrité, nous avons dû procéder autrement: d'une part -- je remercie ici André Lichnerowicz pour ses encouragements constants dans cette voie de recherche -- en écrivant puis en résolvant l'équation décrivant la propagation des ondes d'échelle, montrant en particulier la nécessité de portées d'échelle de l'ordre du nombre d'Avogadro pour le parcours des ondes d'échelle, vérifiable par la conformité à l'expérience des conséquences physiques de cette équation; d'autre part, en transformant ainsi les effets dont nous parlons en la conséquence de faits directement observables tels que les actions métaboliques de synthèses protéiques sur d'autres synthèses musicalement voisines, qu'il est alors possible de vérifier de façon systématique (nous en donnons dans ce qui suit divers exemples); enfin, en confrontant les diagnostics établis par d'autres méthodes à ceux dont une écoute attentive des personnes permet alors de préciser, non seulement le contenu, mais le moment où il est possible d'intervenir.

I. Selon la mécanique quantique, à toute particule est associée une onde, de longueur d'onde inversement proportionnelle au produit de sa masse par sa vitesse. Ceci signifie qu'une particule, même relativement massive, si elle est suffisamment stabilisée vis-à-vis de l'agitation thermique pour que sa vitesse soit assez faible, se comporte en fait comme une onde. Lors du processus de synthèse des protéines, lorsqu'un acide aminé est accroché à son ARN de transfert, il est déjà un peu stabilisé; lorsque de surcroît l'ARN de transfert vient à son tour se fixer sur le ribosome, la longueur d'onde de l'acide aminé devient égale à plusieurs fois sa taille: cette situation constitue un véritable "laboratoire" pour vérifier si, comme de Broglie l'avait suggéré pour toute particule quantique, il est pertinent d'attribuer à cet acide aminé une fréquence propre de vibration proportionnelle à sa masse, et si cette vibration contient, ainsi que nous l'avons nous-même développé (Colloque international "Louis de Broglie, physicien et penseur", Paris 1987, et références citées) des harmoniques impliquant des propriétés d'accord "musical" de ces masses.

1. Les masses des acides aminés tendent effectivement, de façon très significative, à se répartir selon une gamme musicale.



La figure 1 ci-dessus représente les logarithmes des masses des acides aminés (dans la base $2^{1/12}$, prenant ainsi pour unité, en ordonnée, le demi-ton tempéré), rangés (en abscisse) par ordre croissant. La courbe en pointillés est une approximation polynômiale du 5ème degré (dont nous avons pris le centre de symétrie comme origine des ordonnées). Les lignes horizontales correspondent aux fréquences musicales dans la gamme tempérée (c'est-à-dire que leur écartement est un multiple entier de

Log2/12: celles en trait plein sont des lignes de partition musicale en clef de sol, où les valeurs nominales des notes correspondent -- modulo un nombre entier d'octaves -- aux fréquences $\nu = mc^2/h$ associées aux acides aminés de masse m ; c étant la vitesse de la lumière et h la constante de Planck, $c^2/h = 1,49048 \times 2^{77}$ Hertz/Dalton, le la ci-dessus est ainsi situé à 440,0 Hertz modulo 76 octaves). Ainsi qu'on peut le constater, les acides aminés se trouvent soit (sensiblement) à l'intersection de la courbe en pointillés et de ces lignes, soit tendent carrément à se synchroniser sur ces lignes, mettant en évidence la propriété annoncée. (A noter que celle-ci disparaît si l'on retranche une molécule d'eau de l'acide aminé, confirmant qu'elle s'applique bien au processus de synthèse lui-même).

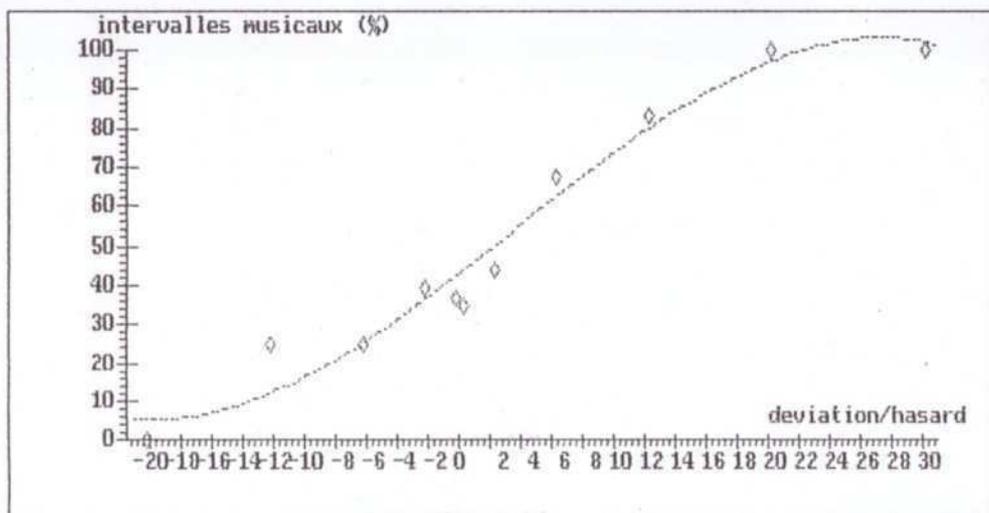
On est alors amené à considérer (pour des raisons qui seront justifiées plus complètement ailleurs) que, de même que les masses des particules élémentaires de durée de vie supérieure à $10^{-12,5}$ s se répartissent toutes (de façon non statistique) sur la gamme tempérée, les fréquences propres des acides aminés, dont les propriétés ondulatoires apparaissent durant le processus de synthèse, se synchronisent effectivement, au bout de $10^{-12,5}$ s environ, sur une même gamme, dès leur stabilisation par l'ARN de transfert. Les fréquences exactes de synchronisation dépendent alors, rigoureusement, de la population des ARN de transfert: si dans tel cas particulier le résultat diffère un peu de la gamme tempérée, celle-ci reste, dans tous les cas, une approximation suffisante pour ne pas modifier les valeurs nominales des notes: celles-ci peuvent s'obtenir en minimisant la distance harmonique globale (qui est une mesure de dissonance généralisant une définition de Y. Hellegouarch, *C. R. Math. Rep. Acad. Sci. Canada*, vol. 4, p. 277, 1982) tenant compte des proportions moyennes d'acides aminés dans les protéines (cf. M. O. Dayhoff, *Atlas of protein sequence and structure*, vol.5, suppl. 3, 1978, p. 363; le résultat reste stable par rapport à de petits déplacements vis-à-vis de ces proportions), ce qui fournit le code suivant, universel à l'approximation de la gamme chromatique, pour les acides aminés (les fréquences étant transposées de 76 octaves vers le bas):

Gly = la grave; Ala = do; Ser = mi; Pro, Val, Thr, Cys = fa; Leu, Ile, Asn, Asp = sol; Gln, Glu, Lys, Met = la; His = si bémol; Phe = si (ainsi que la sélénocystéine); Arg, Tyr = do aigu; Trp = ré aigu.

Il s'agit là des fréquences tempérées les plus proches des fréquences propres initiales des acides aminés, sauf pour la cystéine, la tyrosine et le tryptophane qui baissent d'un tiers de ton, et la sérine qui monte de cinq douzièmes de ton, de façon à minimiser la dissonance globale $\sum_i P_i \log_{\text{sup}}(p, q)_i$ où les $(p/q)_i$ sont les intervalles harmoniques les plus proches des intervalles tempérés pour chaque paire d'acides aminés possible dans la proportion P_i par rapport à l'ensemble, tout en respectant la

condition $\delta f < (1/2)\Delta f$, où Δf est l'écart entre une fréquence et la suivante dans la gamme des fréquences synchronisées et δf le déplacement de la fréquence initiale vers la fréquence synchronisée. (Cette limitation du déplacement, non à un demi-intervalle tempéré mais, comme on le voit, à un demi-intervalle de la gamme obtenue a pour conséquence que les déplacements observés sont parfois supérieurs au quart de ton: elle est nécessaire pour éviter des conséquences illogiques comme une séparation de la cystéine des acides aminés de fréquence initiale voisine de moins d'un sixième de ton). En utilisant ce code, nous pouvons alors traduire les séquences protéiques et vérifier si le résultat obtenu diffère significativement d'une suite de notes au hasard.

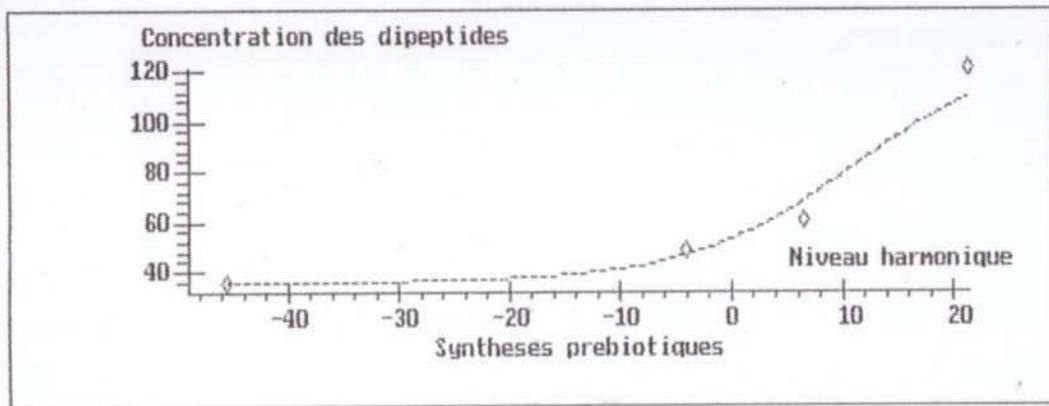
2. Les successions d'acides aminés dans les protéines utilisent, de façon statistiquement très significative, des intervalles "musicaux".



La figure 2 ci-dessus représente la proportion des intervalles les plus musicaux (soit harmoniquement -- octave et quinte, soit mélodiquement -- notes adjacentes de la gamme ci-dessus, soit les deux -- unisson) dans chaque "classe d'écart-type" pour les biais de diaminocides, en fonction du carré du nombre moyen d'écart-types de chaque classe, pour 278 gènes procaryotes (E. Coli + phages). Le résultat donne une courbe sigmoïde, montrant une nette corrélation statistique entre la musicalité des intervalles et la déviation par rapport au hasard du nombre de paires d'acides aminés observées rapporté à la valeur déduite des proportions d'acides aminés.

Ceci met en évidence, non seulement une minimisation de la distance harmonique entre acides aminés successifs, mais aussi des propriétés mélodiques dans leur succession. (Par contraste, les biais jusqu'ici observés dans les proportions de paires d'acides aminés lors de synthèses prébiotiques -- cf. S. Tyagi et C. Ponnamperna, *J. Mol. Evol.* 30, p. 391, 1990 -- suivent l'ordre

des consonances harmoniques, mais non des intervalles mélodiques).



Concentration des dipeptides en fonction de leur niveau harmonique dans les synthèses prébiotiques (d'après les données de Tyagi et Ponnampereuma, loc. cit.)

Pour préciser cette question, nous introduisons la distance mélodique définie de la façon suivante: soient $0, 1, 2, \dots$ des nombres entiers successifs associés aux différentes notes de la gamme, dans l'ordre des fréquences croissantes; la distance mélodique entre deux notes est simplement la valeur absolue de la différence entre les deux nombres entiers correspondants. (Ceci est une simple extension de l'appellation courante "seconde", "tierce", etc... pour les notes d'une gamme). La distance harmonique $\text{Logsup}(p, q)_i$ et la distance mélodique $|n_j - n_i|$ sont alors normalisées en les divisant par leurs valeurs moyennes pour l'ensemble des couples d'acides aminés possibles: la distance mélodique moyenne entre deux acides aminés vaut ainsi $\langle d \rangle = 2,525$ et la distance harmonique moyenne $\langle h \rangle = 1,717498\dots$ avec des écarts-types respectifs $\sigma_d = 1,979741\dots$ et $\sigma_h = 0,995402\dots$ [Ainsi, les intervalles "les plus musicaux" considérés plus haut ont simplement l'une au moins des deux distances normalisées, mélodique ou harmonique, inférieure à $2/3$]. On définit alors une première approche d'un niveau musical en pondérant ces deux distances par les inverses de leurs variances: si δ est la distance pondérée ainsi obtenue, le niveau musical est par définition $100(1 - \delta)$.

Dans ces conditions, il apparaît clairement que les suites de notes obtenues par l'application du code ci-dessus aux séquences d'acides aminés diffèrent nettement du hasard. Certains fragments et certaines séquences entières présentent même un niveau musical particulièrement élevé et nous allons nous consacrer à leur étude.

Signalons toutefois auparavant que les propriétés décrites aux §1 et 2 sont également valables pour les séquences de nucléotides de l'ADN. Si la situation "à trois échelles" acide aminé-tARN-ribosome lors du processus de synthèse des protéines

est en effet exceptionnelle, nous pouvons noter que la situation "à deux échelles", où une petite molécule est stabilisée par une plus grosse tout en conservant une individualité relative, l'est moins: elle se présente en particulier pour les acides aminés à la surface des protéines une fois synthétisées et repliées, sur quoi nous reviendrons; et pour les nucléotides de l'ADN en position "ouverte". Nous pouvons déjà vérifier ici notre hypothèse que la synchronisation des fréquences des amino-acides a lieu dès le rattachement au tARN (autrement dit dès la situation "à deux échelles"): les fréquences propres des nucléotides sont-elles, elles-mêmes, synchronisées sur une même gamme? La réponse est oui: modulo 76 octaves, elles correspondent précisément aux notes mi, fa dièse, sol et la, sur un diapason légèrement plus haut (d'un huitième de ton) du diapason moyen des acides aminés. Notons d'ailleurs que certaines similitudes entre séquences d'ADN et séquences musicales avaient été relevées par divers auteurs: facilité de les mémoriser en les représentant par des suites de notes (K. Hayashi et N. Munakata, *Nature* 310, p. 96, 1984; cf. également M. Pembrey et D. Hart, *Nature* 311, p. 516, 1984; je remercie Marie-Claude Lang pour avoir initialement attiré mon attention sur ces articles), alors qu'une suite de notes au hasard, par exemple, ne se retient pas plus facilement qu'une suite de nombres au hasard [cf. J. E. Marie, *Musique et mémoire*, in *Trois discours sur le musical*, Edisud, Aix-en-Provence 1983]; tendance à la répétition (S. Ohno et M. Jabara, *Chemica Scripta* 26B, p. 43, 1986; S. Ohno et M. Ohno, *Immunogenetics* 24, p. 71, 1986); même avec des codes différents entre eux (et différents des valeurs ci-dessus dont nous avons toutes les raisons de penser qu'elles correspondent bien à la réalité physique), ces observations étaient fort "intrigantes" et certains auteurs les avaient rapprochés de nos propres travaux sur les particules élémentaires (cf. S. Leach et M. Englert, *J. Elec. Spectroscopy*, 41, p. 181, 1986).

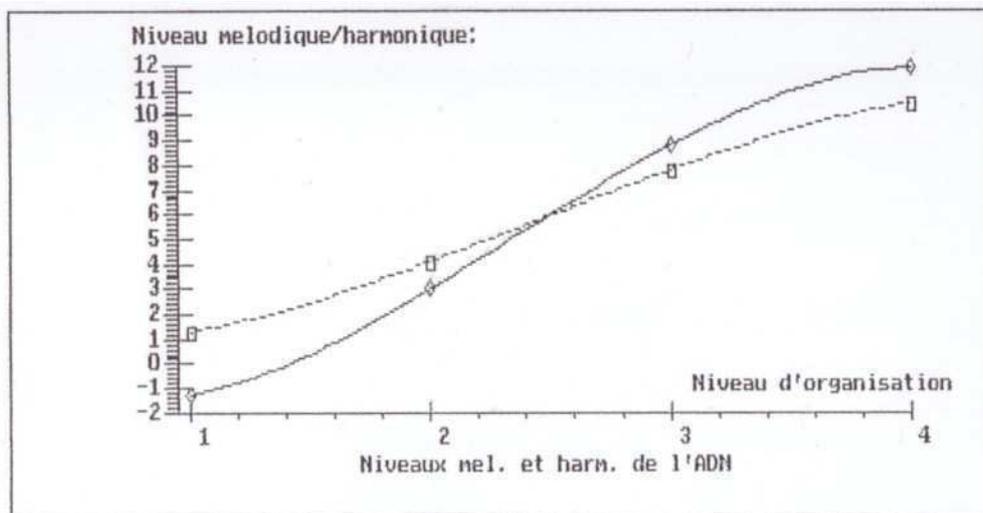
On notera d'ailleurs la précision de la synchronisation des fréquences des nucléotides de l'ADN, qui passe une partie de son temps en position ouverte (liaisons hydrogène ouvertes): dénotant par θ un paramètre qui vaudrait 0 si l'on avait exactement, en fréquences, $CG = AT$ (comme ce serait le cas pour des intervalles, soit tempérés purs, soit harmoniques purs proportionnels à 10/9, 5/4, 4/3 et 3/2, pour des corpuscules isolés) et 1 si l'on avait $C+G = A+T$ (comme si les "paires de base" en position fermée étaient isolées et synchronisées entre elles), l'équation

$$[(C^\theta + G^\theta)/2]^{1/\theta} = [(A^\theta + T^\theta)/2]^{1/\theta}$$

donne avec précision $\theta = 2/3$, indiquant un rythme ternaire pour la "lecture" de l'ADN, comme c'est précisément le cas du fait du code Génétique! (Ceci laisse supposer, en fait, une synchronisation entre la "respiration" de l'ADN et sa transcription dont on peut comprendre l'intérêt biologique. On peut faire ici une analogie dont l'intérêt se précisera plus loin: la rythmicité dans la

musique humaine est basée sur une synchronisation entre la mesure - basée sur les battements du coeur - et le phrasé - basé sur la respiration. C'est en fait là l'idée de base du "décodage" des durées des notes pour les musiques de protéines, sur lequel nous reviendrons).

De fait le niveau musical des séquences de nucléotides de l'ADN (défini à partir des rapports harmoniques de fréquences décrits ci-dessus) diffère nettement du hasard, indiquant là aussi déjà des propriétés mélodiques: les fréquences de dinucléotides tabulées par R. Nussinov, Nucl. Ac. Res. 12, p. 1749, 1984 impliquent un niveau musical tant mélodique qu'harmonique (que l'on peut définir séparément à partir des distances mélodique et harmonique respectivement), croissant régulièrement avec l'évolution des procaryotes aux eucaryotes invertébrés, puis aux vertébrés, puis à l'homme.



Niveaux mélodique (en trait plein) et harmonique (en pointillés) moyens des ADN procaryotes (1), invertébrés (2), vertébrés (3) et humains (4).

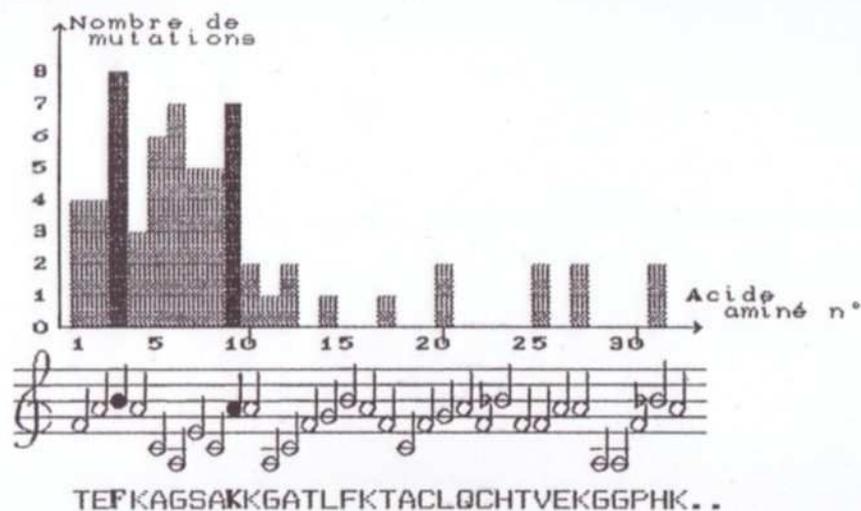
3. Une première fonction de ces propriétés musicales est de protéger les séquences vis-à-vis de mutations.

Si nous regardons les niveaux musicaux des ferrédoxines, protéines bien étudiées dans l'ensemble du règne vivant, nous constatons que les niveaux musicaux les plus faibles (mais quand même supérieurs à 8, contre 0 pour une suite de notes au hasard) sont ceux des bactéries, tandis que les eucaryotes ont entre 16 et 23 -- à l'exception des épinards, "spécialisés" dans le stockage du fer, qui atteignent 25,7 -- et que les algues bleues atteignent 28. Contrairement aux eucaryotes qui ont un noyau et aux bactéries qui ont une coque protectrice, les algues bleues ont un génome fortement soumis aux conditions extérieures et sont de ce fait d'une grande variabilité. On peut en inférer que le mécanisme de contrôle vibratoire que nous étudions -- qui revient à éliminer les "fausses notes", opérant ainsi une sélection -- est un premier

moyen pour elles de contrôler les variations de leur génome.

Cette hypothèse se trouve confirmée par l'étude des "fausses notes", précisément. On peut lire dans "La musique et les musiciens", d'Albert Lavignac (G. Billaudot éditeur, Paris 1956), p. 192: "Un contour mélodique à éviter... est la quarte augmentée en trois notes, dans la même direction... sauf si la dernière note... obéit ensuite à sa tendance attractive" (montant ou descendant d'un demi-ton). Nous n'avons trouvé que très peu d'exemples (une demi-douzaine, alors que la valeur au hasard dépasserait 20, sur plus de 5400 protéines de la banque NBRF 23) où cette règle est violée au début d'une protéine, parmi lesquels 3, qui ont été bien étudiés, présentent en effet un intérêt particulier: il s'agit des cytochromes C2 iso-2 des bactéries *Rhodospirillum fulvum* et *R. Molischianum*, et du cytochrome C iso-1 de la levure de bière (*S. Cerevisiae*). Or, non seulement ces trois organismes -- et eux seulement -- possèdent une deuxième molécule exerçant la même fonction (iso-1 chez les deux bactéries, iso-2 chez la levure), dans laquelle la "fausse note" est corrigée (au prix, apparemment, d'un poids moléculaire plus élevé: deux acides aminés en plus chez les deux procaryotes, quatre chez l'eucaryote), mais la molécule initiale est très mutagène, précisément à l'endroit de la "fausse note" et autour d'elle: plus des deux tiers des mutations provoquées affectant la région amino-terminale du cytochrome C iso-1 de la levure et donnant un résultat viable (cf. D. M. Hampsey et al., *J. Mol. Biol.* 201, p. 471, 1988) soit portent directement sur la fausse note (plus du tiers), soit suppriment sa "fausseté" en changeant l'une des deux qui précèdent, soit en la faisant suivre sa tendance attractive (en montant d'un demi-ton).

Mutations du cytochrome C iso-1 de la levure



La "fausse note" (F-3) est la plus mutagène ; lorsqu'elle est conservée, la mutation la plus fréquente est K-9 changée en R (= do aigu, c-à-d la tendance attractive de F-3 = si).

De façon plus générale, la comparaison des séquences du cytochrome C de l'infusoire *Euglena gracilis* et du cytochrome C humain montre assez comment l'évolution "retient" les corrections qui améliorent la musique: par exemple la proportion des intervalles "les plus musicaux" (d ou $h < 2/3$) passe de 37,5% chez les 32 premiers acides aminés de l'Euglène à 56% chez l'homme; la proportion des intervalles "consonants ou mélodiques" ($h \leq \text{Log}_5$ ou $d \leq 2$) passe de 61% à 90%. Il faut cependant préciser qu'en règle générale, les niveaux musicaux les plus élevés sont non pas nécessairement ceux des molécules humaines, mais ceux des molécules des organismes "spécialisés" (prolactine bovine, ferrédoxine des épinards, etc...).



32 premiers acides aminés du cytochrome C humain (en haut)
 et du cytochrome C de l'Euglène gracile (en bas)
 (les différences sont notées en caractères gras)

4. Une deuxième fonction de ces propriétés musicales est de réguler le rythme de la synthèse.

L'étude des pauses et irrégularités de rythme survenant lors de la synthèse des protéines, menée en particulier par S. Varenne et son équipe (*J. Theor. Biol.* 180, p. 549, 1984), a pour l'essentiel confirmé l'idée que ces pauses sont statistiquement reliées à la rareté de certains codons dans une population d'ARN de transferts: il faut plus de temps pour "installer" dans une chaîne d'acides aminés en élongation un acide aminé correspondant à un codon rare. On peut toutefois être frappé par la plus grande régularité apparente des synthèses réelles, telles que figurées dans l'article précité -- et même si, ainsi que l'a observé M.-C. Lang (communication privée), une électrophorèse bidimensionnelle serait nécessaire pour trancher cette question afin de distinguer les pauses réelles des protéines plus courtes -- en regard des prédictions fondées sur l'usage des codons; de même peut-on remarquer, dans l'exemple de la protéine bactérienne OmpA donné par cet article, un "pic" nettement prédominant de synthèses transitoires de 184 acides aminés, correspondant à une pause plus importante à cet endroit, que la seule rareté des codons ne distingue pas à ce point. Ceci soulève au moins la question de l'existence de mécanismes régulateurs du rythme de la synthèse autres que, tout en s'accommodant de, l'usage des codons.

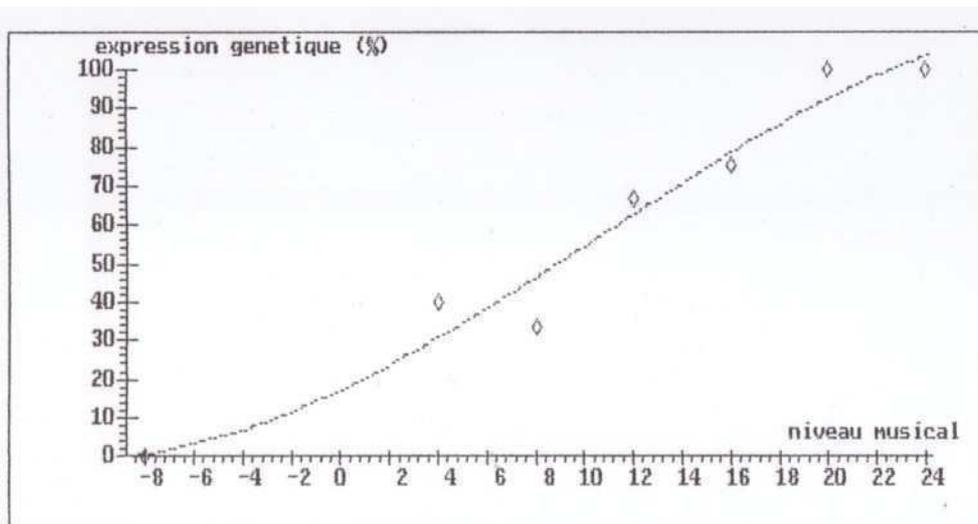
Pour ce qui est de la musique, les expériences de J. Sloboda (*L'esprit musicien*, P. Mardaga éditeur, Liège 1988, et références citées) ont montré qu'un auditeur savait repérer une "cadence", se traduisant par une pause de l'énoncé musical, dans une suite de notes même lorsque celles-ci lui étaient présentées de façon monotone avec des durées identiques. Or dans le cas précité de la protéine OmpA, on a, précisément à l'endroit cité, une "cadence de fin" très nette: ceci amène au moins à faire l'hypothèse que le ou les mécanismes régulateurs supposés existe, l'un au moins étant de nature musicale, c'est-à-dire que les pauses tendent à se placer aux cadences musicales.

Si cela est vrai, il existe un moyen simple de vérifier: en effet l'un des intérêts d'un tel mécanisme serait de pouvoir "synchroniser" des protéines dont la synthèse simultanée est nécessaire à l'organisme. De telles protéines se "superposent"-elles musicalement, pour former des polyphonies?

La réponse est oui. Si nous considérons les deux facteurs d'élongation bactériens EF-T et EF-G, dont la "cosynthèse" est nécessaire à toute synthèse protéique, ils se "superposent" harmoniquement sur une bonne soixantaine d'acides aminés, ce qui exclut toute possibilité de hasard. Ceci est d'autant plus intéressant que leurs usages des codons sont légèrement différents, ce qui impliquerait normalement un "rythme" de synthèse légèrement différent; et que le "décodage" que l'on peut produire à l'oreille pour déterminer les durées des notes de chacun d'eux séparément est également légèrement différent de ce que donne cette superposition. Il apparaît donc bien que l'une des fonctions de ces musiques est de réguler le rythme des élongations polypeptidiques.

5. Une troisième fonction de ces musiques est de réguler le taux de la synthèse.

Si en effet on a des mécanismes de synchronisation à distance entre synthèses protéiques, ceci peut se traduire par des effets de résonance pour des synthèses "en phase" -- ou, comme l'a suggéré Vincent Bargoin (communication privée), d'"opposition de phase" au contraire -- agissant donc sur l'amplitude des vibrations. Or, s'agissant de vibrations quantiques, ces amplitudes sont, par leur carré, proportionnelles à une densité de probabilité de présence: l'accroissement d'une amplitude associée à un ou à une suite d'acides aminés se traduira par davantage d'acides aminés de cette nature -- autrement dit, par une augmentation du taux de synthèse.



Un moyen simple de vérifier cette idée consiste à étudier la variation statistique de l'expression génétique en fonction du niveau musical. La figure 3 ci-dessus représente le pourcentage de protéines "hautement exprimées" de la bactérie E. Coli en fonction du niveau musical des 32 premiers acides aminés. Ici encore, on obtient une corrélation très nette, avec une courbe sigmoïde.

Il s'ensuit que les molécules musicalement homologues doivent se trouver métaboliquement agonistes. Pour vérifier cela, il va nous falloir décoder complètement ces musiques, c'est-à-dire déterminer les durées des notes afin de repérer les homologues musicales. Une indication statistique peut néanmoins être obtenue aisément en comparant les molécules ayant un pourcentage de notes communes non seulement élevé mais nettement supérieur au pourcentage d'acides aminés communs, dont on peut alors présumer que cela proviendrait plus spécialement de la musique. Disons tout de suite que la vérification, non seulement "marche", mais paraît réellement fournir une interprétation à certains résultats contradictoires obtenus avec le génie génétique.

Par exemple, le tPA (tissu plasminogen activator) sur qui l'on fondait de grands espoirs thérapeutiques s'est avéré relativement décevant: les doses auxquelles son injection est efficace sont non seulement bien plus élevées que lorsqu'il est fabriqué par l'organisme, mais sont alors dangereuses. Or le tPA et le plasminogène, qu'il active, ont, sur les 30 premiers acides aminés, 17 notes en commun mais seulement 3 acides aminés: ceci est compatible avec la conclusion qu'ils doivent se stimuler (musicalement) l'un l'autre, ce qui ne se produit donc pas en l'absence de synthèse par l'organisme lui-même de ces protéines.

De tels exemples peuvent être multipliés, mais leur valeur de confirmation, au stade où nous en sommes, est seulement statistique et nous préférons les réserver pour après l'exposé du décodage.

Mais une autre conséquence peut être vérifiée: en effet, s'il y a stimulation de la synthèse protéique par effet de résonance se traduisant par une augmentation d'amplitude, il y a forcément aussi possibilité d'inhibition par opposition de phase: comment cela se traduit-il? Les fréquences des vibrations quantiques sont très élevées (de l'ordre de 1025 Hz), comment une opposition de phase peut-elle se manifester au niveau de la séquence d'acides aminés?

Pour résoudre ce problème, nous sommes obligés de faire un retour à la physique concernée. Toute mesure physique implique en effet d'abord le choix d'une échelle: la prise en compte de cette propriété va nécessairement nous amener à inclure, dans les équations décrivant les phénomènes physiques, un paramètre fixant l'échelle. La question qui se pose alors est celle de la forme de l'opérateur associé: or précisément, pour expliquer le caractère exponentiel (gamme chromatique tempérée) du spectre de masse des particules, nous avons été amenés -- cette idée est issue d'une conversation avec Moshé Flato, avec la participation de Daniel Sternheimer -- à modifier la définition usuelle des impulsions en mécanique quantique relativiste en multipliant celles-ci par l'opérateur $e^{-i\alpha\partial/\partial s}$ où $\partial/\partial s$ est l'opérateur canoniquement conjugué à l'opérateur d'échelle s (c'est-à-dire l'opérateur de dérivation partielle par rapport à ce dernier) et α une constante dont nous verrons la signification plus loin. L'opérateur dalembertien (associé au carré des impulsions relativistes) de l'équation de Klein-Gordon décrivant le comportement d'une particule quantique (pour chaque composante de spin) sera alors lui-même multiplié par le carré de l'opérateur ci-dessus, et nous pouvons ainsi résoudre l'équation obtenue pour une particule libre dont le comportement sera décrit par une fonction d'onde $\Psi(\mathbf{r}, s, t)$ à valeurs dans l'espace-temps et l'échelle :

$$\square e^{-2i\alpha\partial/\partial s} \Psi = m^2 \Psi.$$

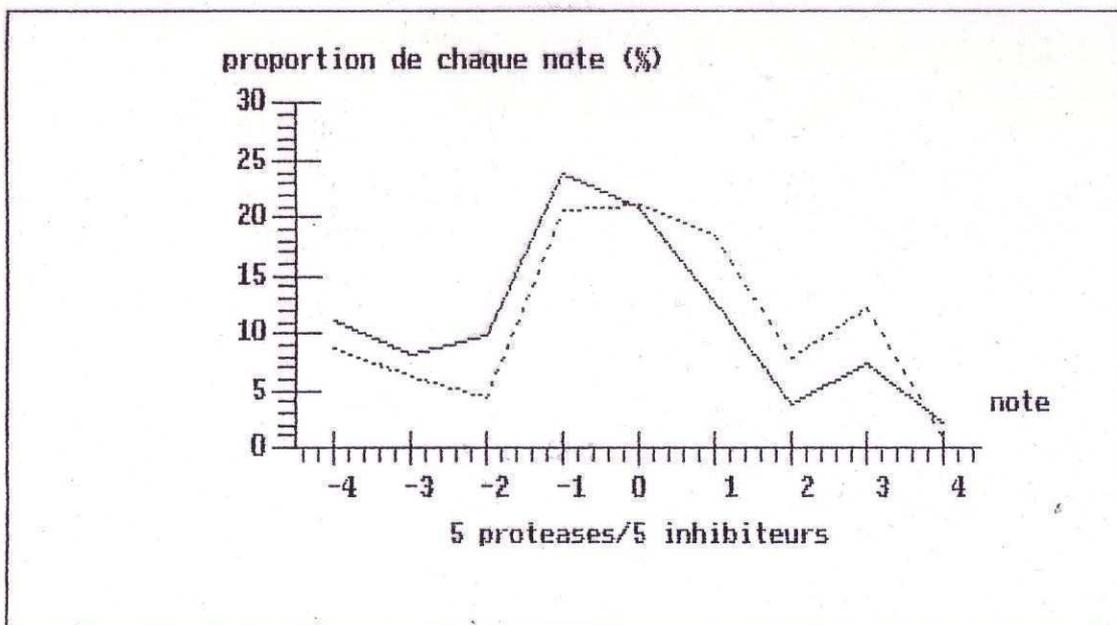
Or les solutions -- que nous appelons ondes d'échelle -- de cette équation possèdent des propriétés remarquables. De forme générale

$$\Psi(\mathbf{r}, s, t) = (\sum_l A_l e^{2\pi l s/\alpha - t/2T - f(\mathbf{r})}) e^{i(\mathbf{k}\cdot\mathbf{r} - 2\pi\nu + \beta ns)}$$

elles apparaissent en effet comme une superposition d'ondes progressives exponentielles dans l'échelle, de vitesses de groupe $\alpha/4\pi l T$, où l est un entier et T la durée de vie de la particule (c'est-à-dire qu'on a supposé une variation temporelle en $e^{-t/2T}$ de l'amplitude de la fonction d'onde), et de vitesses de phase $2\pi\nu/\beta n$, où n est un entier et β égal à 2π divisé par la portée de l'échelle. La valeur entière de n assure que les vibrations aux deux échelles restent constamment en phase, ce qui était l'idée même de de Broglie; en même temps que l'une se trouve être un harmonique de l'autre, ce qui explique les phénomènes d'accord observés. Mais c'est le rapport -- qui est donc un rapport

d'échelle -- entre la vitesse de groupe et la vitesse de phase des ondes d'échelle qui est remarquable ¹ : il atteint en effet des valeurs considérables; si T correspond à la durée pendant laquelle un acide aminé est sur le ribosome (par l'intermédiaire de son tARN), il est même de l'ordre du nombre d'Avogadro! Plus précisément, nous sommes amenés à prévoir l'existence de "périodes" vibratoires dans le processus de synthèse, de l'ordre de 16T environ, au cours desquelles la suite des fréquences va tendre à se répéter en minimisant la dissonance, c'est-à-dire la distance harmonique, avec les fréquences situées une période auparavant. Par rapport à cette modulation des fréquences successives des acides aminés, on aura une opposition de phase en prenant systématiquement la fréquence symétrique par rapport à la fréquence centrale, comme on le voit en décomposant ces "périodes" en composantes sinusoidales.

Précisément, les fréquences propres des acides aminés ont, comme celles des particules élémentaires, une répartition sensiblement symétrique par rapport à un "centre de gravité",



En trait plein, proportion des notes de 5 protéases (Célastase, prothrombine, trypsinogène, chymotrypsinogène et α -lytique protéase -- cf. Atlas NBRF vol. 5, suppl. 2, 1976, p. 302); en pointillés, idem pour 5 inhibiteurs de protéases (α_1 -antitrypsine, anti-thrombine III, "basic trypsin inhibitor", et inhibiteurs de protéase des venins de vipère et mamba).

¹ Ce rapport vaut $8\pi^2 v l T / \alpha \beta n = 8\pi^2 l N v T / n \text{Log} 2 = 113,91. (lN/n). vT$: la valeur du produit $\alpha\beta$ est en effet donnée par le spectre de masse $m = m_0 e^{\alpha\beta n} = m_0.2^{n/N}$ où n est un entier (comme 1) et $N = 1, 12, 72...$

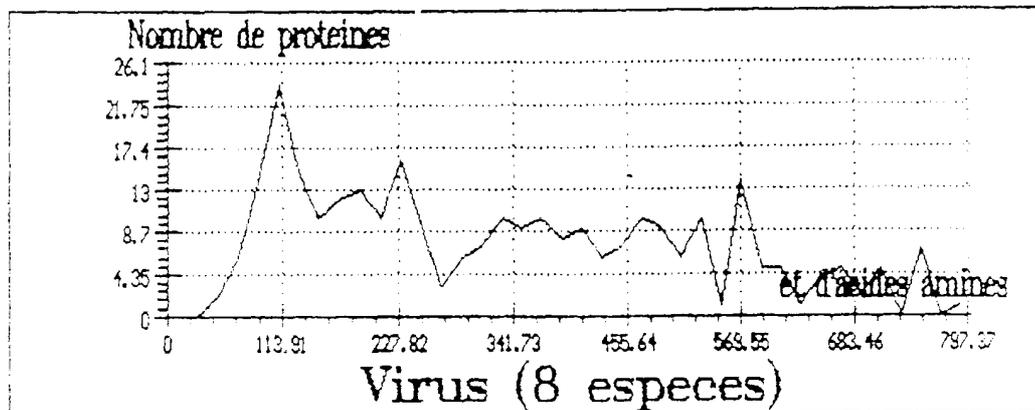
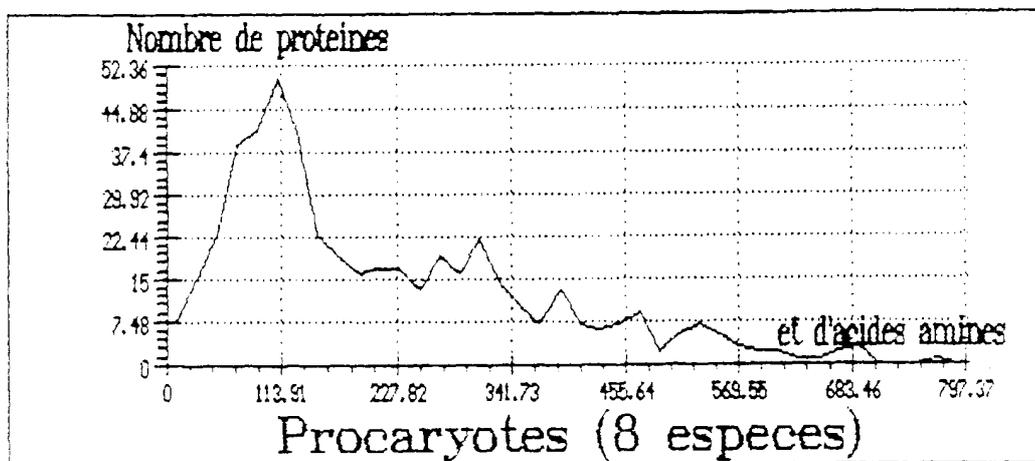
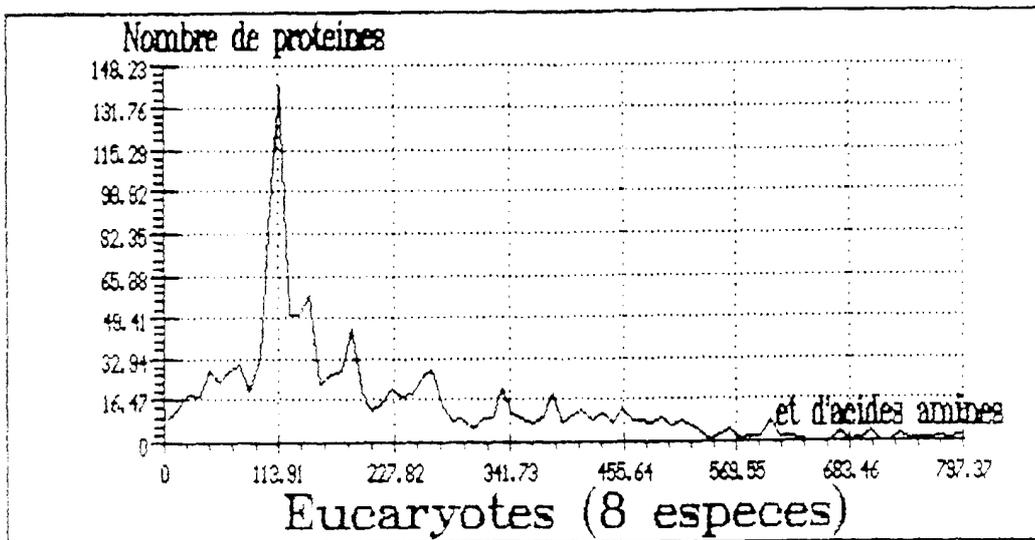
correspondant aux fréquences de la leucine et de l'isoleucine. Nous sommes donc conduits au code suivant, pour produire des séquences de vibration en opposition de phase avec celles d'une synthèse protéique, en prenant le symétrique du code précédent par rapport à sol:

Trp = do; Arg, Tyr = ré; Phe = mi bémol; His = mi; Met, Lys, Glu, Gln = fa; Asp, Asn, Ile, Leu = Sol; Cys, Thr, Val, Pro = la; Ser = si bémol; Ala = ré aigu; Gly = fa aigu.

Nous avons ainsi un test très simple de la validité du précédent raisonnement: les couples de protéines antagonistes ont-ils des séquences vibratoires symétriques? on observe bien, en effet, une telle symétrie de répartition (statistique) entre les protéases et leurs inhibiteurs, entre la prolactine et l'hormone lutéotrope, etc... Un exemple intéressant est celui de la prothrombine et du plasminogène, qui ont un ancêtre commun peu éloigné et des fonctions métaboliques opposées: si l'on compare les "kringle regions" homologues, de séquences d'acides aminés très voisines, le rapport GAS/HFRYW y vaut déjà $1,07 \pm 0,09$ chez la prothrombine (protéase) contre $0,85 \pm 0,07$ chez le plasminogène; sur le plan musical, la première est très rythmée et le second calme, en contraste avec la convergence plasminogène - tPA déjà signalée. Un autre exemple particulièrement net, est celui des protéines G_i et G_s dont les tonalités musicales sont effectivement symétriques, alors même qu'il s'agit là de tonalités rares chez les protéines (respectivement la mineur modulé en sol mineur, et si bémol majeur modulé en do majeur) et qui, toujours respectivement, inhibent et stimulent la même enzyme (l'adénylate cyclase).

Un autre test concerne les échelles de durée, dont la constante caractéristique vaut d'après l'équation ci-dessus, $8\pi^2/\text{Log}2 = 113,91$: considérant les ondes d'échelle reliant chaque acide aminé à une chaîne protéique en formation, la "signature" de la présence d'ondes d'échelle données par cette équation dans les protéines sera la présence de pics situés à cette valeur (et à ses multiples entiers) dans l'histogramme des longueurs de protéines mesurées en nombre d'acides aminés. Comme on peut le constater (figure ci-contre), les statistiques que l'on peut faire à partir des longueurs de protéines (d'après la banque NBRF 23) confirment de façon très remarquable cette prédiction: chez les eucaryotes notamment, on observe la présence d'un pic très étroit de plus de 5 écarts-types, à une valeur qu'un ajustement non-linéaire permet de déterminer précisément à $113,75 \pm 0,25$; ce pic est également présent quoique plus large chez les procaryotes et présente des harmoniques chez les virus, soulignant l'accroissement de l'importance des ondes d'échelle au fur et à mesure de l'évolution (et de leur portée), comme on le constate aussi de par l'évolution du niveau musical, dans l'ADN comme dans les protéines.

La répartition statistique des longueurs de protéines, donnant une mesure du rapport temporel (temps de synthèse protéique):(temps d'adjonction d'un acide aminé), et qui présente un pic dominant à $8\pi^2/\text{Log}2 \approx 113,91$ (ainsi que des harmoniques de ce pic, visibles plus particulièrement chez les virus), fournit une claire évidence de la présence d'ondes d'échelle données par l'équation ci-dessus dans la biosynthèse des protéines.



LONGUEURS DES PROTEINES POUR DIFFERENTS TYPES VIVANTS

Eucaryotes	Procaryotes	Virus
Homme 388	Escherichia Coli 309	Influenza 73
Souris 199	Bacillus 49	Epstein-Barr 68
Bovin 166	Salmonella 31	Adénovirus 51
Rat 126	Pseudomonas 30	V. de souris 35
Levure 97	Rhodopseudomonas 28	Sida 27
Lapin 89	Staphylocoque 24	V. de singe 26
Porc 86	Streptomyces 18	Leucémie 25
Poulet 75	Desulfovibrio 18	Hépatite 23
Total 1226	Total 501	Total 328

Total général 2055 protéines

Les ordonnées sont graduées en écarts-types

Une autre vérification: le biais des alanines

Une autre vérification de la présence des lois "musicales" dans la succession des acides aminés qui composent les protéines découle de la quasi-universalité de la méthionine en première position: c'est le fait que l'alanine doit être alors favorisée en deuxième position, de manière à obtenir un accord parfait entre la méthionine, l'alanine et le début de la protéine en formation. Ce biais, effectivement très marqué (traits pleins sur la figure ci-dessous, en comparaison des pointillés figurant les proportions générales) l'est encore plus dans les protéines ribosomales.

Distances mélodiques et harmoniques

		% dans les peptides (total 617)									
		8.6	8.6	7.0	20.8	21.7	18.2	2.0	2.3	8.3	1.3
		nb codons (total 64)									
		4	4	6	14	13	7	2	2	8	1
h	d	G	A	S	PV TC	LI ND	WK EM	H	F	RY	W
		0	1	2	3	4	5	6	7	8	9
G - Gly	0	0	1	2	3	4	5	6	7	8	9
A - Ala	1	Log 6	0	1	2	3	4	5	6	7	8
S - Ser	2	Log 3	Log 5	0	1	2	3	4	5	6	7
P-Phe V-Val T-Thr C-Cys	3	Log 8	Log 4	Log 16	0	1	2	3	4	5	6
L-Leu I-Ile N-Asn D-Asp (B-Asx)	4	Log 9	Log 3	Log 6	Log 9	0	1	2	3	4	5
Q-Gln K-Lys E-Glu M-Met (Z-Cys)	5	Log 2	Log 5	Log 4	Log 5	Log 10	0	1	2	3	4
H - His	6 [6] [6]	Log 32 [54] [21]	Log 16 [9] [7]	Log 64 [36] [7]	Log 4 [7] [1]	Log 32 [6] [7]	Log 16 [7] [24]	0	1	2	3
F - Phe [SoC]	7	Log 9	Log 15	Log 3	Log 45	Log 5	Log 9	Log 135 [25] [15]	0	1	2
R-Arg Y-Tyr	8	Log 12	Log 2	Log 8	Log 3	Log 4	Log 6	Log 9 [12] [8]	Log 16	0	1
W - Trp	9 [6]	Log 8 [27]	Log 20 [9]	Log 16 [9]	Log 5 [27]	Log 40 [5]	Log 4 [27]	Log 5 [5] [15]	Log 32 [6]	Log 10 [9]	0

Distances mélodiques (d) et harmoniques (h) entre peptides (acides aminés)

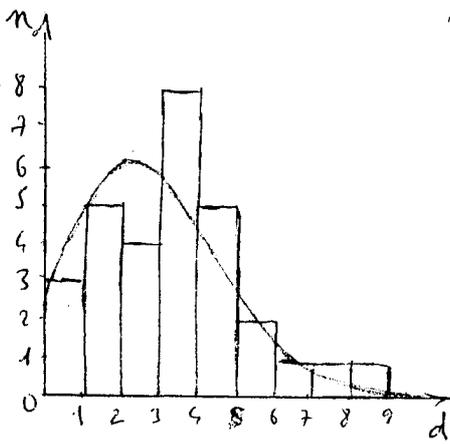
$$\langle d \rangle = 3,30 \pm 2,380476 \quad \text{Avec multiplicateur: } 2,525 \pm 1,999461165$$

$$\langle h \rangle = \dots = \dots \pm \dots [b: 1,772061204]$$

$$\bar{\sigma}_d / \bar{\sigma}_h = \dots \pm \dots$$

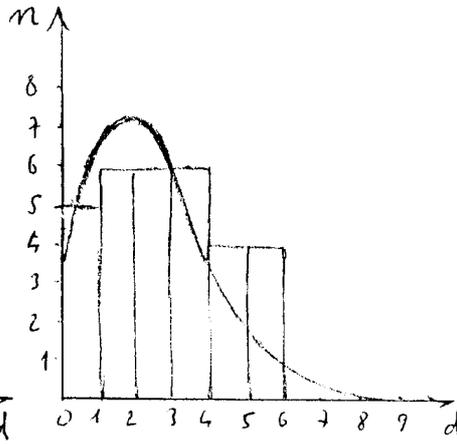
$$\delta = \left[\bar{d} + \left(\frac{\bar{\sigma}_d}{\bar{\sigma}_h} \right)^2 \bar{h} \right] / \left[1 + \left(\frac{\bar{\sigma}_d}{\bar{\sigma}_h} \right)^2 \right] \quad \bar{d} = d / \langle d \rangle, \quad \bar{h} = h / \langle h \rangle$$

$$\nu = 100(1 - \delta) \quad \text{niveau musical}$$



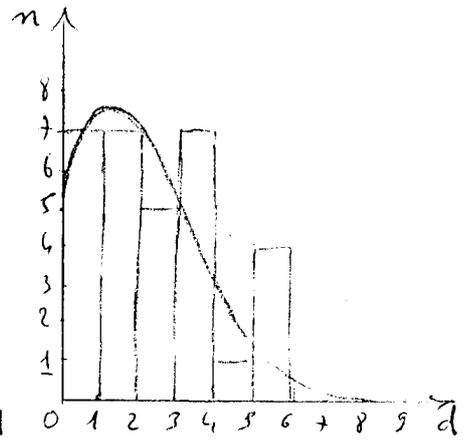
Euglène

$\langle d \rangle = 2,93$, $\lambda = 2,19$
 $P(X > X_0) = 0,9$



Mouche

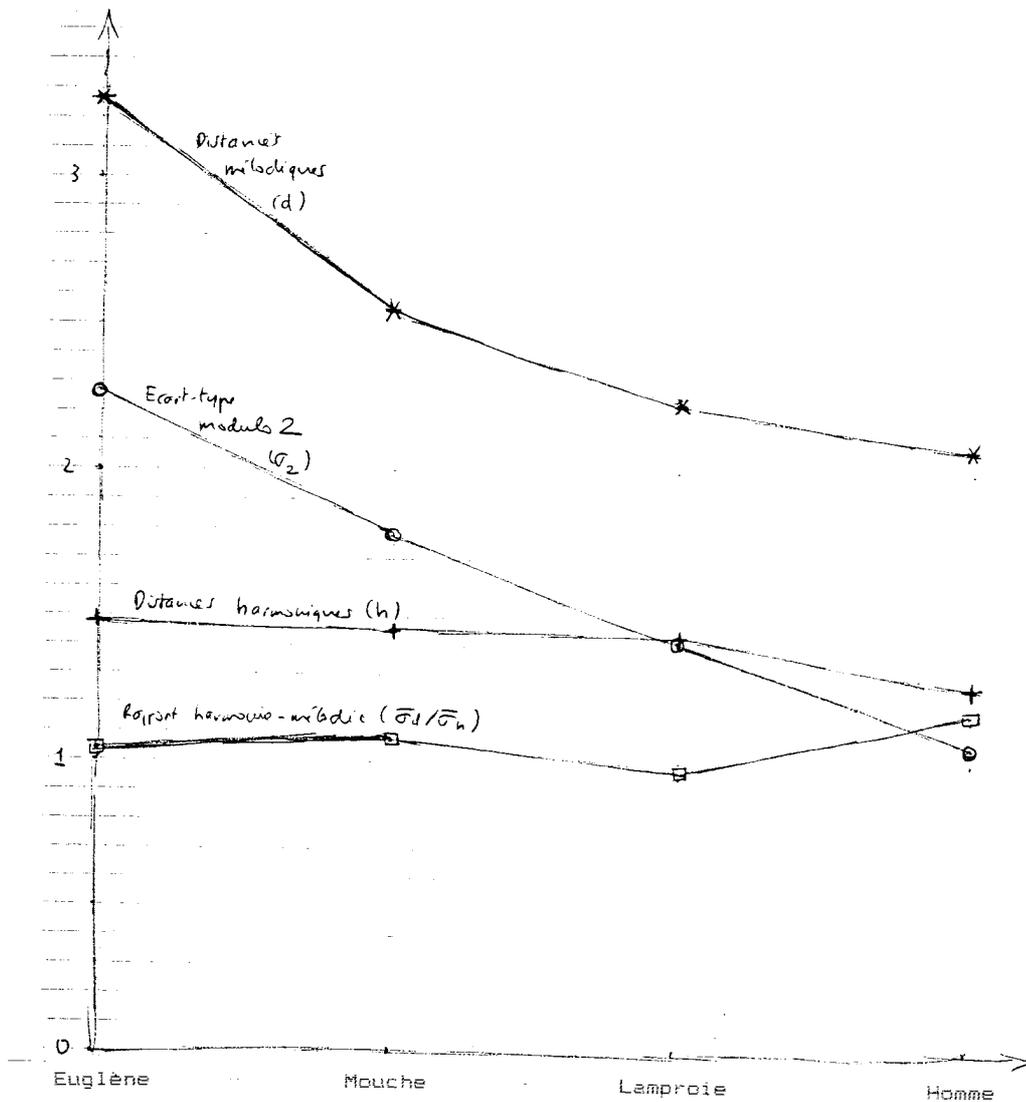
$\langle d \rangle = 2,32$, $\lambda = 2,00$
 $P(X > X_0) = 0,8$



Homme

$\langle d \rangle = 2,00$, $\lambda = 1,46$
 $P(X > X_0) = 0,4$

Distribution des distances mélodiques des 32 premiers acides aminés du cytochrome C rapportées à la courbe d'équation $n = (31 \lambda / \langle d \rangle) e^{-\lambda} \lambda^{d/\langle d \rangle} / \Gamma(\lambda d / \langle d \rangle + 1)$, pour diverses espèces animales.



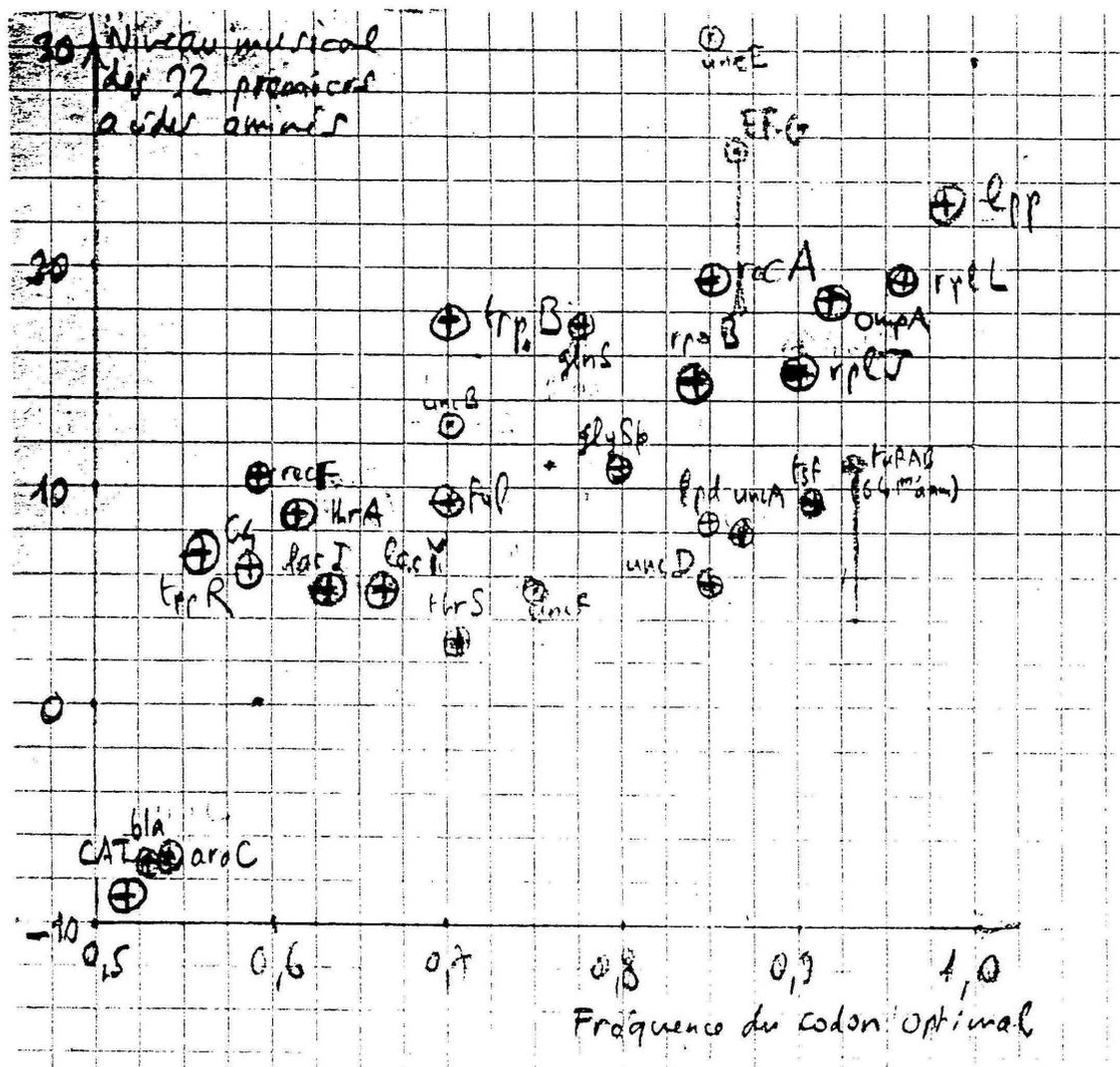
L'usage des codons

Une des énigmes de la biologie moléculaire concerne les biais observés dans l'usage des différents codons correspondant à un même acide aminé. S'il s'agit apparemment d'une forme de stratégie minimisant les essais et erreurs dans les appariements codon-anticodon lors du processus de traduction, la signification réelle de tel choix plutôt que tel autre reste mystérieuse. Et si les pauses dans les synthèses leur apparaissent statistiquement corrélées (M. Varenne et al.), la réalité temporelle du processus d'élongation protéique apparaît cependant sensiblement plus régulière que ce que laisserait attendre un calcul de probabilité simplement basé sur le fait que la rareté d'un codon va statistiquement impliquer un temps plus long pour l'adjonction de l'acide aminé correspondant à la chaîne protéique en formation.

Or comme nous avons pu le vérifier, l'identité des 'styles rythmiques' repérables lors du décodage des protéines, apparaît entraîner celle des usages de codons correspondants. Dans l'exemple ci-contre, on observe une beaucoup plus grande similarité entre les usages de codons de la calpaine chaîne légère humaine (que je qualifierais de rock-boogie) et du fibrinogène gamma de la lamproie (hard rock), qu'entre ce dernier et le fibrinogène gamma humain (que je qualifierais plutôt de quadrille).

USAGE DES CODONS

Acide aminé	Fibrinogène foetus humain	Fibrinogène foetus (au, poie)		Colpaine humaine
Arg	AGA	CAC	o	CAC
Leu	CUG-CUA	CUC		CUG
ser	UCU	UGC		AGC
Thr	ACU	ACG		ACA
Pro	CCU	CCG	.	CCA-CCG-CCG
Ala	GCU	GCC	o	GCC
Gly	GGA	GGC	o	GAC
Val	GUU	GUG		GUC-GUU
Lys	AAA	AAG	o	AAG
Asn	AAU	AAC	o	AAC
Gln	CAG	CAA	o	CAA
His	CAU	CAC	o	CAC
Glu	GAA	GAG	o	GAG
Asp	CAU	CAE	.	CAC-GAU
Tyr	UAU	UAC	o	UAC
Cys	UGU	UGC	o	UGC
Phe	UUC-UUU	UUC	o	UUC
Ile	AUU	AUC	o	AUC
Ter	UAA	UGA	o	UGA
Protéoglycane au 3 ^e position	TAGC	CGAT		CGAT



Données brutes ayant conduit à la courbe du §5

Le taux d'expression génétique étant corrélé au pourcentage de codon optimal (T. Ikemura et al.), nous avons porté la valeur du niveau musical des 32 premiers acides aminés de diverses protéines de E. Coli en fonction de ce pourcentage. Sur la figure résultante ci-dessus, on en remarque quelques-unes qui s'écartent des valeurs moyennes, et singulièrement un couple constitué des facteurs d'élongation EF-G et tufAB = EF-T qui s'en écartent de façon symétrique - propriété conservée si l'on en considère les 64 premiers acides aminés, qui les rapprochent de la valeur moyenne du niveau musical pour ce pourcentage de codon optimal. Ces deux protéines, présentes et actives simultanément dans toute synthèse protéique, possèdent une propriété remarquable : leurs séquences sont harmoniquement superposables sur une soixantaine d'acides aminés. Elles forment ainsi une harmonie à deux voix, dont la probabilité d'occurrence au hasard quasi-nulle fournit une évidence très nette de la présence dans ces séquences des propriétés 'musicales' étudiées ici.

```
(EFT) > M - S K E K F E R T K P H V N - V G T I - G H V - D H G K T T L T A
(EFG) > M A R T T P I A R Y R N I G I S A H - I D A G K T T T T E R I L F Y

(EFT) >> A I T T V - L A K T - Y G G A A R A F D Q I D N - A P E E K A R G
(EFG) >> T G V N H K I G E V H D G - A A - T M D W M E Q E Q E R G I T I T
(EFT) > M S K E K F E - R T K P H V N V
```