ELSEVIER

Biochimica et Biophysica Acta 1458 (2000) 63–87

BIOCHIMICA ET BIOPHYSICA ACTA

BBA

www.elsevier.com/locate/bba

Review

A pragmatic approach to structure based calculation of coupled proton and electron transfer in proteins

M.R. Gunner, E. Alexov

Physics Department City College of New York, 138th St. and Convent Ave., New York, NY 10031, USA

Received 1 November 1999; accepted 1 December 1999

Abstract

The coupled motion of electrons and protons occurs in many proteins. Using appropriate tools for calculation, the three-dimensional protein structure can show how each protein modulates the observed electron and proton transfer reactions. Some of the assumptions and limitations involved in calculations that rely on continuum electrostatics to calculate the energy of charges in proteins are outlined. Approaches that mix molecular mechanics and continuum electrostatics are described. Three examples of the analysis of reactions in photosynthetic reaction centers are given: comparison of the electrochemistry of hemes in different sites; analysis of the role of the protein in stabilizing the early charge separated state in photosynthesis; and calculation of the proton uptake and protein motion coupled to the electron transfer from the primary (Q_A) to secondary (Q_B) quinone. Different mechanisms for stabilizing intra-protein charged cofactors are highlighted in each reaction.

Keywords: Continuum electrostatics; pK_a; Electrochemistry; Reaction center; Proton transfer; Electron transfer; Charge stabilization

1. Introduction

Many transmembrane proteins work to control the passage of charges across cellular membranes. Some proteins use the energy stored in sunlight or reduced substrates to generate electrochemical gradients. Others use the stored energy for vital cellular functions such as substrate transport and ATP synthesis. The reaction sequence and thermodynamics of many energy coupling, electron and proton transfer proteins are well established. However, how each protein carries out each reaction is less well understood. It has been difficult to address this question because few transmembrane protein structures are available. However, slow but steady progress is being made to obtain detailed structures of membrane proteins [1–7]. Each protein's structure should provide in principle all the information needed to determine in detail how it works. No method can currently use an atomic structure to derive all of the rates and thermodynamic parameters of a reaction. However, methods of analysis are being developed that can begin to calculate experimentally measurable properties in individual proteins. These analyses begin to identify protein design features that control reactions.

Much of the initial analysis of proton and electron transfer considers the energy barriers for moving charges out of water, focusing on the stabilization of charges in solvents with a high dielectric constant.

0005-2728/00/$ – see front matter
PII: S0005-2728(00)00060-8

This energy is lost when a charge is moved into the interior of a protein or membrane. Several classic papers describe how charges will be destabilized in a medium with the dielectric response of protein or membrane, highlighting the affect on ion transfer through membranes and on redox reactions in proteins [8–13]. However, proteins are not simple solvents and they have an array of mechanisms to compensate for the lost solvation energy in order to stabilize buried charges. Thus, interactions with specific elements of a protein can balance the cost of removing a charge from water. Electrons or protons moved into proteins can interact with fixed protein charges and dipoles and they can also cause protein conformation changes to stabilize the new charge. Electron and proton transfer reactions can also be coupled together to reduce the net change in charge of a reaction. The balance between these mechanisms for stabilizing buried charges is controlled by each protein's structure. We will describe here a pragmatic approach for obtaining information about how the structure of a protein determines the thermodynamics of intra-protein electron and proton transfer reactions. Detailed analysis of several charge transfer reactions in bacterial photosynthetic reaction centers will be described.

2. How to start to calculate the free energy of electron and proton transfer from the distribution of protein microstates at equilibrium

A calculation of the free energy of a reaction requires being able to estimate the energy of the reactant and product states from the protein structure and the underlying chemistry of the reaction. However, the protein will not have a single structure. Every protein has many degrees of freedom and changes in structure are introduced by fluctuations of atoms around their equilibrium positions. The distribution of structures will change when a reaction occurs. Motions are restricted by bonding constraints and by non-electrostatic and electrostatic forces between non-bonded atoms. Waters, ions, and other ligands distribute themselves between solution and protein binding sites. Acid and basic groups can bind or lose protons, while electron donors and acceptors can be oxidized or reduced. Microstates of a protein can be defined by the positions of all atoms and ionization states of all groups. Each microstate energy (ΔG_i) can be estimated from the potential functions for bond length and angle, electrostatic and van der Waals interactions between non-bonded atoms and the energy for proton, electron, or ligand transfer from solution or reaction partners at ambient pH, E_h, and ligand concentrations. In an ensemble of protein molecules, microstates with lower energy will be occupied more often. At equilibrium the probability of a given microstate (m) in a Boltzmann distribution is:

$$\langle m\rangle = \frac{\exp^{-\Delta G_m/k_B T}}{\sum_{i=1}^{M}\exp^{-\Delta G_i/k_B T}} = \frac{\exp^{-\Delta G_m/k_B T}}{Z} \quad (1)$$

where $\langle m\rangle$ is the fraction of the total population in microstate m, M the total number of microstates and the denominator (Z) is partition function of the system.

Eq. 1 can be used to connect the protein's structure to experimental values. Thus, we might ask if at equilibrium a particular amino acid is ionized, if a ligand is bound, if a group has the appropriate orientation to perturb a spectral marker. The fraction of the protein with a particular property, i.e. the probability of measuring property 'A' (ρ_A) in an equilibrium ensemble of proteins is:

$$\rho_A = \frac{\sum_{i=1}^{M} x_i(A)\exp(-\Delta G_i/k_B T)}{Z} \quad (2)$$

$x(A)$ is a vector where component ($x_i(A)$) is 1 if property 'A' is present in microstate i and 0 if it is absent.

While Eq. 2 is simple, it is of limited utility since it requires calculating the energy for all M microstates. As will be described in more detail below, the number of microstates of any given protein is very large. Monte Carlo sampling allows $\langle m\rangle$ and ρ_A to be estimated for states with low energies and significant population without enumerating all microstates [14–17]. Results can also be obtained from approaches which divide the protein into weakly interacting clusters which are small enough that all microstates in

each cluster can be enumerated [18,19]. However, in order to use the protein's structure to calculate interesting properties it is always necessary to both reduce the number of microstates that must be considered and to limit the number of terms in the analysis of the microstate energy.

The free energy of a reaction is established by the relative energy of the microstates, which satisfy the conditions that make up the observable state. A reaction can be initiated by a perturbation, such as a change in reactant concentration or in the temperature, which changes the most probable observable state of the protein. The free energy of a given observable state (A) can be calculated with respect to a reference energy by [20]:

$$G_A = -k_B T \ln \sum_{i=1}^{M} \exp(-\Delta G_i / k_B T) \equiv k_B T \ln Z \quad (3)$$

in a calculation where only microstates that contain site A are allowed. Eq. 3 can be used to calculate the free energy of a given reaction [17]. For instance, in bacterial photosynthetic reaction centers (RCs) equilibrium is established between the primary (Q_A) and secondary (Q_B) quinone being reduced. The relative free energy of the Q_{A^-} observable state can be calculated with Eq. 3 if only microstates with Q_A reduced and Q_B neutral are included. The relative free energy of macrostate Q_{B^-} is calculated by Eq. 3 keeping only microstates where Q_A is neutral and Q_B reduced (see [17] for a more complete description). The free energy of the reaction $A \rightarrow B$ can be obtained as:

$$\Delta G_{A \rightarrow B} = G_B - G_A \quad (4)$$

where G_A and G_B are the free energies of macrostates A and B, relatively, calculated with Eq. 3. Here the calculation of the free energy of the reaction requires calculation of the two macrostate energies.

The ratio of the probability of the protein being in different observable states (ρ_A or ρ_B) can also be used to determine the free energy difference between the states. This method is conceptually the same as obtaining the reaction free energy from the measured equilibrium constant. Thus:

$$\Delta G_{A \rightarrow B} = -k_B T \ln \frac{\rho_B}{\rho_A} \quad (5)$$

For both experiment and calculation this procedure works best if observable states A and B are present in non-negligible amounts. However, in calculation the accuracy can be improved by the addition of a constant to the energy of all microstates with property A or B. This will artificially bias the distribution of states in Monte Carlo sampling to obtain ΔG_{AB} even when it is large [21,17].

3. What microstates will be included in the analysis

3.1. Reducing the number of microstates considered in a static, continuum electrostatics analysis

Electron and proton transfer reactions change the charge state of protonatable or redox active sites in a protein. The coupling between electron and proton transfers occurs because the charge on one site changes the energy of another. There are few redox cofactors in any protein, However, in an average protein approximately 25% of the residues are ionizable (Asp, Glu, His, Arg, and Lys). Methods have been developed to calculate the equilibrium distribution of ionization states at a given pH assuming that microstates of the protein differ only in the charge on ionizable sites [22,18]. Only minor modifications are required to include changes in charge on redox cofactors in the analysis [17,21,23–25]. If N sites can take two forms (charged and neutral) 2^N microstates will enumerate all possible combinations ranging from all sites neutral to all charged. Limiting the protein to this group of microstates assumes that all other contributions to the microstate energy are unchanged in all ionization states of the protein.

In the static, continuum electrostatic model all microstates have the same atomic positions and so the interaction energies amongst the non-ionizable residues are constant. Additions to microstate energies that are identical in all microstates will not influence equilibrium distributions and so can be ignored. However, when charges are moved in any condensed media things move. Even in a non-polar solvent the molecular electron distribution is polarized by charges inducing microdipoles. In a polarizable medium like water and to a lesser extent protein new charges can also induce changes in atomic positions. Using the equations of continuum electrostatics, the

effect of these perturbations are averaged in the dielectric constant. As will be described below, the polarization of the media stabilizes ensembles of charges through the reaction field energy. In addition, the motion of the surroundings has the effect of screening the interactions between charges.

Thus, a continuum electrostatics methodology has been used to capture some of the responses of the protein to charge changes. Only microstates which differ by the charge state of ionizable residues are explicitly considered. This can still be a very large number of states. For example, in *Rb. sphaeroides* RCs there are more than 100 ionizable residues, resulting in more than 10^{30} microstates. The probability of different residues being ionized is generally determined by Monte Carlo sampling methods. These use random sampling of many microstates with appropriate acceptance criteria. Microstate populations that approach a Boltzmann distribution are found [14].

3.2. Increasing the number of explicit degrees of freedom: hybrid methods

Comparing microstates with no differences in atomic positions limits the information that can be obtained. Thus, no specific, local motions coupled to charge movements can be investigated. One long term goal might be to be able to calculate ionization equilibrium as a function of the solution pH in a molecular dynamics simulation where all atoms are allowed to move. However, calculations of this magnitude are not currently possible. Various more limited methods are being developed that allow some changes in atomic position during calculation of the distribution of ionization states [26–28].

Simultaneous calculation of both ionization and conformation states has been implemented in a number of methods (see [29] and reviews [26–28]). The earliest techniques either averaged interactions between different possible side chain atomic positions [30] or averaged the ionization states calculated in a small number of low energy structures [31,32]. Recently, an iterative mobile cluster approach has been used to calculate multiple site ligand binding to flexible macromolecules [19]. New methods combine conformation changes and calculations of ionizable states [16,17,33].

4. Calculation of microstate energies

Whatever microstates are used in the calculation, the relative energy of each must be determined. Both the energy of the underlying reaction chemistry as well as how changing charges and atomic positions change the protein's energy must be considered.

4.1. Reference states

The calculation described here asks how the free energy of a reaction is changed moving from a reference solvent into the protein, i.e. how the protein perturbs an electron or proton transfer reaction that has been previously characterized. The reference energy may be measured experimentally or modeled computationally. For an acid or base this is the pK_a in solution; for a redox cofactor this is the electrochemical midpoint (E_m); the hexane or vacuum to water partition coefficient can provide the reference free energy for binding affinities [34–37]. Equilibrium constants in the protein are derived from comparison of the energies to transfer reactant and product from a reference solvent into the protein.

Biological electron and proton donor/acceptor chemistry is carried out by a modest number of groups. The acidic and basic amino acids are the most common pure proton donors and acceptors. Polar groups are often involved in proton transfer reactions as links in proton channels. Tyr and Cys also play a role in redox reactions, however, electron transfer reactions generally involve non-amino acid moieties. Thus, iron-sulfur centers and phorphryn derivatives such as hemes and chlorophylls, function as non-protonatable redox sites. Quinones and flavins have more complex electrochemistry as they can be singly and doubly reduced and can bind up to two protons on reduction. The examples discussed in this review will deal with reactions involving quinones, bacteriochlorophylls, and hemes in reaction center proteins of photosynthetic bacteria (RCs).

4.1.1. Experimental measurement of reference reactions

When both reactants and products are stable in a particular solvent the reaction can be measured experimentally. This will provide the equilibrium constant for the reaction of interest under some

standard conditions. Water is preferred for the reference solvent because parameters to model this solvent have been extensively tested [38]. The transferability of the parameters that describe the protein or ligands into other solvents must to be tested for each case [39].

4.1.1.1. Amino acids. When proton transfer is coupled to electron transfer, it is the amino acids of the protein that are often the proton donors and acceptors. The pK_as of the acidic and basic amino acids have been determined for the isolated side chains as well as for the amino acids in short polypeptides in water [40].

4.1.1.2. Hemes. Hemes function as electron transfer donors. The most common redox couple is between a neutral heme with Fe^{+2} and a heme with a plus 1 charge (Fe^{+3}), in each case the porphryn has a charge of −2. Hemes in cytochromes do not exchange ligands during reactions. In other proteins a substrate may be bound to the iron center as it undergoes redox chemistry. To evaluate the electrochemistry of all hemes of biological interest different reference values are needed for hemes with different ligands such as for five coordinate hemes (cytochrome *c′*) or with different oxygen intermediates (cytochrome oxidase). Solution electrochemistry has been measured for hemes with two His ligands (bis-His) and for hemes with His and Met ligands [41–43]. In water, a heme with a His and Met as ligands has an E_m of −70 vs. S.H.E. The His are more electron donating than Met, lowering the redox potential by 150 mV to −220 mV in the bis-His hemes. Systematic studies do provide some information about how different ligands change heme E_ms in solution [44,45].

4.1.1.3. Quinones. Quinones have a more complex chemistry than hemes. The chemistry of quinones in proteins shows the ability of proteins to stabilize reactions that are very unfavorable in solution. In aqueous solution, in the intermediate pH range quinone is reduced directly to the dihydroquinol by the coupled addition of two electrons and two protons [46]. However, quinones in proteins are usually reduced in single electron steps, often without proton binding [47,48]. Extensive data are available for the electrochemistry of quinones in the aprotic solvent dimethyl formamide [49].

4.1.1.4. Chlorophylls, bacteriochlorophylls, pheophytins, and bacteriopheophytins. The electrochemistry of these groups has been determined in a variety of solvents and aggregation states [50] (see [51] for a review).

4.1.2. Ab initio calculations of the reaction

When there are no available experimental reference measurements, state energies calculated in vacuum can also provide a reference free energy. Vacuum is a well characterized 'solvent' which has no influence on the reaction. It is straightforward to calculate the perturbation of the reaction caused by the protein. The problem with using calculated state energies is that they may not be very accurate. Examples of these calculations can be found for the redox potentials of the various redox states in the photosynthetic reaction centers of bacterial [52] and in photosytem II of green plants [53].

4.1.3. Reaction involving the same ligands

The difference in E_m or pK_a of identical cofactors in different sites can be calculated without having a reference reaction free energy in the absence of the protein. This method is particularly useful when identical cofactors are found in different locations in the same protein. For example, there are chemically identical cofactors found along a c_2 symmetry axis in bacterial reaction centers, although only one side of the protein is active [54–57], there are four hemes with different E_ms in one protein subunit in *Rps. viridis* RCs [23], and ubiquinones with different properties in RC Q_A and Q_B sites [17,21,25]. Analysis can help determine how the protein creates the differences in cofactor behavior in each site.

4.2. Parameters used to calculate microstate energies in proteins

In order to calculate how the protein changes the reaction free energy, atomic charges and radii must be assigned to each atom in the amino acids, redox cofactors, substrates, and bound ions and waters in the structure. Additionally, a dielectric constant must be assigned for both protein and solvent.

4.2.1. Charge distribution

In the most basic atomic level description of a protein for computational purposes a radius and a partial charge must be assigned to each atom. In addition, bonded energy terms describe the energies of different bond lengths and torsion and dihedreal angles. The atomic charges are distributed over residues and associated cofactors and substrates to achieve an appropriate net charge. Ionized groups have a non-zero net charge. For example, the net charge is +1 for ionized bases such as Arg, Lys, or His and −1 for acids such as Asp, Glu, or Tyr. Proton transfer adds or removes an atom, the proton, and redistributes the charges on the side chain atoms. Simplified charge distributions have been used to minimize the effects of unknown structural features. For example, early studies often did not explicitly place a proton on neutral acids but used reduced and identical charges on both oxygens [58]. When the positions of these protons can be chosen dynamically, explicit protons with appropriate partial charges can be used [16]. For convenience, parameter sets often keep the same charges on the backbone of all residues except Pro. The net charge on the atoms that makes up the backbone is zero, although there are significant partial charges on the individual atoms. The side chain charges are zero for all but ionized acids and bases. The atomic charges are small for non-polar side chains and significant (as large as ±0.5 e.u. on individual atoms) for polar residues.

4.2.2. Atomic radii and van der Waals energies

The van der Waals forces determine the short range interactions between non-bonded atoms [59]. A short range repulsion and longer range attraction keep atoms at the appropriate distance from each other. In molecular mechanics simulations this energy term is described by:

$$\Delta G_{vdw} = \frac{A}{r^{-12}} - \frac{B}{r^{-6}} \quad (6)$$

where A and B describe the repulsive and attractive interactions. The coefficients used are dependent on the type of atom (e.g. carbon to carbon or oxygen or hydrogen).

In molecular mechanics calculations the interactions between atoms with non-zero charges create forces that influence motion. These motions are influenced by the bonded (bond length and torsion) and non-bonded (van der Waals) interactions. In continuum electrostatics the location of the charges defines the total electrostatic energy of the system and the change in energy caused by change of the charge during a reaction. Pure continuum electrostatics calculations use static structures with no explicit changes in atomic structure in response to proton or electron transfer. Therefore, the non-electrostatic, position dependent energy terms are the same in all ionization states so these have no effect on the reaction. Atomic radii are important in continuum electrostatics calculations since they define regions with different dielectric constants [60]. In hybrid approaches such as the MCCE method, changes in atomic position can accompany ionization changes requiring that electrostatic and non-electrostatic energies must be considered.

4.2.3. Sources of parameters

Parameters for continuum electrostatics calculations have been modified so that they can reproduce water to vapor or hydrocarbon transfer energies for polar molecules [61,62]. The free energy of transfer can be calculated from the energy of forming a cavity and the interaction of the water with the charges in the solute. The first term is from the hydrophobic transfer energy for non-polar solutes; the latter from the reaction field energy described in the next section. The PARSE parameter set optimized atomic charge and radius to fit transfer energies and these provide calculated pK_a values that compare favorably with other parameter sets [58]. Other atomic parameters are also used [63–67].

Knapp and colleagues have used ab initio calculations at the level of the G-31G** basis set to derive quinone partial charges [25]. Partial charges are provided for ubiquinone in the six states: oxidized: UQ, semiquinone: UQ^-, UQH, fully reduced: UQ^{-2}, UQH^-, and UQH_2; as well as for menaquinone in the oxidized (MQ) and semiquinone (MQ^-) states. The atomic partial charges of high spin non-heme iron and its ligands were calculated by a density functional method. Charges for bacteriochlorophylls (neutral, oxidized and reduced) and bacteriopheophytins (neutral and reduced) are available from the work of Parson and Warshel [54].

Non-electrostatic parameters for bonded and non-bonded interactions have been optimized for use in molecular mechanics. This provides energies for bonded atoms as a function of distance and angle and for non-bonded atoms at different distances (e.g. [63–67]).

4.3. Dielectric constant

Perhaps the most contentious parameter used in continuum electrostatics calculations is the dielectric constant [12,28, 68–70]. When there is a change in charge state in a condensed medium the surrounding electrons and nuclei are polarized. This can be called a dielectric response. The dielectric constant is used in continuum electrostatic calculations to average the influence of these small motions on the resultant electrostatic energies. While the dielectric constant of the bulk solution can be obtained experimentally [71,72], its value inside native proteins can not be measured directly. This is because electrostatic interactions are strongly influenced by both the protein and surrounding solvent water. Even if a protein had a uniform, low dielectric constant (ε_{in}), an experimental 'effective dielectric constant' which measured the pairwise interactions between two sites would always be larger than ε_{in} and would vary as the distance between the sites and to the protein surface changed. Continuum electrostatics analysis is a powerful tool for analysis of proteins largely because it allows the significant impact of water on electrostatic energies to be determined without extensive calculation of the water itself. The surrounding water is routinely assigned the bulk dielectric constant of water (78 or 80). As described below waters buried within proteins can be treated explicitly in atomic detail.

There is real difficulty in assigning an average value of the dielectric constant for protein. The measured dielectric constant of dry protein powders, when the solution phase is removed is ≈ 4 [73]. This may underestimate the appropriate value in native, hydrated protein since increased atomic flexibility increases ε. A protein's electrostatic response will also be pH dependent, as the pH changes the ionization state of groups change. The motion of charged residues can provide a substantial increase in the effective dielectric response [74–76].

Several values have been used in electrostatic analysis, each representing a different degree of averaging the response to changes in charge. A dielectric constant of 2 accounts for electronic polarization of the protein atoms. A dielectric constant of 4 adds the polarizability originating from small scale microdipole motions in the protein [68,69]. Much larger values have been suggested by molecular dynamics simulations especially in the vicinity of the protein surface [74–76]. Fluctuations of the ionized side chains of groups outside the protein core make the largest contributions. A high protein dielectric constant has been supported by comparison of calculated and experimental pK_as in several proteins. A value of $\varepsilon = 20$ for the protein provides the best results [77,78].

Using a single number for the dielectric constant of the whole protein averages all of the protein's dielectric response. All atoms, including internal waters are more or less well packed in the protein interior and their flexibility is restricted. However, proteins often undergo functionally important conformation changes [79–81]. Thus, it is important to be able to differentiate between regions with different mobility. For example, in RCs, the Q_A site is much more rigid than the Q_B site and there are very different changes in the protein when each quinone is reduced [17]. Thus, the use of a dielectric constant that is not position dependent is a major source of error in standard continuum calculations. It is therefore important to modify the analysis to capture the spatially varying dielectric response inside a protein [82–84].

One way to account for a non-uniform response of the protein to changes in charge while still using the formalism of continuum electrostatics is to use a non-uniform dielectric constant [84]. Here there are no explicit changes in the protein atomic positions, however averaged motion of different side chains is modeled by giving each a different dielectric constant. For example if side chains are allowed to freely rotate with no restrictions due to neighboring groups it is possible to calculate the average dipole moment and mobility of each residue type and therefore to estimate its local dielectric constant [85]. This type of analysis finds the largest local dielectric constant for Asn and Gln (≈ 72) and smallest for Ile, Leu and Val (≈ 3.1). This range of values highlights how inhomogeneous proteins can be.

Another way to account for inhomogeneity of the dielectric properties of proteins is to use hybrid methods. These consider some protein motions explicitly while the effects of other motions remain averaged in a dielectric constant used for continuum calculations of electrostatic energies [16,17,33]. Some of the dielectric response comes from induced dipoles and from very small displacements of already existing dipoles. An average dielectric constant of 4–6 can capture the effect without needing to explicitly calculate all individual motions [68]. However, larger motions can be explicitly enumerated in microstates (Eq. 1) with different atomic positions. Recent studies find that the local dielectric constant of proteins calculated with hybrid methods depends on the position of interacting sites in the protein as well as on the pH (Alexov and Gunner, unpublished results).

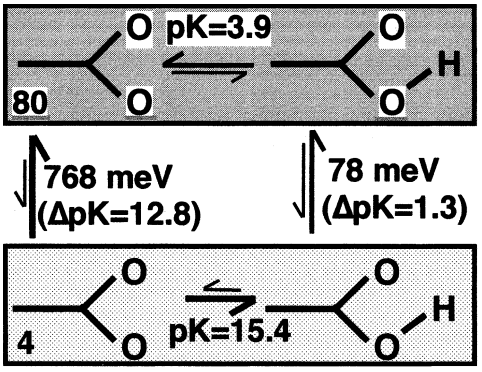

Fig. 1. The free energy cost of burying a charge. The change in pK_a of an Asp side chain ($\varepsilon_{in}=4$) as it is moved from water ($\varepsilon_{out}=80$) to a solvent of $\varepsilon_{out}=4$. The solution pK_a is the reference energy. Combining this with the calculated loss of reaction field energy provides the pK_a in the new environment. The loss of the favorable interaction of the charge with the high dielectric solvent (ΔG_{rxn}) destabilizes the charge, shifting the pK_a by 11.5 pH units, a ΔG of 690 meV or 15.6 kcal/mol. This large, favorable energy is lost in a solvent with restricted dipole flexibility [8,9] as shown by the negligible partition coefficients for charged groups in non-polar solvent [35,141]. Atomic radii and partial charges from [16].

4.4. *Calculating how the reaction free energy changes when the reaction occurs in different media*

4.4.1. *Pure dielectric media with no charges or explicit atomic dipoles*

Water has a high dielectric constant while proteins have a much smaller ability to rearrange around charges and so have a smaller dielectric constant. Membranes also have low dielectric constants. The energy of transferring a charge into an environment with a lower dielectric constant is unfavorable and can be large. Using the continuum electrostatics formalism a simple, analytical expression can be derived to calculate the free energy to transfer a charge from a medium with dielectric constant (ε_1) to another (ε_2) [86]. If ε_2 is larger than ε_1 this process is favorable. If the charged group is a sphere the energy of transfer from ε_1 to ε_2 is:

$$\Delta G_{rxn} = c\frac{q^2}{2r}\left(\frac{1}{\varepsilon_2}-\frac{1}{\varepsilon_1}\right) \quad (7)$$

This energy term ($q^2/2r\varepsilon$) is referred to as the solvation, Born, self or reaction field energy. If q is in multiples of the charge on an electron and r in Å C is 14.4 eV or 331 kcal/mole. Moving a 5 Å spherical charge from a region of dielectric constant 80 to 4 would destabilize the ionized form of the ion by 8 kcal/mole (or 340 meV). Increasing the ion charge, or decreasing the ion radius or the dielectric constant of the low dielectric region increases the penalty. Numerical solutions of the Poisson equation provide ΔG_{rxn} when the dielectric boundary has an arbitrary shape as in a protein. The Poisson–Boltzmann equation provides values when the solvent contains dissolved salt.

The equilibrium constant for reactions involving charge changes will be modified by changing the solvent. Thus, if a reaction is moved into a medium with a lower dielectric constant, the reactants with smaller charge will be destabilized less. The equilibrium constant will shift, favoring the neutral species. The shift in E_m or pK_a can be estimated [8–11]. For example in Fig. 1, the estimated penalty for moving the polar, neutral Asp from water into a medium of $\varepsilon=4$ is 78 meV while the penalty for the ionized Asp is 768 meV. In the low dielectric medium the Asp would not be half ionized until a pH of 15.4, an 11.5 pH unit shift from its value in water. Thus, proteins must have other electrostatic interactions if any charges are to be buried in proteins, especially in transmembrane proteins.

4.4.2. Dielectric media with fixed charges. The contribution of the charge distribution in the protein

The loss of solvation energy when a charged group moves into a protein would seem to ensure that there should never be any charged groups within proteins. If Asp pK_as were 15 these residues would always be protonated. However, an analysis that focuses only on the loss of reaction field energy ignores interactions amongst charged and polar groups within each protein. In fact, these interactions become more important when the dielectric constant is small, just the conditions where the loss in reaction field energy becomes significant. The interactions between the charges localized on fixed atoms can be considered within the continuum electrostatics methodology. The interaction between two charges is:

$$\Delta G_{\mathrm{crg:crg}} = Cq_i(\Psi_{j\rightarrow i}) \underset{\text{uniform } \varepsilon}{\longrightarrow} \frac{Cq_iq_j}{\varepsilon r_{ij}} \tag{8}$$

where q_i and q_j are the charges assigned to the atoms i and j respectively, $\Psi_{j\rightarrow i}$ is the potential at i from q_j, C is 14.4 eV or 331 kcal/mole. When the dielectric constant is uniform the potential at the cofactor and the resulting energy of interaction can be obtained by Coulomb's law where ε is the dielectric constant and r_{ij} is the distance between the charges. When the dielectric constant is not uniform the Poisson equation can be used to determine $\Psi_{j\rightarrow i}$.

The importance of charges depends very much on the dielectric constant. Using Coulomb's law a charge 10 Å from another would modify the free energy of a charge by 720 meV (16.5 kcal/mol) if $\varepsilon = 2$, a very large impact at a long distance. However, if the dielectric constant were high (e.g. 80) the interaction energy would be only a relatively insignificant 19 meV (0.4 kcal/mol).

4.4.3. Dielectric media with permanent dipoles

Homogeneously distributed dipoles will produce no net, static potential in the absence of a charge. The impact of these dipoles rearranging around a charge is included in the dielectric constant. However, in proteins there are dipoles or multipoles with fixed orientation which produce a static, non-zero potential that affects the energy of charges without the dipoles changing position. These dipoles are better treated as groups of two or more charges where the net charge is zero. Thus, the interaction energy can be calculated as in Eq. 8, adding the impact of each partial charge in the polar group. Thus, the interaction energy between a charge (i) and the charges on a polar group where two atoms (j and k) have equal and opposite partial charge is given by:

$$\Delta G_{\mathrm{crg:dip}} =$$

$$Cq_i(\Psi_{j\rightarrow i} + \Psi_{k\rightarrow i}) \underset{\text{uniform } \varepsilon \; r(i\rightarrow j:k)>r(j\rightarrow k)}{\longrightarrow} C\frac{r_{jk}}{\varepsilon r^2_{i(j:k)}} q_iq_j\cos\theta \tag{9}$$

Since i and j have opposite signs, $\Psi_{j\rightarrow i}$ and $\Psi_{k\rightarrow i}$ will also be of opposite sign, diminishing the total interaction energy. r_{jk} is the distance between i and k. $r_{i(j:k)}$ is the distance from i to the midpoint of the line between j and k, and θ is the angle between the vectors along $r_{i(j:k)}$ and r_{jk}. If r_{jk} is small relative to $r_{i(j:k)}$ then the absolute values of $\Psi_{j\rightarrow i}$ and $\Psi_{k\rightarrow i}$ will differ by only a small amount and $\Delta G_{\mathrm{crg:dip}}$ will be small. When $r_{i(j:k)}$ is several times r_{jk} in a medium of uniform dielectric constant, the importance of the dipole falls off as r^{-2} so dipoles have small impact at long distance. For example, a dipole with r_{jk} of 1.5 Å, charges of ±0.5 and a midpoint ($r_{i(j:k)}$) 10Å from the cofactor, aligned for maximum effect ($o = 1$), would yield a $\Delta G_{\mathrm{crg:dip}}$ of 53 meV (1.25 kcal/mol) if the dielectric constant were 2. However, when r_{jk} is large relative to $r_{i(j:k)}$, such as in a hydrogen bond, the interaction energy between a charge and a dipole can be important.

5. Calculation of the distribution of ionization states in a protein

The pH and E_h dependent properties of a protein and the coupling of electron and protein transfer reactions depend on the relative energy of all ionization states. The energy terms described above allow these energies to be estimated for any particular ionization state. If all microstates of a system can be described and their relative energy calculated then in theory all experimental, equilibrium properties

can be calculated. We will show below how Eqs. 1–4 yield exact, and familiar solutions for small systems where all microstates can be enumerated. More generally, statistical techniques are used to approximate the Boltzmann distribution of microstates by sampling a minuscule fraction of all microstates appropriately biased towards the low energy states by Metropolis sampling [14].

5.1. Defining the free energy terms for a single residue in solution

With a single ionizable residue there are $2^1 = 2$ states (charged (c) and neutral (n) i.e. the group can be either charged or neutral). The balance between these two forms is controlled by the solution pH. Any state can be defined as being the reference with an energy of zero. The energy of the neutral state in solution is assigned this role. Then, the microstate energy (in units of $kT = 28.5$ meV = 0.59 kcal/mol) of the ionized form differs from zero by $2.3\gamma(\mathrm{p}K_a - \mathrm{pH})$ (where γ is 1 for an acid and -1 for a base). This $\mathrm{p}K_a$ is the value measured in solution. As required, the site is half ionized when $\mathrm{p}K_a = \mathrm{pH}$ since here the free energy of ionized and neutral states are equal. Thus, in a system with only two microstates, ionized and neutral, the microstate energies can be defined and the partition function (Z) calculated:

$$\begin{aligned} \Delta G_n &\equiv 0 \\ \Delta G_c &= 2.3\ \gamma(\mathrm{p}K_a - \mathrm{pH}) \\ Z &= e^{-\Delta G_c} + e^{-\Delta G_n} \end{aligned} \tag{10}$$

Then, the occupancy of the charged form $\langle c \rangle$, which is simply the fractional ionization of the site, can be calculated as in Eq. 2:

$$\langle \rho_c \rangle = \frac{e^{-\Delta G_c}}{Z} = \frac{10^{-(\mathrm{p}K_a - \mathrm{pH})}}{1 + 10^{-(\mathrm{p}K_a - \mathrm{pH})}} \tag{11}$$

In Eq. 11 the exponent (e) is replaced by the more familiar powers of 10 using $10 = \exp(2.3)$. The ionization of this group in a hypothetical protein, i.e. a protein without other titratable groups, will be perturbed by two extra energy terms: the change in reaction field energy and the interaction of the charge with the explicit dipoles in the system. The free energy of the charged (c) and neutral (n) microstates are now:

$$\begin{aligned} \Delta G_n &= 0 + \Delta G_{n,rxn} + \Delta G_{n,pol} \\ \Delta G_c &= 2.3\gamma(\mathrm{p}K_a - \mathrm{pH}) + \Delta G_{c,rxn} + \Delta G_{c,pol} \end{aligned} \tag{12}$$

where the latter two terms are the changes in energy moving each state into the protein. The ΔG_{rxn} is the difference in desolvation energy in solvent and in the protein (Fig. 1). This value can be calculated by continuum methods or from the interaction with explicit water molecules. ΔG_{pol} is the interaction of the ionizable site with explicit atomic dipoles such as those in the protein backbone. Interactions with specific, unchanging charges are also included in this term. The pH at which the ionized and neutral states have the same energy (the pK) changes from the reference, solution value by $\Delta\mathrm{p}K$:

$$2.3\Delta\,\mathrm{p}K = \Delta G_{c,rxn} - \Delta G_{n,rxn} + \Delta G_{c,pol} - \Delta G_{n,pol} \tag{13}$$

Thus, the $\mathrm{p}K_a$ of this residue in this system with two microstates (often called pK intrinsic [18,22]) is now:

$$\mathrm{p}K_{int} = \mathrm{p}K_{sol} + \Delta\,\mathrm{p}K \tag{14}$$

In the calculation of electrostatic energies, redox centers and acids and bases are formally equivalent. The change in reaction field energy for ionized and neutral forms of reactants and products are calculated in the same manner. The pairwise interactions between charged groups is again the same. For proton transfer the $\mathrm{p}K_a$ is the reference energy, while it is the E_m for electron transfer reactions. In these equilibrium calculations the impact of the solvent on the state energies is contained in the pH or E_h term. Thus, when the distribution of ionization states of electron donors and acceptors are considered, the $\mathrm{p}K_a$ in Eqs. 10–14 is changed to E_m and pH to E_h.

5.2. The distribution of ionization states in systems with several ionizable sites

The coupling between electron and proton transfers occur because surrounding protonatable groups change their ionization state in response to the reaction. This can be determined given the relative energy of the explicit microstates of the system. If there are N sites, which can change only their ionization state then there are 2^N states. Eq. 15 describes the energy

terms that define the energy differences between each state. Again, the neutral form of each site in solution is taken as the reference energy. As there are no degrees of freedom other than ionization, torsion energies and van der Waals interactions are assumed to be unchanged by reaction and so do not add to the difference in state energies. Thus, the ith microstate energy is given in general as:

$$\Delta G_i = \sum_{m=1}^{L} \delta(m)\{2.3\gamma_m\ (\mathrm{p}K_{m,\mathrm{sol}} - \mathrm{pH}) + (\Delta G_{m,\mathrm{rxn}} + \Delta G_{m,\mathrm{pol}}) + \sum_{k=m+1}^{L} \delta(k)\Delta G_{m,k} \quad (15)$$

As written this sum runs over both charged and neutral forms of each ionizable group so $L = 2 \times N$ (see [16] for a more complete description). The neutral forms of ionizable residues have zero net charge but non-zero dipole moment. Here $\delta(m)$ is 1 if m, an ionized or neutral form of a given residue, is present in microstate i, or 0 if it is not; γ_m is 1 for ionized acids, -1 for ionized bases and 0 for neutral forms and $\Delta G_{m,k}$ are pairwise interaction energies (Eqs. 8, 9 or 17 (below)) between (ionized or neutral) residue m and (ionized or neutral) residue k. $\Delta G_{m,k}$ contributes to microstate energy only if $\delta(m)$ and $\delta(k)$ are both 1 in microstate i.

The fraction of a given residue (j) that is ionized at equilibrium can be calculated using:

$$\langle \rho_j \rangle = \frac{\sum_{i=1}^{M} x_i(j)\exp(-\Delta G_i)}{Z} \quad (16)$$

where $M = 2^N$ runs over all microstates and $x_i(j)$ is a vector of length $2N$, with a value of 1 if j is ionized or zero if it is neutral in microstate i. Familiar patterns of pH dependence of coupled acids and bases are recovered from these calculations for small numbers of sites (Eqs. 10, 11). However, the explicit microstate energy description can be extended to any number of ionizable residues.

5.3. The energy of groups of charges: linear combination of self and pairwise energies

The microstate energies described above make an implicit assumption that the state energies can be described as a linear sum of self and pairwise interaction energies. Thus, in Eq. 7 the reaction field energy of a site, be it charged (c) or polar, neutral (n) is calculated without consideration of the ionization state of any other site. The same is true for the polar interactions. Fig. 2 shows the energy of a pair of charges in a spherical 'protein'. Two different pathways for assembly are contrasted. In the first (clockwise) the charges are brought together in a medium with no dielectric boundaries; the dipole is then immersed in solvent. In the second (counterclockwise) assembly path, the charges are individually brought into the 'protein' and their interactions with each other added. The total energy of assembly, going from two separate charges in a uniform medium to the pair of charges in a low dielectric region surrounded by water is the same. This would not be the case if the shape of the low dielectric sphere changed each time a charge was added. Then each added charge would influence the reaction field energy and pairwise interactions of all others. For example, if adding a new charge increases the distance of two charges from the surrounding water, these two sites will loose more reaction field energy and will also have larger pairwise interactions with each other and will have larger interactions with the explicit dipoles.

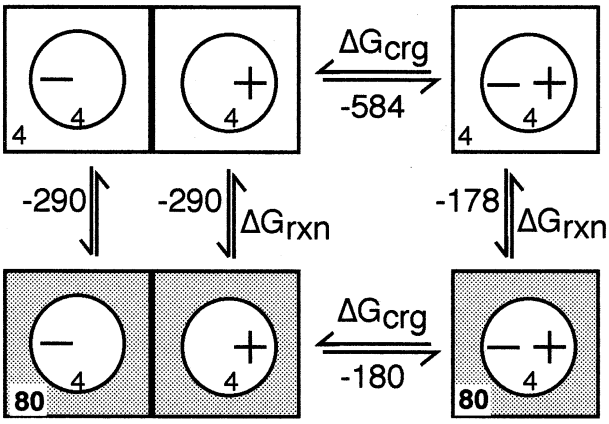

Fig. 2. Electrostatic free energy of assembling two charges. Two charges infinitely far apart ($\varepsilon_{in} = \varepsilon_{out} = 4$) are moved into a 7 Å sphere with $\varepsilon_{in} = 4$ and $\varepsilon_{out} = 80$. The charges are 6 Å apart, 3 Å from the sphere center. Each transfer energy is in meV. The two directions around the thermodynamic box yield the same total energy within the accuracy of the finite difference DelPhi Poisson–Boltzmann calculations (at 0.5 Å/grid) [87].

Thus, Eq. 15 calculates the energy of each ionization state as a linear sum of self energy (ΔG_{rxn}) and pairwise ($\Delta G_{pol}+\Delta G_{k,m}$) terms. No terms where the ΔG_{rxn} of a site depends on the ionization state of another are considered. Nor are terms where the interaction between sites depends on the charge at a third site. Thus, all state energies can be calculated from a look-up table that contain these self and pairwise energy terms [16]. These state energies can be calculated in a compact way. The Poisson–Boltzman equation is solved $2N$ times for N sites. One for the neutral polar form of a residue and one for the ionized form. Each calculation has only a single site with any charges, either polar neutral or ionized. The reaction field of the site is determined from the interaction of the charge with the induced charges at the dielectric boundary between the protein and solvent [87]. At the same time all pairwise interactions between the site with any charges (x) and all other sites (y) can be calculated given:

$$\Delta G_{x:y} = \sum_{\text{atoms in } y} q_y(\Psi_{x\rightarrow y}) \quad (17)$$

Eq. 17 can also provide ΔG_{pol}. Here the sum runs over the explicit dipoles and permanent charges. The series of $2N$ calculations results in vectors of length $2N$ with the reaction field energies and with ΔG_{pol} and a $2N\times 2N$ matrix of pairwise interactions between all combinations of charged or polar, neutral form of each ionizable residue. All 2^N state energies can be determined from appropriate combinations of these energies.

5.4. *Coupling the reaction to the protein. Protein conformation and ionization changes accompanying electron and proton transfer*

In electron and proton transfer reactions there will be changes in electron distribution and nuclear positions as charges move into different locations in a protein. In the calculations of microstate energies described to this point the impact of electronic polarization and averaged nuclear motions are implicitly included in the dielectric constant used in the calculation of the electrostatic free energy terms (ΔG_{rxn} and the pairwise ΔG_{crg}). The result is that the energy of pairwise interactions is smaller in a dielectric medium than in vacuum. This accounts for an increase in the energy of the dielectric medium equilibrated around the charges that reduce the total energy of the pair of charges. Likewise, the calculated reaction field energy accounts for the unfavorable energy stored in the dielectric medium polarized around a charge and the favorable energy gained as this reorganized medium reacts with the charge. Thus, the dielectric constant includes some of the effects of all the microstates generated by the differently polarized medium without enumerating these microstates.

However, hybrid methods are being developed which allow explicit motion of some atoms. The multi-conformation continuum electrostatics procedure (MCCE) [17] allows multiple positions of hydroxyl and water protons, alternative side chain rotamers, water positions and cofactor positions in the calculation of the pH and E_h dependence of ionization equilibria [16,17]. The MCCE method starts with standard protein data bank files, but alternate atomic positions are automatically added. Each alternate position or ionization state of a residue, cofactor, water or ligand is termed a conformer. Each ionizable residue, cofactor, or ligand has ionized and polar, neutral confomers. Alternate hydroxyl protons are placed in their torsion minima and at hydrogen bonding positions. Asn and Gln have their side chains termini rotated by 180°, interchanging the terminal -NH_2 and -O. Crystallographic waters and ions on the protein surface are removed and replaced with the high dielectric surroundings. Buried waters are retained and are given confomers with different positions for their protons. In addition, waters and ions have a conformer with no interactions with the protein allowing them to move into the solvent. A penalty for water and ion binding to the protein can be added to microstate energies [37]. Lastly, residues involved in strong interactions are identified. These sites are most likely to change conformation when the interaction partner changes charge. Therefore, additional conformational flexibility is added for these sites. Studies of protein side chain conformations suggest that all orientations need not be considered. Thus, side chains have preferred torsion minimum [88,89]. Added side chain conformations are therefore restricted to orientations that are found in rotamer libraries and those which make specific, hydrogen bonds with neighboring residues. This reduces the search of conformation space.

Microstates are formed of proteins with one conformer chosen for each residue, cofactors, or ligand. In the MCCE method site energies are calculated as in Eq. 18 with the following differences: (1) In addition to the N ionizable sites there are now additional sites that have explicit degrees of freedom including side chains which can change their position but not their charge; and buried waters and ions that can move within a binding site or leave the protein. Ionizable residues can also change their position as well as their charge and so have more than two available confomers. If there are L sites which can change ionization state and/or position and l_x confomers for each site, the total number of microstates is:

$$\text{states} = \prod_{x=1}^{L} l_x \tag{18}$$

This is an enormous increase in the number of microstates over 2^N. (2) The changing atomic positions means that each site can have a different torsion energy and van der Waals energy in each microstate. The torsion energy is a self energy term which is independent of the position of other sites while the van der Waals energy is a pairwise energy that depends on the distribution of other sites in the microstate. The energy look-up table is now composed of vectors of length Σl_x each for the reaction field and torsion energies representing the electrostatic and non-electrostatic self energies; vectors of this length for pairwise electrostatic and van der Waals interactions with portions of the protein which are held static. Lastly there are two $\Sigma l_x \times \Sigma l_x$ matrices. One for the electrostatic pairwise interactions between conformers and the other for the conformer to conformer van der Waals energies. Microstate energies are then built up by addition of these energy terms. Monte Carlo sampling finds the Boltzmann distribution of microstates and the probability (occupancy) of each possible conformer. Thus, the position or ionization states of particular residues, or occupancy of water sites can be determined. Conformation and ionization changes are coupled through microstate energies which include both electrostatic and non-electrostatic interactions. Thus the MCCE method allows simultaneous changes in atomic position and ionization states as a function of pH and E_h.

As described above, the total energy of a system can be built up as a sum of the reaction field (self) energy of individual sites plus the pairwise interactions between sites if the changes introduced do not change the dielectric boundary (Fig. 2). In static continuum electrostatics calculations there are no explicit motions so there are no boundary changes. In the hybrid MCCE method this assumption can break down. Two limitations on allowed motions are made to minimize the errors. First, the backbone is kept rigid so that the motion of individual residues are not coupled together with changes in backbone position. Second, only side chains in the protein interior are allowed to move, which diminishes changes in dielectric boundaries. In particular for large, intrinsic membrane electron transfer proteins this latter restriction is not severe. Here only a small portion of the protein will be in contact with water. In addition, the presence of the water with its high dielectric constant means that surface residues have limited impact on buried cofactors or other charges so they are relatively unimportant.

6. The balance of free energy terms for charges in proteins: how the protein pays the desolvation penalty

The large loss in reaction field energy appears to make burying charges in protein or membrane very unlikely. However, a protein is not a simple organic solvent, but rather is a very complex environment for charges. Fig. 3 contrasts the basic mechanisms of charge stabilization in polar and polarizable media. A polar medium is one that contains dipoles. In a polarizable medium dipoles can be induced and preexisting dipoles can move in response to a change in charge. A chelator is an example of a polar but nonpolarizable medium. Water is both polar and polarizable because of the freedom of each large water dipole to rearrange. Proteins are quite polar without being very polarizable. Each residue has an amide group with a dipole moment that is larger than that of water and more than half of the side chains in an average protein are polar or ionizable. However, motions are greatly restricted in proteins. Lastly, the energy of a new charge can be reduced by counter ion binding or by a compensating change in ionization of a nearby group. Any charge or di-

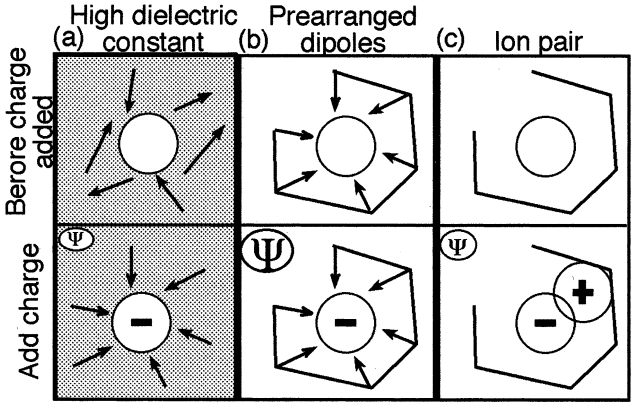

Fig. 3. Stabilizing a buried charge. a: A polarizable solvent like water with flexible dipoles. This is important for residues on the protein surface. The dipoles are in different orientations with and without the charge. Both negative and positive charges are stabilized. The potential from the charge is small at long range (the dielectric response screens the charge). Within proteins reorientation of dipoles or charges stabilize and screen buried charges. b: Pre-oriented dipoles, a polar but not polarizable medium. Buried charges are often stabilized by properly arranged protein backbone dipoles [24,90,142–144]. The dipole positions change little when the charge is added. Interactions with negative and positive charges are of opposite sign. The potential from the charge is large at long range (ε is small, there is little screening). c: Pairing with a second charge of the opposite sign. Many buried charges are found in ion pairs [145,146]. Ion pairs modify protein stability by a modest amount, only sometimes favorably [147–150]. The ion pair forms a dipole, so the net potential from the two charges is small at long range. Coupling binding of opposite charges reduces the net change in charge.

pole is influenced by the surrounding water (Fig. 3a), rigid (3b) and flexible (3a) protein dipoles and charges as well as by coupling to ionization changes at other sites (3c).

Fig. 4 and the accompanying Table 1 describe how each mechanism of stabilization influences the free energy of the available microstates and thus the resulting reaction ΔG. The reduction of an electron acceptor ($A \rightarrow A^-$) is used in this example, but the same mechanisms are found for stabilization of any electron or proton transfer reactions. A new charge generated in a reaction can be stabilized by interaction with fixed charges (a) or dipoles (b). The other mechanisms involve coupling different motions to the reaction. These can be small scale polarization of the protein and surroundings that are averaged in the dielectric response (c); rotation of specific dipoles or preexisting charged groups (d); change in protonation of nearby groups (e); and coupled electron and proton transfer to the cofactor itself.

Any charge or dipole will contribute more stabilization energy if it is in the same position in the reactant and product states. Thus, when (in Fig. 4b) the nearby base is protonated in the initial (A) state this charge stabilizes A^- by the free energy of the pairwise interactions between the two charges. However, if B only binds a proton when A is reduced, then the overall reaction is between BA and BH^+A^-. While the four microstate energies have a number of different terms two are most important: the proton binding by B if A were neutral and the interaction between BH^+ and A^-. The latter, favorable term is the same as is found when B is protonated in the reactant state. However, the proton binding term must be unfavorable or else the proton would be initially bound. Thus, for the same electron acceptor (A) and base (B) the reaction in Fig. 4b will always be more favorable than that in Fig. 4e by the energy of binding the proton. Likewise the reaction free energy of Fig. 4a will be more favorable than Fig. 4d

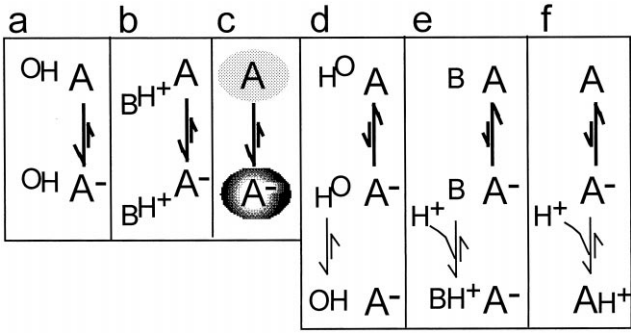

Fig. 4. Five mechanisms for stabilization of a buried charge in the protein. The reduction of an electron acceptor to increase the negative charge. An analogous figure could be drawn for an oxidation that increases the positive charge, or for the gain or loss of a proton. See Table 1 for a more complete description. The first two mechanisms rely on the polarity of the protein: (a) Dipoles in the protein are prearranged to stabilize the product; (b) a nearby charge, present in the initial state, stabilizes the product charge; the latter three columns add the polarizable medium: (c) in any medium except vacuum a new charge induces relatively uniform polarization of its surroundings to stabilize the charge; (d) dipoles in the protein rearrange to stabilize the new charge; (e) a pK_a shift of a nearby base causes proton uptake; (f) the pK_a of the cofactor is lowered when it is reduced in the protein so that a proton is bound to the cofactor itself; here the chemistry of the redox reaction is changed. The last two mechanisms couple two changes in ionization together reducing the net change in charge.

Table 1

	a	b	c	d	e	f
Polar	yes	yes	no	yes	yes	yes
Polarizable	no	no	yes	yes	yes	yes
Net change in charge	yes	yes	yes	yes	no	no
Effect on opposite charge	destabilize by the same amount	destabilize by the same amount	stabilize by the same amount	could stabilize	B could stabilize without H^+ uptake	Depends on A oxidation chemistry
Effect of changes in protein conformation on energy	no change	no change	destabilize	destabilize by ΔG_{oh-rot}	destabilize by $\gamma(pK'-pH)$	destabilize
Stabilization of a charge in a different location	site specific	site specific	the same	site specific	site specific	site specific
final change in equib. $A \rightarrow A^-$	ΔG_{crg}	ΔG_{dip}	ΔG_{rxn}	$\Delta G_{dip}+\Delta G_{oh-rot}$	$\Delta G_{crg}+\gamma(pK'-pH)$	
Microstates to consider	(1) A (2) A^-	(1) A (2) A^-	(1) A (2) A^-	(1) AOH (2) AOH* (3) A^-OH (4) A^-OH*	(1) AB (2) ABH^+ (3) A^-B (4) A^-BH^+	(1) A (2) AH^+ (3) A^- (3) A^-H^+

Polar: requires charges or dipoles in surroundings; polarizable: the distribution of charges or dipoles change; net change in charge: when electron and/or proton transfer reactions are coupled together the net change in charge is reduced. Effect on opposite charge: would a positive charge at the location of A^- be stabilized or destabilized. Effect on protein energy: the energy required to change conformation (c, d) or charge (e) excluding interactions with A^-. Different mechanisms of charge stabilization (see Fig. 4) (a) There is a fixed charge near A. (b) There are fixed dipoles near A. (c) The changes in the protein in response to the charge are treated as an average dielectric response. (d) A dipole changes its orientation. ΔG_{oh-rot} is the difference in energy of the OH in its two orientations in the absence of the charge on A^-. The two orientations are labeled OH and OH* in the enumerated states. By definition this energy is unfavorable or else the OH would be in the other orientation in the ground state. ΔG_{dip} is the energy of interaction between the final OH orientation and A^-. (e) A nearby group changes its protonation state when A is reduced. pK' is the pH that B would be 50% ionized given all interactions in the protein other than those with A^-. $\gamma(pK'-pH)$ must be a positive, unfavorable energy term with the pH above pK' since the site is deprotonated in the reactant state. $\gamma(pK'-pH)$ is the penalty for proton binding. The interaction between BH^+ and A^- (ΔG_{crg}) is sufficiently favorable to overcome the penalty so B binds a proton. (f) The electron acceptor itself binds a proton. Thus, the pK of A in the protein is lower than the pH, while the pK of A^- in the protein is higher than the solution pH.

by the energy penalty to move the dipole to its orientation that has a favorable interaction with A^- but is disfavored by the rest of the protein.

6.1. The relative importance of pairwise and self energy terms

There is some controversy about the role of nonspecific dielectric relaxation and site specific pairwise interaction in the stabilization of charges in protein. These energy terms are contrasted in Figs. 3 and 4. However, Fig. 2 shows that the apparent importance of each term will differ if microstate energies are calculated in different ways. If the dipole is assembled (clockwise) the pairwise interactions are large while the reaction field energy gained transferring the dipole into the higher dielectric solvent is relatively small. In contrast, counterclockwise assembly yields large reaction field energies, but now the pairwise interactions are calculated under conditions where the solvent screens the reaction so these are relatively small. A different balance is seen if like charges are assembled. Here both the favorable reaction field energy (which varies with q^2) and unfavorable pairwise interactions are larger along the clockwise than the counterclockwise path. However, in all cases the free energy of the resultant state is the same. The MCCE method described here assembles microstate energies using the counterclockwise assembly path.

6.2. The distribution of buried charges in proteins

The loss in reaction field energy when charges are

moved out of water is very large. Several mechanisms for stabilizing charges in proteins have been described. There are few redox cofactors in proteins, but the acidic and basic residues Asp, Glu, Arg, and Lys are plentiful in proteins. These provide a large group of sites that can show different ways that proteins can interact with charges. One survey of proteins with different folding motifs and sizes examined the loss in reaction field energy of the Asp, Glu, Arg, and Lys in 300 proteins [90]. Seventy percent of these ionizable residues have lost less than 4.1 kcal/mole of the reaction field energy they would have if free in water, shifting the residue pK_as by less than 3 pH units (Eq. 6). Thus as expected, most of these residues are near the surface. However, 30% of the ionizable residues have $\Delta G_{rxn} > 4.1$ kcal/mol. Half of these residues (15%) have lost sufficient reaction field energy to shift their pK_as by 5 pH units (6.8 kcal/mol). A 5 pH unit shift destabilizes an ionized Asp moving its pK_a from 4 to 9. The same ΔG_{rxn} shifts the pK_a of an Arg from 12.5 to 7.5. Different propensities are found for burying each type of side chain. There are more buried Asp, similar numbers of buried Arg and Glu, and fewer buried Lys. Overall there are more buried acids than bases. This disparity becomes more significant as ΔG_{rxn} increases. For residues where ΔG_{rxn} is 4.1 to 6.8 kcal/mol 56% are acids. Of the residues where ΔG_{rxn} is >6.8 kcal/mol 62% are acids representing 17% of the acids and 12% of the bases.

The protein backbone dipoles are the most prevalent polar group in any protein. The dipoles are often arranged to stabilize buried charges in proteins. Fig. 5 compares the loss in reaction field energy (ΔG_{rxn}) with the pairwise interaction of ionized acids and bases with the backbone ($\Delta G_{crg:dip}$) in *Rps. viridis* RCs. Many residues are found to have lost significant reaction field energy at their position in the protein. However, all calculations of in situ pK_as in RCs suggest that almost all of these residues are ionized at physiological pH [17,21,24,25]. Because of the shape of the amide dipole, with a large carbonyl oxygen at the negative and small proton at the positive end of the dipole the potential from backbone dipoles is, on average, positive within all proteins (see [90] for a more complete description). As a result, there are more acids than bases with negative, favorable interaction energies with the backbone. In addition, a large positive potential from the backbone is occasionally found to significantly destabilize the ionization of bases as seen by their positive $\Delta G_{crg:dip}$ in Fig. 5. However, bases can be stabilized by the backbone when they make specific, favorable hydrogen bonds to backbone amides or are at the negative, C-terminus of an α-helix [91]. Thus, the dipoles of the protein backbone help stabilize many charges in the protein interior. The backbone is only one contributor to the energy of site ionization. For example, the few bases that are found to be destabilized by the backbone dipoles are found near buried acids in regions of positive backbone potential. For these residues it is the ionized acids that help to stabilize the buried bases.

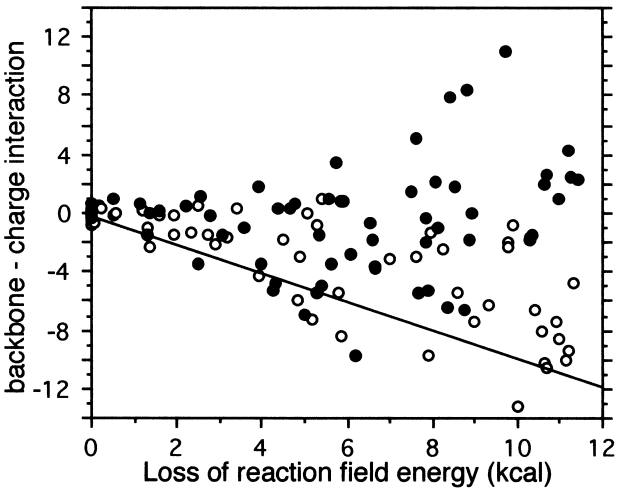

Fig. 5. Comparison of the loss of reaction field energy (ΔG_{rxn}) and interactions with the backbone dipoles for the ionizable residues in *Rps. viridis* RCs. One residue is charged at a time providing its ΔG_{rxn}. Interactions with the backbone are obtained from a sum of (the potential at each backbone atom)× (the partial charge on that atom). A residue along the line of slope −1 has a favorable interaction with the backbone that exactly matches the destabilization of the charge by the ΔG_{rxn}. (○, acids; ●, bases).

7. Analysis of the free energy of several electron transfer reactions in photosynthetic reaction centers

Reaction centers of photosynthetic bacteria were the first intrinsic membrane proteins with a known three-dimensional structure [1,92] (Fig. 6). There are several different reactions that highlight various as-

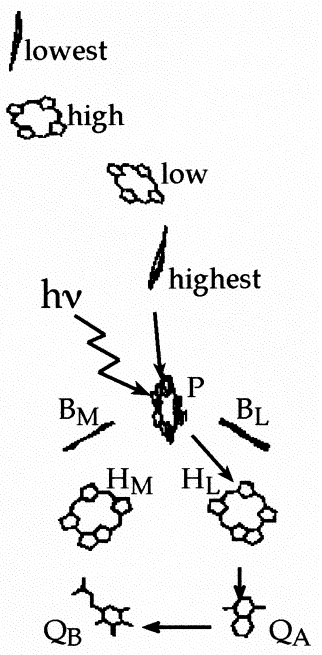

Fig. 6. The position of the redox cofactors in reaction centers of the purple photosynthetic bacteria *Rps. viridis*. The four hemes are labeled by their relative electrochemical redox potential. The subunit binding the four hemes extends into the periplasmic space. The heme subunit is missing from the *Rb. sphaeroides* RCs. The rest of the cofactors are within the membrane in RCs from both bacteria. The arrows show the direction of electron transfer. P is a bacteriochlorophyll dimer, B_L and B_M are bacteriochlorophyll monomers, H_L and H_M bacteriopheophytins, and Q_A and Q_B are quinones.

pects of the movement of charges within an intrinsic membrane protein. Two related issues that have been studied in this protein: The first is how does the protein control in situ cofactor electrochemistry. RCs use the same cofactors in several different locations. There are four hemes in the cytochrome subunit in *Rps. viridis* RCs with E_ms that span 450 mV [93,94]. In addition, the cofactors in the transmembrane portion have a c_2 symmetry in structure but not in function [95]. Differences in the in situ electrochemistry of these groups demonstrates the power of the protein to affect functional cofactor behavior. In addition, the ΔG of two different reactions has been considered in some detail. The first is the charge separation reaction where the excited singlet state of the chlorophyll dimer forms an ion pair (P^+H^-) in 3 ps. The second is the electron transfer between the two quinones occurring on the microsecond time scale. The protein appears to control these reactions in very different ways.

7.1. Cytochromes in the heme subunit

Rps. viridis RCs have four hemes in a bound cytochrome subunit. The hemes are arranged in a line approximately perpendicular to the membrane (Fig. 6). A number of experiments established that the redox potentials vary with the pattern: highest (nearest P), low, high, lowest (furthest into the periplasm) [96–99]. Since the site closest to P has the most positive E_m, it is the last oxidized. Thus, it is kept in the redox state needed to carry out the reduction of P^+. A conventional. static electrostatic analysis obtained E_ms that were within ≈ 60 meV of the experimental values for each of the four hemes. Thus, the information contained in the crystal structure is sufficient to understand how hemes ≈ 14 Å apart (center to center) can have midpoints that differ by more than 400 mV.

Several different electrostatic interactions are responsible for the differences in the heme in situ electrochemistry [23]. There is a significant loss in reaction field energy that makes all hemes harder to oxidize. However, there is not much difference between the reaction field energy of the individual sites. One of the low potential hemes has two His rather than a His and Met as ligands. In solution, bis-His hemes have an E_m 150 meV lower than found for His-Met hemes [41–43]. This difference in reference energy appears to be held in the protein. However, it is one of the three His-Met hemes that has the lowest potential. Ionization of this site, furthest from P, is significantly stabilized by surrounding acids including contributions from the propionic acids associated with each heme. The highest potential site, nearest P, has an Arg nearby which is calculated to raise the potential by >300 mV. Thus, the highest and lowest potentials are produced by pairwise interactions with positively and negatively charged groups. The heme second from the end is another high potential heme. Its two propionic acids are forced into hydrogen bonding distance so one of them is neutral. The removal of this negative charge raises the potential of this heme relative to the others which have both of their propionic acids ionized. Oxidation of this high potential heme is also destabilized by its location

between the two low potential hemes which will become positively charged at lower E_hs. Thus, a relatively simple analysis of the detailed protein structure that considers the distribution of charges, solvent, and ligands yields a relatively accurate estimate of the RC heme E_ms.

7.2. The initial, fast electron transfer reaction

Trapping a photon's energy in photosynthesis requires fast initial electron transfer since the excited singlet reactant states live for nanoseconds or less. The ground state of the protein must therefore have all reactants bound in appropriate conformations in preparation for the reaction. In contrast, in conventional biochemical reactions, reactant states often live for milliseconds providing ample time for conformation changes in the protein or diffusion of reaction partners. The rapid reactions in photosynthetic proteins often have rates as fast at cryogenic temperatures as they do at room temperature. This shows experimentally that few changes in protein, cofactor, or solvent atomic positions can be required for reaction as these would be frozen out as the temperature is lowered.

The first reaction involves moving an electron from a bacteriochlorophyll dimer (P) to a bacteriopheophytin (H) to create the ion pair P^+H^-. The initial charge separation reaction in photosynthesis has been the subject of measurement, mutation, and calculation. The techniques of delayed fluorescence allow charge separated states to remain in equilibrium with the excited singlet P* (Fig. 6.) Values for the ΔG for formation of P^+H^- have been measured to be as little as 90 to as much as 250 meV [100,101]. The midpoints of isolated BChl and BPh show the latter to be the better electron acceptor by $\approx$250 meV [50,102]. This does not appear to be very dependent on solvent or aggregation state, suggesting that the difference between an isolated dimer and BPh could be $\approx$250 meV. Thus, H will be reduced in preference to P. However, in the charge separation process H must be reduced in preference to P^+. In vacuum the ion pair is destabilized by the cost of moving the electron away from P^+ to H by more than 1 eV [52]. The question is how a state that is very unfavorable in vacuum is stabilized in the protein. Various analyses have come to different conclusions about the role of the protein [54–57,103]. These contrast the roles of (1) the dielectric response of the medium; (2) the role of local dipoles [103]; and (3) the role of fields created by more distant buried charged amino acids. Each method can stabilize P^+H^-. Electrons will instantly polarize around the ion pair, stabilizing it by $\approx$600–700 meV even at cryogenic temperature as fast as the electron transfer occurs [57]. As more nuclear, dipolar motions add to the relaxation process the ion pair becomes more stable. Estimates are that an ε of only 5 may be sufficient to make P^+H^- lower in energy than P* without any assistance from specific elements of the protein [104,105]. With an average dielectric constant of 5 to 6 the role of the protein in this reaction would be simply to provide a scaffolding for the cofactors and a close packed media that can polarize a bit to stabilize the rapid formation of P^+H^-.

Pairwise interactions with nearby dipoles are found by all analyses to play a role in stabilization of the ion pair [57,103]. However, the importance of charged residues is less clear. An analysis of the electrostatic potential from acidic and basic residues provides a different view of the role of the protein in the early electron transfer [57]. Charged residues on the protein surface will be solvated by water so they have a negligible impact on the potential at P or H. There are no charged residues within a 15 Å region through the hydrophobic center of the membrane. However, there are many charges that are buried in the region near the top and bottom of the membrane. Analysis of the field from these charges suggests that they contribute to a gradient of the potential arranged appropriately to stabilize P^+H^-, thus more negative near P and positive near H. In *Rps. viridis* this field is due to relatively small individual contributions from more than 20 residues. Electrostatic calculations of *Rps. viridis* RCs suggest that these charges may yield a potential of as much as 30 mV/Å (if $\varepsilon = 2$) which would stabilize the ion pair by almost 800 meV [57].

Thus, both the reaction field energy and a static field can and certainly do stabilize the ion pair. The question is what the balance is. The use of a large dielectric response would be unlikely because the reaction occurs slightly faster at 1 K than it does at room temperature [106]. While electronic polarization will be unaffected by the temperature, the changes in nuclear position required for additional

relaxation should be diminished [107]. In addition, Marcus electron transfer theory describes how the nuclear motions effect the electron transfer rate. A dielectric model is used for these motions [108,109]. As more motion is coupled to the reaction the reorganization energy (λ) increases. The reaction rate is maximal when $-\Delta G = \lambda$. As dielectric relaxation increases λ increases and the rate slows at small driving force. This is clearly found in electron transfer reactions carried out in solvents with different dielectric constants [110,111]. The fast, first electron transfer reaction has a relatively small driving force and a rate that appears to be close to maximal given the electron transfer distance [112]. Thus, if λ is small the role of dielectric relaxation is also small suggesting that there is a small effective dielectric response to the creation of the stable ion pair. A static field provides another means to stabilize the ion pair without increasing the reaction reorganization energy. However, clearly both static field and dielectric relaxation will play important roles in this reaction.

7.3. The electron transfer from Q_{A^-} to Q_B

In RCs Q_A is rapidly reduced by H^- to the semiquinone which in turn reduces Q_B (Fig. 6). Following rereduction of P^+ from bound or soluble cytochrome, a second series of electron transfer reactions results in the second reduction of Q_B [47,113,114]. The doubly reduced Q_B binds two protons, with one proton transfer preceding the second reduction [114]. The protons must be moved in from the cytoplasm through the protein [115–117]. The dihyroquinone is then released into the membrane and replaced with an oxidized quinone. Q_A and Q_B are both ubiquinones in *Rb. sphaeroides* RCs or menaquinone (Q_A) and ubiquinone (Q_B) in *Rps. viridis* RCs [118]. The first electron transfer reaction takes place in microseconds in a process which requires protein conformation change [119–121] and little proton uptake at physiological pH [122,123]. The reaction ΔG and the stoichiometry of proton uptake coupled to the electron transfer is pH dependent. The differences in behavior of Q_A and Q_B highlight the importance of the protein in controlling cofactor electrochemistry.

There have been several continuum electrostatics simulations of the electron transfer from Q_{A^-} to Q_B in RCs of *Rb. sphaeroides* [14,15,17] and *Rps. viridis* RCs [24,25]. All found that at physiological pHs all basic groups (except His) are fully ionized and do not undergo ionization changes upon quinone reduction [17,21,24,25,57]. Most Asp and Glu are ionized. As many of these residues are deeply buried in the protein (Fig. 5), all of the electrostatic calculations on RCs show that these proteins are designed to stabilize buried charges. One reason that many acidic residues are ionized is that the backbone provides large regions of positive electrostatic potential (Fig. 5) [14,15,24]. Remarkably, the potential from the neutral backbone dipoles can reach +500 mV or more. In the Q_B site several residues with large (>10 pH unit) losses of reaction field energy, have even larger, favorable interactions with the backbone.

The protein near Q_A and Q_B shows significant differences in the density and connectivity of ionizable residues. There are fewer near Q_A than Q_B resulting in a less flexible local electrostatic environment for Q_A reduction [24]. In both *Rb. sphaeroides* and *Rps. viridis* RCs Q_B is embedded in a large group of strongly interacting acids and bases. In *Rp. viridis* 24 residues can be connected to Q_B by a chain of interactions of more than $2\Delta pK$ units (2.7 kcal/mol). This web extends over 25 Å [24]. Similar connectivity can also be found in *Rb. sphaeroides* RCs [21]. The large group of interconnected residues make the electrostatic environment near the quinones complex. Also, the protein may have some characteristics of a buffered sponge rather than providing a single path for proton transfer [115–117].

In the Q_B site, residues GluL212 and AspL213 and L210 interact strongly with each other and with Q_{B^-} in *Rb. sphaeroides* RCs [17,21]. These residues have also been recognized as important by site directed mutagenesis studies [124–129]. In *Rps. viridis* L213 is an Asn not an Asp so the groups that make up the cluster must be different. The Glu's L212, H177 (H173 in *Rb. sphaeroides*), and M234 are now identified as the active cluster [24,25]. In the ground state at physiological pH each calculation finds that the cluster is partially protonated. The net charge between -1 and -2 has been estimated in different calculations for the three acids [17]. This provides one to two protons in the cluster that can change position when Q_B is reduced, diminishing the re-

quirement for proton uptake from solution. Most calculations show that many peripheral groups each contribute a little bit to proton uptake following electron transfer at different pHs.

7.3.1. The free energy of the electron transfer reaction

The distribution of atomic positions and ionization states were calculated for the ground, Q_A^- and Q_B^- states using the MCCE method [17]. A cluster of residues (GluL212, AspL213, AspL210, and SerL223) were found to be active participants in the electron transfer from Q_A^- to Q_B (Fig. 7) [17,21]. In the ground and Q_A^- states L213 is ionized forcing L212 and L210 to be protonated. SerL223 is hydrogen bonded to L213 stabilizing the charge on this site. In the cluster, L210 is furthest from Q_B. When Q_B is reduced, a proton is transferred from AspL210 to AspL213. This breaks the hydrogen bond between AspL213$^-$ and SerL223 reorienting the Ser hydroxyl to make a hydrogen bond to Q_B^-. Thus, at physiological pH, the charge on Q_B is stabilized by an internal proton transfer rather than by proton uptake from solution. The structure equilibrated around Q_B^- is poised for the protonation of Q_B^-. The hydroxyl of SerL223 is properly situated to play an active role in transfer from the protonated AspL213. This proton transfer to Q_B^- is energetically unfavorable, but it appears to precede the second reduction of Q_B [114]. Final reduction of Q_B to dehydroquinone requires proton binding from solution [130]. The roles of the residues in the cluster are in agreement with the picture obtained by site directed mutagenesis experiments. These suggested that AspL213 and SerL223 are involved in the first proton transfer to Q_B [124,131–133] which occurs following a second electron transfer to generate $Q_A^-Q_B^-$.

All previous conventional, static calculations have been unable to calculate the measured values for $-\Delta G_{AB}$ and its pH dependence. This is may be because of the presumption of a rigid structure, or because of the use of inappropriate parameters. Rigid protein calculations of Rabenstein et al. [25] do reproduce the measured $-\Delta G_{AB}$ at pH 7. As they note there is significant proton uptake calculated to be coupled to the reaction so the reaction would show significantly more pH dependence than is found experimentally. However, these calculations

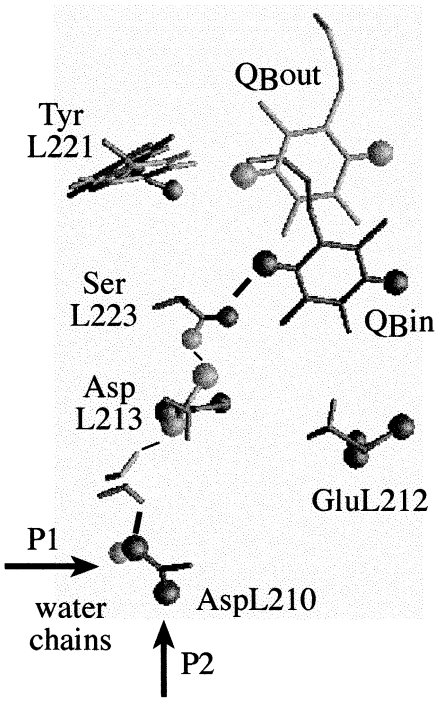

Fig. 7. Residues and a water found to undergo conformation or ionization changes upon the electron transfer from Q_A^- to Q_B in MCCE calculations [17]. The lighter residue is the most probable conformer in the ground state, the darker is found in the presence of Q_B^-. in the ground state the ionized AspL213 makes a hydrogen bond to SerL223 and to a water (shown by light connecting lines). AspL210 is predominantly neutral. In the Q_B^- state, the Ser makes a hydrogen bond to Q_B^- and the water makes a hydrogen bond to the now ionized AspL210 (shown by the dark connecting lines). AspL213 is now neutral. GluL212 is neutral in both ground and Q_B^- states. These changes in protonation and polar proton position provide a possible proton pathway to Q_B^-.

do make use of a more realistic charge set for the quinones with smaller changes in charge on individual atoms, perhaps modifying the pairwise interactions with sites close to the quinone.

Several calculations in a rigid protein obtain a ΔG_{AB} of +170 meV at pH 7 [17,21], while MCCE simulations yield a ΔG_{AB} of −80 meV [17], within 20 meV of the experimental value. The MCCE calculations match the experimental values within 30 meV from pH 5 to 11. Thus conformational flexibility that can be captured in the MCCE method appears to be coupled to the ionization changes to create a favorable reaction free energy. Changes occur in conformation and ionization of the cluster described above as well as at other sites throughout

the protein on reduction of Q_B. Thus, work goes into the protein to create the favorable environment for Q_B^-. The energy of interaction of Q_B^- with the protein during Monte Carlo sampling of microstates with Q_B^- is 260 meV more favorable than the interaction of Q_A^- with the protein when the simulation contains only Q_A^-. However ΔG_{AB} is only −80 meV. Thus, if the interaction with the quinone is removed, the distribution of protein conformations formed with Q_B^- is 180 meV less favorable than with Q_A^-. The less favorable intra-protein interactions, but more favorable protein:quinone interactions in the Q_B^- state yields the final observed free energy. This balance of interaction energies is found whenever conformation changes stabilize a charge (Table 1). Such a response can be described as dielectric relaxation around the new charge. However, in the MCCE method the atomic details of many of these changes can be seen.

7.3.2. *The influence of buried waters on the reaction*

The MCCE calculations were designed to couple the conformational flexibility of waters to electron and proton transfer reactions. Differences in equilibrium water position and site occupancy can be seen when different cofactors are charged [17]. Connected groups of water and polar residues which change conformation may be important in proton uptake. Water channels have been identified in the protein structure near Q_B and Q_A that could transfer protons inward towards the quinones [117,134–137]. For example, mutations at L209 near a water channel perturb proton transfer rates [138]. In the MCCE calculations waters in some but not all channels show changes when the electron is transfer from Q_A^- to Q_B [17].

7.4. *The analysis of mutations in the Q_B site by calculation*

The MCCE method has been used to study RCs with site directed mutations, comparing calculated and experimental free energy of the Q_A^- to Q_B electron transfers [127,132,134,139,140]. Numerical calculations were carried out on the single mutants: L212EQ, L212EA, L213DA, M44ND; double mutant: L212EA/L213DA and suppressors: L212EA/L213DA/M233KL, L212EA/L213DA/M44ND. The

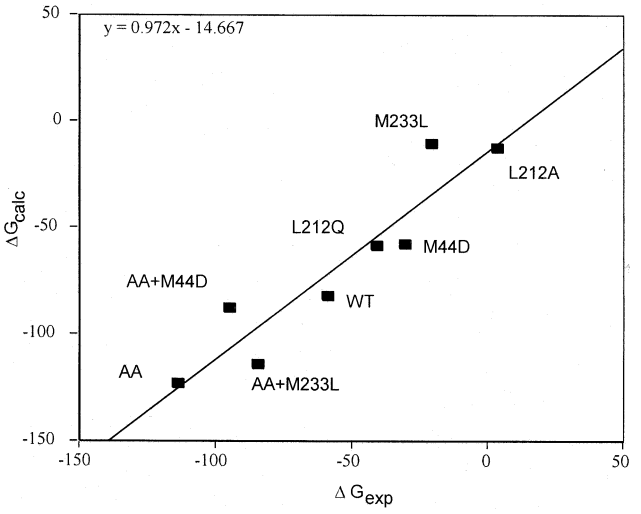

Fig. 8. Comparison of measured and calculated ΔG_{AB} at pH 7 for electron transfer from Q_A^- to Q_B in a series of mutated RCs [140]. The protein structure from *Rb. sphaeroides* [136] was used for the MCCE calculations. Measurements were made in *Rb. capsulatus* RCs [127,134]. The mutated amino acid is given. In the wild type protein, L212 is a Glu, L213 an Asp, M44 an Asn, and M233 an Arg. The AA mutant changes both L212 and L213 to Ala. Thus, each mutant either introduces or removes a ionizable residue.

calculated ΔG_{AB}s agree remarkably well with the experimental values (Fig. 8). Each of these mutations involves changes in ionizable residues in or near the Q_B site. The calculations suggest that RCs responds to these mutation by changes that maintain the same net charge as found in the wild type (WT) protein. In each case the response of the protein reduces the effect of the mutation. For example: (1) Substitution of L212Glu by either Gln or Ala does not change the net charge, because L212Glu is calculated to be neutral in WT RCs. (2) Mutation of M233Arg to Leu is coupled to the nearby H230Glu, ionized in WT RCs, binding a proton in the mutant. (3) The affect of the L213Asp to Ala mutation is reduced by an increase in the ionization of L210Asp in ground and Q_A^- states. L213Asp is neutral in the Q_B^- state in WT RCs so the mutation has little impact on the product state. (4) The M44Asn to Asp mutation has little influence on the net charge as the introduced M44Asp is mostly protonated at physiological pH.

Initial analysis of the mutants generally made the first order assumption that changing residues in the Q_A site would affect the reactant (Q_A–Q_B) state and changes in the Q_B site affect the product ($Q_AQ_B^-$)

state. However, calculations can provide additional clues to the underlying cause of experimental results. One case in point is the mutation of AspL213 to an Asn. As described above this residue plays an important role in the reaction. When the ionizable residue is removed the reaction becomes more favorable, and this is reproduced in the calculations. The experimental result has been interpreted as the mutant removing a negative charge from the vicinity of Q_{B^-} [21,124,133]. However, the calculations show that in the Q_{B^-} state AspL213 is neutral, so replacing it with a neutral amino acid only removes the dipole of the neutral AspH not the charge of Asp-. Thus, despite the mutation being made quite close to Q_B this mutation does not have the largest influence on the product state. Rather, the mutation breaks the hydrogen bond between $AspL213^-$ and SerL233 (Fig. 7), destabilizing the reactant state, increasing the reaction $-\Delta G_{AB}$. In the product state, the Ser makes a hydrogen bond to the quinone so the loss of the Asp produces little change.

Residue mutations can also affect the connectivity within water channels. Following the nomenclature introduced by Abresch et al. [137], it was found that mutations at M44 and L213 sites mostly perturb water channels P2 and P3, while mutations at M233 and L212 sites affect mostly the P1 water channel.

8. Conclusion

Much of this paper has described the thermodynamic consequences of moving charges into different locations in a protein. Although charges in proteins are destabilized by the loss of reaction field (solvation) energy, these can be stabilized by protein charges, dipoles, and conformation changes. Methods are evolving that can start with an atomic structure of a protein and can calculate properties that can be compared with experiment. Calculations of this type begin to use the enormous amount of information contained in atomic protein structures. Each successful calculation shows how different elements of each protein combine to generate the observed, functional behavior.

The analysis of the E_ms of the hemes, the initial charge separation reaction, and the electron transfer from Q_{A^-} to Q_B shows how charges are buried at different sites in the protein. Oxidation of the hemes, formation of the P^+H^- ion pair, and charge shift from Q_{A^-} to Q_B all involve changes in charge at sites that are deeply buried. In these reactions there is no requirement for formal charge compensation to stabilize the buried charge. Rather the electrostatic potential generated by the preexisting distribution of protein charges and dipoles as well as the dielectric response of the protein are important for all reactions at all time scales. The slower Q_A to Q_B electron transfer is coupled not to proton uptake from solution, but to an internal proton transfer reaction and to changes in hydrogen bonding patterns in the Q_B site. Thus, different motifs for charge stabilization can be identified from the analysis of this protein.

Acknowledgements

We would like to thank Dmitri Karpman, Qiang Xu, and Ted Leslie for helpful comments on the manuscript, Peter Brzezinski for his patience, and the financial support of the N.S.F. PECASE award MCB-9629047 and NATO Grant LST.CLG 975754.

References

[1] C.R.D. Lancaster, U. Ermler, H. Michel, in: R.E. Blankenship, M.T. Madigan, C.E. Bauer (Eds.), Anoxygenic Photosynthetic Bacteria, Kluwer, Dordrecht, 1995, pp. 503–526.
[2] G. McDermott, S.M. Prince, A.A. Freer, A.M. Hawthornthwaite-Lawless, M.Z. Papiz, R.J. Cogdell, N.W. Isaacs, Nature 374 (1995) 517–521.
[3] C. Ostermeier, S. Iwata, H. Michel, Curr. Opin. Struct. Biol. 6 (1996) 460–466.
[4] D. Xia, C. Yu, H. Kim, J. Xia, A.M. Kachurin, L. Zhang, L. Yu, J. Deisenhofer, Science 277 (1997) 60–66.
[5] S. Yoshikawa, Curr. Opin. Struct. Biol. 7 (1997) 574–579.
[6] D.A. Doyle, J.M. Cabral, R.A. Pfuetzner, A. Kuo, J.M. Gulbis, S.L. Cohen, B.T. Chait, R. MacKinnon, Science 280 (1998) 69–77.
[7] Z. Zhang, L. Huang, V.M. Shulmeister, Y.I. Chi, K.K. Kim, L.W. Hung, A.R. Crofts, E.A. Berry, S.H. Kim, Nature 392 (1998) 677–684.
[8] A. Parsegian, Nature 221 (1969) 844–846.
[9] R.J. Kassner, Proc. Natl. Acad. Sci. USA 69 (1972) 2263–2267.
[10] R.J. Kassner, J. Am. Chem. Soc. 95 (1973) 2674–2676.

[11] R.J. Kassner, W. Yang, J. Am. Chem. Soc. 99 (1977) 4351–4355.
[12] A. Warshel, S.T. Russell, Q. Rev. Biophys. 17 (1984) 283–422.
[13] A.K. Churg, A. Warshel, Biochemistry 25 (1986) 1675–1681.
[14] P. Beroza, D.R. Fredkin, M.Y. Okamura, G. Feher, Proc. Natl. Acad. Sci. USA 88 (1991) 5804–5808.
[15] M.R. Gunner, B. Honig, in: J. Breton, A. Vermeglio (Eds.), The Photosynthetic Bacterial Reaction Center: Structure, Spectroscopy and Dynamics II, Plenum, New York, 1992, pp. 403–410.
[16] E.G. Alexov, M.R. Gunner, Biophys. J. 72 (1997) 2075–2093.
[17] E. Alexov, M. Gunner, Biochemistry 38 (1999) 8253–8270.
[18] A.-S. Yang, M.R. Gunner, R. Sampogna, K. Sharp, B. Honig, Proteins 15 (1993) 252–265.
[19] V.Z. Spassov, D. Bashford, J. Comp. Chem. 0 (1999) 1–21.
[20] J.A. Schellman, Biopolymers 14 (1975) 999–1018.
[21] P. Beroza, D.R. Fredkin, M.Y. Okamura, R. Feher, Biophys. J. 68 (1995) 2233–2250.
[22] D. Bashford, M. Karplus, J. Phys. Chem. 95 (1991) 9556–9561.
[23] M.R. Gunner, B. Honig, Proc. Natl. Acad. Sci. USA 88 (1991) 9151–9155.
[24] C.R.D. Lancaster, H. Michel, B. Honig, M.R. Gunner, Biophys. J. 70 (1996) 2469–2492.
[25] B. Rabenstein, G.M. Ullmann, E.-W. Knapp, Eur. Biophys. J. 27 (1998) 626–637.
[26] B. Honig, A. Nicholls, Science 268 (1995) 1144–1149.
[27] H. Nakamura, Q. Rev. Biophys. 29 (1996) 1–90.
[28] A. Warshel, A. Papazyan, Curr. Opin. Struct. Biol. 8 (1998) 211–217.
[29] Y.Y. Sham, Z.T. Chu, A. Warshel, J. Phys. Chem. 101 (1997) 4458–4472.
[30] T.J. You, D. Bashford, Biophys. J. 69 (1995) 1721–1733.
[31] D.R. Ripoll, Y.N. Vorobjev, A. Liwo, J.A. Vila, H.A. Scheraga, J. Mol. Biol. 264 (1996) 770–783.
[32] H. Zhou, M. Vijayakumar, J. Mol. Biol. 267 (1997) 1002–1011.
[33] P. Beroza, D. Case, J. Phys. Chem. 100 (1996) 20156–20163.
[34] D.J. Abraham, A.J. Leo, Proteins Struct. Funct. Genet. 2 (1987) 130–152.
[35] A. Radzicka, R. Wolfenden, Biochemistry 27 (1988) 1664–1670.
[36] K. Warncke, P.L. Dutton, Proc. Natl. Acad. Sci. USA 90 (1993) 2920–2924.
[37] R. Wolfenden, A. Radzicka, Science 265 (1994) 936–937.
[38] D.J. Tannor, B. Marten, R. Murphy, R.A. Freisner, D. Sitkoff, A. Nichols, M. Ringnalda, W.A. Goddard, B. Honig, J. Am. Chem. Soc. 116 (1994) 11875–11882.
[39] D. Sitkoff, K.A. Sharp, B. Honig, Biophys. Chem. 51 (1994) 397–409.
[40] R. Richarz, K. Wüthrich, Biopolymers 17 (1975) 2133–21341.
[41] H.A. Harbury, P.A. Loach, J. Biol. Chem. 255 (1960) 3640–3645.
[42] H.A. Harbury, P.A. Loach, J. Biol. Chem. 235 (1960) 3646–3653.
[43] H.A. Harbury, J.R. Cronin, M.W. Fanger, T.P. Hettinger, A.J. Murphy, Y.P. Meyer, S.N. Vinogradov, Proc. Natl. Acad. Sci. USA 54 (1965) 1658–1664.
[44] K.M. Kadish, L.A. Bottomley, Inorg. Chem. 19 (1980) 832–836.
[45] K.M. Kadish, Prog. Inorg. Chem. 34 (1986) 435–605.
[46] P.R. Rich, D.S. Bendall, FEBS Lett. 105 (1979) 189–194.
[47] C.A. Wraight, Photochem. Photobiol. 30 (1979) 767–776.
[48] W.A. Cramer, D.B. Knaff, Energy Transduction in Biological Membranes: A Textbook of Bioenergetics, Springer-Verlag, New York, 1991.
[49] R.C. Prince, P. Lloyd-Williams, J.M. Bruce, P.L. Dutton, Methods Enzymol. 125 (1986) 109–119.
[50] T.M. Cotton, R.P. Van Duyne, J. Am. Chem. Soc. 101 (1979) 7605–7615.
[51] H. Scheer, in: H. Scheer (Ed.), Chlorophylls, CRC Press, Boca Raton, Fl, 1991, pp. 4–30.
[52] M.A. Thompson, M.C. Zerner, J. Am. Chem. Soc. 113 (1991) 8210–8215.
[53] M.R.A. Blomberg, P.E.M. Siegbahn, G.T. Babcock, J. Am. Chem. Soc. 120 (1998) 8812–8824.
[54] W.W. Parson, Z.-T. Chu, A. Warshel, Biochim. Biophys. Acta 1017 (1990) 251–272.
[55] M. Marchi, J.N. Gehlen, D. Chandler, M. Newton, J. Am. Chem. Soc. 115 (1993) 4178–4190.
[56] J.N. Gehlen, M. Marchi, D. Chandler, Science 263 (1994) 499–502.
[57] M.R. Gunner, A. Nicholls, B. Honig, J. Phys. Chem. 100 (1996) 4277–4291.
[58] J. Antosiewicz, J.M. Briggs, A.H. Elcock, M.K. Gilson, J.A. McCammon, J. Comp. Chem. 14 (1996) 1633–1644.
[59] M. Karplus, R.N. Porter, Atoms and Molecules, Benjamin/Cummings, Menlo Park, CA, 1970.
[60] A.A. Rashin, B. Honig, J. Phys. Chem. 89 (1985) 5588–5593.
[61] A. Jean-Charles, A. Nicholls, K. Sharp, B. Honig, A. Tempczyk, T.F. Hendrickson, W.C. Still, J. Am. Chem. Soc. 113 (1991) 1454–1455.
[62] D. Sitkoff, K.A. Sharp, B. Honig, J. Phys. Chem. 98 (1994) 1978–1988.
[63] A.T. Brunger, X-Plor, Yale University Press, New Haven, 1987.
[64] S.N. Rao, U.C. Singh, P.A. Bash, P.A. Kollman, Nature 328 (1987) 551–554.
[65] W.L. Jorgensen, J. Tirado-Rives, J. Am. Chem. Soc. 110 (1988) 1657–1666.
[66] C. Chipot, J.G. Angyan, B. Maigret, H.A. Scheraga, J. Phys. Chem. 97 (1993) 9788–9796.
[67] C. Chipot, J.G. Angyan, B. Maigret, H.A. Scheraga, J. Phys. Chem. 97 (1993) 9797–9807.
[68] M.K. Gilson, B.H. Honig, Biopolymers 25 (1986) 2097–2119.
[69] S. Harvey, Proteins 5 (1989) 78–92.
[70] Y.Y. Sham, I. Muegge, A. Warshel, Biophys. J. 74 (1998) 1744–1753.

[71] R. Pethig, Dielectric and electronic properties of biological materials, John Wiley, New York, 1979.
[72] S. Bone, R. Pethid, J. Mol. Biol. 157 (1982) 571–575.
[73] S. Takashima, H. Schwan, J. Phys. Chem. 69 (1965) 4176–4182.
[74] T. Simonson, D. Perahia, Comp. Phys. Comm. 1995 (1995) 291–303.
[75] T. Simonson, D. Perahia, J. Am. Chem. Soc. 117 (1995) 7987–8000.
[76] T. Simonson, C.L. Brooks, J. Am. Chem. Soc. 118 (1996) 8452–8458.
[77] J. Antosiewicz, J.A. McCammon, M.K. Gilson, J. Mol. Biol. 238 (1994) 415–436.
[78] J. Antosiewicz, J.A. McCammon, M.K. Gilson, Biochemistry 35 (1996) 7819.
[79] E. Bitsmuto, G.B. Strambini, G. Irace, Biophys. J. 55 (1989) 18a.
[80] P.D. Baker, A.G. Mauk, J. Am. Chem. Soc. 114 (1992) 3619–3624.
[81] A.M. Berghuis, G.D. Brayer, J. Mol. Biol. 223 (1992) 959–976.
[82] A. Warshel, S.T. Russell, A.K. Churg, Proc. Natl. Acad. Sci. USA 81 (1984) 4785–4789.
[83] H. Nakamura, T. Sakamoto, A. Wada, Protein Eng. 2 (1988) 177–183.
[84] K. Sharp, A. Jean-Charles, B. Honig, J. Phys. Chem. 96 (1992) 3822–3828.
[85] D. Voges, A. Karshikoff, J. Chem. Phys. 108 (1998) 2219–2227.
[86] J.O.M. Bockris, A.K.N. Reddy, Modern Electrochemistry, Plenum, New York, 1973.
[87] A. Nicholls, B. Honig, J. Comp. Chem. 12 (1991) 435–445.
[88] R.L. Dunbrack, M. Karplus, Nature Struct. Biol. 1 (1994) 334–340.
[89] R.L. Dunbrack, F.E. Cohen, Protein Sci. 6 (1997) 1661–1681.
[90] M.R. Gunner, M. Saleh, E. Cross, A. ud-Doula, M. Wise, Biophys. J. (1999) In press.
[91] W.G.J. Hol, Prog. Biophys. Mol. Biol. 45 (1985) 149–195.
[92] J. Deisenhofer, H. Michel, Annu. Rev. Biophys. Chem. 20 (1991) 247–266.
[93] M.R. Gunner, B. Honig, in: R.A. Scott, A.G. Mauk (Eds.), Cytochrome *c* Sourcebook, University Science Books, Mill Valley, 1996, pp. 347–372.
[94] G.R. Moore, in: D.S. Bendall (Ed.), Protein Electron Transfer, Bios, Cambridge, 1996, pp. 189–248.
[95] B. Heller, D. Holten, C. Kirmaier, Science 269 (1995) 940–945.
[96] S.M. Dracheva, L.A. Drachev, A.A. Konstantinov, A.Y. Semenov, V.P. Skulachev, A.M. Arutjunjan, V.A. Shuvalov, S.M. Zaberezhnaya, Eur. J. Biochem. 171 (1988) 253–264.
[97] G. Fritzsch, S. Buchanan, H. Michel, Biochim. Biophys. Acta 977 (1989) 157–162.
[98] W. Nitschke, A.W. Rutherford, Biochemistry 28 (1989) 3161–3168.
[99] G. Alegria, P.L. Dutton, Biochim. Biophys. Acta 1057 (1991) 239–257.
[100] N.W.T. Woodbury, W.W. Parson, Biochim. Biophys. Acta 767 (1984) 345–361.
[101] A.R. Holzwarth, M.G. Muller, Biochemistry 35 (1996) 11820–11831.
[102] T. Watanabe, M. Kobayashi, Chlorophylls, 1991, pp. 287–315.
[103] R.G. Alden, W.W. Parson, Z.T. Chu, A. Warshel, J. Phys. Chem. 100 (1996) 16761–16770.
[104] R.G. Alden, W.W. Parson, Z.T. Chu, A. Warshel, in: M.-E. Michel-Beyerle (Ed.), The reaction center of photosynthetic bacteria. Structure and dynamics, Springer, Garching, 1995, pp. 105–115.
[105] M.R. Gunner, in: M.E. Michel-Beyerle (Ed.), Proc. Feldefing III Workshop in Reaction Centers of Photosynthetic Bacteria. Structure and Dynamics., Springer-Verlag, New York, in press, 1995.
[106] G.R. Fleming, J.L. Martin, J. Breton, Nature 333 (1988) 190–192.
[107] G.L. Gaines, M.P. O'Neil, W.A. Svec, M.P. Niemczyk, M.R. Wasielewski, J. Am. Chem. Soc. 113 (1990) 719–721.
[108] D. DeVault, Q. Rev. Biophys. 13 (1980) 387–564.
[109] R.A. Marcus, N. Sutin, Biochim. Biophys. Acta 811 (1985) 265–322.
[110] G.L. Closs, L.T. Calcaterra, N.J. Green, K.W. Penfield, J.R. Miller, J. Phys. Chem. 90 (1986) 3673–3683.
[111] J.M. Ortega, P. Mathis, J.C. Williams, J.P. Allen, Biochemistry 35 (1996) 3354–3361.
[112] C.C. Moser, J.M. Keske, K. Warncke, R. Farid, P.L. Dutton, Nature 355 (1992) 796–802.
[113] M.Y. Okamura, G. Feher, Annu. Rev. Biochem. 61 (1992) 861–896.
[114] M.S. Graige, M.L. Paddock, J.M. Bruce, G. Feher, M.Y. Okamura, J. Am. Chem. Soc. 118 (1996) 9005–9016.
[115] J. Miksovska, L. Kalman, M. Schiffer, P. Maroti, P. Sebban, K. Hanson, Biochemistry 36 (1997) 12216–12226.
[116] J. Miksovska, M. Valerio-Lepiniec, M. Schiffer, D. Hanson, P. Sebban, Biochemistry 37 (1998) 2077–2083.
[117] M.L. Paddock, M.S. Graige, G. Feher, M.Y. Okamura, Proc. Natl. Acad. Sci. USA 96 (1999) 6183–6188.
[118] M.Y. Okamura, G. Feher, in: R. Blankenship, M. Madigan, C. Bauer (Eds.), Anoxygenic Photosynthetic Bacteria, Kluwer Academic Publishers, Dordrecht, 1995, pp. 577–593.
[119] D.M. Tiede, J. Vazquez, J. Cordova, A.P. Marone, Biochemistry 35 (1996) 10763–10775.
[120] M.S. Graige, G. Feher, M.Y. Okamura, Proc. Natl. Acad. Sci. USA 95 (1998) 11679–11684.
[121] J. Li, D. Gilroy, D.M. Tiede, M.R. Gunner, Biochemistry 37 (1998) 2818–2829.
[122] P. Maroti, C.A. Wraight, Biochim. Biophys. Acta 934 (1988) 329–347.
[123] P.H. McPherson, M.Y. Okamura, G. Feher, Biochim. Biophys. Acta 934 (1988) 348–368.

[124] E. Takahashi, C.A. Wraight, Biochim. Biophys. Acta 1020 (1990) 107–111.
[125] M.L. Paddock, A. Juth, G. Feher, M.Y. Okamura, Biophys. J. 61 (1992) A153.
[126] E. Takahashi, C.A. Wraight, Biochemistry 31 (1992) 855–866.
[127] P. Maroti, D.K. Hanson, L. Baciou, M. Schiffer, P. Sebban, Proc. Natl. Acad. Sci. 91 (1994) 5617–5621.
[128] R. Hienerwadel, S. Grzybek, C. Fogel, W. Kreutz, M.Y. Okamura, M.L. Paddock, J. Breton, Biochemistry 34 (1995) 2832–2843.
[129] M.L. Paddock, G. Feher, M.Y. Okamura, Biochemistry 36 (1997) 14238–14249.
[130] P.H. McPherson, M.Y. Okamura, G. Feher, Biochim. Biophys. Acta 1144 (1993) 309–324.
[131] M.L. Paddock, P.H. McPherson, G. Feher, M.Y. Okamura, Proc. Natl. Acad. Sci. USA 87 (1990) 6803–6807.
[132] E. Takahashi, C.A. Wraight, Biochemistry 31 (1992) 855–866.
[133] M.L. Paddock, S.H. Rongey, P.H. McPherson, A. Juth, G. Feher, M.Y. Okamura, Biochemistry 33 (1994) 734–745.
[134] D.K. Hanson, L. Baciou, D.M. Tiede, M. Nance, M. Schiffer, P. Sebban, Biochim. Biophys. Acta 1102 (1992) 260–265.
[135] L. Baciou, H. Michel, Biochemistry 34 (1995) 7967–7972.
[136] M.H.B. Stowell, T.M. McPhillips, D.C. Rees, S.M. Soltis, E. Abresch, G. Feher, Science 276 (1997) 812–816.
[137] E.C. Abresch, M.L. Paddock, M.H.B. Stowell, T.M. McPhillips, H.L. Axelrod, S.M. Soltis, D.C. Rees, M.Y. Okamura, G. Feher, Photosynth. Res. 55 (1998) 119–125.
[138] J. Tandori, P. Sebban, H. Michel, L. Baciou, Biochemistry 38 (1999) 13179–13187.
[139] M.L. Paddock, S.H. Rongey, G. Feher, M.Y. Okamura, Proc. Natl. Acad. Sci. USA 86 (1989) 6602–6606.
[140] E. Alexov, M. Valerio-Lepiniec, L. Baciou, M. Schiffer, D.K. Hanson, P. Sebban, M.R. Gunner, Biochemistry, in press.
[141] R. Wolfenden, Science 222 (1983) 1087–1093.
[142] F.A. Quiocho, J.S. Sack, N.K. Vyas, Nature 329 (1987) 561–564.
[143] R. Langen, G.M. Jensen, U. Jacob, P.J. Stephens, A. Warshel, J. Biol. Chem. 267 (1992) 25625–25627.
[144] V. Spassov, A. Karshokoff, R. Ladenstein, Protein Sci. 3 (1994) 1556–1569.
[145] D.J. Barlow, J.M. Thornton, J. Mol. Biol. 168 (1983) 867–885.
[146] D.J. Barlow, J.M. Thornton, Biopolymers 25 (1986) 1717–1733.
[147] Z. Hendsch, B. Tidor, Protein Sci. 3 (1994) 211–226.
[148] Z.S. Hendsch, T. Jonsson, R.T. Sauer, B. Tidor, Biochemistry 35 (1996) 7621–7625.
[149] V. Lounnas, R.C. Wade, Biochemistry 36 (1997) 5402–5417.
[150] Z.S. Hendsch, C.V. Sindelar, B. Tidor, J. Phys. Chem. 102 (1998) 4404–4410.