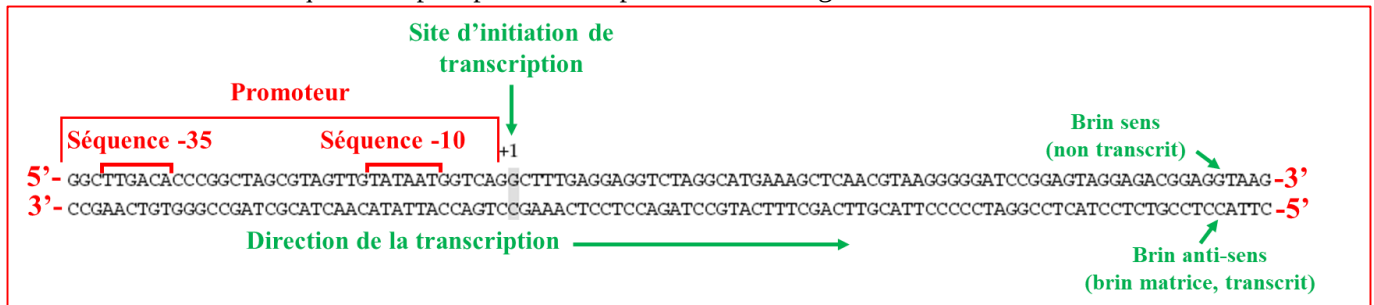


# Corrigé type du TD N°02 de Biologie Moléculaire

## Exercice 1 :

Soit la séquence d'ADN suivante provenant du génome d'*E. coli* et qui code pour un court polypeptide. Le site d'initiation de transcription y est indiqué.

- a. Déterminez la séquence la plus probable du promoteur de ce gène.



- b. Donnez la séquence de l'ARNm transcrit de ce gène.

5'-GCUUUGAGGAGGUCUAGGCAUGAAAGCUAACGUAAGGGGGAUCCGGAGUAGGAGACGGAGGUAAG-3'

- c. Donnez la séquence du polypeptide produit de la traduction de cet ARNm.

5'-GCUUUGAGGAGGUCUAGGCAUGAAAGCUAACGUAAGGGGGAUCCGGAGUAGGAGACGGAGGUAAG-3'

f-Met Lys Ala Gln Arg Lys Gly Asp Pro Glu STOP

Donc, la séquence du polypeptide est la suivante :

NH<sub>2</sub> – f-Met – Lys – Ala – Gln – Arg – Lys – Gly – Asp – Pro – Glu –COOH

- d. Identifiez la position dans la séquence d'ADN au niveau de laquelle une substitution d'une seule paire de base ou une mutation par décalage du cadre de lecture donneraient un peptide dont la séquence est identique mais à demi-taille par rapport au peptide du type sauvage.

**Détermination de la mutation qui donne un peptide identique au sauvage mais à demi-taille :**

- La longueur du peptide sauvage est de 10 aa ; donc le peptide muté devrait être de 5 aa.
- => il faut donc que la mutation ponctuelle touche le sixième codon de la séquence originale et le transforme en codon STOP.
- Le 6<sup>ème</sup> codon est (AAG) correspondant à la lysine, il doit se transformer en codon STOP (UAA, UGA ou UAG). Ceci peut arriver par :
  - 1) Mutation par substitution : AAG => UAG (il s'agit d'une transition où A est muté en U). (puisque la mutation se fait dans l'ADN, A est remplacé par T).
  - 2) Mutation par insertion de U juste avant le codon AAG => UAAG. (dans l'ADN, insertion de T).

**Cette mutation entraîne un décalage du cadre de lecture.**

- e. Identifiez, dans la séquence nucléotidique, la position où une mutation ponctuelle (insertion, délétion ou substitution) pourrait donner un polypeptide dont la séquence est identique à celle du peptide du type sauvage mais ayant quatre résidus arginines additionnels à son extrémité C-terminale.

**Pour que la mutation donne un peptide ayant 4 acides aminés en plus (4 arginines), il faut que la mutation transforme le codon STOP (UAG) de la séquence sauvage en codon codant pour l'arginine, et ce codon doit être suivi de 3 autres codons pour l'arginine, suivis par un codon STOP :**

- Selon le tableau du code génétique, nous avons 6 codons pour l'arginine : CGU, CGC, CGA, CGG, AGA et AGG
- Si une mutation par délétion touche le U du codon STOP (UGA), on aura la séquence suivante d'ARNm :

5'-GCUUUGAGGAGGUCUAGGCAUGAAAGCUAACGUAAGGGGGAUCCGGAGXAGGAGACGGAGGUAAAG-3'

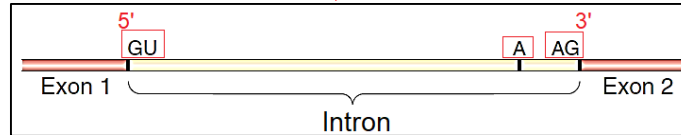
f-Met Lys Ala Gln Arg Lys Gly Asp Pro Glu Arg Arg Arg Arg STOP

## Exercice 2 :

Dans un tube à essai, on met tous les composés nécessaires à l'excision-épissage des ARNm eucaryotes. Une molécule d'ARN contenant un intron et deux exons est ajoutée à la solution, l'intron est éliminé sous forme de lasso et les deux exons sont soudés. Supposant que la molécule d'ARN ajoutée à la solution porte les mutations suivantes, quels produits intermédiaires seront accumulés dans le tube dans chaque cas. Expliquez.

**Pour que l'épissage se produise, il faut que deux coupures se réalisent :**

- la 1<sup>ère</sup> coupure dans la limite entre le 1<sup>er</sup> exon et l'intron. Ceci implique la séquence **GU** au début de l'intron (c'est-à-dire en 5' de l'intron) et la **A** de branchement situé à l'intérieur de l'intron.
- la 2<sup>ème</sup> coupure dans la limite entre l'intron et le 2<sup>ème</sup> exon. Ceci implique la séquence **AG** à la fin de l'intron (c'est-à-dire en 3' de l'intron)



- a. Délétion de GT au niveau du site 5' d'épissage

**La 1<sup>ère</sup> coupure ne se produit pas et on aura dans le milieu un ARN non épissé :**



- b. Délétion du A au niveau du site de branchement

**La 1<sup>ère</sup> coupure ne se produit pas et on aura dans le milieu un ARN non épissé :**



- c. Délétion de AG au niveau du site 3' d'épissage

**La 1<sup>ère</sup> coupure se produit normalement avec formation de lasso, par contre, la 2<sup>ème</sup> coupure ne se produit pas et on aura dans le milieu deux fragments ; l'exon 1 seul et l'exon 2 attaché à l'intron en forme de lasso :**



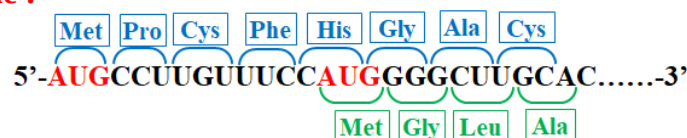
## Exercice 3 :

Des études ont montré que le génome du bactériophage ΦX174 peut coder un nombre de polypeptides plus grand que prévu. Une des raisons de ce nombre élevé est que quelques gènes sont chevauchés ; une séquence d'un gène B par exemple peut être entièrement localisée dans la séquence codante d'un gène A. cependant, la séquence polypeptidiques produites des deux gènes est entièrement différentes. A votre avis, comment ceci est-il rendu possible ?

**La plupart des gènes codent pour une seule protéine (1 gène = 1 protéine), mais on constate que le nombre des protéines dans la cellule est plus grand que celui du nombre des gènes. Donc certains gènes codent pour plus d'une protéine. Chez les eucaryotes, ceci est assuré par deux phénomènes : (1) l'épissage alternatif et (2) la présence de sites multiples de clivage/polyadénylation dans un gène donné (idée développée dans le cours).**

Chez certains virus, on a également constaté que le nombre de protéines est plus grand que le nombre de gènes, comme le cas du bactériophage ΦX174, mais le phénomène qui conduit à ça est complètement différent de celui rencontré chez les eucaryotes. Il s'agit au fait des gènes chevauchés. Le but de cet exercice est de développer l'idée des gènes chevauchés qu'on rencontre dans certains virus. Les gènes chevauchés ont une séquence qui contient plus d'un codon initiateur (ATG) (qui se transcrit et AUG dans l'ARNm). Ces codons (ATG) ne sont pas dans le même cadre de lecture.

Comme dans cet exemple :



**Donc lors de la traduction de cet ARNm, le ribosome peut commencer la traduction au niveau du premier AUG ou au niveau du deuxième AUG, et ça donne deux peptides complètement différents car les deux AUG ne sont pas dans le même cadre de lecture.**

#### Exercice 4 :

Des souches d'*E. coli* cultivées sur glycérol (absence de glucose), en présence et en absence d'IPTG (l'inducteur) sont représentées dans le tableau ci-dessous.

1. Compléter ce tableau en mettant les signes (+) et (-) pour désigner l'expression ou non de la  $\beta$ -galactosidase et la perméase pour chaque souche (suivre l'exemple de la souche 1).

| Souche | Génotype                            | $\beta$ -galactosidase |        | Perméase |        |
|--------|-------------------------------------|------------------------|--------|----------|--------|
|        |                                     | - IPTG                 | + IPTG | - IPTG   | + IPTG |
| 1      | $I^+ O^+ Z^+ Y^+$                   | -                      | +      | -        | +      |
| 2      | $I^+ O^c Z^+ Y^+$                   | +                      | +      | +        | +      |
| 3      | $I^+ O^c Z^+ Y^+ / I^+ O^+ Z^+ Y^+$ | +                      | +      | +        | +      |
| 4      | $I^+ O^c Z^- Y^+$                   | -                      | -      | +        | +      |
| 5      | $I^+ O^c Z^- Y^+ / I^+ O^+ Z^+ Y^+$ | -                      | +      | +        | +      |
| 6      | $I^+ O^c Z^+ Y^-$                   | +                      | +      | -        | -      |
| 7      | $I^+ O^c Z^+ Y^- / I^+ O^+ Z^+ Y^+$ | +                      | +      | -        | +      |
| 8      | $I^- O^+ Z^+ Y^+$                   | +                      | +      | +        | +      |
| 9      | $I^- O^+ Z^+ Y^+ / I^+ O^+ Z^+ Y^+$ | -                      | +      | -        | +      |
| 10     | $I^- O^+ Z^+ Y^+ / I^+ O^+ Z^- Y^+$ | -                      | +      | -        | +      |
| 11     | $I^- O^+ Z^+ Y^+ / I^+ O^+ Z^+ Y^-$ | -                      | +      | -        | +      |

Dans cet exercice, on considère la régulation de l'expression des gènes de l'opéron lactose chez *E. coli* dans un contexte expérimental.

- $I^+$ ,  $O^+$ ,  $Z^+$  et  $Y^+$  désignent les gènes sauvages du répresseur Lac, de l'opérateur, de la  $\beta$ -galactosidase et de la perméases respectivement.
- Le signe (-) désigne un gène muté dont le produit est inactif
- $I^-$  désigne une mutation au niveau du gène du répresseur Lac
- $O^c$  désigne une mutation au niveau de la séquence de l'opérateur (l'opérateur muté n'est plus capable de lier le répresseur Lac, ce qui résulte en l'expression constitutive des gènes de l'opéron même en présence d'un répresseur actif, c'est pour cette raison que cette mutation est désignée par la lettre c)
- $Z^-$  désigne une mutation au niveau du gène de la  $\beta$ -galactosidase
- $Y^-$  désigne une mutation au niveau du gène de la Perméase

L'IPTG a une structure similaire au lactose, c'est un inducteur artificiel de l'opéron lactose.