

Statistical Mechanical Modeling of Chemical Reactions in Complex Systems: The Reaction Free Energy Surface

A. Amadei,^{*,†} M. D'Alessandro,[§] and M. Aschi[‡]

Dipartimento di Scienze e Tecnologie Chimiche, Università di Roma "Tor Vergata", via della Ricerca Scientifica 1, 00133 Roma, Italy, Dipartimento di Chimica, Università di Roma "La Sapienza", P.le Aldo Moro 5, 00185 Roma, Italy, and Dipartimento di Chimica, Ingegneria Chimica e Materiali, Università dell'Aquila, via Vetoio 67010, l'Aquila, Italy

Received: March 18, 2004; In Final Form: June 21, 2004

In this paper, the perturbed matrix method (PMM) is used in combination with basic statistical mechanics, to develop a general theoretical method to model chemical reactions and related molecular processes in complex systems, i.e., liquids, biochemical systems, macromolecules, etc. The main feature of this approach consists of the explicit treatment of the coupling between the reaction center and the fluctuating atomic–molecular environment, providing a rigorous statistical mechanical description of the chemical event. A special attention is dedicated to the approximations and assumptions necessary to use such a theoretical procedure in combination with simulation data.

1. Introduction

The description at atomic level of a chemical event taking place in a complex molecular system is one of the major challenges of theoretical and computational chemistry. In the past few years, a large number of combined quantum–molecular mechanical (QM/MM) approaches have been proposed and successfully applied to different molecular systems.^{1–6} The first problem to be faced when such procedures are considered, is connected to the approximations which must be included for handling systems with a very large number of degrees of freedom. For this reason many efforts have been made, and are still at the center of a certain interest, to obtain force fields of a sufficient degree of reliability but computationally more flexible than any full ab initio description. Nowadays, it is therefore possible to obtain relatively accurate answers to whatever complex question arises concerning the potential energy hypersurface of a many-particle molecular system. On the other hand, the situation appears less promising if one has to deal with the free energy associated with a chemical event in a complex molecular system.⁷ In this case it is in fact well-known that, beyond the accuracy of the force field used, the statistical sampling is the major problem as the presently available QM/MM schemes do not easily provide electronic properties for a large set of molecular configurations. Alternatively semiempirical methods may be used,⁸ but at the price of a less rigorous modeling. We have recently proposed a method, the perturbed matrix method (PMM),^{9,10} which deals with the inclusion of the electronic degrees of freedom into a classical molecular dynamics (MD) or Monte Carlo (MC) simulation, based on quantum mechanical first principles. Such a method seems very promising for treating complex molecular systems^{11–13}

where the usual QM/MM procedures are computationally problematic, and if combined with basic statistical mechanics, this method can provide a powerful, rigorous theoretical approach to obtain a detailed model for chemical reactions occurring in complex systems. In these last few years, modeling of chemical reactions in complex environment has become the center of very active interest, and several theoretical–computational methods have been proposed and successfully applied.⁷ These approaches provided detailed and accurate information for the quantum mechanical process in a reaction event,¹⁴ treating the interacting environment either as a frozen atomic structure, when the chemical event is much faster than the atomic relaxation, or with a simplified mean field description⁷ when the thermodynamics of the reaction is concerned. However, the former do not consider any statistical fluctuations of the molecular environment, which is essential to correctly describe the thermodynamics and the kinetics of a chemical reaction. The latter, although attempting to provide such a coupling, utilizes rather crude approximations, treating the perturbation free energy exerted by environment on the reaction center via macroscopic mean field models typically inaccurate to describe the thermodynamics involved in microscopic systems. In this paper, we address the problem of the calculation of the free energy surface, and related properties, for a reaction defined by a set of classical generalized coordinates (reaction coordinates), exploring in details the statistical mechanical framework and approximations which are to be used.

2. PMM Basic Derivations

Consider a quantum center (i.e., a molecule or a subpart of a molecule) embedded in a classical molecular environment. Defining with r_n the nuclear coordinates of the quantum center and with x the coordinates of the atoms providing the (classical) perturbing field, we can expand¹² the perturbed Hamiltonian matrix \hat{H} of the quantum center on the Born–Oppenheimer surface as

* Corresponding author. E-mail: andrea.amadei@uniroma2.it. Fax: +39-6-72594328

[†] Università di Roma "Tor Vergata".

[‡] Università dell'Aquila.

[§] Università di Roma "La Sapienza".

$$\tilde{H} = \tilde{H}^0 + \tilde{I}q_T \mathcal{V}(\mathbf{r}_0) + \tilde{Z}_1(\mathbf{E}) + \tilde{Z}_2(\tilde{\Theta}) + \dots \quad (1)$$

$$[\tilde{Z}_1]_{ll'} = -\mathbf{E} \cdot \langle \Phi_l^0 | \hat{\boldsymbol{\mu}} | \Phi_{l'}^0 \rangle \quad (2)$$

$$[\tilde{Z}_2]_{ll'} = \frac{1}{2} \text{Tr}[\tilde{\Theta} \tilde{Q}_{ll'}] \quad (3)$$

$$\Theta_{k,k'} = - \left(\frac{\partial E_k}{\partial r_{k'}} \right)_{r=r_0} \quad (4)$$

$$\hat{\boldsymbol{\mu}} = \sum_j q_j (\mathbf{r}_j - \mathbf{r}_0) \quad (5)$$

where \tilde{H}^0 and q_T are the unperturbed Hamiltonian matrix and total charge of the quantum center, $\tilde{Q}_{ll'}$ is the l, l' transition quadrupoles matrix,¹⁰ \mathcal{V} is the perturbing electric potential at the \mathbf{r}_0 position of the quantum center, \mathbf{E} is the perturbing electric field at \mathbf{r}_0 and Φ_l^0 is the l th unperturbed electronic Hamiltonian eigenfunction (we use these eigenfunctions as the basis set). Note that we consider the total charge as invariant in the presence of the perturbation and hence a constant defined by the unperturbed condition.

For a quantum center where the perturbing electric field is almost constant, at least neglecting local atomic interactions typically described by short-range potentials such as the Lennard-Jones one, we can write¹² the perturbed Hamiltonian matrix \tilde{H} of the quantum center on the Born–Oppenheimer (BO) surface as

$$\tilde{H}(\mathbf{r}_n, \mathbf{x}) \cong \tilde{H}^0(\mathbf{r}_n) + q_T \mathcal{V}(\mathbf{r}_0, \mathbf{x}) \tilde{I} + \tilde{Z}_1(\mathbf{E}, \mathbf{r}_n) + \Delta V(\mathbf{r}_n, \mathbf{x}) \tilde{I} \quad (6)$$

where $\Delta V(\mathbf{r}_n, \mathbf{x})$ approximates the perturbation due to all the terms from the quadrupoles on, as a simple short-range potential. For more complex perturbing field we can use a similar expression, although inserting higher order multipole terms, e.g., the quadrupole term \tilde{Z}_2 . The perturbed BO Hamiltonian eigenvalues ϵ_i are, within the approximation given by eq 6

$$\epsilon_i = \epsilon'_i(\mathbf{r}_n, \mathbf{x}) + q_T \mathcal{V}(\mathbf{r}_0, \mathbf{x}) + \Delta V(\mathbf{r}_n, \mathbf{x}) \quad (7)$$

where

$$(\tilde{H}^0 + \tilde{Z}_1) \mathbf{c}_i = \epsilon'_i \mathbf{c}_i \quad (8)$$

and \mathbf{c}_i is the i th perturbed eigenvector. Note that from the set \mathbf{c}_i we can in principle obtain any possible perturbed property, e.g., the perturbed (transition) dipole¹² $\boldsymbol{\mu}_{i,j} = \langle \Phi_i | \hat{\boldsymbol{\mu}} | \Phi_j \rangle$ is

$$\boldsymbol{\mu}_{i,j} = \mathbf{c}_i^{*T} \tilde{\Lambda}_x^0 \mathbf{c}_j + \mathbf{c}_i^{*T} \tilde{\Lambda}_y^0 \mathbf{c}_j + \mathbf{c}_i^{*T} \tilde{\Lambda}_z^0 \mathbf{c}_j \quad (9)$$

$$[\tilde{\Lambda}_x^0]_{ll'} = \langle \Phi_l^0 | \hat{\mu}_x | \Phi_{l'}^0 \rangle \quad (10)$$

$$[\tilde{\Lambda}_y^0]_{ll'} = \langle \Phi_l^0 | \hat{\mu}_y | \Phi_{l'}^0 \rangle \quad (11)$$

$$[\tilde{\Lambda}_z^0]_{ll'} = \langle \Phi_l^0 | \hat{\mu}_z | \Phi_{l'}^0 \rangle \quad (12)$$

where obviously Φ is the perturbed Hamiltonian eigenfunction and \mathbf{c}^{*T} is the transpose of the complex conjugated of \mathbf{c} (typically from quantum chemical calculations \tilde{H} has only real elements, and hence $\mathbf{c} = \mathbf{c}^*$ is a real eigenvector).

3. Statistical Mechanics in the Infinite Dilution Conditions

The partition function of a solute–solvent system for a given electronic state, where for the N solutes we use as (classical) molecular coordinates the center of mass position \mathbf{r}_G , the eulerian

angles θ, ϕ, ψ providing the orientation of the molecular frame and the internal coordinates \mathbf{x}_{in} providing the atom positions in the molecular frame, is^{15,16}

$$Q_T = \sum_I \int \frac{e^{-\beta \mathcal{K}_{T,s}} e^{-\beta \mathcal{K}_T} e^{-\beta \mathcal{U}_{T,I}}}{(1 + \gamma_s)^{N_s} (1 + \gamma)^{N_s} N_s! N! h^{(Nd + N_s d_s)}} \times \int d\mathbf{p}_s d\mathbf{x}_s \prod_{i=1}^N |\sin \theta_i| d\mathbf{x}_i d\mathbf{p}_i \quad (13)$$

$$\mathbf{p}_i = \hat{M}_i \dot{\mathbf{x}}'_i$$

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{r}_{i,G} \\ \theta_i \\ \phi_i \\ \psi_i \\ x_{i,in} \end{pmatrix} \quad \dot{\mathbf{x}}'_i = \begin{pmatrix} \dot{\mathbf{r}}_{i,G} \\ \omega_{i1} \\ \omega_{i2} \\ \omega_{i3} \\ \dot{\mathbf{x}}_{i,in} \end{pmatrix}$$

where $\mathcal{K}_{T,s}$ is the kinetic energy of the N_s solvent molecules, \mathbf{x}_s and \mathbf{p}_s are the classical coordinates and conjugated momenta of the solvent, and the kinetic energy of the N solute molecules is

$$\mathcal{K}_T = \frac{1}{2} \sum_{i=1}^N \mathbf{p}_i^T (\tilde{M}_i)^{-1} \mathbf{p}_i \quad (14)$$

Furthermore, \mathbf{x}_i are the classical coordinates of the i th solute molecule, roto-translational ones plus a subset of the internal coordinates here defined as $\mathbf{x}_{i,in}$, $\mathcal{U}_{T,I}$ is the total potential energy including the l th quantum vibrational term and $1 + \gamma$ and $1 + \gamma_s$ are the symmetry coefficients per molecule for the solute and solvent molecules necessary to correct the partition function from the rotations and intramolecular atomic displacements which correspond to permutations of identical particles which do not change the physical state of the system.^{17,18} Note that quantum nuclear coordinates (responsible of quantum vibrations) are equivalent to classical constrained coordinates, as shown in a previous paper,¹⁰ and hence disappear in the semiclassical partition function. For each i th solute, we used the corresponding transformed momenta, given by the Jacobean $|\sin \theta_i|$, associated with the transformation of the angular velocity from the time derivatives of the Eulerian angles ($\dot{\theta}_i, \dot{\phi}_i, \dot{\psi}_i$) to the projections of the angular velocity vector onto the molecular axes ($\omega_{i1}, \omega_{i2}, \omega_{i3}$), i.e., time derivatives of the rotation angles, defined by

$$\omega_{i1} = \dot{\theta}_i \cos \psi_i + \dot{\phi}_i \sin \theta_i \sin \psi_i \quad (15)$$

$$\omega_{i2} = -\dot{\theta}_i \sin \psi_i + \dot{\phi}_i \sin \theta_i \cos \psi_i \quad (16)$$

$$\omega_{i3} = \dot{\phi}_i \cos \theta_i + \dot{\psi}_i \quad (17)$$

Note that the mass tensor \tilde{M}_i providing the kinetic energy associated with the classical degrees of freedom, is expressed in the form which is appropriate for using $\omega_{i1}, \omega_{i2}, \omega_{i3}$. Finally, h is the Planck constant, d and d_s are the number of classical degrees of freedom in the solute molecules and solvent molecule and $\beta = 1/kT$. For a given configuration we can solve the integral over the solutes momenta using the orthogonal transformations of these momenta which diagonalize each $(\tilde{M}_i)^{-1}$ obtaining

$$\int e^{-\beta(1/2) \mathbf{p}_i^T (\tilde{M}_i)^{-1} \mathbf{p}_i} d\mathbf{p}_i = [(\pi kT)^d \det \tilde{M}_i]^{1/2} \quad (18)$$

Using this last equation, we then have

$$Q_T = \sum_l \int \frac{e^{-\beta \mathcal{K}_{T,s}} e^{-\beta \mathcal{U}_{T,l}} (2\pi kT)^{Nd/2}}{(1 + \gamma_s)^{N_s} (1 + \gamma)^N N_s! N! h^{(Nd + N_s d_s)}} \times \mathbf{d}\mathbf{p}_s \mathbf{d}\mathbf{x}_s \prod_{i=1}^N (\det \tilde{M}_i)^{1/2} |\sin \theta_i| \mathbf{d}\mathbf{x}_i \quad (19)$$

If the N solutes are in the infinite dilution condition, we can simplify the integral in the previous equation considering that there is no interaction between the solute molecules. Hence for the great majority of the solutes roto-translational configurations, the integral on the other coordinates and solvent momenta is a constant (note that in the solute mass tensor no elements depend on the center of mass position and the Eulerian angles). Therefore

$$\int e^{-\beta(\mathcal{U}_{T,l} + \mathcal{K}_{T,s})} \mathbf{d}\mathbf{p}_s \mathbf{d}\mathbf{x}_s \prod_{i=1}^N (\det \tilde{M}_i)^{1/2} |\sin \theta_i| \mathbf{d}\mathbf{x}_i \cong (V 8\pi^2)^N \int e^{-\beta(\mathcal{U}_{T,l} + \mathcal{K}_{T,s})} \mathbf{d}\mathbf{p}_s \mathbf{d}\mathbf{x}_s \prod_{i=1}^N (\det \tilde{M}_i)^{1/2} \mathbf{d}\mathbf{x}_{i,in} \quad (20)$$

where in this last equation we used

$$8\pi^2 = \int_0^\pi |\sin \theta| d\theta \int_0^{2\pi} d\phi \int_0^{2\pi} d\psi \quad (21)$$

$$V = \int d\mathbf{r}_G \quad (22)$$

and $\mathcal{U}_{T,l}$ is now evaluated at fixed roto-translational coordinates of the solute molecules being homogeneously distributed in the volume V of the full system. Note that in the special case we deal with a linear solute molecule where the integral over the Eulerian angles must be performed at ψ fixed and hence reduces to 4π . Finally, considering again that N/N_s is almost zero, we have

$$\int e^{-\beta(\mathcal{U}_{T,l} + \mathcal{K}_{T,s})} \mathbf{d}\mathbf{p}_s \mathbf{d}\mathbf{x}_s \prod_{i=1}^N (\det \tilde{M}_i)^{1/2} \mathbf{d}\mathbf{x}_{i,in} \cong \frac{N_s!}{(n_s!)^N} \left(\int e^{-\beta(\mathcal{U}_l + \mathcal{K}_s)} (\det \tilde{M})^{1/2} \mathbf{d}\mathbf{p}_s \mathbf{d}\mathbf{x}_s \mathbf{d}\mathbf{x}_{in} \right)^N \quad (23)$$

where \mathcal{U}_l is the potential energy including the l th quantum vibrational term of a subsystem defined by a single solute molecule with fixed roto-translational coordinates, in the center of the subsystem, and n_s solvent molecules (n_s is given by the closest integer number to N_s/N) with \mathcal{K}_s the corresponding solvent molecules kinetic energy. The right-hand integral in the last equation is taken over the volume V/N of the subsystem, and \mathbf{x}_s and \mathbf{p}_s are now the coordinates and momenta of the n_s solvent molecules of the subsystem. Clearly \mathbf{x}_{in} and \tilde{M} represent the $3N - 6$ internal coordinates and the mass tensor of the single solute molecule. Note that the factor $N_s!/(n_s!)^N$ is a simple degeneration factor due to the number of ways we can distribute n_s out of N_s solvent molecules in N subsystems. Hence, using $N! \cong N^N e^{-N}$, we can rewrite the whole partition function as $Q_T = (QV)^N/N! \cong (eQV/N)^N$ with Q the partition function of a single subsystem given by

$$Q = \Theta \sum_l \int e^{-\beta(\mathcal{U}_l + \mathcal{K}_s)} (\det \tilde{M})^{1/2} \mathbf{d}\mathbf{x}_s \mathbf{d}\mathbf{p}_s \mathbf{d}\mathbf{x}_{in} \quad (24)$$

$$\Theta = \frac{(2\pi kT)^{d/2} 8\pi^2}{(1 + \gamma)(1 + \gamma_s)^{n_s} n_s! h^{(n_s d_s + d)}} \quad (25)$$

This last result states that we can obtain the exact statistical mechanics for whatever dilute solute (including of course a reactive molecular complex) simply investigating a single solute at a given roto-translational configuration embedded in the solvent (the environment).

4. The Free Energy Reaction Surface

To deal with a chemical reaction it is convenient to express the energy \mathcal{U}_l by the perturbed Hamiltonian matrix as a function of the reaction coordinates $\boldsymbol{\eta}$. Expressing the nuclear coordinates of the quantum center (we consider it as the solute or a part of the solute) as $\mathbf{r}_n = \{\mathbf{x}_q, \boldsymbol{\eta}, \boldsymbol{\xi}\}$ where \mathbf{x}_q are the internal quantum vibrational coordinates, $\boldsymbol{\eta}$ the reaction coordinates (belonging to the solute classical internal coordinates \mathbf{x}_{in}) and $\boldsymbol{\xi}$ the remaining classical coordinates, and defining with $\boldsymbol{\xi}$ all the solute classical internal coordinates except $\boldsymbol{\eta}$, i.e., $\mathbf{x}_{in} = \{\boldsymbol{\xi}, \boldsymbol{\eta}\}$, we have that the free energy change for a chemical transition defined by $\boldsymbol{\eta}_a \rightarrow \boldsymbol{\eta}_b$ is

$$\Delta A = -kT \ln \left[\frac{\sum_l \int e^{-\beta(\mathcal{K}_s + \mathcal{U}_b)} (\det \tilde{M}_b)^{1/2} d\boldsymbol{\xi} \mathbf{d}\mathbf{x}_s \mathbf{d}\mathbf{p}_s}{\sum_l \int e^{-\beta(\mathcal{K}_s + \mathcal{U}_a)} (\det \tilde{M}_a)^{1/2} d\boldsymbol{\xi} \mathbf{d}\mathbf{x}_s \mathbf{d}\mathbf{p}_s} \right] = -kT \ln \langle e^{-\beta(\mathcal{U}_b - \mathcal{U}_a)} (\det \tilde{M}_b / \det \tilde{M}_a)^{1/2} \rangle_a$$

where

$$\langle e^{-\beta(\mathcal{U}_b - \mathcal{U}_a)} (\det \tilde{M}_b / \det \tilde{M}_a)^{1/2} \rangle_a = \frac{\sum_l \int e^{-\beta(\mathcal{U}_b - \mathcal{U}_a)} e^{-\beta(\mathcal{K}_s + \mathcal{U}_a)} (\det \tilde{M}_b / \det \tilde{M}_a)^{1/2} (\det \tilde{M}_a)^{1/2} d\boldsymbol{\xi} \mathbf{d}\mathbf{x}_s \mathbf{d}\mathbf{p}_s}{\sum_l \int e^{-\beta(\mathcal{K}_s + \mathcal{U}_a)} (\det \tilde{M}_a)^{1/2} d\boldsymbol{\xi} \mathbf{d}\mathbf{x}_s \mathbf{d}\mathbf{p}_s}$$

and the subscripts a and b mean that the property is evaluated at $\boldsymbol{\eta}_a$ or $\boldsymbol{\eta}_b$. For those molecular systems where the chemical transition considered is, as usual, not accessible within a simulation (i.e., the Hamiltonian used provides that $\boldsymbol{\eta}$ fluctuations are confined within a very small range around $\boldsymbol{\eta}_a$) the ensemble average of the last equation can be in principle obtained by a molecular simulation taking into account only the configurations where $\boldsymbol{\eta} \cong \boldsymbol{\eta}_a$. However, such a procedure is in general statistically not very efficient since it utilizes only a subpart of the sampling obtained by the simulation, and where molecular dynamics simulations are concerned, often the use of a rather short time step is required, as $\boldsymbol{\eta}$ typically involves bond length coordinates. The use of a constrained simulation where $\boldsymbol{\eta} = \boldsymbol{\eta}_a$ would then be very convenient. However, it must be noted that in the previous equations \tilde{M}_a and \tilde{M}_b are not in general identical to the mass tensors of a constrained system where $\boldsymbol{\eta} = \boldsymbol{\eta}_a$ or $\boldsymbol{\eta} = \boldsymbol{\eta}_b$,^{15,19} and hence we must consider with care when the ensemble average of the previous equations can be obtained using a constrained simulation. In fact we can rewrite the ensemble average as

$$\langle e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})}(\det \tilde{M}_b/\det \tilde{M}_a)^{1/2} \rangle_a = \frac{\langle (\det \tilde{M}_b/\det \tilde{M}_{\xi,b})^{1/2} \rangle_{\eta_b} \langle e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})}(\det \tilde{M}_{\xi,b}/\det \tilde{M}_{\xi,a})^{1/2} \rangle_{\eta_a}}{\langle (\det \tilde{M}_a/\det \tilde{M}_{\xi,a})^{1/2} \rangle_{\eta_a}} \quad (26)$$

$$\langle e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})}(\det \tilde{M}_{\xi,b}/\det \tilde{M}_{\xi,a})^{1/2} \rangle_{\eta_a} = \frac{\sum_l \int e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})} e^{-\beta(\mathcal{K}_{\xi}^+/\mathcal{U}_{a,l})} (\det \tilde{M}_{\xi,b}/\det \tilde{M}_{\xi,a})^{1/2} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s}{\sum_l \int e^{-\beta(\mathcal{K}_{\xi}^+/\mathcal{U}_{a,l})} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s}$$

$$\langle (\det \tilde{M}_b/\det \tilde{M}_{\xi,b})^{1/2} \rangle_{\eta_b} = \frac{\sum_l \int e^{-\beta(\mathcal{K}_{\xi}^+/\mathcal{U}_{b,l})} (\det \tilde{M}_b/\det \tilde{M}_{\xi,b})^{1/2} (\det \tilde{M}_{\xi,b})^{1/2} d\xi dx_s dp_s}{\sum_l \int e^{-\beta(\mathcal{K}_{\xi}^+/\mathcal{U}_{b,l})} (\det \tilde{M}_{\xi,b})^{1/2} d\xi dx_s dp_s}$$

$$\langle (\det \tilde{M}_a/\det \tilde{M}_{\xi,a})^{1/2} \rangle_{\eta_a} = \frac{\sum_l \int e^{-\beta(\mathcal{K}_{\xi}^+/\mathcal{U}_{a,l})} (\det \tilde{M}_a/\det \tilde{M}_{\xi,a})^{1/2} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s}{\sum_l \int e^{-\beta(\mathcal{K}_{\xi}^+/\mathcal{U}_{a,l})} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s}$$

where $\tilde{M}_{\xi,a}$, $\tilde{M}_{\xi,b}$ are the mass tensor (diagonal) blocks corresponding to the ξ velocities, obtained at η_a and η_b . For a non highly flexible solute (where the rotational coupling can be neglected¹⁵), if we consider only a very limited range for the reaction coordinates such that the η transition does not alter significantly the solute structure and therefore its mass tensor determinant, we may assume

$$\begin{aligned} \langle (\det \tilde{M}_a/\det \tilde{M}_{\xi,a})^{1/2} \rangle_{\eta_a} &\cong \langle (\det \tilde{M}_b/\det \tilde{M}_{\xi,b})^{1/2} \rangle_{\eta_b} \\ \det \tilde{M}_a &\cong \det \tilde{M}_b \\ \det \tilde{M}_{\xi,a} &\cong \det \tilde{M}_{\xi,b} \end{aligned}$$

and hence

$$\langle e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})}(\det \tilde{M}_b/\det \tilde{M}_a)^{1/2} \rangle_a \cong \langle e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})} \rangle_{\eta_a} \quad (27)$$

$$\begin{aligned} \langle e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})} \rangle_{\eta_a} &= \frac{\sum_l \int e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})} e^{-\beta(\mathcal{K}_{\xi}^+/\mathcal{U}_{a,l})} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s}{\sum_l \int e^{-\beta(\mathcal{K}_{\xi}^+/\mathcal{U}_{a,l})} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s} \\ \Delta A &\cong -kT \ln \langle e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})} \rangle_{\eta_a} \quad (28) \end{aligned}$$

where $\langle e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})} \rangle_{\eta_a}$ is evaluated in the true constrained ensemble, and so it may be obtained by a constrained molecular simulation with $\eta = \eta_a$.

Moreover, noting that for a small modification of the η coordinates the quantum vibrational energy is likely to be unaffected, we have

$$\mathcal{U}_{b,l} - \mathcal{U}_{a,l} \cong \epsilon_b - \epsilon_a + \mathcal{U}_{\text{env},b} - \mathcal{U}_{\text{env},a} \quad (29)$$

where ϵ , providing the electronic ground state energy, comes from eq 7 and $\mathcal{U}_{\text{env},a}$, $\mathcal{U}_{\text{env},b}$ are the internal energies of the atomic and molecular environment (excluding the interaction with the

quantum center and the quantum vibrational term) obtained when the quantum center is at η_a and η_b respectively. If the environment energy is basically independent of the internal state of the quantum center, as we assume neglecting atomic polarization, then (see eq 7)

$$\begin{aligned} \mathcal{U}_{b,l} - \mathcal{U}_{a,l} &= \epsilon_b - \epsilon_a = \\ &\epsilon'_b - \epsilon'_a + q_T(\mathcal{V}'_b - \mathcal{V}'_a) + \Delta V_b - \Delta V_a \quad (30) \end{aligned}$$

This last equation can be used to simplify our expressions in the usual case when we consider the total partition function as factorized into a quantum vibrational and a semiclassical part. This is equivalent to assume that a set of reference vibrational energies $\epsilon_{\text{ref},l}$ exists such that, for every l th vibrational state^{16,20}

$$\begin{aligned} \int e^{-\beta(\mathcal{U}_{a,l}-\epsilon_{\text{ref},l}+\mathcal{K}_{\xi}^+)} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s &\cong \\ \int e^{-\beta(\mathcal{U}_{a,0}-\epsilon_{\text{ref},0}+\mathcal{K}_{\xi}^+)} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s & \\ \int e^{-\beta(\mathcal{U}_{b,l}-\epsilon_{\text{ref},l}+\mathcal{K}_{\xi}^+)} (\det \tilde{M}_{\xi,b})^{1/2} d\xi dx_s dp_s &\cong \\ \int e^{-\beta(\mathcal{U}_{b,0}-\epsilon_{\text{ref},0}+\mathcal{K}_{\xi}^+)} (\det \tilde{M}_{\xi,b})^{1/2} d\xi dx_s dp_s & \end{aligned}$$

with the zero subscript indicating the vibrational ground-state condition. Hence we can write

$$\langle e^{-\beta(\mathcal{U}_{b,l}-\mathcal{U}_{a,l})} \rangle_{\eta_a} \cong \langle e^{-\beta(\mathcal{U}'_{b,0}-\mathcal{U}'_{a,0})} \rangle_{\eta_a}^0 \quad (31)$$

$$\begin{aligned} \langle e^{-\beta(\mathcal{U}'_{b,0}-\mathcal{U}'_{a,0})} \rangle_{\eta_a}^0 &= \\ \frac{\int e^{-\beta(\mathcal{U}'_{a,0}+\mathcal{K}_{\xi}^+)} e^{-\beta\Delta(\epsilon'+q_T\mathcal{V})} e^{-\beta(\Delta V_b-\Delta V_a)} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s}{\int e^{-\beta(\mathcal{U}'_{a,0}+\mathcal{K}_{\xi}^+)} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s} & \\ \mathcal{U}'_l &= \mathcal{U}_l - \epsilon_{\text{ref},l} \end{aligned}$$

$$\Delta(\epsilon' + q_T \mathcal{V}) = \epsilon'_b - \epsilon'_a + q_T(\mathcal{V}'_b - \mathcal{V}'_a)$$

Finally, a further simplification is possible via

$$\begin{aligned} \langle e^{-\beta(\mathcal{U}'_{b,0}-\mathcal{U}'_{a,0})} \rangle_{\eta_a}^0 &= \\ \frac{\int e^{-\beta(\mathcal{U}'_{a,0}+\mathcal{K}_{\xi}^+)} e^{-\beta\Delta(\epsilon'+q_T\mathcal{V})} e^{-\beta(\Delta V_b-\Delta V_a)} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s}{\int e^{-\beta(\mathcal{U}'_{a,0}+\mathcal{K}_{\xi}^+)} e^{-\beta\Delta(\epsilon'+q_T\mathcal{V})} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s} &\times \\ \langle e^{-\beta\Delta(\epsilon'+q_T\mathcal{V})} \rangle_{\eta_a}^0 & \\ \langle e^{-\beta\Delta(\epsilon'+q_T\mathcal{V})} \rangle_{\eta_a}^0 &= \\ \frac{\int e^{-\beta(\mathcal{U}'_{a,0}+\mathcal{K}_{\xi}^+)} e^{-\beta\Delta(\epsilon'+q_T\mathcal{V})} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s}{\int e^{-\beta(\mathcal{U}'_{a,0}+\mathcal{K}_{\xi}^+)} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s} & \end{aligned}$$

noting that within our approximations

$$\begin{aligned} \frac{\int e^{-\beta(\mathcal{U}'_{a,0}+\mathcal{K}_{\xi}^+)} e^{-\beta\Delta(\epsilon'+q_T\mathcal{V})} e^{-\beta(\Delta V_b-\Delta V_a)} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s}{\int e^{-\beta(\mathcal{U}'_{a,0}+\mathcal{K}_{\xi}^+)} e^{-\beta\Delta(\epsilon'+q_T\mathcal{V})} (\det \tilde{M}_{\xi,a})^{1/2} d\xi dx_s dp_s} &= \\ \frac{\int e^{-\beta(\mathcal{U}'_{b,0}+\mathcal{K}_{\xi}^+)} (\det \tilde{M}_{\xi,b})^{1/2} d\xi dx_s dp_s}{\int e^{-\beta(\mathcal{U}'_{b,0}+\mathcal{K}_{\xi}^+)} e^{-\beta(\Delta V_a-\Delta V_b)} (\det \tilde{M}_{\xi,b})^{1/2} d\xi dx_s dp_s} &= e^{-\beta\Delta A_{\Delta V}} \quad (32) \end{aligned}$$

with $\Delta A_{\Delta V}$ the free energy change between the η_b state and a virtual state identical to the previous except for the short-range

potential which is evaluated at η_a . Because of the small η transition and the weak, short-range nature of ΔV , it is possible to assume $\Delta A_{\Delta V} \cong 0$. Hence,

$$\Delta A \cong -kT \ln \langle e^{-\beta \Delta(\epsilon' + q_T \mathcal{P}')} \rangle_{\eta_a}^0 \quad (33)$$

Moreover, for any electronic property χ of the quantum center, the average at η_b is

$$\begin{aligned} \langle \chi_b \rangle_b &= \frac{\int e^{-\beta(\mathcal{U}_{b,J} + \mathcal{V}_s)} (\det \tilde{M}_b)^{1/2} \chi_b d\xi dx_s dp_s}{\int e^{-\beta(\mathcal{U}_{b,J} + \mathcal{V}_s)} (\det \tilde{M}_b)^{1/2} d\xi dx_s dp_s} \\ &= \frac{\langle e^{-\beta(\mathcal{U}_{b,J} - \mathcal{U}_{a,I})} (\det \tilde{M}_b / \det \tilde{M}_a)^{1/2} \chi_b \rangle_a}{\langle e^{-\beta(\mathcal{U}_{b,J} - \mathcal{U}_{a,I})} (\det \tilde{M}_b / \det \tilde{M}_a)^{1/2} \rangle_a} \end{aligned}$$

which, on the basis of the previous approximations and using derivations similar to the previous ones, can be used to obtain

$$\langle \chi_b \rangle_b \cong \frac{\langle e^{-\beta \Delta(\epsilon' + q_T \mathcal{P}')} \chi_b \rangle_{\eta_a}^0}{\langle e^{-\beta \Delta(\epsilon' + q_T \mathcal{P}')} \rangle_{\eta_a}^0} \quad (34)$$

valid at least for χ properties which are basically independent of the short range interactions (e.g., the perturbed dipoles). Equations 33 and 34 are very suited for calculations based on a constrained simulation, where only the vibrational ground state is considered and the short range energy fluctuations are likely to be not fully sampled because of the fixed subset of classical coordinates. However, in the case where we deal with an unconstrained simulation, the previous approximations would lead to the same expressions although evaluated in the unconstrained ensemble, i.e.

$$\begin{aligned} \Delta A &\cong -kT \ln \langle e^{-\beta \Delta(\epsilon' + q_T \mathcal{P}')} \rangle_a^0 \\ \langle \chi_b \rangle_b &\cong \frac{\langle e^{-\beta \Delta(\epsilon' + q_T \mathcal{P}')} \chi_b \rangle_a^0}{\langle e^{-\beta \Delta(\epsilon' + q_T \mathcal{P}')} \rangle_a^0} \end{aligned}$$

Note that for a rigid or semirigid quantum center $\Delta(\epsilon' + q_T \mathcal{P}')$ can be considered as a function only of η , i.e., independent of the other internal coordinates of the quantum center. Hence, the unperturbed matrices used to construct the Hamiltonian perturbed matrix in PMM can be evaluated at η_a and η_b keeping all the other internal coordinates fixed at a reference position, thus reducing considerably the computational effort needed. Finally, it must be remarked that in the case where a large η transition is concerned and hence some of the previous approximations could be inaccurate, the same approach can be used iteratively to reconstruct the whole reaction path.

5. Conclusions

In this paper, we show that PMM is extremely well suited to construct a general statistical mechanical treatment of chemical reactions in complex systems, based on simulation data. Via the explicit modeling of the coupling between the reactive center and the fluctuating complex environment, it is possible to use rigorous quantum chemical calculations to describe the perturbed reaction path. No special assumptions have been made and the approximations utilized are rather reasonable and usual in the simulation of complex molecular systems. In this paper, we specifically address the problem of the calculation of the free energy reaction surface, and related electronic properties, for chemical reactions which can be defined by a set of classical generalized coordinates (reaction coordinates). Such a free energy surface defines completely the thermodynamics of the reaction and can be used to investigate its mechanism and kinetics. The proposed method is very general and can also be applied to complex reactions where the whole reaction path is too extended to fit within the approximations used. In fact, in such a case, the same procedure can be iteratively applied to the different reaction steps which decompose the whole reaction pathway. These reaction steps could then be used to reconstruct the complete reaction process.

References and Notes

- (1) Bentzien, J.; Muller, R. P.; Florian, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 2293–2301.
- (2) Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298–305.
- (3) Friesner, R.; Beachy, M. D. *Curr. Opin. Struct. Biol.* **1998**, *8*, 257–262.
- (4) Broo, A.; Pearl, G.; Zerner, M. C. *J. Phys. Chem. A* **1997**, *101*, 2478–2488.
- (5) Remsungnen, T.; Rode, B. M. *Chem. Phys. Lett.* **2003**, *367*, 586–592.
- (6) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.
- (7) Straib, M.; Hong, G.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 13333–13344.
- (8) Villa, J.; Warshel, A. *J. Phys. Chem. B* **2001**, *105*, 7887–7907.
- (9) Aschi, M.; Spezia, R.; Di Nola, A.; Amadei, A. *Chem. Phys. Lett.* **2001**, *344*, 374–380.
- (10) Spezia, R.; Aschi, M.; Di Nola, A.; Amadei, A. *Chem. Phys. Lett.* **2002**, *365*, 450–456.
- (11) Spezia, R.; Aschi, M.; Di Nola, A.; Valentin, M. D.; Carbonera, D.; Amadei, A. *Biophys. J.* **2003**, *84*, 2805–2813.
- (12) Amadei, A.; D'Abramo, M.; Zazza, C.; Aschi, M. *Chem. Phys. Lett.* **2003**, *381*, 187–193.
- (13) Aschi, M.; Zazza, C.; Spezia, R.; Bossa, C.; Di Nola, A.; Paci, M.; Amadei, A. *J. Comput. Chem.* **2004**, *7*, 974.
- (14) Warshel, A.; Chu, Z. T. *J. Phys. Chem. B* **2001**, *105*, 9857.
- (15) Amadei, A.; Chillemi, G.; Ceruso, M. A.; Grottesi, A.; Di Nola, A. *J. Chem. Phys.* **2000**, *112*, 9–23.
- (16) Amadei, A.; Apol, M. E. F.; Brancato, G.; Di Nola, A. *J. Chem. Phys.* **2002**, *116*, 4437–4449.
- (17) Landau, L. D.; Lifshitz, E. M. *Statistical physics. Part 1*, 3rd ed.; Pergamon Press: Oxford, England, 1980.
- (18) Amadei, A.; Apol, M. E. F.; Berendsen, H. J. C. *J. Chem. Phys.* **1998**, *109*, 3004–3016.
- (19) Fixman, M. *Proc. Nat. Acad. Sci. U.S.A.* **1974**, *69*, 1527.
- (20) Amadei, A.; Apol, M. E. F.; Berendsen, H. J. C. *J. Chem. Phys.* **1997**, *106*, 1893–1912.