

## **Les boucles de rétroaction, rouages des réseaux de régulation biologiques**

Le fonctionnement approprié des systèmes organisés, tels que des organismes vivants ou des cellules, requiert que le niveau de certains éléments cruciaux puisse être évalué et pris en compte pour déterminer leur futur taux de production. Cette régulation est réalisée par les boucles de rétroaction, qui sont des circuits orientés d'interactions entre éléments du système. Des combinaisons appropriées de boucles négatives, dont le rôle est de maintenir l'homéostasie, et de boucles positives, engageant vers la différenciation, peuvent engendrer des comportements extrêmement complexes. Des méthodes ont été développées qui permettent de dériver la dynamique de réseaux complexes à partir de leur structure ; ces réseaux peuvent être traités, sans perte de rigueur, comme l'ensemble de leurs boucles de rétroaction plutôt que comme l'ensemble de tous leurs éléments.

**René Thomas**  
**Denis Thieffry**

**L**es plus élémentaires des processus biologiques font intervenir de multiples variables. Il nous paraît important de démêler lesquelles jouent un rôle crucial. Convaincus que, dans bien des cas, on pourra arriver à une compréhension des aspects essentiels des processus complexes à partir d'un petit nombre de variables, il nous semble licite, si cela paraît nécessaire, d'approfondir ensuite la manière dont les autres éléments du système modulent le détail de son comportement. On se sent conforté dans cette attitude lorsque l'on voit combien complexes peuvent être les comportements de systèmes dont la structure interactive est parfois étonnamment simple.

Dans un système organisé, qu'il s'agisse d'une cellule vivante, d'un

organisme, d'une société, ou parfois simplement d'une machine (un thermostat par exemple), un fonctionnement efficace requiert que le niveau de certaines variables clés puisse être perçu par le système ; il faut aussi que le système puisse tenir compte de ces indications et régler en conséquence le flux de ces éléments.

Pour bon nombre de biologistes\*, régulation signifie généralement homéostasie, c'est-à-dire, en quelques mots, stabilisation autour d'une valeur optimale. Cependant, il existe un autre type de régulation, bien distinct et tout aussi important, qui, à l'opposé de l'homéostasie, force les

\* A l'exception des embryologistes, qui utilisent le plus souvent « régulation » dans une acception différente.

### ADRESSE

R. Thomas : *professeur de génétique*. D. Thieffry : *docteur es sciences, chercheur*. Laboratoire de génétique, université libre de Bruxelles, rue des Chevaux, n° 67, B-1640 Rhode-Saint-Genèse, Belgique.

variables concernées à un choix stable entre deux valeurs extrêmes. Pour des raisons qui apparaîtront plus loin, nous qualifions ce type de régulation d'épigénétique ou différenciatif.

## Régulations homéostatiques

L'homéostasie est familière à tous les physiologistes depuis Claude Bernard (qui la qualifiait d'« élasticité ») ; elle consiste à maintenir le niveau de certaines variables au voisinage d'une valeur optimale. Les mécanismes qui assurent l'homéostasie fonctionnent comme un thermostat. Par temps froid, la température d'une habitation non chauffée se situe à un niveau « plancher », plus ou moins stable ; celle d'une habitation chauffée au maximum, sans régulation, se maintient à un niveau « plafond » ; le but d'un thermostat est de stabiliser la température autour d'une valeur souhaitée, quelque part entre la valeur plancher et la valeur plafond. On arrive à ce résultat par corrections successives ; un dispositif, le relais, interrompt le chauffage dès que la température dépasse la valeur souhaitée, et le rallume dès qu'elle tombe en dessous de cette valeur. Dans ces conditions, la température décrit une oscillation autour de la valeur souhaitée.

Dans la nature, bien des variables (température des vertébrés homéothermes, composition chimique du milieu intérieur, etc.) sont l'objet d'une régulation homéostatique. Selon les cas, la variable est maintenue à une valeur donnée ou oscille autour d'elle (figure 1A). Dans le cas du thermostat, l'oscillation n'est que le reflet de l'imperfection du système (inertie du dispositif), mais dans d'autres, elle est un élément essentiel du processus homéostatique.

## Régulations épigénétiques ou différenciatives

Imaginons maintenant un thermostat dont le relais aurait été trafiqué de telle sorte que le chauffage fonctionne si la température dépasse la valeur choisie et s'arrête si la température est sub-optimale ! Le fonctionnement d'un tel dispositif paraît absurde, et

l'est certainement dans le cadre de l'économie domestique. On voit cependant apparaître un comportement fort intéressant ; la température de l'habitation peut maintenant évoluer, soit vers le niveau plancher, et y rester bloquée, soit vers le niveau plafond, et y rester bloquée, selon que la température initiale était inférieure ou supérieure à la température de référence. On assiste donc à une décision de caractère durable qui dépend d'un détail du passé plutôt que de l'environnement actuel. Comme nous le verrons sous peu, ce type de situation se retrouve dans les processus de différenciation cellulaire.

La différence fondamentale entre les deux types de régulation qui viennent d'être décrits est illustrée par la comparaison entre les figures 1A et 1B. La régulation homéostatique d'une variable la stabilise (avec ou sans oscillation) entre le niveau plancher et le niveau plafond ; la régulation différenciatrice d'une variable l'oblige à un choix stable, soit du niveau plancher, soit du niveau plafond.

## RÉFÉRENCES

1. Delbrück M. Discussion. In: *Unités biologiques douées de continuité génétique*. Lyon : Éditions du Cnrs, 1949 : 33-5.
2. King TJ, Briggs R. Changes in the nuclei of differentiating gastrula cells, as demonstrated by nuclear transplantation. *Proc Natl Acad Sci USA* 1955 ; 41 : 321-5.
3. King TJ, Briggs R. Serial transplantation of embryonic nuclei. *Cold Spring Harbour Symp Quant Biol* 1956 ; 21 : 271-90.
4. Gurdon JB. Adult frogs derived from the nuclei of single somatic cells. *Dev Biol* 1962 ; 4 : 256-73.
5. Novick A, Weiner M. Enzyme induction as an all-or-none phenomenon. *Proc Natl Acad Sci USA* 1957 ; 43 : 553-66.
6. Cohn M, Horibata K. Inhibition by the glucose of the induced synthesis of the  $\beta$ -galactoside-enzyme system of *Escherichia coli*. Analysis of maintenance. *J Bacteriol* 1959 ; 78 : 601-12.
7. Thomas R. Logical analysis of systems comprising feedback loops. *J Theor Biol* 1978 ; 73 : 631-56.
8. Thomas R. On the relation between the logical structure of systems and their ability to generate multiple steady states or sustained oscillations. *Springer Series in Synergics* 1980 ; 9 : 180-93.
9. Thomas R. The role of feedback circuits : positive feedback circuits are a necessary condition for positive real eigenvalues of the jacobian matrix. *Ber Besenges Phys Chem* 1994 ; 98 : 1148-51.
10. Thomas R, D'Ari R. *Biological feedback*. Boca Raton, Florida : CRC Press, 1990 : 316 p.
11. Thomas R. Regulatory networks seen as asynchronous automata : a logical description. *J Theor Biol* 1991 ; 153 : 1-23.
12. Thomas R. Kinetic logic : a boolean approach to the analysis of complex regulatory systems. *Lect Notes Biomath* 1979 ; 29 : 507 p.

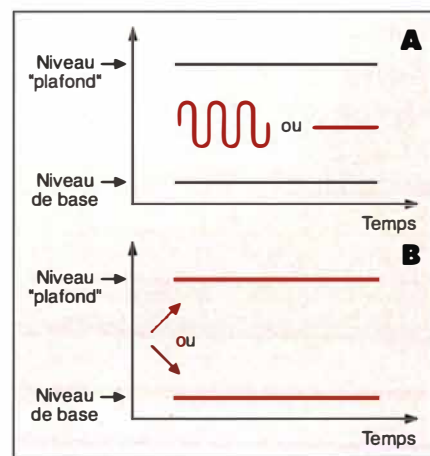


Figure 1. Régulations. A : homéostatique. B : épigénétique.

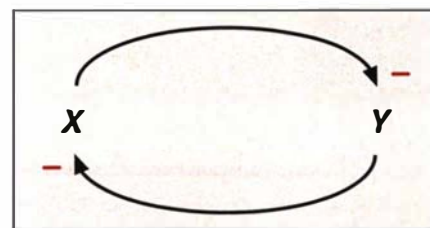


Figure 2. Boucle de régulation (positive) entre deux gènes antagonistes.

A l'aube de ce demi-siècle, la plupart des biologistes étaient convaincus que les différences entre les lignées cellulaires d'un organisme supérieur étaient dues à une répartition différente de particules cytoplasmiques, les « plasmagènes », image chacun d'un gène nucléaire, et « doués de continuité génétique », en d'autres termes, capable de se reproduire au sein même de ce cytoplasme\*. Un symposium mémorable qui s'est tenu en 1948 à l'abbaye de Royaumont marque l'apogée du concept de plasmagène. On trouve cependant dans ses comptes rendus une courte note de Max Delbrück [1] signalant le fait que certains systèmes physiques ou chimiques sont susceptibles de persister dans l'un ou l'autre de deux ou plusieurs états de régime (« multistationnarité »). Delbrück fait remarquer que les différences transmissibles attribuées aux plasmagènes, y compris celles qui interviennent dans la différenciation, peuvent être interprétées en termes de multistationnarité.

Si c'est bien le cas, on est conduit à l'idée que ces différences sont de nature épigénétique, c'est-à-dire que, quoique transmissibles à travers les générations cellulaires, elles ne sont pas liées à des différences au niveau du patrimoine génétique. De fait, il est devenu clair, notamment grâce aux expériences initiées par King et Briggs [2, 3], et il a été irréfutablement démontré par Gurdon [4], que les noyaux de cellules même fortement différenciées ont gardé l'ensemble de l'information génétique de l'individu, et peuvent être substitués au noyau d'un œuf fécondé sans que le développement en soit altéré : la différenciation est de nature essentiellement épigénétique et il ne paraît pas aventureux de dire qu'elle est une modalité biologique du phénomène plus général de multistationnarité.

Pour bien faire sentir la notion de différence épigénétique, nous allons décrire de manière simplifiée un admirable ensemble expérimental dû à Novick et Weiner [5] et à Cohn et

Horibata [6]. Tous les lecteurs savent que le colibacille n'exprime les gènes du catabolisme du lactose qu'en présence d'un inducteur approprié. Point remarquable, il existe une large gamme de concentration extracellulaire d'inducteur (« concentration de maintenance »), insuffisante pour établir l'état induit, mais suffisante pour le maintenir. L'expérience de base, d'une géniale simplicité, consiste à prendre une culture de colibacille non induite, y ajouter une quantité importante d'inducteur, la subdiviser en deux fractions (A et B) que l'on dilue ensuite de manière à faire tomber la concentration finale de l'inducteur au niveau de maintenance. La seule différence entre les fractions A et B réside dans le moment de la dilution : alors que la fraction A est diluée immédiatement, la fraction B est diluée après une dizaine de minutes. Comme on peut s'y attendre, la culture A, qui n'a pas été exposée pendant un temps significatif à la concentration élevée d'inducteur, reste non induite, alors que la culture B est induite. Ce qui, en revanche, n'était pas évident *a priori*, c'est que cet état de choses persiste : après 150 générations (obtenues par dilutions successives toujours dans le même milieu de maintenance), la culture A est toujours non induite, la culture B induite. Ainsi, nous nous trouvons en présence de deux cultures génétiquement identiques, immergées dans un environnement identique, et dont l'une est et reste non induite, l'autre, induite : c'est donc un excellent exemple de différence épigénétique. Dans le cas présent, on sait que cette différence tient au fait que la dilution ait eu lieu immédiatement ou non. Le mécanisme qui explique cette singularité est le suivant. La pénétration de l'inducteur dans la cellule est assurée par une enzyme de surface, la  $\beta$ -galactosidase perméase, dont la synthèse est elle-même dépendante de la présence d'inducteur intracellulaire. Si la concentration externe est très élevée, l'inducteur pénètre même en l'absence de perméase, permettant ainsi la synthèse de perméase. En revanche, dans la gamme de concentration « de maintenance », l'inducteur ne pénètre significativement qu'en présence de perméase. On comprend dès lors que dans la culture A,

dont la concentration d'inducteur a été d'emblée ramenée au niveau de maintenance, l'inducteur ne pénètre pas et ne puisse donc induire la synthèse de perméase. Au contraire, la culture B a été exposée à une concentration élevée d'inducteur pendant un temps suffisant pour que l'induction et, par conséquent, la synthèse de perméase aient eu lieu. Dès lors, les cellules sont induites et le restent même si, comme c'est le cas dans l'expérience, on ramène la concentration externe d'inducteur à son niveau de maintenance. On se trouve donc en présence d'un cercle vicieux : ou il y a déjà de la perméase, et dans ce cas l'inducteur est internalisé et la synthèse de perméase se poursuit, ou il n'y a pas de perméase et, faute d'internalisation de l'inducteur, il n'y a pas de synthèse de perméase. On peut voir facilement que ce schéma simple rend compte des faits expérimentaux en termes de choix entre deux états de régime stables.

### **Boucles de rétroaction**

Il arrive fréquemment que des éléments d'un système interagissent entre eux en formant une boucle. Considérons, par exemple, les produits (x, y et z) de trois gènes de régulation. Si la concentration de x influence le taux de synthèse de y, dont la concentration influence le taux de synthèse de z, dont la concentration influence à son tour le taux de synthèse de x, on voit que le niveau de x influence (indirectement, *via* les éléments y et z) son propre taux d'expression, et donc son propre niveau ultérieur. Cela est vrai également pour y et pour z : le taux de formation ultérieur de chacun de ces éléments est réglé en fonction de sa concentration actuelle. Le petit mécanisme qui vient d'être décrit est une boucle de rétroaction (en anglais, *feedback loop*). On parle de boucle de rétroaction chaque fois que des éléments d'un système forment un circuit orienté, dans le sens utilisé en théorie des graphes. On peut se rendre compte sans difficulté que les boucles de rétroaction peuvent être classées en deux types selon la parité du nombre d'interactions négatives (inhibitrices) qu'elles contiennent.

\* C'est en cela que ces entités hypothétiques se distinguent de l'ARN messager découvert ultérieurement.



## RÉFÉRENCES

13. Thomas R, Thieffry D, Kaufman M. Dynamical behaviour of biological regulatory networks. I. Biological role and logical analysis of feedback loops. *Bull Math Biol* 1995 ; 57 (sous presse).
14. Snoussi EH. Qualitative dynamics of piece-linear differential equations : a discrete mapping approach. *Dyn Stability Syst* 1989 ; 4 : 189-207.
15. Snoussi EH, Thomas R. Logical identification of all steady states : the concept of feedback loop characteristic states. *Bull Math Biol* 1993 ; 55 : 973-91.
16. Nicolis C. A boolean approach to climate dynamics. *Quart J R Met Soc* 1982 ; 108 : 707-15.
17. Ward MD, Moore W H. The growth of the Coercine state. An empirical analysis from 1872-1975, using a Boolean approach. Proceedings of the Annual Meeting of the American Political Science Association, University of Colorado, Institute of behavioral sciences, Boulder, Colorado, 1990.
18. Ciompi L, Ambühl B, Dünki R. Schizophrenie und chaostheorie. Methoden zur untersuchung der nicht-linearen dynamik komplexer psycho-sozio-biologischer systeme. *System Familie* 1992 ; 5 : 133.
19. De Palma A, Stengers I, Pahaut S. Processus de décision et cinétique logique. *RAIRO Rech Oper* 1982 ; 16 : 155.
20. Thieffry D, Thomas R. Dynamical behaviour of biological regulatory networks. II. Immunity control in temperate bacteriophages. *Bull Math Biol* 1995 ; 57 (sous presse).
21. Kaufman M, Urbain J, Thomas R. Towards a logical analysis of the immune response. *J Theor Biol* 1985 ; 114 : 527.
22. Muraille E, Leo O, Kaufman M. The role of antigen presentation in the regulation of class specific (Th1/Th2) immune responses. In: Demongeot J, Auger P, eds. *Mathematics applied to biology and medicine*. Singapore: World Scientific, 1995 (sous presse).
23. Friesen WO, Stent GS. Generation of a locomotory rhythm by a neural network with recurrent cyclic inhibition. *Biol Cybernet* 1977 ; 28 : 40.
24. Thieffry DL, Colet M, Thomas R. Formalization of regulatory networks : a logical method and its automatization. *Math Modelling Sci Computing* 1993 ; 2 : 144-51.

Si une boucle comporte un nombre pair d'interactions négatives, chaque élément exerce sur son propre développement ultérieur un effet positif, et on qualifie la boucle de positive. En revanche, si une boucle comporte un nombre impair d'interactions négatives, chaque élément exerce sur son propre développement ultérieur un effet négatif, et on qualifie la boucle de négative.

Sans entrer dans une énumération fastidieuse, nous prendrons pour exemple une boucle composée de deux éléments, x et y, produits de deux gènes antagonistes. Comme indiqué sur la *figure 2*, le produit x empêche la synthèse du produit y (ou, si l'on préfère, l'« expression » du gène Y) et le produit y empêche la synthèse du produit x (ou, si l'on préfère, l'« expression » du gène X). Si l'on en croit le paragraphe précédent, il s'agit d'une boucle positive, puisqu'elle comprend un nombre pair d'interactions négatives. On peut facilement se rendre compte qu'effectivement chacun des deux éléments de cette boucle exerce bien un effet (indirect) positif sur sa propre évolution ultérieure : x empêche la formation de y qui, s'il était présent, empêcherait la formation de x, et réciproquement.

### Boucles négatives et homéostasie

Lorsqu'on analyse un mécanisme homéostatique, qu'il s'agisse d'un simple thermostat ou du maintien de la concentration intracellulaire d'un métabolite, on s'aperçoit que l'essence même du mécanisme consiste en une boucle de rétroaction négative. Prenons à titre d'exemple la chaîne de réactions enzymatiques responsable de la synthèse d'un acide aminé. Dans bien des cas, le produit final exerce un contrôle négatif sur l'activité ou sur le taux de synthèse d'une ou plusieurs enzymes de la chaîne. Si on symbolise par a, b, c, d les produits successifs de la chaîne métabolique, le circuit complet a une structure du type présenté sur la *figure 3A*. Ce schéma réactionnel signifie que a se transforme en b qui se transforme en c qui se transforme en d, qui, lorsqu'il atteint une concentration suffisante, empêche la synthèse de a, soit en inhibant la réaction, soit en em-

pêchant la synthèse de l'enzyme qui doit la réaliser. Il s'agit bien d'une boucle négative dont le schéma logique est représenté dans la *figure 3B* et comporte un nombre impair d'étapes inhibitrices (une en l'occurrence). Ce qui est commun à tout mécanisme homéostatique, c'est la présence d'une boucle de rétroaction négative : dans une telle boucle, chacun des éléments contrôle son propre niveau ultérieur en « laissant filer » sa production tant qu'elle est suboptimale et en la réprimant dès qu'elle dépasse la valeur fixée. Un tel mécanisme peut conduire, selon les cas, soit à une stabilisation au niveau de la valeur optimale, soit à des oscillations plus ou moins amples autour de cette valeur optimale.

### Boucles positives et différenciation

Revenons à la boucle positive présentée dans la *figure 2*. Si initialement x est présent et y absent, x va empêcher la synthèse de y et, faute de y, x va continuer à être synthétisé et y à ne pas l'être. Inversement, si initialement x est absent et y présent, y restera présent et continuera à être synthétisé, tandis que x restera absent et persistera à ne pas être synthétisé. On voit que ce petit système se stabilisera dans l'un ou l'autre de deux états stables, qu'on peut symboliser (10) et (01) c'est-à-dire, respectivement (x présent, y absent) et (x absent, y présent). Que se passe-t-il si au

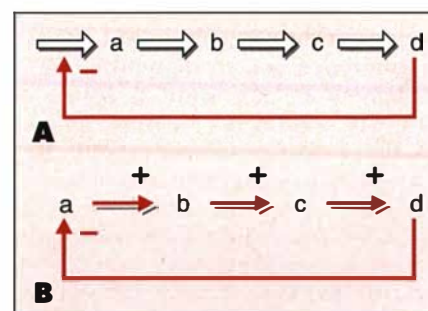


Figure 3. A. Chaîne métabolique à quatre étapes avec répression de la synthèse du premier élément par le produit final. B. Graphe d'interaction correspondant, constitué d'une boucle négative (nombre impair d'interactions négatives).

lieu de partir de l'une des situations précitées, on part de situations où  $x$  et  $y$  sont tous deux absents ou tous deux présents? Dans le premier cas, le système évoluera vers l'état « $x$ » ou vers l'état « $y$ », en fonction des vitesses de synthèse respectives des deux produits; dans le second cas, il évoluera également vers l'état « $x$ » ou vers l'état « $y$ », mais cette fois selon les vitesses de dégradation des deux produits. On pourrait être tenté de croire que ce type de systèmes peut osciller entre les deux états considérés, mais c'est une illusion: en état de régime, soit l'un, soit l'autre s'impose de manière stable.

Ce qui vient d'être exprimé de manière intuitive peut être montré de manière formelle, soit dans le cadre d'une description sous forme d'équations différentielles, soit dans le cadre d'une description «logique». On peut montrer sans difficulté qu'une boucle positive isolée engendre (pour des valeurs appropriées de ses paramètres) un choix entre deux états de régime stables (entre lesquels se trouve un état stationnaire instable).

Dans le cas, examiné plus haut, de deux éléments antagonistes, exclusifs, le choix est: soit l'état « $x$ » noté (10), soit l'état « $y$ » noté (01). Dans le cas d'une boucle positive où les deux éléments,  $x$  et  $y$ , sont interdépendants (s'activent l'un l'autre), il y a aussi deux états stables, mais cette fois (00) ou (11), c'est-à-dire, ni l'un ni l'autre des produits  $x$  et  $y$ , ou tous les deux.

Que se passe-t-il si l'on forme des boucles positives à l'aide d'un nombre croissant d'éléments? Le nombre d'états de régime augmentent-il? Nullement: on peut montrer qu'une boucle positive isolée peut engendrer un choix entre deux états de régimes (plus un état stationnaire instable), quel que soit le nombre de ses éléments.

L'autocatalyse directe (boucle positive à un élément) présente un intérêt particulier. En effet, on trouve un nombre croissant de gènes soumis à un tel type de contrôle positif. Si on la prend dans le sens le plus strict, une telle boucle décrit une situation idéale où le produit d'un gène serait nécessaire (et suffisant) à sa propre synthèse. De deux choses l'une: ou le produit était déjà présent, et dans

ce cas il continuera à être présent et restera présent indéfiniment; ou le produit était initialement absent, et dans ce cas, comme il est nécessaire à sa propre synthèse, il restera indéfiniment absent. Il est clair que nous nous trouvons là dans une situation de cercle vicieux: si le gène était allumé, il le reste, s'il était éteint, il le reste. Mais d'où vient qu'il soit allumé ou éteint? En fait, chaque fois qu'on a trouvé dans la nature un gène soumis à ce type d'«autocontrôle positif», on lui a trouvé associé un «mécanisme d'allumage\*». Le bactériophage lambda, par exemple, possède un gène,  $cl$ , dont le produit est à la fois un répresseur de l'expression de tous les autres gènes du virus et un activateur de sa propre synthèse. Cependant, le gène  $cl$  peut aussi être enclenché par un autre gène viral,  $cII$ . On voit qu'en l'absence des deux produits,  $cl$  ne peut pas s'exprimer. Si le produit du gène  $cII$  apparaît, le gène  $cl$  est enclenché et, comme il s'auto-entretient, il restera enclenché indéfiniment, même si le produit du gène  $cII$  disparaît. C'est en réalité ce qui se passe, car dès que la concentration du produit  $cl$  est suffisante, tous les autres gènes,  $y$  compris  $cII$ , sont bloqués. Le produit du gène  $cII$  fonctionne donc comme un signal transitoire capable d'allumer durablement le gène  $cl$ .

On ne manquera pas de noter que ce type de situation répond à l'une des questions les plus fondamentales de la différenciation cellulaire. Comme nous y avons fait allusion plus haut, la plupart des biologistes admettent aujourd'hui que la différenciation est essentiellement de nature épigénétique. Cela dit, comment comprendre que deux lignées cellulaires génétiquement identiques et plongées dans un même environnement puissent être bloquées de manière durable et transmissible dans des états phénotypiques différents?

Considérons, par exemple, les lignées cellulaires hépatiques et fibro-

blastiques. Alors que les premières synthétisent (entre autres) de la sérum albumine, les fibroblastes ne le font pas. Cette différence se maintient lorsque les deux lignées sont cultivées *in vitro* dans le même milieu. Pourquoi? On peut répondre que, sans doute, le fibroblaste produit un répresseur de la synthèse d'albumine, que ne produit pas la cellule de lignée hépatique et, qu'à l'inverse, la cellule hépatique produit des activateurs non synthétisés par le fibroblaste. Même si ces mécanismes sont bien démontrés, leur principe causal reste à déterminer. Une manière de sortir de cette chaîne explicative consiste à invoquer une causalité circulaire. En effet, il suffit qu'un régulateur de l'expression du gène albumine soit soumis à un autocontrôle positif (direct ou indirect) pour que ce gène régulateur, et par conséquent les gènes qui en dépendent, puisse persister, soit sous l'état allumé, soit sous l'état éteint, et que cet état soit transmissible à travers les générations cellulaires.

Dans cet exemple, comme dans l'hypothèse de Delbrück, dans les expériences de Novick et de Cohn, ou dans la situation du bactériophage lambda, nous trouvons chaque fois une boucle positive au cœur du système. Dès la fin des années 1970, nous avons pensé qu'une boucle positive était le dispositif le plus commode pour permettre à un système d'avoir un comportement multistationnaire [7]. Nous avons ensuite émis l'hypothèse que la présence d'une boucle positive est une condition nécessaire pour avoir multistationnarité [8]. Récemment, cette conjecture a fait l'objet de deux démonstrations formelles (Plathe *et al.*, soumis; Snoussi, en préparation; voir aussi [9]). Dans la mesure où la différenciation est essentiellement une modalité biologique de ce processus de multistationnarité, on peut en inférer que dans la logique sous-jacente à tout processus de différenciation doit se trouver au moins une boucle de rétroaction positive.

## Systèmes à états stationnaires nombreux

La plupart des systèmes multistationnaires étudiés par les théoriciens ont un petit nombre d'états stationnaires

\* Sauf peut-être dans le cas de la scrapie, maladie contagieuse que nous attribuons au recrutement autocatalytique d'une conformation particulière d'une protéine cellulaire. En l'absence du «germe» que constitue cette protéine de conformation anormale, tout se passe bien, mais, si ce produit est apporté de l'extérieur, il enclenche alors la transconformation autocatalytique de la protéine cellulaire.



(souvent trois). Nous sommes loin des 100 à 200 états stationnaires qui seraient requis pour rendre compte des nombreux types cellulaires d'un organisme « supérieur ». Que faut-il pour qu'un système puisse avoir le choix entre de nombreux états stationnaires ? La solution est simple : il faut plusieurs boucles de rétroaction positives. Concrètement, cela veut dire qu'on pourrait rendre compte de nombreux types cellulaires à l'aide d'un nombre modéré de gènes sujets à autorégulation positive, directe ou indirecte. Pour prendre les choses de manière plus quantitative,  $m$  boucles positives peuvent engendrer jusqu'à  $3^m$  états stationnaires, dont  $2^m$  stables\*. Ainsi, il suffirait de 7 gènes autorégulés positivement pour rendre compte de  $2^7 = 128$  types cellulaires.

## Réseaux

Jusqu'ici, nous avons surtout parlé des comportements de boucles de rétroaction individuelles. Il va sans dire qu'on peut rarement se contenter de descriptions aussi simples lorsqu'il s'agit de systèmes biologiques. En effet, on se trouve généralement en présence de réseaux comportant des boucles de rétroaction, non seulement multiples, mais enchevêtrées. L'analyse de tels systèmes requiert impérativement une formalisation. La solution la plus évidente est de tenter une description en termes d'équations différentielles. C'est une démarche que nous utilisons abondamment. Néanmoins, lors de la modélisation différentielle de systèmes biologiques, un problème se pose de manière récurrente. Les interactions à modéliser sont presque toujours non linéaires et il est bien connu que les systèmes d'équations différentielles non linéaires n'ont en général pas de solution analytique. Aussi, une appréciation des comportements possibles du système nécessite-t-elle souvent un traitement numérique assorti

d'un balayage, fastidieux et inesthétique, des valeurs de paramètres. Un outil inestimable est l'analyse de stabilité linéaire, qui permet la détermination de la nature exacte des états stationnaires du système, mais encore faut-il les avoir localisés exactement, ici encore par voie numérique.

C'est l'une des raisons qui ont mené divers auteurs à tenter une analyse plus qualitative, et, plus particulièrement, une analyse logique, caractérisée par le fait que les variables ne peuvent prendre qu'un nombre limité de valeurs. Sous sa forme la plus simple, la description logique n'utilise que deux valeurs : 0 (absent, éteint) et 1 (présent, allumé).

## Motivations et caractéristiques de notre approche logique

En biologie et ailleurs, la plupart des interactions à fonction régulatrice ont une forme sigmoïde. Considérons, par exemple, un gène dont l'expression requiert la présence d'un régulateur positif (figure 4A). En général, le régulateur est pratiquement sans action en dessous d'une valeur seuil ; son effet croît rapidement au voisinage de cette valeur, pour plafonner bientôt. C'est le caractère non linéaire auquel nous avons fait allusion plus haut. Vu la difficulté déjà mentionnée de traiter des systèmes complexes non linéaires, on a tenté (en économie notamment) la « caricature linéaire » (figure 4B) ; elle s'avère désastreuse, sauf au voisinage immédiat des états stationnaires. On peut aussi prendre le contre-pied de cette attitude et raisonner en termes discrets, en remplaçant les sigmoïdes par des « marches d'escalier » ; c'est ce qu'on appelle la description « logique », qui est en quelque sorte « infiniment non linéaire » (figure 4C) ; l'expérience montre que ce type de description, en pratique extrêmement simple, conserve remarquablement l'essentiel des comportements qualitatifs des systèmes non linéaires ; c'est la principale justification de la description logique des systèmes biologiques.

Il serait déplacé de donner ici une description détaillée de la méthode

logique que nous avons élaborée au cours des années écoulées. Nous essaierons cependant de faire ressortir ce qu'elle apporte par rapport aux méthodes existantes. On trouvera tous les détails dans les références [10-13].

Le premier trait saillant de notre méthode logique est son caractère asynchrone (cela signifie simplement que lorsque deux ou plusieurs ordres de changement sont donnés simultanément, ils ne sont pas nécessairement exécutés au même moment). Ce caractère asynchrone est indispensable en biologie : ainsi, il est évident que même si deux gènes sont enclenchés de manière simultanée, leurs produits n'atteindront en général pas en même temps une concentration efficace. En outre, dans la description synchrone qui est généralement utilisée, chaque état du système n'a qu'un seul successeur possible, ce qui empêche tout choix ou différenciation.

Sous sa forme la plus simple, un régulateur est dit « présent » (1) ou « absent » (0) selon que sa concentration excède ou non son seuil d'activité. Mais nous savons qu'un régulateur peut avoir deux sites d'action, et il n'y a pas de raison que les deux seuils impliqués dans ces activités soient égaux ; dans ce cas, nous attribuons à l'élément considéré deux seuils et nous le représentons par une variable logique à trois valeurs (0, 1, 2). Plus généralement, si un élément agit en  $n$  points, nous lui attribuons  $n$  seuils et le décrivons par une variable à  $n+1$  valeurs (0, 1, 2, ...,  $n$ ). Un progrès essentiel a été l'introduction par Snoussi [14] des paramètres logiques. Sans entrer dans les détails, disons que dans la logique classique les termes d'une relation ont tous le même

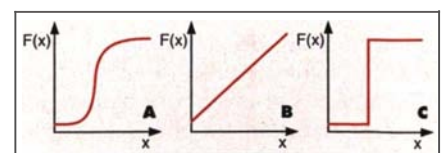


Figure 4. Une interaction sigmoïde (A) ; sa caricature linéaire (B) ; et sa caricature logique (C). En abscisses, la concentration  $x$  d'un régulateur, en ordonnées, le taux d'expression,  $F(x)$ , d'un gène réglé par  $x$ .

\* C'est le cas de  $m$  boucles fonctionnelles disjointes (c'est-à-dire sans éléments communs). Si l'on ajoute des connexions entre les boucles, en général le nombre d'états stationnaires diminue, à moins que les connexions supplémentaires ne fassent apparaître des boucles positives additionnelles.

pois. Les paramètres logiques (discrets, ils peuvent prendre les valeurs 0, 1, 2..., n comme la variable qu'ils concernent) donnent à chaque terme d'une expression un poids approprié. Un raffinement ultérieur nous a permis d'identifier par voie logique tous les états stationnaires qui peuvent être mis en évidence par voie différentielle. Ce raffinement consiste à attribuer explicitement une valeur logique à chaque seuil. L'échelle des valeurs d'une variable logique devient donc 0,  $s^{(1)}$ , 1,  $s^{(2)}$ , 2, ...,  $s^{(n)}$ , n. Au fur et à mesure de ces développements, le hiatus, initialement important, entre les descriptions logique et différentielle, s'est comblé jusqu'à déboucher sur une profonde convergence. Les deux approches sont ainsi devenues concordantes et complémentaires. Si l'on veut qualifier en quelques mots l'originalité de la description logique élaborée, on pourrait dire que dans cette description l'espace des paramètres est subdivisé en un nombre fini de « pavés » dont chacun correspond à une situation qualitativement différente. Il en résulte qu'un système complexe gagne en général à être d'abord démêlé par une analyse logique, quitte à revenir ensuite, s'il y a lieu, à la description différentielle, mais cette fois sans tâtonnements laborieux et inesthétiques.

A ce point de l'exposé, on pourrait se dire, à juste titre, que la beauté de l'analyse logique résidait dans sa simplicité, et que de sophistication en sophistication on a amélioré son acuité, mais aux dépens de l'élégance. Heureusement, un concept nouveau nous a permis de dépasser le stade où la description logique pouvait paraître être devenue compliquée, pour retrouver une pensée unificatrice. Ce concept nouveau est celui d'état caractéristique d'une boucle (ou d'une union de boucles), défini comme le point où la concentration de chacune des variables est égale à la valeur seuil au-delà de laquelle elle agit dans la boucle. Si nous revenons à notre exemple de la *figure 2*, où chaque élément interagit seulement avec l'autre (et nécessite donc chacun un seul seuil, symbolisé  $s^{(1)}$ ), l'état caractéristique de la boucle positive est simplement  $s^{(1)}s^{(1)}$  (c'est-à-dire le lieu où x et y ont chacun leur valeur seuil, quelle qu'elle soit). Cet-

te notion d'état caractéristique a pris une importance particulière avec la démonstration que parmi les états logiques localisés au niveau de seuils, seuls les états caractéristiques de boucles peuvent être stationnaires, et que tout état caractéristique de boucle peut être stationnaire pour des valeurs appropriées de paramètres [15]. Ce théorème simplifie grandement les choses en introduisant une relation simple entre boucles de rétroaction et états stationnaires. En termes plus concrets, au lieu de devoir faire un balayage de tous les états logiques d'un système pour voir lesquels sont stationnaires, il suffit d'identifier les boucles (ou unions de boucles) de rétroaction, de noter pour chacune l'état caractéristique et de voir lesquels de ces états sont stationnaires pour les valeurs de paramètres choisies. Si les valeurs de paramètres ne sont pas choisies d'avance, ce qui est souvent le cas, on cherche quelles sont les contraintes sur les paramètres pour que tel ou tel état caractéristique soit stationnaire. De plus, on peut montrer que, pour qu'une boucle de rétroaction soit fonctionnelle (c'est-à-dire engendre un comportement multistationnaire dans le cas d'une boucle positive, un comportement homéostatique dans le cas d'une boucle négative), il faut que son état caractéristique soit stationnaire (au moins dans le sous-espace des variables de la boucle). Inversement, si l'état caractéristique d'une boucle est stationnaire, celle-ci sera efficiente. Dans notre exemple, la boucle positive engendrera donc un choix possible entre deux états stables si et seulement si les valeurs paramétriques sont telles que l'état caractéristique  $s^{(1)}s^{(1)}$  est stationnaire.

### Applications de la méthode logique

Pour donner une idée du domaine d'application en biologie\*, citons trois applications aussi éloignées que possible.

\* Il est piquant de constater que notre méthode logique a, jusqu'ici, été utilisée dans d'autres disciplines plus qu'en biologie : climatologie [16], politique [17], psychiatrie [18], ou encore prise de décision [19].

Le bactériophage lambda est l'un des systèmes biologiques dont les régulations ont été analysées de la manière la plus approfondie. Comme l'a magistralement montré le regretté André Lwoff, ce virus peut « choisir » entre deux destins. L'un consiste à se multiplier aux dépens de la bactérie-hôte, qui finit par exploser en libérant de nombreux virions. L'autre consiste à insérer son ADN (appelé dès lors « prophage ») dans la continuité du chromosome bactérien ; le génome viral devient ainsi une partie intégrante du patrimoine héréditaire de la bactérie. Malgré la présence de tous les gènes (potentiellement létaux) du virus, cette bactérie dite « lysogène » survit parce que l'un des gènes du virus, *cI*, déjà mentionné plus haut, bloque l'expression de tous les autres gènes du virus. La bactérie est, en outre, immunisée contre toute infection par un virus de même type. L'intérêt majeur de ce système est que le « choix » entre la réponse lytique et la réponse lysogène constitue un mécanisme de différenciation en miniature. De quoi dépend que l'une ou l'autre décision soit prise ? La *figure 5* donne un graphe d'interaction entre cinq gènes régulateurs qui agissent de manière précoce dans le développement du phage. On voit que ce petit réseau, pourtant représenté de manière simplifiée, comporte de nombreuses boucles de rétroaction imbriquées. On peut montrer que certaines de ces boucles jouent un rôle crucial. Deux boucles positives donnent au système sa multistationnarité, c'est-à-dire sa capacité de choix entre deux états de régime, dont l'un mène à la lyse de la bactérie, l'autre à sa perpétuation sous forme de bactérie lysogène. Une boucle négative rend l'expression de plusieurs gènes, dont les gènes impliqués dans la réplication, pratiquement insensible au nombre de copies du génome viral : leur expression, déjà suffisante lorsque la bactérie vient d'être infectée par une seule particule du virus, ne devient pas excessive malgré une réplication intensive du génome viral (jusqu'à environ 200 copies par bactérie). Une analyse détaillée de ces processus peut être trouvée dans [20]. Dans cet exemple, les éléments du



système sont des (macro)molécules, les produits des gènes du virus. Dans le second exemple choisi, les éléments sont des populations de cellules : il s'agit d'une étude des interactions entre différentes populations de cellules (lymphocytes T et B, macrophages, etc.), qui contrôlent le développement d'une réponse immunitaire appropriée lors d'une agression. Les auteurs ont commencé par bâtir un modèle volontairement simplifié à partir d'un petit nombre de données qui paraissaient clairement établies : effet activateur des lymphocytes T<sub>H</sub> sur les lymphocytes T<sub>S</sub> et effet répresseur des lymphocytes T<sub>S</sub> sur les lymphocytes T<sub>H</sub>, comportement autocatalytique de chacun de ces deux types cellulaires, leurs interactions avec les lymphocytes B, etc. [21]. Il est remarquable de constater que le noyau minimal du modèle, consistant en une boucle de rétroaction négative entre les populations cellulaires T<sub>H</sub> et T<sub>S</sub>, flanquée d'une boucle positive sur chacune de ces deux populations, rend compte sans difficulté de la multistationnarité du système : dans des environnements actuels identiques, le système peut persister dans l'un ou l'autre de trois états de régimes stables (« vierge », « immun » et « tolérant »). Lequel de ces états, effectivement établi, dépend du passé du système ? En outre, on peut facilement rendre compte d'un phénomène d'hysteresis, qui consiste en ce qu'un signal transitoire (par exemple l'administration d'un antigène approprié) résulte en un changement d'état durable du système. A la suite de ce premier modèle, d'autres modèles, plus sophistiqués, sont encore en cours d'analyse, parallèlement dans les cadres logique et différentiel [22].

Le troisième exemple diffère des deux autres non seulement par la nature des éléments du système, mais aussi, en partie, par la finalité de la démarche. Une belle série de travaux de Friesen et Stent [23] permet de comprendre certains aspects essentiels de la locomotion de la sangsue à partir de groupes de quatre neurones connectés entre eux, et présents à gauche et à droite de la ligne médiane dans chaque segment de l'animal. Par une subtile dialectique entre expérimentation et

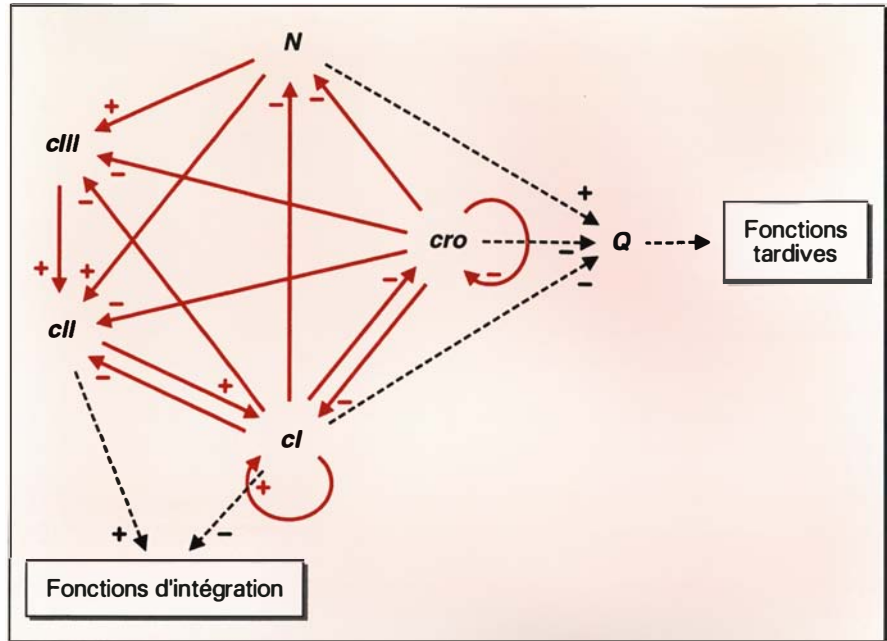


Figure 5. Image simplifiée (!) du réseau de régulation précoce du phage lambda. Les gènes sont indiqués en italique gras ; les flèches rouges indiquent les interactions entre les gènes régulateurs précoces, les flèches pointillées les interactions exercées par les gènes précoces sur d'autres fonctions.

considérations théoriques, ces auteurs sont arrivés à des modèles spécifiant les connexions entre neurones d'un même ganglion et de ganglions appartenant à des segments voisins. Les modèles proposés par Friesen et Stent ne rendaient compte que d'une manière approximative des résultats expérimentaux. Nous avons alors formalisé ces modèles, vérifié avec les auteurs que notre formalisation décrivait correctement les modèles, et analysé leurs prédictions. Il s'est avéré que des modifications mineures des modèles permettaient d'obtenir une concordance parfaite entre leurs prédictions et la séquence expérimentale des allumages et extinctions des neurones.

Nous ne voudrions pas quitter le domaine des applications sans attirer l'attention sur un aspect remarquable et cependant peu connu des boucles de rétroactions positives. Toute boucle positive fonctionnelle constitue en quelque sorte un organe de mémoire. En effet, comme nous l'avons montré plus haut, une boucle positive donne à l'ensemble des variables qui la constituent le choix entre deux niveaux stables ; celle qui

est choisie dépend du passé du système, et il faut une perturbation majeure pour passer de l'une à l'autre. Cette mémoire joue un rôle dans les processus de différenciation, car elle détermine, pour chaque cellule, quels gènes sont bloqués en position active et quels gènes en position inactive, mais le principe peut aussi être appliqué aux réseaux de neurones. En effet, lorsqu'un ensemble de neurones interagissent de manière bouclée, ils forment des boucles de rétroaction qui sont chacune, soit positive\*, soit négative\*, selon la parité du nombre d'interactions négatives. Chaque circuit positif donne à l'ensemble des neurones qui le constituent la possibilité d'un choix entre deux états. A supposer que cet ensemble de neurones soit habituellement dans un état « dormant », l'information liée à l'état alternatif peut être activée par tout signal capable de faire basculer la boucle : le signal a « extrait » la mémoire du petit réseau.

\* Pour autant qu'un neurone donné agisse toujours positivement ou toujours négativement. Dans le cas contraire, un raisonnement de ce type peut encore être tenu, mais devient plus compliqué.



## Conclusions

A travers ce texte, nous nous sommes essentiellement efforcés de faire passer deux messages. D'une part, nous nous demandons si les processus régulateurs, si complexes qu'ils soient, obéissent à des lois fondamentales simples. Il nous semble qu'un pas est franchi dans la compréhension des réseaux complexes de régulation lorsque l'on peut dire que les processus de type homéostatique sont réalisés par des boucles de rétroaction négatives et que les processus de type différenciatif (plus généralement, les processus de multistationnarité) sont réalisés par les boucles positives. On peut dès lors tenter de répondre à des questions du type : quelle est la structure logique minimale requise pour obtenir tel ou tel comportement complexe ? A cette question, il est essentiel d'en adjoindre une autre, tout aussi importante : quels sont les mécanismes effectivement utilisés à cet effet dans la nature ? Par ailleurs, nous pensons avoir indiqué que même un modèle complexe, comportant des boucles de rétroaction multiples et imbriquées, peut être soumis à une analyse qualitative mais néanmoins rigoureuse qui permet de se faire une idée précise des comportements accessibles au système (comportement homéostatique ou différenciatif, nombre et nature des états stationnaires, etc.). Au point où en est arrivée notre analyse, la démarche (actuellement entièrement automatisée [24]) revient simplement à identifier les boucles, noter leurs états caractéristiques et voir dans quelles gammes de valeurs de paramètres ces états caractéristiques sont stationnaires. D'une manière imagée, on peut comparer la dissection d'un réseau régulateur complexe à l'étude d'un mécanisme d'horlogerie : au lieu de considérer individuellement chacune des dents du mécanisme, nous identifions les rouages et la manière dont ils interagissent, sans perte de rigueur ■

## Summary

### Feedback loops, the wheels of biological regulatory networks

Organized systems such as living organisms or cells require that the level of crucial elements somehow be evaluated and taken into account to determine their future rate of production. This regulation is effected by feedback loops, which are oriented circuits of interactions. There are two types of feedback loops, negative and positive, depending on the parity of the number of negative interactions. Their role are radically different : negative loops promote homeostasis while positive loops permit multistationarity, with its biological modality, differentiation. Appropriate combinations of positive and negative feedback loops may generate extremely complex behaviour. Real regulatory systems usually comprise several feedback loops which may be intertwined in complex ways. A major aspect of our work has been the development of methods which permit to derive the dynamics of such complex networks from their structure. Recent progress has shown that a networks can be treated (without loss of rigor) as the set of its interacting loops rather than as the set of all its interacting elements. This is somewhat like considering the wheels of a clock rather than the individual teeth which compose the wheels. This approach is illustrated by various concrete biological examples.

### TIRÉS À PART

R. Thomas.