

# Simulation and Dynamics of Entropy Driven, Molecular Self-Assembly Processes

Bernd Mayer  
Gottfried Köhler  
Steen Rasmussen

SFI WORKING PAPER: 1996-08-058

SFI Working Papers contain accounts of scientific work of the author(s) and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may be reposted only with the explicit permission of the copyright holder.

[www.santafe.edu](http://www.santafe.edu)



SANTA FE INSTITUTE

# Simulation and dynamics of entropy driven, molecular self-assembly processes

Bernd MAYER<sup>a,1</sup>, Gottfried KÖHLER,<sup>a</sup> and Steen RASMUSSEN<sup>b,c,1</sup>

<sup>a</sup>Institut für Theoretische Chemie und Strahlenchemie  
Universität Wien, UZAII  
Althanstraße 14, A-1090 Wien  
Austria  
e-mail: bernd@asterix.msp.univie.ac.at

<sup>b</sup>Santa Fe Institute  
1399 Hyde Park Road  
Santa Fe, New Mexico, 87501  
USA

<sup>c</sup>TSA-DO/SA MS-M997 and CNLS MS-B258  
Los Alamos National Laboratory  
Los Alamos, New Mexico, 87545  
USA  
e-mail: steen@lanl.gov

---

<sup>1</sup>Authors to whom correspondence should be addressed.

## ABSTRACT

Molecular self-assembly is frequently found to generate higher order, functional structures in biochemical systems. One such example is the self-assembly of lipids in aqueous solution forming membranes, micelles and vesicles, another is the dynamic formation and rearrangement of the cytoskeleton. These processes are often driven by local, short range forces and, therefore, the dynamics is solely based on local interactions.

In this paper, we introduce a cellular automata based simulation, the Lattice Molecular Automaton, in which data structures, representing different molecular entities like water, hydrophilic and hydrophobic monomers, share locally propagated force information on a hexagonal, 2D lattice.

The purpose of this level of description is the simulation of entropic and enthalpic flows in a microcanonical, molecular ensemble to gain insight about entropy-driven processes in molecular many-particle systems. Three applications are shown, i.e. modeling structural features of a polar solvent, cluster-formation of hydrophobic monomers in a polar environment, and the self-assembly of polymers. Processes leading to phase-separation on a molecular level are discussed.

A thorough discussion of the computational details, advantages, and limitations of the Lattice Molecular Automaton approach can be found in reference [1].

## Keywords:

*Molecular Self-Assembly, Water Structure, Molecular Dynamics, Cellular Automata, Lattice Gas, Thermodynamics, Constructive Dynamics, Higher Order Emergence*

# 1 Introduction

## 1.1 Biological motivation

Many processes in biomolecular systems lack global, interfering control. The system dynamics is solely based on local interactions, providing, based on immediate reactions on environmental changes, the necessary flexibility and mean stability of the whole system. Considering the prokaryotic cell as a hierarchically structured, dynamical system, it is possible to characterize specific, functionally linked, mesoscopic complexes. One of these compounds is the semipermeable membrane separating space into an inside and an outside [2]. These membranes consist of a certain type of amphiphilic polymers (hydrophilic head, hydrophobic tail) acting in the highly polar environment of water.

A basic component as a membrane is, from the theoretical viewpoint, characterizable as a higher order, emergent structure [3, 4], dynamically formed by interactions between lipids due to an entropy gradient arising from the structured polar (water) environment. Phase-separation of e.g. lipids in water and a concomitant ordering to vesicles and micelles is a spontaneous process lasting from seconds to minutes [5, 6, 7, 8].

The resulting higher order structures have themselves rich dynamics, e.g. turn-over (flip-flop-mechanism) of single lipids within membrane-like structures. This flexibility is of major importance to maintain functionality in a cell membrane, which hosts systems at higher hierarchical levels as e.g. the complex for photosynthesis [9].

The Hydrophobic Effect (HE), describing mainly an entropic effect, seems to be of fundamental importance in the self-organization of such biological systems [10]. Various circumstances lead to the formulation of a hydrophobic *effect* rather than a hydrophobic *force*: The most reliable indication that it is indeed a hydrophobic effect comes from thermodynamics, considering free energy, enthalpy and entropy of solvation processes:

$$\Delta G = \Delta H - T\Delta S \quad (1)$$

$\Delta G$ ...change of free energy

$\Delta H$ ...change of enthalpy

$T$ ...temperature

$\Delta S$ ... change of entropy

It is experimentally known that the free energy change of dissolution of hydrophobic molecules in a polar solvent is positive, although the change in enthalpy (at room temperature) is often zero or even negative [11]. Considering that liquid water has to some extent quasicrystalline features with highly ordered regions [12, 13], the HE is believed to be based on a change of the water structure in the vicinity of hydrophobic surfaces and a concomitant decrease of entropy. Following this model, the solvent is forced to form a “cave around” the hydrophobic surface. This reaction, often referred as hydrophobic solvation, is accomplished by a different dynamics of the solvent in the vicinity of hydrophobic surfaces compared to the bulk solvent phase: The accessibility of microstates decreases and thus the entropy decreases. Due to the high surface tension of such a molecular cave, the solvent tends to “minimize” its contact surface to hydrophobic molecules which leads eventually to the phase-separation between water and hydrophobic molecules.

Entropy gradients and resulting phase separation are, therefore, based on effects *generated* by the system dynamics. They are not observable as explicit interaction forces, but are the result of basic molecular interactions between hydrophobic particles and the polar solvent. They are *emergent* properties.

## 1.2 Simulation of molecular systems

### 1.2.1 Molecular Dynamics and Monte Carlo methods

Recently various new methods are applied to simulate structural and dynamic properties of macromolecular systems. One such example is genetic algorithms [14], implementing formal criteria of Darwinian evolution through a fitness function (which is in analogy to an energy function). With such a method it is for instance possible to determine secondary structure motifs of polymers as proteins. However, the classical tools to simulate large molecular systems are deterministic routines like molecular dynamics, solving Newton’s equations of motion [15], or stochastic algorithms like the Dynamic Monte Carlo method [16]. These tools are based on force field calculations considering the following terms:

$$V_{total} = \sum_{bonds} V^{i,j} + \sum_{bondangle} V^{i,j,k} + \sum_{torsion} V^{i,j,k,l} + \sum_{electrostatic} V^{i,j} + \sum_{V.d.Waals} V^{i,j} \quad (2)$$

$V$  ... potential energy

$i, j, k, l$  ... atoms, atom groups

The total potential energy  $V_{total}$  of a system with  $n$  atoms is calculated as a sum of individual contributions arising from pairwise intra- and intermolecular interactions. Other types of force fields using different intra- and intermolecular potentials, as knowledge based potentials, or mean field minimization methods, are discussed in reference [17]. Recent investigations have proven the importance of weak intermolecular interactions of the Van der Waals type including a polar solvent for general molecular recognition processes [18, 19]. A total potential energy of a molecular system in solution,  $V_{total}$ , has to be calculated as a sum of conformational and solvation energies:

$$V_{total} = V_{conformational} + V_{solvation} \quad (3)$$

Potential energies arising from solvation are calculated as pairwise interactions based on electrostatic and Van der Waals terms represented as Coulomb and Lennard Jones potentials [20] or as changes in free energy in a continuum approximation [19]. The pair-potentials used in a representation of a solvent do not reproduce cooperative effects as they occur in the hydrogen bonded network of water.

There are three major problems associated with the formulation of molecular dynamics as noted above. (i) Using an atomic level of description instead of a molecular (at the monomer) level of description makes a simulation of molecular self-assembly more complicated than it need to be. With such a low-level description it is not possible to simulate, for instance, processes ranging in a time scale up to minutes like the self-assembly of lipid membranes. The second problem with these descriptors on the atomar level is the high complexity of the simulator itself as for instance shown in the protein folding problem [21]. (ii) Using a (pure) mechanistic in stead of Cellular Automata modeling technique makes the updating of the system too cumbersome and slower than necessary, since one in principle needs to make  $\sim n^2$  calculations (every pairwise interaction) in a system with  $n$  paricles instead of  $\sim n$  calculations. (iii) Using

real (continuous) variables in stead of integers or bit operations also slows down the computation when using digital computers. For a further discussion of these issues we refer to [22].

An excellent overview describing lattice models to simulate macromolecular systems is given in reference [23]. Such Ising spin based models are capable of generating generic phenomena of e.g. phase separation.

### 1.2.2 Lattice Molecular Automaton

The Lattice Molecular Automaton (LMA) is an extension of the Lattice Polymer Automata (LPA) [22] and they are both made in the spirit of the Lattice Gas Automata (LGA) [24]. Both the LGA and the LPA have proven to be capable of generating macroscopic effects based on a microscopic, discrete system representation. In the LMA-approach, molecules also interact on a hexagonal lattice with toroidal boundary conditions. This lattice type has proven to be suitable to avoid anisotropic effects [24]. The molecular entities and vacuum are encoded in data structures on each lattice site to ensure optimal parallel processing. Kinetic and potential energy terms are implemented in the LMA via information particles describing an artificial physics within a microcanonical ensemble; constant volume, constant number of molecules, and constant energy. The inner structure of molecules is not considered, only intermolecular interactions model the dynamics of the system. The forces in the system are determined by a propagation of information - or "force"-particles between neighboring data structures. Due to the explicit discrete character of these information particles, discrete state functions of the data structures can be evaluated (counted) to calculate thermodynamic properties as entropy and enthalpy.

## 2 The Lattice Molecular Automaton Concept

### 2.1 Artificial physics in the LMA

The dynamics of a molecular system depends on kinetic energy terms and on the relative molecular position on a potential energy hypersurface. In a non-dissipative system, the basic conservation laws have to be fulfilled. That is keeping mass,  $M$ , momentum,  $P$ , and the total energy,  $E_{total}$ , constant. What is “optimized” in the equilibrium of such a system is the relative position of molecules, leading, depending on the thermal state, to a certain minimum of the sum potential energy.

$$V_{total} = \sum_{i=1}^n \sum_{j=1}^k V^{j,i} \quad (4)$$

$V^{j,i}$  ... local potential energy situation

$n$  ... number of molecules

$k$  ... number of potential energy terms

In the LMA, as we shall see, a Boltzmann distribution of kinetic energies drives the many particle system into locally stable, sum energy configurations.

#### 2.1.1 Implementation of kinetic energy

Each molecule on the hexagonal lattice has six directions of translation. Each direction is, per definition, independent and occupied by a Boltzmann distribution of kinetic energies. For all particles, directions and times, the kinetic energy for each direction of a translation is larger than zero. In the case of a collision process, the kinetic energies are distributed between the respective molecules following a collision model for hard spheres <sup>2</sup>.

The only way to distribute kinetic energies is via collision processes. During a free translation of a molecule, all currently occupied kinetic energy levels are conserved and the molecule is characterized as an isolated particle. The overall (global) thermal state of the system, as the sum kinetic energy, stays constant in time. This implies

---

<sup>2</sup>This formulation describes pairwise collision processes correctly (conservation of momentum and energy), but only approximates more complicated collision situations including more than two molecules or molecules in polymers. See also section 4.4.3 for more details.



also an overall conservation of momentum in time for each of the six independently treated, principle directions on the hexagonal lattice. The kinetic energy distribution is identical for all considered molecular types, further assuming equal mass for water, hydrophobic and hydrophilic monomers.

### 2.1.2 Implementation of potential energy

The implementation of a potential energy term is necessary to characterize special physico-chemical features of the different molecular types:

Let us consider a model-system for a lipid/water mixture consisting of the following molecules (see also the schematic drawings B1 and C1 in figure 1):

solvent: water,  $H_2O$

polymer: fatty acid,  $CH_3 - CH_2 - CH_2 - CH_2 - COOH$

The most important term for calculating potential energies between uncharged model-polymers are electrostatic and Van der Waals terms (see also equation 2) [25]. In the present LMA, three forces of this type are included to characterize hydrophilic and hydrophobic properties of monomers in the polymer and water: (i) Dipole-dipole interactions as well as hydrogen bonds for water-water and water-hydrophilic monomer interactions, (ii) dipole – induced dipole for water-hydrophobic monomer as well as hydrophilic-hydrophobic monomer interactions and (iii) induced dipole – induced dipole for hydrophobic monomer-hydrophobic monomer interactions. One important aspect for the following considerations is the comparable distance-dependence of all three forces: They are short-ranged in aqueous solution. The relative strength of the three different interactions is, however, different, as the potential energy gain from a typical H-bond is around -2 kcal/mol, the other interactions contribute with energies in the range of -0.5 kcal/mol.

The important features of a water model are the highly polar character of water and the ability to form stable hydrogen bonds (H-bonds), i.e. to have three defined interaction directions (two hydrogens and one oxygen). The high degree of order within liquid water is mainly based on these comparably strong H-bonds [12]. Hydrophilic monomers are either charged, or, too, have the capability to form H-bonds, as e.g. the

carboxyl-group in our model polymer. The interactions between hydrophobic and hydrophilic molecules are weaker and mainly based on the polarizability of the hydrophobic molecules. Interactions between hydrophobic moieties are comparably strong, since electrostatic forces are not shielded by a water shell. The dielectric constant  $\epsilon_r$  decreases from 80 in bulk water phase to  $< 10$  between two neighboring hydrophobic surfaces.

The total potential energy  $V_{total}$  of our model system with  $n$  molecules is thus described by:

$$V_{total} = \sum_{i=1}^n \sum_{j=1}^6 V_{dip.-dip.,H-bond}^{j,i} + \sum_{i=1}^n \sum_{j=1}^6 V_{dip.-ind.dip.}^{j,i} + \sum_{i=1}^n \sum_{j=1}^6 V_{ind.dip.-ind.dip.}^{j,i} \quad (5)$$

, where the sum is over all molecules and over the neighborhood (six directions) for each molecule.

In the LMA simulation environment, dipole – dipole interactions and H-bonds account with a (dimensionless) value of -5, all other binding interactions with a value of -1 to the total potential energy. Water molecules can, summing over the six principle directions, be stabilized by a value of -18 (three H-bonds and three dipole – induced dipole interactions), hydrophilic monomers by a value of -14 (two hydrogen bonds, four dipole – induced dipole interactions), and hydrophobic monomers by a value of -6 (six induced dipole – induced dipole and/or dipole – induced dipole interactions). The relative strength of these interactions is chosen according to corresponding experimentally determined values as noted above.

In the LMA approach, these potential energies provide attractive binding or repelling forces. They influence the occupation time of a given molecule on a particular lattice location with a given kinetic energy distribution. The higher the potential energy value, the more likely it is for a molecule to stay at that given location. If the kinetic energy of a molecule exceeds the sum potential energy at a given location, the molecule will not stay, but continue in the direction where it has its highest kinetic energy. If for instance a hydrogen bond is formed, the binding energy of this particular bond is stored as internal energy in the bond molecules. In the case of a collision, the propagated kinetic energy is compared to the internal (binding) energy. If gains from kinetic energy are larger than contributions from binding, the particular bond breaks up.

### 3 Information Dynamics: The LMA update cycle

The discrete field automata are based on the assumption that all molecular interactions can be modeled by mediating particles [22]. Both matter and fields are interpreted as “information particles” that propagate locally along the edges of a lattice and interact with one another at nodes, as in a Lattice Gas (LGA) [26]. Thus, the rules that generate the dynamics are: (i) The rules that propagate the information particles which depend on the current state of the current site, (ii) The rules that evaluate the newly propagated information together with local states, (iii) together with the chosen update schedule. Unlike a standard LGA and as in the Lattice Polymer Automata (LPA) [22] several different types of information particles are used, so the structure of a node is more complicated than the simple six bit register required for a minimal LGA. The molecular model to be discussed here is formulated on a 2D hexagonal lattice (see figure 1.A).

insert figure 1

Scheme B.2 and C.2 in figure 1 depict the abstract LMA representation of water and monomers in a polymer. All stored information in the data structures is propagated to a given neighborhood (NH) on the hexagonal lattice: Excluded volume particles, “Repellons”, are propagated to neighborhood 1. To prevent polymers from breaking up, “Bondons” are propagated to neighboring monomers in the polymer. The bond length between two monomers in a polymer is fixed to the length of one lattice site. The force particles, propagated to neighborhood 1 and 2, represent the Van der Waals properties of the molecular compounds. They also mimic the H-bonding capabilities of water and hydrophilic monomers, as indicated by the arrows.

The transmission of the force particles between the molecules enables an update of each molecule using only local information. After the information particle transport steps, each lattice site can be updated independently. The force-communicating particles propagate locally, that is, between neighboring lattice sites. A variety of molecular interactions may be formulated by choosing the mediating particles properly. For instance a polymer must obey a connectivity constraint between its monomers and all

molecules must obey an excluded volume constraint. The chemical information hydrophobicity and hydrophilicity as well as the structural information on water are characterized by these force particles. As an example, the propagation of force particles of two water molecules in position  $(i, j)$  and  $(i, j + 1)$  on the lattice is shown in figure 2.

insert figure 2

In figure 2.A, two data structures representing water are depicted after the propagation of the excluded volume information particles, the Repellons. These particles mimic a hard sphere collision between two molecular surfaces in close contact. Their propagation is also the basis for the exchange of kinetic energies in a collision process. In figure 2.B, the propagation step of force particles, the “Attractons”, is shown. In a first step, these force particles are propagated to neighborhood 1, and in a second step to neighborhood 2 (see 2.C). These Attractons represent the binding sites for hydrogen bonds as also denoted schematically in B1, B2, C1 and C2 of figure 1. As can be seen in figure 2.C, the NH<sub>2</sub> propagation step of the force particles does not take place in direction 1 for the monomer on the location  $(i, j)$ , and also not in direction 4 for the monomer on  $(i, j + 1)$ . This mimicks the high dielectric constant in bulk water. Force particles (representing an electric field) are shielded by water, but not by hydrophobic monomers (see also section 2.1.2).

In summary, a simulation update consists of the following steps:

- (1) Propagation of molecular types and redistribution of kinetic energies.
- (2) Construction of type specific force fields.
- (3) Calculation of potential energies.
- (4) Calculation of the most proper move direction.
- (5) Readjustment of bonds in polymers according to the move direction.
- (6) Move the molecule and clear the lattice for the new update.

A detailed description of a full LMA update is presented in reference [1].

## 4 LMA Dynamics: Molecular Dynamics and Molecular Self-Assembly

### 4.1 Representation of the polar solvent

A radial distribution function (RDF)  $g_r$  is a quantity that checks basic “geometrical” features of the solvent in a system. This function gives a probability  $P$  to find a molecule in a certain neighborhood distance  $r$ . For liquid water, experimentally determined RDF (by x-ray scattering) show a peak at 0.25 nm (H-bond distance) and smaller peaks for the next theoretical H-bond distances (temperature-dependent). This function reflects the ordered structure of the neighborhood of a particular water molecule in the liquid phase, mainly based on the formation of hydrogen bonds.

insert figure 3

Figure 3 shows a RDF obtained by a LMA simulation of a polar solvent like water (see figure 1, B.1, C.1) after  $10^5$  simulation steps. 50% of the lattice sites are occupied with water. The probability  $g_r$  in the range  $[0, 1]$  is plotted versus the neighborhood distance ( $NH$ ) on the hexagonal lattice. There is an increased probability to find another water molecule for neighborhood 1, 2, and 3 (e.g. the lattice positions  $(i, j + 1)$ ,  $(i, j + 2)$ ,  $(i, j + 3)$  in direction 1), since  $g_r > 0.5$ . This RDF shows the comparably higher local clustering of water molecules based on the H-bond binding energies. Without implementation of these stabilizing energy contributions, the RDF shows no peaks characterizing an ordered neighborhood and the molecules are randomly distributed throughout the lattice ( $g_r = 0.5$  with a 50 % lattice occupation). The turn-over rate (change of molecule position per time step) is in the range of 30 %, but the general shape of the RDF remains unchanged in time.

insert figure 4

Figure 4 shows two snapshots (time-difference of 10 simulation steps between A and B) of a water simulation generating the RDF given in figure 3, where 50% of the lattice sites are occupied with water (denoted by unfilled circles). The formation of local, irregular clusters is shown, but within the next 10 time steps a global rearrangement of cluster

structures takes place. The crucial point to obtain this behavior in the simulation system is an appropriate balance of kinetic and binding (potential) energies.

## 4.2 Balance between kinetic and potential energies

The summed binding energy  $V_{b,total}$  and the summed kinetic energy  $K_{total}$  are given by:

$$V_{b,total} = \sum_{k=1}^n \sum_{l=1}^6 |V_{binding}^{l,k}| \quad (6)$$

$$K_{total} = \sum_{k=1}^n \sum_{l=1}^6 K^{l,k} \quad (7)$$

l ... index over the 6 principle lattice directions

k ... index over  $n$  molecules

In a non-dissipative system, the total inner energy is constant. In the LMA,  $K_{total}$  is by definition constant. The summed binding energy is in the mean constant in the equilibrium situation. The value of  $V_{b,total}$  will in general decrease over time until a (local) minimum value has been obtained which corresponds to equilibrium. The yield of potential energy is implicitly stored in the system (see also section 5).

The mean value of  $V_{b,total}$  is proportional to the fixed value of the thermal state,  $K_{total}$ , of the whole system:

$$K_{total} = a_{global} V_{b,total} \quad (8)$$

$a_{global}$  ... global proportionality factor

The global proportionality factor,  $a_{global}$ , is proportional to the temperature of the system (high kinetic energy corresponds to high temperature) and defines intrinsically the dynamics of the system:  $a_{global} < 1$  fixes the system in a local minimum where the variation of particle position in time tends towards zero (in analogy to a spin glass at low temperature). To obtain the formation of unstable, but locally ordered clusters, as shown in figure 4,  $a_{global}$  has to be in a range of 3-4 (3.7 for the system shown in figure 4). The mean binding energy per lattice site for this water simulation is around 5, summed over the six principle directions. This indicates, that in the mean one H-bond is formed per water molecule. The respective value for the mean kinetic energy, again

summed over the six directions, is 21. High values of  $a_{global} > 6$  enhance the disorder in the system: The structured radial distribution function, as shown in figure 3, vanishes and a random distribution with values near 0.5 is obtained for all neighborhoods. This finding is equivalent to the experimentally determined change of RDF of liquid water at high temperature.

While  $a_{global}$  on average is constant in time in equilibrium, the local proportionality  $a_{local}^{(i,j)}$  between  $V_b^{(i,j)}$  and  $K^{(i,j)}$  at the site  $(i,j)$  exhibits strong fluctuations.

insert figure 5

The values  $a_{local}^{(i,j)}$  of the time series in figure 5, thin line, show the local change of kinetic and binding energy in time (given in simulation steps). They are summed over the six principle directions on the lattice and recorded during 500 time steps in the equilibrated system for one of the water molecules in figure 4. These local characteristics of the molecular dynamics differ completely from the global behavior: Fluctuations between binding and kinetic energies characterize the local situation, whereas the relation between these two terms is on average constant in time for the global system (thick line, figure 5), showing the global proportionality factor of 3.7, as discussed above. This picture represents the local flow of binding (ordering) and kinetic (disordering) energies on a particular water molecule in the global equilibrium situation. In the case  $a_{local}^{(i,j)}$  is larger than  $a_{global}$ , excess kinetic energy drives the molecule into free translation, if  $a_{local}^{(i,j)} < a_{global}$ , the molecule is in the mean fixed on a cluster position in the hydrogen bonded water network. This relationship of the local instability is the cause for the formation of local, unstable clusters in the model for a polar solvent like water.

In this way the system is tunable between a “quasi-crystal” phase and a “gas” phase by altering  $a_{global}$ .

### 4.3 Dynamics of hydrophobic monomers in a polar environment

The main type specific feature of hydrophobic monomers in the LMA is the interaction with other hydrophobic monomers according to the induced dipole – induced dipole type (see section 2.1.2). The relative strength of this force in the present LMA setup

is set equal to the dipole – induced dipole interaction between water and hydrophobic monomers. Hydrophobic monomers have, therefore, no binding preference for water or other hydrophobic monomers.

### insert figure 6

Figure 6 shows snapshots of a water-hydrophobic monomer system,  $10^5$  simulation steps after the initial random mixing of both molecular types. 50 % of the lattice sites are occupied by molecules. The fraction of hydrophobic monomers (black filled circles) is 16% in (6.A) and 25% in (6.B). The remaining fraction (34% and 25%) is filled with water (unfilled circles). In (6.A), the mean binding energy (see equation 6) of a hydrophobic monomer, again summed over all six directions, is 2.5 and the value for water is 5.5. The respective values for (6.B) are 4.2 and 5.7. The mean kinetic energy for both molecular types is again 21 for both simulations. The global proportionality  $a_{global}$  is, after  $10^5$  updates, 4.4 for (6.A) and 4.2 for (6.B). The increase of the global proportionality factor (compared to pure water) is mainly due to the loss of hydrogen bonds between water molecules, whose energy contributions are in this system only partly counteracted by the binding energy based on water-hydrophobic monomer and hydrophobic monomer-hydrophobic monomer interactions.

However, the mean stabilization of a water molecule in a mixture is larger than in the bulk (comparing the values of 5.5 and 5.7 for binding energies denoted above with the value 5.0 discussed in section 4.2).

The hydrophobic monomers (black filled circles) start to form clusters in the polar environment and the mean cluster size depends intrinsically on the fraction of water to hydrophobic monomers. The dynamics to form larger clusters is rather slow after the first formation of small clusters as found in figure (6.A). This behavior is based on nonlinear kinetics: First the comparably fast formation of small clusters, then the slower diffusion dynamics of these clusters to form larger hydrophobic domains.

The energetical basis of the cluster-formation itself is the most interesting part. The clustering takes place, although the binding energy between two hydrophobic monomers is comparably weak - much weaker than the water-water binding (ratio 1:5 in the current setup) and only as strong as the hydrophobic monomer-water interaction. A hydrophobic molecule has therefore no binding energy preference for water or another



hydrophobic monomer, but still clustering occurs. The reason for the starting phase separation is, therefore, based on the properties of the polar solvent, as will be shown in section 5. Complete phase separation can be simulated by increasing the induced dipole - induced dipole interaction to a value  $\leq -2$  instead of  $-1$ . But already the equal interaction strength between water and hydrophobic monomers as well as between different hydrophobic monomers is sufficient for the clustering.

## 4.4 Polymer dynamics in the LMA

### 4.4.1 Hydrophobic and hydrophilic pentamers in a polar environment

The snapshots of figure 7 (after A:  $10^3$ , B:  $5 \cdot 10^4$  and C:  $10^5$  updates) of a simulation of five hydrophobic pentamers in the polar solvent show the progress of cluster formation. Black filled circles denote hydrophobic monomers, unfilled circles denote water molecules.

insert figure 7

This clustering of hydrophobic polymers is not an artifact based on the update rules as can be shown easily when simulating the dynamics of hydrophilic pentamers (which “like” to be in contact with water, based on the formation of strong H-bonds) in the same setup. Hydrophilic monomers are always completely solvated by water molecules and take part in the H-bond network. Free hydrophilic monomers and hydrophilic polymers show no cluster formation in the polar solvent.

The clusters formed by hydrophobic polymers are not ordered, as there is no additional information for orienting or aligning the polymers in a well defined way, as shown in the next paragraph.

### 4.4.2 Lipid-like pentamers in a polar environment

The present LMA setup is already very sensitive to the variation of the molecular types involved in the dynamics. Figure 8 shows intermediates of simulating lipid-like polymers (as schematically shown in figure 1.C) in the polar environment.

insert figure 8

The hydrophilic head-monomers in the pentamer are indicated by the open, large circles, the hydrophobic tail-monomers by the black filled circles. The small circles indicate water molecules. The topology of these clusters (A:  $10^3$ , B:  $5 \cdot 10^4$  and C:  $10^5$  updates) as well as their formation dynamics, especially with respect to a long time scale needed to form clusters (as discussed in section 4.5), differ strongly from the example shown in figure 7. These clusters (the dimeric associates in 8.A and 8.B as well as the quatermeric in 8.C) are ordered in such a way, that the hydrophilic head-monomers always stay in contact with water, whereas the hydrophobic tails try to cluster. This example shows that even a slight variation of the physico-chemical type of only one monomer in the pentamer essentially changes the cluster formation process.

Thus, there is a crucial dependency of the macromolecular aggregation on the chemical type of the molecular entities involved. To ensure proper functionality of higher order structures composed by supramolecular ensembles, specific information has to be present on the basic molecular entities. *The higher the order is and the more specific a task of a biological structure the more information has to be present at the underlying chemical entities to specify the structure and thus its functionality* [4].

#### 4.4.3 Polymer update on a 2D lattice

The treatment of polymers in the present LMA setup is confronted with two basic problems, on the one hand the dimensionality of the lattice, on the other hand the parallel update of an extended object only based on local, discrete rules:

##### *Dimensionality:*

A two dimensional lattice is sufficient to simulate monomeric fluid flow in a qualitative and quantitative correct way, as also demonstrated by the results of the Lattice Gas Automata simulations formulated in two dimensions [24]. Due to the bonds, polymers reduce the degrees of freedom in their vicinity for other monomers significantly. A straight forward simulation of polymers and solvent in the LMA yields typical dimension-based phenomena in the solvent structure around polymers, as e.g. dense clustering. To avoid these effects and to study solvent properties in "local 3D", a second lattice can be introduced to allow solvent molecules to pass positions occupied by monomers in polymers if they are not bond to any other molecule (or more exact, if

their thermal state is higher than their yield of binding energy on the present position). This second lattice corresponds to a solvent reservoir. If a solvent molecule enters the second lattice, an identical solvent molecule, with respect to relative orientation and thermal state, is released onto a free work-lattice position in the move direction of the particular solvent molecule. We used this approximation in the above discussed simulation of polymers.

#### *Parallel polymer update:*

The real time movement of a polymer is an intrinsically parallel process. In a physical system the translational state of a subpart (monomer) of the polymer in time  $t$  is instantaneously (with the speed of light) known to all other subparts. To realize this global information flow, at least  $(l - 1)$  propagation steps in a polymer of the length  $l$  have to be performed in the LMA. To resolve possible conflicts requires additional information propagation steps. For a more detailed discussion of these issues we refer to [22, 1]. A polymer in the LMA is characterized by an elastic deformability of the monomers in a collision process. The information about the energetic state of a monomer is only propagated to the bond neighbors (i.e. monomers in NH 1) in the chain in one time step, assuming that the propagation of the momentum along the chain is slow (1 site per time step) due to the time lack based on the deformation of subunits (librational and vibrational modes of bonds). The same feature also holds for the LMA solvent molecules, where momenta are also only propagated to neighborhood 1 in one time step. This is because a simultaneous fulfillment of (i) a strict parallel update, (ii) strict local interaction rules, and (iii) a strict conservation of momentum, is not possible [27]. For other cellular automata based polymer updating methods we refer to [28] and [29].

## 4.5 Correlation of LMA-updates to a physical time scale

The representation of a LMA simulation in an absolute time scale can only be roughly estimated. The overall relaxation times of polar liquids, determined e.g. by ultrafast fluorescence spectroscopy [30], are in the range of one picosecond. The underlying processes in this time regime are the dynamics of fast, mainly librational and rotational, modes in between 10 and 100 femto seconds, followed by a slower, longitudinal relaxation up in the picosecond range.

The longitudinal relaxation in our lattice model is defined by the translation of a particular water molecule from one lattice site to a site in neighborhood 1. In the mean three full updates of the lattice are needed to translate one water molecule (at the given global proportionality factor of 3.7). These three updates could now be assigned to a one picosecond time step. Following this interpretation, the first two updates correspond to librational and rotational modes of the LMA water molecule (i.e. rotation of a molecule on a given lattice site) and the third update is in the mean the translational contribution to the relaxation time.

The update of a hydrophobic monomer in bulk water is comparably faster, in the mean two updates, which correspond to one translational relaxation. The relaxation time of a hydrophilic monomer is near to the relaxation time of water in our model. The exact values for the system with  $\alpha_{global} = 3.7$  are 3.3 updates for water, 3.0 updates for hydrophilic monomers, and 1.9 updates for hydrophobic monomers in bulk water phase.

However, all calculated monomer systems show a comparable update time in the range of  $< 4$  updates and, related to a time scale, a longitudinal relaxation time should be in the range of one picosecond.

The time for updating the model-pentamers is, due to the constraints introduced by bonds, much longer. A full translational polymer update, and the corresponding translational relaxation, is defined by at least one translation of each monomer in the polymer by a distance one on the lattice. The translational relaxation of a hydrophilic pentamer takes in the mean  $1.5 \times 10^4$  LMA steps, of a hydrophobic pentamer  $6 \times 10^3$  (see figure 7) and of a lipid-like pentamer, as show in figure 8 ,  $1.2 \times 10^4$  LMA update steps. These numbers of update steps correspond, taking the update of a single water molecule as reference (where three updates correspond to 1 ps), to 4-5 ns for the hydrophilic, to 2 ns for the hydrophobic, and to 4 ns for the lipid-like pentamer, respectively.

Surprisingly, the translational relaxation of a lipid-like pentamer is similar to the time range of a hydrophilic pentamer, although only one monomer in the pentamer is hydrophilic. This decreased flexibility of lipid-like polymers, compared to hydrophobic polymers, is also reflected by an increased cluster stability (figure 8) compared to hy-

drophobic clusters (figure 7).

In general, the translational update-times for polymers in the LMA-model seem to be overestimated. This is probably mainly due to the reduced number of degrees of freedom when simulating polymer dynamics on a 2D grid. The simple molecular dynamics and molecular self-assembly simulation discussed here are thus typically simulated up to the order of 0.1 micro second real time and they are easily performed on a PC or a small workstation.

## 5 Thermodynamic Characterization of the LMA

The driving force for dynamics within the LMA is, formally, based on two, in the simulation strictly separated, basins of energy:

- (1) The thermal state, consisting of translational energy and energy distributed over inner degrees of freedom within a molecule as e.g. rotational terms. This thermal state of a molecule is either stored on inner degrees of freedom during the free translation or distributed to other molecules in a collision process.
- (2) The potential energy between specific sites on the three molecular entities recruiting attracting and repelling forces.

Binding of a molecule influences the occupation time of a molecule on a particular lattice location. If the highest kinetic energy in a direction exceeds the sum binding energy of the molecule the intermolecular bond breaks. However, the sum inner energy on the molecule remains unchanged.

Monomers in the LMA can be interpreted as an ideal gas, if all occupied data structures (or lattice sites) are separated by at least a neighborhood of three. In this case, the total energy of the system,  $E_{total}$ , is given by the sum kinetic energy  $K_{total}$ . Compressing this “ideal gas” to the (constant) simulation volume results in a decrease of the potential energy,  $-V_{b,total}$ , as intermolecular interactions are realized and bonds are formed. The decrease of this potential energy during the simulation, based on the “optimization” of the intermolecular interactions and depending on the global proportionality factor  $a_{global}$ , is *implicitly stored as inner (inter-molecular) energy*. Note that the distribution of the kinetic energies in each direction as well as  $K_{total}$  itself is not changed in time. Thus all occurring energy losses are not considered explicitly, but the inner energy provides the necessary energy whenever a bond is broken. Recall that the decrease of this potential energy during the simulation, due to the bond formation and depending on the global proportionality factor  $a_{global}$ , is implicitly stored as inner inter-molecular energy - and that it can always be recovered again as molecular assemblies are broken up.

The total energy of the system is thus given by:

$$E_{total} = K_{total} + V_0 = constant, \quad (9)$$

where  $V_0$  is the potential energy or binding energy term when the simulation is started.

The system has a constant total energy. The total number of grid points on the lattice, i.e. the volume, and the total number of molecules is also constant during the simulation. Thus, the LMA confirms to a *microcanonical ensemble*.

The above interpretation allows the formal definition of the entropy in the following way:

$$S = k_{system} \ln Z, \quad (10)$$

where  $k_{system}$  is a LMA intrinsic constant corresponding to the Boltzmann constant  $k_B$  for a physical system,  $Z$  is a partition function over the states of the molecules in the simulation. Since  $Z$  is not known explicitly (we would have to derive the Hamiltonian for the LMA) an approximation of the entropy,  $S_{LMA}$ , can easily be calculated.  $S_{LMA}$  measures an occupation pointer as an approximation to the partition function of  $n$  molecules on the lattice and does not explicitly take the particular energetical states of a molecule into consideration:

$$S_{LMA}(t) = k_{system} \ln \frac{\sum_{i=1}^n D_i}{n}, \quad (11)$$

where  $D_i$  is an occupation pointer of molecule  $i$  with the value 0, if the molecule does not change its location  $(i, j)$  on the lattice between time  $t$  and  $t + 1$ , and 1 if the molecule changes the position in one update.

This definition of entropy only considers the location parameter of the generalized coordinates  $\Gamma$ , but not explicitly the entropic impact of the distribution of energetic states, as noted for a general partition function in equation (10). But, as we will see below, this *positional* or *structural* entropy is the manifestation of the specific energetic state of monomers in given neighborhood situations.

In the case of the ideal gas-LMA in an infinite simulation volume (no interaction between molecules, no excluded volumes), the sum over all  $D_i$  is equal to  $n$  and the entropy is therefore zero for this reference system. Figure 9 shows the dynamics of entropy as defined in equation (11) for water and water:hydrophobic monomer mixtures. The corresponding energetic implication for the systems shown in figure 6.B (25% water, 25% hydrophobic monomer mixture) compared to a water simulation (shown in

figure 4) is given in figure 10:

### insert figure 9,10

The decrease of entropy ( $S_{LMA}$ ), based on the stabilization of molecules by bond formation, is especially monitored during the first 100 steps of the simulation of a water system, curve *a*, and the two water:hydrophobic monomer mixtures, curve *b* and *c*, as shown in figure 6A (16 % hydrophobic monomers) and 6B (25 % hydrophobic monomers). The mean entropy for the water system is -1.55 and remains constant also after longer simulation times. The entropy is lower for the mixtures, -1.9 for system *b* (16 % hydrophobic monomers) and -2.5 for system *c* (25 % hydrophobic monomers). These values continue to decrease, e.g. the respective values for system *b* are -1.95 after  $10^3$  simulation steps and -2.65 after  $10^5$  steps. This again reflects the comparably slow phase separation process for the mixtures.

The entropy, as defined in equation (11), is generally lower for the mixtures compared to pure water and depends on the concentration of hydrophobic monomers in the system, since the value of  $S_{LMA}$  is lower for *c* compared to *b*. This is especially of interest when considering the  $a_{global}$  values, 3.7, 4.4 and 4.2, for the three systems (as discussed in section 4.2 and 4.3). These values indicate the opposite behavior. A general *destabilization* of molecules in a mixture and a concomitant *increase of entropy*, since an increase in  $K_{total}/V_{b,total} = a_{global}$  is observed.

This increase is mainly based on the fact, that the loss of H-bonds between water molecules is not completely counteracted by dipole - induced dipole and induced dipole - induced dipole interactions (see section 4.3). The reason for clustering and the decrease of entropy is based on the *local* (and not the global, as indicated by  $a_{global}$ ) distribution of energetic states, which is not explicitly considered in equation (11) or figure 9.

Figure 10 denotes the probability  $P_{state}$  to find a water molecule with a particular binding energy ranging from -1 to -18 (three H-bonds, 3 dipole - induced dipole interactions). The unfilled circles denote the  $P_{state}$ -values for water molecules in bulk water phase (as in figure 4). Three main peaks are found at -5, -10 and -15, corresponding to the three possible hydrogen bonds on one particular water molecule. The same overall distribution is also found for the mixture with 25% hydrophobic and 25% water



molecules on the lattice (see also figure 6.B). The  $P_{state}$  values calculated for water molecules neighboring at least one hydrophobic monomer are denoted by the black boxes in figure 10.

The difference in the mixture is the change in the overall probability of water to form a hydrogen bond: Especially water molecules in the neighborhood of hydrophobic monomers are in the mean stabilized to a higher extend, mainly at the binding potential of -5 ( $P_{state} = 0.53$  for water neighboring hydrophobic molecules, 0.42 for water neighboring water).

This inhomogenous, local distribution of binding energies is the reason for the decrease of entropy for the mixture systems. It is the reason for the entropy-decrease shown in curve *b* and *c* in figure 9. This finding indicates the formation of comparably stable structures around hydrophobic clusters and therefore, a slower dynamics in the vicinity of hydrophobic surfaces. This finding corresponds to the experimentally determined decrease of entropy when solvating hydrophobic molecules at room temperature and is in accordance with other models describing hydrophobic effects showing slow dynamics of water molecules in the hydration shell of hydrophobic surfaces [31, 32, 33].

It should be noted, that the phase separation process of hydrophobic monomers in a polar environment as well as the concomitant decrease of entropy, are not explicitly implemented properties of the molecular elements in this simulation. They are the *result* of the system dynamics. These hydrophobic effects are thus *emergent properties* of the molecular dynamics.

## 6 Conclusion

We have presented a new type of molecular dynamics and self-assembly simulation, the Lattice Molecular Automaton (LMA), which is able to handle very large molecular systems over long time (up to the range of seconds). In the LMA all interactions (electrostatic forces) are decomposed and communicated via propagating force particles or “photons”. The monomer-monomer bond forces, the molecular excluded volume forces, the longer range intermolecular forces, the polymer-solvent, and the solvent-solvent interactions are all modeled by propagating information particles.

The concept of lattice data structures, sharing locally propagated information, is the basis for this kind of molecular dynamics and self-assembly simulation. The data structures are interpreted as computational nodes, storing type specific characteristics of molecules as water, hydrophilic and hydrophobic monomers. The update of the state of the data structures at each lattice site is based on the local information and the received (propagated) information from a given neighborhood on the discrete lattice.

The LMA is a tool that enables the study of the principle, physical mechanisms that generate higher order molecular structures:

- (i) A polar solvent like water is characterized by clustering of water molecules in a hydrogen bonded network with fast rearrangement dynamics. The water structures generated in the LMA produce the same radial distribution function (RDF) as experimentally measured.
- (ii) The Hydrophobic Effect, which is not explicitly encoded in the LMA, is correctly *generated* by the interactions between the hydrophobic monomers and the water molecules.
- (iii) Phase separation of hydrophobic monomers in water in the LMA system follows the same dynamical characteristics as have been experimentally determined. This clustering dynamics is — based on the high binding energies of hydrogen bonds — feasible from the enthalpic point of view, but not for the entropic state of the system. The structural features of liquid water are in a sensitive balance of entropy and enthalpy. The dissolution of hydrophobic particles exactly influences this balance.

The consequence of this perturbation, the *generated* Hydrophobic Effect, drives the

phase separation of hydrophobic molecules in a polar environment.

(iv) LMA lipid polymers form ordered, higher order molecular structures in water.

(v) The discrete, microscopic system representation, which enables a direct calculation of thermodynamic properties of a microcanonical ensemble, gives a direct way to compare results from simulation with thermodynamic data from experiments and further determine how different microscopic effects contribute to the macroscopic thermodynamic quantities. The change of entropy during solvation processes can for instance directly be followed in the LMA system.

(vi) The molecular water structures around hydrophobic surfaces can be directly inspected and the entropic changes followed.

Vesicles and micelle-like structures are typical examples of higher order molecular structures that can be generated by lipids:water mixtures. The driving force for the phase separation and ordering of lipids in water is the Hydrophobic Effect, which is a result of the interactions of the hydrophobic and hydrophilic molecules arranged in polymers and the polar water molecules. Minimal physico-chemical properties of molecular entities represented as information in data structures on the 2D lattice allows the LMA system to have constructive dynamics and generate higher order molecular structures, as e.g ordered lipid aggregates.

The data structure concept of the LMA is of course expandable to represent more details of the Physics<sup>3</sup> so that yet higher order interactions and thus structures can be generated. It should for instance be possible to have “membrane proteins” to assemble into the vesicles so that a transport of molecules becomes possible, perhaps to fuel a chemical reaction inside the vesicle that can change the vesicle [8]. Also an assembly of larger molecular units as e.g. whole micellular entities is possible by appropriate expansion of the data structure.

As we in this way may step up into the dynamical hierarchy of molecular structures the computational resource requirements will eventually explode and we will be forced to change the level of description if we want to understand the dynamics of these processes. We would need to go to a higher level of description and now interpret each data

---

<sup>3</sup>We use the term Physics to denote the real world and not our models of the real world which we denote physics.

structure as a molecular aggregate itself and change the data structure appropriately. Science has to a large extent been successful by doing exactly that - chosen appropriate levels of description. With the LMA we have developed a simulation tool that allows us to investigate the dynamics of the generation of molecular hierarchies which are processes that are not well understood.

### **Acknowledgement:**

B.M. and G.K. thank the *Fonds zur Förderung der wissenschaftlichen Forschung* in Austria (project P-09750-CHE) for financial support. The authors also thank Kai Nagel for critical and constructive comments to earlier version of this paper.

## References

- [1] B. Mayer and S. Rasmussen, *Lattice Molecular Automaton: A Discrete Computational System for Constructive Molecular Dynamics.*, submitted to Phys. Rev. E (1996); Preprint available from B. Mayer, Institute for Theoretical Chemistry, UZAll, University of Vienna, Althanstraße 14, A-1090 Vienna, Austria or S. Rasmussen, TSA-DO/SA MS-M997, Los Alamos National Laboratory, Los Alamos, NM 87545, USA (Technical Report number LA-UR 96-1732).
- [2] R. B. Gennis, *Biomembranes* (Springer, Berlin, 1988).
- [3] N. A. Baas, in *Artificial Life III*, SFI Studies in the Science of Complexity, edited by Ch. G. Langton, Proc.Vol. XVII (Addison Wesley, 1993), p.515.
- [4] N. A. Baas, M. W. Olesen and S. Rasmussen, *Generation of Higher Order Emergent Structures*, submitted to Physica D (1996); (Technical Report LA-UR 96-2921) Preprint available from S. Rasmussen, TSA-DO/SA MS-M997, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
- [5] Sh. Chiruvolu, S. Walker, J. Israelachvili, F. J. Schmitt, D. Leckband and J. A. Zasadzinski, *Science* **264**, 1753 (1994).
- [6] H. Hotani, R. Laholz-Beltra, B. Combs, S. Hammerhoff and S. Rasmussen, *Nanobiology* **1**, 67 (1992).
- [7] J. Tabony, *Science* **264**, 245 (1994).
- [8] T. Oberholzer, R. Wick, P. L. Luisi and Ch. Biebricher, *Biochem. Biophys. Res. Comm.* **207**, 1, 250 (1995).
- [9] J. A. Tuszynski, H. Bolterauer and M. V. Sataric, *Nanobiology* **1**, 177 (1992).
- [10] Ch. Tanford, *Science* **200**, 1012 (1978).
- [11] L. R. Pratt, *Ann. Rev. Phys. Chem.* **36**, 433 (1985).
- [12] P. Rossky, *Ann. Rev. Phys. Chem.* **36**, 321 (1985).
- [13] A. Henn and W. Kautzmann, *J. Phys. Chem.* **93**, 3770 (1989).

- [14] J. T. Pedersen and J. Moult, *Curr. Opin. Str. Biol.* **6**, 2, 227 (1996).
- [15] W. F. van Gunsteren, J. F. Luque and P. Timms, *Ann. Rev. Biophys. Biomol. Struct.* **23**, 847 (1994).
- [16] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, *Science* **220**, 671 (1983).
- [17] a: R. L. Jernigan and I. Bahar, *Curr. Opin. Str. Biol.* **6**, 2, 195 (1996).  
b: P. Koehl and M. Delarue, *Curr. Opin. Str. Biol.* **6**, 2, 222 (1996).
- [18] P. Privalov and S. J. Gill, *Adv. Prot. Chem.* **39**, 191 (1988).
- [19] B. Mayer, G. Marconi, Ch. Klein, P. Wolschann and G. Köhler, *Structural Analysis of Host-Guest Systems*, *J. Phys. Chem.*, submitted for publication (1995);  
Ch. T. Klein, B. Mayer, G. Köhler and P. Wolschann, *Influence of Solvation on the Helix Formation of Poly-Alanine Studied by Multiple Annealing Simulations*, *THEOCHEM*, accepted for publication (1996); Both preprints are available from B. Mayer, Institute for Theoretical Chemistry, UZAH, University of Vienna, Althanstraße 14, A-1090 Vienna, Austria.
- [20] T. Solmajer and E. L. Mehler, *Int. J. Quant. Chem.* **44**, 291 (1992).
- [21] G. Böhm, *Chaos, Solitons and Fractals*, **1**, 4, 375 (1991).
- [22] S. Rasmussen and J. R. Smith, *Ber. Bunsenges. Phys. Chem.* **98**, 9, 1185 (1994).
- [23] K. A. Dawson, in *Structure and Dynamics of Strongly Interacting Colloids and Supramolecular Aggregates in Solution*, edited by S. -H. Chen, J. S. Huang and P. Tartaglia, Kluwer Academic Publishers, 1992, p.265, and references cited therein.
- [24] U. Frisch, B. Hasslacher and Y. Pomeau, *Phys. Rev. Lett.* **56**, 1722 (1986).
- [25] S. K. Burley and G. A. Petschko, *Adv. Prot. Chem.* **39**, 125 (1988).
- [26] B. Hasslacher, *Discrete Fluids*, Los Alamos Sciences - Special Issue, (Los Alamos National Laboratory, Los Alamos, NM, 87545, 1987).
- [27] Y. Bar-Yam, Boston University, private communication (1995).
- [28] B. Chopard, *J. Phys. A* **22**, 1671 (1990).

- [29] O. Ostrovsky, M. A. Smith and Y. Bar-Yam, *Ann. Rev. Biophys. Biomol. Struct.*, **24**, 239 (1995).
- [30] M. Marconelli, *J. Mol. Liqu.* **57**, 1 (1993).
- [31] D. A. Zichi and P. J. Rossky, *J. Chem. Phys.* **84**, 2814 (1986).
- [32] P. J. Rossky and M. Karplus, *J. Am. Chem. Soc.* **101**, 1913 (1979).
- [33] H. Tanaka and K. Nakanishi, *J. Chem. Phys.* **95**, 3719 (1991).

## Figure captions:

**figure 1:** (A): Six principle directions on a hexagonal lattice. (B.1) and (C.1): Schematic representation of water and a polymer in the LMA. (B.2) and (C.2): Propagation of information particles to maintain excluded volume (Repellons), bonds between monomers in polymers (Bondons) and force particles, Attractons, generating type specific force fields, to neighborhood (NH) 1 and 2 on the hexagonal lattice, respectively.

**figure 2:** Particle propagation for two water molecules in position  $(i, j), (i, j + 1)$  on the hexagonal lattice: (A): Propagation of Repellons to NH1, ensuring excluded volume. (B,C): Propagation of force particles, Attractons, to NH1 and NH2 describing a type specific force field.

**figure 3:** Radial distribution function  $g_r$  of a LMA water simulation (for definition see text): 50% of the data structures on the lattice are covered with water, 50% are empty. The figure is based on a set of  $10^4$  molecules after  $10^5$  simulation steps

**figure 4:** Two snapshots of the LMA water representation in equilibrium at time  $t = 10^5$  simulation steps, (A), and at time  $t = 10^5 + 10$  steps, (B). Data structures representing water are denoted as circles.

**figure 5:** Evolution of the local proportionality factor  $a_{local}^{(i,j)}$ , thin line, and the global proportionality factor  $a_{global}$ , thick line, (for definition see text) for a LMA water simulation during 500 time steps in equilibrium.

**figure 6:** Snapshots of mixtures of water:hydrophobic monomers after  $10^5$  updates. (A): 16% hydrophobic monomers, 34% water, 50% empty; (B): 25% hydrophobic monomers, 25% water, 50% empty. Water is denoted as unfilled, hydrophobic monomers as black filled circles. Cluster formation (phase separation) is occurring for both mixtures, strongest for the mixture with most hydrophobic monomers.

**figure 7:** Cluster formation in a LMA polymer simulation: (A):  $10^3$  time steps; (B):

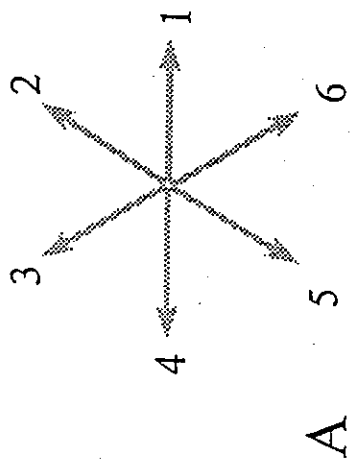


$5 \cdot 10^4$  time steps; (C):  $10^5$  time steps. Hydrophobic monomers in the pentamers are denoted as black filled circles, water molecules as unfilled circles.

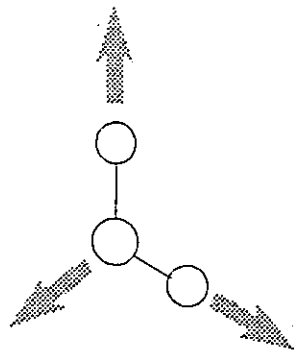
**figure 8:** Snapshots of typical polymer clusters in a simulation of lipid-like pentamers in a polar environment. The details A, B (dimers) and C (quatermer) are taken from a run simulating five pentamers. Unfilled, large circles denote the hydrophilic head monomer in the pentamers, black filled circles denote hydrophobic tail monomers, the small circles denote water molecules (Note that all the molecules have the same “size” in the simulation - they only differ in this graphical representation).

**figure 9:** Evolution of entropy, as formulated in equation (11) (see text), for the first 1000 updates of a water simulation, curve (a), and two mixtures of water and hydrophobic monomers: curve (b) 16% hydrophobic monomers, 34% water, 50% empty, and curve (c) 25% hydrophobic monomers, 25% water, 50% empty.

**figure 10:** Probability  $P_{state}$  to find a particular water molecule with a certain binding energy in the range  $[0, -18]$ .  $P_{state}$  values below 0.02 are not shown. Unfilled circles denote  $P_{state}$  values for molecules in bulk water (as in the system shown in figure 4), black filled boxes denote  $P_{state}$  values for water molecules in a mixture neighboring at least one hydrophobic monomer (equal concentration, see also figure 6.B)

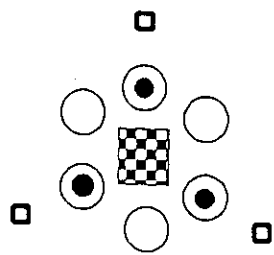


A



1

B



2

← 'H-bond'

type 1, 'water'

type 2, hydrophobic ( $\text{CH}_n$ )

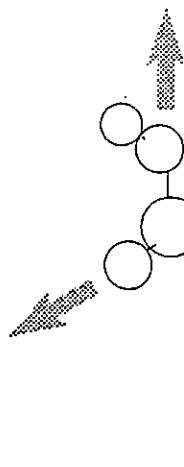
type 3, hydrophilic ( $\text{COOH}$ )

○ Repellon,  $\text{NH}_1$

■ Bondon,  $\text{NH}_1$

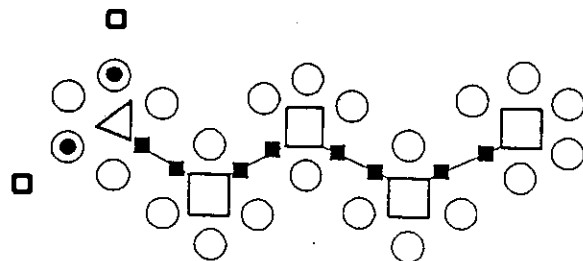
● force particle,  $\text{NH}_1$

◻ force particle,  $\text{NH}_2$



1

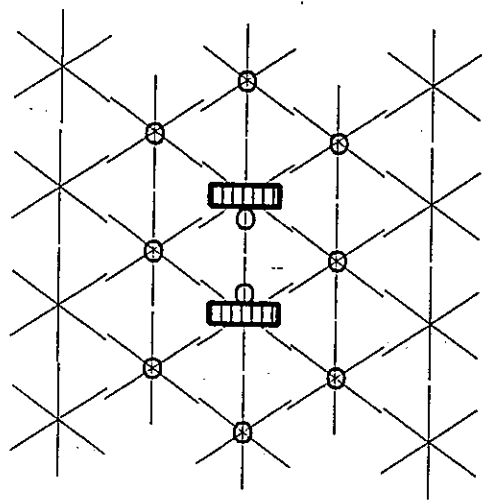
C



2

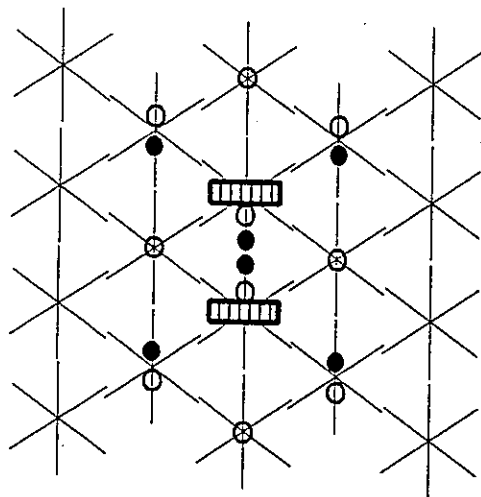
Fig. 1

$j-1 \quad j \quad j+1$

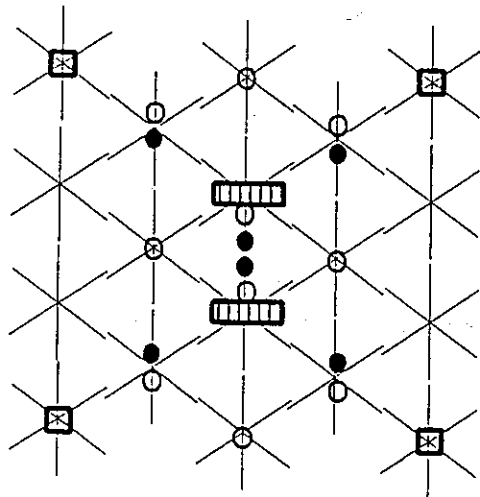


$i-1 \quad i \quad i+1$

*A*



*B*



*C*





-  data structure
-  Repellon
-  NH1 force particle
-  NH2 force particle

Fig. 2

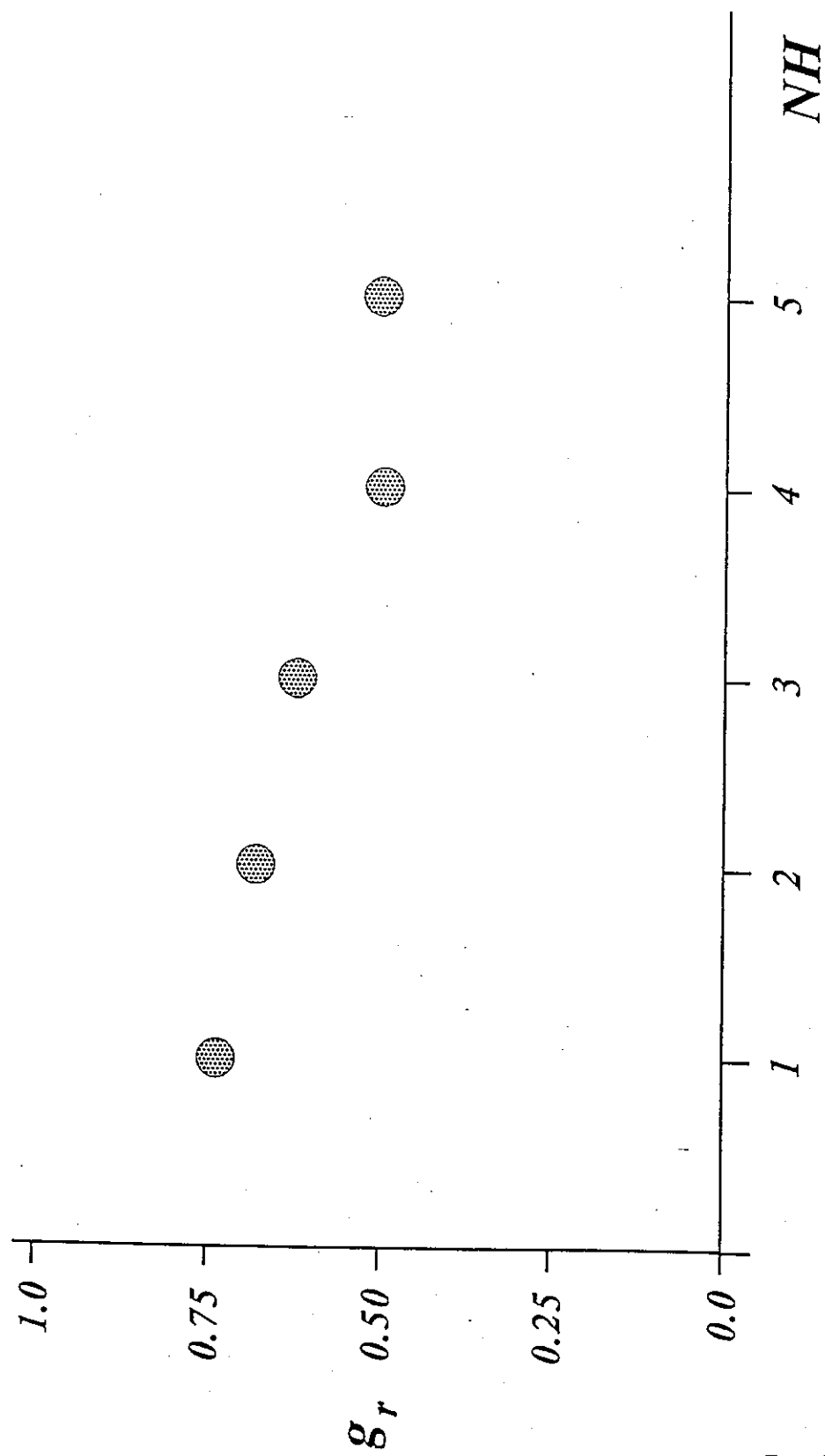
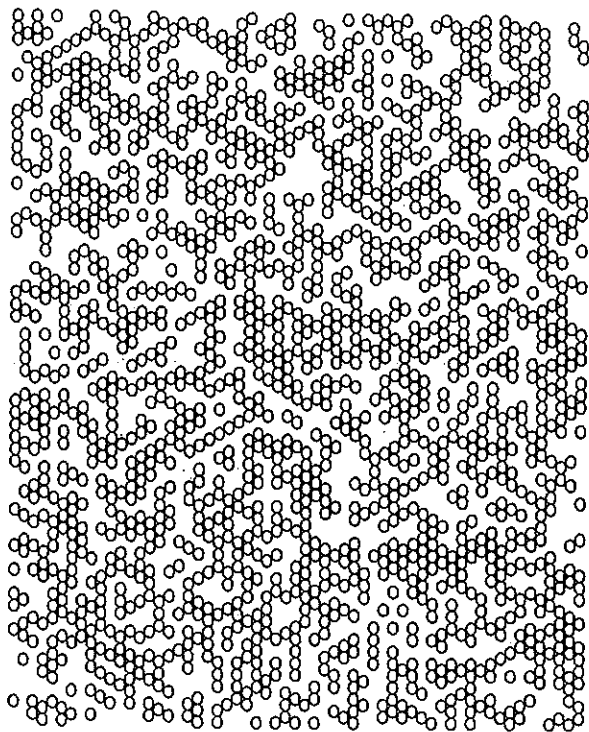
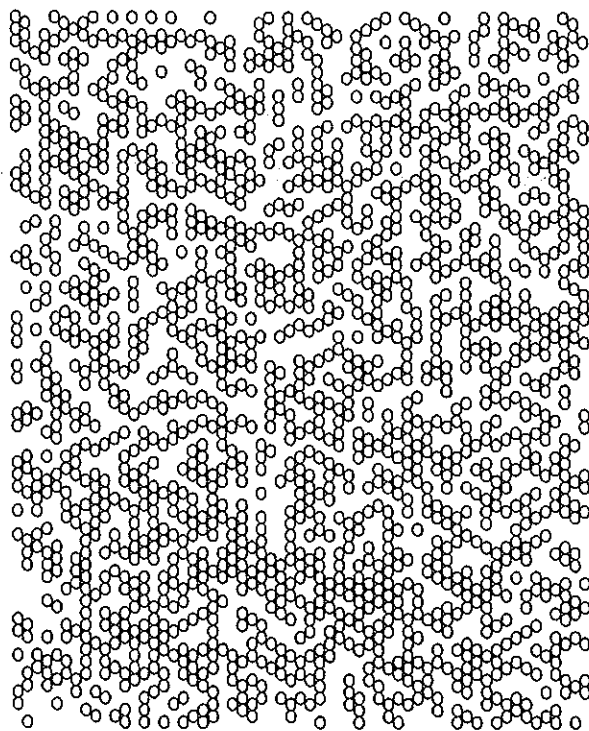


Fig. 3



**A**  $t = 10^5$  steps



**B**  $t = 10^5 + 10$  steps

fig. 4

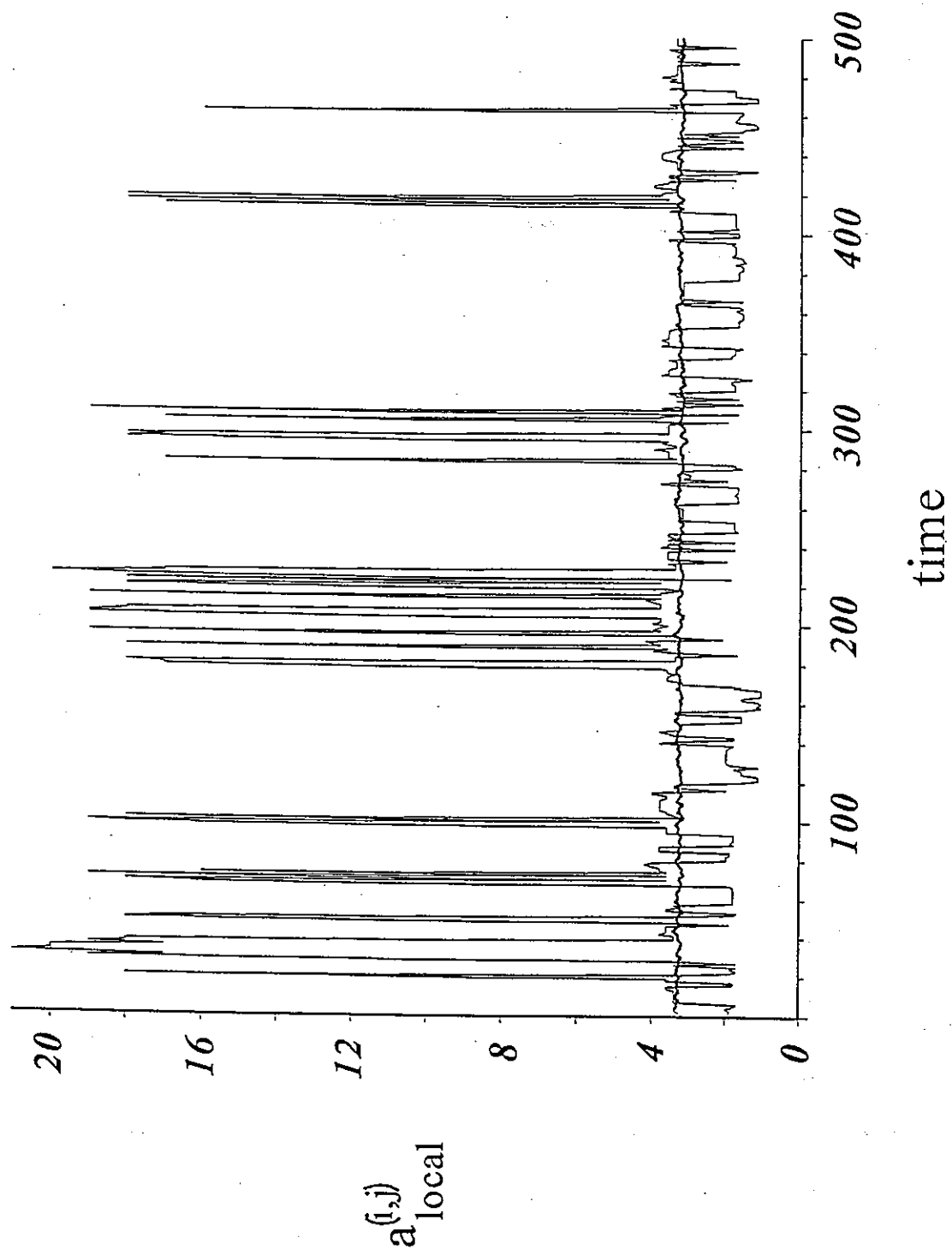
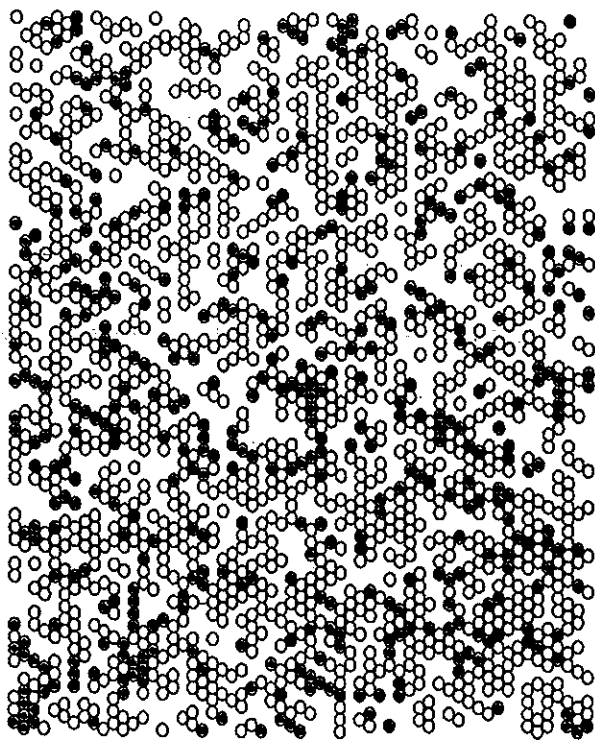
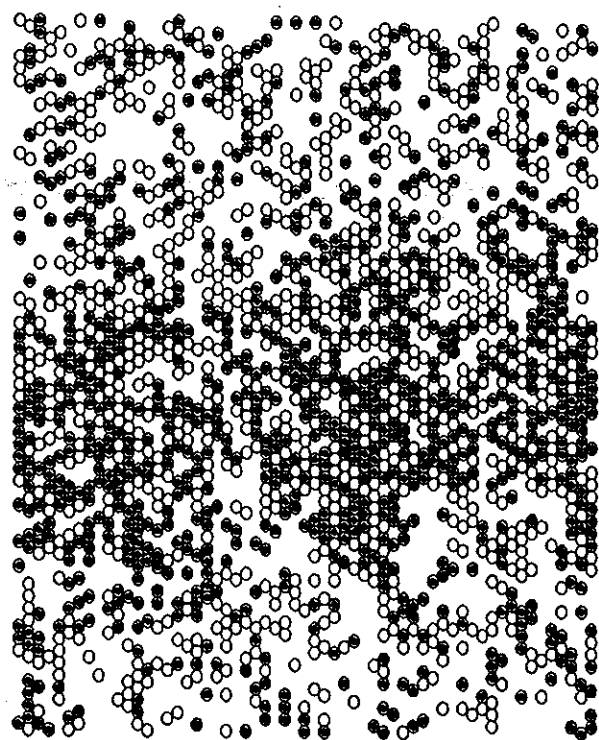


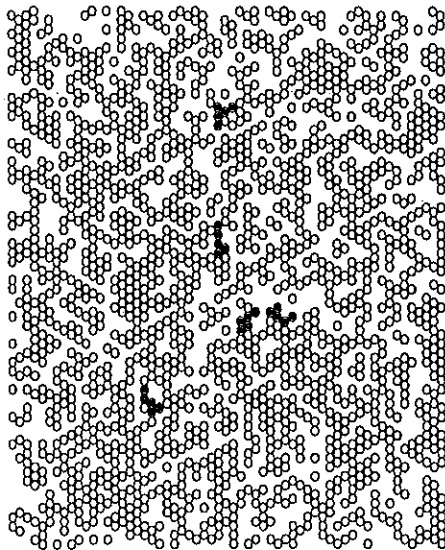
Fig. 5



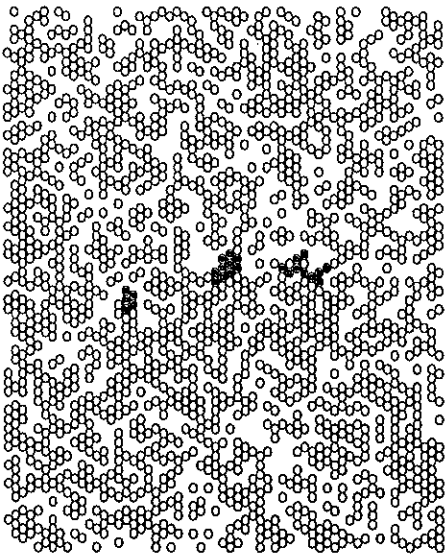
**A**  $t = 10^5$  steps



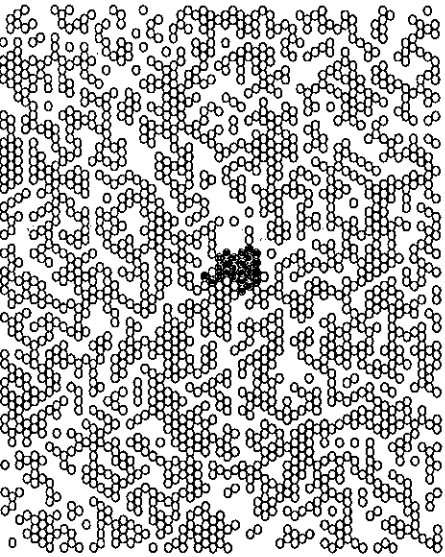
**B**  $t = 10^5$  steps



**A**  $t = 10^3$  steps



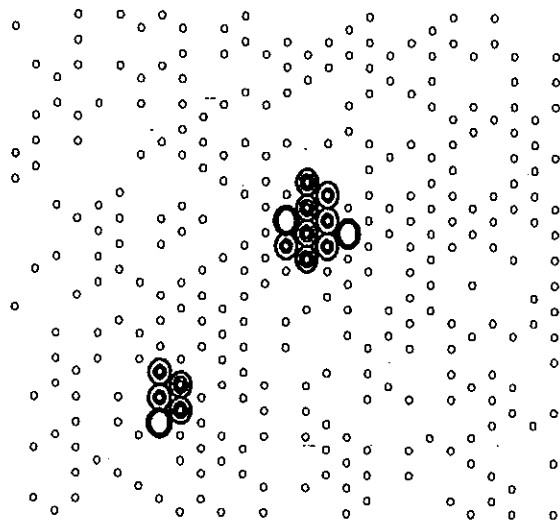
**B**  $t = 5 \cdot 10^4$  steps



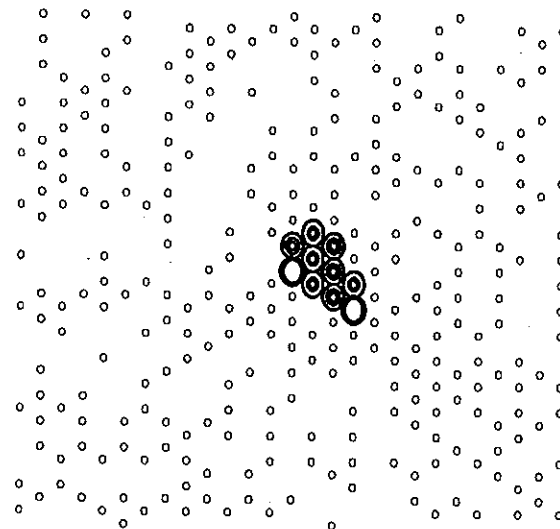
**C**  $t = 10^5$  steps

fig. 7

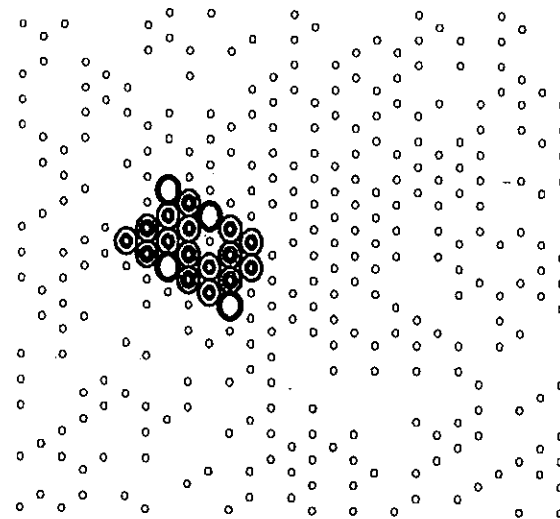




**A**  $t = 10^3$  steps



**B**  $t = 5 \cdot 10^4$  steps



**C**  $t = 10^5$  steps

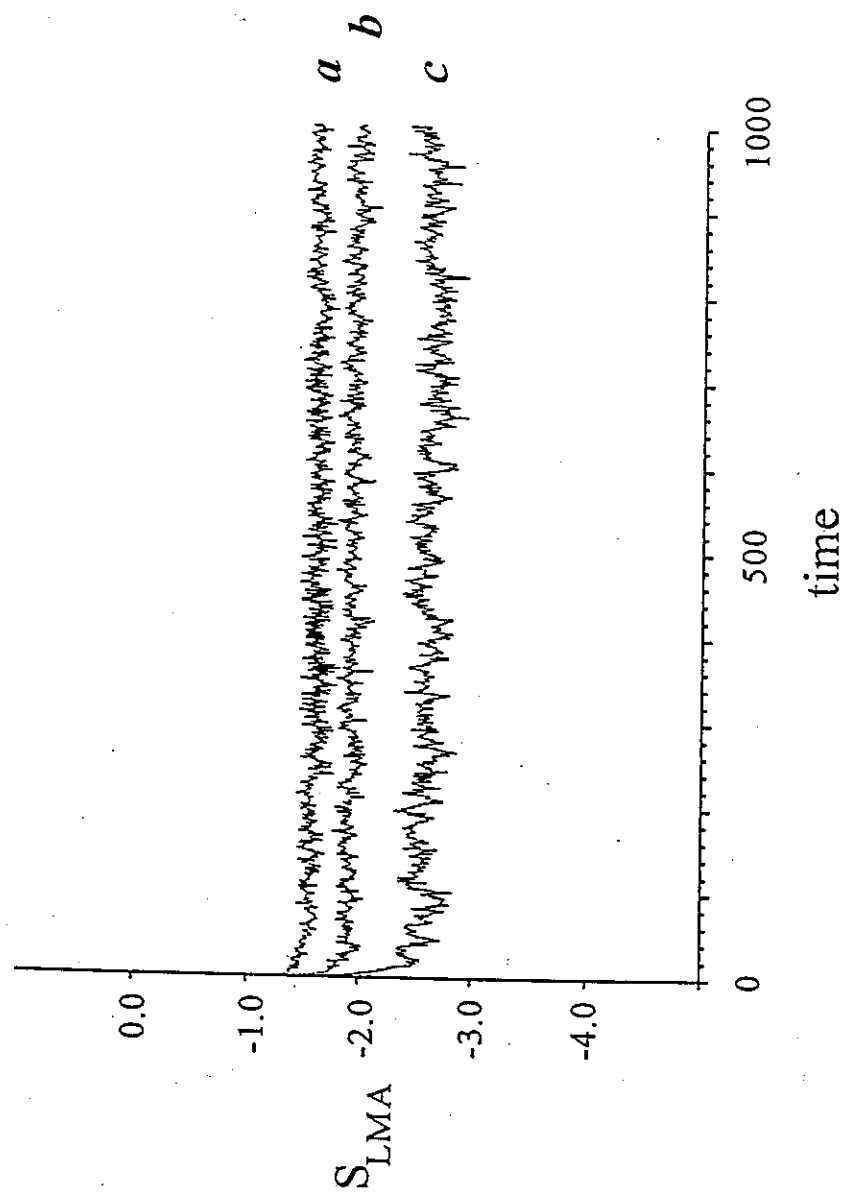


Fig. 9

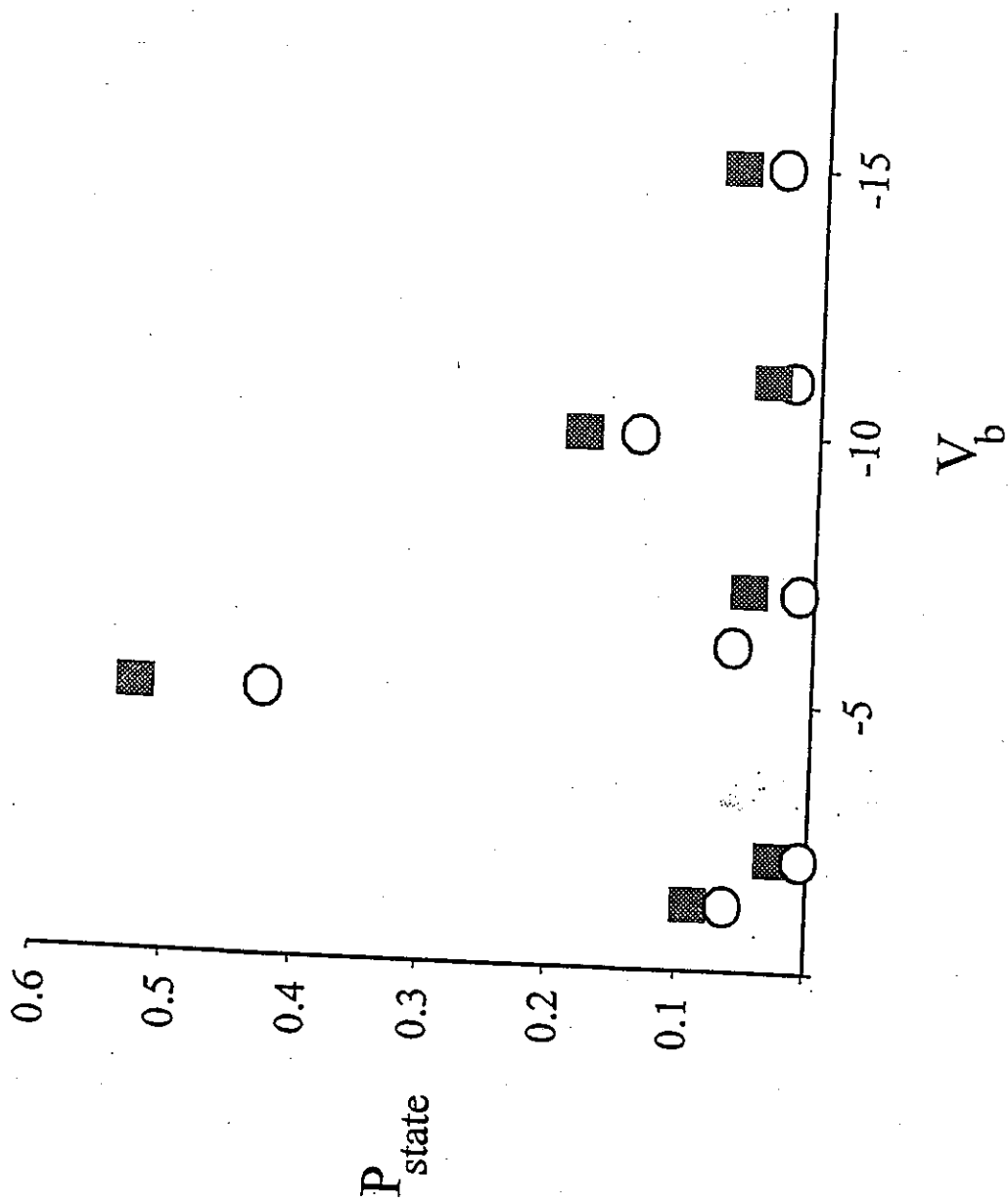


Fig. 10