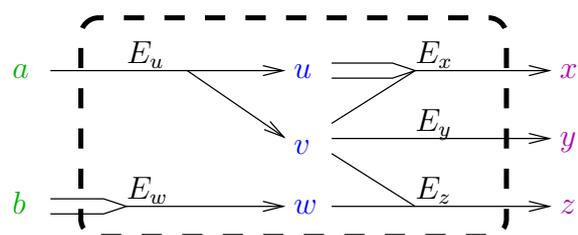


Optimiser la bioproduction par analyse de réseaux métaboliques

La startup *VertBio*, soucieuse de l'environnement, souhaite produire deux molécules pharmaceutiques, que nous appellerons x et z , sans pollution chimique. Elle a isolé une souche bactérienne « presque parfaite » qui possède une bonne croissance *in vitro*, des substrats entrants a et b peu coûteux, et fournit trois substrats en sortie : x , y et z . Les voies métaboliques concernées de cette souche sauvage *Fere perfectum* reposent sur 5 enzymes et 3 métabolites internes supplémentaires. En effet, le réseau enzymatique suivants de *F.perfectum* sauvage est bien connu :



Ici, les substrats entrants sont en *vert*, les produits excrétés par *F.perfectum* en *violet* et les métabolites internes intermédiaires en *bleu*. Les entrées multiples indiquent les stœchiométries ; par exemple E_x transforme deux u et un v en un x .

La première question pour l'entreprise *VertBio* est bien sûr la potentielle rentabilité d'utiliser des cultures de *F.perfectum*, et donc en particulier de calculer le rapport entre les quantités de a et b à fournir et les quantités de x et z obtenues.

Question 1 : D'après le dessin du réseau métabolique, peut-on trouver ce rapport ?

Par exemple, est-il vrai que $2a$ et $2b$ suffisent à produire un x et un z ?

On produirait donc $2u$, $2v$ et $1w$ qui suffisent à produire $1x$ et $1z$.

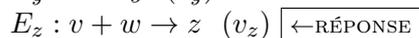
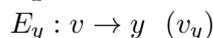
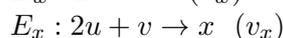
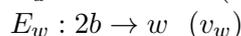
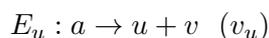
Malheureusement la réaction de E_y risquerait de « voler » des v au passage...

Alors comment calculer ce qui se passe en réalité ?

On voit que la réponse à cette question n'a rien d'évident *a priori* à cause de la concurrence entre les réactions enzymatiques E_x , E_y et E_z ... et pour être complets, à cause aussi des limitations intrinsèques des diverses vitesses de réactions enzymatiques au sein de la bactérie.

On va donc écrire plus complètement les réactions qui forment le réseau métabolique :

Question 2 : Écrire ces réactions avec leur stœchiométrie et en nommant leurs vitesses (pour le moment de valeurs inconnues) v_u , v_w , v_x , v_y et v_z .



L'objectif devient alors d'évaluer un rapport d'efficacité entre les vitesses d'entrée v_u , v_w , et de sortie v_x , v_y , v_z . Les connaissances biologiques fournissent généralement des *plages de vitesses* possibles, par exemple en nombre de molécules transformées par seconde, ou en évolutions de concentrations, ou tout autre unité de mesure : peu importe car ce qui nous intéresse, c'est le rapport entre ces diverses vitesses.

Pour simplifier et ne considérer que des vitesses positives, on va supposer que les réactions sont non réversibles (*i.e.* que leur bilan est toujours orienté comme sur le dessin du réseau). Les expériences *in vitro* sur les réactions enzymatiques fournissent alors les bornes suivantes :

$$v_u \leq 30, \quad v_w \leq 15, \quad v_x \leq 20, \quad v_y \leq 5 \quad \text{et} \quad v_z \leq 8.$$

Malheureusement, ce ne sont que des vitesses maximales envisageables et cela ne donne pas les vitesses de production optimales de x et z .

Toutefois, comme le réseau métabolique de *F. perfectum* est bien connu, on sait qu'à l'état sauvage, aucune autre réaction enzymatique ne consomme ou produit des métabolites u , v ou w .

Question 3 : Est-il possible/raisonnable, selon vous, que les réactions sortantes E_x , E_y et E_z consomment plus, ou moins, de métabolites internes que n'en produisent les réactions entrantes E_u et E_w ? Il est évidemment impossible de consommer plus. Mais consommer moins produirait des accumulations handicapantes pour la bactérie que l'évolution corrigerait. On peut donc considérer, sans que l'hypothèse soit déraisonnable, qu'elles ne consomment pas moins non plus.

On arrive ainsi assez vite à l'idée que les métabolites u , v et w restent en nombre approximativement constant, donc ont des vitesses d'évolution nulles (ni positives ni négatives). Ces vitesses d'évolution (e_u , e_v et e_w) correspondent donc aux productions des réactions E_u et E_w auxquelles on enlève les consommations des réactions E_x , E_y et E_z .

Question 4 : Comment s'écrivent les vitesses d'évolution de ces 3 métabolites internes par rapport aux vitesses des réactions enzymatiques?

$$e_u = v_u - 2v_x$$

$$e_v = v_u - v_x - v_y - v_z$$

$$e_w = v_w - v_z \quad \text{←RÉPONSE}$$

Sachant que ces trois vitesses d'évolution sont nulles ($= 0$), on dispose alors de 3 équations qui relient entre elles les vitesses des 5 réactions enzymatiques. C'est finalement ce que l'on cherchait :

Question 5 : Exprimer les vitesses des réactions enzymatiques de sortie v_x , v_y et v_z en fonction des vitesses des réactions enzymatiques d'entrée v_u et v_w .

$$v_x = \frac{v_u}{2}$$

$$v_z = v_w$$

$$0 = v_u - \frac{v_u}{2} - v_y - v_w \text{ donc } v_y = \frac{v_u}{2} - v_w \quad \text{←RÉPONSE}$$

Question 6 : En supposant que l'on cultive *F. perfectum* de telle sorte que les apports en a et b soient optimaux pour ce réseau métabolique. Quel rapport de vitesse maximiserait la production de x et z ? *A priori* on aimerait $v_x = 20$ et $v_z = 8$ mais

1. $v_x = \frac{v_u}{2}$ limite v_x à 15 car $v_u \leq 30$

2. pire : $v_y = \frac{v_u}{2} - v_w \leq 5$ donc, même si $v_z = 8$, $v_x = \frac{v_u}{2}$ ne peut pas dépasser 13...

Avec $v_x = 13$ et $v_z = 8$ on constate que $v_y = 5$, et dès lors $v_u = 26$ et $v_w = v_z = 8$. Ceci ne contredit aucune limite de vitesse, c'est donc finalement l'optimum.

L'entreprise considère que les rapports entre les vitesses d'entrée et de sortie du réseau métabolique sont rentables; cependant, elle constate que la souche *Fere perfectum* sauvage produirait un excédent de x : en effet la startup *VertBio* a besoin d'autant de x que de z pour son produit pharmaceutique. De plus, le sous-produit y arrive en surplus et il s'avère qu'aucune utilisation rentable de y n'existe actuellement.

Une idée naturelle est de faire un KO du gène produisant l'enzyme E_y . Cela neutraliserait la catalyse de y .

Pour prédire si cette manipulation génétique serait rentable, on doit appliquer la même méthode, sans la catalyse de E_y . Il n'est pas nécessaire de tout refaire : on peut simplement réviser les vitesses maximales en imposant $v_y = 0$.

Question 7 : Quelles sont alors les vitesses qui optimisent la production souhaitée?

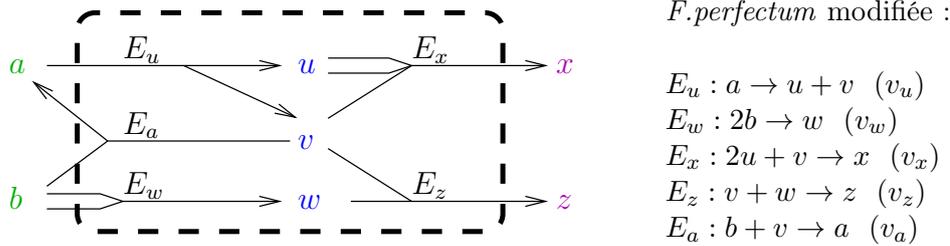
Et produit-on du x ou du z en excès? $v_y = 0$ signifie $\frac{v_u}{2} = v_w$, soit $v_x = v_z$.

Ainsi, *a priori*, on aimerait $v_x = 8$ et $v_z = 8$ et en fait ça convient :

$v_x = 8, v_z = 8, v_u = 16$ et $v_w = 8$ et il n'est donc produit aucun excès.

N.B. : avant d'effectuer cette manipulation génétique sur *F.perfectum* il faudrait effectuer des simulations 3D pour vérifier partiellement nos hypothèses. Comme durant le cours *Biologie, Bio-informatique et I.A.* suivi en GB3, on pourrait utiliser HSIM. On ne le fera pas ici.

Après avoir effectué expérimentalement le KO du gène de E_y , et malgré les simulations effectuées préalablement, *VertBio* constate qu'un nouveau gène, qui était silencieux dans la souche sauvage, s'exprime dans la souche génétiquement modifiée. Ce gène produit l'enzyme E_a et cela modifie le réseau métabolique comme suit :



De plus l'enzyme E_a est très active de sorte que sa vitesse v_a est toujours supérieure à 8. Ainsi les connaissances sur les vitesses sont complétées comme suit :

$$v_u \leq 30, \quad v_w \leq 15, \quad v_x \leq 20, \quad v_z \leq 8 \quad \text{et (nouveau)} \quad 8 \leq v_a \leq 20.$$

Question 8 : En exploitant le fait que les vitesses d'évolution des métabolites internes sont nulles, comment s'expriment v_x, v_z et v_a en fonction de v_u et v_w ...

$$v_u - 2v_x = 0 \text{ donc } v_x = \frac{v_u}{2} \text{ comme avant,}$$

$$v_w - v_z = 0 \text{ donc } v_z = v_w \text{ comme avant,}$$

$$v_u - v_x - v_z - v_a = 0 \text{ donc } v_u - \frac{v_u}{2} - v_w - v_a = 0 \text{ donc } v_a = \frac{v_u}{2} - v_w \quad \text{←RÉPONSE}$$

De même, on va chercher à optimiser les vitesses v_x et v_z pour maximiser l'obtention des produits d'intérêt.

Question 9 : Trouver les vitesses optimales pour la startup *VertBio*. Les deux premières égalités limitent v_z et v_w à 8 et v_x à $\frac{30}{2} = 15$. Cependant on sait aussi que $v_a = \frac{v_u}{2} - v_w \geq 8$, soit $\frac{v_u}{2} - 8 \geq v_w$. et comme $\frac{v_u}{2} \leq 15$ il vient : $v_w \leq 7$. Donc les valeurs optimum deviennent $v_x = 15$ et $v_z = v_w = 7$.

Par conséquent, cette consommation inattendue de v conduit à de moins bons résultats que la bactérie sauvage puisqu'elle produit moins de z (7 au lieu de 8, mais elle économise la quantité de a que *VertBio* doit fournir). Une idée « naturelle » serait alors de dupliquer le gène de E_z afin de tenter de multiplier par deux la vitesse de E_z :

Question 10 : Ce serait une mauvaise idée. Pourquoi? parce que la vitesse minimale de E_a limite toujours autant $v_z = v_w$ donc il ne sert à rien d'augmenter sa vitesse maximale!

L'étape d'après est bien sûr de tenter un KO du gène de l'enzyme E_a ...

Le petit scénario que nous avons développé pendant ce cours est représentatif du travail quotidien des entreprises de biotechnologie : (1) partir de souches considérées comme prometteuses, (2) modéliser les réseaux biologiques d'intérêt de ces souches, (3) en déduire des optimisations souhaitables prédites par la modélisation, (4) valider les prédictions à la paillasse, (5) on n'est pas à l'abri de surprises, ce qui conduit à affiner la modélisation, (6) la modélisation fournit des prédictions qui, en pratique, évitent de nombreuses expériences coûteuses et guide vers les expériences les plus prometteuses, cependant, les validations expérimentales « humides » sont indispensables !