Analytical expressions for the homotropic binding of ligand to protein dimers and trimers

Supplementary Material

Scott T. Lefurgy[1, 2] and Thomas S. Leyh[1, §]

[1]Department of Microbiology & Immunology, Albert Einstein College of Medicine, Bronx, NY 10461

[2]Current Address: Department of Chemistry, Hofstra University, Hempstead, NY 11549

[§] **Corresponding Author:**

Thomas S. Leyh, Ph.D.

Albert Einstein College of Medicine

1300 Morris Park Ave.

Bronx, NY 10461

Phone: 718-430-2857

FAX: 718-430-8711

Email: tom.leyh@einstein.yu.edu

Divalent Interactions

For a dimeric protein binding to ligand [1], the concentrations of species in solution are governed by mass conservation relationships (Eqns. S1 and S2).

$$E_{\mathrm{T}} = E + \Sigma EL_1 + EL_2 \qquad \text{(S1)}$$

$$L_{\mathrm{T}} = L + \Sigma EL_1 + 2EL_2 \qquad \text{(S2)}$$

Each binding step can be described by a site dissociation constant that describes the affinity of ligand for a particular protomer. If the sites are identical, as in a homodimer, the site constants for either protomer at a given step (k_i) will be equivalent. In the case of a heterodimer, each protomer has a different affinity for the ligand, and for the first binding step this difference is accounted for by the coefficient c_0 multiplied by k_1; for the second step, the difference in the second dissociation constant k_2 is given by c_1. The site (k_i) and aggregate (K_i) dissociation constants (Figure 1, main text) are related by statistical factors that take these coefficients into account, either for homomeric proteins with identical subunits ($c_j = 1$) (cf. Eqn. 9, *Theory*) or non-identical ($c_j \neq 1$) (Eqns. S3 and S4).

$$K_1 = \frac{1}{\alpha} k_{1\,(site)} \qquad \text{(S3)}$$

$$K_2 = \alpha k_{2\,(site)} \qquad \text{(S4)}$$

where

$$\alpha = 1 + \frac{1}{c_0}$$

By substituting expressions of the dissociation constants (Eqns. 6, 7 and 8, *Theory* and Eqns. S3 and S4) into the mass conservation equations (S1 and S2), we obtain a cubic polynomial in L (Eqn. S5).

$$L^3 + pL^2 + qL + r = 0 \tag{S5}$$

where

$$p = 2E_{\mathrm{T}} + \alpha k_2 - L_T$$

$$q = k_1 k_2 + \alpha k_2 E_{\mathrm{T}} - \alpha k_2 L_{\mathrm{T}}$$

$$r = -k_1 k_2 L_{\mathrm{T}}$$

The general solution to this cubic is described below in *Trimeric Interactions*.

The value of L is then used to solve for the equilibrium concentrations of all other species (Eqns. S6-S8).

$$E = \frac{k_1 k_2 E_{\mathrm{T}}}{k_1 k_2 + \alpha k_2 L + L^2} \tag{S6}$$

$$\Sigma EL_1 = \frac{\alpha k_2 E_{\mathrm{T}} L}{k_1 k_2 + \alpha k_2 L + L^2} \tag{S7}$$

$$EL_2 = \frac{E_{\mathrm{T}} L^2}{k_1 k_2 + \alpha k_2 L + L^2} \tag{S8}$$

Trivalent Interactions

Fourth-degree polynomials are of the highest order that can be solved analytically [2]. Many methods have been put forth to solve the quartic [3]. Descartes' 1637 method, which factors the quartic into two quadratics is used here [4-6].

The quartic is first reduced through the substitution $L = y - a/4$ to remove the cubic term (Eqn. S9).

$$y^4 + ey^2 + fy + g = 0 \tag{S9}$$

where

$$e = b - \frac{3a^2}{8}$$

$$f = c + \frac{a^3}{8} - \frac{ab}{2}$$

$$g = d - \frac{3a^4}{256} + \frac{a^2 b}{16} - \frac{ac}{4}$$

The reduced quartic is then factored into two quadratics (Eqn. S10)

$$y^4 + ey^2 + fy + g = (y^2 + hy + j)\left(y^2 - hy + \frac{g}{j}\right) \tag{S10}$$

by making the substitutions given by Eqns. S11-S13.

$$e = \frac{g}{j} + j - h^2 \tag{S11}$$

$$f = h\left(\frac{g}{j} - j\right) \tag{S12}$$

$$j = \frac{e + h^2 - \frac{f}{h}}{2} \tag{S13}$$

To obtain h, Eqns. 19-21 are rearranged to obtain a cubic polynomial in h^2 (Eqn. S14).

$$h^6 + 2eh^4 + (e^2 - 4g)h^2 - f^2 = 0 \tag{S14}$$

Eqn. S14 has a general form whose solution is known (Eqn. S15) [7].

$$y^3 + py^2 + qy + r = 0 \tag{S15}$$

where

$$y = h^2$$

$$p = 2e$$

$$q = e^2 - 4g$$

$$r = f^2$$

By making the substitution $y = z - p/3$, the squared term is removed (Eqn. S16).

$$z^3 - sz - t = 0 \tag{S16}$$

where

$$s = \frac{p^2}{3} - q$$

$$t = -\frac{2p^3}{27} + \frac{pq}{3} - r$$

Only one of the three roots of Eqn. S16 corresponds to the physically relevant quantity, L. Due to the constraint that E_T, L_T, k_1, k_2 and k_3 must be positive, the unique root can be selected from among the possible solutions by examining the sign of the discriminant (Δ) (Eqn. S17), allowing automation of the solution in a fitting algorithm (see Supplementary Excel worksheet).

$$\Delta = \frac{t^2}{4} - \frac{s^3}{27} \tag{S17}$$

When $\Delta > 0$, there is one real solution (Eqn. S18) [1].

$$z_1 = \sqrt[3]{\frac{t}{2} + \sqrt{\frac{t^2}{4} - \frac{s^3}{27}}} + \sqrt[3]{\frac{t}{2} - \sqrt{\frac{t^2}{4} - \frac{s^3}{27}}} \tag{S18}$$

When $\Delta < 0$, there are three real solutions (Eqns. S19-S21) [8].

$$z_1 = \frac{2}{3}\sqrt{(p^2 - 3q)}\cos\frac{\theta}{3} \tag{S19}$$

$$z_2 = \frac{2}{3}\sqrt{(p^2 - 3q)}\cos\frac{2\pi - \theta}{3} \tag{S20}$$

$$z_3 = \frac{2}{3}\sqrt{(p^2 - 3q)}\cos\frac{2\pi + \theta}{3} \tag{S21}$$

where

$$\theta = \cos^{-1}\frac{-2p^3 + 9pq - 27r}{2\sqrt{(p^2 - 3q)^3}}$$

The physically relevant root corresponds to z_1 [8].

When $\Delta = 0$, two of the roots are identical, and either of the above equations yields the correct root. When one of the roots is zero, $t = 0$ and Eqn. S16 can be reduced in degree; the identical (positive) roots are given by $z_1 = z_2 = \sqrt{s}$. This is the relevant root, with the exception of the trivial case $L = 0$ when $L_T = 0$.

Once z has been determined using one of the methods above, h^2 is obtained using the substitution:

$$h^2 = z - \frac{p}{3} \tag{S22}$$

To simplify the arithmetic going forward, any positive, real root h^2 is chosen for calculations in Excel; taking the square root, h is obtained, and j is calculated from e, f and h (Eqn. S13).

At this point, the two quadratics in y (Eqn. S10) can be written in L, a, g, h and j (using the substitution $L = y - a/4$) (Eqns. S23-S24) and solved for roots L_1 thru L_4 (Eqns. S25-S28).

$$L^2 + \left(\frac{a}{2} + h\right)L + \left(\frac{a^2}{16} + \frac{ah}{4} + j\right) = 0 \tag{S23}$$

$$L^2 + \left(\frac{a}{2} - h\right)L + \left(\frac{a^2}{16} - \frac{ah}{4} + \frac{g}{j}\right) = 0 \tag{S24}$$

$$L_1 = \frac{-\left(\frac{a}{2} + h\right) + \sqrt{\left(\frac{a}{2} + h\right)^2 - 4\left(\frac{a^2}{16} + \frac{ah}{4} + j\right)}}{2} \tag{S25}$$

$$L_2 = \frac{-\left(\frac{a}{2} + h\right) - \sqrt{\left(\frac{a}{2} + h\right)^2 - 4\left(\frac{a^2}{16} + \frac{ah}{4} + j\right)}}{2} \tag{S26}$$

$$L_3 = \frac{-\left(\frac{a}{2} - h\right) + \sqrt{\left(\frac{a}{2} - h\right)^2 - 4\left(\frac{a^2}{16} - \frac{ah}{4} + \frac{g}{j}\right)}}{2} \tag{S27}$$

$$L_4 = \frac{-\left(\frac{a}{2} - h\right) - \sqrt{\left(\frac{a}{2} - h\right)^2 - 4\left(\frac{a^2}{16} - \frac{ah}{4} + \frac{g}{j}\right)}}{2} \tag{S28}$$

Only one of the four roots corresponds to the equilibrium concentration, L. To determine which root is relevant, it is necessary to inspect the discriminant for both of the quadratic equations, given by Δ_1 for roots L_1 and L_2 and Δ_2 for roots L_3 and L_4 (Eqns. S29-S30).

$$\Delta_1 = h^2 - 4j \tag{S29}$$

$$\Delta_2 = h^2 - \frac{4g}{j} \tag{S30}$$

When both Δ_1 and Δ_2 are positive, there are four real roots, and L_3 is the relevant root. When either Δ_1 or Δ_2 is positive, there are two real roots of opposite sign and the positive root is the relevant root.

References

[1] E.T. Mack, R. Perez-Castillejos, Z. Suo, and G.M. Whitesides, Exact analysis of ligand-induced dimerization of monomeric receptors. Anal Chem 80 (2008) 5550-5.

[2] N.H. Abel, Beweis der Unmöglichkeit, algebraische Gleichungen von höheren Graden als dem vierten allgemein aufzulösen. J reine angew Math 1 (1826) 65-84.

[3] R.S. Ball, Notes on the algebraical solution of biquadratic equations. Quarterly Journal of Pure and Applied Mathematics 7 (1866) 6-9, 358-369.

[4] R. Descartes, La Geometrie, 1637.

[5] K. Manders, Algebra in Roth, Faulhaber, and Descartes. Historia Mathematica 33 (2006) 184-209.

[6] The Math Forum @ Drexel, Solving the Quartic. 2011. Retrieved July 19, 2011 from http://mathforum.org/library/drmath/view/52874.html.

[7] W.D. Lambert, A generalized Trigonometric Solution of the Cubic Equation. The American Mathematical Monthly 13 (1906) 73-76.

[8] Z.X. Wang, An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. FEBS Lett 360 (1995) 111-4.