

Research Notes - CaMKII Project - Version 2.0

Emma Glass, M. Saleet Jafri, Shahid Khan, Greg Conradi Smith

July 28, 2022

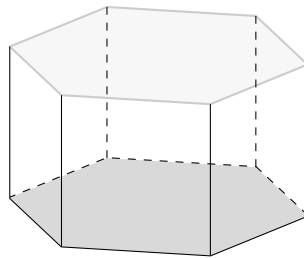
Abstract

A minimal model of CaMKII is constructed. We investigate how cooperativity depends on model parameters; in particular, pairwise interaction energies between CaM-bound (i.e., primed) and phosphorylated subunits.

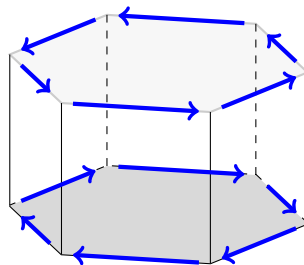
1 Subunit state-transition diagram and symmetry

1.1 CaMKII symmetries

We are assuming 12 CaMKII subunits arranged as a *hexagonal prism*.



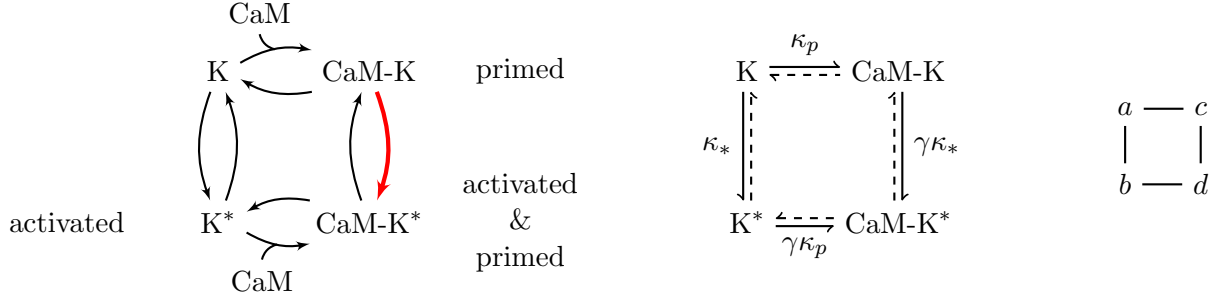
There are several appropriate choices of symmetry groups that could be appropriate, see <http://newton.ex.ac.uk/research/qsystems/people/goss/symmetry/Solids.html>



We are assuming that the symmetries of CaMKII are given by the abstract group D_6 , i.e., the dihedral group of order 12 (see Appendix). The hexagonal prism above has this symmetry group. Note that the orientations of subunits on top and bottom plane are reversed.

1.2 Subunit model

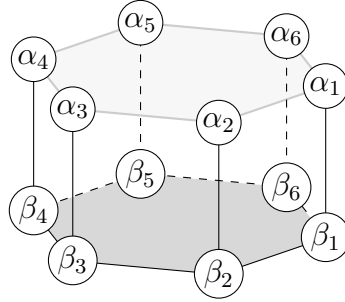
We are using a four state model of calmodulin trapping. The subunit transition state diagram is shown below.



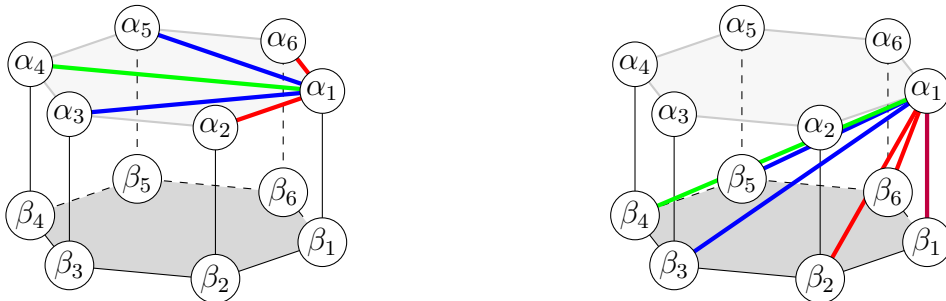
For fixed calmodulin concentration ($[CaM]$), there are three free equilibrium parameters: κ_* (activation) and κ_p (priming), as well as the allosteric parameter γ . The step from *primed* to *activated & primed* (thick red arrow) requires a neighboring kinase (K) to be primed with bound calmodulin (states $CaM-K$ or $CaM-K^*$).

1.3 Interaction between subunits

To understand how we model the free energy of a particular CaMKII configuration, it helps to label the subunits as follows.



There are 6 distinct inter-subunit distances in CaMKII. Three of these are within one plane of 6 subunits: **short** (e.g., $\alpha_1 \mapsto \alpha_2$ and $\alpha_1 \mapsto \alpha_6$), **medium** ($\alpha_1 \mapsto \alpha_3$ and $\alpha_1 \mapsto \alpha_5$), **long** ($\alpha_1 \mapsto \alpha_4$). Four more distinct inter-subunit distances involve one subunit from each plane: **transverse** ($\alpha_1 \mapsto \beta_1$), **transverse-short** ($\alpha_1 \mapsto \beta_2$ and $\alpha_1 \mapsto \beta_6$), **transverse-medium** ($\alpha_1 \mapsto \beta_3$ and $\alpha_1 \mapsto \beta_5$), **transverse-long** ($\alpha_1 \mapsto \beta_4$).



Let us represent a CaMKII state as a tuple

$$\begin{aligned} (\alpha_1, \dots, \alpha_6, \beta_1, \dots, \beta_6) &= (\text{K}, \text{CaM-K}, \text{CaM-K}^*, \text{K}^*, \text{K}^*, \text{CaM-K}, \text{K}, \text{K}^*, \text{CaM-K}, \text{CaM-K}^*, \text{K}^*, \text{CaM-K}^*) \\ &= (a, c, d, b, b, c, a, b, c, d, b, d) \end{aligned}$$

The total free energy of a configuration is the sum of the subunit energies given by

$$\begin{aligned} \kappa_* &= e^{-\beta(G_c - G_a)} \\ \gamma\kappa_p &= e^{-\beta(G_b - G_d)} \\ \gamma\kappa_* &= e^{-\beta(G_d - G_b)} \\ \kappa_p &= e^{-\beta(G_b - G_a)} \end{aligned}$$

Without loss of generality we can assume $G_a = 0$. Then the above equations define G_b , G_c and G_d so,

$$\begin{aligned} G_a &= 0 \\ G_b &= -\log\kappa_p/\beta \\ G_c &= -\log\kappa_*/\beta \\ G_d &= -\log\gamma\kappa_*\kappa_p/\beta \end{aligned}$$

Calculating Pairwise Interaction Energies

We can create a matrix that represents the 3D hexagonal prism geometry based on the distances that subunits 1-12 are from each other. The hexagonal prism has seven different interactions distances as described in section 1.3. 's' represents both the 'short' and 'transverse short' interactions, since we are assuming they are of the same length. The 'm', 'l', 'ts', 'tm', 'tl', symbols represent medium, long, transverse short, transverse medium, and transverse long interaction distances respectively.

$$\left\{ \begin{array}{cccccccccccc} 0 & s & ts & m & tm & l & tl & m & tm & s & ts \\ 0 & 0 & s & ts & m & tm & l & tl & m & tm & s \\ 0 & 0 & 0 & s & ts & m & tm & l & tl & m & tm \\ 0 & 0 & 0 & 0 & s & ts & m & tm & l & tl & m \\ 0 & 0 & 0 & 0 & 0 & s & ts & m & tm & l & tl \\ 0 & 0 & 0 & 0 & 0 & 0 & s & ts & m & tm & l \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & s & ts & m & tm \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & s & ts & m \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & s & ts \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & s \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right\}$$

Since this matrix allow us to represent the organization of the CaMKII subunits in a hexagonal prism, we can then calculate the pairwise interaction energies of a given kinase configuration based on the the quantity of state specific subunit interactions. Here we will show the example of calculating the short

interaction energy of a subunit configuration based upon the matrix representing the 3D hexagonal prism geometry. The short interaction matrix is shown below.

$$I_s = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Which we then left multiply by a matrix which represents a specific configuration of the kinase, and right multiply by its transpose. The *columns* of the matrix represent subunits 1-12 of the kinase. The rows represent subunit states a-d.

$$C = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

If we perform the matrix multiplication of,

$$[C][I_s][C']$$

we will end up with a 4x4 matrix where the *rows* correspond to subunit states a-d, and the *columns* correspond to subunit states a-d. The matrix will then show how many 's' or short interactions of each type are occurring in the specific kinase configuration. This matrix is upper triangular, as b-d interactions are the same as d-b interactions. For example, you may obtain 2 a-a interactions, 1 b-d interaction, 1 c-a interaction, and 1 d-c interaction. If that was the case, the resulting matrix would appear as:

$$\begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Once we have a matrix representing all types of short subunit interactions, we can determine the interaction energies between subunits if we can represent this energy in the form of a matrix. There is a negative energy change in the subunit when an activated state (*) has a primed neighbor or when a primed state has an activated neighbor. The matrix that includes this rule has columns of a, b, c, d , and rows of a, b, c, d .

We use a constant parameter A in the formula to represent the short interaction energies. In order to make the short interaction energy on the same order of magnitude of the individual energies of the monomers as discussed before (G_a, G_b, G_c, G_d), we represent the short interaction energies as: $\frac{-\log(A)}{\beta}$

So, we can represent the short interaction energies as a matrix governed by the rule that $A < 0$ when an activated state has a primed neighbor or when a primed state has an activated neighbor, multiplied by the scalar representing the constant short interaction energy between subunits.

$$\begin{Bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 \end{Bmatrix} \times \frac{-\log(A)}{\text{beta}}$$

Multiplying this scaled short interaction energy matrix by the previous matrix representing the number and type of each short interaction in the given kinase configuration, we obtain a scalar, which is the short interaction energy of the specific kinase configuration.

In order to obtain the total energy of the kinase configuration, we sum the individual subunit energies as given by multiplying the kinase configuration representative matrix by a 4×1 matrix of G_a, G_b, G_c, G_d , and summing the resulting matrix. These resulting monomer energies added to the short interaction energy configuration represent the total energy of the kinase configuration.

Implementing The Metropolis Hastings Algorithm

In order to understand the distribution of which subunits are activated most often given certain parameters, we need to be able to mimic the natural activation and priming of the kinase. By starting with an initial random kinase configuration, represented by a 4×12 matrix like,

$$C = \begin{Bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{Bmatrix}$$

where the rows are subunit states a-d, and the columns represent each individual subunit of the kinase, we can select a new state using the Metropolis Hastings algorithm.

We implement the Metropolis Hastings by selecting one column of the configuration matrix and elect to change the position of the '1' in the column, effectively changing what state that specific subunit of the kinase is in. The S_0 is the original configuration, and S_1 is the configuration with the state change. The Metropolis Hastings algorithm allows us to determine if we want to accept the change from S_0 to S_1 .

For both the S_0 and S_1 states we calculate the energy of that kinase configuration (taking into account both the monomer energies as well as the interaction energies). We then calculate $\alpha = \exp -\beta(GS_1 - GS_0)$, where GS. If the α value is greater than 1, we accept the S_1 as the new configuration. If α is less than 1, we compare it to a randomly generated r value (between 0 and 1), if this r value is less than the calculated r value, we accept S_1 as the new configuration. Otherwise, S_0 stays as the current configuration and the process is repeated. In summary, we accept S_1 as the new state if

$$\begin{aligned} \alpha &\geq 1 \\ r &\leq \alpha \end{aligned}$$

Otherwise, we keep S_0 as the configuration and repeat the process.

Appendix A Enumerating the state space for the CaMKII model

The first task is to enumerate the state space, that is, to determine the number of fundamentally different states in a model composed of n m -state subunits with a given spatial arrangement. For example, for 6 identical 2 state subunits, there are $2^6 = 64$ states. However, if these subunits are arranged in a hexagonal ring, the dihedral symmetry results in 13 essentially different states (shown in Fig. 1).

A.1 Burnside's theorem

When subunits have $m > 2$ states (or more complicated symmetry) it is helpful to calculate the number of essentially different states using **Burnside's theorem**. Let G be a finite group and X and finite G -set. Then

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |X^g|.$$

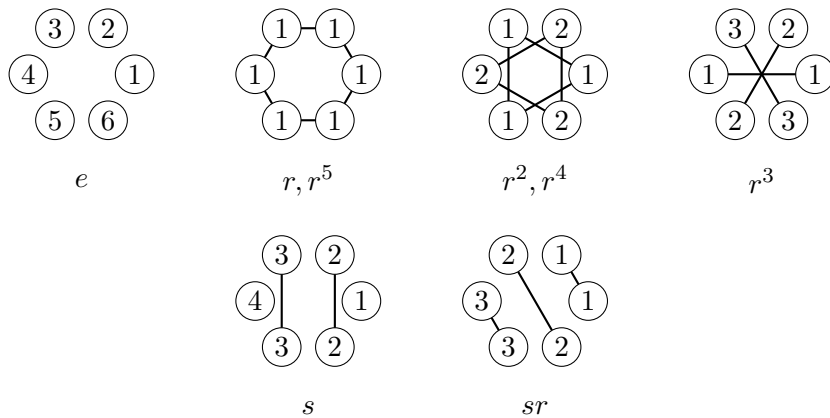
where $|X/G|$ is the number orbits under the action of G , $|X^g|$ is the number of elements stabilized by g , and the right side is the average size of a stabilized set. For example, if we were interested in a single ring of six kinases, the relevant group is the dihedral group of order 12 (usually denoted D_6 , see Appendix)..

A.2 2d hexagonal symmetry

Applying Burnside's theorem we obtain that the number of essentially different states for six hexagonally arranged subunits, each with m states, is

$$\sigma(m; 2d, D_6) = \frac{1}{12} \left[\underbrace{m^6}_e + \underbrace{2m}_{r, r^5} + \underbrace{2m^2}_{r^2, r^4} + \underbrace{m^3}_{r^3} + \underbrace{3m^3}_{sr, sr^3, sr^5} + \underbrace{3m^4}_{s, sr^2, sr^4} \right] \quad (1)$$

where $r^5 = r^{-1}$ (rotate counterclockwise by 5 is same as rotate clockwise by 1) and $r^4 = r^{-2}$. Each term in this expression is of the form am^b where b is the number of distinct groups of elements under one of the 12 the group actions, as follows.



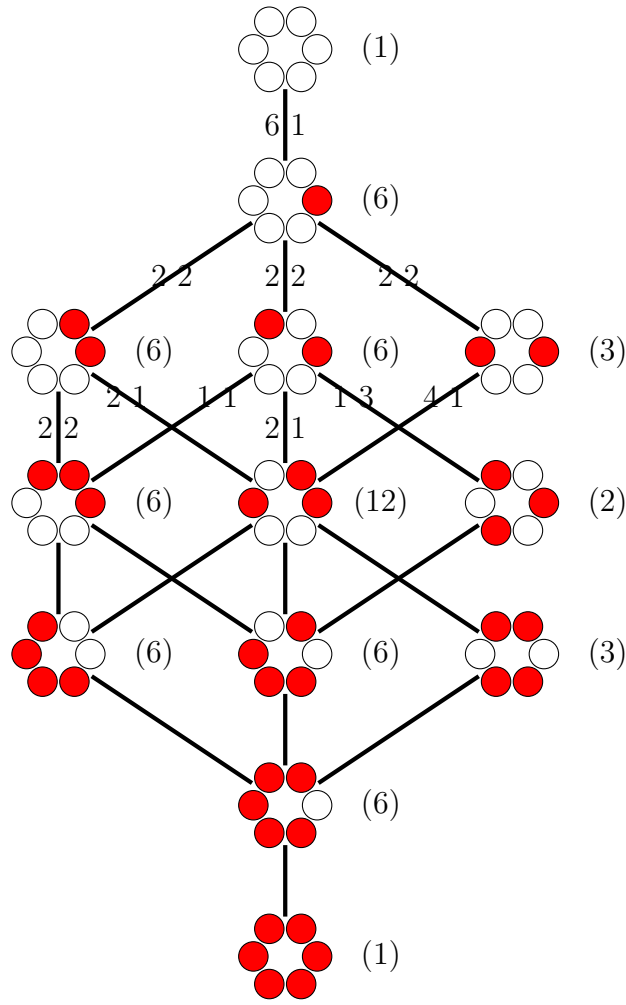


Figure 1: Transition diagram for two state model and hexagon. Summing the (multiplicities) one can confirm that there are a total of 2^6 states: $2(1) + 2(6) + 2(3) + 2(6) + 2(6) + 2 + 12 + 6 = 64$.

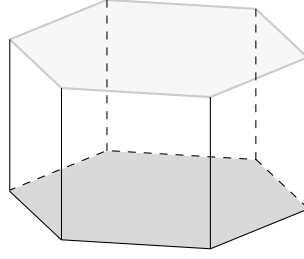
where sr^2 and sr^4 are similar to s , and where sr^3 and sr^5 are similar to sr . The formula simplifies to

$$\sigma(m; 2d, D_6) = \frac{1}{12} [m^6 + 3m^4 + 4m^3 + 2m^2 + 2m] .$$

Note that $\sigma(2; 2d, D_6) = 13$, consistent with the 2-colorings enumerated in Fig. 1; also $\sigma(3; 2d, D_6) = 92$ and $\sigma(4; 2d, D_6) = 430$. That is, if we the 4-state model that is shown in Fig. ??, and assumed six subunits arranged in a hexagon, then there would be 430 essentially different states.

3d hexagonal prism symmetry

The case we are really interested in uses $m = 4$ states (Fig. ??) and 12 subunits arranged as a *hexagonal prism*.

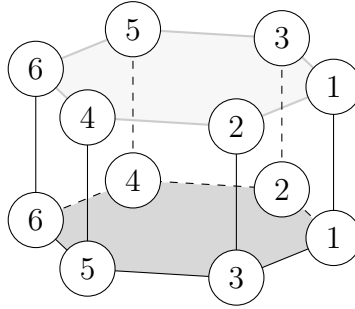


There are several appropriate choices of symmetry groups that could be appropriate. For examples, see

<http://newton.ex.ac.uk/research/qsystems/people/goss/symmetry/Solids.html>

One possible choice—perhaps the easiest—is, as an abstract group, D_6 (discussed above). However, in 3d the action of the element s is to rotate the hexagonal prism around a line in 3d space (as opposed to reflecting it about a plane). **NOTE: The correct choice depends on details of the structural biology. We should discuss this.**

Applying Burnside's theorem for the 3d hexagonal prism we obtain a formula similar to Eq. 1 but with the exponents changed (doubled in every case except s , sr^2 and sr^4 , where it is 6 rather than 8). The six groups for the case of s are shown here:



The other cases are just two copies of the 2d case. Using all 12 group actions, Burnside's theorem yields,

$$\sigma(m, 3d, D_6) = \frac{1}{12} \left[\underbrace{m^{12}}_e + \underbrace{2m^2}_{r, r^5} + \underbrace{2m^4}_{r^2, r^4} + \underbrace{m^6}_{r^3} + \underbrace{3m^6}_{sr, sr^3, sr^5} + \underbrace{3m^6}_{s, sr^2, sr^4} \right].$$

This formula simplifies to

$$\sigma(m; 3d, D_6) = \frac{1}{12} [m^{12} + 7m^6 + 2m^4 + 2m^2].$$

For a 2-state model we get $\sigma(2, 3d, D_6) = 382$ essentially different states (reduced by a factor of 10.7225 from $2^{12} = 4096$). For a 4-state model we get $\sigma(4, 3d, D_6) = 1400536$ essentially different states, that is, a little over 1.4 million(!). This is large, yes, but reduced by a factor of 11.9791 from $4^{12} = 16777216$.

Appendix B Dihedral groups

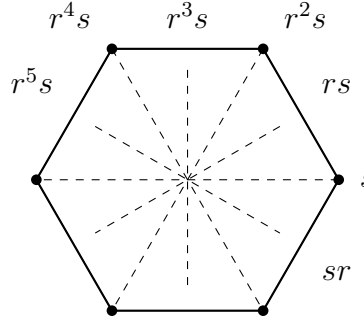
The dihedral group of order $2n$ is the symmetry group of the regular n -gon. D_n has presentation

$$D_n = \langle r, s \mid r^n = s^2 = (sr)^2 = e \rangle$$

That is D_n is generated by a rotation r of order n and a reflection s of order 2 such that $srs = r^{-1}$. The $2n$ elements of D_n can be written as $e, r, r^2, \dots, r^{n-1}, s, sr, sr^2, \dots, sr^{n-1}$. The first n elements are rotations; the second n elements are axis reflections. The product of two rotations or two reflections is a rotation; the product of two reflections is a rotation. Because $s^{-1} = s$, equivalent ways of writing $srs = r^{-1}$ are $sr = r^{-1}s$ and $rs = sr^{-1}$; thus,

$$sr^k = r^{-k}s, \quad r^k s = sr^{-k}.$$

D_6 is the appropriate group for hexagonal symmetry. It is well known that $D_6 = S_3 \times Z_2$. The twelve elements of D_6 are $e, r, r^2, r^3, r^4, r^5, s, rs, r^2s, r^3s, r^4s, r^5s$. Note the identities $rs = sr^5$, $r^2s = sr^4$, $r^3s = sr^3$, $r^4s = sr^2$, and $r^5s = sr$. If s is reflection over the horizontal axis, the axes of the other five reflections are as follows.



The actions are applied/read from right to left, that is, rs means “flip then rotate,” a reflection followed by a rotation.

Appendix C Lumping of states

Depending on the rate constants and allosteric interactions there may be additional lumping of states that may occur. In fact, we hope this is the case!

If 6 2-state subunits had no spatial relationship, and were indistinguishable, then there would be 7 distinct states. The allosteric parameters b, b^2 , etc...

$$[a^6 : a^5b : a^4b^2 : a^3b^3 : a^2b^4 : ab^5 : b^6]$$

$$2^6 = 64$$

Appendix D Cycle Index of D_8

The cycle index for D_8 is

$$P_{D_8}(x_1, x_2, x_3, x_4) = \frac{1}{8} [x_1^4 + 2x_1^2x_2 + 3x_2^2 + 2x_4]$$

In the case of 2 states for each monomer, a and b , this evaluates to a degree 4 polynomial with 5 terms, representing 6 essentially different states (6 is sum of coefficients),

$$P_{D_8}(a + b, a^2 + b^2, a^3 + b^3, a^4 + b^4) = a^4 + a^3b + 2a^2b^2 + ab^3 + b^4$$

In the case of 4 states for each monomer, $a-d$, this evaluates to a degree 4 polynomial with 35 terms, representing 55 essentially different states,

$$\begin{aligned}
P_{D_8}(a+b+c+d, \dots, a^4+b^4+c^4+d^4) = & a^4 + a^3b + 2a^2b^2 + ab^3 + b^4 + a^3c + 2a^2bc + 2ab^2c + b^3c \\
& + 2a^2c^2 + 2abc^2 + 2b^2c^2 + ac^3 + bc^3 + c^4 + a^3d + 2a^2bd \\
& + 2ab^2d + b^3d + 2a^2cd + 3abcd + 2b^2cd + 2ac^2d + 2bc^2d \\
& + c^3d + 2a^2d^2 + 2abd^2 + 2b^2d^2 + 2acd^2 + 2bcd^2 + 2c^2d^2 \\
& + ad^3 + bd^3 + cd^3 + d^4
\end{aligned}$$

For the 4 monomers at the vertices of the square (associated to D_8), the group of short edges of the complete graph corresponds to $y_1y_2+y_1y_4+y_2y_3+y_3y_4$ while the longer edges correspond to $y_1y_3+y_2y_4$. Define

$$Q(y_1, y_2, y_3, y_4) = \left[\frac{y_1y_2 + y_1y_4 + y_2y_3 + y_3y_4}{y_1y_3 + y_2y_4} \right].$$

The state $(y_1, y_2, y_3, y_4) = (a, a, a, b)$ is equivalent to the state (a, a, b, a) because in both cases

$$Q(a, a, a, b) = \left[\frac{2a^2 + 2ab}{a^2 + ab} \right] = Q(a, a, b, a).$$

The state (a, b, a, b) is distinct from state (a, a, b, b) because

$$Q(a, b, a, b) = \left[\frac{4ab}{a^2 + b^2} \right] \neq \left[\frac{a^2 + 2ab + b^2}{2ab} \right] = Q(a, a, b, b)$$

So there are 2 essentially different ways for the states a^2bc to be arranged on the vertices of the square.

$$Q(a, a, b, c) = \left[\frac{a^2 + ab + ac + bc}{ab + ac} \right] \neq \left[\frac{2ab + 2ac}{a^2 + bc} \right] = Q(a, b, a, c)$$

Appendix E Cycle Index of D_{12}

The cycle index for D_{12} is

$$P_{D_{12}}(x_1, x_2, \dots, x_6) = \frac{1}{12} [x_1^6 + 2x_6 + 2x_3^2 + 4x_2^3 + 3x_1^2x_2^2]$$

In the case of 2 states for each monomer, a and b , this evaluates to a degree 6 polynomial with 7 terms, representing 13 essentially different states (13 is sum of coefficients),

$$P_{D_{12}}(a+b, a^2+b^2, \dots, a^6+b^6) = a^6 + a^5b + 3a^4b^2 + 3a^3b^3 + 3a^2b^4 + ab^5 + b^6$$

In the case of 4 states for each monomer, $a-d$, this evaluates to a degree 6 polynomial with 84 terms,

representing 430 essentially different states,

$$\begin{aligned}
P_{D_{12}} = & a^6 + a^5b + 3a^4b^2 + 3a^3b^3 + 3a^2b^4 + ab^5 + b^6 + a^5c + 3a^4bc + 6a^3b^2c + 6a^2b^3c + 3ab^4c \\
& + b^5c + 3a^4c^2 + 6a^3bc^2 + 11a^2b^2c^2 + 6ab^3c^2 + 3b^4c^2 + 3a^3c^3 + 6a^2bc^3 + 6ab^2c^3 + 3b^3c^3 + 3a^2c^4 \\
& + 3abc^4 + 3b^2c^4 + ac^5 + bc^5 + c^6 + a^5d + 3a^4bd + 6a^3b^2d + 6a^2b^3d + 3ab^4d + b^5d + 3a^4cd \\
& + 10a^3bcd + 16a^2b^2cd + 10ab^3cd + 3b^4cd + 6a^3c^2d + 16a^2bc^2d + 16ab^2c^2d + 6b^3c^2d + 6a^2c^3d \\
& + 10abc^3d + 6b^2c^3d + 3ac^4d + 3bc^4d + c^5d + 3a^4d^2 + 6a^3bd^2 + 11a^2b^2d^2 + 6ab^3d^2 + 3b^4d^2 \\
& + 6a^3cd^2 + 16a^2bcd^2 + 16ab^2cd^2 + 6b^3cd^2 + 11a^2c^2d^2 + 16abc^2d^2 + 11b^2c^2d^2 + 6ac^3d^2 + 6bc^3d^2 \\
& + 3c^4d^2 + 3a^3d^3 + 6a^2bd^3 + 6ab^2d^3 + 3b^3d^3 + 6a^2cd^3 + 10abcd^3 + 6b^2cd^3 + 6ac^2d^3 + 6bc^2d^3 \\
& + 3c^3d^3 + 3a^2d^4 + 3abd^4 + 3b^2d^4 + 3acd^4 + 3bcd^4 + 3c^2d^4 + ad^5 + bd^5 + cd^5 + d^6
\end{aligned}$$

In this case

$$Q(y_1, y_2, y_3, y_4, y_5, y_6) = \begin{bmatrix} y_1y_2 + y_1y_6 + y_2y_3 + y_3y_4 + y_4y_5 + y_5y_6 \\ y_1y_3 + y_1y_5 + y_2y_6 + y_3y_5 + y_2y_4 + y_4y_6 \\ y_1y_4 + y_2y_5 + y_3y_6 \end{bmatrix}.$$

So say I'm in state $(a, b, a, c, c, a) \sim a^3bc^2$, then

$$Q(a, b, a, c, c, a) = \begin{bmatrix} a^2 + 2ab + 2ac + c^2 \\ a^2 + ab + 3ac + bc \\ a^2 + ac + bc \end{bmatrix} = Q(a, a, b, a, c, c).$$

where the one on the right is the standard presentation.

If one loops through every 6^4 states then the first time one encounters a new Q , this will be the standard presentation for that Q .

Appendix F Allosteric coupling focusing on interacting transitions or interaction energies of states

The method of enumerating states, transitions, and allosteric parameters in models of receptor oligomers focuses on how one transition affects another. A viewpoint that focuses on states and interaction energies ought to be equivalent.

F.1 Gibbs free energy

The relationship between Gibbs free energy, enthalpy, and entropy is

$$\Delta G = \Delta H - T\Delta S$$

where T is the absolute temperature. Multiplying both sides by the dimensionless inverse temperature $\beta = 1/(k_B T)$ where k_B is Boltzmann's constant, gives

$$-\beta\Delta G = -\beta\Delta H + \Delta S/k_B$$

Rearranging...

$$n\kappa = e^{-\beta\Delta G} = e^{\Delta S/k_B} e^{-\beta\Delta H}$$

so

$$\kappa = e^{-\beta\Delta H} \quad \text{and} \quad n = e^{\Delta S/k_B}$$

This implies $\Delta S = k_B \log n$ (Boltzmann's formula). Also, $-\beta\Delta H = \log \kappa$, so $\Delta H = -k_B T \log \kappa$, $H_{products} - H_{reactants} = -k_B T \log \kappa = k_B T \log(1/\kappa) = \beta^{-1} \log \kappa^{-1} = -(\log \kappa)/\beta$. This is correct, because κ is an equilibrium constant ($\kappa = [\text{products}]/[\text{reactants}]$). When $\Delta H < 0$, $H_{products} < H_{reactants}$, and $-k_B T \log \kappa < 0$, that is, $\log \kappa > 0$ which implies products are favored.

F.2 Example 1

My original approach:

$$a \xleftarrow{\hat{\kappa}_b} b \qquad a^2 \xleftarrow{2\kappa_b} ab \xleftarrow{\frac{1}{2}\kappa_b\eta_{bb}} b^2$$

In terms of interaction energies

$$\begin{aligned} H_{a^2} &= 2H_a + \sigma_{aa} \\ H_{ab} &= H_a + H_b + \sigma_{ab} \\ H_{b^2} &= 2H_b + \sigma_{bb} \end{aligned}$$

$$\begin{aligned} \Delta H_{a^2}^{ab} &= H_{ab} - H_{a^2} = \overbrace{H_b - H_a}^{\Delta H_a^b} + \overbrace{\sigma_{ab} - \sigma_{aa}}^{\Delta \sigma_{aa}^{ab}} \\ \Delta H_{ab}^{b^2} &= H_{b^2} - H_{ab} = \overbrace{H_b - H_a}^{\Delta H_a^b} + \overbrace{\sigma_{bb} - \sigma_{ab}}^{\Delta \sigma_{ab}^{bb}} \end{aligned}$$

$$\begin{aligned} 2\kappa_b &= \exp(-\beta\Delta G_{a^2}^{ab}) = \exp[-\beta(\Delta H_a^b + \Delta \sigma_{aa}^{ab} - \Delta S_{a^2}^{ab})] \\ \kappa_b &= \exp[-\beta(\Delta H_a^b + \Delta \sigma_{aa}^{ab})] \\ \Delta S_{a^2}^{ab} &= k_B \log 2 \end{aligned}$$

$$\begin{aligned} \frac{1}{2}\kappa_b\eta_{bb} &= \exp(-\beta\Delta G_{ab}^{b^2}) = \exp[-\beta(\Delta H_a^b + \Delta \sigma_{ab}^{bb} - \Delta S_{ab}^{b^2})] \\ \kappa_b\eta_{bb} &= \exp[-\beta(\Delta H_a^b + \Delta \sigma_{ab}^{bb})] \\ \Delta S_{ab}^{b^2} &= k_B \log \frac{1}{2} \\ \eta_{bb} &= \frac{\exp[-\beta(\Delta H_a^b + \Delta \sigma_{ab}^{bb})]}{\exp[-\beta(\Delta H_a^b + \Delta \sigma_{aa}^{ab})]} \\ &= \exp[-\beta(\Delta \sigma_{ab}^{bb} - \Delta \sigma_{aa}^{ab})] = \exp[-\beta(\sigma_{bb} - 2\sigma_{ab} + \sigma_{aa})] \end{aligned}$$

$$\pi_{b^2} = \frac{e^{-\beta G_{b^2}}}{e^{-\beta G_{a^2}} + e^{-\beta G_{ab}} + e^{-\beta G_{b^2}}} = \frac{e^{-\beta(\Delta G_{ab}^{b^2} + \Delta G_{a^2}^{ab})}}{1 + e^{-\beta\Delta G_{a^2}^{ab}} + e^{-\beta(\Delta G_{ab}^{b^2} + \Delta G_{a^2}^{ab})}} = \frac{\kappa_b^2\eta_{bb}}{1 + 2\kappa_b + \kappa_b^2\eta_{bb}}$$

and similarly for π_{ab} and π_{a^2} .

Either way there are 2 free parameters: $\kappa_b = \exp[-\beta(\Delta H_a^b + \Delta\sigma_{aa}^{ab})]$ and $\eta_{bb} = \exp[-\beta(\Delta\sigma_{ab}^{bb} - \Delta\sigma_{aa}^{ab})]$. Note that $\kappa_b\eta_{bb} = \exp[-\beta(\Delta H_a^b + \Delta\sigma_{ab}^{bb})]$. If we write $\kappa'_b = \kappa_b\eta_{bb}$ then

$$\pi_{b^2} = \frac{\kappa_b\kappa'_b}{1 + 2\kappa_b + \kappa_b\kappa'_b}$$

and similarly for π_{ab} and π_{a^2} .

F.3 In general?

So now when we consider interactions between various states, how is it that we always get the correct number of free parameters? Do we? We should! But I don't understand why this will happen.

Ignoring physical location of monomers, the number of states is $v^{(k)} = \binom{v+k-1}{k}$. The number of edges in a spanning tree is $\binom{v+k-1}{k} - 1$. The number of edges in a spanning tree of the monomer is $e = v - 1$ and this is equal to the number of free energy differences between states. For a dimer, $k = 2$ and $v^{(2)} = \binom{v+1}{2} = v(v+1)/2$, so the number of free parameters in the dimer is $v(v+1)/2 - 1$. For the dimer, there is only one type of interaction (the interaction distance), so the number pairwise interaction energies is $v(v+1)/2$, and the number of free differences is $v(v+1)/2 - 1$, as it must be. Good.

Appendix G Analytical calculation of cooperativity

Writing $x_i = c_i e^{-\beta h_i} / \sum_j c_j e^{-\beta h_j}$ where $\sum_i x_i = 1$,

$$\begin{aligned} h_0 &= 3\theta_a + 3s_{aa} = 0 \\ h_1 &= 2\theta_a + \theta_b + s_{aa} + 2s_{ab} = \theta_b + 2s_{ab} = \theta + \sigma_1 \\ h_2 &= \theta_a + 2\theta_b + 2s_{ab} + s_{bb} = 2\theta_b + 2s_{ab} + s_{bb} = 2\theta + \sigma_2 \\ h_3 &= 3\theta_b + 3s_{bb} = 3\theta_b + 3s_{bb} = 3\theta + \sigma_3 \end{aligned}$$

where we use $\theta_a = 0$ and $s_{aa} = 0$, and $\theta = \theta_b$ (for notational simplicity). The fraction of activated subunits is

$$\begin{aligned} y &= a_0 x_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 \\ &= x_1 + 2x_2 + x_3 \\ &= \frac{e^{-\beta h_1} + 2e^{-\beta h_2} + e^{-\beta h_3}}{1 + 3e^{-\beta h_1} + 3e^{-\beta h_2} + e^{-\beta h_3}} \end{aligned}$$

where $a_0 = 0$, $a_1 = 1/3$, $a_2 = 2/3$, $a_3 = 1$, and $e^{-\beta h_0} = 1$.

$$\begin{aligned} 1 - y &= 1 - \frac{e^{-\beta h_1} + 2e^{-\beta h_2} + e^{-\beta h_3}}{1 + 3e^{-\beta h_1} + 3e^{-\beta h_2} + e^{-\beta h_3}} \\ &= \frac{1 + 3e^{-\beta h_1} + 3e^{-\beta h_2} + e^{-\beta h_3} - e^{-\beta h_1} - 2e^{-\beta h_2} - e^{-\beta h_3}}{1 + 3e^{-\beta h_1} + 3e^{-\beta h_2} + e^{-\beta h_3}} \\ &= \frac{1 + 2e^{-\beta h_1} + e^{-\beta h_2}}{1 + 3e^{-\beta h_1} + 3e^{-\beta h_2} + e^{-\beta h_3}} \end{aligned}$$

Thus,

$$\frac{y}{1-y} = \frac{e^{-\beta h_1} + 2e^{-\beta h_2} + e^{-\beta h_3}}{1 + 2e^{-\beta h_1} + e^{-\beta h_2}}$$

and

$$\ln \frac{y}{1-y} = \ln (e^{-\beta h_1} + 2e^{-\beta h_2} + e^{-\beta h_3}) - \ln (1 + 2e^{-\beta h_1} + e^{-\beta h_2})$$

We want to calculate

$$z = \frac{d}{d \ln \theta} \left[\ln \frac{y}{1-y} \right]$$

and then find the Hill coefficient $\max_{\theta} z$. Because $d \ln \theta = \frac{d\theta}{\theta}$,

$$z = \theta \frac{d}{d\theta} \left[\ln \frac{y}{1-y} \right]$$

$$z = \theta \frac{d}{d\theta} [\ln (e^{-\beta h_1} + 2e^{-\beta h_2} + e^{-\beta h_3}) - \ln (1 + 2e^{-\beta h_1} + e^{-\beta h_2})]$$

$$z = -\beta \theta \left[\frac{e^{-\beta h_1} + 4e^{-\beta h_2} + 3e^{-\beta h_3}}{e^{-\beta h_1} + 2e^{-\beta h_2} + e^{-\beta h_3}} - \frac{2e^{-\beta h_1} + 2e^{-\beta h_2}}{1 + 2e^{-\beta h_1} + e^{-\beta h_2}} \right] = -\beta \theta A$$

$$0 = \frac{dz}{d \ln \theta} = \theta \frac{dz}{d\theta} = -\theta \beta \left(A + \theta \frac{dA}{d\theta} \right)$$

Because $d \ln \theta = \frac{d\theta}{\theta}$ and $d \ln \frac{y}{1-y} = \frac{1-y}{y} d \left[\frac{y}{1-y} \right] = \frac{dy}{y(1-y)}$,

$$z = \frac{\theta}{y(1-y)} \frac{dy}{d\theta}.$$

$$0 = \frac{dz}{d \ln \theta} = \theta \frac{dz}{d\theta} = \theta \frac{d}{d\theta} \left[\frac{\theta}{y(1-y)} \frac{dy}{d\theta} \right].$$

$$\left[\frac{\theta y'}{y(1-y)} \right]' = \frac{(y' + \theta y'')y(1-y) - (1-2y)\theta y'}{y^2(1-y)^2}$$

Because $\theta \neq 0$, the condition for the maximum is

$$0 = (y' + \theta y'')y(1-y) - (1-2y)\theta y'$$

$$0 = (yy' + \theta yy'')(1-y) - \theta y' + 2y\theta y'$$

Note that $\frac{d}{d\theta}[e^{-\beta h_i}] = -i\beta e^{-\beta h_i}$, so writing $y = N/D$,

$$\frac{dy}{d\theta} = \frac{N'D - ND'}{D^2}$$

where

$$\begin{aligned}
N &= e^{-\beta h_1} + 2e^{-\beta h_2} + e^{-\beta h_3} \\
N' &= -\beta [e^{-\beta h_1} + 4e^{-\beta h_2} + 3e^{-\beta h_3}] \\
D &= 1 + 3e^{-\beta h_1} + 3e^{-\beta h_2} + e^{-\beta h_3} \\
D' &= -\beta [3e^{-\beta h_1} + 6e^{-\beta h_2} + 3e^{-\beta h_3}]
\end{aligned}$$

$$\begin{aligned}
N'D &= -\beta [e^{-\beta h_1} + 4e^{-\beta h_2} + 3e^{-\beta h_3}] [1 + 3e^{-\beta h_1} + 3e^{-\beta h_2} + e^{-\beta h_3}] \\
-N'D/\beta &= e^{-\beta h_1} + 4e^{-\beta h_2} + 3e^{-\beta h_3} \\
&+ 3e^{-2\beta h_1} + 12e^{-\beta(h_1+h_2)} + 9e^{-\beta(h_1+h_3)} \\
&+ 3e^{-\beta(h_1+h_2)} + 12e^{-2\beta h_2} + 9e^{-\beta(h_2+h_3)} \\
&+ e^{-\beta(h_1+h_3)} + 4e^{-\beta(h_2+h_3)} + 3e^{-2\beta h_3}
\end{aligned}$$

$$\begin{aligned}
ND' &= -\beta [e^{-\beta h_1} + 2e^{-\beta h_2} + e^{-\beta h_3}] [3e^{-\beta h_1} + 6e^{-\beta h_2} + 3e^{-\beta h_3}] \\
-ND'/\beta &= 3e^{-2\beta h_1} + 6e^{-\beta(h_1+h_2)} + 3e^{-\beta(h_1+h_3)} \\
&+ 6e^{-\beta(h_1+h_2)} + 12e^{-2\beta h_2} + 6e^{-\beta(h_2+h_3)} \\
&+ 3e^{-\beta(h_1+h_3)} + 6e^{-\beta(h_2+h_3)} + 3e^{-2\beta h_3}
\end{aligned}$$

$$\begin{aligned}
-N'D/\beta &= e^{-\beta h_1} + 4e^{-\beta h_2} + 3e^{-\beta h_3} \\
&+ 3e^{-2\beta h_1} + 15e^{-\beta(h_1+h_2)} + 10e^{-\beta(h_1+h_3)} + 12e^{-2\beta h_2} + 13e^{-\beta(h_2+h_3)} + 3e^{-2\beta h_3} \\
-ND'/\beta &= 3e^{-2\beta h_1} + 12e^{-\beta(h_1+h_2)} + 6e^{-\beta(h_1+h_3)} + 12e^{-2\beta h_2} + 12e^{-\beta(h_2+h_3)} + 3e^{-2\beta h_3}
\end{aligned}$$

$$\frac{dy}{d\theta} = \frac{N'D - ND'}{D^2}$$

$$N'D - ND' = -\beta [e^{-\beta h_1} + 4e^{-\beta h_2} + 3e^{-\beta h_3} + 3e^{-\beta(h_1+h_2)} + 4e^{-\beta(h_1+h_3)} + e^{-\beta(h_2+h_3)}]$$