

# Mass Spectrometry Analysis of Binding to a Dimer

Binding to a dimer (or any molecule containing two sites) gives rise to three molecular species (apo, one ligand bound, two ligands bound) that can be distinguished in mass spectra. The two binding events can be independent, positively or negatively cooperative. Traditional methods of analysis distinguish only two species (unbound and total bound), consequently, MS data provide more information rich data on such processes, which in principle can lead to a better understanding of cooperative binding. This program extends tight-binding analysis to deal with the case of dimer (two-site) binding in order to determine the first and second binding affinities.

The following script is based on concepts and equations introduced by Monod, Wyman & Changeux (1965) J Mol Biol 12, 88-118; Koshland, Nemethy & Filmer (1966), Biochemistry 5, 365-85 and by Koshland in his article 'The molecular basis of enzyme recognition' in 'The Enzymes' Academic Press, New York: 1970. Despite positing different physical mechanisms for cooperativity, both MWC and KNF models have the same fundamental mathematical form and we use a simplified 'mechanism-free' notation closer to that of the review by Hunter & Anderson (2009), Angew Chem Int Ed Engl 48, 7488-99. The fitting procedure itself extends the 'tight-binding' approach for analysis of single-binding in terms of fraction of protein bound (Williams & Morrison (1979) Methods Enzymol 63, 437-67) to deal with the possibility of cooperative binding in systems with two equivalent binding sites.

$K_1$  = association constant for first binding event

$K_2$  = association constant for 2nd binding event

$\alpha = K_2/K_1$  cooperativity

$A_{\text{tot}}$  = total concentration of protein monomer

$S_{\text{tot}}$  = total concentration of ligand

We assume (or know experimentally) that the dimer is constitutive *i.e.* that total concentration of dimer  $D_{\text{tot}} = A_{\text{tot}}/2$ ,

consequently we define

$S$  = concentration of free ligand

$T_0 = D_0/D_{\text{tot}}$  = fraction of unbound dimer

$T_1 = D_1/D_{\text{tot}}$  = fraction of dimer bound to one ligand

$T_2 = D_2/D_{\text{tot}}$  = fraction of dimer bound to two ligands

In these circumstances, we can by inspection write down the concentrations of each species,

```

Equations = {
  S == Stot - T1 * Dtot - 2 * T2 * Dtot,
  T0 == 1 / (1 + 2 * K1 * S +  $\alpha$  * K12 * S2),
  T1 == (2 * K1 * S) / (1 + 2 * K1 * S +  $\alpha$  * K12 * S2),
  T2 == ( $\alpha$  * K12 * S2) / (1 + 2 * K1 * S +  $\alpha$  * K12 * S2)
};

Collect[Eliminate[Equations, {S, T1, T2}], T0]
Collect[Eliminate[Equations, {S, T0, T2}], T1]
Collect[Eliminate[Equations, {S, T1, T0}], T2]

T02 (-4 Dtot K12 Stot + 4 Dtot K12 Stot  $\alpha$ ) + T0 (2 K1 Stot - 4 Dtot K12 Stot  $\alpha$  + K12 Stot2  $\alpha$ ) ==
1 + T0 (-1 - 4 Dtot2 K12  $\alpha$ ) + T03 (4 Dtot2 K12 - 4 Dtot2 K12  $\alpha$ ) + T02 (-4 Dtot2 K12 + 8 Dtot2 K12  $\alpha$ )

-2 K1 Stot + T1 (2 K1 Stot - 2 Dtot K12 Stot  $\alpha$  + K12 Stot2  $\alpha$ ) ==
(-1 - 2 Dtot K1) T1 + T12 (2 Dtot K1 - 2 Dtot K1  $\alpha$  - 2 Dtot2 K12  $\alpha$ ) + T13 (Dtot2 K12  $\alpha$  - Dtot2 K12  $\alpha$ 2)

-K12 Stot2  $\alpha$ 2 + T22 (4 Dtot K12 Stot  $\alpha$  - 4 Dtot K12 Stot  $\alpha$ 2) +
T2 (2 K1 Stot  $\alpha$  - 4 Dtot K12 Stot  $\alpha$  + 4 Dtot K12 Stot  $\alpha$ 2 + K12 Stot2  $\alpha$ 2) ==
T2 (- $\alpha$  - 4 Dtot K1  $\alpha$  - 4 Dtot2 K12  $\alpha$ ) + T23 (4 Dtot2 K12  $\alpha$  - 4 Dtot2 K12  $\alpha$ 2) +
T22 (-8 Dtot K1 + 8 Dtot K1  $\alpha$  + 4 Dtot2 K12  $\alpha$ 2)

```

These can be tidied up to the following three cubic equations  
each describing one of the fractional component molecular species

$$\begin{aligned}
& 4 K1^2 Dtot^2 (1 - \alpha) T0^3 + 4 K1^2 Dtot (-Dtot + 2 Dtot \alpha + Stot - Stot \alpha) T0^2 + \\
& (+ 4 K1^2 Dtot Stot \alpha - K1^2 Stot^2 \alpha - 4 K1^2 Dtot^2 \alpha - 2 K1 Stot - 1) T0 + 1 == 0 \\
& K1^2 Dtot^2 \alpha (1 - \alpha) T1^3 + 2 K1 Dtot (1 - \alpha - K1 Dtot \alpha) T1^2 + \\
& (2 K1^2 Dtot Stot \alpha - K1^2 Stot^2 \alpha - 2 K1 Dtot - 2 K1 Stot - 1) T1 + 2 K1 Stot == 0 \\
& 4 K1^2 Dtot^2 \alpha (1 - \alpha) T2^3 + 4 K1 Dtot (-2 + 2 \alpha - K1 Stot \alpha + K1 Dtot \alpha^2 + K1 Stot \alpha^2) T2^2 + \\
& \alpha (4 K1^2 Dtot Stot - 4 K1^2 Dtot Stot \alpha - 4 K1^2 Dtot^2 - K1^2 Stot^2 \alpha - 2 K1 Stot - 4 K1 Dtot - 1) \\
& T2 + K1^2 Stot^2 \alpha^2
\end{aligned}$$

Although an algebraic solution of each equation is possible the expressions are in the general case rather long and it is not clear which of the three solutions will be real as this changes with the values of the equilibrium constants. Consequently it is much simpler to solve for all the roots numerically in each case. This leaves us with (usually) three possible values of each of T0, T1, T2, however only one combination of values satisfies the following conditions :

- (i) T0 and T1 and T2 each have a real value between 0 and 1 and (ii) T0 + T1 + T2 = 1.

This numerical solution and selection is coded in  
the accompanying *Mathematica* package CooperativeDimer.m,  
together with a Levenberg – Marquardt routine to optimally fit for experimental data on the  
three species for values of K1 and  $\alpha$ . Errors in parameters are estimated from residuals to fit.

## User Definitions

The Mathematica package file CooperativeDimer.m must be in the working directory.

```
In[1]:= workingdirectory =
        "C:\Documents and Settings\ubcg66a\My Documents\Mathematica_Programs";
SetDirectory[workingdirectory];
<< "CooperativeDimer.m";
```

## Experimental Data

Enter concentrations of protein and ligand at each titration point in M and fractions of each molecular species (apo, single ligand bound, two ligands bound) at those points (should add up to 1).  
(error affects data weighting in the fit - values of 0 imply equal weighting of each point)

```
In[4]:= temperature = 298.15;
protein = {44.0, 44.0, 44.0, 44.0, 44.0, 44.0, 44.0, 44.0, 44.0};
ligand = {5.625, 11.25, 16.875, 22.5, 28.125, 33.75, 39.375, 44, 50.625};
apo = {0.831310, 0.635240, 0.46650,
       0.410350, 0.297690, 0.249910, 0.067760, 0.03899, 0.009042};
single = {0.09259, 0.15558, 0.16991, 0.19523, 0.14779,
          0.14248, 0.08781, 0.06352, 0.023901};
double = {0.0761, 0.20918, 0.36359, 0.39442, 0.55452, 0.60761, 0.84443, 0.89749, 0.96706};
error = Table[0.0, {Length[protein]}];
```

## Derived Quantities

(Protein concentrations are multiplied by 0.5 to obtain dimer concentration, in case of two-sites on one protein change this to 1.0)

```
In[11]:= dimer = 0.5 * protein;
freeligand = ligand - single * dimer - 2 * double * dimer;
sitesbound = (single + 2 * double) * dimer;
sitesfree = 2 * apo * dimer + single * dimer;
fractionbound = sitesbound / (sitesbound + sitesfree);
```

## Classic Plots

The 'classic' Hill and Scatchard plots can be produced by estimation of free ligand concentration from Stot, Dtot and measured fractions of apo, singly and doubly bound species. However since the free ligand concentration is not measured directly it is subject to large errors that can lead to erratic plots and numerical errors being thrown

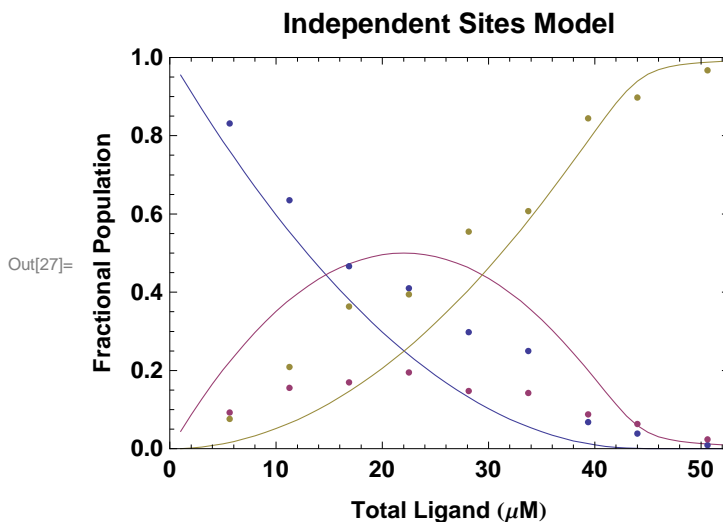
In[16]:=

```
ga = ListPlot[{Transpose[{ligand, apo}], Transpose[{ligand, single}],
  Transpose[{ligand, double}]}], PlotRange → {{0, (Max[ligand] * 1.05)}, {0, 1}},
  Frame → True, FrameLabel → {"Total Ligand ( $\mu$ M)", "Fractional Occupancy"},
  BaseStyle → {FontFamily → "Helvetica", FontSize → 11, FontWeight → "Bold"},
  Axes → False, AspectRatio → 0.7, DisplayFunction → Identity,
  PlotLabel → "Molecular Species"];
ga1 = DataPlot[ligand, apo, {{0, (Max[ligand] * 1.05)}, {0, 1}},
  "Apo", {"Total Ligand ( $\mu$ M)", "Fractional Occupancy"}];
ga2 = DataPlot[ligand, single, {{0, (Max[ligand] * 1.05)}, {0, 1}},
  "Single", {"Total Ligand ( $\mu$ M)", "Fractional Occupancy"}];
ga3 = DataPlot[ligand, double, {{0, (Max[ligand] * 1.05)}, {0, 1}},
  "Double", {"Total Ligand ( $\mu$ M)", "Fractional Occupancy"}];
gb = DataPlot[ligand, fractionbound, {{0, (Max[ligand] * 1.05)}, {0, 1}},
  "Saturation Curve", {"Total Ligand ( $\mu$ M)", "Fraction Bound"}];
gc = DataPlot[fractionbound, (fractionbound / freeligand),
  {{0, 1}, {0, (Max[(fractionbound / freeligand) * 1.05]}},
  "Scatchard Plot", {"Fraction bound", "Fraction Bound/Free Ligand"}];
gd = DataPlot[Log[10, freeligand], Log[10, sitesbound / sitesfree],
  {{(Min[Log[10, freeligand]] - 0.5), (Max[Log[10, freeligand]] + 0.5)}, {Min[
  Log[10, sitesbound / sitesfree] - 0.2, Max[Log[10, sitesbound / sitesfree] + 0.2]}},
  "Hill Plot", {"Log[Free Ligand]", "Log[Bound/Free]"}];
GraphicsGrid[{{ga, gb}, {gc, gd}}]
```

## Independent Sites Model

The cooperative formulation reduces to that of independent binding sites when  $\alpha = 1$ . Fixing this value for  $\alpha$  shows the best fit that can be achieved assuming no cooperativity.

```
In[23]:= bestparameters =
  DimerBindingFit[apo, error, single, error, double, error, ligand, dimer, 1.0];
model = Transpose[Table[FractionalOccupancies[k, Mean[dimer], bestparameters[[1, 1]],
  bestparameters[[1, 2]]], {k, 1, Floor[Max[ligand] + 2], 1}];
ligandmodel = Table[k, {k, 1, Floor[Max[ligand] + 2], 1}]; apomodel = model[[1]];
singlemodel = model[[2]]; doublemodel = model[[3]];
g1 = ListLinePlot[{Transpose[{ligandmodel, apomodel}],
  Transpose[{ligandmodel, singlemodel}], Transpose[{ligandmodel, doublemodel}]},
  InterpolationOrder → 1, PlotRange → {{0, (Max[ligand] + 2)}, {0, 1}},
  Frame → True, FrameLabel → {"Total Ligand ( $\mu\text{M}$ )", "Fractional Population"},
  BaseStyle → {FontFamily → "Helvetica", FontSize → 11, FontWeight → "Bold"},
  Axes → False, AspectRatio → 0.7, DisplayFunction → Identity,
  PlotLabel → "Independent Sites Model"];
Show[g1, ga]
Print[" Fitted parameters: affinity  $K_1 =$ ",
  bestparameters[[1, 1]], "  $\times 10^6$  / mol, Chi-squared ", bestparameters[[2]]];
Print[" Errors in parameters: affinity  $K_1 =$ ", Sqrt[bestparameters[[3, 1, 1]]],
  "  $\times 10^6$  (nonsensical result may occur in case of very poor fit)"];
```



Fitted parameters: affinity  $K_1 = 23.4874 \times 10^6$  / mol, Chi-squared 0.646181

Errors in parameters: affinity  $K_1 = 12.2922$

$\times 10^6$  (nonsensical result may occur in case of very poor fit)

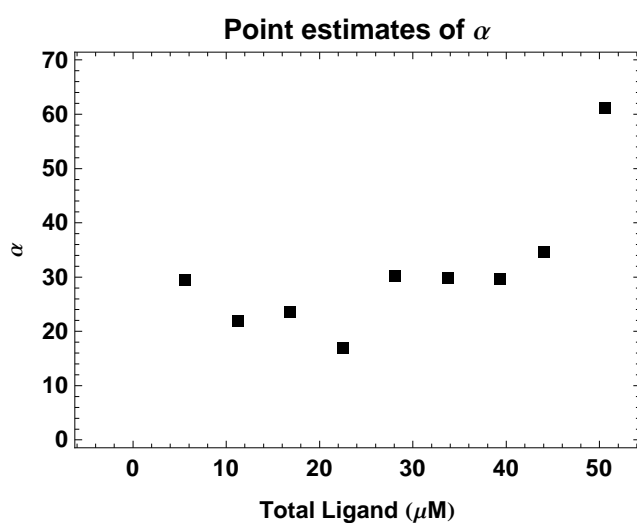
```
Export["independentsites.eps", Show[g1, ga]]
```

```
independentsites.eps
```

## Point estimates of $\alpha$ the coefficient of cooperativity

The cooperativity  $\alpha = 4 \cdot T_0 \cdot T_2 / (T_1 \cdot T_1)$  is not dependent on knowledge of protein or ligand concentrations. Provided both  $\alpha$  and affinities are not too large (i.e.  $T_1$  not too small),  $\alpha$  can be reasonably well estimated at each titration point from the above equation as suggested by Rogniaux et al (2001) Anal Biochem 291, 48-61. In cases where fitting to the cooperative model is possible, this estimate serves as a check for reasonableness of fitting result.

```
DataPlot[ligand, (4 * apo * double) / (single * single), {{-5, (Max[ligand] * 1.05)}, {0, 70}},
  "Point estimates of  $\alpha$ ", {"Total Ligand ( $\mu$ M)", " $\alpha$ "}]
Print[" Mean value of  $\alpha$  = ", Mean[(4 * apo * double) / (single * single)]];
```

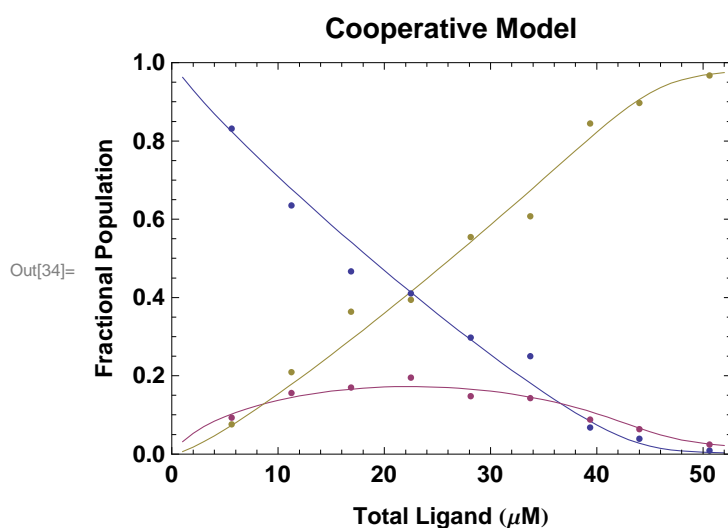


Mean value of  $\alpha$  = 30.8573

## Cooperative Model

Direct fitting of the all data to the cooperative binding equations should produce the most reliable results but requires at least 4 and preferably 8 or more titration points to obtain reliable results for both  $\alpha$  and affinities.

```
In[30]:= bestparameters =
  DimerBindingFit[apo, error, single, error, double, error, ligand, dimer, 0.0];
model = Transpose[Table[FractionalOccupancies[k, Mean[dimer], bestparameters[[1, 1]],
  bestparameters[[1, 2]]], {k, 1, Floor[Max[ligand] + 2], 1}]];
ligandmodel = Table[k, {k, 1, Floor[Max[ligand] + 2], 1}]; apomodel = model[[1]];
singlemodel = model[[2]]; doublemodel = model[[3]];
g2 = ListLinePlot[{Transpose[{ligandmodel, apomodel}],
  Transpose[{ligandmodel, singlemodel}], Transpose[{ligandmodel, doublemodel}]},
  InterpolationOrder → 1, PlotRange → {{0, (Max[ligand] + 2)}, {0, 1}},
  Frame → True, FrameLabel → {"Total Ligand ( $\mu$ M)", "Fractional Population"},
  BaseStyle → {FontFamily → "Helvetica", FontSize → 11, FontWeight → "Bold"},
  Axes → False, AspectRatio → 0.7, DisplayFunction → Identity,
  PlotLabel → "Cooperative Model"];
Show[g2, ga]
Print[" Fitted parameter: binding affinity  $K_1 =$ ",
  bestparameters[[1, 1]], "  $\times 10^6$  / mol, Chi-squared ", bestparameters[[2]]];
Print[" Error in parameter: affinity  $K_1 =$ ",
  Sqrt[bestparameters[[3, 1, 1]]], "  $\times 10^6$  "];
Print[" Fitted parameter: cooperativity  $\alpha =$ ", bestparameters[[1, 2]]];
Print[" Error in parameter: cooperativity  $\alpha =$ ",
  Sqrt[bestparameters[[3, 2, 2]]], "  $\times 10^6$  "];
Print[" Difference in the intrinsic free energy
  between the two binding events (2nd - 1st) = ",
  -0.00831441 * temperature * Log[bestparameters[[1, 2]]], " kJ/mol"];
```



Fitted parameter: binding affinity  $K_1 = 0.447208 \times 10^6$  / mol, Chi-squared 0.0265762

Error in parameter: affinity  $K_1 = 0.0259057 \times 10^6$

Fitted parameter: cooperativity  $\alpha = 23.0571$

Error in parameter: cooperativity  $\alpha = 0.694165 \times 10^6$

Difference in the intrinsic free energy between the two binding events (2nd - 1st) =  
-7.77885 kJ/mol

```
Export["cooperative.eps", Show[ga, g2]]
```

```
cooperative.eps
```