

TUTORAT CHIMIE

Année 2009-2010

Associations protéines-ligand

Fiche préparée par Alexandre LÉBOUCHER (PCEM2 Montpellier).

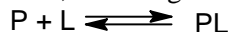
Cette fiche (qui n'en est pas vraiment une...) a pour vocation de réunir la plupart des formules du cours de V.Sieso, et mieux comprendre des points complexes tels que l'allostérie. Les démonstrations ont été volontairement laissées et même plus développées, car elles sont réellement la clé qui vous permettra de comprendre les données que vous manipulez, et peuvent même vous sauver si vous n'arrivez pas à bien mémoriser une des formules particulièrement complexes de ce cours. Cette fiche n'est en fait qu'une version revue et plus élaborée de ce que j'avais moi-même fait en P1, en complément de mon cours.

N'oubliez pas que les questions de calculs ou d'analyses de courbes sont absolument à assurer au concours, car elles sont une source sûre de points contrairement aux questions « par cœur », à la condition que vous fassiez un effort de compréhension (Certes difficile...)

RAPPEL : Les protéines, de part leur structure tertiaire particulière (sites actifs des enzymes, poches hydrophobes...), sont des molécules particulièrement aptes à former des liaisons avec d'autres molécules plus petites, appelées « Ligand » au sens large. Ces liaisons sont le plus souvent de faible énergie (hydrogène, électrostatiques, hydrophobes), très rarement covalentes. Cette association protéine-ligand est le plus souvent spécifique d'un ligand donné. En réalité, cette notion de spécificité est à nuancer, puisque une protéine donnée peut avoir une affinité supérieure pour un ligand donné plutôt qu'un autre (Notion de compétition).

I- Relation d'équilibre : cas simple Michaëlien

Le caractère non covalent de l'association protéine/ligand rend la réaction réversible. Ainsi, on a un équilibre entre la forme « PL » (Protéine liée au ligand) et la forme « P + L » (Protéine non liée au ligand). On peut noter cet équilibre un peu comme une réaction chimique, même si ce n'en est pas une ! (Pas de création de liaisons de covalence, d'échanges électroniques...)



Il faut remarquer qu'à l'instar d'une réaction chimique, il existe une vitesse de formation de PL, et une vitesse de dissociation de PL. Ces deux vitesses s'écrivent différemment.

La vitesse d'une réaction chimique est proportionnelle à la concentration des réactifs et peut s'écrire, pour une réaction de type « n(réactif) → produit » :

$$v = k \times [\text{réactif}]^n$$

Avec v en $M \cdot s^{-1}$, et k constante de vitesse ayant une unité variable selon l'ordre de la réaction

Si $v = k$, alors la réaction est dite d'ordre 0, k est en $M \cdot s^{-1}$.

Si $v = k \times [\text{Réactif}]$, alors la réaction est d'ordre 1, k est en s^{-1}

Si $v = k \times [\text{Réactif 1}] \times [\text{Réactif 2}]$, alors la réaction est d'ordre 2, k est en $M^{-1} \cdot s^{-1}$

Ainsi, par analogie avec la vitesse d'une réaction chimique, on peut écrire la vitesse d'association ou de dissociation du complexe protéine-ligand PL. Si on nomme, k_1 la constante de vitesse d'association du complexe, et k_{-1} la constante de vitesse de dissociation, alors on peut écrire :

- Dans le sens $P + L \rightarrow PL$, $v(\text{association}) = k_1 \times [P] \times [L]$
- Dans le sens $PL \rightarrow P + L$ $v(\text{dissociation}) = k_{-1} \times [PL]$

REMARQUE : La constante de vitesse k d'une réaction chimique peut être calculée par la loi d'Arrhénius, non donnée en cours mais donnée dans le formulaire à la fin, qui est à savoir.

$$k = A \cdot \exp\left(\frac{-\Delta G_a}{RT}\right)$$

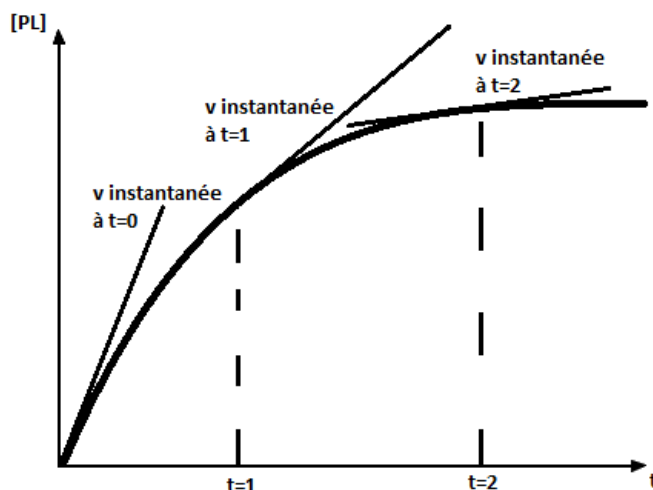
A peut être considéré comme constant si la température varie peu au cours de la réaction.

La loi d'Arrhénius explique ainsi que la cinétique d'une réaction chimique tient compte de la température et de l'énergie d'activation ΔG_a (Que les enzymes, catalyseurs biologiques, abaissent).

Ainsi, on peut déterminer la vitesse instantanée de la formation du complexe PL. Elle est égale à PL qui se forme moins PL qui se dissocie (car il y a équilibre), soit :

$$v_{\text{instantanée}} = v(\text{association}) - v(\text{dissociation})$$

La vitesse instantanée prend donc en compte la concentration de P et L en un temps précis. Si on observe l'évolution de la concentration en complexe PL en fonction du temps $[PL]=f(t)$



En fait, la vitesse instantanée de formation du complexe PL est la variation de la concentration en complexe en fonction du temps. En maths, il s'agit d'une dérivée de type

$$v = \frac{d[PL]}{dt}$$

BONUS COMPREHENSION : La dérivée en un point d'abscisse x étant la pente de la tangente en ce point de la courbe, on peut ainsi déduire la valeur de la vitesse instantanée. Remarquez que cette vitesse instantanée est maximale pour une concentration en $[PL]$ proche de 0, et qu'elle est nulle lorsque $[PL]$ ne varie plus (équilibre). Attention, ne pas confondre la vitesse instantanée de formation du complexe avec les vitesses d'association et de dissociation ! (La vitesse instantanée tenant compte de $v(\text{association})$ et de $v(\text{dissociation})$)

Si on considère maintenant l'équilibre. On a $v_{\text{instantanée}} = 0$ (C'est tout à fait visible sur le schéma ci-dessus, puisque la tangente à la courbe à une pente nulle ! Et donc $d[PL]/dt=0$), d'où :

$$v(\text{association}) - v(\text{dissociation}) = 0$$

$$v(\text{association}) = v(\text{dissociation})$$

$$k_1 \times [P] \times [L] = k_{-1} \times [PL]$$

Il vient alors :

$$\frac{k_{-1}}{k_1} = \frac{[P] \times [L]}{[PL]} = K_d$$

Où **Kd** (Unité = M, c'est une concentration !!) est la constante de dissociation du système. Plus la constante de vitesse de dissociation k_{-1} est élevée, plus **Kd** est fort et plus [PL] se dissocie, on en déduit donc qu'une forte affinité traduit une faible valeur de **Kd**.

On aurait très bien pu aussi définir **Ka**, en tant que constante d'association, qui est l'inverse de Kd.

$$K_a = \frac{k_1}{k_{-1}} = \frac{[PL]}{[P] \times [L]} = \frac{1}{K_d}$$

Où là, plus **Ka** est fort, plus l'affinité est grande. ATTENTION, l'unité de **Ka** n'est pas la même que **Kd**, c'est son inverse ! (L.mol⁻¹)

REMARQUE: On peut faire l'analogie avec la constante de réaction **K** en chimie, que vous avez étudiée au lycée. Pour une réaction d'addition $A+B \rightarrow C$ par exemple, on a $K = \frac{[C]}{[A].[B]}$
C'est le même principe. Ici, **Kd** fait penser au cas inverse, $C \rightarrow A+B$.

II- Signification de Kd

Prenons maintenant le cas où on a une concentration en complexe PL égale à la concentration en protéine P non liée. [P]=[PL] : On peut alors dire que la protéine est saturée à 50% (Attention ce n'est pas forcément l'équilibre !)

On peut donc écrire :

$$\frac{[P] \times [L]}{[PL]} = K_d$$

$$\leftrightarrow [L] = K_d = L_{0,5}$$

La constante de dissociation se définit donc comme la concentration en ligand libre pour laquelle la protéine est saturée à 50% par son ligand.

III- Notion de fraction de saturation

On définit la fraction de saturation **Y**, comprise entre 0 et 1 (100%), qui correspond pour la protéine en solution au pourcentage de forme liée avec son ligand PL sur le nombre total de protéines P_0 .

On peut donc écrire : $Y = \frac{[PL]}{[P_0]}$ Avec : $[P_0] = [PL] + [P]$

Cette formule ne nous permet pas de calculer la saturation en fonction de la concentration en ligand, ce qui est beaucoup plus intéressant.

Or, on connaît une méthode pour exprimer [P] en fonction de [L] : $[P] = K_d \times \frac{[PL]}{[L]}$

On peut donc écrire :

$$Y = \frac{[PL]}{[P] + [PL]} = \frac{[PL]}{K_d \times \frac{[PL]}{[L]} + [PL]}$$

$$Y = \frac{[PL]}{\left(\frac{K_d \times [PL] + [L] \times [PL]}{L} \right)} = \frac{[PL] \times [L]}{K_d \times [PL] + [L] \times [PL]}$$

Par factorisation :

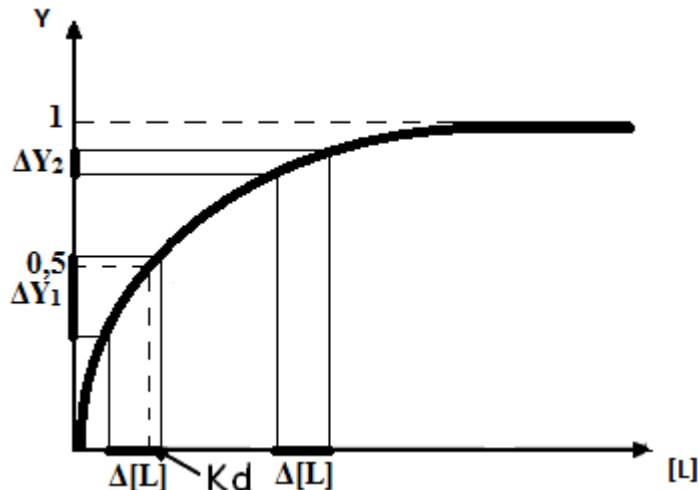
$$Y = \frac{[PL] \times [L]}{[PL] \times (K_d + [L])}$$

Au final, on a donc l'expression de la saturation d'une protéine tenant compte que de la concentration du ligand (Attention : à l'équilibre !)

$$Y = \frac{[L]}{K_d + [L]}$$

Notez qu'elle ne tient compte que de la concentration en ligand, et ce peu importe la concentration en protéine (K_d étant fixe pour un même couple PL)

Si on trace la fonction représentant la saturation en fonction de la concentration en ligand $Y=f([L])$



$Y=f([L])$ a l'allure d'une hyperbole croissante, pour laquelle Y tend vers 1 sans jamais ne l'atteindre. Remarquez qu'une même variation de concentration en ligand ΔL n'a pas les mêmes effets d'augmentation sur la saturation. Sur le schéma, on voit bien que $\Delta Y_1 > \Delta Y_2$ pour la même variation de concentration en ligand libre !

Cependant, ce type de courbe n'est pas toujours pratique, et on peut préférer une représentation en droite, plus simple. (Rappel : équation d'une droite de type $y(x)=ax+b$ avec a = pente ou coefficient directeur, et b =constante qui définit le point en ordonnée que va croiser la droite pour $x=0$)

On a :

$$Y = \frac{[L]}{K_d + [L]}$$

En inversant :

$$\left(\frac{1}{Y}\right) = \frac{K_d + [L]}{L} = \frac{K_d}{L} + \frac{L}{L} = 1$$

Au final, on obtient :

$$\left(\frac{1}{Y}\right) = K_d \times \frac{1}{[L]} + 1$$

$$y = a \quad x + b$$

Equation de la forme :

REMARQUE : Bien qu'à force de l'utiliser vous connaîtrez cette formule par cœur, vous voyez qu'en fait, il n'est pas forcément utile de l'apprendre car elle se retrouve très vite ! Ce qui vous économise de la mémoire (Mot clé = On inverse !)

Assurez-vous d'avoir bien compris tout ce qui a précédé avant de poursuivre !

IV- Détermination de K_d

➤ A partir d'une droite $(1/Y)=f(1/[L])$

- K_d étant le coefficient directeur, on peut le retrouver à l'aide de la formule mathématique permettant de la calculer. Long et difficile mais peut être utile quand on connaît les $1/Y$ ($=y$) pour 2 valeurs de $1/[L]$ ($=x$).

$$a = \frac{(y_b - y_a)}{(x_b - x_a)}$$

Si on prend des valeurs arbitraires (cf schéma ci-après),

- Pour $(1/L)_a = 3$ on a une saturation de 0,25 ou $(1/Y)_a = 4$
- Pour $(1/L)_b = 4$ on a une saturation de 0,20 ou $(1/Y)_b = 5$

Alors on peut déterminer le K_d :

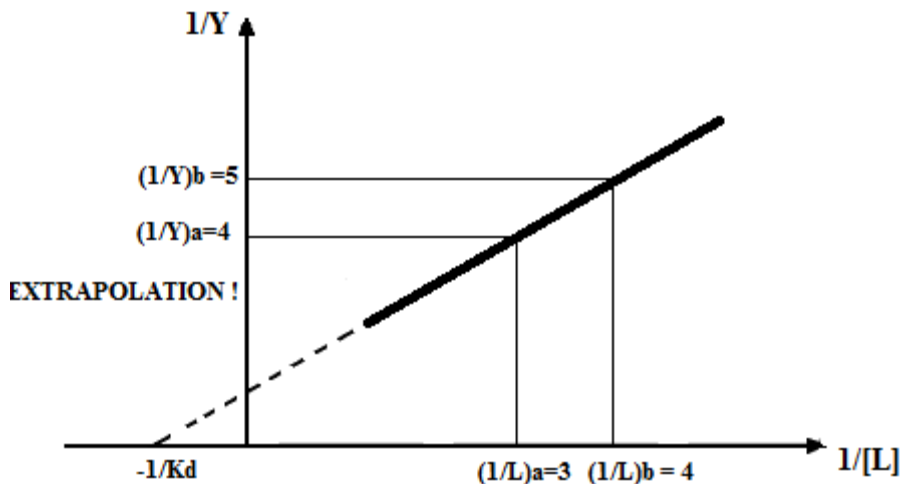
$$K_d = \frac{5-4}{4-3} = 1 \text{ mol/L} \quad \rightarrow \quad \text{La protéine sera saturée à 50\% pour une concentration en ligand libre de 1 mol/L}$$

- Une autre manière, plus simple, de calculer K_d est l'extrapolation mathématique. Elle revient à calculer le point de la droite pour lequel elle coupe l'axe des abscisses, autrement dit, le x pour lequel $y=ax+b=0$.

On a donc $ax = -b$ et donc $x = -b/a$. En l'occurrence, $a=K_d$ et $b=1$.

Ainsi, $x_{y=0} = -1/K_d$

Attention, ce n'est que de l'extrapolation mathématique et ça n'a pas de sens en biologie ! (Correspondrait à une valeur de $x=1/[L]$ négative...). Et même en mathématiques si on tentait de pousser le raisonnement on serait bloqué : $y=1/Y$ ne peut être égal à 0 !



➤ Méthode de la dialyse à l'équilibre

Soient deux compartiments de volumes semblables (1L+1L) séparés par une membrane hémiperméable ne laissant passer que les ligands et non les protéines.

A $t=0$ on place des protéines (Concentration $[P_0]$) et des ligands (Concentration $[L_0]$) dans un des deux compartiments. Les ligands diffusent et s'homogénéisent dans les deux compartiments, contrairement aux protéines.

A l'équilibre, on ne retrouve PL que dans le compartiment qui comportait les protéines, et on dose les concentrations dans ce compartiment. On peut alors écrire, pour ce compartiment:

- $[P_0] = [P]_{eq} + [PL]_{eq}$ (C'est logique)
- $[L_0] = 2 \times [L]_{eq} + [PL]_{eq}$ (On multiplie par 2 la concentration en ligand, car elle a été diluée dans 2 L par diffusion libre à travers la membrane ! Si on avait eu 1L et 2L, on aurait multiplié par 3)

Ces deux formules permettent alors de retrouver toutes les données nécessaires au calcul de K_d .

➤ Méthode de Scatchard

Peut être utile dans les cas où l'on a une protéine avec plusieurs sites !

Soit une collection de N protéines contenant chacune n sites pouvant se fixer chacun à un ligand. On appelle r le nombre de sites occupés.

Que peut-on en déduire pour une seule protéine?

- $n = [S_0]$ (Nombre de sites sur la protéine, et donc de sites de fixation possibles)
- $r = [PL]$ (Car r représente le nombre de site occupés, le nombre de complexes PL !)
- $(n-r) = [P]$ (Les protéines non liées, à l'équilibre)

Ainsi, pour N protéines, on a $[P]=N(n-r)$, $[PL]=N.r$, $[S_0]=n.N$ (S_0 = Nombre de site au total)

On peut transformer :

$$K_d = \frac{[P] \times [L]}{[PL]} \quad \xrightarrow{\text{transformation}} \quad K_d = \frac{N(n-r) \times [L]}{N \times r}$$

D'ores et déjà, on peut simplifier :
$$K_d = \frac{N(n-r) \times [L]}{N \times r} = \frac{(n-r) \times [L]}{r}$$

On inverse :

$$\frac{1}{K_d} = \frac{r}{(n-r) \times [L]}$$

$$\frac{(n-r)}{K_d} = \frac{r}{[L]}$$

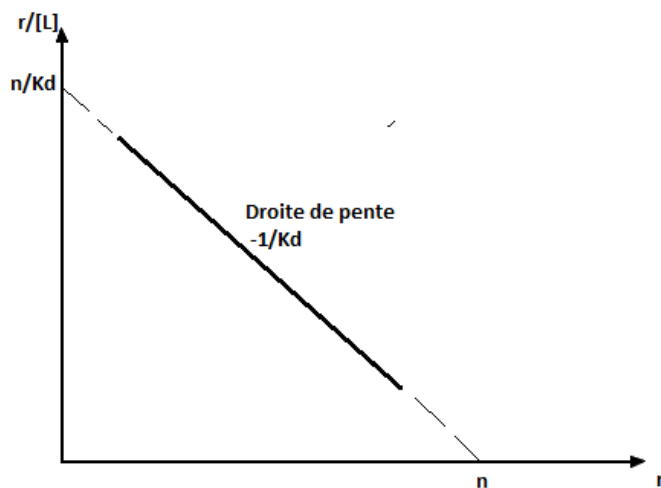
$$\frac{n}{K_d} - \frac{r}{K_d} = \frac{r}{[L]}$$

Enfin, on peut aussi écrire :

$$\frac{r}{[L]} = \left(\frac{-1}{K_d}\right) \times r + \frac{n}{K_d}$$

$$y = a \quad x + b$$

Equation d'une droite décroissante avec r variable. Remarque : r représente le nombre de site occupé sur UNE SEULE protéine. Donc $r = [PL]/N$



Comment comprendre que la droite coupe l'axe des abscisses en un point qui correspond au nombre de site n de la protéine?

- Si $[L]$ tend vers l'infini, alors tous les sites sont occupés et $r=n$ (Saturation 100%)
- Or, si $[L]$ tend vers l'infini, $r/[L]$ tend vers 0, ce qui correspond à l'axe des abscisses.

V- Phénomènes allostériques

Il s'agit du cas où une protéine possède plusieurs sites qui dépendent les uns des autres, pour un même ligand. Autrement dit, le K_d d'un site est déterminé par l'occupation des autres sites. Attention, si une protéine possède plusieurs sites indépendants (Calbindine, Ferritine...), alors on peut considérer que chaque site se comporte comme une protéine michaëlienne classique.

Dans le cas où la fixation d'un premier ligand facilite la fixation d'un autre, on peut alors faire l'approximation suivante



(En réalité, c'est plus complexe, l'équation précédente n'envisage que les cas où la protéine non liée est en équilibre avec la protéine ayant tous ces sites occupés par fixation simultanée, il existe des modèles bioinformatiques qui permettent une plus grande rigueur.)

Ensuite, c'est le même principe, et on peut raisonner par analogie :

Admettons qu'une protéine allostérique possède 2 sites de fixations à un ligand. Ici $n=2$, et on donne alors les vitesses d'association et de dissociation (cf plus haut)

- $v(\text{association}) = k_1 \times [P] \times [L] \times [L]$ **ORDRE 3**
- $v(\text{dissociation}) = k_{-1} \times [PL]$ **ORDRE 1**

Ainsi, à l'équilibre, $v(\text{association}) = v(\text{dissociation})$, et :

$$K = \frac{k_{-1}}{k_1} = \frac{[P] \times [L]^2}{[PL_n]}$$

Si on raisonne avec **n sites de fixations** à un ligand, on a alors :

$$K = \frac{[P] \times [L]^n}{[PL_n]}$$

Attention !! Ici K ne correspond pas à la concentration en ligand pour laquelle on a demi-saturation de la protéine (N'a d'ailleurs pas l'unité d'une concentration : vérifiez pour n=2 par exemple). A demi-saturation, on a $[P] = [PL_n]$, et donc $K = [L]^n$.
K est donc la concentration en ligand élevée à la puissance « n » pour laquelle la protéine allostérique est à demi-saturée.

On peut également écrire la saturation de la protéine :

$$Y = \frac{[PL]}{[PL] + [P]} = \frac{[PL]}{[P_0]}$$

Et comme précédemment (cf Partie III), on peut dire que

$$[P] = \frac{K \times [PL_n]}{[L]}$$

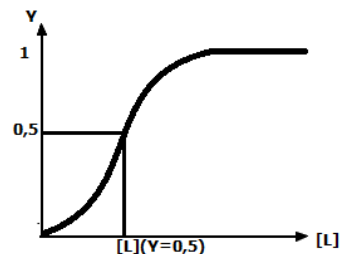
Et en déduire que :

$$Y = \frac{[L]^n}{[L]^n + K} \quad \text{Or,} \quad K = [L]_{0,5}^n$$

Enfin,

$$Y = \frac{[L]^n}{[L]^n + [L]_{0,5}^n}$$

Ce qui correspond à une fonction de type sigmoïde.



Rappelez-vous : On a pris ici le cas d'une coopération parfaite (les 4 molécules de ligand se fixaient à la protéine simultanément).

- Dans ce cas là, on aurait donc eu $n =$ nombre de sites de fixation (ou nombre de protomères, « on parle faussement de n sous-unités en interactions »). On appellera ce nombre n' par la suite, qui correspond donc au nombre de sites de fixation.
- Dans la réalité, ça n'existe pas, et n représente plutôt le nombre de Hill que l'on va appeler n_h par la suite, qui rend compte de l'interaction entre les sites allostériques.
- On définit le nombre de Hill n_h ainsi :

$$\frac{\text{Nombre de sites en interaction } n_h}{\text{Nombre de sites } n'} = \text{Valeur comprise entre 0 et 1}$$

- Ainsi, tout à l'heure nous avons le cas d'une coopération parfaite où le nombre de site en interaction était égal au nombre effectif, donc la valeur était de 1. On aurait par exemple eu pour l'hémoglobine (4 sites de fixations)

$$\frac{n_h = 4}{n' = 4} = 1$$

- Mais expérimentalement, on mesure plutôt un nombre de Hill de 3 (ou 2,8) pour l'hémoglobine, la coopérativité n'est pas parfaite, ceci est dû au fait que les sous-unités ne lient pas simultanément une molécule d'O₂. On a donc :

$$\frac{n_h = 2,8}{n' = 4} = 0,7$$

AINSI, pour l'hémoglobine, on n'écrit pas $Y = \frac{pO_2^4}{pO_2^4 + p50^4}$ mais plutôt $Y = \frac{pO_2^{2,8}}{pO_2^{2,8} + p50^{2,8}}$

car la coopérativité entre les sites n'est pas totale, et le nombre de Hill (2,8) n'est pas égal au nombre de sites ($n'=4$)

La valeur du nombre de Hill n_h , peut prédire la coopérativité des sites :

- $n_h = 1$, la protéine n'est pas allostérique mais Michaëlienne (car $Y = [L]^n / ([L]^n + Kd)$)
- $n_h > 1$, on parle de coopération positive, la fixation d'un ligand facilite la fixation d'un autre ligand sur un autre site.
- $n_h < 1$ (rare), on parle de coopération négative, la fixation d'un ligand rend plus difficile la fixation d'un autre ligand sur un autre site .

Si on reprend notre fonction de saturation : $Y = \frac{[L]^n}{[L]^n + K}$

On applique la fonction « 1-x » : $1 - Y = 1 - \frac{[L]^n}{[L]^n + [L]_{0,5}^n}$

Peut aussi s'écrire : $1 - Y = \frac{[L]^n + [L]_{0,5}^n}{[L]^n + [L]_{0,5}^n} - \frac{[L]^n}{[L]^n + [L]_{0,5}^n}$

Par soustraction : $1 - Y = \frac{[L]_{0,5}^n}{[L]^n + [L]_{0,5}^n}$

On inverse : $\frac{1}{1 - Y} = \frac{[L]^n + [L]_{0,5}^n}{[L]_{0,5}^n}$

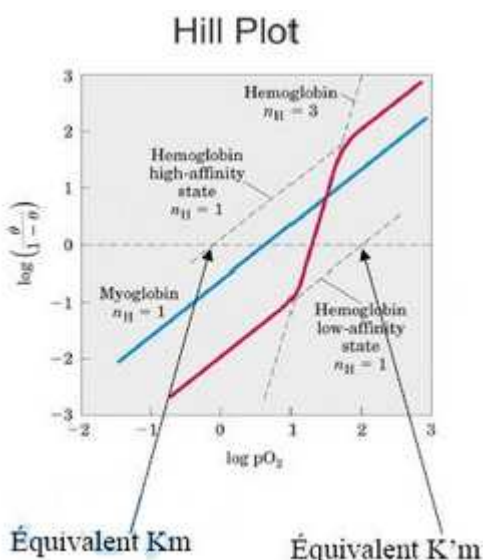
On multiplie par Y et on simplifie: $\frac{Y}{1 - Y} = \frac{[L]^n + [L]_{0,5}^n}{[L]_{0,5}^n} \times \frac{[L]^n}{[L]^n + [L]_{0,5}^n} = \frac{[L]^n}{[L]_{0,5}^n}$

On applique la fonction « ln(x) » $\ln\left(\frac{Y}{1 - Y}\right) = \ln\left(\frac{[L]^n}{[L]_{0,5}^n}\right) = \ln([L]^n) - \ln([L]_{0,5}^n)$

Enfin : $\ln\left(\frac{Y}{1 - Y}\right) = n \times \ln([L]) - n \times \ln([L]_{0,5})$

Equation d'une droite de la forme : $y = a x + b$
Où n est le coefficient directeur de la droite. (remarque: b peut aussi bien être $-\ln(K)$)

VI- Comportement d'une protéine allostérique



La courbe sigmoïde présente une transition allostérique autour de sa $L_{0,5}$, de part et d'autre le comportement s'apparente à un comportement Michaëlien. Ceci est encore plus visible si on linéarise la fonction de saturation comme le montre la courbe ci-contre. Il existe une transition allostérique pour laquelle le nombre de Hill rentre en jeu (autour de $Y=0,5$). De cette linéarisation, on peut déduire pour une protéine allostérique :

- Une plus faible affinité pour les concentrations en ligand faible que pour une protéine Michaëlienne.
- Une plus forte affinité pour les concentrations en ligand forte que pour une protéine Michaëlienne.