Quantitative modeling of sensitivity in bacterial chemotaxis: The role of coupling among different chemoreceptor species

Bernardo A. Mello*† and Yuhai Tu*‡

*IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598; and ‡Physics Department, Catholic University of Brasilia, 72030-170, Brasilia, DF, Brazil

Edited by Howard C. Berg, Harvard University, Cambridge, MA, and approved May 5, 2003 (received for review February 12, 2003)

We propose a general theoretical framework for modeling receptor sensitivity in bacterial chemotaxis, taking into account receptor interactions, including those among different receptor species. We show that our model can quantitatively explain the recent in vivo measurements of receptor sensitivity at different ligand concentrations for both mutant and wild-type strains. For mutant strains, our model can fit the experimental data exactly. For the wild-type cell, our model is capable of achieving high gain while having modest values of Hill coefficient for the response curves. Furthermore, the high sensitivity of the wild-type cell in our model is maintained for a wide range of ambient ligand concentrations, facilitated by near-perfect adaptation and dependence of ligand binding on receptor activity. Our study reveals the importance of coupling among different chemoreceptor species, in particular strong interactions between the aspartate (Tar) and serine (Tsr) receptors, which is crucial in explaining both the mutant and wild-type data. Predictions for the sensitivity of other mutant strains and possible improvements of our model for the wild-type cell are also discussed.

The bacterial chemotaxis pathway is one of the best characterized signal transduction pathways in biology (see refs. 1–3 for recent reviews of the subject). The molecular hardware of the chemotaxis response has been worked out, and we have a rather complete qualitative description of how the signal is received, transduced (to the motor), and regulated. However, at the quantitative level, there are still many unanswered questions. One long-lasting puzzle in bacterial chemotaxis is the problem of gain (4), i.e., a small change in external concentration of attractant or repellent can cause substantial change in the cell’s swimming behavior. One possible source of gain could be the interaction of the signaling molecule CheY-P with the motor complex, in particular the FlIM protein. However, despite the large Hill coefficient for the rotation bias vs. CheY-P concentration relation discovered recently in tethered single-cell experiments (5), a simple calculation quickly shows that the amplification at the motor level can only be part of the story, because quantitatively it simply cannot account for all the gain of the system (6). Therefore, there has to be significant amplification from the ligand concentration change to the change in CheY-P concentration. Recently, this high signal amplification was demonstrated directly in a set of beautiful experiments by Sourjik and Berg (SB) (7), where CheY-P concentration was measured in vivo for the first time by using fluorescence resonance energy transfer. In their study, SB measured the sensitivity of the wild-type and different mutant strains of Escherichia coli in their response to different concentrations of methyl-aspartate (MeAsp). The wild type is found to have extremely high sensitivity, which translates into a high gain of ~36 once the receptor occupancy is inferred by a simple approximation.

Because of their direct, quantitative, and systematic nature, the SB data provide us with a unique opportunity to understand the internal workings of the bacterial chemotaxis signaling pathway. In this paper, we construct a model framework capable of explaining the SB data quantitatively.

Motivation, Models, and Methods

The Motivation. Six mutant strains were studied in ref. 7. The cheR– and cheB– mutants correspond to strains with either CheR or CheB unexpressed. Four cheRcheB mutants correspond to strains missing both CheR and CheB and with the Tar receptor genetically engineered to be in different methylation states (EEEE, QEEE, QEQE, and QQQE). A careful look at the response data for all the mutant strains from the SB paper (reproduced here in Fig. 1a) reveals several suggestive features, which we describe in the following to motivate the construction of our model.

For the four cheRcheB mutants, the response curves have two steep drops as the MeAsp concentration increases: the first one corresponds to the transition of the aspartate receptor (Tar) between its ligand free state and its ligand occupied state; the second corresponds to the same transition for the serine receptor (Tsr) at a higher ligand concentration. The sizes of these two activity drops are different for different cheRcheB mutants. Although the difference for the first drops among different cheRcheB mutants could be explained by the different modification states of Tar in different mutant strains, the difference for the second activity drops is very intriguing, because the Tsr receptors should be in the same methylation state (QEQE) for all cheRcheB mutants. This observation suggests that Tsr activity is affected by Tar receptors, and the interaction depends on the methylation state of the Tsr receptors.

For the cheR– mutant, the receptors (including both Tar and Tsr) are all expected to be in their most demethylated states (EEEE) because of the unbalanced demethylation of receptors by CheB. Therefore, the Tar receptors should be in the same methylation state (EEEE) in both cheR– and cheRcheB(EEEE) mutants. However, the responses of these two strains to MeAsp are extremely different, as shown in ref. 7. The cheR– strain is found to respond to much lower ligand concentration but with much-reduced activity as compared with the cheRcheB (EEEE) strain. To understand this difference, it is important to note that even though the Tar receptors are in the same methylation states in both cheR– and cheRcheB(EEEE) strains, the Tsr receptors are in different methylation states for these two strains: one in its most demethylated state (EEEE) and the other in its unmodified state (QQQE). Therefore, the different behaviors between cheR– and cheRcheB(EEEE) could be understood if we assume again there is strong coupling between Tar and Tsr, and the interaction depends on the methylation state of the Tsr receptor.

These two observations of SB’s response data lead us to construct a theoretical model, in which one of the main ingredients is the coupling among receptors, in particular among different types of receptors. Cooperativeity between receptors as a mechanism for the system’s large gain was first suggested by Bray, Levin, and Morton-Firth (8). Later, Shi and Duke (9) proposed an Ising-type model, where identical receptors on a regular extended lattice interact with...
The half-occupancy ligand concentrations are: experimental data rescaled to their absolute values according to the activity at \( H_1 \). All energies in this paper are in units of the thermal energy \( k_B T \).

The total energy difference \( \Delta E_{qm\lambda} \) contains two different kind of contributions: a self energy \( E_{qm\lambda} \) depending on the properties of the receptor \( T_{qm\lambda} \), and an interaction energy coming from coupling to other receptors (characterized by \( q' m' \lambda' \)):

\[
\Delta E_{qm\lambda} = E_{qm\lambda} + \sum_{q' m' \lambda'} C_{q' q} f_{q' m' \lambda'} (a_{q' m' \lambda'} - 1/2).
\]

Without detailed knowledge of the geometrical nature of the interaction, we adopt the mean-field approach, where each receptor’s local environment is approximated by the overall (fractional) receptor population \( f_{q' m' \lambda'} \). Preferential coupling between different types of receptor can be captured by the coupling strength \( C_{q' q} \) between receptor types \( q \) and \( q' \). For generality, symmetric coupling \( (C_{q' q} = C_{q' q}) \) is not enforced. In principle, coupling strength may depend on the methylation levels or ligand-binding status of the receptors; however, such extra degrees of freedom are not needed for fitting the SB data.

In Eq. 2, we have used the simplest form of interaction, wherein the interaction energy depends linearly on the other receptor’s activity. The interpretation of Eq. 2 is quite transparent. The activity of a given receptor \( T_{qm\lambda} \) is determined by both its own properties characterized by \( E_{qm\lambda} \) and those of its interacting neighbors. If its neighbors’ activity is higher, then through positive coupling implemented by having \( C_{q' q} > 0 \), the total energy difference \( \Delta E_{qm\lambda} \) will be lowered, leading to higher activity for receptor \( T_{qm\lambda} \) from Eq. 1. The activity value at which the interaction is zero is set to 1/2 for convenience, and it can be shown easily that setting it to another value simply amounts to a renormalization of the local energy \( E_{qm\lambda} \).

To complete the description of the system, we need to know how each type of receptor is distributed among its different methylation and ligand-binding states, i.e., the values of \( f_{q m \lambda} \). From the experimental data, particularly the comparison between cheR+ and cheR- (EEE), it is evident that the ligand-binding kinetics may depend on the activity itself. Assuming the time scale for active/inactive transition is faster than that of the ligand-binding/unbinding reaction, the ligand-binding dynamics can be written as:

\[
T_{qm0} + L \xrightarrow{k_{qm}^{t}} T_{qm1}^{t} \quad \text{fast} \quad \uparrow \quad \text{fast},
\]

where \( T_{qm0} \) and \( T_{qm1}^{t} \) are the receptors in the inactive and active states, and \( L \) represents the ligand. The steady-state ratio between ligand-bounded and unbounded receptor populations is given by:

\[
f_{qm1} / f_{qm0} = k_{qm}^{d} (a_{qm} + 1) [L] / k_{qm}^{u} (a_{qm}) [L],
\]
where \([L]\) is the ligand concentration, and \(k_{RA}^{PB} (k_{RA}^{RB})\) are the ligand-binding (unbinding) rates for the active and inactive receptors, respectively. To include the possibility that the ligand-binding kinetics described in Eq. 3 is not an isolated process, no thermodynamics constraints are used to relate the rate constants with the receptor energy differences. However, such constraints are not expected to change the main results of this paper due to the underdetermined nature of the problem.

The Model for the Mutants. For all the mutants studied by SB, the normal methylation process is disabled, because either one or both of the methylation and demethylation enzymes is missing. Therefore, for any given mutant strain, each receptor type \(q(=1, 2)\) is in a unique methylation state \(m_1, m_2\), e.g., \(m_1 = 0\) and \(m_2 = 0\) for the cheR\(^{R}\) mutant; and \(m_1 = 2\) and \(m_2 = 2\) for the cheRcheB(QEQE) mutant. For a given mutant characterized by \(m_1/m_2\), the coupled nonlinear Eqs. 1, 2, and 4 can be solved together at any given ligand concentration to obtain the activities and receptor distribution functions \(\alpha_{m_1, m_2}\) and \(\beta_{m_1, m_2}\) from which the (average) activity of the whole system \(A = \sum_{m_1, m_2} \alpha_{m_1, m_2} \beta_{m_1, m_2}\) can be determined. In this paper, we assume that CheY-P concentration is linearly proportional to the total (kinase) activity \(A\), therefore the behavior of \(A\) from our model is directly compared with that of the (experimentally measured) CheY-P concentrations. The parameters of this model are the self-energies for each receptor type \(E_{m_1, m_2}\), the coupling strengths \(C_{m_1, m_2}\), and the kinetic rates for describing the ligand-binding kinetics. To validate our model quantitatively, our goal is to find a set of parameters such that the behaviors of the model would agree with the experimental data for all six mutant strains in ref. 7.

The Model for the Wild Type. To model the response behavior of the wild type properly, we need to introduce the methylation/demethylation kinetics. At a given ambient ligand concentration, the receptors in a wild-type cell are distributed in different methylation states due to adaptation mediated by the methylation/demethylation kinetics. To make this a tractable problem, we use a coarse-grained description for the Tar receptors with only five methylation states characterized by their total number of methyl groups \(m \in [0, 4]\) instead of considering all 16 possible methylation states. Hence \(m\) has a slightly different meaning here from that in the mutant model. For example, although parameters with \(m = 1\) in the wild-type model represent the average properties of all receptors with one methylated site, i.e., (QEEE), (EQEE), (EQOE), and (EEEQ), they describe the properties of the specific methylation state studied in the current mutant experiments, i.e., (QEEO). Therefore, the parameters from these two models may not be the same, and we choose to fit the mutant and wild-type data separately.

It is known that \(E. coli\) adapts almost perfectly to MeAsp, i.e., the steady-state total activity is independent of the ambient ligand concentration. As shown in previous studies on adaptation (11–14) one of the key conditions for achieving perfect adaptation is that the methylation/demethylation rates depend on the activity of the receptor. Assuming that only inactive receptors can be methylated, and only active receptors can be demethylated (12, 14), we use the following simplified methylation flux balance equation between any two consecutive methylation levels to determine the steady-state distribution of the receptors in different methylation states:

\[
k_{RB} \sum_{\lambda} (1 - a_{1m1}) f_{1m1} = \sum_{\lambda} a_{1m1+1} f_{1m1+1} \quad m = 0, 1, 2, 3, 4
\]

\(k_{RB}\) is the ratio of the methylation/demethylation rates, which depends on the enzyme concentrations and methylation kinetic constants. It is easy to show by summing Eq. 5 over \(m\) that the total activity of the Tar receptors, \(A_1 = \sum_{m=0}^{m=4} f_{1m1} a_{1m1}\) is a constant if the fully methylated receptors are always active, \(a_{14} = 1\), and the least methylated receptors are always inactive, \(a_{10} = 0\). This is also one of the conditions for perfect adaptation found in our earlier work by studying a full adaptation model (14).

The Tsr receptor distribution in different methylation states, \(f_{2m}\), can be determined similarly as in Eq. 5. Because of perfect adaptation and the fact that Tsr does not bind to MeAsp in the concentration range (<10 mM) where the wild-type experiments were performed, the steady-state Tsr distribution \(f_{2m}\) is independent of the ambient MeAsp concentration. To avoid introducing too many parameters describing the details of the different Tsr methylation states, we simplify our model by approximating \(f_{2m} = \delta_{m0}\) (\(m_0\) is set to 2 for bookkeeping purposes), i.e. all Tsr receptors are in the same methylation state. This is a reasonable approximation for studying the MeAsp response, because \(f_{2m}\) is independent of the ambient MeAsp concentration, and \(f_{2m}\) is usually dominated by a single methylation level (see supporting information on the PNAS web site, www.pnas.org, for Tsr distribution \(f_{2m}\)).

Eq. 5, together with Eqs. 1, 2, and 4, defines the steady-state receptor distributions (in different methylation states) of the wild-type cell for any given ambient MeAsp concentration. Because methylation/demethylation is much slower than the ligand-binding process, the short time response of the wild-type cell to different added/removed MeAsp concentrations can be calculated by solving just Eqs. 1, 2, and 4 for the fast processes, while keeping the receptor distribution in different methylation states the same as that of the (prestimulus) steady state. These responses can be directly compared with the wild-type data provided in SB’s study.

The types of parameters in the wild-type model are similar to that of the mutant, except that there is only one Tsr state in the wild-type model. In addition, we have introduced an extra parameter \(k_{RB}\) for describing the methylation/demethylation processes.

The Fitting Method. We construct an error function by summing the square of the difference between the model’s predicted response and the experimental value at each data point. A global scaling factor is also introduced as a fitting parameter in converting the model’s predicted activity into the experimentally observed CheY-P concentration. The Newton–Raphson method (15) is used for solving the self-consistent equations for a given set of parameters, and Powell’s quadratically convergent method (15) is used subsequently in minimizing the error function in the parameter space with different initial parameter values.

Results

The Mutant Strains. Because there is not enough range of activity change for the Tar receptor in \(m = 1, 2, 3\), and 4 states in the experimental data, the ligand-binding/unbinding rates of Tar for \(m \in (1, 4)\) are underdetermined; i.e., there are many sets of ligand-binding/unbinding rates that can produce the same good fit to the data. However, for Tar \((q = 1)\) with \(m = 0\), the best fit to the data always results in \(\kappa_{0m} < \kappa_{1m}\) and \(\kappa_{1m} \gg \kappa_{0m}\). To eliminate the unnecessary degeneracy in our fitting, we have reduced the number of parameters by fixing \(\kappa_{0m} = 0\) and \(\kappa_{1m}\) for all receptor and methylation levels for both the mutant and wild-type models. However, it is important to keep in mind for future studies that this requirement may be necessary only for Tar with \(m = 0\). Now the ligand-binding/unbinding balance equation can be simplified into \(\kappa_{2m0} f_{2m0} = (1 - \delta_{0m}) [L] \gamma_{RB} k_{RB}^{PB} (f_{2m0})\), which depends on only one parameter, \(\kappa_{RB} = k_{RB}^{PB} / f_{0m}^{PB}\), instead of the three ratios between the four kinetic rates in Eq. 4.

From all the initial parameters we tried, the fitting converges to the same point in the parameter space. However, the attraction
basin is very flat, i.e., there are many sets of parameters that can fit the mutant data within the experimental error. Fig. 1a shows one such good fit of our model to the experimental data, the corresponding parameters being given in Table 1 (see supporting information for examples of other good fits). Overall, the model fits the data remarkably well, considering that the number of parameters is rather small. Although more experiments are needed to remove the ambiguity in the parameters, some parameters of the model do seem to be constrained by the existing data, e.g., the interaction strength from Tsr to Tar, $C_{12}$, is found to be always big as compared with other coupling constants.

By using our model, we can also directly measure the fractional ligand occupancy of both the Tar and Tsr receptors, the results being shown in Fig. 1b. Comparing Fig. 1a and b, it is clear that even though the trends of ligand occupancy and the activity curve for a given receptor are correlated, they can be quite different quantitatively. Due to the effect of receptor coupling, high ligand occupancy does not necessarily mean proportionally large suppression of activity. Depending on the details of the model, $K_{1/2}$ (the half-maximum ligand concentration) of the ligand occupancy curve can be either bigger or smaller than that of the corresponding activity curve, as shown in Fig. 1.

### The Wild Type

The fitting for the wild-type data is significantly more difficult than for the mutant data, because of the complexity introduced by multiple Tar methylation states. The best fit to the wild-type data we have found so far within our model is shown in Fig. 2a, with the parameters given in Table 2. There is no guarantee that this is the unique, or globally minimum, solution. The general behavior of the response of the wild-type cell to increasing and reducing ligand concentration predicted by our model is highly consistent with the experimental results. Due to the simplifications made in our wild-type model, the quantitative agreement is not as perfect as for the mutant studies, especially for the large MeAsp concentration changes. However, for a relatively small MeAsp concentration change, our model has captured the essential features for producing high sensitivity and high gain in the wild-type response. At a given ambient ligand concentration $[L]$, the system is characterized by its receptor occupancy $L$ and total activity $A$. On a sudden change of ligand concentration $\Delta [L]$, the resulting (fast) changes (before methylation/demethylation takes place) in receptor occupancy and total activity are represented by $\Delta L$ and $\Delta A$. Following SB (7), sensitivity ($S$) and gain ($G$) are defined as:

$$ S = (\exp(\Delta L)) \Delta A / (\Delta L) $$

$$ G = (1/4) \Delta A / (\Delta L). $$

While $S$ measures the sensitivity of the cell’s kinase activity with respect to the (external) ligand concentration change, $G$ measures the signal amplification within the cell.

In Fig. 2b, we show the sensitivity for the full range of ambient ligand concentrations, in the same way as in ref. 7 (see figure 3b in their paper) for small relative MeAsp concentration changes $\Delta [L]/[L] = 1\%, 10\%, 20\%$. For $>4$ decades of ambient ligand concentrations, high sensitivity is maintained between 1 and 8, consistent with ref. 7. In Fig. 2b, the steady-state activity of our model is also plotted. The system adapts nearly perfectly, i.e., the steady-state activity is a constant independent of ambient ligand concentration, from zero concentration to $\sim 1$ mM. The adaptation becomes imperfect for higher ligand concentrations (75% activity at 10 mM), because one of the perfect adaptation conditions $a_{13} = 1$ is not satisfied, and the $m = 4$ state becomes highly populated at high ligand concentrations $\Delta A_T$. To determine the gain, we have plotted the relative activity change $\Delta A_T / \Delta L$ vs. the change in receptor occupancy $\Delta L$ for several ambient ligand concentrations. As shown in Fig. 2c, most curves for different ambient MeAsp concentrations collapse onto a single one (as in ref. 7). The deviation from common behavior at the extremely high ambient MeAsp concentration ($5$ mM) is probably caused by violation of perfect adaptation at such high MeAsp concentration in our model. For small $\Delta L$, the relative activity change is linearly proportional to the receptor occupancy change with a high slope ($\sim 20$). In SB’s paper (7), receptor occupancy of the wild-type cell was estimated from the data of a single mutant strain; this inaccuracy in determining the receptor occupancy may be responsible for the discrepancy between the gain calculated in our model ($\sim 20$) and that reported from their work ($\sim 36$).

So, where does the high gain come from? In our model, we can easily calculate the contributions to the gain by Tar and Tsr separately. For the parameters in Table 2, the gain from Tsr constitutes a large fraction ($\sim 80\%$) of the total gain for MeAsp response (see supporting information for a detailed derivation). Quantitatively, the gain from Tsr at a given ambient concentration is determined mainly by two factors: the coupling strength from Tar to Tsr, i.e., $C_{21}$; and Tsr’s activity at the ambient ligand concentration $A_T$, which determines the susceptibility of the Tsr activity with respect to change of its neighbor’s activity. From Table 2, we can see that a large value of the coupling constant $C_{21}$ is required to fit the experimental data. Another reason for the high gain from Tsr is that in our model, the Tsr activity $A_T$ is kept around $1/2$, where Tsr activity is most susceptible to the change of its neighbor’s activities, and the perfect adaptation of the system is responsible for keeping the Tsr receptors at their most susceptible states for all ambient ligand concentrations. This observation provides another mechanism in which perfect (or near-perfect) adaptation may be required for the

### Table 1. A set of parameters for which our model fits all the mutant strain data well

<table>
<thead>
<tr>
<th>$m$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{m1}(A_{m0}^*)$</td>
<td>-2.07 (0.89)</td>
<td>-2.51 (0.93)</td>
<td>-2.93 (0.95)</td>
<td>-15.4 (1.0)</td>
<td>-16.0 (1.0)</td>
</tr>
<tr>
<td>$E_{m1}(A_{m1}^*)$</td>
<td>6.22 (0.0)</td>
<td>-1.31 (0.79)</td>
<td>-1.77 (0.85)</td>
<td>-2.27 (0.91)</td>
<td>-2.27 (0.91)</td>
</tr>
<tr>
<td>$E_{m2}(A_{m0}^*)$</td>
<td>$\infty$ (0.0)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$E_{m2}(A_{m1}^*)$</td>
<td>-</td>
<td>-</td>
<td>13.2 (0.0)</td>
<td>-</td>
<td>12.9 (0.0)</td>
</tr>
<tr>
<td>$K_{1m}$</td>
<td>185681</td>
<td>77.5</td>
<td>32.4</td>
<td>0.00018</td>
<td>0.00111</td>
</tr>
<tr>
<td>$K_{2m}$</td>
<td>$\infty$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1. A set of parameters for which our model fits all the mutant strain data well

Next to the self-energy differences $E_{mnm}$, we have included the “bare” activity values $A_{nm}^0 = (1 + \exp(E_{mnm}))^{-1}$ for reference only; the actual receptor activities are always different from their bare values because of receptor coupling.

---

1 Due to the dominance of the extremely large self energy difference for the Tar receptors in our model, large coupling constants do not lead to the “All-or-None” phase transition, as in the standard Ising model.

4 If Tsr distribution in different methylation states is taken into account, the gain will decrease but not by much if distribution is centered around the most susceptible methylation state.
receptor occupancy. For small changes in receptor occupancy, the gain is defined as the ratio between the fractional activity change and the change in decreasing and increasing ligand concentrations, respectively. The gain is initial (ambient) ligand concentrations. The upper and lower curves are for fractional change in activity vs. change in (total) ligand occupancy for different initial (ambient) MeAsp concentrations. Different curves refer to different initial (ambient) MeAsp concentrations (see supporting information for details). Finally, large GL, together with high gain G, leads to persistently high sensitivity.

**Summary and Discussion**

In trying to understand and model recent in vivo experiments on bacterial chemotaxis response by SB, we discover that receptor coupling, in particular among different types of receptors, and dependence of ligand binding on receptor activity are the two most important ingredients in explaining their data. For the mutant data, these two ingredients are essential in explaining some of the intriguing findings in ref. 7, in particular the different activities of the same Tsr methylation state in different cheR-cheB mutants and the significant difference in both activity and sensitivity of the two mutant strains, cheRcheB(EEEEEE) and cheR−. From fitting the wild-type data, the results from our model confirm that the large gain can indeed be explained by receptor coupling quantitatively. More specifically, our model suggests that most of the gain in response to MeAsp could come from Tsr. Even though Tsr does not bind to MeAsp directly except for extremely high MeAsp concentration (>10 mM), its activity is indirectly affected by the change of MeAsp concentration because of its coupling to Tar. The strong coupling between Tar and Tsr within the trimer of dimer complex, as recently discovered by Ames et al. (16). The same sort of large indirect contribution of gain from Tsr (or Tar) may also exist in the response of the system to ligands that bind directly with low-abundance chemoreceptors. Finally, we find that, besides the high gain, the tracking of effective KD with ambient ligand concentration through adaptation and the dependence of ligand binding on receptor activity are crucial in maintaining the system’s high sensitivity.

The wild-type and mutant data are fitted separately in this study, mainly due to the different resolutions used in describing the methylation states for these two data sets. With more experimental data becoming available, the ultimate goal is to find a single set of parameters that can explain both the mutant and the wild-type data within our model. For example, the heterogeneous coupling constants are found to be large (C12 from the mutant study; C21 from the wild-type study); the challenge is whether a unified theory with large values for both C12 and C21 can explain both data sets. Despite the large system to have high gain over a large range of ambient ligand concentrations.

What causes the high sensitivity? From its definition, we can easily relate the sensitivity to the gain:

\[
S = \frac{L}{A} \frac{\Delta A}{\Delta [L]} = \frac{GL}{\Delta[L]} \times G = GL \times G,
\]

where GL = \([L]/\Delta[L]\) represents how receptor occupancy changes in response to (relative) ligand concentration change. From Eq. 6, it is clear that high gain alone does not guarantee high sensitivity unless GL is also kept consistently high for the same range of \([L]\). To have a sense of how GL behaves, we consider the case where ligand binding obeys simple kinetics with a single dissociation constant \(K_d\) for all Tar receptors. In this case, GL reaches its maximum value: max(GL) = 1/12 at \([L] = K_d\) and decays rapidly as \([L]\) moves away from \(K_d\). Fortunately, in bacterial chemotaxis, the receptors exist in different methylation states with different dissociation constants. Due to adaptation, as the ambient MeAsp concentration increases, the Tar receptor population shifts toward higher methylation states with higher dissociation constants; consequently the effective \(K_d\) for all Tar receptors increases with \([L]\). In addition, the dependence of ligand binding on receptor activity, as assumed in our model, increases max(GL), the maximum value of GL, for each individual methylation state. Therefore, the dependence of ligand binding on receptor activity increases GL, and the tracking of the effective \(K_d\) with the ambient ligand concentration keeps the high GL value through a wide range of ambient ligand concentrations (see supporting information for details). Finally, large GL, together with high gain G, leads to persistently high sensitivity.
difference between the mutant and wild-type parameters in the specific solutions shown in this paper (Tables 1 and 2), a unique set of parameters is still possible because of the existence of many equally good solutions when each data set is fitted separately. However, it may also turn out that extra mechanistic ingredients, such as the dependence of the coupling constants on the methylation level, have to be included in the model to explain all the data. Before such a unified theory can be attempted, we need to improve our wild-type model to achieve the same level of quantitative agreement between model and experiment as for Tar receptors. Furthermore, the mean-field-type model used in this paper (Eq. 2) needs to be improved to incorporate the appropriate spatial (maybe even temporal) information on receptor interaction once we understand whether the receptor interaction is mediated by a spatially extended lattice of receptors (17), just by the trimer of dimer complex (18), or through a dynamically changing network of interacting receptors.

In this paper, we have used the in vivo data of ref. 7 for constructing and validating our model. There are several other related recent experimental studies (both in vitro and in vivo), which are interesting to examine in light of our model. Bornhorst and Falke (19) measured the activities of all 16 methylation states of Tar receptors in vitro. Qualitatively, the concentration that induces half-maximal inhibition ($K_{1/2}$) is found to be highly correlated with the activity of the receptor, which is consistent with our model. It is also interesting to note that the $K_{1/2}$ for the same methylation states have different values in refs. 7 and 19, further evidence that the sensitivities of the receptors may depend on receptor interactions and therefore may be different in different systems. The Hill coefficients for the response curves are all found to be between 1 and 3 for Tar receptors (19) but as high as 8–10 for the Tsr receptors (20). In our model, the Hill coefficient depends on interaction strengths among receptors. Therefore, these experiments can be understood within our model, because the interaction strengths can be different for Tar–Tar and Tsr–Tsr interaction, which can also be different from their value in wild-type cells. In a related study, Levit and Stock (21) measured directly the receptor occupancy together with the activity and found that they can be quite different, far from the linear relations that are often assumed. This is again consistent with our model, as seen from the comparison of Fig. 1 a and b. Recently, by using a synthesized multivalent chemoattractant for Trg, Gestwicki and Kiessling (22) discovered a large increase in sensitivity of the cell’s response not only to the synthesized chemoattractant itself but also to serine. This is in support of the notion that the sensitivity of ligand binding is affected by receptor interactions, which might be enhanced by the multivalent chemoattractant.

Specific experiments may also be suggested for validating our model. For example, in our model, the large difference between the behaviors of cheR and cheRcheB(EEEE) mutants is explained by the difference of the Tsr methylation states for these two mutant strains. Another possible scenario is that CheB could inhibit receptor activity, as proposed by Barkai et al. (23). A possible test could be to make a cheRcheB mutant with both Tar and Tsr in their lowest methylation state (EEEE) and measure its response to MeAsp. According to our model, the response to MeAsp for this new mutant should behave more like the cheR strain because of the similarity in the methylation state for the two major receptors, whereas the other theory would predict it to be more like cheRcheB(EEEE), because both of them have no cheB.

We thank Drs. H. Berg and V. Sourjik for explaining their experiments in detail. We also thank Dr. G. Grinstein for a careful reading of the manuscript. The work of B.A.M. is partially supported by a fellowship from the National Council for Scientific and Technological Development of Brazil.

Table 2. A set of parameters for which our model fits the wild-type data well

<table>
<thead>
<tr>
<th>$m$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{m0}(A_{en0}^{\text{m}})$</td>
<td>(0.0)</td>
<td>-12.4 (1.0)</td>
<td>-11.5 (1.0)</td>
<td>-16.0 (1.0)</td>
<td>-30.5 (1.0)</td>
</tr>
<tr>
<td>$E_{m1}(A_{en1}^{\text{m}})$</td>
<td>(0.0)</td>
<td>15.8 (0.0)</td>
<td>14.3 (0.0)</td>
<td>11.0 (0.0)</td>
<td>11.0 (0.0)</td>
</tr>
<tr>
<td>$E_{m2}(A_{en2}^{\text{m}})$</td>
<td>-</td>
<td>-</td>
<td>-1.76 (0.85)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$K_{1m}$</td>
<td>-</td>
<td>486</td>
<td>1679</td>
<td>12.4</td>
<td>$2.32 \times 10^{-4}$</td>
</tr>
<tr>
<td>$K_{2m}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The methylation/demethylation ratio: $k_{8a} = 0.26$.

Nonuniqueness of the parameters. As described in the main text, parameters for our model cannot be uniquely determined from the existing mutant data (the same is true for the WT model). In other words, there are many parameter sets that can be used in our model to fit the experimental MeAsp response data for all the mutant strains within the experimental error. For example, if the local energy differences for the occupied Tar receptors $E_{1m1}$ are decreased, the effects can be absorbed by an increase of the local energy differences $E_{2m0}$ for the unoccupied Tsr receptors. In Table 3, we list such a new set of parameters that are different from those given in Table 2. The fit of our model with this new set of parameters is shown in Figure 4, where the fits are as good as in Fig. 1, and the differences between the fit and the experimental data are well within the experimental error. Such ambiguity in parameters can be fixed only by having more mutant experiments, in particular where the Tsr receptor is also engineered to be in different methylation states (or in the presence of serine), resulting in different activities. However, some parameters of the model seem to be more tightly constrained by the existing data, e.g., the coupling strength $C_{12}$ is always large because the large influence of Tsr on Tar activity/sensitivity is needed to explain the large difference between the cheR− and cheRcheB(EEEE) strains; the interaction strength $C_{11}$ between Tar receptors themselves, is found to be small (practically zero in table 3). This may be related to the fact that all response curves are found to have a rather modest Hill coefficient, which could be true if the Tar receptor interacts preferably with Tsr in the presence of the Tsr receptors.

The gain (in response to MeAsp) due to Tsr. Although Tsr does not respond (bind) directly to MeAsp except at extremely high concentrations of MeAsp, Tsr can contribute to the system’s gain through its coupling to the Tar receptors. Within our model, we can estimate the contribution of gain from the Tsr receptors as compared to that from the Tar receptors. For our WT model, the total energies and therefore the activities from the Tar and Tsr receptors can be determined from the following coupled equations:
\[ \Delta E_{1m\lambda} = E_{1m\lambda} + C_{11}(A_1 - \frac{1}{6}) + C_{12}(A_2 - \frac{1}{3}) \]  
(7)
\[ \Delta E_2 = E_2 + C_{21}(A_1 - \frac{1}{6}) + C_{22}(A_2 - \frac{1}{3}) \]  
(8)
\[ A_2 = \frac{2}{3}(1 + \exp(\Delta E_2))^{-1} \]  
(9)
\[ A_1 = \sum_{m\lambda} f_{1m\lambda} a_{1m\lambda}. \]  
(10)

Summing the methylation balance equations (Eq. 5) over \( m \), and assuming that the Tar receptor population in the \( m = 4 \) state is small, we obtain the total activity due to the Tar receptors:

\[ A_1 \approx \frac{1}{3} \times \frac{k_{RB}^1}{1 + k_{RB}^2}. \]  
(11)

The total activity due to Tsr receptors, \( A_2 \), can then be obtained by Eqs. 8, 9 and 11.

For a small change in ligand concentration \( \Delta[L] \): \( [L] \rightarrow [L] + \Delta[L] \), the corresponding (fast) change (before methylation takes place) in receptor occupancy is represented by \( \Delta L \), and activity changes for the two receptor types are \( \Delta A_1 \) and \( \Delta A_2 \), respectively. From Eqs. 8 and 9, we can relate the activity changes:

\[ \Delta A_2 = -\frac{2}{3} \chi \Delta \Delta E_2 \]  
(12)
\[ \Delta \Delta E_2 = C_{21} \Delta A_1 + C_{22} \Delta A_2, \]  
(13)
where \( \Delta \Delta E_2 \) is the change in total energy difference for the Tsr receptor caused by the change in ligand concentration, and

\[ \chi = \exp(\Delta E_2)/(1 + \exp(\Delta E_2))^2 \]  
(14)

is the measure of susceptibility of Tsr activity with respect to change in its total energy difference, and \( \Delta E_2 \) is the Tsr energy difference before the stimulus.

Solving Eqs. 12 and 13, we obtain the ratio between the changes in activities due to Tsr and Tar, respectively:

\[ \frac{\Delta A_2}{\Delta A_1} \equiv g_{21} = -\frac{2\chi C_{21}}{3 + 2\chi C_{22}}. \]  
(15)

The total gain of the system (as defined by Sourjik and Berg[1]) is then:

\[ G = \frac{1}{A} \frac{\Delta A_1 + \Delta A_2}{\Delta L} = \frac{1}{A} \frac{\Delta A_1}{\Delta L} (1 + g_{21}), \]  
(16)
where it is clear that Tsr’s contribution to the total gain is $g_{21}$ times that of the contribution from Tar.

It is also clear from Eq. 15 that $g_{21}$, i.e., the relative contribution to gain from Tsr is larger when either $|C_{21}|$ or $|C_{22}|$ or $\chi$ is large. The dependence on $C_{21}$ makes perfect sense, because Tsr contributes indirectly to the gain through its coupling with Tar. Quantitatively, the gain from Tsr also depends on the coupling between Tsr receptors themselves, and how susceptible Tsr activity is to changes in its interaction energy with Tar, measured by $\chi$.

For the parameters shown in Table 2, the gain from Tar can be easily estimated. Because the self energies dominate the total energy differences in this case, i.e., $|E_{1m\lambda}| \gg 1$ for all methylation levels except $m = 0$, for the Tar receptor, the activity is either 0 or 1 depending on whether the Tar receptor is ligand bound. Therefore, for the case where the Tar receptor population in $m = 0$ is small, which is true for most of ambient MeAsp concentrations, we have:

$$\Delta A_1 \approx \Delta L. \quad (17)$$

For the parameters in Table 2, the steady-state total activity is nearly constant (near perfect adaptation): $A \approx 0.4$ (see Fig. 2a). Therefore, the gain from Tar receptor:

$$G_1 \equiv \frac{1}{A} \frac{\Delta A_1}{\Delta L} \approx 2.5, \quad (18)$$

which is much smaller than the total gain. The main contribution to the total gain comes from Tsr. The large contribution from Tsr, i.e., the large value of $g_{21} \approx 7$, is a result of large coupling constants $C_{21} = -17.9$. More interestingly, a large value of susceptibility $\chi$, which is at its maximum value of $\frac{1}{4}$ (see Eq. 14). Such maximum susceptibility of the Tsr receptor is maintained through all ambient MeAsp concentrations by having the Tsr total energy difference $\Delta E_2$ fixed at approximately 0, which is possible only because the system adapts (nearly) perfectly. If we take into account all possible methylation states of Tsr, the average susceptibility would be smaller, but not by much if the distribution is centered around the methylation state with the maximal susceptibility. For example, a 50% decrease in susceptibility for half of the Tsr population will still result to only 25% reduction in the average susceptibility. More importantly, this slightly reduced, yet high gain will still be maintained through all ambient MeAsp concentrations by perfect adaptation. Because of this large value of $g_{21}$, the total gain, estimated as $G = G_1 \times (1 + g_{21}) \approx 20$, is also very large, which is consistent with our numerical result, shown in Fig. 2c. Our estimate of the
gain and values of its different sources $G_1$ and $G_2$ is confirmed by direct simulation of our model, as shown in Fig. 4.

Eq. 15 also gives us a sense of the strength of the coupling constants in terms neighboring interactions. Without Tsr-Tsr coupling, the total amount of Tsr activity suppressed by the binding of a MeAsp molecule to a single Tar receptor can be estimated as $-2\chi C_{21}/3 \sim 3$. Considering the maximal change of activity for Tsr is $1/2$, this means about six Tsr receptors are directly affected by a single Tar receptor. The rest of the gain come from the secondary Tsr-Tsr interaction, with a total of $\approx 14$ Tsr receptors being affected.

It is quite clear from our analysis that for parameters such as those listed in Table 2, the total gain depends strongly on $C_{21}$ and $C_{22}$ and is insensitive to the other coupling constants $C_{11}$ and $C_{12}$. Indeed, we can find other parameter sets that have different values of $C_{11}$ and $C_{12}$ from that of table 2, and our model still retains the same high gain and fits the experimental data as well as our model with parameters in Table 2 (data not shown). This, however, does not exclude the possibility that there exist other types of parameter sets, where most gain in response to MeAsp comes from other sources, such as interaction between the Tar receptors alone, so far, however, we have not found any alternative types of parameter sets through our fitting program by using different initial starting parameters.

**The two main ingredients for persistent high $G_L$.** $G_L$ can be written as a sum over the contributions from different methylation states:

$$G_L = \sum_{m \in [1,4]} f_{1m}([L])G_L^{(m)}([L]), \quad (19)$$

where $G_L^{(m)}([L]) = [L] \Delta L_m \Delta[L]$, $L_m$ is receptor occupancy for Tar receptor in the methylation state $m$. In Fig. 5, we show the dependence of $G_L^{(m)}$ for $m \in [1,4]$ and $G_L$ on the ambient ligand concentration $[L]$ with the parameters in Table 2. In Fig. 6, we have plotted the Tar receptor distribution in different methylation states, i.e., $f_{1m}([L])$ for $m \in [0,4]$, together with the total and individual ligand occupancy. As can be seen from Fig. 5, as $[L]$ increases, the dominating $G_L^{(m)}$ shifts toward larger $m$, with the same trend as the receptor population as shown in Fig. 6. This tracking mechanism optimizes the product in Eq. 19 and therefore keeps $G_L$ at high values. However, the matching (tracking) is not exact, because the positions of the peaks for the population and $G_L^{(m)}$ are not at the same $[L]$.

On top of the “tracking” mechanism, another important ingredient in keeping $G_L$ large is the fact that ligand binding is affected by activity. If one assumes simple ligand-binding
kinetics described by just a dissociation constant $K_d^m$ for each methylation level $m$, then the maximum of $G_L^{(m)}$ will be $\frac{1}{4}$ reached at $[L]_0 = K_d^m$. However, for $G_L^{(m)}$ shown in Fig. 5, they are much bigger than $\frac{1}{4}$, which is caused by the fact that the effective $K_d$ in our model depends on the activity.

<table>
<thead>
<tr>
<th>$m$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{1m0}(A_{1m0}^B)$</td>
<td>0.511(0.38)</td>
<td>0.117(0.47)</td>
<td>-0.310(0.58)</td>
<td>-7.10(1.0)</td>
<td>-22.4(1.0)</td>
</tr>
<tr>
<td>$E_{1m1}(A_{1m1}^B)$</td>
<td>15.1(0.0)</td>
<td>1.39(0.2)</td>
<td>0.939(0.28)</td>
<td>0.535(0.37)</td>
<td>0.53(0.37)</td>
</tr>
<tr>
<td>$E_{2m0}(A_{2m0}^B)$</td>
<td>$\infty$(0.0)</td>
<td>–</td>
<td>-0.771(0.68)</td>
<td>–</td>
<td>-0.913(0.71)</td>
</tr>
<tr>
<td>$E_{2m1}(A_{2m1}^B)$</td>
<td>–</td>
<td>–</td>
<td>14.1(0.0)</td>
<td>–</td>
<td>7.68(0.0)</td>
</tr>
<tr>
<td>$K_{1m}$</td>
<td>7.06 $\times 10^7$</td>
<td>103</td>
<td>49.0</td>
<td>0.118</td>
<td>8.16</td>
</tr>
<tr>
<td>$K_{2m}$</td>
<td>$\infty$</td>
<td>–</td>
<td>1.91 $\times 10^{11}$</td>
<td>–</td>
<td>2.84 $\times 10^8$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$C_{qq'}$</th>
<th>$q' = 1$</th>
<th>$q' = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q = 1$</td>
<td>-0.0349</td>
<td>-5.26</td>
</tr>
<tr>
<td>$q = 2$</td>
<td>-7.0</td>
<td>-0.00066</td>
</tr>
</tbody>
</table>