

Année Universitaire 2004-2005

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

P C E M 1

Vendredi 3 juin 2005

EPREUVE DE BIOCHIMIE ET BIOLOGIE MOLECULAIRE

Pr Yves MOREL

Dr Philippe DEVILLER

Durée de l'épreuve : 90 minutes

Nombre de question : 30 questions

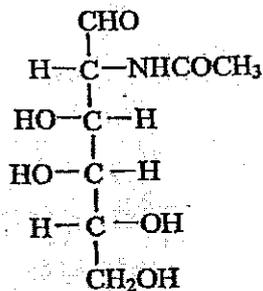
Ce fascicule comprend 16 pages numérotées dont 3 pages de séquences

IMPORTANT : vous devez vérifier au début de l'épreuve que votre livret est complet

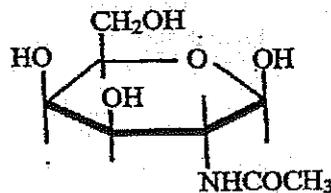
En réponse à chaque question vous pouvez noircir **zéro à cinq cases** sur la grille correspondant à des propositions **justes**

QUESTION N° 7 (3 points)

Soit le dérivé glucidique représenté ci-dessous selon les conventions de la projection de Fischer. On rappelle que l'anomère β correspond à une disposition du groupe hydroxyle de la fonction hémiacétalique au dessus du plan moyen du cycle pour la série D et en dessous du plan moyen du cycle pour la série L.



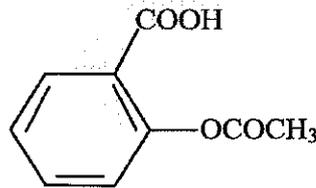
- A Il s'agit d'un dérivé N-acétylé d'aldohexosamine
- B Ce dérivé entre dans la composition de certains glycosaminoglycans
- C La configuration absolue du carbone n°2 est S
- D La représentation selon Haworth de l'anomère β de la forme pyranose de ce sucre est la suivante



- E Il existe une conformation chaise de l'anomère β de la forme pyranose de ce sucre pour laquelle les cinq substituants, autres que l'hydrogène, sont tous en position équatoriale

QUESTION N° 10 (3 points)

De nombreuses spécialités pharmaceutiques contiennent 500 mg d'acide o-acétylsalicylique et de l'hydrogénocarbonate de sodium. On donne les valeurs des pKa suivants : acide carbonique = 6,4 ; acide o-acétylsalicylique = 3,5. La formule de l'acide o-acétylsalicylique est indiquée ci-dessous. Masses atomiques : C = 12, H = 1, O = 16



- A L'acide o-acétylsalicylique est un acide carboxylique à chaîne latérale aromatique
- B L'acide o-acétylsalicylique peut être obtenu par acylation de l'acide 2-hydroxybenzoïque au moyen du chlorométhane
- C L'ion hydrogénocarbonate est une base moins forte que l'ion o-acétylsalicylate
- D A pH 7 le rapport [hydrogénocarbonate]/ [acide carbonique] est inférieur à 1
- E Après dissolution dans 200 ml d'eau, la molarité de la solution d'acide o-acétylsalicylique ainsi obtenue est égale à 13,9 mM

QUESTION N° 11 (2 points)

L'acide aspartique

- A a une charge négative dans une solution aqueuse dont le pH est 8
- B a son point isoélectrique à 5,75
- C se présente dans une solution aqueuse à pH 6 avec une chaîne latérale dont la proportion de la concentration de forme $-COO^-$ par rapport à celle $-COOH$ est de 100
- D est élué en dernier sur une colonne échangeuse de cations
- E est l'acide aminé qui neutralise le proton dans la sous-unité c de l'ATP synthase

QUESTION N° 12 (1 point)

On peut dire que

- A Un acide est une molécule qui accepte un proton H^+
- B Une molécule oxydée est une molécule qui accepte un e^-
- C Le $NADH, H^+$ est une forme réduite du nicotinamide adénine dinucléotique phosphate
- D Si une molécule a un $E'_0 < 0$, cette substance a une affinité plus faible pour un e^- que pour H^+
- E Lorsque la concentration des ions H^+ augmente dans une solution, le pH augmente

QUESTION N° 13 (1 point)

Une sphingomyéline et un ester plasménique ont en commun :

- A de contenir un phosphate
- B de contenir une liaison éther
- C d'être non polaire
- D d'être des lipides membranaires
- E d'être des dérivées de l'acide phosphatidique

QUESTION N° 14 (1 point)

L'oxaloacétate

- A est un acide dicarboxylique comprenant 4 atomes de carbone
- B se condense avec une molécule d'acétyl CoA pour former le citrate et libérer le coenzyme A
- C est régénéré à partir du succinate dans le cycle de Krebs en donnant un FADH_2 et un NADH, H^+
- D est un activateur allostérique de la phosphofructokinase
- E intervient dans la navette malate-aspartate

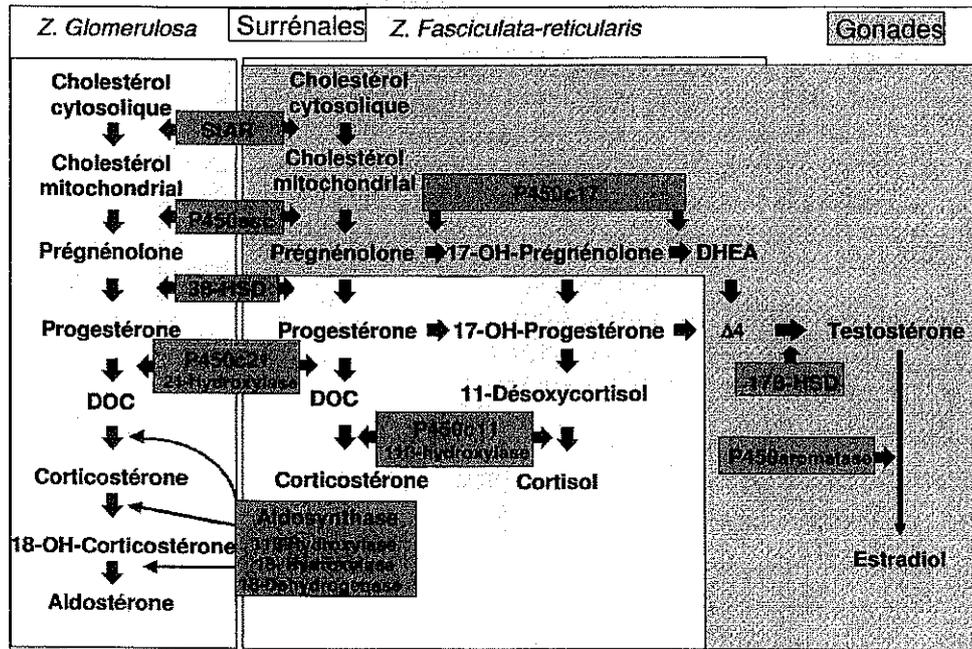
Tous les énoncés des questions suivantes peuvent servir à répondre aux questions 15 à 30.

A la fin du fascicule se trouvent deux séquences dont le début est identique et qui concerne la protéine SF1:

1- la séquence 1

2- la séquence 2 correspond à la séquence complète de l'ADN complémentaire. Le site d'initiation de la traduction se trouve dans la ligne commençant par le nucléotide 181.

QUESTION N° 15 (2 points)



Parmi tous ces enzymes cytochrome P450 participant à la biosynthèse des stéroïdes, lesquelles diminuent la taille d'un stéroïde

- A P450scc
- B P450c21
- C P450c17
- D P450c11
- E P450aromatase

QUESTION N° 16 (2 points)

Tous ces enzymes sont exprimés spécifiquement dans les tissus stéroïdogéniques, en particulier les surrénales, les gonades et le placenta. L'activité de ces enzymes et la transcription de leurs gènes sont augmentées par des hormones trophiques via une voie dépendante de l'AMP cyclique (cAMP). Au sujet de cette voie, on peut dire

- A Les récepteurs transmettant le signal des hormones ACTH et LH sont des récepteurs tyrosine-kinase
- B le cAMP permet d'activer l'adényl-cyclase
- C la protéine-kinase A transforme l'ATP en cAMP
- D le cAMP se lie sur des CREs (cAMP response elements)
- E la protéine qui lie les CRE a un motif « leucine-zipper »

QUESTION N° 17 (2 points)

En fait l'augmentation de la transcription des gènes de ces cytochromes P450 est retardée et nécessite une synthèse protéique, ce qui n'est pas classique pour la voie dépendante de l'AMP cyclique (cAMP). Donc deux équipes en parallèle ont essayé d'isoler une protéine spécifique qui serait un facteur de transcription. Dans un premier temps, elles ont recherché des enhanceurs communs pour tous ces enzymes qui stimulent la stéroïdogénèse. Un enhanceur commun a été isolé. Quelles sont parmi les manipulations suivantes celles qui ont participé à la détermination de cet enhanceur ?

- A Transfection des plasmides construits dans des lignées cellulaires obtenues à partir de tissus stéroïdogéniques.
- B L'ADN cloné dans chacun de ces plasmides correspond à l'ADNc de l'un des gènes codant pour ces cytochromes P450
- C Des délétions partielles de l'ADN cloné permettent de déterminer cette séquence spécifique.
- D Lorsque cette séquence spécifique est présente dans la construction, la transcription du gène « reporter » est diminuée.
- E l'ARN polymérase DNA polymérase est l'un de ces gènes « reporter »

QUESTION N° 18 (2 points)

Un motif commun 5'AGGTCA 3' a été isolé. Pour vérifier que ce motif est bien l'enhancer attendu, deux expériences supplémentaires ont été utilisées :

- 1) l'une montrant sur un gel de séquence le retardement de migration d'un fragment d'ADN marqué contenant ce motif AGGTCA par des extraits cellulaires,
- 2) l'autre peut être déduite des items ci-dessous.

On peut dire que

- A l'extrait cellulaire peut-être un extrait membranaire de cellules de surrénales ou de gonades
- B sur l'autoradiographie de l'autre méthode, la zone contenant la séquence de l'enhancer correspond à une zone peu radioactive
- C l'extrait cellulaire est un extrait nucléaire de cellules qui n'ont pas d'activité stéroïdogénique
- D l'autre méthode utilise la DNase I
- E sur l'autoradiographie de l'autre méthode, la zone contenant la séquence de l'enhancer correspond à une zone très radioactive

QUESTION N° 19 (1 point)(1 seule réponse juste)

Ce motif AGGTCA a permis de purifier cette protéine à partir d'extraits de surrénales bovines. Parmi les méthodes suivantes, laquelle correspond à la dernière étape

- A chromatographie d'affinité avec des billes sur lesquelles des anticorps spécifiques de cette protéine sont attachés
- B chromatographie d'affinité avec une partie de la séquence d'ADNc
- C Western blot utilisant des anticorps spécifiques de cette protéine
- D chromatographie d'affinité avec une partie du promoteur du gène de la 21-hydroxylase, CYP21
- E chromatographie d'affinité avec la séquence codante du gène « reporter »

QUESTION N°20 (1 point)

Le séquençage partiel de cette protéine a permis d'isoler l'ADNc bovin de cette protéine. Secondairement l'ADNc humain de cette protéine a été isolé et correspond à la séquence 2. La sonde d'ADNc bovin a été utilisée. On peut dire.

- A la température d'hybridation de cette banque humaine est plus haute que celle utilisée pour hybrider la banque bovine correspondante
- B la banque utilisée est d'origine génomique
- C la banque utilisée est une banque d'ADNc de surrenale humaine
- D le marquage de cette sonde d'ADNc bovin utilise la méthode du multi-amorçage au hasard avec comme molécule marquée, l' $\alpha^{32}\text{PdCTP}$
- E la longueur du plus grand clone devrait être de 3119 pb si la longueur de la queue polyA est exclue.

QUESTION N°21 (3 points)

L'analyse de la séquence d'ADNc humain permet de déduire la séquence de la protéine qui s'appelle SF1 « steroidogenic factor 1 ». Cette protéine

- A est formé de 461 acides aminés
- B est déduite d'un cadre de lecture qui se termine par le nucléotide 1570 correspond au troisième nucléotide du dernier codon qui code T
- C contient dans sa partie N-terminale plusieurs résidus cystéinyls
- D les résidus d'acides aminés 50 à 54 classiquement favorisent la formation de dimères
- E les résidus d'acides aminés 31, 32 et 34 se trouvent dans une hélice α qui lie la séquence d'ADN 5' AGGTCA 3'

QUESTION N°22 (3 points)

La délétion isolée de certains exons incluant les parties introniques entourant cet exon délété entraîne un décalage de lecture. Lesquels ou lequel ?

- A exon 2
- B exon 3
- C exon 4
- D exon 5
- E exon 6

QUESTION N° 23 (2 points)

La mutation C13X

- A transforme un résidu cystéinyl en résidu asparaginyl
- B peut donner un peptide de 13 acides aminés
- C est due à la mutation du nucléotide G en A
- D peut donner une protéine de 383 acides aminés amputée de sa partie C-terminale
- E peut donner une protéine de 384 acides aminés amputée de sa partie N-terminale

QUESTION N° 24 (2 points)

Pour étudier cette mutation, on réalise une amplification de l'exon la contenant en utilisant des amorces qui permettent de séquencer les sites donneurs et accepteurs. L'amorce en 5' se trouve dans la ligne se terminant par le nucléotide 3960. Quelles sont les bonnes amorces ?

- A 5' CCAGGCTGCCGGTCTCCGCC 3'
- B 5' CAGCGCCTGGGCACAGAGAG 3'
- C 5' GCGCGCGGCAGGGGTGCGGC 3'
- D 5' GCCGCACCCCTGCCGCGCGC 3'
- E 5' CGCGCGCCGTCCCCACGCCG 3'

QUESTION N° 25 (3 points)

Après avoir amplifié ce fragment, on étudie cette mutation.

Enzyme de restriction Mae III 5' T ↓ GTGA 3'. On peut dire

- A le fragment amplifié est de 200 pb
- B le fragment amplifié est de 240 pb
- C le profil d'un patient homozygote pour cette mutation montre une seule bande
- D le profil d'un sujet hétérozygote pour cette mutation montre trois bandes
- E le profil d'un sujet normal est de deux bandes

QUESTION N° 26 (1 point)

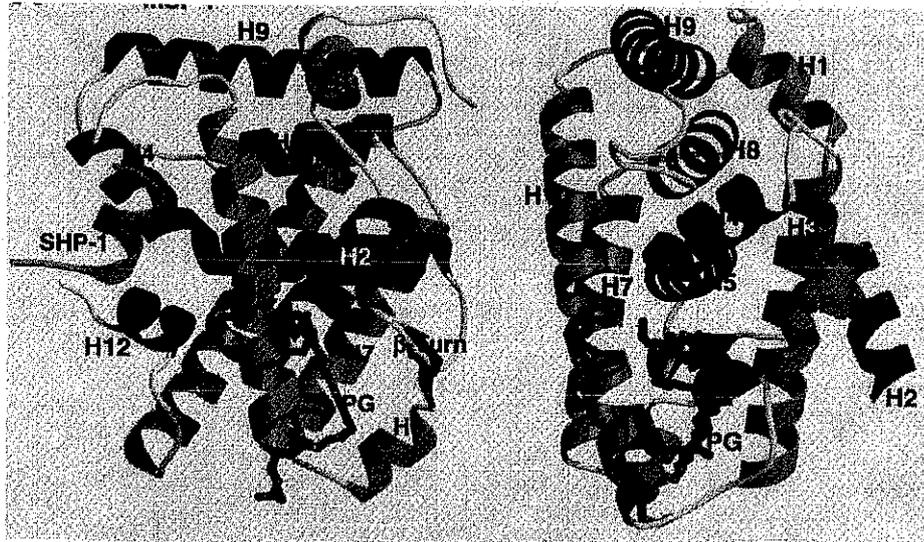
La protéine SF1 est un récepteur nucléaire qui lie un enhancer AGGTCA sous forme monomérique contrairement à la plupart des autres récepteurs nucléaires. Pour stabiliser cette liaison, la séquence GR₉₂NKFGP, appelée boîte A, lie l'ADN. Un patient homozygote pour la mutation R92Q confirme l'importance de cette région car il présente une insuffisance surrénale et une dysgénésie gonadique.

On peut dire que

- A cette mutation est due à la mutation du nucléotide G6736 en A
- B cette mutation est due à la mutation du nucléotide G6737 en A
- C la liaison à l'ADN de cette boîte A de la protéine SF1 mutée est déstabilisée à cause de la perte d'une charge négative
- D la plupart des récepteurs nucléaires sont des hétérodimères grâce à leur domaine leucine-zipper
- E R92 précède un site potentiel de « N-glycosylation »

QUESTION N° 27 (2 points)

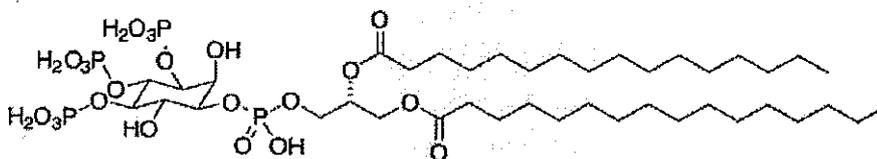
Pour déterminer la structure de l'un des domaines de la protéine SF1, la séquence d'ADN correspondant aux acides aminés 222 à 461 a été cloné dans un plasmide d'expression de sorte que l'extrémité C-terminale se termine par 6 acides aminés « histidine » supplémentaires. Après purification et cristallisation, la structure de ce domaine correspond à la figure suivante (vue du domaine sous deux angles différents). SHP-1 correspond à une partie d'un co-répresseur, PG au ligand.



- A la purification de cette protéine se fait grâce à son étiquette
- B l'extrémité C-terminale de la protéine obtenue après l'expression du vecteur recombinant est H-H-H-H-H-N-V-P-E-L
- C la figure ci-dessus représente la structure quaternaire de la protéine
- D la figure ci-dessus utilise une représentation en ruban
- E le domaine représenté par la figure correspond au domaine de liaison d'un ligand

QUESTION N° 28 (2 points)

Lors de cette modélisation, un ligand PG a été identifié. C'est un lipide dont la formule est la suivante :



On peut dire que

- A il s'agit d'un triglycéride
- B il contient deux chaînes d'acides aminés
- C il contient un ose trisphosphate
- D donne un phosphatidyl inositol trisphosphate après l'action de la phospholipase C
- E il s'agit d'un sphingolipide

Séquence 1

GGAGGACGGACGGACAGGGCCAGCCTGCTGTCCGGCTGCCGCCGCCCGTGGTGTGAGGGGGTTTCTGCCAC 72
CCACAGTCCGCCACCGTCCCACCTGGGCTGCCGGAGCCTCCCCCTGGACCCCTGGTGCCCACTGCCACCTCA 144
TCCGGTGTGAGAGCGCTGCTTCCGCTTCGTAAGTGAAGGCGCTTGGCCAGGCCTCTAATGGCCAGCCTCCTT 216
CTGCACTGCCCTCCCTGCAGCCTGGGCGGGCCACTGCCTGCTCCACTTAGTGACCGGTGCCCCCTGCTCTGGG 288
CAGGGAAGGTGCCAGGGGTCTCCAGACTCACCTCCACCATGCTACCCCTCTTCTTCCCTGGGCTGCCG 360
ATGGGAGGAGCCAGGCCTGTGACCACAAGCCCCATGGAGGTCCGCTGGTGTCCCGTGTCTCTGGCCAGGCA 432
TCCACCAGGCTGTGTGGGTCTGAGCCTAGAGGTCTGTGAGTGTGTCTGGAGGGCTCTGGGCCCTGTCTTGGGG 504
CTCTGTCCCCACCTGAGTTTCTGTGTTGCTCTGTGGCTCTCTGAGTCTGTGTGTATCCGTGTGTCT 576
GTCTCCGGCATCCATCTGTACCTGGGTTCTGCCGCCGGCTGGAACTTGGACCTGTGGGGTGGGTGCCCTCGGAGCCTG 648
GATGAGGGCTGGAACTCCCAGGCGGTGGGCTGGAACTTGGACCTGTGGGGTGGGTGCCCTCGGAGCCTG 720
GAGTGCAGAGGAGCCATGGGCTTCCAGCCGGTGTAGAGTAGAAACCTTGGGGAGAGGGGATTGGAGATAGAG 792
GGAGAAGGGCTGAGCCAGCCCCACCAACCCCGCGGGCTGGGGCTGTGTCTCAGCTCCAGTTCAGTTCCTCA 864
GGAGACTCGGGTGTGCTTACCTTAGAACAGCCAACCTGGGGCACCCAGCCCATATCCACCAACCCCTT 936
CCCTCTGTGAGTCACTGTCTGAGGGCTCCATTGTGGTGGGGAGCTGTTTCAAGAGCGGCAGTGACCCCACT 1008
CCAGCTCGGGCAGATGCTCTTGTCTTCTCAGGGGTTTGGGGTGGGGGGTACGATTAACAGGGTCTCTGGC 1080
TCTTTCTTCCCTCCTCATCCAAGATTCTTGCATGGAAACCGCAGTACCCCATCTGFAAAGCCAACATGTTCC 1152
ACTTTATGCTCGGAGAAACTGAGGCTCTCGTTTCCAGAGAGGAGAAGTACTTGCAGGCTTGCACAGCTA 1224
GTCTGCGGGTCTGTCTCTTCTGTGGGTGAGCCAGCTGAGTCCGGTCTGCGCGGGGGTCTCTGGCCACAA 1296
CTTGCAGGCCCCAGGCCAGCTTCCAACGAGCCCGAGCCCTCCAGGCCGGCGGGAGCCAGATGGAGCCTT 1368
TTGGTGGGTGCGCCGCTGTCAATGCAAAATGGCGTATTTTCTGCTTGAAGATCGCAGCTTTTGGCCCTC 1440
ATGGTCCCAACTCTCTCTGACCCAGGTGCCAGCGGGGGCGGAGGCAGGATGCAAGCAGGCCCTTCCCTT 1512
TCGGGTCCGTTTCTCCAGTGA AAAACAGTGGCCCGCGGATGGTCCCTAGCCAGAGCAACTGCTTGGCCC 1584
ATAGGCCCGGGAACCTTTGTCCTCCAGAGCCCGAGCCCTTCTTCTCGCCCCACCCCCCGTGAGAGC 1656
TCCGAGGGGAGCAGCGGGCCCTGCCCTCCGGGCTCCTCTCCAGCTCCTGCGCCGTAGAGGCTTCCGGTCC 1728
CCCCAGGCCCTCGACATTTGTA AAATGGGATTTCCACGAGTCCCATTTCCAGCCAGCAGCTGAGGCTAGC 1800
GAGCAGAAGGGGTTCCGCCAGATCTCCAGAGGGGGCCACTGAGCGCTTCCGCCAGGCCACACCGGTGC 1872
TTTGGCCCTACTGCTTGGGTTTATCTTCCAGCCACCTTAGCGAGGAAGAGCGGTATCTACCCACTTCA 1944
CAGATGCGGCAACTGAAGACCCGAGACGGGGAGAGAGCTGCTACAGGCTCGAGTGAGGAGATGTGGTGGGC 2016
CCATGGGAGAACTGTGGCTTCCACCCGTGAATCCGAGCCTCGGCCACGGTCCGCACTGCCCCATTACCCACTC 2088
CACGCTGTAGTGCAGCCAGCCACCCCGGTGGCATTGCGCACGACCCAGCTGGGTTTACAGACAGGAGCTGCC 2160
TCCCTCCACTCGAGGCGGGTGACCCAAAGGCCCCCACTGCCCCATACCTGCCAGATCGCGGGGCACTCGGGC 2232
TCCGCACTCTCTGGCCGACTCGGGTCCGGGAGCCGGGCTGCGGCTGGGGGGCTGGAGGATGAGGGCG 2304
AAGACTCGAGCTGACTGCGGAGAGGGTGC CGCGCGGGCCGGCTTGGGGGAGGACACGACCGGGCTCC 2376
GCCCCCGGATACCGCACCGCCGATATAGAGGGCCCGCCCCCGGCCCGCGCGGGTGGCGCAGGGC 2448
GGGGTGGCGGCTGAGGGGCCCCGGTCCAGGTGCGCCTCGGAGCGGCCCGCGCGGGTGGCGGGGGAACA 2520
GGGAGGCAGGATGGGGCCGGGAGGCGGGGAGCGGAGCCCGGCCCGCCCGCGGGTGGCGGGGGAACA 2592
GCTTCCCGTCCCGGCCGAGCGGCCGCTCATCGCGGGCGGGCTGCGCCCGGGTGGCGCTCCTCGCC 2664
GGCGGGCGGGCGCGCGGAGAGGTGCGGGCTCCGCGGGGGCGGACCGGGGGCGCTGCTCGCAGGAGCGAC 2736
CAGGTGGGCGCGGGCCAGCCGAGCCGCGCTCGCGCCCTCTCGGCTCCCGGGCCTGGCTCTTCCGCT 2808
TCCCTCCTCCGACTTCCGCTTCCGCCCCAGGTCCGCTCCTCGCTCGGGCTCGGCTTCCCGGAGCCCGCC 2880
GACCCCGCCCGAGGCTCCTCCCGCCGAGCTCCGGTGGCAGCGCGGCTCGGGCTGCGCTCCCGTTGCCG 2952
TCCGCTGCTGCGGCTTTGCTTCCCGGGGAGGGGTTCCCGCGGCTCTGCGTTTTCGGGCTCTCGGTCGCCG 3024
CGCCGCTGTTTTCCGCTCCTGCGGCCCGGGCCCCCGTCTGGGTCGGACCTTGGGGGCTCGCTTGGCGC 3096
TTGGGCTCCTGCGCGGAGGTTTCCGCGGCTCCCTGTCCCGACCTCGGGCTGGGCGCTTCCGCTCCAG 3168
GTTTTCTCTCCGTTTCTGTAGCTCCCGGCTCGTCTCTCCGGGCTCCTCTGCTCTCGGCTCTCTGTC 3240
TCTGCTATCGTTGTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 3312
TCCGAGTCTCTGATCCAATCCCCACCCCTCGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3384
CTCTCTCTCTCTGCT 3456
GGCGTCTCAGTGTCTGTGCTCTGCGCACTCTCGGACTGGCCTCGTGTCTGCGCCCGGTGGCGTGGTCTG 3528
TCTCTCGCCAGAGTTCTTCCCATCTTGGCTCCCGGACTCTGTCTGCTCTGCTCTGCTCTGCTCTGCT 3600
TCCAGAACACCCCGGGCCGGGCGCCAGCCTCTGGGCTCCCTGGGCTCTTCCCGCGGCTGAGCCCATC 3672
TCTTTTTCTCAGAAATCGGGTTTTGTTCTCAGACAAACGAATCCCAATCGAAGCGCATCGAACCGAGGGCC 3744
CCAGATCGGGCGGACGGGGTCCGCGCGGGAGGAGCCCCCGGACGGGGCGGAGCGTGGGGCCTGGCAGGG 3816
GAAGGCCGGGCGGGGACTGCACCGGACCCACCCGGGACGCTGCGCGGGGGCTCGGTGGGTGCCGGCGG 3888
CAGCGCTTGGGCACAGAGAGGGGATACGCGACGGGCCGGGACAGCGGCTGGGGCGGCGGCGGGGAGC 3960

M 1

CCAGGCTGCCGGTCTCCGCGGCCCTCCCTGACCCGCTGTCCCTCCGAGGCGGACGCGCGGGC ATG 4028
D Y S Y D E D L D E L C P V C G D K 19
GAC TAT TCG TAC GAC GAG GAC CTG GAC GAG CTG TGC CCC GTG TGC GGG GAC AAG 4082
V S G Y H Y G L L T C E S C K 34
GTG TCC GGC TAC CAC TAC GGA CTG CTC ACG TGT GAG AGC TGC AAGGTGAGCGGCGA 4138

G F 36

GCGCGGGCAGGGGTGCGGCAGACGGGCGCTGCGGGCCCTCGCTGACTCTCAGCTCCCGCAGGGC TTC 4206
F K R T V Q N N K H Y T C T E S Q S 54
TTC AAG CGC ACG GTG CAG AAC AAC AAG CAC TAC ACG TGC ACC GAG AGC CAG AGC 4260
C K I D K T Q R K R C P F C R F Q K 72
TGC AAG ATC GAC AAG ACG CAG CGC AAG CGC TGT CCC TTC TGC CGC TTC CAG AAA 4314
C L T V G M R L E 81
TGC CTG ACG GTG GGG ATG CGC CTG GAAGGTGCGCGCCGCGGGGCTCGACCCCTGACCCCC 4376
...AGAGCGGTGAGATTTGTTTGGAAAGGATCTGTGGTGGGGCAGCATGGAAGGGGGCTTAGAGAGGGTGTAG 6680
A V R A D R M R G G R N 93
TCTGACCTCTGCCCTGCCACAGCC GTG CGC GCT GAC CGT ATG AGG GGT GGC CGG AAC 6740

K F G P M Y K R D R A L K Q Q K K A 111
 AAG TTT GGG CCG ATG TAC AAG CGG GAC CGG GCC CTG AAA CAG CAG AAG AAG GCA 6794
 Q I R A N G F K L E T G P P M G V P 129
 CAG ATT CGG GCC AAT GGC TTC AAG CTG GAG ACA GGG CCC CCG ATG GGG GTG CCC 6848
 P P P P P P A P D Y V L P P S L H G P 147
 CCG CCG CCC CCT CCC GCA CCG GAC TAC GTG CTG CCC AGC CTG CAT GGG CCT 6902
 E P K G L A A G P P A G P L G D F G 165
 GAG CCC AAG GGC CTG GCC GCC GGT CCA CCT GCT GGG CCA CTG GGC GAC TTT GGG 6956
 A P A L P M A V P G A H G P L A G Y 183
 GCC CCA GCA CTG CCC ATG GCC GTG CCC GGT GCC CAC GGG CCA CTG GCT GGC TAC 7010
 L Y P A F P G R A I K S E Y P E P Y 201
 CTC TAC CCT GCC TTT CCT GGC CGT GCC ATC AAG TCT GAG TAC CCG GAG CCT TAT 7064
 A S P P Q P G L P Y G Y P E P F S G 219
 GCC AGC CCC CCA CAG CCT GGG CTG CCG TAC GGC TAC CCA GAG CCC TTC TCT GGA 7118
 G P N V P E L I L Q L L Q L E P D E 237
 GGG CCC AAC GTG CCT GAG CTC ATC CTG CAG CTG CTG CAG CTG GAG CCG GAT GAG 7172
 D Q V R A R I L G C L Q E P T K S R 255
 GAC CAG GTG CGG GCC CGC ATC TTG GGC TGC CTG CAG GAG CCC ACC AAA AGC CGC 7226
 P D Q P A A F G L L C R M A D Q T F 273
 CCC GAC CAG CCG GCG GCC TTC GGC CTC CTG TGC AGA ATG GCC GAC CAG ACC TTC 7280
 I S I V D W A R R C M V F K E L E 290
 ATC TCC ATC GTG GAC TGG GCA CGC AGG TGC ATG GTC TTC AAG GAG CTG GAGGTG 7334
 AGTCTCTCCTCCCGCCCGGCTGGCCCTGCATCATGGTCTCCCGGCTGCCCAAGCCTTAGCCCGACTGTCTCT 7406
 V A D Q M T L L Q N C W 302
 ...CCCGCTGCCTCTTCCCTCCAGGTG GCC GAC CAG ATG ACG CTG CTG CAG AAC TGC TGG 14307
 S E L L V F D H I Y R Q V Q H G K E 320
 AGC GAG CTG CTG GTG TTC GAC CAC ATC TAC CGC CAG GTC CAG CAC GGC AAG GAG 14361
 G S I L L V T G Q E 330
 GGC AGC ATC CTG CTG GTC ACC GGG CAG GAGGTGACCAGTCCCTCGGCCCGCCCTGGGCC 14422
 V E L T T V A 337
 ...TCCTCTGACTGTCCCGCCCGTGCCTCTGCCCCAGGTG GAG CTG ACC ACA GTG GCC 16213
 T Q A G S L L H S L V L R A Q E L V 355
 ACC CAG CGG GGC TCG CTG CAC AGC CTG GTG TTG CGG GCG CAG GAG CTG GTG 16267
 L Q L L A L Q L D R Q E F V C L K F 373
 CTG CAG CTG CTT GCG CTG CAG CTG GAC CGG CAG GAG TTT GTC TGC CTC AAG TTC 16321
 I I L F S L 379
 ATC ATC CTC TTC AGC CTGGGTGAGTGCAGCTCCAGTGCAGAGAGGTGGAGACAGCCAGAGGCCG 16386
 D L K F L N N H I 388
 ... GACGGATGGTGACCAGAACCTCCCTTTGCAGAT TTG AAG TTC CTG AAT AAC CAC ATC 24441
 L V K D A Q E K A N A A L L D Y T L 406
 CTG GTG AAA GAC GCT CAG GAG AAG GCC AAC GCC GCC CTG CTT GAC TAC ACC CTG 24495
 C H Y P H C G D K F Q Q L L L C L V 424
 TGC CAC TAC CCG CAC TGC GGG GAC AAA TTC CAG CAG CTG CTG CTG TGC CTG GTG 24549
 E V R A L S M Q A K E Y L Y H K H L 442
 GAG GTG CGG GCC CTG AGC ATG CAG GCC AAG GAG TAC CTG TAC CAC AAG CAC CTG 24603
 G N E M P R N N L L I E M L Q A K Q 460
 GGC AAC GAG ATG CCC CGC AAC AAC CTG CTC ATC GAA ATG CTG CAA GCC AAG CAG 24657
 T * 461
 ACT TGAGCCTGGGCCGGGGCGGGCCGGGACTGGGGCGGGACTGGGGCGGGCCCTGGGGCGGGCCG 24726
 CAGCCACACCGCTGGCTCCGCATGGTTCATTTCTGATGCCACCGAGGAGCCCCAGCCCCGTCCAGAGGC 24798
 CGTGCCCTGAGTTCTGACACTGTGTGTTTGGGAAGGTGGGTGAGGCTGGGCAGGGCCTGGCGAGGTGGA 24870
 GTGGCCACTGGCACTTGCTGCTGCTTGGAGTGCCTCAAGGAGGTGGCTGTTAACCACCCGCCCGCCCT 24942
 CCCGCTCCAGCTCTCTCTCTGAGTCTGAAGCTGCAGGTCCGGGGAGGAGGTTCCGGATTCCCTGGTG 25014
 GGCTCGACGTCCTTGGATCAGAGGTCACTCCCTTCTCTCTCTGAAACAGACAGGGAGAGTTGAGCA 25086
 GGATCAACTAGGGGAGGAGAGGGTCTCCAGTGTTCCTCCCATAGAGACCAGGAGGGAGAGCCTCTGTTT 25158
 TGAAACTAAGGATAACCGAGTTTGTCTAAATTGAGAGGGGCTATTGGCCCTAGAGGACACTAGGAGACTGG 25230
 TTAGGACAAAAGACCTTCTCCCTAGCCCTTCTACCCACCTGACCTCTGCAAGAGGGGGCATTTGATACATC 25302
 ATCGGAAAAAACTTTGCTCCAGGCATCACTGATTCCTCTCCACCCAAGGAGAACGTTTGGTACAATCGA 25374
 CATCCTAGCCCCACCCAGAGGTGGCCCTCCAGGCTGGTATTTATCTGCAAGGTGTAGTCAAGAGGTTT 25446
 CTCGCCCTTTTTGTTTTAAGCTCTAGACACTCCTTGAATGTGTGTGATGGAGGAAGGGACAGAT 25518
 TTGAGGACTGAAGCTGGGGCTTGGGGATTGCCACTAAGTACAGCTGACGGTTTCTCCCGGACTCGCTA 25590
 CTAAGTACCCTTGGGGTGTGCTGGGTCACTTCTGAGCCCCAGCCCAATCCAGAGAAGCGCTGTGTC 25662
 CGCCCTCCACCCACTAGGTGAACAGCAGGATGCCCTGTTGGGGCTTCAGGTCTCTGTGGGTGGGAATGCAA 25734
 GTGAACCTGGGAGGGGGCCAGCCCTGTAGATCAGGGATAGCGCTGTTGATCCCTCTCTGTGGCTCAACC 25806
 CGTTGGGTCCCTTGCTGCAAAACCATGAAGCTGGCCCTCAGCTCCCTGACCCCTGTCTAGTCAATGAAGG 25878
 ACACCTGACAGGTGAAGCACCAGGGAGAGGCCCTCGGCTGTCTCTGTCCTCCCGGGGGTGCCTGCTGTCG 25950
 TCCCGCTTTCATGTTACTGTTGACGCTTGTGCTGAGCCTGCCAGTTGGAGGAGACTGGGCACCCCTGCCTC 26022
 CTGCCTCCCGCTCCCGCCACCTGTCTCAGTACCTCCCTCCCGCCCTGAAACATGTGCCCTGCCAA 26094
 GGCCGGAGACCCACAGCCCTGAAACGAGAGTGCCTTAAGGATCACCCAGCCCCACAGCCCTGGAATAA 26166
 ATTTGCAATTAGTTCC 26184

Séquence 2

1 ggaggacgga cggacagggc cagcctgctg tccggctgcc gcccgccgtg gtgtgagggg
 61 gtttctgccc acccacagtc gccaccgtcc cacctgggct gccggagcct cccctggac
 121 ccctggtgcc cactgccacc ctcatccggt gtgagagcgc tgettccgct tcgcgagcgc
 181 cgcgggcatg gactattcgt acgacgagga cctggacgag ctgtgccccg tgtcggggga
 241 caaggtgtcc ggctaccact acggactgct cacgtgtgag agctgcaagg gcttcttcaa
 301 gcgcaagcag cgaacaaca agcactacac gtgcaccgag agccagagct gcaagatcga
 361 caagacgcag cgaagcgcct gtccctcttg ccgcttccag aatgcctga cggtagggat
 421 gcgctggaa gccgtgcgcg ctgaccgtat gaggggtggc cggaaacaagt ttgggcccgat
 481 gtacaagcgg gaccgggccc tgaaacagca gaagaaggca cagattcggg ccaatggctt
 541 caagctggag acagggcccc cgatgggggt gcccccgcog cccctcccg caccggacta
 601 cgtgctgect cccagcctgc atgggcctga gcccaagggc ctggccgcog gtccacctgc
 661 tgggcccactg ggcgactttg gggcccagc actgcccctg gccgtgcccg gtgcccactg
 721 gccactggct ggctacctct accctgcctt tccctggcct gccatcaagt ctgactacc
 781 ggagccttat gccagcccc cacagcctgg gctgccgtac ggctaccag agccttctc
 841 tggagggccc aacgtgectg agctcatcct gcagctgctg cagctggagc cggatgagga
 901 ccaggtgccc gcccgcatct tgggtgctc gcaggagccc accaaaagcc gccccgacca
 961 gccggcggcc ttggcctcc tgtgcagaat ggcgaccag acctcatct ccaatgggga
 1021 ctgggcacgc aggtgcatgg tcttcaagga gctggaggtg gccgaccaga tgacgtgct
 1081 gcagaactgc tggagcagc tgctggtgtt cgaccacatc taccgccagg tccagcacgg
 1141 caaggagggc agcatcctgc tggctaccgg gcaggaggtg gagctgacca cagtggccac
 1201 ccagggcggc tcgctgctgc acagcctggt gttgcccggc caggagctgg tgctgcagct
 1261 gcttgcgctg cagctggacc ggcaggagtt tgtctgcctc aagttcatca tctcttcag
 1321 cctggatttg aagttcctga ataaccacat cctggtgaaa gacgtcaggt agaaggccaa
 1381 cgcgcctctg cttgactaca cctgtgccca ctaccgcac tgcggggaca aattccagca
 1441 gctgctgctg tgccctggtg aggtgcgggc cctgagcatg caggccaagg agtacctgta
 1501 ccacaagcac ctgggcaacg agatgcccc caacaacctg ctcatgaaa tgctgcaagc
 1561 qaagcagact tgagcctggg gccggggcgg ggcgggact gggggcggga cgtggggcgg
 1621 ggctggggcg gggccgcagc cacacogctg gctctgcatg gttcattttc tgatgccac
 1681 cgaggagccc cagcccctgc ccagagggcg ctgcccctga gttctgacac tgtgtgtttg
 1741 ggaagtgggt gaggctgggc agggcctggc ggaggtggag tggccactgg cacttgctcc
 1801 ctgcttggag tgcccgaagg aggtgctgt taaccaccog ccccgcccc tccctgtcc
 1861 cagctctctc tccctggagt gatcagaggt catcccttcc tccctctctg gaaacagaca
 1921 gtgggcccctg agtcctctg caactagggg agggagaggg gtctccagtg tccccccat
 1981 gggagaagtt gagcaggtat tctgttttgt aaactaagga taaccgagtt tgctaaattg
 2041 agagaccagg agggagagcc tggggccta gaggacacta ggagactggt taggacaaaa agacctctc
 2101 agaggggcta ttgggccta tgacctctgc aagagggggc attgatacat cctcgggaaa
 2161 cctagcccctt ctacccccacc ctgattccct ccccaccca aggagaactg ttggtacaat
 2221 aaactttgct gcccccacca gaggtggccc tcccagctg gtatttatct gcaaggtgtg
 2281 cgacatccta gcttttttgt ttttaagctt ctgacactc ctagacactc cttgaaatgt
 2341 agtcaagagg ttttctccc cgttttttgt aggactgaag ctggggcttg ggggtgcca
 2401 gtgtgtgatg gaggggaagg gacagattg actcgcttac taagtacctg tgggtgctg
 2461 ctaagtacag ctgatggttt cccccggac atccagagaa gcgctgttgc ccgcccctca
 2521 ctgggtcatt acttctgagc cccagcccca atccagagaa ggcgtgttgc ccgcccctca
 2581 cccactaggt gaacagcagg atgccctgt gggggcttca ggtctctgtg ggtgggaatg
 2641 caagtgaact tgggaggggg cacgggctg tagatcaggg atagcgtgtg tgatccccct
 2701 tctgtggctc caaccgctg ggtcccttgc tgcaaaccca tgaagctggc cctcagctcc
 2761 ctgaccccct gtccctaggtc atgaaggaca ctctgcaggg tgaagcacca gggagagggc
 2821 tggctgtctc cctgtccccg ggggggtgcc tgctgtcctg cccgcttcca tgttactgtt
 2881 gcagctgtg ctgagcctgc ccagttggag gagactgggc acccctgctc cctgcctccc
 2941 gcccccgc accctgtctc cctgaaaaga gaagtgcct taaggatcac gtcgccctcc
 3001 caagcccgga gaccacagc cctgaaaaga gaagtgcct taaggatcac gtcgccctcc
 3061 acagcccctg aataaatttc gcaattagtt tccaaaaaaa aaaaaaaaaa aaaaaaaaaa

		2nd base in codon					
		U	C	A	G		
1st base in codon	U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U C A G	3rd base in codon
	C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G	
	A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G	
	G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G	