

Recherche d'une enzyme capable de participer à la production de biocarburant de 2^o génération

Présentation de l'activité

Le bioéthanol est un biocarburant. Il est produit à partir de la transformation de végétaux en éthanol par des micro-organismes fermentaires. Pour assurer une production suffisante, il est nécessaire de cultiver des surfaces importantes de plantes (colza, tournesol) qui servent de substrat aux micro-organismes réalisant la fermentation. La production de ce type de carburant a longtemps été très controversée dans la mesure où elle impliquait l'utilisation de ressources agricoles qui pouvaient être destinées à l'alimentation humaine. De nouvelles recherches ont alors tenté d'améliorer le procédé de fabrication du bioéthanol en permettant sa production à partir de déchets d'origine végétale. Le fait de valoriser ainsi des déchets permet d'éviter la compétition avec la production alimentaire. Le **bioéthanol** ainsi produit appartient à la **2^o génération** de biocarburants.

Les molécules issues des végétaux considérées comme des déchets par l'industrie agroalimentaire sont celles qui ne sont pas digérées par l'Homme. La cellulose, un constituant important de la paroi des cellules végétales, en est un exemple. La cellulose est un polymère de résidus glucose reliés entre eux par des liaisons $\beta 1 \rightarrow 4$. Dans cette activité, nous allons nous intéresser à l'hydrolyse de la cellulose des plantes qui permet de libérer les résidus glucose. Ces derniers font ensuite l'objet de la fermentation qui est à l'origine de la production d'éthanol.

Plan de la recherche :

- 1) Rechercher une enzyme capable d'hydrolyser la cellulose.
- 2) Déterminer par quels micro-organismes cette enzyme est produite.
- 3) Rechercher le gène qui code pour cette enzyme.

1) Rechercher une enzyme capable d'hydrolyser la cellulose

Les termites sont de redoutables parasites capables de détruire des arbres entiers. Cette destruction est attribuée à leur capacité à hydrolyser la cellulose. On pense que l'hydrolyse de la cellulose est en réalité assurée par des bactéries commensales du système digestif des termites. Parmi ces bactéries, certaines appartiennent au genre Pseudomonas. Le document 1 présente l'abstract d'une publication dont sont issues ces informations.

Q1. Relever, dans le document 1, les passages qui justifient les informations soulignées dans le

TRAAM

texte ci-dessus.

Le site Protein Data Base (PDB) est une base de données listant un très grand nombre de protéines ainsi que leurs caractéristiques (origine, structures primaires, secondaire, etc.). La fiche technique 1 explique son fonctionnement.

T1. A l'aide du PDB, rechercher l'enzyme qui hydrolyse la cellulose = cellulase exprimée par les bactéries du genre *Pseudomonas*, puis copier sa structure primaire au format « FASTA » dans un fichier texte que vous nommerez `strut_prim1_cellulase.txt`.

Q2. Relever le nom de l'espèce bactérienne productrice.

Le site Sequence Manipulation Suite (SMS) permet d'analyser des séquences primaires de protéines ainsi que des séquences de nucléotides. La fiche technique 2 explique son fonctionnement.

T2. Utiliser SMS pour faire passer la nomenclature des acides aminés de la cellulase du code à une lettre au code à trois lettres. Copier dans un fichier texte que vous nommerez `strut_prim3_cellulase.txt`.

Q supplémentaire 1. La séquence sauvegardée dans le fichier « `strut_prim3_cellulase.txt` » correspond à la structure primaire de la protéine, c'est à dire l'enchaînement linéaire des acides aminés (cf. cours de CBSV). A l'aide du site PDB, identifier des éléments appartenant aux structures secondaire et tertiaire de la protéine. Préciser, en justifiant, si la protéine possède une structure quaternaire. Relever le nom de la technique qui a permis d'obtenir ces informations supplémentaires.

 **Retour à la paillasse du laboratoire :** la cellulase de la bactérie a été isolée puis purifiée. On détermine son activité enzymatique dans différentes conditions. On découvre ainsi que l'enzyme possède une bonne activité de dégradation de la cellulose dans un intervalle de température assez large. Cette propriété la rend particulièrement intéressante pour produire du bioéthanol.

2) Recherche des micro-organismes qui produisent la même cellulase

On sait qu'une bactérie du genre *Pseudomonas* produit une cellulase et on connaît sa structure. On veut maintenant savoir si d'autres micro-organismes appartenant à d'autres espèces la produisent aussi. Pour ce faire, on réalise un blast, comme cela est décrit dans la fiche technique3.

TRAAM

T3. Comparer la séquence en acides aminés de la cellulase identifiée (strut_prim1_cellulase.txt) avec toutes celles qui ont déjà été séquencées dans le monde grâce à la fonction blast p du site pubMed.

Q3. Vérifier que l'enzyme est bien synthétisée par l'espèce bactérienne *Pseudomonas stutzeri* puis indiquer si d'autres espèces produisent la même enzyme.



Retour à la paillasse du laboratoire : des isolements à partir du tube digestif des termites sont réalisés. La bactérie *Pseudomonas stutzeri* est ainsi isolée, identifiée et mise en culture. Les cultures ainsi réalisées permettent une production relativement modeste de cellulase. On cherche donc à insérer le gène de la cellulase dans une bactérie plus facile à cultiver : *E.coli* pour lui faire exprimer cette enzyme. Mais pour ce faire, il faut connaître ce gène.

3) Recherche de la séquence d'ADN codant la cellulase

il faut que l'on détermine le gène qui code pour la cellulase qui nous intéresse.

Q4. Rappeler, sous forme d'organigramme, le nom et le principe des principales étapes qui permettent la synthèse d'une protéine à partir d'une séquence d'ADN chez les procaryotes.

Nous allons devoir déduire d'une séquence protéique la séquence d'ADN qui correspond, c'est à dire faire l'inverse de ce qu'indique votre schéma de la Q4.

T5. En vous aidant du logiciel SMS (fiche technique 2), rechercher la séquence d'ADN correspondant à la cellulase.

Q5. A partir de l'analyse du document 2, proposer une définition de la séquence « consensus » et de la séquence « most likely ».

Q6. A l'aide de l'analyse conjointe des documents 2 et 3, expliquer l'existence de la séquence « most likely ».

T6. Copier la séquence « most likely » dans un fichier texte « ADNmostlikely_cellulase.txt » et la séquence consensus dans un fichier texte « ADNconsensus_cellulase.txt ».

T7. Faire un BLASTN sur la séquence « consensus » afin de la rechercher dans le génôme de la bactérie *Pseudomonas*.

TRAAM

Q7. Expliquer le résultat.

Q8. Expliquer le traitement effectué par le programme TBLASTN.

T8. Faire un TBLASTN sur la séquence en acides aminés de la cellulase (strut_prim1_cellulase.txt). La page qui s'affiche liste les micro-organismes dans lesquels on trouve la séquence d'ADN qui peut correspondre à la séquence en acides aminés de la protéine. Cliquer sur « view in GBrowse » du micro-organisme qui possède une séquence d'ADN qui a le plus de similitude avec la séquence recherchée (Bit Score le plus élevé). Télécharger la séquence nucléotidique correspondant au gène de la cellulase et le sauvegarder sous le nom « gene_cellulase.txt ».

Q9. A l'aide du document 2, comparer les premiers nucléotides des séquences « gene_cellulase.txt » et « ADNconsensus_cellulase.txt ». Vérifier que les informations de la séquence consensus sont bonnes.

Q supplémentaire 2. Le blast permet de rechercher la séquence d'ADN correspondant à un gène dans le génome d'une bactérie. Nous avons vu que cette recherche pouvait s'avérer complexe. Si on devait rechercher la séquence d'ADN correspondant à un gène chez un organisme eucaryote, la recherche serait encore plus complexe. Justifier cette affirmation.