Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Theoretical Biology 241 (2006) 223–232
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Simulating the evolution of signal transduction pathways

Orkun S. Soyer[a,*], Thomas Pfeiffer[a,b,1], Sebastian Bonhoeffer[a]

[a]Theoretical Biology Group, Ecology and Evolution, CH 8092, Zürich, Switzerland
[b]Computational Laboratories, Swiss Federal Institute of Technology (ETH), CH 8092, Zürich, Switzerland

Received 11 July 2005; received in revised form 4 November 2005; accepted 18 November 2005
Available online 5 January 2006

Abstract

We use a generic model of a network of proteins that can activate or deactivate each other to explore the emergence and evolution of signal transduction networks and to gain a basic understanding of their general properties. Starting with a set of non-interacting proteins, we evolve a signal transduction network by random mutation and selection to fulfill a complex biological task. In order to validate this approach we base selection on a fitness function that captures the essential features of chemotactic behavior as seen in bacteria. We find that a system of as few as three proteins can evolve into a network mediating chemotaxis-like behavior by acting as a "derivative sensor". Furthermore, we find that the dynamics and topology of such networks show many similarities to the natural chemotaxis pathway, that the response magnitude can increase with increasing network size and that network behavior shows robustness towards variations in some of the internal parameters. We conclude that simulating the evolution of signal transduction networks to mediate a certain behavior may be a promising approach for understanding the general properties of the natural pathway for that behavior.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Computational simulation; Signal transduction networks; Evolution; Chemotaxis; Network modeling

1. Background

Conventional analyses of signal transduction networks include experimental approaches concentrating on single proteins or interactions in the network and theoretical approaches concentrating at converting the experimental findings into mathematical models in order to simulate the behavior of these networks under various conditions. The combined application of these approaches has proved crucial to fully describe specific signal transduction networks (see, for example, Bhalla, 2004). A particularly successful case was the signaling network mediating the chemotactic behavior of motile bacteria-like *Escherichia coli*. Both the general properties of chemotactic behavior (Berg and Brown, 1972; Adler et al., 1973; Berg and Tedesco, 1975), and the proteins and interactions mediating it (Blair, 1995; Falke et al., 1997; Bren and Eisenbach, 2000) have been unraveled by experimental studies. This information was then used to develop several mathematical models of the chemotaxis pathway (Segel et al., 1986; Bray et al., 1993; Spiro et al., 1997; Shimizu et al., 2003). In combination, these efforts has led to a thorough understanding of fundamental features of chemotaxis such as sensitivity (Sourjik and Berg, 2002), adaptation (Block et al., 1983), and robustness (Alon et al., 1999) and provided insights on how these are achieved (Segel et al., 1986; Barkai and Leibler, 1997; Spiro et al., 1997; Bray et al., 1998; Yi et al., 2000).

Despite such success stories, the approach of combining experimental and theoretical efforts is not without drawbacks. First of all, the current understanding of chemotactic behavior was achieved over a time span of nearly two decades. This is mainly due to the fact that mathematical modeling approaches rely heavily on knowledge of exact biochemical properties, which can only be gathered through laborious and time-consuming experimental work. Second, it may not always be possible to transfer the specific knowledge gained from one signaling network to another one. A fundamental understanding that could be

*Corresponding author. Tel.: +41 1 6323337; fax: +41 1 6321271.
E-mail address: orkun.soyer@env.ethz.ch (O.S. Soyer).
[1]New address: Program for Evolutionary Dynamics, Harvard University, One Brattle Square, Cambridge, MA-02138, USA.

0022-5193/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2005.11.024

applied to all signal transduction networks usually requires a complete theoretical framework as achieved in the case of chemotaxis (Yi et al., 2000). It would be desirable to devise new theoretical approaches that could help overcome these drawbacks. A crucial point is to make theoretical approaches less dependent on experimental studies. This would lead to faster exploration of theories on signal transduction networks and may direct experimental studies, reversing the information flow between the two approaches.

Here, we propose simulating the evolution of signal transduction networks as a theoretical approach to understand their properties. In brief, this means evolving a certain behavior using a generic model of signal transduction and analysing the resulting network. This approach does not require exact knowledge of biochemical parameters or any other detailed experimental knowledge. Furthermore, it may allow us to address fundamental questions about a signal transduction network such as; what are the minimum requirements to achieve a certain behavior? What are key topological or dynamical features mediating a certain behavior? What are the evolutionary selective pressures resulting in a certain behavior? In these regards, this approach may prove to be useful in achieving a broad understanding of signal transduction networks and may complement current approaches in this quest. Approaches similar to the proposed one have been used in the study of genetic networks successfully. For example, generic models of gene regulation combined with in silico evolution approaches has been used to understand underlying mechanisms behind developmental pattern formation (Salazar-Ciudad et al., 2001) and dynamics of small regulatory networks (Francois and Hakim, 2004). A generic model of regulatory transcription control gave insights on the capabilities of such networks to achieve logic operations (Buchler et al., 2003). In signal transduction networks, Bray et al. attempted to evolve small signal transduction networks starting from a well-defined and simple initial network (Bray and Lay, 1994). Given such a network structure, they let the system evolve through mutations to parameters defining the interactions among proteins and ligand binding to receptor, while keeping the network structure intact. Thus, their simulation differed from the presented approach in that it analysed evolution of network response for a specific network rather than over different network topologies and sizes. Maybe because of this limitation, resulting networks in that study failed to achieve complex responses such as one tracking the derivative of the signal.

In order to validate the use of in silico evolution as an approach to understand signal transduction networks, we use a generic model and evolve chemotactic behavior as observed in bacteria. We then analyse the resulting networks in order to understand the essential network features leading to chemotaxis. The results show that a weak chemotactic performance could be achieved with as few as three proteins, and that a network of 4 and 5 proteins could improve this performance significantly by facilitating response sensitivity and adaptation. Analysis of these networks reveals dynamical and topological features that are similar to the natural chemotaxis pathways, validating this approach for achieving an understanding of signal transduction networks. In the next section we discuss the evolution of chemotactic behavior based on bacterial chemotaxis and present the resulting networks. The generic signal transduction model we use and the details of evolution simulations are explained in Section 4.

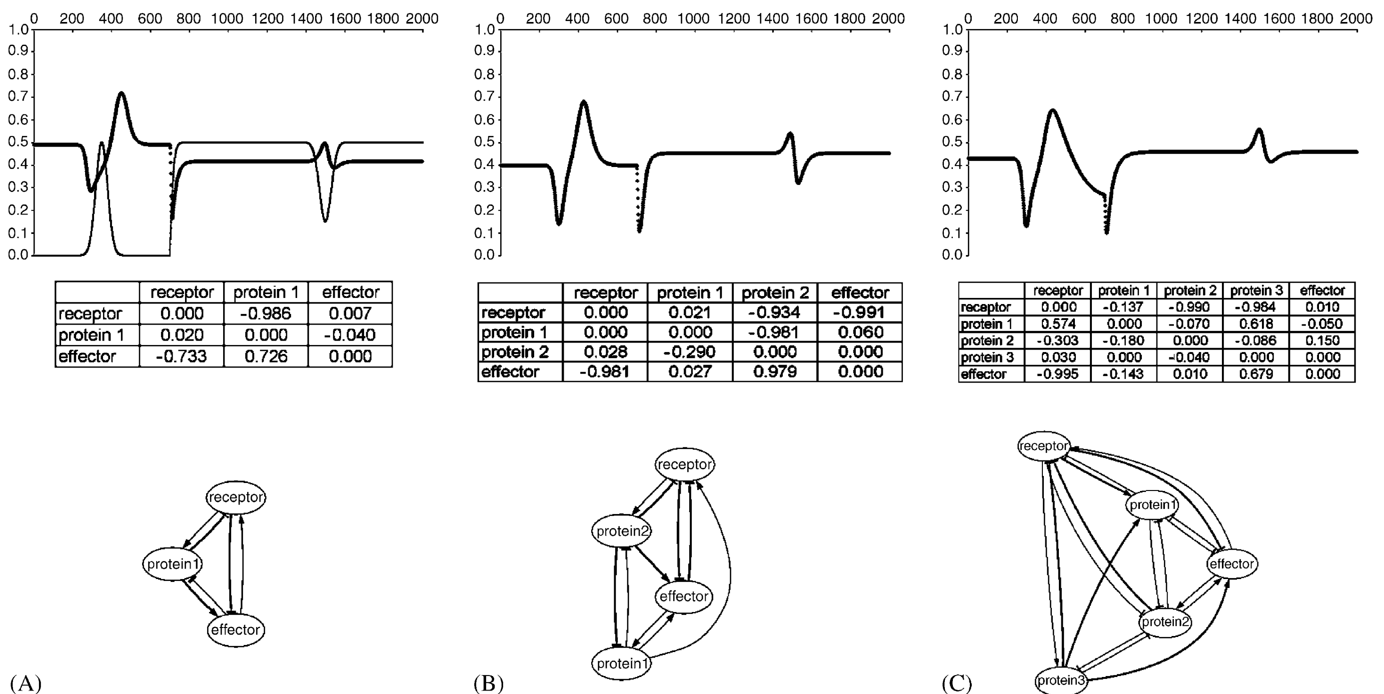
2. Results and discussion

2.1. *In silico evolution of chemotaxis*

Chemotaxis is the ability of motile bacteria to respond and swim up (down) a gradient of attractant (repellent). This behavior and the pathway mediating it, is well-studied in *E. coli* (Mesibov and Adler, 1972; Adler et al., 1973; Tso and Adler, 1974; Berg, 1983). Briefly, the chemotactic behavior is a biased random walk (Berg, 1983). It is achieved by a receptor detecting outside signals and a set of proteins that form a cascade coupling this signal with movement of bacteria. Movement results from rotation of a motor protein; counter clockwise rotation results in smooth swimming of bacteria in a given direction, and clockwise rotation results in abrupt tumbling and random change of direction. Bacteria can bias their random walk towards higher food concentrations by coupling outside signals to tumbling frequency (Macnab and Koshland, 1972; Berg, 1983; Stock et al., 2002). The key properties of bacterial chemotaxis are; response to changes in the concentration of the signal molecule (Macnab and Koshland, 1972; Berg and Tedesco, 1975; Spiro et al., 1997) and a wide range of detection and high sensitivity (Segall et al., 1986; Bray, 2002; Sourjik and Berg, 2002).

Here, we use a generic signal transduction model to evolve such a chemotactic response in silico. The evolution simulation consists of an iterative process of mutation and selection, where mutations involve parameters defining interactions among proteins and selection is based on a fitness function that captures essential features of bacterial chemotaxis (see Section 4). We opted for using such a fitness function as it proved to be computationally very demanding to evolve chemotactic behavior simply by a simulation of bacteria movement (i.e. it requires very long simulations of movement of bacteria on a two-dimensional grid in order to ascertain with sufficient accuracy that a given network outperforms its predecessor, because of stochastic effects inherent in such simulations). Still, we verified the chemotactic ability of resulting networks with such spatial simulations on a two-dimensional grid with a ligand gradient, where the tumbling probability of bacteria was coupled to network response (see Fig. 2).

Fig. 1 shows the topology and interaction parameters of networks with the best chemotactic performance from 10 evolutionary simulations, for models containing 3–5



	receptor	protein 1	effector
receptor	0.000	-0.986	0.007
protein 1	0.020	0.000	-0.040
effector	-0.733	0.726	0.000

	receptor	protein 1	protein 2	effector
receptor	0.000	0.021	-0.934	-0.991
protein 1	0.000	0.000	-0.981	0.060
protein 2	0.028	-0.290	0.000	0.000
effector	-0.981	0.027	0.979	0.000

	receptor	protein 1	protein 2	protein 3	effector
receptor	0.000	-0.137	-0.990	-0.984	0.010
protein 1	0.574	0.000	-0.070	0.618	-0.050
protein 2	-0.303	-0.180	0.000	-0.086	0.150
protein 3	0.030	0.000	-0.040	0.000	0.000
effector	-0.995	-0.143	0.010	0.679	0.000

Fig. 1. Evolved 3- to 5-protein networks. Response, interaction coefficients and topology of evolved networks with 3- to 5-proteins (panels A–C). Network response shows the concentration of the active effector protein vs. simulation time in presence of a signal as changing ligand concentration (thin line in panel A). The network coefficients are shown in a matrix arrangement, with row i showing the effect of other proteins on protein i. The absolute values of negative and positive entries stand for deactivation and activation coefficients (i.e. l_{ij} and k_{ij}), respectively (see Eq. (4) in Section 4). In network topologies, arrows and T-ended lines represent activating and deactivating interactions, respectively. Strong interactions (those with a coefficient larger than 0.5) are shown in bold. See the text for discussion of the dynamics of these networks.

proteins (see Supplementary Fig. 1 and Supplementary Table 1 for information on additional networks). All these networks achieve high fitness values and show the same type of response to changes in ligand concentration. In increasing (decreasing) food concentrations the tumbling frequency (i.e. the active effector concentration) decreases (increases) while returning close to a base value when there is no change in food concentration (compare responses to the signal shown in Fig. 1A). In other words, the response tracks the derivative of the signal and balances the benefit from moving up gradients of food against deviations from a base response. This ensures some adaptive behavior as seen in the response to the sigmoidal part of the signal sequence. To summarize, the presented networks act as a derivative sensor like the natural chemotaxis pathway (Spiro et al., 1997) and their response to changing ligand concentrations is similar to that seen in natural bacterial chemotaxis (Block et al., 1983; Segall et al., 1986) or detailed mathematical models of chemotaxis (see, for example, Rao et al., 2004).

In order to test whether such a response really corresponds to a chemotactic behavior, we implemented these networks in virtual bacteria and performed a stochastic simulation of their movement on a continuous grid in the presence of a ligand distribution (see Section 4). In Fig. 2 we show the behavior of these virtual bacteria along with two others, whose movement was based on a constant tumbling frequency or a detailed model of chemotaxis (as described in Rao et al., 2004). These two additional bacteria would thus correspond to a random movement and perfect chemotaxis, respectively, and act as a reference point. Each panel in Fig. 2 shows the simulation grid, color-coded with the average time spent by the bacterium on each grid location. We quantified the extent of chemotactic behavior by calculating the average amount of ligand (i.e. food) intake during these simulations (see Section 4). Based on the results of this calculation (also shown in Fig. 2) and the performance of a randomly moving bacterium we can conclude that both the 4- and 5-protein networks mediate chemotaxis-like behavior, while the 3-protein network is not capable to outperform a random tumbler with regards to food intake. That the performance of the 3-protein network nevertheless corresponds to a form of chemotaxis is seen in a more detailed analysis of the behavior that it mediates (see Supplementary Fig. 2); the response of this network allows the bacterium to swim up a gradient more often than a randomly moving bacterium would do, but fails to keep it at high ligand concentrations. The ability of the presented networks to show some sort of chemotactic behavior in these stochastic simulations is even more striking given the fact that the signal encountered during these simulations differs significantly from the signal used during the in silico evolution process.

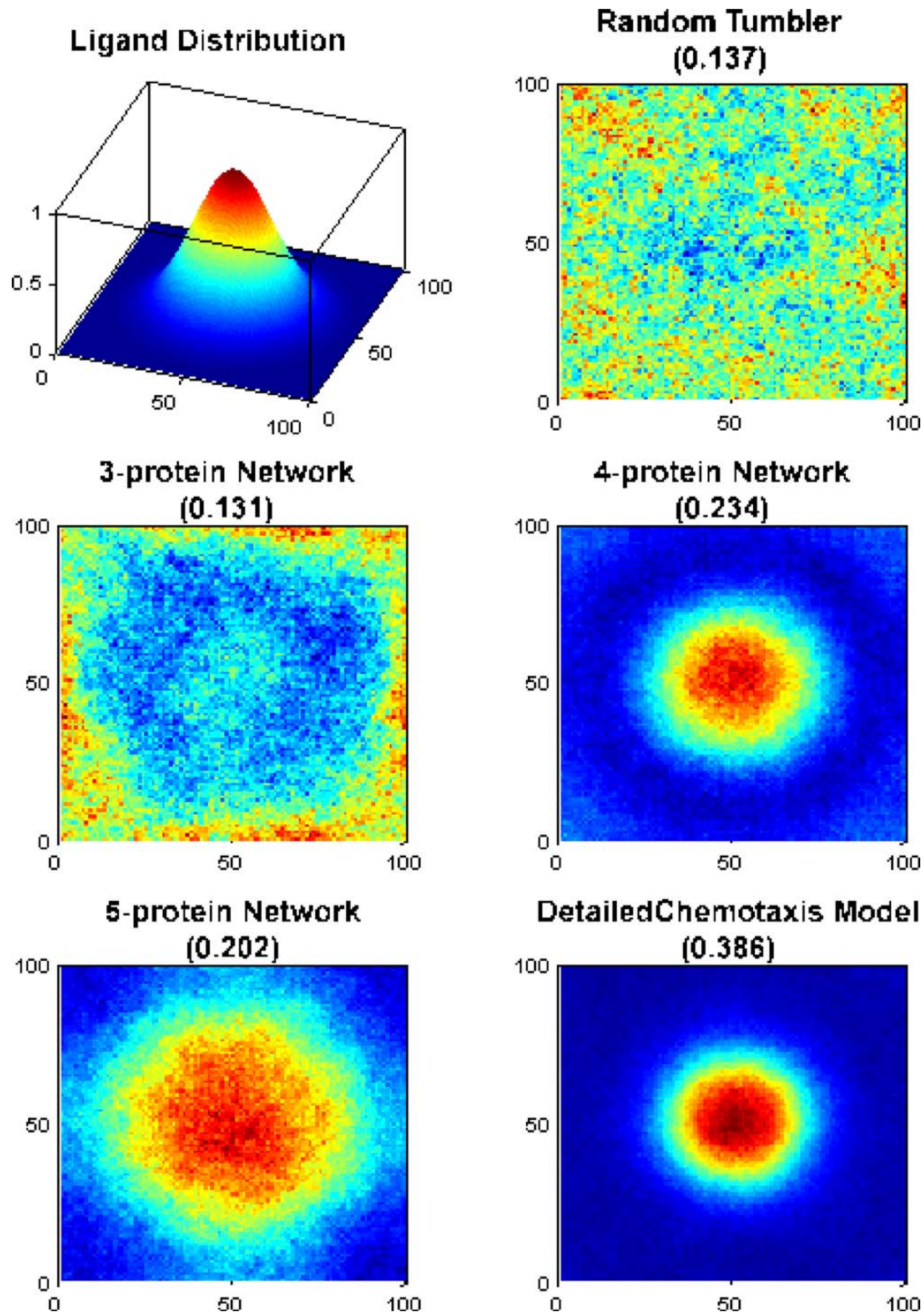


Fig. 2. Chemotaxis using evolved networks. The time spent on each grid location during the stochastic simulations performed on a continuous 100×100 grid (see Section 4 for simulation details). Each panel shows the results of simulations for five different virtual bacteria, whose moving dynamics are dictated by a given model or by a constant tumbling frequency, as a color-coded grid. The chemotactic performance measured as the amount of ligand (i.e. food) intake during the simulations is given in parentheses. Different colors indicate the average number of times that the bacterium was on a given grid location during the simulations. Colors vary from dark blue to dark red with increasing average time spent on a given location (i.e. a location colored dark blue indicates that bacteria was almost never on that location, while dark red indicates that it spent most of its time there). The panel at the top left shows the distribution of the ligand over the grid using the same color-coding scheme for the level of ligand concentration.

2.2. Network dynamics

How do these networks achieve chemotactic behavior and what underlies the differences in their performance? Analyzing the coefficients of the simplest network of three proteins we see a specific network topology. The receptor deactivates the effector, which activates it back forming a negative feedback loop on the effector. The receptor also interacts with the third protein in the system, forming a negative feedback loop on itself. Finally, this regulatory

protein interacts with the effector to form a negative feedback loop on itself. This topology with three feedback loops involving the three proteins results in a rapid response of the system when the signal changes as both the activation of the receptor by the ligand and the deactivation of the effector by the receptor are mediated through strong interactions (see Fig. 1). On the other hand the effects of the regulatory protein (i.e. the deactivation of the receptor and activation of the effector) take place on a much slower time-scale, as the activity of this protein is controlled through weak interactions. Knocking out such weak interactions decreases the response fitness significantly, indicating that the interplay of these fast and slow processes allow the network to act as a derivative sensor.

In larger networks it is more difficult to see how response dynamics and network topology is related. In order to find key interactions in 4- and 5-protein networks we analyzed networks resulting from all evolutionary simulations and that were capable of a derivative sensor type response (see Supplementary Fig. 1). Interestingly, in both 4- and 5-protein networks we were able to find a core structure that resembled closely to the 3-protein topology described above. In networks of both these sizes we note the strong deactivation of effector by the receptor and the presence of an intermediary protein that strongly deactivates the receptor and strongly activates the effector (see Fig. 1). Hence, these interactions seem to be the key in enabling an adaptive response that tracks the derivative of the signal. The additional interactions in the 4- and 5-protein networks facilitate sensitivity and adaptation, allowing them to achieve a much better chemotactic performance then the 3-protein network.

2.3. Comparison to the natural chemotaxis pathways

The network of proteins mediating chemotaxis is well-described in *E. coli* (see, for example, Blair, 1995; Falke et al., 1997). A complex composed of a methyl accepting membrane receptor, the adaptor protein CheW and a kinase (CheA) detects the outside signal and transmits this to the effector CheY via phosphorylation. The phosphorylated form of CheY stimulates tumbling by interacting with the motor. This is the main branch in the chemotaxis pathway that couples outside signals to tumbling frequency of the cell. In addition, proteins CheR and CheB act on the receptor through methylation and demethylation resulting in an increase and decrease in the sensitivity of the receptor complex, respectively. Finally, a phosphatase (CheZ) decreases the binding affinity of CheY for the motor protein, terminating the signal. In concert, these proteins allow an adaptive response underlying chemotactic behavior. Another motile bacterium, *Bacillus subtilis*, has a similar pathway that is less well studied. Despite employing homologous proteins, the two pathways differ in that the phosphorylated CheY increases the probability of tumbling in *E. coli* but smooth swimming in *B. subtilis* (Bischoff et al., 1993) and that the kinase activity is inhibited upon ligand binding in *E. coli* but stimulated in *B. subtilis* (Garrity and Ordal, 1997).

Although the generic model used in this analysis does not allow a direct comparison between presented networks and the natural chemotaxis pathways, there are significant similarities at the fundamental level. First, we observe a direct interaction between the receptor and effector allowing rapid signal transmission just like in natural pathways. Second, there is a need for additional proteins (at least one) to achieve adaptive response and signal transmission. Third, we observe an interesting combination of interaction coefficients in the presented networks that result in an interplay of fast and slow processes, which seem fundamental for a response tracking the derivative of the signal. Indeed, it has been argued that a similar interplay among rapid response and slow adaptation enables natural chemotaxis (Spiro et al., 1997; Sourjik and Berg, 2002). Finally, we note a strong correlation between the chemotactic performance of a network and its ability to adapt to changes in the signal level and have an imbalance in its response to increasing and decreasing ligand gradients (see Supplementary Fig. 1). A mutant 3-protein network (that is created by tweaking the parameters of the evolved 3-protein network to achieve a better adaptation) further supports this observation. These mutations resulted in a significant improvement of the chemotactic behavior mediated by the 3-protein network (see Supplementary Fig. 1), showing the importance of adaptation for achieving proper chemotaxis. This is also observed in natural bacteria: mutants deficient in adaptation were found to be severely defected in their chemotactic ability (Stock et al., 1985; Segall et al., 1986; Weis and Koshland, 1988). Similarly, the imbalance seen in the response of the presented networks is found to occur in natural bacteria, which adapt rapidly to decreases in ligand concentration but slowly to increases (Berg and Tedesco, 1975). These similarities in the response of the presented networks and natural chemotaxis pathway highlight the fundamental features of the system mediating chemotactic behavior.

It is also interesting to note that a recent analysis of chemotaxis pathways from *E. coli* and *B. subtilis* suggested that there could be an ancient core network of three proteins mediating chemotaxis (Rao et al., 2004). The authors speculate that current chemotaxis pathways evolved from such a core network in order to achieve better sensitivity and robustness. The presented analysis is supportive of this hypothesis in several aspects. First, we find three proteins to be the minimum requirement to achieve a chemotactic response, albeit a weak one. Second, although the exact architecture of the presented 3-protein network and the proposed one are different, the presence of feedback loops between receptor and effector seems crucial for both networks to achieve an adaptive response. Finally, the improved chemotactic performance of the presented networks with increasing size support the speculation that additional proteins in the network would lead to better sensitivity and robustness (see below).

2.4. Network sensitivity

The natural chemotaxis pathway achieves an exceptionally high sensitivity resulting in a ratio of 35 between fractional changes of kinase activity and receptor occupancy (Sourjik and Berg, 2002). Like the detailed mathematical models, the networks presented here fail to achieve such a high sensitivity. This observation was expected as the high sensitivity in the natural system is most probably a result of cooperative effects among receptors or other proteins (Bray et al., 1993, 1998; Spiro et al., 1997; Sourjik and Berg, 2004), which are not incorporated in the mathematical models. A more surprising result is that we observe an increase in sensitivity with increasing network size in our simulations. The sensitivity of 3- to 5-protein networks, calculated as the ratio of the change in active effector concentration to that in active receptor concentration during the first part of the signal (time points 225 and 315 in Fig. 1A) is 0.39, 0.51, and 0.63, respectively.

The sensitivity of the larger networks is also evident from stochastic simulations of virtual bacteria (see Fig. 2). The 4- and 5-protein networks are particularly successful in mediating chemotactic behavior, even compared to a detailed mathematical model that captures almost all known features of natural chemotaxis pathway (Rao et al., 2004). As all the networks result from the same evolutionary simulation and differ only in their size, it is plausible to argue that the increased sensitivity results from additional interactions (see Fig. 1) arising from increased network size. One such additional interaction we find corresponds to modulation of CheZ activity by the receptor in the natural chemotaxis pathway, which is shown to partially explain the high sensitivity of chemotaxis (Spiro et al., 1997).

2.5. Network robustness

Robustness towards variations in internal and external parameters is observed in many biological networks (de Visser et al., 2003). The analyses of robustness in natural chemotaxis pathway concentrated only on its adaptive capacity and found it robust towards variations in protein concentrations and kinetic parameters (Barkai and Leibler, 1997; Alon et al., 1999). Further theoretical analysis explained this observation by the presence of an integral

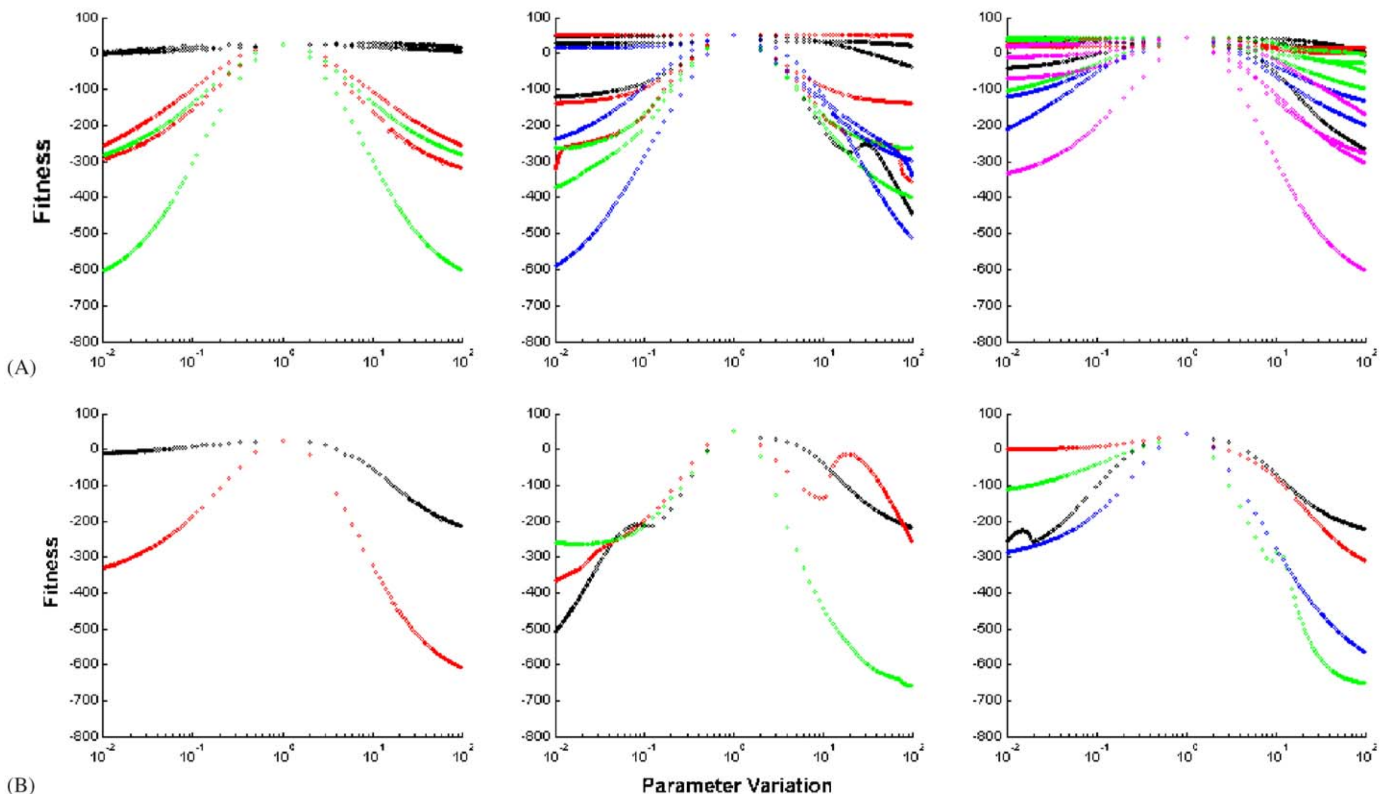


Fig. 3. Robustness of evolved networks. Graphs show the fitness of evolved networks (as calculated using Eq. (5)) under different perturbations to their parameters. The y-axis stands for fitness values, while the x-axis shows the logarithm of the ratio between altered and "wild-type" values of a given parameter. Panels from left to right shows the results for 3- to 5-protein networks, respectively. Each point represents a mutant network created through a random mutation to a specific parameter of the network (see Section 4). Panel A shows mutations to the interaction coefficients describing actions of other proteins on protein i, and panel B shows mutations to the total concentration of protein i (excluding the total concentration of the effector). Mutations to the parameters that are associated with a given protein are color-coded using black, red, green, blue, and magenta moving from the receptor to the effector.

feedback loop and suggested that such a loop is a general requirement for robust perfect adaptation (Yi et al., 2000). Here, we tested robustness in the presented networks by analyzing their chemotactic ability in the face of perturbations to the strength of protein interactions and initial protein concentrations (see Section 4). By the definition of the fitness function, we consider networks that achieve a positive fitness value as capable of some sort of chemotactic behavior.

As seen in Fig. 3, presented networks show considerable robustness towards such perturbations. In biologically plausible ranges of perturbations, networks achieve mostly positive fitness values (in some cases even in spite of 100-fold changes). The level of robustness towards particular perturbations gives an idea about the importance of the parameter being perturbed for the network response. For example, all studied networks are highly robust to perturbations targeting parameters defining the actions of other proteins on the receptor. On the other hand, parameters defining the actions of other proteins on the effector have much more severe effects on network response. The general robustness of the network seems to increase with increasing network size.

Irrespective of the level of robustness, it is very surprising that we find any robustness at all in these networks considering that the evolutionary simulations do not include any selective pressures for robustness. This observation supports the recent theoretical ideas for explaining the robustness of complex systems, stating that robustness is a byproduct of evolutionary processes (e.g. optimization) (Carlson and Doyle, 2000).

2.6. Comparison to mathematical chemotaxis models

The generic signal transduction model is designed for the in silico evolution approach to be applied to any signal transduction system and not only to chemotaxis. In other words it is designed with the assumption that one has no a priori knowledge on the network architecture or proteins involved. Hence, we incorporate only the most fundamental features of such systems in the generic model: proteins can exist only in active or inactive states and only activated proteins can interact with other proteins. In contrast, theoretical approaches concentrating on a specific signal transduction system can make use of all the experimental knowledge and convert this into a detailed and realistic mathematical model. In bacterial chemotaxis for example such mathematical models account for the several methylation states of the receptor, which have different probabilities of being active or inactive, and for the special cases where a protein interacts with another only when in the active or inactive state (Bray et al., 1993; Barkai and Leibler, 1997).

Despite such technical differences in the assumptions they use, the generic and detailed models are still similar at the very basic level: both simulate the change in the concentration of active and inactive proteins as a set of differential equations. Hence, although the topology of the evolved networks and detailed models of chemotaxis (see, for example, Fig. 2 in Rao et al., 2004) differ due to differences in model structure, their response to a given signal sequence and their dynamic properties are very similar as discussed above.

3. Conclusion

Using a simple and generic model of signal transduction networks we demonstrate that in silico evolution could be a powerful approach to understand fundamental features of such networks. Here, we evolved networks with various sizes to achieve chemotactic behavior and analysed their topological and dynamic properties. We find that as little as three proteins could mediate a form of chemotactic behavior and that additional proteins are important to achieve better sensitivity and robustness. The analyses of the resulting networks highlight key dynamic and topological properties that are crucial for the network to mediate a chemotactic behavior. Furthermore, the chemotactic ability of the presented networks shows some robustness to variations in internal parameters.

These findings are in good agreement with those obtained from experimental and theoretical studies of the natural chemotaxis pathway validating the use of the presented approach for understanding signaling networks. Furthermore, this analysis shows that the presented approach complements conventional experimental and theoretical studies to provide a more complete understanding of the signaling network mediating chemotactic behavior. The most important aspect of the in silico evolution approach is its general applicability and independence of knowledge on biochemical parameters. Hence, a similar analysis as the one presented here, could be applied to any signal transduction network to gain a basic understanding of its properties and provide new hypotheses to be tested by experimental studies. The generic signal transduction model can be extended in cases of experimental knowledge on proteins and their various states (such as methylated states). Alternatively, the generic model and the evolution approach could be used to address other fundamental properties of signal transduction networks such as the evolution of modularity.

4. Methods

4.1. A generic model of signal transduction networks

At the most basic level a signal transduction network is a set of interacting proteins that regulate each other's activity. The signal enters the system through a signal sensing protein (i.e. receptor) at the top of the signal transduction cascade, and produces an output through a response protein (i.e. effector). Hence, we model a signal transduction network as a set of proteins, each of which

can exist either in an active P (e.g. phosphorylated) or inactive P state. The equilibrium between the two states of a protein depends on the interactions between this protein and other active proteins in the system, where each interaction is characterized by a specific interaction coefficient (e.g. rate constant). The basic reaction dictating the dynamics of such a system can be written as

$$P_i \underset{\sum_j l_{ij}[P_j^*]}{\overset{\sum_j k_{ij}[P_j^*]}{\rightleftarrows}} P_i^*, \qquad (1)$$

where $[P_i^*]$ represents the concentration of active form of protein i and k_{ij} and l_{ij} represent the strength of the interaction between protein i and j. We assume that proteins do not interact with themselves (i.e. $k_{ii} = l_{ii} = 0$) and their interaction with other proteins can only be activating or deactivating (i.e. $k_{ij} \cdot l_{ij} = 0$), resulting in n^2-n parameters for a network of n proteins. Also, we set the maximum value that an interaction coefficient can attain to one for computational ease.

In order for the system to be able to respond to an incoming signal we define one of the proteins as a receptor and allow its equilibrium state to be influenced by the ligand (i.e. the signal) concentration $[L]$:

$$P_1 \underset{\sum_j l_{ij}[P_j^*]}{\overset{[L]+\sum_j k_{ij}[P_j^*]}{\rightleftarrows}} P_1^*. \qquad (2)$$

We assume that the energy required for the activation and deactivation reactions are provided from outside sources such as high-energy molecules. Finally, we arbitrarily choose another protein in the system as the effector and specify that the concentration of its active form dictate the behavior produced by the network.

Given a set of interaction coefficients for a set of proteins, we can monitor the changes in the active effector concentration in response to a signal, to evaluate the signal processing capacity of the network. The concentrations of active and inactive proteins in the system can be calculated at any time point by solving the set of differential equations resulting from (1):

$$\frac{\mathrm{d}[P_i]}{\mathrm{d}t} = \left[[P_i^*] \cdot \sum_j l_{ij} \cdot [P_j^*]\right] - \left[[P_i] \cdot \left(\delta_{i1} \cdot [L] + \sum_j k_{ij} \cdot [P_j^*]\right)\right], \qquad (3)$$

where $\delta_{i1} = 1$ for $i = 1$ and $\delta_{i1} = 0$ for $i \neq 1$. Note, that the total concentration of each protein $[P_i^{tot}]$ is constant and set to one (i.e. $[P_i^*] = 1 - [P_i]$). The set of differential equations as shown in Eq. (3) allows the simulation of a system of n proteins in presence or absence of a ligand. For example, for the 3-protein network shown in Fig. 1A these equations would look like:

$$\begin{aligned}
\mathrm{d}[receptor]/\mathrm{d}t &= [receptor^*] \cdot 0.986 \cdot [protein1^*] \\
&\quad - [receptor] \cdot ([L] + 0.007 \cdot [effector^*]), \\
\mathrm{d}[protein1]/\mathrm{d}t &= [protein1^*] \cdot 0.04 \cdot [effector^*] \\
&\quad - [protein1] \cdot 0.02 \cdot [receptor^*], \\
\mathrm{d}[effector]/\mathrm{d}t &= [effector^*] \cdot 0.733 \cdot [receptor^*] \\
&\quad - [effector] \cdot 0.726 \cdot [protein1^*].
\end{aligned} \qquad (4)$$

4.2. In silico evolution

Using this generic model we can evaluate the response of a given network to a given signal. This allows us to simulate an evolutionary process in silico and evolve networks that are capable of producing a certain response to an incoming signal. The selection stage of the evolutionary process requires a fitness function that captures the possible selective pressures under which such response has evolved. While it is plausible to use this generic model to evolve various biological responses, here we use bacterial chemotaxis pathway as a model system and evolve a "chemotactic" behavior. To do so, we treat the signal as the concentration of a chemical attractant (i.e. a food source) and the effector in our system as CheY, whose active state concentration dictates the tumbling frequency and hence the movement of the bacteria. We define the fitness function based on the following observations made on bacterial chemotaxis.

The "chemotactic" behavior employed by bacteria is that of a biased random walk (Berg, 1983). Bacteria explore the space by employing a certain tumbling frequency in the absence of any food source. They then decrease (increase) this tumbling frequency in increasing (decreasing) food gradients in order to bias their random walk towards higher food concentrations. The base tumbling frequency of bacteria is not zero, enabling them both to explore the space efficiently and to overcome the issues related with living in an environment with a low Reynolds number (Berg, 1983). Hence, we assume that the main selective pressures on chemotactic behavior are the ability to explore the environment while making maximal gain out of available food sources. Here, we therefore assume that there is an optimal tumbling frequency $R_{opt.}$. A good responder should decrease (increase) its tumbling frequency below (above) $R_{opt.}$ in increasing (decreasing) food gradients and should tumble with a frequency close to $R_{opt.}$ in absence of any food change. These two selective pressures on the motility of bacteria are captured in the following fitness function:

$$Fitness = \sum_t \left[C \cdot \left[\Delta F \cdot (R_{opt.} - R_{avg.})\right] - \left[(R_{opt.} - R_{avg.})^2\right]\right]. \qquad (5)$$

Here, $\Delta F = F_{t+1} - F_t$ and $R_{avg.} = (R_{t+1} + R_t)/2$, where R_t stands for the response of the network at time t (i.e. the

active effector concentration) given the signal F_t (i.e. food concentration). As described above, the response of the network is treated as the tumbling frequency of bacteria in order to simulate its movement. Hence, the first term of Eq. (5) quantifies the gain of a bacterium that behaves according to its signal transduction network over a bacterium that has no signal processing ability and just a constant tumbling frequency of $R_{opt.}$. The second term penalizes deviations from the optimal tumbling frequency $R_{opt.}$. Finally, we use a coefficient C to weigh the relative importance of these two selective pressures on the fitness. Presented results are achieved with $R_{opt.} = 0.45$ and $C = 200$. The value of the former parameter does not affect the presented results, while that of the latter is crucial for evolving a chemotactic behavior. For example, a too small value of C results in a constant response. Note, that the chosen ad hoc fitness function is just one of many possible functions that may capture the above-discussed selective pressures on bacterial chemotaxis.

We start the evolution simulation with a set of non-interacting proteins, each existing in equal amounts of active and inactive forms and a total concentration of one. After equilibrating the system by integrating the set of differential equations until steady state is reached, we introduce a signal in the form of a changing ligand concentration (as shown in Fig. 1A). The integration is then continued in presence of this signal until system reaches steady state or for 2000 time steps and the network response (i.e. active effector concentration) is recorded. This response is then used in Eq. (5) to calculate network fitness. Then, a mutant network is created from the original one through a random or fixed (by an amount of 0.01) incremental change in a randomly selected interaction coefficient in a random direction while keeping it in the interval $[-1,1]$. If an interaction coefficient for proteins i and j was negative (positive) then l_{ij} (k_{ij}) is set to the absolute value of this number and k_{ij} (l_{ij}) is set to zero. The response of the new network is evaluated in the same way and the mutation is accepted or denied based on fitness change. This iterative process is carried until we find an optimum network.

While this optimization process is an approximation to the real process of evolution, where the selective pressures would act on a population of mutant networks and favor offspring from those with higher fitness, it is necessary in this case due to computational burden of simulating networks. Besides such computational costs there are also difficulties with evolving an optimum network through an iterative mutation process. First, the use of a single network (compared to a population of networks) limits the chances of finding successful mutants. Second, the high dimensionality of the parameter space (there are $n^2 - n$ parameters for a network of n proteins) makes the iterative process prone to get stuck in local minima. These challenges can only be overcome by using proper algorithms to search the parameter space. Here, we use a simulated annealing algorithm that is shown to perform well in such problems (Kirkpatrick et al., 1983). In simulated annealing, beneficial mutations resulting in a fitness improvement are always accepted while detrimental mutations are accepted with a certain probability calculated as a function of the change in fitness and a control parameter (this control parameter is donated as temperature based on an analogy between optimization and calculation of the low-temperature state of a physical system). This search strategy allows easier and faster convergence to optima and does not contradict with the evolutionary simulation as it is based on iterative mutations. Here, we start the simulated annealing algorithm with a temperature of 200. After 50 accepted mutations per parameter at a certain temperature level we decrease the temperature exponentially to a lower level. The simulation is stopped when we reach 500 suggested mutations per parameter at a certain temperature level. To overcome inherent stochastic effects in this optimization scheme and to increase the chance of finding global optimum solutions, we perform 10 separate evolution simulations for each network size. Networks with the best chemotactic performance among all 10 resulting networks are shown in Fig. 1. Supplementary Fig. 1 shows all networks that give a response loosely tracking the derivative of the signal.

4.3. Robustness and chemotaxis analyses

In order to see how robust networks are to perturbations to their internal parameters, we created 500 mutants for each parameter and analyzed the fitness of the resulting networks. Mutations are done to interaction coefficients and the total protein concentrations, by increasing or decreasing the original parameter up to 100-fold. Both types of perturbations are biologically plausible; changes in interaction coefficients and protein concentrations could for example result from mutations to proteins and fluctuations in gene expression levels, respectively. For mutations done to the total concentrations, we exclude the effector from the analysis as the fitness value is directly related to the value of this parameter (i.e. $R_{avg.}$ in Eq. (5) is directly related to total effector concentration).

We analysed chemotactic ability by simulating movement of virtual bacteria on a grid with a ligand distribution on it, and coupling movement dynamics to the output of the presented networks. The ligand distribution on the grid had the shape of a two-dimensional Gaussian distribution, placed at the midfield with its peak and having a variance of 15 and a maximum value of one (see Fig. 2). Simulations were performed on a continuous 100×100 grid using periodic boundary conditions, with the bacterium initially located at the lower left corner of the grid and facing towards midfield. At each time step of the simulation, bacterium tumbled to change its move direction or moved on its current direction based on the active effector concentration, as calculated from the network dynamics. To better assess the chemotactic performance of these networks we simulated two additional bacteria, one

moving with a constant tumbling frequency set to 0.45 (i.e. with $R_{opt.}$) and one moving with its tumbling frequency coupled to the output of a detailed model of chemotaxis (we used the model presented in Rao et al., 2004). Each simulation is run for 100,000 time steps and the time spent on each grid location is calculated as an average over 100 separate simulations to reduce stochastic effects. Chemotactic performance is based on the ligand (i.e. food) intake during these simulations and is calculated as the average time spent at a given location times the ligand concentration summed over all locations.

All simulations are written in C++. The ordinary differential equations describing the generic network model are integrated using the Adams method in NAG library Mark 7. Source codes are available from authors upon request.

Acknowledgments

This work is funded by Gebert Ruef Stiftung project grant GRS-031/03. We thank Richard A. Goldstein, Richard R. Neubig, Dennis Bray, Matthew Levin, Karen Lipkow and two anonymous reviewers for very useful comments on the manuscript.

Appendix A. Supplementary materials

The online version of this article contain additional supplementary data. Please visit doi:10.1016/j.jtbi.2005.11.024, for supplemental Figs. 1, 2 and Table 1.

References

Adler, J., Hazelbauer, G.L., et al., 1973. Chemotaxis toward sugars in *Escherichia coli*. J. Bacteriol. 115 (3), 824–847.

Alon, U., Surette, M.G., et al., 1999. Robustness in bacterial chemotaxis. Nature 397 (6715), 168–171.

Barkai, N., Leibler, S., 1997. Robustness in simple biochemical networks. Nature 387 (6636), 913–917.

Berg, H.C., 1983. Random Walks in Biology. Princeton University Press, Princeton, NJ.

Berg, H.C., Brown, D.A., 1972. Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. Nature 239 (5374), 500–504.

Berg, H.C., Tedesco, P.M., 1975. Transient response to chemotactic stimuli in *Escherichia coli*. Proc. Natl Acad. Sci. USA 72 (8), 3235–3239.

Bhalla, U.S., 2004. Models of cell signaling pathways. Curr. Opin. Genet. Dev. 14 (4), 375–381.

Bischoff, D.S., Bourret, R.B., et al., 1993. Purification and characterization of *Bacillus subtilis* CheY. Biochemistry 32 (35), 9256–9261.

Blair, D.F., 1995. How bacteria sense and swim. Annu. Rev. Microbiol. 49, 489–522.

Block, S.M., Segall, J.E., et al., 1983. Adaptation kinetics in bacterial chemotaxis. J. Bacteriol. 154 (1), 312–323.

Bray, D., 2002. Bacterial chemotaxis and the question of gain. Proc. Natl Acad. Sci. USA 99 (1), 7–9.

Bray, D., Lay, S., 1994. Computer simulated evolution of a network of cell-signaling molecules. Biophys. J. 66 (4), 972–977.

Bray, D., Bourret, R.B., et al., 1993. Computer simulation of the phosphorylation cascade controlling bacterial chemotaxis. Mol. Biol. Cell 4 (5), 469–482.

Bray, D., Levin, M.D., et al., 1998. Receptor clustering as a cellular mechanism to control sensitivity. Nature 393 (6680), 85–88.

Bren, A., Eisenbach, M., 2000. How signals are heard during bacterial chemotaxis: protein–protein interactions in sensory signal propagation. J. Bacteriol. 182 (24), 6865–6873.

Buchler, N.E., Gerland, U., et al., 2003. On schemes of combinatorial transcription logic. Proc. Natl Acad. Sci. USA 100 (9), 5136–5141.

Carlson, J.M., Doyle, J., 2000. Highly optimized tolerance: robustness and design in complex systems. Phys. Rev. Lett. 84 (11), 2529–2532.

de Visser, J.A., Hermisson, J., et al., 2003. Perspective: evolution and detection of genetic robustness. Evol. Int. J. Org. Evol. 57 (9), 1959–1972.

Falke, J.J., Bass, R.B., et al., 1997. The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. Annu. Rev. Cell. Dev. Biol. 13, 457–512.

Francois, P., Hakim, V., 2004. Design of genetic networks with specified functions by evolution in silico. Proc. Natl Acad. Sci. USA 101 (2), 580–585.

Garrity, L.F., Ordal, G.W., 1997. Activation of the CheA kinase by asparagine in *Bacillus subtilis* chemotaxis. Microbiology 143 (Pt 9), 2945–2951.

Kirkpatrick, S., Gelatt, C.D., et al., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680.

Macnab, R.M., Koshland Jr., D.E., 1972. The gradient-sensing mechanism in bacterial chemotaxis. Proc. Natl Acad. Sci. USA 69 (9), 2509–2512.

Mesibov, R., Adler, J., 1972. Chemotaxis toward amino acids in *Escherichia coli*. J. Bacteriol. 112 (1), 315–326.

Rao, C.V., Kirby, J.R., et al., 2004. Design and diversity in bacterial chemotaxis: a comparative study in *Escherichia coli* and *Bacillus subtilis*. PLoS Biol. 2 (2), E49.

Salazar-Ciudad, I., Newman, S.A., et al., 2001. Phenotypic and dynamical transitions in model genetic networks. I. Emergence of patterns and genotype–phenotype relationships. Evol. Dev. 3 (2), 84–94.

Segall, J.E., Block, S.M., et al., 1986. Temporal comparisons in bacterial chemotaxis. Proc. Natl Acad. Sci. USA 83 (23), 8987–8991.

Segel, L.A., Goldbeter, A., et al., 1986. A mechanism for exact sensory adaptation based on receptor modification. J. Theor. Biol. 120 (2), 151–179.

Shimizu, T.S., Aksenov, S.V., et al., 2003. A spatially extended stochastic model of the bacterial chemotaxis signalling pathway. J. Mol. Biol. 329 (2), 291–309.

Sourjik, V., Berg, H.C., 2002. Receptor sensitivity in bacterial chemotaxis. Proc. Natl Acad. Sci. USA 99 (1), 123–127.

Sourjik, V., Berg, H.C., 2004. Functional interactions between receptors in bacterial chemotaxis. Nature 428 (6981), 437–441.

Spiro, P.A., Parkinson, J.S., et al., 1997. A model of excitation and adaptation in bacterial chemotaxis. Proc. Natl Acad. Sci. USA 94 (14), 7263–7268.

Stock, J., Kersulis, G., et al., 1985. Neither methylating nor demethylating enzymes are required for bacterial chemotaxis. Cell 42 (2), 683–690.

Stock, J.B., Levit, M.N., et al., 2002. Information processing in bacterial chemotaxis. Sci. STKE 2002 (132), PE25.

Tso, W.W., Adler, J., 1974. Negative chemotaxis in *Escherichia coli*. J. Bacteriol. 118 (2), 560–576.

Weis, R.M., Koshland Jr., D.E., 1988. Reversible receptor methylation is essential for normal chemotaxis of *Escherichia coli* in gradients of aspartic acid. Proc. Natl Acad. Sci. USA 85 (1), 83–87.

Yi, T.M., Huang, Y., et al., 2000. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. Proc. Natl Acad. Sci. USA 97 (9), 4649–4653.