Developing accurate intramolecular force fields for conjugated systems through explicit coupling terms

J. Cerezo, G. Prampolini and I. Cacelli

Abstract The accuracy of molecular mechanics force fields (FF) reveals critical for applications where precise molecular structures along a conformational sampling are required, as in the simulation of electronic spectroscopies. This implies abandoning generalized FFs in favor of specific FFs, with non-transferable parameters able to accurately describe the targeted species. A promising strategy in this direction consists in the so-called quantum mechanically derived FFs, in which the parameters are fitted onto reference data computed through quantum chemistry. However, in order to obtain a global set of parameters able to reliably describe the reference potential energy surface in different regions of the conformational space, the complexity of the analytical expressions of the FF becomes crucial. Regarding intramolecular interactions, the functional form of standard transferable FFs is restricted to terms that depend on only one internal coordinate. It will be shown that such models may reveal insufficient to describe systems as polyenic chains, where complex electronic effects, e.g. conjugation, intrinsically couple different ICs. We propose a functional form for intramolecular FFs, which includes explicit couplings between flexible dihedrals and stiff ICs (bonds and angles), being able to properly describe the geometrical changes arising not only from steric interactions, but also from conjugation effects, i.e. the change of bond orders induced by conformational changes. The parameterization of the coupled FFs is carried out by means of automated and efficient computational

Javier Cerezo*
Departamento de Química Física, Universidad de Murcia, 30100 Murcia, Spain
E-mail: jcb1@um.es

Giacomo Prampolini*
Istituto di Chimica dei Composti Organo Metallici, Consiglio Nazionale delle Ricerche (ICCOM-CNR), Area della Ricerca, via G. Moruzzi 1, I-56124 Pisa, Italy
E-mail: giacomo.prampolini@iccom.cnr.it

Ivo Cacelli
Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via G. Moruzzi 13, I-56124 Pisa, Italy

protocols, purposely developed in the present work. All procedures are tested and validated by generating FFs for the two smallest compounds in the polyenic series (butadiene and hexatriene).

1 Introduction

The most reliable theoretical framework to investigate energy landscapes of atoms and molecules is certainly quantum mechanics (QM). Yet, the computational cost of accurate QM calculations dramatically increases along with the dimensions of the investigated system, and even the cheapest QM techniques, as those based on density functional theory (DFT), rapidly become infeasible when wider portions of the surrounding environment and their dynamics are included in the model. In many cases, a popular solution to circumvent this problem is to resort to classical molecular mechanics (MM), and explore the phase space of the target system through Monte Carlo (MC) or molecular dynamics (MD) methods.[1,2] In such techniques the electron cloud is not explicitly included in the model and its effects are taken into account in a effective manner, considering the ground electronic state. This drastic approximation may undermine the quality of the energy landscape when strong electronic rearrangements occur as a consequence of nuclear geometry distortions, precluding for instance an accurate prediction of most spectroscopic observables.

In the past years, many protocols aimed to combine classical and QM techniques to investigate the spectroscopic behavior of complex systems have been proposed.[3] One possibility is to treat the different regions of the system (e.g.the chromophore, the first neighbor shell, interacting surfaces or substrates, solvent, etc.) at different level of theory, considering their different impact on the spectrum. Such QM/MM integrated strategies can play a significant role in unraveling different contributions to the inherent complexity of large system's spectra. One other possibility is to adopt MD simulations to sample the configurational space of the whole investigated system along its dynamics, and successively employ the extracted snapshots in QM calculations to retrieve the desired spectra. It is though evident that a key ingredient for the success of such strategies is the accuracy of the MM description, and in particular its capability of delivering reliable molecular geometries of both the chromophore and its embedding.

In classical numerical simulations, the accuracy of the results essentially relies in the parameters and complexity of the adopted force field (FF),[1,2,4] which describes the total energy of the system through a collection of analytical functions of the nuclear coordinates. Each FF can be further partitioned in an intramolecular term, which rules the flexibility of each molecule, and in an intermolecular one, which deals with the interaction energy between two or more molecules. Within the frame of computational spectroscopy, since it governs the chromophore internal dynamics, it is the intramolecular FF that certainly plays the most relevant role, and will be the subject of the present work. The vast majority of MD simulations are carried out with empirical FFs,[4–10] parameterized towards experimental and/or quantum mechanical (QM) data for a defined set of similar molecules. On the one hand, this strategy allows for transferable parameters, which can be promptly adopted to describe molecular targets outside the training set (but possibly with similar chemical features). On the other hand, only a coarse description of the internal molecular structure can be obtained,[11,12] because the same parameters are designed to describe in average a class of compounds rather than a specific molecule. As also recently pointed out by Jorgensen's group,[13] a complementary strategy can be followed,[14–17] based on aban-

doning the advantages of transferability in favor of a parameterization specifically tailored on the target molecule, which allows for higher accuracy. In this framework, the bridging role of classical simulations between the QM microscopic description and the macroscopic observables of complex systems can be recovered, provided the FF is able to mimic the quantum description with high accuracy. Such specific and accurate FFs, sometime termed as quantum mechanically derived FFs (QMD-FFs),[15,17] are attracting a growing attention, as testified by the several different parameterization strategies reported in the past decade.[15,18–25] Among other QMD-FFs, the JOYCE parameterization procedure was proposed by some of us,[15] and successfully applied in several fields,[26–33] including computational spectroscopy.[11,12,34–38] As recently pointed out in Ref 12, despite a good adherence to the overall QM description, the main lack of JOYCE's standard approach is the description of the highly delocalized normal modes, which arise from the combination of two or more internal coordinates (ICs).

As a matter of fact, one of the main limitations of standard intramolecular transferable FFs arises from the adopted functional form. Indeed, it is well known[15,39–42] that FFs aiming to spectroscopic accuracy require rather complex analytical expressions and/or the inclusion of coupling terms as in the so called Class II FFs.[42] This is in contrast with the majority of the standard FFs, which include terms involving a single "bonded" coordinate (i.e. a bond, angle or dihedral, described as an harmonic function or a Fourier-like series) and, additionally, non-bonded contributions (described through Lennard-Jones (LJ) and Coulomb expressions) between a number of atom pairs, not directly connected by neither a covalent bond nor a valence angle. In such models, these non-bonded terms are capable of create indirect couplings among all the "bonded" ICs that are close to the interacting atom pair. If energy landscapes are properly parameterized, this strategy can be quite successful to account for steric effects, e.g. the repulsion between different parts of the molecule that might come too close in distorted geometries. However, the parameterization of non-bonded terms is not straightforward and, moreover, in some situations the coupling is due to a change of the electronic structure resulting from the displacement of one internal coordinate, and cannot be properly described with simple non-bonded terms.[12] A paradigmatic situation where the use of non-bonded terms is not able to properly reproduce the quantum mechanical behavior is the geometric distortion in a polyenic chain, when the dihedral angles defined along the chain are displaced. Concretely, as a dihedral angle moves away from planar conformations, the effective conjugation is modified, and the central bond changes its bond order with a consequent change of its length. In order to properly account for these effects, additional intramolecular terms should be included, to explicitly take into account the coupling between the dihedral and the central bond length.

In this paper a popular FF (GAFF[7]), and a set of intramolecular FFs of increasing complexity, all derived by the JOYCE code,[15,21] will be employed to investigate the flexibility and conformational behavior of the smallest prototypes of polyenic chains, namely the 1,3-butadiene (**I**) and 1,3,5-hexatriene (**II**) molecules displayed in Figure 1. It will be shown that a high accuracy can only be gained with the most complex intramolecular FF, obtained including explicit couplings among pairs of ICs. The algorithm has been purposely implemented in JOYCE for the present work.

2 Theoretical models

2.1 Force Fields

As previously discussed, a MM FF represents the Potential Energy Surface (PES) as an analytic function in terms of the nuclear degrees of freedom, which is parameterized so as to reproduce the effect of the electronic Hamiltonian. Usually, covalent bonds are not allowed to be broken (non-reactive FFs), and the contributions to the potential energy (E_{tot}^{FF}) can be partitioned in an intra- (E_{intra}^{FF}) and in an inter-molecular (E_{inter}^{FF}) term:

$$E_{tot}^{FF} = E_{inter}^{FF} + E_{intra}^{FF} \tag{1}$$

The first term of equation (1) is generally computed as a sum of LJ and Coulomb model potentials

$$E_{inter}^{FF} = \sum_{i=1}^{N_{sites}} \sum_{j=1}^{N_{sites}} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{i=1}^{N_{sites}} \sum_{j=1}^{N_{sites}} \frac{q_{ij}}{r_{ij}} \tag{2}$$

where i and j belong to different molecules, N_{sites} is the total number of interacting sites and q_{ij}, σ_{ij} and ε_{ij} are intermolecular parameters which describe the interaction among molecules. Conversely, E_{intra}^{FF} accounts for the intramolecular energy, and can be computed as the sum of the single-molecule contributions (E_k^{intra}).

$$E_{intra}^{FF} = \sum_{k=1}^{N_{molecules}} E_k^{intra}(\mathbf{q}) \tag{3}$$

The E_k^{intra} term drives the flexibility of molecule k and consists in a sum of analytical model potentials, function of a selected set of internal coordinates (ICs) $\mathbf{q}$ and is specified by a set of intramolecular parameters.

Adopting such a partition, intra- and intermolecular parameters can be derived independently. In this work, we focus on the intramolecular potential of polyenic chains, which can in principle be complemented with any type of intermolecular potential.

2.1.1 Uncoupled intramolecular force fields

In the simplest intramolecular FF, each term depends on one IC q_i only, and is therefore often classified as purely diagonal FF. For bonds and angles, such potential terms are commonly described with harmonic potentials. Harmonic functions are also appropriate for stiff dihedrals (e.g., describing the torsion around a double bond or improper dihedrals keeping the co-planarity of the residues in an aromatic ring), whereas flexible dihedrals can be conveniently described by Fourier-like expansions.

$$E(\mathbf{q}) = E_{stretch} + E_{bend} + E_{Rtors} + E_{Ftors} \tag{4}$$

where

$$E_{stretch} = \frac{1}{2} \sum_{\mu}^{N_{bonds}} k_{\mu}^s (r_{\mu} - r_{\mu}^0)^2 \tag{5}$$

$$E_{bend} = \frac{1}{2} \sum_{\mu}^{N_{angles}} k_{\mu}^b (\theta_{\mu} - \theta_{\mu}^0)^2 \tag{6}$$

$$E_{Rtors} = \frac{1}{2} \sum_{\mu}^{N_{Rdihedrals}} k_{\mu}^t (\phi_{\mu} - \phi_{\mu}^0)^2 \tag{7}$$

and

$$E_{Ftors} = \sum_{\mu}^{N_{Fdihedrals}} \sum_{j=1}^{N_{cos\mu}} k_{j\mu}^d \left[1 + cos(n_j^{\mu} \delta_{\mu} - \gamma_j^{\mu}) \right] \tag{8}$$

Note that in the above equations the label *intra* and the subscript k have been dropped, (i.e., $E_k^{intra} \equiv E$) as it is now obvious that the discussion is focused on the internal energy of a single molecule. The same notation will be adopted in the following.

2.1.2 Implicit coupling terms: Lennard-Jones and Coulomb

A simple strategy to account for the coupling between different ICs in the potential is through the inclusion of non-bonded intramolecular potential terms (E_{Nb}).

$$E(\mathbf{q}) = E_{stretch} + E_{bend} + E_{Rtors} + E_{Ftors} + E_{Nb} \tag{9}$$

where

$$E_{Nb} = \sum_{i=1}^{N_{atoms}-1} \sum_{j=i+1}^{N_{atoms}} 4\varepsilon_{ij}^{intra} \left[\left(\frac{\sigma_{ij}^{intra}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}^{intra}}{r_{ij}} \right)^{6} \right] + \sum_{i=1}^{N_{atoms}-1} \sum_{j=i+1}^{N_{atoms}} \frac{q_{ij}^{intra}}{r_{ij}} \tag{10}$$

It is worth noticing that q_{ij}^{intra}, σ_{ij}^{intra} and ε_{ij}^{intra} are intramolecular parameters which are in general different from the analogs of the intermolecular potential q_{ij}, σ_{ij} and ε_{ij} entering equation (2).

The use of implicit coupling terms is a simple yet efficient way to account for couplings among ICs. Nonetheless, it is nearly impossible to exploit this strategem to introduce only few selected couplings between pairs of specific ICs. Indeed, even a single non-bonded term, i.e. defined between only one pair of atoms, often couples more than two ICs. For this reason, non-bonded intramolecular terms are usually employed between almost all possible pairs, thus introducing on overall averaged coupling among all diagonal ICs.

2.1.3 Explicit couplings terms

A more elegant approach consists instead in introducing a separate coupling term for each pair of ICs that is thought to be important for the system under study.

$$E(\mathbf{q}) = E_{stretch} + E_{bend} + E_{Rtors} + E_{Ftors} + E_{Coupl} \tag{11}$$

To adopt a more convenient formalism, the expression for a purely diagonal FF, equation (4), can be re-written as

$$E(\mathbf{q}) = \sum_{a=1}^{N_{func}} p_a f_a(q_\mu) \tag{12}$$

where q_μ indicates one internal coordinate ($q_\mu \equiv r_\mu, \theta_\mu, \phi_\mu, \delta_\mu$), while the p_a and f_a are respectively the force constant (either k_μ^s, k_μ^b, etc.) and the diagonal potential functions ($[r_\mu - r_\mu^0]^2$, $[\theta_\mu - \theta_\mu^0]^2$, $[1 + cos(n_j \delta_\mu - \gamma_j)]$, etc.) assigned to q_μ and entering equations (5)-(8).

Within this formalism, the generalized coupling here proposed consists in a sum of N_{Coupl} pairwise linear coupling terms, defined as the product between two functions, f and g, each depending only on one of the coupled ICs (e.g. q_μ or q_v, respectively):

$$E_{Coupl} = \sum_{a}^{N_{Coupl}} p_a f_a(q_\mu) g_a(q_v) \tag{13}$$

For the sake of clarity, it might be convenient to describe in more detail the effect of such couplings on the involved ICs. If one for instance focuses only on the coupling between a specific dihedral angle δ_v and a neighboring IC q_μ (e.g. a bond r_μ or and angle θ_μ), equation (13) simplifies to

$$E_{Coupl} \equiv E^{Coupl}(q_\mu, \delta_v) = (q_\mu - q_\mu^c) G(\delta_v) \tag{14}$$

where $G(\delta_v)$ is a function depending only on the dihedral δ_v. Considering only the subspace $[q_\mu, \delta_v]$, taking the uncoupled terms depending on q_μ as harmonic, the FF of such a reduced system is given by,

$$E(q_\mu, \delta_v) = \frac{1}{2} k^{harm}(q_\mu - q_\mu^0)^2 + E_{tors}(\delta_v) + (q_\mu - q_\mu^c) G(\delta_v) \tag{15}$$

where $E_{tors}(\delta_v)$, i.e. the diagonal term acting on the dihedral, stands for either expression (7) or (8). It is apparent that such a coupling induces a modification of the equilibrium values of q_μ upon changes of the dihedral angle δ_v. Concretely, taking the minimum of the potential with respect to q_v at its equilibrium value (q_μ^m), i.e., $\frac{\partial E}{\partial q_\mu}\big|_{q_\mu = q_\mu^m} = 0$, implies that q_μ^m is shifted with respect to q_μ^0:

$$q_\mu^m(\delta_v) = q_\mu^0 + \frac{G(\delta_v)}{k^{harm}} \tag{16}$$

2.1.4 Functional form of $G(\delta)$

Equation (16) implies that the equilibrium value of q_μ in terms of the dihedral angle δ_v is dictated by the shape of $G(\delta_v)$. Therefore, it should be flexible enough so as to reproduce the actual reference profiles obtained by QM calculations. In this sense, a Fourier series can provide the required flexibility. Moreover, the functional form should fulfill some additional requisites, such as ensuring that all terms of the series have the same parity as the reference data. Additionally, each term of the series should present a critical point (minima) at the global minima of the function, so to preserve the equilibrium structure. Both these requisites can be verified using the following functional form,

$$G(\delta) = \sum_i k_i^c [1 + \sin(n_i \delta - \gamma_i^c)] \tag{17}$$

The parity and location of the critical points depend on the choice of γ_i^c. Concretely, its value should be $\gamma_i^c = \delta^0 \pm 90°$ in order to preserve the even parity of the $q_\mu^m(\delta_v)$ profile. More specifically, when the energy profiles with minima at 180° (*trans*) obtained around torsions of single bonds in polyenic chains are considered, the adequate values of γ^c (i.e. ensuring that the equilibrium structure is preserved for each term) are -90° and and 90° , for odd values and even values of n_i, respectively.

2.2 Parameterization of the Force Field

2.2.1 Uncoupled FF and implicit couplings

The totally uncoupled FF is derived using the JOYCE program, following the usual 2-step protocol.[12, 15, 21] In the first step all harmonic terms are parameterized with respect to the QM Hessian at the global equilibrium structure, assigning the reference harmonic coordinates to the equilibrium ones. In the second step, the flexible dihedrals are parameterized according to the QM energy at different conformations along the relaxed scan around the torsion associated to the dihedral, adopting the Frozen Internal Rotation Approximation[15] (FIRA) to avoid additional perturbations due to the neglected couplings.

An additional step can be performed, introducing the implicit coupling terms (LJ), that are fitted against the energies obtained at each conformation along the relaxed scan and the gradients along a rigid scan (i.e., varying the value of the dihedral, while keeping the rest of internal coordinates at the equilibrium values in the minimum energy structure).

Finally, it is interesting to use a more general set of parameters for comparison. Concretely, we used GAFF[7] parameters, using antechamber through the acpype parser[43] to set all the parameters for the non-bonded terms (intramolecular LJ from GAFF and Coulomb with bcc-am1 model) and flexible dihedral potentials. For a more fair comparison, within the GAFF harmonic potentials, only the force constants were transferred form the original GAFF values, while the equilibrium harmonic coordinates are specifically taken from the reference QM optimized geometry.

2.2.2 Explicit couplings: direct fitting against equilibrium values, q_μ^m

The most straightforward procedure to parameterize the coupling terms involves the direct fitting of the function $q_\mu^m(\delta_v) - q_\mu^0$ defined from equation (16), where the database, $\mathbf{q_\mu^m}$, is built with the equilibrium values obtained from a relaxed scan of the dihedral angle (i.e. when the torsion is kept frozen at selected values of the dihedral, while the rest of the structure is minimized). Adopting the form of $G(\delta)$ indicated in Eq. (17), the target function is thus

$$q_\mu^m(\delta_v) - q_\mu^0 = \sum_i \frac{k_i^c}{k^{harm}}[1 + \sin(n_i\delta_v - \gamma_i^c)] \tag{18}$$

The expansion is done from $n_i = 1$ to $n_i = 6$, taking the values of $\gamma_i^c = \pm 90°$ as discussed above. The coefficients to be fitted are $\frac{k_i^c}{k^{harm}}$, thus in order to obtain the value of the coupling force constant, k_i^c, the harmonic force constant, k^{harm}, needs to be provided. The latter is obtained from the uncoupled FF , previously obtained through JOYCE, on top of which the couplings are added. Moreover, the complete expression of the coupling potential, equation (13), also includes a zero order term with respect to the rigid coordinate (q_μ^c) that does not appear in the target function (18). One must impose a sensible value which, even though should not affect the equilibrium value of the harmonic coordinate, may impact on the energetics of the flexible dihedral. We use the same zero order term used in the harmonic potential, i.e., $q_\mu^c = q_\mu^0$, which is the equilibrium value for this coordinate. This choice guarantees that the coupling will be small for values of the harmonic coordinate close to one in the equilibrium geometry.

2.2.3 Explicit couplings: global fitting based on the whole PES (JOYCE)

The direct fitting described in the previous section is, by definition, able to accurately reproduce the reference profiles q_μ^m vs. δ (as these constitute the fitting database). However, it does not provide a route to build a complete FF (note that the harmonic values should be provided externally) and, in more general grounds, the independent fitting for each term of the potential would lack any cross-effect between the coupling terms. In this sense, a methodology where all FF parameters, including uncoupled and both implicit and explicit couplings, are fitted altogether provides a more consistent parameterization. In this work, such parameterization is performed using the features of the QM intramolecular PES, *i.e.* energy, gradient and Hessian, at different conformations, following the procedure introduced some years ago by some of us[15] and herein refined and implemented in the JOYCE program. Briefly, the fitting involves the minimization of the following general objective function,

$$I = \sum_g I_g \tag{19}$$

where the sum run over all geometries considered. For each geometry g, the functional take the form,

$$I_g = I_g^E + I_g^G + I_g^H \tag{20}$$

with

$$I_g^E = W_g \left[\Delta U_g - E_g^{FF} \right]^2 \tag{21}$$

$$I_g^G = \sum_K^{3N-6} \frac{W_{gK}'}{3N-6} \left[G_{gK} - \left(\frac{\partial^2 E^{FF}}{\partial Q_K} \right)_g \right]^2 \tag{22}$$

and

$$I_g^H = \sum_{K \leq L}^{3N-6} \frac{2W_{KL}''}{(3N-6)(3N-5)} \left[H_{gKL} - \left(\frac{\partial^2 E^{FF}}{\partial Q_K \partial Q_L} \right)_g \right]^2 \tag{23}$$

In the above equations the indices in the sums (K, L) run over $3N - 6$ orthogonal vibrational coordinates, ΔU_g is the potential energy relative to the absolute minimum evaluated with the QM method, G_{gK} is its first derivative with respect to the K-th vibrational coordinate (Q_K), and H_{gKL} is the second derivative (i.e. a QM Hessian matrix element), with respect to Q_K and Q_L. The FF energy is indicated as E^{FF} as introduced above. Each term of the sum reports the deviation between different descriptors of the PES computed either with a suitable reference QM method and the target FF, weighting each term with W_g (energies), W_{gK}' (first derivatives) and W_{gKL}'' (second derivatives). In practice, not all terms are used at each geometry: for instance, the second derivatives terms are here only evaluated at the equilibrium structure.

The minimization of the objective function (19) allows for fitting the linear terms in the potential, i.e., the force constants associated to each potential term, while for the non-linear parameters sensible values should be given. In the case of uncouple terms, the equilibrium values associated to the harmonic terms, q_μ^0, can be assigned to the values in the equilibrium structure, while γ values in flexible torsions can be also derived from the equilibrium geometry imposing that both the function and its first derivative vanish. As for the coupling terms, the considerations discussed in Section 2.1.4 provide a rationale to select the value of γ^c ($\pm 90°$ in our case). As indicated in Eq. 16, the zero-order term should not influence the constrained equilibrium value of the harmonic coordinate, q_μ^m. However, such term indirectly affects the FF gradients, which are used in the fitting procedure. Consequently, its choice may have a remarkable impact on the performance of the fitting algorithm used by JOYCE (minimization of the merit function), and the actual fitted parameters may vary depending on the selected value. In practice, the choice of q_μ^m, which by default is set to the equilibrium one value, q_μ^0, can actually be used in order to tune the fitting.

The fitting with JOYCE is performed following a 3-step procedure, outlined as follows:

1. As done for the uncoupled FF, only the harmonic terms are fitted against the Hessian at the equilibrium geometry.
2. The couplings with bond lengths are added together with the uncoupled terms related to the flexible dihedrals, and they are fitted simultaneously against the energies along the relaxed scan and the gradients along the rigid scan related to the flexible torsions.

3. The couplings between flexible dihedrals and angles are fitted against the same reference data as in Step 2, but keeping the rest of the parameters fixed at their already derived values. In this last step, it was found that in order to obtain a reasonable fitting, the value of q_μ^c corresponding to the angle should be slightly increased with respect to the equilibrium value, i.e., $\theta_\mu^c = \theta_\mu^0 + \Delta^c$, where $\Delta^c \approx 1°$ turned out to be adequate.

3 Computational Details

Quantum mechanical (QM) data (energy, gradient and Hessian) were computed at DFT level as implemented in Gaussian09 program.[44] We adopted the M06 functional[45] along with the 6-31G(d) basis set. In the present context the quality of the method is not critical for our application, since the only role of QM calculations is to provide a set of reference data. In principle, the fitting procedures can be applied to QM data obtained at any theoretical level. The Hessian is computed at the global energy minimum (*trans* conformer). Further conformations are generated by rotating the flexible dihedral angles while the remaining ICs are either kept frozen (rigid scan) or optimized (relaxed scan). The gradient is evaluated at each point of the rigid scan.

Molecular mechanics (MM) calculations, used to validate the generated FFs, are performed with a modified version of Gromacs 4.5.5[46] that implements the coupling functions introduced in section 2.1.3. The developed parametrization method has been included in the JOYCE package.

4 Results

In this work we aim at providing accurate FFs for polyenic chains, which account not only for the conjugation effects (which result in different lengths of $C - C$ bonds, depending on their position within the chain) but also on the more intricate geometrical variations associated to the torsion around the $C - C$ bond itself. In order to deeply investigate such effects, we here focus on the two smaller members of the polyenic chain series namely 1,3-butadiene (**I**) and 1,3,5-hexatriene (**II**), which are respectively characterized by one and two torsional ICs (δ and $\delta_{1/2}$, see Figure 1), and whose flexibility induces significant changes in the conjugation length.
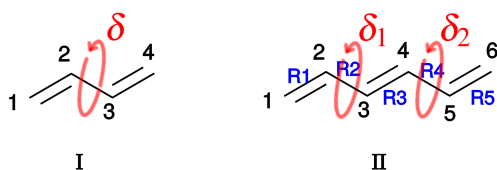

Fig. 1 Structure of 1,3-butadiene (**I**) and 1,3,5-hexatriene (**II**), including the labeling of the carbon atoms used in the text. The principal dihedrals δ, δ_1 and δ_2 connected with the torsional flexibility is evidenced with a red arrow.

4.1 1,3-butadiene

4.1.1 QM features

We first focus on compound **I**, which constitutes the smallest polyenic prototype where conjugation effects are present and can be modulated by the torsion of the central $C-C$ bond. The potential energy and some selected key structural parameters ($C-C$ bond lengths and $C\overset{\frown}{-C-}C$ valence angles) were obtained by a relaxed energy scan around δ (see Figure 1) and monitored as a function of such dihedral in Figure 2. Reference QM results, display an energetic profile with a global minimum at $\delta = 180°$ (*trans* conformer) and two additional local minima at $\delta \sim \pm 30°$, which are 12 kJ/mol above the global minimum. The highest energy barriers (~ 30 kJ/mol) are located at $\delta = \pm 90°$, while at $\delta = 0°$ (*cis* conformer) there is a small barrier connecting the two local minima at $\pm 30°$. Such profile can be rationalized in terms of two main effects: *i)* the changes in the conjugation patterns and *ii)* the steric effects that arise as parts of the molecule get closer upon rotation of the main dihedral δ. On the one hand, the conjugation length indicates the extent of the conjugation of adjacent double bonds, which significantly depends on their coplanarity: the more coplanar they are, the larger is the conjugation length. This in turn leads to a more localized electronic density, thus reducing the energy of the system. If one accounts for this effect only, the most favorable conformations would be *cis* and *trans*. On the other hand, steric effects will prevent that during the torsion two atoms might get too close, as could happen in the *cis* conformation. Indeed, when $\delta=0°$, the external CH_2 groups are constrained at a relatively short distance, and the consequent repulsive contribution increases the potential energy in this conformation.

The profile of the geometrical parameters can also be rationalized as an interplay between conjugation and steric encumbrance. In the case of the central bond (R_{2-3}), the minimum length is reached in the *trans* conformation, where the conjugation between the two double bonds increases the R_{2-3} effective bond order. As the conjugation is broken upon rotation of the δ dihedral, the bond order decreases and the R_{2-3} bond length increases, reaching its maximum at $\pm 90°$. In the vicinity of the *cis* conformation, the equilibrium bond length, intermediate between the values reached at $0°$ and $\pm 90°$, results from the delicate balance between the increased conjugation and the repulsion of the CH_2 groups. The behavior of the external bonds (R_{1-2}) is clearly the opposite: as the conjugation increase, its double bond character decreases and the bond length reaches its maximum value for coplanar chains, *i.e.*, in the *cis* and *trans* conformations. Finally, regarding the valence angle, its value seems not significantly affected by the conjugation of double bonds and its profile is mainly a result of steric effects: the angle drastically increases in the region around the *cis* conformation, pushed toward larger values by the repulsion between the CH_2 groups.

4.1.2 MM models

The first attempted FF parameterization was carried out employing the most simple FF expression, namely the uncoupled FF (**FF0**) provided by equation (4). Explicit or implicit coupling terms will be successively introduced on top of this model. **FF0**
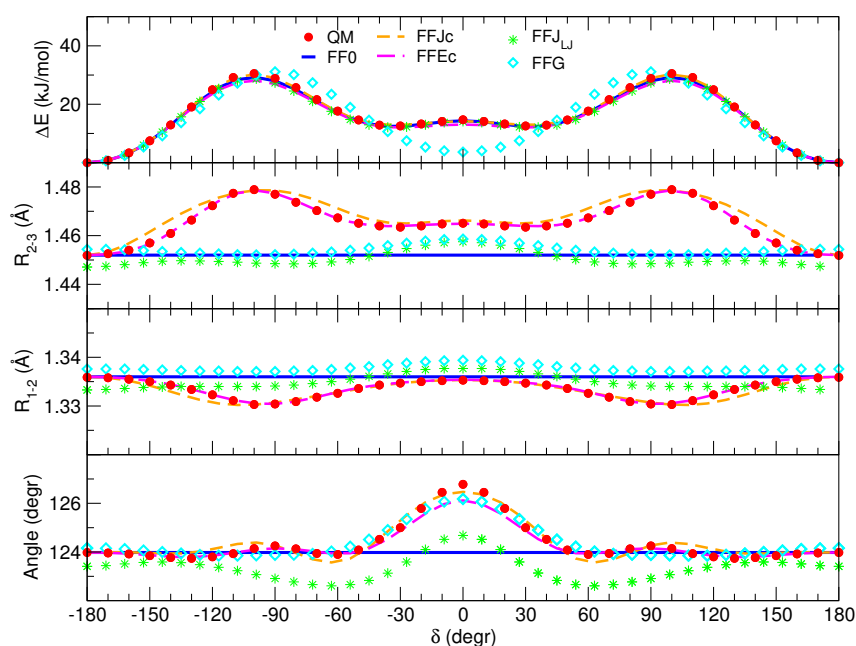
Fig. 2 Comparison of the QM and MM profiles computed during the relaxed torsional energy scan upon the rotation around the central $C-C$ bond. From top to bottom: relative potential energy (ΔE), equilibrium bond lengths of C_2-C_3 (R_{2-3}) and $C_1 = C_2$ (R_{1-2}) and valence angle $C_1 \overset{\frown}{-C_2} - C_3$.

was obtained through the JOYCE code, making use of the QM Hessian matrix in the equilibrium geometries and the relaxed QM energies, *i.e.* including in equation (19) only the terms given by (23) and (21). By looking at the results displayed with blue lines in Figure 2, it appears that **FF0** is able to accurately reproduce the QM relaxed energies upon δ rotation (top panel), but fails in reproducing the variations of the equilibrium values of the monitored structural parameters (lower panels). This was not unexpected, since this behavior arises from the coupling between the torsion with the other investigated ICs, which is missing by construction in the purely diagonal **FF0** model. In fact, it is evident that all the monitored ICs do not depend on the scanned δ coordinate.

The next step is to account for the coupling between different ICs by including it implicitly, through the standard LJ and Coulomb potentials appearing in equation (9). In this case, two sets of FFs were prepared, differing in the parameters entering in their expression: the widely used GAFF description (herein labeled **FFG**), which includes both LJ and Coulomb terms, and the **FFJ**$_{LJ}$ force field, built through the JOYCE program by fitting LJ potentials on top of the previously obtained **FF0**. At difference with **FF0**, **FFJ**$_{LJ}$ was obtained including in the JOYCE objective function (19) the term specified by equation (22): beside the equilibrium Hessian matrix and the previous relaxed torsional energy scan, the FF parameters were fitted also against

QM energy gradients, obtained by a purposely computed rigid energy scan of the δ dihedral. The results obtained with both these FFs are displayed in Figure 2 with cyan and green lines, respectively. The first noticeable feature is that the empirical **FFG** force field is significantly less accurate than the other tested FFs in reproducing the QM reference energetic profile, notwithstanding the harmonic equilibrium bond lengths and angles were taken from the optimized QM geometry. On the contrary, **FFJ**$_{LJ}$ quantitatively reproduces the QM data, as the energetic profile is not affected by the inclusion of the LJ terms. This is a direct consequence of the adopted proto-col, since the QM energy profile is indeed one of the targets of both **FF0** and **FFJ**$_{LJ}$ JOYCE fitting. Turning to the monitored structural parameters (*i.e.* equilibrium bond lengths and angle), it is evident that neither the GAFF model nor **FFJ**$_{LJ}$ are able to re-trieve a very accurate geometry along the relaxed scan. In the first case, where both LJ and Coulomb terms are included, the changes in bond length are not properly repro-duced, as their variation during the δ rotation is negligible. Conversely, it adequately describes the variations of the monitored angle, being in good agreement with the ref-erence one. The **FFJ**$_{LJ}$ shows rather similar results: for both bond lengths, LJ terms are not able to reproduce the reference profiles, while in the case of the angle, the general trend, *i.e* a rapid increase around the *cis* geometry, is qualitatively observed, although **FFJ**$_{LJ}$ is does not provide a quantitative agreement. In global, these results indicate that implicit coupling terms as LJ or charge-charge interactions, are able to account for steric repulsion (main responsible of the changes in the angle), but they completely fail at accounting for the effects related to the conjugation length.

A further refinement of the FFs tested so far involves the explicit inclusion of the coupling between two ICs through the additional terms appearing in equation (11). With this aim, two different sets of FF parameters were derived from the pre-viously computed QM data. The first set was obtained on top of **FF0**, by directly fitting equation (16), with respect to the QM equilibrium values of the harmonic co-ordinates along the relaxed scan. Since this fitting was performed with an "in house" code, external to the JOYCE package, the resulting FF was labeled **FFE**$_c$. The sec-ond set of FF parameters, named **FFJ**$_c$, was obtained through the JOYCE code, using equation (11) for the intramolecular potential terms, and fitted over the usual QM database (*i.e.* Hessian, relaxed energies and gradients from the rigid scan). By look-ing at Figure 2, it is evident that the QM torsional profile is very well reproduced for both sets of parameters. In the **FFJ**$_c$ case, this is not surprising, since the relaxed energies are included in the reference QM database. The good agreement found for **FFE**$_c$ may seem conversely somewhat more unexpected. Yet, it should be reminded that the **FFE**$_c$ force-field differs from **FF0** only for the E_{Coupl} term in equation (11), whereas all the parameters defining the other potential terms were not fitted, and are therefore unchanged with respect to the uncoupled FF. Since in the top panel of Fig-ure 2 the **FFE**$_c$ and **FFJ**$_c$ are practically superimposed, it can be concluded that the importance of the coupling terms in the total energy is only marginal. Turning to the geometrical values, it can be observed that both strategies are able to nicely repro-duce the QM variation of the bond length along the relaxed scan. It is worth noticing that while such a good agreement had to be expected for **FFE**$_c$, where the QM geo-metrical equilibrium values were the target of the external fitting, the results obtained with **FFJ**$_c$ indicate that the inclusion of the gradients in the QM database is able to

properly predict the equilibrium structure consistent with such gradients. From a different point of view, both $\mathbf{FFE}_c$ and $\mathbf{FFJ}_c$ results suggest that a functional form for E_{Coupl} like the one here proposed and shown in equations (13), (14) and (17) is able to reproduce the purely electronic effect related with the variations of the conjugation length. Finally, as far as the variation of the bending angle is concerned the agreement is slightly worse in the case of the external fitting, which may indicate that the coupling between other pairs of internal coordinates may be involved. A somewhat better agreement results for the $\mathbf{FFJ}_c$ force field, where the parameterization of the coupling term between the backbone bending angle and the δ dihedral required the tuning of the θ^c parameter (see equation 14)), which was set $1°$ above θ^0.

4.2 1,3,5-hexatriene

A slightly more complex model can be found in compound $\mathbf{II}$, because 1,3,5-hexatriene includes two rotable $C - C$ bonds, namely $C_2 - C_3$ and $C_4 - C_5$ (see δ_1 and δ_2 in Figure 1). The QM relaxed torsional energy profile, associated with one of the possible rotations is displayed in Figure 4. The energy curve is very similar to the one com-
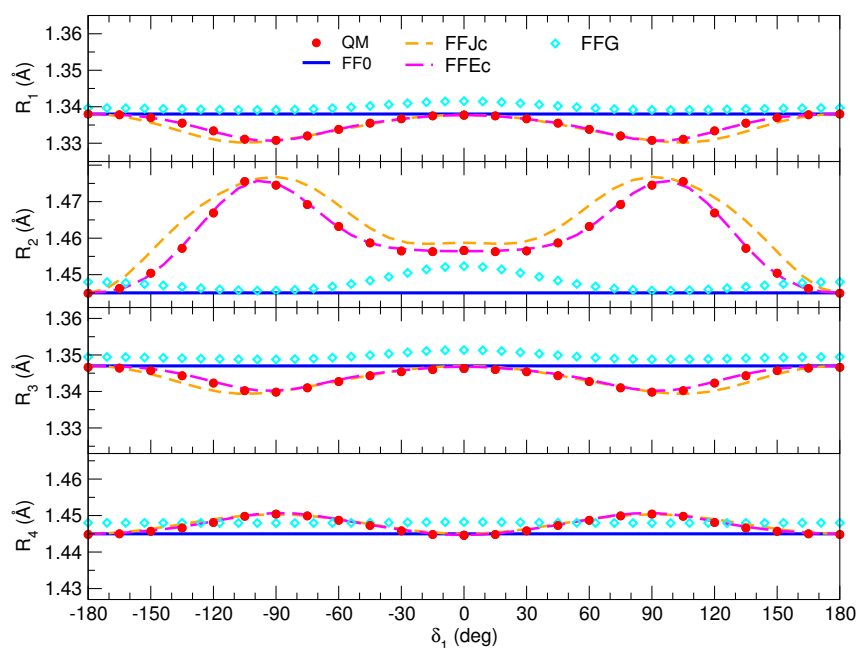

Fig. 3 Changes in the equilibrium bond lengths (R_1 - R_4) obtained in the constraint minima generated from a relaxed scan around the torsion of the R_2 bond.

puted for $\mathbf{I}$ and it can find a very similar rationale, in terms of conjugation and steric

effects. The most relevant difference between the two compounds is that the five different $C - C$ bond length and four $\widehat{C - C - C}$ angles can be affected by the rotation of either δ_1 or δ_2 dihedrals. Since the effect over an IC is expected to be modulated by its distance from the scanned dihedral, the investigated ICs are labeled increasingly starting from the edge closer to the rotated bond. Concretely, assuming that the rotation takes place around the $C_2 - C_3$ bond, we labeled the monitored bonds as R_1 ($C_1 = C_2$), R_2 ($C_2 - C_3$), R_3 ($C_3 = C_4$), R_4 ($C_4 - C_5$), R_5 ($C_5 - C_6$) and the angles as θ_1 ($\widehat{C_1 = C_2 - C_3}$), θ_2 ($\widehat{C_2 - C_3 = C_4}$), θ_3 ($\widehat{C_3 = C_4 - C_5}$) and θ_4 ($\widehat{C_4 - C_5 = C_6}$). Moreover, the analysis of the reference data indicate that the most distant coordinates from the rotation are hardly affected and the profiles for coordinates R_5 and θ_4 (which are nearly flat) are not included in Figures 3 and 4.
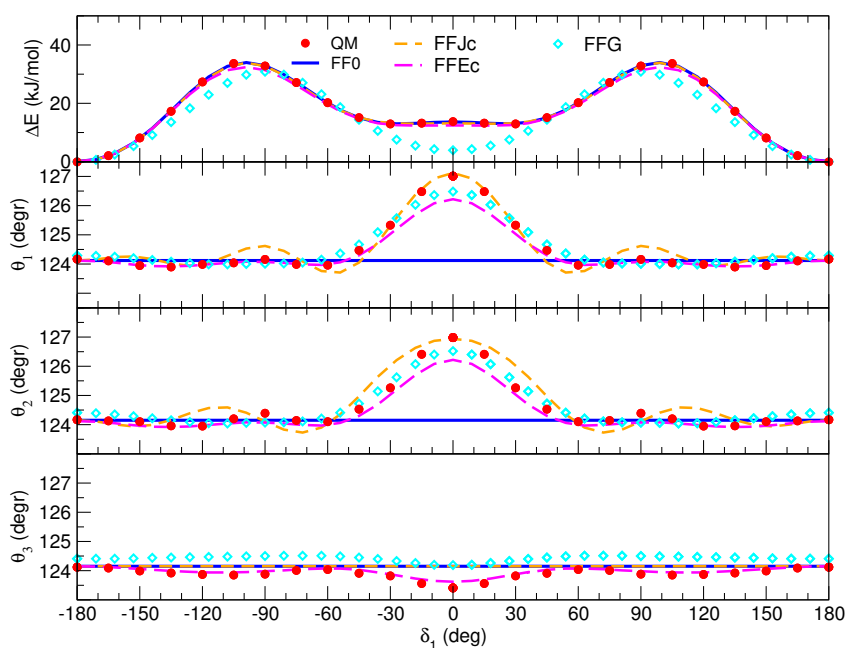

Fig. 4 Changes in the relative potential energy (ΔE) and equilibrium valence angles (θ_1-θ_3), obtained in the constrained minima generated from a relaxed scan around the torsion of R_2 the bond, described by the dihedral δ_1.

In keeping with the analysis for butadiene, we now describe the variations in geometrical parameters along the relaxed scan, evaluated with different FFs. It is worth mentioning that each FF was specifically re-parameterized over the QM data obtained for compound **II**. Nonetheless, since each label refers to the fitting protocol more than to the values of each parameter set, to ease the discussion and the comparison between the two compounds the same FF labels employed for butadiene were employed for hexatriene. As expected, when the uncoupled **FF0** is used, all geometrical
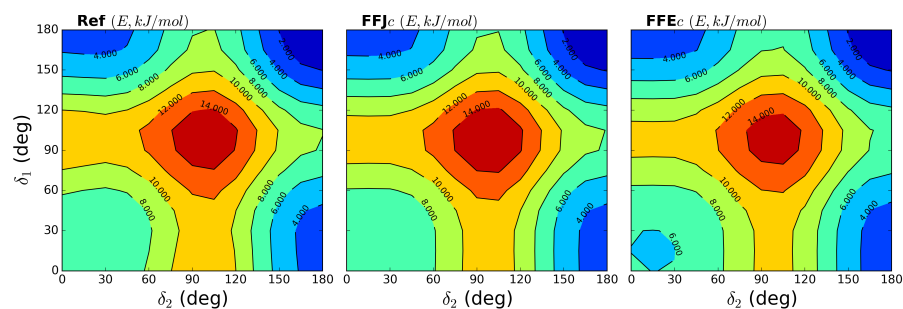
Fig. 5 Torsional energy profiles (ΔE, kJ/mol) at different constrained minima, characterized by fixing the value of both flexible dihedrals δ_1 and δ_2, computed with the reference QM method (Ref, first column), and with the two FFs that include explicit couplings: $\textbf{FFJ}_c$ and $\textbf{FFE}_c$ (second and third column, respectively).

parameters remain constant along the scan. As in the case of butadiene, the use of the implicit coupling model with GAFF leads to an incorrect estimation of the changes observed in the bond lengths in the QM reference, while the change in the angles (mainly related to steric effects) is reproduced with a reasonable agreement. Again, the use of explicit coupling terms ($\textbf{FFE}_c$ and $\textbf{FFJ}_c$), provides a way to include the right behavior of the bond length along the relaxed scan. As commented above, the effect on the bond length is reduced as the bonds involved are more distant from the rotated bond and, in the case of $\textbf{FFJ}_c$ explicit couplings are only considered between the torsion and R_1, R_2 and R_3.

The results in Figures 3 and 4 refer to mono-dimensional scans, obtained by varying only one dihedral and allowing the other one to relax. Hexatriene, however, has two rotable dihedrals, and it is interesting to investigate up to what extent the parameterization performed using the mono-dimensional scans as database is able to provide accurate parameters to describe the effective two-dimensional PES. With this aim further QM calculation were performed, constraining the two dihedrals over a defined grid, and optimizing the rest of the molecule's ICs. Next, the variation of bond lengths along the 2D relaxed scans was monitored using the QM reference method and the two FF with explicit couplings ($\textbf{FFJ}_c$ and $\textbf{FFE}_c$). All results are displayed in Figures 5 and 6. In Figure 5, we first report the energy landscapes. Interestingly, both coupled FFs are able to capture the main features of the reference bi-dimensional plot. More in detail, JOYCE parameterization provides results in better agreement with the reference one, which seems to indicate that a global fitting of all terms in the FF provides a more consistent parameterization, when the number of relevant couplings is increased. Regarding the variation of the bond lengths, displayed in Figure 6, again both FFs are able to reproduce the main trends observed within the reference QM values. In this case, the coupling between the two dihedrals is more evident. It should be noticed that the bond labeled R_1 corresponds to the bond closer to the rotated one taking δ_1 as reference, but it corresponds to R_5 with respect to a rotation of δ_2. Since the coupling between the dihedral and R_5 was not included in the global fitting, this explains why its effect is completely absent when using $\textbf{FFJ}_c$. The changes in this case are however rather limited and the neglection of the coupling would not result

in significant errors along a simulation. Turning to R_2, it becomes R_4 with respect to δ_2. The primary effect is again the one due to the displacements in δ_1, which is the torsion defined around R_2. Here the trends of the cross couplings are well reproduced with both FFs, and a similar behavior is observed for R_3 (which maintains the same label with respect to both dihedrals). In all cases, **FFJ**$_c$ deviates more from the QM reference in the conformations where one of the angles (or both) is close to $90°$, which correspond to unstable repulsive conformations.

Finally, it is interesting to compare how analogous coupling constants vary between the two compounds, as this provides information about the transferability of these parameters. In Table 1, we show the values corresponding to the coupling between the dihedral angle and the central C-C bond, including the results correspond-
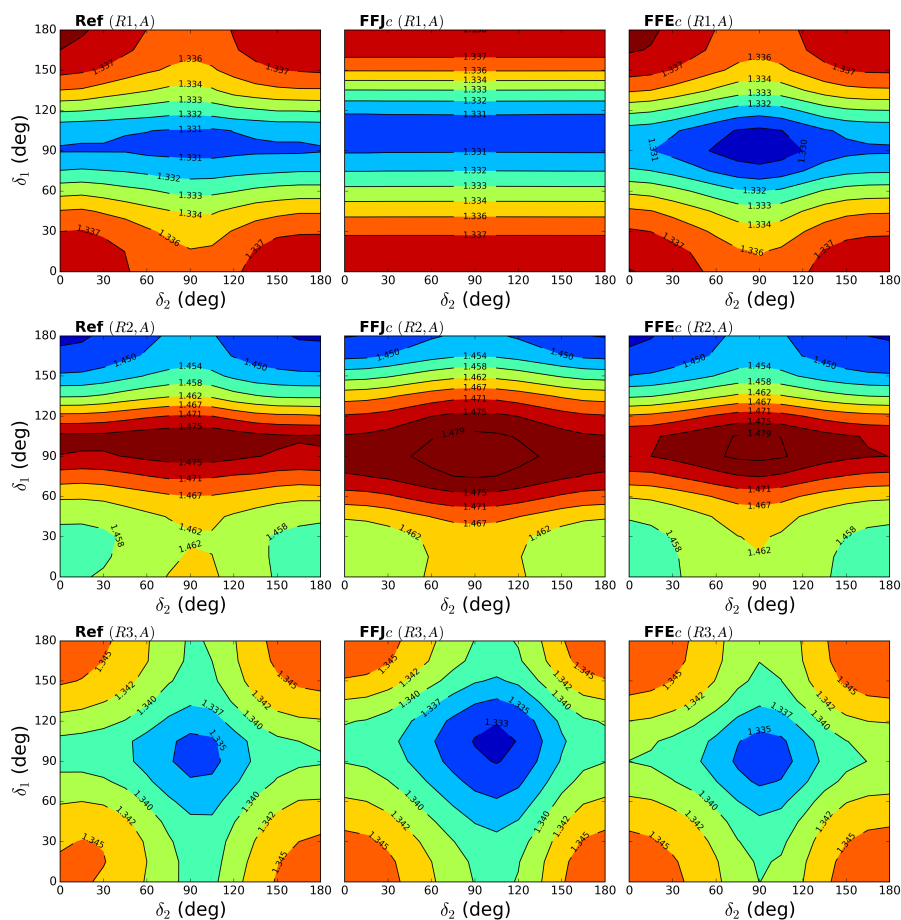

Fig. 6 Changes in bond lengths (R_1, R_2 and R_3, all in Å, in the first, second and third row, respectively) at different constrained minima, characterized by fixing the value of both flexible dihedrals, δ_1 and δ_2, computed with the reference QM method (Ref, first column), and the two FF that include explicit couplings: **FFJ**$_c$ and **FFE**$_c$ (second and third column).

Table 1 Coupling force constants, k_i^c [kJ (mol nm)$^{-1}$], corresponding to the coupling between a dihedral angle and the central single bond, computed for both butadiene and hexatriene with **FFJ**$_c$ and **FFE**$_c$ strategies.

i	FFJ$_c$		FFE$_c$	
	Butadiene	Hexatriene	Butadiene	Hexatriene
1	-75.37	-73.63	-91.09	-85.61
2	-264.88	-336.61	-264.58	-325.98
3	-97.45	-81.89	-110.98	-92.63
4	12.62	7.99	52.50	63.33
5	-21.04	-30.25	24.84	24.07
6	-0.45	-0.29	1.72	-4.11

ing to both **FFJ**$_c$ and **FFE**$_c$ force fields. We first focus on the results obtained with **FFJ**$_c$, and compare the coupling constants (k_i^c, $i = 1 - 6$) obtained for butadiene and hexatriene. From a qualitative level, the same general trend is observed regarding the relative values of each order of the Fourier series: in both molecules the largest value is obtained for $i = 2$, and the sign of the force constant for each order is the same. A closer inspection of the quantitative values reveals however significant differences between the two molecules. For instance, the major coefficient (k_2^c) changes from ~ -265 kJ/(molnm) in butadiene to ~ -337 kJ/(molnm) in hexatriene, i.e., it undergoes an increase (in absolute value) of $\sim 26\%$. Similar effects are observed with **FFE**$_c$. Interestingly, the relative changes on each term of the Fourier series are in the same direction in both cases and, for instance, k_2^c increases in **FFE**$_c$ by $\sim 23\%$. Turning to the C=C bonds adjacent to the torsion, the corresponding force constants are shown in Table 2. Note that for hexatriene the two adjacent C=C bonds are not equivalent and they are marked as outer (R1) and inner (R3). In this case, the relative values between the terms along the Fourier series also follows the same trend for both molecules (and in the case of hexatriene, also for both outer and inter bonds), with all terms of the same order having the same sign. Nevertheless, at a quantitative level, also in this case the variations between both molecules, and even between the non equivalent C=C bonds in hexatriene, are non-negligible. After analyzing the values reported in Tables 1 and 2, the general message is that even though some degree of transferability is observed, it would be more convenient to carry out the parameterization of the couplings for new molecules if we aim at high accuracy. In any case, guessed sets of parameters based on the ones computed for butadiene and hexatriene, may be adequate to account for the conjugation effect in larger polyenic chains at a qualitative level.

5 Conclusions

In this contribution we have investigated the effect of explicit coupling terms added to standard intramolecular force fields, in which only diagonal functions (i.e, depending on one IC) are considered. Concretely, we take into account the coupling between two ICs, where one of them is flexible, such as some molecular torsions, and the other one is stiff, including either stretching of bending degrees of freedom. Such coupling

Table 2 Coupling force constants, k_i^c [kJ (mol nm)$^{-1}$], corresponding to the coupling between a dihedral angle and the adjacent double bonds, computed for both butadiene and hexatriene with **FFJ**$_c$ and **FFE**$_c$ strategies. In the case of hexatriene, two possible adjacent double bonds are possible, either outer or inner, with respect to the polyenic chain.

i	**FFJ**$_c$ Butadiene	Hexatriene (outer)	(inner)	**FFE**$_c$ Butadiene	Hexatriene (outer)	(inner)
1	-18.79	-33.65	-32.07	5.93	2.48	3.38
2	129.91	187.34	159.16	130.00	177.56	143.11
3	40.68	45.63	39.20	14.15	13.17	9.05
4	2.90	2.46	159.16	-23.99	-30.63	-26.79
5	-4.86	-6.46	-6.62	-4.77	-8.62	-7.75
6	-3.40	-4.86	-3.56	4.24	4.97	4.44

terms consist in the product of two functions, each of them depending on one of the ICs only. The stiff coordinates are described by a harmonic function, while the flexible one can be conveniently represented by Fourier-likes series. Namely, a function of the form $\sum_i [1 + \sin(n_i \delta_v - \gamma_i^c)]$, setting γ_i^c to $\pm 90°$, revealed specially well suited for our purposes. The resulting coupled FF was indeed found capable of increased accuracy with respect to diagonal ones. For instance, it is able to describe the change in the equilibrium geometry of the stiff coordinates at different fixed position of the flexible one.

Different parameterization strategies have been presented, either involving the coupled terms only (external fitting) or concerning all terms of the whole FF. In both cases, FF parameters were derived exclusively based on purposely computed quantum mechanical data. In the case of the external fitting, where each coupling term is parameterized independently, the QM database consisted in purely geometrical parameters, such as the profile of the stiff coordinate, along a relaxed scan of the flexible one. As far as the global fitting of the whole FF, is concerned, specific features of the QM PES as relative energies, gradients and Hessian were employed as reference database. This choice is indeed the basis of the strategy developed by some of us and implemented in the code JOYCE,[15] which has been here adapted to properly account for the coupling terms using the same fitting strategy.

In order to investigate the performance of the coupled FFs in practice, we focused on the two simplest members of the polyenic series: butadiene (with one flexible dihedral) and hexatriene (with two flexible dihedrals). In both cases, the torsion along the flexible dihedrals perturbs the conjugation along the chain which, in turn, modifies the equilibrium carbon-carbon bond lengths in the chain. It is indeed a perfect workbench to test our protocols and, moreover, they represent minimal models of more relevant pigments, such as carotenoids, for which the development of coupled FFs is specially attractive within QM/MM methodologies to investigate their photophysical properties (e.g., spectroscopy or excitation energy transfer processes).

Our results on butadiene and hexatriene show that conjugation effects dominate the changes in bond lengths along the relaxed scan of flexible dihedrals, while steric effects played a major role in determining the changes of CCC angles. While standard diagonal FFs with the inclusion intramolecular non-bonded terms are only able

to mimic the latter effect, coupled FFs proved to be able to provide an accurate description also of changes in the PES due to conjugation effects. Yet, non-bonded terms are better suited to describe steric effects, and refined strategies may therefore require a wise combination of both types of terms.

In more general grounds, in this work we showed how the combination of the flexibility of coupled FFs with automatic parameterization procedures, such as the ones implemented in JOYCE, provides a promising strategy to obtain accurate molecular mechanics FFs. In this sense, it is worth to note that, although here the focus was on the conjugation effects, couplings can be inherent to other systems such as constrained cyclic structures. Actually, the protocols presented in this work do not assume any *a priory* specific effect, and the very same methodologies can, in principle, be used to deal with any other type of system.

As perspective, the strategies presented in this work can be applied to parametrize relevant polyenic type compounds, such as carotenoids. The resulting FFs would improve the already successful QMD-FFs already developed for this kind of molecules which, already, can trivially account for the static effects of conjugation, as the change in bond length along the polyenic chain in the equilibrium structure.[12] Coupled FFs can improve the geometries out of equilibrium, such as the distorted dihedral angles along the polyenic chain that can arise along MD simulations. These conformations may be specially relevant within complex biological environments, such as cell membranes[47] or photosystems,[48] where they can be dictated by the specific interactions between the pigment and the heterogeneous surroundings.

References

1. M.P. Allen, D.J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987)
2. D. Frenkel, B. Smith, *Understanding Molecular Simulations* (Academic Press, San Diego, 1996)
3. V. Barone (ed.), *Computational Strategies for Spectroscopy: from Small Molecules to Nanosystems* (John Wiley & Sons, Inc., 2011)
4. W.L. Jorgensen, J. Tirado-Rives, Proc. Natl. Acad. Sci. USA **102**, 6665 (2005). DOI 10.1073/pnas.0408037102
5. A.D. Mackerell, J. Comp. Chem. **25**(13), 1584 (2004)
6. H. Sun, J. Phys. Chem. B **102**, 7338 (1998). DOI 10.1021/jp980939v
7. J. Wang, R.M. Wolf, J.W. Caldwell, P.a. Kollman, D.A. Case, J. Comput. Chem. **25**, 1157 (2004). DOI 10.1002/jcc.20035
8. M. Christen, P.H. Hünenberger, D. Bakowies, R. Baron, R. Bürgi, D.P. Geerke, T.N. Heinz, M.A. Kastenholz, V. Kräutler, C. Oostenbrink, C. Peter, D. Trzesniak, W.F. van Gunsteren, J. Comput. Chem. **26**, 1719 (2005). DOI 10.1002/jcc.20303
9. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A.D. Mackerell, J. Comp. Chem. **31**, 671 (2009). DOI 10.1002/jcc.21367
10. E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J.Y. Xiang, L. Wang, D. Lupyan, M.K. Dahlgren, J.L. Knight, J.W. Kaus, D.S. Cerutti, G. Krilov, W.L. Jorgensen, R. Abel, R.A. Friesner, J. Chem. Theory Comput. **12**, 281 (2016). DOI 10.1021/acs.jctc.5b00864
11. J. Cerezo, F. Santoro, G. Prampolini, Theor. Chem. Acc. **135**(5), 1 (2016). DOI 10.1007/s00214-016-1888-7. URL http://dx.doi.org/10.1007/s00214-016-1888-7
12. O. Andreussi, I. Prandi, M. Campetella, G. Prampolini, B. Mennucci, J. Chem. Theory Comput. **13**, 4507 (2017). DOI 10.1021/acs.jctc.7b00777
13. D.J. Cole, J.Z. Vilseck, J. Tirado-Rives, M.C. Payne, W.L. Jorgensen, J. Chem. Theory Comput. **12**, 2312 (2016)
14. I. Cacelli, G. Cinacchi, G. Prampolini, A. Tani, J. Am. Chem. Soc. **126**(43), 14278 (2004). DOI 10.1021/ja046642u
15. I. Cacelli, G. Prampolini, J. Chem. Theory Comput. **3**(5), 1803 (2007). DOI 10.1021/ct700113h
16. I. Cacelli, A. Cimoli, P.R. Livotto, G. Prampolini, J. Comp. Chem. **33**, 1055 (2012). DOI 10.1002/jcc.22937
17. S. Grimme, J. Chem. Theory Comput. **10**, 4497 (2014). DOI 10.1021/ct500573f
18. A. Vaiana, A. Schulz, J. Wolfrum, M. Saure, J. Smith, J. Comp. Chem. **24**, 632 (2003)
19. T. Verstraelen, D. Van Neck, P. Ayers, V. Van Speybroek, M. Waroquier, J. Chem. Theory Comput. **3**, 1420 (2007)
20. B. Waldher, J. Kuta, S. Chen, N. Henson, A. Clark, J. Comp. Chem. **31**, 2307 (2010)
21. V. Barone, I. Cacelli, N. De Mitri, D. Licari, S. Monti, G. Prampolini, Phys. Chem. Chem. Phys. **15**(11), 3736 (2013). DOI 10.1039/c3cp44179b
22. L. Vanduyfhuys, S. Vandenbrande, T. Verstraelen, R. Schmid, M. Waroquier, V. Van Speybroeck, J. Chem. Theory Comput. **36**, 1015 (2015). DOI 10.1002/jcc.23877. URL http://www.ncbi.nlm.nih.gov/pubmed/25740170
23. Á. Madarász, D. Berta, R.S. Paton, J. Chem. Theory Comput. **12**(4), 1833 (2016). DOI 10.1021/acs.jctc.5b01237
24. F. Zahariev, N. De Silva, M.S. Gordon, T.L. Windus, M. Dick-Perez, J. Chem. Inf. Mod. **57**, 391 (2017). DOI 10.1021/acs.jcim.6b00654
25. J.P. Piquemal, K.D. Jordan, J. Chem. Phys. **147**, 161401 (2017). DOI 10.1063/1.5008887
26. I. Cacelli, C.F. Lami, G. Prampolini, J. Comput. Chem. **30**, 366 (2009)
27. V. Barone, I. Cacelli, A. Ferretti, S. Monti, G. Prampolini, Journal of Physical Chemistry C **115**(10) (2011)
28. V. Barone, I. Cacelli, A. Ferretti, G. Prampolini, G. Villani, J. Chem. Theory Comput. **10**, 4883 (2014). DOI 10.1021/ct500778u
29. N. De Mitri, G. Prampolini, S. Monti, V. Barone, Phys. Chem. Chem. Phys. **16**, 16573 (2014)
30. G. Prampolini, P.R. Livotto, I. Cacelli, J. Chem. Theory Comput. **11**(11), 5182 (2015)
31. G. Prampolini, M. Campetella, N. De Mitri, P.R. Livotto, I. Cacelli, J. Chem. Theory Comput. **12**, 5525 (2016). URL http://pubs.acs.org/doi/abs/10.1021/acs.jctc.6b00705
32. M. Macchiagodena, G. Mancini, M. Pagliai, V. Barone, B.M. Rode, E.C. Meng, T.E. Ferrin, D. Lupyan, M.K. Dahlgren, J.L. Knight, J.W. Kaus, D.S. Cerutti, G. Krilov, W.L. Jorgensen, R. Abel, R.A. Friesner, Phys. Chem. Chem. Phys. **18**, 25342 (2016). DOI 10.1039/C6CP04666E

33. S. Kraner, G. Prampolini, G. Cuniberti, J. Phys. Chem. C **121** (2017). DOI 10.1021/acs.jpcc.7b03923
34. V. Barone, J. Bloino, S. Monti, A. Pedone, G. Prampolini, Phys. Chem. Chem. Phys. **12**, 10550 (2010)
35. A. Pedone, G. Prampolini, S. Monti, V. Barone, Chem. Mater. **23**, 5016 (2011). DOI 10.1021/cm202436b
36. N. De Mitri, S. Monti, G. Prampolini, V. Barone, J. Chem. Theory Comput. **9**(10), 4507 (2013)
37. G. Prampolini, P. Yu, S. Pizzanelli, I. Cacelli, F. Yang, J. Zhao, J. Wang, J. Phys. Chem. B **118**, 14899 (2014)
38. I. Cacelli, A. Ferretti, G. Prampolini, Theor. Chem. Acc. **135**, 156 (2016). URL http://link.springer.com/10.1007/s00214-016-1911-z
39. S. Dasgupta, W.A. Goddard, J. Chem. Phys. **90**, 7207 (1989). DOI 10.1063/1.456250
40. N.L. Allinger, K. Chen, J.H. Lii, J. Comput. Chem. **17**, 642 (1996). DOI 10.1002/(SICI)1096-987X(199604)17:5/6¡642::AID-JCC6¿3.0.CO;2-U
41. S. Dasgupta, K.A. Brameld, C.F. Fan, W.A. Goddard, Spectrochim. Acta Part A **53**, 1347 (1997). DOI 10.1016/S1386-1425(97)00035-8
42. C.S. Ewig, R. Berry, U. Dinur, J.R. Hill, M.J. Hwang, H. Li, C. Liang, J. Maple, Z. Peng, T.P. Stockfisch, T.S. Thacher, L. Yan, X. Ni, A.T. Hagler, J. Comput. Chem. **22**, 1782 (2001). DOI 10.1002/jcc.1131
43. A. Sousa da Silva, W. Vranken, BMC Res. Notes **5**(1), 367 (2012). DOI 10.1186/1756-0500-5-367. URL http://www.biomedcentral.com/1756-0500/5/367
44. 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. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. 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, O. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox. Gaussian 09 Revision D.1 (2009). Gaussian Inc. Wallingford CT
45. Y. Zhao, D. Truhlar, Theor. Chem. Acc. **120**(1-3), 215 (2008). DOI 10.1007/s00214-007-0310-x. URL http://dx.doi.org/10.1007/s00214-007-0310-x
46. S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J.C. Smith, P.M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, Bioinformatics **29**(7), 845 (2013). DOI 10.1093/bioinformatics/btt055
47. J. Cerezo, J. Zuniga, A. Bastida, A. Requena, J.P. Ceron-Carrasco, Phys. Chem. Chem. Phys. **15**, 6527 (2013). DOI 10.1039/C3CP43947J. URL http://dx.doi.org/10.1039/C3CP43947J
48. M. Durchan, J. Tichý, R. Litvín, V. Šlouf, Z. Gardian, P. Hříbek, F. Vácha, T. Polívka, J. Phys. Chem. B **116**(30), 8880 (2012). DOI 10.1021/jp3042796. URL http://pubs.acs.org/doi/abs/10.1021/jp3042796