

A Uniform Framework of Molecular Interaction for an Artificial Chemistry with Compartments

Tooru Watanabe, Kazumasa Koizumi, Koji Kishi, Masaki Nakamura, Keiji Kobayashi,
Mitsuyoshi Kazuno, Yoshikazu Suzuki, Yuuki Asada and Kazuto Tominaga
School of Computer Science
Tokyo University of Technology
1401-1 Katakura, Hachioji, Tokyo 192-0982 Japan
Email: tomi@acm.org

Abstract— Cells have acquired a complex membrane structure through evolution. Each compartment separated by membranes has its specific functions, and the coordination of those functions constitutes the behaviour of the living system at a biochemical level. Therefore, it is necessary for an artificial chemistry to have an ability to express membrane structures if it is to be used to study such behaviour. In this paper, we propose a way of incorporating nested membrane structures in an artificial chemistry. This extension is achieved mainly by adding the following two components to the formalism: one is membranes, which is not only a boundary between two reaction pools but is also a reaction pool itself; the other is membrane molecules, which are embedded in a membrane and change their forms by recombination rules. The introduction of membrane molecules enables the formalism to deal with the movement of molecules across a membrane in a similar way as reaction between normal molecules. Using the extended artificial chemistry, we modelled two functions in natural living systems, namely, signal transduction and intracellular transport, and thereby show the advantages of the present approach to introducing compartments to artificial chemistries.

I. INTRODUCTION

A living system does not only have a variety of molecules inside but has intricate structures at several levels of magnitude. In particular, an eucaryotic cell has intracellular compartments, such as the nucleus, the endoplasmic reticulum, the Golgi apparatus and mitochondria, which have different roles and cooperate with each other. The organization of the compartments is very interesting, and there is no doubt that they must play essential roles in the activity of this kind of living system.

To study such systems, artificial chemistries are one of the good candidates for a research methodology to use. An artificial chemistry is an abstract model of a system whose dynamics is based on chemical reaction [1]. Because a major part of the dynamics of living systems is driven by chemical reaction, artificial chemistries are inherently compatible with them, and various artificial chemistries have been built and studied. Indeed, some of the systems that can be categorized in artificial chemistries, such as P systems [2], have been applied to study the dynamics of systems with compartments.

We proposed an artificial chemistry based on pattern matching and recombination [3], and have been constructed abstract models for some partial systems of life such as DNA

replication, transcription to mRNA, translation to protein [4], and metabolic pathways [5]. While these models capture the mechanisms of the target systems quite well at the given levels of abstraction, each of them is modelled as a system that has only one pool of molecules. The artificial chemistry cannot express structures that comprise compartments, as those found in a cell, surrounded by membrane. In our previous work we proposed a way of expressing multiple reaction pools by connecting them with pipes [6], but the realization is rather for mechanical implementation of nanosystems and is not very compatible with the cell structure. If the artificial chemistry is to be used to model such natural systems, a more plausible way of expressing their structures is desired.

In this paper, we give an extension of our artificial chemistry so that it can model such structures in a more natural manner. Specifically, we introduce the concept of membranes and membrane objects to our artificial chemistry to have it express the functions of compartments naturally.

II. AN ARTIFICIAL CHEMISTRY WITH COMPARTMENTS

In this section, we illustrate an extension to our artificial chemistry [3] that can express multiple compartments in a system.

A. Elements and Objects

An *element*, which corresponds to an atom in natural world, is denoted by a capitalized alpha-numeric string such as Aa, D and Gh1. An *object*, which corresponds to a molecule in natural world, is a stack of sequences of elements (Fig. 1).

The objects in Fig. 1 are denoted by 0#AaBb/ and 0#CcD/1#EGh1/. The number just before # is the starting position of the line which is relative to the first line's starting position. The position is measured by the number of elements. The first line's starting position is always zero.



Fig. 1. Objects

B. Patterns

A *pattern* matches or does not match an object. A pattern can have two kinds of *wildcards*.

One is an *element wildcard*, which matches any element. This type of wildcard is denoted by a number enclosed by angle brackets (i.e., “<” and “>”) such as <0>; we call this number the *ID* of the wildcard. If multiple element wildcards are contiguously placed and their IDs are continuous in ascending or descending order, the description can be abbreviated using dots. For example, a pattern 0#A<0><1><2><3>F/ can be denoted by 0#A<0..3>F/, and 0#Z<5><4>W/ by 0#Z<5..4>W/.

The other type of wildcard is a *sequence wildcard*, which matches any sequence of zero or more elements. This type of wildcard comes only to either end of a line. A sequence wildcard is denoted by its ID and an asterisk enclosed by angle brackets: one at the left end of a line is denoted as <*1>, and one at the right end is denoted as <2*>.

A pattern is denoted in the similar way as that for objects. The pattern 0#Aa<0>/ matches an object 0#AaBb/ (the wildcard <0> matches the element Bb), and the pattern 0#<0*>/1#<*1>E<2>/ matches an object 0#CcD/1#EGh1 (<0*> matches CcD, <*1> matches the null sequence, and <2> matches Gh1). Note that the length of a sequence wildcard is treated as zero in the notation.

C. Recombination Rules

A *recombination rule* transforms a group of objects on its left-hand side into a group of objects on the right-hand side. A recombination rule is denoted like a chemical equation but in terms of patterns as follows.

$$0\#Aa<0>/ + 0\#<1*>/1\#<*2>E<3>/ \rightarrow 0\#Aa<0>><1*>/-1\#<*2>E<3>/ \quad (1)$$

When (1) is applied to objects 0#AaBb/ and 0#CcD/1#EGh1, these objects disappear and the object 0#AaBbCcD/-1#EGh1/ is produced. In this case, <0>, <1*>, <*2> and <3> matches Bb, CcD, an empty sequence and Gh, respectively.

The recombination rule conserves elements just like a chemical reaction does.

D. Cubicles, membranes and membrane objects

Objects are held in a multiset called the *working multiset*, where they are transformed by recombination rules.

A *cubicle* has its own working multiset, and is surrounded by a *membrane* (Fig. 2). Cubicles have parent-child relations. In Fig. 2, Cubicle 1 is the parent of Cubicle 2 and Cubicle 3, and therefore Cubicle 2 and Cubicle 3 are the children of Cubicle 1. A cubicle has its own set of recombination rules. A cubicle corresponds to a cell or a organelle.

A membrane has its own working multiset as well. Objects held in a membrane are called *membrane objects*. Unlike cubicles, a membrane has no recombination rules. Membrane

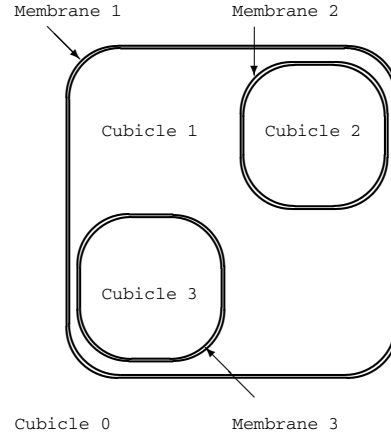


Fig. 2. Structure of a system.

objects in a membrane are transformed by the recombination rules of the adjacent cubicles to the membrane. For example, in Fig. 2, membrane objects belonging to Membrane 2 are transformed by the recombination rules of Cubicle 1 and Cubicle 2.

Membrane objects are referred to by patterns preceded by “^” or “_”. A pattern with “^” in a recombination rule of a cubicle matches membrane objects in the membrane that surrounds the cubicle; a pattern with “_” matches membrane objects in one of the membranes that surround child cubicles of the cubicle. In other words, the symbol “^” refers to membrane objects in the outer membrane, and “_” refers to those in one of the inner membranes. For example, if a membrane surrounds Cubicle 2 in Cubicle 1 (Fig. 3) and the both cubicles have the following rules

$$_0\#\text{Mem}/ + 0\#A/ \rightarrow _0\#\text{Mem}A/ \quad (2)$$

$$\wedge 0\#\text{Mem}/ + 0\#B/ \rightarrow \wedge 0\#\text{Mem}B/ \quad (3)$$

then (2) of Cubicle 1 can be applied to the membrane object 0#Mem/ in the membrane, and (3) of Cubicle 2 can be applied to it.

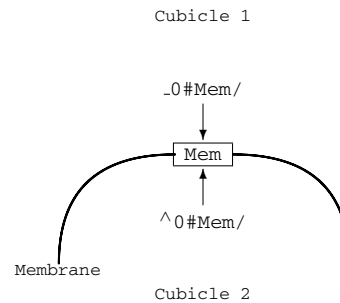


Fig. 3. Reference to a membrane object by a pattern.

E. Dynamics

A *system* consists of a group of cubicles and membranes. When a cubicle is regarded as a node and a membrane as an edge, the cubicles and the membranes of a system constructs a tree structure. The cubicle corresponding to the root is called the *root cubicle*. The root cubicle in Fig. 2 is Cubicle 0. The root cubicle represents the outside of the defined world and has no object and no recombination rule.

A system is interpreted nondeterministically as follows.

- 1) Give cubicles and membranes their initial working multisets.
- 2) Select a cubicle.
- 3) Apply one recombination rule that belongs to the selected cubicle to a collection of objects to which the rule can be applied.
- 4) Go to Step 2.

III. APPLICATIONS OF THE EXTENDED ARTIFICIAL CHEMISTRY

In this section, we show two applications of the extended artificial chemistry. One is a model for signal transduction, and the other is intracellular transport.

A. Modelling signal transduction

Each cell of a multicellular organism cooperates to work with others cells [7, chapter 15]. This cooperation is achieved by sending and receiving *signals* among cells. A signalling cell sends signal molecules to target cells, and a function of the target cells is triggered. When signal molecules reach a target cell, they are received by *receptor proteins* on the surface (i.e., embedded in the plasma membrane) of the cell, and the cell-surface receptors are activated to induce further reaction such as phosphorylation of other molecules inside the cell. This step converts the form of information to another form. In general, the conversion of forms of signal is called *signal transduction*.

The *Ras-MAP-kinase cascade* is a signalling pathway and itself is a signal-transduction process that conveys signals from the surface of a cell to its nucleus; it is used to control cell proliferation and differentiation [7, chapter 15]. When a tyrosine kinase receptor receives a signal molecule called *epidermal growth factor* (EGF), it activates Ras protein. Then the Ras-MAP-kinase cascade is triggered. The active Ras protein activates a MAP-kinase-kinase-kinase (named *Raf*), which in turn activates a MAP-kinase-kinase (*Mek*), which then activates a MAP-kinase (*Erk*). The active Erk goes into the cell nucleus and phosphorylate gene regulatory proteins; this causes changes in gene expression.

Now we model the Ras-MAP kinase cascade with our artificial chemistry. We use three cubicles representing the outside of cell (Cubicle 1), the cell cytoplasm (Cubicle 2) and cell nucleus (Cubicle 3) (Fig. 4). The initial objects are shown in Table I. Since the final product of the cascade, *Erk*, phosphorylates various kinds of molecules in the nucleus, we give a generic object *Grp* (standing for gene regulatory protein) instead of providing specific molecules.

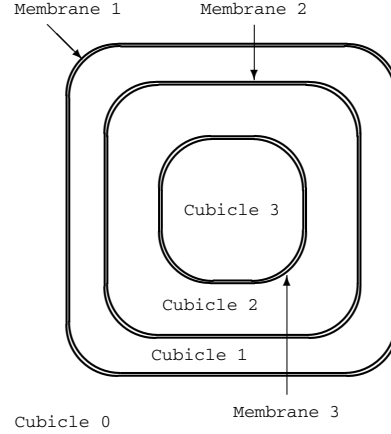
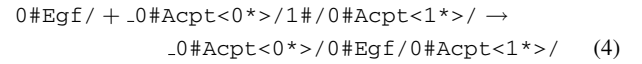


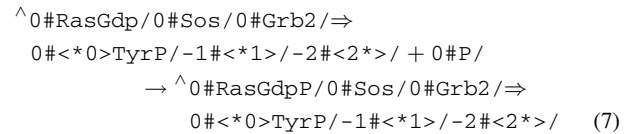
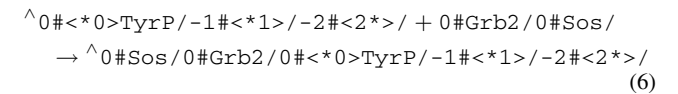
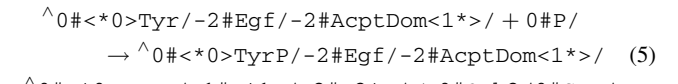
Fig. 4. System structure for modelling the Ras-MAP kinase cascade.

The system works as follows. Note that the whole set of rules is not shown below; the total number of the actual rules is 26.

First, EGF outside the cell binds to a pair of two tyrosine kinase receptors on the membrane (4).



Then the tyrosine kinase receptors get phosphorylated (5) and activate Ras proteins via a Grb-2/SOS complex ((6) and (7)). Expression of some objects are too long to fit in a line, so they are folded (indicated by “ \Rightarrow ”).



The active Ras protein triggers the Ras-MAP-kinase cascade. The cascade activates Raf (8), Mek (9), and Erk (10) in this order.

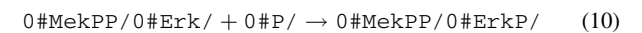
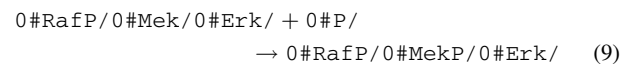
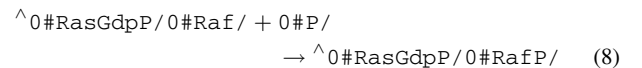
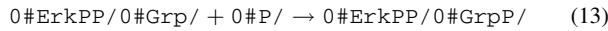
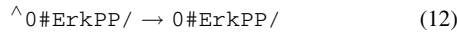
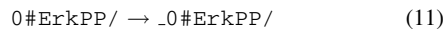


TABLE I
INITIAL WORKING MULTISSETS FOR MODELLING THE RAS-MAP-KINASE CASCADE

site	molecule	object	quantity
cubicle 1	EGF	0#Egf/	10
membrane 2	tyrosine kinase receptors	0#AcptDomTyr/1#/0#AcptDomTyr/	10
	Ras protein	0#RasGdp/	10
cubicle 2	phosphate	0#P/	100
	Grb-2/SOS complex	0#Grb2/0#Sos/	10
	Raf	0#Raf/	10
	Mek binding to Erk	0#Mek/0#Erk/	10
cubicle 3	Grp	0#Grp/	10
	phosphate	0#P/	10

The activated Erk enters the cell nucleus ((11) and (12)) and phosphorylates gene regulatory protein (13).



We developed a simulator for this extended artificial chemistry. Fig. 5 is a screen snapshot from the execution of the above description. The bottom-left window shows receptors, which are ready to be recombined by (7).

B. Modelling Intracellular Transport

A kind of protein functions in a specific compartment of a cell. Therefore, proteins must be transported to the appropriate compartment after they are produced [7, chapter 12]. This *intracellular transport* is achieved by various ways; one of them is based on the sequence of amino acids the protein includes. The sequence that decides where the protein is moved is called *signal sequence*. If a protein has the sequence, it is transported to the site.

A peroxisome is a cell organelle, and is one of the target site of intracellular transport based on signal sequences. Its process is as follows [8, chapter 6].

Proteins carried to peroxisomes have the signal sequence called *PTS1* or *PTS2*. Table II shows the amino-acid sequences

TABLE II
SIGNAL SEQUENCES FOR TRANSPORT TO PEROXISOME

name	sequences of amino acid
PTS1	-(Ser/Ala)-(Lys/Arg/His)-Leu-
PTS2	-(Arg/Lys)-(Leu/Val/Ile)-X-X-X-X-(His/Gln)-(Leu/Ala)-

of PTS1 and PTS2, where X represents any amino acid and “(Aaa/Bbb/Ccc)” means either of Aaa, Bbb and Ccc works. A protein including PTS1 binds to a receptor called *Pex5* and forms a complex, and one including PTS2 forms another complex by binding to a receptor called *Pex7*. Then the complexes *Pex5-PTS1* and *Pex7-PTS2* combine to form one complex, and it reaches to a protein translocator that comprises *Pex14* on the surface of a peroxisome. The translocator takes the proteins in the peroxisome, leaving the receptors outside. After that, PTS2 is detached from the protein.

We modelled the intracellular transport to peroxisome using our artificial chemistry as follows. Three cubicles are provided to represent the cell cytoplasm (Cubicle 1) and two peroxisomes (Cubicles 2 and 3) (Fig. 6). The initial objects are shown in Table III; the underlined portions are signal sequences.

The system operates as follows. First, *Pex5* binds to a protein including PTS1 (14) and *Pex7* binds to one including PTS2 (15).

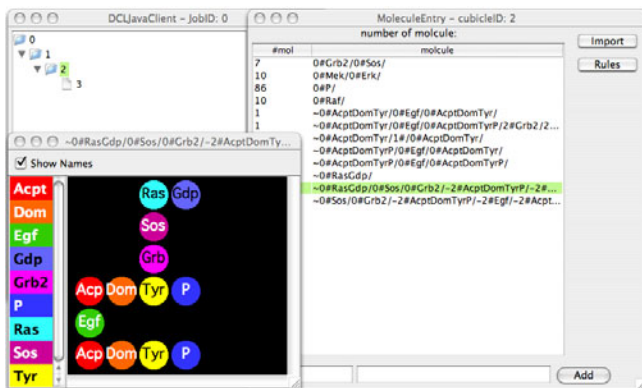


Fig. 5. Execution of the Ras-MAP kinase cascade.

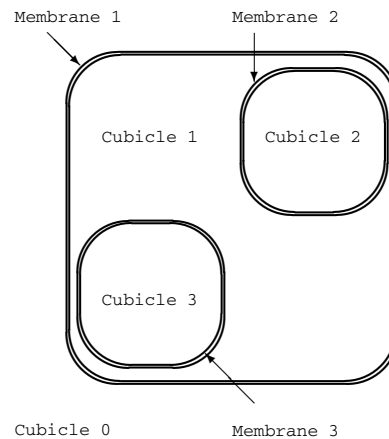


Fig. 6. System structure for modelling intracellular transport to peroxisomes.

TABLE III
INITIAL WORKING MULTISSETS FOR MODELLING THE INTRACELLULAR TRANSPORT TO PEROXISOME

site	molecule	object	quantity
cubicle 1	protein with PTS1	0#MetIleAspAlaArgLeuThrAsn/	10
	protein with PTS2	0#MetProSerPheArgLeuThrTrpTyrTrpThrHisLeuLysLys/	10
	Pex5	0#Pex5/	10
	Pex7	0#Pex7/	10
membrane 2	Pex14	0#Pex14/	10
membrane 3	Pex14	0#Pex14/	10

$$0\#<^*0>AlaArgLeu<1^*>/ + 0\#Pex5/ \\ \rightarrow 0\#<^*0>AlaArgLeu<1^*>/0\#Pex5/ \quad (14)$$

$$0\#<^*0>ArgLeu<1..5>HisLeu<6^*>/ + 0\#Pex7/ \\ \rightarrow 0\#Pex7/0\#<^*0>ArgLeu<1..5>HisLeu<6^*>/ \quad (15)$$

Since there are six (2×3 , see Table II) patterns of sequence that work as PTS1, we give six rules as (14). Similarly, we provide 24 ($2 \times 3 \times 2 \times 2$) rules as (15). Note that the number of rules for PTS2 does not increase due to Xs in Table II because we used element wildcards ($<1..5>$) for those positions.

The Pex5-protein complex combines with the Pex7-protein complex (16), and the product binds to a protein translocator (denoted simply by $0\#Pex14/$) on the membrane of one of the peroxisomes (17).

$$0\#<^*0><1^*>/0\#Pex5/ + 0\#Pex7/0\#<^*2><3^*>/ \\ \rightarrow 0\#<^*0><1^*>/0\#Pex5Pex7/1\#<^*2><3^*>/ \quad (16)$$

$$0\#<^*0><1^*>/0\#Pex5Pex7/1\#<^*2><3^*>/ + _0\#Pex14/ \\ \rightarrow _0\#<^*0><1^*>/-1\#Pex14Pex5Pex7/1\#<^*2><3^*>/ \quad (17)$$

Then Pex5 and Pex7 are removed (18) and the transported proteins are released into the peroxisome (19).

$$_0\#<^*0><1^*>/-1\#Pex14Pex5Pex7/1\#<^*2><3^*>/ \\ \rightarrow 0\#Pex5/ + 0\#Pex7/ \\ + _0\#<^*0><1^*>/0\#Pex14/0\#<^*2><3^*>/ \quad (18)$$

$$^0\#<^*0><1^*>/0\#Pex14/0\#<^*2><3^*>/ \\ \rightarrow 0\#<^*0><1^*>/ + 0\#<^*2><3^*>/ + ^0\#Pex14/ \quad (19)$$

Finally, the PTS2 sequence is excised from the protein in the peroxisome (20); twenty-four rules are given for all the possible PTS2 sequences.

$$0\#<^*0><1>ArgLeu<2..6>HisLeu<7^*>/ \\ \rightarrow 0\#ArgLeu<2..6>HisLeu/ + 0\#<^*0><1><7^*>/ \quad (20)$$

The total number of rules is 58, but most of them are for the multiple possibilities of PTS1 and PTS2 sequences.

Fig. 7 shows a screen snapshot of execution. The window shows a complex of Pex5, Pex7 and two proteins in the cytoplasm, which is ready to bind to a translocator on the surface of a peroxisome.

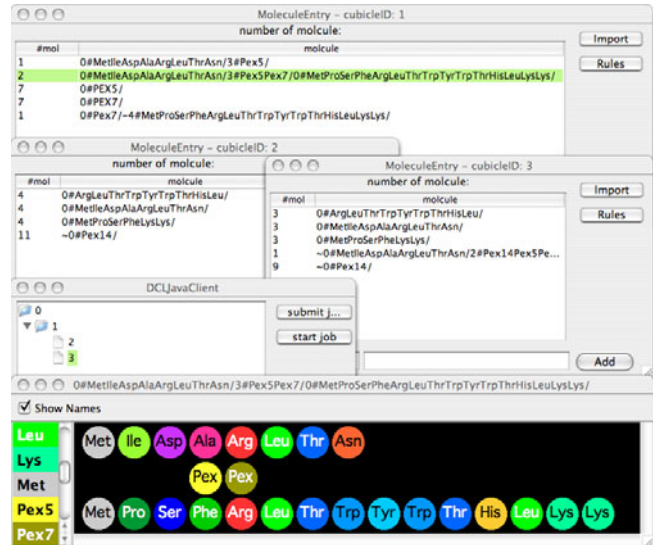


Fig. 7. Execution of the model for intracellular transport for peroxisome.

IV. DISCUSSION

The applications given in the previous section illustrated how some kinds of natural phenomena can be described in our extended artificial chemistry in an abstracted and simple way. It has now compartments that have their own reaction pool, and the objects can move from one pool to another through membranes. Compared to our previous work [6], which used “pipes” to move object from one place to another, the present formulation is more compatible with natural systems that use membranes to separate their compartments.

The modelling that gave the examples was straightforward. Molecules and molecular complexes with whatever functions were identified uniformly as objects, and all the dynamics of a system was described in terms of recombination. We think this view is in a good accordance with natural chemical dynamics: even if a molecule is embedded in a membrane, it interacts with others by chemical reaction.

One notable feature of our extended artificial chemistry is that it allotted a reaction pool to a membrane. This allows any object to be kept in the membrane, as well as being held in a normal compartment (i.e., cubicle). This feature helps modelling a natural system with proteins embedded in a membrane.

Transportation of an object from one cubicle to another

can be implemented either by using membrane objects as mediators (Pex14 in Section III-B) or by permeation without any help of membrane objects (0#ErkPP/ going through Membrane 3). The former can model channel-mediated or carrier-mediated transport, while the latter may be useful to model simple diffusion through a membrane or for the case the mediator can be ignored.

V. COMPARISON WITH RELATED WORKS

Many formalisms have been proposed to model natural membrane structures. Among them, related closely to artificial chemistries and actively studied these days are P systems [2]. P systems are mathematical models for computing using membranes, and some studies have applied them to modelling natural biochemical systems [9]. There are two significant differences between P systems and our extended artificial chemistry in relation to membranes. One is that P systems hold objects only in compartments while ours hold them both in compartments and membranes. The other is that, when objects are to be moved from one compartment to another, P systems specify the target compartment by rules. In contrast, our artificial chemistry does not explicitly specify the destination by rules; what compartment an object moves to is decided by the existence of membrane objects of a specific type. In the example of intracellular transport (Section III-B), Pex5-Pex7-protein complexes formed in the cell cytosol (Cubicle 1) can move to either of the two peroxisomes (Cubicles 2 and 3); it works in the same way if more peroxisomes are added to the system, without giving any additional rules to specify which peroxisome an object should move to. This approach also has an advantage in modelling natural dynamic systems. In our artificial chemistry, it is easy to equip different functions with different membrane objects. This means that functions of a membrane object can be changed by the application of recombination rules. We think the extended artificial chemistry offers useful flexibility for modelling a membrane whose functions change dynamically such as one having complex protein carriers, compared to a variant of P system with symports/antiports [10, chapter 4] for their functions are defined as rules.

Brane Calculi [11] are another formalism for modelling the activities of systems that have membranes. Although they deal with nested membrane structures as ours does, their main focus is on interactions between membranes, not on objects involved. A membrane and its contents are expressed as a formula, and interactions between membranes are represented by a computation, during which the form of the formula changes, performed by a fixed set of generic rules. In contrast, our artificial chemistry regards objects as the primary entities, and multisets of objects — not a single expression — are processed by user-defined rules.

Both P systems and Brane Calculi can handle the structural changes of a system: P systems have a function to dissolve a membrane; the generic rules of Brane Calculi modify the membrane structure in various ways. On the other hand, our artificial chemistry does not yet have functions to change

the structure of a system. Such functions as compartment merging, division, creation and dissolving will be useful to model activities of natural systems. Adding them should be our future work.

Hutton studied the construction of membranes in artificial chemistries [12]. The work focuses on how membranes can be formed from raw materials under given virtual physics, while our study presumes the existence of membranes and deals with no physics.

VI. CONCLUSION

In this paper, we proposed an extension to our artificial chemistry in order to provide it with the ability to express nested membrane structures. The extended artificial chemistry can deal with compartments surrounded by membranes (called cubicles), and objects that are embedded in membranes (called membrane objects). We gave two example applications of this artificial chemistry, namely, a model for signal transduction and one for intracellular transport, and thereby demonstrated that the formalism can model interactions among objects and a membrane quite well. Membrane objects naturally represent proteins embedded in cell membranes; they play their roles by being recombined by recombination rules in the same way as objects in cubicles. This uniformity does not only contribute to simplicity and understandability of the formalism but will also offer good flexibility for modelling systems in which the functions of membranes change dynamically.

REFERENCES

- [1] P. Dittrich, J. Ziegler, and W. Banzhaf, "Artificial chemistries — a review," *Artificial Life*, vol. 7, no. 3, pp. 225–275, 2001.
- [2] G. Păun, "Computing with membranes," *Journal of Computer and System Sciences*, vol. 61, pp. 108–143, 2000.
- [3] K. Tominaga, "A formal model based on affinity among elements for describing behavior of complex systems," Department of Computer Science, University of Illinois at Urbana-Champaign, Tech. Rep. UIUCDCS-R-2004-2413, Mar. 2004.
- [4] K. Tominaga, K. Kobayashi, T. Watanabe, K. Koizumi, and K. Kishi, "An approach to constructing qualitative models in computational cell biology using an artificial chemistry based on pattern matching and recombination," in *Proceedings of the 10th International Symposium on the Simulation and Synthesis of Living Systems (ALIFE X)*, L. M. Rocha, L. S. Yaeger, M. A. Bedau, D. Floreano, R. L. Goldstone, and A. Vespignani, Eds. Cambridge, Massachusetts, USA: The MIT Press, 2006, pp. 172–177.
- [5] Y. Suzuki and K. Tominaga, "Describing metabolic pathways using an artificial chemistry based on pattern matching and recombination," in *Proceedings of the 11th International Symposium on Artificial Life and Robotics (AROB '06)*, M. Sugisaka and H. Tanaka, Eds., 2006, published as CD-ROM.
- [6] K. Tominaga, "Modelling DNA computation by an artificial chemistry based on pattern matching and recombination," in *Proceedings of the Workshop on Artificial Chemistry and Its Applications, part of the 9th International Conference on the Simulation and Synthesis of Living Systems (ALIFE9)*, 2004.
- [7] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, 4th ed. New York, NY: Garland Science, 2002.
- [8] Y. Yoneda, Ed., *Saiboumai Yusou ga Wakaru*. Tokyo: Yodosha, 2002, in Japanese.
- [9] G. Ciobanu, G. Păun, and M. J. Pérez-Jiménez, Eds., *Applications of Membrane Computing*. Berlin: Springer, 2006.
- [10] G. Păun, *Membrane Computing — An Introduction*. Berlin: Springer, 2002.

- [11] L. Cardelli, "Brane calculi, interactions of biological membranes," in *Computational Methods in Systems Biology*, ser. Lecture Notes in Bioinformatics, V. Danos and V. Schachter, Eds., vol. 3082. Berlin: Springer, 2005, pp. 257–278.
- [12] T. J. Hutton, "Making membranes in artificial chemistries," in *Proceedings of the Workshop on Artificial Chemistry and Its Applications, part of the 9th International Conference on the Simulation and Synthesis of Living Systems (ALIFE9)*, 2004.