CHAPTER 1 Introduction

Physicochemical properties of complex molecular systems in solution strongly depend on the protonation pattern and the redox behavior. The fundamental quantities to describe molecular protonation patterns and redox behavior are pK_a values[1-3] and standard redox potentials E^0_{redox} [2, 4, 5] respectively. Titratable residues play key roles in such fundamental processes as protein folding, substrate binding, enzyme reactions and redox behavior.[6-12] The mechanism and kinetics of the hydrolytic reaction catalyzed by lysozyme or serine proteases depend on the protonation characteristics of the titratable amino acids in the catalytic center.[13] Hydrogen bonding is also crucial for the stability and proper behavior of the desoxyribonucleotide acid (DNA). In the latter case hydrogen-bonding patterns between nucleobases stabilize the structure of the DNA double helix.[13]

Redox-active compounds play important roles in chemistry, biochemistry and photochemistry.[11, 12, 14-19] Phenoxyl radicals are major pollutants in aqueous solution that cause severe environmental damages.[16] Their aromatic ring system renders them suitable to sustain redox reactions and electron transport in enzymes.[20] Since carboxyl and benzoyloxyl radicals easily decarboxylate to form CO_2 and alkyl radicals, which is known as the Kolbe reaction, their E^0_{redox} values are difficult to evaluate.[21-23] The free radical chemistry of carbthiyl radicals ($RCOS^\bullet/COS^-$) and benzoylthiyl radicals ($PhCOS^{\bullet}/COS^-$) represents the chemistry of the sulfur counterpart of carboxyl and benzoyloxyl radicals.[23-25] Albeit of being important in organic chemistry, for example polymer chemistry, experimentalists have only recently started to investigate their redox potentials.[23-25]

Despite the amount of performed investigations to understand coupled proton electron transfer in proteins and in solution remains a challenging problem.[11, 12, 19, 20, 26] Quinones are ubiquitous in electron transfer chemistry and biochemistry.[11, 12, 27, 28] In electrochemistry *p*-benzoquinone and *p*-duroquinone serve commonly as aqueous redox indicators.[20, 29] With their ability to bind 2 protons and 2 electrons, quinones have nine different states in terms of protonation and redox potential.[11, 12, 19, 30, 31] In aqueous solution at physiological pH, these quinone states extend from being fully oxidized (Q) to fully reduced (QH_2) in a concerted two-electron and two-proton reaction.[11, 12, 19, 32] The complete reduction cascade of 1,4-benzoquinone from the fully oxidized to the fully reduced state is depicted in Scheme 3.1. Many essential electron transfer processes in biological systems are conveyed by quinones. They are predominantly found in transmembrane proteins with coupled proton electron transfer reactions, including the photosynthetic reaction centers, PSII, PSI, cytochrome *bc*1, and *b*6*f* complexes and in quinol:fumarate and succinate:quinone oxidoreductases.[11, 12, 19, 30-34] The bacterial photosynthetic reaction center is a protein pigment complex in the membrane of purple bacteria.[11, 12, 19, 32] It converts light energy into electrochemical energy by coupling photoinduced electron transfer to proton uptake from cytoplasm. Electronic excitation of the special pair, a bacteriochlorophyll dimer, induces a multistep electron transfer from the special pair to a quinone referred as Q_A.[11, 12, 32] From there the electron moves to a secondary quinone referred as Q_B. After this initial reaction, a second electron transfer from Q_A to Q_B and two protonation reactions resulting in a dihydroquinone Q_BH_2 occur. The dihydroquinone leaves its binding site and is replaced by an oxidized quinone from the quinone pool.[11, 12, 31, 32]

Inspection of the scientific literature reveals that even after years of intense research experimental results of pK_a values and E^0_{redox} for important chemical groups are incomplete or in conflict with each other,[8, 15, 16, 18, 23, 27, 28, 35-41] which indicates principal methodic difficulties to obtain those data. Hence, a reliable computational procedure, which fills the gap would be an important tool for the chemical and biochemical community. Estimated pK_a and E^0_{redox} values could also be applied to analyze microscopic features of molecular systems, which are responsible for their observed macroscopic behaviors. Additionally the determination of the protonation states of novel compounds in aqueous solution is also a challenging and important objective of computational chemistry, because a reliable protocol could be helpful to preview the functionality of pharmacological relevant agents before they are synthesized.[1, 41]

Most theoretical studies on the computation of absolute pK_a values and E^0_{redox} in the last two decades were performed with the help of a suitable thermodynamic cycle (see chapter 2.1.1 and 2.1.2). [1, 20, 35, 40-54] Specific quantum chemical (QC) methods ranging from simple Hartree-Fock[45, 47] to high quality density functional theory (DFT)[41, 45, 46, 49, 51, 54, 55] or CPU time and memory demanding post Hartree-Fock approaches[35, 40, 43, 52, 53] to determine gas phase energetics were combined with a reaction field method[1, 5, 40, 42-46, 48, 50-53] to determine solvation free energies (for more information on the QC methods see chapter 2.2). To determine solvation energies in this way requires to perform QC computations iteratively adjusting the reaction field applied to the wave function until self-consistency is reached, i.e. the reaction field is equal to the electrostatic field that the wave function induces in the dielectric medium. The reaction field is calculated by solving the Poisson equation for the solute molecule in the dielectric medium representing the solvent. Many applications of this approach to estimate pK_a values can be found in the literature.[4, 5, 43, 44, 46, 48, 50, 51, 55]

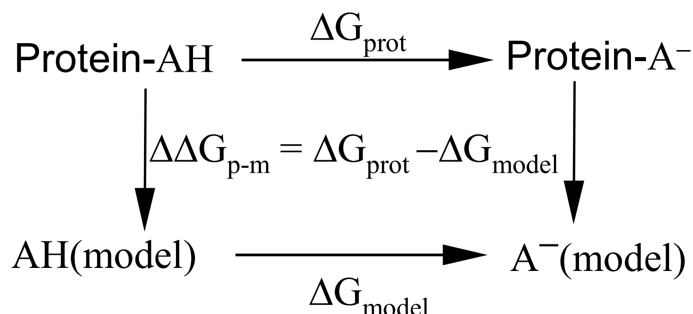
So far theoretical studies on pK_a values and E^0_{redox} focused on a specific chemical group and were therefore generally not transferable to other substance classes.[40, 42-45, 48-51, 54, 55] If data, which included a spectrum of different agents were considered extensive fitting was carried out to reach convergence between measured and estimated results[1] or considerably discrepancies were obtained.[1, 46, 52, 53, 56] Hence, there is need to develop a computational procedure, which enables to predict reliably pK_a values and redox potentials for a large spectrum of organic compounds.

1.1 Determination of pK_a and $E°_{redox}$ shifts by electrostatic energy computation

Because of their outstanding significance, electrostatic interactions in proteins have been investigated intensively in the last decades.[6-9, 11, 12, 57-64] Electrostatic interactions inside a protein prevail for instance between titratable (redox-active) residues as well as between cofactors and titratable (redox-active) residues and the polar aqueous environment.[7, 32, 34, 59, 61] For instance it became more and more certain in the past decade that electrostatic interactions comprise the dominant contribution to reduce the transition state energy in enzymatic reactions.[59, 65] Hence, the function of biological systems is linked strongly to the electrostatic properties of its particles.
The framework to compute pK_a shifts of acid-base amino acids inside a protein was first developed in the pioneering studies by Warshel and coworkers.[6, 66] These studies had an enormous outspread and motivated researcher from many different scientific origin to investigate pK_a values in biological systems. Nowadays, the established procedure to determine pK_a values of

titratable groups in proteins is to compute the electrostatic free energy by solving the Poisson-Boltzmann equation[7, 8, 10-12, 32, 60, 61, 63, 64, 67] or to apply the generalized Born approach.[68] In this approach, the electrostatic free energy difference between the protonated (p) and deprotonated (d) molecular compound is calculated in the protein environment and a suitable model system yielding $\Delta G_{protein} = G^{p}_{protein} - G^{d}_{protein}$ and $\Delta G_{model} = G^{p}_{model} - G^{d}_{model}$, respectively.[7, 11, 12, 67] The model system involves generally the same compound in aqueous solution where the pK_a is known from experiment and is used to evaluate the double free energy difference $\Delta\Delta G_{p\text{-}m} = \Delta G_{protein} - \Delta G_{model}$.[7, 67] The thermodynamic cycle in Scheme 1.1 illustrates the contributions to the double free energy difference $\Delta\Delta G_{p\text{-}m}$. It is important to realize that the proton solvation cancels in the determination of the pK_a shift of the model compound from aqueous solution into the protein environment.

Scheme 1.1. Thermodynamic cycle to analyze the pK_a shift of a titratable compound as it is transferred from the aqueous solution into the protein environment.



In Scheme 1.1 ΔG_{prot} represents the standard free energy of deprotonation of a titratable group inside a protein, whereas ΔG_{model} is the solution-phase standard free energy of deprotonation. $\Delta\Delta G_{p\text{-}m}$ is a double difference in the free energies between the free energy difference of the upper and lower leg: $\Delta\Delta G_{p\text{-}m} = \Delta G_{prot} - \Delta G_{model}$. Assuming that the non-electrostatic intermolecular interactions (van-der-Waals interactions) are independent from the protonation state and that the non-electrostatic intramolecular interactions (covalent interactions) are independent from the influence of the environment (protein or solvent) the double free energy difference $\Delta\Delta G_{p\text{-}m}$ is proportional to the shift in the pK_a value of the titratable compound between the solvent and the protein. Applying the definition of the pK_a value (see eq. 2.6) it can be expressed in the protein according to the subsequent equation:

$$pK_a = -\log(K_a) = \Delta G_R / 2.303\ RT \quad (1.1)$$

In eq. 1.1 ΔG_R is the reaction free energy, R is the gas constant and T the reaction temperature in Calvin. The established framework to estimate pK_a shifts for titratable groups between aqueous solution and the protein environment is transferable for redox–active compounds.[8, 63] To calculate E°_{redox} shifts between solvent and biological systems the above equation has to be reformulated:

$$E_m^{protein} = E_{sol}^{0} + \frac{\Delta G}{F} \quad (1.2)$$

where F is the Faraday constant and ΔG the reaction free energy. Eq. 1.2 is related to the Nernst law, which is the analogue to the Henderson-Hasselbalch (see eq. 2.5) for protonation equilibria of titratable groups.

The success of the method is documented in many applications. The determination of E^0_{redox} for quinones inside the bacterial photosynthetic reaction centres, which are commonly referred to as Q_A and Q_B enabled to elucidate electron transfer processes therein.[11, 12, 32, 34] Based on the described theoretical framework new insights of the radical transfer in the DNA photolyase were also obtained.[8]

Most studies on pK_a and E^0_{redox} shifts were based on solvation free energies evaluated with the linearized Poisson-Boltzmann equation (LPBE)[7, 8, 12, 32, 34, 61, 63, 64, 67] with partial atomic charges derived from a QC calculation.[7, 8, 12, 32, 34, 61, 63, 64, 67] It is commonly accepted to transform the wave function in the vacuum with one of the electrostatic potential based methods[69-74] that will be introduced in the method section (chapter 2.3). Solving the LPBE (see chapter 2.4) in that way ensures that the QC computation is strictly separated from the estimation of solvation free energies. The success of these studies prompted us to exchange the model compound with suitable *ab initio* QC methods to calculate pK_a and E^0_{redox} values in solution without additional experimental data. Under specific circumstances, measured solution pK_a values (pK_a^{model}) are ambiguous or difficult to obtain as it is the case for flavins[8] and quinones,[11, 12, 32, 34] that appear in a number of different enzymes and possess different redox and protonation states. A successful protocol to obtain by means of *ab initio* techniques the physico- chemical properties could also help to obtain more accurate pK_a shifts or E^0_{redox} redox shifts in biological systems. Furthermore, a quantitative agreement of the energetics of solvation and proton uptake of a large number of different compounds can shed new light on the understanding of the underlying protonation/deprotonation equilibria and solute-solvent interactions.

1.2 Ab initio computation of pK_a values and $E°_{redox}$ values

In accordance with the literature thermodynamic assumptions (see chapter 2.1.1 and 2.1.2) guided us to compute absolute pK_a and E^0_{redox} values from a contribution of the reaction free energy in the gas phase and the solvation free energy differences between educts and products. The energetics in the vacuum were evaluated directly with suitable QC methods whereas solvation free energies were estimated within a *two-step* procedure (see below).

Single-molecule studies on absolute pK_a values were performed for a heterogeneous set of organic compounds. Based on the QC DFT functionals Becke-half&half (Becke$^1/_2$)[75] and B3LYP[76, 77] and the two-step procedure for solvation free energies computed pK_a values converged to experimental data (see chapter 3.1). Convergence between experimental and computed one electron reduction potentials for a spectrum of eight different functional groups (phenoxyl, p-benzoquinone, phenylthiyl, p-benzodithiyl, carboxyl, benzoyloxyl, carbthiyl and benzoylthiyl) in protic (water) and aprotic (acetonitrile, N,N-dimethylacetamide) solvents were obtained using the QC method Gaussian3-MP2 (G3MP2)[78] and the procedure to obtain solvation free energies applied in the computation of absolute pK_a values (see chapter 3.1).

Application of thermodynamic cycles (see 2.1.1 and 2.1.2) to compute pK_a values and E^0_{redox} require an accurate evaluation of the gas phase energetics to ensure that accurate final results in

the condensed phase are not due to an error compensation between gas phase and solvation phase energetics. Proton affinity (PA) and electron affinity (EA) describe the enthalpic contributions to ΔG_{gas} of a protonation and a redox reaction respectively.

These energies can be compared to experimental results. It will be shown that computed gas phase results using the QC DFT functionals Becke($^1/_2$)[75] and B3LYP[76, 77] with the cc-pVQZ basis set[79, 80] on optimized molecular geometries converged sufficiently well to experimental data (3.1.2). PA values based on the latter two DFT functionals were compared with the performance of other QC methods and basis sets (see chapter 3.1.4). Even for modern computational quantum chemistry the quantitative description of EA values is especially demanding,[79-81] because the correlation energy of the unpaired electron in the radical state exceeds the correlation energy of the paired electrons. The necessity to use the post Hartree-Fock method G3MP2[78] to estimate accurately EA values is documented in chapter 3.2.1 and 3.2.2.

In the present study solvation free energies were evaluated in two separate steps. First, solute atomic partial charges were determined under vacuum conditions with a QC method of moderate accuracy by matching the electrostatic potential (ESP)[82] based on the QC wave function with the ESP generated by the atomic partial charges using the restrained eletrostatic potential (RESP)[69, 70] procedure (for an explanation of RESP see chapter 2.3). Second, the electrostatic energies of solvation were evaluated from this point charge distribution without further involvement of QC methods by solving the Poisson equation. With this procedure one can account for long range effects of electrostatic energies more easily.[35, 41, 83] Here I refer to this procedure as the *two-step* method as opposed to the generally used direct method, which we consequently called *one-step* procedure. Chapter 2.7 provides a coherent picture of the two-step procedure and compares to the commonly applied one-step procedure.

Neither the solvation free energy of the proton nor the potential of the normal hydrogen electrode, which contribute to computed pK_a and E^0_{redox} values, respectively, were estimated in the scope of the present study (see Scheme 2.1.1 and 2.1.2). Due to inconsistent experimental data[45, 84-89] I used for closed shell organic compounds a proton solvation free energy, within the experimental error range, which then minimized the root-mean-square value between experimental and computed pK_a values. Section 2.5 conveys the framework of important experimental[84, 85, 87, 88, 90] and theoretical studies[89] to determine $\Delta G_{solv}(H^+)$. One-electron reduction potentials were estimated using the recommended literature value of the standard (normal) hydrogen electrode.[35, 55, 90] The derivation of this potential by Reiss and Heller[90] is explained in chapter 2.6.

It will be shown in chapter 3.1.1 and 3.1.2 that the DFT functional B3LYP using the augmented cc-pVTZ performs significantly worse for open shell systems than for closed shell systems. Inspection of the quinone reduction cascade (see scheme 3.1) reveals the concomitant occurrence of proton and electrons transfer processes. To model the reduction cascade of quinones the protonation/deprotonation of open shell systems must be treated accurately. I will show in Chapter 3.2 that the QC method G3MP2 yields excellent result in terms of PA values. Using the latter QC method and the two-step procedure to estimate electrostatic solvation free energies, therefore enabled to evaluate the complete reduction cascade of the fully oxidized quinone to the fully reduced QH_2 in the condensed phase. Computed pK_a values based on G3MP2 no longer employ $\Delta G_{solv}(H^+)$ as an adjustable parameter but uses one value that is recommended in the work of Tissandier et al.[85]