IMPLICIT SOLVENT MODELS*

Benoît Roux†
Department of Biochemistry
Weill Medical College of Cornell University
1300 York Avenue, New-York, NY 10021

Contents

1 Introduction **1**

2 Basic formulation of implicit solvent **2**
2.1 The potential of mean force 2
2.2 Reversible work 3

3 Decomposition of the free energy **4**
3.1 Non-polar free energy contribution 4
3.2 Electrostatic free energy contribution 5

4 Classical continuum electrostatics. **5**
4.1 Poisson equation for macroscopic media . . . 5
4.2 Electrostatic forces and analytic gradients . . 6
4.3 Treatment of ionic strength 7
4.4 Treatment of a transmembrane potential . . . 7

5 Miscellaneous approaches **8**
5.1 Statistical mechanical integral equations . . . 8
5.2 Solvent boundary potentials and implicit/explicit mixed schemes 8
5.3 Solvent accessible surface area models 9
5.4 Knowledge-based potentials 10

6 Summary **10**

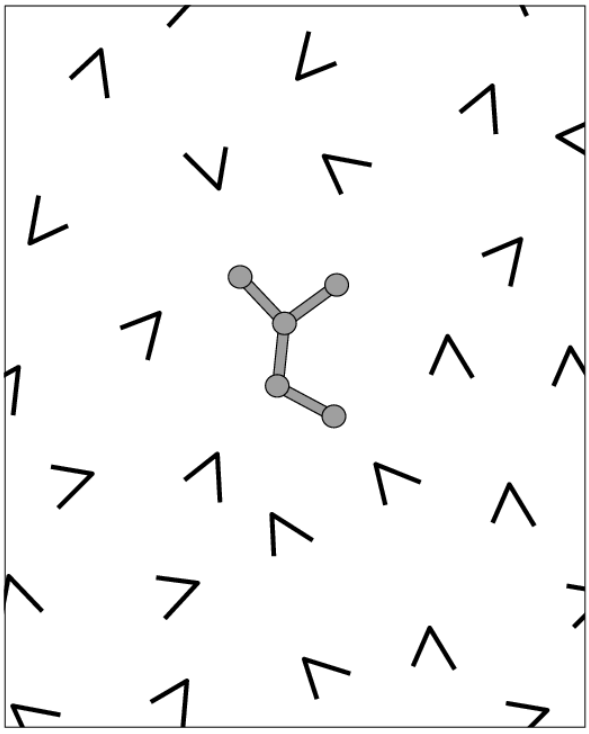

Figure 1: Schematic representation of an atomic model of a biomolecular solute surrounded by explicit water molecules.

1 Introduction

An understanding of a wide variety of phenomena concerning conformational stabilities and molecule-molecule association (protein-protein, protein-ligand and protein-nucleic acids) require the consideration of solvation effects. In particular, a quantitative assessment of the relative contribution of hydrophobic and electrostatic interactions in macromolecular recognition is a problem of central importance in biology. The complexity of the environement in which biomolecules must perform their functions is such that information extracted from simple theoretical models may be helpful to further our understanding of these systems.

There is no doubt that molecular dynamics simulations in which a large number of solvent molecules are treated explicitly represent one of the most detailed approaches to study the influence of solvation on complex biomolecules [1]. The approach, which is illustrated schematically in Figure 1, consists in constructing detailed atomic models of the solvated macromolecular system and, having described the microscopic forces with a potential function, applying Newton's classical equation, $F = MA$, to literally "simulate" the dynamical motions of all the atoms as a function of time [1, 2]. The calculated classical trajectory, though an approximation to the real world, provides ultimate detailed information about the time course of the atomic motions, which is difficult to access experimentally. However, statistical convergence is an important issue because the net influence of solvation results from an averaging over a large number of configurations. In addition, a large number of solvent molecules is required to realistically model a dense system. Thus, in practical situations a significant fraction of the computer time is used to calculate the detailed tra-

*In *Computational Biochemistry & Biopysics*, Marcel Dekker, N.Y. Eds. O. Becker, A.D. MacKerell Jr., B. Roux and M. Watanabe

†Email address: benoit.roux@med.cornell.edu

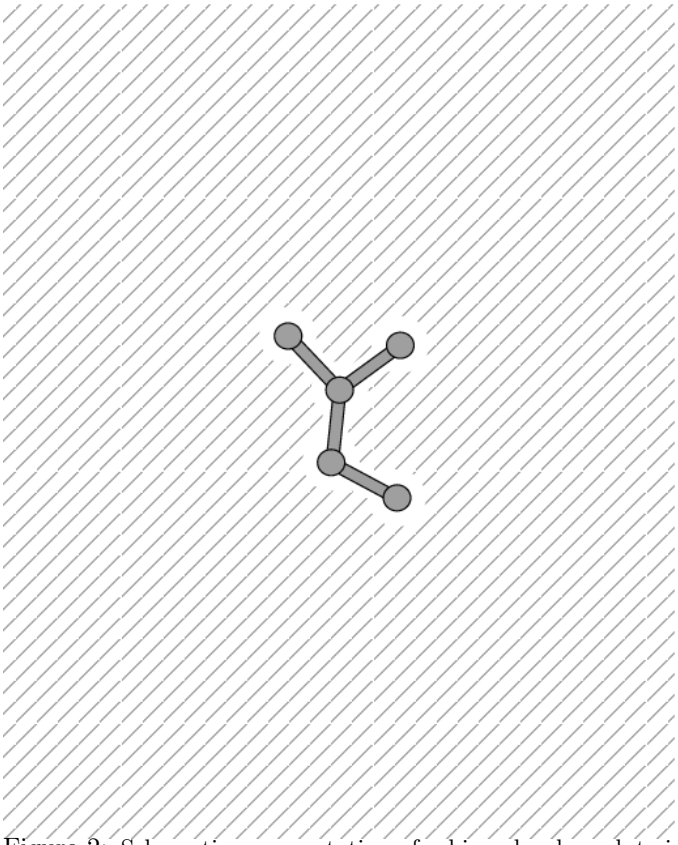

Figure 2: Schematic representation of a biomolecular solute in a solvent environment which is taken into account implicitly.

jectory of the solvent molecules, even though it is often the solute that is of interest.

An alternative approach, illustrated schematically in Figure 2, consists in incorporating the influence of the solvent implicitly. Such approximate schemes can provide useful quantitative estimates of solvation free energies while remaining computationally tractable. Implicit solvent approaches avoid the statistical errors associated with averages extracted from simulations with a large number of solvent molecules. Furthermore, implicit solvent models are sometimes better suited for particularly complex situations. For example, an explicit representation of the cellular membrane potential would require prohibitively large atomic simulation systems and is currently impractical. Finally, implicit solvent representations can be very useful conceptual tools for analyzing the results of simulations generated with explicit solvent molecules and to better understand the nature of solvation phenomena in general.

In this chapter we provide an introductory overview of the implicit solvent models commonly used in biomolecular simulations. A number of questions concerning the formulation and the development of implicit solvent models are addressed. In section 2, we begin by providing a rigorous formulation of implicit solvent from statistical mechanics. In addition, the fundamental concept of the potential of mean force (PMF) is also introduced. In section 3, a decomposition of the PMF in terms of non-polar and electrostatic contributions is elaborated. Owing to its importance in biophysics, section 4 is entirely devoted to classical continuum electrostatics. For the sake of completeness other computational schemes such as statistical mechanical integral equations, implicit/explicit solvent boundary potential, solvent accessible surface area (SASA), and knowledge-based potentials are briefly reviewed in section 5. Finally, the chapter is concluded in section 6 by a short summary of the principal ideas.

2 Basic formulation of implicit solvent

2.1 The potential of mean force

As a first step, it is important to establish implicit solvent models on first principles. For the sake of concreteness, let us consider a solute "u" immersed in a bulk solution "v". The configuration of the solute is represented by the vector $\mathbf{X} \equiv \{\mathbf{x}_1, \mathbf{x}_2, \ldots\}$. All other degrees of freedom of the bulk solution surrounding the solute, which may include solvent molecules as well as mobile counterions, are represented by the vector $\mathbf{Y}$. It is expected that the system is fluctuating over a large number of configurations. It is therefore necessary to consider the problem from a statistical point of view. For a system in equilibrium with a thermal bath at temperature T, the probability of a given configuration $(\mathbf{X}, \mathbf{Y})$ is given by the function $P(\mathbf{X}, \mathbf{Y})$ [3],

$$P(\mathbf{X}, \mathbf{Y}) = \frac{e^{-U(\mathbf{X},\mathbf{Y})/k_{\mathrm{B}}T}}{\int d\mathbf{X} d\mathbf{Y}\ e^{-U(\mathbf{X},\mathbf{Y})/k_{\mathrm{B}}T}} \tag{1}$$

where $U(\mathbf{X}, \mathbf{Y})$ is the total potential energy of the system. For the sake of simplicity, we neglect nonadditive interactions and assume that the total potential energy can be written as

$$U(\mathbf{X}, \mathbf{Y}) = U_{\mathrm{u}}(\mathbf{X}) + U_{\mathrm{vv}}(\mathbf{Y}) + U_{\mathrm{uv}}(\mathbf{X}, \mathbf{Y}) \tag{2}$$

where $U_{\mathrm{u}}(\mathbf{X})$ is the intramolecular potential of the solute, $U_{\mathrm{vv}}(\mathbf{Y})$ is the solvent-solvent potential, and $U_{\mathrm{uv}}(\mathbf{X}, \mathbf{Y})$ is the solute-solvent potential. All observable properties of the system are fundamentally related to averages weighted by the probability function $P(\mathbf{X}, \mathbf{Y})$. For example, the average of any quantity $Q(\mathbf{X})$ depending on the solute configuration is given by

$$\langle Q \rangle = \int d\mathbf{X} d\mathbf{Y}\ Q(\mathbf{X}) P(\mathbf{X}, \mathbf{Y}) \tag{3}$$

An important question is whether one can rigorously express such an average without referring explicitly to the solvent degrees of freedom. In other words: is it possible to avoid explicit reference to the solvent in the mathematical description of the molecular system and still obtain rigorously correct properties? The answer to this question is yes. A reduced probability distribution $\bar{P}(\mathbf{X})$ that depends only on the solute configuration can be defined as

$$\bar{P}(\mathbf{X}) = \int d\mathbf{Y}\ P(\mathbf{X},\mathbf{Y})\ . \tag{4}$$

The reduced probability distribution does not depend explicitly on the solvent coordinates $\mathbf{Y}$, although it incorporates the average influence of the solvent on the solute. The operation symbolized by Eq. (4) is commonly described by saying that the solvent coordinates $\mathbf{Y}$ have been "integrated out". In a system at temperature T, the reduced probability has the form

$$\begin{aligned} \bar{P}(\mathbf{X}) &= \frac{\int d\mathbf{Y}\ e^{-[U_{\rm u}(\mathbf{X})+U_{\rm vv}(\mathbf{Y})+U_{\rm uv}(\mathbf{X},\mathbf{Y})]/k_{\rm B}T}}{\int d\mathbf{X}d\mathbf{Y}\ e^{-[U_{\rm u}(\mathbf{X})+U_{\rm vv}(\mathbf{Y})+U_{\rm uv}(\mathbf{X},\mathbf{Y})]/k_{\rm B}T}} \\ &= \frac{e^{-W(\mathbf{X})/k_{\rm B}T}}{\int d\mathbf{X}\ e^{-W(\mathbf{X})/k_{\rm B}T}}\ . \end{aligned} \tag{5}$$

The function $W(\mathbf{X})$ is called the "Potential of Mean Force" (PMF). The fundamental concept of the PMF was first introduced by Kirkwood to describe the average structure of liquids [4]. It is a simple matter to show that the gradient of $W(\mathbf{X})$ in cartesian coordinates is related to the average force

$$\begin{aligned} \frac{\partial W(\mathbf{X})}{\partial \mathbf{x}_i} &= \left\langle \frac{\partial U}{\partial \mathbf{x}_i} \right\rangle_{(\mathbf{X})} \\ &\equiv -\langle \mathbf{F}_i \rangle_{(\mathbf{X})}\ , \end{aligned} \tag{6}$$

where the symbol $\langle \ldots \rangle_{(\mathbf{X})}$ represents an average over all coordinates of the solvent, with the solute in the fixed configuration specified by $\mathbf{X}$. All solvent effects are included in $W(\mathbf{X})$ and, consequently, in the reduced distribution function $\bar{P}(\mathbf{X})$. The PMF is an effective configuration-dependent free energy potential $W(\mathbf{X})$ making no explicit reference to the solvent degrees of freedom, such that no information about the influence of solvent on equilibrium properties is lost.

As long as the normalization condition given by Eq. (5) is satisfied, an arbitrary offset constant may be added to $W(\mathbf{X})$ without affecting averages in Eq. (3). The absolute value of the PMF is thus unimportant. For convenience, it is possible to choose the value of the free energy $W(\mathbf{X})$ relative to a reference system in which the solute-solvent interactions are absent. The free energy $W(\mathbf{X})$ may thus be expressed as

$$e^{-W(\mathbf{X})/k_{\rm B}T} = \frac{\int d\mathbf{Y}\ e^{-[U_{\rm u}(\mathbf{X})+U_{\rm vv}(\mathbf{Y})+U_{\rm uv}(\mathbf{X},\mathbf{Y})]/k_{\rm B}T}}{\int d\mathbf{Y}\ e^{-U_{\rm vv}(\mathbf{Y})/k_{\rm B}T}}\ . \tag{7}$$

It is customary to write that $W(\mathbf{X}) = U_{\rm u}(\mathbf{X}) + \Delta W(\mathbf{X})$, where $U_{\rm u}(\mathbf{X})$ is the intramolecular solute potential and $\Delta W(\mathbf{X})$ is the solvent-induced influence. In practice ΔW depends on $\mathbf{X}$, the configuration of the solute, as well as on thermodynamic variables such as the temperature T and the pressure p.

2.2 Reversible work

As shown by Eq. (6) the PMF is the reversible work done by the average force. It is possible to express relative values of the PMF between different solute configurations $\mathbf{X}_1$ and $\mathbf{X}_2$ using Eq. (6) and the reversible work theorem [4]

$$\begin{aligned} W(\mathbf{X}_2) &= W(\mathbf{X}_1) + \int_{\mathbf{X}_1}^{\mathbf{X}_2} \sum_i d\mathbf{x}_i \cdot \frac{\partial W(\mathbf{X})}{\partial \mathbf{x}_i} \\ &= W(\mathbf{X}_1) - \int_{\mathbf{X}_1}^{\mathbf{X}_2} \sum_i d\mathbf{x}_i \cdot \langle \mathbf{F}_i \rangle_{(\mathbf{X})}\ . \end{aligned} \tag{8}$$

This relation makes it clear that the PMF is not equal to an average potential energy because one needs to compute a reversible work againts an average force to get the $W(\mathbf{X})$. It is also possible to express the free energy in terms of a thermodynamic integral. Introducing the thermodynamic solute-solvent coupling parameter λ [4], we write the potential energy as,

$$U(\mathbf{X},\mathbf{Y};\lambda) = U_{\rm u}(\mathbf{X}) + U_{\rm vv}(\mathbf{Y}) + U_{\rm uv}(\mathbf{X},\mathbf{Y};\lambda) \tag{9}$$

constructed such that $\lambda = 0$ corresponds to a non-interacting reference system with $U_{\rm uv}(\mathbf{X},\mathbf{Y};0) = 0$, and $\lambda = 1$ corresponds to the fully interacting system. As long as the end-points are respected, any form of thermodynamic coupling is correct. Therefore, we have

$$\Delta W(\mathbf{X}) = \int_0^1 d\lambda\ \left\langle \frac{\partial U_{\rm uv}}{\partial \lambda} \right\rangle_{(\mathbf{X},\lambda)}\ , \tag{10}$$

where the symbol $\langle \ldots \rangle_{(\mathbf{X},\lambda)}$ represents an average over all coordinates of the solvent, for a solute in the fixed configuration $\mathbf{X}$ with thermodynamic coupling λ. It may be noted that $\partial U_{\rm uv}/\partial\lambda$ in Eq. (10) plays the role of a generalized thermodynamic force similar to that of $\partial U/\partial \mathbf{x}_i$ in Eq. (8).

3 Decomposition of the free energy

Intermolecular forces are dominated by short-range harsh repulsive interactions, arising from Pauli's exclusion principle, van der Waals attractive forces arising from quantum dispersion, and long-range electrostatic interactions, arising from the non-uniform charge distribution. It is convenient to express the potential energy $U_{\rm uv}(\mathbf{X},\mathbf{Y})$ as a sum of electrostatic contributions and the remaining non-polar (non-electrostatic) contributions,

$$U_{\rm uv}(\mathbf{X},\mathbf{Y}) = U_{\rm uv}^{\rm (np)}(\mathbf{X},\mathbf{Y}) + U_{\rm uv}^{\rm (elec)}(\mathbf{X},\mathbf{Y}) \; . \qquad (11)$$

Although such a representation of the microscopic non-bonded interactions does not follow directly from a quantum mechanical description of the Born-Oppenheimer energy surface, it is commonly used in most force fields for computer simulations of biomolecules (e.g. AMBER [5], CHARMM [6], OPLS [7]). The separation of the solute-solvent interactions in Eq. (11) is useful for decomposing the reversible work that defines the function $W(\mathbf{X})$. The total free energy of a solute in a fixed configuration $\mathbf{X}$ may be expressed rigorously as the reversible thermodynamic work needed to construct the system in a step-by-step process. In a first step, the non-polar solute-solvent interactions are switched "on" in the absence of any solute-solvent electrostatic interactions; in a second step, the solute-solvent electrostatic interactions are switched "on" in the presence of the solute-solvent non-polar interactions. The solute is kept in the fixed configuration $\mathbf{X}$ throughout the whole processes, and the intramolecular potential energy does not vary during this process. By construction, the total PMF is

$$W(\mathbf{X}) = U_{\rm u}(\mathbf{X}) + \Delta W^{\rm (np)}(\mathbf{X}) + \Delta W^{\rm (elec)}(\mathbf{X}) \qquad (12)$$

where the non-polar solvation contribution is

$$e^{-\Delta W^{\rm (np)}(\mathbf{X})/k_{\rm B}T} = \frac{\int d\mathbf{Y}\; e^{-[U_{\rm vv}(\mathbf{Y})+U_{\rm uv}^{\rm (np)}(\mathbf{X},\mathbf{Y})]/k_{\rm B}T}}{\int d\mathbf{Y}\; e^{-U_{\rm vv}(\mathbf{Y})/k_{\rm B}T}} \qquad (13)$$

and the electrostatic solvation contribution is

$$e^{-\Delta W^{\rm (elec)}(\mathbf{X})/k_{\rm B}T} = \frac{\int d\mathbf{Y}\; e^{-[U_{\rm vv}(\mathbf{Y})+U_{\rm uv}^{\rm (np)}(\mathbf{X},\mathbf{Y})+U_{\rm uv}^{\rm (elec)}(\mathbf{X},\mathbf{Y})]/k_{\rm B}T}}{\int d\mathbf{Y}\; e^{-[U_{\rm vv}(\mathbf{Y})+U_{\rm uv}^{\rm (np)}(\mathbf{X},\mathbf{Y})]/k_{\rm B}T}} . \qquad (14)$$

Combining Eqs. (12), (13) and (14) yields Eq. (7) directly. Although such a free energy decomposition is path-dependent [8], it provides a useful and rigorous framework for understanding the nature of solvation and for constructing suitable approximations to the non-polar and electrostatic free energy contributions.

In the following sections, we describe an implicit solvent model based on this free energy decomposition and which is widely used in biophysics. It consists in representing the non-polar free energy contributions on the basis of the solvent accessible surface area (SASA), a concept introduced by Lee and Richards [9], and the electrostatic free energy contribution on the basis of the Poisson-Boltzmann (PB) equation of macroscopic electrostatics, an idea which goes back to Born [10], Debye and Hückel [11], Kirkwood [12], and Onsager [13]. The combination of these two approximations forms the SASA/PB implicit solvent model. In the next section we analyze the microscopic significance of the non-polar and electrostatic free energy contributions and describe the SASA/PB implicit solvent model.

3.1 Non-polar free energy contribution

To clarify the significance of $\Delta W^{\rm (np)}$ let us first consider the special case of a non-polar molecule solvated in liquid water. We assume that the electrostatic free energy contribution is negligible. Typically, the solute-solvent van der Waals dispersion interactions are relatively weak and the non-polar free energy contribution is dominated by the reversible work needed to displace the solvent molecules to accomodate the short-range harsh repulsive solute-solvent interaction. For this reason, $\Delta W^{\rm (np)}$ is often refered to as the "free energy of cavity formation". The reversible thermodynamic work corresponding to this process is positive and unfavorable. It gives rise to two aspects of the hydrophobic effect: hydrophobic solvation and hydrophobic interaction [14]. The former phenomena is responsible for the poor solubility of non-polar molecules in water; the latter accounts for the propensity of non-polar molecules to cluster and form aggregates in water.

Modern understanding of the hydrophobic effect attributes it primarily to a decrease of hydrogen bonds that can be achieved by the water molecules when they are near a non-polar surface. This view is confirmed by computer simulations of non-polar solutes in water [15]. To a first approximation, the magnitude of the free energy associated with the non-polar contribution can thus be considered to be proportional to the number of solvent molecule in the first solvation shell. This idea leads to a convenient and attractive approximation which is extensively used in biophysical applications [9, 16, 17, 18]. It consists in assuming that the non-polar free energy contribution is directly related to the SASA [9],

$$\Delta W^{\rm (np)}(\mathbf{X}) = \gamma_{\rm v}\mathcal{A}_{\rm tot}(\mathbf{X}) \qquad (15)$$

where $\gamma_{\rm v}$ has the dimension of a surface tension and $\mathcal{A}_{\rm tot}(\mathbf{X})$ is the configuration-dependent SASA (note that both po-

lar and non-polar chemical groups must be included in the SASA for a correct estimate of $\Delta W^{(\mathrm{np})}$). As pointed out by Tanford [19], there should be a close relationship between the macroscopic oil-water surface tension, interfacial free energies, and the magnitude of the hydrophobic effect. However, in practical applications, the surface tension γ_{v} is usually adjusted empirically to reproduce the solvation free energy of alkane molecules in water [18]. Its value is typically around 20 to 30 cal/mol/Å^2, while the macroscopic oil-water surface tension is around 70 cal/mol/Å^2 [19]. The difference between the optimal parameter γ_{v} for alkanes and the true macroscopic surface tension for oil-water interfaces reflects the influence of the microscopic length scale and the crudeness of the SASA model. A simple statistical mechanical approach describing the free energy of inserting hard spheres in water called scaled particle theory (SPT) provides an important conceptual basis for understanding some of the limitations of SASA models [20, 21, 22]. It is clear that the SASA does not provide an ultimate representation of the non-polar contribution to the solvation free energy. Other theories based on cavity distributions in liquid water [23, 24] and long-range perturbation of water structure near large obstacles [25] are currently being explored. A quantitative description of the hydrophobic effect remains a central problem in theoretical chemical physics and biophysics.

3.2 Electrostatic free energy contribution

The electrostatic free energy contribution in Eq. (14) may be expressed as a thermodynamic integration corresponding to a reversible process between two states of the system: no solute-solvent electrostatic interactions ($\lambda = 0$) and full electrostatic solute-solvent interactions ($\lambda = 1$). The electrostatic free energy has a particularly simple form if the thermodynamic parameter λ corresponds to a scaling of the solute charges, i.e., $U_{\mathrm{uv}}^{(\mathrm{elec})}(\mathbf{X}, \mathbf{Y}; \lambda) = \lambda U_{\mathrm{uv}}^{(\mathrm{elec})}(\mathbf{X}, \mathbf{Y})$, and the coupling is linear

$$\Delta W^{\mathrm{elec}}(\mathbf{X}) = \int_0^1 d\lambda \left\langle U_{\mathrm{uv}}^{(\mathrm{elec})} \right\rangle_{(\lambda)} \quad . \tag{16}$$

For this reason, the quantity $\Delta W^{(\mathrm{elec})}(\mathbf{X})$ is often called the "charging free energy". If one assumes that the solvent responds linearly to the charge of the solute, then $\left\langle U_{\mathrm{uv}}^{(\mathrm{elec})} \right\rangle_{(\lambda)}$ is proportional to λ and the charging free energy may be written as

$$\begin{aligned} \Delta W^{\mathrm{elec}}(\mathbf{X}) &= \int_0^1 d\lambda \sum_i q_i \Phi_{\mathrm{rf}}(\mathbf{x}_i; \lambda) \\ &\approx \frac{1}{2} \sum_i q_i \Phi_{\mathrm{rf}}(\mathbf{x}_i; \lambda = 1) \end{aligned} \tag{17}$$

where $\Phi_{\mathrm{rf}}(\mathbf{x}_i; \lambda = 1)$ is the solvent field acting on the ith solute atomic charge located at the position $\mathbf{x}_i$ in reaction to the presence of all the solute charges (in the following, the coupling parameter λ will be omited for the sake of simplicity). The "reaction field" is thus the electrostatic potential exerted on the solute by the solvent that it has polarized. The assumption of linear response implies that $\Delta W^{\mathrm{elec}} = \frac{1}{2}\langle U_{\mathrm{uv}}^{(\mathrm{elec})} \rangle$, a relation which is often observed in calculations based on simulations with explicit solvent [26, 27]. The factor of one half is a characteristic signature of linear solvent response.

The dominant effects giving rise to the charging free energy is often modeled on the basis of classical continuum electrostatics. This approximation, in which the polar solvent is represented as a structureless continuum dielectric medium, was originally pioneered by Born in 1920 to calculate the hydration free energy of spherical ions [10]. It was later extended by Kirkwood [12] and Onsager [13] for the treatment of arbitrary charge distributions inside a spherical cavity. Noawadays, the treatment of solutes of arbitrary shape is possible with the use of powerfull computers and numerical methods. In many cases, this is an excellent approximation. The classical electrostatics approach is remarkably successful in reproducing the electrostatic contribution to the solvation free energy of small solutes [26, 28] or amino acids [27], as shown by comparisons to free energy simulations with explicit solvent. Applications to biophysical systems have been reviewed recently in [29, 30]. Due to its importance the next section is devoted completely to this topic.

4 Classical continuum electrostatics.

4.1 Poisson equation for macroscopic media

The continuum electrostatic approximation is based upon the assumption that the solvent polarization density of the solvent at a position $\mathbf{r}$ in space is linearly related to the total local electric field at that position. The Poisson equation for macroscopic continuum media follows from those assumptions about the local and linear electrostatic response of the solvent [31]

$$\nabla \cdot [\epsilon(\mathbf{r}) \nabla \phi(\mathbf{r})] = -4\pi \rho_{\mathrm{u}}(\mathbf{r}) \tag{18}$$

where $\phi(\mathbf{r})$, $\rho_{\mathrm{u}}(\mathbf{r})$ and $\epsilon(\mathbf{r})$ are the electrostatic potential, the charge density of the solute, and the position-dependent dielectric constant at the point $\mathbf{r}$, respectively. The Poisson equation (18) can be solved numerically by mapping the system onto a discrete grid and using a finite-difference relaxation algorithm [32, 33]. Several programs are available for

computing the electrostatic potential using this approach, e.g., DelPhi [34, 33], UHBD [35], and the PBEQ module [27, 36] incorporated in the simulation program CHARMM [37]. Alternatively, on can use an approach based on finite elements distributed at the dielectric boundary (the boundary element method) [38]. Significant improvements can be obtained with this approach by using efficient algorithms for generating the mesh at the dielectric boundaries [39, 40, 41]. FAMBE is one available program to compute the electrostatic potential using this method [42]. Lastly, a differerent (but physically equivalent) approach to incorporate the influence of a polar solvent, in which the solvent is modeled by a discrete lattice of dipoles that re-orient under the influence of applied electric fields, has been proposed and developed by Warshel and co-workers [43].

It is generally assumed that the dielectric constant is uniform everywhere except in the vicinity of the solute–solvent boundary. If all the solute degrees of freedom are treated explicitly and the influence of induced electronic polarization is neglected, the position-dependent dielectric constant $\epsilon(\mathbf{r})$ varies sharply from one, in the interior of the solute, to $\epsilon_{\rm v}$ in the bulk solvent region outside the solute. Such a form for $\epsilon(\mathbf{r})$ follows rigorously from an analysis based on a statistical mechanical integral equation under the assumption that there are only short-range direct correlations in the solvent [44]. To estimate the electrostatic contribution to the solvation free energy the reaction field $\Phi_{\rm rf}$ used in Eq. (17) is obtained as the electrostatic potential calculated from Eq. (18) with the nonuniform dielectric constant $\epsilon(\mathbf{r})$, minus the electrostatic potential calculated with a uniform dielectric constant of one.

Results obtained using macroscopic continuum electrostatics for biomolecular solutes depend sensitively on atomic partial charges assigned to the nuclei and the location of the dielectric boundary between the solute and the solvent. The dielectric boundary may be constructed on the basis of the molecular surface [34, 35], or the solvent accessible surface (constructed as a surface formed by overlapping spheres) [27]. The parameterization of an accurate continuum electrostatic model thus requires the development of optimal sets of atomic radii for the solutes of interest. Different parameterization schemes aimed at reproducing the solvation free energy of a collection of molecules have been suggested [27, 28]. From a fundamental point of view, the dielectric boundary is closely related to the nearest density peak in the solute-solvent distribution function [45]. As a consequence, the optimal radius of an atom is not a property of that atom alone but is an effective empirical parameter that depends on its charge, neighbors in the solute, and also on the nature of the molecules forming the bulk solvent. In contrast to the radii, the partial charges of the solute are generally taken from one of the standard biomolecular forcefields without modification without further adjustments and are not considered as free parameters.

Continuum electrostatic approaches based on the Poisson equation have been used to address a wide variety of problems in biology. One particularly useful application is in the determination of the protonation state of titrable groups in proteins [46]. For further details the readers are refered to recent reviews [29, 30].

4.2 Electrostatic forces and analytic gradients

In most practical applications of continuum electrostatic, the solute is considered to be in a fixed conformation. However, this procedure has obvious limitations since it ignores the importance of conformational flexibility. To proceed further requires the knowledge of the "electrostatic solvation forces" associated with the continuum electrostatics description of the solvent, i.e., the analytic first derivative of the solvation free energy with respect to the atomic coordinates of the solute. The computation of analytical gradients of the free energy of solvation with respect to nuclear coordinates is important for efficient geometry optimization based on energy minization, conformational searches, and dynamics. Analytic gradient for finite-difference solutions to the Poisson equation have been presented by Gilson et al [47], and Im et al [36]. Boundary element methods can also be used very effectively for computing analytical gradients [48]. Nonetheless, when repeated evaluation of the solvation energy are requested, the solution to the classical electrostatic problem and the calculation of analytical gradient may be too expensive computationally. For this purpose, approximations to the exact continuum electrostatics based on semianalytical functions have been developed. This is possible, in principle, since the free energy can be expressed as a superposition of pairwise additive terms (which depends on the geometry of the solute–solvent dielectric interface) because the equations of continuum electrostatics are linear. The general strategy of semianalytical approaches is to design a suitable closed-form pairwise deshielding function for the charge-charge coupling. One of the most popular approximation is the Generalized Born (GB) [49], although alternative formulations such as the Field Integrated Electrostatic Approach (FIESTA) [50], the Inducible Multipole Solvation Model (IMS) [51], the Analytical Continuum Electrostatics approach (ACE) [52], and the Solvation Models (SMx) of Cramer and Truhlar [53], are also based upon this general idea. Semianalytical approximations such as GB represent a very promising approaches for incorporating the influence of solvent in biomolecular simulations implicitly. Extensions and improvements to the

original form of the GB deshielding function have been proposed and parametrized [54, 55, 56, 57, 58]. The results have been compared with those from numerical continuum electrostatic calculations [57, 59, 58] and explicit solvent simulations [60, 61]. GB has been applied to various problems, e.g., proteins and nucleic acids stability [62, 63, 64] conformational searches [65, 66], macromolecular association [67], and ligand binding [68].

4.3 Treatment of ionic strength

The concentration of salt in physiological system is on the order of 150 mM, which corresponds approximately to 350 water molecules for each cation-anion pair. For this reason, investigations of salt effects in biological systems using detailed atomic models and molecular dynamic simulations become rapidly prohibitive. and mean-field treatments based on continuum electrostatics are advantageous. Such approximations, which were pioneered by Debye and Hückel [11], are valid at moderately low ionic concentration when core-core interactions between the mobile ions may be neglected. Briefly, the spatial density throughout the solvent is assumed to depend only on the local electrostatic potential $\rho_i(\mathbf{r}) = \bar{\rho}_i e^{-q_i\phi(\mathbf{r})/k_{\rm B}T}$, where i refers to a specific ion type (e.g. counterion or co-ion), $\bar{\rho}_i$ is the number density in the bulk solution. The total ion charge density (sumed over the different ion types) is then inserted explicitly in the Poisson equation with the solute charge $\rho_{\rm u}(\mathbf{r})$, resulting in the non-linear form of the Poisson-Boltzmann (PB) equation. Linearization with respect to the potential ϕ yields the familiar Debye-Hückel approximation [11, 69]

$$\nabla\cdot[\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})] - \bar{\kappa}^2(\mathbf{r})\phi(\mathbf{r}) = -4\pi\rho^{(\rm u)}(\mathbf{r}) \qquad (19)$$

where $\bar{\kappa}^2(\mathbf{r})$ is the space-dependent screening factor which varies from zero, in the interior of the solute, to $4\pi\sum_i q_i^2\bar{\rho}_i/k_{\rm B}T$, in the bulk solvent. The spatial dependence of $\bar{\kappa}^2(\mathbf{r})$ is often assumed to be similar to that of $\epsilon(\mathbf{r})$, though that is not necessary. The PB equation (linear and nonlinear) is a particularly simple and powerful approach to address questions about the influence of salt on complex biological systems. In particular, it has been used to examine the salt dependence of the conformational stability of nucleic acids [70, 71] and protein-DNA association [72].

4.4 Treatment of a transmembrane potential

The electrostatic free energy of a macromolecule embedded in a membrane in the presence of a membrane potential V can be expressed as the sum of three separate terms involving the capacitance C of the system, the reaction field $\Phi_{\rm rf}(\mathbf{r})$, and the membrane potential field $\Phi_{\rm mp}(\mathbf{r})$ [73]

$$\begin{aligned}\Delta W^{\rm elec} &= \frac{1}{2}CV^2 + \frac{1}{2}\sum_i q_i\Phi_{\rm rf}(\mathbf{x}_i) \\ &\quad + \left[\sum_i q_i\Phi_{\rm mp}(\mathbf{x}_i)\right]V \end{aligned} \qquad (20)$$

where q_i and $\mathbf{x}_i$ are the solute charges and their position, respectively. Generally, the capacitive energy contribution is negligible. The function $\Phi_{\rm mp}(\mathbf{x}_i)$ corresponds to the fraction of the electrostatic transmembrane potential interacting with a charge of the solute. It is calculated by solving a modified version of the linear PB equation

$$\nabla\cdot[\epsilon(\mathbf{r})\nabla\Phi_{\rm mp}(\mathbf{r})] - \bar{\kappa}^2(\mathbf{r})[\Phi_{\rm mp}(\mathbf{r}) - \Theta(\mathbf{r})] = 0 \qquad (21)$$

where the function $\Theta(\mathbf{r})$ is equal to 1 on the side of the membrane which is contact with the bulk solution set to the reference potential V, and zero otherwise. The Θ function in Eq. (21) insures that the mobile ions are in equilibrium with the bath with which they are in contact. In the case of a perfectly planar system, the electric field across the membrane is constant and $\Phi_{\rm mp}(\mathbf{x})$ is a linear function corresponding roughly to a fraction of the membrane thickness (for this reason, it is often referred to as the "electric distance" [74, 75]). If the shape of the protein-solution interface is irregular the interaction of the solute charges with the membrane potential is more complicated than the simple linear field.

Simple considerations show that the membrane potential cannot be treated with computer simulations and continuum electrostatic methods may represent the only practical approach to address such questions. The capacitance of a typical lipid membrane is on the order of 1 μF/cm^2, which corresponds to a thickness of approximately 25 Å and a dielectric constant of 2 for the hydrophobic core of a bilayer. In the presence of a membrane potential the bulk solution remains electrically neutral and a small charge imbalance is distributed in the neighborhood of the interfaces. The membrane potential arises from a strikingly small accumulation of net charge relative to the bulk ion density. Typical physiological conditions correspond to a membrane potential on the order of 100 mV and a salt concentration of 150 mM. In this situation, the net charge per area is $CV = 10^{-7}$Coul/cm^2, which corresponds to only one atomic unit charge per surface of (130 Å)^2. For molecular dynamics simulations, a minimal salt solution at a concentration of 150 mM with a membrane system of cross-sectional area of (130 Å)^2 containing about 100 ion pairs would require nearly 50000 water molecules and 500 phospholipid, for a total of more than 200000 atoms, which is

computationally prohibitive. At the present time, the modified PB Eq. (21) with membrane potential may provide the only practical way to address questions about the membrane potential and its influence on the configurational free energy of intrinsic protein. The approach has been implemented in the PBEQ module [27, 36, 73] of the biomolecular simulation program CHARMM [37] and has been used to calculate the influence of the transmembrane potential on the insertion of an α-helix into a membrane.

5 Miscellaneous approaches

5.1 Statistical mechanical integral equations

The average solvent structure caused by the granularity, packing, and hydrogen bonding, gives rise to important effects which are ignored by continuum electrostatic approaches. Statistical mechanical theories based on distribution functions and integral equations are sophisticated approaches which can provide a rigorous framework for incorporating such effects into a description of solvation [3, 76]. A complete review of integral equations would be beyond the scope of this chapter, therefore we will only provide a brief overview of this vast field.

One important class of integral equation theories is based on the Reference Interaction Site Model (RISM) proposed by Chandler [77]. These RISM theories have been used to study the conformation of small peptides in liquid water [78, 79, 80]. However, the approach is not appropriate for large molecular solutes such as proteins and nucleic acids. Because RISM is based on a reduction to site-site solute-solvent radially symmetric distribution functions, there is a loss of information about the three-dimensional spatial organization of the solvent density around a macromolecular solute of irregular shape. To circumvent this limitation, extensions of RISM-like theories for three-dimensional space (3d-RISM) have been proposed [81, 82],

$$c_\alpha(\mathbf{r}) = e^{-\beta U_\alpha(\mathbf{r}) + h_\alpha(\mathbf{r}) - c_\alpha(\mathbf{r})} - h_\alpha(\mathbf{r}) + c_\alpha(\mathbf{r}) - 1 \quad (22)$$

and

$$\bar{\rho} h_\alpha(\mathbf{r}) = \int d\mathbf{r}' \sum_\gamma c_\gamma(\mathbf{r}') \chi_{\gamma\alpha}(\mathbf{r} - \mathbf{r}'), \quad (23)$$

where $U_\alpha(\mathbf{r})$ is the solute-solvent interaction on the solvent site α, $h_\alpha(\mathbf{r})$ is the solute-solvent site correlation function $h_\alpha(\mathbf{r}) \equiv [\rho_\alpha(\mathbf{r})/\bar{\rho} - 1]$, $c_\alpha(\mathbf{r})$ is the solute-solvent site direct correlation function, and $\chi_{\gamma\alpha}(\mathbf{r} - \mathbf{r}')$ is the density susceptibility of the uniform unperturbed liquid. The solvent susceptibility (an input in this approach) is related to the equilibrium site-site density susceptibility of the uniform unperturbed liquid. Numerical solutions of the 3d-RISM equation indicate that this approach is able to incorporate important features of hydration such as hydrogen bonding and packing of the solvent molecules in the first solvation shell [81, 82]. Recent advances allow the accurate estimate of the solvation free energy for nonpolar as well as polar biomolecules [83].

Other statistical mechanical theories are also currently being explored. An extension to the mean-spherical-approximation integral equation in three dimensions (3d-MSA), describing the distribution function of a liquid of spherical molecules with an embedded dipole around a polar solute [44] as well as an integral equation describing the structure of water molecules in terms of sticky interaction points [84, 85] were formulated and solved numerically. A theory based on an expansion in terms of two- and three-body correlation functions has been proposed to describe the hydration structure around nucleic acids [86] and proteins [87]. A theory for inhomogeneous fluids in the neighborhood of large non-polar solutes was proposed by Chandler et al [25] to describe the hydrophobic effect.

5.2 Solvent boundary potentials and implicit/explicit mixed schemes

A description in which all atomic and structural details of the solvent molecules are ignored may not always be desirable. In some cases, it may be advantageous to use a mixed scheme which combines an implicit solvent model with a limited number of explicit solvent molecules. An intermediate approach, illustrated schematically in Figure 3, consists in including a small number of explicit solvent molecules in the vicinity of the solute, while representing the influence of the remaining bulk with an effective solvent boundary potential [88, 89, 90, 91, 92, 93, 94, 95]. The first to design such a simulation method appropriate for liquids were Berkowitz and McCammon [88]. In their method, the many-body system was divided into three main spherical regions: a central reaction region, a buffer region and a surrounding static reservoir region. The forces arising from the reservoir region were calculated from fixed atomic centers. Instead of using explicit fixed atomic centers in the bath region, Brooks and Karplus introduced a Mean Force Field Approximation (MFFA) to calculate a soft boundary potential representing the average influence of the reservoir region on the reaction region [89]. In the MFFA treatment, the boundary potential was calculated by integrating all contributions to the average force arising from the reservoir region. The MFFA approach was extended by Brooks, Brunger and Karplus for the simulation of bulk water [90]. A similar potential for water droplets of TIP4P was devel-

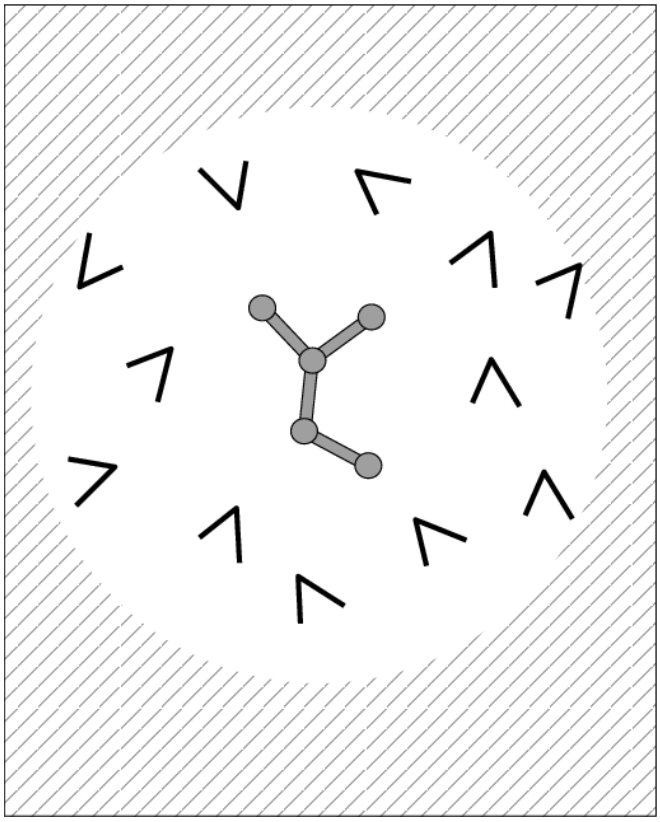

Figure 3: Schematic representation of a mixed explicit/implicit solvent treatment. A a small number of water molecules are included explicitly in the vicinity of the solute while the influence of the remaining bulk is taken into account implicitly.

oped by Essex and Jorgensen [91]. The average electrostatic reaction field was taken into account in the Surface Constrained All-Atom Solvent (SCAAS) treatment of King and Warshel [93], and in the Reaction Field with Exclusion (RFE) of Rullmann and van Duijnen [94].

The problem was reformulated on the basis of a separation of the multidimensional solute-solvent configurational integral in terms of "inner" solvent molecules nearest to the solute, and the remaining "outer" bulk solvent molecules [95]. Following this formulation the solvent boundary potential was identified as the solvation free energy of an effective cluster comprising the solute and inner explicit solvent molecules embedded in a large hard sphere. The hard sphere corresponds to a configurational restriction on the outer bulk solvent molecules; its radius is variable, such that it includes the most distant inner solvent molecule. An approximate Spherical Solvent Boundary Potential (SSBP) based on this formulation has been implemented in the biomolecular simulation program CHARMM [37]. Using computer simulations it was shown that SSBP yields solvation free energies that do not depend sensitively on the number of explicit water molecules [95].

5.3 Solvent accessible surface area models

In section 3 we described an approximation to the non-polar free energy contribution based on the concept of the SASA (see Eq. (15)). In the SASA/PB implicit solvent model, the non-polar free energy contribution is complemented by a macroscopic continuum electrostatic calculation based on the PB equation, thus yielding an approximation to the total free energy, $\Delta W = \Delta W^{\mathrm{np})} + \Delta W = \Delta W^{\mathrm{elec})}$. A different implicit solvent model, which also makes use of the concept of SASA, is based on the assumption that the *entire* solvation free energy of a solute can be expressed in terms of a linear sum of atomic contributions weighted by partial exposed surface area

$$\Delta W(\mathbf{X}) = \sum_i \gamma_i \mathcal{A}_i(\mathbf{X}). \tag{24}$$

Here, $\mathcal{A}_i(\mathbf{X})$ is the partial SASA of atom i (which depends on the solute configuration $\mathbf{X}$), and γ_i is an atomic free energy per unit area associated with atom i. We refer to those models as "full SASA". Because it is so simple, this approach is widely used in computations on biomolecules [96, 97, 98]. Variations of the solvent-exposed area models are the shell model of Scheraga [99, 100], the excluded-volume model of Colonna and Sander [101, 102], and the gaussian model of Lazaridis and Karplus [103]. Full SASA models have been used for investigating the thermal denaturation of proteins [103] and to examine protein-protein assocation [104].

One important limitation of full SASA models is the difficulty in taking into account the dielectric shielding of electrostatic interactions between charged particles in a physically realistic way. The SASA model incorporates, in an average way, the free energy cost of taking a charged particle and burying it in the interior of the protein. In the continuum electrostatic description that corresponds to the self-interaction energy, i.e., the interaction of a charge with its own reaction field. However, as two charged particles move from the solvent to the non-polar core of the protein, their electrostatic interaction should also vary progressively from a fully to an incompletely shielded form. Thus, full SASA approximations require further assumptions about the treatment of electrostatic interactions and dielectric shielding in practical applications. For example, in full SASA models residues carrying a net charge are usually neutralized and a distance-dependent dielectric function is introduced to shield the coulomb potential at large

distances [98, 103].

5.4 Knowledge-based potentials

One of the greatest problems in predicting the three-dimensional fold of a protein is the need to search over a large number of possible configurations to find the global free energy minimum. For extensive configurational searches, it is necessary to use a free energy function $W(\mathbf{X})$ that is as simple and inexpensive as possible. Knowledge-based potentials are the simplest free energy functions that can be designed for this purpose. Such potentials are constructed empirically from statistical analyses of known protein structures taken from structural data bases [105]. The general idea is that the number of residue pairs at a certain distance observed in the data base follows the statistics of a thermal ensemble, in other words a Boltzmann principle [106]. Equivalently, it is assumed that the observed probability of finding a pair of residues at a distance R in a protein structure is related to the Boltzmann factor of an effective distance-dependent free energy. The simplest potentials distinguish only two types of residues: non-polar and polar [105]. Usually no attempts are made to establish a realistic description of the microscopic interactions at the atomic level, though some comparison have been made with explicit solvent simulations [60]. For example, one of the simplest potentials, designed by Sippl [105], is attractive for pairs of non-polar residues and repulsive for pairs of polar residues. Nevertheless, the resulting structures that are obtained via conformational searches, usually with an additional restraint on the protein radius of gyration, are reasonable: the non-polar residues tend to form a hydrophobic core in the center of the structure, while the polar residues tend to be located at the protein surface. There is a growing number of potentials constructed from similar ideas [107, 108, 109, 110, 111]. Recently, Mirny and Shakhnovich have re-examined the methods for deriving knowledge-based potentials for protein folding [112]. Their potential is obtained by a global optimization procedure that simultaneously maximizes thermodynamic stability for all proteins in the database. This field is in rapid expansion, and it is beyond the scope of the present review to cover all possible developments. For more information, see the reviews [113, 114, 115] and references therein.

6 Summary

A statistical mechanical formulation of implicit solvent representations provides a robust theoretical framework for understanding the influence of solvation biomolecular systems. A decomposition of the free energy in terms of non-polar and electrostatic contributions, $\Delta W = \Delta W^{(\mathrm{np})} + \Delta W^{(\mathrm{elec})}$, is central to many approximate treatments. An attractive and widely used treatment consists in representing the non-polar contribution $\Delta W^{(\mathrm{np})}$ by a SASA surface tension term with Eq. (15), and the electrostatic contribution $\Delta W^{(\mathrm{elec})}$ using the finite-difference PB Eq. (19). These two approximations constitute the SASA/PB implicit solvent model. Although SASA/PB does not incorporate solvation effects with all atomic details, it nevertheless relies on a physically consistent picture of solvation. A relationship with first principles and statistical mechanics can be established and the significance of the approximations at the microscopic level can be clarified. The results can be compared with computer simulations including explicit solvent molecules [15, 26, 27, 28]. Implicit solvent models based on the SASA/PB approximation have been used to address a wide range of questions concerning biomolecular systems, e.g., discriminate misfolded proteins [116], assess the conformational stability of nucleic acids [71], and examine protein-ligand [117], protein-DNA [72], and protein-membrane association [73, 118].

It is possible to go beyond the SASA/PB approximation and develop better approximations to current implicit solvent representations with sophisticated statistical mechanical models based on distribution functions or integral equations (see section 5a). An alternative intermediate approach consists in including a small number of explicit solvent molecules near the solute while the influence of the remain bulk solvent molecules is taken into account implicitly (see section 5b). On the other hand, in some cases it is necessary to use a treatment that is markedly simpler than SASA/PB to carry out extensive conformational searches. In such situations, it possible to use empirical models which describe the entire solvation free energy on the basis of the SASA (see section 5c). An even simpler class of approximations consists in using information-based potentials constructed to mimic and reproduce the statistical trends observed in macromolecular structures (see section 5d). Although the microscopic basis of these approximations is not yet formally linked to a statistical mechanical formulation of implicit solvent, full SASA models and empirical information-based potentials may be very effective for particular problems.

References

[1] C. L. Brooks III, Martin Karplus, and B. M. Pettitt. Proteins. A theoretical perspective of dynamics, structure and thermodynamics. In I. Prigogine and S. A. Rice, editors, *Advances in Chemical Physics Vol. LXXI*. John Wiley & Sons, New York, (1988).

[2] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. Oxford Science Publications, Clarendon Press, Oxford, (1989).

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

[4] J.G. Kirkwood. *J. Chem. Phys.* **3**, 300 (1935).

[5] Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, and Kollman PA. *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).

[6] MacKerell Jr. A.D., Bashford D., Bellot M., Dunbrack R.L., Evanseck J.D., Field M.J., Fischer S., Gao J., Guo H., Ha S., Joseph-McCarthy D., Kuchnir L., Kuczera K., Lau F.T.K., Mattos C., Michnick S., Ngo T., Nguyen D.T., Prodhom B., Reiher III W.E., Roux B., Schlenkrich B., Smith J., Stote R., Straub J., Watanabe M., Wiorkiewicz-Kuczera J., and Karplus M. *J. Phys. Chems. B* **102**, 3586–3616 (1998).

[7] Jorgensen W.L., Maxwell D.S., and Tirado-Rives J. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).

[8] Boresch S., Archontis G., and Karplus M. *Proteins.* **20**, 25 (1994).

[9] Lee B. and Richards F.M. *J. Mol. Biol.* **55**, 379 (1971).

[10] M. Born. *Z. Phys.* **1**, 45 (1920).

[11] Debye P. and E. Hückel. *Phys. Z.* **24**, 305–325 (1923).

[12] J.G. Kirkwood. *J. Chem. Phys.* **2**, 351 (1934).

[13] L. Onsager. *J. Am. Chem. Soc.* **58**, 1468–1493 (1936).

[14] Lum K, D Chandler, and JD Weeks. *Adv. Prot. Chem.* **14**, 1 (1959).

[15] J. P. Postma, H. C. Berendsen, and J. R. Haak. *Farad. Symp. Chem. Soc.* **17**, 55 (1982).

[16] Kim Sharp, Anthony Nicholls, Richard Fine, and Barry Honig. *Science* **252**, 106–109 (1991).

[17] Kim Sharp, Anthony Nicholls, Richard Friedman, and Barry Honig. *Biochemistry* **30**, 9696–9697 (1991).

[18] Thomas Simonson and A. T. Brünger. *J. Phys. Chem.* **98**, 4683–4694 (1994).

[19] C. Tanford. *Proc. Natl. Acad. Sci. USA* **76**, 4175–4176 (1979).

[20] H. Reiss. *Adv. Chem. Phys.* **9**, 1–84 (1965).

[21] F. Stillinger. *J. Sol. Chem.* **2**, 141–158 (1973).

[22] Robert A. Pierotti. *Chem. Rev.* **76**, 717–726 (1976).

[23] Hummer G., Garde S., Garcia A.E., Pohorille A., and Pratt L.R. *Proc. Nat. Acad. Sci. USA.* **93**, 8951 (1996).

[24] G. Hummer, L.R. Pratt, and A.E. Garcia. *J. Chem. Phys.* **107**, 9275 (1997).

[25] Lum K, D Chandler, and JD Weeks. *J. Phys. Chem.* (1999).

[26] A. Jean-Charles, A. Nicholls, K. Sharp, B. Honig, A. Tempczyk, T. Hendrickson, and W.C. Still. *J. Am. Chem. Soc.* **113**, 1454–1455 (1991).

[27] M. Nina, D. Beglov, and B. Roux. *J. Phys. Chem.* **101**, 5239–5248 (1997).

[28] Sitkoff D., Sharp K.A., and Honig B. *J. Phys. Chem.* **98**, 1978 (1994).

[29] B. Honig and A. Nicholls. *Science* **268**, 1144 (1995).

[30] K.A. Sharp and B. Honig. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 301–332 (1990).

[31] J.D. Jackson. *Classical Electrodynamics*. John Wiley & Sons, New York, (1962).

[32] Warwicker J. and Watson H.C. *J. Mol. Biol.* **157**, 671–679 (1982).

[33] Klapper I., Hagstrom R., Fine R., Sharp K., and Honig B. *Proteins* **1**, 47 (1986).

[34] Gilson MK, Sharp KA, and Honig BH. *J. Comp. Chem.* **9**, 327–335 (1987).

[35] Davis ME, Madura JD, Luty B, and McCammon JA. *Comp. Phys. Comm.* **62**, 187–197 (1991).

[36] Im W., Beglov D., and Roux B. *Comp. Phys. Comm.* **111**, 59–75 (1998).

[37] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. *J. Comput. Chem.* **4**, 187–217 (1983).

[38] R. Zauhar and R. Morgan. *J. Molec. Biol.* **186**, 815–820 (1985).

[39] Liang J and Subramaniam S. *Biophys. J.* **73**, 1830–1841 (1997).

[40] Cortis CM and Friesner RA. *J. Comp. Chem.* **18**, 1570–1590 (1997).

[41] Cortis CM and Friesner RA. *J. Comp. Chem.* **18**, 1591–1608 (1997).

[42] Vorobjev Y.N. and Scheraga H.A. *J. Comp. Chem.* **18**, 569–583 (1997).

[43] A. Warshel and J. Åqvist. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 267–298 (1991).

[44] D. Beglov and B. Roux. *J. Chem. Phys.* **104**, 8678–8689 (1996).

[45] B. Roux, H.-A. Yu, and M. Karplus. *J. Phys. Chem.* **94**, 4683–4688 (1990).

[46] Antosiewicz J, McCammon JA, and Gilson MK. *Biochem.* **18**, 7819–7833 (1996).

[47] Gilson M.K., McCammon J.A., and Madura J.D. *J. Comp. Chem.* **16**, 1081 (1995).

[48] Friedrichs M, Zhou RH, Edinger SR, and Friesner RA. *J. Phys. Chem. B* **103**, 3057–3061 (1999).

[49] W.C. Still, A. Tempczyk, R.C. Hawley, and T. Hendrickson. *J. Am. Chem. Soc.* **112**, 6127–6129 (1990).

[50] H. Sklenar, F. Eisenhaber, M. Poncin, and R. Lavery. In D. L. Beveridge and R. Lavery, editors, *Theoretical Biochemistry and Molecular Biophysics*. Adenine Press, (1990).

[51] M. E. Davis. *J. Chem. Phys.* **100**, 5149–5159 (1994).

[52] M. Schaefer and M. Karplus. *J. Phys. Chem.* **100**, 1578–1599 (1996).

[53] C. J. Cramer and D. G. Truhlar. In K. B. Lipkowitz, editor, *Reviews in computational chemistry*. VCH, New York, (1995).

[54] Edinger S.R., Cortis C., Shenkin P.S., and Friesner R.A. *J. Phys. Chem. B* **101**, 1190 (1997).

[55] Jayaram B, Sprous D, and Beveridge DL. *Journal of Physical Chemistry B* **102**, 9571–9576 (1998).

[56] Jayaram B, Liu Y, and Beveridge DL. *Journal of Chemical Physics* **109**, 1465–1471 (1998).

[57] Ghosh A, Rapp CS, and Friesner RA. *J. Phys. Chem. B* **102**, 10983–10990 (1998).

[58] Dominy BN and Brooks CL. *J. Phys. Chem. B* **103**, 3765–3773 (1999).

[59] Dixit SB and Jayaram B. *J. Biomol. Struct. Dyn.* **16**, 237–242 (1998).

[60] Mohanty D, Dominy BN, Kolinski A, Brooks CL, and Skolnick J. *Proteins* **35**, 447–452 (1999).

[61] Scarsi M, Apostolakis J, and Caflisch A. *Journal of Physical Chemistry B* **102**, 3637–3641 (1998).

[62] Srinivasan J, Trevathan MW, Beroza P, and Case DA. *Theoretical Chemistry Accounts* **101**, 426–434 (1999).

[63] Jayaram B, McConnell KJ, Dixit SB, and Beveridge DL. *Journal of Computational Physics* **151**, 333–357 (1999).

[64] Luo R, David L, Hung H, Devaney J, and Gilson MK. *Journal of Physical Chemistry B* **103**, 727–736 (1999).

[65] Srinivasan J, Miller J, Kollman PA, and Case DA. *J. Biomol. Struct. Dyn.* **16**, 671–67+ (1998).

[66] Rapp CS and Friesner RA. *Proteins* **35**, 173–183 (1999).

[67] Luo R, Head MS, Given JA, and Gilson MK. *Biophysical Chemistry* **78**, 183–193 (1999).

[68] Bostrom J, Norrby PO, and Liljefors T. *Journal of Computer-Aided Molecular Design* **12**, 383–396 (1998).

[69] R. H. Fowler and E. A. Guggenheim. *Statistical Thermodynamics*. Cambridge University Press, (1939).

[70] Sharp KA and Honig B. *Curr. Opin. Struct. Biol.* **5**, 323–328 (1995).

[71] Misra VK and Honig B. *Biochem.* **35**, 1115–1124 (1996).

[72] Misra VK, Hecht JL, Yang AS, and Honig B. *Biophys. J.* **75**, 2262–2273 (1998).

[73] Roux B. *Biophys. J.* **73**, 2980–2989 (1997).

[74] Hille B. *Ionic Channels of Excitable Membranes, 2nd edition*. Sinauer, Sunderland MA, (1992).

[75] Sigworth F.J. *Quat. Rev. Biophys.* **27**, 1–40 (1993).

[76] J. P. Hansen and I. R. McDonald. *Theory of Simple Liquids, 2nd Edition*. Academic Press, London, (1986).

[77] D. Chandler. The equilibrium theory of polyatomic fluids. In E.W. Montroll and J.L. Lebowitz, editors, *The Liquid State of Matter: Fluids, Simple and Complex*, volume VIII. North-Holland, Amsterdam, (1982).

[78] M. B. Pettitt and M. Karplus. *Chem. Phys. Lett.* **121**, 194–201 (1985).

[79] M. B. Pettitt and M. Karplus. *Chem. Phys. Lett.* **136**, 383–386 (1987).

[80] W.F. Lau. and B.M. Pettitt. *Biopol.* **26**, 1817–1831 (1987).

[81] D. Beglov and B. Roux. *J. Phys. Chem.* **101**, 7821–7826 (1997).

[82] C.M. Cortis, P.J. Rossky, and R.A. Friesner. *J. Chem. Phys.* **107**, 6400–6414 (1997).

[83] Q. Du, D. Beglov, and B. Roux. *J. Phys. Chem. B* (1999).

[84] Y. Liu and T. Ichiye. *Chem. Phys. Lett.* **231**, 380–386 (1994).

[85] J-K Hyun, C.S. Babu, and T. Ichiye. *J. Chem. Phys.* **99**, 5187–5185 (1995).

[86] G. Hummer and D. M. Soumpasis. *Phys. Rev. E* **50**, 5085–5095 (1994).

[87] A. E. García, G. Hummer, and D. M. Soumpasis. *Biophys. J.* **72**(2, Part 2), A104 (1997).

[88] M. Berkowitz and J.A. Mc.Cammon. *Chem. Phys. Lett.* **90**, 215 (1982).

[89] C. L. Brooks III and M. Karplus. *J. Chem. Phys.* **79**, 6312–6325 (1983).

[90] A. Brunger, C. L. Brooks III, and M. Karplus. *Chem. Phys. Letters* **105**, 495–500 (1984).

[91] J.W. Essex and W.L. Jorgensen. *J. Comp. Chem.* **16**, 951–972 (1995).

[92] A. Warshel and G.King. *Chem. Phys. Lett.* **121**, 124 (1985).

[93] G. King and A. Warshel. *J. Chem. Phys.* **91**, 3647 (1989).

[94] J. A. Rullmann and P. Th. van Duijnen. *Mol. Phys.* **61**, 293 (1987).

[95] D. Beglov and B. Roux. *J. Chem. Phys.* **100**, 9050–9063 (1994).

[96] D. Eisenberg and A. McClachlan. *Nature* **319**, 199–203 (1986).

[97] L. Wesson and D. Eisenberg. *Prot. Sci.* **1**, 227–235 (1992).

[98] Fraternali F and Van Gunsteren WF. *J. Mol. Biol.* **256**, 939–948 (1996).

[99] H. A. Scheraga. *Acc. Chem. Res.* **12**, 7–14 (1979).

[100] Y.K. Kang, K.D. Gibson, G. Nemethy, and H. Scheraga. *J. Phys. Chem.* **92**, 4739–4742 (1988).

[101] F. Colonna-Cesari and C. Sander. *Biophys. J.* **57**, 1103–1107 (1990).

[102] P. Stouten, C. Frömmel, H. Nakamura, and C. Sander. *Mol. Sim.* **10**, 97–120 (1993).

[103] T. Lazaridis and M. Karplus. *Science* **278**, 1928–1931 (1997).

[104] Cummings MD, Hart TN, and Read RJ. *Prot. Sci.* **4**, 2087–2089 (1995).

[105] Sippl MJ. *J. Mol. Biol.* **213**, 859–883 (1990).

[106] Sippl MJ. *J. of Computer-Aided Molecular Design* **7**, 473–501 (1993).

[107] Yue K and Dill KA. *Protein Science* **5**, 254–261 (1996).

[108] Bahar I and Jernigan RL. *J. Mol. Biol.* **266**, 195–214 (1997).

[109] Bryant SH and Lawrence CE. *Proteins* **16**, 92–112 (1993).

[110] DeBolt SE and Skolnick J. *Protein Engineering* **9**, 637–655 (1996).

[111] Koretke KK, Luthey-Schulten Z, and Wolynes PG. *Protein Science* **5**, 1043–1059 (1996).

[112] Mirny LA and Shakhnovich EI. *J. Mol. Biol.* **264**, 1164–1179 (1996).

[113] Vajda S, Sippl M, and Novotny J. *Current Opinion in Structural Biology* **7**, 222–228 (1997).

[114] Jones DT and Thornton JM. *Current Opinion in Structural Biology* **6**, 210–216 (1996).

[115] Jernigan RL and Bahar I. *Current Opinion in Structural Biology* **6**, 195–209 (1996).

[116] Vorobjev Y. N., Almagro J. C., and Hermans J. *PROT. Struc. Funct. Gen.* **32**, 399–413 (1998).

[117] Sharp KA. *Prot.* **1**, 39–48 (1998).

[118] Ben-Tal N., A. Ben-Shaul, A Nicholls, and B. H. Honig. *Biophys. J.* **70**, 1803–1812 (1996).