

## Complément d'analyse de la fraction de protéines membranaires

La quantité de protéines retrouvées par cette analyse figure dans le **tableau 4.7**.

**Tableau 4.7** : Nombre de protéines de la souche IFP 2173 identifiées par LC-MS/MS

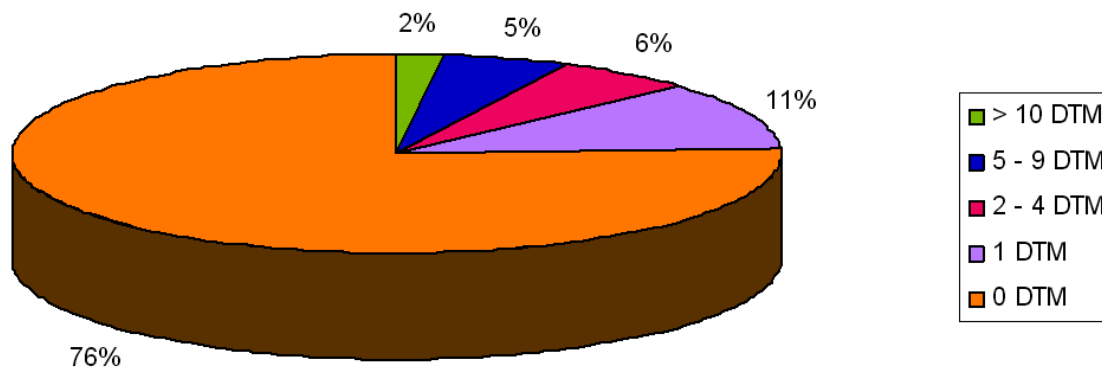
	Nombre de protéines de la fraction membranaire	
	Acétate	2-EHN
Total des protéines identifiées	919	899
Nombre de protéines spécifiques	320	300
Nombre de protéines communes	599 dont 65 induites sur 2-EHN*	

\* la valeur de l'abondance relative calculée à partir des données de spectral count est deux fois supérieure sur 2-EHN que sur acétate.

Les tableaux donnant l'intégralité des résultats de cette analyse protéomique figure en **annexe 1** page 194.

Nous avons recherché dans l'ensemble des protéines identifiées celle qui étaient intrinsèquement membranaires c'est à dire susceptibles de contenir des segments transmembranaires avec une version du logiciel, HMMTOP (Tusnady *et al.*, 2001) modifiée par le Laboratoire d'Etude Dynamique des Protéomes (CEA de Grenoble) pour traiter plusieurs séquences protéiques simultanément.

Les résultats ont montré qu'environ 70 % des protéines identifiées au cours de ces analyses ne présentent pas de domaine transmembranaire (**figure 4.10**). Bon nombre de ces protéines sont cytoplasmiques et sont certainement piégées dans des vésicules membranaires lors de la lyse des cellules. De plus, parmi ces protéines membranaires, il y a beaucoup de protéines liées à la membrane, qui n'ont pas de segments transmembranaires mais qui ont une nature hydrophobe. L'analyse ne nous permet pas de le savoir.



**Figure 4.10** : Répartition des protéines contenant ou non des domaines transmembranaires

### 3.2.2 Limites de la méthode

L'analyse protéomique mise en oeuvre a conduit à l'identification de plus 1300 protéines ce qui est remarquable compte tenu du fait que le génome de la bactérie n'est pas connu. La méthode d'identification et de quantification des protéines a toutefois ses limites :

- Lors de l'identification des protéines par le logiciel MASCOT, les peptides qui ne correspondent à aucune protéine présente dans les bases de données ne sont pas pris en compte. Bien que les bases de données soient de plus en plus riches en séquences de génomes bactériens, on peut rater une protéine d'intérêt au cours de l'analyse. Par exemple, certains spots 2D n'ont pas pu être identifiés.
- Lors du calcul du PAI (Protein Abundance Index), on estime une abondance relative des peptides et on en déduit une abondance relative de protéine. Cette estimation n'est pas tout à fait juste même si les appareils utilisés pour l'analyse sont très performants. Par exemple, il peut y avoir des isoenzymes avec un ou plusieurs peptides en commun. Parmi ces dernières, on ne retiendra que celle qui a le meilleur score et le meilleur taux de recouvrement. Il est alors possible d'écartier, au niveau de la synthèse des résultats, d'autres identifications de ces protéines. Il peut alors y avoir un léger biais.
- Le calcul du spectral count est a priori plus juste puisque nous utilisons le nombre de spectres obtenus dans l'analyse, qui dépend de l'abondance des peptides. Mais nous faisons toujours une approximation entre le résultat de l'analyse des peptides et la quantification des protéines.

### 3.3 La découverte de P450 alcane monooxygénase de type CYP153 chez la souche IFP 2173

Au cours de l'analyse d'une bande à 47 kDa, correspondant à une protéine marquée au <sup>35</sup>S, quatre cytochromes P450 ont été identifiés sur quarante protéines au total, dont un ressemblait à une alcane hydroxylase de type CYP153. Ce cytochrome est différent de la cinquantaine de P450 de *M. vanbaalenii* PYR-1, mais possède une séquence très proche du CYP153A de *Mycobacterium* sp. HXN-1500 trouvée par van Beilen *et al.* (2005b). Ce cytochrome a été largement étudié par Funhoff *et al.* (2006), tous les détails de son analyse figurent dans l'étude bibliographique pages 26 à 31.

Ce cytochrome est soluble mais des prédictions de séquence indiquent une région hydrophobe qui pourrait se lier à la membrane, ce qui expliquerait pourquoi il est retrouvé dans la fraction membranaire.

Ce CYP153 se retrouve parmi les protéines exclusivement produites sur 2-EHN. De par son classement en fonction de son abondance relative (40<sup>ème</sup> position sur un total de 899 protéines), c'est la deuxième protéine spécifique la plus abondante. Son rôle dans la dégradation du 2-EHN a été étudié dans la dernière partie des résultats.

## 4 Étape initiale de dégradation du 2-EHN

L'étude protéomique des enzymes potentiellement impliquées dans la dégradation du 2-EHN a permis de mettre en évidence une alcane mono-oxygénase. Il s'agit d'un cytochrome P450 de type CYP153. Lors d'études antérieures sur *M. austroafricanum* IFP 2173, une alcane hydroxylase de type AlkB a été identifiée. Elle serait impliquée dans la dégradation de l'isooctane (Solano-Serrena *et al.*, 2004). A partir de l'analyse de séquences de génomes de mycobactéries récemment publiées, j'ai cloné et obtenu la séquence complète du gène *alkB1* codant pour l'alcane hydroxylase membranaire de *M. austroafricanum* IFP 2173 et j'ai trouvé un second gène codant pour une alcane hydroxylase appelée *alkB2*. Afin d'étudier les propriétés de ces enzymes, j'ai entrepris leur expression en système recombinant pour savoir si l'une d'elles était impliquée dans l'étape initiale de dégradation du 2-EHN.

### 4.1 Les alcane hydroxylases de type *alkB*

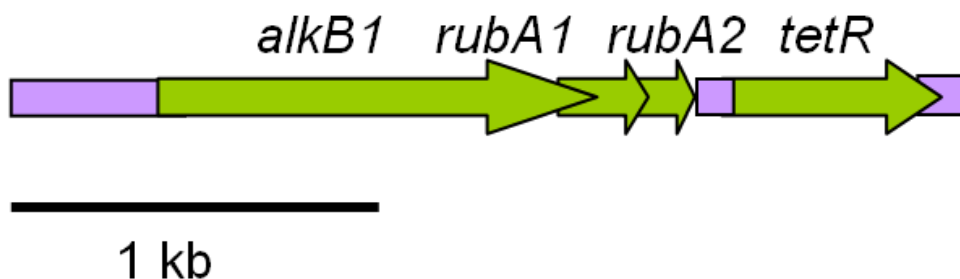
#### 4.1.1 Le gène *alkB1*

A l'origine, ce gène a été trouvé et en partie séquencé par Floriane Solano-Serrena *et al.* lors de l'étude de la dégradation de l'iso-octane par la souche IFP2173. Le numéro d'accession est AAQ88276. Nous disposons donc d'une séquence partielle qui correspondait parfaitement à l'alcane mono-oxygénase *Mvan\_1742* de *M. vanbaalenii* PYR-1. Sur la base de cette information de séquence et des séquences des gènes *Mvan\_1741*, *Mvan\_1743*, *Mvan\_1744*, *Mvan\_1745* et *Mvan\_1747*, j'ai conçu des oligonucléotides pour amplifier les gènes correspondants de la souche IFP 2173, ainsi que les gènes voisins. Les conditions de PCR ont nécessité l'ajout de DMSO à hauteur de 5 %, en raison de la séquence riche en GC et/ou de structures de l'ADN (tiges-boucles) susceptibles de gêner la répllication par l'ADN polymérase. Après avoir purifié les produits PCR de taille attendue, je les ai clonés dans un vecteur commercial (pDrive) et les ai fait séquencés (Génome Express ; Meylan 38). La taille des fragments PCR obtenus ainsi que les gènes qu'ils contenaient sont indiqués dans les **figures 4.11** et **4.12**. Le tableau 2.3b dans Matériel et Méthodes liste les plasmides construits dans cette étude (page 58).

Le gène *alkB1*, composé de 1271 pb, fait partie d'un groupe de gènes comprenant aussi deux gènes de rubrédoxines, nommées *rubA1* (174 pb) et *rubA2* (183 pb). Les gènes *rubA1* et *rubA2* sont dans la même phase de lecture que le gène *alkB1*. Ils sont enchevêtrés au gène qui les précède par une séquence **GTGA**. Dans les deux cas, le codon start est **GTG** pour gènes

*rubA1* et *rubA2* et le codon stop est **TGA** pour le gène précédant. Il n'y a pas de gène codant pour une rubrédoxine réductase comme c'est le cas pour l'opéron de *Pseudomonas putida* Gpo1 (van Beilen *et al.* 2003). Seules 31 bases séparent ces trois gènes de celui codant un facteur de transcription, nommé *tetR* (732 pb). Nous avons inclus dans le fragment amplifié par PCR une séquence non codante de 345 pb en amont du gène *alkB1* susceptible de contenir la région promotrice.

L'organisation et la séquence nucléique de cet ensemble de quatre gènes sont très proches de celles trouvées dans la région correspondante du génome de *M. vanbaalenii* PYR-1 (98 % d'identité de séquence) (**Figure 5.11**).



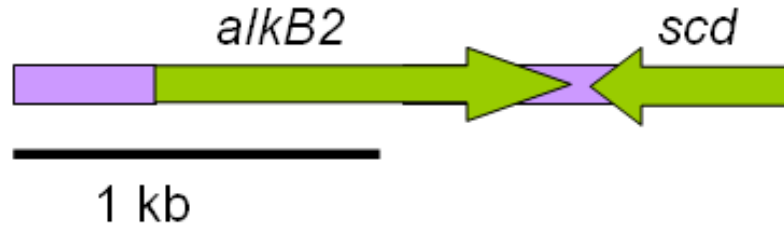
**Figure 4.11** : Carte du locus contenant le gène *alkB1*

Les gènes sont représentés par des flèches vertes, les régions non codantes sont en mauve.

#### 4.1.2 Le gène *alkB2*

*M. vanbaalenii* PYR-1 possède un autre gène qui code pour une autre alcane monooxygénase. J'ai eu accès via la base de donnée NCBI, à la séquence de ce gène *Mvan\_3100* ainsi qu'à celles des gènes voisins *Mvan\_3099* et *Mvan\_3101*, à partir desquelles j'ai conçu des oligonucléotides pour amplifier les gènes correspondant de *M. austroafricanum* IFP 2173. J'ai obtenu par PCR un produit d'environ 2600 pb. L'analyse de la séquence a montré que la séquence du fragment cloné de la souche IFP 2173 était à 99 % identique à la région homologue de *M. vanbaalenii* PYR-1. Ainsi, le gène trouvé code potentiellement pour une alcane mono-oxygénase et je l'ai appelé *alkB2* (1185 pb). L'analyse a aussi montré la présence d'un gène homologue à *Mvan\_3101* en aval de *alkB2*, qui code pour une « short chain dehydrogenase ».

Une illustration du locus contenant le gène *alkB2* est représentée par la **figure 4.12**.



**Figure 4.12 :** Représentation du locus contenant le gène *alkB2* de *M. austroafricanum* IFP 2173

Les gènes sont représentés par des flèches vertes, les régions non codantes sont colorées en mauve.

Les souches de *Mycobacterium* sp. KMS, JLS, MCS et *M. smegmatis* mc<sup>2</sup>155 possèdent un seul gène d'alcane hydroxylase de type *alkB* dans le génome, alors que les souches *M. vanbaalenii* PYR-1 et *M. gilvum* en ont deux. D'après cette analyse de génomes, les mycobactéries n'ont pas plus de deux alcane hydroxylases transmembranaires et nous avons supposé que c'était le cas aussi pour *M. austroafricanum* IFP 2173.

#### 4.1.3 Clonage et expression chez *Escherichia coli*

Dans un premier temps, le clonage et l'expression des gènes *alkB* de *M. austroafricanum* IFP 2173 ont été effectués chez *E. coli*. En effet, bien que selon les prédictions de structure, l'alcane hydroxylase comporte six domaines transmembranaires, certains auteurs ont réussi à exprimer ce type d'enzyme chez *E. coli*. L'alcane hydroxylase de *Pseudomonas putida* Gpo1 a ainsi été exprimé dans une souche de *E. coli* BL21 à l'aide d'un plasmide de type pET portant le promoteur fort du phage T7 (Shanklin & Whittle, 2003).

De même, le gène de l'alcane hydroxylase de *Gordonia* sp. TF6 a été exprimée chez *E. coli* associé aux gènes de deux rubrédoxines et d'une rubrédoxine réductase. La co-expression de ces gènes a conduit à la synthèse d'un système fonctionnel capable de dégrader les *n*-alcane de C<sub>5</sub> à C<sub>13</sub> (Fujii *et al.*, 2004). Sur la base de ces résultats décrits dans la littérature, l'expression chez *E. coli* des gènes *alkB* de la souche IFP 2173 a été entreprise.

##### 4.1.3.1 Stratégie de clonage

Les gènes *alkB* ont été insérés dans les plasmides pET9a et pET15b (Novagen), sous le contrôle du promoteur T7. Le plasmide pET15b a pour particularité de produire une protéine recombinante portant une étiquette poly-histidine en N-ter.

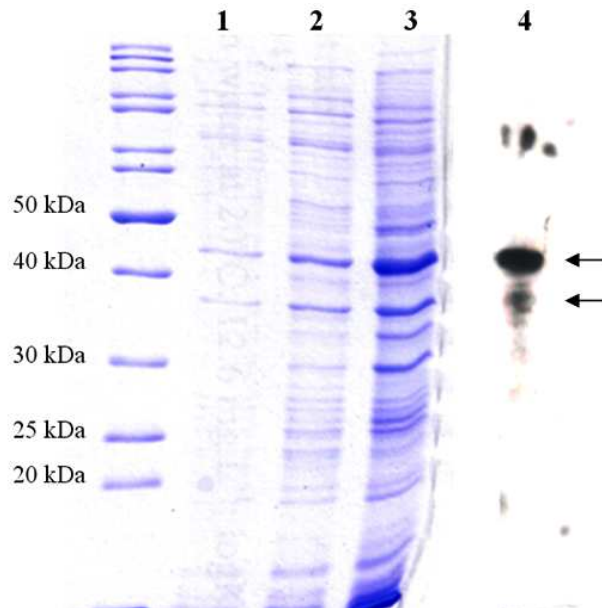
Le gène *alkB1* ainsi que les gènes *rubA1* et *rubA2* ont été amplifiés par PCR, donnant un fragment de 1629 pb qui a été inséré dans chacun des plasmides d'expression, pET9a et pET15b, via les sites de restriction NdeI et BamHI. Les plasmides ainsi obtenus sont

respectivement nommés pEN06 et pEN07. En raison de la présence d'un site de restriction BamHI dans la séquence du gène *alkB2*, la séquence codante de ce gène (1185 pb) a été amplifiée avec un site BglII à l'extrémité 3', compatible avec le site BamHI, puis inséré dans pET15b et pET9a via leurs sites NdeI et BamHI. Les plasmides obtenus sont nommés respectivement pEN15 et pEN16.

#### 4.1.3.2 Surproduction de AlkB1 dans *E. coli*

Les séquences des gènes *alkB1-rubA1-rubA2* et de *alkB2* ont chacun 12 % de codons rares chez *E. coli*. Ainsi, la souche de *E. coli* Rosetta<sup>TM</sup> qui permet la traduction des ARN contenant ces codons rares (AUA, AGG, AGA, CUA, CCC et GGA) a été transformée dans un premier temps par le plasmide pEN07 (pET15b contenant les gènes *alkB1-rubA1-rubA2*).

La souche *E. coli* Rosetta<sup>TM</sup> contenant les gènes *alkB1-rubA1-rubA2* sur le plasmide pEN07, a été utilisée pour étudier l'expression des protéines recombinantes. Afin de limiter la formation de corps d'inclusion et de favoriser la synthèse de protéine inactive, l'induction a été réalisée à 25°C pendant 14 heures. L'extrait protéique total et la fraction protéique membranaire obtenue après ultracentrifugation 30 mn à 230 000 g ont été analysés par SDS-PAGE (**figure 4.13**). L'extrait induit montre que deux bandes polypeptidiques de 46 kDa et de 38 kDa sont plus abondantes (piste 2). La fraction protéique membranaire (piste 3) contient aussi ces 2 protéines, la bande à 46 kDa étant nettement prépondérante. Cette bande est peu visible dans la fraction non induite (piste 1). Comme la protéine recombinante était marquée par une étiquette poly-histidine, nous avons analysé par Western blot anti-His-tag la présence de la protéine AlkB1. Cette dernière est bien détectée et a une taille apparente de 46 kDa correspondant à la bande majeure colorée au bleu de Coomassie. La bande à 38 kDa pourrait résulter d'une protéolyse partielle de la protéine recombinante.



**Figure 4.13 :** Expression de la protéine AlkB1 à l'aide du plasmide pEN07 dans *E. coli* Rosetta<sup>TM</sup>

1 : Extrait protéique total non induit

2 : Extrait protéique total après 14 h d'induction à 25°C.

3 : Extrait protéique membranaire, obtenu après 30 mn d'ultracentrifugation

4 : Détection par Western blot de la protéine marquée par une étiquette poly-histidine.

Ces résultats montrent que la protéine AlkB1 est majoritairement dans la fraction membranaire. La manière dont nous avons construit le système d'expression permet *a priori* de faire des essais de dégradation des alcanes. Il n'y a pas de rubrédoxine réductase pour que le système soit fonctionnel comme cela a été décrit dans la littérature. En revanche, il est possible qu'une réductase aspécifique de *E. coli* prenne le relais.

En ce qui concerne le gène *alkB2*, j'ai effectué le clonage dans les plasmides pET9a et pET15b. Cependant, je n'ai pas testé l'expression de ces constructions car il aurait fallu que je co-exprime ce gène avec au moins un gène de rubrédoxine pour produire un système enzymatique actif vis-à-vis des alcanes. Or je n'ai pas réussi à faire une telle construction.

Etant donné ces résultats, j'ai choisi de changer de système d'expression et de tester la synthèse de l'alcane hydroxylase AlkB dans une bactérie hôte phylogénétiquement plus proche de *M. austroafricanum* IFP 2173, *M. smegmatis* mc<sup>2</sup>155. Afin de pouvoir tester l'action de l'AlkB sur les *n*-alcanes et sur le 2-EHN.

#### 4.1.4 Clonage et expression chez *Mycobacterium smegmatis* mc<sup>2</sup>155

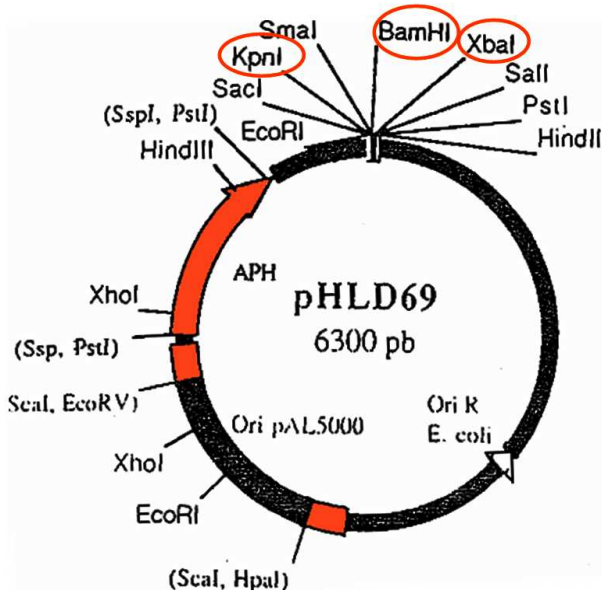
*M. smegmatis* mc<sup>2</sup>155 est une bactérie largement utilisée comme outil génétique pour exprimer des protéines de mycobactéries, provenant en général de souches pathogènes. Son



génomique a été séquencé et annoté en 2006. Elle convient *a priori* mieux à l'expression de protéines généralement retrouvées dans des corps d'inclusion chez *E. coli*. *M. smegmatis* mc<sup>2</sup>155 est équipée d'une seule alcane hydroxylase de type AlkB qui dégrade les alcanes de longueur de chaîne supérieure ou égale à C<sub>12</sub> (Lopes Ferreira *et al.*, 2007). D'après l'analyse de son génome le gène codant pour une alcane hydroxylase est suivi de deux gènes codant chacun pour une rubrédoxine, l'organisation des gènes est semblable à celle du locus *alkB1* de IFP 2173. D'ailleurs, la séquence nucléotidique de *alkB1-rubA1-rubA2* de la souche IFP 2173 présente 87 % d'identité de séquence avec la région homologue de *M. smegmatis* mc<sup>2</sup>155. D'autre part, cette souche ne dégrade pas le 2-EHN, ceci a été montré dans la première partie des résultats (Résultats 1 page 79), mais possède tout le système enzymatique requis pour dégrader des acides gras. C'est donc une souche de choix pour tester l'activité *in vivo* des alcane hydroxylases de *M. austroafricanum* IFP 2173. Ainsi, on peut faire l'hypothèse que si le gène *alkB1* est requis pour l'étape initiale de dégradation du 2-EHN, l'expression de ce gène dans la souche mc<sup>2</sup>155 pourrait rendre la bactérie hôte capable de dégrader cet hydrocarbure.

#### 4.1.4.1 Stratégie de clonage

Les gènes *alkB* de *M. austroafricanum* IFP 2173 ont été insérés dans le plasmide pHLD69 qui possède deux origines de réplication, une reconnue par *E. coli*, l'autre par *Mycobacterium* (**figure 4.14**). pHLD69 n'a pas de promoteur en amont du site d'insertion des gènes d'intérêt. C'est la raison pour laquelle nous devons insérer un fragment englobant la séquence promotrice située en amont du gène. Il a été décrit dans la littérature que le site d'initiation de la transcription est situé 96 bases avant le codon start du gène *alkB* de *Pseudomonas putida* Gpo1 (Nieboer *et al.* 1997).



**Figure 4.14 :** Carte du plasmide pHLD69. Les sites utilisés pour insérer les gènes *alkB* sont entourés en orange. D'après Pascal Poupin, Université de Metz.

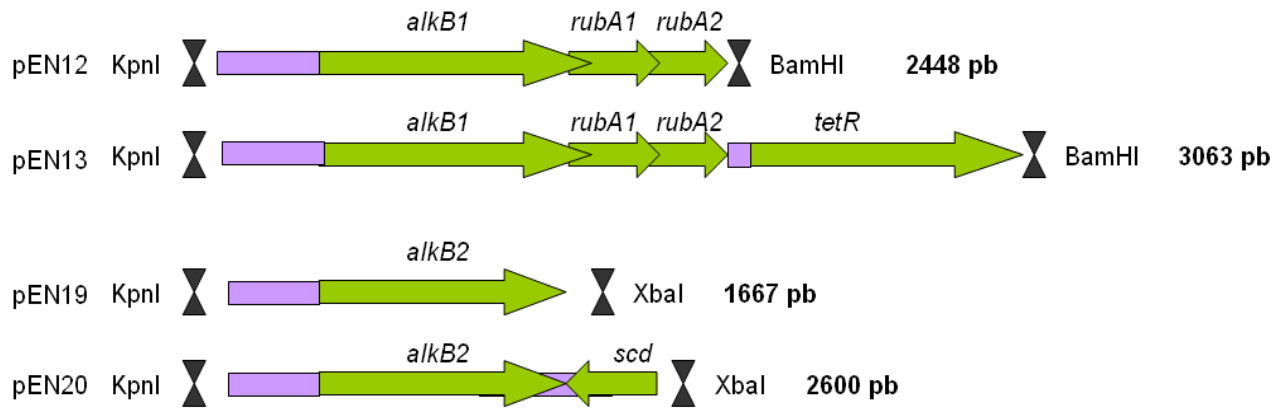
Les séquences amont des gènes *alkB* ont été analysées à l'aide du logiciel Neutral Network Promoter Prediction ([http://www.fruitfly.org/seq\\_tools/promoter.html](http://www.fruitfly.org/seq_tools/promoter.html)). Deux régions situées respectivement à 95 pb et 108 pb en amont du codon d'initiation des gènes *alkB1* et *alkB2* ont ainsi été mises en évidence. Ces séquences sont les suivantes :

```

-95  CGTGTCTTAGACAGGCGACAAAATGCGCTCTTATAGTCAACGGCAGGAGGTG  -55  ...ATG --alkB1-->
-108 ACGAGATTTGCCTGCGGCGGGCACCACCTGGTATGACGGCGTTACCCAATCG  -68  ...ATG --alkB2-->
    
```

Le logiciel repère des éléments conservés dans les promoteurs procaryotes, comme la TATA-box. Deux séquences voisines des séquences consensus TATAAT et TTGCACA des promoteurs de *E. coli*, et généralement situées à 10 et 35 bases en amont du site d'initiation de la transcription, ont en effet été trouvées par cette analyse, suggérant que ces régions font bien partie des promoteurs des gènes *alkB*.

Nous avons cloné les gènes *alkB1-rubA1-rubA-tetR* et *alkB2* dans pHLD69 de telle sorte que les gènes soient précédés par une région de 300 pb en amont des gènes *alkB*, qui d'après notre analyse, devrait contenir le promoteur (**figure 4.15**). Le clonage de gènes *alkB1-rubA1-rubA2* et *alkB1-rubA1-rubA2-tetR* dans pHLD69 a nécessité de faire un clonage intermédiaire dans le plasmide pUC18 permettant d'utiliser les sites de restriction KpnI et BamHI. Ce sont les constructions pEN10 et pEN11.



**Figure 4.15** : Plasmides dérivés de pHLD69 construits pour exprimer les gènes *alkB* de IFP2173 dans *Mycobacterium*.

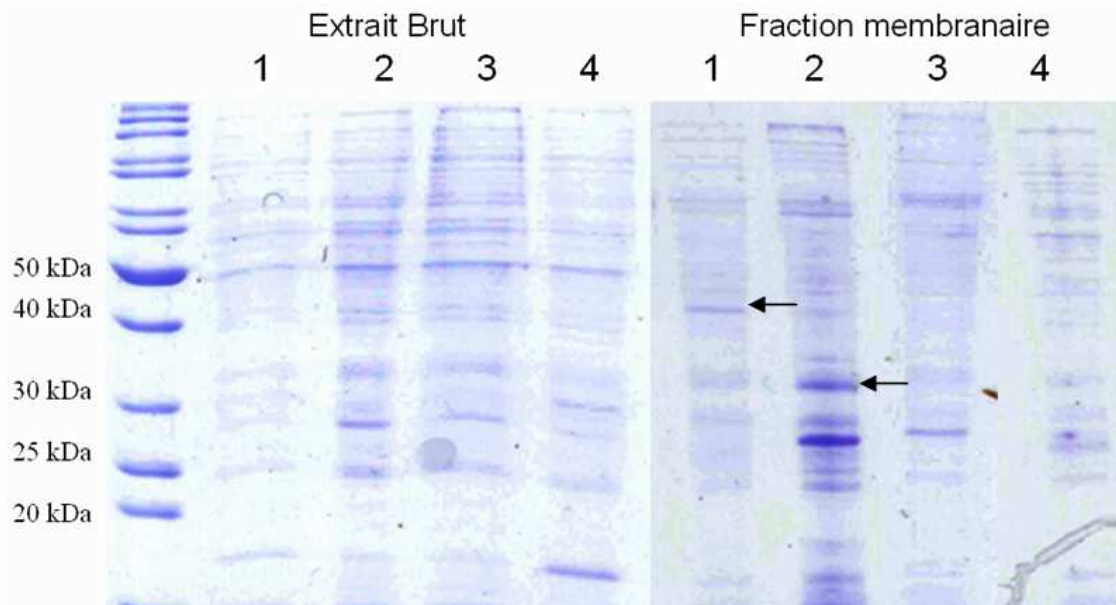
Les zones non codantes sont schématisées en violet, les gènes en vert. Les sites de restriction utilisés pour les constructions sont indiqués.

Quatre constructions ont ainsi été réalisées et transférées chez *Mycobacterium smegmatis* mc<sup>2</sup>155 par électroporation. Des clones transformants ont été obtenus après 10 jours d'incubation à 30°C sur milieu PTYG<sub>20</sub> additionné de kanamycine<sup>20</sup>. Plusieurs clones ont alors été cultivés à 30°C en milieu liquide MMSYE avec du Tween80 comme unique source de carbone. *M. smegmatis* mc<sup>2</sup>155 a été transformée par les plasmides dérivés de pHLD69 avec une efficacité de 3.10<sup>3</sup> transformants.μg<sup>-1</sup> ADN.

#### 4.1.4.2 Tests d'induction de AlkB chez *M. smegmatis* mc<sup>2</sup>155

Les souches transformantes de *M. smegmatis* mc<sup>2</sup>155 décrites précédemment ont été cultivées sur *n*-octane, utilisé comme seule source de carbone. La croissance de ces souches transformées sur *n*-octane à 30°C a duré 19 jours et les D.O.<sub>600</sub> des souches étaient relativement faibles, comprises entre 0,1 et 0,5.

L'extrait protéique total et la fraction membranaire de chaque souche ont été analysés par SDS PAGE afin de détecter la présence des protéines de la souche IFP 2173. La **figure 4.16** représente le gel à 12,5 % d'acrylamide sur lequel toutes les constructions ont été analysées. Dans les extraits bruts de protéines, aucune bande correspondant à l'AlkB1 (46-48 kDa) ou l'AlkB2 (44 kDa) n'a été mise en évidence. Les rubrédoxines RubA1 et RubA2 de masses respectives 6447 Da et 6790 Da, ont une taille trop petite pour être détectables sur ce gel. Quant au facteur de transcription TetR (28,5 kDa), aucune bande majoritaire ne correspond à la taille de cette protéine.



**Figure 4.16 :** Expression des AlkB dans pHLD69.

1. *M. smegmatis* mc<sup>2</sup>155 pEN12 ; 2. *M. smegmatis* mc<sup>2</sup>155 pEN13 ; 3. *M. smegmatis* mc<sup>2</sup>155 pEN19 ; 4. *M. smegmatis* mc<sup>2</sup>155 pEN20.

Dans la fraction membranaire, il y a une légère bande à 46 kDa sur la piste 1. Elle n'est pas retrouvée sur la seconde piste, alors qu'il s'agit dans les deux cas de souches contenant un plasmide portant *alkB1*. Dans l'extrait 2 une bande peut correspondre en taille au facteur de transcription TetR (28,5 kDa). En ce qui concerne les constructions portant le gène *alkB2*, sur les pistes 3 et 4, il n'y a pas de bande clairement détectable aux alentours de 45 kDa. Si l'on ne détecte pas d'enzyme de type AlkB dans ces constructions, cela ne veut pas dire que la protéine n'est pas synthétisée, mais qu'elle est peu visible par SDS-PAGE. En effet, lors de l'analyse protéomique, aucune des protéines AlkB n'a été identifiée dans les fractions membranaires de IFP 2173, alors les transcrits correspondants sont détectables par RT-PCR.

#### 4.1.4.3 Test de dégradation du 2-EHN

Les souches contenant les plasmides pEN12 pEN13, pEN19 et pEN20 ont été testées pour leur capacité à dégrader le 2-EHN. Pour ce faire, des tests de dégradation ont été mis en œuvre selon un protocole décrit dans la partie 1 des résultats page 76. Le suivi de la production de CO<sub>2</sub> est le premier indicateur de dégradation. Le dosage du 2-EHN résiduel, en fin de test, permet de confirmer si le substrat a été dégradé. Les résultats ont montré qu'aucune des souches d'expression testées n'a dégradé le 2-EHN de manière significative. En outre, les souches ont été incapables de dégrader le *n*-octane ou l'hexane. Ces tests de dégradation ont été effectués par Gabriela Della-Giustina au cours de son stage à l'IFP.

Ainsi, il n'a pas été possible de répondre à la question : qui de l'AlkB1 ou de l'AlkB2 intervient dans l'étape initiale de dégradation du 2-EHN ?

## **4.2 Le cytochrome P450 de type CYP153**

Grâce à l'approche protéomique, un cytochrome P450 ayant une séquence très proche de celle du cytochrome P450 de type CYP153A de *Mycobacterium* sp. HXN-1500 (van Beilen *et al.*, 2006) a été découvert.

### **4.2.1 Mise en évidence**

La séquence nucléique du cytochrome P450 de type CYP153 de la souche HXN-1500 est enregistrée dans les bases de données, sous le numéro d'accès AJ783967. Les séquences adjacentes au gène sont connues. Les séquences en amont du CYP153A n'ont pas été décrites dans les bases de données mais il y aurait un système de régulation à deux composants, avec une histidine kinase et un régulateur de transcription à deux composants LuxR. En aval du CYP153A de la souche HXN-1500, il y a une ferrédoxine réductase et une ferrédoxine. Des oligonucléotides ont été conçus en utilisant le début et la fin du gène CYP153A de *Mycobacterium* sp. HXN-1500 (Q65A64) et utilisés comme amorces de PCR avec l'ADNg de *M. australis* IFP 2173 comme matrice. Nous avons obtenu un produit PCR d'environ 1250 pb, taille comparable à celle du gène de référence (1261 pb). Le séquençage de ce produit PCR a révélé qu'il s'agissait bien d'un gène très proche de celui de la souche HXN-1500 (99 % d'identité). La comparaison des séquences protéiques déduites montre que seuls deux acides aminés sont différents, Ser9 est remplacé par Asn et Glu159 par Asp dans le CYP153-1 de IFP 2173. D'après l'étude de van Beilen *et al.* (2006) et celle de Funhoff *et al.* (2006) réalisées sur le CYP153A de *Mycobacterium* sp. HXN-1500, ces acides aminés ne font pas partie du site actif. Le spectre d'action vis-à-vis des alcanes de ce CYP153-1 serait *a priori* le même que celui de la souche HXN-1500, c'est-à-dire qu'il catalyserait l'hydroxylation des alcanes de C<sub>5</sub> à C<sub>16</sub>.

### **4.2.2 Clonage et expression du *cyp153-1* chez *E. coli***

Etant donné que les séquences adjacentes au gène *cyp153-1* n'ont pas été retrouvées, il nous était impossible d'avoir la séquence promotrice de ce gène. De ce fait, nous ne pouvions pas utiliser le vecteur de clonage pHLD69 pour faire l'expression dans *M. smegmatis*. Pour ce qui est de l'expression chez *E. coli*, nous aurions pu utiliser comme précédemment les vecteurs pET9a ou pET15b. Cependant, en cas défaut d'expression, instabilité ou corps d'inclusion, nous n'aurions pas eu d'alternative.

Récemment, un système d'expression chez *M. smegmatis* mc<sup>2</sup> a été mis au point pour l'expression de protéines de souches pathogènes comme *M. tuberculosis*. Ce système rend possible la surexpression de protéines difficiles à produire sous forme recombinante chez *E. coli*, soit parce qu'il s'agit de protéines membranaires, soit parce qu'elles ne se replient pas correctement et forment des corps d'inclusion (Goldstone *et al.*, 2008).

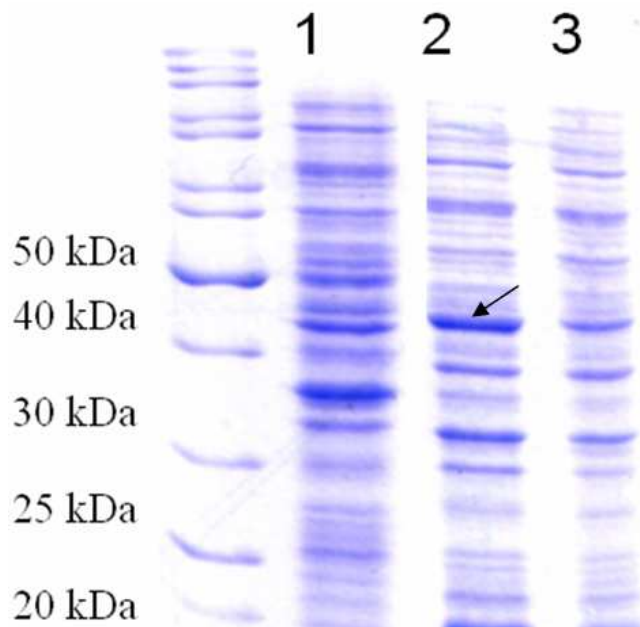
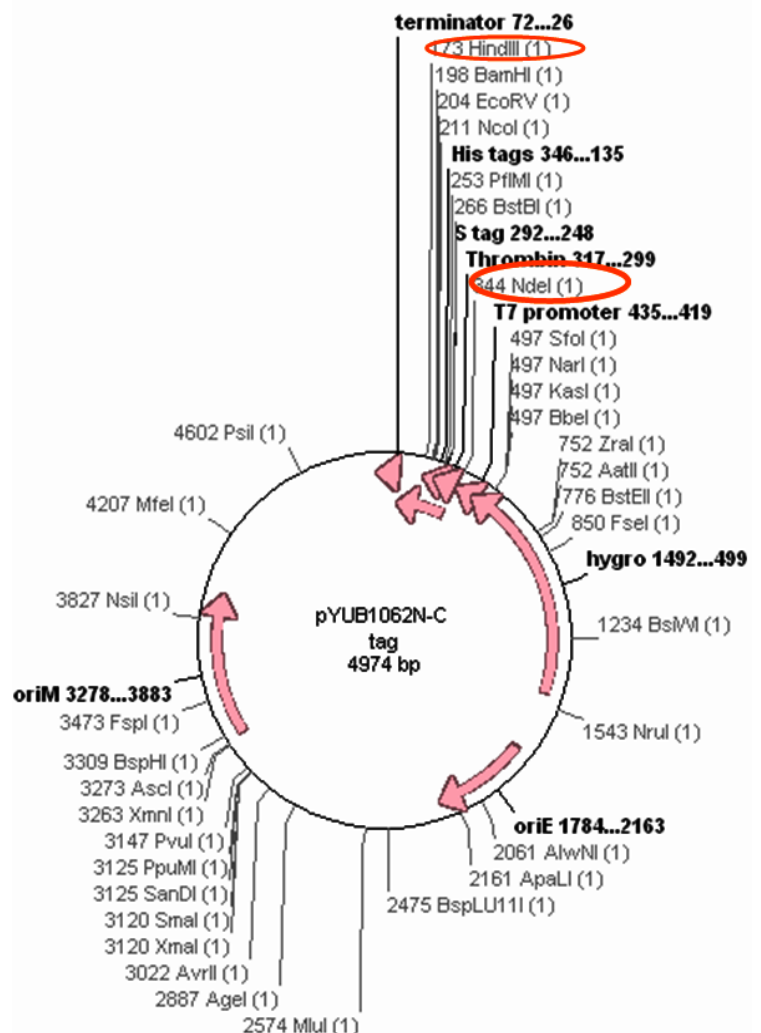
Le plasmide pUYB1062 utilisé dans notre étude est présenté par la **figure 4.17**

Ce plasmide a deux origines de répllication, l'une est reconnue par *E. coli* et l'autre par *Mycobacterium*, et il porte un gène de résistance à l'hygromycine. De plus, il possède un promoteur T7 indirectement inductible par IPTG, qui fonctionne selon le même principe que les promoteurs T7 des vecteurs de type pET dans *E. coli*. La souche de *M. smegmatis* utilisée pour l'expression, appelée mc<sup>2</sup>4517 possède vraisemblablement le gène de la T7 polymérase inséré dans son génome, bien que cela n'ait pas été publié. Nous savons d'autre part que la souche est porteuse d'un marqueur de résistance à la kanamycine, sans doute introduit lors de l'insertion du gène de la T7 polymérase.

Le gène *cyp153-1* a 4 sites de restriction BamHI dans sa séquence. De ce fait, nous avons été obligés de l'insérer dans pUYB1062N via les sites HindIII et NdeI. Le plasmide obtenu a été nommée pEN26.

Dans un premier temps, l'expression a été réalisée chez *E. coli* Rosetta<sup>TM</sup>. L'induction par l'IPTG a été réalisée pendant 14 heures à 25°C. La **figure 4.18** met en évidence qu'une protéine de 46 kDa est produite majoritairement dans l'extrait total induit (piste 2) et dans une moindre mesure dans la fraction soluble (piste 3).

**Figure 4.17 : Carte du plasmide pUYB1062.**  
 Les sites utilisés pour l'insertion du *cyp153-1* sont entourés en orange.



**Figure 4.18 : Contrôle de l'expression du CYP153-1 chez *E.coli* par SDS-PAGE.**

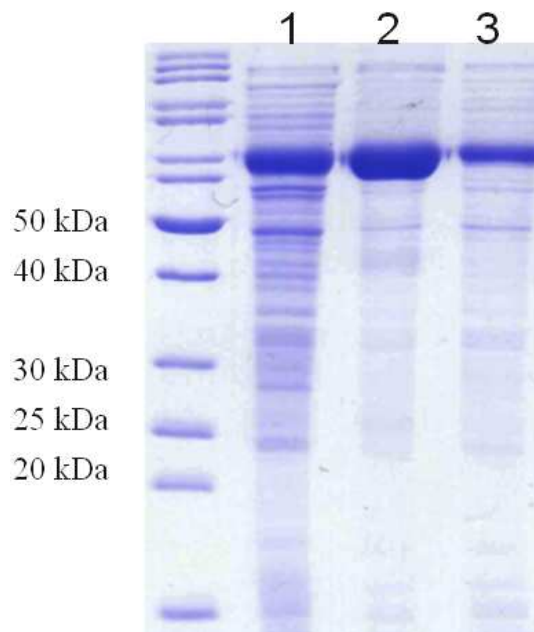
1 : Extrait protéique total de *E. coli* pEN26 non induit  
 2 : Extrait protéique total de *E. coli* pEN26 après 14 h d'induction à 25°C.  
 3 : Extrait protéique soluble de *E. coli* pEN26 après 14h d'induction 25°C.

La taille de la protéine majoritaire est en bon accord avec la masse théorique du CYP153-1 qui est de 47689 Da. Cependant, nous n'avons pas la preuve que c'est la protéine attendue puisque nous n'avons ni anticorps anti-CYP153, ni la possibilité de reconnaître un tag introduit dans la séquence par construction avec des anticorps spécifiques (anti-His-Tag, par exemple).

Puisque ce gène *cyp153-1* a été cloné dans un plasmide d'expression utilisable chez *Mycobacterium*, nous avons tenté de faire l'expression de la protéine chez *M. smegmatis* mc<sup>24517</sup>.

#### 4.2.3 Test d'expression chez *M. smegmatis* mc<sup>24517</sup>

*M. smegmatis* mc<sup>24517</sup> a été transformé avec le plasmide pEN26 par électroporation. Les clones résistant à l'hygromycine ont été sélectionnés et mis à pousser en milieu liquide sur 7H9/ADC/Tween80 avec de l'hygromycine et de la kanamycine. Je n'ai pas réussi à faire pousser cette souche en présence des deux antibiotiques. En revanche, il a été possible de faire pousser *M. smegmatis* mc<sup>2155</sup> transformée par pEN26 en présence d'hygromycine uniquement. Cependant, aucune expression du CYP153-1 n'a pu être observée (**Figure 4.19**). Il y a eu cependant une très forte surexpression d'une protéine de 65 kDa de *M. smegmatis* mc<sup>24517</sup>. Cette protéine a été identifiée par Goldstone *et al.* (2008) comme étant une protéine chaperone nommée GroL.



**Figure 4.19** : Analyse par SDS-PAGE d'extraits de *M. smegmatis* mc<sup>24517</sup> contenant pEN26  
 1 : Extrait protéique total de *M. smegmatis* mc<sup>24517</sup> sans plasmide  
 2 : Extrait protéique total de *M. smegmatis* mc<sup>24517</sup> contenant pEN26 non induit par l'IPTG  
 3 : Extrait protéique total de *M. smegmatis* mc<sup>24517</sup> contenant pEN06 induit par l'IPTG 24 h à 25°C.



Aucune protéine de 47 kDa n'est induite en présence d'IPTG. Ce système n'est donc pas fonctionnel en l'état. Il est dans un premier temps nécessaire de revoir les conditions de transformation de *M. smegmatis* mc<sup>2</sup>4517 par pEN26.

Nos résultats partiels ne permettent pas de savoir pour le moment quelle alcane hydroxylase est impliquée dans l'étape initiale de dégradation du 2-EHN.

## CONCLUSION ET PERSPECTIVES

Cette thèse avait pour objet d'étudier la biodégradation du 2-éthylhexyl nitrate (2-EHN), un composé, employé comme additif procétane des gazoles (Suppes *et al.*, 2001). Principalement produit en France, il est donc très largement utilisé par l'industrie pétrolière. Cependant, l'impact environnemental qu'il est susceptible d'exercer en cas de déversement accidentel est encore méconnu. Dans ce contexte, l'étude que nous avons entreprise présente un double intérêt :

- au point de vue la réglementation, elle répond à la question de la biodégradabilité de ce contaminant potentiel des eaux et des sols, et met en évidence les paramètres clés qui président à sa biodégradation. Elle "renseigne" producteurs et utilisateurs sur les risques environnementaux encourus lors de son utilisation ;
- d'un point de vue scientifique, elle apporte une contribution nouvelle aux connaissances des mécanismes de biodégradation du 2-EHN et des alcanes branchés. A ce jour, la dégradation des alcanes branchés est restreinte aux alcanes méthylés comme le pristane, le phytane et le citronellole. A notre connaissance, il n'existe aucune étude sur la dégradation des hydrocarbures de courtes chaînes, alpha-substitués par un groupement éthyle.

Le critère de biodégradabilité "immédiate" tel que défini par l'OCDE pour les composés chimiques stipule que le composé testé doit être biodégradé à hauteur de 70 % (selon la mesure du COD, équivalent à 60 % selon la mesure de la DThO) dans les 28 jours du test. De surcroît, cette valeur cible doit être atteinte dans un intervalle de temps maximal de dix jours. Tout composé satisfaisant à ces conditions peut être alors considéré comme "facilement biodégradable". Au commencement de notre étude, la biodégradabilité du 2-EHN n'était pas renseignée. L'US EPA (United States, Environmental Protection Agency) a publié les résultats d'un plan d'expérimentation dans lesquels figuraient des éléments préliminaires. Le rapport de l'US EPA indiquait notamment que le 2-EHN n'est pas "facilement" biodégradable (*ready biodegradability*) selon le test normalisé du "CO<sub>2</sub> dans l'espace de tête" (*CO<sub>2</sub> headspace test*). Les auteurs avaient conclu que la faible solubilité du 2-EHN, conjuguée à sa volatilité, rendait le produit très peu biodisponible, ce qui contribuait à l'absence de biodégradation. (The American Chemistry Council Petroleum additives panel, 2006). Par

rapport à ce constat, notre travail a apporté la preuve que le 2EHN est effectivement biodégradable, et a apporté des données inédites sur la biodégradation du 2-EHN.

## 1 Biodégradabilité du 2-EHN

Le recours à des cultures biphasiques dans lesquelles une phase non aqueuse servait de réservoir de substrat (HMN), a été une des clés de la réussite de ce projet, car cela permettait de tenir compte à la fois du caractère volatil, de l'hydrophobicité et du caractère toxique du 2-EHN. Aucune des souches capables d'utiliser le 2-EHN comme seule source de carbone ne le minéralise totalement. Parmi les souches de *Mycobacterium austroafricanum* capables de dégrader le 2-EHN, la souche IFP 2173 est la plus performante, car c'est la seule capable de dégrader le 2-EHN en l'absence de HMN. *M. austroafricanum* IFP 2173 a donc été utilisée pour entreprendre les études cinétiques et rechercher la voie de dégradation du 2-EHN.

## 2 Voie de dégradation du 2-EHN par *M. austroafricanum* IFP 2173

La minéralisation du 2-EHN n'est que partielle puisqu'on assiste à l'accumulation d'un intermédiaire de dégradation. L'intermédiaire présumé, la 4-éthylidihydrofuran-2(3H)-one (4-EDF), plus communément appelé  $\beta$ -éthyl- $\gamma$ -butyrolactone, a été identifié sans ambiguïté par des analyses chromatographiques concordantes et par synthèse chimique (collaboration avec Lucien Kerhoas et Martine Lettere de l'INRA de Versailles).

Grâce à la structure de la 4-EDF, nous avons proposé une voie de dégradation plausible pour le 2-EHN (**figure 5.1**). *M. austroafricanum* IFP 2173 a été isolé pour sa capacité à dégrader l'iso-octane (Solano-Serena *et al.*, 2004) mais ne peut pas métaboliser la 4-EDF. Des essais de dégradation simultanée du 2-EHN et de l'isooctane n'ont pas permis de minéraliser totalement le 2-EHN. Il semble donc que la souche ne dispose pas d'enzyme capable de catalyser l'élimination du groupement éthyle et/ou hydroxyméthyle *antéiso*-terminaux de l'acide 3-(hydroxyméthyl) pentanoïque. La réaction concurrente de lactonisation de l'acide 3-(hydroxyméthyl) pentanoïque est thermodynamiquement très favorable et peut avoir empêché la dégradation complète du 2-EHN.