

# Application of Molecular Dynamics Simulation to Small Systems

Víctor M. Rosas-García and Isabel Sáenz-Tavera  
*Facultad de Ciencias Químicas, Universidad Autónoma de Nuevo León,  
San Nicolás de los Garza, N. L.  
México*

## 1. Introduction

The study of chemical behavior includes answering questions as 'which isomer is the most stable?', 'which relative orientation is the most favorable for such-and-such interaction?', 'which conformer is the global minimum?', 'what are the lowest energy configurations and their relative energies?'. The answer to these questions—and many others—depends on the ability to find and study a variety of configurations of the system of interest. Recently, (Atilgan, 2007) briefly reviewed the use of molecular dynamics simulation for conformational search in the process of drug design, concluding that its use could reduce the errors in estimating binding affinities and finding more viable conformations. In addition, (Corbeil, 2009) considered the need to include ring flexibility in the conformational searches used in flexible docking. Most of the flexible docking algorithms skip searching for conformations in rings, even though a protein may stabilize a conformation other than the most stable one.

The need for a tool to examine the diverse configurations of the constituent particles of a system becomes obvious even as we consider relatively small systems (10-100 atoms). Finding by hand all the conformers of cyclohexane is feasible and maybe even instructive; this is somewhat more complex when doing morpholine and even something as small as an eight-membered heterocycle can be prohibitively complex to analyze by manipulating molecular models by hand. One of the methods available to the researcher to tackle this kind of problem is Molecular Dynamics (MD) simulation. MD allows an exploration of the configurational space of a system, respecting chemical constraints. Chemical constraints—such as atomic connectivities—are needed in cases such as conformations of molecular rings and configurations of molecular clusters, e.g., solvation shells. In both cases, atomic connectivities must be kept intact, otherwise we risk breaking up the ring or the molecular constituents of the cluster. In our research group we routinely employ MD simulation, sometimes in its semiempirical variety, to study small systems (systems with fewer than 100 atoms) whether they be solvation shells, inorganic clusters or heterocycles.

This review is narrowly focused on current software and methods appropriate for doing MD simulations of small heterocycles and clusters composed of 10-100 molecules. In particular, we try to systematize the tools available to tackle the problem of searching for minima in heterocycles and in molecular clusters. We want to use MD simulations as a tool to explore the energy landscape of a small system, so we can locate the global minimum. We do not include the vast literature simulating water solvation by MD. Even though the aggregates of water

molecules around a solute can be considered clusters, we focus our attention on non-water clusters and we only borrow some tools for our purposes.

When MD is based on molecular mechanics force-fields, it cannot model bond breaking-forming processes. For these cases, there are mixed methods such as Quantum Mechanics/Molecular Mechanics (QM/MM), where the bond breaking and formation is taken care of by the quantum mechanical part, while everything else is handled by the molecular mechanics/dynamics part. Another method that considers bond breaking-formation is Car-Parrinello Molecular Dynamics, where the electrons are also considered particles in the dynamics. Such methods are beyond the scope of this chapter.

## 2. Basic concepts of molecular dynamics simulation

Let us imagine a container (an imaginary box) with a finite number of particles. These particles move about the box at varying velocities (both speed and direction can vary), continuously colliding and bouncing off each other. The trajectories of the particles, taken as a set, contain valuable physical information. For example, if the particles interact with each other, making the particles move more slowly (equivalent to lowering the temperature) the interaction may allow the particles to associate and form a liquid. In a condensed phase, collisions happen more often, and the distance a particle can travel before colliding is much shorter than in the gas phase. It follows that, the movements of the particles correspond to the state of the system. It is by analyzing trajectories that we can compute properties. One can also hope that the results of such analysis will yield *physical insight* about the behavior of a system. For a full introduction to MD simulation see the book by Haile(Haile, 1992). It is of the foremost importance to have a correct description of the interactions between the particles.

### 2.1 Is MD ready for general consumption?

For a long time, MD simulation has been the province of specialists, and the literature is still packed with obscure vocabulary and long descriptions of complicated algorithms. All these are necessary when the purpose is the calculation of dynamical macroscopic properties of the system, such as viscosities, surface tensions or rheological properties. Unfortunately, finding such a scenario can put off the non-specialist that does not want to become an expert in the theory of molecular dynamics before attempting some configurational search.

When narrowly restricted to the task of configurational search, we can prescind with many of the details like periodic boundary conditions, equilibration, thermostats and the like. In addition, the field has produced software packages increasingly friendly to the user so, doing a MD simulation becomes just a little harder than using a molecular visualization program.

### 2.2 Description of the interactions

Every MD simulation depends on specifying an interaction model, that is, some analytical function that calculates the energy of interaction between the particles of a system (a potential). The Lennard-Jones (L-J) potential is probably the most popular potential in MD studies, (eq 1)

$$V_{LJ}(r) = 4\epsilon \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \quad (1)$$

where  $\sigma$  is the particle diameter,  $\epsilon$  is the depth of the energy well and  $r$  is the distance between the centers of the particles.  $\sigma$  and  $\epsilon$  are determined empirically, to fit observed properties of the

system of interest, such as boiling point or density, to name a few. The L-J potential illustrates one of the main stumbling blocks faced by the simulator, that is, the need to determine values of empirical parameters, so the simulation has physical meaning.

The L-J potential as such is of little use in chemical systems because it does not consider chemically important bonds, such as covalent or hydrogen bonds. To take into account the *intramolecular* interactions, as well as the *intermolecular* interactions, the potential has to be defined in terms of the bond lengths, bond angles and torsions present in the molecule (its internal coordinates). When the potential energy functions are defined in terms of the molecular internal coordinates, the potential function is called a 'force-field' (Engler, 1973). There are many force-fields available for simulation of chemical systems: AMBER (Case, 2005; 2010; Pearlman, 1995), CHARMM (Brooks, 2009), MM3 (Allinger, 1989), OPLS (Jorgensen, 1988) and TRIPOS (Clark, 1989), among others. In a MD simulation we use the force-field together with the equations of motion to obtain the dynamic behavior of the system (the trajectories of the particles).

### 2.3 Description of the trajectories

The MD algorithm is a way to compute such trajectories and analyze them to extract physical information about a system. How do we describe the particle trajectories? At any given moment, each particle can be described in tridimensional space by 3 numbers, the x-coordinate, the y-coordinate, the z-coordinate. Since we are interested in the trajectories, we need to know the location of a particle and where it is headed at each point in time. So, we also need the momenta along each axis:  $p_x$ ,  $p_y$  and  $p_z$ . The first three numbers (XYZ coordinates) are known as *configuration space*; while the three momenta define the *momentum space*. Taken together they form the *phase space*. In the same way that a set of values for XYZ coordinates defines a point in configuration space, any set of values for these 6 coordinates (three XYZ and three momenta) defines a point in phase space. Inasmuch as both the positions and momenta depend on the time, as time changes the phase space coordinates also change, thus defining a trajectory in phase space. So a MD simulation is about computing the trajectories of particles in phase space.

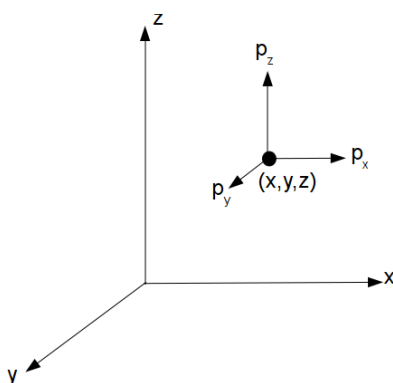


Fig. 1. Phase space coordinates of a particle with position in  $x, y, z$  and momenta  $p_x, p_y, p_z$

## 2.4 Calculation of the trajectories

The forces acting on that particle dictate where a particle is located in space and its direction of movement at each point in time. The classical way to deal with forces uses Newton's laws of motion. The first law states that "every object persists in its state of rest or uniform motion in a straight line unless it is compelled to change that state by forces impressed on it". So, if  $\mathbf{r}$  is a vector that contains the particle coordinates at a given moment, and its first derivative with respect to time (its velocity) is symbolized by  $\dot{\mathbf{r}}$ , then the first law of motion can be mathematically expressed as

$$\dot{\mathbf{r}} = \text{constant} \quad (2)$$

keeping in mind that quantities written in bold are vectors.

The second law of motion states that "force is equal to the change in momentum ( $m\mathbf{v}$ ) per change in time. For a constant mass, force equals mass times acceleration". If  $\mathbf{F}$  is the force,  $m$  is the particle mass, and  $\ddot{\mathbf{r}}$  indicates a second derivative of position with respect to time (or the first derivative of velocity with respect to time), then  $\ddot{\mathbf{r}}$  corresponds to an acceleration. The second law can be mathematically expressed as follows,

$$\mathbf{F} = m\ddot{\mathbf{r}} \quad (3)$$

The third law states that "For every action, there is an equal and opposite re-action". Assuming an isolated system of identical particles, where the total net force,  $\mathbf{F}_{total}$ , is zero,  $\mathbf{F}_{total} = 0$ , then any force exerted by particle 1 on particle 2,  $\mathbf{F}_1$ , is compensated by an equal and opposite force,  $\mathbf{F}_2$ , exerted by particle 2 on particle 1. Using the same notation, this law can be expressed as,

$$\mathbf{F}_1 = -\mathbf{F}_2 \quad (4)$$

To calculate the trajectory in phase space, a MD simulation relies on solving Newton's equations of motion. We just need to use a slightly modified notation. Since we need to keep track of each particle, we use subindices, just as we did for the two particles used in the third law

$$\mathbf{F}_i = m_i\ddot{\mathbf{r}}_i \quad (5)$$

We now turn our attention to an interesting fact: the trajectories depend on time, but the mathematical form of the second law (see equation 3) is time-independent, that is, at any moment the relationship between forces, masses and accelerations is expressed by the same formula. So we expect to find a quantity that remains constant with time for the whole system of particles. In an isolated system, the total energy is constant with time, so this means that the sum of kinetic and potential energies for all the particles in the system is constant. This invariant quantity is also known as the Hamiltonian. The kinetic energy for each particle can be expressed by

$$E_{ki} = \frac{1}{2}m_i\dot{\mathbf{r}}_i^2 \quad (6)$$

while the potential energy is calculated according to the model for description of the interactions, such as the Lennard-Jones potential (see equation 1), although it could be any of the force-fields available in the literature. So the form of the equation for the total energy might be

$$E_{total} = E_{ki} + E_{poti} = \sum_i \frac{1}{2}m_i\dot{\mathbf{r}}_i^2 + \sum_i \sum_j 4\epsilon \left( \frac{\sigma}{r_{ij}^{12}} \right) - \left( \frac{\sigma}{r_{ij}^6} \right) \quad (7)$$

where the new variable  $r_{ij}$  denotes the center-to-center distance between each pair of particles. This formidable-looking equation tells us that the total energy is calculated by adding the contribution of each particle and of each pair of particles to the kinetic and potential energy, respectively.

Once an initial configuration of atoms is specified and values for location and momenta have been assigned to each atom in the system, the system is allowed to evolve as time progresses. This evolution causes a redistribution of energy, and allows the formation of an energy distribution characteristic of the temperature. This step is known as equilibration. The key step in the calculation of an equilibrated distribution is the determination of the time between collisions and the pairs of colliding particles, because the collisions are the ones responsible for the energy redistribution. After the system has achieved equilibration, we can register the trajectories of the particles. This is the simulation step, and it is the only stage when the trajectories have physical meaning. Once the trajectories have been calculated, properties can be estimated, as long as they can be formulated as averages over time. In dealing with small systems, let us say, macrocycles, we do not expect ever to achieve equilibration, because all the atoms that form the system have restrictions on them that preclude an accurate calculation of the parameters that indicate equilibrium.

All the calculations are performed using finite-difference methods, of which Runge-Kutta is probably the best known, although the Runge-Kutta family of methods finds little use in MD simulations because of the large computational demands. One of the most widely used finite-difference methods is Verlet's algorithm, a third-order Störmer algorithm. It is not as stable as a Runge-Kutta, but its computational demands are much lower.

## 2.5 Keeping the system in one piece

Each atom has its own velocity, which could take it in a direction very different from that of the other atoms so, what happens when a particle, atom or molecule, moves far away from the others? The usual way to deal with this problem is to employ *periodic boundary conditions* (PBC). In PBC we formally consider the system as made up by multiple copies of itself along all three X, Y, and Z axes. With this setup, if a particle wanders far enough from the others in one direction as to be located outside the box that contains the particles, another, identical, particle comes into the system from the opposite direction, bearing the same velocity. In general, when dealing with a single molecule—within the molecular mechanics formalism—we do not have to worry about losing atoms, because all the atoms are connected by chemical bonds, and molecular mechanics does not allow for bond breaking. In the case of clusters, it is conceivable that a single group (either a neutral molecule or an ion) might wander off the box limits, but that could give us information about the intensity of the interaction and about the optimum equilibrium geometry.

## 2.6 Simulated annealing

Simulated annealing is a technique able to locate the global minimum of a system of particles. The concept is obtained by analogy with the process of annealing a metal, where the metal is heated to high temperature and then suddenly cooled down by submersion in water. By raising the temperature of the system, it leaves the local minimum where we happened to find it (or build it), and is able to sample the configuration space so it can find another energy minimum when lowering the temperature. Hopefully the new energetic minimum will be

lower in energy than the previous one. In a MD simulation, we can maintain the system at high temperature (even unrealistically high temperatures, like 2000 K) and then the system temperature is reduced. This adds a cooling step to the simulation.

### 3. Software that implements molecular dynamics simulation

AMBER(Case, 2005; 2010; Pearlman, 1995) designates two different things: a force-field and a package for MD simulation. AMBER the package uses AMBER the force-field for its calculations, but it is entirely possible to use AMBER the force-field in a non-AMBER package, such as GROMACS or CHARMM. Beware, though, that using the same force-field in two different packages will not necessarily get identical results. AMBER the package is currently in version 11. Its learning curve is steep. It is possible to simulate small species, with at least one tutorial showing how to do it.

CHARMMBrooks (2009) also shares the situation of AMBER, in that the name designates both a force-field and a computational engine for MD simulation. It is also possible to use CHARMM the force-field in a non-CHARMM engine. Similar caveats apply, although these authors could not find any information on using CHARMM for small molecules.

Gabedit(Allouche, 2011) makes quantum chemistry software accessible to the novice modeller. It presents to the user a rather limited array of options, making for a less confusing experience. On the other hand, this means that to take full advantage of the capabilities of the quantum chemistry software, the user needs to be well versed in the respective manuals. Gabedit can perform MD simulations by itself, using the AMBER99 force-field. It can also setup MD simulations using semiempirical quantum mechanical energy evaluations, and submit them to a variety of computational engines, such as MOPAC2009(Stewart, 2009), ORCA(Radoul, 2010) or FireFly(Granovsky, 2011). When using quantum-mechanical energy evaluations, the user should keep in mind that these methods allow for bond breaking and formation, so it is entirely possible to end up with an isomerized structure.

Ghemical(Hassinen, 2001) can perform MD simulations, and defines a graphical user interface for this. It has a convenient facility to generate water solvation boxes. The user has to make sure that the force-field contains appropriate parameters. New parameters can be added by editing some configuration files. Its graphical user interface makes Ghemical very accessible to the beginning modeller.

GROMACS(Berendsen, 1995; Hess, 2008; Lindahl, 2001; van der Spoel, 2005) is a molecular dynamics software tailored to simulations with hundreds of millions of particles. It is certainly not made with the novice user in mind, and its learning curve is steep. However, it is an extremely fast computational engine. It is not recommended for dynamics of small species because then the advantages of the fast computation are lost, and because it is command-line based. Building the required topology before running the simulation can be daunting to a novice.

AMBER, CHARMM and GROMACS are tailored towards simulation of bio-macromolecules, and the force-fields included reflect this. They contain force-fields highly optimized for aminoacids or nucleotides or carbohydrates.

HyperChem(Hyperchem, 2011) is a commercial software product has a module for MD simulation, in addition to ab initio, density functional and semiempirical capabilities. It can

simulate chemical reactions by molecular dynamics, because it is not limited to molecular mechanics parameters.

MacroModel(Mohamadi, 1990) is a well-developed molecular dynamics package for biomolecules, and it includes a polished interface known as Maestro.

TINKER(Ponder, 2011) is a molecular dynamics package created and maintained by the group of Jay W. Ponder. It employs molecular mechanics and currently lacks a graphical user interface. Its learning curve is somewhat steep.

VASP(Kresse, 1993; 1994; 1996a;b) seems to be very popular in the metal clusters community. It can do MD simulations using density functional theory (considered as part of Ab Initio Molecular Dynamics) and its use does require the skills of an expert computational chemist.

#### 4. Conformational analysis of heterocycles

It is well known that MD is a very inefficient way to search for minima in large rings(Saunders, 1990). However, it should not be underestimated for searches in medium and small rings (rings < 10 atoms). The use of MD simulations for conformational search is in large part favored by convenience, given that many different software programs include it.

Isayev et al.(Isayev, 2007) claim that the pyrimidine ring in nucleic acid bases has a range of effective non-planar conformations under ambient conditions. Saiz et al.(Saiz, 1996) demonstrated that MD simulation with the Tripos force-field was good at reproducing the conformational behavior of a dioxo ring in aqueous solution. Rosas-García et al.(Rosas-Garcia, 2010) studied conformations of fosforinanes using MD simulations in Ghemical, although some parameters had to be determined at the Hartree-Fock level (basis set 6-31G\*). An extensive search by MD found all the conformers for two diastereomers, and a comparison of the global minima allowed to explain why the axial preference of the phenyl ring was linked to the relative configuration of the stereocenters in the molecule.

Sometimes a side chain can modify the conformational behavior of small rings, like the case of Tosco et al. (Tosco, 2008) who used MD simulations in CHARMM to do conformational search on a series of cyclic oxadiazolol, and thiadiazolol isosteres of carboxylate, where most of the conformational freedom came from the side chain attached to the ring. A similar case is that of Bombasaro et al.(Bombasaro, 2008) who used GROMACS in combination with a systematic grid conformational search to study bullacin. Bullacin contains three five-membered rings, with one 11-carbon side chain, and two rings joined by a 12-carbon chain (see Figure 2).

Regarding structures with fused rings, Aleksandrov and Simonson(Aleksandrov, 2006; 2009) reported development of CHARMM22 parameters (employing the TIP3P(Jorgensen, 1983) model for water) for several tetracycline derivatives and a tetracycline/Mg<sup>2+</sup> complex (see the structure in Figure 2). Tetracyclines exist in two conformations, twisted and extended. In this case the interest was not so much the conformational variety but the possibility of several protonation sites, and the uncertainty of the binding sites for metals such as Mg<sup>2+</sup>. They employed TINKER and the MM3 force-field for MD/simulated annealing. On the other hand, Kiliç et al.(Kiliç, 2000) found out by simulations using quantum MD that cubane and their group 14 analogs can convert to eight-membered rings at high temperature.

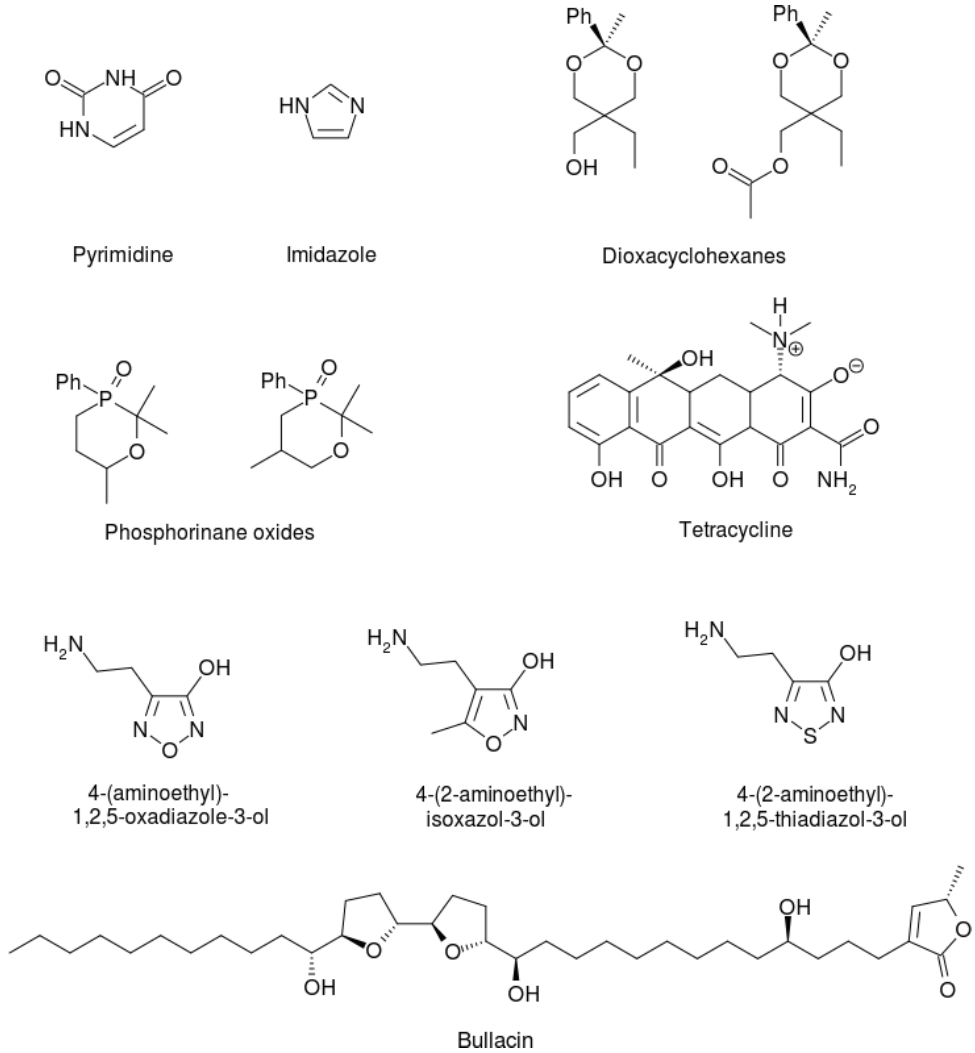


Fig. 2. Some heterocycles and fused cycles studied by molecular dynamics simulations.



## 5. Configurational search of clusters by molecular dynamics simulation

A lot of work on clusters is based on atomic clusters. In this case it may suffice to develop a random-placement algorithm to generate structural variety. Many researchers have taken this route. Chen et al.(Chen, 2011) employed VASP to study clusters of metal carbides. They first generated a variety of  $C_{48}$  clusters and used step-wise addition of carbon atoms and geometry optimization after each step. MD simulations were used to evaluate the thermodynamic stability of several cage structures. The presence of only small distortions in the cage structures at 400 K was taken as evidence of their thermodynamic stability. Fujima and Oda employed VASP to study titanium clusters adsorbed on a single wall nano-capsule. There is no description of the parameters used in the MD calculations (temperature, time step, total simulation time or any others). The only configurational searching took place by putting Ti atoms on different adsorption sites on the carbon wall and doing geometry optimization. Given that they attempted to maximize the contact area between the cluster and the nano-wall, such approach seems justified(Fujima, 2009). Jian-Song and Li(Jiang-Song, 2010) did a configurational search for  $Ga_7As_7$  clusters by randomly choosing points in space from a tridimensional box, cage or sphere, applying distance constraints to keep the atoms at chemically reasonable distances. This is not MD, and they had to generate thousands of structures, although the original paper is sketchy on the details of how many structures were generated. It could well be that the imposed distance constraints biased the resulting structures. This method does have the advantage of not requiring force-field parameters for generating the structures. Jiménez-Sáez(Jiménez-Sáez, 2006) studied the equilibrium structures of copper clusters as a function of the kinetic energy of deposition on a gold surface. For the starting geometries, no configurational search was done. The configurations of the deposited clusters were analyzed in terms of the deformation produced as the kinetic energy of deposition varied. Kuzmin et al.(Kuzmin, 2008) used software developed in-house to study configurations of silver clusters using MD simulations with the embedded atom model on clusters between 13 and 2057 atoms. They used temperatures between 0 and 1300 K. Li et al.(Li, 2007) used full-potential linear-muffin-tin-orbital MD (FP-LMTO-MD) calculations to study the effect of Al impurities on Si clusters. The method is suitable for semiconductor and metal clusters. As for the generation of the initial structures, the authors took the reported ground states of the silicon clusters and added or substituted Al atoms in all possible positions of each cluster. Yang and Xiong(Yang, 2008) used a similar method to generate the initial geometries for  $FeB_n$  clusters. So there appears to be a need for more systematically searching the configurational space of clusters.

In our literature review, we found only one recent example of a configurational search on a cluster using MD simulations, that of Chandrachud et al.(Chandrachud, 2009). These researchers employed VASP to study gold cages using Born-Oppenheimer MD to generate initial configurations. They did MD runs at four different constant temperatures, for 60 ps each, obtaining 600 structures (150 for each constant temperature run). Geometry optimization of these structures yielded 50 distinct isomeric clusters.

Clusters are not limited to groups of separate individual particles like in the case of metal clusters. When we have polyatomic species involved in cluster formation, it becomes important to maintain chemical bonds intact. Shiroishi et al.(Shiroishi, 2005) used Car-Parrinello Molecular Dynamics to study iron oxide clusters. Doll et al.(Doll, 2010) used Ab Initio Molecular Dynamics calculations to study clusters of lithium fluoride. They used a two-part protocol: first generation of candidate structures by means of simulated annealing at

a low level of theory (Hartree-Fock) and, second, optimization of the obtained structures using the Local Density Approximation. Takayanagi (Takayanagi, 2008) studied clusters of solvated glycine using the PM6 Hamiltonian. Their semiempirical MD simulations were performed at 300 K, and the initial geometries were taken from previously reported higher-level results and reoptimized using PM6. They observed dissociation of the proton from a carboxylate group, although could not observe formation of the zwitterion. In our group we have studied calcium carbonate clusters (Rosas-Garcia, 2011), and we have explored the configurational space by means of semiempirical MD simulations, using the PM6 Hamiltonian in MOPAC2009. Pang et al. (Pang, 1994) studied inclusion compounds in cycloalkanes by simulated annealing using HyperChem using the MM+ force-field (a variant of MM2. We cannot recommend the use of MM+ due to the lack of a published description of the modifications) and varying the temperature from 300 to 1000 K in 100 K increments. The dynamics revealed that there was orientational flexibility within the cycle and that the interconversion barriers were as low as 1 kcal/mole.

## 6. How to run a basic simulation

Setting up the run is probably the part of the MD simulation that a novice finds most intimidating. This requires several steps: creation of an initial configuration for the system, choice of physical conditions as defined by an ensemble and a temperature for the run; and some numerical parameters necessary for the integration of the equations: time-step, simulation time, integrator and thermostat.

### 6.1 Creation of the initial geometry

We have to create an initial geometry for the system, that is, to position the atoms and molecules in three dimensions so the calculation can proceed. A simple random placement is not useful, because we must be careful to avoid placing two atoms at the same coordinates, or closer than the sum of their atomic radii. This situations, known as bad contacts, tend to destabilize the numerical algorithms used in the determination of the trajectory (the integration step). Most MD packages, such as GROMACS or AMBER, provide their own utilities for building starting configurations.

For the creation of a starting configuration, we can use programs such as Ghemical, Avogadro (Avogadro, 2011) or Gabedit. Both Avogadro and Ghemical have polished graphic interfaces, and they are very easy to use for building single molecules. Gabedit is still somewhat lacking in this regard, as its interface is harder to use than that of Ghemical or Avogadro. For building clusters, use of a graphical interface quickly becomes tedious and prone to errors. PACKMOL (Martínez, 2009) seems particularly convenient for building any kind of cluster due to its ability to add a given number copies of a molecule, water or any other at the user's choice. For the specific case of water solvation, Ghemical provides a function to build a water box or a water sphere around any compound previously loaded in Ghemical's memory, although the number of water molecules added is less intuitive, because it depends on both the dimensions of the box and of the molecule to be solvated.

Our preferred tools for building systems are packages with graphical user interface, like Ghemical, Avogadro or Gabedit. For running MD simulations we have used, with varying degrees of success Ghemical, MOPAC2009 and Gabedit. PACKMOL is very well suited for the construction of clusters because it lets the user specify the structures of the molecules

of interest and how many of them are to be added. Ghemical is less flexible in this regard, because it has tools only for building solvation shells, and the number of water molecules added is not specified directly by the user, but by the volume specified for the water box. Adding a precise number of water molecules in Ghemical can be a hit-and-miss experience.

## 6.2 Choice of physical conditions

The set of 'physical' information contains: temperature, pressure, relaxation times, compressibilities, whether the simulation will be at constant temperature or constant pressure and—probably most important of all—the force-field used to evaluate the interactions within a molecule and between molecules. Given that the trajectories should have physical meaning, how do we know what kind of experimental conditions are we simulating? This corresponds to the choice of the ensemble. We should be familiar with three ensembles: the microcanonical ensemble, the canonical ensemble and the isothermal-isobaric ensemble.

The microcanonical ensemble maintains constant number of particles ( $N$ ), constant volume ( $V$ ) and constant total energy ( $E$ ), so it is also known as the NVE ensemble. The canonical ensemble keeps a constant number of particles, constant volume and constant temperature ( $T$ ), so it is also called the NVT ensemble.

The isothermal-isobaric ensemble keeps constant number of particles, constant pressure and constant temperature, so it is also known as the NPT ensemble. Physically, the NPT ensemble is the most important in chemistry, because many chemical processes are performed under constant pressure and temperature.

For our particular situation, when we deal with so few molecules that even the concepts of pressure and temperature are not well defined, it suffices to say that these ensembles are different ways to give energy to the system and any one of them can accomplish the task of taking the system out of an energy well and into another one.

In our group we typically choose the program defaults, as we are not interested in the physical meaning of the trajectories, but only in the energetic minima resulting from the dynamic search.

## 6.3 Choice of force-field

Choosing a force-field can be daunting to a novice, because of all the options available. In terms of the specific strengths and weaknesses of each force-field, the reader is referred to the literature. However, the main roadblock in using molecular mechanics force-fields is that, sooner or later, one wants to study a molecule lacking adequate parameters in any force-field. Here, the user of MD software needs to know that some software packages, particularly the most friendly to the user, sometimes allow a dynamics calculation to run substituting default values for the missing parameters. Such calculations have practically no value at all. Ghemical uses the TRIPOS force-field, but the user should be aware of the error messages because usually many parameters are missing, and Ghemical will substitute default values. Gabedit will not run a dynamics unless all the parameters are defined or one decides to use semiempirical methods. The lack of adequate parameters usually requires doing *ab initio* calculations on a model compound, so the parameters can be generated. This route is reasonable when the molecules of interest are large compared to the model molecule, but what are we supposed to do if the molecule and the model compound are the same?

Inasmuch as a MD simulation will evaluate thousands of structures, it may still be worth doing the work of generating molecular mechanics parameters, although this may need a collaboration with a computational chemist. Another option would be to use dynamics not based on molecular mechanics. As previously mentioned, Ab Initio Molecular Dynamics, such as Car-Parrinello, is a specialist technique well beyond the scope of this review but there is a middle ground, in terms of complexity and computational requirements: semiempirical MD simulations. Computational packages exist that are able to use semiempirical methods, such as AM1(Dewar, 1985), PM3(Stewart, 1989) or PM6(Stewart, 2007), for energy evaluation without the need to determine molecular parameters. The user must exercise due care when using semiempirical methods because, at temperatures high enough, the molecules can break apart. This is because semiempirical methods calculate the electronic structure so, even the strongest covalent bonds can break if the temperature is high enough. In our studies of ionic clusters, it was all too easy to destroy the species by giving too much kinetic energy to the system.

#### 6.4 Computational parameters for the run

The set of computational information, includes how long the simulation is supposed to run in picoseconds (the simulation time), the length of simulation time elapsed between energy evaluations in femtoseconds (the time-step), the algorithm for integration, from a variety such as Verlet, Leapfrog-Verlet and Beeman, among others (the integrator) and the thermostat, which is the algorithm that enforces the constancy of temperature.

The duration of the simulation is usually split in three steps: heating, equilibration and cooling of the system. For the purposes of conformational searching, it is advisable to take a simulation length longer than the program default, probably two times or three times the default value, depending on the complexity of the system.

#### 6.5 Choice of software

For the MD-based conformational search, both Gabedit and Ghemical can do it, but Gabedit has a more automated implementation. Gabedit automatically saves a user-defined number of geometries from the trajectory and minimizes their energies, either with molecular mechanics or with a semiempirical method.

### 7. Further efforts

We hope that this brief introduction to molecular dynamics will pique the interest of other researchers in exploring mechanical or semiempirical molecular dynamics as a useful tool for the study of small chemical systems. Recent progress on inclusion of dispersion and H-bonding interactions in semiempirical Hamiltonians, such as the corrections to the PM6 Hamiltonian by Řezáč(Řezáč, 2009) and Korth(Korth, 2010), should prove valuable in this regard. As we have focused on small molecules, the ability to run calculations in parallel over several computers seems unnecessary.

Work remaining to be done includes some benchmarking/calibration of semiempirical MD simulations against higher-level QM/MM simulations or pure ab initio calculations to find out its realm of applicability, and perhaps some further improvement of the semiempirical Hamiltonians to this end. It is also necessary to ascertain the adequacy of including solvation

models, whether continuum or explicit, in these calculations. After all, a global minimum obtained in the gas phase may or may not resemble the global minimum under the influence of aqueous solvation. In addition, the general performance of semiempirical MD simulations to study the behavior of solvated ions is still unknown. We caution the reader against an uncritical application of this technique for conformational, configurational searches or otherwise, e.g., the ability of semiempirical Hamiltonians to model transition structures has been questioned (Schenker, 2011), so modeling the dynamics of chemical reactions by semiempirical means may be dangerous, at best.

## 8. References

- Aleksandrov, A.; Simonson, T. (2006). The tetracycline:  $Mg^{2+}$  complex: A molecular mechanics force field. *Journal of Computational Chemistry* Vol. 27, No. 13, (October 2006), 1517–1533
- Aleksandrov, A.; Simonson, T. (2009). Molecular mechanics models for tetracycline analogs. *Journal of Computational Chemistry*, Vol. 30, No. 2, (January 2009) 243–255
- Allinger, N. L. (1977) Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. *Journal of the American Chemical Society* Vol. 99, No. 25, (December 1977) 8127–8134.
- Allinger, N. L.; Yuh, Y. H.; Lii, J. H. (1989). Molecular mechanics. The MM3 force field for hydrocarbons. 1. *Journal of the American Chemical Society* Vol. 111, No. 23, (November 1989) 8551–8566.
- Allouche, A.-R. (2011). Gabedit-A graphical user interface for computational chemistry softwares. *Journal of Computational Chemistry* Vol. 32, No. 1, (January 2011) 174–182
- Atilgan, C.; Aviyente, V. (2007). Hybrid Usage of Computational Tools in Drug Synthesis. *Current Computer-Aided Drug Design* Vol. 3, No. 2, (June 2007) 149–159
- Avogadro: an open-source molecular builder and visualization tool. Version 1.0.0 URL: <http://avogadro.openmolecules.net/>
- Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. (1995). GROMACS: A message-passing parallel molecular dynamics implementation *Computer Physics Communications* Vol. 91, No. 1-3, (September 1995) 43–56
- Bombasaro, J. A.; Masman, M. F.; Santágata, L. N.; Freile, M. L.; Rodríguez, A. M.; Enriz, R. D. (2008) A Comprehensive Conformational Analysis of Bullacin B, a Potent Inhibitor of Complex I. Molecular Dynamics Simulations and Ab Initio Calculations. *Journal of Physical Chemistry A* Vol. 112, No. 32, (July 2008) 7426–7438.
- Brooks, B. R.; Brooks III, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. (2009). CHARMM: The Biomolecular simulation Program, *Journal of Computational Chemistry*, Vol. 30, No. 10, (July 2009) 1545–1615
- Case, D. A.; Cheatham, T.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, Vol. 26, No. 16, (December 2005) 1668–1688
- Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.;

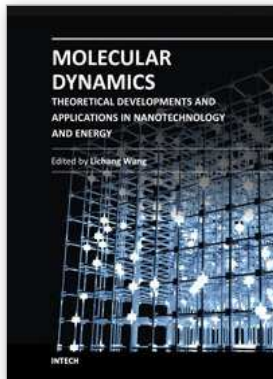
- Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman P. A. (2010). AMBER 11. University of California, San Francisco.
- Chandrachud, P.; Joshi, K.; Krishnamurty, S.; Kanhere, D. G. (2009) Stability of gold cages ( $Au_16$  and  $Au_17$ ) at finite temperature. *Pramana* Vol. 72, No. 5, (August 2009) 845–855
- Chen, G.; Peng, Q.; Kawazoe, Y. (2011). First-principles study on  $Ca_8C_n$  ( $n \leq 12$ ) and  $Ca_mC_{12}$  ( $m \leq 8$ ) metal carbides. *Physics Letters A* Vol. 375, No. 6, (February 2011) 994–999
- Clark, M.; Cramer III, R. D.; van Opdenbosch, N. (1989). Validation of the General Purpose Tripos 5.2 Force Field, *Journal of Computational Chemistry* Vol. 10, No. 8, (December 1989) 982–1012
- Corbeil, C. R.; Therrien, E.; Moitessier, N. (2009). Modeling Reality for Optimal Docking of Small Molecules to Biological Targets. *Current Computer-Aided Drug Design* Vol. 5, No. 4, (December 2009) 241–263
- Dewar, M. J. S.; Zebisch, E. G.; Healy, E. F.; Stewart, J. J. P. (1985). Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *Journal of the American Chemical Society* Vol. 107, No. 13, (June 1985) 3902–3909
- Doll, K.; Schön, J.C; Jansen, M. (2010). Ab initio energy landscape of LiF clusters. *Journal of Chemical Physics* Vol. 133, No. 2, (July 2010) 024107-8
- Engler, E. M.; Andose, J. D.; Schleyer, P. V. R. (2011). Critical evaluation of molecular mechanics. *Journal of the American Chemical Society* Vol. 95, No. 24, (October 2011) 8005–8025
- Fujima, N.; Oda, T. (2009). Bonding properties and structures of titanium clusters on (10,0) single wall carbon nano capsule. *European Physical Journal D* Vol. 52, No. 1-3, (March 2009) 87–90
- Granovsky, A. A.; Firefly version 7.1.G, URL: <http://classic.chem.msu.su/gran/firefly/index.html>
- Haile, J. M. (1992) *Molecular Dynamics Simulation. Elementary Methods*, John Wiley & Sons, Inc. Wiley-Interscience, New York
- Hassinen, T.; Peräkylä, M. (2001). New energy terms for reduced protein models implemented in an off-lattice force field. *Journal of Computational Chemistry*, Vol. 22, No. 12, (September 2001) 1229–1242
- Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. (2008). GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation* Vol. 4, No. 3, (February 2008) 435–447
- HyperChem(TM) Professional 7.51, Hypercube, Inc., 1115 NW 4th Street, Gainesville, Florida 32601, USA
- Isayev, O.; Furmanchuk, A.; Shishkin, O. V.; Gorb, L.; Leszczynski, J. (2007). Are Isolated Nucleic Acid Bases Really Planar? A Car-Parrinello Molecular Dynamics Study. *Journal of Physical Chemistry B* Vol. 111, No. 13, (April 2007) 3476–3480
- Jian-Song, Y.; Li, B.-X. (2010). First-principles study of  $Ga_7As_7$  ionic cluster and influence of multi-charge on its structure. *Chinese Physics B* Vol. 19, No. 9, (September 2010) 097103
- Jimenez-Saez, J.; Perez-Martin, A.; Jimenez-Rodriguez, J. (2006). A molecular dynamics study of atomic rearrangements in Cu clusters softly deposited on an Au(001) surface.



- Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* Vol. 249, No. 1-2, (August 2006) 816–819
- Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics* Vol. 79, No. 2, (July 1983) 926–935
- Jorgensen, W. L.; Tirado-Rives, J. (1988). The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society* Vol. 110, No. 6, (March 1988) 1657–1666.
- Kiliç, Ç.; Yildirim, T.; Mehrez, H.; Ciraci, S. (2000). A First-Principles Study of the Structure and Dynamics of C<sub>8</sub>H<sub>8</sub>, Si<sub>8</sub>H<sub>8</sub>, and Ge<sub>8</sub>H<sub>8</sub> Molecules. *Journal of Physical Chemistry A* Vol. 104, No. 12, (March 2000) 2724–2728
- Korth, M.; Pitoňák, M.; Řezáč, J.; Hobza, P. (2010) *Journal of Chemical Theory and Computation* Vol. 6, No. 1, (January 2010) 344–352
- Kresse, G.; Hafner, J. (1993). Ab initio molecular dynamics for liquid metals. *Physical Reviews B* Vol. 47, No. 1, (January 1993) 558–561
- Kresse, G.; Hafner, J. (1994). Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium. *Physical Reviews B* Vol. 49, No. 20, (May 1994) 14251–14269
- Kresse, G.; Furthmüller, J. (1996). Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science* Vol. 6, No. 1, (July 1996) 15–50
- Kresse, G.; Furthmüller, J. (1996). Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Reviews B* Vol. 54, No. 16, (October 1996) 11169–11186
- Kuzmin, V. I.; Tytik, D. L.; Belashchenko, D. K.; Sirenko, A. N. (2008). Structure of silver clusters with magic numbers of atoms by data of molecular dynamics. *Colloid Journal* Vol. 70, No. 3, (June 2008) 284–296
- Li, B.-X.; Wang, G.-Y.; Ye, M.-Y.; Yang, G.; Yao, C.-H. (2007). Geometric and energetic properties of Al-doped Si<sub>n</sub> (n=2–21) clusters: FP-LMTO-MD calculations. *Journal of Molecular Structure: THEOCHEM* Vol. 820, No. 1-3, (October 2007) 128–140
- Lindahl, E.; Hess, B.; van der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Journal of Molecular Modelling*, Vol. 7, No. 8, (August 2001) 306–317
- Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. (2009). Packmol: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry* Vol. 30, No. 13, (October 2009) 2157–2164
- Mohamadi, F.; Richard, N. G. J.; Guida, W. C.; Liskamp R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. (1990). MacroModel - an Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics, *Journal of Computational Chemistry* Vol. 11, No. 4, (May 1990) 440–467
- Pang, L.; Brisse, F. (1994) Structural and conformational analysis of six-, seven-, and eight-membered cycloalkanes in tris(5-acetyl-3-thienyl)methane inclusion compounds. *Canadian Journal of Chemistry* Vol. 72, No. 11, (November 1994) 2318–2325
- Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham III, T.E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P.A. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular

- dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Commun.* Vol. 91, No. 1-3, (September 1995) 1–41
- Ponder, J. (2011). *TINKER* URL: <http://dasher.wustl.edu/tinker/>
- Radoul, M.; Sundararajan, M.; Potapov, A.; Riplinger, C.; Neese, F.; Goldfarb, D. (2010). Revisiting the nitrosyl complex of myoglobin by high-field pulse EPR spectroscopy and quantum mechanical calculations. *Physical Chemistry Chemical Physics*, Vol. 12, No. 26, (May 2010) 7276–7289
- Řezáč, J.; Fanfrlík, J.; Salahub, D.; Hobza, P. (2009) *Journal of Chemical Theory and Computation* Vol. 5, No. 7, (July 2009) 1749–1760
- Rosas-Garcia, V. M.; Lopez-Cortina, S. T.; Chavarria-Resendez, A. *Journal of Molecular Structure: THEOCHEM* Vol. 958, No. 1-3, (October 2010) 133–136
- Rosas-Garcia, V. M.; Saenz-Tavera, I.; Cantu-Morales, D. accepted manuscript, DOI:10.1007/s10876-011-0420-4
- Saiz, E.; Guzmán, J.; Iglesias, M. T.; Riande, E. (1996). Dynamics and Polarity of Substituted 1,3-Dioxacyclohexanes. *Journal of Physical Chemistry* Vol. 100, No. 47, (January 1996) 18345–18350
- Saunders, M.; Houk, K. N.; Wu, Y. D.; Still, W. C.; Lipton, M.; Chang, G.; Guida, W. C. (1990). Conformations of cycloheptadecane. A comparison of methods for conformational searching. *Journal of the American Chemical Society* Vol. 112, No. 4, (February 1990) 1419–1427
- Schenker, S.; Schneider, C.; Tsogoeva, S. B.; Clark, T. (2011) *Journal of Chemical Theory and Computation* Vol. 7, No. 11, (October 2011) 3586–3595
- Shiroishi, H.; Oda, T.; Hamada, I.; Fujima, N. (2005). Structure and magnetism of anion iron oxide clusters,  $Fe_nO_m$ -( $n=3,4$ ). *Polyhedron* Vol. 24, No. 16-17, (November 2005) 2472–2476
- Stewart, J. J. P. (1989). Optimization of parameters for semiempirical methods I. Method. *Journal of Computational Chemistry* Vol. 10, No. 2, (March 1989) 209–220
- Stewart, J. J. P. (2007). Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular Modeling* Vol. 13, (September 9) 1173–1213
- Stewart, J. J. P. 2008. MOPAC2009. Colorado Springs, CO, USA: Stewart Computational Chemistry. URL: <http://OpenMOPAC.net>
- Takayanagi, T.; Yoshikawa, T.; Kakizaki, A.; Shiga, M.; Tachikawa, M. (2008). Molecular dynamics simulations of small glycine-( $H_2O$ ) $_n$  ( $n = 2-7$ ) clusters on semiempirical PM6 potential energy surfaces. *Journal of Molecular Structure: THEOCHEM* Vol. 869, No. 1-3, (November 2008) 29–36
- Tosco, P.; Lolli, M. L. (2008). Hydroxy-1,2,5-oxadiazolyl moiety as bioisoster of the carboxy function. A computational study on  $\gamma$ -aminobutyric acid (GABA) related compounds. *Journal of Molecular Modeling* Vol. 14, No. 4, February 2008, 279–291.
- Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, Vol. 26, No. 16, (December 2005) 1701–1718
- Yang, Z.; Xiong, S.-J. (2008). Structures and electronic properties of small  $FeB_n$  ( $n = 1-10$ ) clusters. *Journal of Chemical Physics* Vol. 128, No. 18, (May 2008) 184310-8





## **Molecular Dynamics - Theoretical Developments and Applications in Nanotechnology and Energy**

Edited by Prof. Lichang Wang

ISBN 978-953-51-0443-8

Hard cover, 424 pages

**Publisher** InTech

**Published online** 05, April, 2012

**Published in print edition** April, 2012

Molecular Dynamics is a two-volume compendium of the ever-growing applications of molecular dynamics simulations to solve a wider range of scientific and engineering challenges. The contents illustrate the rapid progress on molecular dynamics simulations in many fields of science and technology, such as nanotechnology, energy research, and biology, due to the advances of new dynamics theories and the extraordinary power of today's computers. This first book begins with a general description of underlying theories of molecular dynamics simulations and provides extensive coverage of molecular dynamics simulations in nanotechnology and energy. Coverage of this book includes: Recent advances of molecular dynamics theory Formation and evolution of nanoparticles of up to 106 atoms Diffusion and dissociation of gas and liquid molecules on silicon, metal, or metal organic frameworks Conductivity of ionic species in solid oxides Ion solvation in liquid mixtures Nuclear structures

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Víctor M. Rosas-García and Isabel Sáenz-Tavera (2012). Application of Molecular Dynamics Simulation to Small Systems, Molecular Dynamics - Theoretical Developments and Applications in Nanotechnology and Energy, Prof. Lichang Wang (Ed.), ISBN: 978-953-51-0443-8, InTech, Available from: <http://www.intechopen.com/books/molecular-dynamics-theoretical-developments-and-applications-in-nanotechnology-and-energy/application-of-molecular-dynamics-simulations-to-small-systems>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.