Scanning Reactive Pathways with Orbital Biased Molecular Dynamics

Leonardo Guidoni† and Ursula Rothlisberger*

Ecole Polytechnique Fédérale de Lausanne (EPFL), Institute of Chemical Sciences and Engineering, Laboratory of Computational Chemistry and Biochemistry, BCH - LCBC, CH-1015 Lausanne, Switzerland

Received March 29, 2005

Abstract: To accelerate reactive events in molecular dynamics simulations we introduce a general bias potential scheme which depends only on the electronic degrees of freedom of the reactive system. This electronic reaction coordinate, which is expressed in terms of a penalty function of the one-electron orbital energies, has been applied to study different reaction pathways of s-cis-butaadiene. Three different reactive channels have been identified: the cis/trans isomerization, the s-cis/s-trans isomerization, and the symmetry allowed cyclization. For the latter, despite the fact that the Woodward–Hoffmann rules are guided by the butadiene frontier orbitals, biasing only these orbitals is not enough to drive the system toward cyclization, but a low-lying valence shell orbital needs to be included.

Introduction

Chemical reactions often have activation barriers that are too high to be observed within the typical picosecond time scale of first-principles Molecular Dynamics (MD) simulations. To investigate the dynamical and thermodynamical properties of these rare events, different methods were developed to overcome the activation barriers within the time scale accessible by computer simulations.1–9 The most popular and straightforward methodology consists of performing constraint dynamics along an a priori defined partial reaction coordinate that is expressed as a function of the ionic positions.1,5 Whereas for some chemical reactions the definition of partial reaction coordinates can be a reasonable assumption, for many others involving, for instance, ionic collective motions or dynamic solvent effects, a proper reaction coordinate cannot be easily expressed via a trivial combination of ionic degrees of freedom. Recently, a more general technique, the so-called metadynamics, has been proposed to partially overcome this drawback.8,9 In this method, a set of different reaction coordinates (that can be either local or collective) are defined, and the free energy landscape of the system is efficiently explored through a Car-Parrinello-type dynamics of this set of variables. However, due to the exponentially growing size of the phase space, the exploration of multiple reaction pathways is usually limited to a small number of local or collective variables.

Both methods cited above are based on the definition of one or more reaction coordinates that are functions of the atomic positions of the system (such as bonds, angles, dihedrals, coordination numbers, etc.). As chemical reactions imply rearrangements of the electronic structure, other quantities, which explicitly depend on electronic instead of ionic properties, might provide helpful information about the reactivity of the system and therefore be useful descriptors for the exploration of reaction pathways.

The idea that the electrons themselves could drive a chemical reaction was developed by Fukui in its work on the role of frontier orbitals in organic reactivity.10,11 Fukui’s investigations revealed the importance of frontier orbitals, namely HOMO and LUMO, in regulating the single molecule reactivity toward external oxidation or reduction. When the analysis is extended to the HOMOs and LUMOs of two reacting molecules, the overlaps between the frontier orbitals of the two molecules becomes the principal indicator to predict probable chemical reactions and stereoselective pathways. In this scheme, the frontier molecular orbitals of the noninteracting reactant species are first considered, and

* Corresponding author e-mail: ursula.roethlisberger@epfl.ch.
† Present address: University of Roma “La Sapienza”, Dip. di Fisica, Edificio Fermi, P.le Aldo Moro 5, 00185 Roma, Italy.
their interaction is taken into account as a perturbation. Orbital selection rules obtained from these considerations are equivalent to those based on orbital symmetries,\textsuperscript{12,13} A quantitative analysis can be carried out with the help of the Fukui functions, namely the derivative of the density with respect to the total number of electrons $f^\pm(r) = (\partial \rho(r)/\partial N)^\pm$. These functions are sensitive indicators of the most reactive parts of a molecule, and they can designate the atoms which are likely involved in nucleophilic ($f^-$), electrophilic ($f^+$), and radical ($f^+ - f^-$) attacks of an external reagent.\textsuperscript{14–16} Further investigation on reactivity in molecular fragment representation has been carried out by Nalewajski using principles based on information theory in order to investigate the equilibrium and nonequilibrium electron distribution between molecular subsystems.\textsuperscript{17}

In a paper from the 1970s Nakatsuji made the distinction between ionic movements that are followed by centroids of the electronic clouds and movements where the latter precede the ions.\textsuperscript{18} Inspired by this ‘electron preceding’ picture, we consider the opportunity of forcing the electronic cloud toward certain directions, consequently pushing the ions to follow them, and therefore driving the system to perform a reaction. At variance with the picture emerging from orbital interaction consideration of the single reactants, in this case the reactive electronic quantities refer to the full reactive system (in our case the global one-electron molecular orbitals), without any further partitioning into molecular subsystems.

This type of global electronic variables can also be included as additional collective reaction coordinates within constraint dynamics\textsuperscript{1,5} or metadynamics\textsuperscript{8,9} schemes. The inclusion of these global electronic descriptors has a 2-fold advantage over standard nuclei-based reaction coordinates. First of all, the introduction of a new type of partial reaction coordinates can help in many cases where an appropriate reaction coordinate is difficult to express in a purely geometrical way. In particular, the approach presented here is specially suitable for introducing collective coordinates, since it is based on one-electron molecular orbitals that can be fully delocalized throughout the reactive system. The second and probably most important issue deals with the possibility to select different reactive channels that could not be easily distinguished by ionic degrees of freedom. Other electronic descriptors, such as the electron density, chemical potential, hardness, softness, and Fukui functions (see refs 16, 19, and 20) can in principle be incorporated in the same way into an electronic constraint. Different reactive channels can be therefore selected and explored by biasing different combinations of these chemically sensitive quantities, without the use of a priori nuclear reaction coordinates.

So far, electronic reaction coordinates have only been applied in a few cases.\textsuperscript{21–23} Recently, VandeVondele and Rothlisberger\textsuperscript{21} introduced a finite electronic temperature in ab initio MD so that electrons in the HOMOs can be partially excited into unoccupied orbitals, increasing the reactivity of the system. Following a similar approach, Vuilleumier and Sprik\textsuperscript{22} have applied an external potential that directly controls the HOMO–LUMO energy gap to explore selected reaction channels. The same idea has been used by Mosey et al. within a MD context.\textsuperscript{23}

All these approaches are limited to the exploration of reaction paths that have a decreasing HOMO–LUMO gap towards the transition state. Other reactive channels that do not satisfy this prerequisite are therefore either explicitly or implicitly excluded by these methods.\textsuperscript{21} The construction of a more general electronic reaction coordinate would allow a more extensive exploration of the reactivity of the system and, at the same time, allow the selection among different reaction pathways.

To drive the electronic reaction coordinates toward the desired direction we used a bias potential scheme. The general electronic bias potential form we introduce in the present paper depends on the one-electron molecular orbital energies $\epsilon_i$. As a test case we have chosen the s-cis-butadiene, a system that is small enough to be fully studied in detail but, at the same time, sufficiently complex to perform different chemical reactions in a rather large energy range (6–55 kcal/mol). Moreover, the reactive channels offered by this system (conrotatory and disrotatory cyclization to cyclobutene, cis–trans isomerization around single and double bonds) can be easily rationalized in terms of orbital correlation diagrams and Woodward–Hoffmann rules. Anticipating our results, we demonstrate that the introduction of a small number of ‘critical’ molecular orbital energies in the construction of the bias potentials is enough to drive the system towards the different reactions within a picosecond time scale.

Calculations

Methodology. We developed our electronic bias potential scheme in the framework of density functional theory. The additional bias potential depends exclusively on electronic degrees of freedom and is applied during Born–Oppenheimer molecular dynamics to drive the system toward different reactive channels. The electronic bias energy $E_{EB}$ was chosen to be a harmonic function of the one-electron Kohn–Sham orbital energies $\epsilon_i$, which satisfy the Kohn–Sham equations $H_{KS}/\psi_i = \epsilon_i \psi_i$. $E_{EB}$ can be written in the following general form:

$$E_{EB} = \left[\frac{1}{2}\alpha \sum_i w_i \left(\epsilon_i - E_i\right)^2\right].$$

The harmonic constant $\alpha$ determines the global strength of the constraint and was chosen as a compromise between a large value, that keeps the system close to the desired target, and a small value, that generates small bias force oscillations, and is thus more easily integrated in molecular dynamics. In test runs with $\alpha = 15, 30, 60, 120, 200, 300$, and $400$ a.u. we monitored the fluctuations of the quantity $\sum_i w_i \epsilon_i - E_i$, which turned out to have oscillations of periods ranging from 20 to 6 fs. We have finally chosen the tightest possible value of $\alpha$, which has an oscillation period not smaller than 10 fs, typical of fast molecular vibrational modes such as $C$–$H$ stretching ($\alpha = 200$ a.u.). The index $i$ runs over all the one-electron Kohn–Sham molecular orbitals that are considered in the bias potential, which can be either occupied
or unoccupied states. The weights \( w_i \) are the coefficients of the linear combination of the involved orbital energies \( \epsilon_i \), and \( E_T \) is the target energy which can depend on time. It can be either increased or decreased, starting from its initial value, which is set equal to \( \sum_i w_i \epsilon_i \) in the reactants state. These changes can be done in a discrete number of steps, using the umbrella sampling technique (as used for instance in refs 22 and 25) or in a continuous way by changing smoothly \( E_T(t) = E_T(0) + \alpha t \Delta \) as a function of the simulation time \( t \) up to a final value \( E_T^f \).

The choice of the involved molecular orbitals as well as the weights \( w_i \) is quite arbitrary for a general reaction. Anyway, in many cases, molecular orbital correlation diagrams and chemical intuition can help to guess reasonable electronic reaction coordinates (i.e. \( w_i \) sets) for a specific reaction. In the case of butadiene, which is the subject of the present paper, we will actually show that it is enough to consider only a restricted group of molecular orbitals to provide a good description of the electronic reaction coordinate for many different reactive channels.

The forces on the nuclei \( f_{R_i} \) due to the electronic bias potential \( E_{EB} \) are given by

\[
f_{R_i} = -\frac{\partial E_{EB}}{\partial R_i} = -\alpha \sum_i \left( \sum_j w_j \epsilon_j - E_T \right) \frac{\partial \epsilon_j}{\partial R_i}
\]

The last term of the above expression, namely the derivative of the Kohn−Sham energy \( \epsilon_i \) with respect to the atomic coordinates \( R_i \), is calculated in a perturbative approach within the linear response theory using the same scheme as Villeumier and Sprik. In the special case where \( i \) is a frontier orbital (i.e. \( F \)), \( -\epsilon_i / \partial R_i \) coincides with the nuclear Fukui Function \( \phi_i = (\partial F_i / \partial N)^\uparrow \), namely the response of the force \( F_i \) on the nucleus \( i \) due to oxidation (−) or reduction (+), as shown in ref 22.

Because of the explicit dependence of the electronic bias on selected Kohn−Sham energies \( \epsilon_i \), it turns out to be critical for a proper description to correctly identify the identity of the orbitals along the entire trajectory. Although the density of states of butadiene in vacuo is not high (only 11 occupied valence states), thermal fluctuations as well as the electronic bias cause the orbital energies to cross during the dynamics. To properly recognize these level crossings and identify the orbitals involved in the bias potential, we adopt the following strategy. The Kohn−Sham orbitals at time step \( t+1 \) are compared with the orbitals at time \( t \) by calculating the overlap matrix \( O_{ij} = \langle \psi_i(t) \psi_j(t+1) \rangle \). If no level crossing between orbitals occurred during the time step, the matrix \( O_{ij} \) is close to identity. On the contrary, if the orbital order is changed during the time step, the exchanged orbitals are immediately identified by nonzero off-diagonal elements of \( O_{ij} \), and the correct identity of the orbitals can be restored.

**Computational Details.** The orbital biased Born−Oppenheimer molecular dynamics has been implemented in a modified version of the program CPMD v 3.8 (www.cpmd.org). For the test case studies on butadiene, we used the local density approximation for the exchange and correlation functional. This choice was dictated by the encounter of numerical instabilities in the linear response calculations when GGA functionals are used (see ref 28 for further details). The Kohn−Sham orbitals were expanded in plane waves up to an energy cutoff of 70 Ry, and the interactions with the core electrons were described using Troullier−Martins type pseudopotentials. The temperature was maintained around 300 K by applying a Nose−Hoover thermostat. The computation of the Orbital Biased MD is twice more expensive over conventional ab initio MD, due to the additional energy minimization introduced by the linear response part.

The general form of the bias potential provides a large flexibility to select different reactive channels. Few options are tested here for scanning the reactivity of \( s\text{-}cis\text{-}butadiene \). All dynamics started from the geometry optimized structure. The following three variants of bias potentials (BP1−3) were tested:

**BP1.** The energies of the HOMO and HOMO-1 orbitals are forced to be exchanged. The corresponding electronic bias is \( E_{EB} = (1/2)\alpha (\epsilon_{HOMO} - \epsilon_{HOMO-1} - E_T) \). \( E_T \) ranges from the value \( E_T(t = 0) = (\epsilon_{HOMO} - \epsilon_{HOMO-1}) \) to \( E_T^f = -(\epsilon_{HOMO} - \epsilon_{HOMO-1}) \) calculated on the optimized \( s\text{-}cis\text{-}butadiene \) structure.

**BP2.** The energies of the Lowest Valence Molecular Orbital +1 (LVMO+1) is forced to reach the corresponding value of cyclobutene. The electronic bias potential is \( E_{EB} = (1/2)\alpha (\epsilon_{LVMO+1} - E_T) \). \( E_T \) ranges from the value of \( \epsilon_{LVMO+1} \) of \( s\text{-}cis\text{-}butadiene \) to the corresponding value of cyclobutene.

**BP3.** The energies of the HOMO and the LUMO are forced to exchange. The potential \( E_{EB} = (1/2)\alpha (\epsilon_{LUMO} - \epsilon_{HOMO} - E_T) \) is applied, where \( E_T \) ranges from \( (\epsilon_{LUMO} - \epsilon_{HOMO}) \) to \( -(\epsilon_{LUMO} - \epsilon_{HOMO}) \).

Free energies were calculated for the bias potential BP3 using the umbrella sampling technique. We used 16 windows equally spaced between \( E_T = -3.82 \text{ eV} \) and \( E_T = 0 \text{ eV} \) with \( \alpha = 60 \) to allow enough overlap between two adjacent windows during the dynamics. To better sample the region close to the transition state, 4 additional windows with \( \alpha = 120 \) were added close to \( E_T = 0 \text{ eV} \). For each window 1 ps of MD run was performed, 0.7 ps of which were used for data production.

**Results**

The bias potentials described above are applied to the study of reactive channels of butadiene with different energy barriers \( E_B \): \( s\text{-}cis/s\text{-}trans \) isomerization around the \( C\text{-}C \) single bond (\( E_B \sim 6 \text{ kcal/mol} \)), symmetry allowed cyclization pathway (\( E_B \sim 40 \text{ kcal/mol} \)), and \( cis\text{/trans} \) double-bond isomerization (\( E_B \sim 55 \text{ kcal/mol} \)).

To better understand the following results, it is useful to be reminded that the Woodward−Hoffmann rules for butadiene predict that the conversion to cyclobutene follows a conrotatory instead of a disrotatory pathway. Whereas in the case of the symmetry allowed conrotatory pathway correlation occurs between the \( \pi \)-type HOMO of butadiene and the \( \sigma \)-type HOMO-1 of cyclobutene, in the disrotatory (symmetry forbidden) pathway the \( \pi \)-type HOMO of butadiene correlates with the antibonding \( \pi^* \) LUMO of cyclobutene. The latter pathway, which involves an orbital crossing and the closure of the HOMO−LUMO gap, has therefore a higher
activation barrier and is disfavored by thermal activation. A quantitative correlation diagram between butadiene and cyclobutene during conrotatory cyclization is shown in Figure 1.

**Reaction 1.** Because of the correlation between the two highest occupied orbitals evidenced in Figure 1, it seems, at first glance, reasonable to assume that a bias potential BP1, built from the HOMO and HOMO-1 orbital energies, might be able to drive the reactive butadiene toward conrotatory cyclization. Figure 2 reports the results from a biased dynamics that exchanges $\text{HOMO} \rightarrow \text{HOMO}^{-1}$. Surprisingly, although the orbital energy exchange has been achieved, the applied bias potential does not lead to conrotatory cyclization, as testified by the $C_1-C_2-C_3-C_4$ dihedral angle shifts to about 180 degrees (Figure 2) by rotating around the single bond, leading to $s$-trans-butadiene via a $s$-cis/$s$-trans isomerization. This reaction channel indeed does require the exchange of $\epsilon_{\text{HOMO}}$ and $\epsilon_{\text{HOMO}^{-1}}$, but it has a much lower activation barrier than cyclization ($\sim$6 vs $\sim$40 kcal/mol). The dynamics therefore correctly selected the energy channel with the lowest barrier that is compatible with the applied bias potential.

**Reaction 2.** As shown in the previous paragraph, the exchange between the two highest occupied molecular orbitals is not a sufficient condition to impose conrotatory cyclization. Other low lying MOs seems to be involved in driving this reaction. According to the correlation diagram of Figure 1, the Lowest Valence Molecular Orbital $+1$ (LVMO$+1$) is another orbital that is sensitive to this reactive channel. The shape of this orbital (Figure 1B) explains the reason why the one-particle energy eigenvalue of this low energy orbital is undergoing such a large change during the reaction. The LVMO$+1$ is indeed a delocalized orbital of $\sigma$ character, corresponding to the $C_1-C_2$ and $C_3-C_4$ bond formation, that is therefore sensitive to both of these bond distances (that undergo elongation during cyclization) and to the newly formed $C_1-C_4$ bond. Bias potential forces of BP2 in a representative snapshot of the dynamics show how the effect of the applied potential tends to open the two double bonds (Figure 3B). The applied bias potential
correctly raises $\epsilon_{LVMO+1}$ and leads to the shortening of the distance of the forming C$_1$–C$_4$ bond, as shown in Figure 3A. Upon achievement of the target energy $E_T$, a few steps of room-temperature dynamics were enough to complete the reaction.

**Reaction 3.** In the symmetry forbidden cyclization pathway a HOMO–LUMO inversion occurs. Molecular dynamics of butadiene biased on the potential BP3 that exchanges $\epsilon_{LUMO}$ with $\epsilon_{HOMO}$ is reported in Figure 4. Whereas the energies of the frontier orbitals are correctly exchanged, the $\epsilon_{LVMO+1}$ is not significantly perturbed during this dynamics, in particular it does not rise in energy as expected in the case of cyclization. The C$_1$–C$_4$ distance is not affected by the electronic bias, but the C$_2$–C$_3$–H$_2$–H$_1$ dihedral angle increases to about 180 degrees indicating that the induced reaction is a complete rotation of the –CH$_2$ group around a butadiene double bond. The 90 degree dihedral conformation roughly coincides with the complete closure of the HOMO–LUMO gap.

In addition to the orbital biased dynamics we calculated the free energy profile for the –CH$_2$ rotation using the HOMO–LUMO gap as reaction coordinate. The umbrella sampling results are displayed in Figure 5. The calculated barrier at 0 gap is 68 kcal/mol, ~15 kcal/mol higher than what was previously proposed.$^{33}$ This is due to the use of
The symmetry allowed cyclization occurred preserving the Sprik\textsuperscript{22} carbon plane, as identified previously by Vuilleumier and LUMO gap we found butadiene to assume an asymmetric suggested by orbital correlation diagrams. At zero HOMO and not toward the symmetry forbidden cyclization as potential BP3 drives the system toward double bond rotation.

The results obtained from the application of the electronic bias potential BP1 show that the system can choose the lowest energy pathway compatible with the imposed electronic constraints. In butadiene, both conrotatory cyclization and s-cis/s-trans isomerization reactions have the common feature of exchanging the two highest occupied orbitals, but the latter occurs spontaneously in our biased dynamics because of the lower activation barrier (~6 vs ~40 kcal/mol). To distinguish the two pathways we needed to introduce a low-lying valence molecular orbital in the bias potential BP2 (LVMO+1), which has \( \sigma \) character and is therefore more sensitive to the newly formed \( \text{C}_1-\text{C}_4 \) \( \sigma \) bond. The symmetry allowed cyclization occurred preserving the \( C_2 \) symmetry throughout the path, in agreement with previous calculations.\textsuperscript{32,34,35}

The HOMO–LUMO inversion experienced by bias potential BP3 drives the system toward double bond rotation and not toward the symmetry forbidden cyclization as suggested by orbital correlation diagrams. At zero HOMO–LUMO gap we found butadiene to assume an asymmetric conformation with one \( -\text{CH}_2 \) group perpendicular to the carbon plane, as identified previously by Vuilleumier and Sprik.\textsuperscript{22}

It is interesting to remark that this structure is a transition state of the double bond isomerization and not of the disrotatory cyclization, which is also asymmetric but involves a different geometry.\textsuperscript{33,36,37}

We have also shown, using umbrella sampling in the BP3 case, that thermodynamical quantities, such as the free energy profile, can be computed within the Orbital Bias Molecular Dynamics scheme, using purely electronic reaction coordinates. Extension of the method to local spin density approximation is currently under development to better describe biradicaloid transition states and open shell systems.

The introduction of general electronic reaction coordinates, allows for following different reaction pathways starting from the same reactant, depending on the choice of the bias potential used. Additional improvements can arise using, together with molecular orbitals, other chemical descriptors such as electronegativity, chemical hardness and softness, and population analysis to probe trends in chemical reactivity of large molecular systems.

**Conclusions**

A bias potential which is an explicit function of the Kohn–Sham molecular orbital energies has been implemented to scan the electronic channels of reactive molecules. The first application to the test case system of \( s \)-cis-butadiene was able to identify three different independent reactive channels as well as to drive the reactant molecule to the product states.

The proposed methodology was able to extend the electronic control of reactivity to systems where the HOMO–LUMO gap does not significantly change along the reaction, and thus represents a considerable generalization of the electronic bias methods that have been proposed so far.\textsuperscript{21,22} The use of such ‘electronic reaction coordinates’ seems a promising technique to broaden the computational tools for studying reactivity. First of all, the new electronic coordinates, which are of collective nature, can be a useful complement to the ones based on nuclear geometry. Moreover, they offer the opportunity to directly select the electronic channels through descriptors that have a chemical connotation, such as frontier orbitals, HOMO–LUMO gap, chemical potential, and hardness. The resulting dynamics has the ability, as reported here, of selectively accessing high barrier reactions without exploring those with a low energy barrier. Free energy barriers along the desired electronic reaction coordinates can also be calculated.

Further technical improvements are still necessary to provide the Orbital Biased MD with the generalization necessary to study complex (bio)chemical reactions in condensed phase and in mixed QM/MM calculations. For example, the use of more complex chemical descriptors, such as hardness and softness, and the combination with the metadynamics technique developed by Parrinello and coworkers\textsuperscript{8,9} can be of valuable help in exploring the phase space in electronic coordinates.

**Acknowledgment.** We thank D. Sebastiani, R. Vuilleumier, and I. Tavernelli for the precious help for the implementation of the bias potential in CPMD and M. Sprik and G. B. Bachelet for stimulating discussions.

**References**