**2eme année pharmacie Année universitaire 2019/2020**

**TD Génétique : série 10 génétique moléculaire.**

**Solution**

**Exercice 1**

Soit la séquence d'ADN suivante:

**5'- ATTTACGGGCCTTAATGGCATAACCGCCTAATGGTTAACCGCTAGCGCG - 3'**

* Donner la séquence de l'ADN double brin correspondant.
* A quelle condition (nécessaire) cet ADN double brin serait transcrit *in vivo* ?
* Donner la séquence du transcrit éventuelle.

**Solution 1**

* La séquence de l'ADN double brin correspondant est complémentaire est antiparallèle :

5'ATTTACGGGCCTTAATGGCATAACCGCCTAATGGTTAACCGCTAGCGCG 3'

3’ TAAATGCCCGGAATTACCGTATTGGCGGATTACCAATTGGCGATCGCGC 5’

* Cet ADN double brin serait transcrit *in vivo* s’il existe un promoteur en amont. Bien sûre, il ne faut pas négliger la présence de la machinerie de transcription dont principalement l’ARN pol et tous ses cofacteurs et des NTP.
* Séquence du transcrit éventuelle :

Par convention : Le brin d’ADN transcrit est le **brin antisens**.

Le brin d’ADN non transcrit est le **brin sens**.

Pour avoir un transcrit nous allons donc d’abord supposer la présence d’un promoteur. Ainsi, nous avons deux possibilités :

**Possibilité 1:**



La séquence du transcrit:

**5' AUUUACGGGCCUUAAUGGCAUAACCGCCUAAUGGUUAACCGCUAGCGCG 3'**

**Possibilité 2:**

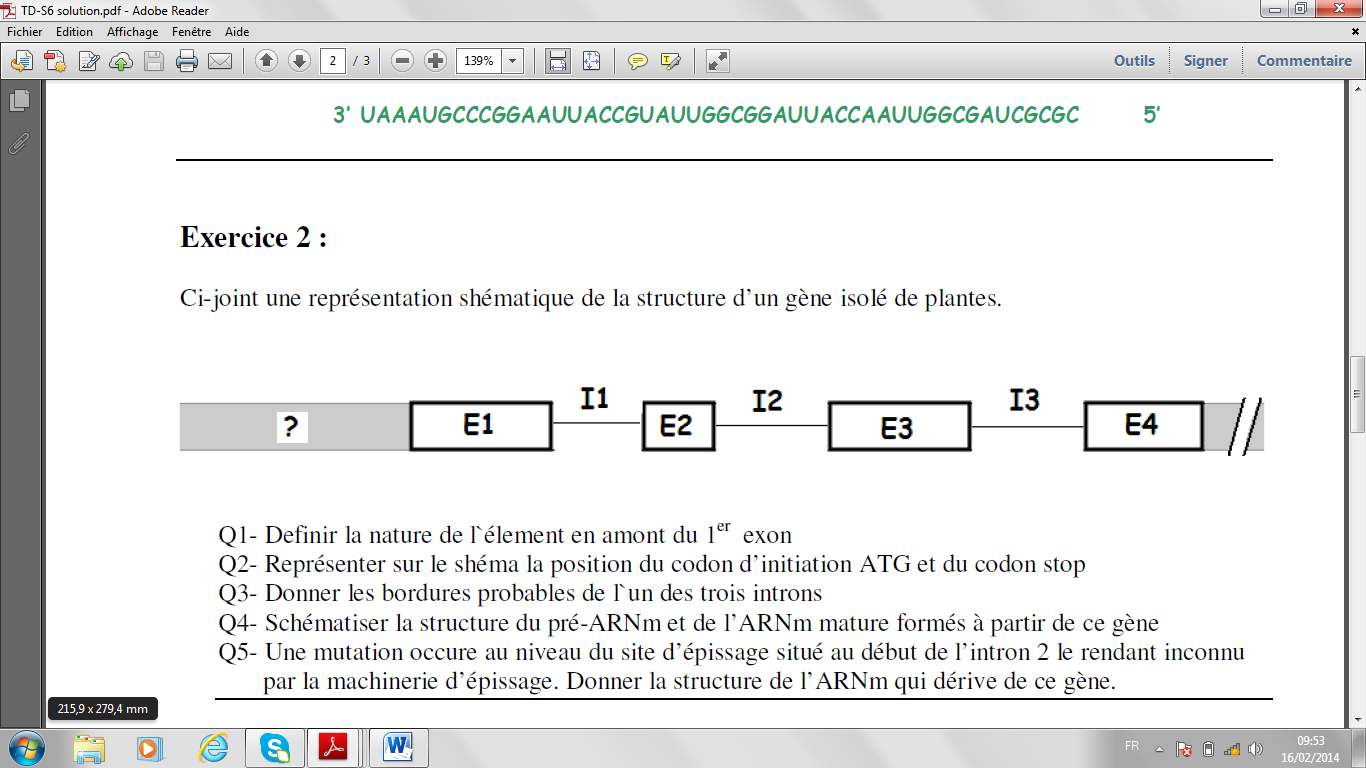
****

La séquence du transcrit:

**3’ UAAAUGCCCGGAAUUACCGUAUUGGCGGAUUACCAAUUGGCGAUCGCGC 5’**

**Exercice 2**

Ci-joint une représentation schématique de la structure d’un gène isolé de plantes.



* Définir la nature de l’élément en amont du 1er exon
* Représenter sur le schéma la position du codon d’initiation ATG et du codon stop. Est-ce les seules positions possibles ? expliquer.
* Schématiser la structure du pré-ARNm (transcrit primaire) et de l’ARNm mature formés à partir de ce gène

Une mutation survient au niveau du site d’épissage situé au début de l’intron 2 le rendant inconnu par la machinerie d’épissage (splicéosome).

* Donner la structure de l’ARNm qui dérive de ce gène.

**Solution 2**

* **Définir la nature de l’élément en amont du 1er exon c’est le p**romoteur. C’est une séquence reconnue par la machinerie de transcription permettant sa fixation et l’initiation de la transcription. Nous citons à titre non exhaustif principalement le motif TATAAA qui définit la TATA box en position -30 à -25 NT, et le motif Py Py A N (T/A) Py Py qui définit le Inr en position -6 à +4 NT.
* **Représenter sur le schéma la position du codon d’initiation ATG et du codon stop**



Cette position théorique simpliste n’est pas la seule possible, au contraire, dans la plupart des gènes l’ATG peut être rencontré dans ou au-delà du premier exon. Et le codon STOP bien avant la fin du dernier exon.

* **Schématiser la structure du pré-ARNm et de l’ARNm mature formés à partir de**

**ce gène**

**Pré- ARNm appelé aussi transcrit primaire :**

Il contient toute la séquence transcrite depuis l’Inr jusqu’au signal de la fin de transcription.



**ARNm mature :**

Il ne contient que les séquence des exons, plus les modifications apportées lors de la maturation : la coiffe, la queue poly A, et d’autres modifications liées à l’édition si elle existe pour ce transcrit.



* **Donner la structure de l’ARNm qui dérive du gène muté :**

Dés lors que le site d’épissage devient inconnu à la machinerie d’épissage (splicéosome) ; le lasso n’est plus formé pour l’intron. Ceci conduit à l’inclusion de la séquence intronique dans le mRNA mature. Notez qu’en plus de ce type de mutations ; il existe des mutations qui créent de nouveaux sites de splicing au niveau des exon ou des introns.



**Exercice 3**

L’ADN génomique présenté ci-dessous contient la totalité de la séquence d’un gène codant une protéine. Les segments nucléotidiques soulignés correspondent aux séquences d’ADN de ce gène retrouvées dans l’ARNm.

1. Quels sont les composants moléculaires nécessaires à la transcription ?
2. Comment se fait l’initiation de la transcription ?
3. Quelle partie de cette séquence peut participer à la régulation de l’expression d’un gène par un signal hormonal ?
4. Parmi les « motifs » de cette séquence marqués en gras ( ggccaatct / tatata / gt / ag / aataaa / tatgtttg ) :

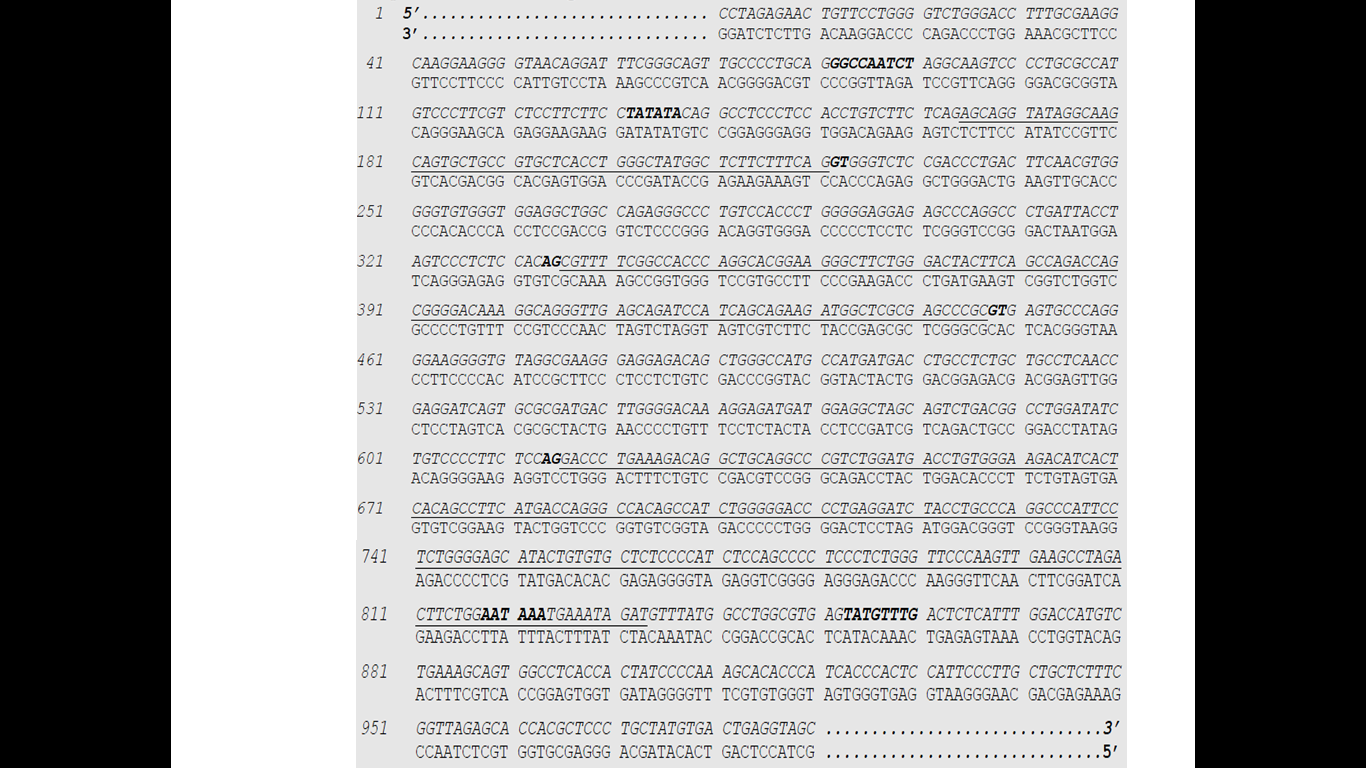
- Quels sont ceux qui définissent l’orientation de la transcription ?

- En quoi définissent-ils le brin sens et le brin matrice ?

- Quels sont ceux qui définissent le début et la fin de la transcription ?

- Quels sont ceux qui interviennent lors de la maturation du transcrit primaire ?

1. A quoi correspondent les segments soulignés dans cette séquence ?
2. Quels « motifs » interviennent dans le processus qui permet de les réunir ?
3. Quel est le rôle du triplet de nucléotides en 206-208
4. Le nucléotide en position 391 est un C qui fait partie de la séquence traduite : quelle est sa position dans le cadre de lecture : N1, N2 ou N3 ?
5. Quelle sera la longueur après transcription et maturation (comprenant l'addition de 1000 AMP du côté 3'-OH) de l'ARN messager ?
6. De combien d'acides aminés se compose votre protéine ?



**Solution 3**

1. Les composants moléculaires nécessaires à la transcription :

Chromatine accessible à la transcription (ex : histones acétylées), séquence promotrice (nécessaire comme la boite TATA, en plus d’autres boites et autres régulateurs en Cis), la RNA polymérase II, la TBP et les facteurs associés, et les looping factors, et n’oubliez pas les NTP.

2. L’initiation de la transcription :

• Le couple **cis-régulateur trans-régulateur** liés entre en contact **avec le complexe d’initiation de la RNA polymérase I**I par l’intermédiaire d’un ensemble de protéines appelé **médiateur**. La liaison du trans-régulateur et du médiateur nécessite un repliement du DNA du promoteur, par les looping factors, pour permettre le rapprochement de ces protéines.

• La RNA polymérase peut démarrer lorsqu’elle est activée par les facteurs de transcription.

• Alors la transcription peut commencer si les ribonucléosides substrats sont présents.

3. La partie de cette séquence qui peut participer à la régulation de l’expression d’un gène par un signal hormonal se trouverait au niveau du promoteur. Un exemple des éléments cis-régulateurs spécifiques le CRE (*cyclic AMP responsive element)* TGACGTCA*.*

4. Parmi les « motifs » de cette séquence marqués en gras ( GGCCAATCT / TATATA / GT / AG / AATAAA / TATGTTTG ) :

- Ceux qui définissent l’orientation de la transcription, et le brin sens et le brin matrice  sont : les boxes : GGCCAATCT, TATATA

- Ceux qui définissent le début et la fin de la transcription sont respectivement : TATATA (-30 NT) et TATGTTTG

- Ceux qui interviennent lors de la maturation du transcrit primaire sont les site donneur et récepteur : GT et AG.

5. Les segments soulignés dans cette séquence correspondent à : des exons

6. Les « motifs » intervenant dans le processus qui permet de lier les exons lors de la maturation du transcrit primaire par épissage (splicing) sont GT et AG

7. Le rôle du triplet de nucléotides en 206-208 ATG est l’initiation de la traduction. Ce codon permet l'association de l'ARNm sur la petite sous-unité du ribosome. Notez que, les nucléotides entourant le triplet initiateur ATG sont aussi bien conservés, ils sont nommés séquence de Kozak, chez les vertébrés cette séquence est 5’- GCCGCCPuCC**ATG**G-3’

8. Le nucléotide en position 391 est un C qui fait partie de la séquence traduite : sa position dans le cadre de lecture est définit de cette façon :

N0 étant 205, pour le codon 1 : N1 206, N2 207, N3 208 ; codon 2 : N1 209, N2 210, N3 211, etc. => Donc N1=N0+3X+1, N2=N0+3X+2, N3=N0+3X+3. X étant le numéro de codon, ATG est le codon 0. Bien sûr, il faut soustraire d’Intron 1.

La position 391=205+186-114=205+3x23+3 => le nucléotide C en position 391 est en N3 dans le cadre de lecture du codon 24.

9. La longueur après transcription et maturation de l'ARN messager (comprenant l'addition de 1000 AMP du côté 3'-OH) : le transcrit primaire s’allonge de 30 NT après la TATA boxe au signal de terminaison : de donc de 102 à 834. Lors de la maturation, le mRNA perd ses 2 introns, une enzyme coupe une dizaine de NT après le signal de poly adenylation AAUAAA, et gagne une queue de 1000 AMP en moyenne et un GMP en coiffe.

834-102=732 taille du transcrit primaire moins le signal de terminaison.

732-113-168+1+1000=1452 NT taille du mRNA « Ce chiffre est approximatif. Car le début de la transcription et la taille de la queue reste théorique dans cet exercice ».

10. Le nombre des Aa dépend du nombre de codons entre le codon de démarrage ATG à la position 206-208 et le codon Stop à la position 713-715.

(713-206) = 507

Les introns en moins !

507-113-168 = 225

La séquence traduite :

ATG GCT CTT CTT TCA GCG TTT TCG GCC ACC CAG GCA CGG AAG GGC TTC TGG GAC TAC TTC AGC CAG ACC AGC GGG GAC AAA GGC AGG GTT GAG CAG ATC CAT CAG CAG AAG ATG GCT CGC GAG CCC GCG ACC CTG AAA GAC AGG CTG CAG GCC CGT CTG GAT GAC CTG TGG GAA GAC ATC ACT CAC AGC CTT CAT GAC CAG GGC CAC AGC CAT CTG GGG GAC CCC

225/3 = 75 Aa

**Exercice 4**

Soit une séquence de bases présente sur un brin d’ADN (brin 1) :

**....-G-A-C-T-T-A-C-A-C-G-C-G-A-T-T-T-T-A-T-A-T-A-G-C-....**

1. Recopiez cette séquence et écrivez la séquence du brin d’ADN complémentaire (brin 2)
2. Sachant que l’ARNm issu de la transcription de ce fragment d’ADN code le début d’une protéine :

- déterminez quel est le brin matrice (justifiez votre réponse) et écrivez la séquence de l’ARNm.

- orientez les séquences des acides nucléiques.

- écrivez la séquence du polypeptide traduit à partir de cet ARNm.

1. A la suite d’une mutation, la 10ème base (G) du brin 1 représenté ci-dessus :

* a été remplacée par une adénine.
* a été remplacée par une cytosine.
* a été remplacée par une thymine.
* a été perdue.

- Écrivez pour chaque mutant les modifications introduites dans les séquences.

- Qu’advient-il de la protéine dans chaque cas ?

**Solution 4**

1. Recopiez cette séquence et écrivez la séquence du brin d’ADN complémentaire (brin 2)

Brin 1 : **-G-A-C-T-T-A-C-A-C-G-C-G-A-T-T-T-T-A-T-A-T-A-G-C-**

Brin 2 : **-C-T-G-A-A-T-G-T-G-C-G-C-T-A-A-A-A-T-A-T-A-T-C-G-**

2. Sachant que l’ARNm issu de la transcription de ce fragment d’ADN code le début d’une protéine :

- Le brin matrice :

Le brin sens est celui contenant le codon ATG, le brin matrice est le brin complémentaire (brin anti-sens).

B. matrice : Brin 1 :  **3’-G-A-C-T-T-A-C-A-C-G-C-G-A-T-T-T-T-A-T-A-T-A-G-C-5’**

B.sens : Brin 2 :  **5’-C-T-G-A-A-T-G-T-G-C-G-C-T-A-A-A-A-T-A-T-A-T-C-G-3’**

- Le brin sens est identique au mRNA en remplaçant le nucléotide T par le nucléotide U :

**5’-C-U-G-A-A-U-G-**U-G-C**-G-C-U-**A-A-A**-A-U-A-**U-A-U**-C-G-3’**

- la séquence du polypeptide traduit à partir de cet ARNm

M-C-A-K-I-Y-R

3. A la suite d’une mutation, la 10ème base (G) du brin 1 représenté ci-dessus :

Brin 1 : **-G-A-C-T-T-A-C-A-C-G-C-G-A-T-T-T-T-A-T-A-T-A-G-C-**

Brin sens : Brin 2 : **-C-T-G-A-A-T-G-T-G-C-G-C-T-A-A-A-A-T-A-T-A-T-C-G-**

Sa position dans le cadre de lecture est N3.

- a été remplacée par une adénine.

Brin 1 : **-G-A-C-T-T-A-C-A-C-A-C-G-A-T-T-T-T-A-T-A-T-A-G-C-**

Brin sens : Brin 2 : **-C-T-G-A-A-T-G-**T-G-T**-G-C-T-A-A-A-A-T-A-T-A-T-C-G-**

la séquence du polypeptide est : M-C-A-K-I-Y-R

cette substitution est synonyme

- a été remplacée par une cytosine.

Brin 1 : **-G-A-C-T-T-A-C-A-C-C-C-G-A-T-T-T-T-A-T-A-T-A-G-C-**

Brin sens : Brin 2 : **-C-T-G-A-A-T-G-**T-G-G-**G-C-T-A-A-A-A-T-A-T-A-T-C-G-**

la séquence du polypeptide est : M-W-A-K-I-Y-R

cette substitution est non-sens

- a été remplacée par une thymine.

Brin 1 : **-G-A-C-T-T-A-C-A-C-T-C-G-A-T-T-T-T-A-T-A-T-A-G-C-**

Brin sens : Brin 2 : **-C-T-G-A-A-T-G-**T-G-A**-G-C-T-A-A-A-A-T-A-T-A-T-C-G-**

la séquence du polypeptide est : M-

cette substitution est non-sens (le nouveau codon est un codon Stop)

- a été perdue.

Brin 1 : **-G-A-C-T-T-A-C-A-C-C-G-A-T-T-T-T-A-T-A-T-A-G-C-**

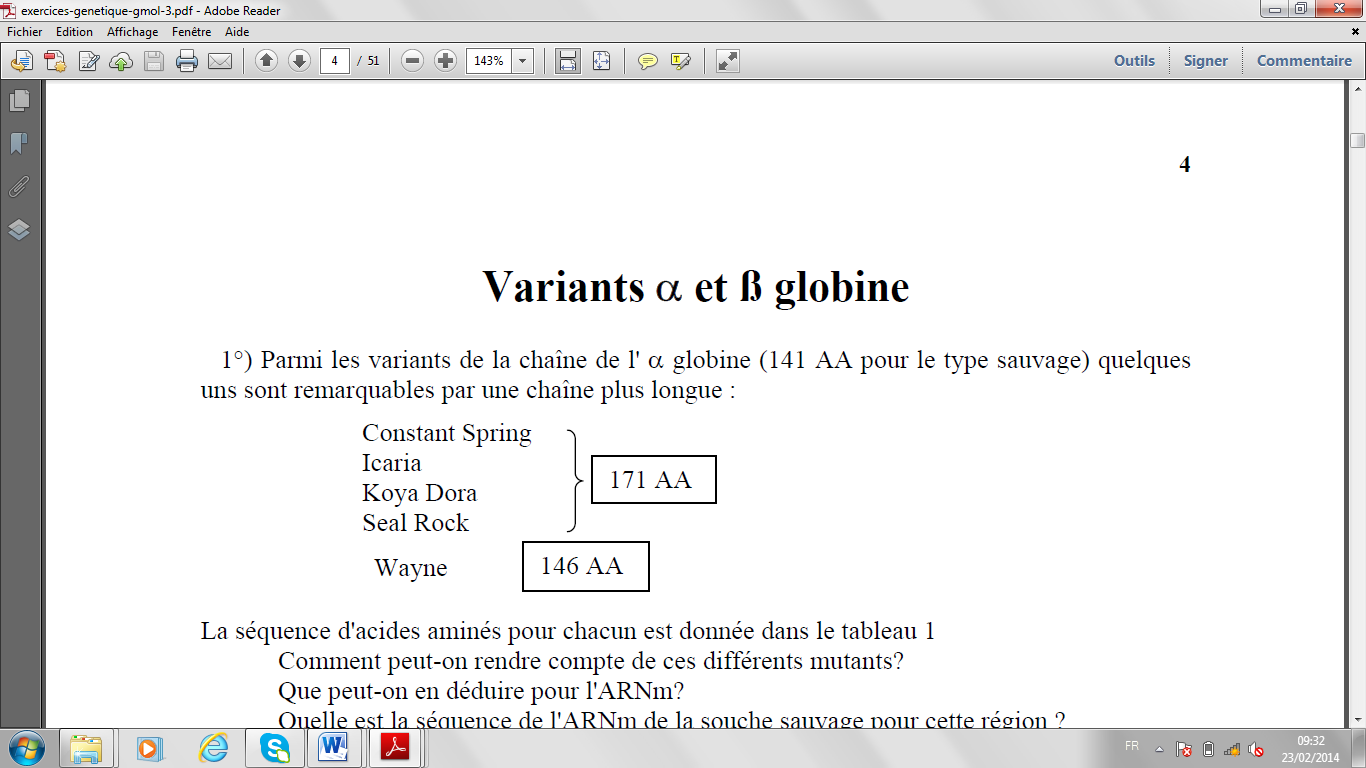
Brin sens : Brin 2 : **-C-T-G-A-A-T-G-**T-G-G**-C-T-A-**A-A-A**-T-A-T-**A-T-C-**G-**

la séquence du polypeptide est : M-W-L-K-Y-I

cette délétion induit un décalage de cadre lecture

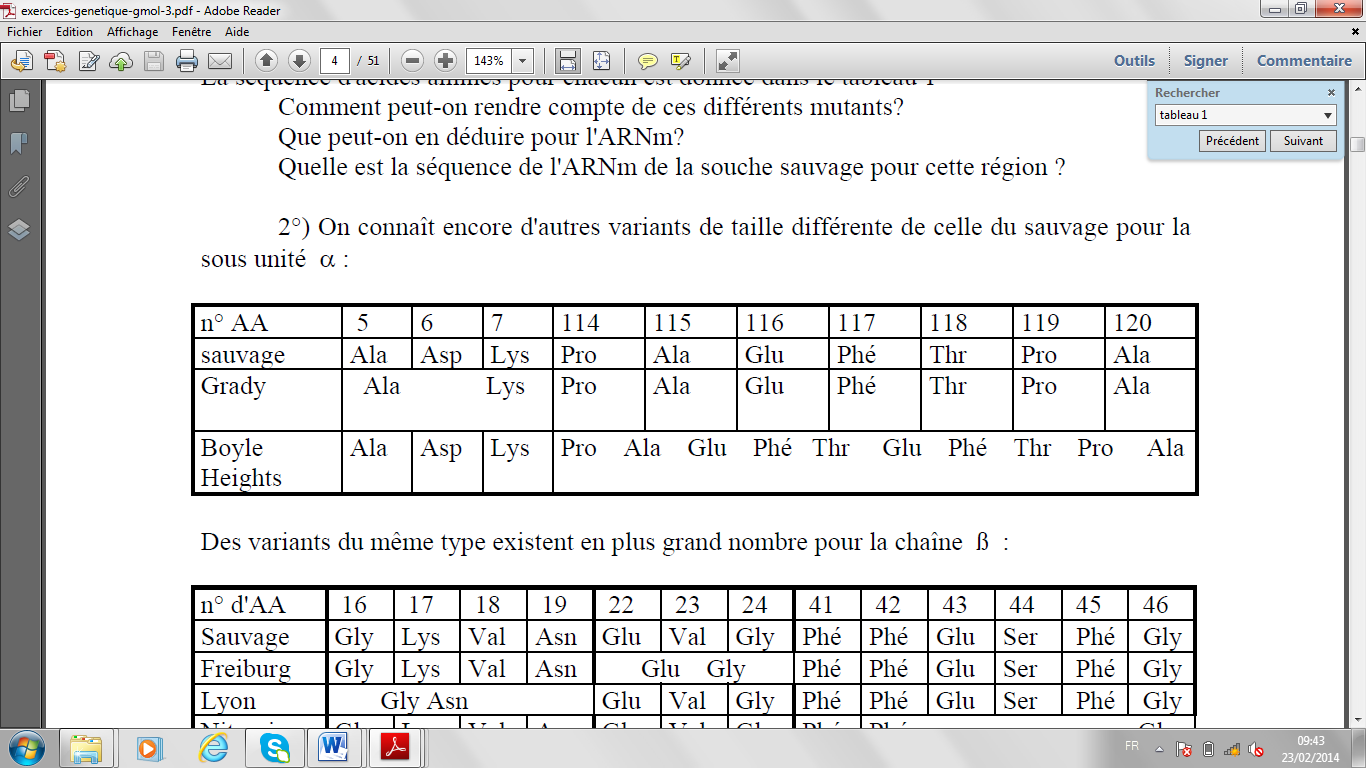
**Exercice 5 :**

1°) Parmi les variantes de la chaîne de l' α globine (141 AA pour le type sauvage) quelques-uns sont remarquables par une chaîne plus longue :



1. Comment peut-on expliquer ces différents mutants?
2. Que peut-on en déduire pour l'ARNm?

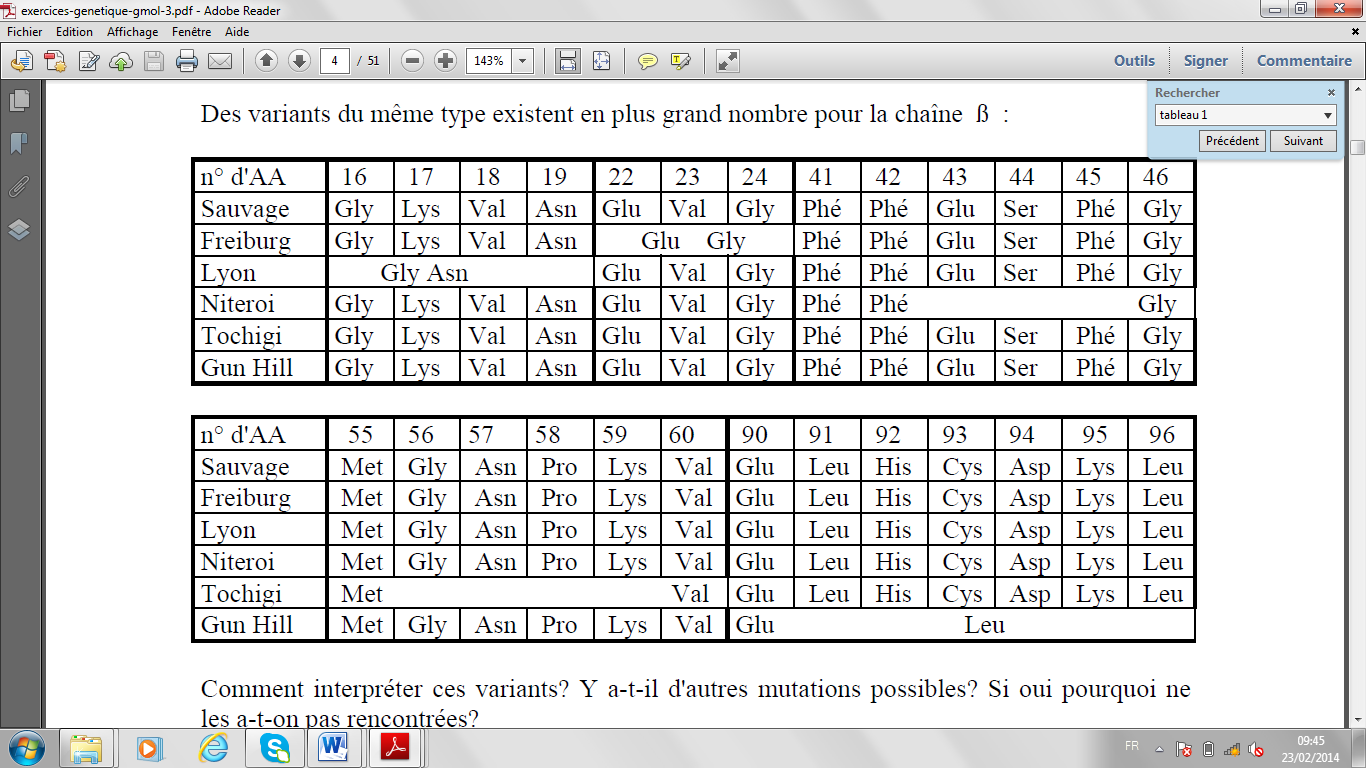
2°) On connaît encore d'autres variants de taille différente de celle du sauvage pour la sous unité α :



Des variantes du même type existent en plus grand nombre pour la chaîne ß :

a- Comment interpréter ces variants?

b- Y a-t-il d'autres mutations possibles? Si oui pourquoi ne les a-t-on pas rencontrées?



**Solution 5**

1°) a-b- les mutations touchent le codon Stop, qui n'est plus reconnu comme tel et la traduction continue jusqu'au prochain codon Stop dans le prolongement de la lecture :

1- Le codon Stop a muté par substitution, de sorte que la lecture s'est prolongée en phase avec ce qui précède, jusqu'au codon Stop suivant.

2- Où le codon Stop a subi une délétion de sorte que la suite de la séquence soit lue dans une phase différente (un cadre de lecture modifié, décalé), jusqu'au codon Stop suivant.

2°)

a- Tous ces mutants ne présentent pas de modification de cadre de lecture. Ils ont seulement un ou plusieurs acides aminés en moins ou en plus. Ils ont donc tous subit des délétions ou des additions d'autant de fois 3 bases (codon) :

délétion de 3 paires de bases : Grady, Freiburg

délétion de 6 paires de bases : Lyon

délétion de 9 paires de bases : Niteroi

délétion de 12 paires de bases : Tochigi

délétion de 15 paires de bases : Gun Hill

addition (duplication ) de 9 paires de bases : Boyle

b- Théoriquement, plusieurs types de mutations peuvent se produire dans la nature, elles ne sont pas observées car elles peuvent être silencieuses (nécessitant un criblage de la population générale à la recherche de polymorphismes), ou létales (leurs apparitions entraînerait un tel bouleversement des fonctions de la protéine qu'elles en deviennent létales, conduisant à l’élimination prématurée).

**Annexe**

