Supplementary Information

Derivation of the modified Scatchard equation for cooperative linear lattice-protein interactions (Eq. 10b in the main text)

As compared to the combinatorial method, the task of the conditional probability approach is to calculate the average number of free protein binding sites (with the site size of n motifs) per lattice. Analogous to the convention applied in the combinatorial method, a protein can bind to three types of free binding sites on a linear lattice, isolated (iso), singly contiguous (sc), and doubly contiguous (dc) sites. Then, each binding is expressed as the Scatchard form with the intrinsic binding constant and cooperativity:

1) \[ \bar{v}_{\text{iso}} = K \cdot \{ \text{Average number of isolated binding sites per lattice} \} \]
2) \[ \bar{v}_{\text{sc}} = K \cdot \{ \text{Average number of singly contiguous binding sites per lattice} \} \]
3) \[ \bar{v}_{\text{dc}} = K \cdot \{ \text{Average number of doubly contiguous binding sites per lattice} \} \]

where \( \bar{v}_{\text{iso}}, \bar{v}_{\text{sc}}, \) and \( \bar{v}_{\text{dc}} \) are the average numbers of bound proteins per lattice in isolated, singly contiguous, and doubly contiguous manners, respectively. Then, the average number of bound proteins per lattice \( \bar{v} \) is the sum of these three terms.

The number of total free binding sites in a \( g \)-motif long gap is \( g - n + 1 \). If the gap length equals the binding site size (\( n \)), then the free binding site is doubly contiguous. If the gap length is longer than the binding site size, there are two singly contiguous binding sites and \( g - n - 1 \) isolated binding sites. Therefore, the average number of each of the three types of free binding sites per gap is given by:

\[ \bar{s}_{\text{iso}} = \sum_{g=n+1}^{M} (g - n - 1) P_g \]  
\[ \bar{s}_{\text{sc}} = 2 \sum_{g=n+1}^{M} P_g \]  
\[ \bar{s}_{\text{dc}} = P_n \]

where \( P_g \) is the probability that any gap is \( g \) free residues long.

In order to calculate the probability \( P_g \), the motifs on a linear lattice are represented by \( f \) for a free motif and \( b_i \) for a bound motif where \( i \) ranges from 1 to \( n \). \( b_1 \) represents the motif being bound by the left end of a protein and \( b_n \) represent the motif being bound by the right end of a protein (Fig. 1A). Then, the following conditional probabilities are defined for a pair of consecutive motifs (Fig. 1A):

1) \( ff \): given a randomly chosen free motif, the probability of the subsequent righthand side motif being free
2) \( fb_i \): given a randomly chosen free motif, the probability of the subsequent righthand side motif being bound by the left end of a protein
3) \( b_{nf} \): given a motif bound by the right end of a protein, the probability of the subsequent motif being free

4) \( b_{nb} \): given a motif bound by the right end of a protein, the probability of the subsequent motif being bound by the left end of a protein

Then, \( P_b \) is given by

\[
P_b = b_n f \cdot ff \theta^{-1} \cdot fb_1
\]  

(A-2)

Since the sum of the probabilities for a given condition equals unity, the following relationship must be satisfied:

\[
ff + fb_1 = 1
\]  

(A-3a)

\[
b_n f + b_n b_1 = 1
\]  

(A-3b)

Moreover, the probability of choosing a free or bound motif on a lattice is simply \( 1 - n\theta \) or \( \theta \), respectively. Then, the following relationship must be satisfied as well:

\[
(1 - n\theta) \cdot ff + \theta \cdot b_n f = 1 - n\theta
\]  

(A-4a)

\[
(1 - n\theta) \cdot fb_1 + \theta \cdot b_n b_1 = \theta
\]  

(A-4b)

Next, the cooperativity term \( \omega \) is defined in terms of the conditional probabilities. This quantity corresponds to an equilibrium constant for moving a bound protein in the isolated manner to the adjacency of any other bound proteins. In this “transfer” process, there is no change in total numbers of free and bound motifs. However, there is a loss of a leftmost bound motif positioned at the right of a free motif but a gain of one positioned at the right of the rightmost bound motif. Simultaneously, there is a loss of one free motif positioned at the right of the rightmost bound motif but a gain of one at the right of the free motif. In short, there are conversions from \( fb_1 \) and \( b_{nf} \) to \( b_{nb} \) and \( ff \), respectively. Therefore, the cooperativity term \( \omega \) is expressed as:

\[
\omega = \frac{b_n b_1 \cdot ff}{ff b_1 \cdot b_n f}
\]  

(A-5)

Now that we have fundamental relationship among conditional probabilities, we provide a step-by-step derivation of the expressions for the conditional probabilities in terms of \( n \), \( \theta \), and \( \omega \).

First, using Eqs. A-3a and A-3b, Eq. A-5 is rearranged into the following form:

\[
\omega t b_n f = \frac{ff}{\omega \cdot \omega - \omega \cdot ff + ff}
\]  

(A-6)

Substituting this result for the \( b_{nf} \) term in Eq. A-4a yields a quadratic equation for \( ff \):

\[
(\omega - 1)(1 - n\theta)ff^2 + (-2\omega - 1)(1 - n\theta - \theta)ff + (1 - n\theta)\omega = 0
\]  

(A-7)
Then, $ff$ is given by:

$$ff = \frac{(2\omega - 1)(1 - n\theta) + \theta - \sqrt{(2\omega - 1)(1 - n\theta) + \theta}^2 - 4\omega(1 - n\theta)^2}{2(1 - n\theta)}$$  \hspace{1cm} (A-8)

The square root term can be further rearranged into the following form:

$$ff = \frac{(2\omega - 1)(1 - n\theta) + \theta - \sqrt{1 - (n + 1)\theta^2 + 4\omega\theta(1 - n\theta)}}{2(1 - n\theta)}$$  \hspace{1cm} (A-9)

For simplicity in expression, we denote $R$ as

$$R = \sqrt{1 - (n + 1)\theta^2 + 4\omega\theta(1 - n\theta)}$$  \hspace{1cm} (A-10)

Other conditional probabilities are readily derived:

$$fb_1 = \frac{(n - 1)\theta - 1 + R}{2(1 - n\theta)}$$  \hspace{1cm} (A-11a)

$$b_nf = \frac{(n - 1)\theta - 1 + R}{2\theta(1 - n\theta)}$$  \hspace{1cm} (A-11b)

$$b_nb_1 = \frac{1 - (n - 2\omega + 1)\theta - R}{2\theta(1 - n\theta)}$$  \hspace{1cm} (A-11c)

Substituting these conditional probability expressions for the corresponding terms in $P_g$ allows calculation of the average number of each of three types of free binding site per gap:

$$s_{iso} = \sum_{g=n+2}^{\infty} (g - n - 1) \cdot P_g \cdot \frac{(b_nf)(ff)^{n+1}}{(fb_1)}$$  \hspace{1cm} (A-12a)

$$s_{sc} = \sum_{g=n+4}^{\infty} 2 \cdot P_g = 2(b_nf)(ff)^n$$  \hspace{1cm} (A-12b)

$$s_{dc} = 1 \cdot P_n = (b_nf)(ff)^{n-1}(fb_1)$$  \hspace{1cm} (A-12c)

In these equations, the assumption of infinite lattice length ($M \to \infty$) was invoked and then the following general expressions for the summation of infinite geometric series and its derivative were used to obtain the final solutions:

$$\sum_{n=k}^{\infty} x^n = \frac{x^k}{1 - x} \hspace{1cm} (for \ - 1 < x < 1, k \geq 1)$$  \hspace{1cm} (A-13a)
\[ \sum_{n=k}^{\infty} nx^n = \frac{kx^k(1-x) + x^{k+1}}{(1-x)^2} \quad \text{for } 1 < x < 1, k \geq 1 \] \hspace{1cm} (A-13b)

Multiplying Eqs. A-12a to 12c by the total number of gaps per lattice \( n + 1 \sim n \) under the assumption of infinite lattice length, we obtain the average number of each of three types of free binding sites per lattice. Since \( \nu \) is the sum of \( \nu_{iso}, \nu_{sc} \) and \( \nu_{dc} \), the overall binding equilibrium is expressed as:

\[ \nu = \nu_{iso} + \nu_{sc} + \nu_{dc} = K[P](\nu \cdot \bar{s}_{iso}) + K\omega[P](\nu \cdot \bar{s}_{sc}) + K\omega^2[P](\nu \cdot \bar{s}_{dc}) \] \hspace{1cm} (A-14)

In the following procedure, this expression is rearranged and simplified to get the final Scatchard form of binding density:

\[ \nu = \nu_{iso} + \nu_{sc} + \nu_{dc} = K[P](\nu \cdot \bar{s}_{iso}) + K\omega[P](\nu \cdot \bar{s}_{sc}) + K\omega^2[P](\nu \cdot \bar{s}_{dc}) \]

\[ = K[P]v(\nu \cdot f)^{(n-1)} + \omega \cdot 2(b_n f)(f)^n + \omega^2 \cdot (b_n f)(f)^{n-1}(f b_i) \]

\[ = K[P]v(\nu \cdot f)^{(n-1)}(b_n f)(f b_i) \]

\[ = K[P]v(\nu \cdot f)^{(n-1)}(b_n f)(f b_i) \]

\[ \therefore \theta[K(P)] = \theta[K(1-n\theta)] \]

\[ \frac{1}{P} = \frac{2(\omega - 1)(1-n\theta) + \theta - \omega + (2\omega - 1)(1-n\theta) + \theta + \omega}{2(\omega - 1)(1-n\theta)} \]

\[ \text{Derivation of the modified Scatchard equation for noncooperative linear lattice-protein interactions (Eq. 11 in the main text)} \]

In the case of \( \omega = 1 \) (noncooperative binding), by using L'Hopital's rule, Eq. 10b can be reduced to the following:
\[
\lim_{\omega \to 1} R = 1 - (n - 1)\theta \quad \text{(A-16a)}
\]
\[
\lim_{\omega \to 1} \frac{(2\omega - 1)(1 - n\theta) + \theta - R}{2(\omega - 1)(1 - n\theta)} = 0 \quad \text{(A-16b)}
\]
\[
\lim_{\omega \to 1} 2(\omega - 1)(1 - n\theta) = 0 \quad \text{(A-16c)}
\]

By L'Hospital's rule, the first squared bracket term in Eq.10b is expressed as:
\[
\lim_{\omega \to 1} \frac{(2\omega - 1)(1 - n\theta) + \theta - R}{2(\omega - 1)(1 - n\theta)} = \frac{2(1 - n\theta)}{2(1 - n\theta)} = \frac{1 - n\theta}{1 - (n - 1)\theta} \quad \text{(A-17)}
\]

Moreover, the second squared bracket term in Eq.10b is expressed as:
\[
\lim_{\omega \to 1} \frac{1 - (n + 1)\theta + R}{2(1 - n\theta)} = 1 \quad \text{(A-18)}
\]

Thus, the modified Scatchard equation for noncooperative linear lattice-protein interactions is expressed as:
\[
\frac{\theta}{[P]} = K(1 - n\theta) \left[ \frac{1 - n\theta}{1 - (n - 1)\theta} \right]^{n-1} \quad \text{(A-19)}
\]

Derivation of the partial derivatives of the modified Scatchard equation with respect to \( \theta \) (Eq. 13 in the main text)
\[
\frac{\partial (\theta/[P])}{\partial \theta} \bigg|_{\theta=0} = \left( K(-n) \left[ \frac{(2\omega - 1)(1 - n\theta) + \theta - R}{2(\omega - 1)(1 - n\theta)} \right]^{n-1} \frac{1 - (n + 1)\theta + R}{2(1 - n\theta)} \right) + K(1 - n\theta)(n - 1) \left[ \frac{(2\omega - 1)(1 - n\theta) + \theta - R}{2(\omega - 1)(1 - n\theta)} \right]^{n-2} \frac{\partial}{\partial \theta} \left( \frac{(2\omega - 1)(1 - n\theta) + \theta - R}{2(\omega - 1)(1 - n\theta)} \right) \frac{1 - (n + 1)\theta + R}{2(1 - n\theta)}^2 
\]
\[
+ K(1 - n\theta) \left[ \frac{(2\omega - 1)(1 - n\theta) + \theta - R}{2(\omega - 1)(1 - n\theta)} \right]^{n-1} \cdot 2 \left[ \frac{1 - (n + 1)\theta + R}{2(1 - n\theta)} \right] \frac{\partial}{\partial \theta} \left( \frac{1 - (n + 1)\theta + R}{2(1 - n\theta)} \right) \bigg|_{\theta=0} \quad \text{(A-20)}
\]

The individual partial derivatives in the squared brackets are given by:
\[
\frac{\partial}{\partial \theta} \left( \frac{(2\omega - 1)(1 - n\theta) + \theta - R}{2(\omega - 1)(1 - n\theta)} \right) = \frac{2(\omega - 1)(1 - n\theta) (1 - n(2\omega - 1) - \frac{\partial R}{\partial \theta}) + [(2\omega - 1)(1 - n\theta) + \theta - R]2n(\omega - 1)}{(2(\omega - 1)(1 - n\theta))^2} \quad \text{(A-21a)}
\]
\[
\frac{\partial}{\partial \theta} \left( \frac{1 - (n + 1)\theta + R}{2(1 - n\theta)} \right) = \frac{2(1 - n\theta) \left[ \frac{\partial R}{\partial \theta} - (n + 1) \right] + 2n\{1 - (n + 1)\theta + R\}}{(2(1 - n\theta))^2} \quad \text{(A-21b)}
\]
\[
\frac{\partial R}{\partial \theta} = \frac{(n + 1)(n + 1)\theta - 1 + 2\omega(1 - 2n\theta)}{R} \quad (A-21c)
\]

The overall partial derivative is then significantly simplified to the following expression:

\[
\frac{\partial (\theta/[P])}{\partial \theta}_{\theta=0} = \frac{\partial (v/[P])}{\partial v}_{v=0} = K(2\omega - 2n - 1) \quad (A-22)
\]
Supplementary Fig. S1. Schematic illustration of the conditional probabilities for linear lattice-protein interactions.