where ΔG_{PB} represents the polar contribution and ΔG_{SA} represents the nonpolar contribution to the solvation free energy.

The polar term in Equation (5.8) represents the energy stored in the continuum dielectric in response to the presence of the solute's charge distribution and is typically obtained by solution of the Poisson–Boltzmann (PB) equation. The PB equation provides a rigorous framework for representing discrete solute molecules embedded in a uniform dielectric continuum and has been shown to be capable of producing relatively robust predictions of electrostatic contributions to solvation free energies of small molecules as well as biological macromolecules.[25,26] The PB solutions are obtained in separate calculations for the ligand, protein, and bound protein/ligand complex, and the final solvation free-energy values are assembled using the thermodynamic cycle for association in solution.[27,28]

For any PB calculation, one must choose a particular representation of the dielectric boundary between solute and solvent, which can involve a number of subtleties.[29,30] In addition to the boundary representation, dielectric functions for the solute and solvent must also be chosen. For typical protein/ligand systems, constant values of 1.0 for solutes and 80.0 for solvent are most commonly used,[31] though there are also other arguments that using 2.0, 4.0, or a residue-based dielectric for the solute may give superior performance.[32,33] It should be noted that most force fields have been parameterized using an internal dielectric of 1.0.

Finally, the last term in Equation (5.8) is the nonpolar component of solvation free energy, which is usually treated as being proportional to the solvent exposed surface area[34] of the solute,

$$G^{SA} = \gamma \Delta SA, \tag{5.9}$$

where ΔSA is the change in accessible molecular surface area on binding, and γ is a microscopic surface free energy for formation of a cavity in water.[35] The form of this equation derives from empirical data on transfer free energies for linear, cyclic, and branched hydrocarbons.[36,37] The precise value of γ depends on the particular method used to probe the solvent-accessible surface of the solute.[25] The equation implicitly assumes that the nonpolar component has negligible contributions from dispersion interactions between solute and solvent relative to the energy required in displacing solvent molecules to create the cavity. A number of objections to this expression point out its oversimplification,[38–40] and a number of models have been proposed to attempt to address these shortcomings with more sophisticated frameworks.[39–41]

The last term on the right-hand side of Equation (5.6) is the entropic cost of confining the free ligand, which represents a significant fraction of the total change in solute entropy ΔS_{solute} for formation of the bound complex. Additional estimates of solute entropy can be performed, which typically use some form of normal-mode analysis and that are very computationally expensive to perform.[42,43] Alternatively, one could use empirical estimates of average entropic costs, such as the entropy required to constrain rotation around any given torsional degree of freedom.[44] However, neither of these approaches produce quality estimates of solute entropy; instead, they tend to add a significant random scatter to results.[21]

Because of the complications in dealing with entropy, it is often neglected for computational convenience. This approximation may be reasonable in cases where we are only interested in rank-ordering, and the amount of entropy/enthalpy compensation remains roughly constant across ligands. It will certainly be unreasonable for any case where absolute comparisons of free energy are desired across protein targets and in situations for which non-negligible perturbations in binding modes and pocket geometries occur across a ligand set. Recent developments for treating the entropy more properly show significant promise.[45]

In systems where there are relatively few populated states, it may be sufficient to perform PB calculations alone to generate robust affinity estimates. In a number of situations, PB solutions have been successfully used to estimate affinities,[28,46–48] although some implementations begin to resemble empirical scoring methods.[49] A major criticism of these approaches is their potential inaccuracy in situations where conformational flexibility plays a significant role.[50]

In generating the dynamics trajectories for the MM-PBSA analysis explicit representation of water molecules are typically used. Although explicit water molecules give the most detailed glimpse into structural dynamics, it has been shown that there can be pathologies in certain situations when using implicit-solvent theory to "score" explicit water configurations,[21,51] because the ensemble average energies should be computed with the same energy function used to generate the ensemble structures. An alternative to this is to sample in implicit solvent directly.[52]

As PB solutions are in general computationally demanding calculations, a number of groups have put significant efforts into developing faster approximations, such as the suite of generalized Born (GB) approaches.[53–56] However, there are numerous examples of pathologies using GB methods.[57–59]

Other implicit solvent methods

Another approach requiring only simulations of the bound and unbound states is to compute the partition function directly. The partition function of a molecular system can be computed as the sum of the integral of Boltzmann factors over neighborhoods of only the low-energy states, which are a relatively small fraction of the total configurations of the molecules.[60]

With the full partition function for the protein, ligand, and the complex in the case of absolute free energies, or for two ligands and two complexes in the case of relative free energies, the binding free energy or changes in binding