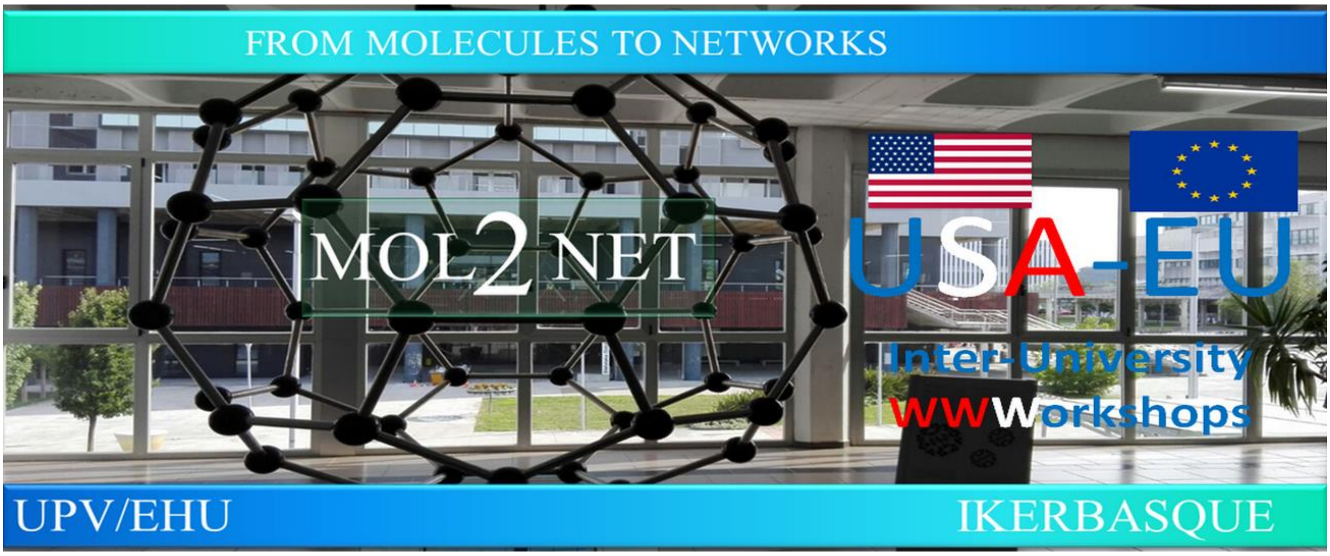
MOL2NET, International Conference Series on Multidisciplinary Sciences



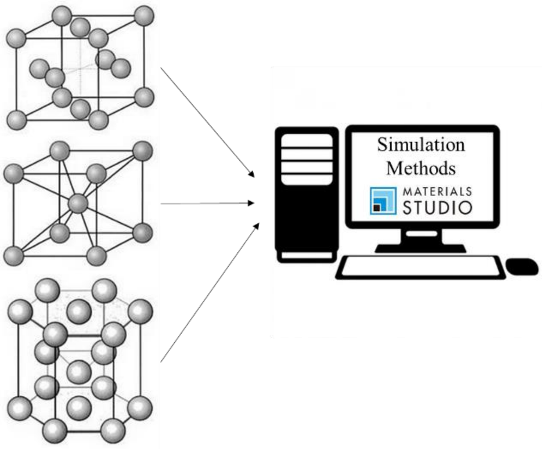
Review: Study of Crystal Structures by Simulation and Modeling

Harbil Bediaga

[a] *Department of Physical Chemistry, University of Basque Country UPV/EHU, 48940, Leioa, Spain.*

[b] *Department of Organic and Inorganic Chemistry, University of Basque Country UPV/EHU, 48940, Leioa, Spain.*

Graphical Abstract



Abstract.

This work is a review of techniques and models for the simulation of crystalline structures. On the one hand, molecular simulation techniques are explained, including quantum techniques such as *ab initio*, semi-empirical and DFT, and on the other hand classical techniques such as minimization, molecular dynamics and Monte Carlo. Applications of molecular simulation techniques are also analyzed: determination of crystalline structures, how to proceed to create a model and Software example.

Introduction

The prediction of crystal structures from the beginning has been one of the great challenges for computational methods in chemistry and materials science. The goal of being able to reliably predict crystal structures at an atomic level in detail with only the chemical composition as input presents several challenges. A solution to this challenge for the prediction of crystal structure requires advances in several areas of computational chemistry. Theoretical chemists have been attracted to these challenges from an academic perspective, while the development of methods to solve the problem of crystal structure prediction has also been motivated for study and application in which a prediction of the tight structure that could guide experimentation.

Crystal structure predictions have been used to study organic molecules such as the polymorphism of pharmaceutical molecules, where changes in crystalline form can lead to changes in important physical and chemical properties, which must be strictly controlled in a pharmaceutical product, or materials. inorganic where the exact determination is necessary the computational design of new materials with specific properties, such as porosity, electronic or mechanical properties. However, the communities that address methods and applications in predicting organic and inorganic crystal structures have remained largely separate, due to the different approaches that have been used in these two areas.

Molecular simulation techniques

To simulate a physical system, in the first place, it is necessary to clearly define the problem to be analyzed, that is, to identify the type of property interesting for the study, the range of parameters that should be considered and the precision. Based on this, the details of the model and the most suitable simulation method for the study can be selected.

1. Quantum techniques (*ab initio*, Semi-empirical and DFT)

Techniques based on quantum mechanics are the most used in molecular modeling. Computational techniques based on quantum mechanics with which they offer more confidence and rigor. Thanks to these methods, properties such as the geometry or energy of chemical systems can be reliably obtained, since the properties that are related to spectroscopic parameters, charge distributions, dipole moment, etc. can be quantified. These advantages are obtained in exchange for very high computational requirements, which is why they are only applicable to systems made up of few atoms.

Quantum mechanics postulates the existence of a wave function or state function, which contains all the possible information of the system. To obtain this function it is necessary to solve the well-known Schrödinger equation (**Equation 1**).

$H\psi = E\psi$ (**1**)

In this **Equation 1** H corresponds to the Hamiltonian operator that includes the potential energy of nuclei and electrons and the kinetic energy, and E to the energy of the system. Although the equation seems simple, the resolution only for simple systems such as hydrogen. In order to be able to solve this equation for more complex and larger systems, several simplifications have been introduced, such as the Born Oppenheimer approximation. This approach argues that the motion of the nuclei is much slower than electrons. Other simplifications that have been used in the representation of the wave function in terms of molecular orbitals (the orbital approximation) and the expression of the orbitals as the linear combination between them (MO-LCAO approximation).

In practice, three methodologies are used to obtain the wave function: *ab initio*, the semi-empirical and the one based on the Density Functional Theory (DFT). Each of these methodologies have a completely different philosophy. In summary, *ab initio* methods rigorously use quantum chemical formalism without employing empirical parameters that go beyond the most elementary physical quantities [1].

It is necessary to take into account the keys of the *ab initio* technique for quantum mechanical calculations and the method that should be used in practice. Quantum mechanics can be used to calculate a wide spectrum of properties. In addition to thermodynamic and structural properties, quantum mechanics can be used to calculate properties dependent on electronic distribution. These properties cannot be quantified by any other method.

The Hartree-Fock (HF) method or self-consistent field is the simplest among the *ab initio* methods. It is based on the variational theorem that seeks to find the wave function that minimizes the energy of the system. This wave function is calculated considering that each electron moves in a field due to the nuclei and an average field created by all the electrons. This method does not include the interaction between each electron and the electronic correlation of the other electrons individually, which causes an overestimation of the electronic repulsion. The most common and simple way to include electronic correlation is the interaction of configurations. For this, the ground state of the molecules is represented with a mixture of all possible states. Even so, many times only a few states are accessible and the expansion must be truncated.

In general, IC provides the most complete possible non-relativistic description of a molecular system. However, this method is prohibitive for medium-sized systems. Another way to deal with electronic correlation are methods based on disturbance theories. Undoubtedly, the most used is the Møller-Plesset (MP), which is based on Rayleigh and Schrödinger's theory.

In *ab initio* methods the total energy of the system is equal to the sum of the electronic energy and the nuclear repulsion energy. A more useful way to quantify the comparison is enthalpy of formation, which is defined as the change in enthalpy when a mole of compound is formed from its elements and ground states. Therefore, the heat of formation can be calculated by subtracting the heats of atomization of the elements and the atomic ionization energies from the total energy. Unfortunately, *ab initio* calculations that do not include electron correlation provide uniformly poor estimates of formation hearts with errors in bond dissociation energies, even at the HF cutoff for diatomic molecules. When combined with an energy minimization algorithm, quantum mechanics can be used to calculate equilibrium geometries of molecules. The results of such calculations can be compared with structures obtained from gas phase experiments using microwave spectroscopy, electron spectroscopy, and electron diffraction. Agreement between theory and experiment has been studied for *ab initio* calculations, generally improving as a ratio is the size of the set of bases [2].

Semi-empirical methods were born with the purpose of providing quantum mechanical techniques for the study of molecular properties, in such a way that being sufficiently precise and reliable to have a practical value in chemical research, they do not present the disadvantages of *ab initio* procedures, that is, that they are applicable to large molecules. For this, a series of simplifications are introduced in the expression of the Hamiltonian HF, which are compensated by the inclusion of a set of parameters.

Most of the time for *ab initio* HF calculation is invariably spent calculating and manipulating integrals. The most obvious way to reduce computational effort is therefore to ignore or approximate some of these integrals. Semi-empirical methods accomplish this in part by explicitly considering only the valence electrons of the system; the electrons of the nucleus are included in the nucleus. The basis for this approach is that the electrons involved in chemical bonds and other phenomena that we might want to investigate are those of the valence shell. By considering all valence electrons, semi-empirical methods differ from those theories (for example, Hückel's theory) that explicitly consider only the π electrons of a conjugated system and are therefore limited to specific classes of molecules. Semi-empirical calculations invariably use sets of bases comprising Slater- type s, p, and sometimes d orbitals. The orthogonality of such orbitals allows the equations to be further simplified.

Hückel's theory can be considered the "ancestor" of molecular orbital approximation methods, as it was formulated in the early 1930s. Hückel's theory is limited to conjugated π systems and was originally devised to explain non-natural nature. additive of certain properties of aromatic compounds. Although Hückel's theory, as originally formulated, is used relatively little in current research, extensions of it, such as extended Hückel's theory, are still used and can provide qualitative insights into the electronic structure of important classes of molecules. Hückel's theory is also widely used for educational purposes to introduce a "real" theory that can be applied to relatively complex systems with little more than pencil and paper or a simple computer program.

Parameterization of semi-empirical approaches typically includes geometric variables, dipole moments, ionization energies, and heats of formation. The performance of successive semi-empirical methods has gradually improved from method to method, although it must always be remembered that abnormal results can be obtained for certain types of systems. Some of these limitations were outlined in the analysis of the various semi-empirical methods. One of the main drawbacks of semi-empirical methods arises when trying to calculate properties that have not been given sufficient importance when making the parameterization process. Also, to achieve optimal performance for specific classes of molecules or specific properties, it would be appropriate to include representative systems during the parameterization procedure.

DFT applied to adsorption is a classical statistical mechanics technique. The calculations at this time can be useful for modeling, but are questionable for analysis with unknown surfaces. The reason for this is that the specific forces, or input parameters, required for a calculation depend on the atoms that are supposed to be present on the surface. For unknown surfaces, a reversion to the use of the Brunaver, Emmett and Teller (BET) equation is often employed.

DFT and, for that matter, Monte Carlo techniques are methods for calculating adsorption modeling given certain assumptions. These assumptions generally include site attractions between the surface atoms and adsorbate molecules and attractions between the latter. Interaction potentials and surface spacings are assumed. The configuration of the adsorbate molecules is adjusted to produce a minimum in the total free energy of the system. In DFT, this adjustment in the settings is done by adjusting the numerical density mainly based on the distance from the surface [3].

It is difficult to find a complete explanation of how DFT works in the literature, so an attempt is made to explain the technique below.

2. Classic techniques (Minimization, Molecular Dynamics and Monte Carlo)

A Molecular Simulation, either by the Molecular Dynamics or Monte Carlo method, has a series of common stages. First, an initial configuration of our model is chosen which, although it may be arbitrary, is not chosen too far from the typical to the simulation conditions (for example, starting from a regular network at the prescribed density, or in the case of Molecular Dynamics, with velocities obtained from a Maxwell distribution). The core of the simulation corresponds to the generation of a series of configurations by means of rules that depend on the type of simulation, and that we can generally divide into two sets: the initial configurations, in which the configurations are not yet typical in the simulation conditions, which is called the equilibrium period, and a larger set of configurations on which the physical properties of interest will be evaluated. These methods are reviewed below.

The minimization problem can be formally stated as follows: given a function f that depends on one or more independent variables x_1, x_2, ..., x_i find the values of those variables where f has a minimum value (**Equation 2**). At a minimum point, the first derivative of the function with respect to each of the variables is zero and the second derivatives are all positive.

$\partial f / \partial xi = 0; \partial^2 f / \partial x_i^2 > 0$ (**2**)

The most interesting function will be the energy of quantum mechanics or molecular mechanics with the variables x_i, being the Cartesian or internal coordinates of the atoms. The minimization of molecular mechanics is almost a function of 3N variables; it is more common to use internal coordinates with quantum mechanics. For analytical functions, the minimum of a function can be found using standard calculation methods. However, this is generally not possible for molecular systems due to the complicated way in which energy varies with coordinates to produce lower energy configurations until the minimum is reached.

We can classify the minimization algorithms into two groups: those that use derivatives of energy with respect to the coordinates and those that do not. Derivatives can be useful because they provide information about the shape of the energy surface and, if used properly, can significantly improve the efficiency with which the minimum is located. There are many factors that must be taken into account when choosing the most appropriate algorithm (or combination of them) for a given problem; the ideal minimization algorithm is the one that provides the answer as quickly as possible, using the least amount of memory.

Energy minimization is widely used in molecular modeling and is an integral part of techniques such as conformational search procedures. Energy minimization is also used to prepare a system for another type of calculation. For example, energy minimization can be used prior to a Monte Carlo or molecular dynamics simulation to eliminate or decrease any unfavorable interactions in the initial system setup. This is especially recommended for simulations of complex systems such as macromolecules or sets of large molecules.

Energy minimization and normal mode analysis have an important role to play in solid state study. Similar algorithms to those discussed above are employed, but an additional feature of such systems, at least when they form a perfect cell, is that the symmetry of the cell's space group can be exploited to speed up calculations. It is also important to take proper account of the interactions with atoms in neighboring cells.

The simplest type of network minimization is done at constant volume, where the dimensions of the basic unit cell do not change. A more advanced type of calculation is performed at constant pressure, in which case there are forces in both the atoms and the unit cell as a whole. Cell vectors are considered additional variables along with atomic coordinates. The laws of elasticity describe the behavior of a material when it is subjected to stress. An obvious source of stress is any external pressure, but stress can also arise from another source, especially interatomic forces within the cell, resulting in "internal stress." The concept of tension is also key in this matter; deformation is the fractional change in dimension.

In the Molecular Dynamics method [4], what is analyzed is the temporal evolution of the considered model. Therefore, the central part of this method constitutes the algorithm of integration of Newton's equations of motion, which are coupled ordinary differential equations, non-linear and of the second order, which must be solved numerically. Given the initial positions and velocities at an initial time, the algorithm obtains the positions and velocities in you.

Molecular dynamics calculates the "real" dynamics of the system, from which the time averages of the properties can be calculated. Sets of atomic positions are derived in sequences by applying Newton's equation of motion. Molecular dynamics is a deterministic method as mentioned above. Using this method, the state of the system at any future time can be predicted from its current state. The first molecular dynamics simulations were performed using very simple potentials such as the hard sphere potential. The behavior of the particles moves in a straight line at a constant speed between collisions. These collisions are perfectly elastic and occur when the separation between a pair of spheres is equal to the sum of the radii. After a collision, the new velocities of the colliding spheres are calculated using the principle of conservation of linear momentum. The hard sphere model has provided many useful results, but it is not the best method for simulating atomic or molecular systems.

The other method of Molecular Simulation is the so-called Monte Carlo method [5]. This is a probabilistic or stochastic method whose objective is to evaluate the averages of the dynamic magnitudes in a certain group. The Monte Carlo technique is a stochastic methodology in which the structural facts and thermodynamic properties are obtained from averages. Thus, the Monte Carlo method tries to obtain a Boltzmann sampling of a system according to an algorithm that consists of generating structures at random, evaluating their energy and deciding from an appropriate criterion whether the new structure generated is accepted or discarded. In Monte Carlo simulation, each setting depends only on its predecessor and not on any other of the previously visited settings. The Monte Carlo method generates configurations randomly and uses a special set of criteria to decide whether or not to accept each new configuration. These criteria ensure that the probability of obtaining a given configuration is equal to its Boltzmann factor. Therefore, low-energy states are more likely to be generated than higher-energy configurations.

In a Monte Carlo simulation, each new configuration of the system can be generated by randomly moving a single atom or molecule. In some cases, new configurations can also be obtained by moving several atoms or molecules or by rotating around one or more bonds. Then the energy of the new configuration is calculated using the potential energy function. If the energy of the new setting is less than the energy of its predecessor, the new setting is accepted. If the energy of the new configuration is greater than the energy of its predecessor, the Boltzmann factor of the energy difference is calculated.

The simplest Monte Carlo formalism is the Metropolis algorithm, in which the generation of new structures takes place randomly, ensuring the microscopic reversibility of the system. However, in very dense chemical systems, such as polymers, the Metropolis method is not very efficient and more sophisticated algorithms must be used.

Applications of molecular simulation techniques: determination of crystalline structures

Computational methods that automatically extract knowledge from data are critical to enabling data-driven materials science. A reliable identification of cell symmetry is a crucial first step in material characterization and analysis. Current methods require a user-specified threshold and cannot detect average symmetries for defective structures.

Crystals play a crucial role in materials science. In particular, knowing the chemical composition and crystal structure (the way in which atoms are arranged in space) is an essential ingredient in predicting the properties of a material. In fact, it is well known that the crystal structure has a direct impact on the properties of materials. Just to give a concrete example: in iron, the solubility of carbon (important for the formation of steel) increases almost forty times when going from body-centered cubic α-Fe (ferrite) (BCC) to γ-Fe (austenite) face-centered cubic (FCC)). From a computational point of view, the identification of crystalline symmetries allows, for example, to construct grids of k-points appropriate for sampling the Brillouin zone, to generate paths between points of high symmetry in band structure calculations or to identify distortions for calculations of finite displacement phonons. Given the importance of atomic arrangement in both theoretical and experimental materials science, an efficient way to classify crystals is to find the group of all transformations under which the system is invariant; In three dimensions, these are described by the concept of spatial groups. Currently, to determine the space group of a given structure, the allowed symmetry operations are first determined and then compared with all possible space groups to obtain the correct label; this is implemented in existing symmetry packages. For idealized crystalline structures, this procedure is exact, but in most practical applications, atoms are displaced from their ideal symmetry positions due to intrinsic (unavoidable) defects or impurities or experimental noise. To address this, it is necessary to establish thresholds to define how flexible you want to be in the classification (that is, up to what deviations from the ideal structures are acceptable); different thresholds can lead to different classifications [6].

Finding the stable crystalline structure for a given chemical composition, mathematically formulated as global (free) energy minimization, is finally, although with certain important caveats, a solved problem [7]. The solution to this problem is a crystal structure (the global minimum). There are at least three major problems in which the solution is not a single structure, but a set of structures or a set of materials. These issues that can be considered resolved are described below:

(I) Prediction of the variable composition structure, looking for all stable compounds (and their crystal structures) formed by given elements. Let's take three simple examples. The Fe-S system has two stable iron sulfides, FeS and FeS_2. The HO system has two well-known compounds, H_2O and H_2O_2, but the first is stable and the second is metastable. In the Na-Cl system, only one compound is known, NaCl. The stability or instability of different compounds can be conveniently determined by the convex (or Maxwell) hull construction. This construction is a convenient representation of the free energies of all possible reactions in the system. All stable compounds (that is, those that have a lower free energy than any isochemical assemblage) form a convex figure on the graph, whose y-axis is the free energy of formation (normalized per atom) and the x-axis is the composition. The advantage of evolutionary algorithms is that the different sampled stoichiometries compete, exchange structural information (greatly speeding up the search) and evolve, producing new stoichiometries. This is much more efficient than sampling all possible compositions independently.

(II) Prediction of optimal materials in two or more properties, that is, multi-objective optimization or Pareto optimization. Without realizing it, we use Pareto optimization in our own decision making every day to optimize price and performance. The set of optimal solutions, known as Pareto front (or first Pareto front), is made up of all non-dominated solutions, that is, those solutions that cannot be surpassed by any other solution in all the objective properties at the same time.

Materials scientists also often want to find materials that meet several criteria: for example, for many applications, materials with the highest possible hardness and fracture resistance are needed. In general, it is advisable to have stability as one of the optimized properties: nobody wants to predict materials that are so unstable that they cannot be synthesized.

(III) Search for the highest performing materials among all the possible compounds of all the elements. It is possible to optimize the desired properties and predict the best performing materials among all possible compounds. Recently, a coevolutionary method was developed to do this, called Mendelevian Search [8]. Here, the key is to organize the chemical space in such a way that the neighboring points are chemically similar (for example, the Na-Cl and K-Cl systems are similar) and have similar properties: then the target property could be globally optimized.

Proceed to create a model

A typical development process incorporates four stages, from setting requirements through product design, development, and testing. In this section, we explore each step in the context of computational materials design [9].

1. Requirements

The fundamental properties required are dictated by the specific application for which the study is to be conducted. In most cases where multiple criteria must be met, a hierarchy of needs must be established. This could be in the form of a merit function, constructed from a combination of weighted descriptors that favor low-level needs, such as thermodynamic stability, over high-level requirements, such as cost and complexity.

2. Design

The greatest challenge in design is based on addressing how to formulate a material, as defined by a chemical composition and a crystalline structure, to meet the stated requirements. The almost infinite number of possibilities must be reduced to a manageable set using the tools described above. Introducing constraints is useful to reduce the physical search space. These could be imposed by limiting the search to a smaller number of items based on cost, availability, and toxicity, as required for the intended application. It is also possible to limit the structural space, for example, to combinations of metal oxide octahedra and tetrahedra as structural building blocks. The efficient search of the available phase space could be facilitated by a combinatorial optimization algorithm, such as the set of branching and linking methods. There is an important distinction between screening and design. The first refers to the search for a solution in a large phase space, while the second involves the use of existing knowledge or the vision of the future. An effective materials design procedure must employ known chemical principles: in magnetism, for example, the connectivity necessary to promote electron exchange interactions is well understood, while in ionic solids the electronegativity of the components is key to determining stability and chemical toughness. It is difficult to avoid the influence of existing archetypes, for example, for the photoelectrochemical division of water, TiO_2 and its derivatives, such as $SrTiO_3$, have been intensively studied, yielding four decades of information on the performance and limitations of the materials. From these studies, it is known that a d_0 cation can be beneficial for reduction processes, due to the long useful life of the photoexcited electrons, which is required due to the slow kinetics of the electron transfer reactions [10]. Even with a highly optimized screening algorithm, a better set of inputs will more efficiently and reliably provide a better set of solutions. The application of design to a restricted physical search space is explored in the section on solar cells.

3. Development

In the development stage, ideas that emerge from the design procedure can be translated into actual materials. Ultimately, there must be a cost-benefit analysis in terms of the calculations. Often several rounds of screening based on the established hierarchy of requirements is a beneficial approach, that is, if a material is calculated to be highly thermodynamically unstable, no further calculations are required to establish its properties. If the calculation of a certain property requires thousands of hours of processing and cannot be implemented at the design stage, then it could be used for secondary selection of candidates emerging from the design procedure. Candidate materials that reach the development stage must be carefully analyzed in terms of chemical bonds and crystal structure. If similar characteristics evolve independently in multiple systems, this may mean a convergence towards a transferable design principle. To facilitate comparison between different studies, it is necessary to have reliable reference points. With the exception of simple properties like cell constants and cohesive energies, there is thus far a lack of consistency and open data in the field of computational materials science. To reproduce a result, the crystal structure and program input must be provided, and both rarely are.

4. Tests

To ensure that candidate materials meet the design requirements, it is essential to test and validate as many features as possible. A feedback loop may be required with the experiment, which, rather than simply iterating the design process, also modifies it to maximize the overlap between theory and measurement. This step could include providing spectral signatures, insights into finite temperature behavior through molecular dynamics simulations, and data on preferred crystal terminations and the effect of morphology on physical properties.

Software: materials studio®

MaterialsStudio® (**Figure 1**) is a comprehensive environment for modeling and simulation designed to allow materials science and chemistry researchers to predict and understand the relationships of a material's atomic and molecular structure with its properties and behavior.

MATERIALS
STUDIO

Figure 1. Materials Studio®

Materials Studio® offers a wide range of methods based on classical interactions between atoms and molecules. These include molecular dynamics, cell dynamics, and various Monte Carlo-based methods, as well as the provision of force fields.

Various ways can be used to obtain a crystal model in Materials Studio®, having detailed information on the crystal structure [11]. Crystal structure information and parameters can be found in references, such as Acta Crystallographica, or in software, such as Inorganic Crystal Structure Database (ICSD). From these websites or software, we can export the crystallographic information files, such as the ".cif", which can be imported directly into Materials Studio® to obtain the atomic model graph. On the other hand, Materials Studio® also contains many common structures. You need several parameters to build a crystal structure, such as:

- Cell parameters (or unit cell constants), such as a, b, c, α, β, γ
- Space group of this crystal structure, or space group number
- Atomic position of unit cell, such as x, y, z, and atom type

Here, we will be able to select diamond, graphite, nanorope and fullerite as an example to display the atomic models. The model building process is illustrated below for diamond:

First, we obtain the structural parameters of the diamond structure, as listed in **Table 1**. Next, we open the Materials Studio® software and look for the "Build" button in the menu. Using the "Build Glass" button to build a relative glass box in Materials Studio®.

Table 1. Structural parameters of diamond

Structure type	Parameters								
Unit cell	3.5668	3.5668	3.5668	90	90	90			
Spacial Group	F d -3 m S								
SG number	227								
Pearson	cF8								
Wyckoff	a								
Position:	*Atom*	*#*	*OX*	*SITE*	*x*	*y*	*z*	*SOF*	
Atom type:	C	1	+0	8a	0	0	0	1.	

Structure information, such as space group number and cell parameters, must be added manually, as shown in Table 1. After construction of the glass box, atoms can be added via position values (x, y, z), respectively. The fully constructed model, as shown in **Figure 2**, exhibits a perfect diamond unit cell.

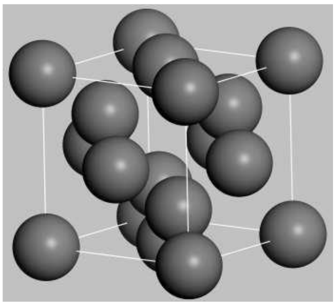


Figure 2. Structure of the diamond.

Similar to the diamond construction procedure, a complicated model can also be constructed in MS, forming the regular crystal structure, as shown in **Figure 3**.

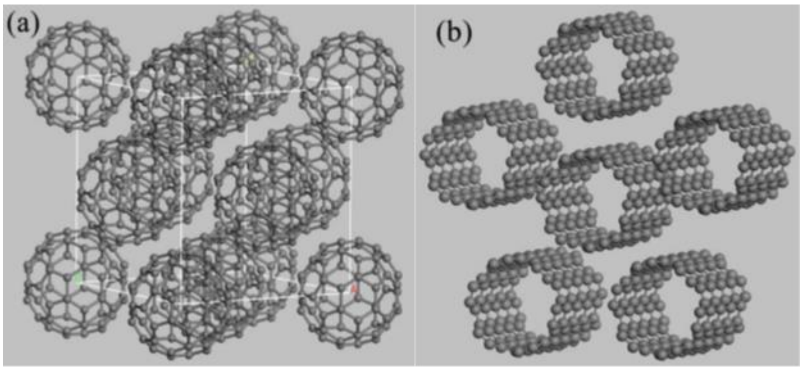


Figure 3. Atomic structure model of a) fullerite and b) nanorope.

This structure model can easily and clearly describe the characteristics of these complicated glass unit cells, arousing the enthusiasm and interest of students. In addition to the common illustration function, MS can also be operated to change the viewing angle and appearance of the atom, as shown in **Figure**

4. For this reason, we can use MS to analyze and describe the different characteristics of different planes. lenses changing angle view. This function is important for describing the orientation and plane of the crystal by changing and rotating the 3D atomic models. **Figure 4(a)** shows the common 3D side view, while **Figure 4(b)** shows the atom packing characteristic of the (0001) plane as the hexagon.

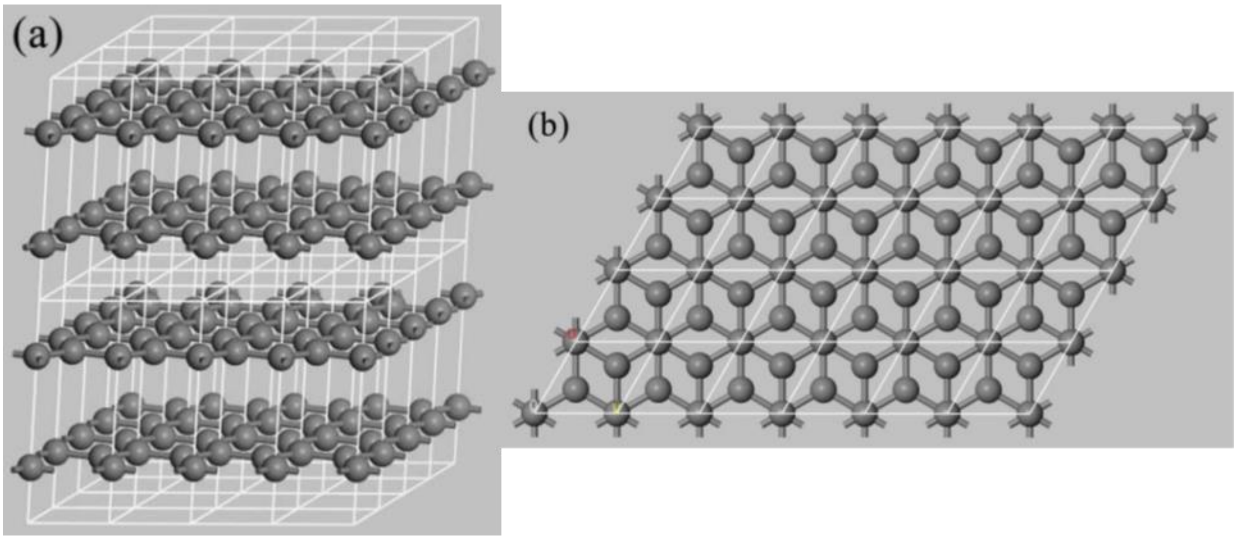


Figure 4. Atomic structure model for graphite, a) side view and b) view from the 0001 plane.

References

[1] Leach, A. R., Molecular modelling : principles and applications. 2nd ed. ed.; Prentice Hall: Harlow, England; 2001.

[2] Hehre, W. J.; Radom, L.; Von Ragué Schleyer, P., Ab initio molecular orbital theory. New York (N.Y.): Wiley: 1986.

[3] Condon, J. B., Chapter 7 - Density Functional Theory. In Surface Area and Porosity Determinations by Physisorption, Condon, J. B., Ed. Elsevier Science: Amsterdam, 2006; pp 207-233.

[4] Jacobs, P. W. M.; Rycerz, Z. A., 4 - Molecular Dynamics Methods. In Computer Modeling in Inorganic Crystallography, Catlow, C. R. A., Ed. Academic Press: London, 1997; pp 83-115.

[5] McGreevy, R. L., 6 - Reverse Monte Carlo Methods for Structural Modelling. In Computer Modeling in Inorganic Crystallography, Catlow, C. R. A., Ed. Academic Press: London, 1997; pp 151-184.

[6] Ziletti, A.; Kumar, D.; Scheffler, M.; Ghiringhelli, L. M., Insightful classification of crystal structures using deep learning. Nature Communications 2018, 9 (1), 2775.

[7] Oganov, A. R., Crystal structure prediction: reflections on present status and challenges. Faraday Discussions 2018, 211 (0), 643-660.

[8] Allahyari, Z.; Oganov, A. R., Coevolutionary search for optimal materials in the space of all possible compounds. npj Computational Materials 2020, 6 (1), 55.

[9] Butler, K. T.; Frost, J. M.; Skelton, J. M.; Svane, K. L.; Walsh, A., Computational materials design of crystalline solids. Chemical Society Reviews 2016, 45 (22), 6138-6146.

[10] Kudo, A.; Miseki, Y., Heterogeneous photocatalyst materials for water splitting. Chemical Society Reviews 2009, 38 (1), 253-278.

[11] Fan, X. H.; Xu, B.; Xu, Y.; Li, J.; Shi, L.; Wang, F. M.; Lin, J. P., Application of Materials Studio Modeling in Crystal Structure. Advanced Materials Research 2013, 706-708, 7-10.