# A Cellular Automata Model of an Aqueous Solution

Lemont B. Kier[*,†] and Chao-Kun Cheng[‡]

Department of Medicinal Chemistry and Department of Mathematical Sciences,
Virginia Commonwealth University, Richmond, Virginia 23298-0540

A cellular automata model of a solute in a solvent (water) has been created based on our previous model of water. By changing the rules for the ingredients of the solution, we observe emergent properties of (1) an increase in "solubility" with increased "temperature", (2) the disruption of water structure with increasing solute concentration, and (3) the "structure-making" effect in water due to solutes which have a decreased affinity for water. This latter configuration is proposed as a model of a nonpolar solute producing a hydrophobic effect.

Our knowledge of solutions has been primarily derived from a reductionist treatment or analysis to give us disconnected parts for study. Considerable information has been produced, but we must recognize that the structure of water and its properties are emergent in a complex system. Of necessity, any model must be a dynamic one since the molecular-level structure of water is believed to be a rapidly changing system. Some insight has been gained recently by dynamically simulating water from individual molecules as with the technique of molecular dynamics. Although not without limitations, this category of approach seems destined for further exploitation. It is in this realm of study that we have focused attention in our research on water and solution phenomena.

In an earlier study we modeled liquid water using the dynamic simulation method known as cellular automata (CA).[1] The initial success of that model prompted us to introduce the capability to employ rules governing the dynamics of substances in addition to water. Our objective is to develop the capability to model a system which would to some extent simulate an aqueous solution. The success of our CA model of water can be demonstrated by its ability to simulate some of the attributes of water such as the average cluster size relating to the viscosity, the percent of unbound water molecules relating to the vapor pressure, and the similarity of the graphics with other simulation models such as molecular dynamics. These achievements prompt us to continue this research to see how extensive the ability of a CA model is to simulate a system such as an aqueous solution. Accordingly, we have pursued this goal and report here our progress.

## CELLULAR AUTOMATA

Cellular automata is a dynamic method for the modeling of the emergent behavior of a complex system.[2-5] It is an infinitely extended lattice or grid of sites called cells in one, two, or three dimensions. The intrinsic quality or status of any cell, i, is the state of that cell. Each cell, i, is in a local situation or map called a neighborhood. This neighborhood

of tesselated cells influences the changes that cell i may experience. These changes or transition functions are encoded into a set of rules. With the successive response of each cell i to the transition functions, an iteration of the entire ensemble of cells occurs producing a configuration of the system and a unit of "time". The rules of a cellular automata are local, hence, there is no action at a distance and the laws are uniform, that is the rules are the same for every cell. The discreteness in time, space, and rules results in an error-free specification of the dynamics. The functioning of a computer is by the same process.

## THE RULES

The model is made up of a domain composed of tesselated cells on the surface of a torus. Each cell, i, has four tesselated neighbors, j, and the cells two places beyond i, labeled k in Figure 1. The cells j constitute the von Neumann neighborhood for cell i, and the cells j and k constitute the extended von Neumann neighborhood for cell i. The state of each cell reflects whether or not it is occupied with a water molecule. The configuration of the entire system at any time is defined by the set of state values. At discrete intervals of time, discrete changes in each cell occur according to rules that constitute the state and transition functions. These rules are based upon probabilities, hence the initial state does not uniquely specify the dynamics or any configuration. Only the parameters, not the initial data, need be varied.

Two parameters are adopted for our model to control the probabilities for moving molecules in the grid. The breaking probability, $P_B$, used in our previous study (1) is the probability for a molecule to break away from the molecule in j cells when there is exactly one occupied j cell. The value for $P_B$ lies in the closed unit interval. The second parameter, $J$, describes the movement of the molecule at cell i toward or away from the molecule at a k cell in the extended von Neumann neighborhood when the intermediate j cell is vacant. It represents the ratio of the probability that a molecule at i cell will move toward an occupied k cell while the intermediate j cell is vacant, and the probability that a molecule at i cell will move toward a vacant k cell while the intermediate j cell is vacant. $J$ is a positive real number. When $J = 1$, it indicates that the molecule i has the same

A Cellular Automata Model of an Aqueous Solution

*J. Chem. Inf. Comput. Sci., Vol. 34, No. 6, 1994* **1335**
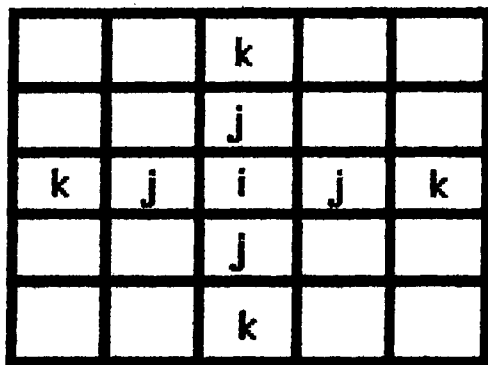


**Figure 1.** The extended von Neumann neighborhood.

**Table 1.** Effect of $P_B(W)$ Change on Solute Attributes

| $P_B(W)$ | $f_0(L)^a$ | av$^a$ cluster size | av no. of$^a$ L bonds |
|---|---|---|---|
| 0.3 | 0.11 | 4.5 | 2.00 |
| 0.6 | 0.33 | 2.0 | 1.25 |
| 0.9 | 0.44 | 1.5 | 1.00 |

$^a$ 10 000 iteration result.

probability of movement toward k as for the case when k is empty. When $J > 1$, it indicates that i has a greater probability of movement toward k than when k is empty. When $J < 1$, it indicates that i has a lesser probability of movement toward k relative to when k is empty.

In our model, the moving probability, $p_m$, for a molecule in a cell is defined as follows:

$$p_m = \text{sum of } p_m(d) \text{ over all possible d if the sum} < 1,$$
$$\text{otherwise 1} \quad (1)$$

where d is a direction in which i can move. The formula for computing the $p_m(d)$ in a certain direction d is as follows:

Let $n$ be the number of occupied j cells for a given i cell. There are $4 - n$ many directions that i can move. We define

$$p_m(d) = 1/(1 + ((4 - n)/(P_B)^n - 1)/J(d))$$
$$\text{if } n < 4 \text{ and } P_B > 0$$

$$= 0 \text{ otherwise} \quad (2)$$

where $J(d) = 1$ if the k cell in the direction d is vacant, otherwise $J(d) = J$.

The choice of $p_m(d)$ and $p_m$ allows for a simple computation of moving probabilities and also limits the influence from molecules in k cells to a portion of the moving probability, thus the influence from one direction does not overshadow the influence from other directions. Furthermore, when $J = 1$, $p_m = (P_B)^n$; this agrees with the intuitively reasonable assumption that if the movement of a molecule is not influenced by molecules not bound with it, then its moving probability is the joint probability of probabilities for it to break away from molecules bound to it.

## A CA MODEL OF AN AQUEOUS SOLUTION

It is our purpose in this study to continue our dynamic syntheses begun earlier with water.[1] In this study we introduce an ingredient into water and vary the joining and breaking parameters of each and the corresponding cross-term parameters. The attributes calculated for each parameter set study are compared with solution properties to assess relationships. This comparative study determines the quality of the solution model.

## INITIAL CONDITIONS

The size of the grid used is $55 \times 55$ cells as the surface of a torus, thus there are no boundary conditions. As in our earlier study, we have designated 69% of the cells as water molecules. For the simulation of a solution, we replace some

number of water cells (usually 100) with an equal number of designated solute cells. The expectation of some contraction in the volume is not simulated, thus the total volume percent remains the same. At the beginning of each dynamic simulation, the water and the solute cells are randomly distributed over the grid.

The transition functions or moving rules for water (W) and the solute (L) are designated thus

$P_B(W) = \text{water}-\text{water breaking probability}$
$P_B(L) = \text{solute}-\text{solute breaking probability}$
$P_B(WL) = \text{water}-\text{solute breaking probability}$
$J(W) = \text{water}-\text{water joining parameter}$
$J(L) = \text{solute}-\text{solute joining parameter}$
$J(WL) = \text{water}-\text{solute joining parameter}$

## CALCULATED ATTRIBUTES

From each dynamic simulation we can capture a set of attributes with each iteration. After allowing for a stabilization in these attribute values to occur we may describe a configuration with the following:

(1) Fraction of unbound water or solute molecules: $f_0(W)$ or $f_0(L)$. In the earlier study, this was reported as the percent unbound water molecules $\%\langle N_u \rangle$.

(2) Fraction of water or solute molecules bound to one up to four other molecules: $f_1(W)$ or $f_1(L)$, $f_2(W)$ or $f_2(L)$, $f_3(W)$ or $f_3(L)$, and $f_4(W)$ or $f_4(L)$. The values range from zero to one.

(3) Average bonded (tesselated) state of water or solute: $n_{HB}$ (for water only). The values range from zero to four.

## STUDIES AND RESULTS

**1. Influence of $P_B(W)$ "Water Temperature" on $f_0(L)$ "Solubility".** In this synthesis, we draw two correlations between CA attributes and solution phenomena. First, we adopt the relationship discovered in our first paper,[1] that the $P_B(W)$ relates to the temperature of water when the water/grid ratio is about 69%. Specifically the relationship is

$$P_B(W) = T(°C)/100$$

The second relationship involves the modeling of the "solubility" of a solute (L) in water by calculating the fraction of free solute molecules, $f_0(L)$. This model is based on the assumption that isolated solute molecules are in solution, while clusters of some size (here taken to be dimers or higher) are not in solution.

The parameters $P_B$ and $J$ are held constant in this study except for the $P_B(W)$ value. This varies between 0 and 1.0. In table 1 are shown three $P_B(W)$ values ("temperatures") and the corresponding values of $f_0(L)$, (fraction of "soluble"

**Table 2.** Effect of Concentration of Solute on Water Structure

| % solute | $f_o(W)$ | $n_{HB}$ |
|---|---|---|
| 0 | 0.0041 | 3.013 |
| 1.8 | 0.0054 | 3.008 |
| 3.5 | 0.0063 | 2.993 |
| 5.2 | 0.0069 | 2.928 |
| 6.9 | 0.0081 | 2.906 |

**Table 3.** Influence of $P_B(WL)$ on Water Structure

| $P_B(WL)$ | $f_o(W)$ | $n_{HB}(W)$ | $f_4(W)$ |
|---|---|---|---|
| 0.1 | 0.0074 | 2.85 | 0.268 |
| 0.2 | 0.0062 | 2.89 | 0.269 |
| 0.3 | 0.0065 | 2.93 | 0.301 |
| 0.4 | 0.0065 | 2.93 | 0.298 |
| 0.5 | 0.0056 | 2.93 | 0.300 |
| 0.6 | 0.0058 | 2.97 | 0.313 |
| 0.7 | 0.0060 | 2.95 | 0.308 |
| 0.8 | 0.0058 | 2.97 | 0.320 |
| 0.9 | 0.0050 | 2.97 | 0.320 |

solute molecules). These results are based on 100 solute molecules and 2000 water molecules in a $55^2$ cell grid. The results show that by increasing the $P_B(W)$ parameter while keeping the other parameters constant, there is a change in the configuration of the solute molecules. This change is an increase in the isolated solute molecules, a decrease in the solute cluster size, and a decrease in the number of tessellated solute faces. These changes occur as emergent properties of the system since there is no change in any parameter directly governing solute molecules. There is a parallel between this CA model and the effect of increasing the solubility by increasing the temperature of a solvent.

**2. Influence of Solute on Water Structure.** In this study we evaluated the influence of the increase of the concentration of a solute on the attributes of water. The ratio of the total water + solute cells to the grid size is kept constant at 69% by increasing both concentrations and the grid size. All $P_B$ and $J$ parameters are held constant. At several concentrations of solute in this study, the $f_o(L)$ values are about 1.0, thus this models complete solubility in this range. The attributes are recorded after achieving stable values after 5000 iterations.

The results are recorded in Table 2. When the concentration of the solute increases, there is an emergent effect on the water structure since no $P_B$ or $J$ parameters involving water or solute have changed. The increased concentration of solute produces an increase in the number of single water molecules and a modest decrease in the average tessellation of water molecules (interpreted as hydrogen bonding[1]). This model is of a solute disrupting water structure when the concentration is increased.

**3. Influence of $P_B(WL)$ Water–Solute Breaking Parameter and Water Structure.** This study was designed to reveal the influence of the water–solute breaking parameter $P_B(WL)$ on properties of water. The parameters for this synthesis were $P_B(W) = 0.25$, $P_B(L) = 0.10$, $J(W) = 1$, $J(L) = 1$, $J(L) = 1$, $J(WL) = 1$, and $P_B(WL)$ varied between 0.1 and 0.9. Stability in the calculated attributes was achieved after 5000 iterations. The results shown in Table 3 reveal that the $P_B(WL)$ parameter, encoding the affinity between solvent and solute, has a significant influence on the structure-making ability of the solute on water. Higher values of this parameter correspond to a decreased affinity between solute and water, a model of a nonpolar solute. Values of $P_B(WL)$ at the higher end of the scale produce the emergent properties of reduced single water molecules, greater tesselation (interpreted as greater hydrogen bonding), and a higher frequency of four-bonded water molecules. In essence, higher $P_B(WL)$ values lead to the synthesis of a more structured water. This model coincides with current ideas about the hydrophobic effect.[6-11]

## GRAPHIC SIMULATION OF THE HYDROPHOBIC EFFECT

In the third study described above, we have recorded in Table 3 the water attributes obtained by varying $P_B(WL)$.

The variation of this parameter through the closed unit interval produce an increased structuration of the water at higher $P_B(WL)$ values. While we gain some insight into this effect by recording the calculated attributes, we must actually view the dynamic simulation to effectively understand the changing architecture of the water continuum. We describe here the changing landscape of the solution with changing $P_B(WL)$ values.

**1. $P_B(WL) = 0.10$.** Most of the solute molecules are dispersed as single molecules within the patches of water-occupied cells. The cavities among the water molecules are mostly free of any solute molecules. Figure 2a shows this configuration.

**2. $P_B(WL) = 0.25$.** A few of the solute molecules are within the water patches and a few are within the cavities. About half of the solute molecules are positioned at the boundaries between these two features. Figure 2b reveals this configuration.

**3. $P_B(WL) = 0.40$.** No solute molecules are within the water patches. Most are within the cavities with a few forming a boundary between the patches and the cavities. Figure 2c shows this configuration.
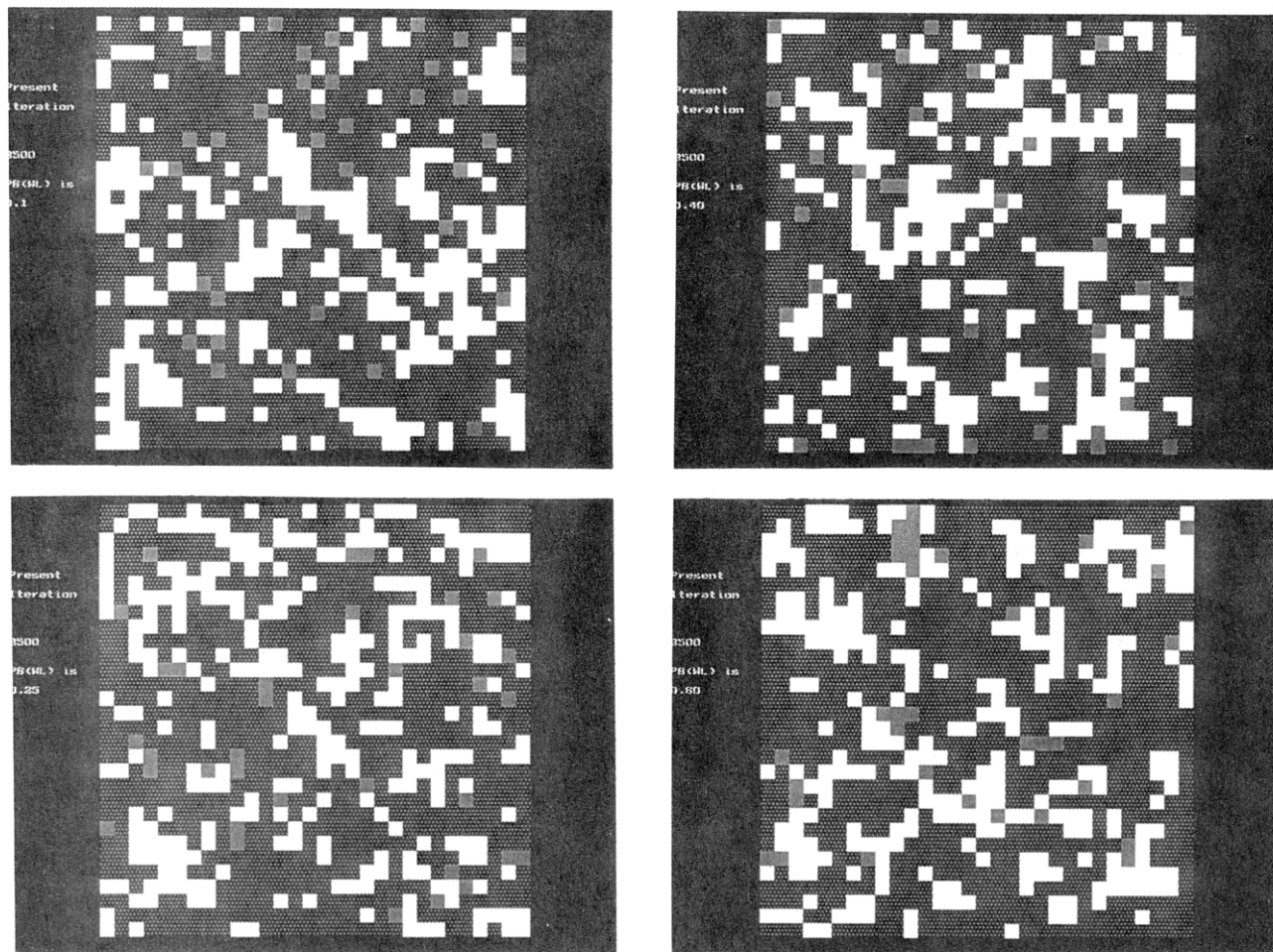
**4. $P_B(WL) = 0.80$.** All of the solute molecules are within the cavity areas. See Figure 2d for this configuration.

## DISCUSSION

The behavior of the dynamics shown here may be evaluated as a reasonable model of solution phenomena by comparing the stable configurations with experimental observations. In the first study, the simulations, based upon these initial conditions, produced a set of configurations varying with the water breaking probability, $P_B(W)$. This parameter, related to the temperature of the water, influences the disintegration of the solute molecules into nonbonded species. If we equate that species with solutes in solution, then we can regard this model as simulating the increased solubility of a solute as the solvent temperature increases. This is an emergent property not inherent in the initial conditions of our CA.

The second study produced a series of configurations in which the structure of the water changes when the concentration of the solute changes. Increasing the concentration of solute leads to a modest increase in the number of free water molecules and a modest decrease in the average number of tesselations, interpreted as the number of hydrogen bonds. This result is a realistic simulation of the disruption of water structure when the concentration of a solute increases.

The third study is an interesting simulation that reveals the important role of the $P_B(WL)$ parameter. With an increasing value of this parameter, corresponding to a decreased affinity between water and solute thus simulating

A CELLULAR AUTOMATA MODEL OF AN AQUEOUS SOLUTION

*J. Chem. Inf. Comput. Sci., Vol. 34, No. 6, 1994* **1337**



**Figure 2.** Dynamic simulations of a solute in water setting $P_B(W)$ = (a, top left) 0.10, (b, bottom left) 0.25, (c, top right) 0.40, and (d, bottom right) 0.80. The water cells are black, the solutes are gray, and the cavity areas are white.

a nonpolar substance, the solute moves from an interpatch position in the water to an intercavity position. As a consequence, the departure of the solute from a position among the water molecules leaves the water molecules more intact. The integrity of the water patches is more complete. We can say that the water is more structured. This phenomena has been postulated to explain the negative entropy effect produced by dissolving nonpolar solutes in water.[6-11] Referred to as the hydrophobic effect it is considered to be an important effect influencing biological phenomena. These CA simulations appear to produce a reasonable model of the hydrophobic effect.

## ACKNOWLEDGMENT

## REFERENCES AND NOTES

(1) Kier, L. B.; Cheng, C. K. A Cellular Automata Model of Water. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 647−652.

(2) von Neumann, J. In Theory of Self-Reproducing Automata; Burks, A. W., Ed.; University of Illinois Press: Urbana, IL, 1966.

(3) Wolfram, S. Statistical Mechanics of Cellular Automata. *Rev. Mod. Phys.* **1983**, *55*, 601−644.

(4) Jackson, E. A. *Perspectives of Non-Linear Dynamics*; Cambridge University Press: Cambridge, England, 1991; Chapter 10.

(5) Ermentrout, G. B.; Edelstein-Keshet, L. Cellular Automata Approaches to Biological Modeling. *J. Theor. Biol.* **1993**, *160*, 97−113.

(6) Franks, F. *Water—A Comprehensive Treatise*; Plenum: New york, 1973; Vol. 2, Chapter 5.

(7) Ben-Naim, A. *Water Aqueous Solutions*; Plenum: New York, 1974.

(8) Pratt, L. R. Theory of Hydrophobic Effects. *Ann. Rev. Phys. Chem.* **1985**, *36*, 433−449.

(9) Ohime, I.; Tanaka, H. Fluctuation, Relaxations, and Hydration in Liquid Water. Hydrogen Bond Rearrangement Dynamics. *Chem. Rev.* **1993**, *93*, 2545−2566.

(10) Kauzmann, W. Some Factors in the Interpretation of Protein Denaturation. *Adv. Protein Chem.* **1959**, *14*, 1−63.

(11) Blokzijl, W.; Engberts, J. B. F. N. Hydrophobic Effects. Opinions and facts. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1545−1579.