molUP: A VMD Plugin to Handle QM and ONIOM Calculations Using the Gaussian Software

Henrique S. Fernandes, Maria J. Ramos, and Nuno M. F. S. A. Cerqueira

The notable advances obtained by computational (bio)chemistry provided its widespread use in many areas of science, in particular, in the study of reaction mechanisms. These studies involve a huge number of complex calculations, which are often carried out using the Gaussian suite of programs. The preparation of input files and the analysis of the output files are not easy tasks and often involve laborious and complex steps. Taking this into account, we developed molUP: a VMD plugin that offers a complete set of tools that enhance the preparation of QM and ONIOM (QM/MM, QM/QM, and QM/QM/MM) calculations. The starting structures for these calculations can be imported from different chemical formats. A set of tools is available to help the user to examine or modify any geometry parameter. This includes the definition of layers in ONIOM calculations, choosing fixed atoms during geometry optimizations, the recalculation or adjustment of the atomic charges, performing SCANs or IRC calculations, etc. molUP also extracts the geometries from the output files as well as the energies of each of them. All of these tasks are performed in an interactive GUI that is extremely helpful for the user. MolUP was developed to be easy to handle by inexperienced users, but simultaneously to be a fast and flexible graphical interface to allow the advanced users to take full advantage of this plugin. The program is available, free of charges, for macOS, Linux, and Windows at the PortoBioComp page https://www.fc.up.pt/PortoBioComp/database/doku.php?id=molup. © 2018 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.25189

Introduction

Computational (bio)chemistry is today an area of research with applications in several fields of chemistry, biochemistry, and material sciences. It exploits the advances in computer hardware and software, as well as new mathematical and theoretical approaches, to provide accurate and detailed qualitative/quantitative descriptions of chemical problems.

One of the main advantages of computational (bio)chemistry relies on its ability to tackle, with atomic detail, the cleavage and formation of chemical bonds from complicated multistep processes that often go beyond the limits of detection of the experimental capabilities. All of these features, taken together, have prompted the use of computational (bio)chemistry in situations where experiments are difficult to perform, expensive or when the chemistry is difficult to predict. Computational (bio)chemistry simulations have thus become a popular approach in many research fields, and it is revolutionizing the way we do science. Currently, it is widespread in academia and also in industry where there is an urgent need to decrease the high cost associated with the development of new chemical compounds, investigate and predict the properties of new materials, or even to interpret, analyze, and predict the results of chemical reactions.

Computational (bio)chemistry can be used to model a wide variety of (bio)chemical processes, ranging from very accurate studies of small molecules to complex simulations of macromolecular systems. Different theoretical methodologies are used in different types of applications. Small molecules are treated using sophisticated and very accurate models of the Schrödinger equation. Larger molecules are treated with more approximate methods such as density functional theory (DFT). Macromolecular systems are treated at an even simpler level, where only the "reactive center" is studied by quantum chemistry (QM) while the surroundings are modeled by classical mechanics or molecular dynamics (MD) simulations. The combination of the accuracy of a QM method with the low computational cost of molecular mechanics (MM), has become very popular in the past decade. The incorporation of quantum mechanics into molecular mechanics can be accomplished in various ways, and one of them is the combined QM and MM (QM/MM).

In 2017, there are over 5000 articles in the Web of Science database where the QM/MM term is used. The generalized use of these methodologies started in 2002, and since then it has experienced exponential growth, accounting today for more than 50% of the articles where the term QM is employed (Fig. 1). From this subset, the subtractive QM/MM method ONIOM, developed by Morokuma and coworkers,[1–9] is collecting many followers and until 2017 it has been cited by more than

H. S. Fernandes, M. J. Ramos, N. M. F. S. A. Cerqueira
UCIBIO@REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre s/n, Porto 4169-007, Portugal
E-mail: nscerque@fc.up.pt

Contract grant sponsor: European Union; Contract grant number: POCI/01/0145/FEDER/007728; Contract grant sponsor: Fundação para a Ciência e Tecnologia and Ministério da Educação e Ciência; Contract grant numbers: PT2020; UID/MULTI/04378/2013; NORTE-01-0145-FEDER-000024; Contract grant sponsor: Norte Portugal Regional Operational Programme; Contract grant number: NORTE 2020; Contract grant sponsor: Fundação para a Ciência e a Tecnologia; Contract grant number: IF/01310/2013; Contract grant sponsor: the Fundação para a Ciência e Tecnologia; Contract grant number: SFRH/BD/115396/2016

© 2018 Wiley Periodicals, Inc.

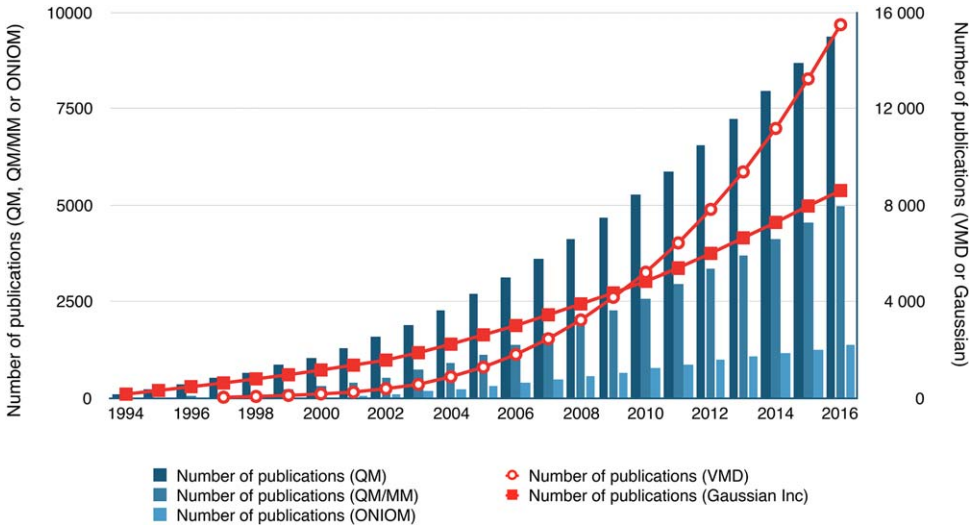

Figure 1. Number of citations found in the Thomson Reuters Web of Science database using the keywords QM, QM/MM, ONIOM, Gaussian, and VMD, from 1994 to 2016. The data displayed in the graphic is cumulative. [Color figure can be viewed at wileyonlinelibrary.com]

25% of the QM/MM articles that are described in the literature (Fig. 1).

Despite the methodological developments of the QM/MM methods, and the good results obtained through their application, they are still very challenging.[10–12] The two major bottlenecks of these methods are i) the computational cost of the QM method that is employed, which turns the full process very time and resource demanding and ii) the way through which the coupling between the QM and MM parts is accomplished. With the advent of faster computers, the computational cost has been minimized, and now it is possible to run QM/MM calculations faster, with QM regions reaching over 200 atoms, and to fully characterize transition states (TS). The coupling between the QM and MM parts continues to be the heart of these methods and the most discussed issue. However, in the last 5 years, several algorithms have been developed and used with relative success to improve the treatment of the artificial linkage that connects the QM and MM regions.[2,7,13]

From a practical point of view, these methods were and continue to be very difficult to work with. The model systems used in QM/MM calculations often have more than 10 thousand atoms, and therefore the input files are very complex and laborious. The analysis and extraction of useful information from the output files are also cumbersome. They often lie in one or two lines from the millions of lines that composed these output files that normally occupy several gigabytes of disk space. For this reason, the visualization of these files is also difficult, either because there are few suites of software that can handle this sort of information or because most of them crash due to memory limitations.

We believe that the manipulation of the input and output files from the QM/MM calculations is currently the major limitation that prevents their generalized use in science. Taking this into account, we developed molUP, an extension of the popular Visual Molecular Dynamics software (VMD)[14] that allows VMD to create, manipulate, and analyze the input and output files from the ONIOM calculation files that are generated with the Gaussian software.[15]

We have chosen the ONIOM method because it is becoming a very popular QM/MM, QM/QMMM, or QM/QM method and it is the only one that is currently available in the versatile and very popular Gaussian software.[15] The inclusion of molUP in the VMD as a plugin was chosen because VMD is a versatile molecular visualizer that can handle thousands of atoms and represent them in several and different ways. Both Gaussian and VMD are also, today, among the most popular software applications used in computational (bio)chemistry. Therefore, we thought of taking advantage of this molecular visualizer software and add several new features to make it ready to prepare and analyze QM/MM, QM/QM, or QM Gaussian-related files.

In this paper, we summarize the main features of molUP including some explanations on the type of methods and protocols that are commonly used in computational (bio)chemistry. In addition, we provide, in Supporting Information, a complete and detailed guide to study the catalytic mechanism of an enzyme using molUP. We will use as an example the catalytic mechanism of L-Histidine Decarboxylase (HDC).[10]

The molUP Plugin

molUP is a VMD extension that provides a full-featured graphical user interface (GUI) (Fig. 2) to the computational chemistry software Gaussian.[15] It helps users to prepare the laborious input files that are required to run the QM or ONIOM calculations and also to analyze and visualize the complex output files that result from these calculations using the advanced visualization facilities from VMD.

The molUP GUI is composed by five main sections: a menubar, a molecule selector, a central section, a quick representations chooser and, at the bottom of the GUI, an icon bar where frequently used tools were made more accessible (Fig. 2).

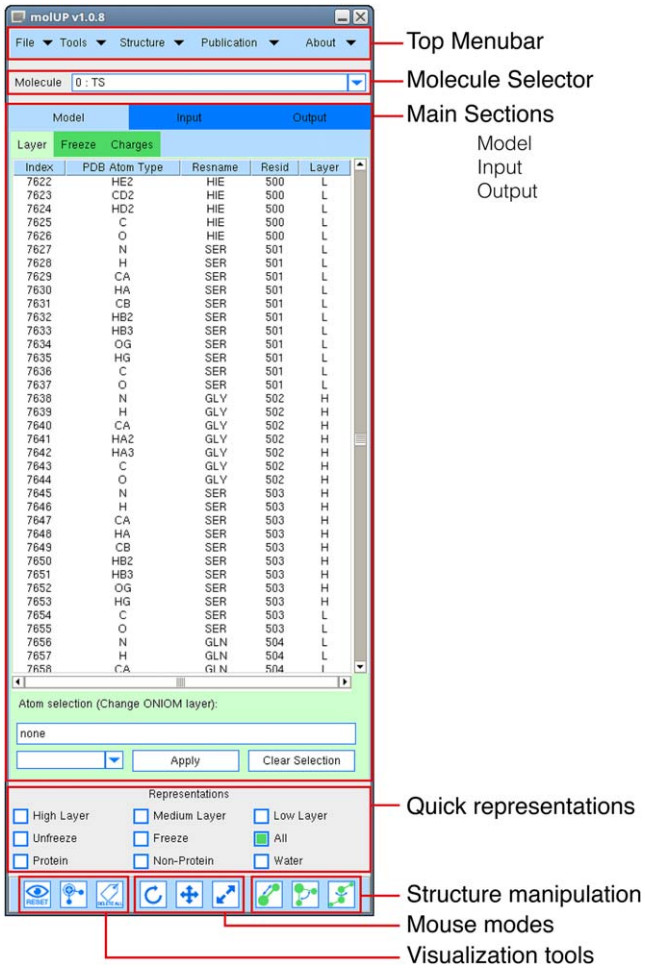

Figure 2. Overview of the molUP GUI. [Color figure can be viewed at wileyonlinelibrary.com]

The top menubar has five dropdown menus, File, Tools, Structure, Publications, and About. Each of these menus includes several tools that allow the user to perform a variety of functions. The "File" menu allows to open/save input files and open output files from the Gaussian software. The "Tools" menu provides shortcuts for useful VMD commands, such as reset view, center the view in a specific atom, delete all labels, and alternate between mouse modes. The "Structure" menu provides useful tools to modify the bond, angles or dihedral angles of the atoms present in the VMD main windows. The "Publication" menu presents a methodological overview of the calculation that has been set up in molUP. Finally, the "About" menu provides general information on the molUP extension.

The molecular selector section has a combobox widget that allows the user to select the molecular system that is loaded in VMD and that will be analyzed by molUP. This widget is very useful when several molecules are loaded in VMD, and the user can swap between them to prepare or analyze different Gaussian input or output files.

The main section of molUP is where the main operations will take place. This region is where all the details about the model system, which will be used to create the Gaussian input or visualize the data from the output files, will be handled and visualized. Further details about this section will be given below.

The quick representation section provides a set of switches that can be used to easily show/hide different molecule representations of the model system. This tool is very helpful when analyzing QM/MM files where the different layers can be easily represented and zoomed in.

The bottom section of the GUI includes shortcuts to some of the most frequently used tools present in the molUP GUI. Each of these options can also be accessed in the "Tools" and "Structure" menus of molUP.

The workflow of a computational study addressed to a molecular system can be divided into three main steps: setup of input files, running calculations, and the analysis of the results (output files). In the following section of the manuscript, each of these steps will be converted into different sections and will be used to describe the main features of molUP (Fig. 3). The three sections all together can be viewed as a general workflow that can be used by any user to prepare and analyzed Gaussian files.

Input section

Gaussian input file is where the user chooses the type of calculation that will be run and sets the coordinates of the system that will be studied.

Initial Structure. molUP supports a wide range of file formats. As molUP is built as a plugin of VMD, it benefits from the capabilities of VMD to load structures from a wide range of different chemical formats, such as AMBER, CHARMM, Autodock, Gromacs, MOL2, Molden, PDB, XYZ, and MOL2. This means that the files can be imported directly to molUP not only from simple coordinate files but also from Molecular Dynamics simulations, Monte Carlo Simulations, Molecular Docking calculations, among others. These structures can be uploaded directly in VMD, and molUP will automatically recognize them and use them to generate Gaussian input files.

VMD lacks the capability to read Gaussian input or output files from the most recent versions (Gaussian09). In addition, it is also not prepared to handle the files derived from the ONIOM calculations. For this reason, we developed new procedures in molUP that allow VMD to read QM, QM/MM, QM/MM, or even QM/QM/MM files derived from the Gaussian software. These files can be loaded onto VMD using the "File > Open" menu from the menubar placed on the top of the molUP GUI. The user can upload multiple Gaussian input Files to molUP. Once loaded they will be listed in the "Molecular Selector" combobox that can be used by the user to interchange their visualization in VMD and analysis in molUP.

Model Preparation. *Choosing the model.* When we plan to study the catalytic mechanism of an enzyme by computational means, one of the first thing that needs to be defined is the model that is going to be used during the calculations. A good model must represent the chemistry of the system that is being studied and must provide a balanced compromise between model size and computational cost.

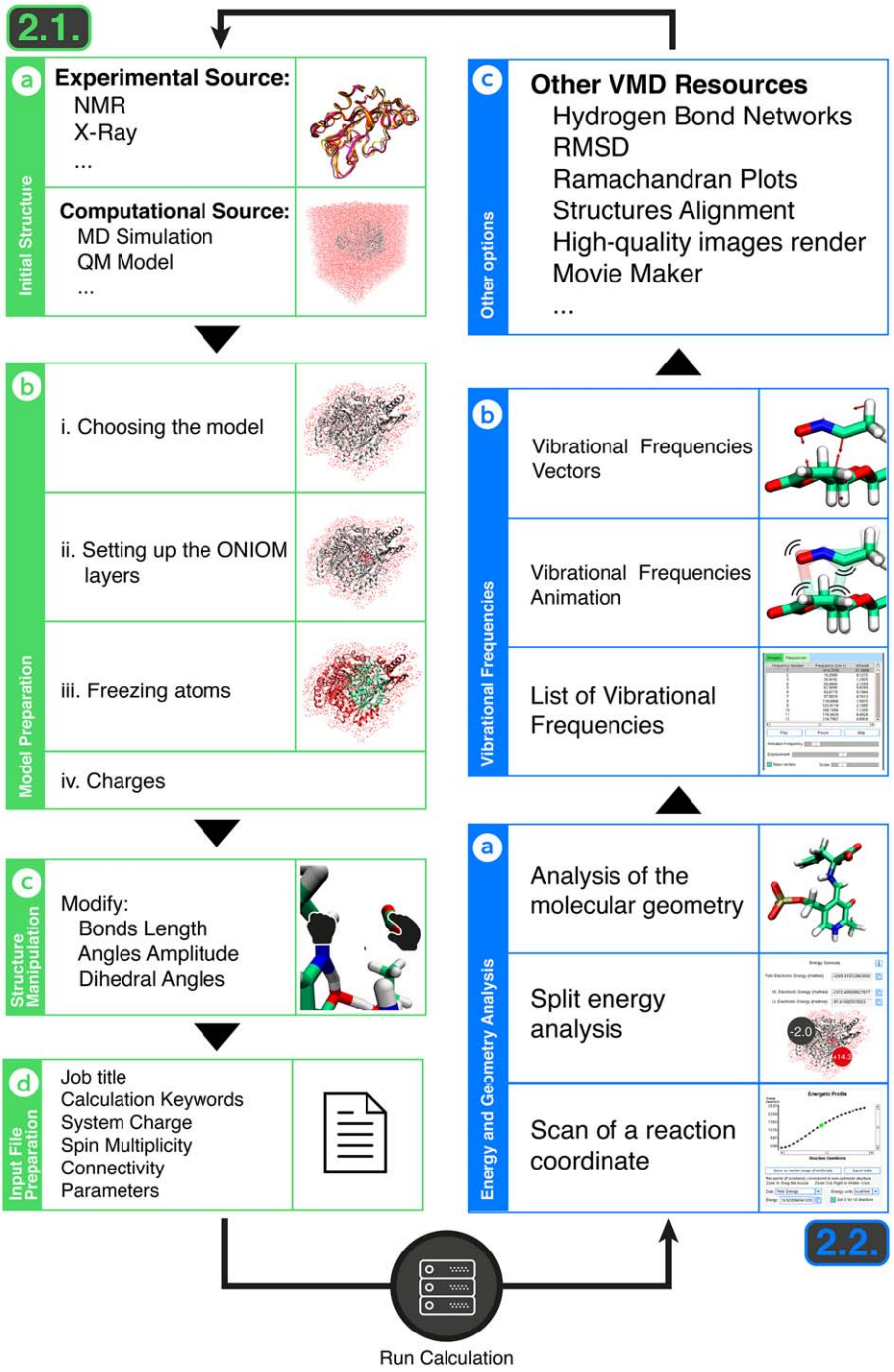

Figure 3. Schematic overview of the protocol used to carry on a QM/MM study using Gaussian, VMD, and molUP. [Color figure can be viewed at wileyonlinelibrary.com]

The chemical structure that is loaded onto molUP should therefore contain all the atoms that the user wants to include in the theoretical calculations. The coordinates of the atoms can be modified with molUP (as it will be described later on) but the addition or elimination of atoms is not supported.

Afterwards it is necessary to set up the model system for the computational calculations. This may require the division of the atoms in different layers that will be calculated with different theoretical levels of theory, which atoms will be optimized and/or frozen during geometry optimizations, modify the charge of the atoms, etc. These tasks can be easily accessed through the "Model" tab of molUP that are briefly described in the following sections.

Setting up the ONIOM layers. The division of atoms in different layers can be done in the "Model > Layer" tab of molUP and should be only used when the user wants to perform an ONIOM calculation. The ONIOM acronyms sets for "our own n-layered integrated molecular orbital and molecular mechanics," which means that the system is split into different layers and

each one of them is treated with a different theoretical level or method. If the user wants to perform a simple QM calculation, all the atoms should be included in a unique layer, such as the high-layer, or simply keeping the layer column blank.

When the user wants to perform an ONIOM calculation, the model system can be divided into two or three layers. The most common one is the two-layered partition, such as the QM/QM or QM/MM partition, the "high-layer" (HL) contains the atoms that will be treated with the higher level of theory (such as QM, in a QM/MM system), whereas the "low-layer" (LL) contains the remaining atoms that are treated with the lower level of theory (such as MM, in a QM/MM system). A three layered-systems can contain three layers, the HL, the "medium-layer" (ML), and the LL, that will be calculated with three different levels of theory.

The selection of the atoms that will be included in each layer can be done in the "Model > Layer" tab. This tab also contains all the attributes of the atoms and residues of the molecules (such as resid and resname) that can help the user to choose which atoms will be included in each layer. The selection of the atoms can be interactively visualized in the VMD interface, which also helps the user in the selection process. At the bottom of the molUP interface, it is also available a set of check buttons that can be used to visualize or hide the atoms present in each layer.

Freezing atoms. The spatial constrain (freeze) of some atoms during the geometry optimizations is often employed in Gaussian calculations. The selection of the atoms that the user wishes to freeze can be quite cumbersome, specially in QM/MM calculations that contain thousands of atoms. In molUP this can be easily done using the "Model > Freeze" tab. The atoms that the user wishes to freeze are commonly designated by "frozen atoms" and are often assigned the value "−1." The atoms that are free have a "0" value. In addition, negative integers can be attributed to atoms, such as "−2" and "−3." These values are used to freeze the coordinates of those atoms that move as rigid blocks during geometry optimizations. The attribution of these values to each atom can be done directly on the molUP interface through the selection of a group of atoms and from which a specific value can be attributed. This feature can be used in QM or ONIOM calculations. The selection-mode is particularly useful in ONIOM calculations where the positions of hundreds of atoms are often employed to decrease the number of conformation rearrangements in the "low-level" layer.

Charges. The "Model > Charges" tab of molUP is particularly important for Molecular Mechanics and ONIOM calculations, in particular when the latter employs a QM/MM partition. The atomic charges of each atom must be carefully included to run the MM calculations. These values are presented in the form of a table together with the atom type used by Gaussian to perform the MM calculations. Users can edit the charge values of each atom or obtain this information directly from files that already includes it, such as Gaussian input or output files or even AMBER output files that result from standard minimization or MD simulations.

Structure Manipulation. molUP cannot be used as a general molecular editor to build new molecules because VMD is not a molecular editor but rather a molecular visualizer. This means that molUP cannot be used to add or delete atoms. However, it can be used to manipulate the position of one atom or a group of atoms of the molecule that is displayed by VMD. This can be done through a set of tools that can be accessed directly from the last three icons placed on the bottom bar of the molUP GUI or through the "Structure" menu (top menubar). Here, it is possible to manipulate bond length, angles, and dihedral angles of atoms or groups of atoms (Fig. 4).

Preparation of the Gaussian Input File. The input file for Gaussian calculations is composed of several sections where different types of information must be inserted, such as the level of the theory that is going to be used in the calculation, the basis set, the type of calculation, atom coordinates, connectivity, and atom parameters. In molUP, this information can be found in the "Input" tab and follows a similar organization of a Gaussian input file, as illustrated in Figure 6.

The first section of molUP input tab is the job title. This section can be used by the user to assign notes to the input file and it is a good way to identify the type of calculation that is prepared.

The second section of molUP input tab is labeled as "Calculations" and it is where all the keywords that are required to run Gaussian calculations are set up. It includes all the information from the "Route 0 section" where a set of technical parameters such as the RAM and CPU limits and/or the path of scratch files that are generated during the calculation are defined. It also includes all the keywords that determine the type of calculations that is going to be performed (Gaussian Route section). This part must start with a "#" character and it is where the level of theory, the basis set, the type of calculations (single-point energy calculation, geometry optimization, etc.), frequency analysis, among others are defined. Since the keywords that are used by Gaussian are never trivial, molUP includes a color flag system, in which the most common keywords are displayed in different colors. For example, the basis set keywords are written in purple, the functionals in green, the ONIOM in blue, etc. We believe that this feature will prevent the preparation of input files with mistakes that normally crash or provide wrong results. The file containing these keywords can be changed by the user to add additional keywords or even to change the default color scheme. This file can be found in the "/user/references.txt" file within the molUP installation directory. Since the same Gaussian input files are routinely used several times, molUP has also a dropdown menu (called "Calculations") in the input tab that contains some of the most commonly used Gaussian input files. This means that the user can setup very easily a Gaussian input file just choosing an input file type from the template menu. The user can also save his/her own template and include it in the dropdown menu to have it accessible when future input files will be prepared. For example, in ONIOM calculations, the user can easily interchange between mechanical and electrostatic

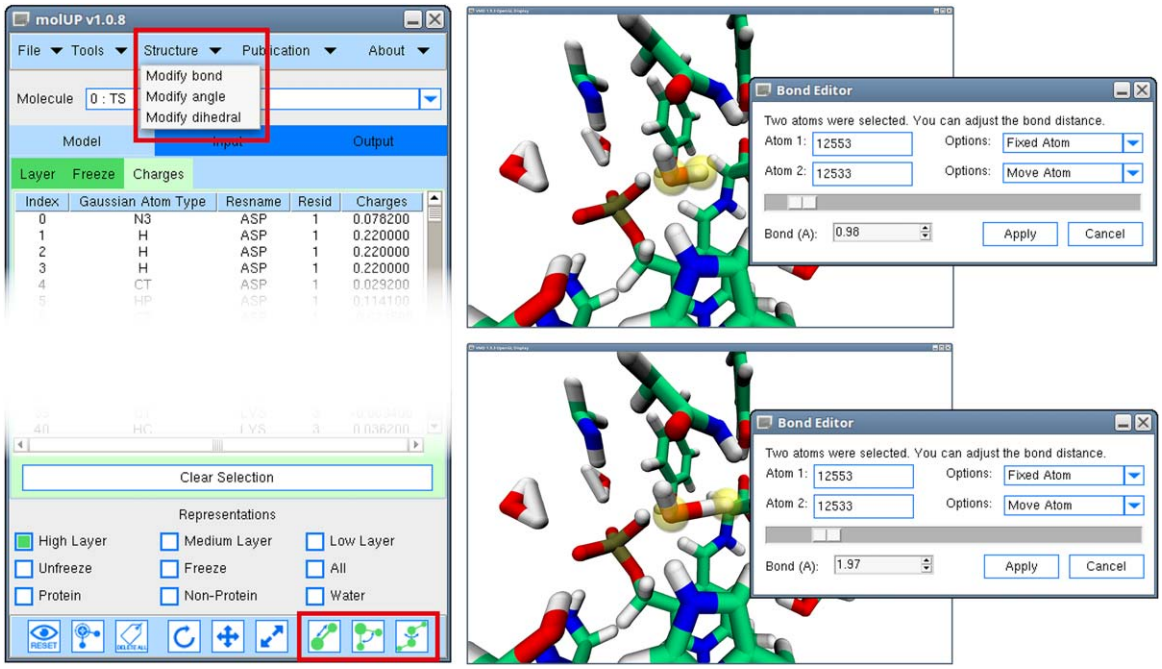

Figure 4. (Left) Tools available in molUP to manipulate bonds, angles, and dihedral angles. (Right) A chemical bond being modified using molUP tools. [Color figure can be viewed at wileyonlinelibrary.com]

embedding type of calculations, which only differ in a single keyword (for more details see Supporting Information).[16]

The following section of the "Input" tab of molUP is where the total charge and spin multiplicity of the model system are defined. In the case of ONIOM multi-layered calculations, this section also includes the charge and spin multiplicity of the atoms included in different layers. In these cases, molUP also provides a handy function that automatically calculates the total charge of each layer, based on the atomic charges of the atoms present in each of those regions, and allows to highlight the positively and/or negatively charged residues present in each layer. This information is however only available when the atomic charges are available, which can be imported from previous QM/MM calculations or even from MM simulations.

The third section of the molUP "Input" tab includes the connectivity of the atoms present on the model system. This section can be left empty in QM calculations, but it is particular important when MM methods are used, as for example in ONIOM calculations. molUP always tries to load the connectivity of the atoms from the files where the coordinates of the model system were obtained from. If this information is not available, then molUP will automatically generate them based on the connectivity predicted by TopoTools,[17] a third-party plugin included in every VMD distribution. It also offers the option to load the connectivity information from other Gaussian input files or from other file containing this information. This feature is very useful in ONIOM calculations to maintain the connectivity of the atoms in reactant, TS, and product of the reaction constant and minimize accordingly the energy fluctuations in MM layers, which can be produced from different connectivities among the atoms from the different minima.

The last section of the "Input" tab of molUP includes additional information that may be required for the Gaussian input file. It often includes the information about ModRedundant commands that freezes one or more coordinates, bond lengths, angles, or dihedral angles during geometry optimizations, information about SCANs, information on the basis set when the "gen" keyword is used, etc.

molUP places the information about the freeze state of the atoms in an extra column between the atom information and its coordinates (Fig. 5 Right) in the "Model > Freeze" tab, as described before. However, Gaussian also allows the user to fix the position of atoms, angles and dihedral angles using a ModRedundant entry placed after the connectivity data. For example, we can fix the angle between atoms 1, 2, and 3 by adding line "A 1 2 3 F" to the "Other information" section of the input file.

Since the SCANs are commonly employed to search for the reaction coordinate of a chemical reaction, we included a "ModRedundant editor" tool that allows setting up size and step for the scan as well as which atoms will be tested across the scan.

In the case of ONIOM calculations, this section can be used to include the MM parameters of the atoms. Since the MM parameters used by ONIOM are based from the AMBER force field, molUP also offers the option to import these parameters directly from the AMBER parameters file (.prmtop file), which turns this process more straightforward and easy.

By default, molUP adds the MM parameters to the Gaussian input file in a "softfirst" manner. This means, that during the ONIOM calculation, Gaussian will read these parameters from the input file and if any parameter is missing it will look for it in the Gaussian built-in library. This procedure is often chosen

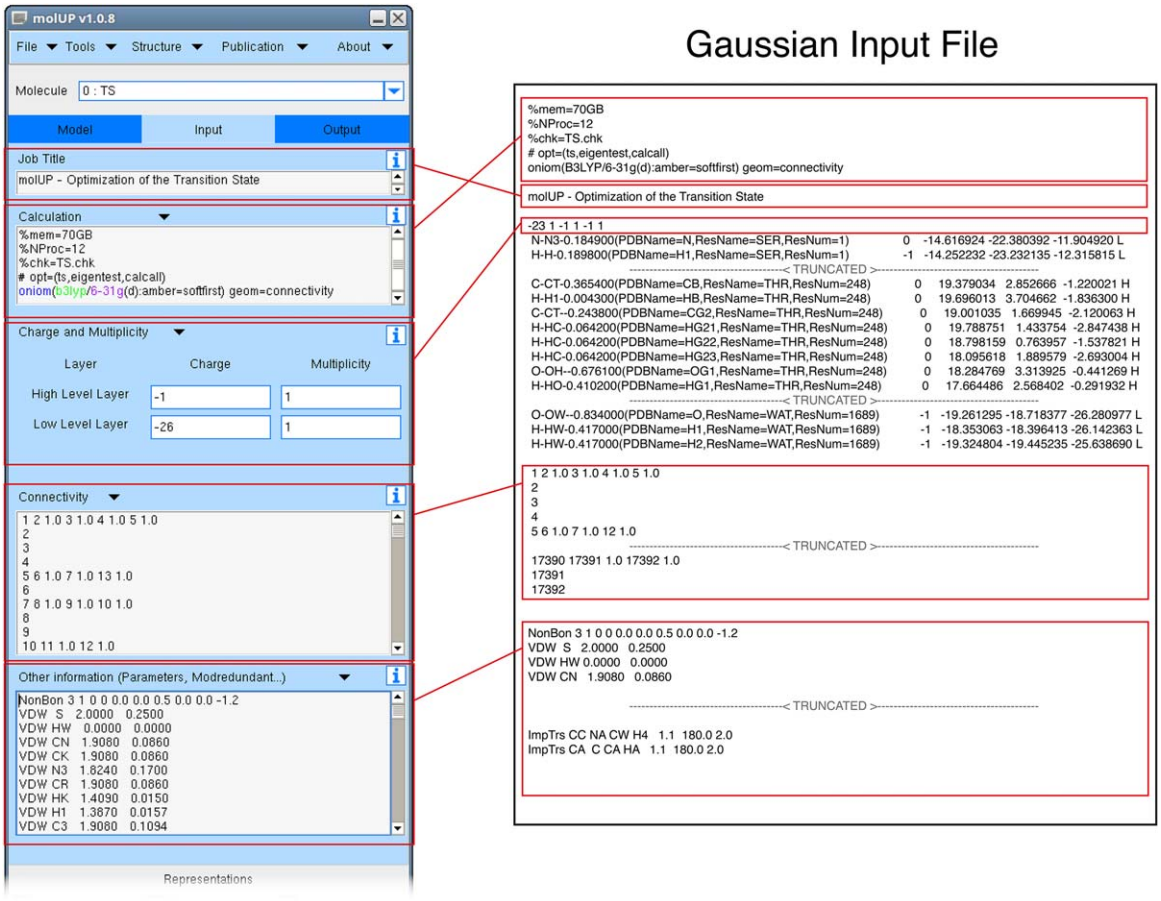

Figure 5. Input section of molUP and the associated section of a Gaussian input file. Boxed regions of the Gaussian input file include information that has the same organization of the adjacent lines. [Color figure can be viewed at wileyonlinelibrary.com]

since it provides a better control of the parameters that are used in ONIOM calculations (for more details see Supporting Information).

Once all the fields from the "Input" tab have been completed, the Gaussian input file can be saved using the menu "File > Save." A new window will pop up, where the user can choose the name of the file as well as the directory where the file will be saved. In ONIOM calculations, the software will automatically add the information concerning the link atoms (hydrogen atoms) that are used to divided the atoms in different layers. When saving ONIOM files, the user can choose to save a small subset of atoms in the input file. For example, only the atoms present in the "high-level" layer. In these cases, molUP will automatically add hydrogens as link atoms to the region where the cleavage between the high- and low-layer regions has been done (in a QM/MM model). This is particularly useful when the user wants to calculate the energy of each layer in a different software without using the ONIOM partition directly.

Output section analysis

After running the Gaussian calculations, a complex output file is generated, which can also be viewed and analyzed by molUP. Routinely performed tasks, such as the analysis of the molecular geometries, energy data, vibrational frequencies, can be easily performed in molUP (Fig. 3). Each of these features will be described in the following sections.

molUP also provides a warning feedback if some error in the Gaussian output file is found. This feature is very useful to warn the user for potential errors in the calculations or problems in running the file.

Energy and Geometry Analysis. The Gaussian output files can be opened in the "File" menu of the molUP interface. These files are very complex and are often composed by several gigabytes of data, which can turn the loading time very long. To accelerate this process, molUP offers several options to turn this process more efficient. For example, only the optimized and non-optimized geometries of the calculation or the last geometry can be loaded.

Independently of the option that is selected, molUP will extract the energy of each structure that is found. If the user loads solely the last structure of the output file, a single energy value is presented. In the case of ONIOM calculations, molUP presents the total energy of the system plus the energy of each layer (Fig. 6A), e.g. in a QM/MM system, the energy of the HL and the LL. If the output file contains thermal corrections, the zero-point energy, enthalpy, and Gibbs free energy will be shown too (Fig. 7A). All of these energy values can be found in the "Energies" section of the "Output" tab.

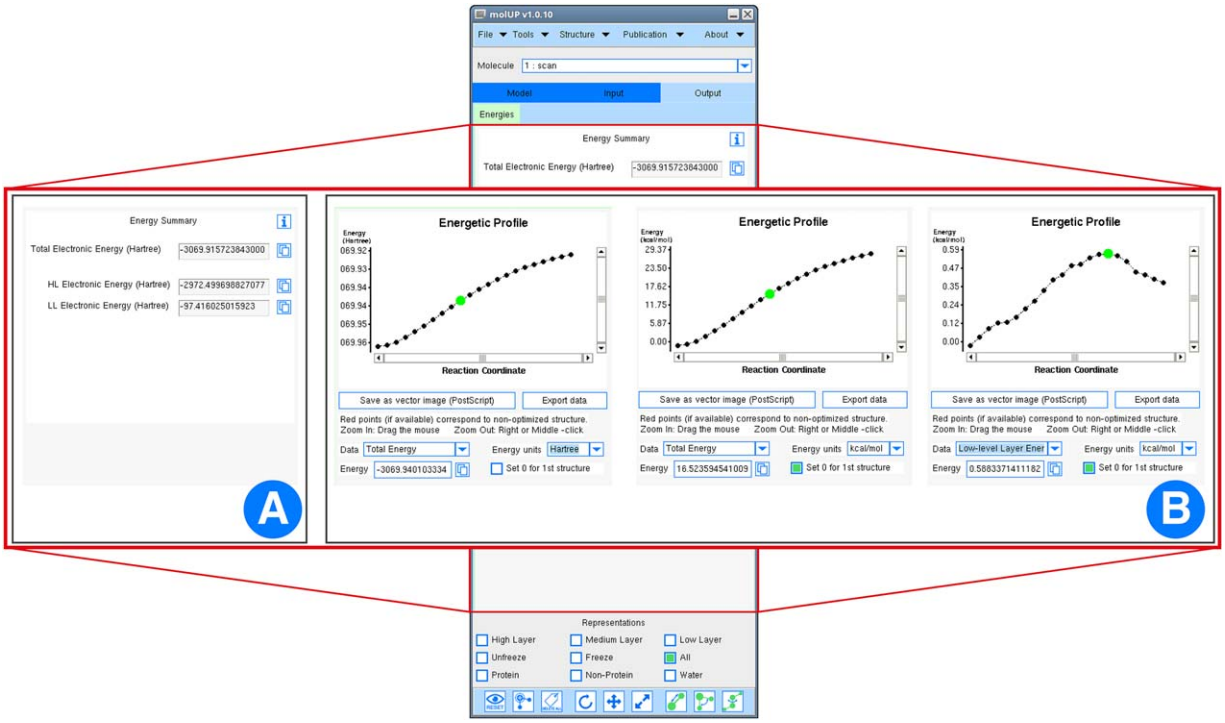

Figure 6. Different layout of the "Energies" sections according to the option used during the loading of an output file. A) Energy values collected from the last structure of a geometry optimization of an ONIOM QM/MM system. B) Energy plot obtained from a scan of a reaction path using an ONIOM QM/MM system: (Left) Energy values in Hartree; (Middle) Energy values in kcal/mol setting the energy of the first structure to zero; (Right) Energy plot considering only computed energies for the LL in an ONIOM QM/MM system. [Color figure can be viewed at wileyonlinelibrary.com]

When more than one geometry is available in the output file and they are loaded onto molUP, they will be plotted into a chart. The user can use the chart to interactively interchange each geometry that is displayed on the VMD main window by selecting the different dots that correspond to different geometries. If optimized and non-optimized geometries are available, they will be plotted in the chart with different colors (black and red, respectively). molUP also allows to interchange the units of the energies that are plotted in the chart from Hartree to kcal/mol, kJ/mol, and eV. An additional switcher allows to set the energy of the first structure as zero by subtracting the energy of that structure to all the others. This feature is particularly useful when analyzing SCANs or intrinsic reaction coordinates (IRC) (Fig. 6B). When ONIOM calculations are analyzed, the values of the total energy or the energy of each layer can be independently analyzed in this chart.

The data that is available in the chart can also be exported to a text file and later on imported in a spreadsheet for further analysis, or exported to a vector image (PostScript).

Vibrational Frequencies. Vibrational frequencies calculations are used to characterize all the minima of a potential energy surface. From each calculation results a list of frequencies with the respective magnitude and a list of vibrational vectors for the involved atoms. In the case of a TS, one imaginary frequency should be present that should be correlated with the bond that is being formed or broken, when a chemical bond is studied.

molUP automatically checks if an output file contains data regarding the frequencies and, in that case, presents the results in a new tab called "Frequencies," where a list of all the computed vibrational frequencies, organized into three columns: Frequency number, Frequency value (cm^{-1}), and Wavelength in the infrared spectrum are provided. Users can click on each frequency and observe the movement of the atoms related to that vibrational frequency. Users can customize the animation of the frequency, namely, the animation speed and the amplitude. Furthermore, molUP has three buttons to control the animation, to play, pause or even stop it. In addition, molUP allows representing all the computed vectors associated with the vibrational frequencies.

In the example present in Figure 7B, it is displayed an imaginary frequency that characterizes the TS of a reaction. As we can observe in Figure 7B, the bigger vectors are related to the two atom pairs involved in the reaction. Users can custom the size and color of the vectors, and also a threshold that limits the minimum vector norm to be represented in the image.

Other Options. Since molUP is integrated in VMD it can take advantage from all the built-in features present in VMD including several plugin extensions. For example, all the structures loaded into molUP can be analyzed with tools that are normally used in molecular dynamics simulations, such as hydrogen bond networks, RMSDs, Ramachandran plots, and the alignment of different amino acid sequences.

Moreover, molUP includes a new tool to help the user to measure bond lengths and the amplitude of angles and

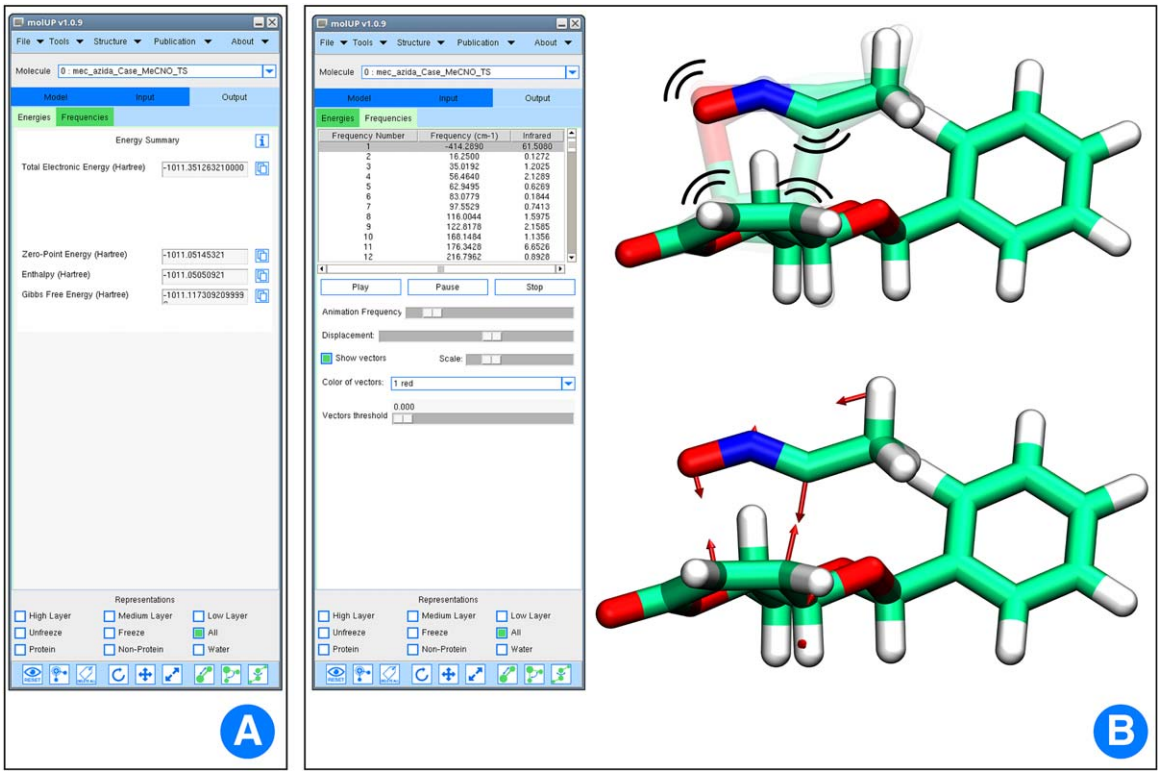

Figure 7. Output of a vibrational frequency calculation. A) Energy of the system including thermal correction for thermochemistry variables; B) Analysis of all vibrational frequencies, animation of those vibrational frequencies, and representation of the computed vectors assigned to each atom involved in a certain vibrational mode. [Color figure can be viewed at wileyonlinelibrary.com]

dihedrals. This tool is available under the name "Measure Bonds, Angles, and Dihedral" under the "Tools" menu of the molUP GUI. Once the tool is activated, a new window shows up, and the user has to click on desired atoms and, at the same time, the information concerning the index, type, and resid are filled in as well as the distance and angles that are measured. In addition, the atoms that are selected in this process are highlighted in the main window of VMD (Fig. 8).

molUP can also be used to produce videos from frequencies, geometry optimizations, etc. To this end, it uses the "Movie Maker" extension of VMD to generate a movie containing all the frames corresponding to each structure that is stored in molUP.

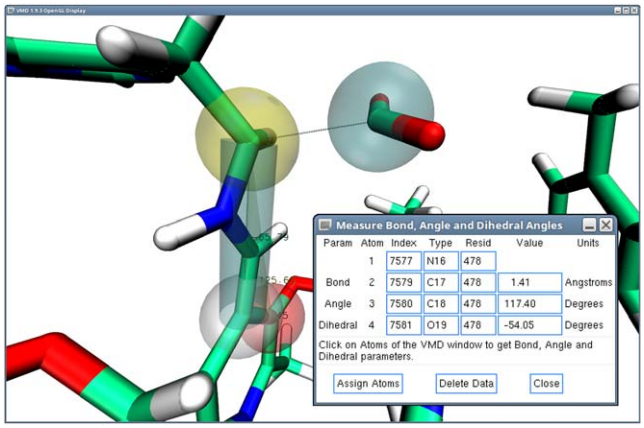

Figure 8. Measure of the bond length, angle and dihedral amplitudes using a molUP tool. [Color figure can be viewed at wileyonlinelibrary.com]

Protocol details

All the calculations that are setup in molUP, are based on flag keywords that are automatically recognized by molUP and colored by a specific color. molUP also uses these keywords to generate a list of references that should be included in the method that is used in the calculation. This information is available in the "references.txt" file within the "user" directory of the molUP installation path. This file can be edited by users to add more keywords and the respective reference. All the references occupy two lines in the file, one corresponding to its contents (authors name, title, journal, volume, issue, year, and pages) and another with the DOI or the URL for the publication.

Technical Implementation

All software parts of molUP were coded in Tcl/Tk. No additional libraries are needed. All the required ones are available in the last version of VMD (1.9.3). molUP can be directly integrated into the extensions menu of VMD from where it opens as a GUI. The program comes with an installation software that turns the incorporation of the plugin on VMD very easy. Since molUP uses built-in tools only available on UNIX systems, molUP is only supported by macOS, Linux, and Windows operating systems.

Gaussian software is required for the calculations. All the other procedures involving the preparation of input files and analysis of output files are only dependent on VMD and molUP.

Conclusion

molUP is a VMD plugin that was developed to enhance handling the input and output files from the Gaussian software.

The development of this plugin was based on the vast experience that our research group has in QM and ONIOM calculations. The main idea was not to develop a competitive application but rather to complement the existing ones. Particular attention was given to the problems that often arise when analyzing or preparing Gaussian files that occupy several gigabytes of disk space and contain molecules with thousands of atoms. With molUP we try to overcome most of these problems and turn all of these tasks easy and straightforward. Particular attention was given to the preparation of ONIOM calculations and subsequent analysis, which are the most problematic ones.

From our own experience, molUP has boosted the preparation and analysis of this sort of files, increasing our workflow and leaving more time for science. This was an essential objective in the development of molUP that was achieved with success.

We believe that molUP will be of interest not only for computational chemists and biochemists without technical background but also for expert users that will take the full advantage of this plugin. The simplicity of the software makes it an attractive tool for education purposes too.

Keywords: molUP · visual molecular dynamics (VMD) · Gaussian · ONIOM · QM · QM/MM

How to cite this article: H. S. Fernandes, M. J. Ramos, N. M. F. S. A. Cerqueira *J. Comput. Chem.* **2018**, DOI: 10.1002/jcc.25189

Additional Supporting Information may be found in the online version of this article.

[1] T. Vreven, K. Morokuma, In *Annual Reports in Computational Chemistry*; C. S. David, Ed.; Elsevier, **2006**, pp. 35–51.
[2] T. Vreven, K. S. Byun, I. Komáromi, S. Dapprich, J. A. Montgomery, K. Morokuma, M. J. Frisch, *J. Chem. Theory Comput.* **2006**, *2*, 815.
[3] F. R. Clemente, T. Vreven, M. J. Frisch, In *Quantum Biochemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, **2010**, pp. 61–83.
[4] M. Orozco, I. Marchán, I. Soteras, T. Vreven, K. Morokuma, K. V. Mikkelsen, A. Milani, M. Tommasini, M. D. Zoppo, C. Castiglioni, M. A. Aguilar, M. L. Sánchez, M. E. Martín, I. F. Galván, H. Sato, In *Continuum Solvation Models in Chemical Physics*; John Wiley & Sons, Ltd, **2007**, pp. 499–605.
[5] S. Humbel, S. Sieber, K. Morokuma, *J. Chem. Phys.* **1996**, *105*, 1959.
[6] M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, K. Morokuma, *J. Phys. Chem.* **1996**, *100*, 19357.
[7] F. Maseras, K. Morokuma, *J. Comput. Chem.* **1995**, *16*, 1170.
[8] T. Matsubara, S. Sieber, K. Morokuma, *Int. J. Quantum Chem.* **1996**, *60*, 1101.
[9] T. Vreven, K. Morokuma, *J. Comput. Chem.* **2000**, *21*, 1419.
[10] H. S. Fernandes, M. J. Ramos, N. Cerqueira, *Chemistry* **2017**, *23*, 9162.
[11] N. M. F. S. A. Cerqueira, H. Moorthy, P. A. Fernandes, M. J. Ramos, *Phys. Chem. Chem. Phys.* **2017**, *19*, 12343.
[12] C. E. A. Sousa, A. M. P. Ribeiro, A. Gil Fortes, N. M. F. S. A. Cerqueira, M. Alves, *J. Org. Chem.* **2017**, *82*, 982.
[13] G. A. Cisneros, J.-P. Piquemal, T. A. Darden, *J. Phys. Chem. B* **2006**, *110*, 13682.
[14] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 27, 33.
[15] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc.: Wallingford, CT **2009**.
[16] N. S. Cerqueira, P. Fernandes, M. J. Ramos, *ChemPhysChem* **2017**.
[17] A. Kohlmeyer, J. Vermaas, **2017**.

Received: 15 November 2017
Revised: 12 January 2018
Accepted: 30 January 2018
Published online on 00 Month 2018