



## A Cellular Automata Model of Ligand Passage Over a Protein Hydrodynamic Landscape

LEMONT B. KIER\*<sup>†</sup>, CHAO-KUN CHENG<sup>‡</sup> AND BERNARD TESTA<sup>§</sup>

*\*Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, VA 23298, U.S.A. ‡Department of Mathematical Sciences, Virginia Commonwealth University, Richmond, VA 23298, U.S.A. and §Institute de Chemie Therapeutique, Universite de Lausanne, CH-1015 Lausanne, Switzerland*

*(Received on 24 January 2001, Accepted in revised form on 6 December 2001)*

The subject of ligand passage to an active site on a protein is addressed. Current views on the mechanism and the possible role of surface water are discussed. A theory is presented in which the pattern of hydrophobic states of protein surface amino acid side chains is invoked as the influence on the relative hydrophobic effects of nearby water. The theory describes a ligand passage through the hydrodynamic near-surface water which exhibits temporary organized cavities resembling the chreodes introduced by Waddington. The passage of the ligand to the active site is facilitated by this dynamic mechanism. Cellular automata models of preferential directional diffusion through these chreodes support the theory. The theory may be invoked to explain a number of ligand–active site observations and serves as an idea for further studies.

© 2002 Elsevier Science Ltd. All rights reserved

### The Path of a Ligand to its Target

A major focus of attention in ligand–target site studies is the structural influence on the encounter of these two molecules. A variation in the structure of the ligand is often related, through models, to the binding affinity with a receptor or the catalytic rate with an enzyme. This leads to progress in the design of new, active biomolecules and some understanding of the processes underway. Less attention has been paid to the journey of the ligand to the target site. How does a ligand get to a target site; what conditions facilitate this; what is the effect of interference with the mechanism? These questions are central to our recent studies and are the subject of this article.

<sup>†</sup>Author to whom correspondence should be addressed.

### The Role of Diffusion

An early view described the approach of a ligand through bulk water to a target site. This three-dimensional random walk is postulated to be too slow to permit coordination of the heterogeneous metabolic processes in living systems (Welch, 1996). The current view invokes some period of residence of the ligand in some form of a two-dimensional manifold, on the way to an encounter with a target site. The nature of this two-dimensional passage is not generally agreed upon and is the subject of several models. Indeed, there may well be more than one mechanism in play. The concept of diffusion-controlled reactions is built around the model of an approach of a ligand to a target site. The central idea is that the approach of a ligand to a target site is a process assumed to be diffusion,

acting as a limiting condition to the rate of the reaction. How this diffusion takes place has been the subject of a number of studies and models over the years. There is agreement, however, that a two-dimensional surface diffusion to the active site can enhance the rate of a diffusion-controlled reaction (Adam & Delbruk, 1968; Eigen, 1974; Welch, 1996).

Several models of a two-dimensional passage have been described. An early model (Berg & Purcell, 1977) describes the path of a ligand passing over the surface of a cell, alternately touching the surface and then passing briefly into the bulk water in its random movement. These excursions ultimately lead to an encounter with a protein molecule containing the target site. Their calculations show that there is a higher probability of a ligand–target site encounter occurring in this model than an encounter with a ligand directly from the bulk water. A major part of the two-dimensional diffusion in this model is across a landscape of membrane surface, a passage that must be a random walk since there are no distinguishing features on this surface that would produce any variation. A model by Rhodes *et al.* (1985) and colleagues invokes diffusion through or across a membrane surface and demonstrates the advantages in reaction rates via these models. Other models consider the protein surface as a field for diffusion of a ligand to a target site. A model by Hasinoff (1982) describes the non-specific binding of a ligand to the charged surface of an enzyme. The ligand then follows a two-dimensional walk to the active site. This effectively increases the size of the target on the enzyme.

A comparison of models led Chou & Zhou (1982) to prefer a model described for an enzyme molecule with multiple receptor sites. A ligand molecule encounters the enzyme surface and diffuses to a nearby receptor to initiate the reaction. The role of van der Waals interactions are prominent in this model. The expectation of a significant accumulation of ligands on the enzyme surface is an outcome of the calculations based on this model. The entire surface of the enzyme is regarded as a “sink”, where the ligand diffuses to an active site. The conclusion was drawn that amino acid side chains outside the

active site play a major role as a “promoter”, facilitating a rapid diffusion to the active site.

The participation of the water near the protein surface was discussed by Welch (1977) who theorized that the microviscosity of near water was influential on the rate of catalytic reactions. The possibility that the viscosity of water near the protein surface is significantly higher than bulk water creates an ordering in a layer that might facilitate faster diffusion of a solute near the protein surface. A test of the influence of a solute on the organization of nearby solvent molecules was shown by Fidler *et al.* (1994). Using circular dichroism studies and molecular dynamics simulations, they showed that a chiral solute could create asymmetric pockets of solvent in the immediate vicinity. Another study by Muegge & Knapp (1995) revealed a variable influence of protein side chain hydrophobicity on the diffusion of water near the surface.

A topological concept of ligand passage over a protein to the active site was proposed by Blum *et al.* (1988). In this model, a structural transformation may occur among  $\alpha$ -helix and  $\beta$ -sheet forms of a protein: the Bonnet transformation. The substrate is thought to pass through bulk water and then encounter the protein through van der Waals attraction. Because of the varying curvature of the protein surface presented during the Bonnet transformations, the substrate is “focused” on the active site in a concerted manner. Several studies have proposed electrostatic forces to be operating over the protein surface, guiding the ligand to the active site gorge (Radic *et al.*, 1997), producing hydrodynamic torque effects (Antosiewicz *et al.*, 1996), ionic tethering near the active site (Wade *et al.*, 1998) and temporary binding (Brunori *et al.*, 1999). Van Belle *et al.* (2000) modeled the ligand clearance from an enzyme gorge using molecular dynamics in a solvent-free system. This showed the importance of some non-polar side chains.

A functional model was proposed by Birch & Latymer (1980), who did not specify the nature of the ligand approach to the receptor but described a model of the residence of several ligands near the active site. An experimental observation was made that upon administration of sweet-tasting molecules to observers followed

by a washout of the compound, a persistence of the sweet taste ensued. Birch considered several possible models including a non-specific localized pooling of the compound in the vicinity of the receptor. Based on experimental work, he proposed a model in which the sweet molecules are located in a queue held in place on the surface of the protein. After washout, the molecules are retained in the queue and so they continue to move toward the receptor, activating it as long as the queue is occupied. When the queue is empty, the sweet taste vanishes. This model invokes a directional effect focused on the receptor as well as the possibility of sustained residence of several molecules in the queue.

In summary, experimental and modeling evidence may be interpreted as describing a facilitation of reaction rates and receptor activation due to the rapid diffusion of ligands to target sites in essentially a two-dimensional domain. This diffusion most likely occurs across the surface of a protein, involving structural features not part of the target site. The forces guiding the ligand to the target site would be expected to result in several effects. These include some period of retention on the protein surface, a minimum extent of binding delaying the passage, and some influence carrying the ligand to the target site. These considerations and the demonstrated role of water lead us to consider a model involving the immediate layers of water enshrouding the protein.

### The Role of Water

It is well known that water is an essential ingredient in the reactions of biological phenomena. Indeed, the complex system of ligand–target site–water is a triad that exists at the core of biological transformations (Kier & Testa, 1992). Any model of two-dimensional diffusion of a ligand across a membrane or protein must include the participation of water. Water is an evanescent substance, constantly changing its architecture by making and breaking hydrogen bonds. Diffusion through water is a passage through the voids between clusters of hydrogen-bonded water molecules. The presence of solute molecules has a varying influence on water molecules in their vicinity depending on the

interactions possible between the different species. Polar solutes form hydrates, collections of water molecules in intimate contact with the solute molecule. In contrast, non-polar solute molecules are not effectively bound to water molecules, allowing the local organization of water to occur, driven by the preference of water to bind to water rather than water binding to solute. This phenomenon is called the hydrophobic effect.

The influence of water on solute diffusion and that of solute structure on water organization become issues in the attempt to model ligand diffusion across a surface. In one study of these effects, cellular automata models of diffusion and the hydrophobic effect were created (Kier *et al.*, 1997). The model and evidence from comparisons with literature values, led to the conclusion that hydrophobic solutes diffuse faster in water than hydrophilic solutes of the same size. Another study compared the diffusion rate of a solute through water-containing randomly placed stationary molecules. It was found that the solutes diffused faster when the stationary molecules were hydrophobic. This differential influence of stationary ingredients on solute diffusion in a two-dimensional cellular automata model and supporting experimental data have led us to propose a general theory of the approach of a ligand to a target site across a protein surface.

### Ligand Passage and the Hydrophobic Effect

We propose that ligand molecules encounter the surface of a protein molecule and are captured within the first few layers of water on the surface. They are then guided to the active site over a series of cavity paths created by the relative hydrophobic effect responding to the hydrophobic state of each amino acid side chain. These paths are preferences reminiscent of the chreodes envisioned by Waddington (1957) in his description of an epigenetic landscape. Waddington coined the word from the Greek words for “necessary” and “route” or path. He defined it as a representation of a temporal succession of states of a system, which is characterized by a property that a dynamic system will tend to respond to perturbations

by returning to the chreode. There are two characteristics encoded in this concept. The first is the presence of a degree of progress as one proceeds from the initial to the final state of the system. In some periods of the traverse through the chreode, there is a great deal of progress, while in other periods, it is less. The second characteristic is the relative strength of the tendency of the system to return to the trajectory created as a chreode. Because this definition is close to the phenomenon we invoke here, we have adopted this term to characterize the system.

In this report, we evaluate the possibility of the existence of these chreodes and reflect on the consequences of their function. The modeling of this system is done with cellular automata, which we have used many times to study aqueous systems (Kier *et al.*, 1999, 2000, 2001).

### Cellular Automata

#### THE METHOD

Cellular automata are dynamical computational systems that are discrete in space, time and state and whose behavior is specified completely by rules governing local relationships. They are an attempt to simplify the often numerically intractable dynamic simulations into a set of simple rules that mirror intuition and that are easy to compute. As an approach to the modeling of emergent properties of complex systems, it has a great benefit in being visually informative of the progress of dynamic events. From the early development by von Neumann (1966), a variety of applications ranging from gas phenomena to biological applications have been reported (Ermentrout & Edelstein-Keshet, 1993; Choppard & Droz, 1998).

Our model is composed of a grid of square spaces called cells on the surface of a torus to remove boundary conditions. Each cell  $i$  has four tessellated neighbors,  $j$ , and four extended neighbors,  $k$ , in what is called an extended von Neumann neighborhood, Fig. 1. Each cell has a state governing whether it is empty or is occupied by a water or other molecule. The contents of a cell move, join with another occupied cell or break from a tessellated relationship according to probabilistic rules.

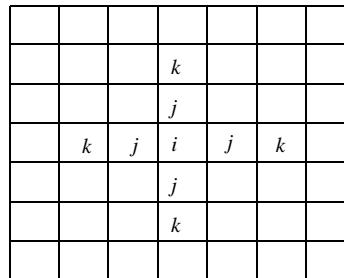


FIG. 1. The von Neumann neighborhood focused on cell  $i$ .

These rules are established at the beginning of each simulation. The rules are applied one after the other to each cell at random, the complete application of the rules to all cells constituting one iteration. The rules are applied uniformly to each cell type and are local; thus, there is no action at a distance. Our cellular automata model is kinematic, asynchronous and stochastic. The initial conditions are random, hence they do not determine the ultimate state of the cells called the configuration. The same initial conditions do not yield the same set of configurations after a certain number of iterations except in some average sense. The configurations achieved after many iterations reach a collective organization that possesses a relative constancy in appearance and in reportable counts of attributes. What we observe and record from the cellular automata simulations are emergent attributes of a complex system.

#### The Rules

The cellular automata model is made up of a square grid of cells  $(i, j)$  on the surface of a torus. A cell's four nearest neighbors  $(i - 1, j)$ ,  $(i + 1, j)$ ,  $(i, j - 1)$ ,  $(i, j + 1)$  constitute its von Neumann neighborhood ( $vNn$ ). Its extended von Neumann neighborhood ( $evNn$ ) are the four cells  $(i - 2, j)$ ,  $(i + 2, j)$ ,  $(i, j - 2)$ ,  $(i, j + 2)$ . Let  $(i', j')$  denote any cell in the  $vNn$ , and  $(i'', j'')$  denote any cell in the  $evNn$ .

On this grid of cells, each cell can be either unoccupied or occupied by a molecule. Each molecule has a fixed type. Let  $\{1, 2, 3, \dots, t\}$  be the set of types involved in the simulation, then the state of a cell  $(i, j)$  could be expressed by a single integer  $s(i, j) = k$ ,  $k = 0$  if unoccupied,  $k > 0$  if occupied by a molecule of type  $k$ . The

configuration of the entire system is defined by the state values of the cells in the grid. Molecules are allowed to move to a vacant  $vNn$  cell, and leave its original cell vacant, thus the configuration of the entire system changes accordingly. The movement of a molecule is determined by the states in its  $vNn$  and  $evNn$  following the general rules. The rules depend on two parameters, which allow the moving probabilities of a cell to be computed, and then a random choice according to those probabilities determines where a molecule will move if it would move at all.

Two parameters are selected for a simulation to control the probabilities for movement of the molecules in the grid. The breaking probability,  $P_B(X, Y)$ , is the probability for a molecule, of type  $X$ , in cell  $(i, j)$ , to break away from a molecule, of type  $Y$ , in its  $vNn$ , when there is exactly one occupied  $vNn$  cell. The value of  $P_B(X, Y)$  lies between 0 and 1 (inclusive). The second parameter  $J(X, Y)$  describes the movement of the molecule of type  $X$  in cell  $(i, j)$  toward or away from the molecule of type  $Y$ , in its  $evNn$  when the intermediate  $vNn$  cell is vacant.

Given that  $(i, j)$  cell is occupied by a molecule of type  $X$ , and all  $vNn$  cells of  $(i, j)$  cell are vacant,  $J(X, Y)$  is defined to be the ratio  $A/B$ : where  $A$  is the probability that the molecule in cell  $(i, j)$  will move toward an  $evNn$  cell occupied with a molecule of type  $Y$ ; and  $B$  is the probability that the molecule in cell  $(i, j)$  will move toward a vacant  $evNn$  cell. The value of  $J(X, Y)$  is a positive real number. When  $J(X, Y) = 1$ , the molecule in cell  $(i, j)$  has the same probability of movement toward an occupied  $evNn$  with molecule of type  $Y$ , as when that cell is empty. When  $J(X, Y) > 1$ , it indicates that the molecule in cell  $(i, j)$  has a greater probability of movement toward an occupied  $evNn$  cell with molecule of type  $Y$ , than when that cell is empty. When  $J(X, Y) < 1$ , it indicates that the molecule in cell  $(i, j)$  has a lower probability of such movement.

The simulation is controlled as follows: let the simulation starting time be 1, and let  $n$  be the number of molecules in the system. The  $i$ -th iteration of the simulation is the time  $i \times (n - 1) + 1$  till  $i \times n$ . At each time step in any iteration, a molecule is randomly selected to

act, once and only once, according to the state of the configuration in the previous time step. Its action could be “stay” or “move to one of its vacant  $vNn$  cell”. In one iteration, every molecule in the system is given the chance to act.

During a time step, if the molecule at cell  $(i, j)$  is selected to act, the following probabilities are calculated so that they can be used to decide where the molecule will move if it would move at all:

1. The probability that this molecule will move, denoted by  $P_m$  and
2. the conditional probability that this molecule moves into an empty cell  $(i', j')$  in its  $vNn$  given that it will move, denoted by  $P_m(i', j')$ .

We first decide randomly whether the molecule will move using  $P_m$ . If it would move, then we use  $P_m(i', j')$ s to decide randomly where it will move. This approach allows for a simple computation, and also limits the influence from molecules in  $evNn$  cells to a portion of the moving probability, thus the influence from one direction does not overshadow the influence from other directions. Furthermore, when  $J(X, Y)$  parameters are set to 1,  $P_m$  turns out to be the product of  $P_B$  values associated with molecules bound to it. This agrees with the intuitively reasonable assumption that if the movement of a molecule is not influenced by a molecule not bound to it, then its moving probability is the joint probability of probabilities for it to break away from molecules bound to it. The detail of the calculation can be found in the pseudo-code program in Appendix A.

### A Cellular Automata Model of the Hydrophobic Effect

The hydrophobic effect arises from the greater attraction and binding of water to itself, rather than water binding to another molecule (a solute or a stationary molecular fragment). The relative hydrophobic state of the other molecule influences the degree of aggregation of the water molecules in the vicinity. On the surface of a membrane or protein, this has been referred to as the microviscosity (Welch, 1977). To reveal this effect, we have previously calculated the relative hydrophobicity of a solute in water in a cellular

automata simulation (Kier *et al.*, 1999). The hydrophathic state is encoded in the rule,  $P_B(WS)$ , where a high probability reflects a hydrophobic solute and a low value reflects a hydrophilic solute. This earlier study illustrated the ability of the hydrophathic state of a solute to organize the water in its vicinity. From this, we infer that the amino acid side chains, with variable hydrophathic states, may organize the water and the cavities into some pattern which could function as our postulated chreode.

#### ENCODING THE HYDROPATHIC STATES OF THE AMINO ACID SIDE CHAIN

The hydrophathic state of the amino acid side chains is of importance in order to model their influence on ligand diffusion in our model. The  $P_B(WS)$  rules equivalent to the hydrophathic states of the amino acid side chains,  $S$ , are derived by considering the side chains in subsets. The hydrophathic states reported in the literature vary to some degree and so this approximation is intended to reflect trends. The  $P_B(WS)$  values adopted for five subsets of the amino acid side chains are shown in Table 1.

#### The Hydrophathic State and Ligand Diffusion

Our previous studies have shown that the diffusion of a solute through a solution is influenced by the hydrophathic states of other solutes (Kier *et al.*, 1997). The diffusion of a solute was demonstrated to be faster if another solute is hydrophobic. In our model of a chreode on a hydrodynamic landscape, it is necessary to consider the influence on diffusion of multiple stationary ingredients representing the side chains on a protein surface. This multiple influence can be simplified by considering the side chain surrogates in pairs.

In this study, we have created a series of square zones on the cellular automata grid surface, Fig 2. At the diagonal corners of the zones are cells with the same hydrophathic states. Vertical and horizontal cell pairs operate as gates. The hydrophathic state of each corner cell is expressed by the letter code used in defining the subsets of side chains as shown in Table 1. There are 15 possible pairings of the five subsets from Table 1. The remainder of the grid is filled

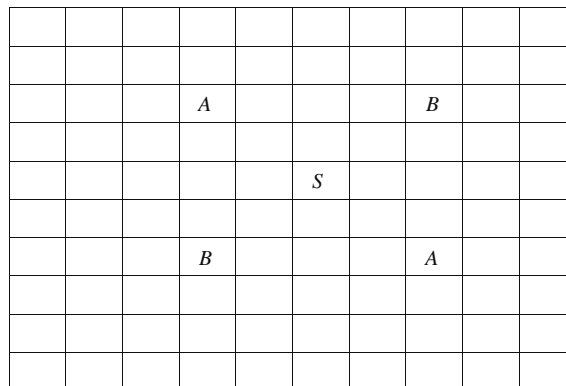


FIG. 2. Example of side chain “gates” with three-cell separation (here  $A$  and  $B$ , see Table 1) with diffusant  $X$ .

TABLE 1  
*Rules equivalent to the hydrophathic states of amino acid side chains*

Code	Side chain	$P_B(WS)/J$	Hydrophathic state
$A$	Arg, Asp, Lys	0.1/2.80	Very hydrophilic
$B$	Asn, Glu, Gln, Gly, Ser	0.3/1.40	Hydrophilic
$C$	Ala, His, Pro, Thr	0.5/0.71	Intermediate
$D$	Cys, Met, Tyr, Val	0.7/0.36	Hydrophobic
$E$	Ile, Leu, Phe, Trp	0.9/0.18	Very hydrophobic

with water cells that move freely and interact with the  $S$  cell and the gate cells. The solute,  $S$ , in the center of the zone is allowed to move randomly until it crosses the row or column of grid cells connecting any pair of gate cells. The number of iterations for this event to occur is recorded. The average number of iterations (“time”) over 200 runs is recorded. These values had a standard error of the mean of 2.5%. The shortest path to the exit position divided by the average “time” is the rate of diffusion for that set of conditions. The example shown in Fig. 2 has a minimum path of three cells to exit the zone. The observed diffusion rate is a consequence of the hydrophathic state of each pair of gate cells, acting in concert.

The consequences of the hydrophathic effects on the diffusion of molecule,  $S$ , are shown in a series of matrices, Tables 2–6. In Table 2, the side chain gates are represented by letters

TABLE 2  
Diffusion rates, cells per iteration ( $\times 10^{-2}$ ) of an intermediate hydrophobic state diffusant through gates with a three-cell separation\*

	Gate types				
	A	B	C	D	E
A	2.3	2.5	2.8	2.9	3.1
B		2.8	3.0	3.2	3.2
C			3.2	3.2	3.2
D				3.3	3.4
E					3.4

\*See Table 1 for hydrophobic states corresponding to the letter codes.

designating the five subsets of side chains classified according to their approximate hydrophobic states. This study reveals the relative diffusion rates of molecule *S* with a neutral hydrophobic state, passing through gates with a three-cell separation. The gate structures that are most hydrophobic permit a faster diffusion than those that are most hydrophilic, a rate increase of almost 50%. This model reveals a difference among the modeled amino acid side chains as regards their influence on diffusion of ligands near them.

Table 3 shows an extension of this study where the gate structures are placed farther apart. The same general result occurs as described above. The more hydrophobic pairs of structures produce a faster diffusion than the most hydrophilic pairs but in this case by an increase of less than 20%. Finally, Table 4 shows the effect on a neutral diffusant when it confronts the gate structures separated by five cell spaces. Here the increase in diffusion through hydrophobic gates relative to hydrophilic gates is virtually nil.

The influence on the diffusion rate due to the hydrophobic state of the diffusant, *S*, was modeled using the three-cell gate pattern. Using a hydrophilic, neutral and a hydrophobic diffusant, shown in Tables 5, 2 and 6, respectively, the diffusion rate was compared for the two *C*-type gates, separated by three cells. The results are shown in Table 7. The influence of the hydrophobic state of the diffusant on its rate of diffusion is significantly greater than the effect due to separation of the gate structures. The maximum diffusion portrayed in this model is

TABLE 3  
Diffusion rates, cells per iteration ( $\times 10^{-2}$ ) of an intermediate hydrophobic state diffusant through gates with a four-cell separation\*

	Gate types				
	A	B	C	D	E
A	2.4	2.5	2.6	2.7	2.7
B		2.7	2.7	2.8	2.9
C			2.7	2.9	2.8
D				2.7	2.8
E					2.8

\*See Table 1 for hydrophobic states corresponding to the letter codes.

TABLE 4  
Diffusion rates, cells per iteration ( $\times 10^{-2}$ ) of an intermediate hydrophobic state diffusant through gates with a five-cell separation\*

	Gate types				
	A	B	C	D	E
A	2.4	2.3	2.4	2.4	2.3
B		2.3	2.2	2.3	2.2
C			2.3	2.2	2.4
D				2.2	2.2
E					2.2

\*See Table 1 for hydrophobic states corresponding to the letter codes.

TABLE 5  
Diffusion rates, cells per iteration ( $\times 10^{-2}$ ) of a hydrophilic diffusant through gates with a three-cell separation\*

	Gate types				
	A	B	C	D	E
A	0.21	0.25	0.28	0.29	0.32
B		0.30	0.33	0.33	0.36
C			0.35	0.38	0.39
D				0.41	0.43
E					0.42

\*See Table 1 for hydrophobic states corresponding to the letter codes.

for the case of a hydrophobic ligand passing through a relatively close array of hydrophobic side chains.

### Creation of a Model Chreode

With the information gleaned from the gate studies and hydrophobic influences, we next

TABLE 6  
Diffusion rates, cells per iteration ( $\times 10^{-2}$ ) of a hydrophobic diffusant through gates of a three-cell Separation\*

	Gate types				
	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>
<i>A</i>	5.1	5.6	5.8	5.9	5.9
<i>B</i>		5.6	5.9	6.0	6.4
<i>C</i>			5.9	5.9	6.1
<i>D</i>				6.3	6.0
<i>E</i>					6.0

\*See Table 1 for hydrophobic states corresponding to the letter codes.

TABLE 7  
Comparison of diffusion rates, cells per iteration ( $\times 10^{-2}$ ) for different hydrophobic states and gate sizes

Gate size	Diffusant hydrophobic states		
	Hydrophilic	Intermediate	Hydrophobic
3	0.35	3.2	5.9
4		2.7	
5		2.3	

explored the possibility of a guided or directed trajectory for the diffusant; in essence, a model of a chreode on the hydrodynamic landscape. For this study, we used a cellular automata grid located on the surface of a torus with a central region representing a target for a ligand as shown in Fig. 3. A random distribution of the five cell types representing the five groups of amino acid side chains based on hydrophobic state (Table 1) were introduced onto a cellular automata grid, Fig. 3. All of these cells, (*A, B, C, D, E*) were separated from each other by three cell spaces, not explicitly shown in Fig. 3 but present in the calculation. The system contains water in the same proportion as used in the previous studies. The water is allowed to move freely and to interact with the ligand and stationary gate cells. The diffusant was started at a position 38 cells from the target and endowed with a neutral hydrophobic state in this study. The count of iterations necessary for the ligand, *S*, to traverse the grid and touch the central target was averaged over 200 runs. These values

<i>S</i>	<i>C</i>	<i>C</i>	<i>D</i>	<i>D</i>	<i>B</i>	<i>E</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>D</i>	<i>B</i>	<i>E</i>	<i>B</i>	<i>C</i>	<i>C</i>	<i>A</i>	<i>C</i>	<i>D</i>	<i>C</i>
<i>B</i>	<i>D</i>	<i>A</i>	<i>D</i>	<i>A</i>	<i>C</i>	<i>D</i>	<i>B</i>	<i>B</i>	<i>D</i>
<i>A</i>	<i>C</i>	<i>B</i>	<i>A</i>	<i>D</i>	<i>A</i>	<i>A</i>	<i>C</i>	<i>E</i>	<i>A</i>
<i>B</i>	<i>D</i>	<i>C</i>	<i>E</i>	<i>B</i>	<i>E</i>	<i>C</i>	<i>B</i>	<i>A</i>	<i>C</i>
<i>E</i>	<i>C</i>	<i>A</i>	<i>D</i>	♥	<i>A</i>	<i>B</i>	<i>E</i>	<i>B</i>	<i>B</i>
<i>C</i>	<i>A</i>	<i>C</i>	<i>A</i>	<i>C</i>	<i>E</i>	<i>A</i>	<i>C</i>	<i>D</i>	<i>C</i>
<i>C</i>	<i>A</i>	<i>B</i>	<i>D</i>	<i>A</i>	<i>E</i>	<i>C</i>	<i>B</i>	<i>A</i>	<i>D</i>
<i>A</i>	<i>D</i>	<i>A</i>	<i>B</i>	<i>B</i>	<i>A</i>	<i>B</i>	<i>A</i>	<i>C</i>	<i>B</i>
<i>C</i>	<i>A</i>	<i>C</i>	<i>A</i>	<i>D</i>	<i>B</i>	<i>A</i>	<i>E</i>	<i>A</i>	<i>C</i>

FIG. 3. A cellular automata model of a random surface: ♥ is the active site position; *S*, the diffusant with an intermediate hydrophobic state; *A, B, C, D*, and *E* are the different subsets of side chains shown in Table 1. All side chains are three spaces apart with water molecules moving freely among them.

and those in the following study had a standard error of the mean of 6%. The average number of iterations necessary in this study to traverse the distance to the center was calculated to be 12429.

A second model was created in which the same number of *A, B, C, D*, and *E* structures were randomly scattered as in Fig. 3. Each cell was separated from another by a three-cell space. Eighteen of these were organized to form a chreode pattern shown in underlined italics in Fig. 4. The dynamics were run as before and the time for the *S* cell to traverse the spaces to the central target was averaged over 200 runs. In this study, the average time to diffuse to the center was calculated to be 8609 iterations. This model simulates the possible diffusion of a ligand across a grid surface with specifically positioned stationary cells coordinating their hydrophobic states to facilitate diffusion toward the center of the grid. This simulates a generalized model of a chreode as we have defined it in this study. It was found that the rate of diffusion of a ligand is faster in the ordered stationary cell model as compared to a random distribution of these same cells on the grid. The diffusion observed in

S	C	C	D	<u>A</u>	E	B	C	D	B
D	B	E	B	<u>B</u>	C	A	C	D	C
B	D	C	D	<u>C</u>	C	D	B	B	D
A	C	B	A	<u>D</u>	A	A	C	E	A
B	A	C	E	<u>E</u>	A	C	B	A	C
<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	♥	<u>E</u>	<u>D</u>	<u>C</u>	<u>B</u>	<u>A</u>
C	B	A	A	<u>E</u>	A	D	A	D	C
C	A	B	D	<u>D</u>	E	C	B	A	D
A	D	A	B	<u>C</u>	A	B	A	C	B
C	A	C	A	<u>B</u>	B	A	E	A	C

FIG. 4. A cellular automata model of a generalized chreode (three-cell gates): ♥ is the active site position; S, the diffusant with an intermediate hydrophatic state; A, B, C, D, and E are the different subsets of side chains shown in Table 1. All side chains are three spaces apart with water molecules moving freely among them.

the second study is a consequence of the existence of a pattern of cells and their hydrophatic states, a possibility existing on the surface of a protein.

### Experimental Evidence of the Chreode Role

A number of recent reports describe experimental evidence supporting the role of amino acid side chains in the passage of ligands to active sites, the key ingredient in our chreode model. A study by Radic *et al.* (1997) explored the role that surface and active center amino acid side chains play in the attraction of ligands across the surface of the enzyme acetylcholinesterase (AChE) to the active center located in a gorge in the protein molecule. They studied the kinetics of association of cationic and neutral ligands with the active center and the peripheral region of AChE. By neutralizing the surface residues by chemical modification, they achieved a lower ionic set of residues, presented to the ligands. Their measured association rates and their contribution to the reaction rates were then compared to those of side chains on native AChE. The differences measured in these rates attest to the influences of the relative ionic

strengths of surface residues, and hence the role of their hydrophatic states on the processes involved. Brunori and colleagues (1999) reported studies on ligand diffusion on myoglobin using stopped-flow laser photolysis. Their measurements supported a conclusion that the ligand migrates on the protein surface to a secondary binding site under the influence of side chain dynamics. These side chains may or may not allow access to the secondary site based on their properties. From the secondary site, the ligands proceed to the primary binding site. The side chains are concluded to exercise a control over the regulation of the ligand presence and binding at the sites on myoglobin. Boyd *et al.* (2000) have studied the influence of individual amino acid side chains in governing rates of entry into the active center gorge of acetylcholinesterase. Using spectroscopic and kinetic measurements, they observed that variation in the polarity, hence the hydrophatic state achieved by selective mutagenesis, influenced the binding and catalysis by limiting access to the active site.

### Discussion

The diffusion of ligands across protein surfaces has been considered as a route for these molecules to reach the active sites. This model is offered to explain the rapid response of receptors and the fast rates of enzyme catalysis. Two general mechanisms have been proposed in the past to define the facilitation of the diffusion. In the first case, the forces of interaction between ligands and the side chains are suggested to be electrostatic. If this were true, the attraction between ligand and certain side chains might be strong enough to ensnare the molecule in regions of the protein surface, retarding diffusion. These side chains would, in essence, function like an active site or like a binding site rather than a diffusion-promoting feature as described by Chou and Zhou. In contrast, van der Waals forces were proposed between ligands and side chains, serving to facilitate the diffusion to the active site. These forces require a very close approach of the ligand and side chain, presenting a deterrent to rapid diffusion because of steric entanglement.

A possibility, proposed in this study, is to invoke the participation of the layers of water molecules immediately adjacent to the protein surface. Each amino acid side chain intruding into the bulk water exercises an influence on the water architecture. This takes the form of a hydrophobic effect from hydrophobic side chains and side chain hydration with hydrophilic side chains. The cavities between clusters of hydrogen-bonded water molecules are in a dynamic state, joining with other cavities, breaking away, reforming, all in a manner resembling a dynamic peristaltic pump. These cavities form the proposed chreodes facilitating the passage of diffusants across the protein surface.

To study the possibility of selective side chain diffusion influences, we have simulated, using cellular automata, pairs of amino acid side chains using stationary cells in a field of water. These side chains were endowed with rules giving them hydrophobic character, typical of five classes of amino acid side chains. A diffusant was allowed to escape past paired side chains to evaluate the effect on diffusion rate due to their hydrophobic state. The study showed that the more hydrophobic side chains permitted a faster diffusion than the hydrophilic side chains. Another observation concerned the distance between the stationary cells, surrogate for a pair of side chains. A separation distance of three cells between side chains produced the most selective effect on diffusion rate, while a separation distance of four cells was less discriminating and a distance of five cells did not produce any discrimination in the rate of diffusion. The implication is that the side chains forming a chreode must have some optimum distance between them in order to influence the water and the subsequent chreode formation. Another observation from these studies was that more hydrophobic molecules diffused faster than those that are hydrophilic. We have reported on this observation in an earlier study using cellular automata wherein we cited supporting experimental evidence.

To establish the possibility that side chains might form chreodes directing the diffusion of molecules to an active site, we have modeled a pattern of occupied cells simulating the side

chains in such a configuration. The pattern was constructed so that the more hydrophobic side chains were located near the active site and the more hydrophilic side chains further away. The model produced faster diffusion to the active site than that of a pattern of side chains that was randomly organized. The cellular automata models indicate that the hydrophobic state of the side chains can function to influence a guided and facilitated diffusion to a specific location. This model is supported by experimental evidence described above. This facilitation occurs because of a proposed dynamic process in water forming a chreode in which a molecule traverses the surface to an active site. An important aspect of the side chain role is the pattern that must exist on the protein surface to form a chreode. If this role for a side chain exists, then we can expect to find a pattern of side chains on a protein surface that constitutes a code, operating as we have proposed.

Several consequences of the theory may be derived from a consideration of known phenomena. The measurement of the affinity of a molecule may reflect a more complex system than just a ligand-active site encounter. It may be that the measurement is including some residence of the ligands in chreodes. Another observation that may have a connection with the chreodes is the phenomenon of lag, where some time must pass before an effect is observed in a pharmacological test system. The lag may reflect the time needed for the drug molecule to displace the transmitter from the chreodes leading to the active site. The sequel to that observation is persistence, where the measured effect continues for some time after the washout of the ligand from the test system. The velocity of enzyme reactions may, in part, be explained by the facilitated trajectory of the ligand through chreodes to the active site and the velocity of departure of the product through chreodes. The chreodes may exhibit some selectivity of their occupants and may possibly reject some molecules that are detrimental to fitness of the system. Finally, the mechanism of non-specific anesthetic agents may depend on their ability to interfere with the chreodes carrying the normal transmitter to the receptor.

## REFERENCES

- ADAM, G. & DELBRUK, M. (1968). Reduction of dimensionality in biological diffusion processes. In: *Structural Chemistry and Molecular Biology* (Rich, A., Davidson, N., eds), pp. 198–215. San Francisco, Freeman.
- ANTOSIEWICZ, J., BRIGGS, J. M. & McCAMMON, J. A. (1996). Orientational steering in enzyme–substrate association: ionic strength dependence of hydrodynamic torque effects. *Eur. Biophys. J.* **24**, 137–141.
- BERG, H. C. & PURCELL, E. M. (1977). Physics of chemoreception. *Biophys. J.* **20**, 193–215.
- BIRCH, G. G. & LATYMER, Z. (1980). Intensity/time relationships in sweetness: evidence for a queue hypothesis in taste chemoreception. *Chem. Senses* **5**, 63–78.
- BLUM, Z., LIDIN, S. & ANDERSSON, S. (1988). Enzymes and curvature. *Angew. Chem. Int. Ed. Engl.* **27**, 953–956.
- BOYD, A., MARNETT, A. B., WONG, L. & TAYLOR, P. (2000). Probing the active center gorge of acetylcholinesterase by fluorophores linked to substituted cysteines. *J. Biol. Chem.* **275**, 22401–22408.
- BRUNORI, M., CUTRUZZOLA, F., SAVINO, C., TRAVAGLINI-ALLOCAPELLI, C., VALLONE, B. & GIBSON, Q. H. (1999). Structural dynamics of ligand diffusion in the protein matrix: a study on a new myoglobin mutant Y(B10)Q(E7)R(E10). *Biophys. J.* **76**, 1259–1269.
- CHOPPARD, B. & DROZ, M. (1998). *Cellular Automata Modeling of Physical Systems*. Cambridge: Cambridge University Press.
- CHOU, K.-C. & ZHOU, G.-P. (1982). Role of the protein outside active site on the diffusion-controlled reaction of enzyme. *J. Am. Chem. Soc.* **104**, 1409–1413.
- EIGEN, M. (1974). Diffusion control in biochemical reactions. In: *Quantum Statistical Mechanics in the Natural Sciences* (Mintz, S. L., Widemeyer, S. M., eds), pp. 37–61. New York: Plenum Press.
- ERMENTROUT, G. B. & EDELSTEIN-KESHET, L. (1993). Cellular automata approaches to Biological modeling. *J. theor. Biol.* **160**, 97–113.
- FIDLER, J., RODGER, P. M. & RODGER, A. (1994). Chiral solvent structure around chiral molecules: experimental and theoretical study. *J. Am. Chem. Soc.* **116**, 7266–7273.
- HASINOFF, B. B. (1982). Kinetics of acetylthiocholine binding to electric eel acetylcholin esterase in glycerol/water solvents of increased viscosity. *Biochim. Biophys. Acta* **704**, 52–58.
- KIER, L. B. & TESTA, B. (1992). Complex systems in drug research II. The ligand–active site–water confluence as a complex system. *Complexity*, **1**, 37–42.
- KIER, L. B., CHENG, C.-K., TESTA, B. & CARRUPT, P.-A. (1997). A cellular automata model of diffusion in aqueous systems. *J. Pharm. Sci.* **86**, 774–778.
- KIER, L. B., CHENG, C.-K. & TESTA, B. (1999). Cellular automata models of biochemical phenomena. *Fut. Gen. Comput. Sys.* **16**, 273–289.
- KIER, L. B., CHENG, C.-K. & SEYBOLD, P. G. (2000). Cellular automata models of chemical systems. *SAR and QSAR* **11**, 79–102.
- KIER, L. B., CHENG, C.-K. & SEYBOLD, P. G. (2001). Cellular automata models of aqueous solution phenomena. *Rev. Comput. Chem.* **17**, 205–254.
- MUEGGE, I. & KNAPP, E. W. (1995). Residence times and lateral diffusion of water at protein surfaces. *J. Phys. Chem.* **99**, 1371–1374.
- RADIC, Z., KIRCHHOFF, P. D., QUINN, D. M., McCAMMON, J. A. & TAYLOR, P. (1997). Electrostatic influence on kinetics of ligand binding to acetylcholinesterase. *J. Biol. Chem.* **272**, 23265–23277.
- RHODES, D. G., SARIMENTO, J. G. & HERBETTE, L. G. (1985). Kinetics of binding of membrane-active drugs to receptor sites. *Molec. Pharmacol.* **27**, 612–623.
- VAN BELLE, D., DE MARIA, L. IURCU, G., & WODAK, J. (2000). Pathways of legand clearance in acetylcholinesterase by multiple copy sampling. *J. Mol. Biol.*, **298**, 705–726.
- VON NEUMANN, J. (1966). *Theory of Self-reproducing Automata* (Burks, A., ed). Champaign, IL: University of Illinois Press.
- WADE, R. C., GABDOULLINE, R. R., LUDEMAN, S. K. & LOUNNAS, V. (1998). Electrostatic steering and ionic tethering in enzyme-ligand binding: insights from simulations. *Proc. Natl Acad. Sci. USA* **95**, 5942–5949.
- WADDINGTON, C. H. (1957). *The Strategy of the Genes*, pp. 29–38. London: George Allen & Unwin Ltd.
- WELCH, G. R. (1977). On the free energy “cost of transition” in intermediary metabolic processes and the evolution of cellular infrastructures. *J. theor. Biol.* **68**, 267–279.
- WELCH, G. R. (1996). The enzymatic basis of information processing in the living cell. *BioSystems* **38**, 147–153.

## Appendix A

## The Pseudo-code for the Implementation of our System

DoSimulation( $n$ )

The argument  $n$  is the number of iterations to be performed in the simulation.

```
{
    set initial state of the configuration; //usually done randomly
    for ( $j = 1; j \leq n; j^{++}$ ) DoAnIteration();
}
```

DoAnIteration()

In the procedure,  $BB(i',j')$  is used to keep  $P_B$  value between molecules in  $(i,j)$  and  $(i',j')$ , and  $JJ(i',j')$  is used to keep  $J$  value between molecules in  $(i,j)$  and  $(i'',j'')$ , where  $(i',j')$  and  $(i'',j'')$  lie in the same side of  $(i,j)$ .

```

{   Open = all molecules in the grid;
    while Open is not empty
    {    $m$  = randomly selected molecule from Open;
        remove  $m$  from Open;
         $(i,j)$  = the cell in which  $m$  resides;
         $X$  =  $m$ 's type;
        for each  $(i',j')$  in the  $vNn$  of  $(i,j)$  do
        {    $JJ(i',j') = 1$ ;
            if  $(i',j')$  is not occupied then
            {    $BB(i',j') = 1$ ;
                if the  $evNn$  cell along the same direction of  $(i',j')$  cell is occupied with
                molecule of type  $Z$  then  $J(i',j') = J(X,Z)$ ;
            }
            else  $BB(i',j') = P_B(X,Y)$ ;
        }
        }
         $PBB$  = the product of  $PB$  of all  $BB(i',j')$ ;
         $n$  = number of vacant  $vNn$  cells;
        for each  $(i',j')$  in the  $vNn$  of  $(i,j)$  do
        {   if  $n > 0$  and  $PBB > 0$  then
            {    $temp = (n/Q-1)/JJ(i',j')$ ;
                 $ApprP\_m(i',j') = 1/(1 + temp)$ ;
            }
            else  $ApprP\_m(i',j') = 0$ ;
        }
        }
         $temp1$  = sum of all  $ApprP\_m(i',j')$  over all  $vNn$ ;
        //Now the  $P\_m$  is defined
        if  $temp1 > 1$  then  $P\_m = 1$ ;
        else  $P\_m = temp1$ ;
        //Now the  $P\_m(i',j')$ s are defined
        for each  $(i',j')$  in the  $vNn$  of  $(i,j)$  do
        {   if  $temp1 > 0$  then  $P\_m(i',j') = ApprP\_m(i',j')/temp1$ ;
            else  $P\_m(i',j') = 0$ ;
        }
        }
        use  $P\_m$  to decide randomly whether  $m$  is to move;
        if  $m$  is to move
        {   use  $P\_m(i',j')$ s to decide randomly where  $m$  would move;
            move  $m$  accordingly, set the configuration to reflect this action;
        }
    }
}

```