Quantum Chemical Calculation of Chemical Shifts in the Stereochemical Determination of Organic Compounds: A Practical Approach

10

Simone Di Micco, Maria Giovanna Chini, Raffaele Riccio, and Giuseppe Bifulco

Contents

Abstract

In this chapter, we report an integrated approach of NMR and quantum mechanical calculation for the determination of the relative configuration of natural products. The entire protocol is described starting from building the investigated compound to the calculation of NMR properties at quantum theory level and the interpretation of the results. Each step of the protocol is described, and the main applied methods are reported. We report, as case studies, the determination of the relative configuration of two natural products: bonannione B isolated from *Bonannia graeca*, and callipeltin A isolated from the sponges *Callipelta* sp. and *Latrunculia* sp. Through the analysis of these natural products, we show the use of ^{13}C chemical shift and homo and hetero *J* coupling constants, respectively, as

S. Di Micco • M.G. Chini • R. Riccio (✉) • G. Bifulco (✉)
Department of Pharmaceutical and Biomedical Sciences, University of Salerno, Via Ponte Don Melillo, Fisciano (Salerno), Italy
e-mail: sdimicco@unisa.it, mchini@unisa.it, riccio@unisa.it, bifulco@unisa.it

E. Fattorusso, W. H. Gerwick, O. Taglialatela-Scafati (eds.),
Handbook of Marine Natural Products, DOI 10.1007/978-90-481-3834-0_10,

an important tool in the interpretation of the experimental data for the determination of the relative configuration of organic compounds.

10.1 Introduction

In the last decade, quantum chemical approaches have shown their potential in solving chemical problems. This has been mostly due to the ever increasing computational capabilities at a relatively affordable cost and to a parallel development of user-friendly software. While many aspects of molecular structure and dynamics may be solved with the use of classical methods based on empirical force fields, quantum chemistry allows the comprehension of many problems related to the electronic density distribution. In particular, quantum chemistry methods can be applied for calculating the spectroscopic properties of molecules, and the efficient prediction of UV, IR, CD, and NMR spectra have been extensively reported in the last 15 years. NMR chemical shift calculation by quantum mechanical methods has attracted the interest not only of the theoretical chemists but also of the experimental NMR spectroscopists. In fact, this kind of approach has been used by our and other research groups as a contribution to the structure elucidation of natural products. We have presented two original methodologies, based on GIAO (gauge including atomic orbitals) quantum mechanical ^{13}C chemical shift calculations, that have been efficiently employed as a support in the analysis of the NMR data of organic molecules. The first methodology regards the structure validation of natural products by means of GIAO ^{13}C chemical shift calculations, while the second one, based on the same methodology, has been directed to the determination of the relative configuration of flexible compounds. For the interested readers, the most significant applications of quantum chemical calculations of NMR parameters in the resolution of stereochemical problems have been recently reported by us in two reviews. On the other hand, the scope of this chapter is to present a step-by-step guide for the NMR parameter calculation of organic compounds in the determination of their relative configuration.

10.2 Calculation Protocol for Molecular Frameworks

Different experimental approaches have been proposed for the determination of the configuration of organic compounds [1, 2]. One of the most used is the total synthesis, which is highly demanding in terms of human and economical resources. On the other side, there are analytical methods, such as nuclear magnetic resonance (NMR), X-ray crystallography, circular dichroism (CD), and mass spectrometry, which allow also to preserve the investigated compound. For the stereostructural determination of natural products, it is possible to apply a protocol based on the calculation of ^{1}H and/or ^{13}C of chemical shifts as a support in the experimental data analysis (Fig. 10.1) [3, 4]. Such a protocol consists of up to six fundamental steps: (a) building the molecules by dedicated software; (b) conformational search at the empirical theory level [5, 6], generally through molecular dynamics (MD) or by Monte Carlo multiple minimum

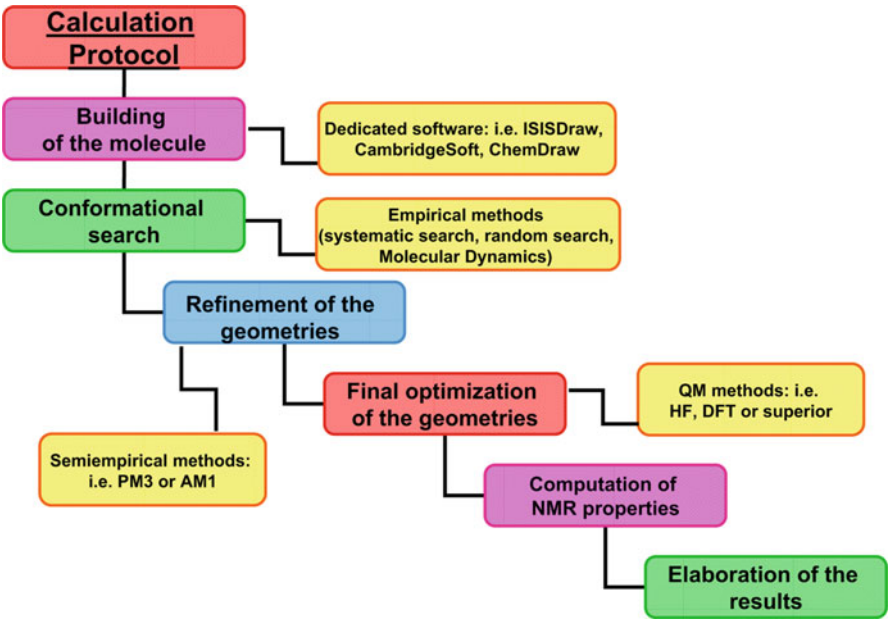

Fig. 10.1 General protocol used for the determination of the relative configuration of organic compounds

methods (MCMM) [7]; (c) preliminary geometry optimization at semiempirical level (e.g., AM1 [8], PM3 [9]) of all the significantly populated conformers of each stereoisomer; (d) final geometry optimization of all the species at the quantum mechanical (QM) level; (e) GIAO (gauge including atomic orbital) ^{13}C and/or ^{1}H NMR calculations of all the structures obtained by the optimization at QM level of all diastereomers, taking into account the Boltzmann distribution; and (f) elaboration of the results comparing the Boltzmann-averaged NMR parameters calculated for each stereoisomer with those experimentally measured for the compound under examination. In order to facilitate the understanding of the protocol, we will show an application of the QM/NMR methodology to a real case: the assignment of the relative configuration of the bonannione B (**1**, Fig. 10.2), a flavanone isolated as minor compounds from the aerial parts of *Bonannia graeca* (*Umbelliferae*) [10].

The example of the flavonoid bonannione B (**1**, Fig. 10.2) represents a straightforward application of ^{13}C NMR chemical shift calculation, since the configuration of stereopair (C-2″ and C-3″) under investigation could not be deduced by simple analysis of the 2D ROESY spectra, and the small amount of the isolated compound was not sufficient to apply Mosher's method [11].

10.2.1 First Step: Building the Molecules

The generation of the molecule model represents the starting point for any computational chemistry study. The tridimensional model (3D) defines the relative

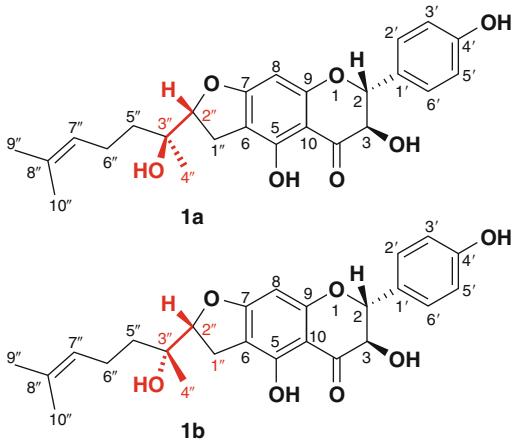

Fig. 10.2 Molecular structure of two possible diastereoisomers of bonannione B (**1a** and **1b**)

position of the atoms in the space through a set of Cartesian coordinates obtained from several sources. The X-ray crystallographic database [12], the building from a preexisting fragment library, the conversion of 2D structural data (e.g., .cdx, .skc, .sk2, and smiles files) into 3D form (e.g., .mol, .pdb, .mol2, .mae, and .gjf files) with a dedicated software – ACD/ChemSketch [13], ISISDraw [14], and ChemDraw [15] – represent some examples for the achievement of the model. In the illustrated case of bonannione B (Fig. 10.3), it is possible to build the 2D structure with a molecular editor software, for example, by using the ChemOffice package, and then converting it into a file containing the Cartesian coordinates (Fig. 10.4).

The choice of method to adopt depends on the case and the resources of the user, but in any case, a reasonable and reliable initial geometry is essential for the quality of the following investigation.

10.2.2 Second Step: Conformational Search

The QM calculation of properties belonging to the investigated compound, such as energy, UV, NMR parameters, must be performed on models that better represent the real conformations of the molecule: the minima of the potential energy surface. In fact, differently from rigid molecular systems, flexible molecules may exist in more than a single conformation. The changing among the conformers, derived from the variation of dihedral angles around single bonds, corresponds to different points in the potential energy surface. The diverse energy points, coinciding with local minima, may represent the different conformations adopted at the thermal equilibrium by the molecule. As the spectroscopic properties strictly depend on the compound geometry, their calculation needs to take into account all the possible representative local minima in equilibrium because each conformer will have a specific weight in the global spectroscopic properties. (These concepts can be better understood analyzing the different conformation of ethane and n-butane.

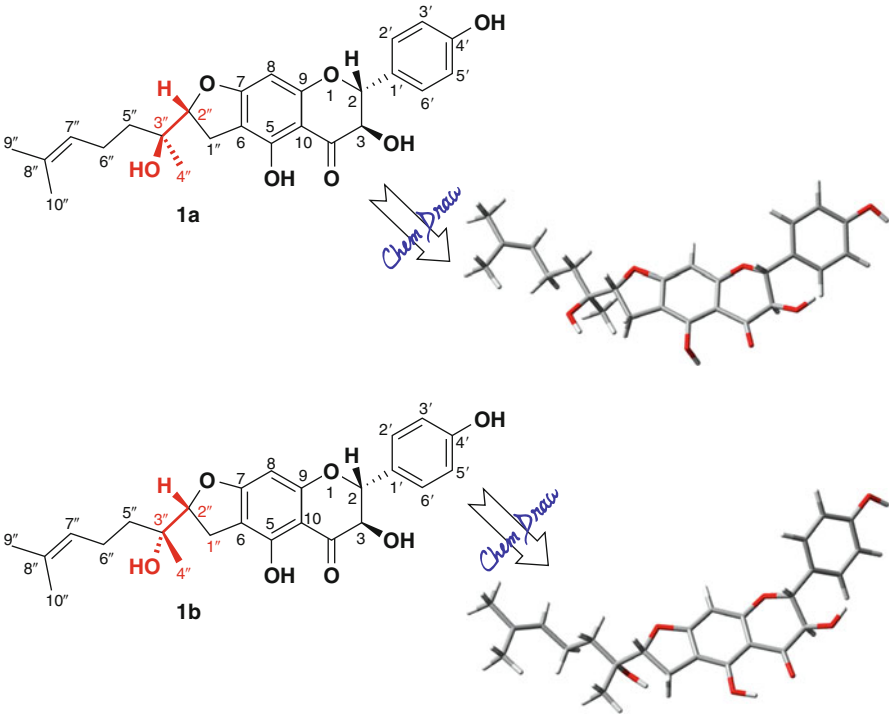

Fig. 10.3 Conversion of 2D structure into a 3D model of bonannione B

Increments of 60° for torsion angle around the C–C bond in ethane translate in two extreme cases: three eclipsed conformations at 0°, 120°, and 240° (global maxima), and three staggered conformations at 60°, 180°, and 300° (global minima), and so for the study of its physical and chemical properties, only the staggered arrangements must be considered (Fig. 10.5). Even if the rotation around carbon–carbon bond is not completely free owing to the energy difference between the two conformer families, in the case of ethane, the difference of 2.8 kcal/mol is small enough to be overcome, and thus, the conformers are interconvertible at room temperature. In the case of n-butane, increments of 60° for torsion angle around the bond C2–C3 give rise to one maximum global at 0° (eclipsed form), two local minima at 60° and 300° (*gauche* form), two local maxima at 120° and 240°, and one global minimum at 180° (*anti* form). At room temperature, the *gauche* and *anti* rotamers (energy difference of 0.9 kcal/mol) are in conformational equilibrium accounting for a 30:70 ratio, respectively, and so, both the conformations must be considered when the physical and chemical properties are studied.)

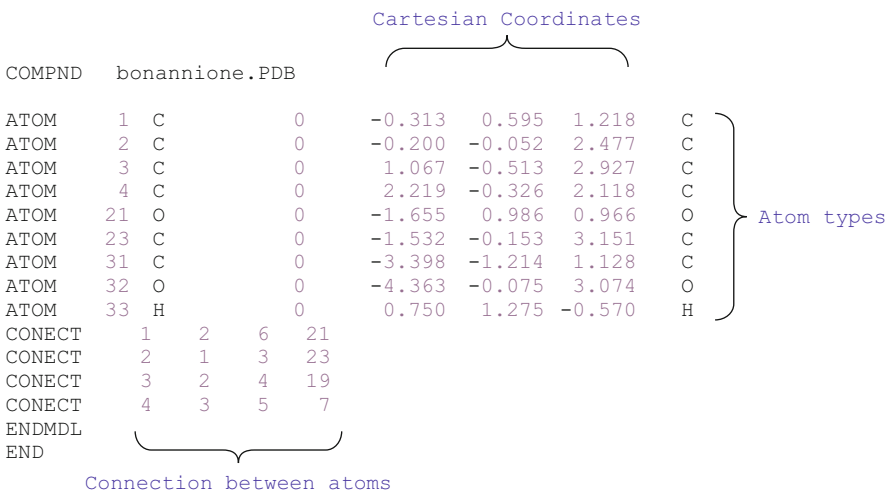

Fig. 10.4 The PDB file representation of the bonannione B

Consequently, the conformational search and the analysis of relative stability of different conformers are mandatory to obtain significative results after the QM calculation of NMR parameters.

This fundamental step for the search of the most representative rotamers can be performed in several ways (systematic search or grid search, random search or Monte Carlo methods, and molecular dynamics) by different dedicated softwares (Discover [16], Insight Package [17, 18], Gromacs [19, 20], Spartan [21], Macromodel [22], etc.).

10.2.2.1 Conformational Analysis Methods

The systematic search (grid search) consists in the step-by-step changes for the all possible dihedral angles of the entire and/or a part of molecule for all 360°, taking in account all the possible combinations between them. Each generated conformation must be subsequently minimized. One simple example of this conformational search through this algorithm can be the analysis of the poly-L-alanine peptide (Fig. 10.6).

Considering planar and in *trans* conformation the amide bonds, and the bonds' angles and distances fixed, the only two variables that regulate the different conformations are represented by φ and ψ angles (Fig. 10.6). The systematic increments of 15–30° for both the dihedral angles in the region between 180° and −180° result in changes of the conformation and energy of the peptide. The variation of the energy as function of φ and ψ, in fact, can be visualized in the Ramachandran plot (Fig. 10.7), where the systematic angle changes are equivalent to examine all the significant points of a grid.

In the case of poly-L-alanine peptide, in the grid, few minima points are identifiable and so the most of conformations, after a minimization step, often converge toward the same rotamer. The main inconvenience of the systematic

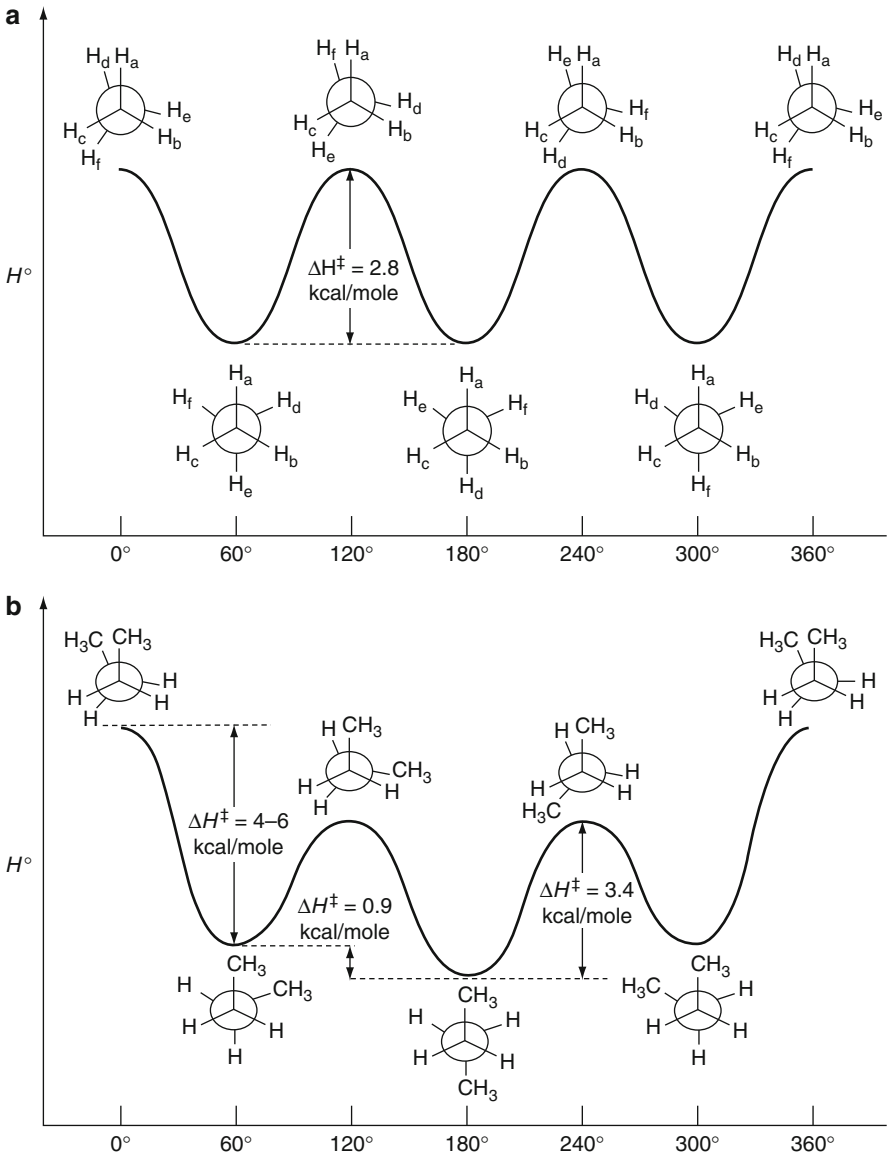

Fig. 10.5 Variation of energy for ethane (**a**) and n-butane (**b**) as function of dihedral angle

search is represented by the number of the generated conformations that exponentially increase with the number of rotatable bonds. This method is also difficult to apply to cyclic compounds where the dihedral angles are not independent between each other. On the other hand, the application of the grid search provides a comprehensive exploration of the conformational space, differently from the methods that will be discussed further.

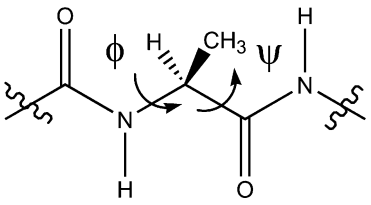

Fig. 10.6 Molecular structure of poly-L-alanine peptide

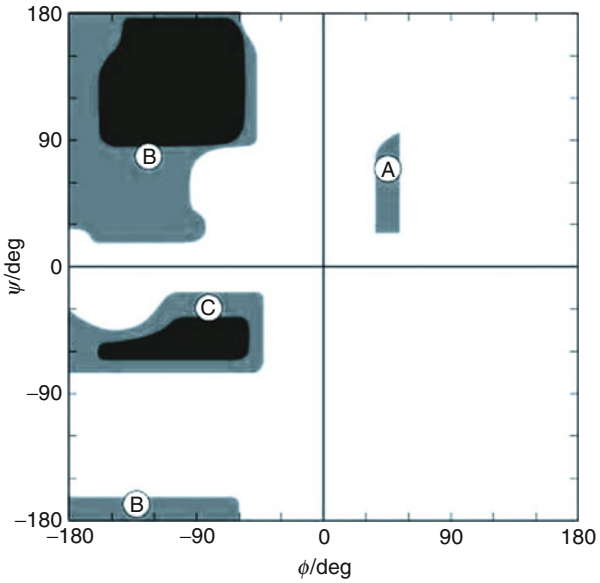

Fig. 10.7 A Ramachandran plot for poly-L-alanine peptide. The *black*-and-*gray* areas (numbered **a**–**c**) represent the allowed conformations; the white regions represent the forbidden rotamers

Differently from the systematic search described above, by the Monte Carlo method, the conformational space is explored in a random mode. The conformations are achieved by varying randomly the torsional angles, and then minimizing the so-obtained casual conformations. In particular, the new conformation is generated in two possible ways: varying or the Cartesian coordinates of each atom, or the dihedral angles of random amount. The resulting conformation is minimized, and its energy is compared against those of all structures previously found, and, if unique, it is stored (criterion based on an energy cut off). After this cycle, another conformation is chosen, and the proceeding is resumed. In order to compensate intrinsic limits of the random search, it is appropriate to perform several parallel runs starting from diverse rotamers. In addition to the casual choice, different approaches exist to select the starting structure, and one of these methods consists in using the local minimum found for each search. The main advantage for the Monte Carlo method is the possibility to sample cyclic systems that are difficult to treat by systematic methods. Another advantage is the potential treatment of molecular systems of any size, even though very large and flexible molecules may not provide converging results due to their wide conformational space.

The molecular dynamics (MD) represents another efficient method for the exploration of the molecular conformational space. It may be used for linear compounds with multiple torsions, when both the systematic and the random search are unadvisable, or for cyclic compounds, when the conformational analysis by the random search is complicated by the required computational time. The fundamental aim of the MD is the reproduction of the real behavior of the molecule, mimicking the time-dependent motions of the atoms. The generation of the new structure represents the principal difference of this approach with respect to the other two described above because the molecular dynamics is a deterministic approach, where the new generated structure is causally determined by previous conformer. In particular, given a set of initial coordinates and velocities with a force (F_i), the evolution of the system, all atoms will be moved to new positions following the Newton's second law (Eq. 10.1)

$$F_i(t) = m_i a_i(t) \tag{10.1}$$

where F is the force on atom i at time t, m_i is the mass of atom i, and a_i is the acceleration of atom i at time t. The so obtained conformation will be used as starting point for a new step, the cycle will be repeated for a predefined number of times, and the conformers will be collected (conformational ensemble). In particular, since significant conformational changes do not always happen at room temperature, especially when more atoms are involved, the MD simulations are performed at different increasing temperatures for supplying a sufficient kinetic energy to overcome the energetic barriers between the possible rotamers. During the MD simulations, in order to increase the probability to explore the entire conformational space, it is convenient to vary some fundamental parameters during the calculation as for example the simulation temperature, the time step, the equilibration, and the simulation time. However, if converged results are desired, different parallel MD simulations must be performed starting from different conformations, at diverse temperature – it is preferable to use a gradient of temperature from 450 to 650/700 K – and long time simulations should be used. Thanks to the kinetic energy of molecules it is possible that some collected conformers represent the local and/or global maximum points, and so, after the MD simulation, a minimization process of this ensemble of rotamers is always necessary.

A particular MD simulation is represented by simulated annealing. This methodology consists in a molecular dynamics at high temperature where temperature is gradually decreased to 0 K. During the simulation at high temperature, the molecules are able to explore conformations very different among each other, and thanks to the temperature dropping the molecule may converge in a local minimum. In order to obtain a collection of low energy conformations, the cycle will be repeated several times. The main advantage of simulated annealing with respect to the classic MD approach is represented by the possibility to avoid the minimization step since the resulting structures represent energy minima.

For all the above considerations, the choice of the method is exclusively dictated by the molecule under study, e.g., if it is linear, or cyclic, how many torsions are

present, if some experimental data are available, etc. In conclusion, it should be kept in mind that each method has its advantages and/or limitations.

In addition to the exploration method, the efficiency of the conformational search depends on a series of other factors such as the force fields and the algorithms of minimization.

In the molecular mechanics, the calculation of the energy is obtained through the force fields, functions that approximate the electronic potential surface using the specific parameters and mathematical equations of classic physics. The most simple force field equation is described in Eq. 10.2:

$$E_{total} = E_{tors} + E_{vdw} + E_{elec} + E_{Hbond} + E_{oop}... \tag{10.2}$$

where the total energy (E_{total}) of the molecule is obtained by the sum of energy terms associated with the internal coordinates. E_{vdw}, for example, represents the energy related to the van der Waals interactions between atoms, and it is described by the Leonard–Jones equation (Eq. 10.3):

$$E_{vdw} = \varepsilon\left[(r_0/r)^{12} - 2(r_0/r)^6\right] \tag{10.3}$$

The electrostatic term (E_{elec}) is calculated by Coulomb's law; E_{tors} is a cosine periodic function (Eq. 10.4), etc.

$$E_{tors} = k\omega(1 - \cos 3\omega) \tag{10.4}$$

The parameter sets of the force field, e.g., the length and angles of the bonds, the hybridization of the atoms, the charges of the atoms, derive from the experimental data, and so the different force fields are tailored for a particular molecular system and are not able to describe all types of structures under investigation. For example, MM2 [23] is parameterized for the organic molecules, MM3 [24] for conjugated systems, AMBER [25, 26] for the peptides and nucleic acids, and OPLS* [27, 28] for peptides.

In this context, it is clear that the obtained geometries by specific method of conformational search must be minimized. Also in this step, different methods exist for the minimization process, but the most used algorithms are the steepest descent, conjugate gradient, and Newton–Raphson. These utilize the first and second derivative values of energy function in order to choose the direction to undertake. The steepest descent algorithm uses the first derivative value to move the atoms toward the local minimum; this algorithm works very well with conformation far from the equilibrium, but it is less efficient in the conformations close to the energy minimum. The steepest descent direction is also considered in the conjugate gradient algorithm, but in this case, the new direction is dependent from the previous; this permits a dramatic improvement of the efficiency of the method in proximity to the local minima points. The Newton–Raphson algorithm is able to discern between the minima and the maxima points, considering also the value of the second derivative.

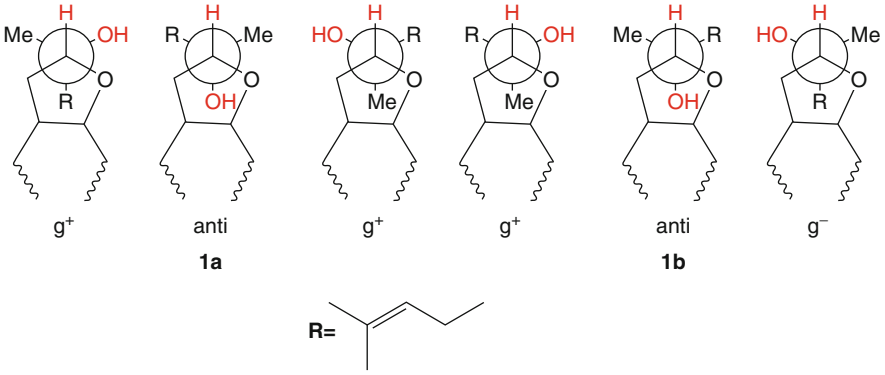

Fig. 10.8 The gauche (g^+ and g^-) and anti rotamers around C-2″–C-3″ for compounds **1a** and **1b**. The gauche and anti positions are referred to the –OH group

Moreover, in addition to the gradient, the curvature of the function allows to identify the search direction. The only disadvantage of this method is represented by the computational resource required (N^2 for N internal coordinates), but it is efficient enough in the conformation close to the minima points. On consequence, for small/medium molecules far from the minimum, it is preferable to apply as first the steepest descent and then the Newton–Raphson method; on the other hand, for biggest molecules, the Newton–Raphson method can be replaced by the conjugate gradient.

Taking into account the above-reported arguments, in the real illustrated example of bonannione B (**1a** and **1b**, Fig. 10.2), the several free molecular dynamics calculations are performed using the MMFFs [29] force field at different temperatures (between 400 K and 650 K) for 10 ns (simulation time), with a time step and an equilibration time of 1.5 fs and 1.0 ps, respectively (MacroModel software package) [22]. All the structures so obtained (numbering 100) are minimized by using the Polak–Ribier conjugate gradient algorithm [30] (PRCG, 1,000 steps, maximum derivative less than 0.05 kcal/mol). This approach leads to the selection of the lowest energy minimum rotamers (g^+, *anti* and g^-) around the C-2″/C-3″ bond for each diastereoisomers (Fig. 10.8) that can be submitted to the next step: the final energy and geometry optimization.

10.2.3 Final Energy and Geometry Optimization

Following the scheme of the protocol reported in Fig. 10.1, after the conformational search and the preliminary geometry optimization of all found conformers, a further refinement of the three-dimensional coordinates is performed by using a more sophisticated theoretical level. As described above, the force fields are based on classical mechanics laws, and they include, in the description of a given chemical system, parameters derived from experiments or from quantum mechanical calculations. Molecular mechanics methods do not explicitly model the electrons in the

description of a molecular system, but the electrons are treated in the force field by parameterization. This approximation renders the molecular mechanics approaches faster than other more sophisticated theories and allows its application to very large systems. For a well-parameterized general force field, many molecular properties, such as geometry and relative conformational energies, can be calculated with high accuracy for a broad chemical space. Otherwise, where a force field is not able to describe well certain chemical species, the use of more sophisticated theories, such as quantum mechanics, is needed. It is noteworthy that the prediction of molecular properties, where the electron effects are predominant, such as transition state, the formation or breaking of a bond, and spectroscopic parameters, cannot be afforded by molecular mechanics. The quantum mechanics explicitly model the electron system of a molecule, and thus it can be used to study chemical issues.

The geometry optimization is a crucial step to study chemical problems, such as transition state, or to predict spectroscopic properties, such as NMR parameters, IR, and so on. In our case, the aim is to predict chemical shifts or coupling constants for the relative configuration assignment of organic compounds [1, 2, 31]. Moreover, the treatment of the investigated structure by quantum mechanical theory allows to get more accurate energy values for each considered conformers. These energy values are useful to calculate the Boltzmann distribution, which is a function of the energy (Eq. 10.11, see below). Taking into account the Boltzmann distribution is important to forecast molecular properties of flexible chemical systems because the predicted values result from the energy-weighted contribution of each conformer in solution.

Following the general protocol, the conformers' geometry obtained and optimized by using molecular mechanics force fields should be refined through semiempirical methods [32] before the application of quantum theory. The semiempirical methods differ from molecular mechanics because their theory combines the Schrödinger equation and some experimental derived parameters to simplify the computation. The most popular methods applied to organic compounds are AM1 [8] and PM3 [9]. The semiempirical methods present the advantage of being faster than quantum chemical methods, but it is not always that they can correctly predict a molecular property. This is due to the fact that the experimental derived parameters are not representative of all chemical systems. Compared to the quantum mechanical calculation of NMR parameters, the geometry optimization is more time-consuming. Thus, the aim to run a geometry optimization by semiempirical methods is to speed up the following computation at more time-consuming quantum mechanical theory level. Nowadays, thanks to the development of more potent computers this step is not necessary. Indeed, also by a desktop computer, it is possible to optimize the geometry with a modest theoretical level in a reasonable time.

The quantum mechanics is not based on classical mechanics laws and on the use of experimental parameters; indeed, the molecular properties are calculated by solving the Schrödinger equation:

$$\hat{\mathrm{H}}\Psi = \hat{\mathrm{H}}\mathrm{E} \tag{10.5}$$

where $\hat{H}$ is the Hamiltonian operator, Ψ is the wave function, and E is the energy. The wave function Ψ is a mathematical function of the electron and nuclear positions, and it represents a probabilistic description of electron behavior. In details, it can describe the probability of electrons being in certain locations, but it cannot predict exactly where electrons are located. As the quantum mechanics can describe mathematically the correct behavior of the electrons, it is possible to predict chemical properties of molecular systems, by applying relative mathematical operator corresponding to a particular chemical physical observable. For example, in Eq. 10.5, the Hamiltonian operator allows to calculate the energy of the electron, whereas the application of the spin-Hamiltonian can predict NMR parameters. (In order to obtain a physically relevant solution of the Schrödinger equation, the wave function must be continuous, single-valued, normalizable, and antisymmetric with respect to the interchange of electrons.

The Hamiltonian operator $\hat{H}$ is, in general,

$$\hat{H} = -\sum_{i}^{\text{particles}} \frac{\nabla_i^2}{2m_i} + \sum_{i<j}^{\text{particles}} \sum \frac{q_i q_j}{r_{ij}} \tag{10.6}$$

where ∇_i^2 is the Laplacian operator acting on particle i. Particles indicate electrons and nuclei. The symbols m_i and q_i are the mass and charge of particle i, and r_{ij} is the distance between particles. The first term represents the kinetic energy of the particle, and the second term is the energy due to Coulomb interactions of particles. This formulation is the time-independent, nonrelativistic Schrödinger equation. Additional terms can appear in the Hamiltonian when relativity or interactions with electromagnetic radiation or fields are taken into account.)

Only in few simple cases the Schrödinger equation can be exactly solved, such as particle in the box, harmonic oscillator. The solution of this fundamental equation requires the use of mathematical approximations, and different quantum theories have been developed to this aim. The starting point of all quantum mechanical methods is the Born–Oppenheimer approximation (Eq. 10.6). This approximation treats the electrons and nuclei motions as separated and considers the nuclei fixed, thus neglecting the kinetic contribution of the nuclei.

$$\hat{H} = -\sum_{i}^{\text{electrons}} \frac{\nabla_i^2}{2} - \sum_{i}^{\text{nuclei}} \sum_{j}^{\text{electrons}} \frac{Z_i}{r_{ij}} + \sum_{i>j}^{\text{electrons}} \sum \frac{1}{r_{ij}} \tag{10.7}$$

Here, the first term is the kinetic energy of the electrons only, the second one is the attraction of electrons to nuclei, and the third term is the repulsion between electrons called correlation. The repulsion between nuclei is added onto the energy at the end of the calculation. The motion of nuclei can be described by considering this entire formulation to be a potential energy surface on which nuclei move.

One of the most applied theories is the Hartree–Fock (HF) [32], which is based on the central field approximation. It considers the Coulomb repulsion between two electrons by integrating the repulsion term. This approximation does not give

the exact effect of the electrons repulsion (correlation) but only an average of this repulsive interaction, and it represents an important limitation of HF application. This effect is important because including correlation generally improves the accuracy of computed energies and molecular geometries, as well as NMR parameters [1, 2, 31]. The wave function results from a linear combination of atomic orbitals, which are described by a set of functions called basis set.

One of the most used post-HF methods is the Møller–Plesset [33], which accounts the correlation by adding to the Hartree–Fock wave function the perturbation term. Different Møller–Plesset methods have been developed and the mostly used is the second order (MP2) [34]. Compared to HF theory, the Møller–Plesset methods are more accurate but very computationally expensive; thus, their application is limited to small molecular systems.

In the last years, the density functional theory (DFT) [35, 36] has been gaining a great success due to the lower demanding computational costs but with results accuracy similar to the post-HF methods, such as Møller–Plesset. DFT considers that the energy of a molecule can be determined from the electron density instead of the wave function. The electron density is expressed as a linear combination of basis functions similar in mathematical form to HF orbitals. In particular, a determinant is then formed from these functions, obtaining the Kohn–Sham orbitals. From the electron density of these orbitals, the energy is calculated applying a density functional. Different functionals [37, 38] have been developed, and between them, the most used is B3LYP in the geometry optimization of investigated compounds. The MPW1PW91 [35, 36] has shown to produce adequate structural geometry, and many application of this functional are reported in literature [1, 2, 31]. Recently, new promising functionals have been introduced: M05 [39] and M05-2X [40, 41]. The DFT takes into account the electron correlation, giving better results than the HF-based methods [1, 2, 31].

It is noteworthy that independently from the QM theory applied, the choice of the basis set is important in the prediction of the molecular properties. The basis set should be large enough, for example, to take into account the electron correlation effects or to describe the molecular charge distribution. At the same time, the basis set should be small enough to be applied to the investigated molecule, especially for larger chemical systems. In our contribution, where different methods have been compared in the prediction of chemical shift of organic compounds, we observed that the 6–31 G with DFT is the right trade-off between accuracy and rapidity of calculation [31].

Different softwares dedicated to the quantum mechanical (QM) calculation of molecular properties are available: Gaussian [42], HyperChem [43], Jaguar [44], and Spartan [21]. These softwares allow the application of different theoretical methods to carry out the forecast of different molecular properties.

Let us continue with the case study of the determination of the relative configuration of bonannione B (**1**), to show the application of QM approach for geometry refinement and the prediction of ^{13}C NMR chemical shift (Sect. 10.2.4). From the conformational search, three main conformers are found for two considered stereostructural hypothesis, and all of them have to be treated by the software

Gaussian 03 [45]. For all six staggered rotamers, six input files are prepared. Each file contains the coordinates derived from molecular mechanics level and the instruction for the optimization of the geometry.

In this case, the geometry refinement is performed by applying the DFT theory, the B3LYP as functional, and 6–31 G(d) basis set. After the calculation, a new geometry for each staggered rotamer of **1a** and **1b** is obtained, and it is used for the calculation of the chemical shifts (Sect. 10.2.4). Besides the new spatial arrangements of the conformers, associated energy values are achieved by the computation. This value is expressed in Hartree, and it can be converted in kJ: one Hartree is 2625.5 kJ mol^{-1}.

10.2.4 ^{13}C NMR Chemical Shifts Calculation

The nuclear magnetic resonance (NMR) has been playing a crucial role in solving molecular structural aspects thanks to the ability of some NMR parameters (coupling constant, chemical shift) to provide fundamental information on the configurational and conformational arrangement of organic molecules. The configuration of cyclic compounds with three- to six-membered rings, and of compounds with predictable conformational behavior, can be easily determined by analyzing NMR parameters, such as proton–proton *J* coupling values, chemical shifts, and/or nuclear Overhauser effect intensities. The relative configurational assignment of flexible systems, such as polysubstituted open chains and macrocycles, is more difficult to study, due to geometrical uncertainty associated with these types of molecules.

In the last years, great advances have been made in developing quantum mechanical (QM) methods of chemical interest able to predict molecular properties [1, 2, 32]. In particular, quantum mechanical calculation of NMR parameters has been used as an emerging strategy for the assignment of the relative configuration of flexible organic molecules on the basis of the high accuracy in the reproduction of the experimental NMR properties also achieved at a low level of theory [1, 2, 31, 46].

The final step of the protocol described in this chapter is the calculation of an NMR parameter: ^{1}H, ^{13}C, $^{2,3}J_{H\text{-}H}$, and $^{2,3}J_{C\text{-}H}$.

Different approaches have been devised to calculate the chemical shift, such as IGLO [47], LORG [48], and CSGT [49], but the most popular is GIAO (gauge including atomic orbital) [50, 51], which is based on the perturbation theory, resulting suitable for HF or DFT theories. In particular, GIAO has shown to provide more reliable results compared with other methods at the same basis set [1, 2, 31, 32, 50, 51]. The electron correlation effect is not negligible in the chemical shift calculation, especially for ^{13}C; in fact, DFT or MP theories have shown to give more accurate results than HF approach [31, 52]. In this field, DFT is successfully emerging due to its accuracy in NMR properties prediction and its ability to handle large molecular systems not easily treatable by post-HF methods [1, 2, 52]. For this reason, DFT is widely used in the calculation of NMR parameters

[1, 2]. The most popular functionals applied in the calculations of NMR parameters are B3LYP, MPW1PW91, and PBE1PBE [53, 54].

It is also important that sufficiently large basis sets are used [31]. Of course, the choice of the basis set to use is based on the accuracy and on the size of the molecule under study. The 6-31 G (d) basis set should be considered the absolute minimum for reliable results, but better outcomes are obtained by using 6-31 G(d,p) [31]. The choice of the basis set, thus, depends both on the accuracy and on the size of the molecule and also on the number of the chemical species to take into account.

Continuing with the reported case study, all conformers for both the diastereoisomers under investigation have been optimized by using DFT/B3LYP approach. The new coordinates, derived by quantum mechanical optimization, are now used to calculate the ^{13}C chemical shifts by GIAO method. Thus, six input files are set up, indicating the instruction for the chemical shift calculation. The theory level to apply is the same of that one used in the geometry optimization (DFT/B3LYP), but the basis set is 6-31 G(d,p).

In particular, the quantum mechanical methods calculate the shielding tensor and not directly the chemical shift. One of the mostly used methods to obtain the chemical shift is to subtract the isotropic shielding value of TMS from the shielding tensor of the conformer. This can be easily done by collecting the data in a spreadsheet and subtracting the isotropic shielding tensors of the conformer and TMS in absolute value. Of course, the shielding tensor of TMS has to be calculated at the chosen theoretical level for the investigated compound. We observed for sp^3 carbons referenced to TMS a good fitting between experimental and calculated, whereas for the sp^2 carbons, this agreement is lower [31]. For this reason, the sp^2 carbons could not be considered in the comparison with the experimental chemical shifts. Recently, the multi-standard approach (MSTD) has been reported [55] which is demonstrated to perform better than the application of TMS as reference compound. The MSTD consists on the use of two different reference compounds and calculates the chemical shift through the following equation:

$$\delta^{i}{}_{\text{calc}} = \sigma_{\text{ref}} + \sigma_i + \delta^{\text{ref}}{}_{\text{exp}} \tag{10.8}$$

where σ_{ref} and σ_i are the shielding tensors of atom i and the reference compound, respectively, and the $\delta_{exp}{}^{ref}$ is the experimental value of the reference. In particular, the calculation of sp^3 carbon atoms is referenced to methanol, whereas the chemical shifts of sp^2 and sp carbon are calculated from benzene.

Another method is based on linear regression analysis. In particular, calculated chemical shift of a collection of organic compounds is plotted against their experimental values. The intercept and the slope of the obtained straight line can be used to scale the calculated values:

$$\delta_{\text{scaled}} = \frac{\delta_{\text{calc}} - \text{intercept}}{\text{slope}} \tag{10.9}$$

In the reported case study, the chemical shifts are referenced to TMS. It is important to keep in mind that the experimental chemical shift is a weighted value from the contribution of conformers in equilibrium among them:

$$\delta_{\text{exp}} = \delta_i p_i + \delta_j p_j + \ldots \tag{10.10}$$

where p_i is the population fraction of N_i molecules in a determined conformation and δ_i the associated chemical shift. The same is for the number of molecules j and all other conformers in solution. The number of molecules in a specific three-dimensional arrangement depends on the energy of that conformation. The ratio of the number of conformers N_i with energy E_i, to the number of molecules in state j is given by the Boltzmann distribution:

$$\frac{N_i}{N_j} = e^{-\left(E_i - E_j\right)/k_{\text{B}}T} \tag{10.11}$$

where k_{B} is the Boltzmann constant (1.38066×10^{-23} J/K) and T the temperature. Following the Eq. 10.11, two conformers differing for 3 kcal/mol ($\approx$ 10 kJ/mol) in energy present a ratio of 99:1. Thus, each conformation differing from the global minimum by $\geq$ of 3 kcal/mol does not effectively contribute to the averaged chemical shift and can be neglected in the calculation. From the computed energy values for each conformer at quantum theory level, it is possible to calculate the population fractions contributing to the weighted chemical shift, by using the following equation:

$$\delta = \sum_i \left[\delta_i \exp(-E_i/RT) / \sum_i \exp(-E_i/RT) \right] \tag{10.12}$$

where δ_{i} is the chemical shift of conformer i, R is the molar gas constant (8.3145 J K^{-1} mol^{-1}), T is the temperature (298 K), and E_i is the calculated energy of conformer i relative to the energy of the global minimum.

In the output file, along with isotropic shielding tensors, an energy value at the basis set 6-31 G(d,p) is obtained and is taken into account in the Boltzmann distribution calculation.

From the calculated energies, it is obtained for isomer **1a** that the *anti* and g^+ conformers represent 70% and 28% of the total population, respectively, due to the hydrogen bond between the OH group and the oxygen atom of the ring for both these rotamers. The remaining 2% is attributed to the g^- conformer. A similar trend of the energies and the population distribution is observed for **1b**, where the *anti* rotamer accounts for 71% and the g^+ rotamer for 28%, and 1% is represented by the rotamer g^- (Table 10.1)

The calculated chemical shifts, according to the Boltzmann distribution, are compared with the experimental ones in order to find which of the two considered diastereoisomers (**1a** and **1b**) presents the best fitting with the experimental set of

Table 10.1 Comparison of calculated for stereoisomers **1a** and **1b** *vs*. experimental ^{13}C NMR chemical shifts in $CDCl_3$

^{13}C Chemical shifts			
Carbons	**1a**	**1b**	Experimental
2	81.4	81.5	79.0
3	46.2	46.1	42.7
4	187.5	187.5	196.3
5	155.4	155.4	158.3
6	103.2	103.2	105.8
7	161.7	161.7	169.5
8	86.3	86.4	90.5
9	158.4	158.4	163.6
10	101.9	101.9	103.1
1′	127.2	127.2	129.6
2′	122.4	122.5	127.6
3′	107.4	107.4	115.8
4′	150.5	150.5	156.8
5′	111.0	110.9	115.8
6′	123.2	123.3	127.6
1″	29.1	29.5	25.9
2″	94.0	94.0	91.3
3″	73.9	73.9	73.8
4″	**23.1**	**19.7**	22.4
5″	**37.3**	**40.6**	36.6
6″	24.8	25.0	21.8
7″	122.1	122.5	123.7
8″	128.7	128.0	132.0
9″	26.5	26.4	25.4
10″	17.5	17.6	17.5

resonances. The comparison is made, calculating the difference between the predicted and experimental values (Table 10.1). Other parameters could be taken into consideration to understand which hypothesized structure best reproduce the experimental. The mean absolute error is a statistical parameter obtained from the following equation:

$$\mathrm{MAE} = \sum [|(\delta_{\mathrm{exp}} - \delta_{\mathrm{calcd}})|]/n \tag{10.13}$$

where $\delta_{\mathrm{exp}} - \delta_{\mathrm{calcd}}$ are respectively the experimental and calculated chemical shifts, and n is the total number of considered resonances. A variant of the MAE is the corrected mean absolute error (CMAE), where the calculated values are scaled with the experimental chemical shifts. In detail, the predicted set of resonances are plotted in function of the experimental values, and the intercept and the slope obtained by linear regression are used to correct the forecast chemical shifts, by using Eq. 10.9. From the scaled values, the MAE is calculated obtaining the CMAE.

$$\mathrm{CMAE} = \sum[|(\delta_{\mathrm{exp}} - \delta_{\mathrm{scaled}})|]/n \tag{10.14}$$

The correlation coefficient, r, is also used as parameter to compare calculated and experimental chemical shifts. Recently, Smith and Goodman have introduced the DP4 probability which performs better than the mean absolute error and correlation coefficient [56].

It is possible to observe that the resonances of almost all carbon atoms for both diastereoisomers (**1a** and **1b**) are very similar to the experimental chemical shift, differing by a maximum of 0.2 ppm. Large differences are found around the couple of stereocenters under study, in particular, for the C-4″ and C-5″, which are bound to C-3″ (Table 10.1). The predicted values of C-4″ are 23.1 ppm and 19.7 ppm for **1a** and **1b**, respectively. Concerning the C-5″, the **1a** resonance is predicted at $\delta = 37.3$ ppm and at $\delta = 40.6$ ppm for **1b** (Table 10.1). Isomer **1a** shows a variation from the experimental of 0.7 ppm for C-4″ and C-5″, whereas for **1b** larger $\Delta\delta$ values are found: 2.7 ppm and 4.0 ppm for C-4″ and C-5″, respectively. It is noteworthy that the obtained discrepancies reflect the different magnetic environment of the methyl group in the 4″-position and the methylene group in the 5″-position in **1a** and **1b** due to the different configuration of C-3″. Thus, the ^{13}C chemical shift calculation is a useful tool to probe the different 3D spatial arrangement of stereogenic carbon substituents, suggesting the correct relative configuration of the investigated compound. The relative configuration of bonannione B can be assigned as **1a**.

10.2.5 *J* Coupling Constants Calculation

In the previous section, it has been shown that ^{13}C NMR chemical shifts have been predicted by quantum mechanics method to assign the relative configuration of the compound under study. The chemical shift is not the only NMR parameter that can be calculated by QM approach and used in the stereostructure analysis. Homo and heteronuclear coupling constants can be predicted, and they could be integrated with experimental values to shed light on the relative configuration. The relative configuration assignment of the natural product callipeltin A (**2**, Fig. 10.9), isolated from the sponges *Callipelta* sp. and *Latrunculia* sp. [57], is reported here as an example. In particular, the compound is a peptide (Fig. 10.9) constituted by nine amino acids, but only the following portion has been determined by QM-*J* method: the two units (named $\mathrm{AGDHE}_{2,3}$ and $\mathrm{AGDHE}_{3,4}$) contained in the AGDHE fragment, the two threonine residues (named D-*a*Thr1 and D-*a*Thr2), and the β-OMeTyr amino acid.

Behind the use of *J* coupling constants to assign the relative configuration, through this case study, it is shown how to afford the analysis of large molecule presenting more than one couple of stereocenters. In detail, the strategy consists in dividing the entire molecule in small C_2 fragments, following a rational building: at least two heavy groups (carbon, oxygen atoms) substitute the main chain, and every branched chain is replaced by at least one heavy atom [58]. For each C_2 fragment,

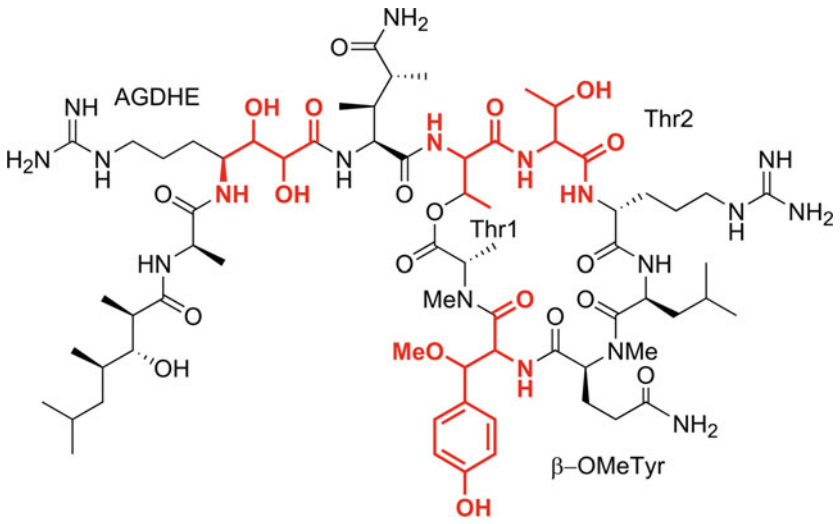

Fig. 10.9 Callipeltin A (**2**) with amino acid residues still stereochemically undetermined represented in red

the *erythro* and *threo* configuration are considered, and for each diastereoisomer, three main staggered rotamers are built (*anti*, g^+ and g^-) (Fig. 10.10).

As the analysis is based on simple molecular fragments derived from the whole molecule, directly the main three staggered rotamers are built and considered in the calculations. The basic idea to simplify the entire structure with more than one couple of stereocenters is to limit the number of conformers to analyze at the quantum mechanical level. Indeed, the number of combinations of all staggered conformers from C_2 arrangements is $N_r = 6^n$, where n is the number of pairs of stereocenters [58]. The simplification is possible because local atomic environment mainly affects the coupling constants. Usually, further than two atoms away from the nuclei involved in the scalar coupling, the effects are negligible [58].

Let us apply this strategy to the callipeltin A. The following C_2 fragments are built and investigated (Fig. 10.11):

For sake of simplicity, only the analysis of D-aThr1 fragment is described, but the approach is repeated for all the investigated simplified systems. The *anti* and two *gauche* rotamers are built for *erythro* and *threo*, and the geometry optimization is performed at DFT theory level by using the functional MPW1PW91 and the basis set 6-31 G(d). The calculations are performed by taking into account the contributions of the following interactions: Fermi contact (FC), paramagnetic spin–orbit (PSO), diamagnetic spin–orbit (DSO), and spin–dipole (SD). As made for the bonannione B (see previous section), on the optimized geometry at quantum mechanics level, the calculation of the NMR parameter is performed. In this case, the prediction concerns the $^{2,3}J_{C\text{-}H}$ and $^{2,3}J_{H\text{-}H}$ through the same functional of the optimization step (MPW1PW91) and by the 6-31 G(d,p) basis set. Both the quantum mechanical calculation steps (optimization and NMR property prediction)

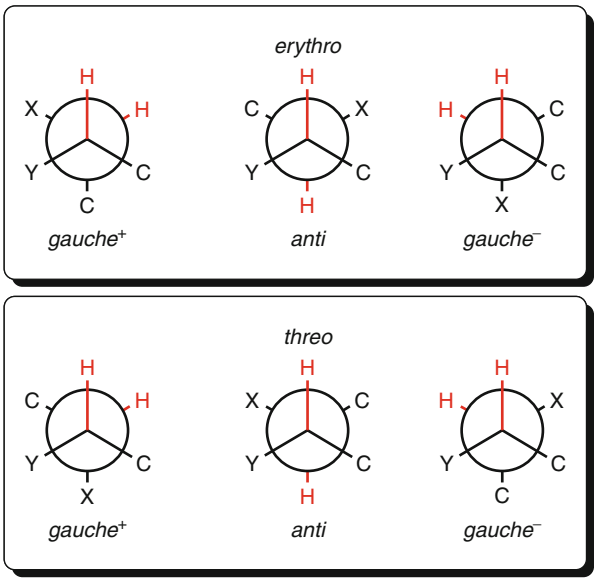

Fig. 10.10 Main staggered rotamers (*anti*, g^+, and g^-) for each relative stereochemical arrangement *erythro* and *threo*

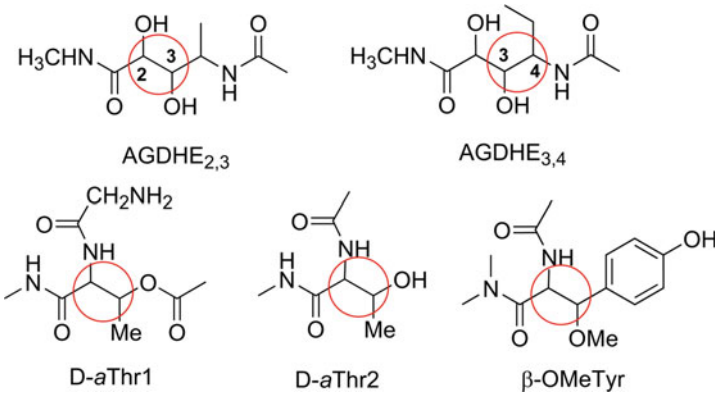

Fig. 10.11 Molecular structures of C_2 fragments from callipeltin A, considered in the QM calculations. The couples of the investigated stereocenters are indicated by red circles

are made by using the IEF-PCM solvent continuum model, to mimic the presence of the methanol. The choice to use the solvent is due to the presence of hydrogen donors and acceptors, which influence the conformational arrangement in absence of the solvent and could cause misleading information for the assignment of relative configuration.

The set of calculated coupling constants for all staggered rotamers of *erythro* and *threo* is compared with the experimental one (Table 10.2). The comparison is made through the total absolute deviation (TAD) values: $\Sigma|J_{calc}-J_{exp}|$. The conformer

Table 10.2 Calculated *J* values for the six conformational arrangements belonging to *erythro* and *threo* series of C_2 fragments from callipeltin A, in comparison with the experimental data: single deviations and TAD ($\Sigma|J_{calc}-J_{exp}|$) values are reported

	Calc						Exp
	Erythro			*Threo*			
	g^+	*Anti*	g^-	g^+	*Anti*	g^-	
D-aThr1							
$^3J_{H2\text{-}H3}$	3.7	8.6	2.6	4.5	9.2	1.3	2.8
$^2J_{H2\text{-}C3}$	0.9	−4.4	−4.9	−4.5	−4.2	0.6	−5.7
$^3J_{H2\text{-}Me}$	0.7	2.3	4.0	0.3	2.7	1.9	5.0
$^2J_{H3\text{-}C2}$	−1.4	−2.2	0.3	−0.6	−1.8	0.6	−1.4
$^3J_{H3\text{-}C=O}$	7.2	2.4	1.4	7.4	1.0	2.1	2.0
TAD	17.0	11.0	4.3	13.8	11.6	13.0	
AGDHE$_{2,3}$							
$^3J_{H2\text{-}H3}$	4.6	8.0	3.2	1.8	5.9	6.6	9.0
$^2J_{H2\text{-}C3}$	−3.0	−3.7	−0.3	2.2	−2.7	−3.0	−4.0
$^3J_{H2\text{-}C4}$	6.7	2.3	0.3	0.4	4.6	5.3	2.6
$^2J_{H3\text{-}C2}$	1.9	−0.8	−1.3	3.8	0.3	−2.4	−3.5
$^3J_{H3\text{-}C=O}$	0.1	1.8	6.6	0.2	4.0	7.1	1.2
TAD	16.0	4.9	19.4	23.9	13	13.1	
AGDHE$_{3,4}$							
$^3J_{H3\text{-}H4}$	3.3	8.7	3.1	4.0	5.0	0.9	1.7
$^3J_{H3\text{-}C5}$	0.4	1.5	4.9	3.8	0.0	2.6	1.7
$^2J_{H4\text{-}C3}$	−4.3	−4.4	0.5	−3.9	−5.3	0.7	1.2
$^3J_{H4\text{-}C2}$	6.2	4.0	0.6	5.9	0.1	2.5	1.2
TAD	13.4	15.6	5.9	14.2	12.6	3.5	
D-aThr2							
$^3J_{H2\text{-}H3}$	3.7	8.6	2.8	4.6	7.7	1.7	3.5
$^2J_{H2\text{-}C3}$	−0.4	−4.9	−4.9	−4.2	−3.6	−1.5	−4.1
$^3J_{H2\text{-}Me}$	0.5	1.9	4.1	4.3	4.4	1.6	1.1
$^2J_{H3\text{-}C2}$	−1.6	−0.3	2.4	−1.2	1.1	0.9	−2.8
$^3J_{H3\text{-}C=O}$	7.3	3.2	0.7	7.0	1.5	2.0	7.3
TAD	5.7	13.2	16.2	6.2	17.6	13.8	
β-OMeTyr							
$^3J_{H2\text{-}H3}$	1.5	8.7	3.9	8.3	10.7	3.7	9.1
$^2J_{H2\text{-}C3}$	−5.4	−3.2	0.0	−3.4	−4.3	−3.4	−3.9
$^3J_{H2\text{-}Ph}$	3.8	2.5	0.7	5.6	1.8	1.1	1.3
$^2J_{H3\text{-}C2}$	−0.8	−2.6	−2.7	−1.2	−2.3	0.4	−3.2
$^3J_{H3\text{-}C=O}$	1.4	2.0	6.4	6.6	1.6	1.5	1.7
TAD	14.3	3.2	14.9	12.5	3.5	9.1	

showing the lowest difference with the experimental values represents the right relative configuration. Concerning the D-aThr1 residue, the lowest sum of absolute errors (4.3 Hz) is observed for g^- *erythro* arrangement, whereas the TAD for all other conformers ranges between 11 and 17 Hz (Table 10.2).

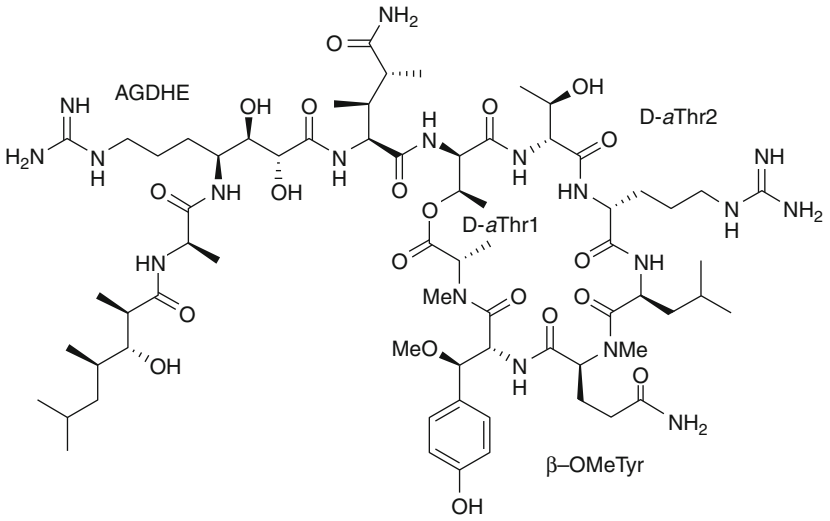

Fig. 10.12 Molecular structure of callipeltin A, with all residues stereostructurally determined

The same analysis is made for the other considered C_2 fragments. In particular, for the residue $AGDHE_{3,4}$, the best agreement between calculated and experimental is found for the g^- *threo* arrangement, characterized by a total deviation of only 3.5 Hz (compared to TAD values of 5.9–15.6 Hz for the reminder conformers). Also, for the $AGDHE_{2,3}$ fragment, the *anti erythro* model displays the lowest TAD value of 4.9 Hz, much below the other deviations (13.0–23.9 Hz). The analysis of D-aThr2 and β-OMeTyr residues is more complicated because two possible conformers present similar TAD values. For the D-aThr2, the g^+ of *erythro* presents a deviation of 5.7 vs. 6.2 Hz of g^+ of *threo*. For the β-OMeTyr residue, the *anti* arrangements of *erythro* and *threo* show comparable TAD values: 3.2 and 3.5 Hz, respectively. Thus, for these residues, the investigation is integrated by an analysis of ROESY spectra, which confirms the QM-based results. In details, a strong ROE effect between the H-2 proton and the methyl group confirms the g^+ *erythro* arrangement for the D-*a*Thr2. This observation is also consistent with the small $^3J_{\text{H-2,Me}}$ value and, consequently, with a *gauche* relationship between these groups. Concerning the β-OMeTyr residue, the observed ROE cross-peak between the aromatic and amide protons suggests the *erythro* configuration as representative arrangement. By using the presented strategy, the fully stereostructural determination of callipeltin A has been carried out (Fig. 10.12).

10.2.6 Solvent Effects

Many of the organic compounds are soluble in nonpolar solvents, and the prediction of NMR parameters can be fairly well conducted in vacuum [1–4, 31, 41]. However, NMR parameters are very sensitive to the surrounding environment. Thus, solvents

are not negligible in the calculations, especially for the polar ones [59]. The solvation affects the solute geometry and the electronic structure, which along with solute–solvent interactions induces variation in the NMR properties [60–62].

In general, the methods to model the solvent fall in two main categories: the implicit and explicit solvent models. The first model represents the solvent as continuous medium with a uniform dielectric constant, surrounding the solute molecular cavity. The most commonly used methods are the polarized continuum model (PCM) [63], the conductor-like screening model (COSMO) [64], and the generalized born surface area (GBSA) [65]. The explicit model treats the solvent molecules as discrete entities, and it is mostly applied in molecular dynamics. Both methods present advantages and shortcomings, but they could improve the prediction of NMR parameters where solvent effects are not negligible. As reported in literature [60–62], the solvent effects limited to conformational changes of the investigated chemical system are well described by using continuum models. Where the solute–solvent interactions play a crucial role, the use of explicit solvent molecules give better predictions. One strategy to model the discrete solvent molecules around the solute is to run molecular dynamics or Monte Carlo simulations, in order to obtain different configurations of the investigated compound and the surrounding solvent molecules. On the obtained arrangements, the NMR calculation at quantum mechanical levels is performed. However, it is not trivial to model the orientation and the number of solvent molecules around solute counterparts and not always perform well in prediction of NMR properties. This method, based on molecular mechanics to model compound embedded by the solvent, is not feasible for the configurational assignment due to the large number of chemical systems to treat at quantum mechanical level. To date, there is not a well-defined strategy to account the effects of the solute–solvent interactions in the stereostructural studies.

10.3 Conclusions

The protocol described in this chapter shows how the calculation of NMR parameters at the QM theoretical level is a useful tool for conformational and configurational analysis of natural products. The great success of QM prediction of molecular properties is due to its ability to reproduce the experimental data at affordable computational expense.

In particular, two case studies have been reported regarding the determination of the relative configuration of two natural products: bonannione B isolated from *Bonannia graeca*, and callipeltin A isolated from the sponges *Callipelta* sp. and *Latrunculia* sp. With these two examples, we respectively show the use of ^{13}C, $^{2,3}J_{H\text{-}H}$, and $^{2,3}J_{C\text{-}H}$ values calculation as a tool for the determination of the relative configuration of organic molecules. It is noteworthy that both calculated NMR properties at QM theory level can be combined for the stereostructural investigation depending on the organic compound under examination. Moreover, besides the use of QM calculation of NMR parameters in the conformational and

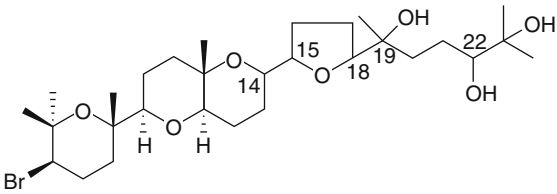

Fig. 10.13 Molecular structure of aplysiol B

configuration analysis, the fast and convenient quantum chemical approach can be applied to lead the total synthesis of complex natural compounds toward the correct stereoisomers, saving time and resources.

It should be highlighted that QM methods have been shown to accurately reproduce experimental molecular properties in vacuum, especially in cases where nonpolar solvents are used for acquiring the experimental spectra. On the other hand, solvent effects limited to conformational changes of the solute can be taken into account by continuum solvent models; in cases where the solute–solvent interactions have a significant weight, an explicit solvent treatment should be applied.

10.4 Study Questions

1. What are the molecular features to be considered in the choice of the force field for the conformational search?
2. Why are quantum mechanical methods required to study chemical problems, such as transition state, or to predict spectroscopic properties?
3. Why DFT or MP may be more accurate than HF in the prediction of NMR parameters?
4. In which cases is DFT approach preferred to MP methods?
5. Try to establish a protocol to determine the relative configuration of the stereocenters C-14, C-15, C-18, C-19, and C-22 of aplysiol B isolated from *Aplysia dactylomela* [66] (Fig. 10.13).
6. What are the main features of implicit and explicit solvent models?
7. What are the criteria to select the appropriate solvent model?
8. Why is it important to get an accurate energy value for each conformer?
9. What is a basis set?

References

1. Di Micco S, Chini MG, Riccio R, Bifulco G (2010) Quantum mechanical calculation of NMR parameters in the stereostructural determination of natural products. Eur J Org Chem 8:1411–1434. doi:10.1002/ejoc.200901255
2. Bifulco G, Dambruoso P, Gomez-Paloma L, Riccio R (2007) Determination of relative configuration in organic compounds by NMR spectroscopy and computational methods. Chem Rev 107:3744–3779. doi:10.1021/cr030733c
3. Barone G, Gomez-Paloma L, Duca D, Silvestri A, Riccio R, Bifulco G (2002) Structure validation of natural products by quantum-mechanical GIAO calculations of ^{13}C NMR

chemical shifts. Chem Eur J 8:3233–3239. doi:10.1002/1521-3765(20020715)8:14 < 3233::AID-CHEM3233 > 3.0.CO;2-0/abstract/
4. Barone G, Duca D, Silvestri A, Gomez-Paloma L, Riccio R, Bifulco G (2002) Determination of the relative stereochemistry of flexible organic compounds by ab initio methods: conformational analysis and Boltzmann-averaged GIAO ^{13}C NMR chemical shifts. Chem Eur J 8:3240–3245. doi:10.1002/1521-3765(20020715)8:14 < 3240::AID-CHEM3240 > 3.0.CO;2-G/abstract
5. van Gunsteren WF, Berendsen HJC (1990) Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry. Angew Chem Int Ed 29:992–1023. doi:10.1002/anie.199009921
6. Höltje HD, Sippl W, Folkers G (2003) Molecular modeling basic principles and applications. Wiley-VCH, Weinheim
7. Chang G, Guida WC, Still WC (1989) An internal-coordinate Monte Carlo method for searching conformational space. J Am Chem Soc 111:4379–4386. doi:10.1021/ja00194a035
8. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. J Am Chem Soc 107:3902–3909. doi:10.1021/ja00299a024
9. Stewart JJP (1989) Optimization of parameters for semiempirical methods I. Method. J Comput Chem 10:209–220. doi:10.1002/jcc.540100208
10. Rosselli S, Bruno M, Maggio A, Bellone G, Formisano C, Mattia CA, Di Micco S, Bifulco G (2007) Two new Flavonoids from Bonannia graeca: a DFT-NMR combined approach in solving structures. Eur J Org Chem 15:2504–2510. doi:10.1002/ejoc.200600969
11. Dale JA, Mosher HS (1973) Nuclear magnetic resonance enantiomer regents. Configurational correlations via nuclear magnetic resonance chemical shifts of diastereomeric mandelate, O-methylmandelate, and.alpha.-methoxy-.alpha.-trifluoromethylphenylacetate (MTPA) esters. J Am Chem Soc 95:512–519. doi:10.1021/ja00783a034
12. (a) Protein Data Bank (PDB) See http://www.rcsb.org/pdb/home/home.do; (b) Nucleic Acid Databank (NDB) See http://ndbserver.rutgers.edu/; (c) Cambridge Structural Database (CSD) See http://www.ccdc.cam.ac.uk/products/csd/; (d) Crystallography Open Database (COD) See http://www.crystallography.net/. Accessed 04 Nov 2011
13. http://www.acdlabs.com/products/draw_nom/draw/chemsketch/. Accessed 04 Nov 2011
14. MDL Information Systems, Inc. (MDL ISIS™) http://mdl-isis-draw.software.informer.com/. Accessed 04 Nov 2011
15. http://www.cambridgesoft.com/software/ChemDraw/. Accessed 04 Nov 2011
16. Discover molecular modeling software (1993) Biosym Technologies Inc., San Diego, CA
17. Insight INSIGHT II molecular modeling package (2000) Accelrys, San Diego, CA
18. http://accelrys.com/. Accessed 04 Nov 2011
19. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) GROMACS: fast, flexible, and free. J Comput Chem 26:1701–1718. doi:10.1002/jcc.20291
20. http://www.gromacs.org/. Accessed 04 Nov 2011
21. Wavefunction, Inc., Irvine. http://www.wavefun.com/products/spartan.html. Accessed 04 Nov 2011
22. Mohamadi F, Richard NGJ, Guida WC, Liskamp R, Lipton M, Caufield C, Chang G, Hendrickson T, Still WC (1990) MacroModel – an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. J Comput Chem 11:440–467. doi:10.1002/jcc.540110405
23. Allinger NL (1977) Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V_1 and V_2 torsional terms. J Am Chem Soc 99:8127–8134. doi:10.1021/ja00467a001
24. Allinger NL, Yuh YH, Lii JH (1989) Molecular mechanics. The MM3 force field for hydrocarbons. 1. J Am Chem Soc 111:8551–8566. doi:10.1021/ja00205a001
25. Weiner SJ, Kollman PA, Case D, Singh UC, Alagona G, Profeta S, Weiner P (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. J Am Chem Soc 106:765–784. doi:10.1021/ja00315a051

26. Weiner SJ, Kollman PA, Nguyen NT, Case DA (1986) An all atom force-field for simulations of proteins and nucleic-acids. J Comput Chem 7:230–252. doi:10.1002/jcc.540070216
27. Jorgensen WL, Tirado-Rives J (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. J Am Chem Soc 110:1657–1666. doi:10.1021/ja00214a001
28. Jorgensen L, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935. doi:10.1063/1.445869
29. Halgren TA (1999) MMFF VI. MMFF94s option for energy minimization studies. J Comput Chem 20:720–729. doi:0.1002/(SICI)1096-987X(199905)20:7<720::AID-JCC7>3.0.CO;2-X/abstract
30. Polak E, Ribiere G (1969) Note sur la convergence de methods de directions conjuguès. Revenue Francàise Informat Recherche Operationnelle 16:35–43
31. Cimino P, Gomez-Paloma L, Duca D, Riccio R, Bifulco G (2004) Comparison of different theory models and basis sets in the calculation of ^{13}C NMR chemical shifts of natural products. Magn Reson Chem 42:S26–S33. doi:10.1002/mrc.1410
32. Young DC (2001) Computational chemistry. Wiley, New York
33. Møller C, Plesset MS (1934) Note on an approximation treatment for many-electron systems. Phys Rev 46:618–622. doi:10.1103/PhysRev.46.618
34. Head-Gordon M, Pople JA, Frisch MJ (1988) MP2 energy evaluation by direct methods. Chem Phys Lett 153:503–506. doi:10.1016/0009-2614(88)85250-3
35. Hohenberg P, Kohn W (1964) Inhomogeneous electron gas. Phys Rev B 136:B864–B871. doi:10.1103/PhysRev.136.B864
36. Kohn W, Sham LJ (1965) Self-consistent equations including exchange and correlation effects. Phys Rev A 140:A1133–A1138. doi:10.1103/PhysRev.140.A1133
37. Adamo C, Barone V (1998) Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: the mPW and mPW1PW models. J Chem Phys 108:664–675. doi:10.1063/1.475428
38. Adamo C, Barone V (1999) Toward reliable density functional methods without adjustable parameters: the PBE0 model. J Chem Phys 110:6158–6170. doi:10.1063/1.478522
39. Zhao Y, Schultz NE, Truhlar DG (2005) Exchange-correlation functionals with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions. J Chem Phys 123:161103-1–161103-4. doi:10.1063/1.2126975
40. Zhao Y, Schultz NE, Truhlar DG (2006) Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. J Chem Theory Comput 2:364–382. doi:10.1021/ct0502763
41. Zhao Y, Truhlar DG (2008) Density functionals with broad applicability in chemistry. Acc Chem Res 41:157–167. doi:10.1021/ar700111a
42. www.gaussian.com. Accessed 04 Nov 2011
43. www.hyper.com. Accessed 04 Nov 2011
44. http://www.schrodinger.com/. Accessed 04 Nov 2011
45. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V. Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Came R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng C Y, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian 03, Revision E.01. Gaussian, Inc., Wallingford CT

46. Cheeseman JR, Trucks GW, Keith TA, Frisch MJ (1996) A comparison of models for calculating nuclear magnetic resonance shielding tensors. J Chem Phys 104:5497–5509. doi:10.1063/1.471789
47. (a) Kutzelnigg W (1980) Theory of magnetic susceptibilities and NMR chemical shifts in terms of localized quantities. Isr J Chem 19:193–200; (b) Schindler M, Kutzelnigg W (1982) Theory of magnetic susceptibilities and NMR chemical shifts in terms of localized quantities. II. Application to some simple molecules. J Chem Phys 76:1919–1933. doi:10.1063/1.443165
48. Hansen AE, Bouman TD (1985) Localized orbital/local origin method for calculation and analysis of NMR shieldings. Applications to ^{13}C shielding tensors. J Chem Phys 82:5035–5047. doi:10.1063/1.448625
49. Keith TA, Bader RFW (1993) Calculation of magnetic response properties using a continuous set of gauge transformations. Chem Phys Lett 210:223–231. doi:10.1016/0009-2614(93)89127-4
50. Ditchfield RJ (1972) Molecular orbital theory of magnetic shielding and magnetic susceptibility. J Chem Phys 56:5688–5691. doi:10.1063/1.1677088
51. Wolinski K, Hinton JF, Pulay P (1990) Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations. J Am Chem Soc 112:8251–8260. doi:10.1021/ja00179a005
52. Helgaker T, Jaszuński M, Ruud K (1999) Ab initio methods for the calculation of NMR shielding and indirect spin − spin coupling constants. Chem Rev 99:293–352. doi:10.1021/cr960017t
53. Ernzerhof M, Perdew JP, Burke K (1997) Coupling-constant dependence of atomization energies. Int J Quantum Chem 64:285–295. doi:10.1002/(SICI)1097-461X(1997)64:3<285::AID-QUA2>3.0.CO;2-S/abstract
54. Ernzerhof M, Scuseria GE (1999) Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional. J Chem Phys 110:5029–5036. doi:10.1063/1.478401
55. Sarotti AM, Pellegrinet SC (2009) A multi-standard approach for GIAO ^{13}C NMR calculations. J Org Chem 74:7254–7260. doi:10.1021/jo901234h
56. Smith SG, Goodman JM (2010) Assigning stereochemistry to single diastereoisomers by GIAO NMR calculation: The DP4 probability. J Am Chem Soc 132:12946–12959. doi:10.1021/ja105035r
57. Bassarello C, Zampella A, Monti MC, Gomez-Paloma L, D'Auria MV, Riccio R, Bifulco G (2006) Quantum mechanical calculation of coupling constants in the configurational analysis of flexible systems: determination of the configuration of callipeltin A. Eur J Org Chem 604–609. doi:10.1002/ejoc.200500740
58. Bifulco G, Bassarello C, Riccio R, Gomez-Paloma L (2004) Quantum mechanical calculations of NMR J coupling values in the determination of relative configuration in organic compounds. Org Lett 6:1025–1028. doi:10.1021/ol049913e
59. Cramer CJ (2004) Essentials of computational chemistry. Wiley, Chichester
60. Bagno A, Rastrelli F, Saielli G (2005) NMR techniques for the investigation of solvation phenomena and non-covalent interactions. Prog Nucl Magn Reson Spectrosc 47:41–93. doi:10.1016/j.pnmrs.2005.08.001
61. Aidas K, Møgelhøj A, Kjær K, Nielsen CB, Mikkelsen KV, Ruud K, Christiansen O, Kongsted J (2007) Solvent effects on NMR isotropic shielding constants. A comparison between explicit polarizable discrete and continuum approaches. J Phys Chem A 111:4199–4210. doi:10.1021/jp068693e
62. Dračínský M, Bouř P (2010) Computational analysis of solvent effects in NMR spectroscopy. J Chem Theory Comput 6:288–299. doi:10.1021/ct900498b
63. Tomasi J, Mennucci B, Came R (2005) Quantum mechanical continuum solvation models. Chem Rev 105:2999–3093. doi:10.1021/cr9904009
64. Klamt A, Schüürmann G (1993) COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. J Chem Soc Perkin Trans 2:799–805. doi:10.1039/P29930000799

65. Qiu D, Shenkin PS, Hollinger FP, Still WC (1997) The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate born radii. J Phys Chem A 101:3005–3014. doi:10.1021/jp961992r
66. Manzo E, Gavagnin M, Bifulco G, Cimino P, Di Micco S, Ciavatta ML, Guoc YW, Cimino G (2007) Aplysiols A and B, squalene-derived polyethers from the mantle of the sea hare Aplysia dactylomela. Tetrahedron 63:9970–9978. doi:10.1016/j.tet.2007.07.055