P systems under uncertainty: the case of transmembrane proteins

Stefano Aguzzoli[a], Ioan I. Ardelean[b], Daniela Besozzi[c], Brunella Gerla[d], Corrado Manara[d]

[a]Università degli Studi di Milano
Dipartimento di Scienze dell'Informazione
Via Comelico 39, 20135 Milano, Italy
E-mail: aguzzoli@dsi.unimi.it

[b]Institute of Biology of the Romanian Academy
Centre of Microbiology
Splaiul Independenţei 296
PO Box 56–53, Bucharest 060031, Romania
E-mail: ioan.ardelean@ibiol.ro

[c]Università degli Studi di Milano
Dipartimento di Informatica e Comunicazione
Via Comelico 39, 20135 Milano, Italy
E-mail: besozzi@dico.unimi.it

[d]Università degli Studi di Salerno
Dipartimento di Matematica e Informatica
Via Ponte don Melillo, 84084 Fisciano (SA), Italy
E-mail: {bgerla,cmanara}@unisa.it

DRAFT

1 Introduction

P systems, initially proposed in [9], are a class of distributed and parallel computing devices inspired by the architecture and functioning of living cells. It has to be stressed that these systems were not intended to be a model of the cell, instead their purpose was to investigate some computational features which can be abstracted from the cellular biology. Anyway, in several recent works the framework of P systems has been used to define models of specific cellular processes or structures (see an updated bibliography at http://psystems.disco.unimib.it).

In this work we are mainly interested in investigating the modelling power of P systems, and in the extension of the framework, in order to deal with *imprecise* biological information. Indeed, many aspects of the cell functioning are still unknown to biologists and source of *uncertainty*. In these cases

uncertainty does not emerge not only because of the lack of knowledge about the occurrence of some event (a mathematical model for this is provided by the probability theory), but it is due to the *vagueness*, that is the capability to use imprecise information to describe the cell functioning. *Fuzzy set theory* and *fuzzy logic* could be useful in this framework. Hence, we propose to approach the problem of integrating fuzzy techniques in P systems to deal with uncertainty in biological systems.

In particular, we will briefly analyze the case of transmembrane proteins. First we consider the P model proposed in [2] for simulating the activity of mechanosensitive channels, where the uncertainty is related to some relevant parameters of the model, and we give some suggestions and open problems concerning how to use fuzzy tools within that model.

Then we propose the investigation of the global behavior of populations of (equal or different) transport proteins, such as ATP-powered pumps, ion channels and transporters, and the study of flux dynamics of the corresponding transported molecules. In this case, the source of uncertainty is concerned with the local position in the cellular membranes of transport proteins and with the local distribution of substances, since the rate and extent of transport is also influenced by the concentrations of substances and by the electric potential that exists across the membrane.

The motivation of this research is to use P systems for modelling the functioning of specific cellular structures and phenomena, having as final goal the production of useful and relevant tools for biologists, and hence motivating further cooperations between scientists working in the areas of P systems and Microbiology. Indeed, the design of appropriate software simulators, based on the corresponding P models, would provide an easier way to check both the effectiveness and the correctness of the models, and hopefully become a tool for testing known data, predicting unknown scenarios and returning meaningful information to biologists.

2 The case of mechanosensitive channels

In this section we first give a biological description of mechanosensitive channels with large conductance (MscL, in short), then we report a sketch and a brief overview of the P system presented in [2] for modelling the functioning of MscL analyzed during patch clamping experiments. The reader is referred to [2] fur further notions and details, as well as for the references therein. Finally, we propose and motivate a fuzzy extension of that model.

2.1 Biological description of mechanosensitive channels

Mechanosensitive channels are homopentameric transmembrane proteins gated by mechanical forces. Their physiological function consists in the protec-

tion against severe osmotic downshifts, since they allow the rapid exit of different chemicals and the sudden decrease of the osmotic pressure inside the cell. This event is fundamental for bacterial cell because, when the turgor pressure is too large, the integrity of the cell can be damaged by disruption of cell wall and plasma membrane, followed by cell death. The increase in the pressure exerted against the cellular membrane may be due to natural environmental conditions (e.g., rain falling) or to a suction applied during artificial patch clamping experiments. In correspondence to these distinct situations, in [2] two models for the description of the activity of MscL in *E. coli* and in other prokaryotes are presented. In Section 2.2 we will only consider the *in vitro* model, corresponding to patch clamping experiments.

When the cellular membrane is submitted to a mechanical stretch, it experiences an increase in the *membrane tension*, which causes the progression of the channel from the steady-state closed conformation to an expanded – but still closed – conformation and, through some subconducting open states (which correspond to the breaking away of the sections in the homopentameric structure), to the fully open state (see Figure 1). According to a biological model proposed in [11], we consider the following conformations and their relative notations:

- the closed conformation, denoted by **C**;
- the expanded closed conformation, denoted by **CE**;
- the first subconducting open conformation, denoted by **SO1**, where only one subunit (out of five) is open;
- the second subconducting open conformation, denoted by **SO2**, where two (out of five) subunits are open;
- the third subconducting open conformation, denoted by **SO3**, where three (out of five) subunits are open;
- the fourth subconducting open conformation, denoted by **SO4**, where four (out of five) subunits are open;
- the fully open conformation, denoted by **O** (where all five subunits are open).

Data collected from patch clamping experiments on *E. coli* [12] correspond to the following real values, or interval of values, for the membrane tension (measured in dyne/cm):

(*i*) $t_C \in [0, 10)$, when no suction is applied to the patch membrane;

(*ii*) $t_{CE} = 10$, when a suction is applied to the patch membrane, the membrane tension increases and MscL is in the closed expanded substate;

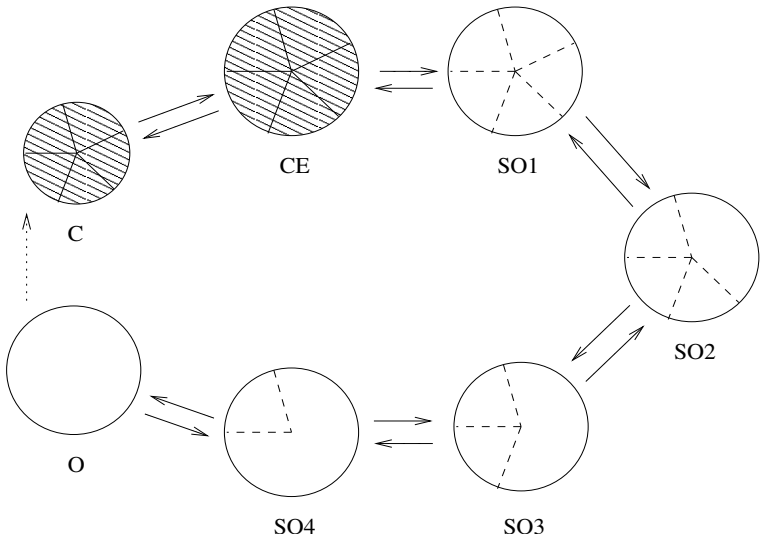

Figure 1: Transitions in MscL: from the closed to the fully open conformation, via subconducting states.

(*iii*) $t_{SO1}, t_{SO2}, t_{SO3}, t_{SO4} \in (10, 13)$, when the channel is partly open (solutes and water pass from the internal region to the external medium) and shows a flickering through subconducting states;

(*iv*) $t_O = 13$, when MscL is fully open, chemicals and water continue to pass from the internal region to the external medium;

(*v*) $t_L \geq 14$, when the applied suction is so high to cause the membrane lysis.

Similarly, we can consider the following conductivity values of the subconducting and open states:

- the conductivity of the subconducting state **SO1** is $0.25 \cdot 3.5$nS, that is 0.875nS;
- the conductivity of the subconducting state **SO2** is $0.56 \cdot 3.5$nS, that is 1.96nS;
- the conductivity of the subconducting state **SO3** is $0.74 \cdot 3.5$nS, that is 2.59nS;
- the conductivity of the subconducting state **SO4** is $0.89 \cdot 3.5$nS, that is 3.115nS;
- the conductivity of the subconducting state **O** is 3.5nS.

2.2 A P system model for mechanosensitive channels

As said before, MscL act as transmembrane mechanoelectrical switches, opening in response to lipid bilayer stretch and deformations and converting a mechanical stress of the membrane into gating transitions. The channel open probability, as well as the dynamic of close–to–open transitions, are functions of the membrane tension, an essential parameter described by means of a variable label attached to the membrane [2]. Hence, the evolution rules not only intervene in the transformation and communication of objects, but also in the modification of the label, which is to be interpreted as a key descriptor of the channel status. This is a new interpretation of the membrane label, which becomes a fundamental component of the system used to describe a biological significant counterpart (the status of the channel, in this case), and not just an identifier of a membrane in the membrane structure.

The definition of the variable membrane parameter (the *tension*) in the *in vitro* model is based on the real data reported in Section 2.1. The tension label assumes real positive values in the finite set of labels $Tension = \{t_C, t_{CE}, t_{SO1}, t_{SO2}, t_{SO3}, t_{SO4}, t_O, t_L\}$.

The solutions inside and outside the cell are described by considering an external *environment* (in short, Env) and an inner *region* (in short, Reg): the environment is made of solutes (symbols from a given alphabet V_{chem}) and water molecules (each denoted by a symbol $w \notin V_{chem}$); the internal region consists of objects over the same alphabet of the environment, and we assume that no other processes take place inside the cell. The semi-bracket notation is used to denote the membrane (labelled with the tension parameter $t \in Tension$) which separates the external environment and the internal region, that is: $Env\ [_t\ Reg$.

Transitions among tension values are due to the changes in the pressure applied to the patch membrane and simulated in P systems by means of a new type of evolution rules. Namely, an *in vitro environmental rule* describes a change in the pressure parameter p due to external actions, which can happen at any time in the environment and cannot be controlled by any component of the system. We write $\langle p, apply\rangle[_t \longrightarrow_{prob} [_{t'}$ for some $p \in \mathbf{R}, t, t' \in Tension, prob \in [0,1] \subset \mathbf{R}$, to denote any environmental rule which introduces the action of the parameter p, has consequences on the membrane tension value and is applied according to the associated probability value $prob$.

We give an example of *in vitro* evolution rules and their meaning:

$$\langle p, apply\rangle[_{t_C} \longrightarrow_{prob=0.01} [_{t_{CE}} \quad \text{for some } p \ll 40$$

$$\langle p, apply\rangle[_{t_C} \longrightarrow_{prob=0.99} [_{t_C} \quad \text{for some } p \ll 40$$

If the membrane tension is equal to t_C and the applied pressure has a value $p \ll 40$mmHg, then the conformation of the MscL is more likely to

remain unchanged (second rule) because the applied suction is not enough to trigger the channel activation; though, we also model the passage to the expanded state with a very low probability (first rule).

The complete set of evolution rules and corresponding explanation can be found in [2]. Moreover, there are reported the simulations *in silico* performed by means of the complex systems simulators *EdnaCo* [5]. The observed quantities emerging from the simulations (that is, they were not explicitly programmed in the simulation) are the tension, the conductance and the current) . Obtained results appear to be in line with the general biological phenomena and thus offer biologists a challenge to verify results by actual laboratory experiments. See [2] for output pictures of simulations and for further details.

Moreover, note that the P model presented in [2] was constructed for a single MscL, anyway it can be considered consistent with the analysis of a population of mechanosensitive channels. In this case, it suffices to consider the same model for many channels, but using different sets of probability values associated to rules, as well as, possibly, to different membrane tension values.

2.3 Towards a fuzzy P model for mechanosensitive channels

Thanks to the promising results of simulation *in silico*, and in order to give a more realistic and fine description of the functioning of MscL, we propose to extend the *in vitro* P model by including some fuzzy tools. From Section 2.1 we know real values or interval of values of the parameter tension (and conductivity) when the channel is closed, closed expanded, and so on, and in Section 2.2 we have seen how to attach a label t to the membrane to denote the current state of the channel. The values of t in the set $Tension$ have been considered consistent with the real values of membrane tension measured in dyne/cm, but no continuous transition has been assumed among channel conformation. More precisely, one takes care only of discrete time steps, hence the channel state is initially closed (with membrane label t_C and some real value of t_C in $[0, 10)$) and, at the next step, it might be closed expanded (with membrane label t_{CE} and value equal to 10). Actually, in the cell there is a gradual transition from one conformation of the channel to the next one, hence it might be better (and closer to reality) to study these situations together with the introduction of *membership functions*, which describe the parameters tension and conductivity. Two possible sets of membership functions for the case study are depicted in Figure 2 and 3.

For simplicity, here we used triangular and trapezoid membership functions. In general, a membership function is a function $\mu_F : U \to [0, 1]$ that assigns to every $u \in U$ a degree of membership $\mu_F(u) \in [0, 1]$ to F, where U is a universe of objects and F is a fuzzy set. The latter is completely determined by the set of tuples $F = \{(u, \mu_F(u)) \mid u \in U\}$. Like a crisp set, a

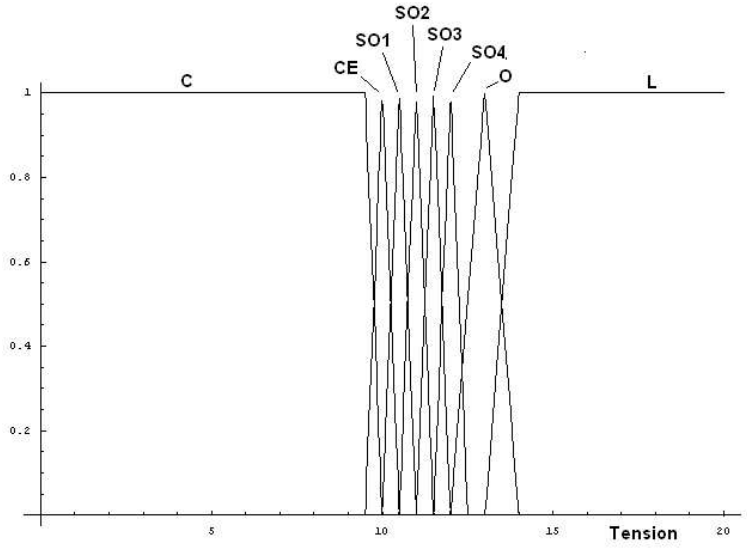

Figure 2: An example of membership functions for the tension.

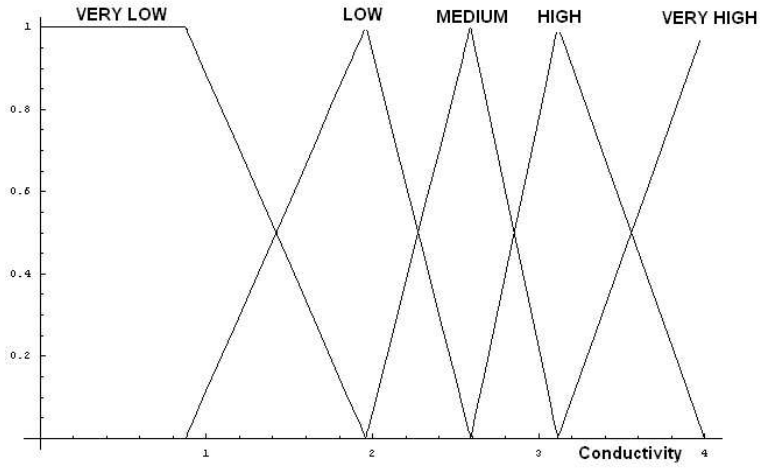

Figure 3: An example of membership functions for the conductivity.

fuzzy set can be used to describe the value of a variable. In fuzzy set theory, the variable is linguistic and its values are described both qualitatively by a linguistic term (the symbols $C, CE, SO1, SO2, \ldots$) and quantitatively by the corresponding membership function. This knowledge representation is used in fuzzy rule-based inference. In order to define this technique and to relate it with a P model, we need to introduce the following definitions.

An *atomic fuzzy proposition* has the form

$$X \text{ is } F$$

where X is the linguistic variable, 'is' stands for 'has the property of being' and F is a fuzzy set that describes a property. Based on this definitions and the use of *linguistic connectives* as 'and', 'or' and 'not' one can construct more complex fuzzy propositions.

A *fuzzy rule* is symbolically expressed as

$$\text{if } <\text{fuzzy proposition}> \text{ then } <\text{fuzzy proposition}>.$$

The first proposition (the antecedent) describes an observed condition, while the second one (the consequent) describes a conclusion that depends on the antecedent. An example of fuzzy rule is:

$$\text{if } X \text{ is } H \text{ then } Y \text{ is } A.$$

The basic steps of fuzzy rule-based inference are the following (see [13, 4] for more details):

1. Fuzzy matching: Calculate the degree to which the input data match the condition of the fuzzy rules.

2. Rule base processing: Calculate the rule's conclusion base on its matching degree and combination in a final conclusion.

3. Defuzzification: For applications that need a crisp output this step converts a fuzzy conclusion in a crisp one.

Usually, in the theory of fuzzy control, given a set of input and output linguistic variables we obtain a fuzzy set as the result of the inference (that is, the conclusion). In the case of mechanosensitive channels, the tension can be considered as an input linguistic variable, while the conductivity can be considered as an output variable. Moreover, another expected output should be the control action that triggers the application of evolution rules defining the P model for MscL. Hence we are looking for an higher level of description: how to integrate the standard fuzzy approach with P systems?

3 The case of membrane transport proteins

The phospholipid bilayer of cellular membranes is essentially impermeable to most water-soluble molecules and ions. Hence, the passage across membranes of many biochemical substances (amino acids, glucide, ions) has to be mediated by transmembrane proteins, which are usually selective with respect to the transported substances. On the contrary, there exist other substances (gases and small uncharged molecules) which can directly cross the phospholipid bilayer by passive diffusion, down their concentration gradients. In this section we are only interested in transmembrane proteins.

The three major types of transport proteins are *ATP-powered pumps*, *ion channels* and *transporters* (uniporter, symporter, antiporter), which all exhibit a high specificity for the transported substances and differs with respect to the rate of transport and to the mechanism of action. ATP-powered pumps use the energy of ATP hydrolysis to move ions or small molecules against a chemical concentration gradient or electrical potential (the process is known also as active transport). Ion channels simultaneously transport multiple water molecules or many (specific) ions down their concentration or electric potential gradients, at a very rapid rate. Some of them are usually open (for instance, the potassium-specific channel), others are usually closed and open only in response to specific signals. In contrast, transporters bind only one (or a few) molecules at the same time, then a conformational change of the protein allows the transport of such molecules across the membrane. Among transporters, uniporters move one molecule at a time down its concentration gradient, while symporters and antiporters couple the passage of one type of molecule (or ion) against its concentration gradient to the passage of a different type of molecule (or ion) down its concentration gradient. For more notions about membrane transport proteins the reader can consult [8, 1].

In all cases of transport proteins, the rate and extent of ion transport is influenced by the ion concentrations on the external and internal sides of the membrane, and by the electric potential that exists across the membrane, as well as by the biological structure and chemical properties of the proteins. On the other side, the ionic gradients and electric potential across the membrane drive many biological processes. For instance, the conduction of an electric impulse down the axon of neurons is mediated by opening and closing of sodium, potassium and calcium channels; in most cells, an increase in cytosolic calcium concentration is a fundamental regulatory signal, while sodium concentration gradient power the uptake of amino acids and other molecules.

From these considerations it is clear how important would be, from the biological point of view, an investigation of the global behavior of a population of transport proteins and the dynamics of transported molecules. For instance, consider the possibility of analyzing a population of channels of the

same type (e.g., a population of sodium-potassium ATPases) or of different types, possibly "competing" for the transported molecules or ions (e.g., a population of calcium ATPases, calcium-sodium antiporters, sodium-potassium ATPases, potassium channels). The natural phases of investigation for this problem consist in first defining a good model for each transport protein of interest, and then designing a software simulator which allows the study of flux dynamics of transported molecules. This kind of study is consistent with the guidelines of Systems Biology [6, 7].

In this work we are mainly interested in finding a good framework for modelling biological structures and processes. As reported in Section 2, P systems have been proved valid for the modelling and the simulation of a very particular type of transmembrane channel, and our aim in Section 2.3 consisted in extending the known P model with fuzzy tools and techniques, in order to gain the highest resemblance to reality. In the case of transport proteins, whose functionality depends on concentration gradients and voltage, a fundamental aspect of the biological reality has to considered: the *"locality"*, which stands for the local physical conditions and for the notion of nearness. Indeed, in a cell the concentration of ions or molecules is not uniformly distributed all over the cellular membrane, but there can exist small local zones with a higher concentration with respect to others with a lower concentration. Thus, some elements in the population might be more active, others working at lower rates, others even resting, according to (1) the position of each (type of) transport protein in such areas (characterized by high or low concentration gradients and voltage), (2) the respective position of the surrounding transport proteins (of the same or different type) and (3) the distance existing between each transport protein and the protein-specific molecules to be transported.

To define appropriate models for simulating the behavior of channels under different local conditions, it might be useful to approach the analysis with the help of fuzzy techniques, in the same direction of Section 2.3.

References

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter. *Molecular biology of the cell.*, 4th edition, Garland Science, New York, 2002.

[2] I.I. Ardelean, D. Besozzi, M.H. Garzon, G. Mauri, S. Roy, P system models for mechanosensitive channels, submitted, 2004.

[3] D. Besozzi, G. Ciobanu, A P system description of the sodium-potassium pump, submitted, 2004.

[4] D. Driankov, H. Hellendoorn and M. Reinfrank. *An Introduction to Fuzzy Control.* Springer-Verlag, Berlin, 1993.

[5] M. Garzon, D. Blain, A. Neel, Virtual Test Tubes for Biomolecular Computing *Journal of Natural Computing*, 3:4 (2004), in press.

[6] T. Ideker, T. Galitski, L. Hood, A new approach to decoding life: systems biology, *Annual Reviews Genomics Hum. Genet.*, 2, 2001, 343–372.

[7] H. Kitano, Systems biology: a brief overview, *Science*, 295, 2002, 1662–1664.

[8] H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, J. Darnell. *Molecular cell biology*, 4th edition, W.H. Freeman and Co., New York, 2000.

[9] Gh. Păun, Computing with membranes, *Journal of Computer and System Sciences*, 61(1), 2000, 108–143 (See also *Turku Center for Computer Science-TUCS Report* No 208, 1998, www.tucs.fi).

[10] Gh. Păun, *Membrane Computing. An Introduction*, Springer-Verlag, Berlin, 2002.

[11] S.I. Sukharev, M. Betanzos, C.S. Chiang, H.R. Guy, The Gating Mechanism of the Large Mechanosensitive Channel MscL, *Nature*, 409, 720–724, 2001.

[12] S.I. Sukharev, W.J. Sigurdson, C. Kung, F. Sachs, Energetic and Spatial Parameters for Gating of the Bacterial Large Conductance Mechanosensitive Channel, Mscl, *J. Gen. Physiol.*, 1999, 113, 525–539.

[13] J. Yen and R. Langari. *Fuzzy Logic. Intelligence, Control, and Information.* Prentice-Hall, Upper Saddle River, 1999.