SUPPLEMENTARY INFROMATION

Experimental and mathematical evidence that thrombin-binding aptamers form a 1 aptamer:2 protein complex

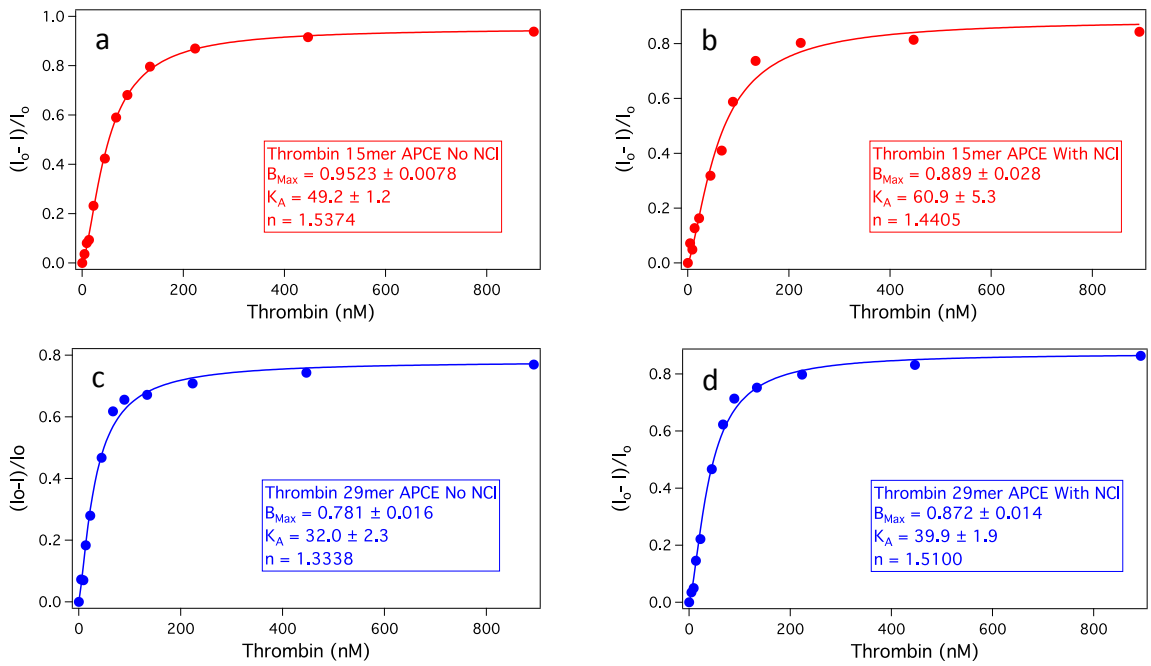
Kepler S Mears[1], Daniel L Markus[1], Oluwadamilare Ogunjimi[1] and Rebecca J Whelan[1,2]

[1]Department of Chemistry and Biochemistry, Oberlin College, USA

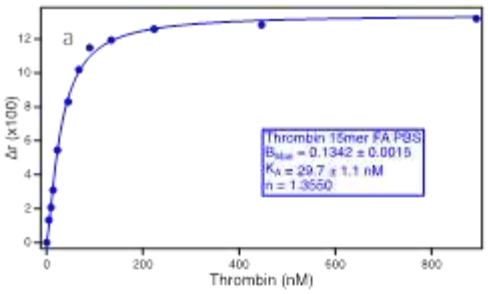
[2]Department of Chemistry and Biochemistry, University of Notre Dame, USA

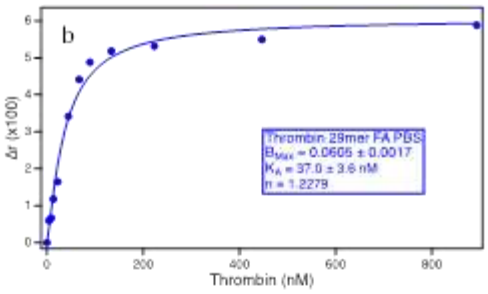
Correspondence to: Rebecca Whelan, Email: rwhelan1@nd.edu, Tel: +574-631-1853, Fax: 574-631-6652

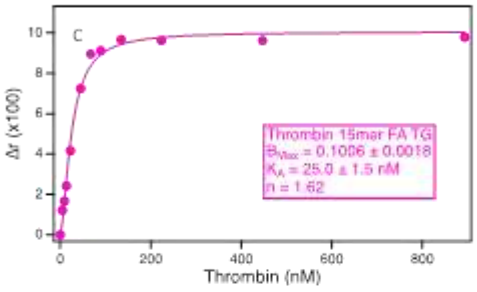
Received: 07 September 2018 | **Revised:** 04 October 2018 | **Accepted:** 07 October 2018 | **Published:** 10 October 2018

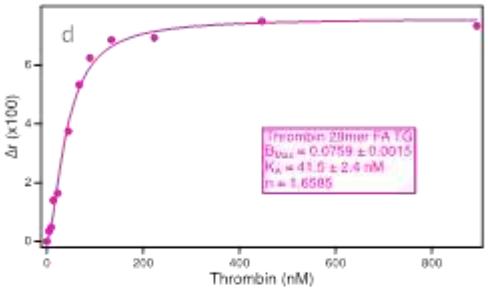


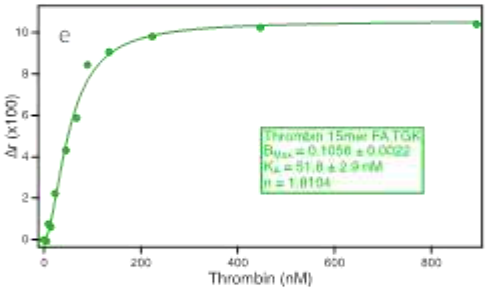
Supplementary Figure 1: Isotherms generated from APCE assays conducted without (left; a and c) and with (right; b and d) the non-cooled capillary inlet. The isotherms for the 15mer are on the top (a and b); the isotherms for the 29mer are on the bottom (c and d). All data were fit with the Hill equation as described in the text.

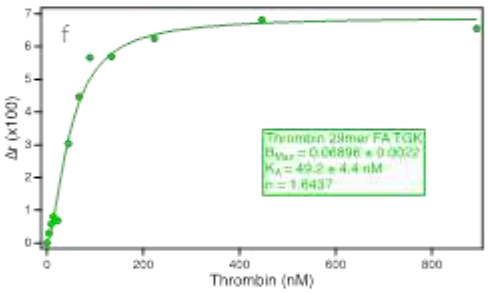

a
Thrombin 15mer FA PBS
B_{Max} = 0.1342 ± 0.0015
K_A = 29.7 ± 1.1 nM
n = 1.3550
Δr (x100)
Thrombin (nM)

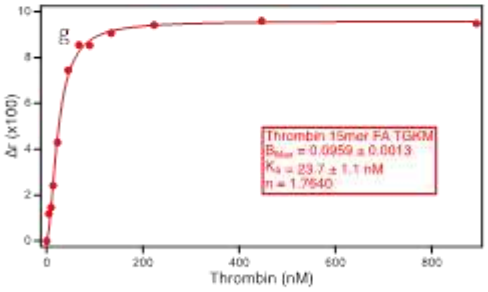

b
Thrombin 29mer FA PBS
B_{Max} = 0.0605 ± 0.0017
K_A = 37.0 ± 3.6 nM
n = 1.2279
Δr (x100)
Thrombin (nM)

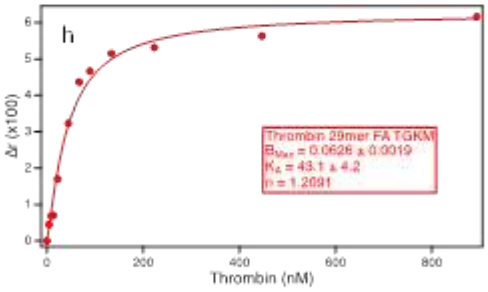

c
Thrombin 15mer FA TG
B_{Max} = 0.1006 ± 0.0018
K_A = 25.0 ± 1.5 nM
n = 1.62
Δr (x100)
Thrombin (nM)

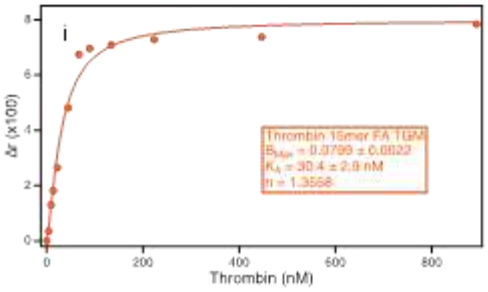

d
Thrombin 29mer FA TG
B_{Max} = 0.0759 ± 0.0015
K_A = 41.5 ± 2.4 nM
n = 1.6585
Δr (x100)
Thrombin (nM)

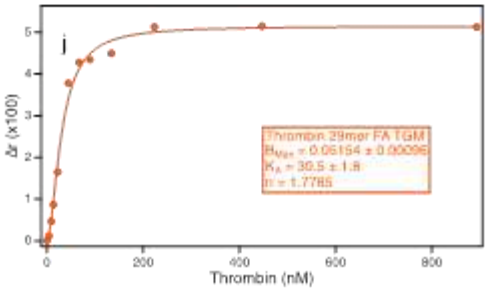

e
Thrombin 15mer FA TGK
B_{Max} = 0.1056 ± 0.0022
K_A = 51.8 ± 2.9 nM
n = 1.8104
Δr (x100)
Thrombin (nM)


f
Thrombin 29mer FA TGK
B_{Max} = 0.06896 ± 0.0022
K_A = 49.2 ± 4.4 nM
n = 1.6437
Δr (x100)
Thrombin (nM)


g
Thrombin 15mer FA TGKM
B_{Max} = 0.0959 ± 0.0013
K_A = 23.7 ± 1.1 nM
n = 1.7640
Δr (x100)
Thrombin (nM)


h
Thrombin 29mer FA TGKM
B_{Max} = 0.0626 ± 0.0019
K_A = 43.1 ± 4.2
n = 1.2091
Δr (x100)
Thrombin (nM)


i
Thrombin 15mer FA TGM
B_{Max} = 0.0799 ± 0.0022
K_A = 30.4 ± 2.9 nM
n = 1.3556
Δr (x100)
Thrombin (nM)


j
Thrombin 29mer FA TGM
B_{Max} = 0.05154 ± 0.00096
K_A = 30.5 ± 1.8
n = 1.7785
Δr (x100)
Thrombin (nM)

Supplementary Figure 2: Isotherms generated from FA assays in five different buffers: PBS (a and b), TG (c and d), TGK (e and f), TGKM (g and h), and TGM (i and j). The isotherms for the 15mer are on the left (a, c, e, g, i); the isotherms for the 29mer are on the right (b, d, f, h, j). All data were fit with the Hill equation as described in the text.

DERIVATION OF BINDING MODELS

Simple Hyperbola

To derive the simple square hyperbola isotherm, we start with a 1:1 equilibrium of aptamer, A, and protein, P, forming a complex, AP and the associated K_d expression:

$$AP \Leftrightarrow A + P \tag{S-1}$$

$$K_d = \frac{[A][P]}{[AP]} \tag{S-2}$$

We define the fraction bound, f_b, as the fraction of aptamer bound to protein over total aptamer, $[A]_t$ were $[A]_t = [A] + [AP]$:

$$f_b = \frac{[AP]}{[A]_t} = \frac{[AP]}{[A] + [AP]} \tag{S-3}$$

Next, equation 2 is solved for $[AP]$, substituted into equation 3 and simplified:

$$f_b = \frac{\frac{[A][P]}{K_d}}{[A] + \frac{[A][P]}{K_d}}$$

$$f_b = \frac{\frac{[P]}{K_d}}{1 + \frac{[P]}{K_d}} * \left(\frac{K_d}{K_d}\right)$$

$$f_b = \frac{[P]}{K_d + [P]} \tag{S-4}$$

Which is our desired result. In the case where $[A]_t \ll [P]_t$ where $[P]_t$ is the total protein added, after equilibrium is reached $[P]_t \approx [P]$ so we can substitute into equation 4:

$$f_b = \frac{[P]_t}{K_d + [P]_t} \tag{S-5}$$

Which is the useful version of equation 4. This result is important because it allows us to assume the protein added to samples is the protein concentration after equilibrium. Alternatively, equation 5 is often rewritten as:

$$f_b = \frac{1}{1 + \frac{K_d}{[P]_t}} \tag{S-6}$$

Expanded Hyperbola

In the case of equation 5 where we cannot assume $[A]_t \ll [P]_t$ we must use the expanded isotherm, obtained by substituting $[P] = [P]_t - [AP]$ for $[P]$ in equation 4 and solving for $[AP]$ using the quadratic formula:

$$\frac{[AP]}{[A]_t} = \frac{[P]_t - [AP]}{K_d + [P]_t - [AP]} \tag{S-7}$$

$$[AP](\, K_d + [P]_t - [AP]) = \, [A]_t([P]_t - [AP])$$

$$[AP]K_d + [AP][P]_t - [AP]^2 = \, [A]_t[P]_t - [A]_t[AP]$$

$$[AP]K_d + [AP][P]_t - [AP]^2 - [A]_t[P]_t + [A]_t[AP] = 0$$

$$[AP]^2 - ([A]_t + [P]_t + K_d)[AP] + [A]_t[P]_t = 0$$

$$[AP] = \frac{([A]_t + [P]_t + K_d) \pm \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2} \tag{S-8}$$

We now plug equation 8 back into equation 7 and simplify:

$$f_b = \frac{[P]_t - \frac{([A]_t + [P]_t + K_d) \pm \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2}}{K_d + [P]_t - \frac{([A]_t + [P]_t + K_d) \pm \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2}}$$

$$f_b = \left(\frac{[P]_t - \frac{([A]_t + [P]_t + K_d) \pm \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2}}{K_d + [P]_t - \frac{([A]_t + [P]_t + K_d) \pm \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2}} \right) * \left(\frac{\frac{1}{[P]_t - \frac{([A]_t + [P]_t + K_d) \pm \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2}}}{\frac{1}{[P]_t - \frac{([A]_t + [P]_t + K_d) \pm \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2}}} \right)$$

$$f_b = \frac{1}{1 + \left(\frac{K_d}{[P]_t - \frac{([A]_t + [P]_t + K_d) \pm \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2}} \right)} \quad \text{(S-9)}$$

To determine the case of "$\pm$" we take the limit as $[P]_t$ approaches infinity. As $[P]_t$ approaches infinity eventually we should be able to make the assumption that $[A]_t << [P]_t$ and equation 8 should reduce to the simple isotherm i.e. equation 6. As $[P]_t$ approaches infinity equation 8 becomes:

$$f_b = \frac{1}{1 + \left(\frac{K_d}{[P]_t - \frac{[P]_t \pm [P]_t}{2}} \right)}$$

Which is indeterminate in the case where "$\pm$" is "+." On the other hand when "$\pm$" is "-':

$$f_b = \frac{1}{1 + \left(\frac{K_d}{[P]_t - \frac{[P]_t - [P]_t}{2}} \right)}$$

$$f_b = \frac{1}{1 + \left(\frac{K_d}{[P]_t} \right)} \quad \text{(S-10)}$$

Which is the same as equation 6. Thus the final form of the expanded hyperbola is:

$$f_b = \frac{1}{1 + \left(\frac{K_d}{[P]_t - \frac{([A]_t + [P]_t + K_d) - \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2}} \right)} \quad \text{(S-11)}$$

This form is of the expanded hyperbola is similar to the version of the simple hyperbola in equation 6 and clearly shows a correction for the protein concentration. However there is a simpler version of equation 11. Alternatively, once we have reached equation 7, given that $f_b = [AP]/[A]_t$ we can divide by $[A]_t$ to get:

$$\frac{[AP]}{[A]_t} = \frac{([A]_t + [P]_t + K_d) \pm \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2[A]_t}$$

And from there take the limit as $[P]_t$ goes to infinity to, knowing that as $[P]_t$ approaches infinity f_b must approach 1 to get the result:

$$f_b = \frac{([A]_t + [P]_t + K_d) - \sqrt{([A]_t + [P]_t + K_d)^2 - 4[A]_t[P]_t}}{2[A]_t} \quad \text{(S-12)}$$

Which is an equivalent form of the expanded hyperbola.

Hill Equation

The Hill equation models the association of "n" protein molecules to 1 aptamer molecule depicted in equation 13. It is important to distinguish that in equation 13 that K_{d^n} is the dissociation for higher order interactions, it has units of M^n, different than the 1:1 K_d in equation 2. To obtain a function a similar analysis as the simple hyperbola is then taken.

$$A + nP \overset{K_{d^n}}{\Longleftrightarrow} AP_n \quad \text{(S-13)}$$

$$K_{d^n} = \frac{[A][P]^n}{[AP_n]} \quad \text{(S-14)}$$

$$f_b = \frac{[AP_n]}{[A]_t} = \frac{[AP_n]}{[A] + [AP_n]} \quad \text{(S-15)}$$

Substituting equation 14 into 15:

$$f_b = \frac{[AP_n]}{[A]_t} = \frac{\frac{[A][P]^n}{K_{d^n}}}{[A] + \frac{[A][P]^n}{K_{d^n}}}$$

$$f_b = \frac{\frac{[P]^n}{K_d}}{1 + \frac{[P]^n}{K_d}} * \frac{K_{d^n}}{K_{d^n}}$$

$$f_b = \frac{[P]^n}{K_{dn} + [P]^n} \quad \text{(S-16)}$$

Units of M^n returned by the K_{d^n} are impractical for analysis. In order to simplify the interpretation of the Hill equation the K_A is substituted for the K_d:

$$K_{d^n} = (K_A)^n \quad \text{(S-17)}$$

$$f_b = \frac{[P]^n}{{K_A}^n + [P]^n} \quad \text{(S-16)}$$

The K_A always has units in M and, like the K_d for 1:1 interactions in the simple square hyperbola, is the concentration of protein at half maximum binding. K_A is a more practical measure of affinity. Theoretically, "n" refers to the number of binding sites on a target but practically is interpreted as cooperativity that reflects how multiple binding sites interact. An n>1 reflects positive cooperativity and n<1 reflects negative cooperativity. Similarly with the simple square hyperbola we assume that $[P]_t \approx [P]$ with the final form of the Hill equation being:

$$f_b = \frac{{[P]_t}^n}{{K_A}^n + {[P]_t}^n} \quad \text{(S-17)}$$

Sequential complex formation

Equation 13 implies a simultaneous multibody interaction which is rare for higher order complexes (when n>1). As such, complex rarely forms through equation 13 and the Hill equation is an approximation of multibody binding. A more rigorous and realistic examination of complex formation would be a stepwise formation of individual 1:1 interactions such as the following:

$$A + P \overset{K_{d1}}{\Longleftrightarrow} AP + P \overset{K_{d2}}{\Longleftrightarrow} AP_2 + P \overset{K_{d3}}{\Longleftrightarrow} \ldots \overset{K_{dn-1}}{\Longleftrightarrow} AP_{n-1} + P \overset{K_{dn}}{\Longleftrightarrow} AP_n \quad \text{(S-18)}$$

For the case of n=2 equation 18 can written explicitly as equation 19 and easily modeled by the same approach as with the Hill equation and square hyperbola. Again we use the fundamental relations established by equations 20, 21 and 22:

$$A + P \overset{K_{d1}}{\Longleftrightarrow} AP + P \overset{K_{d2}}{\Longleftrightarrow} AP_2 \quad \text{(S-19)}$$

$$K_{d1} = \frac{[A][P]}{[AP]} \quad \text{(S-20)}$$

$$K_{d2} = \frac{[AP][P]}{[AP_2]} \quad \text{(S-21)}$$

$$f_b = \frac{[AP] + [AP_2]}{[A]_t} = \frac{[AP] + [AP_2]}{[A] + [AP] + [AP_2]} \quad \text{(S-22)}$$

To achieve a function we substitute equations 20 and 21 into equation 22 and simplify:

$$f_b = \frac{\frac{[A][P]}{K_{d1}} + \frac{[AP][P]}{K_{d2}}}{[A] + \frac{[A][P]}{K_{d1}} + \frac{[AP][P]}{K_{d2}}}$$

$$f_b = \frac{\frac{[A][P]}{K_{d1}} + \frac{[A][P]^2}{K_{d1}K_{d2}}}{[A] + \frac{[A][P]}{K_{d1}} + \frac{[A][P]^2}{K_{d1}K_{d2}}}$$

$$f_b = \frac{\frac{[P]}{K_{d1}} + \frac{[P]^2}{K_{d1}K_{d2}}}{1 + \frac{[P]}{K_{d1}} + \frac{[P]^2}{K_{d1}K_{d2}}} * \left(\frac{K_{d1}K_{d2}}{K_{d1}K_{d2}}\right)$$

$$f_b = \frac{K_{d2}[P] + [P]^2}{K_{d1}K_{d2} + K_{d2}[P] + [P]^2} \quad \text{(S-23)}$$

Equation 23 is the functional isotherm that represents the binding in equation 19. A similar process can be repeated for any number of "n". In doing so a pattern becomes evident that can allow a generalized isotherm for one aptamer binding "n" proteins but this will not be covered here. Equation 22 also operates under the assumption that $[P]_t \approx [P]$ to be practical:

$$f_b = \frac{K_{d2}[P]_t + {[P]_t}^2}{K_{d1}K_{d2} + K_{d2}[P]_t + {[P]_t}^2} \quad \text{(S-24)}$$

It is possible to expand equation 24 like that of the expanded hyperbola, but it cumbersome and yields a non-practical result. Furthermore, as there only exists closed form expressions to solve for up to 4th order polynomials, the expandability of isotherms only works when there are 4 or fewer binding proteins involved in binding.

Independent binding sites

It is possible that instead of complex forming sequentially that there exist multiple binding sites independent of one another. In this case each site has an individual associated binding isotherm. For the case of two protein ligands this would look like:

$$f_b = \frac{\frac{1}{2}[P]_t}{K_{d1} + [P]_t} + \frac{\frac{1}{2}[P]_t}{K_{d2} + [P]_t} \quad \text{(S-25)}$$