INTRODUCTION TO MOLECULAR MODELING OF MATERIALS IN AN UNDERGRADUATE ENGINEERING DEGREE

J. Torras [1]
Universitat Politècnica Catalunya
Barcelona, Spain
ORCID 0000-0001-8737-7609

D. Zanuy
Universitat Politècnica Catalunya
Barcelona, Spain
ORCID 0000-0001-7704-2178

Conference Key Areas: *Physics and Engineering Education*
Keywords: *exercise & project-based learning; computer laboratory; chemistry; physics; practical courses*

ABSTRACT

Molecular modeling is a chemistry tool that has been widely used in the last decades to mainly support the basic concepts of general chemistry and organic chemistry, in both undergraduate programs of basic sciences and some technological careers. Despite its use, except in some very specific cases, it has been extensively employed as illustrative examples of the chemical concepts that were being demonstrated. Despite the numerous existing applications to comprehend the phenomena behind the development of new materials and biomedicine, it is difficult to find a conceptual introduction of molecular modeling applied to specific problems on the modern engineeries within the undergraduate programs. In the present work, it will be shown the introduction and adaptation of molecular modeling concepts within a new optional course for students coming from materials engineering, chemical engineering and biomedicine engineeries. Different approaches to problem-based and small project-based learning are presented to encourage the scientific spirit of students using techniques of molecular modeling that had not been visited throughout their studies and, thus, to discover their potential appliacation in a more specialized context.

[1] *Corresponding Author*
J. Torras
joan.torras@upc.edu

doi:10.5821/conference-9788412322262.1121

1 INTRODUCTION

Interdisciplinary teachings such as molecular modeling, which involves mathematics, physics, chemistry, computer science and biology, are still scarce both at high school and at the general College level [1]. However, it is observed a shy wave of change in the last decade due to easier access to powerfull desktop installed in the computer classrooms, as well as to a more accessible simulation software [2]. Indeed, subjects where the use of computer programs is predominantly present, such as molecular visualization, bioinformatics, molecular modeling, and computer simulations, are almost absent from high schools, and in the University are offered only to very specific science majors or as a complement to other disciplines, in the most advanced years [1]. However, there is some simple activities such as those devoted to introduce fundamental chemical concepts regarding molecular shape and atomic orbitals, which can be easily included in the curricula of secondary schools able to involve an students' profile with very little background in chemistry [3].

On the other hand, in the University it is common to find undergraduate physical chemistry programs integrating molecular modeling into their quantum mechanics curriculum [4]. Also, undergraduate life science programs such as the biochemistry [5], pharmaceutical [6], and medicinal chemistry [7, 8], among others, have been gradually emboiding most of the molecular modeling techniques to predic biological outcomes and elucidate details at atomistic level. Indeed, currently molecular modeling is being incorporated as an important part of the research in areas ranging from physics to chemistry to structural biology, as well as the drug design in the pharmaceutical industry.

Regarding undergraduate engineering programs molecular modeling usually is not explicitly included in the syllabus even though implicitly is present in many topics developed in the curricula. Altarawneh and Dlugogorski [9] show the importance of strategic applications of quantum chemistry in chemical industries and present an easy way to introduce chemical engineering students to illustrative case studies that deploy molecular modeling in the design of reactors and derivation of thermochemical functions. In fact, there are reported in the bibliography different experiences incorporating the molecular modeling into the chemical engineering curriculum with specialized curses presenting a comprehensive overview of the use of simulation techniques, including thermochemistry and reaction rate data, physisorption equilibria and diffusion rates, and transport properties [10]. Other initiatives were more in the way to build specialized front-end application to deal with quantum mechanical calculations exploring applications more focalized to the thermodynamic properties of molecules, thermochemistry of chemical reactions, and the optical and chemical properties of doped nanodiamonds [11].

In the present work a new optional course in molecular modeling of materials is offered to students coming from materials engineering, chemical engineering and biomedicine engineeries is shown as an ongoing project. Tuvi-Arad [12] presents three different models for including computational chemistry in the undergraduate program: the specialized course model, the augmented course model, and the

islands of computations model. Here, an specialized course is the way chosen to incorporate the theoretical and practical concepts. Along the course several problem-based and small research project-based learning were conducted. Students participating in an undergraduate research (e.g., in a computational molecular modeling lab) have shown to be able to develop better novel and independent projects later [5]. The development of critical thinking and the communication of scientific ideas in oral and written formats that are expanded in these types of projects are essential to ensure that students evolve into a greater understanding of basic scientific knowledge and the research process. Currently the molecular modeling research/training is less dependent on 'home-written' programs than it used to be, existing a large number of well-written and well-supported open source packages available to the user. Thus, facilitating and shortening the initial learning curve. In fact, Allen [2] shows a succesfully experience using such software packages for teaching molecular modeling to non-specialists students in a relatively short time.

2 IMPLEMENTATION OVERVIEW

2.1 Material and Methods

The main practical challenges to develop this course lay on both the computing environment and the computational tools that are involved in the computation chemistry applications. Independently of the science branch in which the course should be inserted, the main challenge is to shift the minds of the student from an application's-based environment (either *Windows* or *Apple*) to an intensive computing-based environment, i.e., *Linux*. Most students have never had contact with Linux Operating System (OS) and require building up the proper computing tools to emulate Linux OS within a Windows or Apple (*OS X*) environment. Two potential options may arise to achieve this goal, either building a dual mode configuration, with the inherent problematic of running two different and often incompatible, OS or building a virtual machine able to run in any given home-focused OS. In that regard, our university developed a universal web-based platform [13], which is a desktop virtualization solution developed from free software. This platform provides the user with a personal virtual desktop that runs on the remote server independently of the device (Windows, OS X, or Linux computer) from which the connection is made. Through this platform the students can emulate a Linux environment inside any personal platform that they may be using, either classroom desktops, personal laptops or even tablets.

Within this context, each virtual desktop contains all the Linux tools that the students will require to achieve the course goals. Moreover, their virtual machine allows an easy connection to our small computer cluster via ssh protocol (security shell connection [14]). The use of a computer cluster is a necessity in order to carry out simulations of actual meaning. One of the main drawbacks of computation chemistry, especially when performing Quantum Chemistry simulations (QM), is the necessity of using multiple computing threats to finish those simulations in acceptable time

frames. Solving thousands of complex non-liniar equations at the time implies the need of huge amount of RAM memory and many CPUs working in concert to achieve calculations speeds suitable for acceptable time costs.

The computational tools used for the proper development of the course will be distributed between the virtual desktop used by the students as main terminal and the advanced programs used in the cluster. The virtual machines will have all the resources of *Debian GNU/Linux10*, which includes *Mozilla Firefox* for internet access and scientific tools such as *Python* scripting, *GNUplot* for results graphical plots, *jmol* [15] for visualizing chemical structures, and *VMD* [16] for analyzing simulated trajectories. In the computing cluster, students will find all the scientific programs and suits necessary to learn and explore the basic approaches used in computational chemistry. This pool includes computer programs to perform Quantum Chemistry (QM) calculations, which consist in *NWChem* program [17] and Molecular Mechanics (MM) based programs. For the latter case, *AMBER* suite is preferred, version 20 [18], because not only performs all kinds of simulations within MM approach but also allows performing hybrid QM/MM simulations.

2.2 Course development

The structure and contents of the course are organized on active learning and solving problems. Each lesson is based on learning and assimilating new concepts through solving practical problems or performing realistic simulations. While unfolding the contents of each chapter, the students are engaged to solve specific problems and exercises that illustrate each new explained concept and at the end of each part the students must solve a final realistic simulation problem organized as an individual directed work (hereafter “practice”) that applies the acquired knowledge to a practical exercise, which at the same time would constitute a part of their academic evaluation. The final grade will be averaged over all the accumulated of individual grades, complemented with an oral presentation of a chosen subject (commonly a review of a scientific article), performed in small groups of 2 or 3 people. This methodological approach allows a progressive and integrative acquisition of new concepts based on a global vision of the subject and a direct practical application into the academic fields from which the students come from, since this subject is a multidisciplinary class and the general field of molecular simulations is used in a great variety of scientific applications.

3 RESULTS AND DISCUSSION

3.1 Subject implementation

The overall structure of the course and the timeframe used for each section is listed in Table 1.

In the initial sessions of the first section the students are introduced and familiarized with Linux Operating system, its basic commands, and the use of scripting as programming tools and execution of simulations. At the end of this section, the students are required to solve practical applications by using scripting, all gathered in

Practice 1: *Introduction to Linux. Based on scripting programming*. Most of the focus lies on a practical approach, in which the students are engaged in applied challenges to be solved using the commands they are introduced. For instance, the students are tought to use bash scripting to manage different commands within real simulation environment, which allows for multiple solutions to solve the same problem.

Table 1. Course structure and the timeframe section

#	Section	Time[a]	Assessment[b]
1	Introduction to molecular modeling	12	P
2	Classic model: conformational studies of materials and proteins	12	P
3	Quantum model: heterogeneous & enzymatic catalysis studies	12	P
4	Applications. Advanced simulation of materials	22	PJ
5	Research papers. Real cases under study.	2	RO

[a] Time in hours; [b] P, Problem-based; PJ, Project-based; RO, Report & Oral exposition

In the second section, the classical model applied to chemical structures is presented, showing the computation of structural energies using a Classical Force Filed, followed by the definition of simulation concept and the two main kinds of simulations used within the framework of *Molecular Mechanics* (MM), *Molecular Dynamics* (MD) simulations and *Monte Carlo* (MC) simulations. It is also at this point where the students are introduced to visual tools that allow the representation and visualization of crystals, amorphous materials and proteins. The final activity of this chapter is summarized in Practice 2: *Structural study of a polymer, crystal and/or protein*. In this final practice the students are challaged to study the effect of several enviromental factors, such as ionic strength or temperature, over the conformational stability of an small protein using AMBER program (Fig. 1)

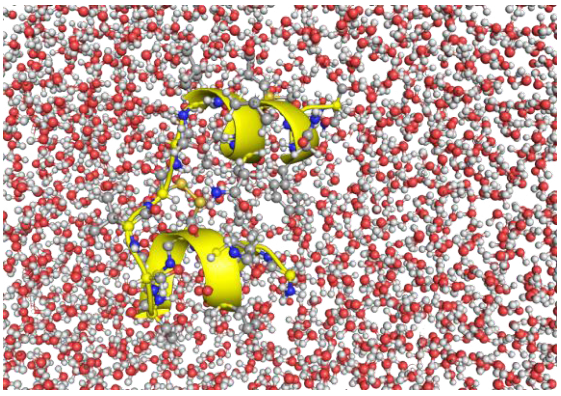

Fig. 1.- Human Insuline immersed in water filled box

In the third section of the course the students are presented with the challenges of modeling of thermodynamic properties, chemisorption on different materials and chemical reactions, which implies solving Schrödinger Equation. A quick review to its exact solution for the hydrogen atom is used to present the conundrum of setting the

proper Hamiltonian in polyelectronic systems, leading to the exploration of the different quantum simulation techniques, going from Hartree-Fock based methodologies, their approximated solution through Semiempirical methods and finally the introduction to Density Functional Theory and its applications. This section is completely self-guided through different exercises of growing complexity starting with the atomic properties such as the atom multiplicity, solving the IR spectrum of some molecules, studying the geometry of organometallic complexes, and ending by obtaining the transition state of a reaction in gas phase. The chapter is closed by solving Practice 3. *Simulation of a heterogeneous and/or catalytic process* (Fig. 2).

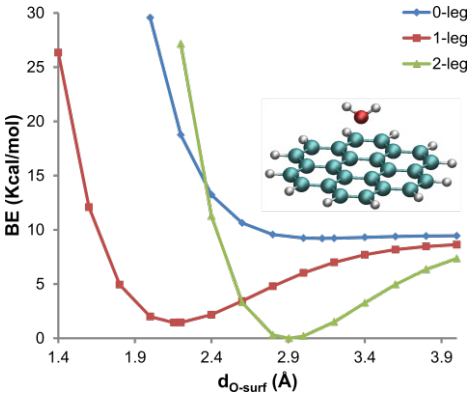

Fig. 2. - Chemisorption of water in graphene[19]

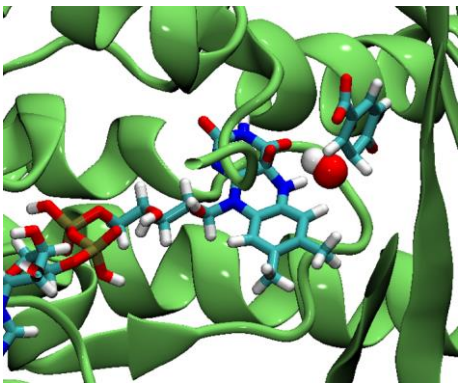

Fig. 3. - Oxygenation reactions catalysed by the enzyme p-hydroxybenzoate hydroxylase[20]

The fourth section of the course is divided in two different parts. A first section showing other computational tools used in the Biological and Material sciences, e.g., polymer and biopolymer simulations, enzyme catalysis (Fig. 3) and structural bioinformatics. This part is organized in demo-pills with all the material and files made available to the students, thus allowing a later reproducibility. Structural bioinformatics is emphasizing the techniques involved in homology modeling, fast threading (Fig. 4), and computational aided drug design. A second and last part, that revolves around solving a practical scientific case using the tools and techniques learnt throughout the course. This application is organized in a final simulation project (see section 3.2) that constitutes Practice 4.

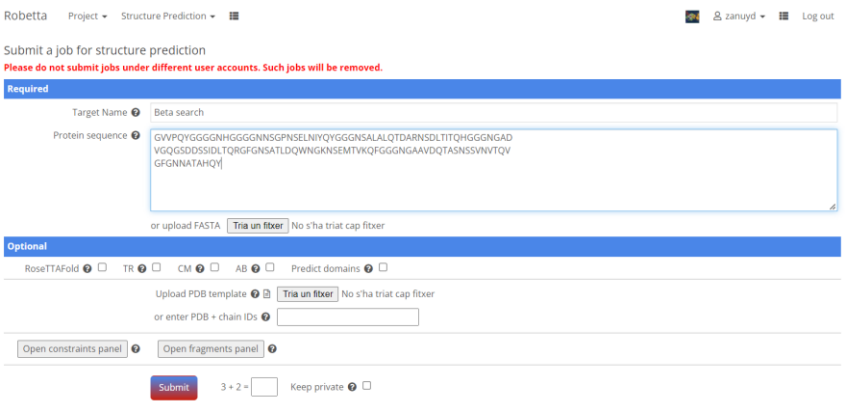

Fig. 4. - Online interface for Robetta project computations for fast threading.

The last section is devoted to studying different research works on molecular modeling, published in peer-reviewed journals, to complement the learning of those new methodologies not explored throughout the course. The students, organized in very small groups, choose a research project at the beginning of the course among several according to their own scientific interest. They try to understand the methodology used, the results obtained, as well as the conclusions reached by the authors of the article. On the last day, together with a report, it is exposed to the rest of the classmates, which allows a small debate to be opened about it.

3.2 Directed Learning Activity: Simulation Project

The final simulation project is the longest directed learning activity of the course and it is conducted after the students have gained most of the theorethical and practical concepts. Each student choses one topic among several projects (all with different development) and laying in different fields within molecular modeling, e.g., material science, structural biology, and bioinformatics. All the methodology to be used in this section has been acquired within the subject and the main difficulty developing the project is laying in the comprehension of the studied concepts. Usually, at this point the students get much involved in the topics, thus gaining a strong command of molecular modeling. Indeed, an improvement in critical thinking and a greater knowledge of the research process by students was observed at the end of the course, which was reflected in the work developed in Practice 4. *Advanced simulation of materials*.

3.3 Course Assessment

The subject has been conducted along three academic years involving a total amount of 29 students. One of the most difficult point that students have to face in this multidisciplinary course, beside of revive the basic buried concepts of chemistry and physics in the first years of the undegraduate program, is the missing unawareness on basic computational science. The student is very familiar with the visual informatics as common user but fails in the command and scripting languagues, which are necessary in this kind of simulation, where most of the applications are running in Linux OS. Thus, facing the Linux OS is one of the first challenges that the student has to deal with. The continous learning curve observed is mainly based in the computer lab experience, adquired along the subject development. Indeed, the subject assessment is trying to reinforce the working lab through proposing to the student different problems to be solved, as well as a small project that will be evaluated individually, coinciding with the five sections in which the subject is divided (see Table 1).

Finally, an equitable evaluation of the five sections of the subject is carried out. Considering this way, the part of individual development, both of the short problems at the end of each main topic (sections 1 to 3) and the application project (section 4). The assessment ends with the commentary in group on the research article (section 5). Spite of most of the developed work is evaluated independently, there is a common background in the problem description with a different problem

development that would allow to the students to discuss and work in their own project at the same time that are working in group to obtain a final solution. Differences between their scientific background and origin (10 % students were in an Erasmus program) reinforced the collaborative work at the time that increased the group motivation.

We assessed the degree of satisfaction of our students using a enquiry that we specifically designed for this porpouse. The general level of satisfaction for this new class was between high or very high. Some students had some missgivings regarding the depth of the theoretical contents and suggested a more applied approach. However, this particular question becomes quite unfeseable due to the lack of background in statistical mechanics and quantum mechanics that they had from previous years. There is though consensus about both the utility of the applied examples and the direct applications practical of the newly accuaried knowledge. Throughout these three years most of the suggestions raised by our former students have been incorporated. However, it is really tough to only focus on practical problems when there are important lacks on their background.

4 SUMMARY AND ACKNOWLEDGMENTS

In this work the introduction of a multidisciplinary optional subject straddling different teaching programs of different engineering careers is presented. The whole subject is supported by short problems- and projects-based learning that the student carries out to complete their learning curve. The constant challenge designed to foster scientific discussions among them from different students backgrounds, working together to solve a problem, increases their motivation for the subject. As a result of these first three years has been the presentation at the end of the course of elaborate simulation projects with a high dose of independence, development of basic scientific knowledge and research process.

REFERENCES

[1] Taly, A., Nitti, F., Baaden, M., S., P. (2019) Molecular modelling as the spark for active learning approaches for interdisciplinary biology teachingInterface. *Interface Focus* **9**, 20180065-20180065.

[2] Allen, M. P. (2007) Educational aspects of molecular simulation. *Mol. Phys.* **105**, 157-166.

[3] Ochterski, J. W. (2014) Using Computational Chemistry Activities To Promote Learning and Retention in a Secondary School General Chemistry Setting. *J. Chem. Educ.* **91**, 817-822.

[4] Johnson, L. E., Engel, T. (2011) Integrating Computational Chemistry into the Physical Chemistry Curriculum. *J. Chem. Educ.* **88**, 569-573.

[5] Brown, A. M., Lewis, S. N., Bevan, D. R. (2016) Development of a structured undergraduate research experience: Framework and implications. *Biochem. Mol. Biol. Educ.* **44**, 463-474.

[6] Tsai, C. S. (2007) Using Computer Applications and Online Resources To Teach and Learn Pharmaceutical Chemistry. *J. Chem. Educ.* **84**, 2019.

[7] Bisht, N., Sah, A. N., Bisht, S., Joshi, H. (2021) Emerging Need of Today: Significant Utilization of Various Databases and Softwares in Drug Design and Development. *Mini Reviews in Medicinal Chemistry* **21**, 1025-1032.

[8] Hayes, J. M. (2014) An Integrated Visualization and Basic Molecular Modeling Laboratory for First-Year Undergraduate Medicinal Chemistry. *J. Chem. Educ.* **91**, 919-923.

[9] Altarawneh, M., Dlugogorski, B. Z. (2018) Introducing Quantum Chemistry in Chemical Engineering Curriculum. *J. Chem. Educ.* **95**, 1562-1571.

[10] Baldwin, R. M., Ely, J. F., Way, J. D., Daniel, S. R. (2000) Incorporating Molecular Modeling Into the Chemical Engineering Curriculum. *Chemical Engineering Education* **34**, 162-167.

[11] Simeon, T., Aikens, C. M., Tejerina, B., Schatz, G. C. (2011) Northwestern University Initiative for Teaching NanoSciences (NUITNS): An Approach for Teaching Computational Chemistry to Engineering Undergraduate Students. *J. Chem. Educ.* **88**, 1079-1084.

[12] Tuvi-Arad, I. (2021) Computational Chemistry in the Undergraduate Classroom – Pedagogical Considerations and Teaching Challenges. *Isr. J. Chem.* **61**, 1–10.

[13] RaVaDa: https://cbl.upc.edu/ca/intranet/pas/area-tecnica/serveis/descripcio-intranet-dels-serveis/escriptoris-virtuals-ravada

[14] Ylonen, T., USENIX Association.

[15] Jmol: an open-source Java viewer for chemical structures in 3D.

[16] Humphrey, W., Dalke, A., Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Molec. Graphics* **14**, 33-38.

[17] Aprà, E., Bylaska, E. J., Jong, W. A. d., Govind, N., Kowalski, K., Straatsma, T. P., Valiev, M., Dam, H. J. J. v., Alexeev, Y., Anchell, J., Anisimov, V., Aquino, F. W., Atta-Fynn, R., Autschbach, J., Bauman, N. P., Becca, J. C., Bernholdt, D. E., Bhaskaran-Nair, K., Bogatko, S., Borowski, P., Boschen, J., Brabec, J., Bruner, A., Cauët, E., Chen, Y., Chuev, G. N., Cramer, C. J., Daily, J., Deegan, M. J. O., DunningJr., T. H., Dupuis, M., Dyall, K. G., Fann, G. I., Fischer, S. A., Fonari, A., Früchtl, H., Gagliardi, L., Garza, J., Gawande, N., Ghosh, S., Glaesemann, K., Götz, A. W., Hammond, J., Helms, V., Hermes, E. D., Hirao, K., Hirata, S., Jacquelin, M., Jensen, L., Johnson, B. G., Jónsson, H., Kendall, R. A., Klemm, M., Kobayashi, R., Konkov, V., Krishnamoorthy, S., Krishnan, M., Lin, Z., Lins, R. D., Littlefield, R. J., Logsdail, A. J., Lopata, K., Ma, W., Marenich, A. V., Campo, J. M. d., Mejia-Rodriguez, D., Moore, J. E., Mullin, J. M., Nakajima, T., Nascimento, D. R., Nichols, J. A., Nichols, P. J., Nieplocha, J., Otero-de-la-Roza, A., Palmer, B., Panyala, A., Pirojsirikul, T., Peng, B., Peverati, R., Pittner, J., Pollack, L., Richard, R. M., Sadayappan, P., Schatz, G. C., Shelton, W. A., Silverstein, D. W.,

Smith, D. M. A., Soares, T. A., Song, D., Swart, M., Taylor, H. L., Thomas, G. S., Tipparaju, V., Truhlar, D. G., Tsemekhman, K., Voorhis, T. V., Vázquez-Mayagoitia, Á., Verma, P., Villa, O., Vishnu, A., Vogiatzis, K. D., Wang, D., Weare, J. H., Williamson, M. J., Windus, T. L., Woliński, K., Wong, A. T., Wu, Q., Yang, C., Yu, Q., Zacharias, M., Zhang, Z., Zhao, Y., Harrison, R. J. (2020) NWChem: Past, present, and future. *J. Chem. Phys.* **152**, 184102.

[18] Case, D. A., Aktulga, H. M., Belfon, K., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham-III, T. E., Cisneros, G. A., Cruzeiro, V. W. D., Darden, T. A., Duke, R. E., Giambasu, G., Gilson, M. K., Gohlke, H., Goetz, A. W., Harris, R., Izadi, S., Izmailov, S. A., Jin, C., Kasavajhala, K., Kaymak, M. C., King, E., Kovalenko, A., Kurtzman, T., Lee, T. S., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Machado, M., Man, V., Manathunga, M., Merz, K. M., Miao, Y., Mikhailovskii, O., Monard, G., Nguyen, H., O'Hearn, K. A., Onufriev, A., Pan, F., Pantano, S., Qi, R., Rahnamoun, A., Roe, D. R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C. L., Skrynnikov, N. R., Smith, J., Swails, J., Walker, R. C., Wang, J., Wei, H., Wolf, R. M., Wu, X., Xue, Y., York, D. M., Zhao, S., Kollman, P. A., (2021) Amber 2021, University of California, San Francisco.

[19] Brandenburg, J. G., Zen, A., Fitzner, M., Ramberger, B., Kresse, G., Tsatsoulis, T., Grüneis, A., Michaelides, A., Alfè, D. (2019) Physisorption of Water on Graphene: Subchemical Accuracy from Many-Body Electronic Structure Methods. *J. Phys. Chem. Lett.* **10**, 358-368.

[20] Billeter, S. R., Hanser, C. F. W., Mordasini, T. Z., Scholten, M., Thiel, W., van Gunsteren, W. F. (2001) Molecular dynamics study of oxygenation reactions catalysed by the enzyme p-hydroxybenzoate hydroxylase. *Phys. Chem. Chem. Phys.* **3**, 688-695.