An allosteric model for heterogeneous receptor complexes: Understanding bacterial chemotaxis responses to multiple stimuli

Bernardo A. Mello*† and Yuhai Tu*‡

*IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598; and †Department of Physics, Catholic University of Brasiliam, 71966-700, Brasilia-DF, Brazil

Edited by Howard C. Berg, Harvard University, Cambridge, MA, and approved September 30, 2005 (received for review August 11, 2005)

The classical Monod-Wyman-Changeux model for homogeneous allosteric protein complex is generalized in this article to model the responses of heterogeneous receptor complexes to multiple types of ligand stimulus. We show that the recent *in vivo* experimental data of *Escherichia coli* chemotaxis responses for mutant strains with different expression levels of the chemo-receptors to different types of stimulus [Sourjik, V. & Berg, H. C. (2004) *Nature* 428, 437–441] all can be explained consistently within this generalized Monod-Wyman-Changeux model. Based on the model and the existing data, responses of all of the strains (studied in this article) to the presence of any combinations of ligand (Ser and MeAsp) concentrations are predicted quantitatively for future experimental verification. Through modeling the *in vivo* response data, our study reveals important information about the properties of different types of individual receptors, as well as the composition of the cluster. The energetic contribution of the nonligand binding, cytoplasmic parts of the cluster, such as CheA and CheW, is also discussed. The generalized allosteric model provides a consistent framework in understanding signal integration and differentiation in bacterial chemotaxis. It should also be useful for studying the functions of other heterogeneous receptor complexes.

cooperativity | signal transduction

Most biological functions are carried out by multiprotein complexes instead of a single protein molecule. The combinatorics of possible protein complex give the cell tremendous flexibility and specificity in its functions. Receptors have been found to form aggregates in a wide range of systems, including T cell receptors in the immune system (1), ryanodine receptors in skeletal and cardiac muscle cells (2), tumor necrosis factor receptors regulating apoptosis (3), and neurotransmitter receptors in neurons (4). The cooperative interaction within these receptor complexes is crucial in controlling sensitivity in different signaling pathways (5). In bacterial chemotaxis, for example, the methyl-accepting chemotaxis proteins (MCPs) are the membrane-bound chemo receptors, to which the stimulus ligand can bind. The MCPs are observed to form large clusters near the cell pole with other cytoplasmic proteins, in particular CheA, a histidine kinase, and CheW, a linker protein (6). The main function of the MCP cluster is to regulate the kinase activity of the histidine kinase CheA, which, in turn, controls the flagella motor through a response regulator CheY (for details about the signaling pathway, see refs. 7 and 8). It is generally believed that cooperativity caused by the receptor clustering contributes to the high gain in signal transduction observed in *Escherichia coli* (9).

For the past several years, much work has been devoted to understanding the structure of the MCP cluster in *E. coli* (10–13). However, because of its complexity, a comprehensive picture of the cluster formation and its *in vivo* structure is still missing. On the functional side, significant progress has been made in *in vivo* measurement of the kinase activity of the cell by using FRET (14, 15). The quantitative measurements have made modeling possible beyond the conceptual level. Indeed, with quantitative modeling, we were able to infer from these functional data (alone) not only the existence of receptor interaction, but also more subtle properties of the system such as the existence of strong interaction between different types of receptors (16).

In modeling the cooperativity within the receptor cluster, most of the recent modeling efforts (16–19) adapted a nearest-neighbor interaction scheme where the receptors are located on a regular lattice and the activity of each individual receptor is affected by the activities of its nearest neighbors. The analogy between these models and the Ising model for magnetism in physics was explored to gain useful insight about the cooperativity of the receptor cluster. Conceptually, the Ising-type model relies on the definition of an "activity" for each individual receptor (in the context of bacterial chemotaxis, the receptor homodimer is referred to as receptor in this article), which is difficult to define and measure. Furthermore, the strength of the nearest-neighbor interaction is also hard to determine. It depends on the nature of the bond between the neighboring receptors, i.e., whether it is a weak bonding between different trimers of dimers or between members of the same trimer of dimers, or if they are connected through CheW and CheA. Without a detailed cluster formation model, it is not clear how the interaction strength would depend on other cellular conditions, such as the expression levels of the receptors and other relevant chemotaxis proteins. Given the lack of information on the cluster structure, the allosteric model for cooperativity proposed some 40 years ago by Monod, Wyman, and Changeux (MWC) (20) and reviewed recently by Changeux and Edelstein (21) seems rather appealing. In general, cooperativity can be characterized by the correlation length of the system, which depends on the strength of the nearest-neighbor interactions in Ising-type models. In the MWC model, the correlation length is effectively set by the size of the cluster (22), therefore bypassing all of the complexity in determining the local interactions between receptors. Also, the MWC model can be solved algebraically, making the analysis easier and more intuitive.

However, the classical MWC model was only applicable for systems composed of identical subunits and is inadequate for describing functions of heterogeneous clusters. For bacterial chemotaxis, the MCP cluster contains five different types of receptors and other cytoplasmic proteins. Among the five types of receptors, Tar and Tsr, which bind to Asp and Ser, respectively, are the most abundant, and the interaction between different types of receptors is highly relevant (14, 16, 23,

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: MCP, methyl-accepting chemotaxis protein; MWC, Monod, Wyman, and Changeux; HMWC, heterogeneous MWC.

‡To whom correspondence should be addressed. E-mail: yuhai@us.ibm.com.

www.pnas.org/cgi/doi/10.1073/pnas.0506961102

24). Our interest is to understand the behaviors of a heterogeneous functional cluster within which the receptors are strongly coupled and therefore could be described by the MWC-type model. (The large polar receptor cluster seen experimentally could contain many such smaller functional clusters.) Recent attempts to use MWC-type models to describe chemotactic response (25) used independent energy parameters for different mixed receptor clusters and therefore missed the connection between different receptor configurations. In this article, we construct a general model for heterogeneous allosteric protein complexes and apply it to studying the recent *in vivo* FRET data (15) for different mutant strains in responses to different stimulus (Ser and/or MeAsp). Our goal is to understand all of the data within a coherent model where the interrelation between the responses of these different strains can be revealed and further quantitative predictions can be made to be compared with future experiments.

An Allosteric Model for Heterogeneous Protein Complex

We first reformulate the classical MWC model by using an energy function (Hamiltonian) approach, from which generalization to the more general case of heterogeneous complexes emerges naturally. The complex is made of N identical subunits, each of which can bind to a ligand molecule. The ligand occupancy of the ith subunit is given by σ_i: $\sigma_i = 0, 1$ for vacant and occupied receptor, respectively ($i = 1, 2, \ldots, N$). In the all-or-none MWC model, the activity s of the complex is either active ($s = 1$) or inactive ($s = 0$). For the MWC model, the energy of the complex depends on s and σ in the following way:

$$H = \left(E + \varepsilon \sum_i \sigma_i\right)s + \mu \sum_i \sigma_i, \qquad [1]$$

E is the energy difference between the active and inactive state in the absence of ligand; each occupied receptor suppresses the activity by increasing the energy of the active state by $\varepsilon > 0$; μ is the energy for ligand binding for the inactive state and depends on the ligand concentration and a dissociation constant, K_i, for the inactive state. All energies are in units of the thermal energy k_BT. The correspondence between the energy parameters used here and that of the original MWC model can be summarized in the following:

$$e^{-E} = L, \quad e^{-\varepsilon} = C, \quad e^{-\mu} = \frac{[L]}{K_i}, \qquad [2]$$

where $[L]$ is the ligand concentration. The dissociation constant for the active state, K_a, is simply given by: $K_a = K_i/C$. L is the equilibrium constant.

Given the Hamiltonian (1), the partition function Z is given by:

$$Z = \sum_{\text{all states}} \exp(-H) = (1 + e^{-\mu})^N + e^{-E}(1 + e^{-(\varepsilon+\mu)})^N. \qquad [3]$$

From the partition function, all of the steady-state (equilibrium) properties of the model can be easily calculated. In particular, the average activity $\langle s \rangle$ can be determined:

$$\langle s \rangle = -Z^{-1}\frac{\partial Z}{\partial E} = \frac{L(1 + C[L]/K_i)^N}{(1 + [L]/K_i)^N + L(1 + C[L]/K_i)^N}, \qquad [4]$$

recovering the familiar MWC results when the parameters relations in Eq. **2** are used.

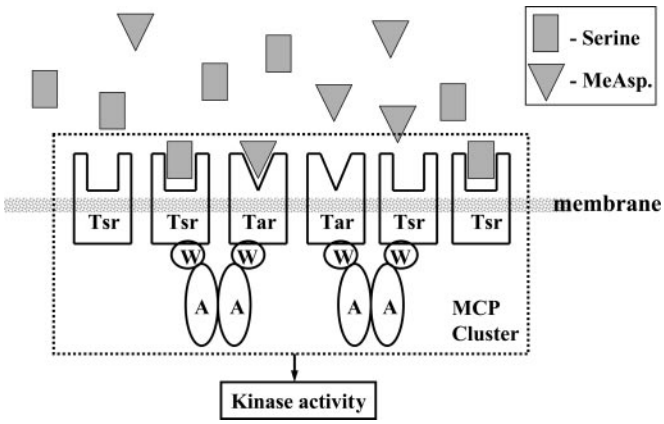

Fig. 1. Illustration of a functional MCP complex, consisting of different types of membrane-bound receptors, Tar and Tsr, and other important cytoplasmic components, CheW (W) and CheA (A), in response to two stimuli, Ser and MeAsp, which bind to Tsr and Tar, respectively. The whole polar receptor cluster may contain many such functional complexes.

From the Hamiltonian formulation of the MWC model, we can now generalize it to complexes with multiple types of receptors, each responding to different types of ligand, as illustrated in Fig. 1. For simplicity, we will derive the simplest extension of the MWC model with two types of receptors, Tar and Tsr, responding to two different types of ligand, Ser and Asp. Further generalization to more than two species of receptors or for the case where multiple receptors can bind with the same ligand with different affinities can be developed similarly.

Consider a cluster consisting of N_1 Tar receptors and N_2 Tsr receptors. The binding of Tar (Tsr) to MeAsp (Ser) is described by $K_1(K_2)$, the dissociation constants for the inactive state, and $C_1(C_2)$, the ratio between the dissociation constants in the inactive and the active states (we neglect the binding of Tsr to MeAsp at high concentration in this article). The Hamiltonian H_m for the mixed cluster can be written as:

$$H_m = \left(E + \varepsilon_1 \sum_{i_1=1}^{N_1} \sigma_{i_1} + \varepsilon_2 \sum_{i_2=1}^{N_2} \sigma_{i_2}\right)s + \mu_1 \sum_{i_1=1}^{N_1} \sigma_{i_1} + \mu_2 \sum_{i_2=1}^{N_2} \sigma_{i_2}, \qquad [5]$$

where subscripts i_1 and i_2 represent the Tar and Tsr receptor, respectively. E is the energy difference between the active and the inactive states for the whole cluster in the absence of any types of ligand. The relations between energy parameters and the MWC parameters are given the same way as in Eq. **2**: $\exp(-E) = L$, $\exp(-\varepsilon_1) = C_1$, $\exp(-\mu_1) = [L]_1/K_1$, $\exp(-\varepsilon_2) = C_2$, $\exp(-\mu_2) = [L]_2/K_2$, where $[L]_1$ and $[L]_2$ are the MeAsp and Ser concentrations, respectively.

The partition function Z_m of the mixed cluster can be determined:

$$Z_m \equiv \sum_{\text{all states}} \exp(-H_m) = \left(1 + \frac{[L]_1}{K_1}\right)^{N_1}\left(1 + \frac{[L]_2}{K_2}\right)^{N_2} + L\left(1 + C_1\frac{[L]_1}{K_1}\right)^{N_1}\left(1 + C_2\frac{[L]_2}{K_2}\right)^{N_2}, \qquad [6]$$

from which the activity $\mathcal{A}_m$ in the presence of a mixture of ligand concentrations $[L]_1$ and $[L]_2$ can be calculated:

BIOPHYSICS

Table 1. Strain-specific parameters from the fitting to the experimental data

j(strain)	$f_{j,1}$	$f_{j,2}$	$N_{j,1}$	$N_{j,2}$	$A_j^{(0)}$
1	0.6	2	4.95	16.5	0.0806
2	1	2	4.00	8.00	0.0933
3	2	2	4.39	4.39	0.118
4	6	2	18.7	6.24	0.0875
5	1	0	14.0	0	0.0323
6	2	0	29.8	0	0.0645
7	6	0	73.5	0	0.0872
8	0	0.6	0	9.85	0.0133
9	0	1.4	0	15.2	0.0365
10	0	10	0	32.3	0.0983

The strain-specific parameters determined by fitting the HMWC model to the experimental response data from ref. 15 for the 10 mutant strains studied here. Also listed are the expression levels of Tar and Tsr, $f_{j,1}$ and $f_{j,2}$ (in units of Tar expression level in the WT cells) for strain $j(\in [1, 10])$ taken from ref. 15 (the expression level of Tsr is set to be two times that of Tar in the WT cells). The receptor-specific parameters are found to be $l_1 = 1.23$, $C_1 = 0.449$, $K_1 = 49.2(\mu M)$ for Tar; and $l_2 = 1.54$, $C_2 = 0.314$, $K_2 = 34.5(\mu M)$ for Tsr. The average equilibrium constant for the cytoplasmic components of the cluster $\bar{l}_0 = 0.826$. All the model parameters given here are obtained by setting $N_{2,t}(\equiv N_{2,1} + N_{2,2})$, the total number of receptors in a functional cluster for strain 2, to be 12. See text for details on other possible parameter sets.

$$\mathcal{A}_m = \frac{L\left(1 + C_1\frac{[L]_1}{K_1}\right)^{N_1}\left(1 + C_2\frac{[L]_2}{K_2}\right)^{N_2}}{\left(1 + \frac{[L]_1}{K_1}\right)^{N_1}\left(1 + \frac{[L]_2}{K_2}\right)^{N_2} + L\left(1 + C_1\frac{[L]_1}{K_1}\right)^{N_1}\left(1 + C_2\frac{[L]_2}{K_2}\right)^{N_2}}. \quad [7]$$

In the presence of only one type of ligand, i.e., when either $[L]_1 = 0$ or $[L]_2 = 0$, the activity of the mixed cluster follows exactly the same form as the original MWC model. However, the responses of the mixed MCP complex to pure Ser and pure MeAsp are not independent, they are related by having the same equilibrium constant L, whose value is affected by the presence of different types of receptors and other cytoplasmic proteins in the cluster. In fact, the responses of a given mixed MCP complex to any combinations of Ser and MeAsp concentrations all are connected by the simple unified formula (Eq. **7**).

Application of the Heterogeneous MWC (HMWC) Model to Understanding Bacterial Chemotaxis Responses

In a recent study by Sourjik and Berg (15), a number of strains of *E. coli* CheRB$^-$ mutant with different induced Tar and Tsr expression levels were constructed, and their responses to either Ser or MesAsp were measured quantitatively *in vivo* by using FRET (15). In this article, we focus on the 10 strains studied in ref. 15, which only differ in their expression levels of Tar and/or Tsr receptors. As shown in Table 1, for the two major receptors (Tar and Tsr), strains 5–7 have only Tar expressed at different levels, strains 8–10 have only Tsr expressed at different levels, and strains 1–4 have both. One of the primary motivations for our study is to explain the FRET response data for all of these strains together consistently within the HMWC model, which would then enable us to understand the intrinsic connection between the behaviors of all of the different strains.

One important constraint for developing a coherent description is that the intrinsic parameters, such as the ligand binding constants for a given receptor (type and methylation level) should be the same in all of the strains. The different behaviors observed in the aforementioned different strains should only be caused by the differences in the size and composition of the functional receptor cluster. For a particular strain $j \in [1, 10]$, using the HMWC model Eq. **7**, the FRET measurement of kinase activity $F_j([L]_1, [L]_2)$ in the presence of Ser and MeAsp concentrations $[L]_1$ and $[L]_2$ can be written as:

$$F_j([L]_1, [L]_2) = A_j^{(0)} \cdot \frac{L_j\left(1 + C_1\frac{[L]_1}{K_1}\right)^{N_{j,1}}\left(1 + C_2\frac{[L]_2}{K_2}\right)^{N_{j,2}}}{\left(1 + \frac{[L]_1}{K_1}\right)^{N_{j,1}}\left(1 + \frac{[L]_2}{K_2}\right)^{N_{j,2}} + L_j\left(1 + C_1\frac{[L]_1}{K_1}\right)^{N_{j,1}}\left(1 + C_2\frac{[L]_2}{K_2}\right)^{N_{j,2}}}, \quad [8]$$

the receptor-specific parameters K_1, C_1 for Tar and K_2, C_2 for Tsr, are the same for each strain, i.e., independent of j. The four strain-dependent parameters $A_j^{(0)}$, L_j, $N_{j,1}$, and $N_{j,2}$ correspond to an overall scaling factor, the equilibrium constant, and the number of Tar and Tsr receptors in the cluster, respectively. The scaling factor $A_j^{(0)}$, aside from an overall factor converting kinase activity to FRET signal strength, should be proportional to the amount of CheA bound to the clusters. The other three strain-dependent parameters, $N_{j,1}$, $N_{j,2}$, and L_j can be related to each other based on simple assumptions.

First, assuming Tar and Tsr are well mixed within a cluster, their relative abundance within a cluster should be the same as the ratio of their expression levels:

$$N_{j,1} : N_{j,2} = f_{j,1} : f_{j,2}, \quad [9]$$

where $f_{j,1}$ and $f_{j,2}$ are the expression levels for Tar and Tsr; the expression level of Tar in the WT cells is used as the unit for expression levels throughout this article. Because $f_{j,1}$ and $f_{j,2}$ were measured experimentally (15) as given in Table 1, Eq. **9** eliminates one free parameter from our model and provides another link between experiments and our theory.

Second, the equilibrium constant L_j is determined by the energy difference E_j between the active and inactive state in the absence of any ligand: $L_j = \exp(-E_j)$. If we assume that E_j is the sum of contributions from all of the constitutive parts of the cluster, we have $E_j \approx N_{j,1}e_1 + N_{j,2}e_2 + N_{j,0}e_0$, where e_1 and e_2 are the energy contributions from each Tar and Tsr receptor, and $N_{j,0}$ and e_0 are the number and average energy contribution of the other cluster components, such as the minor receptors and the cytoplasmic proteins in the cluster, respectively. The validity of this additive energy assumption is addressed in *Summary and Discussion*. The corresponding expression for L_j:

$$L_j \approx l_0^{N_{j,0}} l_1^{N_{j,1}} l_2^{N_{j,2}}, \quad [10]$$

where $l_1 = \exp(-e_1)$, $l_2 = \exp(-e_2)$ are the equilibrium constants for a single Tar or Tsr receptor, respectively, and $l_0 = \exp(-e_0)$ is the average equilibrium constant for the rest of the cluster. Like the K_1, C_1 and K_2, C_2 parameters, l_1 and l_2 are the intrinsic parameters for the receptors and should be the same for all of the strains. For the 10 strains considered here, the expression levels of cytoplasmic components of the MCP cluster, including CheA and CheW, are kept constant, so we assume $N_{j,0}/N_{j,1} = f_{j,0}/f_{j,1}$, where $f_{j,0}$ is the total expression level of all of the cytoplasmic components of the cluster. The contribution to the equilibrium constant from the cytoplasmic proteins is: $l_0^{N_{j,0}} \equiv \bar{l}_0^{(N_{j,1}/f_{j,1})}$, where $\bar{l}_0 \equiv l_0^{f_{j,0}}$ is a constant parameter because $f_{j,0}$ is the same for all of the strains studied here. Eq. **10** eliminates one more strain-specific parameter from our model, but more importantly, it establishes another connection between different strains in terms of their equilibrium constants.

Taken together, our HMWC model has three parameters for each receptor (K_1, C_1, l_1 for Tar; K_2, C_2, l_2 for Tsr), a (modified) average equilibrium constant $\bar{l}_0$ for the rest of the cluster and two

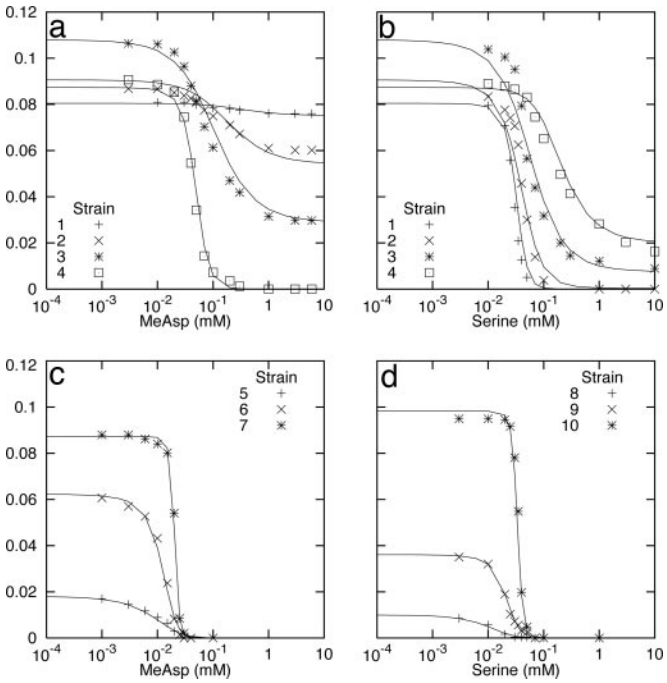

Fig. 2. Fitting of our model (lines) to the 14 experimental response data sets (symbols) for the 10 CheRB$^-$ mutant strains reported in ref. 15. (*a*) Response to MeAsp for strains 1–4. (*b*) Response to Ser for strains 1–4. (*c*) Response to MeAsp for strains 5–7. (*d*) Response to Ser for strains 8–10. The parameters of our model are given in Table 1.

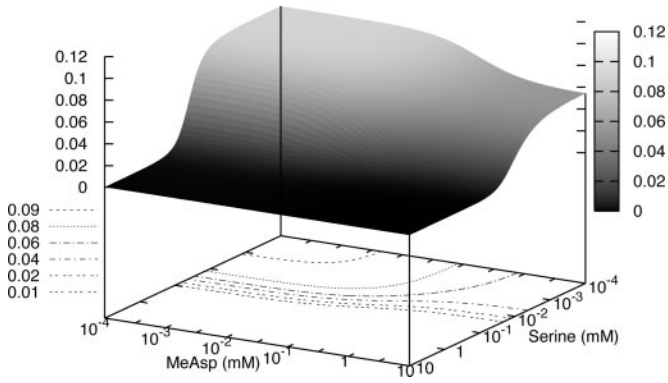

Fig. 3. The kinase activity (vertical axis) predicted by our model for strain 2 in the presence of both MeAsp and Ser concentrations. The contour lines for different activities in the ligand concentration space are also plotted.

independent strain-specific parameters $A_j^{(0)}$ and $N_{j,1}$. Our strategy is to fit all of the response data from the 10 strains together by using this unified model. To successfully fit such a simple model with a relatively small set of parameters to large diverse sets of data not only justifies our model, it is probably the only way to determine the *in vivo* parameters for such a complex biological system.

The fitting is done by minimizing an error function $\chi^2 = \Sigma_{j=1}^{10} (\tilde{D}_j - \tilde{F}_j)^2/\delta_j^2$, where $\tilde{D}_j$ is the measured response values at different stimulus level for strain j, and the corresponding model prediction is given by $\tilde{F}_j$ obtained by Eq. **8** with a given choice of the parameter values. δ_j ($\approx$10%) represents the experimental error in strain j.

We find that the HMWC model with the constrains given here can be used to fit excellently with all 14 response curves for the 10 different mutant strains studied by Sourjik and Berg (15). In Fig. 2, the results from our model (lines) are plotted together with the experimental data (symbols). The parameters for the model are given in Table 1 together with the gene expression levels for Tar and Tsr measured experimentally for each strains. We find that there is a family of possible parameters that could generate equally good fit to the data. Such nonuniqueness in parameters is related to an approximate symmetry of the MWC model, wherein a change in the cluster size parameter N (above certain lower bound) can be compensated by corresponding changes of the other parameters (L, K, and C) to render the response curve unchanged. Ideally, this degeneracy can be removed by direct measurement of receptor occupancy, which would provide extra information about K and C. Without the ligand occupancy data, we break this symmetry (arbitrarily) by fixing the functional cluster size in one of the 10 strains. For example, we set the total number of receptors in a functional cluster for strain 2, $N_{2,t} \equiv N_{2,1} + N_{2,2} = 12$, to reach the unique parameter set shown in Table 1. The details of this symmetry and the dependence of the parameters on the choice of N_2 are given in *Supporting Text* and Figs. 6 and 7, which are published as supporting information on the PNAS web site. Overall, the fitting error for each curve is well within 10%. To estimate error in the fitting parameters, we have generated 100 sets of test data by randomly shifting the experimental data by 10% of their measured values and fitted these test data sets with our model. The resulting spread in the parameters are all within 10% except for N_7 and N_{10} that have $\approx$30–50% error, probably related to the aforementioned symmetry in MWC-type model where for large N, a large change in N can be compensated by a relatively small change in C and l.

Quantitative Predictions: Responses to Arbitrary Mixture of Ser and MeASP

For an *E. coli* cell with mixed receptor cluster, such as the WT cell and the mutant strains 1–4 studied in the previous section, one interesting question is how they respond to a mixture of multiple stimuli, e.g., in the presence of both Ser and MeAsp. With all of the parameters determined from fitting our model to the existing data, we can readily predict quantitatively the response of any one of these strains to any combination of Ser and MeAsp concentrations by using Eq. **8**. In Fig. 3, the predicted response of strain 2, which has the same Tar/Tsr expression levels as the WT cell, is shown for arbitrary values of Ser and MesAsp concentrations. The details of the predicted responses are shown in Fig. 4 *a* and *b*, where the predicted responses to Ser (MeAsp) in the presence of various concentrations of MeAsp (Ser) are plotted. The predicted responses for the other mixed strains are given in Figs. 8–11, which are published as supporting information on the PNAS web site.

From Fig. 4 *a* and *b*, it is obvious that the presence of one ligand affects the response to another ligand strongly in a nonadditive fashion. For a given background MeAsp concentration $[L]_1$, the normalized response to the Ser concentration $[L]_2$, $F_2([L]_2|[L]_1)/A_2^{(0)}$, can be approximated by the Hill function: $(S_{max} - S_{min})(1 + ([L]_2/K_{eff})^h)^{-1} + S_{min}$. The range of kinase activity S_{min}, S_{max}, the apparent dissociation constant K_{eff}, and the steepness of the response h, depend on $[L]_1$ as plotted in Fig. 4*c*. The corresponding parameters for response to MeAsp in the presence of Ser concentration $[L]_2$ are plotted in Fig. 4*d*. It is clear from Fig. 4 *c* and *d* that the presence of one ambient attractant suppresses the response to another attractant. However, the response to these two ligands is not additive, as first noted in ref. 15, evidenced by the change in the shape of the response curve (instead of just an uniform shift). Quite interestingly and perhaps counterintuitively, the presence of one ambient attractant enhances the sensitivity of the system toward the other attractant as indicated in the monotonic decreasing of

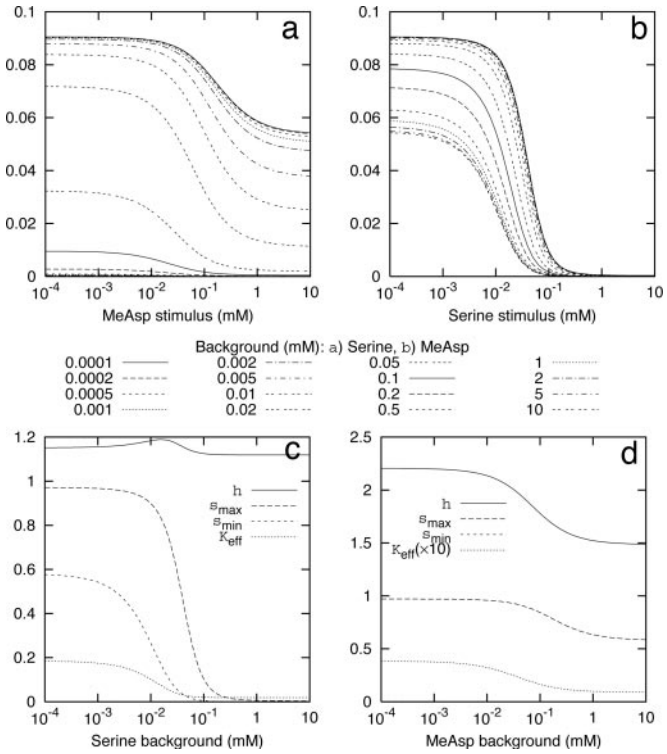

Fig. 4. The predicted kinase activity and response characteristics for strain 2 in the presence of two stimuli, Ser and MeAsp. (*a* and *b*) The responses of strain 2 to MeAsp (*a*) and Ser (*b*) in the presence of different background levels of Ser or MeAsp. (*c* and *d*) Each normalized response curve is fitted by a Hill function $(S_{max} - S_{min})(1 + ([L]/K_{eff})^h)^{-1} + S_{min}$, and the resulting parameters, S_{max}, S_{min}, K_{eff}, and h for MeAsp (*c*) and Ser (*d*) responses are plotted against the concentrations of the Ser and MeAsp background concentrations, respectively.

K_{eff} versus the background ligand concentration. Any large values of steepness of the response is in general suppressed by the presence of the other attractant. All of these qualitative behaviors can be understood naturally from Eq. **8**. Overall, the HWMC model provides a valuable starting point for answering questions on signal integration and response to complex environments in bacterial chemotaxis.

Summary and Discussions

In this article, extending the classical work of MWC, we propose a model to describe allosteric interactions within a heterogeneous cluster. This extension is necessary to account for the activities of protein complexes with multiple types of receptors and other constitutive components. Applying the heterogeneous allosteric model to bacterial chemotaxis, we can explain quantitatively the recent *in vivo* response measurements for a diverse set of strains to different stimuli in a coherent framework. Based on our model and the fitting to the existing data, we predict the responses of these strains to any combination of multiple ligand concentrations, which would hopefully stimulate further experiments to verify our theory and refine the parameters. There are two types of parameters in our model, the receptor-specific parameters and the strain-specific ones. The determination of these parameters through fitting of our model to the existing data provides valuable insights into the properties of individual receptor as well as the function and formation of the MCP cluster in bacterial chemotaxis. We discuss these findings in the following.

The Properties of Tar and Tsr. As shown in Eq. **2**, the values of $\varepsilon_i = -\ln(C_i)$ and $e_i = -\ln(l_i)$ are the changes in activation energy caused by a ligand binding to one receptor and addition of one receptor (in the absence of any ligand), respectively. These energies (in the unit of thermal energy k_BT) are found to be around order unity in our study, consistent with the small conformational changes each receptor can have. The parameters for Tar and Tsr are not drastically different, consistent with their structural similarity. However, we found $\varepsilon_2 > \varepsilon_1$, indicating that Tsr is a slightly more dominating receptor than Tar in the sense that an equal amount of receptor occupancy would lead to more suppression in activity for the Ser response than for the MeAsp response. Unique determination of the *in vivo* parameters, as shown here, may require direct ligand binding measurements, which are not available *in vivo*. However, it would be interesting to test whether the *in vitro* ligand binding measurements (26) could be explained consistently with the *in vivo* kinase activity measurements within the HMWC model.

The Possible Origin of l_0. Besides chemo receptors, the MCP cluster contains other important cytoplasmic components; among them, CheW and CheA both are shown to be critical to the formation of the cluster *in vivo*, and CheA is directly responsible for the kinase activity of the cluster. Within the allosteric model proposed here, the effects of these nonligand binding components are included in their modification of the equilibrium constant by a factor $L_0 = l_0^{N_{j,0}}$. If we ignore this effect by setting $l_0 = 1$ in our model, we could not get a good fit to the experimental data (see Fig. 12, which is published as supporting information on the PNAS web site), implying the importance of the energetic contributions from the cytoplasmic parts of the cluster. Furthermore, if the stoichiometry of the MCP cluster was fixed for all of the strains, say TA_XW_Y (T, A, and W represent receptor, CheA, and CheW, respectively) with fixed values of X and Y, we could eliminate L_0 from our model by redefining l_1 and l_2 as the equilibrium constant for the complex subunit TA_XW_Y instead of the receptor T alone. The fact that we needed strain-dependent L_0 in our model (i.e., $l_0 \neq 1$) to fit the experimental data supports the notion that the MCP cluster complex has variable stoichiometry, as suggested in ref. 15. Interestingly, for the strains we studied, we can get good agreement with the experimental data by having $L_{j,0} = l_0^{N_{j,0}}$ with a constant l_0, indicating that the stoichiometry depends on the relative expression levels of CheA, CheW, and the receptors. The constancy of $\bar{l}_0 = l_0^{f_{j,0}}$ in our study reflects the fixed expression levels of CheA and CheW in the strains we studied. When we tried to fit our model to strains with different levels of CheW or CheA (data not shown), we found we needed to have different values of $\bar{l}_0$ for these strains, confirming the origin of l_0 and the dependence of the equilibrium constant on the CheW and CheA expression levels.

The Validity of the Additive Energy Assumption. For the energy responsible for the equilibrium constant, we used the simplest linear approximation wherein the energy depends linearly on the number of each type of receptors. The validity of this linear approximation is justified by the good fit of our model to the data. In an independent analysis by Sourjik and Berg (15), the data from strains 5–7 and 8–10 with a single type of receptor were fitted with the classical MWC model, and the resulting values of L and N also show exponential dependence, i.e., $\ln(L)$ versus N is linear (see Fig. 13, which is published as supporting information on the PNAS web site). However, with the current limited amount of data, we could not rule out the existence of corrections to this linear relation, in particular for clusters with both Tar and Tsr.

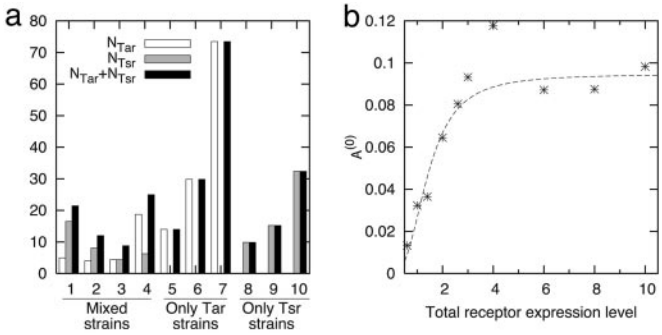

Fig. 5. The strain-dependent parameters from our model (see Table 1). (*a*) The number of receptors in the cluster in the 10 different strains. (*b*) The scaling factor $A^{(0)}$ against the total receptor expression level T_t (in units of WT Tar expression level) for different strains; most of the values of $A^{(0)}$ from the 10 strains collapse onto a curve that is fitted by function $A^{(0)} = 0.0947/[1 + (1.43/T_t)^{2.61}]$ shown as a dotted line. The outlier at $T_t = 4$ corresponds to strain 3 where the Tar and Tsr are equally expressed as $f_{3,1} = f_{3,2} = 2$.

The Dependence of Cluster Size on Receptor Gene Expressions. The cluster size parameter N_j in our model is best understood as the number of highly correlated receptors in the system. It could come from the finite size of a functional cluster with highly coupled receptors or be limited by the strength of the local receptor–receptor interaction within a large extended lattice of receptors. For the strains with only one type of receptor, i.e., strains 5–7 for Tar and strains 8–10 for Tsr, $N_j = N_{j,1} + N_{j,2}$ is a monotonically increasing function of $T_t(\equiv f_{j,1} + f_{j,2})$, the total expression level for both Tar and Tsr, as shown in Fig. 5*a*. However, interestingly, we found for the strains with mixed receptors, i.e., strains 1–4, the dependence of N_j on T_t, as shown in Fig. 5*a*, is not monotonic. In fact, N_j is the smallest for the strain $j = 3$, the strain with equal amounts of Tar and Tsr, indicating possible dependence of the cluster size on the heterogeneity of the receptor population.

The Dependence of Cluster-Bound CheA on the Total Number of Receptors. The overall scale in our model $A_j^{(0)}$ should be proportional to the amount of cluster-bound CheA. Assuming Tar and Tsr have roughly equal affinity to bind with CheA (through CheW), we have plotted in Fig. 5*b* the $A^{(0)}$ versus T_t for all 10 strains. Overall, the dependence for the 10 different strains seems to collapse onto one curve, which can be approximated by a simple binding curve with a characteristic dissociation constant of $K_d \approx 1.43$. The collapse of $A^{(0)}$ versus T_t onto a single curve confirms the origin of $A^{(0)}$. However, there are also obvious outliers, such as strain $j = 3$ with $T_t = 4$; interestingly, it corresponds to the case where the cluster has the most mixed population: $f_{3,1}$:f_3 = 1:1, which may indicate possible correction to the CheA binding curve caused by the difference between Tar and Tsr. These last two findings should serve as guidance for further development of cluster formation models.

Ising-Type Models Versus MWC-Type Models. The MWC-type model can be considered as a coarse-grained version of a generalized Ising-type model (22). The differences between the two types of models are most evident only in scales smaller than the correlation length (or MWC cluster size). Measurements of total kinase activity alone, therefore, may not be sufficient to distinguish the two types of models, e.g., earlier FRET response data for mutants with different methylation levels (14) can be explained by an Ising-type model (16) and the HMWC model developed here (data not shown). For the response data discussed here, the (practical) disadvantage of the Ising-type model is that different sets of coupling strength parameters would have to be introduced for different mutant strains in an ad hoc way, and the greater number of parameters may lead to data overfitting. However, to unambiguously decide whether all of the polar receptors are connected in a continuous lattice (Ising-type models) or form many smaller all-or-none functional clusters (MWC-type models) probably requires measurements that can probe the local structure (tightness) of the cluster under different internal/external conditions, such as receptor expression levels and ligand concentrations.

In summary, the heterogenous allosteric model proposed here provides a consistent, intuitive modeling framework for understanding signal integration and response to complex environment in bacterial chemotaxis. The model is also general so it should be useful for studying the functions of other heterogeneous protein complexes in biology.

Note. Upon finishing our work, we learned that Dr. Tom Shimizu of Harvard University has been working on a similar HMWC model independently.

We thank Dr. Howard Berg and Dr. Victor Sourjik for explaining their experiments to us and Dr. Gustavo Stolovitzky for careful reading of the manuscript.

1. Monks, C. R., Freiberg, B. A., Kupfer, H., Sciaky, N. & Kupfer, A. (1998) *Nature* **395,** 82–86.
2. Marx, S. O., Ondrias, K. & Marks, A. R. (1998) *Science* **281,** 818–821.
3. Ashkenazi, A. & Dixit, V. M. (1998) *Science* **281,** 1305–1308.
4. Liu, F., Wan, Q., Pristupa, Z. B., Yu, X. M., Wang, Y. T. & Niznik, H. B. (2000) *Nature* **403,** 274–280.
5. Bray, D. & Duke, T. (2004) *Annu. Rev. Biophys. Biomol. Struct.* **33,** 53–73.
6. Maddock, J. R. & Shapiro, L. (1993) *Science* **259,** 1717–1723.
7. Bren, A. & Eisenbach, M. (2000) *J. Bacteriol.* **182,** 6865–6873.
8. Falke, J. J. & Hazelbauer, G. (2001) *Trends Biochem. Sci.* **26,** 257–265.
9. Bray, D., Levin, M. D. & Morton-Firth, C. J. (1998) *Nature* **393,** 85–88.
10. Shimizu, T. S., Novère, N. L., Levin, M. D., Beavil, A. J., Sutton, B. J. & Bray, D. (2000) *Nat. Cell Biol.* **2,** 792–796.
11. Kim, S., Wang, W. & Kim, K. K. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 11611–11615.
12. Weis, R. M., Hirai, T., Chalah, A., Kessel, M., Peters, P. J. & Subramaniam, S. (2003) *J. Bacteriol.* **185,** 3636–3643.
13. Francis, N. R., Wolanin, P. M., Stock, J. B., DeRosier, D. J. & Thomas, D. R. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 17480–17485.
14. Sourjik, V. & Berg, H. C. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 123–127.
15. Sourjik, V. & Berg, H. C. (2004) *Nature* **428,** 437–441.
16. Mello, B. A. & Tu, Y. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 8223–8228.
17. Duke, T. A. J. & Bray, D. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 10104–10108.
18. Shi, Y. (2001) *Phys. Rev. E* **64,** 21910.
19. Mello, B. A., Shaw, L. & Tu, Y. (2004) *Biophys. J.* **87,** 1578–1595.
20. Monod, J., Wyman, J. & Changeux, J. P. (1965) *J. Mol. Biol.* **12,** 88–118.
21. Changeux, J. P. & Edelstein, S. J. (2005) *Science* **308,** 1424–1428.
22. Duke, T. A. J., Novere, N. L. & Bray, D. (2001) *J. Mol. Biol.* **308,** 541–553.
23. Ames, P., Studdert, C. A., Reiser, R. H. & Parkinson, J. S. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 7060–7065.
24. Ames, P. & Parkinson, J. S. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 2117–2122.
25. Rao, C. V., Frenklach, M. & Arkin, A. P. (2004) *J. Mol. Biol.* **343,** 291–303.
26. Levit, M. N. & Stock, J. B. (2002) *J. Biol. Chem.* **277,** 36760–36765.

Corrections and Retraction

CORRECTIONS

BIOPHYSICS. For the article "An allosteric model for heterogeneous receptor complexes: Understanding bacterial chemotaxis responses to multiple stimuli," by Bernardo A. Mello and Yuhai Tu, which appeared in issue 48, November 29, 2005, of *Proc. Natl. Acad. Sci. USA* (**102,** 17354–17359; first published November 17, 2005; 10.1073/pnas.0506961102), the authors note that the citation given for ref. 24 [Ames, P. & Parkinson, J. S. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 2117–2122] was incorrect. The citation should have read as follows:

24. Studdert, C. A. & Parkinson, J. S. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 2117–2122.

www.pnas.org/cgi/doi/10.1073/pnas.0600530103

GENETICS. For the article "Mutagenic specificity of endogenously generated abasic sites in *Saccharomyces cerevisiae* chromosomal DNA," by Paul Auerbach, Richard A. O. Bennett, Elisabeth A. Bailey, Hans E. Krokan, and Bruce Demple, which appeared in issue 49, December 6, 2005, of *Proc. Natl. Acad. Sci. USA* (**102,** 17711–17716; first published November 28, 2005; 10.1073/pnas.0504643102), the authors note that ODG was erroneously listed as an abbreviation for "oxanine-DNA glycosylase." In fact, ODG represents the empty vector for parallel experiments to the CDG and TDG vectors. In some but not all places, ODG was corrected to read "pODG"; this correction should have been made throughout the paper. These errors do not affect the conclusions of the article.

www.pnas.org/cgi/doi/10.1073/pnas.0600853103

IMMUNOLOGY. For the article "A pilot clinical trial of a recombinant ricin vaccine in normal humans," by Ellen S. Vitetta, Joan E. Smallshaw, Elaine Coleman, Hasan Jafri, Callie Foster, Robert Munford, and John Schindler, which appeared in issue 7, February 14, 2006, of *Proc. Natl. Acad. Sci. USA* (**103,** 2268–2273; first published February 3, 2006; 10.1073/pnas.0510893103), the authors note the following. "E.S.V. and J.E.S. are patent holders on RiVax, a recombinant ricin vaccine, and would be eligible to accept royalties if the vaccine is commercialized. In the case of E.S.V., royalties would be signed over to the University of Texas Southwestern Medical Center."

www.pnas.org/cgi/doi/10.1073/pnas.0601105103

PHYSIOLOGY. For the article "Regulation of sodium-proton exchanger isoform 3 (NHE3) by PKA and exchange protein directly activated by cAMP (EPAC)," by Katharina J. Honegger, Paola Capuano, Christian Winter, Desa Bacic, Gerti Stange, Carsten A. Wagner, Jürg Biber, Heini Murer, and Nati Hernando, which appeared in issue 3, January 17, 2006, of *Proc. Natl. Acad. Sci. USA* (**103,** 803–808; first published January 9, 2006; 10.1073/pnas.0503562103), the authors note that the concentrations for EPAC1 (40) and PKA (1) affinities were published incorrectly due to a printer's error. On page 803, in lines 12–14 of the first full paragraph, right column, "EPAC1 and the B domain of EPAC2 bind cAMP *in vitro* with a K_m of ≈40 mM, whereas the K_m of PKA is ≈1 mM (17)" should read: "EPAC1 and the B domain of EPAC2 bind cAMP *in vitro* with a K_m of ≈40 μM, whereas the K_m of PKA is ≈1 μM (17)." These errors do not affect the conclusions of the article.

www.pnas.org/cgi/doi/10.1073/pnas.0600995103

RETRACTION

BIOPHYSICS. For the article "Packing defects as selectivity switches for drug-based protein inhibitors," by Ariel Fernández, Ridgway Scott, and R. Stephen Berry, which appeared in issue 2, January 10, 2006, of *Proc. Natl. Acad. Sci. USA* (**103,** 323–328; first published December 30, 2005; 10.1073/pnas.0509351102), the editors note that there is substantial overlap in the figures and text of this PNAS article with the article by A. Fernández that appeared in the December 2005 issue of *Structure* (**13,** 1829–1836) titled, "Incomplete protein packing as a selectivity filter in drug design." The latter article, which is copyrighted by *Structure*, is not cited in the PNAS article, and all panels of the figures appearing in the *Structure* article are reproduced in the PNAS article without reference to the *Structure* article. PNAS policy states that articles must not be previously published and that permissions must be obtained for any previously published portions of the work prior to publication.

Because this article does not meet these requirements, PNAS is withdrawing it.

Nicholas R. Cozzarelli, *Editor-in-Chief*

www.pnas.org/cgi/doi/10.1073/pnas.0601034103