Coupling density functional based tight binding with class 1 force fields in a hybrid QM/MM scheme

Maysa Yusef-Buey, Tzonka Mineva, Mathias Rapacioli

Theor Chem Acc manuscript No.
(will be inserted by the editor)

Coupling Density Functional based Tight Binding with Class 1 Force Fields in a hybrid QM/MM scheme

Maysa Yusef-Buey, Tzonka Mineva, Mathias Rapacioli

Abstract QM/MM (Quantum Mechanics/Molecular Mechanics) hybrid methods have become very popular schemes to incorporate environmental effects in the calculation of molecular properties, when it is mandatory to have both a quantum description of electrons to compute these properties and an atomistic description of the environment. However, even Density Functional Theory/MM schemes may become timecosting when a large part of the system should be treated at the QM level or when plenty of single point energy calculations are intended to be done. We report a new implementation, within the deMonNano code, of a hybrid QM/MM scheme combining the density functional based tight binding with class 1 force fields. Two types of additive couplings can be chosen, namely the mechanical coupling, consisting of a Lennard-Jones potential and the electrostatic coupling, in which the MM part of the system is also polarizing the region described at the QM level. As first test-case application, the harmonic infrared spectra of simple molecules in the gas phase and in water clusters are computed and compared to those obtained at the DFT/MM level. Binding energies are also compared. Similar trends are obtained with the two levels of calculations and the main differences are discussed.

Keywords Density Functional based Tight binding · hybrid QM/MM

M.Yusef-Buey · M. Rapacioli (✉)
Laboratoire de Chimie et Physique Quantique (LCPQ/IRSAMC), UMR5626, Université de Toulouse (UPS) and CNRS, 118 Route de Narbonne, F-31062 Toulouse, France
E-mail: mathias.rapacioli@irsamc.ups-tlse.fr

T. Mineva (✉)
ICGM, Univ. Montpellier, CNRS, ENSCM, Montpellier, France
E-mail: tzonka.mineva@enscm.fr

1 Introduction

The hybrid quantum mechanics/molecular mechanics (QM/MM) computational schemes owing their origin in the works of Warshel and Karplus [1] and Warshel and Levitt [2] incorporate simultaneously the efficiency of molecular mechanics algorithms and the accuracy of *ab initio* methods. The QM/MM approach is attractive with its simple idea to combine different levels of atomistic approaches and beneficial for large size systems requiring to explicitly consider the electrons only around a reactive center embedded in the electrostatic field of the environment, treated at molecular mechanics levels. The total energy of a QM/MM system ($E^{QM/MM}$) is obtained as sum of the energies of QM (E^{QM}) and MM (E^{MM}) subsystems, corrected by the energy term (E^{QM-MM}), accounting for the interaction between the two subsystems:

$$E^{QM/MM} = E^{QM} + E^{MM} + E^{QM-MM} \quad (1)$$

The main question is how to compute the interaction energies between QM and MM regions. Among the different implementation schemes, two approaches emerged: the subtractive and the additive [3]. In the former approach, the QM-MM interaction is obtained from the difference between the potential energy of the entire (QM + MM) system and the potential energy of the QM subsystem, both calculated at the MM level. An example of subtractive scheme is ONIOM [4]. The additive scheme is an embedding scheme considering the QM subsystem embedded by the MM subsystem and the QM-MM interactions are calculated explicitly at different levels of complexity by including mechanical, electrostatic and/or polarisation embedding techniques. Here, we have chosen to implement the additive QM/MM approach. The computations of energy gradients of MM atoms and their coupling to the energy

gradients of QM atoms is straightforward. Geometrical optimisation and dynamics simulations based on both subtractive and additive DFT/MM schemes have been employed with success to study reactivity, spectroscopic features [5–7], to explore conformational space or to compute free energy profiles [8, 9].

In DFT/MM schemes, most of the computational time is related to the calculation of the DFT energy, and its gradients when they are needed. When many single point energy calculations are intended to be done, like in global exploration searches or in long molecular dynamics (MD) simulations, or when the region treated at the QM level contains many atoms, it is appealing to replace the DFT QM scheme by less computationally expensive methods. In this domain, the Density Functional based Tight Binding (DFTB) is a well suited scheme [10–14]. DFTB can be seen as an approximated DFT, whose computational efficiency arises from the use of a minimal valence atomic basis (matrix diagonalisation scales as the cube of the number of basis functions) and the use of tabulated values (the so-called Slater Koster tables [15]) to avoid explicit calculations of integrals. With respect to force fields and empirical based tight binding schemes, a main advantage of DFTB is it's *ab initio* grounding as it is formally derived from DFT allowing for a greater transferability. As a consequence, it was shown to be a powerful method to simulate various processes, like chemical reactivity [16] or fragmentation [17], to compute different spectral properties, like electronic [18], harmonic[19, 20] and anharmonic [21] IR and Raman [22] spectroscopy, to derive ionisation and electronic affinity energies [23], and to be valuable method to address various scientific fields ranging from material science [24] to biology [25].

Several implementations of DFTB/MM schemes within various MM packages such as AMBER [26, 27], GROMACS[28], CHARMM [29] or the coupling of DFTB with the Universal Force Field (UFF) scheme in the deMonNano code [30] have been previously reported. Let us also mention the coupling of DFTB with a polarisable force field to account for Argon matrix environment [31]. Other strategies have been used to consider the perturbation of a system described at the DFTB level by its environment, such as implicit models based on a continuous description of the medium [32, 33] or models based on a coarse-graining strategy [34]. DFTB has also been included in QM/QM' schemes, acting as the lower level QM' method when combined with DFT [35].

In the present work, we report the coupling of DFTB with various force fields of class 1 (inter-molecular interactions are obtained by harmonic terms only and Lennard-Jones 6-12 potential), its implementation within the deMonNano code [36] and first applications to test cases. The present implementation is based on an additive coupling scheme, differing from the previous DFTB/UFF scheme within the same deMonNano code based on a substractive scheme [30]. In addition, its implementation allows the use of any kind of parametrized class 1 force fields, as long as the user provides the appropriate parameters file. Furthermore, unlike some DFTB/MM coupling schemes combining different softwares [37, 38], we report, a DFTB/MM hybrid implementation in a single code, as previously done in some other softwares [29, 39, 40]. It avoids the use of any interface, reducing the loss of computational efficiency associated to communication processes. We compare harmonic vibrational spectra obtained at the DFT/MM and DFTB/MM levels for simple molecules, isolated or trapped in water clusters. In the next section, the basics of DFTB and MM methods are briefly reviewed before presenting the coupling strategy employed. Details about DFT/MM calculations are also given as well as the strategy used to compute energetic properties and harmonic spectra at the DFTB/MM level and to provide cross-comparisons with DFT/MM results. The laters are shown and discussed in section 5.

2 Incode additive DFTB/MM implementation

2.1 Density Functional based Tight Binding

An extended description of the DFTB method can be found in several review papers [13, 14, 41, 42]. This theory was build in the 90's [10, 11] following the pioneering works of Foulkes and Haydock [43]. In this scheme, the molecular orbitals (MOs) $\{\phi_i\}$ are expanded on a non-orthogonal minimal atomic valence basis $\{\varphi_\mu\}$: $\phi_i = \sum_\mu c_{i\mu}\varphi_\mu$, where $c_{i\mu}$ are the MO's decomposition coefficients. The energy expression relies on a Taylor expansion of the effective Kohn-Sham potential around a reference density ($\rho = \rho_0 + \delta\rho$) which can be performed up to the first, second or third order, leading to the various levels of DFTB, known as zeroth-order DFTB[10, 11], self-consistent charges (SCC)-DFTB [44] and third order DFTB (also called DFTB3) [45, 46]. In the widely used SCC-DFTB scheme, the energy is expressed as :

$$E = \sum_{a,b} E_{rep}(R_{ab}) + \sum_{i,\mu,\nu} n_i c_{i\mu} c_{i\nu} H^0_{\mu\nu} + \frac{1}{2}\sum_{a,b} \gamma_{ab} q_a q_b \tag{2}$$

In this expression, $E_{rep}(R_{ab})$ is a two-body repulsive contribution depending on the distance between atoms

a and b $H^0_{\mu\nu}$ is a matrix element of the Kohn-Sham operator taken at the reference density ρ_0, n_i is the occupation of the ith MO, q_a is the atomic charge of atom a computed from the MO's coefficients within the Mulliken scheme and γ_{ab} is a diatomic matrix accounting for the long range coulomb interaction between fluctuating charges plus some short range contributions of exchange correlation energy, the onsite term being equal to the atomic chemical hardness. The neglect of 3-centers integrals allows to parameterise the matrix elements of H^0 from atomic and diatomic DFT calculations and to write the second order contribution as a function of atomic charges. The electronic problem is then solved by applying the variational principle to this energy expression leading to the following secular equation :

$$(H^0 + H^1(q))C_i = \epsilon_i S C_i \tag{3}$$

Where S is the atomic orbital overlap matrix and C_i the vector containing the coefficients of the i^{th} MO. H^1 is the contribution to the operator arising from the second order energy contribution (last term in Eq. 2) and can be written as :

$$H^1_{\mu\nu}(q) = \frac{1}{2} S_{\mu\nu} \sum_c q_c(\gamma_{ac} + \gamma_{bc}) \tag{4}$$

with $\mu \in a$ and $\nu \in b$. In the zeroth-order DFTB, the second order term is neglected and the operator is restricted to H^0. In the DFTB3 scheme [45, 46], including third order terms in the DFT Taylor expansion allows to consider the fluctuation of the Hubbard parameters with the electronic density expressed again as a function of the atomic charges q. This results in adding a new contribution to the DFTB energy expression (Eq. 2) and its corresponding perturbation $H^2(q)$ to the DFTB operator (Eq. 3). In the following, E^{DFTB} and H^{DFTB} can refer to either zeroth-order DFTB, SCC-DFTB or third order DFTB.

In practice, the results presented in section 5 were obtained from the third order DFTB combined with a correction for the γ matrix as recommended in reference [46]. We used the 3ob parameter set [47] downloaded from the *www.dftb.org* website.

2.2 Molecular Mechanics (MM)

In this section we briefly recall the fundamentals of the MM approach. The MM potential energy function is computed from the classical pairwise bonded and non-bonded inter-atomic interactions [48]. The general expression is given by:

$$E_{MM} = \sum_i^{N_{bond}} V_i^{N_{bond}} + \sum_j^{N_{angles}} V_j + \sum_l^{N_{tors}} V_l^{N_{tors}} + \sum_l^{N_{imp}} V_l^{N_{Imp}} + \sum_c^{N_{MM}} \sum_{d>c}^{N_{MM}} V_{cd}^{Coul} + \sum_c^{N_{MM}} \sum_{d>c}^{N_{MM}} V_{cd}^{LJ} \tag{5}$$

The first four terms correspond to the bond stretching and angle bending contributions as well as to the dihedral and improper torsion (also known as local or short-range) contributions. In class 1 FFs the bond stretching and angle bending terms are computed from harmonic functions [49]:

$$V^{bond} = \frac{1}{2} k_b (r - r_0)^2 \qquad V^{angles} = \frac{1}{2} k_a (\theta - \theta_0)^2 \tag{6}$$

where r -r_0, $\theta - \theta_0$ are the deviations from the equilibrium bond distances and valence angles, and k_b,k_a are the bond and angle force constants, respectively [50]. The torsion (dihedral and improper) potentials are modelled from periodic cosines functions :

$$V^{tors} = \frac{V_n}{2}[1 + cos(n\phi - \delta)] \tag{7}$$

where ϕ is the dihedral or torsional angle, n is an integer, non-negative constant, indicating the periodicity, δ is the phase and V_n is the height of the rotational potential barrier [49]. The last two terms in Eq.5 correspond to the Coulomb and the Van der Waals (vdW) interactions (non-bonding interactions). VdW are modelled by 12-6 Lennard-Jones potential as a combination of Pauli short-range repulsion and dispersion long-range attraction, given by the respective C_{12}^{cd} and C_6^{cd} parameters, and R_{cd} the distance between atoms c and d. [3, 49]:

$$V_{cd}^{LJ} = \left(\frac{C_{12}^{cd}}{R_{cd}}\right)^{12} - \left(\frac{C_6^{cd}}{R_{cd}}\right)^{6} \tag{8}$$

An advantage of MM approach is that the bonding and non-bonding interactions are computed using different parameter sets, obtained from *ab intio* or semi-empirical QM calculations, as well as from fitting experimental data. [49]. Because we have coupled DFTB with class 1 FFs we presented above the energy potential terms only for this class of force fields. Some examples of class 1 FFs are OPLS, AMBER, CHARMM, GROMOS, etc [3].

The general expression for the MM potential can be tuned, adding new terms in order to increase the level of accuracy of the molecular interactions. This is for example, the case of Urey-Bradly interactions [49]

$$V_{UB} = \sum_{angles} \frac{1}{2} k_{UB}(s - s_0)^2, \tag{9}$$

where s is the distance between two non-bonded atoms (1, 3) and k_{UB} is the force constant parameter.

As a final appreciation, in MM approaches, the electronic degrees of freedom are not considered explicitly. As a consequence, the impact of the E_{MM} energy contribution on the secular equation can be, thus, ignored and its contribution to the total energy can be derived straightforwardly. [51]

2.3 DFTB/MM couplings

In this section our implementations of DFTB/MM mechanical and electrostatic embedding schemes are presented.

2.3.1 Mechanical coupling

The simplest mechanical embedding comprises only the non-bonding interactions between QM and MM regions, thus the QM subsystem is not polarized by the MM subsystems and vice versa [3]. The E^{QM-MM} term in the DFTB/MM coupling, referred throughout this work as $E^{QM-MM/mech}$ is :

$$E^{QM-MM/mech} = \sum_{a\in QM}\sum_{c\in MM}\left(\frac{C_{12}^{ac}}{R_{ac}}\right)^{12} - \left(\frac{C_{6}^{ac}}{R_{ac}}\right)^{6} \tag{10}$$

2.3.2 Electrostatic coupling

The Coulomb interaction between the MM point charges and the DFTB density should in principle involve Coulomb integrals over the DFTB density described in the atomic basis. However, following the DFTB approximation used to compute the coulomb term in the second order energy correction, the DFTB density is approximated for this term by atomic charges using the Mulliken scheme formalism to replace the electronic density by atomic monopoles. It leads to the following expression for the QM/MM Coulomb potential energy :

$$E^{QM-MM/elec} = \sum_{a\in QM}\sum_{c\in MM}\frac{q_a q_c}{R_{ac}} \tag{11}$$

arising from the interaction between MM atomic charges (q_c) and DFTB atomic charges (q_a), derived from the electronic density in the DFTB region, is added to the Lennard-Jones potential (Sec. 2.3.1). As the atomic charges from the QM region are optimised self-consistently, the DFTB operator used to solve the secular Eq. 3 contains an additional term leading to the following expression for the matrix elements in the atomic basis :

$$H_{\mu\nu} = H_{\mu\nu}^{DFTB} + H_{\mu\nu}^{QMMM/elec}$$

with

$$H_{\mu\in a;\nu\in b}^{QMMM/elec} = \frac{1}{2}S_{\mu\nu}\sum_{c\in MM} q_c\left(\frac{1}{R_{ac}} + \frac{1}{R_{bc}}\right) \tag{12}$$

The analytical gradients have been derived and implemented for both mechanical and electrostatic embedding schemes. Benchmarks are reported in the supplementary material, comparing the calculation of the energy gradients from the analytical implementation and from numerical finite differences as well as the conservation of the total energy during MD micro-canonical dynamics. We note that the present implementation benefited in part from the former SCC-DFTB/UFF implementation within the deMonNano code [30] and from the MM implementation within the deMon2k code [52] Actual version of DFTB QM/MM hybrid method implementation is already available in the deMonNano website [1][36]

3 DFT computational details

The DFT calculations were carried out with the deMon2k code [53] including the incode DFT/MM [52]. The PBE exchange correlation functional [54] was used in combination with the DZVP basis set [55]. Automatically generated auxiliary basis sets were employed [56]. An adaptive grid for numerical quadrature was used. The SCF and gradient convergence tolerances were set to 10^{-7} and 10^{-5} a.u, respectively.

4 Computational Strategy

To validate the implemented scheme, we have compared DFT/MM and DFTB3/MM calculations, with the electrostatic coupling scheme. OPLS-AA (TIP3P for water molecules) has been used for the MM potential in both cases. The test systems consist of small molecules, namely ammonia, glycine, pentane and water, isolated (gas phase) or trapped in a water cluster (between 205-215 water molecules). The goal here is to evaluate the effects of a water cluster environment, treated at MM level, on the infrared (IR) spectra of the solute molecules using either DFT/MM or DFTB3/MM

[1] The present version of the code corresponding to this work can be downloaded from deMonNano website (10.12.2021) and corresponds to git hash commit f3707842dd69b9629954a55a28693189d572f425

schemes and to compute the corresponding binding energies.

A complete study of such systems would require global explorations of the complex potential energy surfaces, which is out of the scope of the present work with the focus on the in-code coupling between DFTB3 and MM schemes. Therefore, the procedure to generate the geometrical configurations prevents the use of the present results beyond methods comparison. In practice, our benchmark approach has two steps. First, the selected small molecules were optimized at both the DFT and DFTB3 levels in gas-phase. In the second step, each molecule was introduced in a water cluster. A short molecular dynamics (50 ps) was performed at the DFTB3/MM level in the canonical ensemble (temperature of 60 K maintained by a chain of 5 Nose Hoover thermostats [57–59] with energy exchange frequency of 800 cm^{-1}). The final geometry was optimised at the DFTB3/MM level. Finally, these geometries were further optimised at the DFT/MM level. All the cluster's water molecules were treated at the MM level. For these optimised geometries in gas-phase and in water clusters, the IR spectra were computed in the harmonic approximation, involving the diagonalisation of the mass weighted Hessian matrix. In both, DFT/MM and DFTB3/MM implementations, only the dipole moments arising from the solute molecule, therefore the QM dipole moments, were considered in the calculations of the IR intensities. The relevant information is therefore the solvation effects on the IR spectra of the solute molecules and not the spectrum of the full cluster.

5 Applications to test cases

5.1 Ammonia

Figure 1-top compares the ammonia IR spectra computed for the gas phase ammonia at the DFT and DFTB3 levels. At both the DFT and DFTB3 levels, the most intense band corresponds to N-H wagging vibrational mode and the frequencies are in good agreement (1022 cm^{-1} for DFTB3 vs 1030 cm^{-1} for DFT). A second band, by far less intense corresponds to H-N-H scissoring modes and the even less intense bands (almost no intensity at the DFT level) located around 3400 cm^{-1} result from N-H stretching modes. When NH_3 is inside the cluster, the main band (wagging mode) computed at the DFT/MM level is shifted towards higher frequencies (figure 1-bottom). This trend is captured by the DFTB3/MM spectra (figure 1-top), although the shift is smaller (93 cm^{-1} vs 187 cm^{-1} at the DFT/MM level). The intensity of this band is increased at both the DFT/MM and DFTB3/MM levels with respect to the isolated molecule spectrum. Very few changes are observed for the scissoring bands whatever the potential used. The intensities of the stretching modes are increased with either DFT/MM or DFTB3/MM. The last effects of the environment is the appearance of bands below 500 cm^{-1} visible in the DFT/MM spectrum and present, although much less intense, in the DFTB3/MM spectrum. These frequencies correspond to soft intermolecular modes. In some of them, the ammonia presents rotation and translation as a rigid body, whereas some others only correspond to intermolecular vibrations of water molecules without the involvement of displacements of ammonia atoms. In the latter modes, only the ammonia electronic density is perturbed. A little contribution from water environment is also observed at both the DFTB3/MM and DFT/MM levels around 3855 cm^{-1} and corresponds to combinations of the water stretching modes, inducing ammonia dipole fluctuations. Similarly, at both levels of theory, the small feature around 2100 cm^{-1} is associated to combinations of bending modes of many water molecules.

5.2 Neutral Glycine

The results obtained for the neutral glycine molecule are presented in Figure 2. Focusing first on the isolated molecule (top), it can be seen that the DFT spectrum presents an intense band at 3242.5 cm^{-1} assigned to the O-H stretching mode. This band is also among the most active ones at the DFTB3 level, although presenting a smaller intensity as well as a blueshift (3523.2 cm^{-1}) with respect to the DFT results. Apart from this band, the DFT neutral glycine spectrum is dominated by two modes corresponding mostly to the displacement of a carbon atom from the carboxyle group at 1795 cm^{-1} and to a C-O-H bending mode at 1427 cm^{-1}. At the DFTB3 level, these two modes are the most active ones located at 1702 and 1252 cm^{-1}, respectively. Finally, several other modes are observed at lower frequencies in both spectra, the main difference being related to the modes around 900 cm^{-1} whose intensities are underestimated at the DFTB3 level in comparison to their DFT intensities.

When glycine is inside the water cluster, both methods lead to the same trends for the O-H streching mode, namely an increase of the intensity and a shift toward lower frequencies, the latter being stronger at the DFT/MM (395 cm^{-1}) level than at the DFTB3/MM level (162 cm^{-1}). The band at 1795 cm^{-1} (DFT value) is also redshifted in the cluster at both DFT/MM and

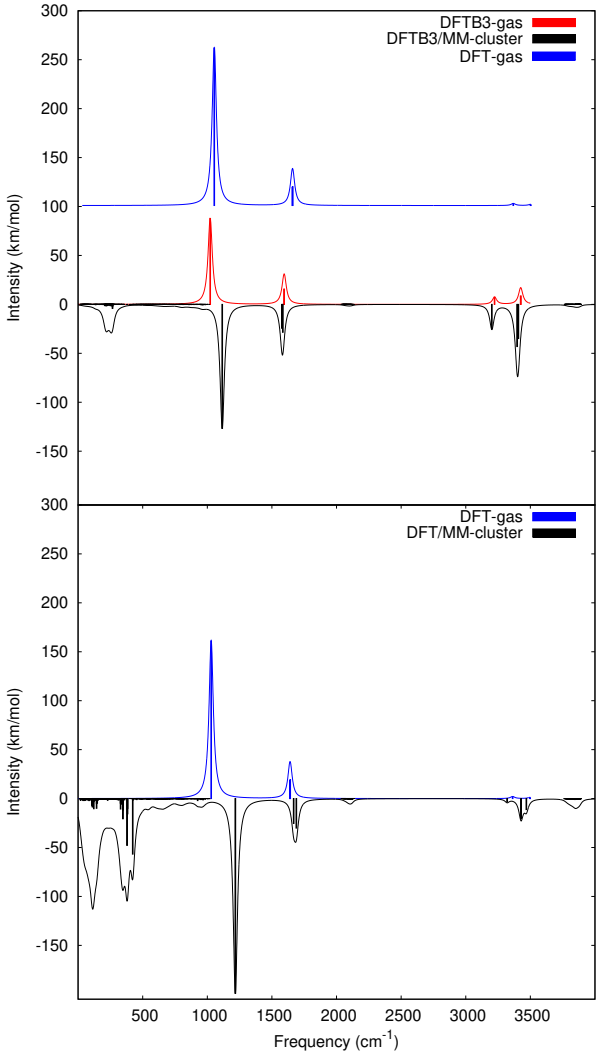

Fig. 1 Harmonic IR spectra for ammonia. Top : gas phase molecule at the DFT and DFTB3 levels and solvated ammonia at DFTB3/MM levels; Bottom : gas phase and solvated ammonia at the DFT and DFT/MM levels. Harmonic stick spectra were convoluted with Lorentzian function (FWMH=20 cm^{-1}) to produce the line spectra. For further analysis find in supplementary material comparison between DFTB3/MM and DFT/MM spectra for ammonia system

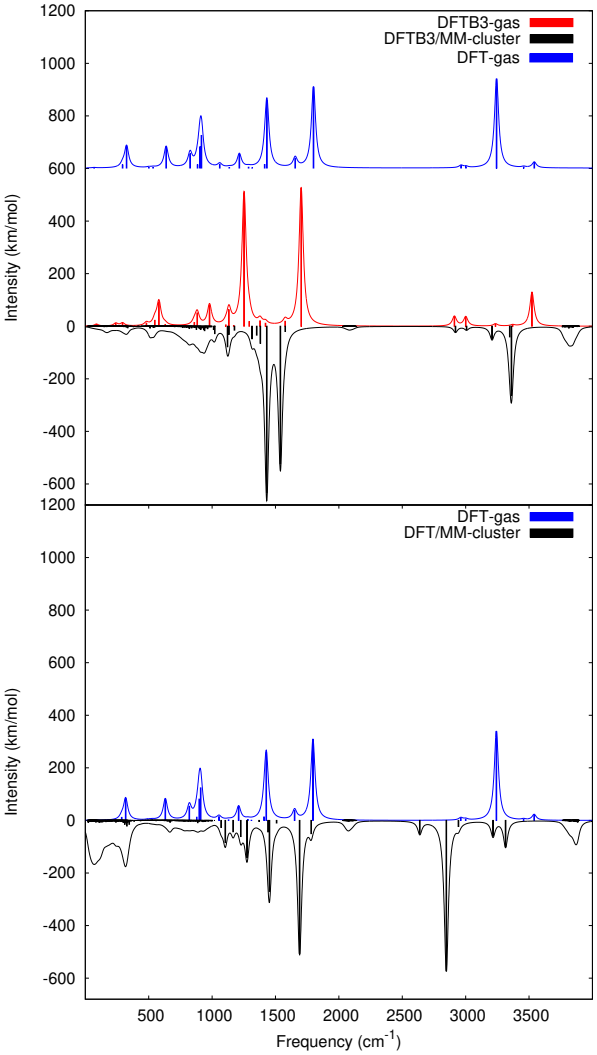

Fig. 2 Harmonic IR spectra for neutral glycine. Top : gas phase molecule at the DFT and DFTB3 levels and solvated glycine at DFTB3/MM levels; Bottom : gas phase and solvated glycine at the DFT and DFT/MM levels.. Harmonic stick spectra were convoluted with Lorentzian function (FWMH=20 cm^{-1}) to produce the line spectra.

DFTB3/MM levels, whereas a slight blue shift of the C-O-H bending mode is only observed at the DFTB3/MM level. Let us finally mention that, as reported in the case of ammonia, modes associated to either intermolecular vibrations, combinations of water bending or stretching modes are observed, in the low frequency domain and in the 2100 and 3850 cm^{-1} regions, respectively.

5.3 Pentane

The harmonic IR spectra obtained for the pentane molecule are reported on Figure 3. For the isolated molecule, the two levels of theory agree on the main spectral features, namely the presence of two distinct active regions : the [2900-3000] cm^{-1} region corresponding to various combinations of C-H streching modes and the [1300-1500] cm^{-1} region, corresponding to various combinations of C-H bending modes (essentially nodding and shearing modes). At lower frequencies, small contributions arise from modes associated to global deformations of the carbonaceous skeleton as well as some bending modes (torsion and swinging modes). From Figure 3, both top and bottom pannels, it can be seen that, in the opposite to what was observed in the case of ammonia and glycine, the main features of the har-

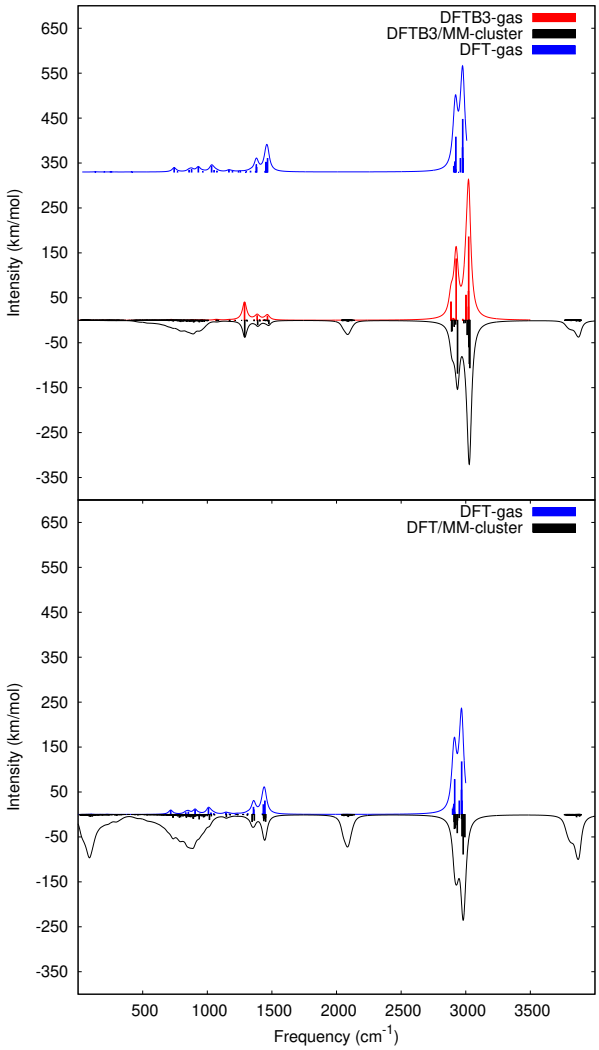

Fig. 3 Harmonic IR spectra for pentane. Top : gas phase molecule at the DFT and DFTB3 levels and solvated pentane at DFTB3/MM levels; Bottom : gas phase and solvated pentane at the DFT and DFT/MM levels.. Harmonic stick spectra were convoluted with Lorentzian function (FWMH=20 cm^{-1}) to produce the line spectra.

monic spectrum do not present significant changes due to the cluster environment, whatever the level of theory. This is well in line with the known hydrophobic nature of pentane[60], which does not form strong H-bonds with the surrounding water molecules. The major observed changes are the appearance of frequencies in the 2100 and 3850 cm^{-1} regions, assigned to water H-O-H bending modes and O-H stretching vibrations, as well as in the low frequencies region, associated to intermolecular modes.

5.4 Water

Our last example is the analysis of water spectra which are reported on Figure 4. Focusing first on the isolated molecule, the DFT spectrum indicates a bending mode at 1635 cm^{-1} and two stretching modes at 3721 and 3854 cm^{-1} for the symetric and antisymetric modes, respectively. These bands are also observed at the DFTB3 level at 1461, 3604 and 3877 cm^{-1}, respectively. The intensities are in agreement with the exception of the symmetric stretching mode which is active in the DFTB3 spectrum but presents very low intensity in the DFT spectrum. Inside the water cluster, the investigated water molecule presents similar modifications when computed either at the DFT/MM or DFTB3/MM levels, namely an increase of the absorption intensities, in particular for the two stretching modes, a red-shift of the stretching modes and a blue-shift of the bending mode. One can also observe, in both spectra, broad features below 1100 cm^{-1} which originate from global intermolecular modes. The main difference is that, in the DFT/MM spectrum, one of these modes, located at 1085 cm^{-1} presents a strong intensity, whereas in the DFTB3/MM spectrum no intermolecular mode is observed with such a strong intensity.

We finally discuss the influence of the size of the region treated at the QM level. To do so, we have optimized and computed the harmonic spectrum for the water cluster treating all the water molecules (i.e. those previously described either at the QM or at the MM level) at the DFTB3 level. To be consistent with the previous calculations, only the dipole moment of the molecule previously described at the QM level is included in the intensity calculation. This spectrum is compared with the one of the isolated water molecule on Figure 5. As in the calculation performed at the DFTB3/MM level, the O-H stretching modes are red-shifted and their intensities are increased. Let us note that the intensity is now spread over many almost degenerated modes. We also recover the appearance of a broad pattern at low frequencies. The main difference relies in the bending modes, which was blue shifted in the DFTB3/MM spectrum, but is almost not impacted by the environment, apart from the fact that its intensity is now spread over many modes with very similar frequencies. This can clearly be related to the fact that the frequency of the water bending mode is 1471 cm^{-1} when computed at the DFTB3 level and 2028 cm^{-1} when computed at the MM level. As a consequence, at the DFTB3/MM level the coupling of the bending mode of the QM molecule with the bending modes of the MM molecules at higher frequency are responsible

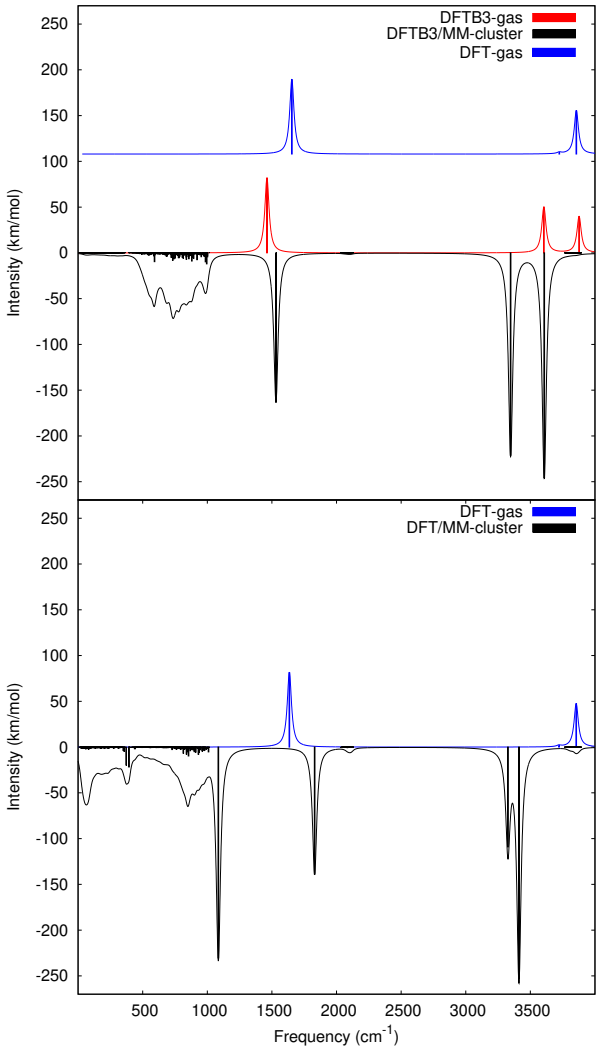

Fig. 4 Harmonic IR spectra for water. Top : gas phase molecule at the DFT and DFTB3 levels and solvated water molecule at DFTB3/MM levels; Bottom : gas phase and solvated water molecule at the DFT and DFT/MM levels. Harmonic stick spectra were convoluted with Lorentzian function (FWMH=20 cm^{-1}) to produce the line spectra.

for the observed blue shift. To conclude on this point, we would like to stress that the size of the subsystem treated at the QM level should be large enough to limit such artifacts originating from the difference between the water molecules frequencies in the two levels of calculations and that force fields of higher accuracy (polarisable force fields or TIP4P, or SPC, or TIP3F) could provide better descriptions of water frequencies.

6 Binding energies

We finally compare the binding energies between the two subsystems, i.e. between the studied molecule

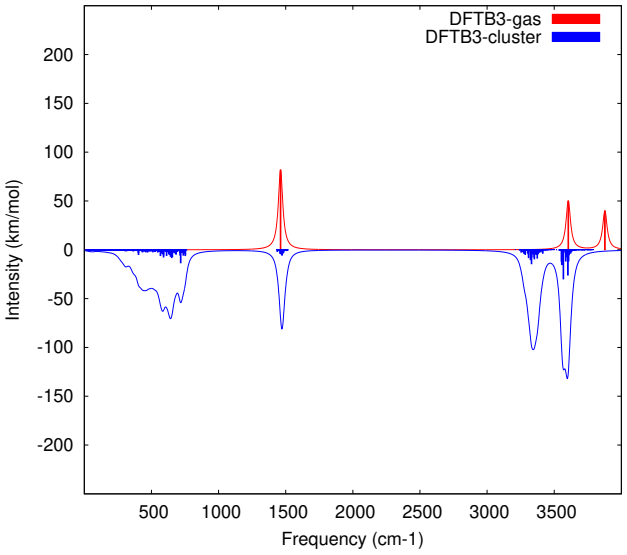

Fig. 5 Harmonic spectra for water molecule in the gas phase and solvated in a water cluster computed at the DFTB3 level.

(NH_3, $C_2H_5NO_2$, C_5H_{12} or H_2O), treated at the QM level, and the surrounding water cluster, treated at the MM level). This binding energy is computed from the energies of the full system minus the sum of energies of the two subsystems :

$$E_{bind} = E(QM + MM) - E(QM) - E(MM) \qquad (13)$$

The geometries were not relaxed to compute the energies E(QM) and E(MM), i.e. the coordinates of the QM molecule for E(QM) (resp. of the water cluster for E(MM)) having the same values as in the full system. Table 1 reports the binding energies computed for the geometries optimized either at the DFTB3/MM or DFT/MM levels. Small differences (less than 0.13 eV) can be observed between the binding energies obtained at the DFT/MM level for the geometries resulting from local optimisation at the DFTB3/MM (3^{rd} column) or DFT/MM ($2nd$ column) levels. This indicates small structural relaxation when changing the computational scheme. The binding energies computed at the DFT/MM and DFTB3/MM levels for the same geometries (optimised at the DFTB3/MM level, columns 3 and 4) are in very good agreement for the glycine and pentane molecule, whereas for ammonia and water, the values obtained at the DFT/MM level (-0.649 and -1.380 eV, respectively) are larger than those obtained at the DFTB3/MM level (-0.394 and -1.001 eV, respectively). In order to track the origin of this difference, we report the contribution of the Lennard-Jones potential to the coupling energy, this value being the same in

	$E_{Binding}$ (eV)			
Geometry	DFT/MM Geo opt.	DFTB3/MM Geo opt.		
Theory	DFT/MM	DFT/MM	DFTB3/MM	LJ contrib.
NH_3	-0.778	-0.649	-0.394	0.065
$C_2H_5NO_2$	-2.877	-2.759	-2.798	0.064
C_5H_{12}	-0.639	-0.631	-0.640	-0.643
H_2O	-1.498	-1.380	-1.001	0.317

Table 1 Binding energies computed at the DFT/MM or DFTB3/MM levels from Eq. 13 for the four investigated systems at the geometries optimized either at the DFT/MM or DFTB3/MM levels of theory. Lennard-Jones contributions to the binding energies are also reported.

both the DFT/MM and DFTB/MM schemes. It appears that, appart from the pentane case where it almost equals the binding energy, this contribution is repulsive. It means that most of the stabilising interaction arises from the interaction between the charges of the MM subsystem and electronic density of the QM subsystem. The difference could originate in a lack of polarisation at the DFTB/MM level. In the DFTB and DFTB/MM operator expressions (Equations 4 and 12), the atomic charges of the QM and MM regions induce couplings between the orbitals of different atoms but not between orbitals of the same atom, due to the fact that the overlap matrices elements $S_{\mu\nu}$ are zero if μ and ν correspond to two different orbitals of the same atom. As a consequence, an MM region would not induce polarisation of an isolated QM atom. This neglect of atomic polarisation energy could be, at least partially, at the origin of an underestimation of the binding energy at the DFTB3/MM level. Let us note in particular that the atomic polarisabilities of some elements [61], like nitrogen (1.100 $Å^3$) or oxygen (0.802 $Å^3$), are significant with respect to molecular polarisabilities [62] of ammonia (2.103 $Å^3$) or water (1.501 $Å^3$) .

7 Conclusion

We have reported a new implementation for an hybrid QM/MM scheme where the QM part of the system is treated at the DFTB level and the MM part of the system using class 1 force fields approaches. The DFTB level can vary from zeroth-order DFTB, SCC-DFTB to third order DFTB. Only the class 1 force fields AMBER and OPLS-AA have been used up to now but the implementation allows to use any class 1 Force Filed, as long as the user provides the corresponding parameters. The additive scheme has been implemented and two levels of coupling can be chosen by the user, namely the mechanical and the electrostatic coupling. In both cases, the interactions between the MM and the QM subparts of the system are computed from a Lenard-Jones potential. In the electrostatic coupling, an additional term describes the coulomb interactions between the atomic charges of the two regions. The atomic charges from the MM domain remain constant but those of the QM region are optimised self-consistently within the DFTB variational scheme, including the polarisation of the QM region by the charges of the MM region.
The energy gradients, (mandatory to perform molecular dynamics) of compute vibrational spectra , have also been implemented for this hybrid scheme. Benchmarks have been performed by comparing the latter computed either from the analytical formula or from finite differences and by checking that the total energy was conserved during micro-canonical simulations.
For a set of selected small molecules in water clusters, harmonic spectra have been computed at the DFTB3/MM and DFT/MM levels, making use of the same force field. The geometrical configurations were arbitrarily generated, , they do not exemplify any real system. Any analysis of the properties and characteristic of this geometries can not be extracted. Nevertheless, these test cases showed that the main environment effects computed at the DFT/MM level are reproduced at the DFTB3/MM level.
In the future, the combination of the present implementation with schemes similar to the one developed by Iftner *et al.* [31] to couple polarisable Argon atoms description at the MM level with the DFTB scheme would allow us to mix DFTB with polarisable force fields similarily to what was done at the DFT level[52], [63]

Acknowledgements The authors acknowledge the ANR agency (project ANR-19-CE29-0011-02 RUBI) for financial support and the supercomputing facilities of CALMIP for generous allocation of computer resources (project P0059 and P18009).

Supplemental Material

We provide, in Supplemental Material, benchmarks for the implementation of DFTB3/MM energy and gradi-

ents, namely a Table S1 comparing energy gradients computed for an arbitrarily chosen geometry of a cluster consisting of an ammonia molecule (QM level) and a water molecule (MM level) computed either from the analytical implementation and through finite differences and a Figure S1 showing the conservation of the total energy during an MD simulation in the microcanonical ensemble for an ammonia molecule embedded in a water cluster (215 molecules).

Furthermore, a comparison of IR Harmonic spectra of soluted ammonia in water cluster at DFTB3/MM and DFT/MM levels of theory are included in the supplemental material (Figure S2). A brief analysis from both plots is extracted, showing that the effect of water environment on the system treated by DFT/MM method is reproduced by the DFTB3/MM level of theory.

References

1. A. Warshel, M. Karplus, Journal of the American Chemical Society **94**(16), 5612 (1972)
2. A. Warshel, M. Levitt, Journal of Molecular Biology **103**(2), 227 (1976)
3. G. Groenhof, Introduction to QM/MM Simulations (Humana Press, Totowa, NJ, 2013), pp. 43–66
4. F. Maseras, K. Morokuma, Journal of Computational Chemistry **16**(9), 1170 (1995)
5. H.M. Senn, W. Thiel, Current Opinion in Chemical Biology **11**(2), 182 (2007)
6. J.P. Marcolongo, A. Zeida, J.A. Semelak, N.O. Foglia, U.N. Morzan, D.A. Estrin, M.C. González Lebrero, D.A. Scherlis, Frontiers in Chemistry **6**, 70 (2018)
7. P.K. Biswas, S. Chakraborty, Nucleic Acids Research **47**(6), 2757 (2019)
8. M. Klähn, J. Schlitter, K. Gerwert, Biophysical Journal **88**(6), 3829 (2005)
9. H.R. Zhekova, V. Ngo, M.C. da Silva, D. Salahub, S. Noskov, Coordination Chemistry Reviews **345**, 108 (2017)
10. D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, R. Kaschner, Phys. Rev.B **51**, 12947 (1995)
11. G. Seifert, D. Porezag, T. Frauenheim, Int. J. Quantum Chem. **58**, 185 (1996)
12. M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, G. Seifert, Phys. Rev. B **58**, 7260 (1998)
13. M. Elstner, G. Seifert, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **372**(2011), 20120483 (2014)
14. F. Spiegelman, N. Tarrat, J. Cuny, L. Dontot, E. Posenitskiy, C. Martí, A. Simon, M. Rapacioli, Advances in Physics: X **5**(1), 1710252 (2020)
15. J.C. Slater, G.F. Koster, Phys. Rev.**94**, 1498 (1954)
16. M. Gruden, L. Andjeklovič, A.K. Jissy, S. Stepanovič, M. Zlatar, Q. Cui, M. Elstner, J. Comput. Chem. **38**(25), 2171 (2017)
17. L. Zheng, S. Zamith, M. Rapacioli, Theoretical Chemistry Accounts **140**(2), 19 (2021)
18. T.A. Niehaus, S. Suhai, F. Della Sala, P. Lugli, M. Elstner, G. Seifert, T. Frauenheim, Phys. Rev. B **63**, 085108 (2001)
19. H. Witek, K. Morokuma, A. Stradomska, J. Theor. Comput. Chem. **4**, 639 (2005)
20. L. Dontot, F. Spiegelman, S. Zamith, M. Rapacioli, The European Physical Journal D **74**(11), 216 (2020)
21. S. Chakraborty, G. Mulas, M. Rapacioli, C. Joblin, Journal of Molecular Spectroscopy **378**, 111466 (2021)
22. H. Witek, K. Morokuma, A. Stradomska, J. Chem. Phys. **121**(11), 5171 (2004)
23. J. Cuny, N. Tarrat, F. Spiegelman, A. Huguenot, M. Rapacioli, Journal of Physics: Condensed Matter **30**(30), 303001 (2018)
24. M. Wahiduzzaman, A.F. Oliveira, P. Philipsen, L. Zhechkov, E. van Lenthe, H.A. Witek, T. Heine, J. Chem. Theor. Comp. **9**(9), 4006 (2013), pMID: 26592396
25. M. Elstner, D. Porezag, G. Seifert, T. Frauenheim, S. Suhai, MRS Online Proceedings Library **538**, 541 (1998)
26. R. Salomon-Ferrer, D.A. Case, R.C. Walker, WIREs Comput Mol Sci **3**(2), 198 (2013)
27. G.d.M. Seabra, R.C. Walker, M. Elstner, D.A. Case, A.E. Roitberg, The Journal of Physical Chemistry A **111**(26), 5655 (2007), pMID: 17521173
28. T. Kubar, K. Welke, G. Groenhof, J. Comput. Chem. **36**(26), 1978 (2015)
29. Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim, M. Karplus, J. Phys. Chem.B **105**(2), 569 (2001)
30. G.F.d. Lima, T. Heine, H.A. Duarte, J.R. Sabin, E. Brändas, Chapter 5 - Molecular Dynamics of Polypeptides and Their Inclusion Compounds with Cyclodextrin in Aqueous Solution Using DC–SCC–DFTB/UFF Approach (Academic Press, 2010), vol. 59, pp. 145–180
31. C. Iftner, A. Simon, K. Korchagina, M. Rapacioli, F. Spiegelman, J. Chem. Phys. **140**(3), 034301 (2014)
32. G. Hou, X. Zhu, Q. Cui, Journal of Chemical Theory and Computation **6**(8), 2303 (2010), pMID:

20711513
33. Y. Nishimoto, J. Phys. Chem.A **120**(5), 771 (2016)
34. T. Kubař, P.B. Woiczikowski, G. Cuniberti, M. Elstner, J. Phys. Chem.B **112**, 7937 (2008)
35. M. Lundberg, Y. Sasakura, G. Zheng, K. Morokuma, J. Chem. Theo. Comp. **6**(4), 1413 (2010)
36. T. Heine, M. Rapacioli, S. Patchkovskii, J. Frenzel, A. Koster, P. Calaminici, H.A. Duarte, S. Escalante, R. Flores-Moreno, A. Goursot, J. Reveles, D. Salahub, A. Vela, deMonNano, http:demonnano.ups-tlse.fr (2009)
37. N. Gillet, M. Elstner, T. Kubař, The Journal of Chemical Physics **149**(7), 072328 (2018)
38. B. Hourahine, B. Aradi, V. Blum, F. Bonafé, A. Buccheri, C. Camacho, C. Cevallos, M.Y. Deshaye, T. Dumitrică, A. Dominguez, S. Ehlert, M. Elstner, T. van der Heide, J. Hermann, S. Irle, J.J. Kranz, C. Köhler, T. Kowalczyk, T. Kubař, I.S. Lee, V. Lutsker, R.J. Maurer, S.K. Min, I. Mitchell, C. Negre, T.A. Niehaus, A.M.N. Niklasson, A.J. Page, A. Pecchia, G. Penazzi, M.P. Persson, J. Řezáč, C.G. Sánchez, M. Sternberg, M. Stöhr, F. Stuckenberg, A. Tkatchenko, V.W.z. Yu, T. Frauenheim, The Journal of Chemical Physics **152**(12), 124101 (2020)
39. G. de M Seabra, R.C. Walker, M. Elstner, D.A. Case, A.E. Roitberg, The journal of physical chemistry. A **111**(26), 5655 (2007)
40. A.W. Götz, M.A. Clark, R.C. Walker, Journal of computational chemistry **35**(2), 95 (2014)
41. A. Oliveira, G. Seifert, T. Heine, H. duarte, J. Braz. Chem. Soc. **20**, 1193 (2009)
42. P. Koskinen, V. Makinen, Comput. Mat. Sc. **47**(1), 237 (2009)
43. W.M.C. Foulkes, R. Haydock, Phys. Rev.B **39**, 12520 (1989)
44. M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, G. Seifert, Phys. Rev. B **58**(11), 7260 (1998)
45. Y. Yang, H. Yu, D. Uork, Q. Cui, M. Elstner, J. Phys. Chem. A **111**, 10861 (2007)
46. M. Gaus, Q. Cui, M. Elstner, Journal of Chemical Theory and Computation **7**(4), 931 (2011)
47. M. Gaus, A. Goez, M. Elstner, J. Chem. Theo. Comp. **9**(1), 338 (2013), pMID: 26589037
48. L. Monticelli, D.P. Tieleman, Force Fields for Classical Molecular Dynamics (Humana Press, Totowa, NJ, 2013), pp. 197–213
49. M.A. González, JDN **12**, 169 (2011)
50. K. Vanommeslaeghe, O. Guvench, J. MacKerell, Alexander D, Current pharmaceutical design **20**(20), 3281 (2014)
51. M.J. Field, P.A. Bash, M. Karplus, Journal of Computational Chemistry **11**(6), 700 (1990)
52. A. de la Lande, A. Alvarez-Ibarra, K. Hasnaoui, F. Cailliez, X. Wu, T. Mineva, J. Cuny, P. Calaminici, L. López-Sosa, G. Geudtner, I. Navizet, C. Garcia Iriepa, D.R. Salahub, A.M. Köster, Molecules **24**(9) (2019)
53. A. Koster, G. Geudtner, A. Alvarez-Ibarra, P. Calaminici, M. Casida, J. Carmona-Espindola, V. Dominguez, R. Flores-Moreno, G. Gamboa, A. Goursot, T. Heine, A. Ipatov, A. de la Lande, F. Janetzko, J. del Campo, D. Mejia-Rodriguez, J.U. Reveles, J. Vasquez-Perez, A. Vela, B. Zuniga-Gutierrez, D. Salahub, deMon2k Version 6 (2018)
54. J.P. Perdew, K. Burke, M. Ernzerhof, Physical Review Letters **77**(18), 3865 (1996)
55. N. Godbout, D.R. Salahub, J. Andzelm, E. Wimmer, Canadian Journal of Chemistry **70**(2), 560 (1992)
56. A.M. Köster, J.U. Reveles, J.M. del Campo, The Journal of Chemical Physics **121**(8), 3417 (2004)
57. S. Nosé, J. Chem. Phys. **81**(1), 511 (1984)
58. W.G. Hoover, Phys. Rev. A **31**(3), 1695 (1985)
59. A.F. Oliveira, P. Philipsen, T. Heine, J. Chem. Theo. Comp. **11**(11), 5209 (2015), pMID: 26574316
60. V. Gutmann, G. Resch, Lecture Notes on Solution Chemistry, vol. p.97 (World Scientific Co. Pte.Ltd, 1995)
61. T.M. Miller, B. Bederson, D.R. Bates, Atomic and Molecular Polarizabilities-A Review of Recent Advances (Academic Press, 1978), vol. 13, pp. 1–55
62. T.N. Olney, N.M. Cann, G. Cooper, C.E. Brion, Chemical Physics **223**(1), 59 (1997)
63. X. Wu, J.M. Teuler, F. Cailliez, C. Clavaguéra, D.R. Salahub, A. de la Lande, Journal of Chemical Theory and Computation **13**(9), 3985 (2017)