We wish to develop a rule-based model for a system in which autophosphorylation of a receptor tyrosine kinase (RTK) can generate a multitude of receptor phosphoforms and phosphorylation-dependent adapter-bound receptor states. The model should capture the interactions a cytosolic adapter protein (or set of such proteins) with a (tightly or loosely associated) dimer of RTKs. The adapter protein is comprised of a Src homology 2 (SH2) domain, and each receptor in a dimer is comprised of an active catalytic subunit and n autophosphorylation sites. When one of these sites is phosphorylated, we will assume that it can bind the SH2 domain of an adapter protein. Using pseudo BioNetGen Language (BNGL), one can specify rules for the system of interest as follows:

$$\mathrm{D(Y}_{ij}\mathrm{\sim U) \rightarrow D(Y}_{ij}\mathrm{\sim P)} \quad (1a)$$

$$\mathrm{D(Y}_{ij}\mathrm{\sim P) \rightarrow D(Y}_{ij}\mathrm{\sim U)} \quad (1b)$$

$$\mathrm{D(Y}_{ij}\mathrm{\sim P) + A}_{k}\mathrm{(SH2) <-> D(Y}_{ij}\mathrm{\sim P!1).A}_{k}\mathrm{(SH2!1)} \quad (1c)$$

where 'D' denotes a receptor dimer, 'Y_{ij}' (i=1,…,n; j=1,2) denotes the ith tyrosine of the jth receptor in a receptor dimer, 'U' denotes an unphosphorylated tyrosine, 'P' denotes a phosphorylated tyrosine, 'A_k' denotes the kth adapter protein, and 'SH2' denotes the SH2 domain of an adapter protein. (As usual in BNGL, the internal state label of a molecular component is prefixed by a tilde, and a bond name is prefixed by an exclamation mark. Sharing of a bond name indicates that two molecular components are connected.) The above rules represent autophosphorylation of receptor tyrosines (Eq. (1a)), dephosphorylation of receptor tyrosines via phosphatases not explicitly included in the model (Eq. (1b)), and reversible adapter-receptor binding via SH2 domain recognition of phosphotyrosine (Eq. (1c)). What assumptions have been made in writing these rules?

For the case of a single adapter protein, the total number of rules defined in Eqs. (1) is $6n$, and as can easily be confirmed, the number of chemical species implied by a rule set, N, is given by

$$N = 1 + 3^n + \binom{3^n}{2} = 1 + 3^n(1 + 3^n)/2 \quad (2)$$

Thus, the rules of Eqs. (1) tend to imply a large reaction network. Consider the case of m adapter proteins, where adapter A_k (k=1,…,m) interacts with a subset s_k of the n autophosphorylation sites of a receptor. How many species are possible in this case? Although N tends to be large (Eq. (2)), each receptor tyrosine is independent according to Eqs. (1). As a result, the network implied by Eqs. (1) can be characterized by a number of coupled ordinarily differential equations (ODEs) derived from the law of mass action that is much smaller than N.

Assuming a single adapter protein, we can write the following mass-action equations for i=1,…,n and j=1,2:

$$d[U_{ij}]/dt = -\phi_i[U_{ij}] + \delta_i[P_{ij}] \quad (3a)$$

$$d[P_{ij}]/dt = \phi_i[U_{ij}] - \delta_i[P_{ij}] - k_{+i}[P_{ij}] + k_{-i}[AP_{ij}] \quad (3b)$$

$$d[AP_{ij}]/dt = k_{+i}[P_{ij}][A] - k_{-i}[AP_{ij}] \quad (3c)$$

where $[U_{ij}]$ is the concentration of the ith tyrosine in the jth receptor in unphosphorylated form, $[P_{ij}]$ is the concentration of the ith tyrosine in the jth receptor in phosphorylated form and unbound to adapter, $[AP_{ij}]$ is the concentration of adapter protein bound to the ith tyrosine in the jth receptor, $[A]$ is the concentration of free adapter protein, ϕ_i is the apparent first-order rate constant for autophosphorylation of the ith tyrosine in a receptor (we assume that autophosphorylation is substrate limited), δ_i is the apparent first-order rate constant for dephosphorylation of the ith tyrosine in a receptor (we assume that phosphatases are present in excess), k_{+i} is the rate constant for binding of the adapter protein to the ith tyrosine in a receptor, and k_{-i} is the rate constant for dissociation of the adapter protein from the ith tyrosine in a receptor. If we assume that mass is conserved on the time scale of interest, we can also write the following equation:

$$[A] = [A_T] - \sum_{j=1}^{n} [AP_{ij}] \quad (4)$$

where $[A_T]$ is total concentration of adapter protein. In writing Eqs. (3), we have assumed that $[U_{ij}]_{\mathrm{local}} \ll K_{\mathrm{M}}$, where $[U_{ij}]_{\mathrm{local}}$ is the local concentration of tyrosine i in receptor j in the vicinity of the protein tyrosine kinase of receptor {1 or 2}$\neq j$, and K_{M} is the Michaelis-Menten constant for the protein tyrosine kinase. How would the model need to be changed to account for $[U_{ij}]_{\mathrm{local}} \gg K_{\mathrm{M}}$ for all i and j (i.e., for saturation of the protein tyrosine kinases in a receptor dimer)?

Eqs. (3) and (4) can be solved using standard numerical integration methods. System dynamics can also be found by using BNGL to specify a rule-based model and then simulating this model using the various methods available within the BioNetGen framework.