Atom Condensed Fukui Function in Condensed Phases and Biological Systems and its Application to Enzymatic Fixation of Carbon Dioxide

Javier Oller, David Adrian Saez, and Esteban Vöhringer-Martinez *

Departamento de Físico-Química, Facultad de Ciencias Químicas, Universidad de Concepción, Concepción, 4070386, Chile

E-mail: evohringer@udec.cl

Abstract

Local reactivity descriptors such as atom condensed Fukui functions are promising computational tools to study chemical reactivity at specific sites within a molecule. Their applications have been mainly focused on isolated molecules in their most stable conformation without considering the effects of the surroundings. Here, we propose to combine QM/MM Born-Oppenheimer molecular dynamics simulations to obtain the microstates (configurations) of a molecular system using different representations of the molecular environment and calculate Boltzmann weighted atom condensed local reactivity descriptors based on conceptual DFT. Our approach takes the conformational fluctuations of the molecular system and the polarization of its electron density by the environment into account allowing us to analyze the effect of the molecular environment on reactivity. In this contribution, we apply the method mentioned above to the catalytic fixation of carbon dioxide by crotonyl-CoA carboxylase/reductase and study if the enzyme alters the reactivity of its substrate compared to an aqueous solution. Our main result is that the protein environment activates the substrate by the elimination of solute-solvent hydrogen bonds from aqueous solution in the two elementary steps of the reaction mechanism: the nucleophilic attack of a hydride anion from NADPH on the α, β unsaturated thioester and the electrophilic attack of carbon dioxide on the formed enolate species.

Introduction

Predicting the chemical reactivity of molecular systems is the main focus of conceptual DFT and its various global and local descriptors.[1–10] These descriptors have provided a theoretical basis to quantify a broad range of concepts and reactivity trends in chemistry. Often the reactivity of a specific site within a molecule is of interest to obtain a particular product as in organic synthesis and local descriptors are used to assign a certain type of reactivity as electrophilicity or nucleophilicity to a specific atom or region within a molecule, predicting

the most probable product. This regioselectivity within molecules has been described with the Fukui function[11–13] and its condensed form.[14–16]

Most research on chemical reactivity descriptors has focused on small molecular systems without considering the molecular environment and different conformations accessible to the systems. Specific interactions of the molecular system with the environment can alter the reactivity, especially in condensed phases. These variations may also depend on the different conformations of the molecule of interest and the surrounding molecules. Biological systems present additional challenges to predict chemical reactivity because molecules adopt different conformations induced by their environment, and the conformational changes occur on different time scales. Previous studies have made first steps to address the conformation and environment-induced changes on chemical reactivity descriptors considering different binding poses in affinity studies or protein structures in their analysis.[17–21] From a statistical mechanics perspective, reactivity descriptors should consider all conformations (microstates) contributing to the ensemble defined by the macroscopic thermodynamic variables as the number of nuclei, volume, and temperature (NVT). Depending on the complexity of the molecular systems this ensemble-averaged descriptors are more or less difficult to obtain, but they certainly would be more representative of the chemical reactivity of molecular systems in condensed phases.

The effect of the molecular environments on global reactivity descriptors as the chemical potential and hardness was addressed by Pearson.[22] Fuentealba *et al.* studied the Fukui function in homogeneous environments as solvents from a theoretical point of view and using the Onsager solvation model.[23,24] More recently, Padmanabhan *et al.* applied more advanced solvation models accounting for the reaction electric field on the solvent-specific cavity surface with a specific dielectric constant.[25,26] These methods account for the polarization of the electron density due to the macroscopic dielectric properties of the environment but neglect local interactions as hydrogen bonds and their intrinsic dynamics which might affect the local reactivity.

Other studies considered the conformational dependence of reactivity descriptors in a vacuum. Parthasarathi *et al.* analyzed global reactivity descriptors along normal modes and internal rotations.[27] The conformational relationship of global reactivity descriptors was addressed with molecular dynamics simulations under the Born-Oppenheimer approximation in vacuum by Liu *et al.*[28,29] More recently, the conformational dependence of the solute and the surrounding solvent molecules have been studied with sequential Monte Carlo sampling and electronic structure calculations by Jaramillo *et al.*[30] Their main finding was that explicit water molecules change global reactivity descriptors mainly when these molecules are described by quantum mechanics together with the solute. In a similar approach Safi *et al.* had also studied local and global reactivity descriptors of small diatomic molecules with the effective fragment potential method to describe the effect of surrounding water molecules and their orientation.[31] The authors created random initial solvent orientations around the different solutes and optimized them to the next local minimum to derive global and local reactivity descriptors. They conclude that hydrogen bonds between solutes and water molecules increase the condensed atomic softness of the electronegative atoms of Lewis bases as NH_3, H_2O or the halides in protic acids, accompanied with changes in the geometry of the solutes.

Here, we propose to combine sampling techniques based on molecular dynamics simulations to obtain the representative microstates (configurations) of the molecular system and then use the Quantum Mechanics / Molecular Mechanics (QM/MM) methodology to calculate local condensed reactivity descriptors in heterogeneous environments with application to different systems as enzymes and solutes surrounded by explicit solvent molecules. These descriptors consider the conformational dependence of the molecular system and the polarization of its electron density by the environment. Ensemble-averaged reactivity descriptors are derived based on statistical mechanics, which may be more representative of the experimentally measured reactivity in the laboratory. This approach is only valid if electrons are not coupled to nuclear motion (Born-Oppenheimer approximation) that is described by

classical mechanics, and if the electrons in the molecule of interest can be separated from the rest of the molecular environment. Under these approximations, this approach is based on all microstates of the configurational phase space defined by nuclei positions weighted by the Boltzmann distribution.

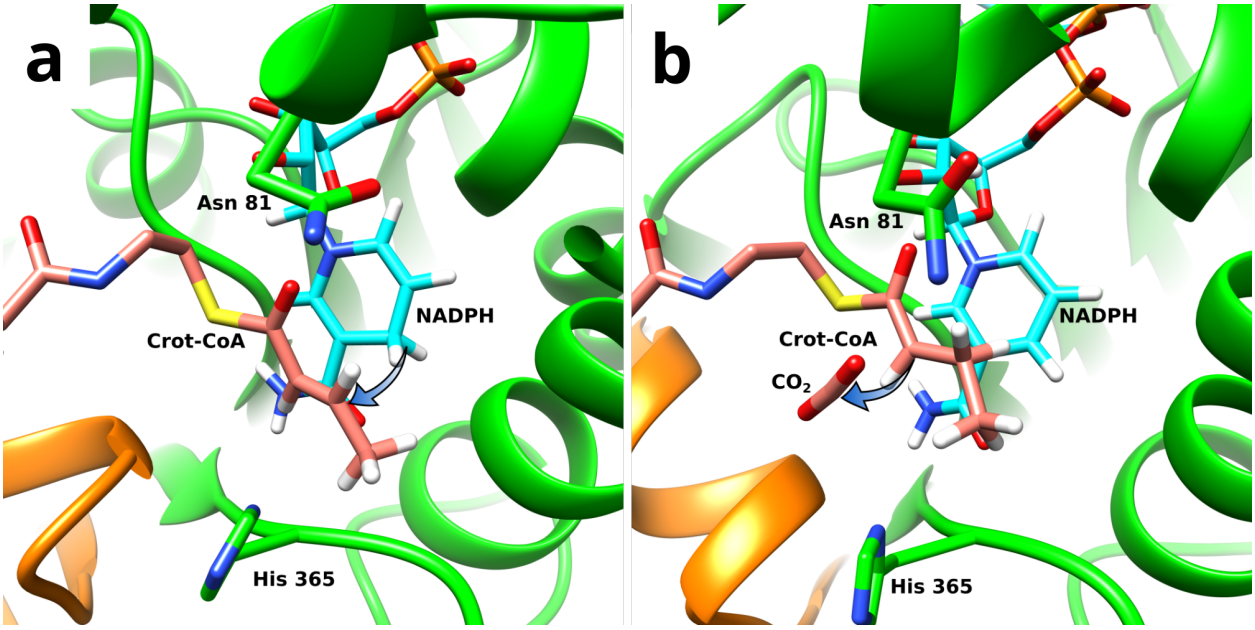

Figure 1: Representative snapshots of the two elementary steps in CO_2 fixation by Crotonyl carboxylase/reductase: a. Nucleophilic attack of the hydride anion from NADPH to the crotonyl-CoA α, β unsaturates thioester. b. Electrophilic attack of carbon dioxide on the the formed enolate species.

As a test system, we study the nucleophilic attack of a hydride anion from NADPH on crotonyl-CoA in the crotonyl-CoA carboxylase-reductase enzyme and the reactivity of the formed anion towards the electrophilic attack of carbon dioxide[32] (see Figure 1). Using the approach mentioned above a possible change of the substrate's reactivity due to a conformational change induced by the enzyme or variations in the molecular environment compared to an aqueous solution has been be analyzed. In our previous study, we have shown that the condensed Fukui functions under the fragment molecular response (FMR) approach with the Hirshfeld-I partitioning method correctly describes the reactivity of α, β unsaturated systems as the crotonyl thioester.[33] Recently, our findings were further confirmed independently also by others.[34,35] Here, we use the validated local reactivity descriptor to analyze the crotonyl test system in different conformations and molecular environments as a vacuum, aqueous

solution, and in the enzyme. The obtained results provide additional insights on how the enzyme is able to enhance the reactivity of the substrate by isolation from solute hydrogen bonds and will allow to study alternative α, β unsaturated substrates able to react with carbon dioxide in the future.

Methods

The reactivity of the crotonyl thioester as the enzyme substrate model was studied with local reactivity descriptors based on conceptual DFT as the Fukui function[11–13] within the linear energy model. The chosen approximations for the Fukui function and the method to condense it to atoms were validated in our previous study on the electrophilicity and nucleophilicity of α, β unsaturated esters, amides, and thioesters[33] and are summarized here briefly.

Condensed Fukui Function

Based on the linear energy model, the Fukui function describing a molecular region more prone to a nucleophilic $f^+_{N_0}(\mathbf{r})$ or electrophilic attack $f^-_{N_0}(\mathbf{r})$ is obtained as the derivative of electron density from above and below with respect to the number of electrons, N, at constant external potential $v(\mathbf{r})$

$$\left(\frac{\partial \rho_N(\mathbf{r})}{\partial N}\right)_{v(\mathbf{r})} |_{N=N_0^+} = \rho_{N_0+1}(\mathbf{r}) - \rho_{N_0}(\mathbf{r}) = f^+_{N_0}(\mathbf{r}) \tag{1}$$

$$\left(\frac{\partial \rho_N(\mathbf{r})}{\partial N}\right)_{v(\mathbf{r})} |_{N=N_0^-} = \rho_{N_0}(\mathbf{r}) - \rho_{N_0-1}(\mathbf{r}) = f^-_{N_0}(\mathbf{r}) \tag{2}$$

where $\rho_N(\mathbf{r})$ is the electronic density of the molecule with N electrons. $\mathbf{r}$ describes a position in the molecular coordinate frame defined by nuclei's position.

Neglecting the orbital relaxation due to the change of the total number of electrons under the frontier molecular orbitals approach(FMO), the Fukui function is approximated by the electronic density of the lowest unoccupied molecular orbital (LUMO) and the highest molecular orbital (HOMO) as follows:

$$f_{N_0}^{+}(\mathbf{r}) = \rho_{N_0+1}(\mathbf{r}) - \rho_{N_0}(\mathbf{r}) = \rho^{LUMO}(\mathbf{r}) \tag{3}$$

$$f_{N_0}^{-}(\mathbf{r}) = \rho_{N_0}(\mathbf{r}) - \rho_{N_0-1}(\mathbf{r}) = \rho^{HOMO}(\mathbf{r}) \tag{4}$$

The linear Fukui function was condensed to atoms with the Hirshfeld-I partitioning method[36] using the fragment of molecular response approach (FMR).[36] This combination of methods reproduced correctly the reactivity of α, β unsaturated systems as reported in our previous study.[33]

The weight function of atom A used to partition the electron density is

$$w_A^{HI}(r) = \frac{\rho_A^0(r)}{\Sigma_{A=1}^{M} \rho_A^0(r)} \tag{5}$$

where $\rho_A^0(r)$ is the spherically averaged proatomic electronic density of atom A obtained as a linear combination of atoms with integer populations closest to the population of the atom in the molecule in an iterative self-consistent approach.

Finally, condensed Fukui function of atom A in the molecule is obtained with the HI weight function as:

$$f_A^{\pm} = \int w_A(r) f^{\pm}(r) dr \tag{6}$$

The atom condensed Fukui function is independent of the molecular coordinate frame used for its calculation, but its derivation depends on the external potential provided by the nuclear position.

Boltzmann Weighted Atom Condensed Fukui function.

Under the Born Oppenheimer approximation where the electrons are uncoupled from the nuclear degrees of freedom, the Fukui function will depend on the positions of the nuclei defining the external potential. We neglect temperature effects on the electronic degrees of freedom and describe nuclei's position and momenta classically to define a canonical ensemble at constant number of distinguishable nuclei (N), temperature T, and the volume V by the canonical partition function

$$Q(N,V,T) = \frac{1}{\lambda^{3N}} \int_{D(v)} d\mathbf{R}_1 ... d\mathbf{R}_N e^{-\beta U(\mathbf{R}_1,...,\mathbf{R}_N)}. \tag{7}$$

Independent momenta and positions of nuclei are assumed, the integration of the momenta is given by $\frac{1}{\lambda^{3N}}$, and the integration extends over the whole volume accessible by the nuclei positions $D(v)$. $U(\mathbf{R}_1, ..., \mathbf{R}_N)$ is the potential energy of the system depending on the position of the nuclei $\mathbf{R}_i$ and can be obtained from ab-initio electronic structure calculations, from properly parametrized force fields or by hybrid methods as QM/MM, which describe the reactive part of the system with ab-initio methods and the rest with force fields. By this approach we explicitly assume that electrons are at 0 K and uncoupled from the nuclear degrees of freedom.

Based on this approximation a configurational partition function

$$Z(N,V,T) = \int_{D(v)} d\mathbf{R}_1 ... d\mathbf{R}_N e^{-\beta U(\mathbf{R}_1,...,\mathbf{R}_N)} \tag{8}$$

is defined to obtain the Helmholtz free energy difference ΔA between two states with the same number of atoms and configurational partition functions, Z_i and Z_j, differing in the atomic interactions as

$$\Delta A_{i \to j} = -kTln\frac{Z_j}{Z_i} \tag{9}$$

The configurational partition function Z in statistical mechanics enables the calculation of an observable of the system as an ensemble average of the corresponding phase space function $a(\mathbf{R}_1, ..., \mathbf{R}_N)$ associated with the observable, which depends only on the position of the nuclei $\mathbf{R}_i$.

Assuming the condensed Fukui function of an atom as an observable representing its reactivity, we define an ensemble-averaged condensed Fukui function as:

$$\langle f_A^+ \rangle = \frac{1}{Z(N,V,T)} \int_{D(v)} d\mathbf{R}_1 ... d\mathbf{R}_N f_A^+(\mathbf{R}_1, ..., \mathbf{R}_N) e^{-\beta U(\mathbf{R}_1, ..., \mathbf{R}_N)} \tag{10}$$

with $\beta = \frac{1}{k_B T}$, T being the temperature and k_B Boltzmann constant. $f_A^+(\mathbf{R}_1, ..., \mathbf{R}_N)$ is the atom condensed Fukui function obtained at one specific configuration of nuclei given by $\mathbf{R}_i$ as described in the subsection above.

For molecular systems, the challenge consists in finding methods capable of providing a complete description of the phase space accessible to the nuclei. As the first step in this direction, here, we first explore the different conformational degrees of freedom of the crotonyl thioester which may show variations at room temperature in vacuum and implicit solvent, represented by rotation around single bonds using DFT methods (see Figure 2). The two resulting minima accessible at room temperature on the potential energy surface (PES) for the crotonyl thioester serve as starting points to study their interconversion with the Umbrella sampling method in different molecular environments (implicit and explicit description of water) combined with the Weighted Histogram Analysis Method(WHAM).[37] All microstates characterized by the atom position of the molecule for each window in the Umbrella Sampling method[38] were used to obtain the averaged local atom-condensed f_A^+ value of each atom for each value of the torsional angle. The ensemble-averaged Fukui function was then obtained with the Boltzmann weight of each window using the calculated free energy difference.

In the enzyme, the crotonyl thioester substrate is constrained by the interactions with the protein residues and strong binding of the Coenzyme-A portion, therefore QM/MM molecu-

lar dynamics simulations of sufficient length to sample the local reorientation of interacting residues and water molecules were used to provide a representation of the accessible phase space of the substrate.

Computational details

A conformational study of free rotable single bonds in the crotonyl thioester was performed to find the local minima on the potential energy surface. Geometries were optimized at the ωB97X / def2-svp level of theory, and free energies of the local minima were calculated under the rigid rotor – harmonic oscillator approximation including zero point energies at 298 K using the Orca package.[39] Two local minima were identified which are accessible at room temperature and 37 configurations with evenly distributed values along the torsional coordinate connecting the two minima were optimized. The same procedure was performed for the calculations using the CPCM SMD implicit solvent model in the Orca package. These 37 conformations served as starting conformations for molecular dynamics simulations in the NVT ensemble employing semiempirical DFTB3 Hamiltonian as implemented in the Amber16 package[40] applying a Langevin thermostat with the following parameters $\Delta t = 1$ fs, $\gamma = 4$ ps^{-1}. The torsional angle was restrained with a harmonic potential and a force constant of k_{restr}= 50 kcal/mol rad^{-2}. To guarantee the correct conservation of energy in the simulations the convergence criteria of the SCF calculation was varied between 10^{-8} and 10^{-11} kcal/mol in the microcanonical ensemble assigning initial velocities from a Maxwell-Boltzmann distribution at 300K. The convergence criteria of 10^{-10} kcal/mol presented no significant energy drift in the employed simulation time and was used for all simulations.

The WHAM method[37] was used to calculate the free energy along the torsional coordinate for the 37 windows. Also the same method was used to obtain the respective probabilities of each window. The corrected free energies at the ωB97X / def2-svp level were obtained by subtracting the average difference of potential energy between the DFTB3 and ωB97X

/ def2-svp level for each window assuming that the entropic contribution remains the same $A(DFT) = (\langle E_{\mathrm{DFT}} \rangle - \langle E_{\mathrm{DFTB3}} \rangle) + A_{\mathrm{DFTB3}}$. Energies at the DFT level were calculated with the ωB97X / def2-svp level and the Terachem program[41] (using the PCM implicit model) and the obtained electron densities served as input to calculate the local condensed Fukui function using the Chemtools program.[42]

In total 200 ps of molecular dynamics simulations in a vacuum and using an implicit or explicit solvent model per window were carried out for converged values of the reactivity descriptors. The MD simulations with implicit solvent were carried out with the Amber program using the Born solvent model at the DFTB3 level of theory[43] with the Umbrella sampling methodology at constant temperature. Finally, 400 structures every 0.5 ps were taken per window (velocity autocorrelation time in explicit solvent was 0.2 ps) for the calculation of the DFT wavefunction using the PCM solvent model in the Terachem program.[41]

The simulations involving an explicit solvent model were also carried out with the Amber program using periodic boundary conditions in an electrostatic embedding, an electrostatic cut-off radius of 8 Å applying the same Langevin thermostat and the TIP3P water model[44] maintaining a minimum distance from the solute to the wall of the cubic simulation box of 15 Å in the setup of the simulation system.

To take into account the enzymatic environment, the reactive complex formed by crotonyl-CoA and NADPH was modelled inside the active cavity of Ccr using the CHARMM22 force field, based on the crystallographic structure and modelling procedure described by Stoeffel *et al.*[45] MM equilibration of the system was carried out (500 ps NVT, 5 ns NPT, 100 ns NVT), and four structures were randomly extracted from the final 20 ns. Using these configurations as a starting point, four independent QM/MM simulations were set up. The QM region, spanning the thioesther tail of the substrate and the nicotinamide ring of the cofactor, was modelled with the DFTB3 Hamiltonian. After minimization of energy, four independent QM/MM molecular dynamics simulations were performed during a minimum of 2ns each.

Uncorrelated structures of all four trajectories were then used in combination with the Amber-Terachem interface including only the *S*-methyl-2-butenethioate part of the Crotonyl-CoA substrate in the QM region applying the link atom scheme and an electrostatic embedding. QM/MM calculations used the ωB97X/def2-svp method for the QM region and the reactivity descriptors were calculated with the Chemtools software[42] from the obtained polarized electron density.

Results

Conformational analysis of the substrate

As a substrate model of the enzyme, we used *S*-methyl-2-butenethioate. This choice neglects the coenzyme-A fragment of the molecule which is primarily responsible for the binding to the protein but not involved in the catalytic reaction. We analyzed the different conformations this molecule may adopt through single bond rotations of dihedrals containing heavy atoms. Figure 2 shows the stable conformers associating the color of the respective rotation of the single bond with the same colored arrow representing the conformational transition. The relative free energy at 298 K is shown below each conformer in kcal/mol calculated in a vacuum at the ωB97X/def2-svp level of theory under the rigid rotor-harmonic oscillator approximation and below using the SMD implicit solvation model representing an aqueous solution. Only two conformers obtained by rotation of the C_α–$C_{C=O}$ bond show a significant population based on the calculated free energy values ($\simeq$3% for the less stable conformer compared to $\simeq$97% for the most stable one). This population distribution also holds in aqueous solution represented with the implicit solvation model although the population of the less stable conformer increases to $\simeq$9%.

Based on this initial conformational analysis, we focused on the dependence of local reactivity descriptors on the conformational transition between the two highlighted conformers in Figure 2.

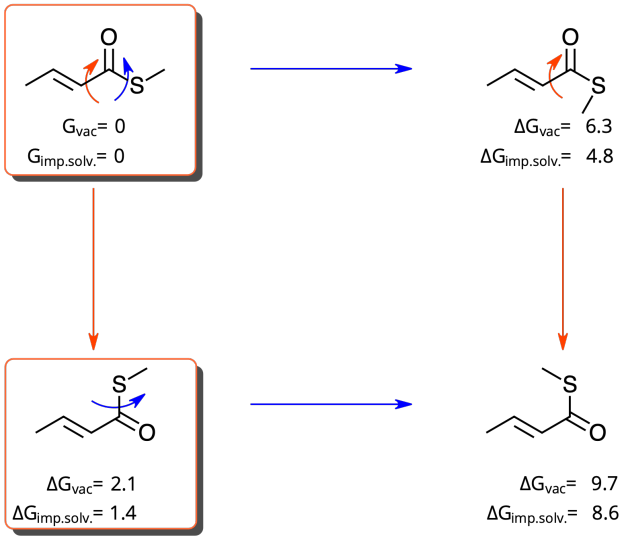

Figure 2: Scheme representing the most stable conformers of the α, β unsaturated thioester with the relative free energies in kcal/mol calculated under the rigid rotor harmonic oscillator approximation at the ωB97X/def2-svp level of theory in a vacuum and with an implicit solvent model representing an aqueous solution.

Effect of the environment and dynamics on atom condensed Fukui functions

From the obtained two conformers, we optimized structures following the conformational transition described by the dihedral angle. The potential energy barrier is 5.1 kcal/mol at the ωB97X/def2-svp level in a vacuum and 4.9 kcal/mol using the implicit solvation model. These values match within the chemical accuracy of 1 kcal/mol the value obtained with the DFTB3 Hamiltonian (5.4 kcal/mol) used for the molecular dynamics simulations. The DFTB3 semiempirical method, however, fails in reproducing the energy difference between the two conformers (180 and 0) making the less stable conformer more favored than the global minimum at the DFT level(Δ E_{DFTB3}(0°- 180°) = 0.5 kcal/mol, see Fig. S1 in the Supporting Information).

Molecular dynamics simulations restraining the dihedral angle with the umbrella sampling method included the effect of nuclear motion and entropy in different environments. Considering nuclear degrees of freedom classically reduced the free energy barrier to 3.9 kcal/mol in a vacuum and 4.3 kcal/mol using the implicit solvation model, maintaining the

overall free energy difference between the conformers. For the explicit solvent the free energy barrier is 5.0 kcal/mol at 300 K and 4.8 kcal/mol at 360 K (see free energy profiles in Fig. S2 in the Supporting Information).

To study the conformational dependence for the nucleophilic attack by the hydride anion on the substrate (see Figure 1) we analyzed the averaged condensed Fukui functions in the conformational transition between the two most stable conformers. The two most prone atoms for the nucleophilic attack are the C_β and the carbonyl carbon atom. Figure 3 shows the averaged condensed Fukui values of these two atoms obtained from dihedral restraint molecular dynamics simulations in a vacuum and with an implicit or explicit representation of the aqueous solvent at 300K. Additionally, for the simulations with an atomistic description of the solvent, we also studied the effect of increasing the temperature to 360 K on the reactivity descriptor.

The condensed Fukui functions assigns the largest reactivity to the C_β carbon atom over the carbonyl one in the region of the most stable conformers (close to 0 and 180°). Inclusion of the implicit solvent reduces the reactivity difference between the two atoms. In the transition state region (aprox. from 60° to 120°) the reactivity changes and the carbonyl carbon atom becomes the most reactive site in the molecule, which is enhanced in the presence of the implicit solvent model possibly due to a more polarized electron density by the reaction electric field of the aqueous solution. The change in reactivity upon rotation is associated with a loss in conjugation of the π electron system involving the $C_\beta - C_\alpha$ and the carbonyl bond. Comparing the averaged local reactivity descriptor from the dynamics to the ones from the optimized structures (see Fig. S4 Supporting Information) we observe that the inclusion of the nuclear dynamics does not change the overall reactivity, although it reduces slightly the minima and maxima of the condensed Fukui values at the transition states.

At the bottom of Figure 3 the condensed Fukui function on the same two most reactive atoms are shown but using an explicit description of the water molecules in QM/MM molec-

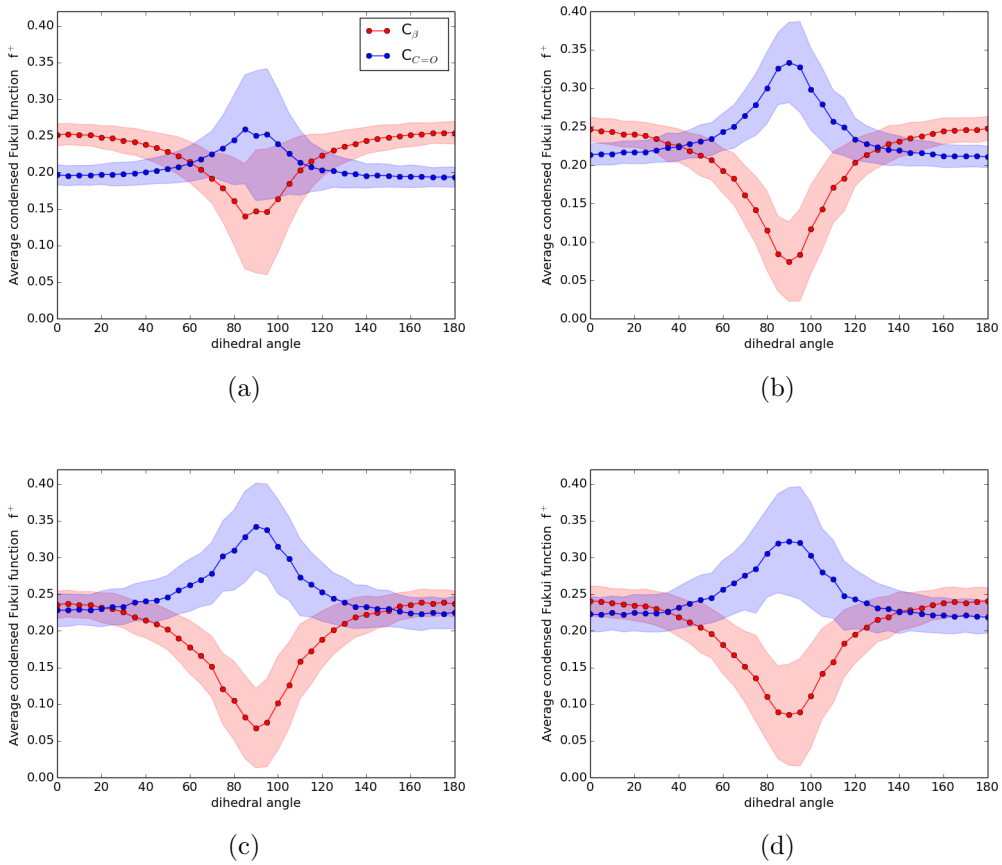

Figure 3: Average condensed Fukui function on the C_β (red) and carbonyl carbon atom (blue) in the α, β unsaturated methyl thioester per umbrella sampling window along the single bond rotation highlighted in Figure 2 in different environments: (a) vacuum (b) using an implicit representation of the aqueous solution (c) QM/MM simulations with explicit water molecules at 300K (d) QM/MM simulations with explicit water molecules at 360K.

ular dynamics simulations at 300K and at 360K on the right. This methodology introduces the electrostatic effect of hydrogen bonds represented by atomic charges together with their dynamics to obtain average local reactivity descriptors. Interestingly, the explicit description of the aqueous environment make the two carbon atoms, within the error, equally prone to a nucleophilic attack in the two most stable conformers, enhancing the reactivity difference in the transition state region. The introduced method to obtain ensemble averaged local descriptors enables studying the effect of the temperature. Increasing the temperature to 360K changes the reactivity for different values of the dihedral angle making it more similar to the one in vacuum due to the decrease in the difference (max - min) of the values in the

transition state region . This change in reactivity is directly related to the larger number of average hydrogen bonds of the carbonyl oxygen atom at lower temperatures as function of the temperature for dihedral angles (see Fig. S3 Supporting Information).

We can conclude that the presence of hydrogen bonds is the key factor which modulates reactivity in the substrate together with the polarizing effect of the reaction electric field of the solvent, and that by a conformational distortion the reactivity of the two most reactive atomic sites for nucleophilic attack in the molecules can be inverted being more feasible in explicit solvent due to their almost equal local reactivity descriptors. This is in line with general organic chemistry knowledge which assigns both atomic sites similar reactivities resulting in different products.

Substrate activation by the Ccr enzyme

Based on the analysis of the environment effects and the conformational variations, enzymes which catalyze the reaction as crotonyl-CoA carboxylase/reductase (Ccr) have different possibilities to make the substrate more reactive on the preferred C_β atom. Upon binding to the protein, a conformational change could be induced as shown recently by us in another enzyme[46] or it could change the molecular environment eliminating hydrogen bonds.

Starting from the crystal structure of the enzyme with bound crotonyl-CoA substrate and solvated by water and ions, we equilibrated the system at 300 K using molecular mechanics to allow proper relaxation of the protein and a correct distribution of the water molecules. In a second step we simulated the reactive region involving NADPH and the crotonyl substrate with QM/MM molecular dynamics (DFTB3 Hamiltonian). Representative structures (7400) were obtained from 2 ns QM/MM trajectories starting from four different initial structures of the equilibration, yielding all four similar average values (see Fig. S6 Supporting Information). Only the α, β unsaturated methyl thioester was included in the QM region to calculate the condensed Fukui function of the reactive C_β and $C_{C=O}$ atom as a Boltzmann weighted average of the different conformations in the enzyme.

Table 1 summarizes the Boltzmann Weighted Atom Condensed Fukui Function (BW-AC-FF) in four different molecular environments accounting for conformational changes as described above. For the non-enzymatic environments we calculated the BW-AC-FF values based on the averages along the dihedral angle applying the Boltzmann weights per window obtained from WHAM analysis. In all four environments the most prone site for a nucleophilic attack is the C_β atom, although in the simulations which account for the hydrogen bonds the reactivity of the $C_{C=O}$ atom becomes almost the same. Comparing the values in different environments it becomes evident that in the Ccr enzyme the reactivity is almost the same as in vacuum and different from the reactivity observed in aqueous solution, described by either implicit or explicit solvent model. This is accompanied with a markedly reduced average number of hydrogen bonds of the oxygen carbonyl atom of the substrate in the enzyme (0.35) compared to aqueous solution (1.62). Hydrogen bonds present in the enzyme mostly involve amino groups of the Asn85 residue which has been identified previously by us as key residue in the catalytic mechanism.[45]

There is no significant conformational change associated with the studied dihedral angle for the substrate in the enzyme because the Coenzyme-A fragment of the substrate fixes its position in the cavity and other residues of the enzyme (Asn85) together with NADPH impede its rotation (see dihedral angles and associated atom condensed Fukui values of the substrate in the enzyme cavity in the Fig. S5 in the Supporting Information).

From this analysis we conclude that the Ccr enzyme increases the reactivity of the substrate effectively shielding it from the aqueous phase and their characteristic hydrogen bonds, and the surrounding residues provide a favorable electrostatic environment which enhances the reactivity. The enzyme, therefore, assigns the most favourable reactivity site within the substrate on the β carbon and increases its reactivity compared to the aqueous solution. Experimentally, it is known that in the enzyme the hydride anion is transferred to this atom prior to the fixation of the CO_2 molecule.

After the formation of the enolate species by the hydride transfer from NADPH an

Table 1: Boltzmann weigthed atom condensed Fukui function $\langle f^+ \rangle$ and $\langle f^- \rangle$ at 300 K (ωB97X / def2-svp) of the most reactive atomic sites in α, β unsaturated methyl thioester in different molecular environments. (Standard deviations in parentheses)

Nucleophilic attack of hydride anion from NADPH				
Atom	vacuum	implicit solvent	explicit solvent	enzyme
C_β	0.250 (0.016)	0.242 (0.017)	0.234 (0.019)	0.248 (0.017)
$C_{C=O}$	0.195 (0.014)	0.214 (0.015)	0.228 (0.022)	0.200 (0.020)
Electrophilic attack of CO_2				
Atom	vacuum	implicit solvent	explicit solvent	enzyme
C_α	0.445 (0.062)	–	0.359 (0.033)	0.448 (0.045)
$O(C=O)$	0.161 (0.037)	–	0.145 (0.047)	0.157 (0.049)

electrophilic attack of CO_2 takes place (see Figure 1). The addition of the hydride anion on the C_β atom changes the reactivity of the substrate with the most probable site for an electrophilic attack being the α carbon atom followed by the carbonyl oxygen atom. Possible effects of the environment on the reactivity of this species was addressed calculating the atom condensed Fukui function f^- as an Boltzmann average from molecular dynamics simulations in the enzyme, in aqueous solution and in a vacuum discussed above. The simulation in the aqueous solution serves as comparison for the reactivity in the enzyme and has no experimental verification since the enolate species would react instantaneously with the solvent.

In Table 1 Boltzmann weighted atom condensed Fukui function for electrophilic attack $\langle f^- \rangle$ of the most reactive sites in the substrate are shown. The enzyme environment presents a reactivity descriptor similar to that observed in a vacuum and different from the aqueous solution, especially on the most reactive α carbon atom. Therefore, also for the second step associated with the carboxylation of the enolate, the enzyme provides a suitable molecular environment, increasing the reactivity on the most reactive site compared to aqueous solution.

Conclusion

We introduced a methodology to account for changes in reactivity at specific atoms in molecules induced by conformational transitions or the molecular environment. Boltzmann weighted atom condensed Fukui function indices (BW-AC-FF) were used to analyze reactivity changes in a conformational transition of the substrate of the CO_2 fixing enzyme CCr considering an explicit description of the surrounding water molecules and enzymatic environment and comparing to the vacuum or the implicit solvation model obtained reactivity. The rotation of a single bond in the α, β unsaturated thioester increases the reactivity at the carbonyl carbon atom at the cost of the β carbon atom in all molecular environments. In principle, enzymes may alter their substrate's reactivity from the aqueous solution by reactant state conformational destabilization and the introduced methodology presents a valuable tool to address these effects in the future.

Hydrogen bonds between the α, β unsaturated thioester and the solvent modulate the reactivity of its two most reactive sites, the β and the carbonyl carbon atom. To address the effect of the hydrogen bonds on the reactivity we used QM/MM simulations accounting for conformational changes of the substrate and the surrounding water molecules or the enzyme. From the QM/MM simulations BW-AC-FF reactivity descriptors are obtained suitable to analyze reactivity in condensed and biological systems.

In their application on the Ccr enzyme we observe enhanced substrate reactivity by desolvation eliminating hydrogen bonds to water molecules. The enzyme's catalytic cavity isolates the substrate and increases its reactivity at the β carbon atom for the experimentally observed hydride transfer from NADPH. Furthermore, the subsequent fixation of CO_2 is also facilitated by an increased reactivity of the enolate species as evidenced by the BW-AC-FF on the α carbon atom.

In summary, the presented methodology provides a new approach to study local reactivity descriptors in condensed phases and biological systems as enzymes.

Supporting Information

Fig. S1 shows the potential energies in vacuum and using the implicit solvent model in the conformational transition between the most stable conformers at different levels of theory. Fig. S2 summarizes the free energy profiles in the conformational transition between the two most stable conformers in vacuum, implicit solvent and explicit solvent. Average number of hydrogen bonds at different temperatures are shown in Fig. S3, and in Fig. S4 the condensed Fukui function from optimized structures along the main dihedral angle rotation in vacuum and implicit solvent model. Fig. S5 and S6 exhibit the fluctuations of dihedral angle and condensed Fukui function in the CCr enzyme simulations and histograms of the condensed Fukui function for different representative conformations of CCr.

Acknowledgement

JO acknowledges Beca Doctorado Nacional 2015 CONICYT folio 21150596 and EVM acknowledges Proyecto Fondecyt 1160197 and financial support by the Max-Planck Society as Partner Group. The authors thank Paul Ayers for carefully reading the manuscript.

References

(1) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.

(2) Ayers, P. W.; Yang, W. *in Density Functional Theory*; P. Bultinck, H. de Winter, W. Langenaeker and J. P. Tollenaere Dekker: New York, 2003.

(3) Geerlings, P.; Proft, F. D.; Langenaeker, W. Conceptual Density Functional Theory. *Chem. Rev.* **2003**, *103*, 1793–1873.

(4) Geerlings, P.; Proft, F. D.; Langenaeker, W. Density Functional Theory : A Source

of Chemical Concepts and a Cost-Effective Methodology for Their Calculation. *Adv. Quantum Chem.* **1998**, *33*, 303–328.

(5) Chermette, H. Chemical reactivity indexes in density functional theory. *J. Comp. Chem.* **1999**, *20*, 129–154.

(6) Ayers, P. W.; Anderson, J. S. M.; Bartolotti, L. J. Perturbative perspectives on the chemical reaction prediction problem. *Int. J. Quantum Chem.* **2005**, *101*, 520–534.

(7) Parr, R. G.; Yang, W. T. Density-Functional Theory of the Electronic Structure of Molecules. *Annu. Rev. Phys. Chem.* **1995**, *46*, 701–728.

(8) Liu, S. B. Conceptual Density Functional Theory and Some Recent Developments. *Acta Phys. Chim. Sinica* **2009**, *25*, 590–600.

(9) Proft, F. D.; Geerlings, P.; Ayers, P. W. *The conceptual density functional theory perspective of bonding. In The chemical bond: Fundamental aspects of chemical bonding*; Eds. Wiley: Darmstadt, 2014.

(10) Johnson, P. A.; Bartolotti, L. J.; Ayers, P. W.; Fievez, T.; Geerlings, P. *Charge density and chemical reactivity: A unified view from conceptual DFT. In Modern Charge Density Analysis*; Eds. Springer: New York, 2012.

(11) Parr, R. G.; Yang, W. T. Density functional approach to the frontier-electron theory of chemical reactivity. *J. Am. Chem. Soc.* **1984**, *106*, 4049–4050.

(12) Yang, W. T.; Parr, R. G.; Pucci, R. Electron density, Kohn–Sham frontier orbitals, and Fukui functions. *J. Chem. Phys.* **1984**, *81*, 2862–2863.

(13) Ayers, P. W.; Levy, M. Perspective on "Density functional approach to the frontier-electron theory of chemical reactivity". *Theor. Chem. Acc.* **2000**, *103*, 353–360.

(14) Bultinck, P.; Alsenoy, C. V.; Ayers, P. W.; Carbó-Dorca, Critical analysis and extension of the Hirshfeld atoms in molecules. *J. Chem. Phys.* **2007**, *126*, 144111.

(15) Ayers, P. W.; Morrison, R. C.; Roy, R. K. Variational principles for describing chemical reactions: Condensed reactivity indices. *J. Chem. Phys.* **2002**, *116*, 8731–8744.

(16) Bulat, F. A.; Chamorro, E.; Fuentealba, P.; Toro-Labbé, A. Condensation of Frontier Molecular Orbital Fukui Functions. *J. Phys. Chem. A* **2004**, *108*, 342–349.

(17) Roos, G.; Geerlings, P.; Messens, J. Enzymatic Catalysis: The Emerging Role of Conceptual Density Functional Theory. *J. Phys. Chem. B* **2009**, *113*, 13465–13475.

(18) Maynard, A. T.; Huang, M.; Rice, W. G.; Covell, D. G. Reactivity of the HIV-1 nucleocapsid protein p7 zinc finger domains from the perspective of density-functional theory. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11578–11583.

(19) Faver, J.; Merz, K. M. Utility of the Hard/Soft AcidBase Principle via the Fukui Function in Biological Systems. *J. Chem. Theory Comput.* **2010**, *6*, 548–559.

(20) Roos, G.; Loverix, S.; Brosens, E.; Van Belle, K.; Wyns, L.; Geerlings, P.; Messens, J. The Activation of Electrophile, Nucleophile and Leaving Group during the Reaction Catalysed by pI258 Arsenate Reductase. *ChemBioChem* **2006**, *7*, 981–989.

(21) Roos, G.; Loverix, S.; De Proft, F.; Wyns, L.; Geerlings, P. A Computational and Conceptual DFT Study of the Reactivity of Anionic Compounds: Implications for Enzymatic Catalysis. *J. Phys. Chem. A* **2003**, *107*, 6828–6836.

(22) Pearson, R. G. Ionization potentials and electron affinities in aqueous solution. *J. Am. Chem. Soc.* **1986**, *108*, 6109–6114.

(23) Fuentealba, P.; Perez, P.; Contreras, R. On the condensed Fukui function. *J. Chem. Phys.* **2000**, *113*, 2544–2551.

(24) Fuentealba, P.; Cedillo, A. The variations of the hardness and the Kohn–Sham Fukui function under an external perturbation. *J. Chem. Phys.* **1999**, *110*, 9807–9811.

(25) Padmanabhan, J.; Parthasarathi, R.; Subramanian, V.; Chattaraj, P. K. Chemical Reactivity Indices for the Complete Series of Chlorinated Benzenes: Solvent Effect. *J. Phys. Chem. A* **2006**, *110*, 2739–2745.

(26) Padmanabhan, J.; Parthasarathi, R.; Sarkar, U.; Subramanian, V.; Chattaraj, P. Effect of solvation on the condensed Fukui function and the generalized philicity index. *Chem. Phys. Lett.* **2004**, *383*, 122–128.

(27) Parthasarathi, R.; Elango, M.; Subramanian, V.; Chattaraj, P. K. Variation of electrophilicity during molecular vibrations and internal rotations. *Theor. Chem. Acc.* **2005**, *113*, 257–266.

(28) Liu, S. Dynamic behavior of chemical reactivity indices in density functional theory: A Bohn-Oppenheimer quantum molecular dynamics study. *J. Chem. Sci.* **2005**, *117*, 477–483.

(29) Zhong, A.; Rong, C.; Liu, S. Structural and Dynamic Properties of $(SiO_2)_6$ Silica Nanostructures: A Quantum Molecular Dynamics Study. *J. Phys. Chem. A* **2007**, *111*, 3132–3136.

(30) Jaramillo, P.; Pérez, P.; Fuentealba, P.; Canuto, S.; Coutinho, K. Solvent Effects on Global Reactivity Properties for Neutral and Charged Systems Using the Sequential Monte Carlo Quantum Mechanics Model. *J. Phys. Chem. B* **2009**, *113*, 4314–4322.

(31) Safi, B.; Balawender, R.; Geerlings, P. Solvent Effect on Electronegativity, Hardness, Condensed Fukui Functions, and Softness, in a Large Series of Diatomic and Small Polyatomic Molecules: Use of the EFP Model. *J. Phys. Chem. A* **2001**, *105*, 11102–11109.

(32) Erb, T. J.; Brecht, V.; Fuchs, G.; Muller, M.; Alber, B. E. Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase. *Proc. Nati. Acad. Sci. USA* **2009**, *106*, 8871–8876.

(33) Oller, J.; Pérez, P.; Ayers, P. W.; Vöhringer-Martinez, E. Global and local reactivity descriptors based on quadratic and linear energy models for α,β-unsaturated organic compounds. *Int. J. Quantum. Chem.* **2018**, *118*, e25706.

(34) Hoffmann, G.; Tognetti, V.; Joubert, L. Can molecular and atomic descriptors predict the electrophilicity of Michael acceptors? *J.Mol. Model.* **2018**, *24*, 281.

(35) Hoffmann, G.; Tognetti, V.; Joubert, L. On the influence of dynamical effects on reactivity descriptors. *Chem. Phys. Lett.* **2019**, *724*, 24–28.

(36) Bultinck, P.; Fias, S.; Alsenoy, C. V.; Ayers, P. W.; Carbó-Dorca, Critical thoughts on computing atom condensed Fukui functions. *J. Chem. Phys.* **2007**, *127*, 034102.

(37) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(38) Kästner, J. Umbrella sampling: Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 932–942.

(39) Neese, F. The ORCA program system. *Rev. Comput. Mol. Sci.* **2011**, *2*, 73–78.

(40) Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N. et al. 2016; Amber 2016; University of California: San Francisco.

(41) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2619–2628.

(42) Heidar-Zadeh, F.; Richer, M.; Fias, S.; Miranda-Quintana, R. A.; Chan, M.; Franco-Pérez, M.; González-Espinoza, C. E.; Kim, T. D.; Lanssens, C.; Patel, A. H. et al. An

explicit approach to conceptual density functional theory descriptors of arbitrary order. *Chem. Phys. Lett.* **2016**, *660*, 307–312.

(43) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948.

(44) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(45) Stoffel, G. M. M.; Saez, D. A.; DeMirci, H.; Vögeli, B.; Rao, Y.; Zarzycki, J.; Yoshikuni, Y.; Wakatsuki, S.; Vöhringer-Martinez, E.; Erb, T. J. Four amino acids define the CO $_2$ binding pocket of enoyl-CoA carboxylases/reductases. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, 201901471.

(46) Vöhringer-Martinez, E.; Dörner, C. Conformational Substrate Selection Contributes to the Enzymatic Catalytic Reaction Mechanism of Pin1. *J. Phys. Chem. B* **2016**, *120*, 12444–12453.

Graphical TOC Entry