

Année Universitaire 2008-2009

**Université Claude Bernard Lyon 1
UFR de Médecine Lyon- RTH LAENNEC**

**CONCOURS de PCEM 1
Mercredi 20 mai 2009**

**EPREUVE DE BIOCHIMIE ET BIOLOGIE MOLECULAIRE
Pr Yves MOREL**

Nombre de points sur 250 : 40

Durée de l'épreuve : 90 minutes

Nombre de questions : 27 questions

Ce fascicule comprend 15 pages numérotées dont 3 pages de séquences.

IMPORTANT : vous devez vérifier au début de l'épreuve que ce fascicule est complet

Les calculatrices sont interdites

En réponse à chaque question vous pouvez noircir **zéro à cinq cases** sur la grille correspondant à des propositions **justes**

QUESTION N° 1 (1 point)

L'azote

- A est un élément de la même famille que le soufre
- B est plus électronégatif que l'élément ayant Z égal à 15
- C dans la molécule d'ammoniac, NH_3 , le doublet libre de l'azote est dans une orbitale 2s
- D dans la molécule d'acide nitrique, HNO_3 , l'arrangement des orbitales moléculaires autour de l'azote est de type AX_4
- E est un élément métallique

QUESTION N° 2 (1 point)

L'électronégativité

- A pour une période donnée augmente avec Z
- B pour une famille donnée augmente avec la période
- C dans une liaison de covalence entre deux atomes, le moins électronégatif porte une charge partielle δ^-
- D correspond, pour un atome lié, à sa tendance à attirer vers lui les électrons de liaison
- E la polarité d'une liaison entre deux atomes dépend de la somme des valeurs d'électronégativité de ces atomes

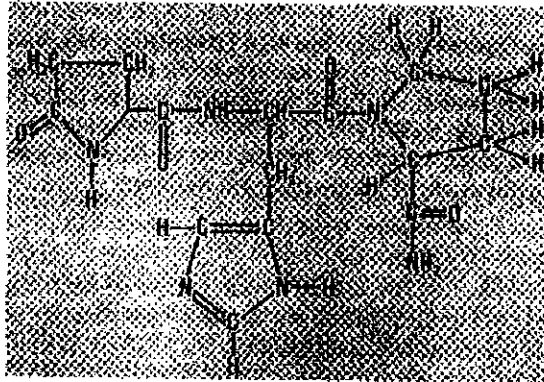
QUESTION N° 3 (1 point)

Pour préparer 50 mL de glucose à 10% (v/v) à partir d'une solution de glucose à 50%

- A Il faut faire une dilution au $1/10^{\text{ème}}$
- B Il faut faire une dilution au $1/5^{\text{ème}}$
- C prendre 50 mL de la solution à 50% et lui ajouter 50 mL d'eau
- D prendre 10 mL de la solution à 50% et lui ajouter 40 mL d'eau
- E prendre 20 mL de la solution à 50% et lui ajouter 80 mL d'eau

QUESTION N° 4 (2 points)

- A La réduction d'une fonction aldéhyde donne un alcool primaire
- B L'oxydation d'une fonction aldéhyde donne un acide carboxylique
- C La réduction du dioxyde de carbone donne l'acide carboxylique
- D Le glycérol est un trialcool avec deux fonctions alcool primaire
- E L'interaction entre un aldéhyde et un acide donne une liaison anhydride

QUESTION N° 5 (1 point)

Cette molécule

- A est l'Aspartam (L-aspartyl-L-phénylalanine-méthylester)
- B contient deux liaisons peptidiques
- C a son extrémité N-terminal modifiée par une amidification interne
- D est formée de 3 acides aminés aromatiques
- E donne 3 acides aminés (acide glutamique, histidine et proline) après hydrolyse acide

QUESTION N° 6 (2 points)

Le glucose-6-phosphate est converti

- A en glucose-1-phosphate dans le cerveau après un repas
- B en glucose dans le muscle lors d'une hypoglycémie
- C en fructose-6-phosphate dans le muscle lors d'un effort prolongé
- D en glucose dans le foie lors d'un jeûne
- E en glucose-1-phosphate dans le rein si la charge énergétique est basse

QUESTION N° 7 (2 points)

Le pyruvate est converti

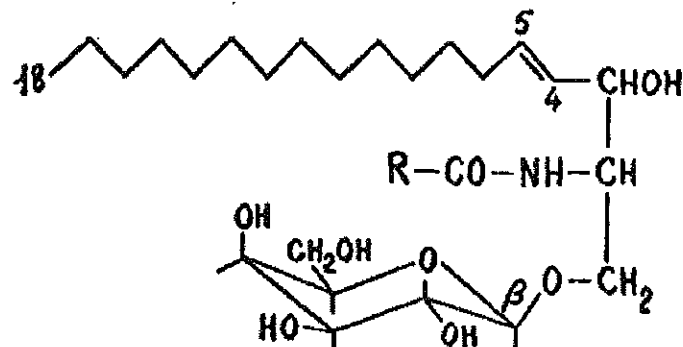
- A en lactate par la lactate deshydrogénase suite un apport insuffisant d'O₂ lors d'une anesthésie
- B en oxaloacétate par la pyruvate carboxylase dans le muscle lors d'un jeûne
- C en phospho-énol-pyruvate par la pyruvate kinase dans le foie lors du jeûne
- D en acétyl CoA par la pyruvate deshydrogénase dans le cerveau après un repas
- E en acétaldéhyde par la pyruvate décarboxylase lors d'un effort prolongé

QUESTION N° 8 (2 points)

Le lactose

- A contient un épimère en 4 du glucose
- B est un disaccharide non réducteur
- C a un de ses composés qui doit passer par une étape d'activation par l'UTP-glucose
- D produit 6 molécules de NADPH lors de sa dégradation par la glycolyse
- E doit être éliminé dans le régime des enfants atteints d'un déficit en glucose-1-phosphate-uridyl transférase

QUESTION N° 9 (1 point)



R est une chaîne aliphatique saturée.

A propos de cette molécule,

- A elle contient deux acides gras
- B elle se nomme glucosyl-céramide
- C c'est un ganglioside
- D elle est caractéristique des membranes plasmiques des cellules nerveuses
- E elle contient une liaison éther

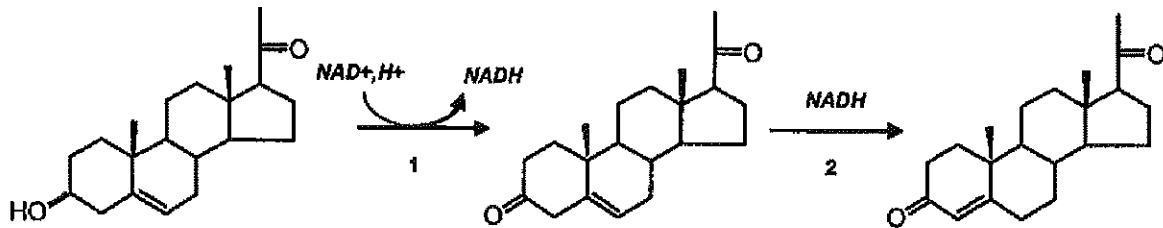
QUESTION N° 10 (2 points)

A propos de l'atome de Fer,

- A la couche de valence de l'ion ferrique, Fe^{3+} , contient 5 électrons
- B l'atome de Fe passe à l'état d'ion ferreux, Fe^{2+} , par une réaction d'oxydation
- C dans l'hème de l'oxyhémoglobine, il accepte 6 liaisons de coordination avec les atomes N et O
- D dans l'hème de l'hémoglobine, il est sous forme d'ion ferrique avec sa couche 3d incomplètement remplie
- E dans l'hème de l'hémoglobine lorsqu'il lie l'oxygène, il diminue de taille par saturation de toutes ses cases quantiques

Tous les énoncés des questions suivantes peuvent servir à répondre aux questions 11 à 27.

Le déficit en 3β -hydroxystéroïde déshydrogénase (3β -HSD), maladie à transmission autosomique récessive, est responsable chez le sujet 46,XY dans sa forme sévère d'une hypomasculinisation et d'un syndrome de perte de sel. Le premier gène 3β -HSD, appelé HSD3B1, dont l'ADNc a été isolé à partir d'une banque d'ADNc de placenta, code une enzyme catalysant deux activités dans son site actif. Le NADH formé lors de la première réaction permet le déroulement de la deuxième. HSD3B1 s'exprime dans les tissus périphériques comme la prostate, la glande mammaire, la peau,....



QUESTION N° 11 (1 point)

On peut dire que

- A le substrat de cette enzyme est la progestérone
- B la réaction 1 est une réaction d'oxydation
- C la réaction 2 est due à une activité isomérase
- D le produit formé est la 17-hydroxyprégnénolone
- E le produit intermédiaire a une liaison Δ^4

QUESTION N° 12 (1 point)

Comme aucune mutation du gène HSD3B1 n'a été trouvée chez des patients atteints d'un déficit en 3β -HSD, la première étape de la recherche du gène responsable du déficit en 3β -HSD a été faite en utilisant un anticorps anti-HSD3. Cet anticorps peut dans ce cas servir pour cribler

- A une banque d'ADN génomique humaine
- B une banque d'expression d'ADN complémentaire placentaire
- C une banque d'expression d'ADN complémentaire de cortico-surrénale
- D une banque d'expression d'ADN complémentaire de prostate
- E une banque d'expression d'ADN complémentaire de testicule

QUESTION N° 13 (3 points)

A la fin du fascicule se trouvent les deux séquences qui concernent le gène HSD3B2, responsable du déficit en 3 β -HSD. La séquence 1 (p.13) est complète. Dans la séquence 2 (p.14-15), plusieurs parties du gène, notamment introniques, ont été enlevées, mais la numérotation initiale a été maintenue.

Le gène HSD3B2

- A code une protéine de 372 acides aminés (incluant la méthionine)
- B se termine au nucléotide 9877 (à 3 nucléotides près)
- C contient 2 introns
- D ne présente que des sites accepteurs forts
- E contient une partie codante dans son premier exon

QUESTION N° 14 (2 points)

La mutation c.220G>A donne un acide aminé qui

- A potentiellement peut être N-glycosylé
- B devrait modifier la carte peptidique de l'enzyme purifiée en donnant un fragment se dirigeant plus vers le pôle négatif et migrant plus avec le solvant hydrophobe
- C s'écrit p.D74Q
- D possède une fonction amide dans sa chaîne latérale
- E peut lier l'iode 125

Enoncé pour les QUESTIONS N° 15 à 17

Cette mutation s'avère être un SNP. Si on voulait sur une large population rechercher la fréquence allélique de cette mutation, plusieurs méthodes pourraient être utilisées. Certaines de ces méthodes nécessitent, au préalable, pour être suffisamment sensibles, une amplification de l'exon en utilisant des amorces dont le début est situé à environ 30 nucléotides en amont ou en aval des sites d'épissage. Pour les trois questions suivantes, tous les oligonucléotides nécessaires à la méthode utilisée doivent être cochés.

QUESTION N° 15 (2 points)

Pour la mise en évidence de ce SNP par la méthode de séquençage, l'ensemble des oligonucléotides utilisés sera

- A 5' TTGTTTTATCTCCTCCCCAG 3'
- B 5' GCCTGCCAGGACGTCTCGGT 3'
- C 5' GGTGCGCCACGCAGGCTGGC 3'
- D 5' CTGGGGAGGAGATAAAACAA 3'
- E 5' TTTCCAATGACCTGACCTGT 3'

QUESTION N° 16 (2 points)

Pour la méthode MLPA, seulement une partie des sondes est mentionnée. Elle correspond à des séquences de l'exon.

- Oligo 1 5' TCCTGAAAAGAGCCTGCCAG 3'
 Oligo 2 5' GCCTGCCAGGACGTCTCGGT 3'
 Oligo 3 5' GACGTCTCGGTCGTCATCCA 3'
 Oligo 4 5' CTGGGGAGGAGATAAAACAA 3'
 Oligo 5 5' AACGTCTCGGTCGTCATCCA 3'

Quels sont les couples d'oligonucléotides qui permettront de détecter l'un des deux allèles ?

- A Oligo 1 et Oligo 2
 B Oligo 1 et Oligo 3
 C Oligo 1 et Oligo 4
 D Oligo 1 et Oligo 5
 E Oligo 3 et Oligo 4

QUESTION N° 17 (2 points)

Pour la mise en évidence de ce SNP par la méthode de Dot Blot, l'ensemble des oligonucléotides utilisés sera

- A 5' GCCTGCCAGGACGTCTCGGT 3'
 B 5' TTTCCAATGACCTGACCTGT 3'
 C 5' GGTGCGCCACGCAGGCTGGC 3'
 D 5' TTGTTTTATCTCCTCCCCAG 3'
 E 5' GCCTGCCAGAACGTCTCGGT 3'

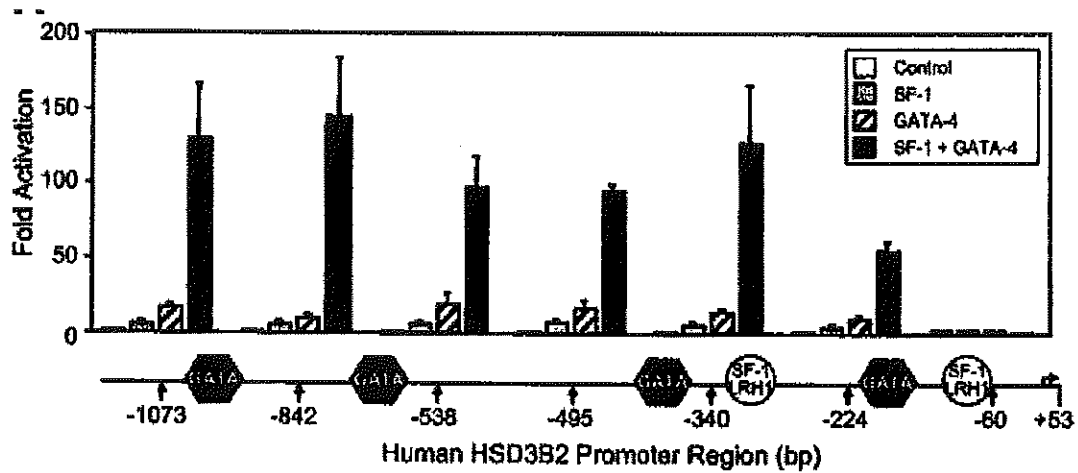
Enoncé pour les QUESTIONS N° 18 à 22

Pour essayer de comprendre la spécificité d'expression du gène HSD3B2 qui se limite aux surrénales, aux cellules de Leydig dans le testicule et aux cellules thécales dans l'ovaire, des séquences consensus pour des facteurs de transcription spécifiques des gonades ont été recherchées. Vous avez, page suivante, l'emplacement de certaines de ces séquences dans la région 5' en amont du gène s'étendant du nucléotide -1073 au nucléotide +53 après le début de la transcription. Les protéines GATA6, mais surtout GATA4 sont exprimées de façon importante dans les gonades tout comme SF1 appelé aussi NR5A1 et une protéine LRH-1 ou NR5A2.

L'expérience ci-dessous est le résultat de cotransfection dans une lignée de fibroblastes de rein de singe

- de différents vecteurs exprimant soit SF1 soit GATA-4.
- et d'un plasmide contenant une partie variable de la séquence promotrice du gène HSD3B2 et une autre séquence non variable. Le chiffre qui se trouve sous la flèche correspond à l'extrémité 5' de la région promotrice insérée. Toutes se terminent en 3' par le nucléotide +53.

Les résultats de l'expérience pour chaque séquence promotrice insérée sont au dessus de la flèche et représentés par 4 colonnes.



QUESTION N° 18 (1 point)

On peut dire que

- A les séquences nucléotidiques marquées GATA contiennent la séquence TATA
- B SF1 et LRH-1 appartiennent à la famille des récepteurs nucléaires
- C la séquence nucléotidique placée en aval de la région promotrice pourrait coder la luciférase
- D les valeurs en ordonnées correspondent à l'augmentation de l'expression de SF1 ou de GATA-4
- E l'expérience dont le résultat correspond aux colonnes noires nécessite seulement la cotransfection de deux vecteurs différents dans la lignée de fibroblastes

QUESTION N° 19 (1 point)

Si on considère que la hauteur des colonnes correspondant à GATA4 (hachurées) et à SF1 (en gris) ne sont pas statistiquement différentes, on peut dire que

- A la région située de -1073 à -842 ne contient pas d'enhancer pour la protéine GATA-4
- B la région située de -60 à +53 ne contient pas d'enhancer pour les protéines GATA-4 et SF1
- C SF1 seule a une activité faible de stimulation de la transcription
- D l'activité de SF1 et GATA-4 est fortement augmentée lorsqu'ils sont associés
- E un seul enhancer SF1 est suffisant pour obtenir le maximum d'effet synergique avec GATA-4 sur la transcription

QUESTION 20 (1 point)

SF1 est donc une protéine importante pour la régulation de la stéroïdogénèse. Elle contient un domaine contenant deux doigts de Zinc. L'atome de Zinc le plus fréquent a un nombre de masse de 64 et un nombre atomique de 30. A propos de cet atome, on peut dire que

- A il fait partie des métaux de transition
- B sous forme Zn, l'occupation des orbitales est $1s^2 2s^2 2p^6 3s^2 3p^6 3d^8 4s^2$
- C l'un de ses isotopes stables, l'atome ^{66}Zn , a 26 protons
- D il contient 34 neutrons
- E à l'état de Zn dans la protéine SF1, il forme des liaisons de coordinance avec les atomes S

Complément d'énoncé pour les QUESTIONS N° 21 et 22

Ces deux figures représentent des autoradiographies de retardement de migration sur gel

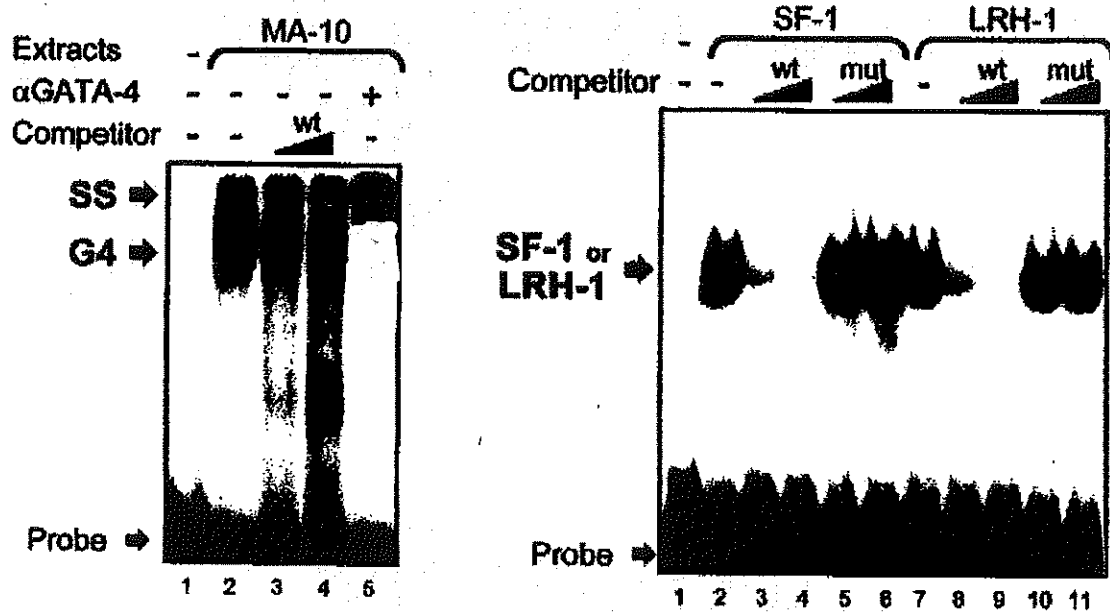


Figure 1

Figure 2

Sur la figure 1, MA-10 est une lignée de cellules de Leydig qui sécrète la testostérone ; G4 correspond à GATA-4

QUESTION 21 (1 point)

A propos de ces expériences,

- A elles servent à démontrer que GATA-4 et SF1 sont des enhanceurs
- B la sonde (« probe ») est un fragment d'ADN marqué supposé être un enhanceur
- C le « competitor wt » est identique à la sonde mais non marqué
- D α GATA-4 est la protéine GATA-4 radioactive
- E la séquence du « competitor wt » est identique pour SF1 et GATA-4

QUESTION 22 (1 point)

A propos de la spécificité de ces résultats,

- A les 2 expériences sont différentes parce que toutes les protéines, qui lient l'ADN, ne sont pas purifiées
- B la ligne 5 de la figure 1 montre que la séquence d'ADN est encore plus retardée sur cette électrophorèse et est un argument pour affirmer que GATA-4 lie spécifiquement la sonde
- C le « competitor » muté est radioactif et rentre en compétition avec la sonde
- D dans les expériences de la figure 2, les protéines SF1 et LH-R1 sont pures et ont été produites par des vecteurs d'expression contenant l'ADNc de ces protéines
- E dans les expériences de la figure 1, ligne 2, il est possible que d'autres molécules que GATA-4 lient la sonde

QUESTION 23 (3 points)

Une patiente 46,XX présente un syndrome de perte de sel à la naissance avec une élévation de 17-OHprogesterone suffisante pour évoquer un déficit en 21-hydroxylase. Le diagnostic de déficit en 3 β -HSD est soupçonné par l'absence de virilisation des organes génitaux externes et des taux très élevés de 17-OHPrégnénone avec rapport 17-OHPrégnénone/17-OHprogesterone > 10 (normalement autour de 1). La séquence du gène HSD3B2 montre que la patiente est hétérozygote composite pour la mutation p.Gly15Asp et une mutation dans le dernier exon.

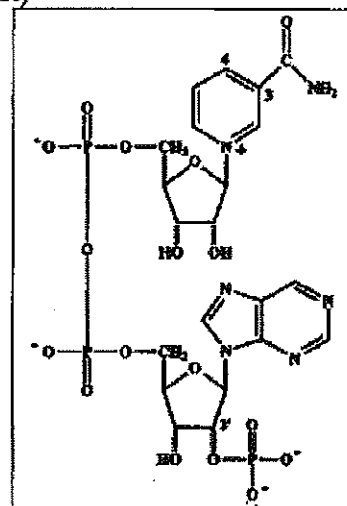
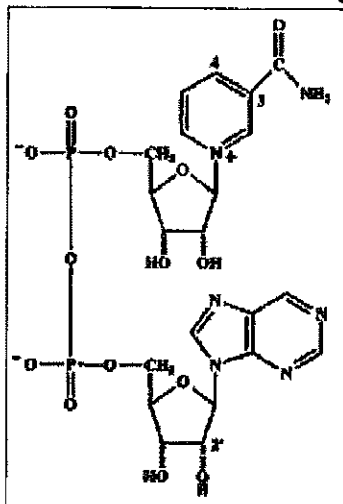
Ci-dessous une partie du résultat du séquençage du dernier exon de l'ADN de la patiente (numérotation de la séquence 1 à quelques nucléotides près : début 5'nucléotide 989 – fin 3'nucléotide 1060)

5'..... GNNCNTNCNTTANCCNNATGTACTGGATTGGCTTCCTGCTGGAA
GTAGTGAGCTTCCTACTCAGCCCA.....3'

La lettre N signe l'impossibilité du logiciel du séquenceur de déterminer le nucléotide. Le plus souvent, il existe une double séquence.

A propos de la seconde mutation

- A elle entraîne un décalage du cadre de lecture
- B elle donnerait théoriquement une protéine de 298 acides aminés en l'absence de NMD (dégradation des ARNm contenant des mutations « non sens »)
- C il s'agit de la mutation c.866delT
- D une amorce située dans le dernier intron est utilisée pour le séquençage
- E ce séquençage permet de dire que, s'il existe un phénomène de NMD pour l'ARNm provenant de cet allèle muté, il est partiel.

QUESTION N° 24 (1 point)

A propos de ces deux molécules

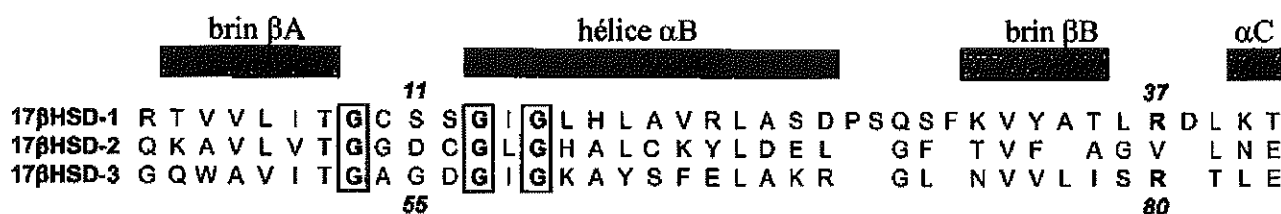
- A elles sont des transporteurs d'électrons
- B elles sont sous la forme réduite
- C la partie réactive dans l'oxydo-réduction est constituée par des phosphates
- D le cycle nicotinamide provient d'une vitamine hydrosoluble
- E elles ont les propriétés des groupements prosthétiques

Enoncé pour les questions suivantes 25 et 26

L'étude *in vitro* des 2 gènes HSD3 et du mutant p.Gly15Asp donne les résultats suivants pour la première activité de ces enzymes en prenant comme substrat celui de la question n° 11.

	HSD3B1	HSD3B2	p.Gly15Asp
K_M molécule1	3,7 μ M	47 μ M	188 μ M
K_M NAD ⁺	34 μ M	86 μ M	435 μ M

La séquence consensus T-G-X-X-X-G-X-G permet l'ancrage du coenzyme NADH ou NADPH dans la superfamille des Hydroxystéroïdes deshydrogénases (HSD). Le *domaine de liaison du cofacteur* comprend un brin β , une hélice α et un brin β . Le coude entre le 1^{er} brin β et l'hélice α contient la séquence consensus T-G-X-X-X-G-X-G. Ci-dessous, vous avez l'alignement des séquences partielles de ce domaine de 3 protéines appartenant à cette superfamille des HSD



Classiquement, la présence d'un acide aminé chargé dans la séquence consensus au niveau des trois X et la nature des acides aminés de la boucle entre le brin β B et l'hélice α C favorisent le choix du cofacteur et donc de la réaction. Comme vous le voyez sur la figure ci-dessus, la préférence de l'enzyme 17 β -HSD de type 1 pour le NADPH est due à l'absence d'acide aminé acide dans la séquence consensus et à la présence de R37 et de K195 (non montré).

La Gly15 des deux protéines HSD3 correspond bien à la Gly15 de 17 β -HSD1. Des études structure – fonction de ces HSD3 ont montré que le mutant p.D35A utilise plutôt NADP que NAD comme cofacteur.

QUESTION N° 25 (2 points)

En utilisant tous les renseignements en votre possession depuis la question 11, on peut dire que l'enzyme HSD3B2.

- A a un domaine de liaison du cofacteur presque similaire à celui de la famille des 17 β HSD
- B mutée (p.Gly15Asp) a plus d'affinité pour le cofacteur
- C utilise plutôt le NAD à cause de la présence du résidu aspartyl-35 qui repousse la charge positive du phosphate supplémentaire du NADP
- D n'a pas d'acide aminé acide dans la séquence consensus qui contient les résidus glycyll
- E mutée (p.Gly15Asp) n'affecte que l'interaction de l'enzyme avec le cofacteur.

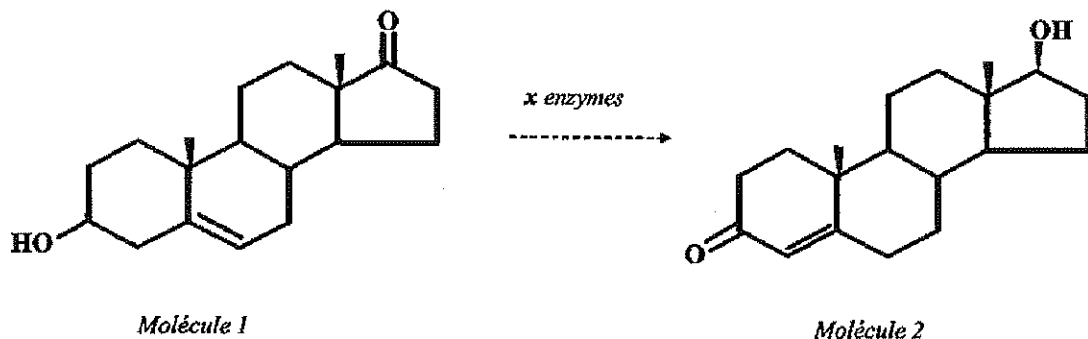
QUESTION N° 26 (2 points)

Quelles sont les hypothèses pouvant expliquer l'augmentation paradoxale de la 17-OHProgestérone chez ce patient atteint d'un déficit sévère en 3 β -HSD ?

- A Persistance d'une activité résiduelle HSD3B2
- B Conversion périphérique de la 17OHPrégénolone par l'enzyme HSD3B1 qui a une meilleure affinité pour ce stéroïde que l'enzyme HSD3B2 normale ou mutée.
- C L'anticorps anti-17OHprogestérone utilisé pour le dosage reconnaît un peu (ou croise avec) la 17-OHPrégénolone qui est très élevée.
- D Conversion périphérique de la 17OHPrégénolone par l'action successive des enzymes HSD3B1 et 21-hydroxylase.
- E Conversion périphérique de DHA par l'enzyme HSD3B1 qui n'est pas mutée

QUESTION N° 27 (2 points)

Des fortes concentrations non physiologiques de la molécule 1 ont été découvertes récemment chez plusieurs athlètes français. Cette molécule 1 est considérée dans ce cas comme une molécule dopante car elle se convertit dans les tissus périphériques (cutanés, musculaires,.....) en la molécule 2 grâce à l'action de plusieurs enzymes.



On peut dire que

- A la molécule 1 est la DHEA
- B l'une des enzymes qui catalyse cette conversion périphérique s'appelle la protéine HSD3B2
- C après l'action de la protéine HSD3B1, il reste plusieurs étapes enzymatiques pour obtenir la molécule 2
- D la molécule 2 lie un récepteur membranaire
- E l'une des enzymes qui catalyse cette conversion périphérique s'appelle la 17 β -hydroxystéroïde-deshydrogénase.

Séquence 1

GAGGCAGTAAGGACTTGGACTCCTCTGTCCAGCTTTTAAACAATCTAAGTTACGCCCTCTTCTGGGGTCACGCCT 72
 AGAATCAGATCTGCTCTCCAGCATCTTCTGTTTCTGCTGGCAAGTGTTCCTGCTACTTTGGATTGGCCACG 142
 M G W S C L V T G A G G L L G Q R I 18
 ATG GGC TGG AGC TGC CTT GTG ACA GGA GCA GGA GGG CTT CTG GGT CAG AGG ATC 196
 V R L L V E E K E L K E I R A L D K 36
 GTC CGC CTG TTG GTG GAA GAG AAG GAA CTG AAG GAG ATC AGG GCC TTG GAC AAG 250
 A F R P E L R E E F S K L Q N R T K 54
 GCC TTC AGA CCA GAA TTG AGA GAG GAA TTT TCT AAG CTC CAG AAC AGG ACC AAG 304
 L T V L E G D I L D E P F L K R A C 72
 CTG ACT GTA CTT GAA GGA GAC ATT CTG GAT GAG CCA TTC CTG AAA AGA GCC TGC 358
 Q D V S V V I H T A C I I D V F G V 90
 CAG GAC GTC TCG GTC GTC ATC CAC ACC GCC TGT ATC ATT GAT GTC TTT GGT GTC 412
 T H R E S I M N V N V K G T Q L L L 108
 ACT CAC AGA GAG TCC ATC ATG AAT GTC AAT GTG AAA GGT ACC CAG CTA CTG TTG 466
 E A C V Q A S V P V F I Y T S S I E 126
 GAG GCC TGT GTC CAA GCC AGT GTG CCA GTC TTC ATC TAC ACC AGT AGC ATA GAG 520
 V A G P N S Y K E I I Q N G H E E E 144
 GTA GCC GGG CCC AAC TCC TAC AAG GAA ATC ATC CAG AAC GGC CAC GAA GAA GAG 574
 P L E N T W P T P Y P Y S K K L A E 162
 CCT CTG GAA AAC ACA TGG CCC ACT CCA TAC CCG TAC AGC AAA AAG CTT GCT GAG 628
 K A V L A A N G W N L K N G D T L Y 180
 AAG GCT GTG CTG GCG GCT AAT GGG TGG AAT CTA AAA AAT GGT GAT ACC TTG TAC 682
 T C A L R P T Y I Y G E G G P F L S 198
 ACT TGT GCG TTA AGA CCC ACA TAT ATC TAT GGG GAA GGA GGC CCA TTC CTT TCT 736
 A S I N E A L N N N G I L S S V G K 216
 GCC AGT ATA AAT GAG GCC CTG AAC AAC AAT GGG ATC CTG TCA AGT GTT GGA AAG 790
 F S T V N P V Y V G N V A W A H I L 234
 TTC TCT ACA GTC AAC CCA GTC TAT GTT GGC AAC GTG GCC TGG GCC CAC ATT CTG 844
 A L R A L R D P K K A P S V R G Q F 252
 GCC TTG AGG GCT CTG CGG GAC CCC AAG AAG GCC CCA AGT GTC CGA GGT CAA TTC 898
 Y Y I S D D T P H Q S Y D N L N Y I 270
 TAT TAC ATC TCA GAT GAC ACG CCT CAC CAA AGC TAT GAT AAC CTT AAT TAC ATC 952
 L S K E F G L R L D S R W S L P L T 288
 CTG AGC AAA GAG TTT GGC CTC CGC CTT GAT TCC AGA TGG AGC CTT CCT TTA ACC 1006
 L M Y W I G F L L E V V S F L L S P 306
 CTG ATG TAC TGG ATT GGC TTC CTG CTG GAA GTA GTG AGC TTC CTA CTC AGC CCA 1060
 I Y S Y Q P P F N R H T V T L S N S 324
 ATT TAC TCC TAT CAA CCC CCC TTC AAC CGC CAC ACA GTC ACA TTA TCA AAT AGT 1114
 V F T F S Y K K A Q R D L A Y K P L 342
 GTG TTC ACC TTC TCT TAC AAG AAG GCT CAG CGA GAT CTG GCG TAT AAG CCA CTC 1168
 Y S W E E A K Q K T V E W V G S L V 360
 TAC AGC TGG GAG GAA GCC AAG CAG AAA ACC GTG GAG TGG GTT GGT TCC CTT GTG 1222
 D R H K E T L K S K T Q * 373
 GAC CGG CAC AAG GAG ACC CTG AAG TCC AAG ACT CAG TGATTTAAGGATGACAGAGATG 1280
 TGCATGTGGGTATTGTTAGGAAATGTCATCAAACCTCCACCCACCTGGCTTCATACAGAAGGGCAACAGGGGCA 1352
 CAAGCCCAGGTCTGCTGCCTCTCTTTCACACAATGCCCAACTTACTGTCTTCTTCATGTCATCAAATCTG 1424
 CACAGTCACTGGCCCAACCAGAACTTTCTGTCCATAATCATAACACCAGAAGACAAACAATATGATTTGCTGTT 1496
 ACCAAATCTCAGTGGCTGATTCTGAACAATGTGGTCTCTCTTAACTTGAGGTTCTCTTTTACTAATAGAG 1568
 CTCCATTTCCCCTCTTAAATGAGAAAGCATTTCTTTCTCTTAAATCTCCTATTCTTTCACACAGTTCAACA 1640
 TAAAGAGCAATAAATGTTTTAATGCTT 1667

Séquence 2

CAGCTGAAGGAACATTTGGTCCCACCCCTGAAGGTTCATGGGGTCTGTAAGGTGCACTCCAGCATCTGATATG 72
GCAGGCAAAGGTGTGTGATGCACCTGGGTGGTTTCAGATAGTGTCCACAAGTCCCAGGGCCCGTGTCTAGT 144
TCTGAGAGCTTCAGAAGACAAGCTGCTACACCTTGCATCACAGGACCAACCTTTCCCTAGGCAAGTGGGTTTG 216
AAGCATTGGCAGCCAGGTTCCAAGTAGAAGGTGAATGTCTACACACAGGAGAAAGGCAAGACCAGAATGGG 288
GAGTCAAGCTAAATCATTTGTAGAGTATGACTTCCAACCACTTCAGAAAGCATTCTTTTGAAGTGTCTATGG 360
CATTTCCACCAACATAGACCATAGGCTGGACCATAAAAACAAATCTCAATATATATCAAAGAAATCAAAATCAT 432
ACAGAGTATATGCTCTGGCCACAATAATATTTAATTAGAAATACATAACAATAAGAAAATCCCAGTTTATAG 504
AAAACATGCAACATGCTTCTCAATAACACTTGGGTCAAAGAAGACATCACAAAGAAAATTTCAAAACGTCAG 576
TAACTGAATAATGATAACCCAATGCATCAAAATGTTGGATACAGCTAACATAGTGTTTAGAGAGAAATTTAC 648
AACTTTAAATGCCTATGGTGGAAATGAGCAAAGAAGTAAAGTCAATTATCTATGCTTTTATCATAGAACC 720
TGAAAAGAAGAGTAAAATAAACCCAAAGTTAATAAACATTTAAGCCAATAATAAAAATAATGAAAATCATG 792
AAATAGAAAATGAACAAAACATAGAGACAATCAGTGGTCAATTTTTCTGTACTCCATGAACAAAACCTTCAG 864
TTCAGAATGATTTTATATTTAAATATATATCCAGAATCACAAGATCTTAGAAGAAAATTTAGGAAAATAT 936
AGGAATAAAGTGGGAACCTCTGTGGGAATAAAAATCAGAAAGTAAATTGCCTACCAGGTAAACTCGGGGAGGG 1008
CAGGCAATTGACTAAGAACAGGCAGATAAAACTTTCTGGAGTGAAGGAAATGGTGGTCTCTTTTCGTTGGT 1080
ACTTACATGGACATATATAACTGTCAAGGTTCACTTGACAAATGAACACTTAAGTGTACACTGAACGCTTAAGA 1152
TCTGGGCCCTTGTAAATGCCAGATTACATCTACTTAAAAAACTACATCTCTATTTTTTTAAGTAGAGAAAAG 1224
TGTCTTGGGGAGTTTATGGCAGTGATGTTTGGGCAGACTGGGCATAGGTTTCAAGGCATTCATAGCATCT 1296
TCAGTTCAACAATTTGCTAAGAATGGTTTTGCAGCCAAAATGTAAGATGAGTGCAAATATCTATGATTAGA 1368
GTCATTTGGGCATACAACTACAAATCCCCACACTTGTATTACAGGCCATATGATTAATTCGTTTACACTAGAT 1440
ACCTTTCCCTCAATGAGGTTCTTCCATTAGGAACCCAGAGCTCTCCAGGGAAAATTCGAATGAGGTTTTG 1512
GATATATTTGGTGGAAAAGAAAGTACCTCACATAAACTTAGTGATTGGAGCTGTCACCATTGAAAATTC 1584
TAACTTTGCAAGACAGCAGTGCAGTTCATTAATGACACTCTCAGAGAGTTGATAATGGGCTAAAATAGATCT 1656
CCCTCCAGGTGGATTTACTGTACAAGGACAACATTTACATTGCACTTGGAGACTTCTCCAGTTTGGTTTAA 1728
GTTACAGATTGCAGATCCCAGACAGCTGGTATCAACTGACCAGTGTCTGTTAAGGCTAAAGCCAGACTC 1800
TTTATCACACTGTGGCCTTAAGATTGGATTTCTCTTCTGTTCTTGGGAAGAATTAGAGATATAACCTAAAG 1872
GTCACTATTATTCTGAGAAAAGGGATTCTGGAGGAGGGCAATGAGTATGTTGGCAGGAGTTCAAGGTA 1944
ATAAGGGCTGAGACACAAGCCACAGAGCATAAAGCTCCAGTCTTCTTCCAGGGATGAGGAGTAAGGACTT 2016
GGACTCCTCTGTCAGCTTTTAAACAATCTAAGTTACGGTTAGAGCTTTCTCCTTTTTCTTCACTTACTCTG 2088
GCAGTTGTGGGTCATGGAATTTTTGTAAAAAATGGGGTGGAGGAAAATAAGGCATCTGCTGAGTGTATAAC 2160
CATTTACCTCTTGTTTTTAGCCCTCTTCTGGGTACGCTAGAATCAGATCTGCTCTCCAGCATCTTCTGTT 2232
M G W S C L V T 8
TCCTGGCAAGTGTTCCTGCTACTTTGGATTGGCCACGATG GGC TGG AGC TGC CTT GTG ACA 2294
G A G G L L G Q R I V R L L V E E K 26
GGA GCA GGA GGG CTT CTG GGT CAG AGG ATC GTC CGC CTG TTG GTG GAA GAG AAG 2348
E L K E I R A L D K A F R P E L R E 44
GAA CTG AAG GAG ATC AGG GCC TTG GAC AAG GCC TTC AGA CCA GAA TTG AGA GAG 2402
E F S 47
GAA TTT TCTAGTAAAGTAAACTTGAGTTCATGGGTCTGTGGCTCCATCTTAAACTCTGCATGGGTGTGGG 2470
GAGGTTGACCTTGTCTAGCAAGTTAAGGAAAGTTGTAGCCAAATGAAAGCCAGTCACACATCTAAAGTCATC 2542
AGAAAGGAAATAGAATAAAATGGTATAGTGTGAAAGATACTGGATGGGGTGTCCAGAGACTGGATTCTGGCC 2614
(...)TTTCCCAATAAACTGCTACACATGCTGATTTCTGTGCTTTTTGTTTACTTGTTCCTTTTATGGAATGT 6142
AGTACACCCCTCACTCTAATACCCACACTCTAATCTCTTGGAGACCTATGTAACATCGCCTTTATCAGAAA 6214
K L Q N R 52
ACTFCCCAGCCAGATCCAGAAATCTTTCCAATGACCTGACCTGTGTTTACACACAGAG CTC CAG AAC AGG 6282
T K L T V L E G D I L D E P F L K R 70
ACC AAG CTG ACT GTA CTT GAA GGA GAC ATT CTG GAT GAG CCA TTC CTG AAA AGA 6336
A C Q D V S V I H T A C I I D V F 88
GCC TGC CAG GAC GTC TCG GTC GTC ATC CAC ACC GCC TGT ATC ATT GAT GTC TTT 6390
G V T H R E S I M N V N V K 102
GGT GTC ACT CAC AGA GAG TCC ATC ATG AAT GTC AAT GTG AAAGGTACAGTAGCCTGG 6447
GGAGGAGATAAAAACAAAGTTGGTTAAATGAGGATCAGAAAGAAGGACAAGAAAGGGAAGAGAAGTCACTCCAT 6519
TGAACACCTGCTGAGCTCTTGGCCAAGTGCCTTTGCTGATCACTACTGACTGGGGAGTTCAAGGCTGGTAAAC 6591
(...)TCTTTTATTTTTGTGTTTTTCTACGGCTGTATCATGACCCAATCTCAGTCAGAGCCACAGAAGAA 8535
TGCACCCCTGAGTCTGTTATAACCACTGCACCTTGGGAGTGGGGAGTGGGGCACATGGATCTGTGATGTGTT 8607
G T Q L L 107
GCAGCTCCTTTGGGATATTTCTGACACTGTCATCATGCTCTTTCGTGGGCAGGT ACC CAG CTA CTG 8673
L E A C V Q A S V P V F I Y T S S I 125
TTG GAG GCC TGT GTC CAA GCC AGT GTG CCA GTC TTC ATC TAC ACC AGT AGC ATA 8727
E V A G P N S Y K E I I Q N G H E E 143
GAG GTA GCC GGG CCC AAC TCC TAC AAG GAA ATC ATC CAG AAC GGC CAC GAA GAA 8781
E P L E N T W P T P Y P Y S K K L A 161
GAG CCT CTG GAA AAC ACA TGG CCC ACT CCA TAC CCG TAC AGC AAA AAG CTT GCT 8835

E K A V L A A N G W N L K N G D T L 179
 GAG AAG GCT GTG CTG GCG GCT AAT GGG TGG AAT CTA AAA AAT GGT GAT ACC TTG 8889
 Y T C A L R P T Y I Y G E G G P F L 197
 TAC ACT TGT GCG TTA AGA CCC ACA TAT ATC TAT GGG GAA GGA GGC CCA TTC CTT 8943
 S A S I N E A L N N N G I L S S V G 215
 TCT GCC AGT ATA AAT GAG GCC CTG AAC AAC AAT GGG ATC CTG TCA AGT GTT GGA 8997
 K F S T V N P V Y V G N V A W A H I 233
 AAG TTC TCT ACA GTC AAC CCA GTC TAT GTT GGC AAC GTG GCC TGG GCC CAC ATT 9051
 L A L R A L R D P K K A P S V R G Q 251
 CTG GCC TTG AGG GCT CTG CGG GAC CCC AAG AAG GCC CCA AGT GTC CGA GGT CAA 9105
 F Y Y I S D D T P H Q S Y D N L N Y 269
 TTC TAT TAC ATC TCA GAT GAC ACG CCT CAC CAA AGC TAT GAT AAC CTT AAT TAC 9159
 I L S K E F G L R L D S R W S L P L 287
 ATC CTG AGC AAA GAG TTT GGC CTC CGC CTT GAT TCC AGA TGG AGC CTT CCT TTA 9213
 T L M Y W I G F L L E V V S F L L S 305
 ACC CTG ATG TAC TGG ATT GGC TTC CTG CTG GAA GTA GTG AGC TTC CTA CTC AGC 9267
 P I Y S Y Q P P F N R H T V T L S N 323
 CCA ATT TAC TCC TAT CAA CCC CCC TTC AAC CGC CAC ACA GTC ACA TTA TCA AAT 9321
 S V F T F S Y K K A Q R D L A Y K P 341
 AGT GTG TTC ACC TTC TCT TAC AAG AAG GCT CAG CGA GAT CTG GCG TAT AAG CCA 9375
 L Y S W E E A K Q K T V E W V G S L 359
 CTC TAC AGC TGG GAG GAA GCC AAG CAG AAA ACC GTG GAG TGG GTT GGT TCC CTT 9429
 V D R H K E T L K S K T Q * 373
 GTG GAC CGG CAC AAG GAG ACC CTG AAG TCC AAG ACT CAG TGATTTAAGGATGACAGA 9486
 GATGTGCATGTGGGTATTGTTAGGAAATGTCATCAAACCTCCACCCACCTGGCTTCATACAGAAGGCAACAGG 9558
 GGCACAAGCCCAGGTCTGCTGCTCTCTTTACACAATGCCCAACTACTGTCTTCTTCATGTCATCAAAA 9630
 TCTGCACAGTCACTGGCCCAACCAGAACTTTCTGTCTAATCATACACCAGAAGACAAACAATATGATTTGC 9702
 TGTTACCAAATCTCAGTGGCTGATCTGAACAATTGTGGTCTCTCTTAACTTGAGGTTCTCTTTTGACTAAT 9774
 AGAGCTCCATTTCCCTCTTAAATGAGAAAGCAATTTCTTTCTCTTAAATCTCCTATTCCTTCACACAGTTC 9846
 AACATAAAGAGCAATAAATGTTTAAATGCTTAAACATGGAGAGAAGCATGTTTTCTGTTAATACATACTTTCA 9918
 TCCTTTTCCTTTCAATTTCCAACATACCCATCTCCAACCTAGAAGAGTAGCATGAAACCAGGGTTGGGGGGT 9990
 GGCAGAGTAGTAGGAGGTACAGAAAGGGCAGATGAATACTGAATAAGCTCTACTATGTTTTAGGCATGTAT 10062
 GAAATGAACTACATATGACATATATTTACTCAATTAAGTACTCATCAGAAACCTGTCAATTGAGGAAATGAC 10134
 AAGACACGTCTCAATCATTTTAGGAGATTCACTTGAAGATGTGCCAGGAGACAGGCCTATGCCTTTCTCCA 10206
 AAGATGATTTTGAGGGCTCCAATTCAAAGGAGGGAAAGAGAGGGATATTGAGAAGCACACAGTTTTTTATGTA 10278
 ACGGGATGGTAGGAAAAAATAGTCATTCATGCCTTTGTCTGGCTCAGTGAATCTGCATTTTTTACATAAGA 10350
 TGACATAAACAGTAGGGGCAGAGGAAATGCAGATTCCCTGCATTTTACATAAAAATAACAGACAAAATGGGG 10422
 TAGGGGAAAAATCAGATAGGCATTTGTGTCTGGGGTAACTGCACCGGTAGTGATAAGCTATCAATTTGCAT 10494
 GCCGTGGTGAAGTTTAAACAGCTCACTAGGAAATTCCTTGTGGGCAAAATATGGGGGAGGCATGTAGCTTT 10566
 CATCTTGTAGCCATCTGTTTAGGCACCAAAGGGGGGAGGCAGGTTTGCTTGACCCAGTTCCAGCTTGAT 10638
 TTTTCCCTTTGGCTAAATGAGATTGGGGCCCCATAATTTAATTTTTTTTACAGTTTTTCCCTTTTTTTCTCT 10710
 AAACCTTTTTGAAGTACTGCATTTTGAAGAAATATGCATTCCTGGCCCTCGGTTGCTTTTTCTTTTTTTTT 10782
 TTTTTTGCTGGTTTTCTCATGGCTAGGATGTTTTATTCCTAGAAGTTCAAGTCCACATCACTAGGAAGGCTC 10854
 ATTCCTGGGAGGTTAATGTCACCTTGCTAGGAAGGCTTATTCCTAGGAAATCTTGTCCCATGAAGAAAAA 10926

