Journal of Mathematical Chemistry 10(1992)139–181

SOLVENT EFFECT THEORIES: QUANTUM AND CLASSICAL FORMALISMS AND THEIR APPLICATIONS IN CHEMISTRY AND BIOCHEMISTRY

O. TAPIA
Department of Physical Chemistry, University of Uppsala, Box 532, S-75121 Uppsala, Sweden

Abstract

The quantum chemical and mathematical background and some new approaches to the modeling of solvent effects are described.

1. Introduction

Solvent effects modulate a number of phenomena in physics, chemistry and biology [1–7]. Their theoretical treatment has followed two different but complementary lines of thought. One blend of theory targets observable quantities, e.g. a direct evaluation of the solvation free energy and/or excitation energy; this domain, known as solvation theory, has recently been discussed in a book edited by Dogonadze et al. [5]. In the other type of scheme, known as the solvent effects theory, the attention focuses on the system's electronic wave function or its classical (intra) molecular degrees of freedom to determine the extent the surrounding medium affects either of them; the changes produced on the structure, molecular and/or dynamical properties of the small system by its embedding in the medium are usually measured with respect to the system in vacuo; the approach has quantum chemical connotations that makes it suitable for studying chemical reacting systems at a microscopic level [6,7].

Solvent effects theory has advanced along two different, albeit convergent, lines. In one of them, emphasis is placed on quantum mechanical aspects to represent the solute (or system of interest), while the surrounding medium is modeled at a classical physics level. In the other one, the representation of the solute and surrounding medium is made with the help of classical statistical mechanics techniques; this approach has benefited from the development of computer-assisted Monte Carlo (MC) and molecular dynamics (MD) simulation schemes; quantum mechanics, when it is employed, serves to generate information on intermolecular potentials, charge distributions or any other relevant molecular information. Here, the fundamentals of these two brands of theory are examined and selected applications discussed.

The quantum mechanical theory of solvent effects on the electronic structure is presented from a unifying viewpoint: the generalized self-consistent reaction field (GSCRF) theory. Related schemes are derived from GSCRF theory. The classical

statistical mechanics scheme is used together with the quantum mechanical approach. The former employs projected Liouvillians for describing the system of interest coupled to a thermal bath [8,9]. Such a formulation is conceptually adequate to treat solvent effects [10] in conjunction with the GSCRF theory.

2. Quantum approach to the theory of solvent effects

For a molecule, or any well-defined subsystem, e.g. a model active site in an enzyme, the solute and surrounding medium are represented by the particles' (nuclei and electrons) Hamiltonians $H_s(\boldsymbol{r}_s, \boldsymbol{R}_s)$ and $H_m(\boldsymbol{r}_m, \boldsymbol{R}_m)$, respectively, and the interaction operator $V_{sm} = V(\boldsymbol{r}_s, \boldsymbol{r}_m, \boldsymbol{R}_s, \boldsymbol{R}_m)$, and are well known [7,12–14]. The latter, in the Coulomb gauge [14], is the electrostatic interaction between all charges in the system and can be written with the help of charge density operators (Ω) as follows:

$$V_{ms} = \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}' \Omega_s(\boldsymbol{r}) T(\boldsymbol{r}-\boldsymbol{r}') \Omega_m(\boldsymbol{r}'), \tag{1}$$

where $T(\boldsymbol{r}-\boldsymbol{r}') = 1/|\boldsymbol{r}-\boldsymbol{r}'|$ is the Coulomb kernel [15], and the subsystem (solute) charge density operator in atomic units is defined by:

$$\Omega_s(\boldsymbol{r}) = -\sum_i \delta(\boldsymbol{r}-\boldsymbol{r}_i) + \sum_{si} Z_{si} \delta(\boldsymbol{r}-\boldsymbol{R}_{si}), \tag{2}$$

with a similar expression for the solvent charge density operator $\Omega_m(\boldsymbol{r})$. In these equations, $\boldsymbol{r}_i$ stands for the ith electron position vector operator: $\boldsymbol{R}_{si}$ is the position vector of the solute sith nuclei; Z_{si} its corresponding nuclear charge; $\delta(\boldsymbol{r})$ is Dirac's delta distribution and $\mathrm{d}\boldsymbol{r}$ is a volume element in real space ($\mathbb{R}^3$). The electron coordinate operators for each subsystem $\boldsymbol{r}_s$ and $\boldsymbol{r}_m$ and the nuclear vector position coordinates of each subsystem $\boldsymbol{R}_s$ and $\boldsymbol{R}_m$ are embedded in the definitions of Ω.

The starting electrostatic Hamiltonian H describing the total system of electrons and nuclei is then:

$$H = H_s(\boldsymbol{r}_s, \boldsymbol{R}_s) + H_m(\boldsymbol{r}_m, \boldsymbol{R}_m) + V(\boldsymbol{r}_s, \boldsymbol{r}_m, \boldsymbol{R}_s, \boldsymbol{R}_m); \tag{3}$$

the wave function $\Psi(\boldsymbol{r}_s, \boldsymbol{r}_m; \boldsymbol{X})$ of the global system and its energy E are obtained, in principle though not in practice, by solving the Schrödinger equation:

$$H\left|\Psi(\boldsymbol{r}_s, \boldsymbol{r}_m; \boldsymbol{X})\right\rangle = E\left|\Psi(\boldsymbol{r}_s, \boldsymbol{r}_m; \boldsymbol{X})\right\rangle. \tag{4}$$

The dynamics of the nuclei is driven by a Hamiltonian $H(\boldsymbol{R}_s, \boldsymbol{R}_m)$ which is obtained after quantum averaging over the electron coordinates with $\Psi(\boldsymbol{r}_s, \boldsymbol{r}_m; \boldsymbol{X})$:

$$H(\boldsymbol{R}_s, \boldsymbol{R}_m) = H_s(\boldsymbol{R}_s) + H_m(\boldsymbol{R}_m) + V(\boldsymbol{R}_s, \boldsymbol{R}_m). \tag{5}$$

The electrons provide now with a potential energy function for the nuclear motions and the inter-system potential energy $V(\boldsymbol{R}_s, \boldsymbol{R}_m)$ is given by:

$$V(R_s, R_m) = V_{sm}$$
$$= \langle \Psi(r_s, r_m; X) | \int dr \int dr' \Omega_s(r) T(r - r') \Omega_m(r') | \Psi(r_s, r_m; X) \rangle. \quad (6)$$

When the separability hypothesis holds, this equation can be rearranged to show, explicitly, the classical electrostatic interactions between both subsystems and the Coulomb exchange interactions which are responsible for short-range repulsive forces between closed shell electronic systems.

The problem resides in the calculation of the wave function for the solute in the presence of the surroundings. Assuming the wave function for the solute, $\Psi_s(r_s; R_s, R_m)$, and surrounding medium, $\Psi_m(r_m; R_m, R_s)$, are known at a given instant t and with a given nuclear configuration $X = (R_m, R_s)$, an approximation to the total wave function can be written down: $\Psi(r_s, r_m; X) = A_{sm} \Psi_s(r_s; R_s, R_m) \Psi_m(r_m; R_m, R_s)$, where A_{sm} is an antisymmetrizer operator between electrons from these two groups s and m. The total Hamiltonian is symmetric to any electron permutation. The expectation value of H with the particular ansatz for $\Psi(r_s, r_m; X)$ is given as:

$$\langle A_{sm} \Psi_s(r_s; R_s, R_m) \Psi_m(r_m; R_m, R_s) | H | A_{sm} \Psi_s(r_s; R_s, R_m) \Psi_m(r_m; R_m, R_s) \rangle$$
$$= \langle \Psi_s(r_s; R_s, R_m) \Psi_m(r_m; R_m, R_s) | H | A_{sm} \Psi_s(r_s; R_s, R_m) \Psi_m(r_m; R_m, R_s) \rangle, \quad (7)$$

since H and A_{sm} commute and $A_{sm} A_{sm} = A_{sm}$. Bearing in mind that the antisymmetrizer can be written as a sum of the identity operator **1** and non-trivial permutations P_{ms}, the interaction term V_{sm} can be cast into:

$$V_{sm} = \langle \Psi_s \Psi_m | \int dr \int dr' \Omega_s(r) T(r - r') \Omega_m(r') | \Psi_s \Psi_m \rangle$$
$$+ \langle \Psi_s \Psi_m | \int dr \int dr' \Omega_s(r) T(r - r') \Omega_m(r') | P_{ms} \Psi_s \Psi_m \rangle, \quad (8)$$

which contains in its first term the complete electrostatic interactions between both subsystems including inductive (polarization) effects, since the wave functions for the subsystems are in principle obtained as solutions of effective Schrödinger equations where both parts are interacting. The second term contains electron exchange effects.

The ansatz used above for the wave function does not contain electron correlation and charge transfer effects among both subsystems. The former are second-order effects, while the latter are first-order effects in perturbation theory language. In the construction of effective Schrödinger equations, both are neglected [6, 7]. This type of second-order effect is responsible for the universal Van der Waals attractive term which has to be added to $V(R_s, R_m)$ in eq. (6) when treating the dynamics of the nuclei.

If solute to solvent charge transfer effects are important, the quantum subsystem has to be defined in such a way so as to include those solvent molecules participating in the transfer. In this manner, for this approach to be valid, one assumes that there is no charge transfer between the subsystems.

2.1. GENERALIZED SCRF THEORY [16]

The construction of an effective Hamiltonian $\mathcal{H}_s$ implies separability of the total wave function Ψ into an antisymmetrized product of wave functions for each subsystem since electrons from both groups are indistinguishable. For closed shell electronic systems, exchange interactions due to the Pauli exclusion principle, which are included in the antisymmetrization of the total wave function, are responsible for molecular shapes. Explicit antisymmetrization of the total wave function is avoided by including the resultant repulsive forces as interatomic potential functions to the configurational space potential energy term $V(\boldsymbol{R}_s, \boldsymbol{R}_m)$ in eq. (6). Analogously, the instantaneous electron dipole–dipole corrections between the subsystems are included as standard Van der Waals potentials in the statistical mechanical treatment. Under these assumptions, the product wave function can be given a Hartree form: $\Psi \approx \Psi_s(\boldsymbol{r}_s; \boldsymbol{R}_s, \boldsymbol{R}_m)\Psi_m(\boldsymbol{r}_m; \boldsymbol{R}_m, \boldsymbol{R}_s)$. By taking the quantum average over $\Psi_m(\boldsymbol{r}_m; \boldsymbol{R}_m, \boldsymbol{R}_s)$ of the total Hamiltonian H and neglecting the self-energy of the surrounding medium, an effective Hamiltonian $\mathcal{H}_s$ is obtained:

$$\mathcal{H}_s(\boldsymbol{r}_s; X) = H_s(\boldsymbol{r}_s, \boldsymbol{R}_s) + \int d\boldsymbol{r} \int d\boldsymbol{r}' \, \Omega_s(\boldsymbol{r}) \langle \Psi_m | T(\boldsymbol{r} - \boldsymbol{r}') \Omega_m(\boldsymbol{r}') | \Psi_m \rangle. \tag{9}$$

The expression for the interaction between both subsystems can be cast into a form where the charge density of the surrounding medium $\rho_m = \int d\boldsymbol{r}' \langle \Psi_m | \Omega_m(\boldsymbol{r}') | \Psi_m \rangle$ appears explicitly. The interaction Hamiltonian now describes the coupling of the solute charge density operator with the electrostatic potential created by the surroundings at fixed X: $V_m(\boldsymbol{r}) = \int d\boldsymbol{r}' \langle \Psi_m | T(\boldsymbol{r} - \boldsymbol{r}') \Omega_m(\boldsymbol{r}') | \Psi_m \rangle$. This potential $V_m(\boldsymbol{r})$ fulfills the classical electrostatics Poisson equation: $\nabla_{\boldsymbol{r}}^2 V_m(\boldsymbol{r}) = \rho_m(\boldsymbol{r})$.

The effective Hamiltonian (9) acquires the simple physical form:

$$\mathcal{H}_s(\boldsymbol{r}_s; X) = H_s(\boldsymbol{r}_s, \boldsymbol{R}_s) + \int d\boldsymbol{r} \, \Omega_s(\boldsymbol{r}) V_m(\boldsymbol{r}; X). \tag{10}$$

For each nuclear configuration, the solute wave function Ψ_s and the effective energy $E_s(X)$ are obtained as a solution of the effective Schrödinger equation:

$$\mathcal{H}_s | \Psi_s \rangle = E_s(X) | \Psi_s \rangle. \tag{11}$$

Equation (10) is not very useful unless the solvent charge density could be given a tractable form. This goal can be attained if one realizes that the solute can be taken as a classical external electrostatic source to the surrounding medium. For this approximation to be accurate, the solute wave function must be fairly well localized in the volume assigned to the solute system; overlap with the surrounding medium must be minimal. The condition is fulfilled when the system of interest is defined as discussed above.

Let $\rho_m^0(\boldsymbol{r})$ be the classical charge density of the surrounding medium in the absence of the solute field. The total solvent charge density is obtained by adding to $\rho_m^0(\boldsymbol{r})$ the polarization charge density $\rho'(\boldsymbol{r})$ set up by the solute electric field $\boldsymbol{e}(\boldsymbol{r})$. This latter density is the divergence of the polarization vector $\boldsymbol{p}(\boldsymbol{r})$ [17]:

$$\rho_m(\boldsymbol{r}) = \rho_m^0(\boldsymbol{r}) - \nabla_r \cdot \boldsymbol{p}(\boldsymbol{r}), \tag{12}$$

where ∇_r is the gradient operator at $\boldsymbol{r}$. Note that the polarization density $\boldsymbol{p}$ at the boundaries and outside the macroscopic volume occupied by the solvent is zero. The effective Hamiltonian eq. (10) acquires an implicit nonlinear structure via the polarization density term:

$$\mathcal{H}_s = H_s + \int \mathrm{d}\boldsymbol{r}\, \Omega_s(\boldsymbol{r}) \left[V_m^0(\boldsymbol{r}) + \int \mathrm{d}\boldsymbol{r}' \boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}') \cdot \boldsymbol{p}(\boldsymbol{r}') \right], \tag{13}$$

where $\boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}') = \nabla_r T(\boldsymbol{r} - \boldsymbol{r}')$ is a unit electric field at $\boldsymbol{r}'$ produced by a unit charge at $\boldsymbol{r}$. The term

$$\Pi(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}' \boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}') \cdot \boldsymbol{p}(\boldsymbol{r}') \mathrm{d}\boldsymbol{r}' \tag{14}$$

is the reaction field (RF) potential at the point $\boldsymbol{r}$ in the volume occupied by the solute. V_m^0 is the potential acting on the quantum system that is generated by the surrounding medium charge density. $V_m^0(\boldsymbol{r})$ and $\Pi(\boldsymbol{r})$ must now be operationally determined.

2.1.1. *Surrounding potential*

$V_m^0(\boldsymbol{r})$ is obtained from model charge densities representing the atoms or molecules in the solvent or surrounding medium. There are several model charge densities currently available. For proteins, Warme–Scheraga effective charges for the amino acid residues have been used in studies of enzyme catalyzed reactions [7]. Current MD or MC program packages [18,19] have model point charges to calculate electrostatic interactions. They result from fittings of empirical data [18–21]. Potential functions for water based on ab initio calculations have also been produced [22]. The construction of intermolecular potentials is still a very active field [22–24].

2.1.2. *Reaction field potential*

The polarization vector can be obtained from model solvent response functions without resorting to a multipolar expansion. In this way, the effective Hamiltonian can be cast as a functional of the solute charge density, thereby expliciting the nonlinear structure of the Hamiltonian eq. (13).

In the classical electrodynamics framework, let $\mathbf{X}(\boldsymbol{r})$ be the static response tensor to the external electric field $\boldsymbol{e}(\boldsymbol{r})$. In the present case, the field is produced by the solute charge density $\rho_s(\boldsymbol{r}') = \langle \Psi_s | \Omega_s(\boldsymbol{r}') | \Psi_s \rangle$ and is given by:

$$\boldsymbol{e}(\boldsymbol{r}) = -\nabla_r \int \mathrm{d}\boldsymbol{r}' \rho_s(\boldsymbol{r}')\, T(\boldsymbol{r} - \boldsymbol{r}') = -\int \mathrm{d}\boldsymbol{r}' \rho_s(\boldsymbol{r}') T(\boldsymbol{r} - \boldsymbol{r}'). \tag{15}$$

The polarization density is then:

$$\boldsymbol{p}(\boldsymbol{r}) = \mathbf{X}(\boldsymbol{r}) \cdot \left[\boldsymbol{e}(\boldsymbol{r}) + \int \mathrm{d}\boldsymbol{r}'\, \mathbf{T}(\boldsymbol{r} - \boldsymbol{r}') \cdot \boldsymbol{p}(\boldsymbol{r}') \right], \tag{16}$$

where $\mathbf{T}(r-r') = \nabla_r T(r-r') = \nabla_r \nabla_r T(r-r')$ is the dipole–dipole interaction tensor. The integration can be carried out over the whole volume occupied by the surrounding medium while avoiding the singularity by an appropriate cutoff; in practical schemes, this is never a real problem [7]. Equation (16) is iteratively solved and, by introducing the surrounding medium response tensor $\mathbf{C}(r, r')$ defined by:

$$\mathbf{C}(r,r') = \mathbf{X}(r)\cdot\left[\mathbf{1}\,\delta(r-r') + \mathbf{T}(r-r')\cdot\mathbf{X}(r')\right.$$
$$\left. + \int dr''\mathbf{T}(r-r'')\cdot\mathbf{X}(r'')\cdot\mathbf{T}(r''-r')\cdot\mathbf{X}(r') + \ldots\right], \tag{17}$$

where $\mathbf{1}$ is the unit tensor, the polarization density is cast into a compact form resembling the one used in the electrodynamics with spatial dispersion [7]:

$$p(r) = \int dr'\mathbf{C}(r,r')\cdot e(r'). \tag{18}$$

This equation emphasizes the non-local response of the medium towards the solute (external) field.

The reaction field contribution to the interaction Hamiltonian is given by:

$$V_{\mathrm{RF}} = \int dr\,\Omega_s(r)\int dr'\int dr''\int dr'''T(r-r'')\cdot\mathbf{C}(r''-r''')\cdot T(r'''-r')\rho_s(r'); \tag{19}$$

the integration variables r and r' are points associated only with the solute system, while integration over r'' and r''' is over points associated with the solvent system only. Thus, defining an RF kernel function,

$$G(r,r') = \int dr''\int dr'''T(r-r'')\cdot\mathbf{C}(r''-r''')\cdot T(r'''-r'), \tag{20}$$

the effective Hamiltonian of eq. (13) is now cast into its final form by using eqs. (16) to (20):

$$\mathcal{H}_s = H_s + \int dr\,\Omega_s(r)\left[V_m^0(r) + \int dr' G(r,r')\rho_s(r')\right]. \tag{21}$$

Note that this Hamiltonian depends upon the solute charge density.

The solution of eq. (11) with the Hamiltonian eq. (21) has to be made self-consistent. This condition is independent of the quantum chemical approach used for constructing the wave function.

Once the solution has been found, at the end of the self-consistent calculation, the total energy of the solute W_s is obtained by adding to $E_s(X)$ the energy required to polarize the surrounding medium:

$$W_s = E_s(X) - (1/2)\left\langle\Psi_s\left|\int dr\,\Omega_s(r)\int dr' G(r,r')\rho_s(r')\right|\Psi_s\right\rangle. \tag{22}$$

The general form of the RF attained in the present theory, namely,

$$\Pi(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}' G(\boldsymbol{r},\boldsymbol{r}')\rho(\boldsymbol{r}'), \tag{23}$$

shows an explicit functionality on the solute charge density. However, eq. (14) is better adapted for algorithmic developments.

The theory described thus far [16,25] is well suited for designing computation schemes coupled to MC or MD simulations of the surrounding medium. MC and MD provide information on the global configuration $\boldsymbol{X}$. Once $\boldsymbol{X}$ is known, $V_m^0(\boldsymbol{r})$ and C can be evaluated provided a solvent charge density and external field susceptibility are given. These problems have already been solved in the framework of the inhomogeneous SCRF theory [7,25]. At the quantum chemical level, matrix elements over one-electron operators are required. However, there is more in this approach, since different types of theories can be obtained for varied choices of C-tensors. This is the point where a microscopic approach to the surrounding medium can be used. Another possibility is open to the practitioners if the dielectric theory is chosen. The latter procedure allows for derivations of the continuum models as particular cases of a generalized dielectric theory.

2.2. STATISTICALLY SIGNIFICANT GSCRF THEORY

The connection with standard solvent effect theories must be made after a statistical mechanical averaging over the solvent configurations. Temperature enters here in a natural manner.

Since the solute–solvent interaction Hamiltonian does not depend upon particle momenta, the statistical averaging is carried out over the configurational space; it will be designated by angular brackets $\langle \ldots \rangle$. A fixed molecular reference frame is assigned to the solute by using, for instance, its nuclear equilibrium configuration $\boldsymbol{R}_s$. Following Kirkwood, an averaging over the entire solvent configuration is performed while the solute is kept at a fixed orientation defined by the Euler angles with respect to a fixed molecular frame [6]. This averaging is designated by the angular brackets with a subscript m to indicate this fact.

2.2.1. Polarization density

Let us consider first the statistically averaged polarization density:

$$\langle \boldsymbol{p}(\boldsymbol{r};\boldsymbol{X})\rangle_m = \int \mathrm{d}\boldsymbol{r}' \langle \mathrm{C}(\boldsymbol{r},\boldsymbol{r}';\boldsymbol{X}) \cdot \boldsymbol{e}(\boldsymbol{r}';\boldsymbol{X})\rangle_m; \tag{24}$$

the integral over spatial coordinate $\boldsymbol{r}'$ can be rigorously taken out of the averaging integration. Introduce the statistically averaged response tensor $\langle \mathrm{C}(\boldsymbol{r}, \boldsymbol{r}';\boldsymbol{X})\rangle = \langle \mathrm{C}\rangle_m$ and electric field $\langle \boldsymbol{e}(\boldsymbol{r}';\boldsymbol{X})\rangle_m = \langle \boldsymbol{e}\rangle_m$. One calculates the fluctuation correlation vector function:

$$\langle \delta \mathbf{C} \cdot \delta e \rangle_m = \langle (\mathbf{C}(r,r';X) - \langle \mathbf{C} \rangle_m) \cdot (e(r';X) - \langle e \rangle_m) \rangle_m. \tag{25}$$

At thermodynamic equilibrium (in Feynman's sense, namely, all fast things have happened while the slow ones have not [26]), $\langle \delta \mathbf{C} \cdot \delta e \rangle_m$ fades away and we have:

$$\langle (\mathbf{C}(r,r';X) \cdot e(r';X) \rangle_m = \langle \mathbf{C} \rangle_m \cdot \langle e \rangle_m, \tag{26}$$

and eq. (24) can be uncoupled:

$$\langle p(r;X) \rangle_m = \int \mathrm{d}r' \langle \mathbf{C}(r,r';X) \rangle_m \cdot \langle e(r';X) \rangle_m. \tag{27}$$

This relationship indicates that, at thermal equilibrium, the surrounding medium will react to the averaged solute electric field. Note that if we average over the orientation of the solute, the r.h.s. of (24) cancels out. The theory described in this subsection is of the mean field type.

The statistically averaged reaction field susceptibility $\langle G(r, r') \rangle_m$ can also be uncoupled following a reasoning similar to the one above. Thus,

$$\langle G(r,r') \rangle_m = \int \mathrm{d}r'' \int \mathrm{d}r''' T(r-r'') \cdot \langle C(r'',r''') \rangle_m \cdot T(r'''-r'), \tag{28}$$

and the RF becomes a response of a statistically averaged susceptibility tensor $\langle \mathbf{C}(r'', r''') \rangle_m$. This enitity is temperature and pressure dependent.

2.2.2. *Solvent potential*

The averaged solvent electrostatic field $\langle V_m^0(r) \rangle_m$ is important for describing inhomogeneous media, such as enzymes, membranes, miscelles and perturbed crystalline environments. Due to the existence of strong correlations, such a field may not cancel out after averaging. This factor becomes an important contribution to solvent effects at a microscopic level. In a study of non-rigid molecules in solution, Sese et al. [27] constructed a $\langle V_m^0(r) \rangle_m$ by using the solute–solvent atom–atom radial distribution function. Electrostatic interactions in three-dimensional solids were treated by Ángyán and Silvi [28] in their self-consistent Madelung potential approach; such a procedure can be traced back to a calculation of $\langle V_m^0(r) \rangle_m$. In an earlier application of the ISCRF theory to the study of proton mechanisms in crystals of hydronium perchlorate, both $\langle V_m^0(r) \rangle_m = \langle V_m^0(r; \langle X \rangle)$ and the RF potential [29] were computed. Another example was provided by the calculation of the electric field produced at the active site of alcohol dehydrogenase by the protein atoms surrounding it [25].

2.2.3. *Averaged GSCRF equation*

In the mean field approach, a statistically significant effective Hamiltonian is obtained by m-averaging ($\langle \ldots \rangle_m$) [6] eq. (21):

$$\mathcal{H}_s = \left\{ H_s + \int \mathrm{d}r\, \Omega_s(r)[\langle V_m^0(r)\rangle_m + \int \mathrm{d}r''\langle G(r,r')\rangle_m \langle \rho_s(r')\rangle_m] \right\}. \tag{29}$$

The key step here for constructing a simple SCRF Hamiltonian is to replace the m-averaged solute charge density ($\langle \rho_s(r')\rangle_m$) by the expectation value of the solute charge density operator with respect to a new wave function Φ_s, i.e. $\langle \rho_s(r')\rangle_m = \langle \Phi_s | \Omega_s(r') | \Phi_s \rangle$, this wave function is now an object replacing an ensemble, and it is obtained as a solution of the effective Schrödinger equation with Hamiltonian eq. (30):

$$\mathcal{H}_s = \left\{ H_s + \int \mathrm{d}r\, \Omega_s(r)[\langle V_m^0(r)\rangle_m + \int \mathrm{d}r''\langle G(r,r')\rangle_m \langle \Phi_s | \Omega_s(r') | \Phi_s \rangle] \right\}. \tag{30}$$

This effective Hamiltonian represents a solute system in thermal equilibrium with respect to all the degrees of freedom except its orientation. This assumption is congruent to the mean field approach. Under this hypothesis, the effective Schrödinger equation has the same form as in the generalized SCRF theory presented in the preceding section. The interpretation is fundamentally different in both cases.

An explicit formulation of Hamiltonian (30) is obtained if $\langle G(r, r')\rangle_m$ is replaced by eq. (28):

$$\mathcal{H}_s = \left\{ H_s + \int \mathrm{d}r\, \Omega_s(r)[\langle V_m^0(r;X)\rangle_m + \int \mathrm{d}r' \int \mathrm{d}r'' \int \mathrm{d}r'''\, T(r-r'') \right.$$
$$\left. \cdot \langle \mathbf{C}(r'',r''';X)\rangle_m \cdot T(r'''-r') \langle \Phi_s | \Omega_s(r') | \Phi_s \rangle] \right\}, \tag{31}$$

where the dependence of both the charge density and the averaged response tensor is explicitly shown. This is the fundamental Hamiltonian of the statistically significant GSCRF approach. The effecive Schrödinger equation:

$$\mathcal{H}_s | \Phi_s \rangle = E_s(R_s;T) | \Phi_s \rangle \tag{32}$$

depends upon the surrounding medium temperature T.

The generalized SCRF theory can now be used in several different ways. One makes use of the knowledge provided by MD and/or MC statistical mechanical simulation procedures on the statistical distribution of X. Thus, for systems having a well-defined average structure, as for instance a native enzyme, $\langle \mathbf{C}(r'', r'''; X)\rangle_m = \mathbf{C}(r'', r'''; \langle X\rangle_m)$. It is also possible to replace the theory into the framework of the electrodynamics of dielectric materials. In this case, $\langle \mathbf{C}(r'', r'''; X)\rangle_m$ is represented with the help of the static permittivity tensor. In the next section, the microscopic approach is examined first and thereafter the dielectric approach.

2.3. MICROSCOPIC APPROACH

In the discrete microscopic model, the theory is completed with a recipe to calculate the surrounding medium response to an external field, i.e. C and the

permanent electrostatic potential $V_m^0(r)$, which is different from zero for an arbitrary X-configuration.

The solute electric field at a given instant of time, at a fixed geometry, is coupled via the static polarizability to the surrounding medium. Thus, if the environment of the solute is represented as a set of discrete atomic and group polarizabilities $\{\alpha_{mi}\}$, the tensor $\mathbf{X}(r)$ can be written as [7,16,17,23,29]:

$$\mathbf{X}(r) = \sum_{mi} \alpha_{mi} \delta(r - R_{mi}), \tag{33}$$

where the summation is carried over the atoms, molecules or functional groups in the molecular environment. For a solute surrounded by a sample of solvent molecules, e.g. water, configurations X generated in Monte Carlo or molecular dynamics simulations can be used to calculate explicitly eq. (18) for the polarization density p at each point in the sample. Taking supervector notation, $\mathbf{p}$ is in this case a vector of dimension equal to the number of points (m) representing the solvent system: $\mathbf{p}^{\mathrm{T}} = (p_1, p_2, p_3, \ldots, p_m)$. $\mathbf{p}^{\mathrm{T}}$ is the transpose of the column vector $\mathbf{p}$. The tensor $\mathbf{C}$ becomes a supermatrix: $\mathbf{C}(r_i, r_j')$, where i and j run from 1 up to m. Each matrix element is a second-order tensor given by:

$$\mathbf{C}(r_i, r_j') = \boldsymbol{\alpha}_i \cdot [\delta_{ij} + \mathbf{T}(r_i - r_j') \cdot \boldsymbol{\alpha}_j + \sum_k \mathbf{T}(r_i - r_k) \cdot \boldsymbol{\alpha}_k \cdot \mathbf{T}(r_k - r_j) \cdot \boldsymbol{\alpha}_j + \ldots]. \tag{34}$$

The terms in square brackets can be formally summed up since they form a geometric progression:

$$p_i = \sum_j \mathbf{C}(r_i, r_j) \cdot e(r_j) = \boldsymbol{\alpha}_i \cdot \sum_j [\delta_{ij} - \mathbf{T}(r_i - r_j) \cdot \boldsymbol{\alpha}_j]^{-1} \cdot e(r_j), \tag{35}$$

or, using a supermatrix notation, one can rewrite eq. (35) as follows:

$$\mathbf{p} = \boldsymbol{\alpha} \cdot [\mathbf{1} - \mathbf{T} \cdot \boldsymbol{\alpha}]^{-1} \cdot \mathbf{e} = [\mathbf{1} - \boldsymbol{\alpha} \cdot \mathbf{T}]^{-1} \cdot \boldsymbol{\alpha} \cdot \mathbf{e}, \tag{36}$$

where $\mathbf{e}^{\mathrm{T}} = (e_1, e_2, e_3, \ldots, e_m) = (e(r_1), e(r_2), e(r_3), \ldots, e(r_m))$; the nonzero elements of the supermatrix $\boldsymbol{\alpha}$ are the diagonal group matrices $\boldsymbol{\alpha}_i$, the supermatrix $\mathbf{T}$ has the diagonal elements equal to zero.

Equation (36) is a well-known formula. This equation can be implemented to obtain the induced dipole at each point in the surrounding medium as it is produced by an electric field whose source is the solute system. The many-body polarizability supermatrix $\mathbf{A}$ is defined as

$$\mathbf{A} = [\mathbf{1} - \boldsymbol{\alpha} \cdot \mathbf{T}]^{-1} \cdot \boldsymbol{\alpha}, \tag{37}$$

and the polarization density is again given a form similar to eq. (18), namely,

$$\mathbf{p} = [\mathbf{1} - \boldsymbol{\alpha} \cdot \mathbf{T}]^{-1} \cdot \boldsymbol{\alpha} \cdot \mathbf{e} = \mathbf{A} \cdot \mathbf{e}. \tag{38}$$

Matrix inversion is a cumbersome procedure, in particular for matrices having a very large dimension. To avoid this problem, an iterative procedure can be devised to calculate the many-body polarizability supermatrix **A** which follows from the self-consistency relationship applied to eq. (37), namely,

$$\mathbf{A} = \boldsymbol{\alpha} - \boldsymbol{\alpha} \cdot \mathbf{T} \cdot [\mathbf{1} - \boldsymbol{\alpha} \cdot \mathbf{T}]^{-1} \cdot \boldsymbol{\alpha} = \boldsymbol{\alpha} + \boldsymbol{\alpha} \cdot \mathbf{T} \cdot \mathbf{A}. \tag{39}$$

Defining now $\mathbf{A}^{(0)} = \mathbf{O}$, one obtains:

$$\mathbf{A}^{(n)} = \boldsymbol{\alpha} + \boldsymbol{\alpha} \cdot \mathbf{T} \cdot \mathbf{A}^{(n-1)}, \tag{40}$$

from which one can obtain the second identity in eq. (38) as an iterative scheme:

$$\mathbf{p}^{(n)} = (\boldsymbol{\alpha} + \boldsymbol{\alpha} \cdot \mathbf{T} \cdot \mathbf{A}^{(n-1)}) \cdot \mathbf{e}, \tag{41}$$

and **e** is the field external to the polarizable domain. These equations are solved iteratively to obtain the induced dipole at each point in the surrounding medium as it is produced by an electric field whose source is the solute system. For the nth iteration, one has:

$$\mathbf{p}^{(n)} = \boldsymbol{\alpha} \cdot (\mathbf{e} + \mathbf{T}\, \mathbf{p}^{(n-1)}), \tag{42}$$

with $\mathbf{p}(0) = \mathbf{O}$. In this case, there is no need to invert large matrices and the procedure is computationally more efficient. Convergency is usually attained in three to five iterations.

Note that in the quantum mechanical treatment, **e** is self-consistent in the sense that it is calculated with a self-consistent wave function. As emphasized above, this field is external to the polarizable domain. For a classical set of fixed charges and induced dipoles, the components of **e** are electric fields whose sources are the fixed charges, each one of them being external to the remaining polarizable domain; it is the supermatrix **A** which takes care of the full induction.

2.4. GSCRF THEORY IN THE DIELECTRIC APPROACH

In the electrodynamics of dielectrics, an electric induction vector $\boldsymbol{D}(r)$ and a dielectric polarization vector $\boldsymbol{P}(r)$ are introduced for describing the fields acting in the medium [17]. Correspondences between these objects and those coming from the GSCRF theory are established below.

The averaged polarization density $\langle \boldsymbol{p}(r) \rangle_m$ corresponds, by construction, to $\boldsymbol{P}(r)$. The assignment of $\langle \boldsymbol{e}(r) \rangle_m$ is more delicate. In standard electrodynamics, the external sources are not perturbed by the dielectric medium. If such were the case here, there would be no effective Schrödinger equation including reaction fields; first-order perturbation theory would be enough to calculate interaction energies between the solute and the medium. Since $\langle \boldsymbol{e}(r) \rangle_m$ results from reciprocal inductive

effects, the polarization effects are built into the "external" charge density. Thus, it is more natural to assign this field to the electric induction $\boldsymbol{D}(\boldsymbol{r})$ when the self-consistent wave function is used to calculate it. Thus, at self-consistency, eq. (16) becomes a relationship between the polarization density and the electric induction:

$$\boldsymbol{P}(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}' \langle \mathbf{C}(\boldsymbol{r}, \boldsymbol{r}') \rangle_m \cdot \boldsymbol{D}(\boldsymbol{r}'). \tag{43}$$

The statistical mechanical averaging of $\mathbf{C}$ introduces correlations related to all molecular motions in the solvent that are allowed to relax until attaining thermal equilibrium. It is natural to seek a function of the static permittivity tensor $\boldsymbol{\varepsilon}(\boldsymbol{r}, \boldsymbol{r}')$ to represent the averaged response function. In the absence of polarizable matter in the surroundings, the averaged tensor $\langle \mathbf{C}(\boldsymbol{r}, \boldsymbol{r}') \rangle_m = \mathbf{1}\, \delta(\boldsymbol{r}, \boldsymbol{r}')$, the presence of a dielectric screens the interactions with a functional dependency $\boldsymbol{\varepsilon}^{-1}(\boldsymbol{r}, \boldsymbol{r}')$ so the averaged susceptibility tensor is:

$$\langle \mathbf{C}(\boldsymbol{r}, \boldsymbol{r}') \rangle_m = \mathbf{1}\, \delta(\boldsymbol{r}, \boldsymbol{r}') - \boldsymbol{\varepsilon}^{-1}(\boldsymbol{r}, \boldsymbol{r}'), \tag{44}$$

where $\boldsymbol{\varepsilon}^{-1}(\boldsymbol{r}, \boldsymbol{r}')$ is the inverse of the static permittivity tensor, the statistical mechanical averaged polarization density now satisfies standard electrostatic equations for media having spatial dispersion, namely,

$$\boldsymbol{P}(\boldsymbol{r}) = \boldsymbol{D}(\boldsymbol{r}) - \int \mathrm{d}\boldsymbol{r}' \boldsymbol{\varepsilon}^{-1}(\boldsymbol{r}, \boldsymbol{r}') \cdot \boldsymbol{D}(\boldsymbol{r}'). \tag{45}$$

Note that the last term corresponds to the electric field of the sources as if it were in vacuo: $\boldsymbol{E}(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}' \boldsymbol{\varepsilon}^{-1}(\boldsymbol{r}, \boldsymbol{r}') \cdot \boldsymbol{D}(\boldsymbol{r}')$. Thus, the standard electrostatic relationship between the displacement and external field is recovered: $\boldsymbol{D}(\boldsymbol{r}) = \boldsymbol{E}(\boldsymbol{r}) + \boldsymbol{P}(\boldsymbol{r})$. The internal consistency of the present theory is therefore ensured.

The generalized self-consistent reaction field equation in the dielectric approach now reads:

$$\left\{ H_s + \int \mathrm{d}\boldsymbol{r}\, \Omega_s(\boldsymbol{r}) \left[V_m^0(\boldsymbol{r}) \right\rangle_m + \int \mathrm{d}\boldsymbol{r}' \boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}') \cdot \int \mathrm{d}\boldsymbol{r}'' [\mathbf{1}\, \delta(\boldsymbol{r}' - \boldsymbol{r}'') - \boldsymbol{\varepsilon}^{-1}(\boldsymbol{r}', \boldsymbol{r}'')]\, \boldsymbol{D}(\boldsymbol{r}'') \right] \right\} |\Phi_s\rangle$$

$$= U_s(\boldsymbol{R}_s) |\Phi_s\rangle. \tag{46}$$

For homogeneous environments only, $\langle V_m^0(\boldsymbol{r}) \rangle_m$ averages out to zero. For liquids, if a very structured solvation shell remains around the solute, it is convenient to introduce this region (cybotactic region) into the definition of the solute and treat the surroundings as homogeneous.

The effective Schrödinger eq. (46) has the most general form among those proposed in the literature to represent solvent effects in the dielectric approach. As a matter of fact, the effective Hamiltonian eq. (46) contains those as special cases.

2.4.1. Continuum approach

In the continuum approach to the surrounding medium, one has to set up by definition $\langle V_m^0(r)\rangle_m$ equal to zero. Medium effects are therefore represented by a reaction field term in eq. (46). Three types of environment can be represented in this framework: (i) an anisotropic medium without spatial dispersion, where the permittivity tensor is defined with the ansatz: $\varepsilon(r-r') = \varepsilon(r)\,\delta(r-r')$; (ii) an isotropic medium which is characterized by $\varepsilon(r-r') = \varepsilon(r)\,\mathbf{1}\,\delta(r-r')$; (iii) a homogeneous and isotropic medium, the permittivity tensor is the unit tensor multiplied by the static dielectric constant ε. All these models lead to a distance-dependent dielectric representation. Now, the effective Schrödinger equation for each case is obtained from eq. (46) after integration over the r'-variable with the corresponding ansatz for the permittivity tensor.

2.4.2. Cavity models

The cavity immersed in a continuum dielectric continues to attract attention from quantum chemists trying to incorporate solvent effects in their in vacuo calculations [30–44]. Cavity models in a continuum dielectric follow in a simple manner by directly modelling $\langle \mathbf{C}(r, r')\rangle_m$.

Let r^c be the cavity radius and V' the sample volume where the spherical cavity containing the solute has been cut out. For a homogeneous isotropic dielectric in eq. (23), $\langle \mathbf{C}(r'', r''')\rangle_m$ can be written as: $(1-\varepsilon^{-1})\Theta(r''-r^c)\Theta(r'''-r^c)\,\delta(r''-r''')$, where $\Theta(r''-r^c)$ is the Heasivide step function, i.e. it is zero inside the cavity and one outside and at its boundary. The RF potential can then be cast into:

$$(1-\varepsilon^{-1})\int \mathrm{d}r''\, T(r-r'')\cdot\Theta(r''-r^c)\int \mathrm{d}r'''\,\Theta(r'''-r^c)\,\delta(r''-r''')\cdot D(r''').$$

The inner integral can be carried out straightforwardly as r'' and r''' run outside the cavity, thus,

$$\Pi(r) = (1-\varepsilon^{-1})\int \mathrm{d}r''\,\Theta(r''-r^c)\,T(r-r'')\cdot D(r''), \tag{47}$$

and introducing the quantum mechanical definition of the solute field, one finally obtains:

$$\Pi(r) = (1-\varepsilon^{-1})\int \mathrm{d}r'\Big[\int \mathrm{d}r''\,\Theta(r''-r^c)\,T(r-r'')\cdot T(r''-r')\Big]\rho_s(r'). \tag{48}$$

This equation contains the basic model used to study the solvated electron (see ref. [7] for a discussion). If the charge density is replaced by a classical unit charge at the origin of the sphere, the RF potential obtained after integration of eq. (48) corresponds to the Born model for a metallized sphere immersed in an isotropic continuum.

The model using spherical harmonics expansions for the RF potential can be derived from eq. (48) by introducing spherical boundary conditions. The procedure has already been outlined by the present author [6] and will not be repeated here.

2.4.3. Semi-continuum models

In this type of approach, the first solvation shell is represented in the supermolecule and, consequently, enters into the quantum chemical description. Basically, the radius of the sphere embedded in the continuum dielectric is much larger than for the desolvated solute. This model has been used on several ocassions, e.g. solvated electron, electron transfer in solution and in the study of solvent effects on the transition structure for the hydride transfer step in the catalytic mechanism of liver alcohol dehydrogenase [39].

Recently, two developments have taken place in this field: one concerns the quantum chemical level used to solve open shell electronic structures [40], the other is an extension of the homogeneous model to treat an anisotropic surrounding medium [41].

The level of electronic structure theory used by Mikkelsen et al. [40] corresponds to the multiconfigurational (MC) self-consistent field (SCF) where the wave function is fully optimized with respect to all variational parameters; these include orbital and configurational parameters. The main deficiency of standard SCF ab initio procedures, namely, lack of intramolecular correlation effects, is overcome in this MCSCF approach. The level of solvent effects theory is the standard spherical cavity immersed in a continuum dielectric; an early formalism proposed by Rinaldi and Rivail [31] was used (see ref. [6] for a more extensive analysis).

The contribution by Hoshi et al. addresses the formulation of a theory for the estimation of a molecular electronic structure surrounded by an anisotropic medium [41]. It is assumed there that the medium surrounding the solute system is composed of more than two polarizable dielectrics with different dielectric constants; the different dielectric regions make contact with each other through arbitrarily shaped boundaries. As usual, the solute charge distribution interacts with the dielectrics via a reaction field.

In a different vein, Claverie and coworkers [45] have made a significant effort to improve the continuum model used for calculation of solvation thermodynamics quantities of a molecule embedded in a cavity formed by the intersecting Van der Waals spheres of the solute in a polarizable medium.

At a more elementary level, Gersten and Sapse have investigated solvent effects through the use of an embedded Born equation [48]. Rashin and Namboodiri [47] have also contributed with a simple method for the calculation of hydration enthalpies of polar molecules with arbitrary shapes.

An analysis of discrete and continuum dielectric models as they have been applied to calculate protonation energies in solution has been presented by Rullman and Van Duijnen [48]. The model proposed by these authors combines a discrete

molecular description of the first two or three solvation layers with a continuum description of the bulk solvent. The solute is described quantum chemically at an HF level. The accuracy of the results depends on the dielectric model as well as on the details of the electrostatic potential and on inductive interactions.

The simple virtual charge model discussed by Constanciel and Tapia [6,49,50] has been developed into an extended generalized Born (EGB) approach. Different approximations have been proposed. Constanciel [43] has analyzed the theoretical basis used as a foundation for empirical reaction field approximation through the continuum model to the surrounding medium. Artifacts in the EGB scheme have been clearly identified. The new approximate formulation derives from an exact integral equation of classical electrostatics following a well-defined procedure. It is shown there how to compute the wave function of solvated species embedded in cavities formed by interlocking spheres in a polarizable continuum.

2.4.4. *Generalized cavity models*

Equation (48) can be generalized for a cavity of arbitrary shape. Let the former sphere radius become a parametric function of s describing the surface embodying the system of interest: $r^c(s)$, including this function in the form (48) and interpreting the tetha function in a more general manner, the reaction field can be written as:

$$\Pi(r) = (1-\varepsilon^{-1}) \int \mathrm{d}r' \Big[\int \mathrm{d}r'' \Theta(r''-r^c(s)) T(r-r'') \cdot T(r''-r') \Big] \rho_s(r'). \tag{49}$$

If the system of interest occupies the volume say υ, the tetha function changes the total volume of integration for the variable $\mathrm{d}r''$ to be $V' = V - \upsilon$. We then integrate over $\mathrm{d}r''$ to obtain:

$$\int_{V'} \mathrm{d}r'' T(r-r'') \cdot T(r''-r') = \int \mathrm{d}r'' \operatorname{div}(T(r-r'')T(r''-r')) - \int_{V'} \mathrm{d}r'' \operatorname{div} T(r''-r'). \tag{50}$$

The divergence of $T(r''-r')$ in the volume V' is actually zero in this model, and according to Gauss' theorem, the first integral is equal to the integral of $T(r-r'')T(r''-r')$ over the surface bounding the volume of integration S'. Since the field is zero at infinity, this integral reduces to a surface integral over precisely the surface boundary u, i.e. S_υ, and with the normal unit vector to the surface at r'' designated by $n(r'')$, eq. (49) can now be written as:

$$\Pi(r) = (1-\varepsilon^{-1}) \int \mathrm{d}r' \left[\int_{S_\upsilon} \mathrm{d}s'' n(r'') \cdot T(r''-r') T(r-r'') \right] \rho_s(r'). \tag{51}$$

The electric field set up by the solute charge density can be written directly as:

$$e(r'') = \left[\int dr' T(r'' - r') \right] \rho_s(r'). \tag{52}$$

The reaction field potential can be cast in terms of the electric field created by the solute at the surface delimiting its volume:

$$\Pi(r) = (1 - \varepsilon^{-1}) \int_{S_v} ds'' n(r'') \cdot e(s'') T(r - r''), \tag{53}$$

which is the Scrocco–Tomasi–Miertus theory of solvent effects [51,52]. Tomasi's group have extensively used this blend of theory to treat a number of interesting systems [52]. In particular, the GEPOL scheme developed for calculating the surface and volume of any molecule [53] has been adapted to actually calculate the reaction field potential of eq. (53).

2.5. MUTUALLY CONSISTENT FIELD THEORIES

In the GSCRF theory discussed above (cf. section 2.1), only the solute is described by an effective Schrödinger equation, namely, eq. (11). Instead of taking a classical electrodynamics representation for the solvent charge density as we did above, one can treat both subsystems at a quantum mechanical level. Thus, one has for each of them the following Hamiltonians:

$$\begin{aligned} \mathcal{H}_s(r_s; X) &= H_s(r_s, R_s) + \int dr \int dr' \, \Omega_s(r) \langle \Psi_m | T(r - r') \Omega_m(r') | \Psi_m \rangle \\ &= H_s(r_s, R_s) + J^m \end{aligned} \tag{54}$$

and

$$\begin{aligned} \mathcal{H}_m(r_m; X) &= H_m(r_m, R_m) + \int dr \int dr' \, \Omega_m(r) \langle \Psi_s | T(r - r') \Omega_s(r') | \Psi_s \rangle \\ &= H_m(r_m, R_m) + J^s, \end{aligned} \tag{55}$$

where J^s and J^m are the Coulomb operators used by Ángyán and Náray-Szabo [54] in their mutually consistent field theory.

One must now solve the coupled effective Schrödinger equations:

$$\mathcal{H}_s | \Psi_{s\sigma} \rangle = E_{s\sigma}(X) | \Psi_{s\sigma} \rangle \tag{56}$$

and

$$\mathcal{H}_m | \Psi_{m\mu} \rangle = E_{m\mu}(X) | \Psi_{m\mu} \rangle, \tag{57}$$

where σ and μ are quantum numbers characterizing the solute and solvent eigenstates.

Consider the hypothetical situation where we separate out both subsystems while still keeping the global nuclear configuration X. One can write down a Schrödinger equation for each subsystem:

$$H_s|\Phi_{s\sigma}\rangle = E_{s\sigma}(X_s)|\Phi_{s\sigma}\rangle \tag{58}$$

and

$$H_m|\Phi_{m\mu}\rangle = E_{m\mu}(X_m)|\Phi_{m\mu}\rangle. \tag{59}$$

These sets of eigenfunctions are now used to expand the interaction operators of the mutually interacting system described by eqs. (56) and (57). Focusing on the solute system again, and to second-order perturbation level, one has for $|\Psi_{m\mu}\rangle$ the following expansion [6,14,54]:

$$|\Psi_{m\mu}\rangle = |\Phi_{m\mu}\rangle + \sum_{\mu'\neq\mu} |\Phi_{m\mu'}\rangle\left[E_{m\mu} - E_{m\mu'}\right]^{-1}\langle\Phi_{m\mu'}|J^s|\Phi_{m\mu}\rangle, \tag{60}$$

which can now be introduced in J^m to obtain an explicit form for this operator in terms of the solvent unperturbed electronic properties:

$$\begin{aligned} J^m = {} & \int dr \int dr' \, \Omega_s(r)\langle\Phi_{m\mu}|T(r-r')\Omega_m(r')|\Phi_{m\mu}\rangle \\ & + \sum_{\mu'\neq\mu}\langle\Phi_{m\mu}|T(r-r')\Omega_m(r')|\Phi_{m\mu'}\rangle\left[E_{m\mu} - E_{m\mu'}\right]^{-1}\langle\Phi_{m\mu'}|J^s|\Phi_{m\mu}\rangle \\ & + \sum_{\mu'\neq\mu}\langle\Phi_{m\mu}|J^s|\Phi_{m\mu'}\rangle\left[E_{m\mu} - E_{m\mu'}\right]^{-1}\langle\Phi_{m\mu'}|\Omega_m(r')T(r-r')|\Phi_{m\mu}\rangle \\ & + \text{higher-order terms}, \end{aligned} \tag{61}$$

introducing now the definition of J^s, neglecting higher-order terms and using Ángyán's charge density response function, namely,

$$\begin{aligned} C^m(r,r') = {} & \sum_{\mu'\neq\mu}\langle\Phi_{m\mu}|\Omega_m(r)|\Phi_{m\mu'}\rangle\left[E_{m\mu} - E_{m\mu'}\right]^{-1}\langle\Phi_{m\mu'}|\Omega_m(r')|\Phi_{m\mu}\rangle \\ & + \sum_{\mu'\neq\mu}\langle\Phi_{m\mu}|\Omega_m(r')|\Phi_{m\mu'}\rangle\left[E_{m\mu} - E_{m\mu'}\right]^{-1}\langle\Phi_{m\mu'}|\Omega_m(r)|\Phi_{m\mu}\rangle, \end{aligned} \tag{62}$$

the interaction operator set up by the surrounding medium acquires the form:

$$\begin{aligned} J^m = {} & \int dr \int dr' \, \Omega_s(r)\langle\Phi_{m\mu}|T(r-r')\Omega_m(r')|\Phi_{m\mu}\rangle \\ & + \int dr \int dr' \int dr'' \int dr''' \Omega_s(r)T(r-r'')C^m(r''-r''')T(r'''-r')\rho_s(r'); \end{aligned} \tag{63}$$

the second line in this equation can be recast in terms of the reaction field response function $G^m(r, r') = \int dr'' \int dr''' T(r-r'')C^m(r''-r''')T(r'''-r')$, which is the analog of our reaction field response function $G(r, r')$, eq. (20). Actually, if the Dirac delta functionals appearing in the definition of the charge density operator $\Omega_m(r')$

(see eq. (2) which is the counterpart for the solute) is developed around $\boldsymbol{R}$, e.g. $\boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{R}$ [14]:

$$\delta(\boldsymbol{r}+\boldsymbol{R}) = \delta(\boldsymbol{R}) + \boldsymbol{r}\cdot\nabla\delta(\boldsymbol{R}) + \dots , \tag{64}$$

using adequate expansions in (62) one finally obtains eq. (19), thereby showing the intimate relationships between both schemes.

On the other hand, Weinstein and coworkers [55] start from the fully antisymmetrized product of the total wave function

$$\Psi(r_s, r_m; X) = A_{sm}\Psi_s(r_s; R_s, R_m)\Psi_m(r_m; R_m, R_s)$$

and consider the solute and solvent wave function as individually antisymmetrized group functions. The equation to be solved is then the full Schrödinger equation:

$$H\,|\,\Psi(r_s, r_m; X)\rangle = E\,|\,\Psi(r_s, r_m; X)\rangle. \tag{65}$$

Then the total energy of the system is written in terms of the self-energies of the component groups in the complex, together with an interaction term:

$$E = \langle\Psi_s(r_s; X)\,|\,H_s(r_s, R_s)\,|\,\Psi_s(r_s; X)\rangle + \langle\Psi_m(r_m; X)\,|\,H_m(r_m, R_m)\,|\,\Psi_m(r_m; X)\rangle$$
$$+ \langle\Psi(r_s, r_m; X)\,|V(r_s, r_m; X)\,|\,\Psi(r_s, r_m; X)\rangle. \tag{66}$$

The interaction between the quantum and its surroundings, e.g. protein environment, is calculated with an iterative computational scheme developed on the basis of McWeeny's variation-perturbation theory of group functions. The iterative procedure neglects exchange interactions between the two subsystems; furthermore, the perturbation is represented by linear terms in the Hamiltonian, unlike the formulation presented by us in the GSCRF theory.

2.6. PROTEIN'S ACTIVE SITE MODEL

Chemical events in enzyme catalyzed reactions usually take place in small volumes (active site) compared to the full extent of the biosystem. Main and side chain functional groups, e.g. imidazole ring of histidines, together with other molecules provide the material basis for the chemical events to take place. To construct molecular models for the active site components, bonds are broken in order to keep the size of the quantum system within reasonable bounds. The dangling bonds are usually saturated with hydrogen atoms. This procedure is not fully satisfactory when a C–C bond is replaced by a C–H bond. Actually, embedding a model molecular system in its enzyme environment is not yet a solved problem. Depending on the level of theory, i.e. a semiempirical or an ab initio MO scheme, the embedding problem may admit different solutions.

The theory and procedure discussed until now apply to fixed nuclear coordinates. In the following section, we discuss methods used for treating nuclear movements.

3. Classical statistical mechanical scheme

The dynamics in the X-space, namely, the configurational space, is normally assumed to be driven at the classical mechanical level. Once the quantum mechanics of electronic motion has been solved, the nuclei are submitted to intra- and intermolecular potentials. The total Hamiltonian H is written as a sum of molecular Hamiltonians (H_m and H_s) and the intermolecular interactions V_{ms}:

$$H(\boldsymbol{X},\boldsymbol{P}) = H_m(\boldsymbol{P}_m,\boldsymbol{R}_m) + H_s(\boldsymbol{P}_s,\boldsymbol{R}_s) + V_{ms}(\boldsymbol{R}_m,\boldsymbol{R}_s), \tag{67}$$

where $\boldsymbol{P}$ is the linear momentum in Cartesian coordinates of the ensemble of particles; $\boldsymbol{P}_m$ and $\boldsymbol{P}_s$ identify the momenta associated to each subsystem. The solute subsystem, in the classical statistical mechanical treatment, may contain a number of molecules larger than those entering in the quantum subsystem treated in the preceding section. To simplify notation, and without loss of generality, the solvent molecules are taken as atoms. Then,

$$H_m(\boldsymbol{P}_m,\boldsymbol{R}_m) = \sum_i \left[H\left(\boldsymbol{P}_{m_i}\right) + \sum_j V\left(\boldsymbol{R}_{m_i},\boldsymbol{R}_{m_j}\right) \right],$$

where $\boldsymbol{R}_{m_i}$ is the position vector and $\boldsymbol{P}_{m_i}$ is the canonically conjugated momentum of the ith solvent atom; $H(\boldsymbol{P}_{m_i})$ is the kinetic energy. The atoms interact via the total charge densities; $V(\boldsymbol{R}_{m_i}, \boldsymbol{R}_{m_j})$ is the interatomic potential between solvent atoms.

The solute is a polyatomic system characterized by the intramolecular potential $V(\boldsymbol{R}_s)$. The solute Hamiltonian in the laboratory coordinate frame can be cast into a matrix form as follows:

$$H_s = (1/2)\boldsymbol{P}_s^{+} \cdot \mathbf{M}^{-1} \cdot \boldsymbol{P}_s + V(\boldsymbol{R}_s). \tag{68}$$

$\mathbf{M}$ is the mass diagonal matrix, $\mathbf{M}^{-1}$ its inverse; $\boldsymbol{P}_s^{+}$ is the row vector obtained by Hermitian conjugation of the column vector $\boldsymbol{P}_s$. The potential $V(\boldsymbol{R}_s)$ represents the force field of the solute.

The solute–solvent interaction V_{ms} is represented as a pairwise potential energy function:

$$V_{ms} = \sum_{i=1}^{N_m} \sum_{j=1}^{N_s} V\left(\left| \boldsymbol{R}_{m_i} - \boldsymbol{R}_{s_j} \right| \right), \tag{69}$$

where N_m and N_s are the number of atoms in the solvent and solute, respectively. As discussed in section 2, $V(|\boldsymbol{R}_{m_i} - \boldsymbol{R}_{s_j}|)$ contains exchange repulsive terms,

Van der Waals attractive interactions and electrostatic interaction terms. The dependence of this potential function with the internal geometry of the solute is neglected in standard treatments. It is usually assumed that the solute is found at its equilibrium nuclear configuration when $V(|\boldsymbol{R}_{m_i} - \boldsymbol{R}_{s_j}|)$ is either calculated from ab initio quantum chemical calculations, or fitted from experimental data.

The problem resides in the construction of the time evolution equation for the dynamical variables of the system of interest embedded in its surroundings. One of its results is the trajectory $\boldsymbol{X}(t)$ for the particles of the subsystem. To obtain such equations, a projection technique applied to the Liouville equation for the full system leads to the formal equations.

The time evolution of the dynamical variables is controlled by the Liouvillian superoperator $\mathbf{iL}$. Let $\boldsymbol{A} = (A_1, A_2, \ldots)$ be a row vector gathering the dynamical variables of our system. According to classical statistical mechanics, the equation of motion is given by the Liouville equation:

$$\mathrm{d}\boldsymbol{A}(t)/\mathrm{d}t = \dot{\boldsymbol{A}}(t) = \mathbf{iL}(t)\boldsymbol{A}(t), \tag{70}$$

where the Liouvillian operator is given by:

$$\mathbf{iL}(t) = \sum_k (\partial H/\partial \boldsymbol{P}_k)\partial[\ ]/\partial R_k - \sum_k (\partial H/\partial \boldsymbol{R}_k)\partial[\ ]/\partial \boldsymbol{P}_k; \tag{71}$$

the sum over k goes over the $3N$ degrees of freedom including both the solute and its surroundings. From this definition, it is obvious that the Liouvillian can be written in a way similar to the Hamiltonian, i.e.:

$$\mathbf{iL} = \mathbf{iL}_m + \mathbf{iL}_s + \mathbf{iL}_{ms}. \tag{72}$$

The interaction Liouvillian acquires a simple form:

$$\mathbf{iL}_{ms} = -\sum_i \sum_j (\partial V\left(\left|\boldsymbol{R}_{m_i} - \boldsymbol{R}_{s_j}\right|\right)/\partial \boldsymbol{R}_{mi}\partial[\ ]/\partial \boldsymbol{P}_{sj}$$

$$-\sum_j \left\{\sum_i \partial V\left(\left|\boldsymbol{R}_{m_i} - \boldsymbol{R}_{s_j}\right|\right)/\partial \boldsymbol{R}_{mj}\right\}\partial[\ ]/\partial \boldsymbol{P}_{mi}, \tag{73}$$

describing the coupling between the solvent and solute, at time t, in terms of forces.

In molecular dynamics simulations, the fundamental dynamical variables are the coordinates and canonically conjugated moments. The time evolution of the pdf (probability distribution function) $f(\Gamma)$, where Γ is a point in phase space at time t, i.e. $\Gamma = (\boldsymbol{P}_{m1}, \boldsymbol{P}_{m2}, \ldots, \boldsymbol{R}_{m1}, \boldsymbol{R}_{m2}, \ldots, \boldsymbol{P}_{s1}, \boldsymbol{P}_{s2}, \ldots, \boldsymbol{R}_{s1}, \boldsymbol{R}_{s2}, \ldots)$, is described by the Liouville equation:

$$\partial f/\partial t = -\mathbf{iL}\, f(\Gamma, t). \tag{74}$$

The formal solution is $f(t) = \exp(-\mathbf{i}\mathbf{L}\,t)f(\Gamma, 0)$. The initial condition is taken to be an ensemble in thermal equilibrium which, for a number of situations of chemical interest, is represented by the canonical distribution:

$$f(\Gamma,0) = f(\Gamma_0) = \exp(-\beta H) / \int \mathrm{d}\Gamma_0 \exp(-\beta H), \tag{75}$$

with $\beta = 1/k_B T$; k_B stands for the Boltzmann constant and T is the absolute temperature. This distribution function is used to carry out the statistical averages ($\langle \ldots \rangle$) in what follows.

The first problem faced is the construction of an effective pdf with which thermal averages could be carried out for the subsystem of interest. This latter may be a rather large system if only classical mechanical degrees of freedom are considered. Such is the case in molecular dynamics simulations of liquids and biomolecules.

3.1. PROJECTED EQUATION OF MOTION

For a theory of solvent effects on a given subsystem, it is important to define its reduced (effective) pdf, $f_s(\Gamma_s)$. This can be achieved by using the theory of projection operators [9]. The operator projecting onto the dynamical variable space of the subsystem of interest can be defined as: $f_s(\Gamma_s, t) = f_m(\Gamma_m, 0)\int \mathrm{d}\Gamma_m f(\Gamma, t) = \mathbf{P}f$; the density for the subsystem m is obtained as: $f_m(\Gamma_m, t) = (1 - \mathbf{P})f(\Gamma, t)$. Multiplying eq. (74), in turn, by $\mathbf{P}$ and the projection operator for the orthogonal complement $\mathbf{Q} = 1 - \mathbf{P}$, two differential equations are obtained. Solving the one for $f_m(\Gamma_m, t)$ first, and introducing this solution into the one for $f_s(\Gamma_s, t)$, the density of the subsystem s can be obtained as a solution of the master equation:

$$\partial f_s(\Gamma_s,t) / \partial t = \mathbf{i}\mathbf{P}\mathbf{L}\mathbf{P} f(T,t) + \mathbf{i}\mathbf{P}\mathbf{L}\mathrm{e}^{\mathbf{i}\mathbf{Q}\mathbf{L}t} f(\Gamma,0) + \int \mathrm{d}u\, \mathbf{i}\mathbf{P}\mathbf{L}\mathrm{e}^{\mathbf{i}\mathbf{Q}\mathbf{L}u}\mathbf{Q}\mathbf{L}\mathbf{P} f(\Gamma,t-u). \tag{76}$$

This equation describes the time evolution of the projected density $f_s(\Gamma_s, t)$. The first term in (76) evolves in the subspace of the s-system; the second term acts as a force produced by the evolution in the m-system which has a projection onto the s-system; the last term has a non-Markovian character, as can be seen from the memory kernel $\mathbf{i}\mathbf{P}\mathbf{L}\,\mathrm{e}^{\mathbf{i}\mathbf{Q}\mathbf{L}u}\mathbf{Q}\mathbf{L}\mathbf{P}$.

In principle, eq. (76) solves the problem of finding a projected pdf. As a result of the partitioning accomplished above, an effective Liouvillian can be written,

$$\langle \mathbf{i}\mathbf{L}_s \rangle_m = \mathbf{i}\mathbf{L}_s + \mathbf{i}\langle \mathbf{L}_{ms} \rangle_m + \mathbf{i}\mathbf{L}_{nc}\,, \tag{77}$$

where $\mathbf{i}\langle \mathbf{L}_{ms} \rangle_m$ describe the conservative force effects between both subsystems, and $\mathbf{i}\mathbf{L}_{nc}$ stands for the non-conservative effects. This latter factor is neglected in standard treatments and the corresponding Liouvillian $\langle \mathbf{i}\mathbf{L}_s \rangle_m$ is written as iL_s.

It is a commonly found situation that the solvent molecules in a neighborhood of the solute are strongly perturbed by it. The subsystem of interest is therefore

defined so as to include sufficient solvent molecules in the simulation. This extended system is usually referred to as cybotaxis. In this way, the correlations between this redefined subsystem and the remainder may be weak enough to neglect them. The effective Liouvillian $i\mathcal{L}_s$ describes this latter type of subsystem. Still, the problem remains as to where the border should be drawn.

The projection operator technique is also used for evaluating the dynamical variables of the cybotaxis. In order to proceed, it is necessary to define a scalar product among linearly independent dynamical variables. One useful definition is via a correlation function matrix between two vector variables, say $\boldsymbol{A}$ and $\boldsymbol{B}$:

$$\langle \boldsymbol{A}(t)\boldsymbol{B}(0)\rangle = \int \mathrm{d}\Gamma_0 f(\Gamma_0)\boldsymbol{A}(t)\boldsymbol{B}^+(0) = (\boldsymbol{A}(t),\boldsymbol{B}^+(0)). \tag{78}$$

In terms of matrix elements, this equation reads:

$$\langle A_i(t)B_j(0)\rangle = \int \mathrm{d}\Gamma_0 f(\Gamma_0)A_i(t)(t)B_j^*(0) = (A_i(t),B_j^*(0)); \tag{79}$$

as usual, the star indicates complex conjugation. The last identity is a commonly used notation to indicate a scalar product.

The time evolution of the momentum variables for the subsystem $\boldsymbol{P}_s$ are obtained after introducing a projection operator for the variables of interest into the global Liouville eq. (76):

$$\dot{\boldsymbol{P}}_s = (\mathrm{i}\mathbf{L}\boldsymbol{P}_s(0),\boldsymbol{P}_s^+(0))(\boldsymbol{P}_s(0),\boldsymbol{P}_s^+0))^{-1}\boldsymbol{P}(t) + \boldsymbol{F}_s(t) + \int \mathrm{d}t'\Phi(t-t')\boldsymbol{P}(t'). \tag{80}$$

The first term is nothing but the fully projected part of the dynamics $\mathbf{iPLP}\boldsymbol{P}(t)$, the second term, $\boldsymbol{F}_s(t) = \mathbf{iPLQ}\,\mathrm{e}^{\mathrm{iQL}t}\boldsymbol{P}(0)$, is a stochastic force put up by the coupling of the system of interest with the remaining system, and the third term is another way of writing the non-Markovian effects.

The autocorrelation matrix of the stochastic force $\boldsymbol{F}_s(t)$ is related to the kernel of the non-Markovian term by:

$$\Phi_s(t) = (\boldsymbol{F}_s(t),\boldsymbol{F}_s^+(0))(\boldsymbol{P}_s(0),\boldsymbol{P}_s^+(0))^{-1}, \tag{81}$$

which is the second fluctuation–dissipation theorem. The time integral of $\Phi_s(t)$ is a generalized friction function:

$$\gamma(t) = \int \mathrm{d}t'\,\Phi(t-t'). \tag{82}$$

Approximate time evolution equations for the dynamical variables of the system of interest that are used in molecular dynamics simulations can now be obtained. If the total Liouvillian is replaced by $i\mathcal{L}_s$ in eq. (80) and use of eqs. (71) and (73), a generalized Langevin equation follows:

$$\dot{\boldsymbol{P}}_s = -\partial V(\boldsymbol{R}_s)/\partial \boldsymbol{R}_s - \partial\langle V_{ms}\rangle_m/\partial \boldsymbol{R}_s + \boldsymbol{F}_s(t) + \int \mathrm{d}t'\,\Phi(t-t')\boldsymbol{P}_s(t'). \tag{83}$$

The first term describes the intramolecular force field, the second one is important if the medium surrounding the subsystem of interest is structured: such a situation is met in enzyme's active site simulation (see below); otherwise, this term cancels out. The stochastic forces and memory term are the subject of particular modelizations.

Equation (83) is the fundamental equation behind molecular dynamics simulation of small systems immersed in a thermal bath.

3.1.1. Force fields

The potential energy function $V(\boldsymbol{R}_s)$ describes the interactions between the atoms in the system of interest. For a protein or nucleic acid, the potential is composed of terms representing covalent bond stretching, bond angle bending, quadratic dihedral bending which takes care of out-of-plane and out-of-tetrahedral configuration motion, sinusoidal dihedral torsion, interatomic repulsive forces and Van der Waals attractive interactions which make up for the Lennard–Jones potential used in normal liquid simulations, and Coulomb interactions. Empirical parameter sets have been developed and are now embodied in Rehovot (M. Levitt), Harvard (CHARMM) [18], Groningen (GROMOS) [19] and AMBER [24] computer programs.

Water–water potential functions have been reported from different laboratories. The procedure leads either to analytic functions fitted to empirical data, e.g. SPC (single point charge) [20] and TIPS (transferable intermolecular potential functions) family [21], or to ab initio quantum chemical calculations, e.g. Clementi's MCY potential [22]. At this point, it is interesting to examine the performance of the MCY potential. This water model has been extensively used in the literature. Rice and coworkers have tested the accuracy of water–water potentials [56–59]. Predictions of densities, lattice energies and lattice geometries of the proton ordered ices have been used to test this potential. The MCY potential predicts the correct ice lattice structures but not their density. Frequency-dependent properties have also been examined. As pointed out by Beveridge [60], excellent agreement is obtained with experiments on the oxygen–oxygen radial distribution function. The model fails to give reasonable pressure. In a recent release of this potential (CPC model), this drawback has been overcome. Three-body interactions have been computed and used in MC simulations by Clementi and Corongiu [61]. Later, fourth-body interactions were included [62]. Improvements in the oxygen–oxygen radial distribution function and enthalpy have been reported. Computer simulations with the MCY potential of the dielectric constant of water have been reported by Neumann [63].

3.1.2. Electric field of water

Colonna et al. [64] have recently studied the modelling of the electric field of water obtained from accurate SCF wave functions. Model representations are constructed with the distributed multipole approach [65–71] and effective point charges. Benchmark computations with extended basis sets were made for the

dipole, the quadrupole and the octupole moments of a water molecule, together with the electric field on several grids around it. Comparisons were made with fields obtained from standard basis sets (6-31G, 6-31G*, 6-31G**, 3-21G, 3-21G* and the minimal basis set STO-3G) and point charge models commonly used in molecular simulations of water, namely, TIP3P [72], ST2 [73], SPC [20], TIP4P [72], Kollman's 5-charges [75], Dacre's model [74], MCY [22] and CPC [76] models.

This theoretical study showed that at 1 Å distance, TIP3P [72], ST2 [73] and SPC [20] model charges yield a rather poor fit to the reference field. TIP4P [72] and one of Kollman's 5-charge models [75] are somewhat better. Dacre's representation [74] yields results that are within acceptable bounds to qualify it as accurate. MCY [22] charges, as expected, do not reproduce the field correctly, while the more recent CPC model [76], which is designed to incorporate polarization effects explicitly, works well.

From a more general perspective, modelling the electric field outside an anisotropic charge density basically defining the active site (cybotaxis) is a problem of current interest both in classical molecular dynamics and in microscopic quantum chemical theories of solvent effects. In the past, electrostatic interactions between highly charged species have been represented with rough and sometimes unrealistic dielectric models (for a review, see [77]) that would fail to describe the anisotropic response of the microscopic polarizable media. An adequate description of interactions between ions, zwitterions, highly polar molecules, proteins and other types of biomacromolecules requires a proper handling of reaction field effects. In principle, such effects can be rigorously incorporated in MD simulations by using analytical first- and second-energy derivatives, which have recently been reported by us for the polarization model of many-body interactions of molecular aggregates [78]. Bearing in mind that it is the electric field which is the source of polarization, its accurate determination is enforced both in classical and quantum chemical contexts.

3.2. ACTIVE SITE MODELS

The theory of solvent effects is constructed on a spatially localized model for the events of interest. By such techniques, as it was illustrated in the quantum mechanical section, only a small part of the system is included in the explicit simulation and the effects of the remainder of the system are treated implicitly. Equation (83) reflects this idea.

For liquids, Adelman, Karplus, McCammon and coworkers [79–83] have proposed a method for simulating a localized region and replacing distant atoms by a suitably constructed boundary, including a stochastic heat bath. Warshel's surface constraint compressible dipole model pioneered this type of approach [84]. Basically, the system is partitioned into a reactant subsystem and a boundary or reservoir region. The former is further partitioned into a reaction region and a buffer region. The molecules in the buffer are treated as Langevin particles and, in the reaction region, they are simulated by standard molecular dynamics. The reservoir provides

a static force field that helps ensure that correct structural and dynamic properties will be maintained within the reaction zone.

This type of approach has been extended to treat active site dynamics of enzymes by us and several other authors [85–87].

3.3. MOLECULAR DYNAMICS WITH POLARIZATION FIELDS

Molecular dynamics simulations including polarization require, in addition to standard terms in the potential energy function, a modified expression for the electrostatic interaction including reaction field (polarization) effects. This issue has been thoroughly discussed by Ángyán et al. in a recent communication [78].

In the microscopic approach, the total electrostatic energy for a system formed by a set of point charges (Q) and polarizabilities (α) is given by:

$$U = \frac{1}{2}\sum_i \sum_j Q_i Q_j T(\boldsymbol{R}_i - \boldsymbol{R}_j) - \frac{1}{2}\sum_i \sum_j Q_i \boldsymbol{T}(\boldsymbol{R}_i - \boldsymbol{R}_j)\cdot \boldsymbol{p}(\boldsymbol{R}_j). \tag{84}$$

$\boldsymbol{p}(\boldsymbol{R}_j)$ is the induced dipole moment at the point $\boldsymbol{R}_j$ and the Coulomb kernels have already been defined in section 2. Note that self-energy terms are avoided in the summations by taking the corresponding kernels equal to zero; this convention facilitates practical computations.

In eq. (84), the first term is the direct electrostatic interaction between bare charges and the second term is the induction energy. This expression can be recast in terms of the field produces by the charges $\boldsymbol{e}(\boldsymbol{R}_i)$ at $\boldsymbol{R}_i$, and the electric field of the induced dipoles. $\pi(\boldsymbol{R}_j) = -\nabla_{R_i}\Pi(\boldsymbol{R}_j)$ is the reaction field:

$$\boldsymbol{e}(\boldsymbol{R}_i) = -\sum_j Q_j \nabla_{R_i} T(\boldsymbol{R}_i - \boldsymbol{R}_j) = -\sum_j Q_j \boldsymbol{T}(\boldsymbol{R}_i - \boldsymbol{R}_j) \tag{85}$$

and

$$\pi(\boldsymbol{R}_j) = -\nabla_{R_i}\Pi(\boldsymbol{R}_j) = \sum_{j\neq i} \boldsymbol{T}(\boldsymbol{R}_i - \boldsymbol{R}_j)\cdot \boldsymbol{p}(\boldsymbol{R}_j), \tag{86}$$

since the induced dipoles are determined by eq. (16) translated into a discrete representation,

$$\boldsymbol{p}(\boldsymbol{R}_i) = \alpha_i(\boldsymbol{e}(\boldsymbol{R}_i) + \sum_{j\neq i} \boldsymbol{T}(\boldsymbol{R}_i - \boldsymbol{R}_j)\cdot \boldsymbol{p}(\boldsymbol{R}_i)), \tag{87}$$

and after introducing the supermatrix representation of section 2.3, one obtains eq. (38), namely,

$$\mathbf{p} = [\mathbf{1} - \boldsymbol{\alpha}\cdot\mathbf{T}]^{-1}\cdot\boldsymbol{\alpha}\cdot\mathbf{e} = \boldsymbol{\alpha}\cdot[\mathbf{1} - \boldsymbol{\alpha}\cdot\mathbf{T}]^{-1}\cdot\mathbf{e} = \mathbf{A}\cdot\mathbf{e}.$$

The induction energy may now be written either in terms of the permanent field interacting with the induced dipoles or as a bilinear form in the permanent field including the many-body polarizability matrix:

$$U_{\text{ind}} = -\frac{1}{2}\sum_i p(R_i)\cdot e(R_i) = -\frac{1}{2}\sum_i \sum_j e(R_i)\cdot A(R_i, R_j)\cdot e(R_i), \tag{88}$$

where $A(R_i, R_j)$ is an element of the supermatrix of many-body polarizabilities.

The calculation of the force acting at the ith nucleus is obtained after taking the gradient:

$$F(R_i) = -\nabla_{R_i} U = Q_j e(R_i) + \tfrac{1}{2}(p(R_i)\cdot \nabla_{R_i} e(R_i) + Q_j \Pi(R_i) + \sum_j e(R_j)\cdot \nabla_{R_i} p(R_j). \tag{89}$$

The problem is now shifted to the calculation of the gradient of the induced dipole. This can be done as shown by Ángyán et al. by noting that the derivative of the supermatrix A with respect to an arbitrary coordinate x can be written as:

$$\partial \mathbf{A}/\partial \mathbf{x} = \mathbf{A}(\partial \mathbf{T}/\partial x)\mathbf{A}. \tag{90}$$

Then it can be shown that

$$\begin{aligned}\nabla_{R_i} p(R_j) = -A(R_i, R_j)\cdot \Bigg(& \sum_k \nabla_{R_j} T(R_j - R_k) Q_k + \sum_k \sum_l [\nabla_{R_i} \nabla_{R_i} T(R_i R_j)] \\ & \cdot A(R_j, R_k)\cdot T(R_k - R_l) Q_l \Bigg) \\ + \sum_j A(R_i, R_j)\cdot \Bigg(& \sum_k \nabla_{R_j} T(R_j - R_k) Q_k + \sum_k \sum_l [\nabla_{R_i} \nabla_{R_i} T(R_i - R_j)] \\ & \cdot A(R_j, R_k)\cdot T(R_k - R_l) Q_l \Bigg).\end{aligned} \tag{91}$$

The first terms in both sums are the electrostatic field gradients of the permanent charges at points R_j and R_k, respectively. The second terms are the field gradients created by induced dipole moments. Thus, the dipole moment derivative can be recast into:

$$\begin{aligned}\nabla_{R_i} p(R_j) = & -A(R_i, R_j)\cdot (\nabla_{R_j} e(R_j) + \nabla_{R_j} \Pi(R_j)) \\ & + \sum_k A(R_i, R_k)\cdot (\nabla_{R_k} T(R_k - R_j) Q_j - [\nabla_{R_k} \nabla_{R_k} T(R_k - R_j)]\cdot p(R_j)).\end{aligned} \tag{92}$$

This result is based upon the self-consistency of eq. (39) and it does not necessarily hold for approximately calculated induced moments.

Interestingly, when first-order induced moments are used, i.e. $p^{(1)}(R_i) = \alpha_i e(R_i)$, terms involving derivatives of the many-body polarization matrix vanish and $A(R_i, R_j)$ is just the polarizability tensor at R_i: $A(R_i, R_j) = \delta_{ij}\alpha_i$, and eq. (91) becomes

$$\nabla_{R_i} p^{(1)}(R_j) = -\delta_{ij}\alpha_j \cdot \nabla_{R_j} e(R_j) + \alpha_i \nabla_{R_i} T(R_i - R_j) Q_j. \tag{93}$$

The total force in eq. (89) can be obtained if one substitutes eq. (92) into (89). A simple form can be obtained if one uses the total electric field, namely $\boldsymbol{E}(\boldsymbol{R}_i) = \boldsymbol{e}(\boldsymbol{R}_i) + \Pi(\boldsymbol{R}_j)$, and its field gradient:

$$\boldsymbol{F}(\boldsymbol{R}_i) = -\nabla_{\boldsymbol{R}_i} U = Q_i \boldsymbol{E}(\boldsymbol{R}_i) + \boldsymbol{p}(\boldsymbol{R}_i) \cdot \nabla_{\boldsymbol{R}_i} \boldsymbol{E}(\boldsymbol{R}_i). \tag{94}$$

In the particular case of first-order induced dipoles, one obtains:

$$\boldsymbol{F}^{(1)}(\boldsymbol{R}_i) = Q_i \boldsymbol{E}(\boldsymbol{R}_i) + \boldsymbol{p}^{(1)}(\boldsymbol{R}_i) \cdot \nabla_{\boldsymbol{R}_i} \boldsymbol{e}(\boldsymbol{R}_i). \tag{95}$$

Second-order derivatives have also been obtained by Ángyán et al. [78].

Finally, we give the expression for the induction energy in first-order approximation:

$$U_{\text{ind}} = -\frac{1}{2} \sum_i \boldsymbol{p}^{(1)}(\boldsymbol{R}_i) \cdot \boldsymbol{e}(\boldsymbol{R}_i) = \frac{1}{2} \sum_i \sum_j Q_i Q_j \boldsymbol{T}(\boldsymbol{R}_i - \boldsymbol{R}_j) \cdot \alpha_j \cdot \boldsymbol{T}(\boldsymbol{R}_j - \boldsymbol{R}_i). \tag{96}$$

In this case, the total electrostatic energy is given by

$$U = \frac{1}{2} \sum_i \sum_j Q_i Q_j T(\boldsymbol{R}_i - \boldsymbol{R}_j) + \frac{1}{2} \sum_i \sum_j Q_i Q_j \boldsymbol{T}(\boldsymbol{R}_i - \boldsymbol{R}_j) \cdot \alpha_j \cdot \boldsymbol{T}(\boldsymbol{R}_j - \boldsymbol{R}_i). \tag{97}$$

The interaction energy without including charging effects reads:

$$U_{\text{int}} = \sum_{i<j} Q_i Q_j T(\boldsymbol{R}_i - \boldsymbol{R}_j) + \sum_{i \neq j} Q_i Q_j \boldsymbol{T}(\boldsymbol{R}_i - \boldsymbol{R}_j) \cdot \alpha_j \cdot \boldsymbol{T}(\boldsymbol{R}_j - \boldsymbol{R}_i). \tag{98}$$

In all summations over centers, self-energy terms are set equal to zero.

Polarization forces have been implemented in CHARMM by F. Colonna. His simulation of simple model systems shows energy conservation if the self-consistent approach is used. The use of non-converging induced dipoles leads to trajectories where energy is not conserved. Interestingly, the first-order approximation works well in the sense that there is energy conservation during the MD simulation [78].

3.4. MICROSCOPIC POLARIZATION EFFECTS: SIMPLE ANALYTICAL MODELS

Polarization effects at the microscopic level are not reproducible by standard dielectric functions. This can be appreciated when the theory is applied to simple models that can be analytically solved. In fig. 1, two model situations are presented; they were discussed previously by Buckingham [88,89].

For the case depicted in fig. 1(a), eq. (98) can easily be specialized, giving for the total electrostatic energy the formula:

$$E_{12} = (q_1 q_2 / Q)\,(1\text{-sgn}(q_1 q_2) 32\, \alpha / Q^3), \tag{99}$$

which is the equation given in ref. [88].

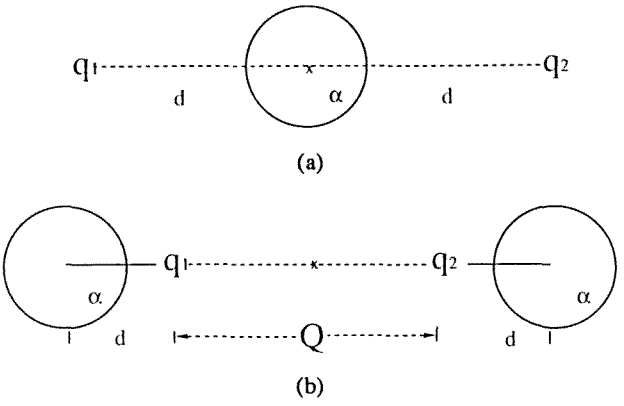

Fig. 1.Buckingham's model illustrating microscopic polarization effects. *a* represents polarizability; *d* is the distance between a charge and the polarizable center; Q is the distance between charges q_1 and q_2.

Thus, the attractive interaction between oppositely charged particles is enhanced, contrary to the expected result if a standard dielectric function were used, while the repulsive interaction is decreased by the polarizability, in agreement with the use of dielectric functions.

Let us calculate the forces between the charges as specified for the cases depicted in fig. 1(b). Since the procedure has its own interest, it will be presented in some detail.

First, we calculate U_{rf}, the potential energy in a reaction field. For a set of point charges q as discussed in the preceding section, this energy is:

$$U_{rf} = \sum_i q_i \Pi(Q_i), \tag{100}$$

where $\Pi(Q_i)$ is the reaction field potential at point Q_i. Consider now the charge pair q_1 and q_2 at a distance Q_{12} and calculated the force between them:

$$-\delta U_{rf}/\delta Q_{12} = -\{(\delta U_{rf}/\delta Q_1)(\delta Q_1/\delta Q_{12}) + (\delta U_{rf}/\delta Q_2)(\delta Q_2/\delta Q_{12})\}, \tag{101}$$

which can be recast in terms of the reaction field $\pi(Q)$:

$$-(\delta U_{rf}/\delta Q_{12}) = q_1\pi(Q_1) - q_2\pi(Q_2) = F_{rf}(12). \tag{102}$$

Specializing this equation to describe the case of fig. 1(b), one obtains after some simple algebraic rearrangements:

$$F_{rf}(12) = F_{rf}(\text{self}) + F_{rf}(\text{int}), \tag{103}$$

where

$$F_{rf}(\text{self}) = 2\alpha(q_1^2 + q_2^2)\{1/d^5 - 1/(Q+d)^5\} \tag{104}$$

and

$$F_{rf}(\text{int}) = (q_1q_2)/Q^2\{(4\alpha/d^3)(1+2d/Q+(d/Q)^2)(1+d/Q)^{-5}\}. \tag{105}$$

Now, by adding the in vacuo forces and neglecting the self-reaction force, one finally obtains:

$$F_{12} = ((q_1q_2)/Q^2)\{1+\text{sgn}(q_1q_2)(4\alpha/d^3)(1+2d/Q+(d/Q)^2)(1+d/Q)^{-5}\}, \tag{106}$$

which is the final formula.

Equation (106) with $\text{sgn}(q_1q_2) = -1$ has to be compared with Buckingham's equation for $q_1 = -q_2 = q$, which looks like:

$$F_{12} = ((q_1q_2)/Q^2)\{1-(4\alpha/d^2Q)(1+2d/Q)(1+d/Q)^{-5}\}, \tag{107}$$

while for $q_1 = q_2 = q$, it reads:

$$F_{12} = ((q_1q_2)/Q^2)\{1+(4\alpha/d^2Q)(1+2d/Q)(1+d/Q)^{-5}\}, \tag{108}$$

Equations (107) and (108) are slightly different when compared to those obtained from eq. (106). The difference stems from the fact that once the interaction energy U_{12} is calculated, the field produced by, say q_1, at the place the charge q_2 occupies, follows from:

$$F_{12} = \lim_{q_2=0} -(1/q_2)(\delta U_{12}/\delta Q). \tag{109}$$

Thus, the test charge does not alter the field created by the source q_1. The reaction field approach, on the contrary, includes the perturbing effects of the "test" charge. The point is that both schemes yield the same qualitative effects.

The force strength between two charges of opposite sign decreases according to eq. (106) where $\text{sgn}(q_1q_2) = -1$ or eq. (107), in agreement with the standard dielectric theory for macroscopic bodies, while the force strength between two charges of equal sign increases in the presence of polarizable matter, which is a result that cannot be reproduced with the standard models of dielectric continuum.

4. Solvent effect studies. Selected examples

Progress in computer technology has profoundly influenced the field of solvent effects. From small solutes up to proteins in water, a number of studies have been reported in the literature.

4.1. SOLVENT EFFECTS ON MOLECULAR PROPERTIES

4.1.1. Monte Carlo simulations

Monte Carlo simulation techniques have been extensively used to study solvent effects on molecular properties and equilibrium points. Jorgensen has summarized theoretical work of condensed-phase effects on conformational equilibria [90].

Monte Carlo studies of a dilute aqueous solution of benzene have been reported by Beveridge and coworkers [60] and by Linse et al. [91]. Intermolecular pairwise potential functions determined from quantum mechanical calculations were used; for water–water interactions, both groups have used the pairwise Matsuoka–Clementi–Yoshimine potential. Interesting solvation patterns around benzene have been found.

Solvent effects on the relative energies for the planar and perpendicular allyl cations in liquid hydrogen fluoride have been studied with MC simulations. The intermolecular potential functions describing the solute–solvent interactions have been obtained from ab initio molecular orbital calculations with a 4-31G basis set. The conformers were represented by different charge distributions but the same Lennard–Jones parameters. The solvent–solvent interactions were described by a function of the TIPS form including one Lennard–Jones term and three charged sites for each hydrogen fluoride monomer [92]. Significant differences in solvation are detected. The more localized (perpendicular) conformer is found to be better solvated, in agreement with traditional notions of ion solvation.

Clementi and coworkers have been deriving intermolecular potentials from ab initio computations and using these potentials in statistical mechanical computer experiments on pure solvents and solutions, the goal being the derivation of an ordered set of approximations leading to an increasingly realistic description of water and aqueous solutions. Extensive work on solvation of biomolecules has been carried out [93–96] by this group. The reader is referred to the MOTECC-90 book, edited by Clementi (ESCOM, Leiden).

Free energy simulations using the Metropolis MC method and a coupling parameter approach with umbrella sampling have been performed for a number of systems ranging from liquid water to chemical reaction in solution. Applications to the study of liquid water, hydrophobic interactions and solvent effects on conformational stability have been reported by Beveridge et al. [97]. The effect of hydration on the torsional energy surface of butane using statistical perturbation theory has been studied by Jorgensen and Buckner [98]. The methodology therein proposed is shown to yield results with high precision and to have significant advantages over umbrella sampling.

4.1.2. Molecular dynamics simulations

Dynamics of proteins and nucleic acids and their solvent surroundings have been reviewed recently by McCammon and Harvey [99]. Van Gunsteren [100] has surveyed methods for simulation of molecular systems on a computer. Free energy

perturbation calculations have become a standard technique to study binding and catalysis after site directed mutagenesis, relative binding free energies of different substrates onto the same receptor, and other applications [101–103].

Solvent effects on biochemical solutes studied by MD techniques have also been discussed by Jorgensen. In this section, recent studies on protein solvation are examined.

Only a few MD simulations of proteins in aqueous solution have been carried out. Such studies provide data for examining conformational differences between crystal structures and solution structures [19, 104–108]. Thus, Ahlström et al. have studied parvalbumine in vacuo and in aqueous solution. Parvalbumine is a Ca^{+2} binding protein. Simulations in vacuo of Ca^{+2}-free protein were also reported. Considerable structural changes were detected relative to the initial coordinate in the two in vacuo simulations. Water surroundings, although they help to hold the structure in a neighborhood of the crystallographic structure a little better, are not sufficient to keep a correct pattern. It is not clear from the analysis presented by these authors whether their results are mere artifacts deriving from the potential function used or if the results are a fair representation of the protein behavior in solution.

Simulations of BPTI (bovine pancreatic trypsin inhibitor) in Van der Waals solvents have been reported [105, 106]; the density and molecular size were chosen to simulate those of water. More realistic water representations were used in further simulations [19, 107, 108]. Levitt and Sharon have recently reported accurate simulations of protein dynamics in solution for BPTI [109]. Avian pancreatic polypeptide hormone in crystal and in aqueous solution have been reported by Kruger [110]. These studies tend to indicate that the calculations in vacuo represent fairly correctly the motion of the protein core, while exposed side chains represent more strongly solvent effects.

A 500-picosecond MD simulation in water of the Manα1 $\rightarrow$ 2Manα glycosidic linkage present in Asn-linked oligomannose-type structures on glycoproteins is reported in [111]. Significant dampening of the molecular fluctuations was found when comparisons of in vacuo and in water simulations were made. The in-water simulation showed only occasional short-lived deviations from the minimum energy conformation; this result is more consistent with the carbohydrate "breathing" mode rather than flexiblity of the global structure. Recently, a simulation covering 100 ps at 25 °C of the third domain of silver pheasant ovomucoids in aqueous solution has been reported [112].

Collective motions of secondary and supersecondary structures in proteins are important for understanding functionality. A unique collective motion has been detected by molecular dynamics simulations of the carboxy terminal fragment (CTF) of the L7/L12 ribosomal protein [113]. In the crystal state, the unit cell embodies eight units of CTF. A dimer has been proposed as a functional significant structure. MD simulations of the dimer [114] and of CTF immersed in a bath of 2352 SPC water molecules, carried out in our laboratory, have shown that protein collective

motions are not damped [115,116]. For this particular case, the simulations in vacuo are good enough for reproducing structural features and dynamical behavior of the whole protein.

Since the number of systems studied so far is rather small, general conclusions concerning model water effects on protein structural and dynamic properties cannot be drawn.

4.2. SOLVENT EFFECTS ON CHEMICAL REACTIONS

A Brownian reactive dynamics method has been developed by McCammon and coworkers [101] to calculate the rate at which reactant molecules diffusing in solution would collide with appropriate orientations for a reaction to take place. The method has been applied to study the reaction rate of superoxide dismutase.

A significant application of computer simulation techniques to solvent effects evaluation is the study of nucleophilic addition and the bimolecular nucleophilic substitution S_{N^2} reactions.

The S_{N^2} reaction, $X^- + RY \rightarrow XR + Y^-$, has been simulated with the MC equilibrium method by Jorgensen and coworkers [117,118]. The procedure used by these authors involves three steps: (i) the lowest energy reaction path is determined for the in vacuo system by using ab initio molecular orbital calculations; (ii) intermolecular potential functions are obtained to describe the interactions between the substrate and a solvent molecule; these potentials depend on the internal structure of the substrate; (iii) MC simulations are carried out to determine the free energy profile for the reaction in solution. This is a heavy computational job, since importance sampling methods are required to explore all values of the reaction coordinate. A similar technique was used by Madura and Jorgensen [119] in simulating the nucleophilic addition of the hydroxide ion to formaldehyde in the gas phase and in aqueous solution.

An integral equation approach to obtain free energy surfaces was reported by Rossky and coworkers [120] in the study of the aqueous phase S_{N^2} reaction of chloride with methyl chloride. It was shown that in the limiting case of fast reaction dynamics and rapid charge transfer near the transition state, the correction to the TS theory rate constant can be estimated within the integral equation framework.

Molecular dynamics simulations of the S_{N^2}-type reaction were reported by Bergsma et al. [121]. The technique is used to explore the role of polar solvent dynamics and configurations in modulating the reaction trajectories and the ratio (k) between the true value of the rate constant k and the one obtained from the transition state theory k^{TST}. Important results concerning solvent effects on the dynamics are obtained.

In a more chemical vein, Tapia and coworkers [122–124] have studied hydration effects for the rate limiting step of the acid catalyzed rearrangement of α-acetylenic alcohol to α,β-unsaturated carbonyl compounds. Water intervenes in the reaction mechanism. Ab initio MO studies of the energy hypersurface were carried out at

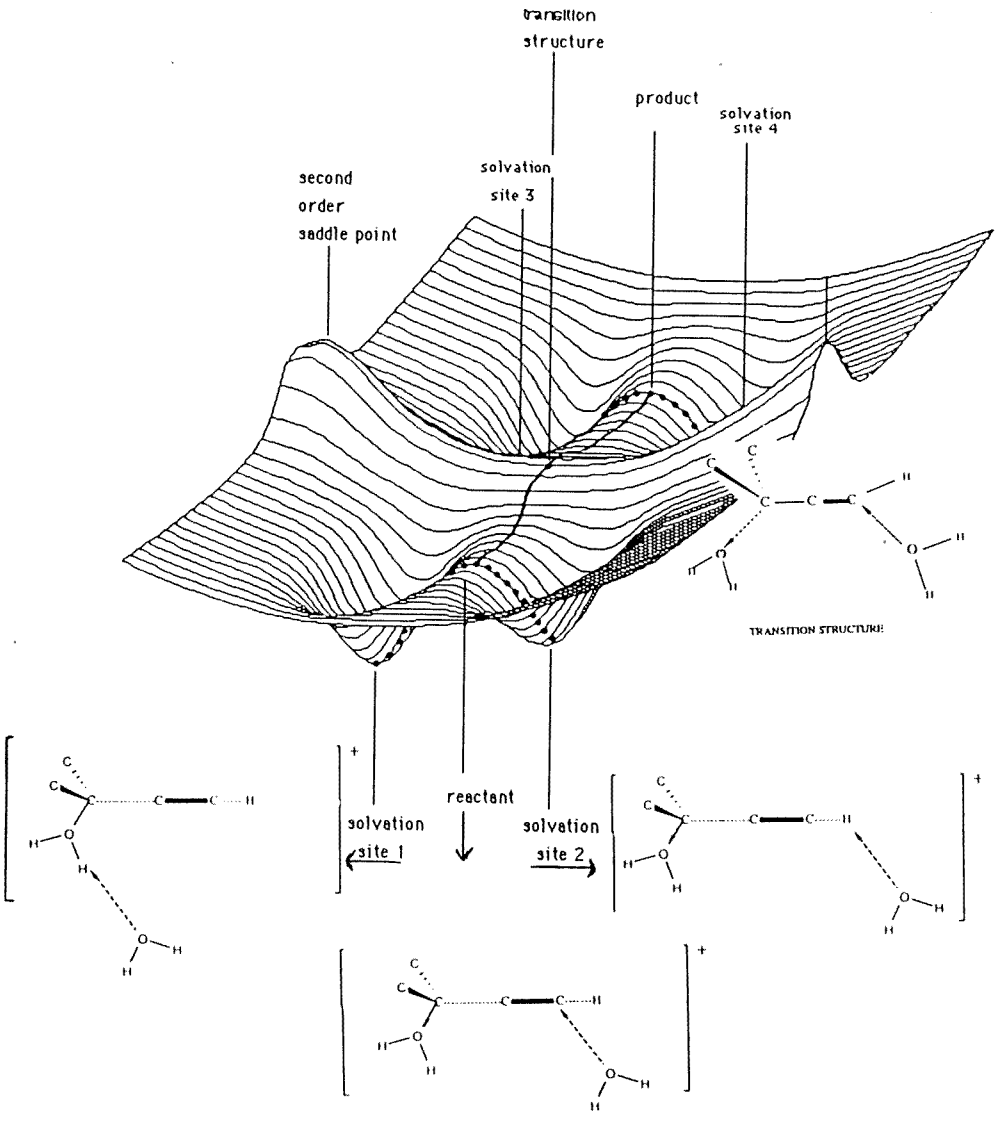

Fig. 2. Schematic representation of the reactive hypersurface for the rate limiting step in the acid catalyzed rearrangement of α-acetylenic alcohol to α, β-unsaturated carbonyl compounds. Note that the reactant in the RLS is a saddle point along the solvation geodesic. As was shown in ref. [124], solvent water molecules occupy solvation sites one and two in a way opening the reactive channel. This work suggests one of the difficulties a theoretical following of a chemical reaction in water may encounter: the curvature of the reactive hypersurface may crucially depend upon many-body solvation effects.

a 4-31G basis set level [123]. The topography of this energy hypersurface presents two solvation sites for the protonated alcohol in the reactant basin and two for the protonated allenol in the product basin; the solvation minima of the protonated

alcohol and allenol are connected via saddle points, which are schematically represented in fig. 2. The surprising result is that the reactant and the product of the RLS (rate limiting step) are also saddle points of first order in the supermolecule hypersurface. MC simulations [124] using the stationary point structures on the energy hypersurface show that the solvation sites are occupied by solvent water molecules, and due to solvent caging effects, the reactant and product of the RLS in solution become stable species. Mechanistically, the transition state for the RLS derives from the solvated reactant by the jump of one solvent molecule towards its nucleophilic center. Since the unrelaxed solvation shell is less efficient than the equilibrium one, the TS for this reaction is better described as a poorly solvated saddle point structure: relaxation of this shell opens the channel leading to final products. This set of studies illustrates the difficulties simulation methods will encounter when water takes an active part in the molecular mechanism.

4.3. INTEGRATED QUANTUM/STATISTICAL MECHANICS STUDIES

In the MC and molecular dynamics studies overviewed above, the key issue of incorporating solvent effects in the quantum mechanical calculation has not been satisfactorily solved. Warshel's empirical valence bond approach, Van Duijnen's direct reaction field method [48], and Tapia's ISCRF theory [25], by including them, are steps forward in this direction. Although the key theoretical issue cannot be considered satisfactorily solved, the applications made are most interesting.

The energies and dynamics of the S_{N^2} class of reactions in aqueous solution have been studied by a combination of the empirical valence-bond method and a free energy perturbation technique. The solvent is represented by a surface constrained all-atom solvent (SCAAS) model, and many-body interactions are taken into account with a solvent parameter set that includes atomic polarizabilities [125]. In this work, activation free energies for the $X^- + CH_3Z \rightarrow XCH_3 + Z^-$ reactions are computed and the general relationship between the reaction free energies and the solvent contribution to the activation free energies are examined. The dynamical aspects of the S_{N^2} charge transfer reaction are explored by propagating trajectories downhill from the transition state by using linear response theory. The simulations suggest solvent fluctuations play a central role in driving the system towards the transition state; the relaxation time for the reactive fluctuations is determined by both the polarization time of the solute dipole moment and the dielectric relaxation time of the solvent. Effects of solute–solvent coupling and solvent saturation effects on solvation dynamics of charge transfer reactions were pursued in another work by Hwang et al. [126].

Evaluation of free energies in genetically modified proteins was made by Warshel et al. [127]. A combination of the EVB method and free energy perturbation approach similar to the one described above was applied to study the activity of genetically modified enzymes: trypsin and subtilisin. The importance of dynamics aspects was sensed by using autocorrelation functions of the protein reaction field

on the reacting substrate. Important free energy perturbation calculations of the catalytic effects associated with substitutions of the active site calcium ion in staphylococcal nuclease were reported by Åqvist and Warshel [128,129]. This procedure shows the enormous power simulation techniques have in helping to predict important biological changes in enzyme functions.

Earlier theoretical work on charge-relay catalysis in the function of serine proteases has been critically evaluated by Schowen [130]. The catalytic steps for acylation and hydrolysis of a model ester by chymotrypsin have been studied with the ISCRF theory. The results are in good qualitative agreement with experimental facts. Since the semiempirical MO method was not calibrated to reproduce in vacuo properties, no quantitative agreement was reached [131]. Náray-Szabó [132–134] has emphasized the role of electrostatics in enzyme catalysis. Progress in recombinant DNA techniques has made it possible to analyze the structural bases of enzyme action and substrate specificity via directed replacement of residues of interest in a given enzyme.

The mechanism of liver alcohol dehydrogenase (LADH) has been extensively studied. For a recent overview, the reader is referred to ref. [135]. Reaction field effects on the transition structure of model hydride transfer systems have been calculated at an ab initio 4-31G basis set level [135,136]. The active sites of enzymes are usually assumed to be designed to receive transition states for the reaction they catalyze. This special sort of surrounding medium effects has been computationally documented recently [137]. From the reaction geodesic passing through the transition state for hydride transfer in the pyridinium cation/methanolate model system, *only the TS-structure could be fitted into the LADH active site.* The normal mode analysis carried out on the TS showed an excellent agreement with isotropic substitution experiments [137]. Reaction field calculations on the model system have also been performed. For an overview of biomolecular interactions, the reader is referred to ref. [138].

An interesting solvation–desolvation effect put in evidence via MC simulations [139,140] has been confirmed in a quantitative structure–activity relationship (QSAR) study [141]. The analysis of the inhibition constants of pyrazoles, phenilacetamides, formylbenzylamines and acetamides acting on LADH yielded QSARs having a linear dependency on octanol–water partition coefficients. One of the tenets of the hydride transfer hypothesis required that the water had to be evacuated from the active site due to substrate binding. This QSAR documents the solvation–desolvation process present along the mechanistic steps in LADH.

Until now, most of the combined quantum/statistical mechanical calculations including realistic models of the surrounding medium have been carried out at a low level of basis set representation or, in most cases, with semiempirical MO approaches. This trend is being changed.

Weinstein and coworkers [55] have carried out ab initio calculations in the iterative computational scheme by using extended and minimal basis sets. The quantum motif, $NH_3 \ldots H^+ \ldots H_2S$, models the active site residues His162 and

Cys25 in the low pH form of actinidin, a sulfhydryl protease. The polarization energy contribution from the quantum motif turned out to be negligible in comparison to both the unperturbed electrostatic interaction energy and to the polarization of the macromolecule. Not surprisingly, the polarization energy remains nearly constant along the proton transfer path. The calculated polarization of the quantum motif by the collection of point charges representing the macromolecule V_m^0 is significantly increased by going from a minimal to a split valence basis set; however, addition of a polarization function to the latter does not produce any significant improvements in the calculation of the polarization energy. For this particular case, minimal and extended basis sets produce similar results for the electrostatic interactions between the unperturbed systems [55].

In our group, the GSCRF theory [16] has been implemented at an ab initio level [142]. Water-in-water calculations have probed the quality of diverse basis sets. The electrostatic potential $V_m^0(r)$ is represented with sources having a permanent dipole moment of 2.4 D, and spherical polarizability of 1.24 $Å^3$ situated at the oxygen atom and zero for the hydrogen atoms. Individual water molecules with their respective environments included in a sphere of radius 15 Å were selected from an MC sample. Such a sample comprises about 460 molecules around the central water molecule. The permanent and reaction potentials, fields and induced dipole–induced-dipole interactions were calculated with a cutoff of 15 Å. Statistical analysis was done on 60 different samples. STO-3G, 4-31G, 4-31G** and 6-31G** basis sets were used. Fluctuations of Mulliken populations, dipole moments, permanent solvent potentials, permanent electric fields, reaction potential and reaction field components were reported for these basis sets [142].

In the usual homogeneous reaction field model (e.g. in the case of cavity models), the reaction field energy is always stabilizing (negative). It is remarkable that in the present non-homogeneous reaction field model, one can also have a destabilizing contribution of the reaction field. This can be explained if one considers that the non-homogeneous distribution of the surrounding water molecules around a particular target may give rise to negative reaction potentials on a negatively charged oxygen atom [142].

The solvent field affects considerably the charge distribution of the water molecule. This is illustrated in fig. 3, where the dipole moment distribution is depicted. Overpolarization of the molecular charge distribution (even for the isolated molecules) is shown by the dipole moments obtained for the 4-31G, 4-31G** and the 6-31G** basis sets. The effective charges at the oxygen and hydrogen atoms in some cases also present extreme fluctuations [142].

Experimentally, the dipole moment of the water molecule in vacuo is increased by about 33% when it is immersed in liquid water at 25 °C. The low quality of the minimal STO-3G basis set is clearly displayed in the sense that only 10% increase is obtained. Pictorially speaking, the STO-3G charge density appears stiff with respect to its response against external fields, i.e. it has a small polarizability. The 4-31G result is a 24% increase, the distribution peaks at 3.04 D with a large population

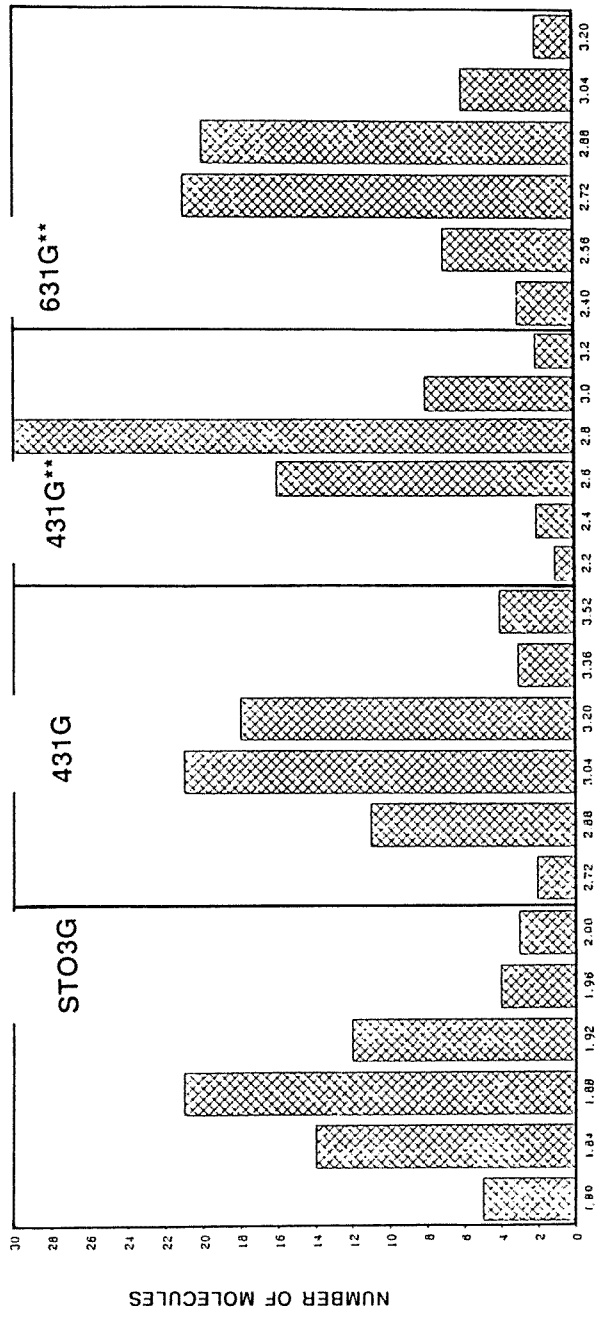

Fig. 3. Histograms representing the dipole moments in Debyes calculated with the set of basis sets indicated in the insets. (Figure prepared by F. Colonna.)

corresponding to 3.2 D; this result is unrealistic since it is the dipole moment of water in vacuo, namely 2.5 D. It was noted that both the 4-31G** and 6-31G** results showed distributions centered around 2.8 D and 2.72 D. Very few molecules have dipole moments larger than 3 D. As one would have expected, the polarization functions are essential to describe a more realistic coupling between the solute and the surrounding medium [142]. Actually, the increase in dipole moment was 29.5% and 27% for the 4-31G** and 6-31G** basis sets, respectively. This is in fairly good agreement with experimental results, except for the poor in vacuo dipole moments. Thus, insofar as the theory of solvent effects is concerned, the numerical testing can be considered as satisfactory.

5. Conclusions and perspectives

The theory of solvent effects and some of its applications have been overviewed. The generalized self-consistent reaction field framework has been used to present a unified approach to the theory underlying the quantum chemical calculations of subsystems embedded in a given environment. The treatment of the configurational space $X(t)$ for a subsystem coupled to a given surrounding was done in the framework of the statistical mechanical theory of projected equation of motions. This theory underlies applications of molecular dynamics simulations to the study of solvent and thermal bath effects on carefully defined subsystems of interest. Relationships have been established between different approaches used so far to calculate solvent effects and the general approach advocated by this reviewer. Applications to molecular properties in a time-independent framework have been presented.

The embedding of a quantum subsystem into a classical one is not yet a fully solved problem. In a semiempirical context, pseudo atoms can be introduced at boundary atoms. This procedure is advocated by Ostlund in his graphics-driven approach presented at the WATOC-II Symposium–1990. In ab initio MO contexts, pseudo potential representations of the boundary atoms have been evoked by Weinstein's group and in our group by Ángyán. To which extent the sigma core of the remaining protein may affect the electronic events at the active site is a matter of debate.

The applications described in this paper have been selected to illustrate the power of the solvent-effects approach to biochemical and physicochemical problems. The techniques are also applied in other fields. For example, molecular dynamics simulations are being used to study solvent effects on static and dynamic properties of linear and star polymers [143, 144]. A generalized Langevin equation has been used to represent a solvent around a polymer with stochastic forces; the solution of this equation agrees well with corresponding MD simulations [145]. A molecular dynamics study of one methane molecule in a cavity of NaA zeolite has recently been reported [146], as well as of Na ions in an A-type zeolite framework [147]; simulations of alkali-silicate glasses have been made in an effort to identify compositional fluctuations which could act as precursors of phase separation [148]. Finally,

a molecular dynamics simulation of a model reverse micelle in an apolar solvent was reported by Brown and Clarke [149]. Intermolecular interaction calculations with the direct reaction field method have recently been reported [150].

There has been progress in the study of solvent effects on chemical reactions in liquids and reactivity in solids such as zeolites and enzymes. For carbonic anhydrase, recent ab initio SCF-MO calculations including the zinc cation and a model for the coordination shell in the enzyme have yielded a detailed molecular mechanism [151].

The full statistical mechanical approach to solvent effects on dynamic processes epitomized by eq. (83) provides the starting point to build up model systems and develop non-equilibrium statistical mechanical approaches. Non-equilibrium solvent effects are now the subject of intensive experimental investigations. Different blends of time-resolved spectroscopies are contributing to the knowledge of dynamical processes in liquids, solutions and bioenvironments [152,153]. The electronic properties of the solute system become affected by fluctuations of the solvent configuration $X(t)$. Subpicosecond phenomena are involved in the response of the solvent to a newly created charge or dipole. The reciprocal effects can be followed theoretically by implementing, for instance, the generalized SCRF scheme, as shown in ref. [142]. Such developments have been anticipated by Warshel and coworkers in their studies with the EVB approach. Banacky and Zajac [154,155] have derived the theory of particle dynamics in solvated molecular complexes. A time-dependent nonlinear equation of motion for the probability density of a proton in a solvated symmetric H-bond system was derived. Earlier work has been overviewed by the present author [10]. Simple models with complex formalisms have been used in theoretical studies of chemical dynamics [11,156–162].

Brute force MD and MC simulations of solvent effects of the dynamics and static properties of proteins summarized in this work will serve as benchmark calculations to gauge model representations of solvent effects on biomacromolecules. The non-specialist reader should be warned of the limitations of present-day computer simulation methods. They do not give a complete answer to solvent effects. Sampling difficulties are always there. The application to proteins is mostly reduced to the study of fluctuations around a native state conformation(s). The hypothesis of broken ergodicity [163] is required for these methods to work. The problem of multiple minima is still an unsolved problem. The use of realistic dielectric models can also be seen as a complementary approach to represent solvent effects on biomolecules. Thus, in spite of all limitations, the conditions are mature for including more sophisticated ab initio studies into the description of time-dependent physicochemical phenomena.

Acknowledgements

The author would like to thank NFR for financial support. J.G. Ángyán and F. Colonna have decisively contributed in developing the field of solvent effects; the author is most grateful to them for valuable discussions and partnership.

References

[1] E.S. Amis and F.H. Hinton, *Solvent Effects on Chemical Phenomena* (Academic Press, New York, 1973).

[2] N. Tanaka, H. Ohtaki and R. Tamamushi (eds.), *Ions and Molecules in Solution* (Elsevier, Amsterdam, 1983).

[3] H.L. Strauss, G.T. Babcock and C.B. Moore (eds.), *Annual Review of Physical Chemistry*, Vol. 37 (Ann. Rev. Inc., Palo Alto, CA, 1986).

[4] R.R. Dogonadze, E. Kalman, A.A. Kornyshev and J. Ulstrup (eds.), *The Chemical Physics of Solvation. Part B: Spectroscopy of Solvation* (Elsevier, Amsterdam, 1986).

[5] R.R. Dogonadze, E. Kalman, A.A. Kornyshev and J. Ulstrup (eds.), *The Chemical Physics of Solvation. Part A: Theory of Solvation* (Elsevier, Amstersdam, 1985).

[6] O. Tapia, in: *Quantum Theory of Chemical Reactivity*, Vol. 2, ed. R. Daudel, A. Pullman, L. Salem and A. Veillard (Reidel, Dordrecht, 1980).

[7] O. Tapia, in: *Molecular Interactions*, Vol. 3, ed. H. Ratajczak and W.J. Orville-Thomas (Wiley, Chichester, 1982), p. 47.

[8] D.A. McQuires, *Statistical Mechanics* (Harper and Row, New York, 1976).

[9] R. Serra, M. Andretta and M. Compiani, *Physics of Complex Systems* (Pergamon, Oxford/New York, 1986).

[10] O. Tapia, in: *Molecules in Physics, Chemistry and Biology*, Vol. 3(Kluwer Academic, Holland, 1988), p. 405.

[11] J.T. Hynes, in: *The Theory of Reactions in Solution*, Vol. 4, ed. M. Baer (CRC Press, Boca Raton, FL, 1985), p. 171.

[12] A. Szabó and N.S. Ostlund, *Modern Quantum Chemistry* (MacMillan, New York, 1982).

[13] R. McWeeny and B. Pickup, Rep. Prog. Phys. 43(1980)1065.

[14] D.P. Craig and T. Thirunamachandran, *Molecular Quantum Electrodynamics* (Academic Press, London, 1984).

[15] A.D. Buckingham, Adv. Chem. Phys. 12(1967)107.

[16] O. Tapia, J. Mol. Struct. (THEOCHEM) 226(1991)59.

[17] J.D. Jackson, *Classical Electrodynamics*, 2nd ed. (Wiley, New York, 1962).

[18] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan and M. Karplus, J. Comp. Chem. 4(1983)187.

[19] W.F. van Gunsteren, H.J.C. Berendsen, J. Hermans, W.G.J. Hol and J.P.M. POstma, Proc. Nat. Acad. Sci. USA 80(1983)4315.

[20] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren and J. Hermans, in: *Intermolecular Forces*, ed. B. Pullman (Reidel, Dordrecht, 1981), p. 331.

[21] W.L. Jorgensen, J. Chem. Phys. 77(1982)4156.

[22] O. Matsuoka, E. Clementi and M. Yoshimine, J. Chem. Phys. 64(1976)1351.

[23] S.L. Price, Mol. Simul. 1(1988)135.

[24] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta, Jr. and P. Weiner, J. Amer. Chem. Soc. 106(1984)765.

[25] O. Tapia and G. Johannin, J. Chem. Phys. 75(1981)3624.

[26] R.P. Feynman, *Statistical Mechanics* (Benjamin, Reading, MA, 1972).

[27] L.M. Sese, V. Botella and P.C. Gomez, J. Mol. Liquids 32(1986)259.

[28] J. Ángyán and B. Silvi, J. Chem. Phys. 86(1987)6957.

[29] J. Ángyán, M. Allavena, M. Picard, A. Potier and O. Tapia, J. Chem. Phys. 77(1982)4723.

[30] M. Karelson, T. Tamm, A.R. Katritzky, M. Szafran and M.-C. Zerner, Int. J. Quant. Chem. 37(1990)1.

[31] E.S. Marcos, B. Terryn and J.L. Rivail, J. Phys. Chem. 89(1985)4695; for an earlier reference, see J.L. Rivail and D. Rinaldi, Chem. Phys. 18(1976)233.

[32] R. Contreras and A. Aizman, Int. J. Quant. Chem. Symp. 20(1986)573.
[33] D. Rinaldi, M.F. Ruiz-López, M.T.C. Costa and J.L. Rivail, Chem. Phys. Lett. 128(1986)177.
[34] M.F. Ruiz-López, D. Rinaldi and J.L. Rivail, Chem. Phys. 110(1986)403.
[35] K.V. Mikkelsen and M.A. Ratner, Int. J. Quant. Chem. S21(1987)341.
[36] H. Ågren, C. Medina-Llanos and K.V. Mikkelsen, Chem. Phys. 115(1987)43.
[37] G. Karlström, J. Phys. Chem. 92(1987)1315.
[38] M. Karelson, A.R. Katritzky, M. Szafran and M.-C. Zerner, J. Chem. Soc. Perkin 2(1990)195.
[39] O. Tapia, R. Cardenas, J. Andres, J. Krechl, M. Campillo and F. Colonna, Int. J. Quant. Chem. (1990), in press.
[40] K.V. Mikkelsen, J. Ågren, H.J.A. Jensen and K.V. Helgaker, J. Chem. Phys. 89(1988)3086.
[41] H. Hoshi, M. Sakurai, Y. Inoue and R. Chujo, J. Chem. Phys. 87(1987)1107.
[42] F.M.L. Stamato and O. Tapia, Int. J. Quant. Chem. 33(1988)187.
[43] R. Cobstanciel, Theor. Chim. Acta 69(1986)505.
[44] J.S. Gomez-Jeria and R. Contreras, Int. J. Quant. Chem. 30(1986)581.
[45] J. Langlet, P. Claverie and A. Pullman, J. Phys. Chem. 92(1988)1617.
[46] J.I. Gersten and A.M. Sapse, J. Amer. Chem. Soc. 107(1985)3786.
[47] A.A. Rashin and K. Namboodiri, J. Phys. Chem. 91(1987)6003.
[48] J.A.C. Rullman and P.Th. van Duijnen, Mol. Phys. 61(1987)293.
[49] R. Constanciel and O. Tapia, Theor. Chim. Acta (Berl.) 48(1978)75.
[50] C. Lamborelle and O. Tapia, Chem. Phys. 42(1979)25.
[51] S. Miertus, E. Scrocco and J. Tomasi, Chem. Phys. 55(1981)117.
[52] G. Alagona, C. Ghio, J. Igual and J. Tomasi, J. Amer. Chem. Soc. 111(1989)3417.
[53] J.L. Pascual-Ahuir and E. Silla, in: *Quantum Chemistry: Basic Aspects, Actual Trends*, ed. R. Carbó (Elsevier, Amsterdam, 1989), p. 597.
[54] J.G.Ángyán and G. Náray-Szabó, in: *Theoretical Models of Chemical Bonding*, Part 4, ed. Maksic (Springer, Berlin, 1990).
[55] G.A. Mercier, Jr., J.P. Dijkman, R. Osman and H. Weinstein, in: *Quantum Chemistry: Basic Aspects, Actual Trends*, ed. R. Carbó (Elsevier, Amsterdam, 1989), p. 577.
[56] M. Morse and S.A. Rice, J. Chem. Phys. 74(1981)6514.
[57] M. Morse and S.A. Rice, J. Chem. Phys. 76(1982)6514.
[58] M. Townsend, M. Morse and S.A. Rice, J. Chem. Phys. 79(1983)2496.
[59] G. Nielson and S.A. Rice, J. Chem. Phys. 80(1984)4456.
[60] G. Ravishanker, P.K. Mehrotra, M. Mezei and D.L. Beveridge, J. Amer. Chem. Soc. 106(1984)4102.
[61] E. Clementi and G. Corongiu, Int. J. Quant. Chem. 10(1983)31.
[62] J. Detrich, G. Corongiu and E. Clementi, Chem. Phys. Lett. 112(1984)426.
[63] M. Neumann, J. Chem. Phys. 82(1985)5663.
[64] F. Colonna, J.G. Ángyán and O. Tapia, Chem. Phys. Lett. (1990), in press.
[65] F. Vigne-Maeder and P. Claverie, J. Chem. Phys. 88(1988)4934.
[66] A.J. Stone, Chem. Phys. Lett. 83(1981)233.
[67] A.J. Stone and M. Alderton, Mol. Phys. 56(1985)1047.
[68] S.L. Price, R.J. Harrison and M.F. Guest, J. Comp. Chem. 10(1989)552.
[69] V. Magnasco and G. Figari, Mol. Phys. 67(1989)1261.
[70] A. Goldblum, D. Perahia and A. Pullman, Int. J. Quant. Chem. 15(1979)1.
[71] A. Wallqvist, P. Åhlström and G. Karlström, J. Phys. Chem. 94(1990)1649.
[72] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey and M.L. Klein, J. Chem. Phys. 79(1983)927.
[73] F.H. Stillinger and A: Rahman, J. Chem. Phys. 60(1974)1545.
[74] P.D. Dacre, Mol. Phys. 50(1983)1132.
[75] U.C. Singh and P.A. Kollman, J. Comp. Chem. 5(1984)129.
[76] E. Clementi, G. Corongiu, G.C. Lie, U. Niesar and P. Procacci, MOTECC, ed. E. Clementi (1989), Ch. 8, p. 363.

[77] A. Warshel and S.T. Russel, Quart. Rev. Biophys. 17(1984)283.
[78] J.G. Ángyán, F. Colonna and O. Tapia, Chem. Phys. Lett. 166(1990)180.
[79] S.A. Adelman, J. Chem. Phys. 73(1980)3145.
[80] M.W. Balk, C.A. Brooks and S.A. Adelman, J. Chem. Phys. 79(1983)804.
[81] M. Berkowitz and J.A. McCammon, Chem. Phys. Lett. 90(1982)215.
[82] C.L. Brooks and M. Karplus, J. Chem. Phys. 79(1983)6312.
[83] A. Brunger, C.L. Brooks and M. Karplus, Chem. Phys. Lett. 105(1984)495.
[84] A. Warshel, J. Phys. Chem. 102(1979)6218.
[85] C.L. Brooks, A. Brunger and M. Karplus, Biopolymers 24(1985)843.
[86] A. Brunger, C.L. Brooks and M. Karplus, Proc. Nat. Acad. Sci. USA 82(1985)8458.
[87] W.F. van Gunsteren and H.J.C. Berendsen, in: *Molecular Dynamics and Protein Structure*, ed. J. Hermans (University of North Carolina Printing Deptartment, Chapel Hill, 1985), p. 5.
[88] A.D. Buckingham, in: *Molecular Interactions: From Diatomics to Biopolymers*, ed. B. Pullman (Wiley, New York, 1978), Ch. 1.
[89] A.D. Buckingham, in: *Structure and Motion: Membranes, Nucleic Acids and Proteins*, ed. E. Clementi, G. Corongiu, M.H. Sarma and R.H. Sarma (Adenine Press, Guilderland, 1985).
[90] W.LK. Jorgensen, J. Phys. Chem. 87(1983)5304.
[91] P. Linse, G. Karlström and B. Jönsson, J. Amer. Chem. Soc. 106(1984)4096.
[92] M.E. Cournoyer and W.L. Jorgensen, J. Amer. Chem. Soc. 106(1984)5104.
[93] E. Clementi, F. Cavallone and R. Scordamaglia, J. Amer. Chem. Soc. 99(1977)5531.
[94] E. Clementi, G. Corongiu and G. Ranghino, J. Chem. Phys. 74(1981)578.
[95] E. Clementi and G. Corongiu, in: *Studies in Theoretical Chemstry*, Vol. 27, ed. N. Tanaka, H. Ohtaki and R. Tamamushi (Elsevier, Amsterdam, 1982), p. 397.
[96] E. Clementi, G. Corongiu, J.H. Dietrich, H. Khanmohammadbaigi, S. Chin, L. Domingo, A. Laaksonen and H.L. Nguyen, in: *Structure and Dynamics: Nucleic Acids and Proteins*, ed. E. Clementi, G. Corongiu, M.H. Sarma and R.H. Sarma (Adenine Press, New York, 1985), p. 49.
[97] D.L. Beveridge, M. Mezei, G. Ravishanker and B. Jayaram, Int. J. Quant. Chem. 29(1986)1513.
[98] W.L. Jorgensen and J.K. Buckner, J. Phys. Chem. 91(1987)6083.
[99] A. McCammon and S.C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, Cambridge/New York, 1987). (Initial(s) for 2nd author?)
[100] W.F. van Gunsteren, Mol. Simul. 3(1989)187.
[101] J.A. McCammon, Science 238(1987)486.
[102] P.A. Bash, U.Ch. Singh, F.K. Brown, R. Langridge and P.A. Kollman, Science 235(1987)574.
[103] S. Hirono and P.A. Kollman, J. Mol. Biol. 212(1990)197.
[104] P. Ahlström, O. Teleman, B. Jönsson and S. Forsen, J. Amer. Chem. Soc. 109(1987)1541.
[105] W.F. van Gunsteren and M. Karplus, Biochem. 21(1982)2259.
[106] S. Swaminathan, T. Ichiye, W.F. van Gunsteren and M. Karplus, Biochem. 21(1982)5230.
[107] W.F. van Gunsteren and H.J.C. Berendsen, J. Mol. Biol. 176(1984)559.
[108] M. Levitt and R. Sharon, in: *Crystallography in Moleculat Biology*, ed. D. Moras, J. Drenth, B. Strandberg, S. Dietrich and K. Wilson (Plenum, New York, 1987), p. 197.
[109] M. Levitt and R. Sharon, Proc. Nat. Acad. Sci. USA 85(1988)7557.
[110] P. Kruger, W. Strassburger, A. Wollmer and W.F. van Gunsteren, Eur. Biophys. J. 13(1985)77.
[111] C.J. Edge, U.C. Singh, R. Bazzo, G.L. Taylor, R.A. Dwek and T.W. Rademacher, Biochem. 29(1990)1971.
[112] J. Tirado-Rives and W.L. Jorgensen, J. Amer. Chem. Soc. 112(1990)2773.
[113] J. Åqvist, W.F. cvan Gunsteren, M. Leijonmark and O. Tapia, J. Mol. Biol. 183(1985)593.
[114] J. Åqvist, M. Leijonmark and O. Tapia, Eur. Biophys. J. 16(1989)327.
[115] J. Åqvist and O. Tapia, Biopolymers 30(1990)205.
[116] O. Tapia, O. Nilsson, M. Campillo, J. Åqvist and E. Horjales, in: *Structure and Methods: DNA Protein Complexes and Proteins*, Vol. 2 (1990), p. 147.

[117] J. Chandrasekhar, S. Smith and W.L. Jorgensen, J. Amer. Chem. Soc. 106(1984)3049.
[118] J. Chandrasekhar, S. Smith and W.L. Jorgensen, J. Amer. Chem. Soc. 107(1985)154.
[119] J.D. Madura and W.L. Jorgensen, J. Amer. Chem. Soc. 108(1983)2517.
[120] S.E. Huston, P.J. Rossky and D.A. Zichi, J. Amer. Chem. Soc. 111(1989)5680.
[121] J.P. Bergsma, B.J. Gertner, K.R. Wilson and J.T. Hynes, J. Chem. Phys. 86(1987)1356.
[122] O. Tapia and J.M. Lluch, J. Chem. Phys. 83(1985)3970.
[123] J. Andres, R. Cardenas, E. Silla and O. Tapia, J. Amer. Chem. Soc. 110(1988)666.
[124] O. Tapia, J.M. Lluch, R. Cardenas and J. Andres, J. Amer. Chem. Soc. 111(1989)829.
[125] J.-K. Hwang, G. King, S. Creighton and A. Warshel, J. Amer. Chem. Soc. 110(1988)5297.
[126] J.-K. Hwang, S. Creighton, G. King, D. Whitney and A. Warshel, J. Chem. Phys. 89(1988)859.
[127] A. Warshel, F. Sussman and J.-K. Hwang, J. Mol. Biol. 201(1988)139.
[128] J. Åqvist and A. Warshel, J. Amer. Chem. Soc. 112(1990)2860.
[129] J. Åqvist and A. Warshel, Biochem. 28(1989)4680.
[130] R.L. Schowen, in: *Principles of Enzyme Activity*, Vol. 9, ed. J.F. Liebman and A. Greenberg (VCH Publ., FL, 1988).
[131] F.M.L.G. Stamato, E. Longo, R. Ferreira and O. Tapia, J. Theor. Biol. 118(1986)45.
[132] G. Náray-Szabó, J. Mol. Catal. 47(1988)281.
[133] G. Náray-Szabó and A. Warshel, in: *Highlights of Modern Biochemistry*, ed. A. Kotyk, J. Skoda, V. Paces and V. Kostka (VSP Int. Sci. Publ., Zeist, 1989), pp. 169–176.
[134] K. Gráf, A. Jancsó, L. Szilágyi, G. Hegyi, K. Pintér, G. Náray-Szabó, J. Heoo, K. Medzihradszky and W. Rutter, Proc. Nat. Acad. Sci. USA 85(198)4961.
[135] O. Tapia, J. Mol. Catal. 43(1988)199.
[136] O. Tapia, J. Andres, J.M. Aullo and R. Cardenas, J. Mol. Struct. 167(1988)395.
[137] O. Tapia, R. Cardenas, J. Andres and F. Colonna-Cesari, J. Amer. Chem. Soc. 110(1988)4046.
[138] G. Náray-Szabó and K. Simon (eds.), *Steric Aspects of Biomolecular Interactions* (CRC Press, Boca Raton, FL, 1987).
[139] O. Tapia and H. Eklund, Enzyme 36(1986)101.
[140] O. Tapia, H. Eklund and C.-I. Brändén, in: *Steric Aspects of Biomolecular Interactions*, ed. G. Náray-Szabó and K. Simon (CRC Press, Boca Raton, FL, 1987).
[141] C. Hansch, T. Klein, J. McClarin, R. Landridge and N.W. Cornell, J. Med. Chem. 29(1986)615.
[142] O. Tapia, J.G. Ángyán and F. Colonna, J. Chim. Phys. 87(1990)875.
[143] B. Smit, A. van der Put, C.J. Peters and J. de Swaan Arons, J. Chem. Phys. 88(1988)3372.
[144] B. Smit, A. van der Put, C.J. Peters and J. de Swaan Arons, Chem. Phys. Lett. 144(1988)555.
[145] S. Toxvaerd, J. Chem. Phys. 86(1987)3667.
[146] E. Cohen De Lara, R. Kahn and A.M. Goulay, J. Chem. Phys. 90(1989)7482.
[147] J.M. Shin, K.T. No and M.S. Jhon, J. Phys. Chem. 92(1988)4533.
[148] A.A. Tesar and A.K. Varshneya, J. Chem. Phys. 87(1987)2986
[149] D. Brown and J.H.R. Clarke, J. Phys. Chem. 92(1988)2881.
[150] P.Th. van Duijnen and J.A.C. Rullman, Int. J. Quant. Chem. 37(1990)181.
[151] O. Jacob, R. Cardenas and O. Tapia, J. Amer. Chem. Soc. 112(1990)8692.
[152] G.R. Fleming, Ann. Rev. Phys. Chem. 37(1986)81.
[153] A.P. Demchencko, Ess. Biochem. 22(1986)120.
[154] P. Banacky and A. Zajac, Chem. Phys. 123(1988)267.
[155] P. Banacky and A. Zajac, Int. J. Quant. Chem. 37(1990)209.
[156] B.M. Ladanyi and J.T. Hynes, J. Amer. Chem. Soc. 108(1986)585.
[157] B.J. Gertner, J.P. Bergsma, K.R. Wilson, S. Lee and J.T. Hynes, J. Chem. Phys. 86(1987)1377.
[158] D. Borgis and J.T. Hynes, in: *The Enzyme Catalysis Process*, NATO ASI Serie (Plenum, New York, 1989).
[159] B.J. Gertner, K.R. Wilson and J.T. Hynes, J. Chem. Phys. 90(1989)3537.
[160] S.-B. Zhu, J. Lee, G.W. Robinson and S.H. Lin, J. Chem. Phys. 90(1989)6335.
[161] S.-B. Zhu, J. Lee, G.W. Robinson and S.H. Lin, J. Chem. Phys. 90(1989)6340.
[162] G.W. Robinson, S. Singh, R.S. Krishnan, S.-B. Zhu and J. Lee, J. Phys. Chem. 94(1990)4.
[163] R.G. Palmer, Adv. Phys. 31(1982)669.