

Computational chemistry tools for atomic level investigation of clay composites

Francesco Ferrante

Dipartimento di Fisica e Chimica “E. Segrè” - Università degli Studi di Palermo,
Viale delle Scienze Ed. 17, I-90128 Palermo, Italy

e-mail: francesco.ferrante@unipa.it

Abstract

The most common computational methods used for the investigation of molecular and periodic systems will be briefly described, with particular emphasis on those approaches that could be employed for the study of clay structures at the atomistic level. The first part of the chapter is mainly dedicated to the conceptual basis of density functional theory and its implementation for molecular and periodic systems. The tight binding approximation to density functional theory and its modern variants, particularly suitable for atomistic studies of large systems, are treated as well. Classical molecular mechanics and molecular dynamics methods, as well as the definition of force fields suitable for clay materials, are discussed. In the second part, case studies of application of computational approaches for the characterization of structures and properties of clay materials (in particular, the halloysite nanotube) are reported.

keywords: Density functional theory, Exchange-correlation functionals, Geometry optimization, Density functional tight binding, Molecular dynamics, Force fields, Halloysite nanotubes

Introduction

The extraordinary evolution of computational methods [1-2], in particular those based on Density Functional Theory (DFT), combined with the power and technology of modern computers, as well as with the implementation of efficient algorithms into optimized codes, allows today the atomistic level investigation of systems having dimensions that were unimaginable even twenty years ago [3-4]. Despite all this, those who want to study complex systems with a certain degree of confidence still must face are still many challenges, the first of them being to gain the experience needed to discriminate reliable computational results from useless numbers. Clearly this experience cannot be acquired by merely reading this chapter but I hope it will still be a pleasant reading that can stimulate the interested scientists to deepen the many aspects and formalisms that will be just mentioned, so that, if he/she wants, he/she can use them in the context of his own research or to better understand the research of others in this constantly evolving field. In what follows, after a brief introduction to the definition and calculation of molecular properties, to wavefunction-based computational methods and to the concepts behind density functional theory, some emphasis will be given to the computational approaches that can be used for the investigation of large systems, which could be of particular interest for the material scientists.

The energy and the Born-Oppenheimer surface

In the Born-Oppenheimer approximation the total energy of a molecular system formed by N electrons and M nuclei is given by the energy of its electronic state plus the Coulomb repulsion energy between its nuclei

$$E(\{\mathbf{R}\}) = E_{\text{el}}(N, \{\mathbf{R}, \mathbf{Z}\}) + \sum_{A=1}^M \sum_{B \neq A}^M \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \quad (1)$$

where Z_I is the atomic number of nucleus I and atomic units are used. The value of the molecular energy depends on the position in space of all its nuclei, $\{\mathbf{R}\}$, i.e. what we commonly call the

molecular geometry; equation (1) defines thus a hypersurface, known as Born-Oppenheimer (BO) potential energy surface, which forms not only the basis for the calculation of (not too exotic) molecular properties, but also the scenario from where extremely important chemical concepts originate, including just the concept of molecular geometry.

Years of research have shaped the abstract chemical intuition: chemists can get an idea of the geometry even for the weirdest compounds. Unfortunately, an idea is not enough in order to calculate the vast majority of chemical and chemico-physical properties of molecular species; for this the researcher needs the exact values of distances, plane and dihedral angles between the atomic nuclei in the molecule, where the “exact” term was used with the meaning “accurate enough, such that a larger accuracy does not affect the conclusion”. As a matter of fact, the first thing to do before tackling any calculation is to find the possible molecular geometries. In practice, by starting from some educated guess of the geometry, which correspond to a point in the BO surface, one can find the stationary point in the surface which is closer to that guess. If this point is a local or global minimum, a molecular conformation was achieved; a stationary point of the first order (a maximum along one direction, a minimum along all the others) is interpreted as a transition state (TS), while still no conventional meaning has been given to higher order stationary points. The search for minima and TSs is conceptually simple: they are points of the surface where no resulting forces act on the nuclei, so one can merely find those points where the nuclear gradients of the energy (1) is below a certain threshold close to zero. The algorithm, called geometry optimization, proceeds according to the flowchart reported in Figure 1: the energy is calculated in the point $\{\mathbf{R}\}$ corresponding to the geometry starting guess, followed by the calculation of the nuclear gradient on the same point; if the gradient along all direction of minimization is smaller than a predefined value, the geometry is optimized; otherwise, the geometry is relaxed following the information on the forces acting on the nuclei obtained from the gradients and a new iteration starts. Once an optimized molecular geometry has been achieved, one must be sure that the nature of the stationary point is the one searched for. This information is readily given by the nuclear Hessian, i.e. the matrix of the second derivatives of the energy with respect to the nuclear coordinates: if the Hessian is definite positive, the geometry is a minimum on the BO surface; if the Hessian has only one negative eigenvalue, the stationary point is a transition state. Once calculated the Hessian, the harmonic vibrational normal modes frequencies are available, which have a threefold role: i) must be used to add the zero-point vibrational energy (E_{ZPV}) to the molecular energy (1); ii) can be used to simulate the IR spectrum of the investigated molecule (IR intensities are easily obtained) and iii) can be used to calculate the vibrational partition function (Q_V), which together with the translational and rotational ones (Q_T , Q_R), allows to evaluate the absolute values of the thermodynamic quantities of formation, referred to the ideal state where all the nuclei and all the electrons forming the molecule are at infinite distance one from the other. The value of E_{ZPV} in the case of TSs is used to estimate the activation energy barrier of the elementary steps of chemical reactions, that could be transformed to rate constants by means of the Eyring model (but beware of large error propagation). Those listed here and reported in Figure 1 are only routine molecular properties available from calculations; many others can be obtained as derivative of the molecular energy with respect to perturbing physical entity.

Needless to say that, in order to use equation (1) for the calculation of molecular properties, one must have some way to evaluate the energy of, at least, the ground electronic state of the system: this is just the problem that quantum chemistry aims to solve.

A (very) brief survey of wavefunction-based methods

The elementary unit of modern computational methods based on the wavefunction is the Slater determinant (SD); it is the determinant of a matrix whose elements are mono-electronic functions called spinorbitals, each spinorbital being the product of a molecular orbital and a spin function. For an N-electron system, a SD is formed by N occupied spinorbitals, with the electron label representing the row indexes and the spinorbital label representing the column indexes of a matrix.

In the Full Configuration Interaction (FCI) method the wavefunction is written as a linear combination of all the unique, same-spin, same-symmetry $N \times N$ SD which can be obtained by distributing N electrons in K spinorbitals. If K tends to infinity the FCI wavefunction tends to the exact solution of the Schrodinger equation, and the associated energies would represent the exact non-relativistic energy values of the electronic states for the investigated system. Due to the extremely unfavourable scaling of the FCI computational cost with respect to N and K , up to date this method can be applied to nothing but the smallest molecules (diatomics of the first period elements) and it is commonly used as benchmark for approximated approaches. There are essentially two ways to obtain approximations to the FCI wavefunction. In the first series of methods, the number of electrons and of spinorbitals to which the FCI is applied is reduced by the appropriate definition of an active space (CAS-based methods); in the second, which includes historical methods like CISD, the number of determinant is reduced instead by truncation of the FCI expansion. One of the approaches with the largest accuracy/cost ratio, the CCSD(T) method, is within this second series; it is based on the so-called coupled cluster ansatz and is dubbed as the golden rule of computational chemistry. In this second series we can find also the most drastic approximation, the Hartree-Fock (HF) method, according to which the wavefunction is expressed as a single SD. The method due to Douglas Hartree and Vladimir Fock is essentially the “father” of all *ab initio* approaches and, even if it is described here as an approximation to FCI, it is actually the method which gives the reference SD upon which all other approximations to FCI are constructed. Due to its monodeterminantal nature, which essentially means that the wavefunction is a antisymmetrized product of mono-electronic functions, the HF method suffers of the mean field approximation, where electrons motion is not correlated (the probability of finding an electron in a certain point of the space does not depend on the probabilities associated to the other electrons); the difference between the exact non-relativistic energy and the HF one, always negative, is called correlation energy. HF method can be applied to very large systems (hundreds of atoms) but, because of the lacking of electron correlation, the HF results have an accuracy that is out of the modern standards. So a problem, common in computational science, appears: accurate method can be applied only to small systems, while large systems can only be treated with low accuracy. Density functional theory (DFT) solves this problem to some extent, allowing to obtain satisfactory results even on moderately large systems.

Density functional theory

Following the DFT formalism, the system energy is obtained as a sum of contributions determined from the integral over all space of some function of the ground state electron density, $\rho(\mathbf{r})$, which, according to the first Hohenberg-Kohn theorem, contains all the information about the electronic states of the system at hand. The energy is exactly and rigourously defined as the sum of a contribution due to the interaction of the N -electron system with the nuclear charges, $V[\rho(\mathbf{r})]$, plus another contribution called Hohenberg-Kohn functional, $F^{\text{HK}}[\rho(\mathbf{r})]$. The form of F^{HK} is unknown and the strategy for DFT implementation is to extract from it all contributions that can be expressed as an integral involving $\rho(\mathbf{r})$, i.e. the Coulomb repulsion energy between two electron charge densities, $J[\rho(\mathbf{r})]$, and the kinetic energy of an idealized reference system of N not-interacting electrons which has, in every point of the space, the very same electron density of the real system, $T_s[\rho(\mathbf{r})]$. All other unknown contributions, namely the Coulomb and exchange correlation energies and the correlation part of the kinetic energy, are collected in a new functional of the density, called the exchange-correlation functional, $E^{\text{XC}}[\rho(\mathbf{r})]$. The final expression for DFT energy is

$$\begin{aligned}
 E^{\text{DFT}}[\rho(\mathbf{r})] &= V[\rho(\mathbf{r})] + T_s[\rho(\mathbf{r})] + J[\rho(\mathbf{r})] + E^{\text{XC}}[\rho(\mathbf{r})] \\
 &= - \sum_{A=1}^M Z_A \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_A|} d\mathbf{r} + T_s[\rho(\mathbf{r})] + \frac{1}{2} \int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 + E^{\text{XC}}[\rho(\mathbf{r})] \quad (2)
 \end{aligned}$$

In the implementation of DFT due to Walter Kohn and Lu Jeu Sham, the electron density of a N -electron system is written as a sum of one-electron charge probabilities

$$\rho(\mathbf{r}) = \sum_{i=1}^N |\theta_i^{\text{KS}}(\mathbf{r})|^2 \quad (3)$$

where the $\theta^{\text{KS}}(\mathbf{r})$ functions, called Kohn-Sham orbitals, are the DFT counterpart of the molecular orbitals of the Hartree-Fock theory. With this choice

$$T_s[\rho(\mathbf{r})] = -\frac{1}{2} \sum_{i=1}^N \int \theta_i(\mathbf{r}) \nabla^2 \theta_i(\mathbf{r}) d\mathbf{r} \quad (4)$$

In order to find the best Kohn-Sham orbitals, the variational theorem is used, according to which the best $\theta^{\text{KS}}(\mathbf{r})$ functions are those giving an electron density to which corresponds the minimum energy for the investigated system in its electronic ground state. Then, the problem is solved by applying a minimization of the energy (2) with respect to the θ^{KS} functions, with the constraints that these functions remain always orthogonal to each other. This constrained functional minimization procedure lead to the following pseudo-eigenvalue Kohn-Sham equations

$$\left[-\frac{1}{2} \nabla^2 - \sum_{A=1}^M \frac{Z_A}{|\mathbf{r}-\mathbf{R}_A|} + \int \frac{\rho(\mathbf{r}_2)}{|\mathbf{r}-\mathbf{r}_2|} d\mathbf{r}_2 + \frac{\delta E^{\text{XC}}}{\delta \rho(\mathbf{r})} \right] \theta_i^{\text{KS}}(\mathbf{r}) = \varepsilon_{ki} \theta_i^{\text{KS}}(\mathbf{r}) \quad (5)$$

which, since the operator within square brackets (called Kohn-Sham operator, f^{KS}) depends on the electron density, must be solved iteratively. Therefore, a form for the Kohn-Sham orbitals is guessed and the electron density is calculated from equation (3); knowing $\rho(\mathbf{r})$, f^{KS} is defined and Kohn-Sham equations can be solved to give updated forms for θ^{KS} . The iterations proceed until the variations in all the Kohn-Sham orbitals from a cycle to the next are under a given threshold. In practice, this procedure involving the solution of a integro-differential equation is substituted by an algebraic one simply by expanding the Kohn-Sham orbitals in terms of linear combinations of known functions, those collectively forming the basis set,

$$\theta_i^{\text{KS}}(\mathbf{r}) = \sum_{\mu=1}^K c_{\mu i} \varphi_{\mu}(\mathbf{r}) \quad (6)$$

The basis set $\{\varphi\}$ is the second, fundamental ingredient which characterizes a quantum chemistry calculation. It is outside the scope of the present chapter to discuss deeply of basis sets, nor in treating the concepts defining their quality neither to give hints about the system characteristics one must take into account for a proper choice between one set or another. It suffices to say that, in the vast majority of molecular calculations, the basis set is formed by a certain number of contracted Gaussian-type orbitals (CGTO), consisting of linear combinations of normalized atom-centered real solid harmonic Gaussian-type functions (called primitive), whose angular part defines the ‘‘angular momentum’’ of the CGTO and is taken from the corresponding atomic one-electron wavefunction

$$\varphi(\mathbf{r}; \mathbf{R}_A, l, m) = \sum_{i=1}^{L_c} d_{\mu i} g_i(\mathbf{r}; \mathbf{R}_A, l, m) \quad g_i(\mathbf{r}; \mathbf{R}_A, l, m) = N_i |\mathbf{r}-\mathbf{R}_A|^l e^{-\alpha_i |\mathbf{r}-\mathbf{R}_A|} \tilde{Y}_{l,m} \quad (7)$$

Here N_i is the normalization factor, \mathbf{R}_A is the position vector of the nucleus A and \tilde{Y} is the real or imaginary part (depending on m) of a spherical harmonic. The contraction coefficients $d_{\mu i}$ and the best ‘‘contraction scheme’’ (i.e. the number L_c of primitives for each function in the basis set) are commonly determined by means of dedicated atomic calculations, from where also the optimal Gaussian exponents, α_i , defining the radial part of each primitive, are obtained. Basis set are catalogued depending on the number and the role of the functions they contain. In strictly

variational methods a larger number of functions in the basis set assures a more accurate value of the molecular energy (an infinite number is needed to reach the “limit” of a given method), but this assertion is not rigorously true in DFT. Large basis sets are essentially employed only for highly accurate post-Hartree-Fock calculations, while in common usage DFT approaches are joined to the so-called double zeta plus polarization (DZP) basis sets. Even if smaller basis sets exist, DZP ones represent the minimum level accepted in today standards; if we use the concept of atomic orbital as can be found in standard chemistry textbooks, the DZP basis set are formed by one function for every core orbital and two functions for every valence orbital, plus one or more sets of additional functions having high angular moment, which are needed for the correct description of the electron cloud deformation (polarization) when an atom is within a molecular context. The most common DZP basis sets are those included in the families developed by Pople, Dunning, Ahlrichs, but many basis sets were proposed also for specific purposes, like the calculation of electric or magnetic properties, NMR shielding tensors, and so on. In material science it is common to deal with heavy elements, like those from the fourth period onwards: in this case, based on the rationale that chemistry properties are mostly determined by valence electrons, the common choice is to substitute the inner electron density with an effective core potential (ECP) and to use an explicit basis set only for the outermost electrons; the ECP will screen the nuclear charge and their effect is included in the nuclear attraction one-electron integrals (see expressions (11) below).

Returning to DFT, the linear combination coefficients $c_{\mu i}$ of the expansion (6) now become the variational parameters and the Kohn-Sham equations (5) reduce to the corresponding matrix equation form

$$\mathbf{F}^{\text{KS}} \mathbf{C} = \mathbf{S} \mathbf{C} \boldsymbol{\varepsilon} \quad (8)$$

where \mathbf{F}^{KS} is the representative matrix of the Kohn-Sham operator in the $\{\varphi\}$ basis

$$F_{\mu\nu}^{\text{KS}} = \int \varphi_{\mu}(\mathbf{r}) f^{\text{KS}} \varphi_{\nu}(\mathbf{r}) d\mathbf{r} \quad (9)$$

\mathbf{S} is the overlap matrix

$$S_{\mu\nu} = \int \varphi_{\mu}(\mathbf{r}) \varphi_{\nu}(\mathbf{r}) d\mathbf{r} \quad (10)$$

\mathbf{C} is the matrix whose i -th column contains the linear combination coefficients and $\boldsymbol{\varepsilon}$ is the diagonal matrix of the Lagrange multipliers ε^{KS} . Equation (8) can be solved just like an eigenvalue equation (but iteratively) by using well defined and very efficient algorithms, involving the calculation of one-electron integrals

$$-\frac{1}{2} \int \varphi_{\mu}(\mathbf{r}) \nabla^2 \varphi_{\nu}(\mathbf{r}) d\mathbf{r} \quad \int \varphi_{\mu}(\mathbf{r}) \frac{Z_A}{|\mathbf{r}-\mathbf{R}_A|} \varphi_{\nu}(\mathbf{r}) d\mathbf{r} \quad \int \varphi_{\mu}(\mathbf{r}) \frac{\delta E^{\text{xc}}}{\delta \rho(\mathbf{r})} \varphi_{\nu}(\mathbf{r}) d\mathbf{r} \quad (11)$$

i.e. the kinetic energy, nuclear attraction and exchange-correlation integrals, respectively, and the more expensive Coulomb repulsion two-electron integrals

$$\int \varphi_{\mu}(\mathbf{r}_1) \varphi_{\nu}(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1-\mathbf{r}_2|} \varphi_{\lambda}(\mathbf{r}_2) \varphi_{\sigma}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (12)$$

Who already knows the Hartree-Fock theory, or who is willing to study it from standard textbooks, will recognize that, even if they are based on different concepts, there is a strict analogy between the Kohn-Sham implementation of DFT and the Hartree-Fock formalism, and this is why the old computer programs able to calculate the HF energy were readily extended to DFT. As a matter of

fact, the Kohn-Sham equations are very similar to the HF ones, with an important difference: the exact exchange energy contribution appearing in the Fock operator is replaced, in the Kohn-Sham operator, by the functional derivative of the exchange-correlation functional, introducing de facto Coulomb electron correlation, hence higher accuracy, in DFT. The equivalent of the matrix equation (8), in HF, are the famous Roothaan equations (of utmost importance for the historical development of computational chemistry) while, as a consequence of what said above, the third kind of one-electron integrals in (11) are replaced in HF by the exchange integrals, i.e. two-electron integrals that must be obtained from the Coulomb repulsion integrals by means of specific permutations. Just about the exchange-correlation integrals in DFT, it must be said that it has to be evaluated by means of some numerical quadrature technique, based on atom-centered grids of points and employing the numerical values of the basis set functions in these points.

At this point it should be clear that the DFT problem can be solved only after having found a suitable description of E^{xc} , an issue that is at the heart of density functional theory; but how can one achieve a suitable description of something unknown? Apart from the knowledge of some characteristics the exact exchange-correlation functional must have, its form is commonly chosen and adapted to reproduce well defined experimental or highly accurately calculated data for various property. This introduces some empiricism in DFT, that can be applied ab initio only after having defined possible parameters inside E^{xc} , but did not prevent this approach to become the most widely used in computational chemistry, being it applicable with confidence to large size systems (more or less those treatable with the Hartree-Fock method). The forms for E^{xc} , firstly proposed in the early 60s, were collected in two families: those functionals depending only on the electron density (local density approximation, LDA) and those including also a dependence on the gradient of the electron density (generalized gradient approximation, GGA). However, the routine use of DFT within the framework of molecular structure calculations can be traced back to the beginning of the 90s when Becke found the way to introduce a portion of the exact Hartree-Fock exchange within the formalism, devising the so-called hybrid exchange-correlation functionals. Since then researchers in the field proposed a lot of new functionals, most of them dedicated to alleviate some problems within DFT. As a matter of fact, DFT is plagued by two major problems: it does not reduce the accuracy problem on the description of dispersion interaction, already present in the HF method and whose solution would require the use of expensive highly correlated methods, and it fails to describe charge transfer electronic excitations. While some attempt was proposed to mitigate the second problem (e.g. range-separated functionals), new generation exchange-correlation functionals (meta-GGA, with empirical corrections, etc.) seem to be fruitful on solving the first one, that can be important in computational material science. In particular, the very cheap and efficient correction due to Grimme is becoming an essential ingredient of all functionals.

The trend today is to relegate the use of post-Hartree-Fock method to very small systems, in particular when one wants to investigate properties requiring high (chemical or more) accuracy, such as spectroscopic constants, atomization energies, etc.; coupled cluster is the method of choice if the static correlation (that due to the fact that two or more electron configurations are equivalent or nearly-equivalent) in the investigated molecule can be regarded as negligible, otherwise some CAS-based approach, followed by a multireference method, must be applied. For medium-large molecules, reaction kinetics, periodic calculations, ab initio molecular dynamics simulations and such, density functional theory, preferably with a dispersion-corrected exchange-correlation functional, is the only formalism to employ, at least until some new coupled cluster variant becomes usable for large molecules.

Density functional tight binding

But what if the system size are so large that even DFT becomes not affordable? This is indeed common in material science and it is clear that in these cases some compromise must be accepted in terms of the desired accuracy. This said, one can wonder what computational approach could be

used for very large systems. In the first years of quantum chemistry, when calculations were prohibitive also for medium size molecules, a plethora of empirical and semiempirical methods arise. Starting from the simple Huckel method, from where chemists presumably developed their, sometimes too naive, vision of molecular electronic structure, in these kind of approaches, together with the introduction of crude approximations within the Hartree-Fock theory, there is the substitution of the expensive computation of most electronic integrals, in particular the two-electron ones, with some kind of atomic parameters. The implementation of these methods, therefore, required tables containing the atomic and atomic-type parameters, obtained in a number of ways but always outside from the method itself (which was therefore never *ab initio*). The ideas behind semiempirical methods developed enormously and those approaches which still survive (most of them are used no more) reached a remarkable degree of sophistication, which allows some confidence for their application to large size systems. However, just as Hartree-Fock method was superseded by density functional theory, so semiempirical methods based on approximate HF treatments have left the place to approximate DFT descriptions, one and for all the so-called density functional tight binding (DFTB) approach. In the DFTB formalism [5], a reference electron density is defined as a superposition of the electron densities of the isolated neutral atoms forming the investigated molecule and the molecular energy is expanded at a given order in the electron density fluctuation, $\delta\rho(\mathbf{r})$. As happened for DFT, also the DFTB method was for a long mainly confined within applications to solid state, until a newly proposed variant made it successfully applicable for molecular systems. This variant is the self-consistent-charge density functional tight binding (SCC-DFTB) approach, according to which the system energy is expressed as the following expansion to second order in $\delta\rho(\mathbf{r})$

$$E^{\text{DFTB}} = \sum_i n_i \int \theta_i(\mathbf{r}) H_0^{\text{KS}} \theta_i(\mathbf{r}) d\mathbf{r} + E^{\text{REP}} + \frac{1}{2} \int \frac{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \frac{1}{2} \int \delta\rho(\mathbf{r}) \frac{\delta^2 E^{\text{XC}}}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} \delta\rho(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \quad (13)$$

where n_i is the occupation number of the i -th θ orbital and H_0^{KS} is the tight binding Hamiltonian. A number of approximations leads to the parameterized expression

$$E^{\text{DFTB}} = \sum_{i\mu\nu} n_i c_{\mu i} c_{\nu i} H_{\mu\nu}^0 + \sum_{ab} U_{ab} + \frac{1}{2} \sum_{ab} \gamma_{ab} \Delta q_a \Delta q_b \quad (14)$$

where the first term is obtained from the corresponding term in (13), after the θ functions have been substituted by a linear combination of atomic orbitals (obtained by solving the so-called pseudo-atom equation, containing a confinement parameter). The sum of U_{ab} terms approximate the repulsive E^{REP} energy contribution; the U_{ab} pair potentials, depending on the distance between the centers a and b , are parameters specific to every possible couple of atoms in the investigated molecule and are obtained by a fitting procedure based on the difference between E^{DFTB} and DFT reference energies calculated as a function of the distance of atoms a and b in simple interacting molecules. Finally, the last term in (14) collects the last two terms in (13) and describes them as fluctuations of the atomic Mulliken charges q as a function of molecular geometry, with the γ_{aa} parameters correlated to the difference between the ionization potential and the electron affinity of atom a , and the γ_{ab} parameters expressed in analytic form by a damped Coulomb potential.

In further developments, the energy is expanded to third order, which seems to give better results for dispersion interactions. Apart this, however, the DFTB approach is plagued by the need of too specific parameters, whose transferability is hardly guaranteed. One could get around this obstacle by generating parameters appropriate only for specific systems and purposes, which is what is generally done, but this generation is too involved and arbitrary to some extent. It is true that the Extended Tight Binding (XTB) formalism and all its variants, recently proposed by Grimme, is free from atomic-pair-specific parameters, and therefore is the optimal candidates to achieve the leading

position between the semiempirical approaches, but this method still need some testing for general applicability.

Hybrid methods

If a molecular system is formed by hundreds or thousands of atoms, it is in some cases possible that the property we are interested in is localized in a determined portion of the molecule, while the influence of the remaining part can be viewed as a substituent effect. Clearly, this is not the case if one wants to investigate the secondary structure of a small protein, but some useful informations can be acquired on the local interactions determining that secondary structure. The approach to use in such cases was pioneered by Morokuma in the 90s, when, by collecting and improving earlier ideas, he proposed the ONIOM method. According to this approach, the system to investigate is divided in two or more well defined regions which are related one to the other by an inclusion relationship. One of these region (the “model system” M) is the molecular portion considered the most responsible of a given property of interest; in the 2-layer version of ONIOM, M is included in the entire molecule (indicated as “real system”, R), while in the 3-layer version, M is included in an intermediate system, I, which is in turn embedded in R. Now, a high accuracy computational method H is chosen which can be applied to the model system, followed by a medium accuracy method, m, that can be applied to the intermediate system, if any, and a low level method, L, that can treat the entire real system. By indicating with $E(A,B)$ the energy of a system A treated with the computational method B, the energy $E(R,H)$ which is the uncomputable energy of the whole molecule at high accuracy, is extrapolated as $E(M,H)+E(R,L)-E(R,L)$, for Oniom-2layer, or as $E(M,H)+E(I,m)+E(R,L)-E(M,m)-E(I,L)$ for the 3-layer variant, i.e. $E(R,H)$ is evaluable by means of affordable only calculations. The expression of the target energy as a sum of independent contributions, joined to a proper treatment of the link atoms, i.e. those atoms (generally hydrogen) introduced on the edge of model and intermediate systems in order to saturate the dangling bonds originated from the cut and make them suitable for a standalone calculation, make possible the evaluation of ONIOM energy nuclear derivatives, essential for the analytical geometry optimization and vibrational frequency calculation, thus allowing a very fruitful employment of this approach.

A specific, somehow crafty usage of the ONIOM concept is the building of a necessary framework around a portion of a very large molecule or even of a periodic system. The issue is better explained with an example. Consider one wants to investigate the adsorption of a given molecule on the inner wall of a zeolite cage, let's say the wall of the larger channel in the beta zeolite. Taking into account the wall thickness, the region of this channel is defined by ca. 80 silicon tetrahedra, giving a total of ca. 250 atomic centers. Such a supercell hardly can be the subject of accurate calculations; but the problem becomes prohibitive if one wants to avoid spurious interactions of the adsorbed molecule with its periodic images, so that the supercell must be doubled. A way to treat the problem could be to define a minimal inner surface, i.e. the smallest surface that can accomodate the molecule whose adsorption characteristic are desired, and cut it from the zeolite structure, having care to saturate with hydrogen the dangling bonds originated from the cutting. The system so created should be subjected at least to geometry optimization, to discover the adsorption geometry. Doing so, however, the adsorbing surface would relax, and it will be no more representative of the zeolite largest channel wall. Indeed, that portion of the surface has the correct geometry only if it is embedded in the zeolite framework. Being the use of a larger supercell prohibitive, it could be a natural choice to couple the DFT approach employed for the portion of interest to a semiempirical, or better a molecular mechanics force field, according to ONIOM prescription, the embedding assuring to some extent no spurious distortion of the channel surface. It is trivial to say that, just as in all other cases in which computational chemistry approaches are used, preliminary investigations must be afforded to test the reliability of the model and have some confidence with the results obtained.

The approaches for periodic systems

Let's take a look now to the quantum chemical approaches that can be used to treat the solid state. According to Bloch theorem, a function which approximately describes the probability amplitude of finding an electron in a certain point of the space, when it is subjected to a the periodic potential generated by an infinite lattice of positive charges and to the mean field repulsion of all other electrons, can be expressed in the form

$$\xi_{\mu k}(\mathbf{r}) = U_{\mu k}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \quad (15)$$

where the U functions have the same periodicity of the crystal lattice, represented by the vector $\mathbf{L} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$, being $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ the three lattice vectors, and \mathbf{k} is a reciprocal space vector inside the so-called first Brillouin zone (BZ). In the DFT context, this means that the solutions of the Kohn-Sham equations (5) must satisfy the following constraint

$$\xi_{\mu k}(\mathbf{r} + \mathbf{L}) = e^{i\mathbf{k}\cdot\mathbf{L}} \xi_{\mu k}(\mathbf{r}) \quad (16)$$

The most natural choice for the expansion of the ξ functions is a linear combination of plane waves

$$\xi_{\mu k}(\mathbf{r}) = \sum_{\mathbf{G}} a_{\mathbf{k}}(\mathbf{G}) e^{i(\mathbf{k} + \mathbf{G})\cdot\mathbf{r}} \quad (17)$$

where the summation is extended to all the reciprocal space vectors \mathbf{G} selected from a grid and includes all the plane waves having kinetic energy below a certain cutoff. Since in order to reproduce the local form of the ξ functions prohibitively high values of the cutoff should be necessary, only the valence regions of the atoms are explicitly considered, inner regions being described by appropriate pseudopotentials. Finally, the electron density needed in the KS equations is defined and calculated on a grid of points belonging to the first Brillouin zone, according to the sampling proposed by Monkhorst and Pack:

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{\mu=1}^n \int_{\text{BZ}} \xi_{\mu k}^*(\mathbf{r}) \xi_{\mu k}(\mathbf{r}) d\mathbf{k} \simeq \frac{1}{V} \sum_{\mu=1}^n \sum_{\mathbf{k} \in \text{BZ}} \xi_{\mu k}^*(\mathbf{r}) \xi_{\mu k}(\mathbf{r}) \quad (18)$$

The expansion in terms of plane waves according to equation (17) is certainly the most used, and is implemented in the most common softwares for periodic DFT calculations, such as VASP and AbInit, but is not the only possible choice. In the Crystal program, gaussian basis sets are employed while the SIESTA code, which implements the homonym computational method and in the first years of 2000s represented a breakthrough in this field, efficiently uses appositely generated numerical atomic orbitals whose quality can be tuned. From a periodic calculation on a crystal lattice (whose symmetry characteristics commonly come from the experiment) useful informations can be obtained: electronic bands, energy gap and conducting properties, cell geometries, density of states (from which one can extract information such as the number of mobile states for electric conductivity), phonon dispersion relations, elastic properties, etc. However there are very important chemical processes (one and for all, heterogeneous catalysis) which do not occur in the infinitely self-replicating bulk of the crystal, but on regions where lattice periodicity is broken: the crystal surfaces. What is the approach to use in order to investigate a portion which is located at the interface between a semi-infinite periodic crystal and the vacuum? In the most intuitive approach, the system is still periodic in three dimensions, but the cell vector normal to the surface object of investigation is artificially elongated to assure a vacuum (typically 30 Å); along the other two directions and the direction opposite to the vacuum the unit cell is replicated, with unchanged lattice constants, a number of times sufficient to create a slab, which is treated as a supercell to replicate periodically. In this slab approach care should be taken to have some issues always under control: i) the slab should be large enough to avoid spurious, unwanted interactions between a molecule adsorbed or reacting on the surface and its periodic replicas; ii) the slab should be thick enough to

mimic the semi-infinite bulk below the surface and its polarizing effect; iii) for semiconductors it is necessary to saturate the dangling bonds in the underneath surface; iv) energy should be corrected to eliminate the fictitious dipole-dipole interaction originated by slab creation. Alternatively, the cluster, not periodic, approach can be employed. In this case, the vacuum above the surface is guaranteed and lateral interactions cannot occur; further, being outside from the field of periodic calculations, one can take advantage of the wider choice of methods proper of molecular calculations (e.g. hybrid exchange-correlation functionals and tailored basis sets), not to mention the availability of energy gradients which greatly facilitates the quest for transition states in chemical reactions (for transition states in periodic investigations numerical approaches must be used instead, such those based on the Nudged Elastic Band method). However, the cluster approach has its own problematic deficiencies, such the possible occurrence of strong border effects and the unnatural structural relaxation of the whole system. Obviously, the last sentence is true if one wants to describe a crystal surface by using a cluster as model, but those mentioned become physical characteristics of the system if the object of the investigation is a true, physical cluster of atoms.

Molecular dynamics simulations

Just some rows above, the word "force field" comes into play. Even if it could be applied in a wider context, this term was used here to introduce the models of classical molecular mechanics (CMM) and molecular dynamics (CMD) and explore if there is some room for their use in material science. By using the term "classic" it is intended that the treatment of the molecular systems occurs outside the realm of quantum mechanics; so, electrons are no more considered explicitly but their presence is incorporated inside fictitious parameters, representing e.g. the force of a given chemical bond. Energy is here calculated as a sum of parameterized contributions, commonly grouped in internal (bond stretch, bending, etc.) and external (Coulomb, van der Waals interactions and so on) to the molecular entity. The union between i) the nature of the contribution to the energy, ii) the formula chosen for the calculations of each contribution and iii) the set of all the atom-type and contribution-type specific parameters, form a force field. As an example, one of the most diffuse force field is the AMBER one, which defines the energy of a M-atoms molecule as

$$E = \sum_{\text{bonds}} \frac{1}{2} K_b (r - r_0)^2 + \sum_{\text{angles}} \frac{1}{2} K_a (\theta - \theta_0)^2 + \sum_{\text{torsions}} \frac{1}{2} V_t [1 + \cos(n\omega - \gamma)] \\ + \sum_{i=1}^{M-1} \sum_{j=i+1}^M \frac{q_i q_j}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^{M-1} \sum_{j=i+1}^M \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{|\mathbf{r}_i - \mathbf{r}_j|} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{|\mathbf{r}_i - \mathbf{r}_j|} \right)^6 \right] \quad (19)$$

where the first term defines the contribution of bonds stretching, written as a harmonic potential depending on the interatomic distance r and containing two parameters: the harmonic force constant for the given bond, K_b , and its equilibrium length, r_0 . Similarly, the second contribution is that due to bond angles bending, with obvious symbols, while the third collects the energetic terms related to variations in the torsional angles. Electrostatic interactions between atoms in the same or in different molecules are described by a simple Coulomb potential between fictitious atom-centered partial charges q , and dispersion interactions are entrusted to a 12-6 Lennard-Jones potential, with its famous pair parameters ϵ and σ . These last two contributions run over all possible couples of atomic centers in the investigated system. The atom-type and pair parameters are to be determined by some procedures which could involve well defined experimental data, equation of state, dedicated experiments, ab initio calculations, and so on. Transferability is sometimes assured between parameters for molecules belonging to classes of homologue systems (e.g. proteins, DNA, sugars), but commonly a force field cannot be considered of general applicability, so that a very large number of new force fields (as well as extension/variation/recalibration of existing ones) suitable for particular purposes are continuously proposed to the scientific community. For example, based on the AMBER force field but without the torsional term, the CLAYFF is parameterized on purpose for the investigation of the clay-water interface.

Force field based approaches can be used for geometry optimization by simply calculating the energy of a certain molecular geometry and relaxing it according to the force on every atom, and this is the way it is used for assuring the embedding of the molecular portion in ONIOM calculation. More often, on the other hand, the force fields are used for the time evolution of systems (particularly in the condensed phase), by discretizing the physical time in appropriate timesteps (in the order of femtoseconds) and exploiting the forces obtained from the force field in efficient algorithms (e.g. velocity Verlet) which propagate the system according to classical Newton or Lagrange equations. One speaks in these cases of classical molecular dynamics simulations. Dedicated calculation centers can afford CMD simulation for relatively long times (μs) or very large system size (billion of atoms), but it clear that the results must be trusted in the view of ensemble properties, being useless here to search for accuracy on the details. The quantum-mechanics counterparts, like Born-Oppenheimer molecular dynamics or the Carr-Parrinello approach, where the forces between atomic centers can be calculated ab initio from the potential due to the electrons, are conversely full of details, but are applicable only to small systems for very short simulation times.

Case studies: halloysite nanotubes

Halloysite nanotubes (HNTs) are natural clay materials first described in 1826 by Berthier. They have peculiar spiral morphologies conferred by the rolling up of kaolinite-like sheets, composed by a layer of SiO_4 tetrahedra (which will form the outer surface of the nanotube) joined to a layer of AlO_6 octahedra (representing the inner surface). The HNT stoichiometry, $\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4 \cdot n\text{H}_2\text{O}$, allows to distinguish a hydrated form ($n = 2$, indicated as HNT-10Å), bearing interarms water molecules, and a dehydrated one ($n = 0$, HNT-7Å). Relegated for a long time to a use almost exclusively relevant to the ceramics industry, in the last fifteen years HNTs started to become important as a cheap and eco-friendly material for a number of applications, ranging from controlled release to nanocarrier for active compounds, from design of polymeric nanocomposites to catalysis.

The computational characterization at atomistic level of HNT composite must face many challenges. Let's imagine one wants to investigate the hydrogen bonds network which realizes in the region comprised between the spiral arms of a halloysite-10Å nanotube. Here there are H-bond interactions between the water molecules and between them and the two surfaces of HNT. The spiral structure of HNTs gives rise to serious problems for the simulation and the first of them is merely the construction of the input system geometry. It may sound trivial, but every calculation approach needs a molecular geometry to begin with: many free or commercial softwares are indeed devoted to the graphical construction of molecules to be used as input geometry but when one treats large size systems it is clear that some automatism must be devised. A vast number of codes are dedicated to the building, visualization and analysis of unit cells of given symmetry for periodic calculations, or for the construction of polypeptides given the list of the aminoacid residuals in the chain, or the generation of carbon nanotubes and alike given the chirality. HNTs, however, cannot be builded using codes for periodic systems, since they are periodic only in the direction of the rolling axis. As a matter of fact, the spiral has a curvature which changes from point to point, so no exact periodicity of a unit cell can occur along its arms. One artifact could be to forget the spiral structure and use a multiwalled cylindrical model for HNT, but, as will be discussed below, this could cause the missing of some features related just to the spiral architecture. It is better to assume that HNT can be generated by starting from a reference regular lattice of points arranged on the arms of an Archimedean spiral, whose polar equation is $r(\theta) = a + b\theta$. The lattice points can be placed on the spiral by exploiting the formula for the spiral arc length

$$s(\theta) = \frac{1}{2}b \left[\theta\sqrt{1+\theta^2} + \ln(\theta + \sqrt{1+\theta^2}) \right] \quad (20)$$

An appropriate repetition unit is now convoluted with every lattice point (see Figure 2), opportunely rotated and slightly adjusted according to a correction angle determined from the spiral curvature in that point (which decreases smoothly along the arms)

$$\kappa(\theta) = \frac{2+\theta^2}{b(1+\theta^2)^{3/2}} \quad (21)$$

This approach works well unless the curvature is too pronounced, but this should be never the case since the HNT model must have a relatively large inner cavity, hence the phase of the spiral is never too small. Clearly, the model constructed this way is only the starting point for geometry optimization. As a matter of fact, after having generated the HNT supercell, one must face the problems related to its dimension: even a small supercell corresponding to the section of a HNT-10Å having a modest 5 nm inner cavity diameter, and formed by spiral arms that overlap for one half, has ca. 1400 atoms if a armchair rolling is chosen, a number which increases to more than 2000 when the opposite limiting zig-zag rolling direction must be considered. Such a supercell can be periodically replicated along the nanotube axis by using the appropriate cell vector length, while vacuum must be assured in the other two perpendicular directions. This system is too large to be treated with *ab initio* methods, but its dimension is within the applicability range of semiempirical approaches like, e.g., SCC-DFTB. If this is the case, the *matsci* set of Slater-Koster parameters could be used, which was realized to treat systems containing common elements in material science. Concerning the code to use to perform the calculation, as a guideline I suggest to employ whenever possible softwares dedicated to the implementation of only one kind of computational approach instead of general ones; so in this case the choice could be the DFTB+ program developed by Aradi and coworkers.

Geometry optimization of the HNT-10Å supercell reveals some features connected to the intrinsic disorder of the spiral architecture. In order to discuss these ones, let make a comparison with a hypothetical system formed by water molecules sandwiched between two planar kaolinite layers, i.e. let unroll the HNT-10Å. In this case a perfect periodicity is present in the system: all tetrahedral silicon centers have equal geometry, the same holds true for the octahedral aluminum centers and, from a structural point of view, a regular network of hydrogen bonds would occur between the water molecules, and between them and the upper and lower surfaces. The transformation of this system to the spiral HNT-10Å causes non periodic distortions of the SiO₄ and AlO₆ geometries along the arms and calculation revealed that (at least for the idealized system investigated) the less amount of energy required seems to occur in spiral arrangements where relatively large distortions are localized in a limited number of points along the arms, joined to very small distortions in all other points. This sort of polygonal spiral would be preferred to a smoother one where distortions would be distributed between all the SiO₄ and AlO₆ centers. Further, the characteristics of the hydrogen bonds network in the unrolled HNT originated from the fact that the water layer was comprised between a periodic SiO surface and a periodic AlO ones. In the spiral HNT-10Å, on the other hand, the water molecules experience the disordered environment in which they are enclosed; as a consequence, a large variety of hydrogen bond configurations can take place and the water molecules would adapt in order to provide the best interactions. Still, at a given temperature, the H-bonds configurations can change over time, giving rise to a dynamic linkage between the arms of the halloysite nanotube, whose characteristics could be investigated in terms of pair correlation functions after that a dedicated molecular dynamics simulation is performed. The occurrence of the adaptive and dynamic H-bond network, originated from the natural disorder related to the spiral structure, could be one of the reasons why HNT-10Å exists instead of a lamellar kaolinite with water between adjacent layers.

After having investigated the HNT model in its integrity, one could be interested to nanocomposites involving modified or functionalized HNTs for specific applications, and this constitutes another challenge, to tackle which one must resort to the usual compromises. To alleviate the burden, it can be considered that many applications of halloysite-based composites depend on the interactions or

reactions of its outer and inner surface with molecules of interest. Then, most of the cases can be investigated by using local models of HNT instead of the whole nanotube. Conversely, it must be taken into consideration that now we must go into the details of interactions, adsorption geometries, local distortions, bonds breakings and formations, activation barriers, etc., therefore quantitative informations are needed if one wants to be predictive, and resorting to accurate DFT approaches cannot be avoided. Since this is expensive (or even impossible) both in terms of time and resources, a suitable choice must be done. A common way to proceed could be to tailor a portion of surface from the nanotube model. What could be a representative portion in terms of lateral extension and thickness would heavily depend on the nature of the system and on the properties the investigator is interested on but, again, the choice is sometimes restricted to what can be treated by DFT. One could assume that a coronene-like surface portion formed by 24 silicon tetrahedra and the same number of aluminum octahedra (depicted in Figure 3) could be a good, minimax, choice for e.g. the adsorption of one molecule, or the surface modification with one molecular fragment. It could be treated as the model system of a ONIOM calculation or as a standalone molecule; in both cases the DFT calculation would be performed using a HNT portion with $\text{Al}_{24}\text{Si}_{24}\text{O}_{126}\text{H}_{48}$ stoichiometry, where the additional hydrogen atoms would come from the need to saturate the dangling bonds originated by the cut (which is the role of the linking atoms in ONIOM). In order to obtain a reliable description of dispersion forces one could use the B3LYP-D3 exchange correlation-functional, i.e. a hybrid one with Grimme correction; however, if one wants to take advantage of the resolution of identity approximation, a pure exchange-correlation functional should be used, in which case the choice could fall on some Minnesota functional, such as the M06-L. The basis set must be obviously of double zeta quality plus polarization. If an ONIOM approach is employed, the high level DFT method can be joined to a molecular mechanics force field (as CLAYFF, or the Universal Force Field, UFF) as low level method. If instead the portion is investigated as a standalone unit, the user should pay attention to freeze the positions of some non influent atomic centers (e.g. those on the borders), otherwise the portion will distort and the spiral local curvature will be lost. Taking into account that a molecule of, say, 20-30 atoms is to be adsorbed on the HNT portion, the DFT geometry optimization of the entire system will take from 2 to 6 days on 32 shared-memory new generation processors.

Final remarks

Computational chemistry is an extremely important resource of modern science, which can work in synergy with the experiment. It can be useful for developing basic knowledge, for the interpretation of data and trends, for the prediction of the characteristics of new molecules or of the behavior of already known molecules in new or unusual conditions. It can be applied in essentially all fields of chemistry. However, before starting his/her computational work, the researcher who intends to practice calculations should ask himself/herself the following set of questions: why do I want to make calculations on this system? Have I well identified the problem I want to investigate? Can I obtain the answers I need? What accuracy do I need? The one required to reproduce, for example, the rotational spectrum of a molecule is far greater than that commonly used to describe a reaction mechanism. In addition to being necessary to establish a correct way to proceed, having these issues clear helps to understand which methods, among the many available, should be used. It must be borne in mind that some types of computational investigations may be impossible, at least today, and that sometimes one could be satisfied with evaluating trends: the exact value of a particular quantity that refers to a physical property could be only a detail.

References

- 1) C. J. Cramer, "Essentials of Computational Chemistry - Theories and Models", Second edition, Wiley
- 2) D. Young, "Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems", Wiley-interscience
- 3) J. G. Lee, "Computational Materials Science - An Introduction", Second edition, CRC Press
- 4) K. Ohno, K. Esfarjani, Y. Kawazoe, "Computational Materials Science - From Ab Initio to Monte Carlo Methods", Springer
- 5) Spiegelman, F, Tarrat, N, Cuny, J, Dontot, L, Posenitskiy, E, Marti, C, Sikmon, A, Rapacioli, M 2019 'Density-functional tight-binding: basic concepts and applications to molecules and clusters', *Advances in Physics: X*, vol. 5, no. 1, pp. 1710252

Figure captions

Figure1: Flowchart representing the cycle of the geometry optimization algorithm and the subsequent calculation of routine molecular properties

Figure 2: (a) An Archimedean spiral lattice of points defined by spiral parameters such that the inner diameter-like distance is close to 10 nm and the interarms spacing is enough to give HNT-10Å. (b) An appropriate repetition unit formed by a Si-O layer connected to a Al-O one and four water molecule, with stoichiometry $H_{15}Si_4Al_4O_{24}$; color code: yellow = Si, cyan = Al, orange = O, white = H. (c) The HNT-10Å input supercell as the result of the convolution of the repetition unit with the spiral lattice, followed by a rotation whose entity depends on the lattice point. (d) The HNT-10Å spiral nanotube obtained by periodic repetition of the supercell along the rolling axis.

Figure 3: The nanotube portion with $Al_{24}Si_{24}O_{126}H_{48}$ stoichiometry, which could be used as a model to investigate the possible modification of the outer and inner surfaces of halloysite. Color code: yellow = Si, cyan = Al, orange = O, white = H.