

Development of Complex Curricula for Molecular Bionics and Infobionics Programs within a consortial* framework**

Consortium leader

PETER PAZMANY CATHOLIC UNIVERSITY

Consortium members

SEMMELWEIS UNIVERSITY, DIALOG CAMPUS PUBLISHER

The Project has been realised with the support of the European Union and has been co-financed by the European Social Fund ***

**Molekuláris bionika és Infobionika Szakok tananyagának komplex fejlesztése konzorciumi keretben

***A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

INTRODUCTION TO BIOPHYSICS

(Bevezetés a biofizikába)

MULTIPLE EQUILIBRIA

(Többszörös egyensúlyok)

GYÖRFFY DÁNIEL, ZÁVODSZKY PÉTER

Introduction

- Multiple equilibria occur when small molecules bind to large molecules with multiple binding sites, such as hormones to receptors, substrates to enzymes, antigens to antibodies, or oxygen to haemoglobin

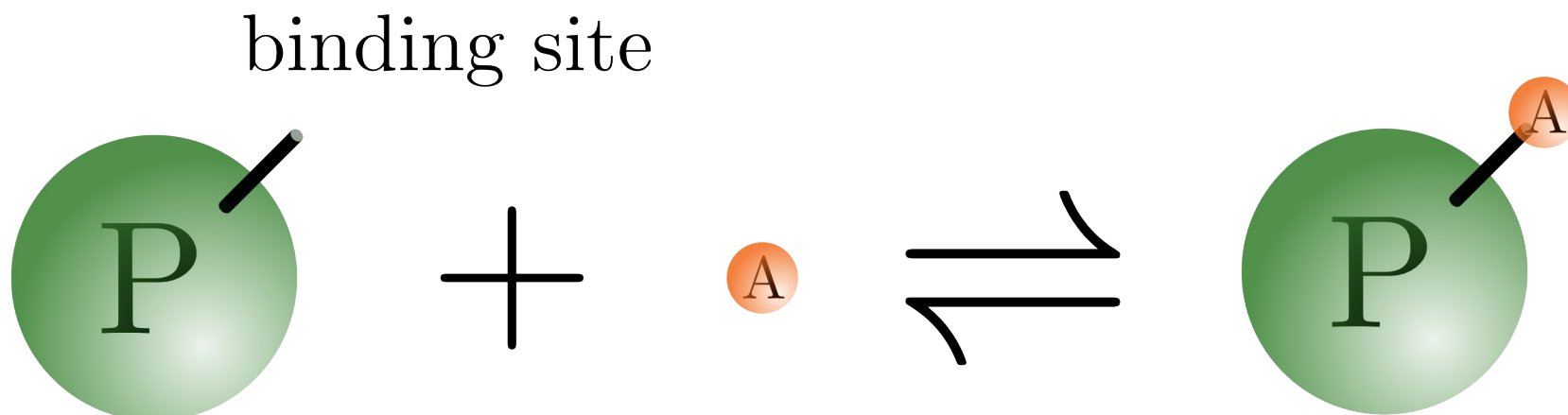
Law of mass action

- Multiple equilibria are based on the *law of mass action*
- Let us consider a reaction



where P is a macromolecule, say a protein with one binding site, A is an unbound ligand, PA is a complex of the protein and the ligand and K_a is the *association constant*

Ligand binding of a protein with one binding site



- The *law of mass action* for this reaction is

$$K_a = \frac{[P - A]}{[P][A]}$$

where $[P]$, $[A]$ and $[P-A]$ are the concentration of the protein, the unbound ligand and the complex, respectively

- Let us introduce a new quantity, $\bar{v}$, as the ratio of the number of moles of bound A and the total number of moles of protein P, that is

$$\bar{v} = \frac{[P - A]}{[P - A] + [P]}$$

- We can combine the equations for $\bar{v}$ and K_a to eliminate all quantities with a P in them
- This is because [P-A] and [P] always appear as ratios

$$K_a = \frac{[P - A]}{[P][A]}$$

and after rearrangement

$$K_a[A] = \frac{[P - A]}{[P]}$$

- For $\bar{v}$

$$\bar{v} = \frac{[P - A]}{[P] + [P - A]}$$

and dividing by $[P]$ both the numerator and denominator on the right hand side we get

$$\bar{v} = \frac{[P - A]/[P]}{1 + [P - A]/[P]}$$

- So

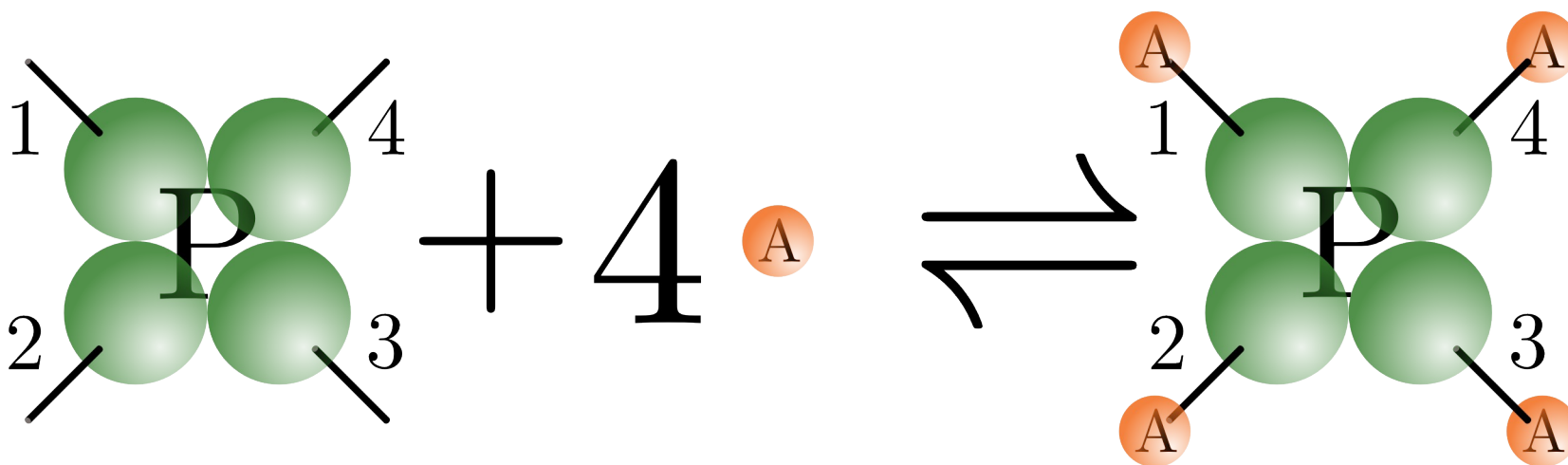
$$\bar{v} = \frac{K_a [A]}{1 + K_a [A]}$$

- This means that the amount of bound ligand per protein molecule depends only on the concentration of the ligand and not on the concentration of protein
- Let us note, however, that it is the unbound concentration of A which is important

Proteins with several binding sites

- In most cases, proteins have several sites for binding
- Now, let us consider proteins having several identical subunits
- We also assume that these are independent binding sites, i.e. site 1 does not sense whether site 2 is occupied

Protein with four subunits and binding sites



- Let $\bar{v}$ denote again the ratio of the total number of moles of the bound ligand and the total number of moles of the protein
- The total number of moles of bound ligand is the sum of the number of moles of ligand bound to the different binding sites

$$\bar{v} = \bar{v}_1 + \dots + \bar{v}_4 = \frac{[P_1 - A]}{[P] + [P - 4 \cdot A]} + \dots + \frac{[P_4 - A]}{[P] + [P - 4 \cdot A]}$$

where

$$[P - 4 \cdot A] = [P_1 - A] = [P_2 - A] = [P_3 - A] = [P_4 - A]$$

- Making use of the equation for K_a

$$K_a[A] = \frac{[P_i - A]}{[P]}$$

where

$$i = 1, 2, 3, 4$$

we get

$$\bar{v} = \frac{K_a[A]}{1 + K_a[A]} + \dots + \frac{K_a[A]}{1 + K_a[A]} = \frac{4 K_a[A]}{1 + K_a[A]}$$

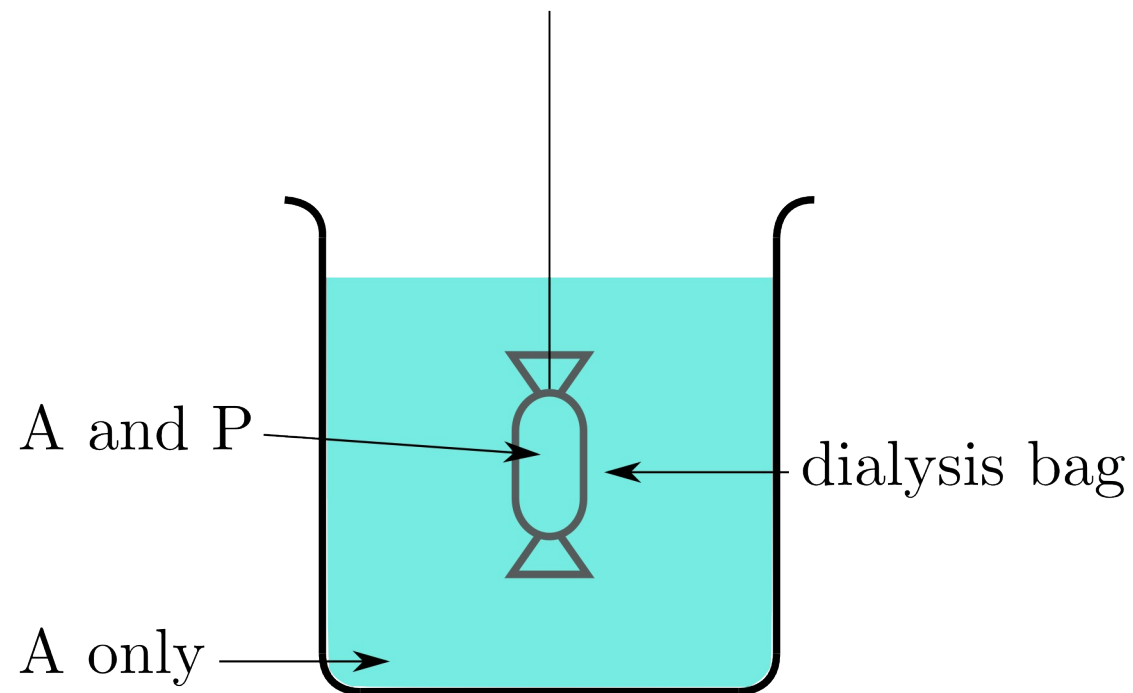
- In general for a protein with n equivalent, independent binding sites

$$\bar{v} = \frac{n K_a [A]}{1 + K_a [A]}$$

- Let us note that the amount of A bound to P is dependent only on n , K_a and the concentration of unbound A, $[A]$

- Now the question arises how we can evaluate n and K_a
- We can get these if we measure how much A is bound to P
- Let us consider an equilibrium dialysis experiment

Dialysis experiment



- Place a solution of the protein, P, in the bag and the bag in a solution of the ligand, A
- A can pass through the pores of the bag but P cannot
- What do we know with precision?
 - The mass of the protein, m_p , and that of the ligand, m_A
 - The molecular weight of the protein, M_p , and that of the ligand, M_A
 - The number of moles of the protein, $n_p = m_p / M_p$, and that of the ligand, $n_A = m_A / M_A$

- Let us assume that

$$[A]_{inside} = [A]_{outside}$$

- Now we only need to measure $[A]_{out}$
- We can do this by spectrophotometry or radioactive labelling

- Based on these measurements, we can calculate $\bar{v}$

$$\bar{v} = \frac{\text{moles of bound } [A]}{\text{total moles of } P} = \frac{\text{total } A - \text{free } A}{\text{total } P} = \frac{n_A - [A]V}{n_P}$$

where V is the total volume of the solution

Scatchard plot

- The Scatchard plot is a plot helping us to obtain n and K_a
- Let us set out from the expression

$$\bar{v} = \frac{n K_a [A]}{1 + K_a [A]}$$

- Performing rearrangements we obtain

$$\bar{v} \left(1 + K_a [A] \right) = n K_a [A]$$

so

$$\bar{v} + \bar{v} K_a [A] = n K_a [A]$$

and

$$\bar{v} = n K_a [A] - \bar{v} K_a [A]$$

- And finally we get the linear relationship

$$\frac{\bar{v}}{[A]} = n K_a - \bar{v} K_a$$

- Let us plot $\bar{v}/[A]$ against $\bar{v}$
- The slope of the curve will be $-K_a$
- The $\bar{v}$ intercept is n : at

$$\bar{v}/[A] = 0,$$

$$n K_a = n \bar{v}$$

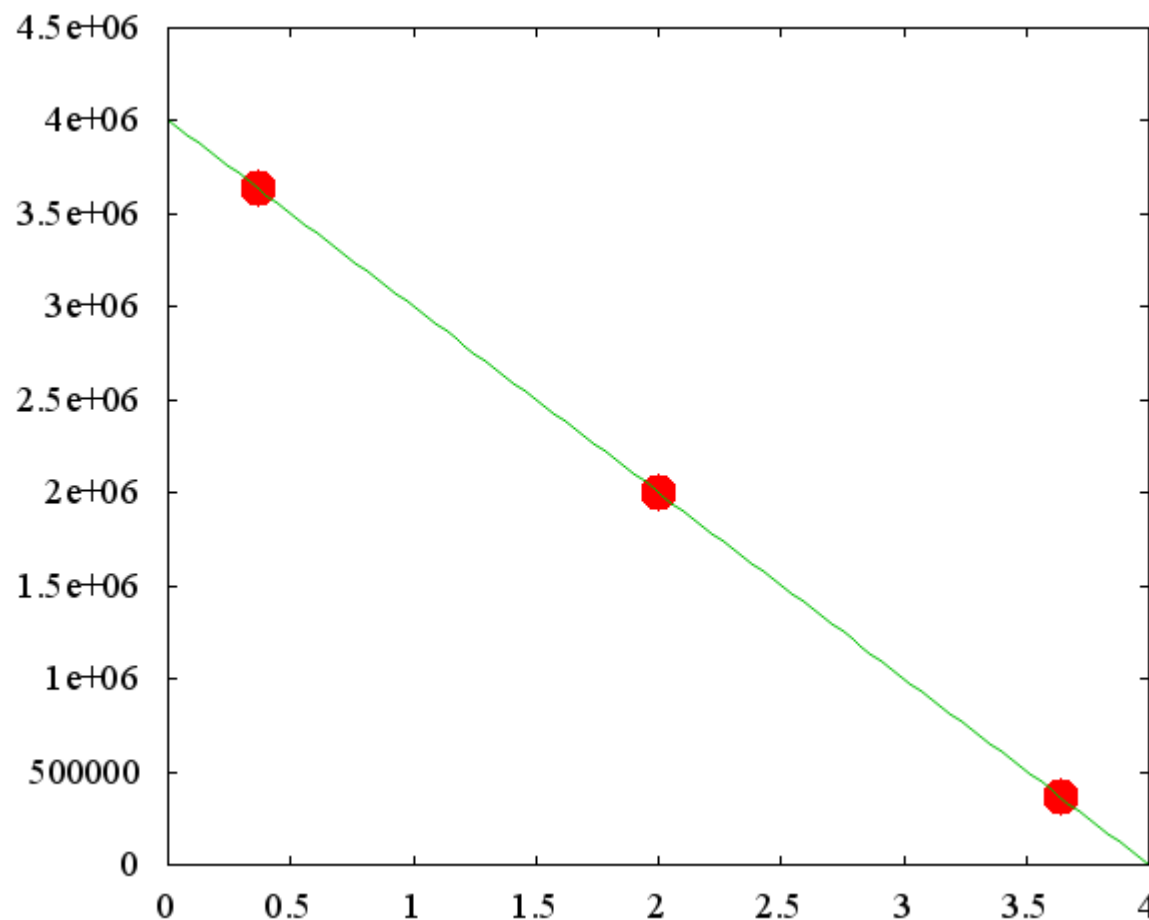
so

$$n = \bar{v}$$

- Let us consider an example

$[A]$ (M)	$\bar{v}$	$\bar{v}/[A]$ (M ⁻¹)
10^{-7}	0.364	$3.64 \cdot 10^6$
10^{-6}	2.000	$2.00 \cdot 10^6$
10^{-5}	3.64	$3.64 \cdot 10^5$

Scatchard plot



- In this example:

$$-slope = K_a = 10^6 M^{-1}$$

and

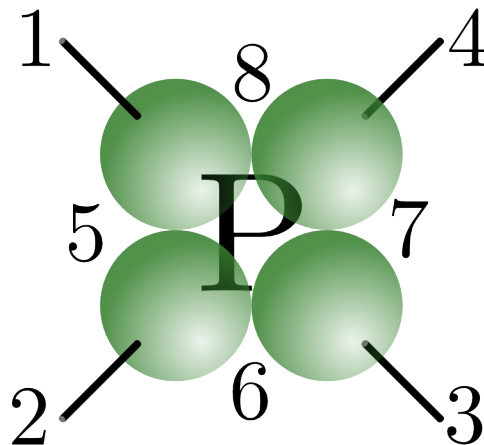
$$n = 4$$

- But we encounter a problem:
 - Scatchard plots are never linear in reality
- There are four reasons for the non-linearity of Scatchard plots:
 - Binding site heterogeneity: more than one class of binding sites with different K_a 's
 - Donnan potential: if A is charged, $[A]_{in} \neq [A]_{out}$. Most biological molecules are charged
 - Debye-Hückel effect: if P and A are charged, then when A binds to P there is an electrostatic interaction in addition to the normal binding site
 - Binding site cooperativity: binding sites are not independent

Binding site heterogeneity

- Let us suppose that there are two classes of binding sites:
 - Class 1 with K_{a1} , n_1
 - Class 2 with K_{a2} , n_2
- Let us consider a protein with two classes of binding site, each having four actual sites

Binding site heterogeneity



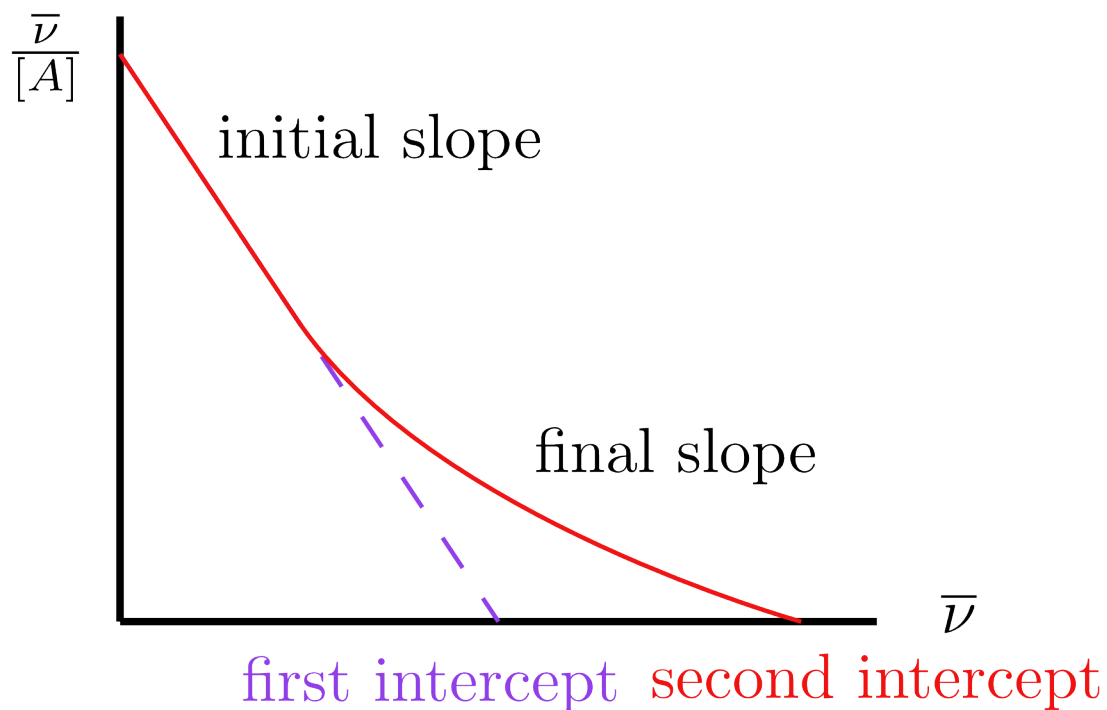
- Class 2 binding sites might be adventitious binding sites which frequently exhibit non-specific weak binding
- These class 2 sites affect the total binding

- If the binding sites are still independent, then

$$\begin{aligned}\bar{v} &= \bar{v}_1 + \bar{v}_2 + \dots + \bar{v}_8 \\ &= \frac{n_1 K_{a1} [A]}{1 + K_{a1} [A]} + \frac{n_2 K_{a2} [A]}{1 + K_{a2} [A]}\end{aligned}$$

- Now, let us plot $\bar{v}/[A]$ vs. $\bar{v}$ again

Scatchard plot for binding site heterogeneity



- The initial slope is

$$\text{initial slope} = \frac{-n_1 K_{a1}^2 + n_2 K_{a2}^2}{n_1 K_{a1} + n_2 K_{a2}}$$

- If $K_{a1} \gg K_{a2}$, then

$$\text{slope} \approx -K_{a1}$$

- But if K_{a1} is only one order of magnitude greater than K_{a2} , this will give a 10% error

- The final slope is

$$\text{final slope} = \frac{-(n_1 + n_2) K_{a1} K_{a2}}{n_1 K_{a1} + n_2 K_{a2}}$$

- If $K_{a1} \gg K_{a2}$, then

$$\text{slope} \approx \frac{-(n_1 + n_2)}{n_1} K_{a2}$$

- The first intercept is

$$\text{first intercept} = \frac{(n_1 K_{a1} + n_2 K_{a2})^2}{n_1 K_{a1}^2 + n_2 K_{a2}^2}$$

- If $K_{a1} \gg K_{a2}$, then

$$\text{intercept} \approx n_1$$

- The second intercept is

$$\text{second intercept} = n_1 + n_2$$

where n_1 and n_2 must be integers

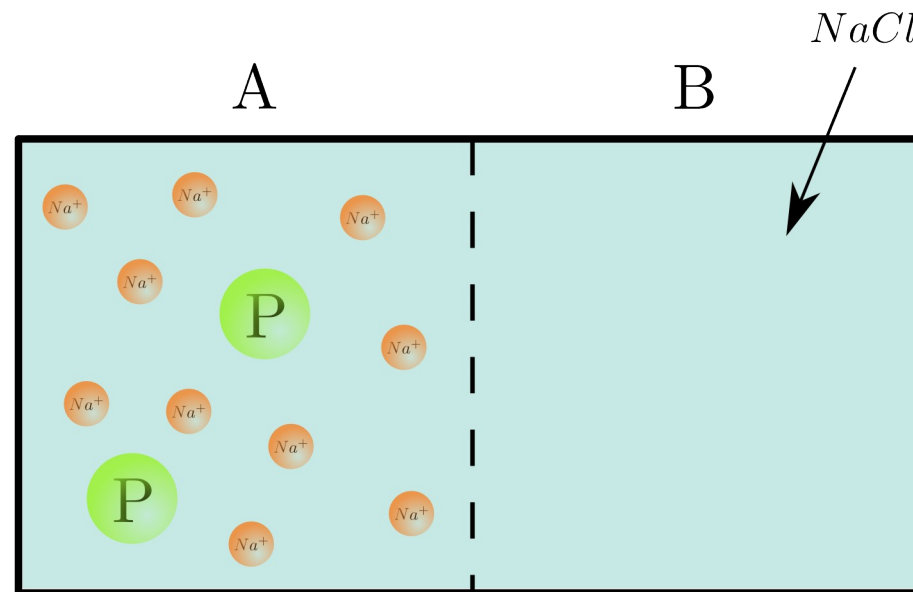
- Thus we have four unknowns but we also have four equations
- Thus, we can calculate n_1 , n_2 , K_{a1} and K_{a2}

Donnan effects

- Let us consider the system shown below
- The protein on the left-hand side releases five Na^+ ions:



Donnan effect



Concentrations of ion species

Ion species	Concentration on the A side	Concentration on the B side
P	a	0
Na^+	$5a+x$	$b-x$
Cl^-	x	$b-x$

- Let us recall that we added NaCl to side B to avoid generating a large Donnan potential

$$(c_+ c_-)_A = (c_+ c_-)_B$$

SO

$$(5a + x)(x) = (b - x)^2$$

- Thus x is from this equation

$$x = \frac{b^2}{5a + 2b}$$

- Thus the Donnan potential is

$$\Delta \Phi = \frac{RT}{zF} \ln \frac{c_{+A}}{c_{+B}} = \frac{RT}{zF} \ln \frac{(5a + x)}{b - x}$$

- How does this Donnan potential influence the binding of small molecule A which has two positive charges?
- Since

$$[A^{2+}] \ll [Na^+]$$

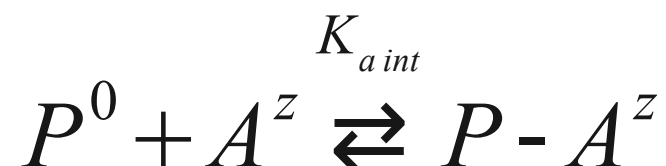
to a first approximation, we can neglect the contribution of A^{2+} to the Donnan potential

- But A^{2+} still feels the Donnan potential.

- If we ignored the Donnan potential, it would appear that the protein binds one A^{2+} molecule although it does not
- We can set the pH of the solution to the isoelectric point of protein to avoid this effect

Debye-Hückel effects

- Let us consider the case when the protein is uncharged (is at its *isoelectric point*)

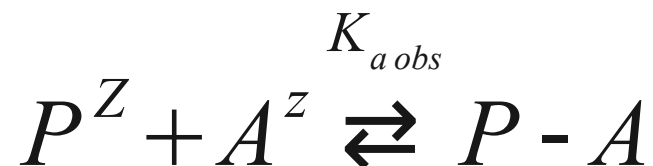


where z is the charge on A

- There is no electrostatic interaction so $K_{a \text{ int}}$ is the intrinsic association constant

$$\Delta G_{int}^0 = -RT \ln K_{a \text{ int}}$$

- Now, let us compare the case when the protein is charged



where Z and z are the charge on the protein and on A , respectively

- $K_{a\,obs}$ is the observed association constant which includes electrostatic interactions

$$\Delta G_{obs}^0 = -RT K_{a\,obs}$$

- The difference between these standard free energies is equal to the electrical work necessary to bring the charges together

$$\Delta G_{obs}^0 - \Delta G_{int}^0 = \Phi(r) z e N$$

where r is the radius of the protein

- Substituting the expression with the association constants into the equation above, we get

$$RT \ln \frac{K_{a\,obs}}{K_{a\,int}} = -\Phi(r) z e N$$

- Thus after rearrangement

$$K_{a\,obs} = K_{a\,int} e^{-z e N \Phi(r) / RT}$$

- For an ion atmosphere

$$\Phi(r) = \frac{Z e}{D r} \frac{e^{-\kappa r}}{RT}$$

so it is plausible to write the equation above in the following form

$$K_{a\,obs} = K_{a\,int} e^{-2wZz}$$

where

$$w = \frac{N e^2}{Z D r} \frac{e^{-\kappa r}}{RT}$$

- For the binding of a small molecule A with charge z to a protein P with charge Z

$$\bar{\nu} = \frac{\text{number of moles of bound } A}{\text{total number of moles of protein } P}$$

may be written in terms of an observed association coefficient $K_{a \text{ obs}}$

- $K_{a\text{ obs}}$ is affected by the charge interactions and may be expressed in terms of the intrinsic association constant $K_{a\text{ int}}$
- The term

$$e^{-2wZz}$$

has to do with bringing the charged molecule to the protein surface from infinity

- Let us consider the following example



so the charge, z , on A is 2

- If we begin at the isoelectric point, $Z = 0$, initially

$$K_{a\text{obs}} = K_{a\text{int}}$$

for the first A^{2+}

- However, the second A^{2+} to bind will see a charge of +2 on the protein
- How does this affect $K_{a\text{ obs}}$

A to bind	Z	$K_{a\text{ osb}}$
1.	+2	$1 \cdot K_{a\text{ int}}$
2.	+4	$0.45 \cdot K_{a\text{ int}}$
3.	+6	$0.2 \cdot K_{a\text{ int}}$
4.	+8	$0.09 \cdot K_{a\text{ int}}$

- Let us note that this was calculated using a value of $w=0.1$ (w is always on the order of 0.1)
- It is dimensionless and varies with the protein radius and the ionic strength

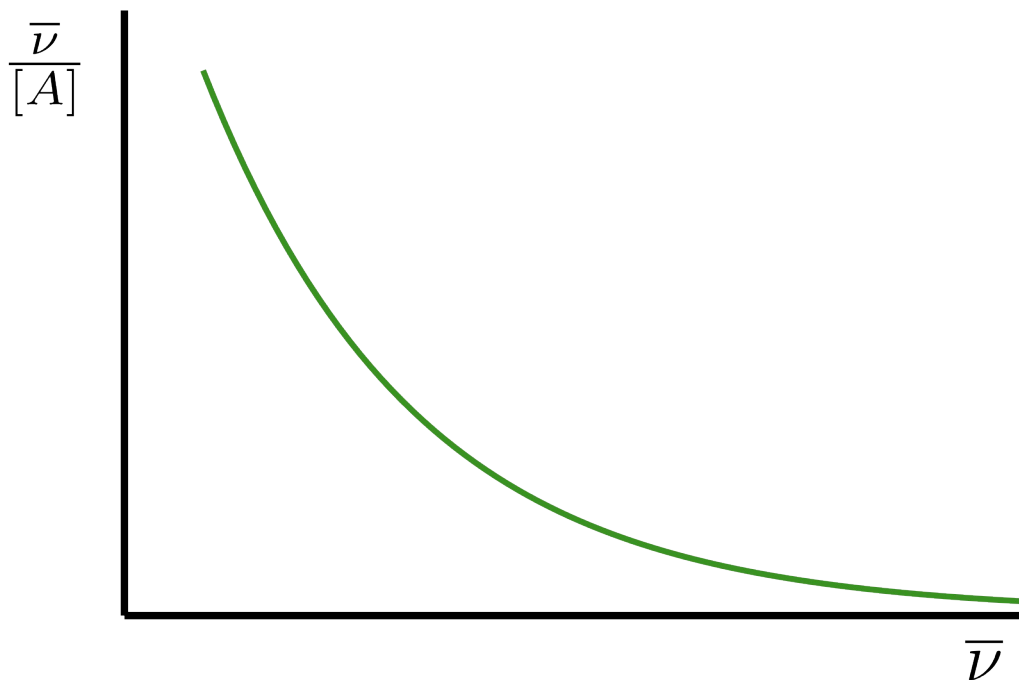
$$w = \frac{N e^2}{Z D r} \frac{e^{-\kappa r}}{RT}$$

and from the Debye-Hückel theory

$$\kappa^2 = \frac{8 \pi N e^2 I}{1000 D k_B T}$$

- Let us note that $K_{a\text{ obs}}$ decreases by more than an order of magnitude from $K_{a\text{ int}}$ in this example
- Even beginning at the isoelectric point, the Scatchard plot is badly curved if we do not take the Debye-Hückel theory into account
- Let us plot $\bar{v}/[A]$ as a function of $\bar{v}$

$\bar{\nu}/[A]$ vs. $\bar{\nu}$



- When we are binding more than one charged molecule to a protein we get a curved plot
- Instead:

$$\bar{v} = \frac{n K_{a \text{ int}} e^{-2 w Z z} [A]}{1 + K_{a \text{ int}} e^{-2 w Z z} [A]}$$

- From this equation we get

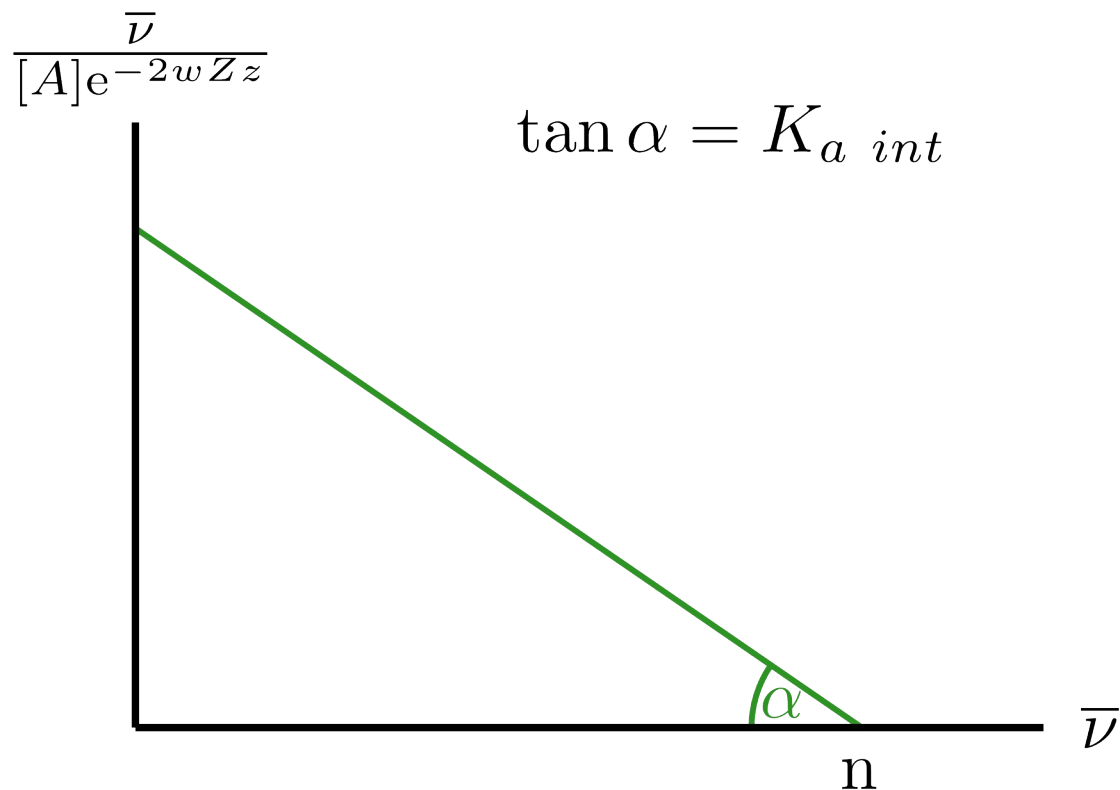
$$\frac{\bar{v}}{[A] e^{-2 w Z z}} = n K_{a \text{ int}} - K_{a \text{ int}} \bar{v}$$

- Now, let us plot

$$\frac{\bar{v}}{[A] e^{-2wZz}}$$

as a function of $\bar{v}$ to get a linear relationship

$$\bar{\nu}/[A]e^{-2wZz} \text{ vs. } \bar{\nu}$$



- This type of analysis may be applied to protein titration, that is the binding of protons to proteins
- Protons are small charged particles, and proteins have specific binding sites for them
- These are the acidic and basic side chains of amino acids
- They can be classified by their chemical character and characterized by their $pK_{a \text{ int}}$

- All negative charges on the surface of a protein are due to ionized acidic groups
- All positive charges on the protein surface are due to ionized basic groups
- The titration curve for a protein can be generated from the amino acid composition

- Deviations from this behaviour:
 - When the protein has a large positive or negative charge (due to gain or loss of protons) it begins to bind anions or cations from the solution
 - At very low or very high pH, the protein denatures exposing buried groups and changing the radius of the protein, thus changing w

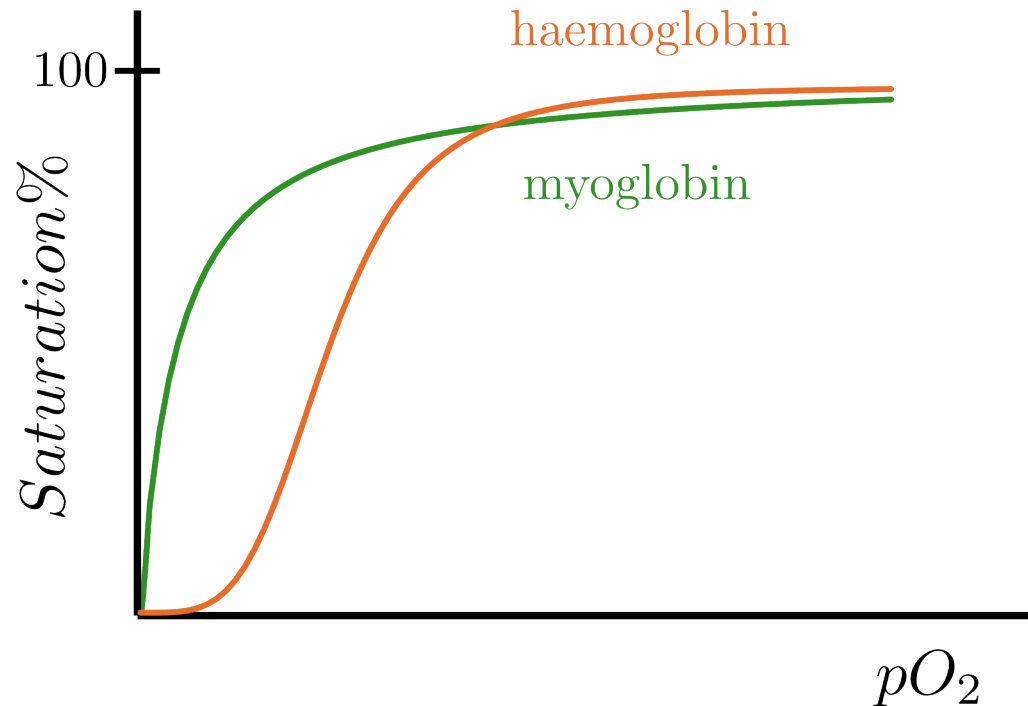
Cooperativity

- Let us suppose that the binding sites are not independent of each other, that is there is communication between binding sites
- As an example, let us compare haemoglobin and myoglobin

Comparison of haemoglobin and myoglobin

Haemoglobin	Myoglobin
Cooperativity	No cooperativity
4 subunits	1 polypeptide chain
4 hem groups	1 heme group
Binds 4 O ₂ 's	Binds 1 O ₂
O ₂ storage	O ₂ transport in the blood

Comparison of saturation curves of myoglobin and haemoglobin



- Its simple to explain the saturation curve for myoglobin

$$\bar{v} = \frac{n K_a [O_2]}{1 + K_a [O_2]}$$

- In the case of myoglobin

$$n = 1$$

- Now, let us switch to the use of partial pressure instead of molar concentration of oxygen

$$p O_2 = \beta [O_2]$$

where β is a constant

- So

$$\bar{v} = \frac{K_a' p O_2}{1 + K_a' p O_2}$$

where

$$K_a' = K_a / \beta$$

- The experimental curve fits the calculated curve for myoglobin almost perfectly
- For haemoglobin, things do not work this well
- We could try lots of different equations of the form

$$\bar{v} = \sum_{i=1}^n \frac{n_i K_{ai} [O_2]}{1 + K_{ai} [O_2]}$$

but this will always give a curve with decreasing slope, it will not give the sigmoidal curve

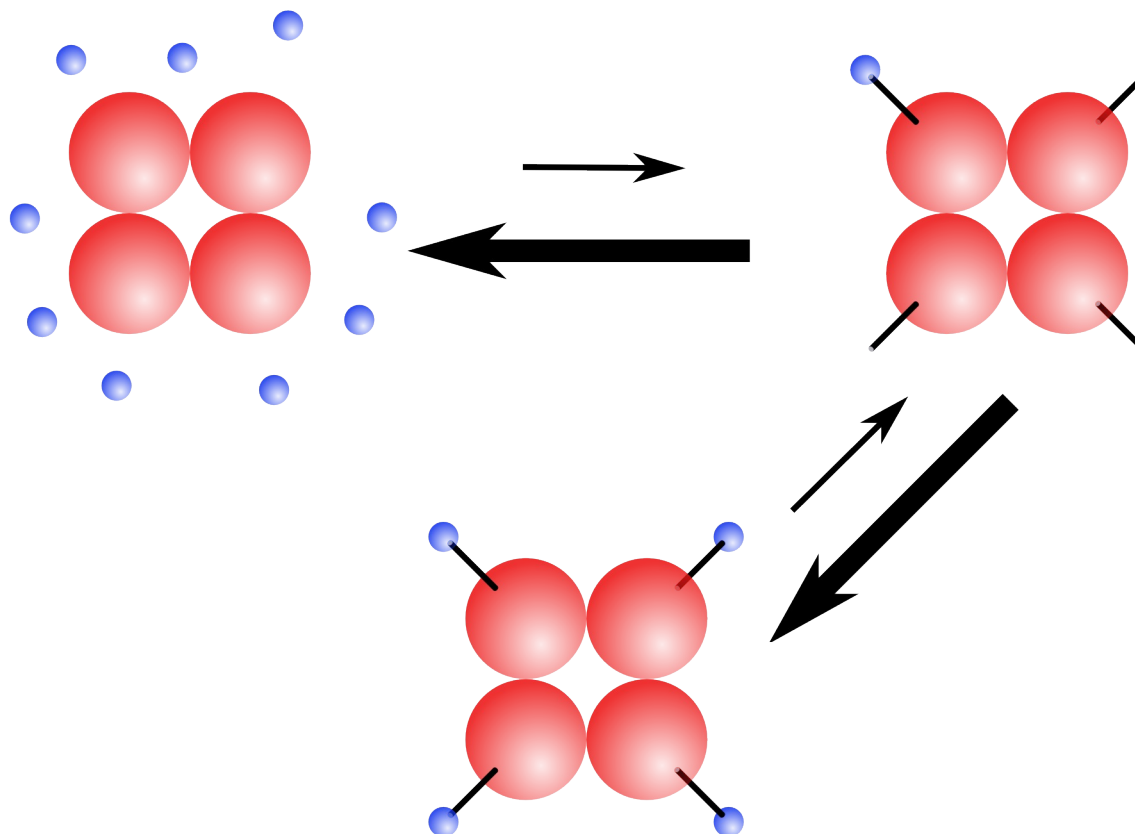
- This can also not be due to a Donnan or Debye-Hückel effect because O_2 is not charged
- Something new is going on at molecular level – this is the cooperativity

- How does cooperativity work at molecular level?
- Let us consider two kinds of subunits, α and β
- Binding sites on α subunits are stronger
- O_2 binds to an α subunit which induces a conformational change and thus an increased affinity of β to O_2 (positive cooperativity)

- As a consequence, it allows haemoglobin to release O_2 at higher pO_2 than myoglobin
- Haemoglobin dumps O_2 into the tissues over a very narrow range of pO_2
- This keeps the pO_2 more constant throughout the body

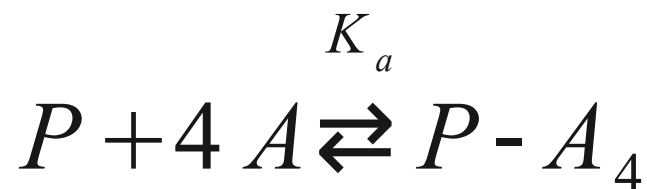
- Now, let us devise a model for extreme positive cooperativity
- Let us consider a protein with four subunits with four hidden binding sites
- A high concentration of A is necessary to bind the first O_2 to the first site

Positive cooperativity



- This model predicts only two kinds of large molecule
 - With no O_2 is bound
 - With 4 O_2 's are bound
- There is a negligible amount of the forms 1,2 and 3 O_2 's bound

- The model predicts that we will have only extremes of saturation: 0% and 100%



where

$$K_a = \frac{[P - A_4]}{[P][A]^4}$$

- The saturation is

$$\bar{v} = \frac{\text{moles of bound } A}{\text{moles of total } P} = \frac{4 [P - A_4]}{[P] + [P - A_4]}$$

and after substitutions and rearrangement

$$\bar{v} = \frac{4 K_a [A]^4}{1 + K_a [A]^4}$$

- This equation gives an S-shaped curve

- Let us recall that for independent binding sites

$$\bar{v} = \frac{4 K_a [A]}{1 + K_a [A]}$$

- What if we do not have extreme positive cooperativity?
- More realistically, we cannot neglect the other partially saturated species
- There is an intermediate case

Hill plot

- The Hill equation is

$$\bar{v} = \frac{n K_a [A]^x}{1 + K_a [A]^x}$$

where n is the number of binding sites and x is the cooperativity and

$$1 \leq x \leq n$$

- If

$$x = n$$

an extreme cooperativity exists

- If

$$x = 1$$

there is no cooperativity, so the binding sites are independent of each other

- Hill did not prove this equation, he just showed that it works pretty well
- Subsequently, it has been shown why it works well

- Now, we would like to know how we get x
- Let us take the Hill equation and solve it for $k[A]^x$, then take the logarithm of both sides

$$k[A]^x = \frac{\bar{v}}{n - \bar{v}}$$

$$\log \left(\frac{\bar{v}}{n - \bar{v}} \right) = \log k + x \log [A]$$

- Hill plot is a plot of

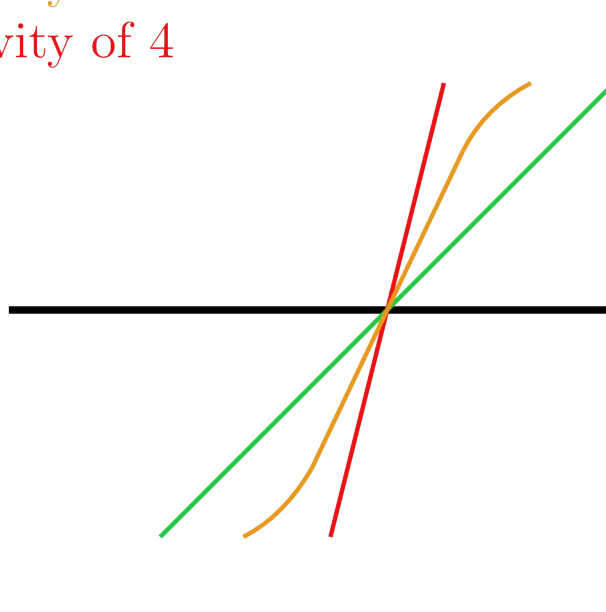
$$\log \left(\frac{\bar{v}}{n - \bar{v}} \right)$$

VS.

$$\log [A]$$

Hill plot

no cooperativity
cooperativity of 3
cooperativity of 4



- For independent binding sites, we get a slope of 1
- For extreme positive cooperativity, we get a slope of n (in this case, 4)
- For intermediate positive cooperativity, the plot is straight in the centre with a slope of x
- It tapers off at the ends to a slope of 1
- The cooperativity, x , is the slope of the curve at the centre

Archibald Vivian Hill (1886-1977)

