Electronic Supplementary Material (ESI) for Molecular BioSystems

Electronic Supplementary Information (ESI)

Exploring higher-order EGFR oligomerisation and phosphorylation—a combined experimental and theoretical approach

N. Kozer, D. Barua, S. Orchard, E.C. Nice, A.W. Burgess, W.S. Hlavacek and A.H.A. Clayton

Appendix S1

The results of Figs. S3 and S4 were generated as described herein. The general methodology used is standard and is presented for the sake of completeness and clarity. We are careful to note problem-specific aspects of the methodology.

Inspired by recent uses of Bayesian methods for analysis of models of cell signaling systems (Klinke, 2009; Klinke et al., 2012; Eydgahi et al., 2013), we used a Bayesian approach and the data presented in Fig. 2B to estimate the free parameters of our model (k_{cx}, k_{cr}, k_u, k_v, χ, and α). Other model parameters were fixed at the values given in Table 1. In the estimation procedure, we introduced log-transformed parameters $\Theta = (\theta_1, \ldots, \theta_6)$, where each θ_i is the logarithm (base 10) of one of the free parameters identified above. (The parameters and the elements of Θ are in one-to-one correspondence.) A change of a log-transformed parameter from θ_i to $\theta_i + \Delta\theta_i$ corresponds to a $10^{\Delta\theta_i}$-fold change in the value of the corresponding model parameter.

To describe the estimation procedure, let us use $y = (y_1, \ldots, y_5)$ to represent the set of measured average receptor cluster densities reported in Fig. 2B and $M(\Theta) = (M_1(\Theta), \ldots, M_5(\Theta))$ to represent the set of corresponding receptor cluster densities predicted by our model (of Fig. 1) when the free parameters are assigned the values consistent with Θ.

The procedure is based on Bayes' theorem, which provides the following relationship amongst certain probabilities and conditional probabilities:

$$P(\Theta|y) = \frac{P(y|\Theta)P(\Theta)}{P(y)}$$

The term on the left-hand side can be interpreted as quantifying the degree of belief in Θ after considering y. In the parlance of Bayesian statistics, $P(\Theta|y)$ is called the posterior, $P(\Theta)$ is called the prior, $P(y|\Theta)$ is called the likelihood, and $P(y)$ is called the evidence. By taking the logarithm of both sides, the above expression can be rewritten as follows: $\ln P(\Theta|y) = -\ln P(y) + \ln P(\Theta) + \ln P(y|\Theta)$. In this expression, $-\ln P(y)$ is a constant term that can be ignored (for reasons that will become clear later), $\ln P(\Theta)$ can be specified arbitrarily to reflect prior opinion about Θ (although particular functional forms are commonly used), and $\ln P(y|\Theta)$ can be equated with $-\chi^2$, where

$$\chi^2 = \sum_{i=1}^{5} \frac{1}{2\sigma_{y_i}^2}(y_i - M_i(\Theta))^2$$

In the χ^2 function (which is the objective function minimized in nonlinear least squares fitting), σ_{y_i} is the standard deviation that characterizes measurement noise for y_i. We specified values of σ_{y_i} consistent with the error bars shown in Fig. 2B.

We assumed a normal prior:

$$P(\Theta) = \exp\left(-\sum_{j=1}^{6} \frac{1}{2\sigma_{\theta_j}^2}(\mu_j - \theta_j)^2\right)$$

where μ_j represents an initial guess for the value of θ_j and each σ_{θ_j} is a hyperparameter of the prior. A small (large) value for $|\sigma_{\theta_j}|$ represents high (low) confidence in the initial guess μ_j. We set $\sigma_{\theta_j} = 1\ \forall j$, and we set each μ_j to the logarithm of the best-fit value listed in Table 1 for the corresponding free parameter. These choices bias the estimation procedure toward the nominal parameter values found through nonlinear least squares fitting and reflect uniform confidence in our initial guesses.

Given the above considerations, it follows that $\ln P(\Theta|y)$ is equal to the following expression up to an additive constant:

$$-\sum_{i=1}^{5} \frac{1}{2\sigma_{y_i}^2}(y_i - M_i(\Theta))^2 - \sum_{j=1}^{6} \frac{1}{2\sigma_{\theta_j}^2}(\mu_j - \theta_j)^2$$

This expression was used in the formula given next to estimate the posterior via a Markov chain Monte Carlo (MCMC) algorithm. The algorithm involves a random walk in the log-transformed parameter space. At each step along the walk, a new position in the parameter space, Θ^*, is proposed based on the current position Θ. The proposed move to a new position is accepted or rejected with probability κ according to the Metropolis-Hastings criterion:

$$\kappa = \min\left\{1, \frac{P(\Theta^*|y)}{P(\Theta|y)}\right\}$$

If the proposed move is accepted, the current Θ is replaced by the proposed Θ^*; if the proposed move is rejected, Θ remains unchanged. A stationary distribution is reached after an initial burn-in period, and the stationary distribution is sampled to estimate the posterior. A single MCMC run was performed to obtain the results of Figs. S3 and S4. After a burn-in period of 500,000 steps, we thereafter sampled (i.e., recorded Θ) every 100th step and collected a total of 5,000 samples. The run started from a randomly selected point in log-transformed parameter space by setting $\theta_j = \mu_j + (1 - 2r_j)$ for $j = 1, \ldots, 6$, where each r_j is a uniform random deviate between 0 and 1. At each MCMC step, the proposed move from Θ to Θ^* was found by using the following equation:

$$\Theta^* = \Theta + \gamma\frac{\Delta}{||\Delta||}$$

where γ is an algorithmic parameter that tunes the probability of accepting a proposed move and $\Delta = \{\rho_1, \ldots, \rho_6\}$, where each $\rho_j = N(0, 1)$ for $j = 1, \ldots, 6$ is a standard normal deviate. We set $\gamma = 0.1$, which yielded an acceptance probability of approximately 35%.

From the MCMC procedure described above, we obtained an estimate of the posterior (or joint posterior distribution) in the form of the 5,000 sampled Θ sets and the marginal posterior distributions for each of the six free parameters (Fig. S3). Each histogram in Fig. S3 characterizes the independent variation in the estimate of a particular parameter value; these histograms do not reflect covariation of parameter values. As described in the caption of Fig. S4, the joint posterior distribution (which does reflect covariation) was sampled—a Θ set was chosen randomly with uniform probability from the 5,000 Θ sets—and the corresponding sets of values for the free model parameters were used to quantify confidence in the prediction of our model that relative EGFR cluster density decreases with increasing EGF concentration (Fig. S4).

References

1. Eydgahi H, Chen WW, Muhlich JL, Vitkup D, Tsitsiklis JN, Sorger PK (2013) Properties of cell death models calibrated and compared using Bayesian approaches. *Mol Syst Biol* **9**:644.

2. Klinke DJ 2nd (2009) An empirical Bayesian approach for model-based inference of cellular signaling networks. *BMC Bioinformatics* **10**:371.

3. Klinke DJ 2nd, Cheng N, Chambers E (2012) Quantifying crosstalk among interferon-γ, interleukin-12, and tumor necrosis factor signaling pathways within a T_H1 cell model. *Sci Signal* **5**:ra32.