

## ADN, protéines et transcription

L'activation d'un gène, c'est-à-dire le déclenchement de sa transcription, nécessite la constitution d'un complexe nucléoprotéique comportant l'ARN polymérase et de nombreuses autres protéines. La première étape de la constitution de ce complexe, dénommé « complexe d'initiation de la transcription », est probablement la fixation sur le promoteur de protéines reconnaissant les motifs « TATA box » ou d'autres séquences situées à proximité immédiate du site d'initiation de la transcription [1, 2]. Puis, correctement positionnée grâce à la

mise en place préalable des facteurs précédents, l'ARN polymérase se fixe, suivie d'autres protéines qui pourraient être requises pour l'activation de cette enzyme [1, 2]. Certains des partenaires du complexe d'initiation sont des protéines reconnaissant des séquences particulières d'ADN, d'autres se lient à ces protéines fixées à l'ADN plutôt qu'à l'ADN lui-même (figure 1). La formation ou l'activation de ce complexe sont soumises à l'influence régulatrice d'autres protéines se liant soit à proximité immédiate, soit à distance sur un

élément *enhancer* (lexique *m/s*, supplément au n° 7, vol. 1, p. 3 et *m/s* n° 7, vol. 2, p. 410). Dans ce dernier cas (figure 2), il semble que les deux complexes nucléoprotéiques (complexe d'initiation et complexe *enhancer*) puissent néanmoins interagir par contact direct, via la constitution de boucles d'ADN [3]. L'interaction entre des protéines liées à l'un et à l'autre de ces complexes, éventuellement entre molécules de la même protéine s'il existe un site de fixation équivalent au niveau du promoteur et du *enhancer*, peut constituer le mécanisme stabilisant

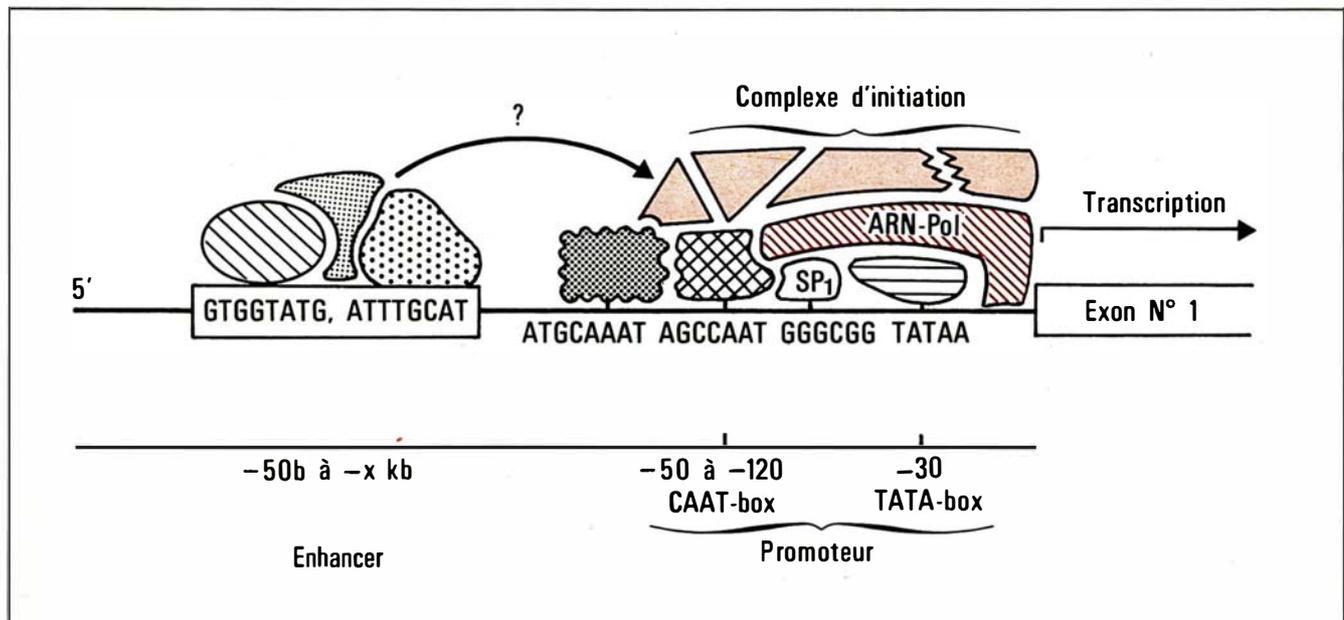


Figure 1. **Complexe d'initiation de la transcription et complexe enhancer.** Les distances sont exprimées, sous le schéma, en nombre de bases en amont du site d'initiation de la transcription. Le enhancer est situé à une distance très variable du site d'initiation, de quelques bases à plusieurs kilobases. Outre la « TATA box » (TATAA), la « CAAT box » (AGCCAAT), d'autres éléments fixant des facteurs transcriptionnels ont été représentés : le facteur SP1 se fixe sur une « GC box » (GGGCGG) ; le facteur NF-A reconnaît la séquence ATGCAAAT du promoteur et son inverse ATTTGCAT localisé au niveau du enhancer ; la séquence GTGGTATG du enhancer est un motif (le core enhancer) très fréquemment trouvé au niveau de ce type d'élément. Les mécanismes de l'interaction entre le complexe enhancer et le complexe d'initiation restent hypothétiques.

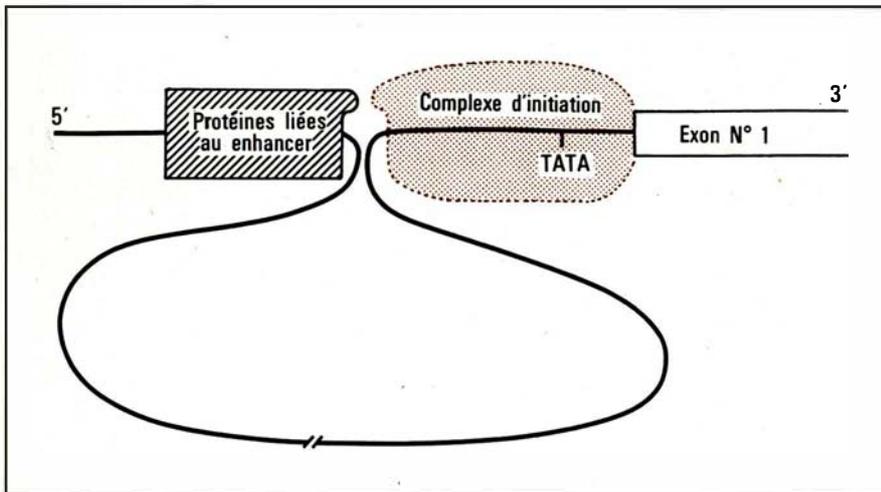


Figure 2. **Schéma de l'interaction directe** entre les éléments du complexe enhancer et ceux du complexe d'initiation.

ces boucles, comme nous l'avons rapporté pour le répresseur du phage lambda (*m/s* n° 7, vol. 3, p. 428). Une telle hypothèse constituerait une explication rationnelle de l'efficacité particulière de l'enhancer du gène des chaînes lourdes d'immunoglobuline sur les promoteurs situés en amont des segments codant pour les parties variables de ces gènes. Ces deux éléments, promoteurs et enhancers, contiennent en effet, dans l'un ou l'autre sens (*figure 1*), un motif dénommé octamère qui lie une protéine spécifique NF-A (il existerait une forme ubiquitaire de cette protéine [NF-A<sub>1</sub> ou IgNF-A<sub>1</sub>] et une forme spécifique des lymphocytes T [IgNF-A<sub>2</sub>]). Deux molécules de NF-A, fixée l'une sur le promoteur et l'autre sur le enhancer pourraient ainsi conduire à la formation de boucles d'ADN rapprochant les complexes enhancer et d'initiation [4, 6]. Une autre question concerne les relations existant entre une protéine fixée à l'ADN et la stimulation (ou l'inhibition) transcriptionnelle qui s'ensuit. Dans un certain nombre de cas, particulièrement bien illustrés par les protéines GAL-4 et GCN-4 de levure, on a pu démontrer que les domaines protéiques responsables de la fixation

à l'ADN et de l'activation transcriptionnelle étaient différents [7-9]. A condition que la séquence reconnue soit insérée à proximité du promoteur contrôlé, il est possible de remplacer GAL-4 par une protéine hybride dont la partie N-terminale est constituée du domaine de liaison à l'ADN d'un répresseur d'*E. coli*, seule la partie C-terminale appartenant à GAL-4. Des expériences de mutagenèse dirigée ont permis de démontrer que les contraintes de séquence pour que les protéines GAL-4 et GCN-4 restent activatrices étaient extrêmement faibles, nécessitant seulement la présence d'une séquence composée d'acides aminés dans la moitié C-terminale des molécules. Cette grande variabilité des structures peptidiques susceptible d'assurer l'activation transcriptionnelle suggère que le mécanisme n'est pas ici de nature catalytique (action de topo-isomérase, nucléase, méthylase, etc.) mais implique bien, plutôt, une interaction physique directe entre macromolécules via des liaisons ioniques, hydrogène ou hydrophobes (*figure 3, partie 2*). Dans le cas des récepteurs des hormones stéroïdes (*m/s* n° 3, vol. 3, p. 172), plusieurs auteurs ont démontré qu'il était

possible de déléter largement les régions N-terminales (domaines A et B) et C-terminales (régions D et E) de la protéine. Le récepteur tronqué, réduit à moins de 100 résidus codés par la région C, a alors une activité « constitutive », c'est-à-dire indépendante de l'hormone [10, 11]. Ces résultats indiquent que la région E correspond probablement à un domaine inhibiteur de la fixation du récepteur à l'ADN, l'inhibition étant levée lorsque l'hormone se fixe à son site situé dans ce domaine E. Les domaines A et B ont un rôle encore mal compris ; ils pourraient augmenter la spécificité de la fixation du récepteur aux séquences d'ADN responsables de la réponse transcriptionnelle à l'hormone, ou encore participer à l'activation transcriptionnelle. Quoi qu'il en soit, il n'est pour l'instant pas possible de séparer fonctionnellement les régions du récepteur responsables de la fixation à l'ADN de celles indispensables à la stimulation de la transcription. Une tierce protéine, différente du récepteur, pourrait constituer un intermédiaire transmettant l'information activatrice depuis la région du récepteur interagissant avec l'ADN jusqu'au complexe d'initiation (*figure 3, partie 2*). Un autre mécanisme, d'ailleurs non exclusif du précédent, pourrait impliquer des modifications de la torsion et de la courbure de l'ADN secondaires à la fixation d'une ou de plusieurs molécules de récepteurs. Le caractère coopératif de la réponse transcriptionnelle, lorsque plusieurs sites sont occupés [12], pourrait militer en faveur de l'intervention d'un tel phénomène, puisque de telles fixations multiples sont susceptibles d'augmenter les contraintes induites dans la double hélice (*m/s* n° 7, vol. 3, p. 428). Dans cette hypothèse, schématisée dans la *figure 3, partie 3*, la stimulation de la transcription serait transmise à distance, via ces modifications topologiques de l'ADN. La période actuelle connaît une

activité fébrile de purification et de caractérisation des facteurs transcriptionnels spécifiques de motifs particuliers d'ADN, reconnaissant des séquences localisées dans les promoteurs ou les *enhancers* et parfois dans l'un et l'autre. Le groupe de Tjian à Berkeley (Californie, USA) est probablement le plus actif dans ce domaine ; il a notamment déjà cloné l'ADN complémentaire codant pour le facteur SP1 qui se fixe sur des régions riches en GC

des promoteurs (« GC boxes », figure 1) ([13] et résultats non publiés). Au moins six autres facteurs ont été purifiés totalement par ce groupe ou par d'autres et il est probable que leurs gènes spécifiques seront bientôt isolés. Tous ces résultats permettent enfin d'entrevoir le mécanisme du contrôle de l'expression des gènes au cours de la différenciation, ainsi que nous l'envisagerons dans une prochaine nouvelle.

Axel Kahn

## RÉFÉRENCES

1. Reinberg D, Rooder RG. Factors involved in specific transcription by mammalian RNA polymerase II. *J Biol Chem* 1987 ; 262 : 3310-37.
2. Hawley DK, Rooder RG. Functional steps in transcription initiation and reinitiation from the major late promoter in a HeLa nuclear extract. *J Biol Chem* 1987 ; 262 : 3452-61.
3. Ptashne M. Gene regulation by proteins acting nearby and at a distance. *Nature* 1986 ; 322 : 697-701.
4. Staudt LM, Singh H, Sen R, Wirth T, Sharp PA, Baltimore D. A lymphoid specific protein binding to the octamer motif of immunoglobulin genes. *Nature* 1986 ; 323 : 640-3.
5. Doyen N, Leblond-Francmillard M, Holm I, Dreyfus M, Rougeon F. Analysis of promoter and enhancer cell type specificity and the regulation of immunoglobulin gene expression. *Gene* 1986 ; 50 : 321-31.
6. Dreyfus M, Doyen N, Rougeon F. The conserved decanucleotide from the immunoglobulin heavy chain promoter induces a very high transcriptional activity in B-cells when introduced into an heterologous promoter. *EMBO J* 1987 ; 6 : 1685-90.
7. Brent R, Ptashne M. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 1985 ; 43 : 729-36.
8. Ma J, Ptashne M. Deletion analysis of GAL-4 define two transcriptional activating segments. *Cell* 1987 ; 48 : 847-53.
9. Struhl K. Promoters, activator proteins and the mechanism of transcriptional initiation in yeast. *Cell* 1987 ; 49 : 295-7.
10. Hollenberg SM, Giguere V, Segui P, Evans RM. Co-localization of DNA binding and transcriptional activation functions in the human glucocorticoid receptor. *Cell* 1987 ; 49 : 39-46.
11. Miesfeld R, Godowski PJ, Maler BA, Yamamoto K. Glucocorticoid receptor mutants that define a small region sufficient for enhancer activation. *Science* 1987 ; 236 : 423-7.
12. Jantzen HM, Strähle U, Gloss B, et al. Cooperativity of glucocorticoid response element located far upstream of the tyrosine aminotransferase gene. *Cell* 1987 ; 49 : 29-38.
13. Briggs MR, Kadonaga JT, Bell SP, Tjian R. Purification and biochemical characterization of the promoter specific transcription factor SP1. *Science* 1986 ; 234 : 47-52.

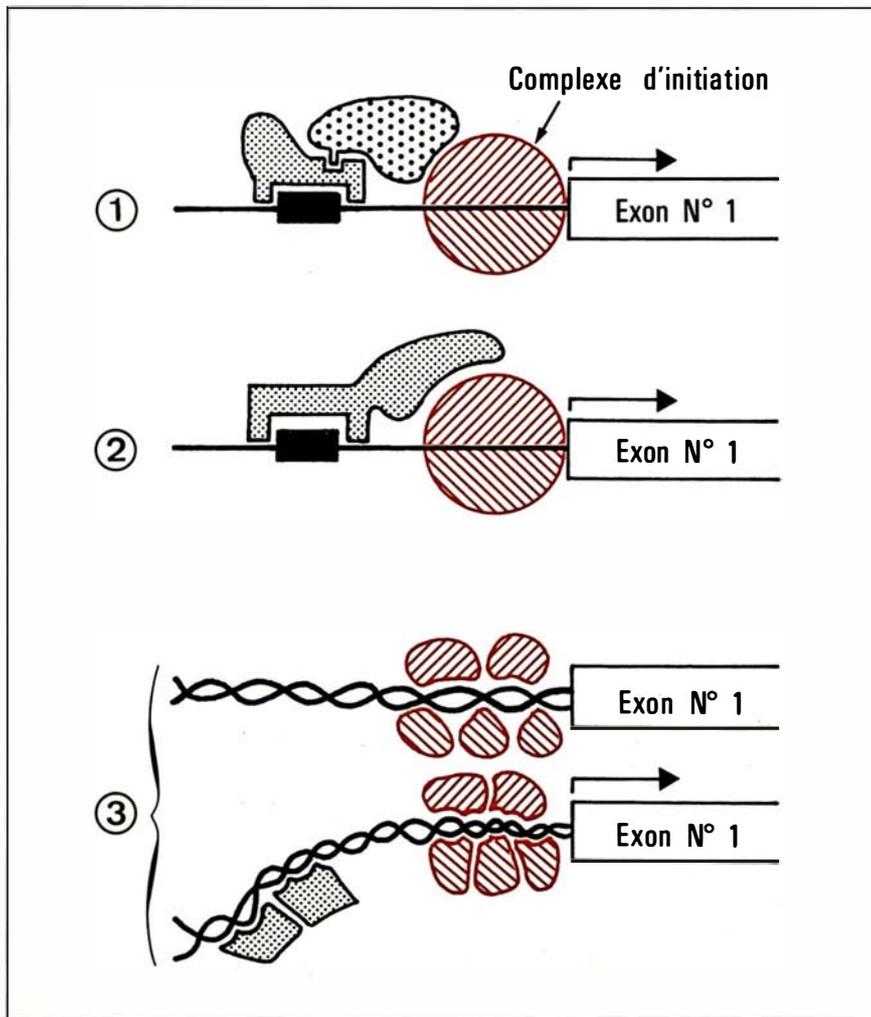


Figure 3. **Modèle de la transmission du message d'activation transcriptionnelle entre une protéine liée à l'ADN et le complexe d'initiation.** (1) transmission par l'intermédiaire d'une tierce protéine. (2) facteur transcriptionnel bifonctionnel, un domaine du facteur assurant la liaison à l'ADN alors qu'un autre domaine est responsable de l'effet transcriptionnel. (3) modification topologique de l'ADN (courbure, modification de l'hélicité) induite par la fixation d'une protéine. Cette modification de la courbure et de l'hélicité transmet, à distance, le message au complexe d'initiation.