Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 333 (2005) 1–4
BBRC
www.elsevier.com/locate/ybbrc

Breakthroughs and Views

Modeling directed ligand passage toward enzyme active site by a 'double cellular automata' model

Sayed-Amir Marashi, Reza Behrouzi *

Department of Biotechnology, Faculty of Science, University of Tehran, Enghelab Avenue, Tehran, Iran

Received 11 April 2005
Available online 28 April 2005

Abstract

It is recently proposed that directed passage of ligand on the surface of enzymes may play an important role in the process of enzyme activity, as a result of decreasing the required steps of random walking of the ligand toward the active site. Here, we revisited the approach applied by others, where a cellular automaton is designed to simulate the behavior of a ligand molecule traveling toward the active site of an enzyme. Since a cellular automaton plane is topologically equivalent to a torus surface, we recommended the use of a 'double cellular automata' to model globular proteins. With the boundary conditions applied, our model is topologically identical to a sphere. It was shown that using this model, even fewer steps are needed for a molecule to attend the active site. This assumption can lead to more realistic results in the modeling of surfaces with spherical topology.

Keywords: Cellular automata; Protein surface; Perfect enzyme; von Neumann neighborhood; Moor neighborhood

The centerpiece of enzyme-substrate interaction is the proper collisions of the two molecules. Therefore, the rate of an enzymatic reaction is affected by the translational diffusion coefficient of the smaller molecule, normally the substrate. This effect is only revealed if the catalytic efficiency of the enzyme exceeds the delivery rate of the substrate to the enzyme active site, so that the reaction becomes "diffusion controlled." These so-called perfect enzyme, possess such a sophisticated biochemical machinery that they can instantly transform the delivered substrate to the product [1].

Interestingly, there are reports of enzymes whose association constants exceed the values that are predicted by Smoluchowski equation [2]. This equation defines the maximum rate constant for the enzyme-substrate encounter. On the other hand, it has been shown that diffusion of the substrate from the bulk solution to the enzyme active site is too slow to permit the coordination of heterogeneous metabolic processes in living systems [3]. Such observations imply the need for improvements in the classic view of active site and substrate collisions.

To date, several models have been proposed to simulate this behavior. Some of them try to find the answer in the surface properties of proteins in that they assume the existence of some special "trajectories" on the surface of the proteins which direct (guide) the substrate to the enzyme active site [4,5]; the substrate may spend some time interacting with the protein surface, before it is guided to the active site. Therefore, a much larger area of the protein surface, than the active site itself, may act as a substrate collector.

To evaluate the consequences of these theories, cellular automata have been used [4–6]. Cellular automata (CA) are mathematical models, first introduced by J. von Neumann as formal models for self-reproducing organisms. Today, they are widely used as numerical

* Corresponding author. Fax: +98 21 649 1622.
E-mail address: behroozi@khayam.ut.ac.ir (R. Behrouzi).

0006-291X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2005.04.107

solutions to partial differential equations and as models for chemical, physical, and biological systems [7,8]. They are regular arrangement of discrete cells that any of them may have several but finite states. The state of each cell in any given time depends only on its previous state and its neighbor cells. Change of a cell's state may be based on probabilistic or deterministic rules.

In simple simulations which have been used for modeling ligand passage toward the active site [4–6], the state of each cell is either "occupied" or "empty." The cells may represent single amino acids or areas of the protein surface. This is the basis of "update rules" of the cells, i.e., the rules that affect the ligand movement to each cell. In these models, a ligand in the (i, j) cell can go to one of $\{(i+1, j), (i-1, j), (i, j+1), (i, j-1)\}$ cells. In case $i-1$ (or $j-1$) is not a positive value, it will be replaced by n, i.e., the ligand 'jumps' to the right (or bottom) border of the grid because of the infinity of the surface. A similar replacement approach (but with the value one) is applied to the situations where $i+1$ (or $j+1$) exceeds n.

With all these movement and boundary rules taken into account, the cellular automata will have the topological properties of the surface of a torus (Fig. 1). This impairs the efficiency of this model to mimic the surface properties of proteins, which in most known cases, are closed surfaces (like a sphere).

To obtain the properties of a globular protein (as of a sphere), we made use of a 'double cellular automata.' By this model we mean two simultaneously updated CA with special movement rules as described below. The substrate may move to each cell in these cellular automata, which is an immediate neighbor of the current cell. In a cellular automaton of size n, whenever $i-1$ (or $j-1$) is less than zero, it is replaced by i (or j), but the cellular automaton on which the ligand is placed is changed; i.e., the ligand moves from one cellular automaton to another. The same procedure is performed whenever $i+1$ (or $j+1$) exceeds n (Fig. 2).

A software program was written to perform the ligand passage simulations using the above-mentioned model (the program was developed in Managed C++ and is available upon request). For the results to be comparable to previously published data, we used the surface parameters as those that had been proposed by Ghaemi et al. [6]. Briefly, in this model, each cell represents one surface amino acid, and it holds a hydropathy index between 0 and 8, corresponding to most hydrophobic and most hydrophilic side chains, respectively. This scheme has been adopted from Kyte and Doolittle [9]. The simulator program assigns a value between 0 and 99 to each cell, directly proportional to its hydropathy and its neighborhood situation. This value represents the number of water molecules in interaction with that amino acid. The procedure is that first a random number between 0 and 99 is assigned to each cell. Then the assigned number of nine cells in a 3×3 block is added up and then redistributed over the cells, based on their hydropathy indices. Then the 9-cell block is shifted one cell down and right, and the procedure of summation and redistribution is repeated until a static pattern of water molecules is reached.

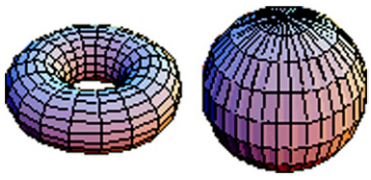

Fig. 1. Topology of a torus (left) and a sphere (right).

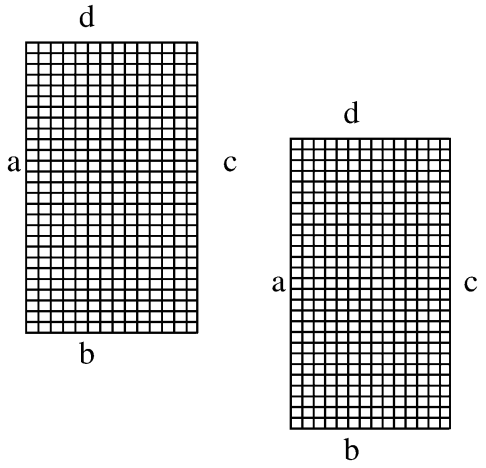

Fig. 2. Boundary conditions in a double cellular automata model. Whenever the moving element passes through each boundary (i.e., a, b, c or d), it moves to the equivalent cell in the other cellular automaton on the same boundary.

For the number of cells to be relatively the same, we used either a double CA, each of size 19×19 or 27×14. These double cellular automata are, respectively, 7 and 27 cells larger than the original 27×27 single cellular automata used by the Ghaemi et al. [6].

We also extended our simulation to the so-called Moor neighborhood, where a ligand in the cell (i, j) may move to any of $\{(i-1, j-1), (i-1, j), (i-1, j+1), (i, j-1), (i, j+1), (i+1, j-1), (i+1, j), (i+1, j+1)\}$ (Fig. 3).

The ligand passage was simulated 5000 times using each of neighborhood types. In each round of simulation, a different set of random numbers were generated and the proper hydrodynamic scale was created for it. The similarity of all the reported procedures and variables in [6] to ours was strictly watched. All the four types of protein surface models that were introduced in [6] were created analogously by the double CA. These forms are uniform with the hydropathy index of 4, random distribution, uniform with the value of 4, and random distribution but with diagonal, mid-vertical, and mid-horizontal lines with the value of 8 (see Fig. 4).

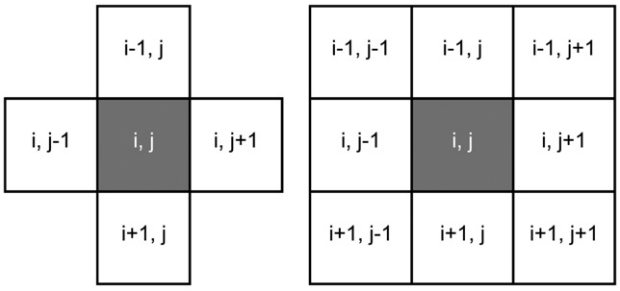

Fig. 3. von Neumann neighborhood (left) and Moor neighborhood (right).

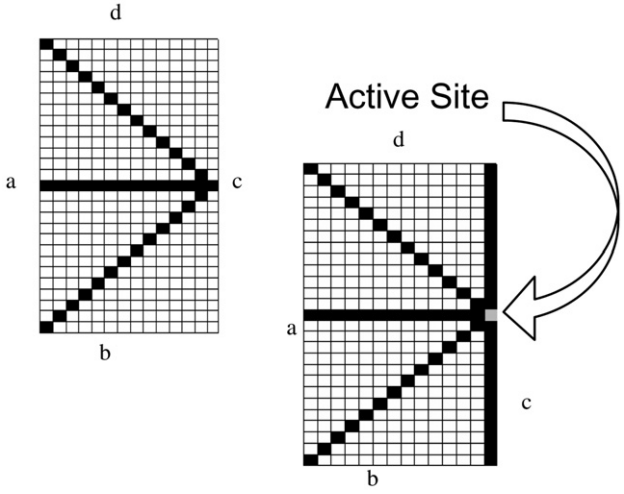

Fig. 4. Directed passage in the uniform and random context. The cells depicted with dark color possess the hydropathy index 8, while other cells are assigned 4 in the uniform context, or random integers from the 0 to 8 interval in the random context.

Table 1 represents the comparison of the results of simulations and the original results reported in [6]. It must be noted that due to the random nature of ligand movement on the cellular automata, the travel lengths will be skewed toward large values. This behavior affects the mean of travel length simulations. Although this effect is negligible in movement in uniform and directed passage cellular automata, it becomes a concern in randomly filled ones.

To maintain the reproducibility of the results, we found all the "outlier" and "suspected outlier" values of travel length simulations using the method described by Hogg and Tanis [10]. In this method, we obtain the first and third quartiles of the travel length simulations and define the inter quartile range (IQR) as the difference between the first and the third quartiles. Any value which is greater than 1.5 times the IQR is a suspected outlier and any one greater than three times the IQR is an outlier. These values were both omitted and not used to calculate the mean travel lengths. One can neglect substrate molecules with very large travel lengths toward enzyme active site and consider them as inert. Therefore, it is rational to remove the effect of these paths from our modeling, because we are just concerned about those substrate molecules whose paths toward the enzyme active site are facilitated by the enzyme surface.

As it may be seen, there is a significant decrease in travel length in the double cellular automata model compared to a single cellular automaton. The results are even more remarkable when we note that the number of cells in our 'double CA' is more than those used in [6].

Bearing in mind the increased similarity of the double CA model to the protein surface, it can be judged that the effect of the existence of ligand passage directions on the protein surface is significantly more than what had been thought before. One can observe that in our model there is a 2.9 decrease in travel length from uniform distribution to directed passage in uniform context, while the decrease in travel time in single cellular automaton model was only 1.2-fold [6]. The same pattern is seen in random distribution relative to directed passage in random context: the decrease in our model is 4.0-fold while it is 1.3-fold in single cellular automaton model [6]. The observed differences in the two double CA models (i.e., 19×19 and 27×14) may be interpreted as a simple consequence of the difference in the number of cells, since the results are approximately the same in the simulation results of the two models.

It is also noteworthy that increasing the neighborhood size from 4 to 8 cells (Moor rather than von Neumann neighborhood [8]) decreases the travel length even

Table 1
Mean travel lengths averaged over 5000 simulations

Experiment	Uniform	Random	Directed in uniform context	Directed in random context
Single cellular automaton [6] (27×27)	2142	2522	1778	1984
Double CA/von Neumann neighborhood (19×19)	1151	2281	—	—
Double CA/von Neumann neighborhood (27×14)	1236	2337	429	579
Double CA/Moor neighborhood (19×19)	812	1309	—	—
Double CA/Moor neighborhood (27×14)	884	1402	389	431

To maintain the similarity between our double CA and the CA model presented in [6], we simulated directed passage only in double CA of size 27×14.

more. We may assume that as the 8-cell neighborhood is a better model of a circle around a center point than the 4-cell model, it is a more realistic approach toward the movement of a ligand on the protein surface.

Acknowledgments

We are grateful to Dr. A. Shafei-Deh-Abad from Department of Mathematics, Faculty of Science, and N. Rezaei-Ghaleh from Institute of Biochemistry and Biophysics, University of Tehran, for their useful comments regarding CA.

References

[1] T.T. Simopoulos, W.P. Jencks, Alkaline phosphatase is an almost perfect enzyme, Biochemistry 33 (1994) 10375–10380.
[2] C.J. Camacho, Z. Weng, S. Vajda, C. DeLisi, Free energy landscapes of encounter complexes in protein–protein association, Biophys. J. 76 (1999) 1166–1178.
[3] G.R. Welch, The enzymatic basis of information processing in the living cell, Biosystems 38 (1996) 147–153.
[4] L.B. Kier, C.-K. Cheng, B. Testa, A cellular automata model of ligand passage over a protein hydrodynamic landscape, J. Theor. Biol. 215 (2002) 415–426.
[5] L.B. Kier, C.K. Cheng, B. Testa, Studies of ligand diffusion pathways over a protein surface, J. Chem. Inf. Comput. Sci. 43 (2003) 255–258.
[6] M. Ghaemi, N. Rezaei-Ghaleh, M.N. Sarbolouki, Directed ligand passage over the surface of diffusion-controlled enzymes: a cellular automata model, Lect. Notes Comput. Sci. 3305 (2004) 719–724.
[7] P. Sarkar, A brief history of cellular automata, ACM Comput. Surv. 32 (2000) 80–107.
[8] D.A. Wolf-Gladrow, Lattice-gas cellular automata and lattice Boltzmann models: an introduction, Lect. Notes Math. 1725 (2000) 1–308.
[9] J. Kyte, R.F. Doolittle, A simple method for displaying the hydrophobic character of a protein, J. Mol. Biol. 157 (1982) 105–132.
[10] R.V. Hogg, E.A. Tanis, Probability and Statistical Inference, sixth ed., Prentice-Hall, New Jersey, 2001.