CHAPTER 4 Discussion and Conclusion

Discussion

In the previous chapters I introduced consistent protocols to calculate pK_a values and redox potentials in different solvents. Chapter 3.1 described the computation of pK_a values for closed shell organic molecules. Chapter 3.2 showed the calculation of one-electron reduction potentials in different solvents. In chapter 3.3 I extended the estimation of absolute pK_a values to open-shell systems, so that the energetics of proton-electron coupled reactions can now be simulated.

It is a common feature in theoretical studies on the computation of pK_a values and redox potentials to derive the reaction free energy ΔG_R with the help of suitable thermodynamic cycles (see Scheme 2.1.1 and 2.1.2) from contributions in the gas phase of ΔG_{gas} and from solvation free energies $\Delta\Delta G_{solv}$. This procedure was successfully applied in the current thesis as documented in chapter 3.1, 3.2 and 3.3. Comparison of computed PA and EA values with experimental data (see chapter 3.1.2 and 3.2.2) ensured that the obtained convergence between measured and estimated pK_a and E^0_{redox} values is not due to an error compensation between computed ΔG_{gas} and $\Delta\Delta G_{solv}$ values.

As documented in chapter 3.1.1 I computed successfully absolute pK_a values for a spectrum of six different functional groups, which contained O-atom and N-atom titratable compounds. The substances of four groups (carboxylic acids, benzoic acids, phenols and imides) changed upon deprotonation from the neutral to the anionic state, whereas the members of substituted pyridines and imidazoles changed upon deprotonation from the cationic to the neutral state. In spite of the heterogeneous data set the computed pK_a values were accurate with RMS deviations of 0.53 pK_a and 0.57 pK_a units and maximum errors of 1.0 and 1.3 pK_a units based on QC computations with the DFT functionals Becke($^1/_2$)[75] and B3LYP,[76, 77] respectively.

Although the considered molecules in chapter 3.1.1 belonged to a heterogeneous data set the accurateness of obtained pK_a values is competitive to results presented in other theoretical studies, where only a distinct chemical functional group or subset thereof was studied. Using QC methods and the PCM reaction field method[155, 156] to estimate solvation free energies Schüürmann et al[48] concentrated on the computation of substituted carboxylic acids, without obtaining a quantitative agreement between measured and computed data. Liptak et al.[43] computed pK_a values for substituted carboxylic acids yielding a RMS deviation of only 0.4 pK_a units when compared to experimental data. They evaluated accurately the deprotoantion reaction in the gas phase using the post-Hartree-Fock QC method CBS-QB3.[161] They computed solvation free energies with the CPCM reaction field method[157, 158] on vacuum optimized geometries using a $\Delta G(H^+)$ value of −264.61 kcal/mol, which is close to the value proposed recently by Tissandier et al. (see chapter 2.5.3).[85] In a subsequent study they estimated pK_a values for substituted phenols.[40] To obtain agreement with experimental data in this study a modified geometry optimization in the presence of the reaction field was performed. Using otherwise the same parameters than before they obtained agreement with experimental data for substituted phenols with an RMS deviation of 0.4 pK_a units. The modified procedure for geometry optimization compensated for relative errors in the solvation free energy differences between the two functional groups. To compare the results of Liptak et al.[40, 43] with the one of the current

thesis one should note that the CPCM method[157, 158] can not compute solvation free energies for the carboxylic and benzoic acids consistently. In the present thesis a coherent protocol was given to estimate absolute pK_a values for carboxylic acids and substituted phenols (see Table 3.1). Here I computed pK_a values for carboxylic acids with RMS deviations of 0.56 and 0.90 kcal/mol for Becke($^1/_2$)[75] and B3LYP,[76, 77] respectively. The results obtained with Becke($^1/_2$) are of comparable accurateness to the results of Liptak et al..[40] Computed pK_a values for substituted phenols amounted to an RMS deviation of 0.54 and 0.35 pK_a units based on Becke($^1/_2$) and B3LYP, respectively. Both values are competitive to the results calculated by Liptak et al.[40] Goddard III and coworkers[49] concentrated on the estimation of pK_a values of 5-substituted uracils only, which converged to experimental data with high accuracy. In these computations the largest discrepancy from a measured value amounted to 0.67 kcal/mol for 5-fluorouracil. Solvation free energies were computed using the reaction field approach that is implemented in the QC software package Jaguar.[162, 163] Although I did not include substituted uracils in this study computed pK_a values for imides (see Table 3.1), which are structurally related to the pyrimidines, match measured pK_a values. Goddard III and coworkers[49] evaluated gas phase free energies on optimized geometries using B3LYP with the aug-cc-pVTZ basis set.[79, 80] Usage of aug-cc-pVTZ instead of the cc-pVQZ basis set provided PA values that are systematically lower than the gas-phase results presented in chapter 3.1.2 particularly for molecular compounds with a titratable oxygen (see Table 3.2 and 3.3 and Figure 3.6). To obtain agreement between computed and calculated results for substituted uracils GoddardIII and coworkers applied a questionable solute dielectric constant that is lower than one. Chen and MacKerell Jr[51] computed relative instead of absolute pK_a values, but restricted themselves to substituted pyridines only. They computed accurate PA values using the QC method MP2, but did not obtain overall convergence with experimental data, because tested reaction field models failed to reproduce acceptable $\Delta\Delta G_{solv}$ values for all of the considered pyridines.[51] Therefore a consistent protocol to calculate relative pK_a values for substituted pyridines was not obtained. In the current thesis it was shown that the less demanding DFT functionals Becke($^1/_2$) and B3LYP were sufficient to compute accurate PA values. Computed pK_a values for substituted pyridines match accurately with experimental data independently of the applied QC method. Topol[45] calculated absolute pK_a values for a series of four substituted imidazoles (imidazole, 2-methyl, 2-amino and 2-chloro imidazole). Their RMS deviation amounted to 0.80 pK_a units regarding the four types of imidazoles. The largest error occurred for 2-amino imidazole, which deviated from the measured value by 1.30 pK_a units. Inspection of Table 3.1 reveals that our computed data agree within 0.82 pK_a units when Becke($^1/_2$) and 0.98 pK_a units when B3LYP is used. The RMS deviation for the five considered imidazoles amounted to 0.56 and 0.54 pK_a units with Becke($^1/_2$) and B3LYP, respectively.

Guida and coworkers[1] computed accurate pK_a values for a spectrum of functional groups, which comprised a larger data set than the one considered in chapter 3.1.1, but needed a number of adjustable parameter to reach convergence with measured data. The vdW radii of solvation were adjusted for each functional group. The final pK_a value was then computed from a raw pK_a values according to:

$$pK_a = A(pK_a) + B \qquad (2.4.1)$$

using different sets of A and B for each considered functional group. The proton solvation free energy was lumped into the factor B. Richardson et al.[46] computed pK_a values for a heterogeneous set of organic compounds. They used a similar set of vdW radii as in this study

(see chapter 2.7), but applied a reaction field method, which is implemented in the ADF software package. In this study only qualitatively pK_a values could be reproduced.

Computations of absolute pK_a values (in contrast to relative pK_a values) requires the inclusion of the Gibbs free energy of proton solvation $\Delta G_{solv}(H^+)$ (see chapter 2.1.1). Unfortunately, this quantity is difficult to determine computationally as well as experimentally.[40, 43, 45, 46, 54, 84, 85, 87, 89, 90, 120]It was shown in chapter 2.5.2 that Noyes[84] derived the proton solvation free energy from thermodynamic measurements. With this approach he deduced $\Delta G_{solv}(H^+) = -260.5$ kcal/mol.[84] This value was applied in the studies on absolute pK_a values of Richardson et al. and Quennville et al.[42, 46] Quennville et al.[42] investigated recently the pK_a value of His-291, in the catalytic center of cytochrome c oxidase that binds a heme-iron cofactor, using the DFT functional B3LYP with the 6-311+G* basis set and the same reaction field method used by GoddardIII and coworkers.[49] They used a $\Delta G_{solv}(H^+) = -258.32$ kcal/mol, which minimized the discrepancy between measured and computed pK_a values in their study. The basis set 6-311+G* is a triple zeta basis set that leads to similar PA values than the aug-cc-pvTZ basis set, that was applied by GoddardIII and coworkers (see above). The discrepancy between the two $\Delta G_{solv}(H^+)$ values of almost 2.5 kcal/mol suggests that the procedure by Quennville et al.[42] is not transferable to substituted pyrimidines. It was shown in chapter 2.5 that $\Delta G_{solv}(H^+)$ can be derived from measurements of the SHE electrode.[84, 87, 88, 90, 121, 124, 125] The average f –259.5 kcal/mol derived from five independent measurements of the standard hydrogen electrode was recommended in pK_a computations of Lim et al.[Lim, 1991 #164] It is documented in chapter 2.5.2 that four of the five measurements[88, 90, 124, 125] were close to each other, whereas one deviated significantly.[121] Taking the value of Reiss and Heller[90] $\Delta G_{solv}(H^+)$ would be equal to –263.15 kcal/mol (see chapter 2.5.2). The cluster-base-pair approximation by Tissandier et al. yielded $\Delta G_{solv}(H^+) = -263.96$ kcal/mol (see chapter 2.5.3). Liptak et al. employed a value of $\Delta G_{solv}(H^+) = -264.61$ kcal/mol, which is close the value suggested by Tissandier. To compute accurate pK_a values for closed-shell organic molecules I used a $\Delta G_{solv}(H^+)$, that minimized the RMS deviation between measured and computed pK_a values (see chapter 3.1).[41] Using the QC method G3MP2[78] for the computation of pK_a values allowed to apply $\Delta G_{solv}(H^+) = -264.00$ kcal/mol, which was suggested by recently Tissandier et al.[85]

Chapter 3.1.2 proofs that PA values for closed-shell organic compounds computed with the DFT functional Becke($^1/_2$) are as accurate as values calculated with the DFT functional B3LYP. It was shown that the B3LYP functional has a hybridization, exchange and correlation functional whereas Becke($^1/_2$) has only a hybridization and an exchange functional (see chapter 2.2.3). Comparison of absolute electronic energies for a protonated and a deprotonated compound revealed that B3LYP yields lower energies than Becke($^1/_2$) (data not shown). PA values are enthalpic energy differences between the protonated and deprotonated states. Therefore, one might assume that the effects of electron correlation vanish in the energy difference of a PA value. Computed EA values based on Becke($^1/_2$) did not converge with experimental data (data not shown). Upon electron transfer the electronic state changes from the radical state to the non-radical state or in opposite direction. EA values describe the energy difference between these two states. The failure of Becke($^1/_2$) to compute accurately EA values indicates that for one-electron reduction potentials the correlation effects are of utmost importance.

The procedure to compute accurate pK_a values was successfully adapted to compute the energetics of one-electron reduction processes in the condensed phase as documented in chapter

3.2. Compared to experimental data computed redox potentials yielded for 21 considered organic compounds in different solvents overall RMS deviations of 0.058 V and 0.131 V for G3MP2 and B3LYP, respectively (see Table 3.15).[35] For both QC methods the atomic partial charges were derived using the RESP[69, 70] method from a vacuum QC ESP[109] calculated for the optimized geometries at the B3LYP/6-31G** level of theory.

To compute E^0_{redox} in acetonitrile Guo and coworkers[53] recently combined the DFT functional B3LYP[76, 77] using a triple zeta basis set with the PCM reaction field method[155, 156] for a number of structurally unrelated organic compounds. They obtained an averaged discrepancy of 0.170 V between measured and computed results. Truhlar and coworkers[44, 50] tested the performance of their homemade reaction field model (SM5.42R) to reproduce experimental redox potentials for substituted anilines and phenols. As Guo and coworkers[53] they applied DFT methods to evaluate gas phase energetics. But instead of the B3LYP functional they applied the BPW91[76, 100] functional. Considerable deviations between measured and computed redox-potentials resulted in both studies[44, 50].

As for absolute pK_a values (see chapter 3.1) the convergence between measured and calculated E^0_{redox} values was based on the exact evaluation of ΔG_{gas} (see chapter 3.1.2 and 3.2.2) and the electrostatic solvation free energies using the *two-step* procedure (see chapter 2.7). As already mentioned in chapter 3.2 the considered redox-active compounds change upon reduction from the open-shell to the closed-shell state or in opposite direction for benzoquinone and benzodithiyl (see chapter 3.2.1 and 3.2.2). This change in the electronic state required the application of more elaborate QC methods to calculate EA values for the considered compounds (see chapter 3.2.1 and 3.2.2). Comparison of computed EA values based on the post-Hartree-Fock method G3MP2,[78] DFT computations using the B3LYP[76, 77] functional with the aug-cc-pVTZ basis set[79, 80] and measured values reveals that G3MP2 is the QC method of choice to determine the energetics of one-electron transfer reactions in the gas phase (see Table 3.15 and Figure 3.16). The RMS deviation between measured and computed redox-potentials amounts to 54 meV and 210 meV for G3MP2 and B3LYP, respectively. Consequently, the correlation diagram (Figure 3.16) shows EA values computed with G3MP2 on the diagonal or close to it, whereas EA values estimated with B3LYP deviate significantly.

To evaluate the reliability of the procedure (see chapter 3.1) I restricted the computation of pK_a values to the reproduction of known and accepted experimental data. The considered redox–active compounds play important roles in a variety of different fields as for instance organic, polymer and environmental chemistry. The need of a reliable protocol to estimate one-electron reduction potentials is evident as for the considered 42 reduction potentials only 27 could be compared to measured data.[15, 16, 18, 23-25, 27, 28, 38, 39, 141-143] In chapter 3.2 the introduced protocol already enables to fill in experimental data. To get an idea of uncertainties and difficulties in measured redox potentials it is instructive to consider the E^0_{redox} values of substituted phenoxyls in water, which were determined by several groups. Phenoxyl radicals rapidly dimerize in aqueous solution yielding biphenylic compounds, which constitutes an obstacle for a reliable experimental determination of E^0_{redox}.[16, 38] Pulse radiolysis tries to circumvent this caveat by establishing rapidly an equilibrium between the phenoxyl radical/phenolate ion couple and using a reliable redox pair as reference.[16, 38] Except for the p-amino phenoxyl radical, the E^0_{redox} values of all five p-substituted phenoxyl radicals considered here were measured in water by Lind et al.[38] in 1990 using pulse radiolysis. In these experiments the error margin was between ±10 mV

and ±20 mV except for p-nitro-phenoxyl radical where an error of ±60 mV was given.[38] The same five phenoxyl radicals were studied in 1999 by Li & Hoffman[16] by means of cyclic voltammetry yielding an error of ± 19 mV.[16] Inspection of Table 1 reveals that the E^0_{redox} values of the two groups deviate by 30 mV to 70 mV, which partially exceeds the internal error ranges, albeit it was considered as good agreement. Applying cyclic voltammetry Harriman[15] obtained for the p-methoxy-phenoxyl radical in aqueous solution E^0_{redox} = 0.44 V, which deviates significantly from the data of Lind et al.[38] (E^0_{redox} = 0.54 V) and Li & Hoffman (E^0_{redox} = 0.58 V).[16] These deviations are presumably due to differences in solvent composition and an associated different dimerization behavior of phenoxyl radicals. A theoretical approach possesses the formal advantage to avoid the dimerization problem explicitly.[35] As for the computation of pK_a values the inspected set of molecular compounds comprises a variety of different functional groups.

Experimental E^0_{redox} values of redox-active compounds depend on the solvent. For the 10 redox-active compounds where experimental E^0_{redox} values are available they are larger by 0.28 V to 0.56 V, if solvated in water rather than in AcN/DMAc (Table 3.15). Surprisingly, the $E°_{redox}$ value of p-benzoquinone in methanol ($E^{o\ methanol}_{redox}$ = 100 mV)[154] is much higher than the corresponding value in AcN (E^0_{redox} = −270 mV, Table 1) but nearly identical to the value in water ($E^{o\ water}_{redox}$ = 110 mV, Table 1). This contrasts with the values of the dielectric constant, which are close for methanol[132] and AcN[132] ($\varepsilon_{methanol}$ = 33.0 and ε_{AcN} = 37.5), while for water it is ε_{H2O} = 80. This puzzle is likely related to another characteristic of these solvents and not related to the dielectric constant. Water and methanol have in common that they are both protic solvents and possess polar hydrogens, while AcN and DMAc are aprotic solvents, where polar hydrogens are absent. From all atoms, hydrogens possess the smallest vdW radius and can therefore come closer to solute atoms as the larger non-hydrogen atoms. This invokes large solvent-solute electrostatic interactions with polar hydrogens available in protic but not aprotic solvents. These polar hydrogens contribute significantly to solvation energies of solute molecules in protic solvents, stabilizing the anionic state and thus up-shifting the redox potential. Describing solvent molecules explicitly in atomic detail this effect would automatically be considered by the small vdW radius of the solvent hydrogen atoms. But, in the present study solvent molecules are modeled implicitly using a single solvent radius that refers to the molecule as a whole. Varying this solvent radius cannot account appropriately for such specific effects. Hence, to consider this effect in computations of solvation energies with implicit solvent models the vdW radii of the solute atoms need to be enlarged for aprotic solvents as is done for instance in the PCM[155, 156] or CPCM[157, 158] methods, which are implemented in GAUSSIAN03.[131] Enlargement of solute atom radii was also applied in the present approach.

Theoretical evaluations of redox potentials or pK_a values of molecular compounds in different solvents are generally performed by computing solvation energies directly in a *one-step* procedure solving the molecular wave functions by QC methods in a dielectric medium self-consistently.[1, 4, 5, 43-46, 48-56] Analysis of scientific literature on pK_a and E^0_{redox} values revealed the limitations of those studies. They all faced severe difficulties to compute accurately by means of *ab initio* techniques a broad spectrum of compounds within one consistent protocol.[1, 4, 5, 43-46, 48-56] In this study I applied a *two-step* procedure to evaluate electrostatic free energies that were needed to derive ΔG_R. First, atomic partial charges were determined from the QC ESP[109] by the

quantum chemical method B3LYP/6-31G(d,p) by means of the potential based method RESP[69, 70] (see chapter 2.4) Second, solvation energies were determined, solving the Poisson equation with these atomic partial charges. To calculate one-electron reduction potentials in solution successfully by this procedure, two assumptions must hold. (i) The intramolecular (covalent) interactions do not depend on the environment (solvent or protein). (ii) The nonelectrostatic intermolecular vdW interactions do not depend on the redox state.[35, 41, 83] The main difference between the present *two-step* and the *one-step* procedure is that in the former case the wave function used to determine the atomic partial charges is computed under vacuum conditions, while in the latter case the wave function is evaluated in a dielectric medium. Another more technical point is that in contrast to the *two-step* procedure, the *one-step* procedure is iterative. It may look counterintuitive that the simpler *two-step* procedure should be superior to the elaborate *one-step* procedure in computing redox potentials, but with the *two-step* procedure the electron leakage problem can be better controlled, which is crucial for the quality of results obtained in the present study. The electron leakage effect is particularly strong for the loosely bound excess electron of an anionic molecule if placed in high dielectric medium, where as a consequence the electronic wave function is exploring empty space more excessively than under vacuum conditions. Thus, the electronic energy of the anionic molecular state is lowered considerably. In the real world, the empty space of dielectric medium is filled with molecules whose electrons repel each other by the Pauli exclusion principle, thus preventing excessive electron leakage.[35, 41, 83] That is why in the present study the electronic wave function is computed purposely under vacuum conditions to avoid this electron leakage effect.

Conclusion

In the present study, I provide a consistent procedure to compute absolute pK_a values and one-electron reduction potentials for a broad spectrum of chemically important functional groups (see Figure 3.1a,b and 3.10a,b) comprising oxygen- and sulfur-centered radicals in protic and aprotic solvents. The thermodynamic cycle in Scheme 2.1.1 and 2.1.2 served as a scaffolds to split the energetics of the redox reaction into a gas-phase contribution and solvation free energies. The results presented in Chapter 3.1 documents that the DFT functionals Becke($^1/_2$)[75] and B3LYP[76, 77] give sufficient accurate gas phase results for closed shell organic molecules. The presented results in chapter 3.2 and 3.3 revealed that QC *ab initio* computations based on G3MP2[78] are suitable to evaluate accurately gas-phase reaction energies in terms of electron affinities, while the less demanding QC B3LYP methods sometimes failed to provide reliable electronic energies for the open-shell anionic states.

I used the *two-step* procedure to compute electrostatic energy differences between vacuum and solvent in the computation of pK_a values and one-electron reduction potentials. Other methods that were often applied to calculate $\Delta\Delta G_{solv}$ are based on a reaction field method PCM,[155, 156] CPCM,[157, 158] or JAGUAR's Poisson equation solver.[162] A critical parameter of these approaches is the cavity size defining the boundary between solute and solvent.[45, 54] The definition of the cavity size strongly depends on empirical optimization of several parameters and was often adjusted when new functional groups were considered. Herein,the cavity size is defined by the vdW radii of the solute atoms, whose values were taken from the RESP approach of Kollman and coworkers[69, 70] with minor adjustments (see chapter 2.7). Herein, I demonstrate that for the considered set of chemically quite different titratable organic compounds a uniquely defined protocol can provide accurate pK_a values and one-electron reduction potentials.

Combining computed PA and EA values with solvation energies yielded accurate pK_a values and redox potentials. Estimated pK_a values based on DFT computations agreed with measured values with an RMS deviation of 0.53 and 0.57 pK_a units.[41] Figure 3.2a and 3.3a illustrate the general agreement between computed and experimental results for pK_a values. Computed E^0_{redox} generally agreed with measured values within the experimental error range corresponding to an RMS deviation between a measured and computed E^0_{redox} value of about 50 mV.[35] Figure 3.11 illustrates the general agreement between computed and experimental results. Only the computed E^0_{redox} value of the *p*-aminophenylthiyl radical in acetonitrile deviates considerably from the measured E^0_{redox} value (closed triangle in the lower left of Figure 3.11).

In summary, I would like to point out that with the present study I opened an avenue to evaluate accurate pK_a values and redox potentials of medium-sized organic compounds of arbitrary composition.