Mathematical modeling of biological systems

D. Holcman
Weizmann Institute of Science, Rehovot, 76100 Israel

January 11, 2006

Abstract
In the past 50 years, major discoveries in biology have changed the direction of science. From the study of the sexual life of oysters, which was in some sense boring for the previous generations, biology has become today the Queen of Science. All hardcore fields, such as physics, mathematics, chemistry, and computer science are now necessary for the big adventure of unraveling the secrets of life and conversely, the mathematical sciences are all now enthusiastically inspired by biological concepts, to the extent that more and more theoreticians are interacting with biologists. What is today the role of a theoretician among the biologists, eager to incorporate new concepts? An important part of biology, besides amassing new experimental information, is the explanation of new phenomena. In order to explain how a pure theoretician, can contribute to the analysis of biological systems, I would like to discuss some selected open questions.

1-Biology is now driving science

In the past 50 years, major discoveries in biology have changed the direction of science. From the study of the sexual life of oysters, which was in some sense boring for the previous generations, biology has become today the Queen of Science. All hardcore fields, such as physics, mathematics, chemistry, and computer science are now necessary for the big adventure of unraveling the secrets of life and conversely, the mathematical sciences are all now enthusiastically inspired by biological concepts, to the extent that more and more theoreticians are interacting with biologists. Actually, it is not an understatement to say that biology has a Viagra effect on the old classical fields.

What is today the role of a theoretician among the biologists, eager to incorporate new concepts? An important part of biology, besides amassing new experimental information, is the explanation and prediction of new phenomena by applying the quantitative laws of physical chemistry, that is, by quantifying phenomena in mathematical terms, not by merely fitting curve with Numerical Recipes in Matlab. Theory is not a painting of the real but it gives the framework for quantitative computations, analysis and prediction. Data analysis is only small fraction of statistics.

The putting together of the pieces of the puzzle of life begins with the understanding the life of a protein, a microstructure, a cell, a network, and finally, the life of a living organism. In order to explain how a pure theoretician, can contribute to the analysis of biological systems, let us review some selected open questions.

2-Molecular trafficking and cellular organization

The smallest living unit in biology is the cell and the central questions are: how it functions, how it is organized and what are the rules and the machineries involved at a molecular level to make it works? The efficient working of organisms indicates that cells in the body of any animals are very well ordered, organized and orchestrated. But, what kind of molecular mechanisms control the cell function? Certainly, one has to search first at the principle of cell generation. In particular, in pluricellular organism such as mammalians, the cell specialization depends on its location, which is also part of its identity. The field of morphogenesis [12,7,9] consists precisely in identifying the rules used to orchestrate the construction and the organization of a complex organism. Each cell is dedicated to a precise job. For example in the brain, some cells participate to build the cortex, while other groups of cells are devoted to the skin layer, and so on. Small errors can occur in the task distributions and if for example, not enough cells can be dedicated to build a region such as the brain in the early stages of the development, then it may result in building a monster or a mutant with a smaller cortex. It is a challenging question to understand how cells get their instruction to build a given region and not another with an accurate precision. Obviously, positional information needs to be exchanged between cells in order to help a given cell to know where it is located and thus to activate the correct genetic code. Not only positional information tells the cell what to do, but a cell is also labeled, which provides its identity. The principle of ordering an ensemble of cells can be generated by morphogenetic gradients (figure 1a and 1b) where a substance travels from cell to cell, and each time a cell is passed, the concentration decreases. The city hall uses similar tricks to label houses in a street by plugging an increasing sequence of number at each house; as a result one knows immediately where to go. Isn't it an optimal method to know immediately where we are? Thus not only cells can be labeled by gradients, but probably specific genes are activated at a specific morphogen concentration. However how a cell can read a gradient remains an open problem. The expression of specific portions of the genetic code gives to cell it specialization its function, but the principle of the basic organization is still a mystery.

Let us go back to the cellular level where the elementary unit is the protein or a molecule. To maintain a cell functional, molecules have to go to the correct location and when they become not functional, they must be replaced. It has indeed been observed that proteins are constantly being replaced. Regulation of proteins is a fundamental process in the functioning of the cell, but it is not known how comes so few mistakes can be made, given that our entire body is being totally replaced every month?

Indeed, if the lifetime of a protein is of the order of one day, it is conceptually challenging to understand how the physiological function can be maintained for years. The protein turnover is a central question in cell biology and no clear mathematical framework for its description has been constructed so far. Actually, since proteins move from compartments by random motion, sometimes guided by filaments, it is unclear what mechanisms control the replacement of nonfunctional proteins and how proteins are guided to their correct locations.

The description of the protein trajectory is a problem in stochastic analysis and should help to describe cellular regulation. A typical problem can be to compute the

mean time a receptor takes to find a specific location [4]. A mathematical model of this scenario is needed.

Another example where the importance of protein trafficking can be found is in the neural cell communication. Neuronal communication relies on micro-contacts called synapses and is based on changes in the electrical activity, which are controlled by only few channels 50 to 100. A fundamental constraint is that the neuronal signal must be stable over time, especially if the synapses are understood as a memory checkpoints [3] (See figure 2, about random movement of receptor). In that context, how such a small number of receptors can be controlled precisely? And more important, how the synaptic connections can be maintained for years, if the lifetime of the receptor is about 24 hours? The proteins have to be replaced constantly and correctly. New scenarios inevitably come with new mathematical models to explain the accuracy of such processes [4]. But this is only the starting point.

From ionic channels (the doors of cells) to abstract coding

Ionic channels are probably the most fascinating doors [8] (figure 3), much more sophisticated than car or apartment doors. They open under specific conditions. A channel can select specific ions to let through (e.g., big or small). They can select an ion of a given charge and not another. This discrimination process is still largely unexplained.

To estimate the ionic currents flowing through a channel, a stochastic description of the ionic motion has been used to compute the ionic current for a given potential [1]. However, understanding ionic channels is insufficient for the understanding of how an action potential is generated at the cellular level. A new branch of statistical physics has emerged to integrate channels into a neuronal dendrite: the pioneers were Hodgkin and Huxley, about fifty years ago. Today, new experimental data have revealed that channel distributions and the membrane potential fluctuate, due to the synaptic activity. Thus we are inching our way closer to fundamental questions, without touching them. For example, how is sensory information encoded and processed at the level of dendrites and sensory inputs influence the neural organization? No mathematical framework for answering these questions exists at present. We need much more than Shannon's Information Theory (developed about fifty years ago) or Kolmogorov's theory of most efficient approximation of functions with a given accuracy. Certainly, the sensory information has to be encoded in a very efficient way and retrieved very quickly. At the same time, redundancy guarantees that this information is not lost by small perturbations.

The lesson drawn from the Hodgkin-Huxley model is that propagation equations of action potentials can be derived at the molecular level from channel dynamics. Preciously few models of biological process, with this quality, have been derived so far. A second lesson is that it takes talented biologists to derive such a mathematical model.

Neural networks and the representation of the external world

What about the scale of a neural network, where thousands, millions, even billions of neurons work together? At the network level, it is still unclear what is the relevant mathematical framework necessary to study how the neuronal signal is

processed, as well as how summations, multiplications, or any elementary operations underlie the fact that 2+2=4.

But let's stick closer to reality. Some circuitry has been detected experimentally. One would like to know the principles of how the brain is wired; are there any rules? The brain is not a telecommunication-wiring box, where the blue wire goes to blue and red-to-red. The problem is that nobody knows where the red wires are.

Combination of theory and experiments has revealed some of the complex organization of the visual cortex, where specific neurons in the visual cortex fire in responses to a visual stimuli, which is a bar presented with a given orientation (one neuron responds to a vertical bar and another to a horizontal one). If a bar rotates, the activity of neurons rotates around points, which appear as topological singularities, called pinwheel [10] (figure 4). Not only topology is necessary, but even finite dynamical systems are not enough to characterize the activity. These equations have to be replaced by integro-differentiable systems and even this is still not enough. Noise seems also to play a crucial role in the maintenance of the activity and in coding spatial information. Finally, some of the neurons which receive many inputs seem to work not in a linear regime, but in a state where the synapses seem always tired [11], phenomena known as synaptic depression (figure 5). This phenomena resembles to a filtering process which adapts to high frequency stimulations.

This is only the very beginning. Neurons communicate and generate specialized cortical areas, driven by external inputs and once again, no theoretical frameworks have been proposed to describe cortical plasticity [2] and predict the role of activity in changing the cortical maps. It is probably here that learning and memory are hidden. Words are missing to complete the description.

Microstructures and the limit of instrumental resolution

The theoretical approach has become a challenging tool, once the scientific questions have reached the border of instrumental resolution. For example, many interesting processes, such as cellular calcium dynamics, occur in a space of the order of 1-2 um and on a time scale of the order of milliseconds. Why this time scale is important? it is exactly at this time scale that most of cellular processes operate: transduction of the signal such as phototransduction, induction of synaptic plasticity and so on. Quantitative questions arise and for example, it turns out that we would like to know what is the quantity of calcium that crosses a microstructure such as a dendritic spine [6], after receptors are activated: how many chemical bonds, calcium ions will formed during their journey inside the spine. Now comes the problem of experimental limitations: experimentally exogenous buffers or fluorescent dye molecules are usually added to visual calcium, but as a result, the medium is significantly perturbed and it becomes very difficult to estimate the amount of calcium necessary to induce new changes that produce cellular changes underlying memory. Thus, it has become a very interesting and challenging problem to build the correct mathematical framework to be able to study, model and compute calcium in microdomains. At the present moment, a mathematical description of chemical reactions at a molecular level in micro-domains [5] can give only partial insight about the physiology of a cell at this level. What is exactly needed? First the geometry of the cell has to be modeled, which requires to use in general the theory of

partial differential equations. Then because few molecules might be involved, it is also important to use a stochastic description. Building good and convincing models requires in general a strong collaboration between biologists and mathematicians. If one succeeds in producing realistic models and convincing simulations, not only we can understand the cellular response to a stimuli but also how the cell function is generated at a molecular level. The quantitative role of each participating molecules can then be evaluated. Moreover, knocking out proteins or adding a drug can be analyzed, as well as understood.

Building such models and analyzing the new equations cannot be achieved by a fast reconversion of mathematicians, bored with the classical problems, on the contrary, the best mathematicians are needed. The situation is comparable to the effort made eighty years ago in quantum mechanics: Many physicists, theoretical chemists and mathematician formulated fundamental questions, they wrote the associated equations and obtained under approximations, the solutions. It is true that the price to pay to jump into modeling biological processes would be to renounce to the old Hilbert problems or more recently to the one million dollar problems, but it is the effort needed to formulate some of the new problems of our time.

Future and Perspectives

The time of mathematics in biology has just arrived. But the written bible is not in our bag yet, neither is the Civil Code of biology ready. We have to find the rules and write the laws of biology. This is a piece of work, which will give pleasure and color to good brains.

New mathematical models are needed, as well as a new conceptual framework for biological questions. A new generation of theoretical biologists is needed, trained in statistical physics, mathematical analysis, stochastic processes, differential geometry, partial differential equations, with a good overview of electrical engineering, chemical physics, and who had spent time in a real lab. Biology cannot be learnt from textbooks, and certainly not from mathematicians, but only from biologists. New methods and concepts have to be created to solve the new problems. Sticking to old concepts is a waste of time and money.

Pharmaceutical companies, and the industrial world in general, have to be ready to talk to some of the flakes called mathematicians. Mathematicians tend to think about new problems for hours, days, months, and even years, before they come up with something worthwhile. But when they do, it is a blast. Their environment has to be patient with them, because everybody benefits when mathematicians embrace science.

Acknowledgements: D. H thanks Z. Schuss and M Segal for comments.

References

[1] Barkai E, Eisenberg RS, Schuss Z.Bidirectional shot noise in a singly occupied channel. Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics. 199654(2):1161-1175

[2] Buonomano DV, Merzenich MM. Cortical plasticity: from synapses to maps. Annu Rev Neurosci. 1998;21:149-86. Review.

[3] Choquet D, Triller A.The role of receptor diffusion in the organization of the postsynaptic membrane.Nat Rev Neurosci. 2003 Apr;4(4):251-65.

[4] D. Holcman, Z. Schuss. 2004. Diffusion of receptors on a postsynaptic membrane: exit through a small opening, J. of Statistical Physics, 117, 5/6 191-230

[5] D. Holcman, Z. Schuss. 2005. A theory of stochastic chemical reactions in confined microstructures, Journal of Chemical Physics 122, 114710, 2005

[6] E. Korkotian, D. Holcman, M. Segal. 2004. Dynamic Regulation of Spine-Dendrite Coupling in Cultured Hippocampal Neurons, European Journal of Neuroscience, 20:2649-2663

[7] Kathie M. Bishop1, John L. R. Rubenstein2, and Dennis D. M. O'Leary, Distinct Actions of Emx1, Emx2, and Pax6 in Regulating the Specification of Areas in the Developing Neocortex, The Journal of Neuroscience, 2002, 22(17):7627-7638

[8] R. MacKinnon Science, Vol 306, 5700, 1304-1305 , 2004

[9] McLaughlin T, O'Leary DD, Molecular gradients and development of retinotopic maps.Annu Rev Neurosci. 2005;28:327-55. Review.

[10] Shmuel A, Grinvald A. Coexistence of linear zones and pinwheels within orientation maps in cat visual cortex. Proc Natl Acad Sci U S A. 2000;97(10):5568-73.

[11] M. V. Tsodyks H. Markram The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc Natl Acad Sci U S A ;94(2):719-23. Erratum in: Proc Natl Acad Sci U S A 1997;94(10):5495.

[12] Turing AM.The chemical basis of morphogenesis. 1953. Bull Math Biol. 1990;52(1-2):153-97; discussion 119-52.

Figure 1a: Morphogenetic gradient in the cortex

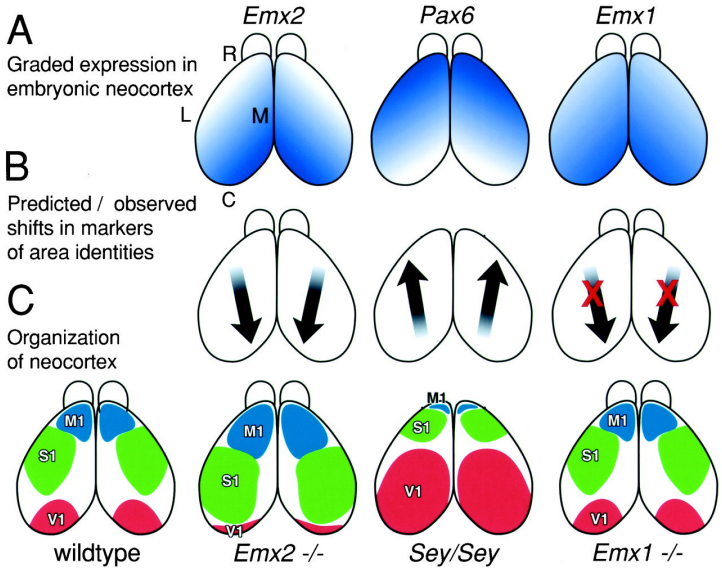

Figure 1a. Morphogenetic gradients are generated in cortical areas. In this paper, the authors demonstrated that over-expression or deletion of specific transcription factors modify the size of the cortical areas.

Diagrams are of dorsal views of the mouse neocortex. A, Graded expression patterns of the transcription factors Emx2, Pax6, and Emx1 across the embryonic neocortex. Emx2 and Emx1 are expressed in a high caudomedial to low rostrolateral gradient, whereas Pax6 is expressed in an opposing gradient.

B, Arrows indicate the direction of the predicted shifts in markers of area identity in Emx2, Pax6 (Sey/Sey), and Emx1 loss-of-function mutants, if these genes are involved in regulating arealization of the neocortex. The predicted shifts are observed in Emx2 and Pax6 mutants but not in Emx1 mutants (indicated by red X marks).

C, Organization of the mouse neocortex into areas predicted by our findings. These diagrams are not intended to show the exact sizes and shapes of the primary neocortical areas but rather to depict the disproportionate changes in area size and positioning, or no changes, in arealization in the different mutants. These predicted organizations suggested by analyses of gene markers and area-specific thalamocortical projections are limited because the Emx2 and Pax6 mutants die on the day of birth, before areas become anatomically and functionally distinct, and thalamocortical projections do not develop in Pax6 mutants. (adapted from [7]).

Figure 1b: Gradient in cells

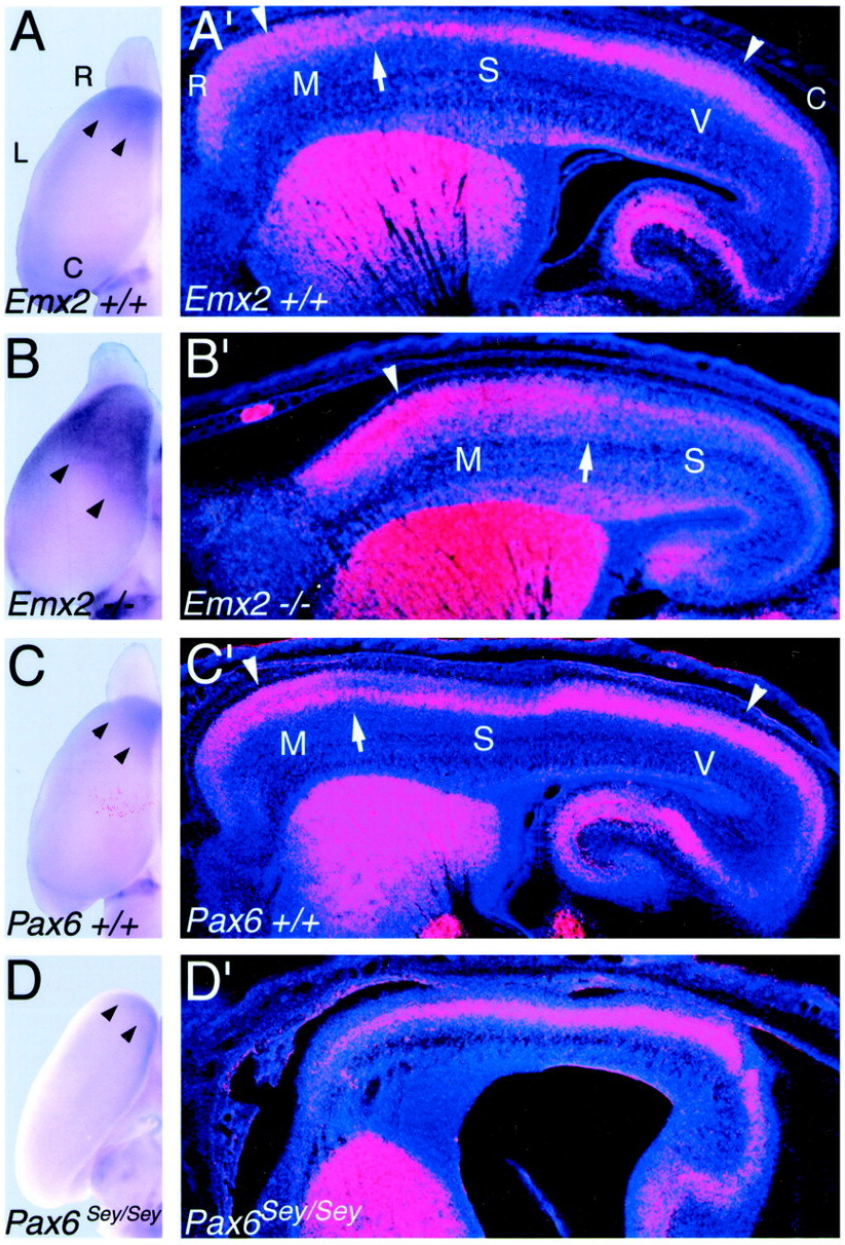

Figure 1b Gradients in the cortex: Opposing changes in the expression domains of the cadherin, Cad8 in Emx2, and Pax6 (Sey/Sey) mutants. A-D, Dorsal views of whole mounts of P0 cortical hemisphere of Emx2 wild-type (+/+) (A), Emx2 mutant (/) (B), Pax6 wild-type (+/+) (C), and Pax6 (Sey/Sey) mutant (D) processed for in situ hybridization using digoxygenin-labeled

riboprobes for Cad8. Arrowheads mark the caudal limit of the rostral expression domain of Cad8 in the superficial layers, which is characteristic of motor areas. A'-D', Sagittal sections through E18.5 brains of mice of the corresponding genotypes as in A-D, processed for in situ hybridization using S35-labeled riboprobes for Cad8. Marked are the approximate locations of the motor (M), somatosensory (S), and visual (V) areas in the wild-type cortex and their shifted locations in the Emx2/ cortex suggested by the expansion and caudal shift in patterns of Cad8 expression, which are unique in each of these areas in wild-type mice. Arrowheads in A'-C' mark rostral and caudal expression domains in the superficial layers characteristic of motor and visual areas; in comparison, expression is substantially diminished in the superficial layers of the intervening somatosensory area.

Arrows in A'-C' mark the presumed border between motor and somatosensory areas. The Cad8 expression rostral to these arrows is the expression that is evident in the whole mounts shown in A-C. This superficial rostral expression domain characteristic of motor areas is essentially absent in Pax6 (Sey/Sey) mutants (D, D'). The wild type and mutants in each pair are age-matched littermates. (adapted from [7]).

Figure 2

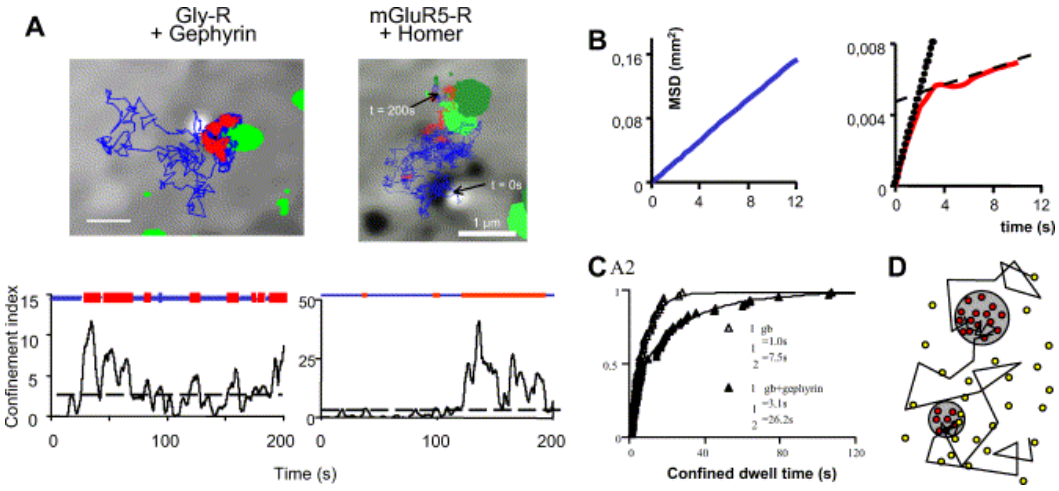

Figure 2 Movement of a receptor on the surface of a neuronal cell. Receptor diffusion in neurons measured by single particle tracking.
This review reports recent findings about single receptor trafficking at the surface of neurons. Receptor trajectory are made of intermittent piece which corresponds to the transition from free to confined Brownian motion. Confinement might be due to the entrance of the receptor inside microdomain, where bindings with some molecules can occur. This view suggested that receptor are highly motile and thus the number of receptors at a synapse fluctuates, which raises the question of the reliability of synaptic transmission.

a | Superimposed image of the trajectory of 500 nm beads bound to glycine receptors (GlyRs; left panel) and the metabotropic glutamate receptor (mGluR5; right panel) with the fluorescent image (green) of green fluorescent protein (GFP)-tagged gephyrin (left) and Homer (right). Periods of free diffusion (blue lines) and confinement (red lines) in the trajectories are detected using a confinement index (lower panels). **b** | Plots of the average mean squared displacement (MSD) function during periods of free diffusion (left panel) and confinement (right panel) for GlyRs. Note the difference in both shape and amplitude of the MSDs. The curved shape of the MSD is characteristic of movement in a confined space. **c** | Plots of the cumulative distributions of dwell times for GlyR in the confined state in the presence (red triangles) and absence (orange triangles) of gephyrin. Distributions are fitted with the sum of two exponentials, the time constants of which increase in the presence of gephyrin. **d** | Superimposed image (upper panel) of the trajectory of a 500 nm bead bound to the glutamate type II receptor (GluR2) and of presynaptic terminals stained with FM1-43 (green). The lower panel shows a plot of the confinement index versus time for the corresponding trajectory. **e** | Model for the exchange of receptors between dispersed (light pink circles, freely diffusing) and

clustered (dark pink circles, confined movement) states. Receptors might also diffuse within the clusters. The dashed line shows the trajectory of a given receptor. Clustered receptors at postsynaptic domains are circled red. **f** | Model for the exchange of receptors (dark pink) between postsynaptic domains through lateral diffusion (arrows). Part a modified, with permission, from Ref. 120 © (2002) Society for Neuroscience. Parts a–c modified, with permission, from *Nature Neuroscience* Ref. 123 © (2001) Macmillan Magazines Ltd. Part d modified, with permission, from *Nature* Ref. 122 © (2002) Macmillan Magazines Ltd. (origin [3]).

Figure 3

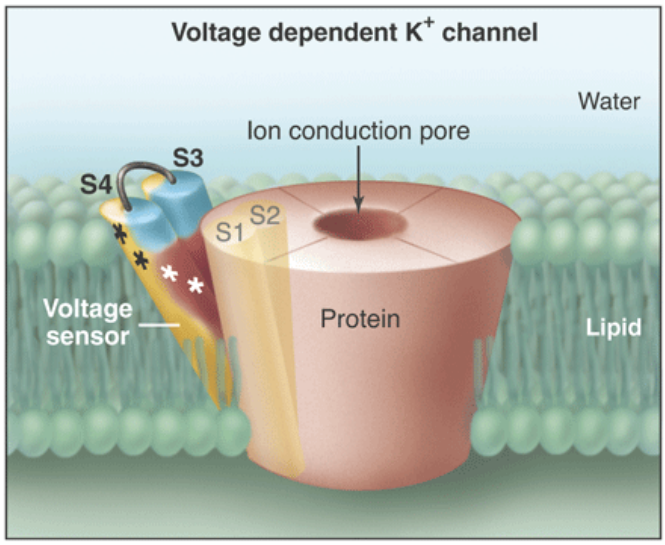

Figure 3. A model of the KvAP K+ channel with a voltage-sensor at the protein-lipid interface.

Understanding the mechanism of channel gating is fundamental in the cellular communication. Channels and exchangers regulate most of the cellular exchange between the cell and the extracelullar medium. The figure represent the relation of the K+ channel's voltage sensor to the lipid membrane, water surfaces, and pore (origin [8]).

Figure 4: Pinweels

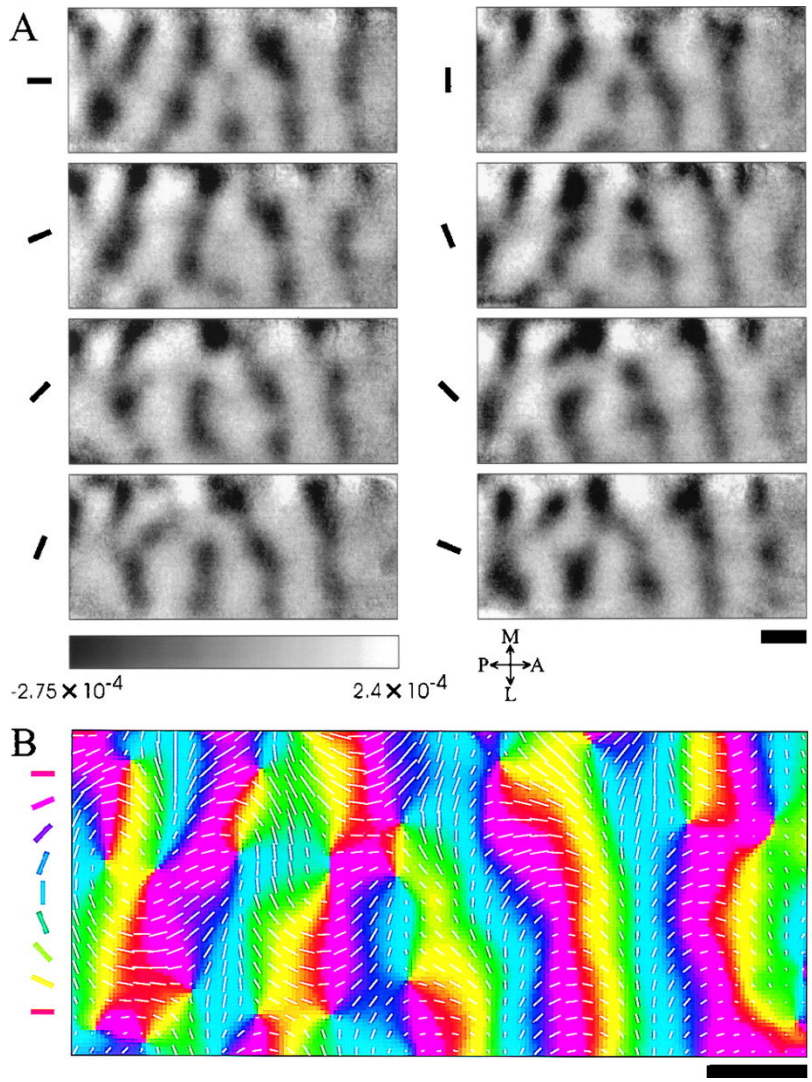

Figure 4. Relationship between the patterns of response to individual oriented gratings and the organization of orientation preference.

(A) Patterns of response to individual grating patterns. Within the posterior part of the presented image, the cortical zones activated by a single oriented grating have a beaded appearance. In contrast, more anteriorily, the pattern of activated domains is elongated approximately parallel to the ML axis.

(B) Angle representation of the orientation map from the same part of cortex. Within the posterior part of the presented image, the preferred orientation is organized in elongated iso-orientation domains, located around orientation centers. More anteriorily, zones with iso-orientation preference are organized as slabs elongated parallel to the ML axis. The values of preferred orientation change continuously in an approximately linear manner along the PA axis. (origin [10]).

Figure 5 Neural code

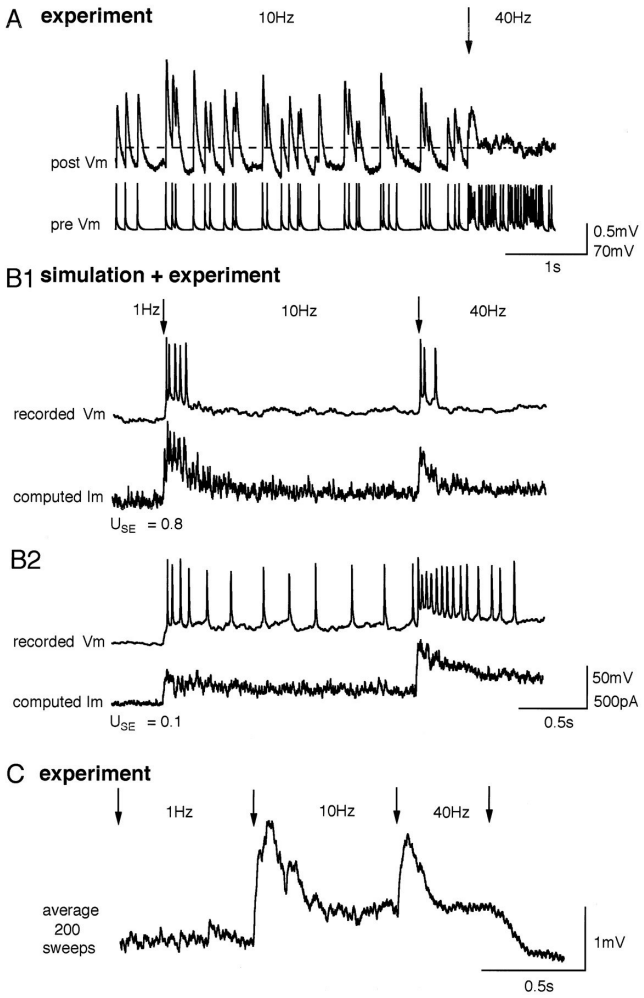

Figure 5. Neural code. Signaling of synchronized transitions in the activity of a population of presynaptic neurons.

(A) Experimentally recorded EPSPs generated by a Poisson spike train undergoing transition (indicated by arrow) from 10 to 40 Hz. The average membrane potentials before and after the transition (indicated by dashed line) were equal to the third decimal point.

(B1) Simulated postsynaptic current, generated by Poisson spike trains, of a population of 500 presynaptic neurons with synchronous transitions from 1 to 10 Hz and then to 40 Hz, together with the response of a pyramidal neuron when the simulated synaptic current was injected into the soma. A population signal emerged as the number of neurons in the presynaptic pool was increased.

Parameters of the model are the same as in Fig. 3. (B2) The same as B1 but with lower value of USE and twice as large ASE. (C) Average voltage response recorded from a postsynaptic neuron after stimulating the presynaptic neuron with the sequence of 200 different Poisson spike trains undergoing the same transitions as in B (origin [11]).