Supporting Information for

A nonequilibrium allosteric model for receptor-kinase complexes: The role of energy dissipation in chemotaxis signaling

David Hathcock, Qiwei Yu, Bernardo A. Mello, Divya N. Amin, Gerald L. Hazelbauer and Yuhai Tu

Yuhai Tu.
E-mail: yuhai@us.ibm.com

This PDF file includes:

Supporting Information Text

Testing equilibrium allosteric models

MWC model: fitting ligand binding leads to inconsistent kinase activity prediction. In similar spirit to Fig. 1B of the main text, we use the MWC model to fit the binding curves and plot the resulting kinase activity curves. The results are shown in Fig. S1, which is another demonstration that the MWC model is inconsistent with the data from Amin and Hazelbauer (1).

Linear dependence between additional degrees of freedom. In the main text [Eq. (7)], we argued that coupling an additional equilibrium degree of freedom (e.g. kinase activity) to the MWC model produces a response that is linearly related to the MWC activity. Here we show this result is quite general and holds for any chain of binary variables coupled via the following Hamiltonian,

$$H = (-\mu + E_{0,1}s_1)\sum_{i=1}^{N}\sigma_i + \sum_{i=1}^{M-1}(E_i s_i + E_{i,i+1}s_i s_{i+1}) + E_M s_M. \quad [1]$$

As in the main text, $\sigma_i = 0, 1$ denotes the receptor occupancy, which depends on ligand concentration $[L]$ and dissociation constant K_i through the chemical potential $\mu = \log([L]/K_i)$. The variables $s_i = 0, 1$ are activities for various components of the system. For example, these might represent different conformational changes along the receptor body or kinase-controlling tip as well as conformational changes of the kinase itself, each coupled in a chain via equilibrium mechanisms. For each s_i, E_i is the energy difference between the active and inactive states, and $E_{i,i+1}$ is the coupling energy between neighboring conformations.

We can relate the average states of neighboring conformational degrees of freedom using conditional probability:

$$\begin{aligned}\langle s_{j+1}\rangle &= \langle s_{j+1}|s_j = 0\rangle P(s_j = 0) + \langle s_{j+1}|s_j = 1\rangle P(s_j = 1)\\ &= \langle s_{j+1}|s_j = 0\rangle(1 - \langle s_j\rangle) + \langle s_{j+1}|s_j = 1\rangle\langle s_j\rangle \quad [2]\\ &= C_0 + C_1\langle s_j\rangle,\end{aligned}$$

where C_0, C_1 are constants independent of μ. The final step relies on the interactions between activities being equilibrium, so that the conditional expectation $\langle s_{j+1}|s_j = 0, 1\rangle$ is independent of μ. For nonequilibrium models, including the model in the main text, this is not generally true: the conditional expectation can be a function of $\langle s_j\rangle$ and therefore maintain μ-dependence. Given the linear relationship between $\langle s_{j+1}\rangle$ and $\langle s_j\rangle$ in equilibrium models, the response curves must be identical after normalization,

$$\frac{\langle s_{j+1}\rangle_{\max} - \langle s_{j+1}\rangle}{\langle s_{j+1}\rangle_{\max} - \langle s_{j+1}\rangle_{\min}} = \frac{\langle s_{j+1}\rangle_{\max} - \langle s_{j+1}\rangle}{\langle s_{j+1}\rangle_{\max} - \langle s_{j+1}\rangle_{\min}}. \quad [3]$$

Since linear proportionality applies to any pair of neighboring activities, it extends to the entire chain. Therefore, if a system is governed by equilibrium interactions between neighboring conformational states, any measurement of activity can be captured by an effective MWC model.

Ising models fail to explain both ligand binding and kinase response. Ising models are also frequently used to study the behavior of chemoreceptor clusters (2, 3). In contrast to the MWC model, which assumes all-or-none activity within a cluster, Ising models allow each receptor to be independently active or inactive, based on its occupancy. The receptors interact cooperatively via equilibrium mechanisms (perhaps reflective of the lattice spatial structure present in real chemoreceptor complexes). A general Ising Hamiltonian for a chemoreceptor cluster is given by,

$$H = \sum_{i=1}^{N}(E_s s_i + E_0 s_i\sigma_i - \mu\sigma_i) - \sum_{i\neq j} J_{ij}s_i s_j. \quad [4]$$

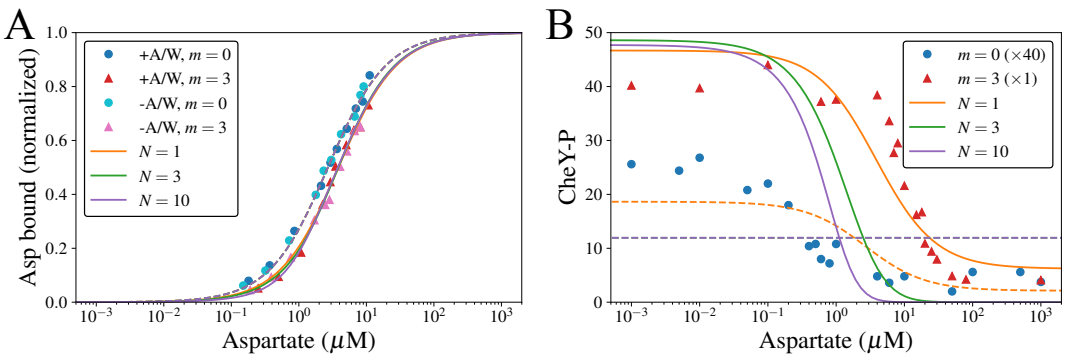

Fig. S1. The measurements by Amin and Hazelbauer (1) are inconsistent with the MWC model. Here, we fit the binding curves and use the fitting parameters to predict kinase activity.

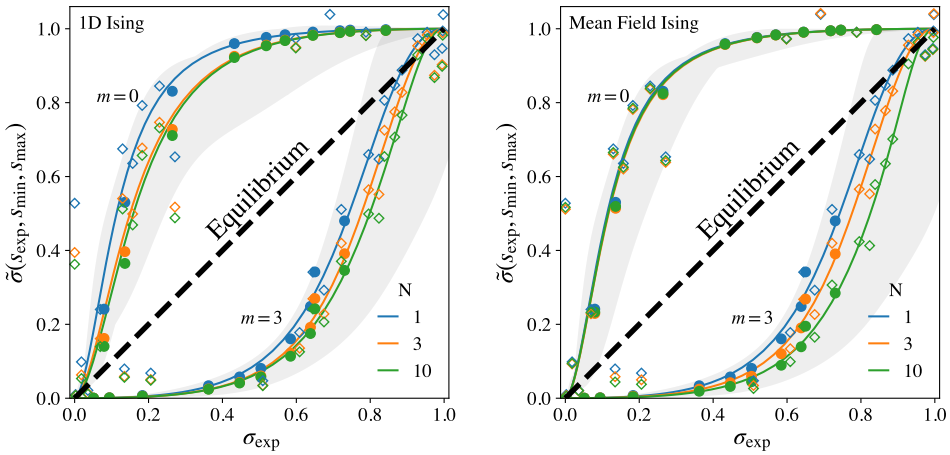

Fig. S2. Measured aspartate binding and kinase activity of Tar-containing signaling complexes (1) are inconsistent with equilibrium Ising models. Parametric test plotting the occupancy inferred from kinase activity measurements $\tilde{\sigma}(s_{\rm exp}, s_{\rm max}, s_{\rm min}|J, N)$ against aspartate binding measurements $\sigma_{\rm exp}$ for (A) 1D Ising models and (B) Mean-field Ising models. In each panel, we show the test for $N = 1$,3, and 10. Aspartate binding measurements (●), CheY-P concentration measurements (◇), and Hill function fits (solid lines) are transformed using coupling $J = 5$. The gray shaded area shows the envelope of 95% confidence regions from Hill function fits for each N and for J in the range $[0, 10]$. Curves are identical to the MWC parametric test for $N = 1$ and qualitatively similar for $N > 1$ (compare to Fig. 1, main text). The transformed data lie well off the diagonal, indicating the presence of nonequilibrium driving.

The energy parameters have the same meanings as their MWC counterparts: $\mu = \log([L]/K_i)$ is the chemical potential (which depends on ligand concentration $[L]$ and dissociation constant K_i) and E_s is the energy difference between active and inactive states, which is increased by E_0 when the receptor is occupied. Each receptor s_i reacts individually to its occupancy σ_i, but the receptors interact cooperatively via the coupling $J_{ij} > 0$.

Integrating out the binding degrees of freedom leads to a standard Ising model with an effective field,

$$H_{\rm eff} = -\sum_{i\neq j} J_{ij} s_i s_j + h_{\rm eff} \sum_i s_i + N \log(1 + e^{\mu}), \qquad h_{\rm eff} = E_s + \log \frac{1 + e^{\mu}}{1 + e^{\mu - E_0}}. \tag{5}$$

We will consider the 1D (periodic) $J_{ij} = J(\delta_{i,j-1} + \delta_{i,j+1})/2$ and mean field $J_{ij} = J/(2N)$ Ising models. Given Eq. (5), the partition function $Z = \sum_{\{s_i\}} \exp(-H_{\rm eff})$ is well known for each of these cases, from which the binding and activity can be readily computed $\langle s\rangle = -\,\mathrm{d}\log Z\,/\,\mathrm{d}E_s$ and $\langle\sigma\rangle = \mathrm{d}\log Z\,/\,\mathrm{d}\mu$. We omit the expressions here due to their complexity.

Given expressions for $\langle s\rangle$ and $\langle\sigma\rangle$, we can derive a parametric relation between the two $\langle\sigma\rangle = \tilde{\sigma}(\langle s\rangle, s_{\rm min}, s_{\rm max}|J, N)$, similar to that obtained in the main text for the MWC model. For the Ising models, $\langle s\rangle$ is generally not invertible, so we determine the inferred occupancy $\tilde{\sigma}$ numerically for a given receptor-receptor coupling J and system size N. Fig. S2 shows the parametric test for 1D and mean field Ising models applied to the Amin and Hazelbauer measurements (1). For any choice of system size and receptor coupling, the data lie well off the diagonal, indicating these equilibrium models cannot simultaneously explain the binding and activity measurements.

Given the similarity between the parametric tests for the MWC model and both Ising models, we expect equilibrium Ising models with more complex spatial structures (e.g. finite-dimensional lattices) will also fail to resolve the kinase-binding discrepancy.

Details of the nonequilibrium allosteric model

Summary of parameters and fitting procedure for the nonequilibrium allosteric model. Here, we briefly describe the procedure of fitting both the binding and kinase activity data of the chemoreceptor. The fitting parameters are summarized in Table. S1. First, we determine the dissociation constants $K_i(m)$ by fitting the binding data to $\langle\sigma\rangle = \frac{[L]}{[L]+K_i(m)}$, which is the large E_s limit of Eq. 3 in the main text. This is consistent with previous results (1) which also found that the binding curves are well described by Hill functions with $n_H = 1$. Since binding measurements were only available for $m = 0$ and $m = 3$, we further assume the binding energy to be linear in m. Hence, the methylation-dependence of the dissociation constants takes the form $K_i(m) = K_i(0)[K_i(3)/K_i(0)]^{m/3}$. Next, we fit the kinase activity to Eq. (11) in the main text (namely, in the infinite dissipation limit) to determine E_0, E_p, ϵ_m, and A_0. Finally, we determine $G_1 = \log(k_1/k_1')$ by fitting the kinase activity in the finite dissipation limit [Eq. (10) in the main text] while keeping all the other parameters fixed. The fittings are done by minimizing the mean squared difference, normalized by the mean activity at each methylation level.

The number of fitting parameters is equivalent to previous works (4) using MWC models to describe kinase activity, with the addition of $G_{\rm ATP}$ and G_1 which do not affect the quality of fit as long as $G_{\rm ATP}$ is sufficiently large.

Parameter	Interpretation/Definition	Value
$K_i(m)$	Dissociation constants at different methylation levels	[2.69, 3.04, 3.42, 3.86, 4.35] (μM)
E_0	Binding energy difference between active and inactive receptor states	1.80
E_p	$\ln \frac{k_1}{k_z} = E_p + \epsilon_m N m$	1.03
ϵ_m		0.492
G_1	$G_1 = \log(k_1/k_1')$	13.7
A_0	Conversion factor from activity to measured CheY-P level: [CheY-P]=$A_0 \langle a \rangle$	37.7 (pmoles)

Table S1. Parameters for fitting the nonequilibrium allosteric model to experimental data from Amin and Hazelbauer (1). We fix $N = 10$, $E_s = 5$, and $G_{\mathrm{ATP}} = 20$ while fitting. Energies are in units of k_BT.

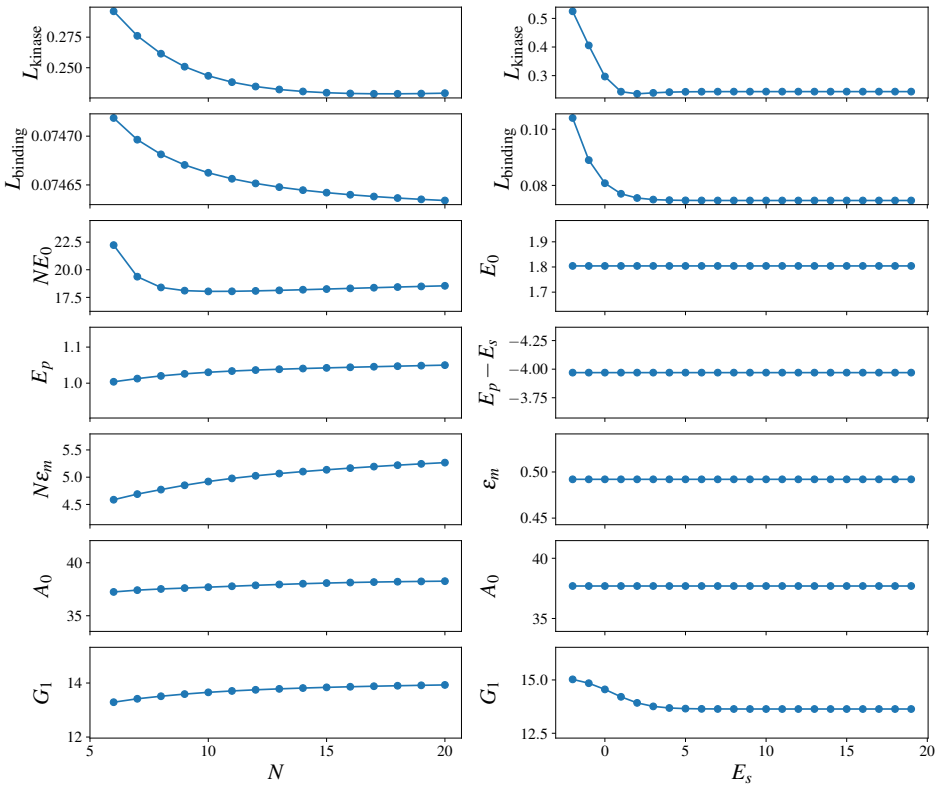

Fig. S3. The loss and fitting parameters converge at large N and large E_s. Left: dependence on N with $E_s = 5$ fixed; right: dependence on E_s with $N = 10$ fixed. $G_{\mathrm{ATP}} = 20$. L_{kinase} is the root mean squared fitting error normalized by the mean, averaged over all methylation levels. L_{binding} is the root mean squared error averaged over $m = 0$ and $m = 3$.

Dependence on the MWC cluster size N and receptor conformation energy bias E_s. Both the loss and the fitting parameters converge at large N, as demonstrated in Fig. S3 (left panel). Therefore, the results do not rely on the particular choice of N. Similarly, the fitting converges at large E_s (Fig. S3, right panel).

Effect of measuring CheY-P at finite time. In the main text, we study the steady state behavior of the nonequilibrium allosteric model, which simultaneously captures the ligand binding and kinase activity measurements by Amin and Hazelbauer (1). In these experiments, the CheY-P level was measured 10 seconds into the reaction, during which the system might have not reached the steady state. Here, we take this possibility into account and fit the data with the finite-time kinase activity of the model. The results demonstrate that taking into account the transient dynamics does not affect our main findings.

The dynamics of CheY-P concentration [CheY-P] can be described by the following equation:

$$\frac{\mathrm{d}[\text{CheY-P}]}{\mathrm{d}t} = k_p[\text{CheY}] - k_{dp}[\text{CheY-P}], \qquad [6]$$

where $[\text{CheY}] = [\text{CheY}]_{\mathrm{tot}} - [\text{CheY-P}]$ is the unphosphorylated CheY with $[\text{CheY}]_{\mathrm{tot}}$ being the total CheY concentration; $k_p = \frac{\alpha}{1+\alpha}k_1 + k_z'$ is the total phosphorylation rate (see main text), and $k_{dp} = \frac{\alpha}{1+\alpha}k_1' + k_z$ is the total dephosphorylation rate ($\frac{\alpha}{1+\alpha}$ is the fraction of time spent in the ON-state). If all CheY starts in the unphosphorylated form ($[\text{CheY-P}](t=0) = 0$), the CheY-P concentration at time t is

$$[\text{CheY-P}](t) = [\text{CheY-P}]_\infty\left(1 - e^{-(k_p+k_{dp})t}\right) = A_0\langle a\rangle\left(1 - e^{-(\tilde{k}_p+\tilde{k}_{dp})k_z t}\right), \qquad [7]$$

with $[\text{CheY-P}]_\infty$ being the steady-state CheY-P concentration and $\langle a \rangle$ being the steady-state activity given in Eq. 10 of the main text. Here, $\tilde{k}_p = \frac{k_p}{k_z}$ and $\tilde{k}_{dp} = \frac{k_{dp}}{k_z}$ are the dimensionless rates which are determined by the nonequilibrium allosteric model parameters (defined in Table. S1):

$$\tilde{k}_p = \frac{k'_z}{k_z} + \frac{\alpha}{1+\alpha}\frac{k_1}{k_z} = e^{-G_z} + \frac{\alpha}{1+\alpha}e^{E_p+\epsilon_m N m}, \quad \tilde{k}_{dp} = 1 + \frac{\alpha}{1+\alpha}\frac{k'_1}{k_z} = 1 + \frac{\alpha}{1+\alpha}e^{-G_1+E_p+\epsilon_m N m}. \tag{8}$$

The new dimensionless parameter $k_z t$ captures the transient dynamics in finite time measurements. It quantifies how far the system is from the steady state: the fit in the main text is in the steady-state limit $k_z t \to \infty$.

To investigate the effect of finite-time measurements, we fix $k_z t$ at different values and fit the measurements to Eq. 7. We find that tuning $k_z t$ across a large dynamical range has virtually no impact on the quality of fit, indicating that the conclusions are robust whether or not the CheY-P concentration has reached the exact steady state in this experiment. As shown in Fig. S4A, the fitting loss had very weak dependence on $k_z t$ with a shallow minimum near $k_z t \approx 0.84$. The corresponding ligand binding and kinase activity are shown in Fig. S4B, which has no significant improvement from the steady state fit ($k_z t \to \infty$; Fig. 2B of the main text). At the small $k_z t$, the experiments at low methylation level have not reached the steady state, while those at high methylation level have due to their high phosphorylation rate k_1.

Taking $k_z t = 0.84$ (minimum of the loss) and $t = 10$s [as in Amin and Hazelbauer (1)] leads to half decay time $\tau_{1/2} = \log 2 \cdot \frac{10s}{0.84} = 8.3s$, which is consistent with the half life for spontaneous dephosphorylation reported in Ref. (5). However, we emphasize that this is just a crude estimate since the loss minimum in $k_z t$ is very shallow and other values of $k_z t$ would also work. In order to better resolve the transient dynamics and to determine the rates, it will be useful to carry out time trace measurements of CheY-P concentration at different methylation levels and fit the results with the time-dependent model described above.

The nonequilibrium allosteric model captures binding and activity of other signaling proteins

***E. coli* oxygen-sensing protein DosP.** *E. coli* DosP is a c-di-GMP phosphodiesterase, whose enzymatic activity can be significantly enhanced by the binding of oxygen to DosP. Previous experiments measured the ligand binding and phosphodiesterase activity at different oxygen concentrations and demonstrated that the results are incompatible with an equilibrium model (6). The same study proposed that the inconsistency with the equilibrium model could be attributed to "memory effects" (6), which suggests that the system operates out of equilibrium. Here, we show that these measurements can be consistently explained under the framework of the nonequilibrium allosteric model presented in this work.

First, we carry out a parametric test to examine whether the measurements are consistent with an equilibrium MWC model with $N = 4$ [DosP are known to form tetramers (7)]. Due to the type of data available, the parametric test used here is slightly different from the one for the chemoreceptor measurements. Specifically, we infer the mean ligand occupancy $\tilde{\sigma}$ from the mean activity $\langle s \rangle$ using the MWC model:

$$\langle \sigma \rangle = \tilde{\sigma}\left(\langle s \rangle, t, [L]_{1/3}, [L]_{2/3}\right), \tag{9}$$

where $[L]_{1/3}$ and $[L]_{2/3}$ are the ligand concentration required to achieve 1/3 or 2/3 of normalized activity and $t = \frac{s_{\max}}{s_{\min}}$ is the ratio of maximum and minimum activities, which was measured to be 17 (6). The inferred occupancy $\tilde{\sigma}$ can be obtained by solving Eqs. (2) and (3) of the main text. As shown by Fig. 6A of the main text, the parametric test found all data points deviating from the diagonal $\langle \sigma \rangle_{\text{exp}} = \tilde{\sigma}$, revealing that the measurements are inconsistent with the equilibrium MWC model.

The parametric test implies that the binding and activity data cannot be simultaneously explained under the framework of the MWC model, no matter how the fitting is done. To demonstrate this, we show representative fits of the MWC model to the data in Fig. S5. When fit to ligand binding (Fig. S5A, blue line), the model incorrectly predicts the half-maximal concentration for activity (orange line). Conversely, the model does not capture the half-maximal concentration for binding when fit to the activity curve (Fig. S5B). When fit to both curves, the model predicts the correct half-maximal concentrations but completely misses the sharpness of the curves, which describes cooperativity (Fig. S5C). There exist other ways to fit the data, for example, by assigning different weights to activity and binding, respectively. Nonetheless, none of them will be consistent with the measurements as demonstrated by the parametric test in Fig. 6A of the main text.

Next, we show that the nonequilibrium allosteric model is capable of simultaneously capturing binding and activity (Fig. 6B in the main text). Here, we use a slightly generalized version of the model shown in the main text by allowing hydrolysis for both receptor conformation states $s = 0$ and $s = 1$, whose rates are given by $k_0 = k_z e^{E_{p0}}$ and $k_1 = k_z e^{E_{p1}}$, respectively. The model is shown in Fig. S6. In the large dissipation limit, the mean activity is

$$\langle a \rangle = \frac{1}{1+\mathcal{P}}, \quad \text{where } \mathcal{P} = \frac{\alpha_0 + 1}{\alpha_0 e^{E_{p1}} + e^{E_{p0}}}, \ \alpha_0 = e^{-E_s}\left(\frac{1+[L]/K_i}{1+e^{-E_0}[L]/K_i}\right)^{-N}. \tag{10}$$

The mean ligand occupancy is given by the solution to the MWC model,

$$\langle \sigma \rangle = \frac{[L]}{K_i + [L]}(1 - \langle s \rangle) + \frac{[L]}{e^{E_0}K_i + [L]}\langle s \rangle. \tag{11}$$

where $\langle s \rangle = \frac{\alpha_0}{1+\alpha_0}$ is the mean receptor activity. The model successfully fits the data as shown in Fig. 6B in the main text.

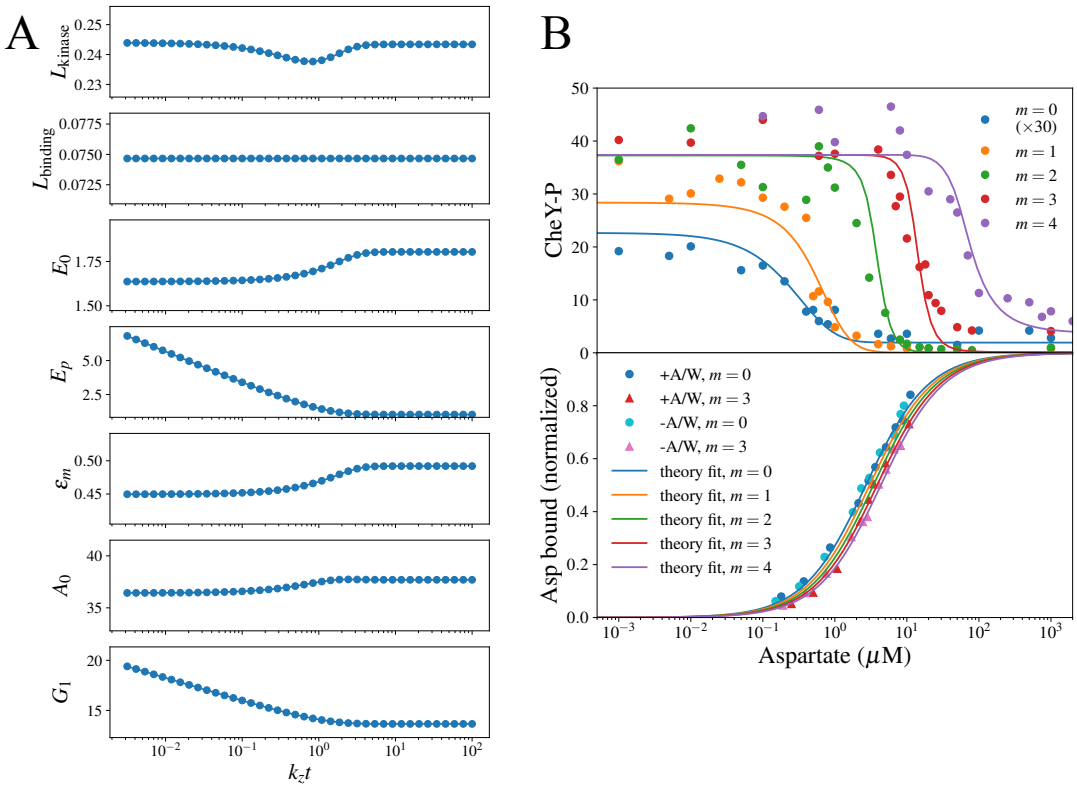

Fig. S4. Fitting CheY-P measurements at finite time does not affect the main results. (A) The best fits (loss and fitting parameters) as functions of $k_z t$. The fitting loss only depends weakly on $k_z t$. L_{kinase} is the root mean squared fitting error of the kinase activity normalized by the mean phosphorylation level, averaged over all methylation levels. L_{binding} is the root mean squared error of binding averaged over $m = 0$ and $m = 3$. (B) The activity and binding at the $k_z t$ value (0.84) that minimizes the loss. The quality of fit shows no discernible improvement from the steady-state fit in Fig. 2 of the main text. Fixed parameters: $G_{\mathrm{ATP}} = 20$, $N = 10$, $E_s = 5$.

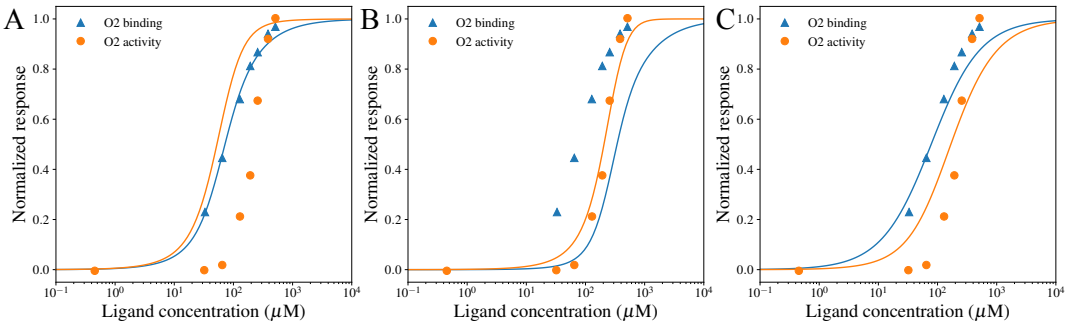

Fig. S5. The MWC model cannot capture both the ligand binding and phosphodiesterase activity response of *E. coli* DosP. (A) Fitting the binding curve results in significant inconsistency with the activity curve. (B) Fitting the activity curve results in inconsistency with the binding curve. (C) Fitting both curves simultaneously results in failure to capture their cooperativity (sharpness). For all fits, $N = 4$ and $s_{\max}/s_{\min} = 17$.

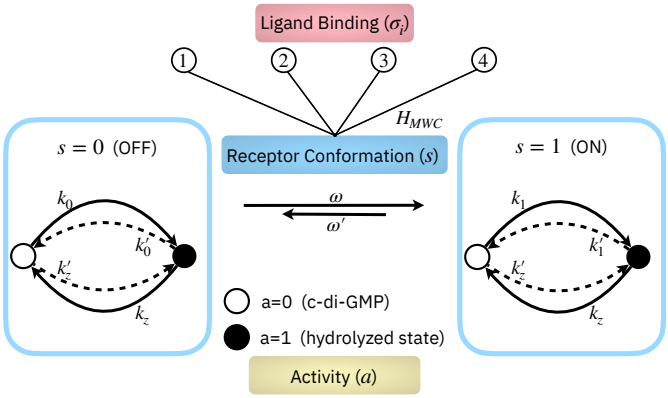

Fig. S6. The nonequilibrium allosteric model with $k_0, k_0' > 0$ (namely, c-di-GMP hydrolysis is allowed both active and inactive states). For the DosP system, $a = 0$ means c-di-GMP, and $a = 1$ stands for its hydrolyzed state. $k_{0,1}$ are the rates of c-di-GMP hydrolysis, which leads to GMP through the intermediate pGpG. k_z is the rate of c-di-GMP synthesis by diguanylate cyclases. The nonequilibrium driving is provided by GTP hydrolysis.

References

1. DN Amin, GL Hazelbauer, Chemoreceptors in signalling complexes: shifted conformation and asymmetric coupling. *Mol. Microbiol.* **78**, 1313–1323 (2010).
2. BA Mello, Y Tu, Quantitative modeling of sensitivity in bacterial chemotaxis: The role of coupling among different chemoreceptor species. *Proc. Natl. Acad. Sci. USA* **100**, 8223–8228 (2003).
3. G Lan, S Schulmeister, V Sourjik, Y Tu, Adapt locally and act globally: strategy to maintain high chemoreceptor sensitivity in complex environments. *Mol. systems biology* **7** (2011).
4. T Shimizu, Y Tu, H Berg, A modular gradient-sensing network for chemotaxis in escherichia coli revealed by responses to time-varying stimuli. *Mol. systems biology* **6** (2010).
5. V Sourjik, HC Berg, Binding of the *E. coli* response regulator CheY to its target is measured in vivo by fluorescence resonance energy transfer. *Proc. Natl. Acad. Sci. USA* **99**, 12669–12674 (2002).
6. JR Tuckerman, et al., An Oxygen-Sensing Diguanylate Cyclase and Phosphodiesterase Couple for c-di-GMP Control. *Biochemistry* **48**, 9764–9774 (2009).
7. T Yoshimura, I Sagami, Y Sasakura, T Shimizu, Relationships between Heme Incorporation, Tetramer Formation, and Catalysis of a Heme-regulated Phosphodiesterase from Escherichia coli. *J. Biol. Chem.* **278**, 53105–53111 (2003).