On the computation of pK_A values of organic and organometallic molecules in different solvents

Dissertation zur Erlangung des akademischen
Grades des Doktors der Naturwissenschaften

(Dr. rer. nat.)

eingereicht im Fachbereich
Biologie, Chemie, Pharmazie
der Freien Universität Berlin

vorgelegt von

Emanuele Rossini
aus Lucca, Italien

Berlin, Dezember, 2016

Die vorliegende Dissertation wurde unter Anleitung von Prof. Dr. E. W. Knapp im Zeitraum 05/2013 – 12/2016 am Institut für Biologie, Chemie und Pharmazie der Freien Universität Berlin durchgeführt.

1. Gutachter:	Prof. Dr. Ernst Walter Knapp, Freie Universität Berlin
2. Gutachter:	Prof. Dr. Maria Andrea Mroginski, Technische Universität Berlin

Disputation am 03.02.2017

TO FRANCESCO

Abstract

The present thesis comprises four related topics.

(1) An ab-initio method for computing pK_A values of organic molecules involving quantum chemical and electrostatic energy computations is updated and refined. The resulting proton solvation free energies of -255.1, -265.9, -266.3, and -266.4 kcal/mol are determined by matching computed and measured pK_A values for a set of organic compounds in acetonitrile (MeCN), methanol (MeOH), water, and dimethyl sulfoxide (Me_2SO), respectively. These energy values are very accurate and suitable to update the ab-initio method to predict the pK_A values of organic molecules in the four solvents [1, 2].

(2) The revised ab-initio method is applied to predict the pK_A values of di-oxo-manganese complexes inspired by the oxygen evolving complex (OEC) of photosystem II. The obtained results have an accuracy that ranges between 1.19 and 3.69 pH units. These results justify the application of the ab-initio method to titrate the -oxo oxygen atoms of the oxygen evolving complex in photosystem II [3].

(3) A purely electrostatic method to transform known pK_A values from one solvent to another is introduced and tested on 30 organic compounds belonging to 10 different molecular families in MeCN, MeOH, water, and Me_2SO [4]. With this method, pK_A values of small molecules in different solvents can be computed with an accuracy of about 0.7 pH units, albeit it requires only moderate computational effort.

(4) An empirical method is parameterized that converts known pK_A values in aqueous solution to the corresponding values in MeCN, MeOH, and Me_2SO with an accuracy of about 0.50 pH units. This method is very general and can be applied to organic compounds that belong to 20 different molecular families.

In summary, the methods developed in this thesis represent novel and robust approaches to compute the pK_A values of titratable molecules in protic and aprotic solvents with high accuracy.

Zusammenfassung

In der vorliegenden Doktorarbeit wurden vier verwandte Themen bearbeitet.

(1) Die ab-initio Methode zur Berechnung von pK_A-Werten wurde weiterentwickelt und verfeinert. Für eine Serie organischer Verbindungen wurden durch den Abgleich berechneter und experimenteller pK_A-Werte sehr genaue Werte für die Solvationsenergie von Protonen in Acetonitrile (MeCN), Methanol (MeOH), Wasser und Dimethyl Sulfoxid (Me_2SO) bestimmt. Diese haben die Werte -255.1, -265.9, -266.3 und -266.4 kcal/mol [1, 2].

(2) In Anlehnung an den Sauerstoff entwickelnden Komplex (OEC) in Photosystem II wurde die weiterentwickelte ab-initio Methode angewandt, um pKa-Werte von Di-oxo-mangan-Komplexen zu ermitteln. Mit einer Genauigkeit zwischen 1,19 und 3,69 pH-Einheiten rechtfertigen diese Ergebnisse die Verwendung der ab-initio Methode, um die Protonierung der -oxo Sauerstoffatome des OEC in Photosystem II zu bestimmen [3].

(3) Es wurde eine rein elektrostatische Methode entwickelt, um mithilfe bekannter pK_A-Werte in einem Lösungsmittel die jeweiligen Werte für andere Lösungen zu bestimmen. Die Methode wurde mithilfe von 30 organischen Verbindungen aus zehn verschiedenen Molekülfamilien getestet, für die insgesamt 77 pK_A – Werte für MeCN, MeOH, Wasser und Me_2SO verfügbar sind [3]. Dabei ergibt sich eine durchschnittliche Genauigkeit von 0,7 pH-Einheiten bei geringem Rechenaufwand.

(4) Es wurde eine empirische Methode entwickelt, um auf Basis der pK_A-Werte organischer Verbindungen aus 20 verschiedenen Molekülfamilien in Wasser die entsprechenden pK_A-Werte für die Lösungen MeCN, MeOH und Me_2SO zu ermitteln. Die Genauigkeit der Methode liegt bei 0,50 pH-Einheiten [4].

Mithilfe der in dieser Arbeit entwickelten methodischen Ansätze können pK_A-Werte titrierbarer Moleküle sowohl für protische, als auch aprotische Lösungen effizient und mit hoher Genauigkeit bestimmt werden.

List of publications

First author publications:

- E. Rossini and E. W. Knapp. "Proton solvation in protic and aprotic solvents". In: J. Comput. Chem. 37 (2016), pp. 1082–1091.
- E. Rossini and E. W. Knapp. "Erratum: Proton solvation in protic and aprotic solvents [J. Comput. Chem. 2015, 37, 1082-1091]". In: J. Comput. Chem. 37 (2016), pp. 1082–1091.
- E. Rossini, R. R Netz, and E. W. Knapp. "Computing pK_A values in different solvents by electrostatic transformation". In: J. Chem. Theory Comput. 12 (2016), pp. 3360–3369.

Additional publications:

- P. Chernev, I. Zaharieva, E. Rossini, A. Galstyan, H. Dau, and E.-W. Knapp. "Merging Structural Information from X-ray Crystallography, Quantum Chemistry, and EXAFS Spectra: The Oxygen-Evolving Complex in PSII". In: J. Phys. Chem. B 120 (2016), pp. 10899–10922.
- D. Huang, E. Rossini, S. Steiner and A. Caflisch. "Structured water molecules in the binding site of bromodomains can be displaced by cosolvent" In: ChemMedChem 9 (2014), pp. 573-579

Acknowledgements

First, I would like to thank my family: my parents, and my brother and grandparents for all the unconditional support they gave me since ever.

Next, I would like to acknowledge my supervisor Prof. Ernst-Walter Knapp for giving me the opportunity to work in his group where I've learnt something new each day. His advice helped me in all the time of research and writing of this thesis. It has been a joy and honor for me to be his student.

Thanks to doctors Galstyan Gegham and Artur for all the help they gave me in countless situations. I like to thank doctor Art Bochevarov for fruitful discussions and doctor Nadia Elghobashi-Meinhardt for her invaluable help in proofreading this thesis.

I gratefully acknowledge the financial support from the Sonderforschungsbereich, SFB 1078.

I further need to thank all my other colleagues for the great atmosphere in the lab and all my friends in Berlin, particularly Federico, Stefan and Matteo for all the fun we have had in the last years.

A big thank to all my friends in Lucca, particularly Fabio, Jonathan, Alessandro, Anthony and the two Simone because they have been like a second family to me.

Last and most important, I would like to thank Anne for the wonderful answer she gave me a few months ago: *"JA!"*

Contents

Chapter 1

Theoretical background

The theoretical background related to the techniques applied in this thesis is here presented. A punctual and rigorous introduction to the density functional theory applied in this thesis can be found consulting [5]. Therefore, the quantum mechanical background will be not discussed in this chapter. The issues tackled here are, the Poisson and Poisson-Boltzmann equations, the continuum electrostatics, and the determination of the atomic partial charges of molecules.
The environment of the considered molecular system plays a central role in chemistry influencing the solubility, the structural stability, and the reactivity of molecules. The solvation is accomplished generating into the bulk solvent a cavity hosting the solute compound. Such a physical phenomenon is thermodynamically allowed if it is energetically favorable. It may also be entropically favored if the solute-solvent mixture is less ordered than its alternative pure phases. Moreover, also the equilibrium constant of chemical reactions is affected by the molecular environment. The solvent may stabilize the structure of reactants and products differently thus pushing the equilibrium of reaction toward the more favorable substance. This is due to the non-covalent interactions, which are hydrogen bonds and van der Waals interactions between the solute and the solvent molecules. In addition to this, the rate of a chemical reaction is influenced by the solvent either through stabilizing the transition states or by friction, density and viscosity effects. For all these reasons, it is clear that the molecular environment needs to be carefully described to characterize the properties of a chemical compound in solution faithfully.
The process of solvation can be modeled by using two different approaches, the discrete and the continuum solvation methods.
(1) The discrete solvation is based on explicit solvent molecules placed to surround a compound of interest, modeling the specific solute-solvent interactions. Such an atomistic approach describes very detailed the physics of a compound in solution. On the other hand, an accurate discrete model of a solvated system

requires a high number of explicit solvent molecules and averaging over many relevant conformations. The description of such a system at high level of theory is very demanding and may comport computational cost which are not affordable. Therefore, a compromise about the applied level of theory needs to be accepted in order to enable the fully atomistic strategy. The solution to this problem is offered by multi-layered approaches that combine two different levels of theory to depict the system of interest and model its environment. In this hybrid approach, high level of theory is used to describe the solute while a lower one covers all the solvent molecules. An example of such strategy is given by the so-called QM/MM methodologies, where quantum and molecular mechanics theories are exploited as high and low levels, respectively.
(2) The second type of solvation models describes the influence of the solvent on the solute by continuum methods. These approaches are based on the atomistic description of the solute compound only, whereas the environment is represented as a dielectric medium. Although limited by the omission of all the possible explicit solute-solvent interactions, such solvation models are extensively used by the community due to their low computational cost and efficiency. In this doctoral work the effect of the solvent on organic and organometallic compounds is studied through the application of a continuum electrostatic method. In order to properly present such an approach, the theoretical background of this technique is introduced in the following.

1.1 Poisson-Boltzmann equation

1.1.1 Derivation

The Poisson-Boltzmann equation (PBE) plays an important role in many different fields of science, particularly in physical chemistry. The key equation of electrostatics, the PBE is a very useful tool when describing processes of molecular solvation. Moreover, it can be truncated to the simpler Poisson equation when the ionic strength of the medium vanishes. The electrostatic potential at position r in a homogeneous dielectric continuum with dielectric constant ε generated by a point charge q_1 at position r_1 is given by

$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0\varepsilon}\frac{q_1}{|\mathbf{r} - \mathbf{r_1}|} \tag{1.1}$$

Here, the $\mathbf{r}$ and $\mathbf{r_1}$, shown also in Figure 1.1, are the two vectors that locate the point P and the charge q_1, respectively.

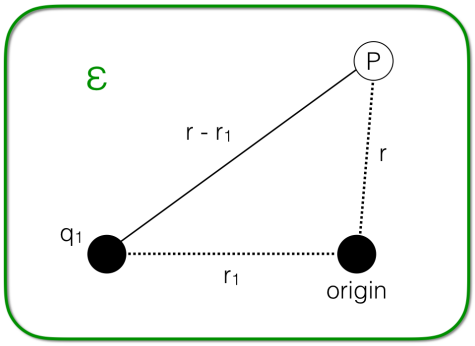

Figure 1.1: Model representing a single point charge placed in the homogeneous dielectric continuum with dielectric constant ε.

The equation 1.1 can also be converted to atomic units, where the factor $1/4\pi\varepsilon_0$ is replaced by unity. For a set N point charges q_i the electrostatic potential in r reads

$$\phi\left(\mathbf{r}\right) = \sum_{i=1}^{N} \frac{q_i}{\varepsilon\left|\mathbf{r}-\mathbf{r_i}\right|} \tag{1.2}$$

To describe the complex scenario of molecular systems further steps of generalization need to be taken. If the point charge is replaced by a charge distribution the electrostatic potential will be defined by the charge density ρ as:

$$\phi\left(\mathbf{r}\right) = \int_V \frac{\rho\left(\mathbf{r_i}\right)\mathbf{d^3r_i}}{\varepsilon\left|\mathbf{r}-\mathbf{r_i}\right|} \tag{1.3}$$

Considering that the Green's function for the three-dimensional Laplacian operator ∇^2 corresponds to $\left(-4\pi\left|\mathbf{r}-\mathbf{r_i}\right|\right)^{-1}$, its application to both the right and left sides of equation 1.2 leads to the Poisson equation. This expression is an essential equation of electrostatics and is defined as

$$\nabla^2\phi\left(\mathbf{r}\right) = \sum_{i=1}^{N} -\frac{4\pi q_i}{\varepsilon}\delta(\mathbf{r}-\mathbf{r_i}) \tag{1.4}$$

where $\delta(\mathbf{r})$ is the Dirac delta function. In order to obtain a more convenient form of the Poisson equation the expression eq. 1.4 can be combined to equation 1.5

$$\rho\left(\mathbf{r}\right) = \sum_{\mathbf{i=1}}^{\mathbf{N}} q_i\delta(\mathbf{r}-\mathbf{r_i}) \tag{1.5}$$

Now, the Poisson equation can be re-written as

$$\nabla^2\phi\left(\mathbf{r}\right) = -\frac{4\pi\rho(\mathbf{r})}{\varepsilon} \tag{1.6}$$

Only the embedment of a set of point charges in a homogeneous dielectric has been taken into consideration so far. Such case represents still a mere theoretical assumption; in nature the molecule is in fact surrounded by an inhomogeneous medium. For this reason, the environment ε is not anymore constant but becomes dependent on the position r as in Figure 1.2. The Poisson equation adopts now a more general form

$$\nabla \cdot \left[\varepsilon\left(\mathbf{r}\right)\nabla\phi\left(\mathbf{r}\right)\right] = -4\pi\rho\left(\mathbf{r}\right) \tag{1.7}$$

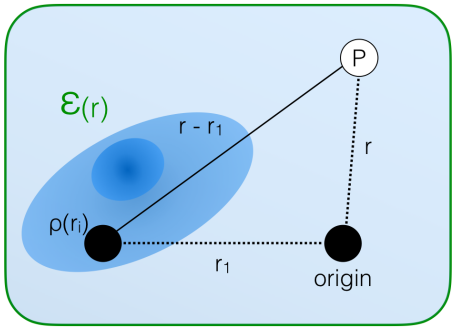

Figure 1.2: The electrostatic potential determined by a charge distribution in the inhomogeneous dielectric continuum.

In order to meticulously describe the environment of a solution the concept of implicit solvent model must be now introduced. As extension of the Deybe-Hückel theory [6] the implicit solvent model contains three regions. An internal region X, populated by immobile point charges, is assumed to represent the solute. Such region is surrounded by a second one Y, called "ion-exclusion layer". Finally, the model is completed by a third and more external region Z representing the actual solvent as a dielectric continuum that may contain also ions. The electrostatic potential in X ϕ_X is described by the charge density $\rho_X(\mathbf{r})$ and the ε_X dielectric constant through the Poisson equation. In Y and Z the electrostatic potential is determined knowing the charge density of the ions and the dielectric constant $\varepsilon\left(\mathbf{r}\right)$ of the medium.
According to the assumption of the Deybe-Hückel theory, the solvated mobile ions distribution follows the Boltzmann distribution law. For this reason, the ratio of

the ion concentration near the solute molecule (c_s^r) to the one in the remote Z (c_s) is given by

$$\frac{c_s^r}{c_s} = e^{-\frac{W_i(\mathbf{r})}{\mathrm{kT}}} \tag{1.8}$$

here k is the Boltzmann constant and T the absolute temperature. $W_i(\mathbf{r})$ is the work required to move an ion from infinity where ϕ(r)=0 to the position r. Such a model considers only a 1:1 electrolyte solution populated by ions $+q_s$ and -q_s. For this reason, the $W_+(\mathbf{r})$ and $W_-(\mathbf{r})$ will be equal to +/-$q_s\phi_Z(\mathbf{r})$, respectively. Therefore, the positive and negative ion concentrations at r are equal to $c^r_{+/-} = c_s e^{-/+\frac{q_s\phi_C(r)}{\mathrm{kT}}}$, with c_s the ion concentration at infinite distance from the region X where the electrostatic potential vanish to zero.

The ionic charge density at **r** in the Z region is calculated as,

$$\rho_C(\mathbf{r}) = \sum_{s=1}^{N_s} c_s q_s e^{-\frac{q_s\phi_Z(r)}{\mathrm{kT}}} \tag{1.9}$$

with $\mathrm{N_s}$ the number of different ionic types. Therefore, according to the 1:1 electrolyte composition, the electrostatic potential of Z is given as

$$\nabla^2\phi(\mathbf{r}) = -\frac{4\pi\rho_Z(\mathbf{r})}{\varepsilon_Z} = \tag{1.10}$$
$$\frac{4\pi}{\varepsilon_Z}\sum_{s=1}^{N_s} c_s q_s e^{-\frac{q_s\phi_Z(r)}{\mathrm{kT}}} = \left(\frac{8\pi c_s q_s}{\varepsilon_Z}\right)\sinh\left(\frac{q_s\phi_Z(r)}{\mathrm{kT}}\right)$$

When the ionic charge density is added to the Poisson equation, the Poisson-Boltzmann equation (PBE) is obtained. The PBE governs the electrostatic potential over the three considered regions X, Y, and Z.

$$\nabla\cdot[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] + 4\pi\sum_{s=1}^{N_s} c_s(\mathbf{r})\, q_s e^{-\frac{q_s\phi(\mathbf{r})}{\mathrm{kT}}} = -4\pi\rho(\mathbf{r}) \tag{1.11}$$

This equation is non-linear in the unknown ϕ(r). Such non-linear equations are difficult to solve. The PBE can be linearized for small ionic strength expanding the exponential term

$$\sum_{s=1}^{N_s} c_s\left(\mathbf{r}\right) q_s e^{-\frac{q_s\phi(\mathbf{r})}{\mathrm{kT}}} \approx \sum_{s=1}^{N_s} c_s\left(\mathbf{r}\right) q_s - \frac{1}{\mathrm{kT}} \sum_{s=1}^{N_s} c_s\left(\mathbf{r}\right) q_s^2 \phi\left(\mathbf{r}\right) \tag{1.12}$$

Due to the assumed electro-neutrality of the system the first summation term vanishes, while by using the ionic strength definition $I\left(r\right) = \frac{1}{2}\sum_{s=1}^{N_s} c_s(r) q_s^2$, the second term can be written more sophisticated, yielding the final form of the linearized Poisson-Boltzmann equation:

$$\nabla \cdot \left[\varepsilon\left(\mathbf{r}\right) \nabla\phi\left(\mathbf{r}\right)\right] - \frac{8\pi}{\mathrm{kT}} I\left(\mathbf{r}\right) \phi\left(\mathbf{r}\right) = -4\pi\rho(\mathbf{r}) \tag{1.13}$$

The ion-screening constant or Deybe-Hückel term k_{D}, is defined as

$$k_D^2 = \frac{4\pi I}{\varepsilon \mathrm{kT}} \sum_{s=1}^{N_s} c_s q_s^2 = \frac{8\pi I}{\varepsilon \mathrm{kT}} = \frac{1}{l_D^2} \tag{1.14}$$

with l_{D} the Debye length. Therefore, considering that $\overline{k} = k_D\sqrt{\varepsilon}$, the linearized Poisson-Boltzmann becomes finally:

$$\nabla \cdot \left[\varepsilon\left(\mathbf{r}\right) \nabla\phi\left(\mathbf{r}\right)\right] - \overline{k}^{\mathbf{2}}\left(\mathbf{r}\right) \phi\left(\mathbf{r}\right) = -4\pi\rho(\mathbf{r}) \tag{1.15}$$

An analytical solution of the Poisson-Boltzmann equation exists only for idealized systems. For complex geometries and charge distributions the equation can only be solved numerically. Several different methodologies are available to efficaciously solve the Poisson-Boltzmann equation. Among these, frequently applied techniques are versions of the finite difference method, FDM. The basic idea behind this methodology is the discretization on grid elements of all the relevant physical quantities such as dielectric constant, electrostatic potential, ionic strength, and molecular charges. Such grid mapping strategy allows to replace the differential operators by grid value differences. Alternative approaches often applied to solve the Poisson-Boltzmann equation are boundary element BE and the finite element FE methods. While the BE methods are for example used in the PCM approach [7], the FE has been recently used also in our group to develop the software mFES (molecular Finite Element Solver) introduced in [8, 9].

1.1.2 Numerical solution: The finite difference method

As previously mentioned, a common way to numerically solve the Poisson-Boltzmann equation is represented by the finite difference FD method. This approach discretizes the solute geometry on a regular simple cubic grid with lattice constant h. The linearized Boltzmann equation is integrated over the grid volume yielding

$$\iiint \nabla\cdot\left[\varepsilon\left(\mathbf{r}\right)\nabla\phi\left(\mathbf{r}\right)\right]d^3\mathbf{r} - \iiint \overline{k}_{\mathbf{0}}^{\mathbf{2}}\left(\mathbf{r}\right)\phi\left(\mathbf{r}\right)d^3\mathbf{r} = -\iiint 4\pi\rho\left(\mathbf{r}\right)d^3\mathbf{r} \tag{1.16}$$

This form can be simplified converting the first term on the left hand side to a surface integral by applying the Gauss theorem. Additionally, the second term can be approximated to $\overline{k}_{\mathbf{0}}^{\mathbf{2}}\phi_0 h^3$ where the k_0 are the modified Debye-Hückel parameter, whereas ϕ is the electrostatic potential at the grid point, respectively. The usual Debye-Huückel parameter k is defined by:

$$k = \sqrt{\frac{2FI}{\varepsilon \mathrm{RT}}} \tag{1.17}$$

where F is the Faraday constant, I the ionic strength, and ε the permittivity of the medium. The modified parameter k_0 takes the values $k_0 = \sqrt{\varepsilon_{Y,Z}}k$ in the solution region and $k_0 = 0$ in the molecule region. It is therefore dielectric independent and proportional to the ionic strength of the solution. The right hand side integral of equation 1.16 is equal to $4\pi q_0$, where q is the total charge of the volume element. Consequently, the equation can be re-written as

$$\iint \varepsilon\left(\mathbf{r}\right)\nabla\phi\left(\mathbf{r}\right)\cdot d\mathbf{A} - \overline{k}_{\mathbf{0}}^{\mathbf{2}}\phi_0 h^3 = -4\pi q_0 \tag{1.18}$$

where d$\mathbf{A}$ is the surface normal vector of the cubic volume element. For each of the six sides of the cubic volume, the surface integral can then be solved separately

$$\iint \varepsilon\left(\mathbf{r}\right)\nabla\phi\cdot d\mathbf{A} = \sum_{\mathbf{i=1}}^{\mathbf{6}} \varepsilon_{\mathbf{i}}\left(\phi_i - \phi_0\right)h \tag{1.19}$$

here the ϕ_0 and ϕ_i are the two electrostatic potential values of the central grid point and the average over the four corner points of one of the six faces of the cube,

respectively. The dielectric constant ε_i in eq. 1.19 and below refers to the midpoint between the point i and the reference point 0. So that

$$\sum_{i=1}^{6} \varepsilon_i \left(\phi_i - \phi_0\right) h - \overline{k}_{\mathbf{0}}^{\mathbf{2}} \phi_0 h^3 = -4\pi q_0 \tag{1.20}$$

Such equation can be rearranged into equation 1.21 in order to calculate the ϕ_0, yielding

$$\phi_0 = \frac{\sum_{i=1}^{6} \varepsilon_i \phi_i + \frac{4\pi q_0}{h}}{\sum_{i=1}^{6} \varepsilon_i + \overline{k}^{\mathbf{2}} h^{\mathbf{2}}} \tag{1.21}$$

This is the final equation that is iteratively solved until the convergence criterion is met.

As introduced before, the electrostatic potential ϕ related to any central grid point and to its six neighbors has to be estimated. An important limitation of this technique concerns the calculation performed on points placed at the grid boundary. Such points have less than six neighbors and it would be impossible to determine all the ϕ_i values of equation 1.19. In order to solve this problem, the electrostatic potentials could be estimated on a grid that is much larger than the studied molecule. In this case the ϕ related to any edge point will then either be set to zero or more rigorously calculated with the Deybe-Hückel approximation

$$\phi_i = \sum \frac{q_i e^{-k r_{\mathrm{ij}}}}{\varepsilon r_{\mathrm{ij}}} \tag{1.22}$$

where r_{ij} is the distance separating the grid point i where the charge q_i is localized from the grid point j at the boundary. However, the computing time required to properly describe simple cubic grids with a sufficiently small grid constant that are large enough that at the boundary the Deybe-Hückel approximation can be applied represents a critical limitation. In order to overtake the problem, and finally be able to compute the electrostatic potential, a multi-focusing approach is usually applied. A first calculation on a large grid scheme at low resolution is performed to minimize the computational time. This is then followed by a series of steps where only smaller grid regions closer to the molecule are taken into consideration. The electrostatic potential at the boundary of the small grids can therefore be

taken from the interpolated values from the previous steps. While a many focusing step approach can be used, a two-step process is adequate for the small molecules considered in this thesis. Such process is shown in Figure 1.3.

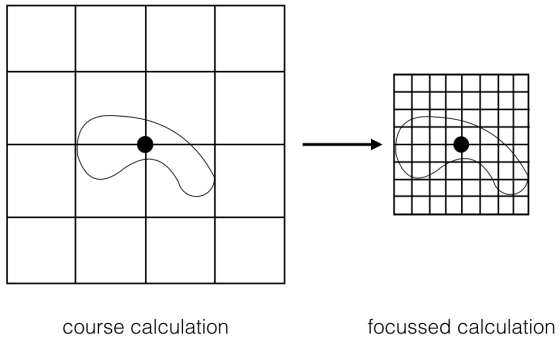

Figure 1.3: A two-dimensional depiction of the procedure of focusing is shown.

1.2 Continuum electrostatics

1.2.1 Characteristics of the reaction field

In the classical electrostatics approach, the compound is solvated by a dielectric polarizable medium. In this environment, the charges of the solute generate an electric field, which moves the solvent molecules (or ions) to develop a non-zero dipole moment. Consequently, such dipoles determine an induced external field acting against the one of the solute. The induced field is called reaction field and its strength is directly proportional to the magnitude of its induction. The reaction field is generated by two mechanisms; the electronic and the orientational polarizations. The first one is due to distortions of electronic wave function of molecules that are generated by the inducing field. The latter, known also as nuclear polarization, is relevant for all the molecules with a permanent dipole moments and results from their tendency to orient along the electric field lines. The electronic polarization alone would yield a dielectric constant of 1.5-2.5, while the nuclear polarization can have a much larger effect on the value of the dielectric constant, which for water at room temperature is as large as 80. A minor additional contribution to the reaction field is generated by intramolecular vibrations, since the atomic polarizability has a minimal influence on the dielectric constant (between 0.05 and 0.30).

1.2.2 Electrostatic energy

The electrostatic potential is used to compute the electrostatic energy that, in the trivial case of two point charges q_i and q_j is calculated as

$$G^{\mathrm{el}} = \frac{q_i q_j}{\varepsilon r_{\mathrm{ij}}} = q_i \phi_i \tag{1.23}$$

where $r_{ij} = |r_i - r_j|$ is the distance between the two point charges and ϕ_i is the electrostatic potential at position r_i generated by the charge q_j at position r_j.
This expression is valid only when describing a homogeneous dielectric medium and cannot be applied for molecules embedded in a solvent. In the latter case two regions are considered. The first one represents the solute and is characterized by a low dielectric constant. This region is then surrounded by a second one, characterized by a specific dielectric constant that represents the solvent. The total electrostatic free energy G^0 of such a system is given by the summation of different contributions: The Coulomb interactions between the charges in the system G^0_{C}, the reaction field energy caused by the interaction between charges and the polarizable solvent G^0_{R}, and the interaction between the charges and the induced ion distribution in the solvent G^0_{I}. The expression can be alternatively summarized into the equation 1.24:

$$G^0 = G^0_{\mathrm{ij}} + G^0_{\mathrm{ii}} \tag{1.24}$$

here, the total electrostatic energy is calculated as sum of G^0_{ij}, containing all the pairwise interaction energies and G^0_{ii} representing the self-interaction energies. The first term includes the Coulomb interaction energy of two point charges (i and j), and the interaction energies of the charge i with the reaction field and the ionic distribution induced by j. The second term collects all the energy terms of self-interaction and includes the Coulomb, the reaction field, and the ionic energy contributions. The Coulomb self-interaction is the electrostatic energy of a classical atomic point charge in its own electrostatic potential. This contribution, known as self-energy, is infinite for point charges. In practical applications this singularity results in a finite but large contribution when solving the lPBE on a finite grid where the atomic point charges are smeared over cubic grid elements. This is the so-called grid artifact whose contribution cancels in appropriately computed energy differences. Namely, the energy terms are then computed for systems with the same coordinates in both the vacuum and solvated phases. The second term of eq. 1.24 includes also the interaction energy of the i charge with its reaction field, and with the ionic distribution, both induced by the same solvated charges.

1.2.3 Solvation energy

The solvation energy is the free energy to transfer a molecule or an ion from the gas-phase to the solvent. The electrostatic solvation energy is computed according to equation 1.25

$$\Delta G^{\circ}_{\mathrm{sol}} = G^{\circ(\varepsilon_{\mathrm{ext}};\varepsilon_{\mathrm{mol}})} - G^{\circ(1;\varepsilon_{\mathrm{mol}})} \tag{1.25}$$

where $\varepsilon_{\mathrm{ext}}$ is the dielectric constant of the solvent and ε=1 is the corresponding gas-phase value. If the inter-atomic space in the solute $\varepsilon_{\mathrm{mol}}$ is set to the value in vacuum, then

$$\Delta G^{\circ}_{\mathrm{sol}} = G^{\mathrm{elect},(\varepsilon_{\mathrm{ext}};1)} - G^{\mathrm{elect},(1;1)} \tag{1.26}$$

However, the total solvation energy $\Delta G^{\circ,tot}_{\mathrm{sol}}$ does not only depend on the electrostatic contribution. Further factors contribute to the solvation energy. Hence, the total solvation energy reads

$$\Delta G^{\circ tot}_{\mathrm{sol}} = \Delta G^{\circ}_{\mathrm{sol}} + \Delta G^{\circ}_{\mathrm{vdW}} + \Delta G^{\circ}_{\mathrm{cav}} \tag{1.27}$$

where $\Delta G^{\circ}_{\mathrm{sol}}$ is the electrostatic solvation energy, and $\Delta G^{\circ}_{\mathrm{vdW}}$ the energy of van der Waals interactions between the molecules of the solvent and the solute. $\Delta G^{\circ}_{\mathrm{cav}}$ is the energy term required to form a cavity and host the molecule within the solvent, such positive energy term includes the change in entropy due to the reorganization of the solvent around the solute. This term strictly depends on the number of molecules populating the first shell of solvation and it is then proportional the solvent accessible surface area (SASA) of the solute. Also $\Delta G^{\circ}_{\mathrm{vdW}}$ depends on the number of solvent molecules present in the first shell of solvation, the vdW interaction energy decreases indeed with the distance as $1/r^6$. For this reasons, also the $\Delta G^{\circ}_{\mathrm{vdW}}$ shows a proportionality to the SASA. Therefore, the definition of the solvent accessible surface plays a critical role. The SASA is generated by rolling a solvent probe sphere over the vdW volume of the molecule of interest [10], as shown in Figure 1.4. Both contributions, $\Delta G^{\circ}_{\mathrm{vdW}}$ and $\Delta G^{\circ}_{\mathrm{cav}}$ follow the expression

$$\Delta G^{\circ}_{\mathrm{vdW}} + \Delta G^{\circ}_{\mathrm{cav}} = \alpha + \beta SAS \tag{1.28}$$

where the constants α and β can be taken from experiment and the SAS is the solvent accessible surface area. The energy contribution related to these energy terms is relevant particularly for neutral and non-polar solutes.

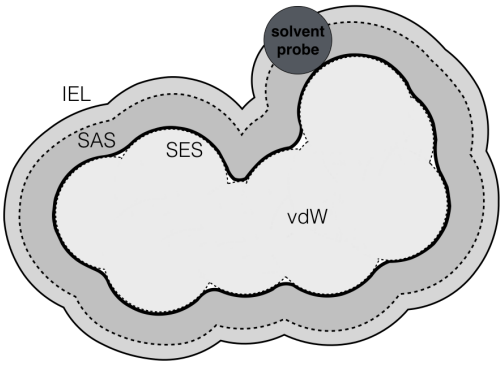

Figure 1.4: The solvent exluded surface SES (depicted as a bold line), solvent accessible surface SASA (the dashed line), and ion exclusion layer IEL are generated by rolling a solvent probe sphere over the vdW volume of the molecule.

1.3 Atomic partial charges derivation

The atomic partial charges of a molecule occur due to the negative charges of electronic wave function and the positive charges of the atomic nuclei. They are crucial parameters when computing electrostatic energies. The atomic partial charges are not directly measurable. Nevertheless, they can be estimated by quantum chemical computations or molecular properties related to the electron density distribution. There are various methods to obtain the atomic partial charges from measurements. They can be derived from spectroscopic data, population analysis of wave functions, partitioning the electron density distribution, electron density dependent properties and the most popular electrostatic potentials ESP. The molecular electrostatic potential is a quantity, which can be determined from the wave function of the molecule

$$\phi\left(\mathbf{r}\right) = \phi_n\left(\mathbf{r}\right) + \phi_e\left(\mathbf{r}\right) \tag{1.29}$$

where $\phi_n(\mathbf{r})$ and $\phi_e(\mathbf{r})$ are the nuclear and electronic electrostatic potential, respectively. These contributions are calculated according to equations 1.30 and 1.31, respectively.

$$\phi_n(\mathbf{r}) = \sum_{A=1}^{M} \frac{Z_A}{|\mathbf{r}-\mathbf{R_A}|} \tag{1.30}$$

$$\phi_e(\mathbf{r}) = -\int \frac{d\mathbf{r}'\rho_e(\mathbf{r})}{|\mathbf{r}'-\mathbf{r}|} \tag{1.31}$$

where Z_A is the nuclear charge, and $\rho_e(\mathbf{r})$ the electron density. A least-squares fitting procedure is then applied to derive a set of partial charges that reproduce the electrostatic potential in the neighborhood of the molecule. In such procedure, the equation 1.32 is minimized

$$\chi^2_{\mathrm{ESP}}(q_i, \lambda) = \sum_{i=1}^{M} w_i(\phi_i^0 - \phi_i^{\mathrm{model}})^2 + \lambda \left| Z - \sum_{i=0} q_i \right| \tag{1.32}$$

where ϕ_i^0 is the electrostatic potential at the i-th grid point, ϕ_i^{model} the electrostatic potential of the charge model, N is the number of grid points, and w_i the statistical weighting factor of the i-th grid point. Therefore, ϕ_i^{model} at the i grid point and due to all the q_j charges is calculated as

$$\phi_i^{\mathrm{model}} = \sum_{j=1}^{N} \frac{q_j}{|\mathbf{r_i}-\mathbf{r_j}|} \tag{1.33}$$

Although widely used to compute atomic partial charges, the normal ESP methodologies are not trivial. The generated charges are not simply transferable between common functional groups. Moreover, they strictly depend on the geometry of the molecular conformation. The reason for such a behavior is connected with the nature of the fitting procedure, which can be quite insensitive to the charges of all buried atoms.
A popular approach used to generate charges from the quantum chemically derived electrostatic potentials is RESP, the Restrained ElectroStatic Potential method. The RESP is based on the application of a penalty function to derive the charges. Although helping to minimize the problems of the normal ESP methods such

restrains implicate a price to pay. The technical limitation depends on a minor fitting quality of the matched electrostatic potential compared the one which is determined quantum chemically. In RESP the penalty function χ^2_{pen} is introduced as additional term to the χ^2_{ESP} in the charge fitting procedure

$$\chi^2 = \chi^2_{\text{ESP}} + \chi^2_{\text{pen}} \tag{1.34}$$

χ^2_{pen} controls the absolute value of the charges not to become to large and has the form

$$\chi^2_{\text{pen}} = a \sum_j (\left(q_j^2 + b^2\right)^{\frac{1}{2}} - b) \tag{1.35}$$

where the scale factor a defines the asymptotic limit of the strength of the constraint and b the narrowing of the hyperbola around the minimum. As smaller is b and as larger is a as more are the charges controlled to be small. Within the RESP approach is also possible to impose the symmetry of geometrically non-equivalent atoms that are equal though rapid exchange (e.g. hydrogen atoms in methyl groups). However, such constraints could lead to ESP reduction at the molecular polar regions. In order to solve this additional problem, a two stage approach has been suggested. The fitting procedure is first performed with weak restraints and no symmetry and in a second stage the fitting will be carried out only for the groups that need to be symmetrized. It has turned out that such a procedure yields atomic charges which are most suitable for the computation of solvation and conformational energies.

Chapter 2

Introduction

2.1 Acidic dissociation constant

The acidity dissociation constant measures the strength of an acid in solution and is usually reported as pK_A value. The two reactions of protonation AH $\overset{K_A}{\Longleftrightarrow}$ A^- + H^+ and B + H^+ $\overset{K_B}{\Longleftrightarrow}$ BH^+ have equilibrium constants expressed as

$$K_A = \frac{[A^-]\,[H^+]}{[AH]} \tag{2.1}$$

and

$$K_B = \frac{[B]\,[H^+]}{[BH^+]} \tag{2.2}$$

The pK_A value of an acid is defined as the negative decimal logarithm of the equilibrium constant K_A. The determination of this dimensionless quantity for chemical substances represents an important issue in chemistry and plays a relevant role in drug discovery. The pK_A value influences the solubility, the permeability, and the clearance of a drug molecule. About two-thirds of all approved pharmaceutical compounds contain ionizable groups, usually in the range of 2-12 pH units [11]. The states of protonation of the ionizable compounds are strictly associated to their pK_A values. Charged compounds are generally more soluble in water, whereas neutral molecules have a stronger lipophilic nature. In order to perform their function in the cell environment, drugs must penetrate the lipid bilayer cytoplasmic membrane. Such processes can be efficiently accomplished only by uncharged molecules, which for titratable molecules depends on the acidity constant. The

pK_A value of a compound also regulates its passive renal tubular re-absorption. Many drugs are either weak bases or acids. The urine pH can vary from 4.5 to 8.0. Consequently, in an acidic environment weak acid drugs are reabsorbed while in alkaline urine, weak bases tend to be re-internalized. Hence, the influence of the pH variation on the process of tubular re-absorption of a molecule can be effectively predicted by considering its pK_A value. For all these reasons, fast and accurate methodologies to compute the pK_A value of a molecule in protic environment as well as in apolar non-aqueous solvents can be very useful for the design of successful drug lead compounds.

2.1.1 Prediction of pK_A values in water

Many methodologies are available for computing or predicting the pK_A value of organic molecules solvated in water. These methods range from purely empirical to essentially ab-initio approaches, including mixed semi-empirical strategies. The empirical methodologies are usually based on machine learning in combination with molecular descriptors. Such descriptors are either dependent on structural information alone (e.g. QSAR/QSPR methods) [12, 13, 14], based on a combination of structure and semi-empirical quantum chemistry [15, 16, 17, 18, 19, 20], or alternatively related to other quantities as group-philicities [21]. The correlation between pK_A value and molecular properties is often restricted to specific families of compounds. Therefore, the application of such empirical schemes is practically limited to specific data sets of molecules. The ab-initio methods, on the other hand, represent the most elegant approach to the pK_A prediction as they are usually based on a combination of quantum chemistry (QC), molecular dynamics (MD), and electrostatics [7, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31]. In this spirit, the computation can be performed by using QM/MM methods. Such methods describe the region of the system in which the protonation process takes place at quantum mechanical level, while the remainder is treated with a classical molecular mechanics force field [32, 33, 34, 35, 36, 37]. Alternative pK_A prediction schemes based on discrete solvation models use only MD simulations and are very often applied to protein systems [38, 39, 40, 41, 42, 43, 44, 45]. Although explicit solvation enables a fully atomistic depiction of the system, many ab-inito methodologies are based on the implicit representation of the solvent environment. The influence of the solvent on the solute molecule can, for instance, be described by Langevin dipoles as in [46, 47, 48, 49]. Nevertheless, the great majority of implicit solvent models use a dielectric continuum to calculate the solvation energies of organic compounds in water, solving the Poisson equation [50, 51, 52] or using the generalized Born approximation and variations [53, 54, 55, 56, 57] . The third class of methodologies, the so-called mixed approaches, are essentially based

on semi-empirical schemes. In those predictive strategies, an approximate pK_A value is computed through a simplified ab-initio method and further adjusted via a linear regression function [58, 59, 60, 61, 62, 63, 64, 65]. An example of a mixed approach is given by the pK_A prediction module implemented in the software package Jaguar (Schrödinger, LLC)[66] and will be described in section 2.3.2.

2.1.2 Prediction of pK_A values in organic solvents

So far, the discussion has been focused only on the determination of pK_A values of organic compounds solvated in water. A different problem concerns the computation of these values in non-aqueous solvents such as acetonitrile (MeCN), dimethyl-sulfoxide (Me_2SO), and methanol (MeOH).
Solvent molecules equipped with polar hydrogen atoms are called protic solvents; water is the most important protic solvent. In contrast, solvent molecules that carry only non-polar hydrogens, such as MeCN and Me_2SO, are called aprotic solvents. Other solvents are characterized by a hybrid nature, such as methanol (CH_3OH). In this case, the solvent molecules carry not only the hydroxyl polar group but also three non-polar hydrogens attached to the carbon. Therefore, CH_3OH has a mixed nature placed on the gradient between the protic and aprotic extremes. As explained before, these non-aqueous solvents play a central role in pharmacology. The ability to predict pK_A values with high accuracy for both protic and aprotic solvents is then crucial. Nevertheless, only a few methodologies are currently able to predict the pK_A of non-aqueous solvents [1, 18, 56, 59, 67, 68, 69, 70, 71]. Furthermore, such methods, which are often semi-empirical, have been tested only to work on a limited set of molecular families. In 2002, Chipman applied an ab-initio approach combined with a linear regression scheme to compute pK_A values of a series of alcohols, carboxylic acids, and ammonium ions in water, acetonitrile, and dimethyl sulfoxide [59]. In 2004, Almerindo et alia performed an ab-initio prediction of pK_A values for a set of organic acids in dimethyl sulfoxide using the PCM method. But, they obtained maximum deviations higher than 4 pH units from experiment [68]. In a very similar spirit, Li et al. in 2006 [70] and Poliak in 2014 [72], respectively, performed ab-initio computations on amines and phosphitines in acetonitrile, and on derivatives of anilinium and pyridinium in seven different solvents obtaining overall accuracies within 1.1 pH units. In 2014, also the SM8 and SMD models of solvation were tested on the pK_A prediction of carboxylic acids, phenols, and amines in ethanol by Miguel et al. [69]. However, that approach gave results with a high level of accuracy only if finally a linear fitting procedure was applied. The SM solvation models will be introduced in section 2.3.1.3 . Moreover, Jover et al. applied a fully empirical approach QSPR

designed to predict the pK_A values of a series of phenols in many different solvents including water, methanol, and dimethyl sulfoxide [18]. Although several different approaches have already been proposed, an ab-initio method widely applicable in different non-aqueous solvents is nevertheless still needed.

2.2 Methodological approaches developed in this thesis

This section introduces:

- The ab-initio methodology update with the derivation of the free energies of proton solvation in different solvents.
- The application to the revised ab-initio method computing the pK_A values of a series of di-oxo-manganese clusters solvated in water and acetonitrile;
- The electrostatic transformation method applied for different protic and aprotic solvents;
- The empirical conversion scheme applied to shift the pK_A of organic molecules from water to other solvents.

2.2.1 Ab-initio method: Update and applications

The general scheme of the ab-initio pK_A computational methodology used in this thesis is presented here. In the ab-initio method, the computation of the pK_A values is performed combining quantum mechanics and electrostatics in order to exploit the thermodynamic cycle shown in Figure 2.1.

The pK_A value of a molecule can be derived from the free energy of proton dissociation, $\Delta\Delta G_{deprot}$. This energy cannot be directly computed in solution using the rigid rotor-harmonic oscillator approximation. However, for larger organic molecules the rigid rotor energy nearly cancels between the protonated and deprotonated species of the considered molecule, such that the vibrational energy remains to be evaluated. Hence, by taking advantage of the thermodynamic cycle, one can derive $\Delta\Delta G_{deprot}$ as

$$\Delta\Delta G_{\mathrm{deprot}} = \Delta G_{\mathrm{gas}} + \Delta G_{\mathrm{solv}}\left(A^{-}\right) + \Delta G_{\mathrm{solv}}\left(H^{+}\right) - \Delta G_{\mathrm{solv}}\left(\mathrm{AH}\right) \qquad (2.3)$$

where ΔG_{gas} and ΔG_{solv} are the gas-phase proton affinities and the solvation energies of the two molecular species (and the proton), respectively. The $\Delta\Delta G_{deprot}$

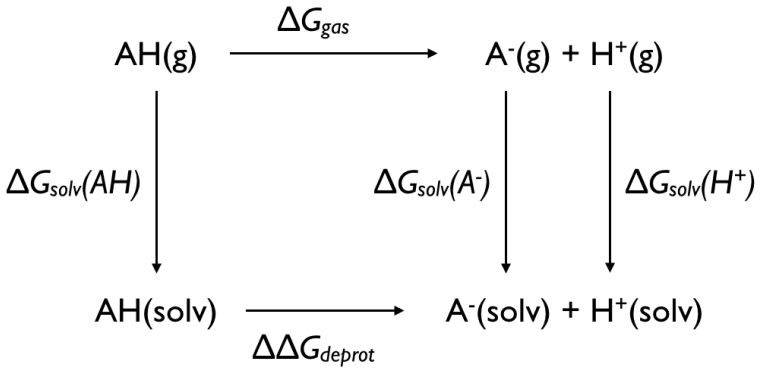

Figure 2.1: The thermodynamic cycle connecting the gas-phase and solvent-phase of proton dissociation is shown.

can be finally related to the macroscopic measurable pK_A value by the following thermodynamic relation:

$$pK_A = \frac{\Delta\Delta G_{\text{deprot}}}{(2.303RT)} \pm \log_{10}(n) \tag{2.4}$$

Here, the free energy term quantifies the microscopic pK_A value whereas the parameters R and T are the ideal gas constant and the absolute temperature, respectively. Equation 2.4 also takes into account of the number of multiple sites of protonation; the additive logarithmic term $\log_{10}(n)$ is added to cover the n equivalent possibilities to protonate the A^- molecular species, and subtracted if there are n equivalent chances to deprotonate the species AH.

The next two sections illustrate the theoretical framework applied in order to compute the gas-phase proton affinities and the solvation energies of the deprotonated and protonated molecular states as well as of the proton.

2.2.1.1 Gas-phase proton affinities and solvation energies

The gas-phase free energy computation represents the lowest source of error in pK_A predictions. Reliable ΔG_{gas} values within an accuracy of 1.1-1.6 kcal/mol are produced by level of theory such as CBS-QB3[73, 74] and CBS-APNO[75]. Moreover, CCSD(T) calculations on small molecules can return gas-phase free energies within the experimental accuracy[76]. Nevertheless, such computationally expensive methodologies cannot be applied for a large number of molecules. A combination of different methods, such as density functional theories (DFTs), model

chemistry, and ab-initio theories with different basis sets, have been applied to achieve high accuracy at low computational cost [77, 78, 79, 80].
In this thesis, the gas-phase reactivity is investigated through a quantum chemical approach at a DFT level of theory [1]. The vacuum contribution to the energy of proton dissociation is the so-called free energy of gas basicity, defined as

$$\Delta G_{\text{gas}} = G_{\text{gas}}(\text{AH}) + G_{\text{gas}}\left(H^+\right) - G_{\text{gas}}\left(A^-\right) \tag{2.5}$$

Here, the terms $G_{\text{gas}}(\text{AH})$ and $G_{\text{gas}}(A^-)$ are the free energies of the protonated and the deprotonated molecular species under vacuum conditions, respectively. These energy contributions are derived according to:

$$G_{\text{gas}} = E^0 + E_{\text{ZPVE}} + G^{298}_{\text{vib}} \tag{2.6}$$

where E^0 is the ground state electronic energy, E_{ZPVE} the zero-point vibrational energy and G^{298}_{vib} the thermal vibrational energy. The energy term $G_{\text{gas}}(H^+)$ is the free energy spent to create a mono-atomic particle, for instance the proton in the gas-phase. Assuming a mono-atomic ideal gas behavior of the proton, such free energy is then computed at room temperature according to equation 2.7

$$G_{\text{gas}}\left(H^+\right) = E_{\text{gas}}\left(H^+\right) - T_{298}S_{\text{gas}}\left(H^+\right) \tag{2.7}$$

The free energy term $G_{\text{gas}}(H^+)$ cannot be calculated using quantum mechanics because the proton does not contain electrons. The problem can be elegantly solved by combining the translational kinetic energy of the three degrees of freedom (1.5RT) and the excess mechanical energy required to generate the additional free proton from the acidic species (RT = pV). In this way the free enthalpy of a proton, $E_{\text{gas}}(H^+)$, is given as 5/2RT or 1.48 kcal/mol at room temperature. According to the Sackur-Tetrode equation [81, 82, 83, 84] the entropy contribution of a free atom in the gas-phase at room temperature is $T_{298}S_{\text{gas}}(H^+) = 7.76$ kcal/mol. Finally, the free energy to create a proton in the gas-phase is then equal to -6.28 kcal/mol. As outlined before, the zero-point vibrational energy E_{ZPVE}, and the thermal vibrational free energy G^{298}_{vib} are summed to calculate the free energy of a molecule under vacuum conditions. The zero-point vibrational energy E_{ZPVE} is calculated as in equation 2.8,

$$E_{\text{ZPVE}} = \frac{1}{2}\sum_i \text{hc}v_i = 1.429 \cdot 10^{-3}\sum_i v_i \tag{2.8}$$

Here, h is Planck's constant, c the speed of light and v_i denotes the frequency of the i-th normal mode in units of cm^{-1}.

The thermal vibrational energy in kcal/mol is calculated as

$$G_{\text{vib}}^{298} = H_{\text{vib}}^{298} - 298\text{K}S_{\text{vib}}^{298} \tag{2.9}$$

The H_{vib}^{298} and S_{vib}^{298} are the thermal vibrational enthalpy and entropy, respectively, at the absolute temperature T = 298K. The H_{vib}^{298} term is computed as

$$H_{\text{vib}}^{298} = \sum_i hcv_i \frac{1}{\exp\left(\frac{\Theta_i}{298K}\right) - 1} = 2.857 \cdot 10^{-3} \sum_i \frac{v_i}{\exp\left(4.826 \cdot 10^{-3} \cdot v_i\right) - 1} \tag{2.10}$$

where the vibrational temperature Θ_i of the i-th normal mode is:

$$\Theta_i = \frac{\text{hc}v_i}{k_b} = \ 1.438777 \cdot v_i \tag{2.11}$$

with k_b the Boltzmann constant and v_i the frequency of the i-th normal mode in units of cm^{-1}.
The thermal vibrational entropy is computed according to

$$\begin{aligned} S_{\text{vib}}^{298} = R \cdot \sum_i \Big(\frac{\frac{\Theta_i}{298K}}{\exp\left(\frac{\Theta_i}{298K}\right) - 1} - \ ln[1 - \exp\left(-\frac{\Theta_i}{298K}\right)]\Big) = \\ 1.986 \cdot 10^{-3} \sum_i \Big(\frac{4.826 \cdot 10^{-3} \cdot v_i}{\exp\left(4.826 \cdot 10^{-3} \cdot v_i\right) - 1} - \\ ln[1 - \exp\left(-4.826 \cdot 10^{-3} \cdot v_i\right)]\Big) \end{aligned} \tag{2.12}$$

Here, R is the molar gas constant equal to $1.986 \cdot 10^{-3}$ kcal/(mol K) [85].
In order to compute the $\Delta\Delta G_{\text{deprot}}$, eq. 2.3, besides the vacuum contribution, the free energies of solvation of both protonated and deprotonated molecular states, as well as of the proton, are also required. The two energy terms $\Delta G_{\text{solv}}\left(A^-\right)$ and $\Delta G_{\text{solv}}\left(\text{AH}\right)$ are computed by applying an electrostatic continuum methodology. The Restrained ElectroStatic Potential (RESP) procedure is applied to calculate the atomic partial charges in the gas-phase and the Poisson equation is solved by using the finite difference method as implemented in SOLVATE from the energy program suite MEAD (macroscopic electrostatics with atomic detail) [86, 87] (See sections 1.2.3 and 1.3 for theoretical background). On the other hand, the free energy of proton solvation $\Delta G_{\text{solv}}\left(H^+\right)$ represents a constant parameter in the thermodynamic cycle. Many investigations have proposed several different free energy values of proton solvation for the solvents investigated in this thesis. Various approaches are presented in the next section, including the strategy applied in this work.

2.2.1.2 Proton solvation energies in different solvents

One of the most relevant challenges in physical chemistry is the correct estimation of the absolute free energy of proton solvation, $\Delta G_{\text{solv}}(H^+)$. This task becomes particularly complicated for a proton solvated in water. In fact, in aqueous solution the particle is not attached to a single solvent molecule but dynamically solvated by a water cluster. Even less is known about the solvation of protons in non-aqueous solvents.
As mentioned in section 2.1.2, hydrogen atoms play a crucial role in solvation and the solvents can be divided into different families depending to the types of hydrogen atoms (polar and non-polar) that they carry. Usually solvents are divided into two families: protic and aprotic.
Proton solvation energies are difficult to measure directly and their determination is often based on the energetics of protonation reaction data combined with theoretical assumptions [88, 89, 90, 91]. For many years, the value of the free energy of proton solvation in water was rather uncertain, several different values were published, ranging between -252.6 and -271.7 kcal/mol [89, 90, 91, 92]. The major uncertainty in any determination of ion solvation free energies is that ions are never isolated from the surroundings. The calculation of an anionic solvation energy can be achieved only if the corresponding value for the cation is known. Therefore, in practice, all ionic solvation energy values are referenced against the one for the proton [27]. If the proton solvation free energy varies, the value for all the other ions will change as well. In 1998, Tissandier et al. [93] applied the cluster-pair approximation (CPA) to analyze the ion-pair solvation in small water clusters, this approach yielded a free energy of solvation equal to -264 kcal/mol. However, in 2005, a correction of -1.9 kcal/mol was applied to adjust for appropriate standard conditions and since then the resulting free energy of solvation of -265.9 kcal/mol has been accepted as the consensus value [94]. The CPA approach employs ion-water cluster data and bulk conventional ion hydration free energies (referenced to the proton) in order to deduce the proton hydration enthalpy and free energy. More recently, two new studies based on the application of the same CPA procedure resulted in the values of -266.1 (very close to the consensus) and 263.4±1.9 kcal/mol [88, 94, 95]. Finally, in 2015 Seybold published a slightly different energy equal to -265.6 kcal/mol [91].
The estimation of the free energy of proton solvation is required in order to match computed and measured pK_A values. As introduced in section 2.1.1, many of the methods to compute pK_A values are based on the application of quantum chemistry combined with the solvation energies of proton and molecular species. For this reason, the energy can also be derived from the computation of the acidic dissociation constants. Generally, the solvation process requires an implicit modeling of the solvent, using a dielectric continuum either alone or, alternatively,

adding one to three explicit water molecules in the ionic state [71, 92, 96, 97]. In 2004 the pK_A values for a selection of 26 organic compound were computed in water by applying the ab-initio method also used in this thesis [92]. The corresponding derived proton solvation energy was estimated to be -265.7 kcal/mol, which is very close to the accepted consensus value. In a more specific approach, the proton solvation energy has been directly computed solvating neutral and protonated water molecules embedded in an Eigen water-cluster geometry surrounded by a dielectric continuum [98]. This strategy yielded an extrapolated value of -262.4 kcal/mol [99]. The same approach yields -266.7 kcal/mol when more generally applied to up to 14 explicit water molecules [100]. Although the value -253.4 kcal/mol was more recently obtained by using the same water cluster [101]. It is therefore clear that even for water, the exact value of the proton solvation energy is subject to a high degree of uncertainty.

Table 2.1: Experimental and theoretical proton solvation energies are listed in kcal mol^{-1}. The table is adapted from [1, 4]. a: [102, 103] b: [38] c: [71] d: [104] e: [41] f: [105] g: [40] h: [106] i: [18] j: [43]

measured		
MeCN	Me_2SO	MeOH
-255.2[a]	-270.5[a], -268.6[b]	-263.8[a]
-260.2[c]	-273.3[c]	-263.5[c]
-254.2[d]	-268.55[d]	-261.9[d]
computed		
-252.3 [e]	-273.2 [f], -267.0 [g]	-263.4 [h]
-253.2 [i]	-267.6 [e], -261.1 [i]	-253.6 [j]

Proton solvation in water has been investigated for many years. However, there are only a few studies dedicated to other solvents. These investigations are summarized here in Table 2.1. In 2000 Kalidas et al. used the tetraphenylarsonium tetraphenylborate (TATB) assumption in order to obtain the transfer free energies of proton solvation from water to acetonitrile, methanol and dimethyl sulfoxide [102]. The TATB approach assumes the size of large and spherical ions such as tetraphenylarsonium and tetraphenylborate to be equal. Consequently, these two systems show very similar solute-solvent interactions and therefore practically equal solvation free energies [105, 107]. The assumption is widely applied in the formulation of absolute free energies of solvation of single-ions between different solvents. A few years later, Kelly et al. applied the CPA approximation to derive

the free energy of proton solvation for the same solvents, while Fawcett obtained the energy values by the measurement of electrochemical potentials [104]. Finally, other studies applied a combination of quantum chemical and electrostatic approaches to compute pK_A values and determine the solvation energies in the three solvents by matching with experimental measurements [105, 108, 109, 110, 111, 112].
In this thesis, the free energies of proton solvation in water, acetonitrile, methanol and dimethyl sulfoxide are determined by fitting measured and ab-initio computed pK_A values. The computational methodology used in this work is detailed in section 3.1 and the related publication is presented in section 4.1.

2.2.2 Investigation of di-oxo-manganese complexes

Multinuclear organometallic complexes play a central role in chemistry and biology [113] and their key feature is the ability to perform the electroneutral transition between different oxidation states. This transition is achieved by decreasing the free energy needed to reach high valence redox states through the coupling of electron and proton transfer reactions [114, 115, 116, 117]. The experimental characterization and theoretical description of these important processes remain challenging. Such metal ion clusters have been recently discussed by Bruijnincx and Sadler as anti-cancer agents, by Suh and Chei as artificial proteases, and by Barnham and Bush for their involvement in Alzheimer's and Parkinson's diseases [118, 119, 120, 121]. Organometallic compounds clearly have a relevant impact in biotechnological applications. Therefore, theoretical models to describe the redox and protonation patterns of such bio-inorganic complexes can be very useful in understanding their physical-chemistry. The most eminent organometallic compound in nature is the oxygen evolving complex (OEC) in photosystem II (PSII) of plants, algae, and cyanobacteria. The OEC, whose structure is shown in Figure 2.2 , is composed of four manganese atoms connected by five bridging oxygens and a calcium atom. Moreover, two pairs of water molecules are coordinated to the calcium ion and to the so-called Mn4 atom, respectively.

The OEC catalyzes the splitting of water during the photosynthesis in which electrons and protons are produced as an energy supply for the organism. In the process, the excess oxygen atoms are released to the atmosphere in form of molecular oxygen as a vast product. Many questions are still open and common to both the isolated and protein-included redox active manganese clusters. Among these issues, a very interesting one concerns the quantification of the pK_A values of the oxo-briges and terminal water relative to each other. However, an accurate investigation of the protonation processes performed in such a complex system is particularly difficult and requires a high level of quantum chemical theory. Several problems including spin state, antiferromagnetic coupling of the manganese

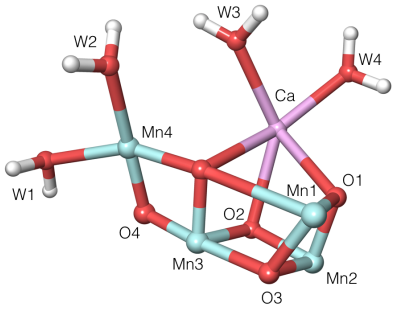

Figure 2.2: The OEC core (involving Ca, Mn1-Mn4, and O1-O5) is depicted, including the four waters ligated to manganese (Mn4) and calcium (Ca). Oxygen, and manganese atoms are denoted in red and blue, respectively. The single calcium atom is denoted in violet and the carbon and hydrogen atoms in light gray. The figure is adapted from [122].

atoms, the Jahn-Teller distortion of the oxygens bridging the manganese, and the charge delocalization and transfer between the different groups need to be taken into account. An additional challenge is represented by the ability of the cluster to switch between a combination of possible oxidation states. The four high-valent manganese ions can be found in the oxidation state III or IV. Taking into account that the five oxygen bridges can be either μ-hydroxo or μ-oxo, a total of $16X32 = 512$ microstates are possible.

Figure 2.3: The combination of the OEC microstates are here depicted taking into account the possible oxidation states of the manganese atoms and the protonation of the μ-oxo-oxygens.

Moreover, thousands of other potential microstates need to be considered when introducing the deprotonation of the four terminal water molecules. For all these reasons, the direct application of any predictive methodology to the full PSII manganese cluster needs to be tested on model systems first.
The ab-initio methodology has been in this context applied in order to predict the pK_A values of a series of model compounds inspired by the OEC. The basic structure of these models is shown in Figure 2.4 . They belong to a series of compounds

where the di-oxo-manganese core cluster is surrounded by an organic scaffold. The investigated series is referenced to as SOZMUP, structurally organized as [$Mn_2(\mu$-O)$_2$((3,5-di(R) or 5-R)-N,N'-bis(salicylidene)-1,3-propanediamine)$_2$], where R can be H, Cl, NO_2, or OCH_3.
Only di-oxo-manganese complexes (and their states) for which measured pK_A values are available were considered. The pK_A values of the investigated molecules were measured by Pecoraro, et al in acetonitrile (MeCN) by applying spectrophotometric and electrochemical techniques [123, 124].
The models considered in this thesis are simpler than the OEC. Nevertheless, they contain the main structural features to capture the transition between several different microstates. Consequently, the investigation of these compounds necessitates the application of methodologies that can properly consider all the oxidation and protonation states. The investigation performed in this thesis aims to predict the pK_A values of bridging oxygens in model compounds, and to test the ab-initio methodology recently used to titrate the -oxo oxygen atoms of the oxygen evolving complex in photosystem II [3].

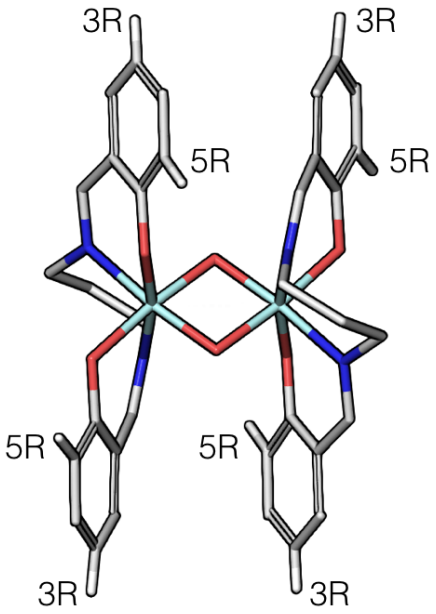

Figure 2.4: Model manganese complexes. Oxygen, nitrogen, manganese atoms are denoted in red, blue, violet, respectively. SOZMUP [$Mn_2(\mu$-O)$_2$((3,5-di(R) or 5-R)-N,N'-bis(salicylidene)-1,3-propanediamine)$_2$], where R can be H, Cl, NO_2, or OCH_3.

2.2.3 Electrostatic transformation method

The thermodynamic cycle of proton dissociation can be meticulously explored only through an ab-initio methodology describing both the vacuum and the solvated phases. As discussed in section 2.2.1, such an investigation is based on the calculation of the free energy of proton dissociation, which can be converted to the macroscopic pK_A value of a molecule. As introduced in section 2.1.2 the computation of pK_A values of molecules in non-aqueous solvents represents an important issue not yet fully understood. The ab-initio methodology has proven to be robust and precise in predicting pK_A values of organic compounds in different solvents [1]. A related but computationally much less expensive methodology to compute the pK_A values of organic molecules is the electrostatic transformation method [4]. According to this novel ab-initio approach, the pK_A value of a titratable residue is transformed from one solvent to another by computing the electrostatic solvation energies (ESE) of the protonated and deprotonated molecular species. The ESEs of organic molecules can be computed with a relatively high accuracy. However, the proton solvation energies in the corresponding solvents must be available in order to compute pK_A values.
The key equation that relates a molecule's pK_A value in one solvent (i) to the molecule's pK_A value in a different solvent (j) is

$$\left[pK_A\left(i\right) - pK_A\left(j\right)\right] k_B T \ln(10) \; = \Delta G^{\text{solv}}\left(i\right) + G_H\left(i\right) - \Delta G^{\text{solv}}\left(j\right) - G_H\left(j\right) \quad (2.13)$$

where $k_B T \ln(10)$ converts the dimensionless pK_A value to energies, the G_H are the free energies of proton solvation, and the $\Delta G^{\text{solv}}(i)$ are the free energies of solvation for the molecule in solvent i. The free energy terms can also be expressed as the difference between the protonated (BH) and deprotonated molecular species (B^-)

$$\Delta G^{\text{solv}} = G^{\text{solv}}\left(\text{BH}\right) - G^{\text{solv}}\left(B^-\right) \quad (2.14)$$

The application of equations 2.13 and 2.14 allows the transformation of any experimentally measured or theoretically estimated reference pK_A, $pK^{\text{ref}}_{A,k}(j)$, of a specific molecule k in a solvent j to the value, $pK^{\text{th}}_{A,k}(j \rightarrow i)$, of the same molecule in a second solvent i according to

$$pK^{\text{th}}_{A,k}\left(j \rightarrow i\right) = pK^{\text{ref}}_{A,k}\left(j\right) \quad (2.15)$$
$$+\left[k_B T \ln(10)\right]^{-1}\left[\Delta G^{\text{solv}}\left(i\right) + G_H\left(i\right) - \Delta G^{\text{solv}}\left(j\right) - G_H\left(j\right)\right]$$

In this thesis the solvation energies are computed electrostatically for compounds surrounded by a continuum dielectric medium. The electrostatic potential based RESP approach and the FDM (finite difference method)-SOLVATE package from MEAD (macroscopic electrostatics with atomic detail) [86, 87] are applied to compute atomic partial charges and to solve the Poisson equation, respectively. As introduced in section 1.2.2, a technical restriction limits the application of such approach to molecules with the same structural conformation in both the gas- and solvated phases. Moreover, the atomic partial charges are computed in vacuum, assuming no medium influence on their determination. For this reason the strategy applied in the electrostatic transformation method is valid only if the binding energy of a proton to a solute does not depend on the environment of the molecule. This assumption concerns a broad range of organic molecules belonging to diverse molecular families. In fact, as shown by the typical pK_A ab-initio method [1], the environment can even be absent (e.g. in gas-phase), and the proton binding energy to the solute is virtually the same as in a solvent [4]. On the other hand, the van der Waals (vdW) interaction between the solvent and the solute in both the protonation states is very similar because the two molecular species differ only by one proton. For this reason, the only interactions that affects the pK_A value of a molecule in different solvents are the electrostatic interactions and the proton solvation energies. Consequently, only the electrostatic solvation energies ΔG_k^{solv} of a titratable molecule k in the two solvents i and j are taken into consideration in equation 2.15. To obtain absolute solvation energies for individual molecular species vdW energies are necessary, however they are not needed for the pK_A computations.
The electrostatic transformation method presented in this doctoral thesis has been used to predict the pK_A values of 30 organic molecules solvated in water, methanol, acetonitrile and dimethyl sulfoxide. The proton solvation free energies derived in the study and introduced in section 4.1 are used here. The computational methodology applied in this work and the results are presented in section 4.2 and [4], respectively.

2.2.4 Empirical conversion method

Empirical approaches gather many popular methods to compute the pK_A values of molecules [12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 125] and they can be based on many different schemes as introduced in section 2.1.1. Moreover, it has been shown that combinations of parameters can be applied converting pK_A values between different solvents [126]. Usually, such a conversion is valid for a limited set of similar compounds and based on the use of a combination of multiplicative and additive parameters [126]. The key empirical function to convert a pK_A value of a

molecule from a reference solvent (*ref*) to a different one (X) is then

$$pK_A^{\mathrm{X}}\left(Y_{mol}\right) = f\left[pK_A^{\mathrm{ref}}\left(Y_{mol}\right), A\right] \tag{2.16}$$

where Y_{mol}, is the solute molecule that can be either in the protonated (YH) or deprotonated (Y) form, and A is the set of adjusted parameters applied to the $\mathrm{pK_A}$ value of reference in order to obtain the as result of the conversion scheme.
In this thesis it is suggested that a $\mathrm{pK_A}$ value of any molecule can be converted between different solvents by using a single additive empirical parameter. The empirical approach is then exploited to condense in a single parameter all the energy contributions to shift the $\mathrm{pK_A}$ value. Such parameters show a hybrid dependency based on the solvents and the molecular families of interest. The families of compounds considered here are selected according to the degree of homogeneity between the chemical functional groups of their components and trace the categorization scheme proposed in [58]. Any measured or computed $\mathrm{pK_A}$ value can then be converted from a reference solvent to any other one according to

$$pK_A^{\mathrm{X}}\left(Y_{mol}\right) = pK_A^{\mathrm{ref}}\left(Y_{mol}\right) + A^X_{(Y_{mol})} \tag{2.17}$$

where $A^X_{(Y_{mol})}$ is the empirical conversion parameter applied to shift the $pK_A^{\mathrm{ref}}\left(Y_{mol}\right)$ value of a molecule in the solvent of reference to the one in the new solvent $pK_A^{\mathrm{X}}\left(Y_{mol}\right)$.
In the present study the empirical conversion is performed from water to the other three solvents acetonitrile (MeCN), dimetyl sulfoxide ($\mathrm{Me_2SO}$), and methanol ($\mathrm{CH_3OH}$). The reference $\mathrm{pK_A}$ values taken into consideration to optimize the set of parameters are computed in water by using the predictions of Friesner et al. [58]. Alternatively, for the considered compounds not included in [58], the water references are obtained by using the $\mathrm{pK_A}$ prediction module as implemented in Jaguar 8.0 [127]. In both the cases, the values are provided in water by applying a semi-empirical scheme in which a set of empirical parameters adjust the so-called "raw" $\mathrm{pK_A}$ value obtained, thereby exploiting a version of the thermodynamic cycle introduced in section 2.2.1. The set of parameters are then applied to take into account any close range effect not adequately covered by both the solvation method and the thermodynamic cycle description [58]. The set of additive shifting parameters used in this project therefore strictly dependent on the predictive scheme of [58]. However, they can be easily corrected to convert any measured or differently computed water $\mathrm{pK_A}$ of reference.
The additive parameters are optimized by fitting the empirical results to benchmark

pK_A values in the three solvents. The pK_A values of many organic compounds are available as measured in pure Me_2SO [128]. In MeCN and CH_3OH, on the other hand, the measurements are often performed in aqueous mixtures. In these cases, the titratable atoms are often biased by polar water clusters, so the pK_A value in acetonitrile and methanol cannot be used as a benchmark. For this reason, in order to enrich the database of benchmarks obtaining a representative set of pK_A values in the three solvents the electrostatic transformation method introduced in section 2.2.3 is applied. The computational methodologies applied and the obtained results are detailed in sections 3.4 and 5.2, respectively.

2.3 Alternative methods and investigations

2.3.1 Solvation models

In this paragraph alternative approaches to model the solvation are introduced, although not explicitly used in this doctoral work. However, methodologies such as the Polarizable Continuum Model, PCM, the conductor-like screening model, COSMO, and the minnesota solvation model, SM6, are widely applied in the electrostatic community and then worth to be mentioned here.

2.3.1.1 Polarizable continuum model

The PCM describes the electrostatic potential generated by a charge distribution ρ_M representing the solute located in a cavity and surrounded by a dielectric medium with dielectric constant ε. The total electrostatic potential is defined as the sum of two different contributions shown in equation 2.18

$$V(x) = V_M(x) + V_\sigma(x) = \int_R^3 \frac{\varrho_M(y)}{|x-y|} dy + \int_\Sigma \frac{\sigma(s)}{|x-y|} \mathrm{ds} \tag{2.18}$$

Accordingly, the total electrostatic potential is the sum of the potential determined by the solute charge distribution ρ_M and by the cavity surface charge σ representing the reaction field. The first term is an integral over the volume of the solute, while the second term is an integral running over the solute surface. The $\sigma(s)$ is the charge density of polarization at the surface and is computed using the boundary element method subdividing the cavity surface in small K tesserae tiles with constant $\sigma(s)$. The solvent polarization is consequently described through point charges q_k placed on the surface of the cavity. The definition of the cavity is critical for a correct description of the reaction field. In contrast to the method used in this thesis, in

the PCM method the cavity is directly defined by the surfaces of the van der Waals spheres of the atoms of the molecule.

2.3.1.2 Conductor-like screening model

The conductor-like screening model COSMO includes the solute in a cavity and treats the solvent as a dielectric medium with dielectric constant ε. The cavity, approximated by segments, is here usually defined by the van der Waals spheres of the solute atoms where the radii are increased by an empirical factor (up to 120%). The method derives the polarization charges of the continuum from a scaled-conductor approximation. The charges of the solute are known from quantum chemistry and the charges at the surface segments are calculated and reduced by an empirical factor. The solute-solvent interaction energy is calculated from the charge distribution of the molecule and the determined surface charges as in the PCM model.

2.3.1.3 Minnesota solvation models

In the SMx semi-empirical solvation model series [129], the solvation energy G_S^* is computed as

$$\Delta G_S^* = \Delta E_{\text{elec}} + \Delta E_{\text{relax}} + \Delta G_{\text{conc}}^* + G_P + G_{\text{CDS}} \tag{2.19}$$

where ΔE_{elec} is the change in the internal electronic energy of the solute in moving from the gas-phase to the liquid phase at the same geometry, ΔE_{relax} is the change in the internal energy of the solute due to the geometry relaxation accompanying the solvation process, ΔG_{conc}^* is the change in concentration between the standard states of the two phases, G_P is the free energy of polarization associated with the solvation process, and G_{CDS} accounts for all the contributions to the solvation free energy related to the size of the first solvation shell of the solute. In this model, $\Delta E_{\text{elec}} + G_P$ is defined as ΔG_{EP}. The same concentrations are used in both the phases and all calculation are based on vacuum geometries, therefore ΔG_{conc}^* and ΔE_{relax} are equal to 0. The generalized Born approximation is applied to calculate the polarization contribution to the total free energy and the ΔG_{EP} is computed in a self-consistent molecular orbital calculation [130]. In SM6, the Charge Model 4 (CM4) [129] is used to obtain the partial atomic charges. Moreover, the solvent accessible surface (SASA) is used to compute the lacking G_{CDS} energy contribution to the solvation energy. The SASA is here generated by rolling a solvent probe

sphere of 0.4Å over the solute volume defined by a combination of Bondi and van der Waals radii [129].

2.3.2 Jaguar pK_A prediction module

The pK_A prediction module implemented in the software package Jaguar (Schrodinger, LLC) [66] is fully based on [58]. This semi-empirical method, proposed by Friesner et al. in 2001, is based on a self-consistent reaction field (SCRF) approach that is parameterized to match experimental data. The optimization of empirical parameters is performed on a training set of 17 different molecular families of small organic compounds [58]. The methodology is based on equation 2.20 and the thermodynamic cycle already introduced in this thesis in section 2.2.1.

$$pK_A = A\,(pK_A^{\mathrm{raw}}) + B \tag{2.20}$$

Here, A and B are the molecular family dependent empirical parameters applied on the preliminary pK_A^{raw} value obtained as follows. As for the ab-initio method introduced in this thesis, the method applicability is limited to rigid compounds geometrically optimized with DFT. The gas-phase free energy of deprotonation is computed as

$$\Delta G_g = H_g - T\Delta S = E^{A-} + E_{\mathrm{vib}}^{A-} + \frac{5}{2}RT - E^{\mathrm{AH}} - E_{\mathrm{vib}}^{\mathrm{AH}} - T\Delta S^{H+} \tag{2.21}$$

where E^{AH} and E^{A-} contributions are the ab-initio energies of protonated and deprotonated molecules in vacuum, and $E_{\mathrm{vib}}^{\mathrm{AH}}$ and E_{vib}^{A-} are their zero-point energies, respectively. $T\Delta S^{H+}$ is the H^+ entropic term, and $\frac{5}{2}$RT is the H^+ enthalpy term. The entropic terms of protonated and deprotonated species are assumed to cancel out. The gas-phase deprotonation energies are therefore computed with DFT and the B3LYP[131, 132] functional and the cc-pVTZ basis set [133]. However, contrary to the method introduced by this thesis, the zero point energies are not explicitly included in the pK_A computation but rather incorporated into the parameterization of each molecular family. The solvation of the compounds is performed using the SCRF-PBF method [134, 135] optimized for neutral molecules [58]. For the ionic species, the dielectric radii are adjusted to fit experimental pK_A values. Moreover, additional empirical corrections are made to include first shell hydrogen bonding corrections for the ionic groups [58]. The test of the methodology on 19 aromatic nitrogen heterocycles resulted in a total deviation of 0.5 pH units as compared to experimental data [58].

2.3.3 Semi-empirical prediction of di-oxo-manganese pK_A values

The pK_A values of a group of mono and di-manganese model compounds, including the compounds investigated in this thesis, were computed in 2013 by Amin, et al. using a continuum electrostatics semi-empirical approach [136]. In this investigation, the di-Mn complexes are modeled considering each metal ion, μ-oxygens, as well as each of the asymmetric organic ligands as individual fragments with integer charges. The software MCCE [137] was used to perform Monte Carlo sampling and obtain the Boltzmann distribution of the possible microstates as a function of either the solution electron potential E_h or the pH. In such an approach, the continuum electrostatic interactions are computed by solving the Poisson-Boltzmann equation using the software Delphi [138].
The microstate free energy (ΔG^x) is computed relative to the free energy of the isolated fragments in the dielectric medium of reference as

$$\Delta G^x = \sum_{i=1}^{M} \delta_{x,i} \left[2.3 m_i k_b T \left(pH - pK_{A,sol,i}\right) + n_i F (E_h - E_{m,sol,i})\right] \qquad (2.22)$$
$$+ \Delta\Delta G_{solv,i} + \sum_{j=i+1}^{M} [\Delta G_{\mathrm{ij}}]$$

here, M is the total number of states of all the fragments, $\delta_{x,i}$ is equal to 1 or 0 if the fragment-state is present in the microstate or not, respectively; T is the temperature in K; n_i is the number of gained electrons considering the most oxidized state as a reference; and F is the Faraday constant. The $pK_{A,sol,i}$ and $E_{m,sol,i}$ are the reference pK_A and E_m for the fragment I surrounded by a dielectric continuum. The $\Delta\Delta G_{\mathrm{solv}}$ is the energy required to move a fragment from the pure solvent into the complex (e.g. SOZMUP) in the same solvent. The ΔG_{ij} energy term collects the pairwise electrostatic and Lennard-Jones interaction between the fragments i and j in the microstate x of the cluster. In this semi-empirical approach, formal integer charges are considered for each fragment and kept fixed during the computation. Thus, depending on the oxidation state the Mn atoms carry formal charges of +2, +3, or +4, while the charge of the bridging oxygens in the deprotonated state is -2. In contrast, due to their interactions with the metal core, protonated μ-hydroxo groups are not assumed to carry the formal charge -1. The atomic charges are therefore further empirically adjusted to reproduce the experimental shift of 8.7 pH units between the two oxidation states MnIII-MnIII and MnIV-MnIII measured in water for a different series of manganese compounds [136]. These empirical charges are then used for all the μ-hydroxo ligands in MeCN. Moreover, the ligand charges

are assigned by a two-step DFT computation based on the B3LYP functional in combination with the LANL2DZ and 6-31G* basis sets. In the first step, the ESP charges are computed for the geometrically optimized system including the manganese atoms in the oxidation state MnIV. The final partial charges of the isolated organic ligand molecule are then obtained by setting an appropriate total net charge and performing a second single point computation. The coordinated Mn is represented by a frozen fractional ESP charge at the optimized position. The resulting atomic partial charges are used to parameterize the actual ligands. In this approach, the predictions are performed by shifting reference measured pK_A values using the free energy of a microstate computed according to equation 2.22. The reference pK_A value of 6.5 was taken for the first protonation SOZMUP with 3,5-R = H. The results obtained by Amin, et al. are summarized in table 2.2.

Table 2.2: The calculated and experimental pK_A values of the SOZMUP derived complexes for the different oxidation states, MnIV-MnIV and MnIII-MnIV are shown.

SOZMUP	exp	theory	dev
MnIV-MnIV R=H	13.4	12.0	-1.4
MnIII-MnIV R=H	24.5	21.3	-3.2
MnIV-MnIV R=Cl	10.8	11.2	0.4
MnIII-MnIV R=Cl	20.2	21.2	1.0
MnIV-MnIV R=NO_2	5.0	2.1	-2.9
MnIII-MnIV R= NO_2	13.3	12.7	-0.6

The results of Amin, et al. yield a pK_A-RMSD of about 2 pH units. However, such a method can be applied only when experimental pK_A reference values are available. Moreover, the quality of the reference pK_A value used directly affects the accuracy of all predictions. Finally, since the proton solvation free energy is not included in equation 2.22, the shift of the pK_A values from the reference pK_A value cannot be performed between different solvents.

2.4 Spin state of transition metal complexes

This section presents a brief introduction of ligand and crystal field theories. The electronic structure of an isolated transition metal ion is organized in five degenerate d-orbitals. The coordination of this ion to ligands results in a loss of degeneracy that induces energy splitting. The nature and magnitude of such a splitting depends

on several factors: (i) ligand strength, (ii) arrangement of the ligands around the ion, (iii) coordination number, (iv) nature of the metal ion, and (v) oxidation state of the ion.

Figure 2.5 shows the orbital splitting in the octahedral and tetrahedral ligand

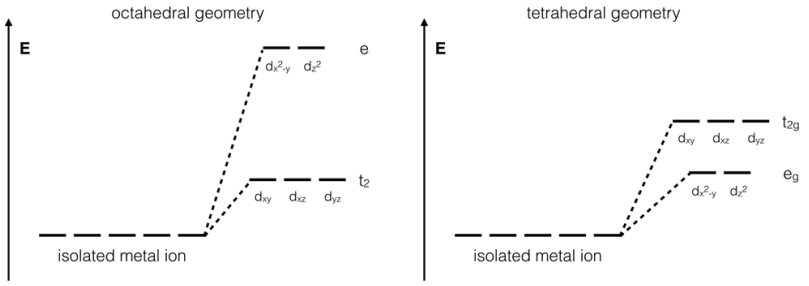

Figure 2.5: The energy splitting of the d-orbitals in transition metal complexes with identical ligands in octahedral and tetrahedral geometry are depicted.

fields. In the octahedral geometry, the orbitals e_g ($d_x{}^2{}_{-y}{}^2$ and $d_z{}^2$) are involved directly in the repulsive interaction with the ligands and their energy increases. In contrast, in tetrahedral geometries the coordination to ligands decreases the energy and the orbitals t_{2g} (d_{xy}, d_{xz}, and d_{yz}) are then shifted up with respect to the e_g orbitals. For the manganese ion in the oxidation states III and IV, the energy splitting for both the low- and high-spin electronic ground states of the d-orbitals are possible and shown in Figure 2.6.

In weak ligand fields, the ground state spin multiplicity is maximal and the

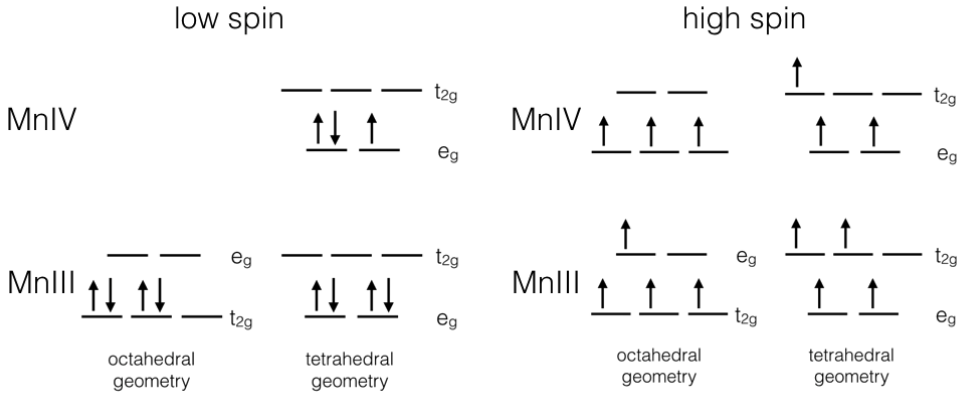

Figure 2.6: The electronic configuration of MnIV and MnIII in (left) the low- and (right) high-spin states are shown for the octahedral and tetrahedral geometries.

transition metal assumes a high-spin state. On the other hand, a strong ligand field stabilizes the transition metal in the low-spin state with a minimum multiplicity. Alternative configurations are determined by ligand fields of intermediate strength, where the lowest vibronic levels may be very close to each other. In such systems, minor perturbations could result in a change of spin state called spin transition (or spin-crossover), which can occur in both the solid and liquid states. The spin-crossover effect is very important for transition metal complexes in biological systems and can be induced by thermal energy, pressure, or light. Other relevant effects are ferromagnetism and antiferromagnetism. In nature, ferromagnets have a spontaneous magnetic moment determined by a regular arrangement of electron spins and magnetic moments. A ferromagnet is generated by internal interactions (called exchange field), which organize the magnetic moments parallel to each other. Conversely, in antiferromagnets the spins are ordered in an antiparallel arrangement as shown in Figure 2.7.

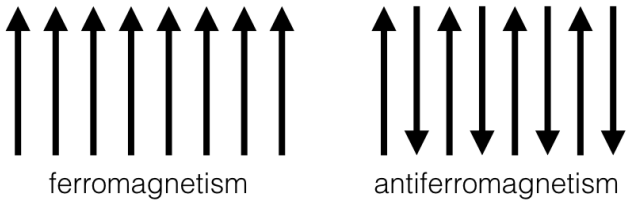

Figure 2.7: The spin ordering in ferromagnets and antiferromagnets.

Chapter 3

Computational methods

The computational methodologies applied in the four related topics of this thesis are presented here in four different sections. The first section details the standard techniques applied for the *ab-initio* computation of the pK_A values of organic molecules. The investigation aims to determine the free energies of proton solvation in water, methanol, acetonitrile, and dimethyl sulfoxide. The second section details the application of the *ab-initio* methodology in predicting the pK_A values of di-oxo-manganese clusters. The third and fourth sections explain the computational methodologies behind the two new approaches to calculating pK_A values: electrostatic transformation and empirical conversion.

3.1 Ab-initio compiutation of pK_A values

The computation of pK_A values can be performed through an *ab-initio* method based on the combination of quantum mechanical and electrostatic approaches. As introduced in section 2.2.1, the pK_A value of a titratable group is proportional to the free energy of proton dissociation computed by exploiting a thermodynamic cycle and summing up the free energies of gas basicity and solvation. This approach is valid for compounds with identical geometry and atomic partial charges in vacuum and solvent environment. The solute-solvent van der Waals interactions in both protonated and deprotonated molecular species differ by a negligible amount, therefore these energy contributions cancel in the computation of the energy $\Delta\Delta G_{\text{deprot}}$ in equation 2.3. Hence, as detailed in section 2.2.1, only the electrostatic contribution to the solvation energy is required to translate the proton affinity into the free energy of proton dissociation, except for the proton solvation energy. Such a contribution can be indirectly estimated by comparing computed

and measured pK$_A$ values. The solvation free energy of the proton (G_H) in water, methanol, acetonitrile and dimethyl sulfoxide is determined here by obtaining the best match between computed and measured pK$_A$ values for a reference set of 19 organic titratable compounds shown in figure 3.1

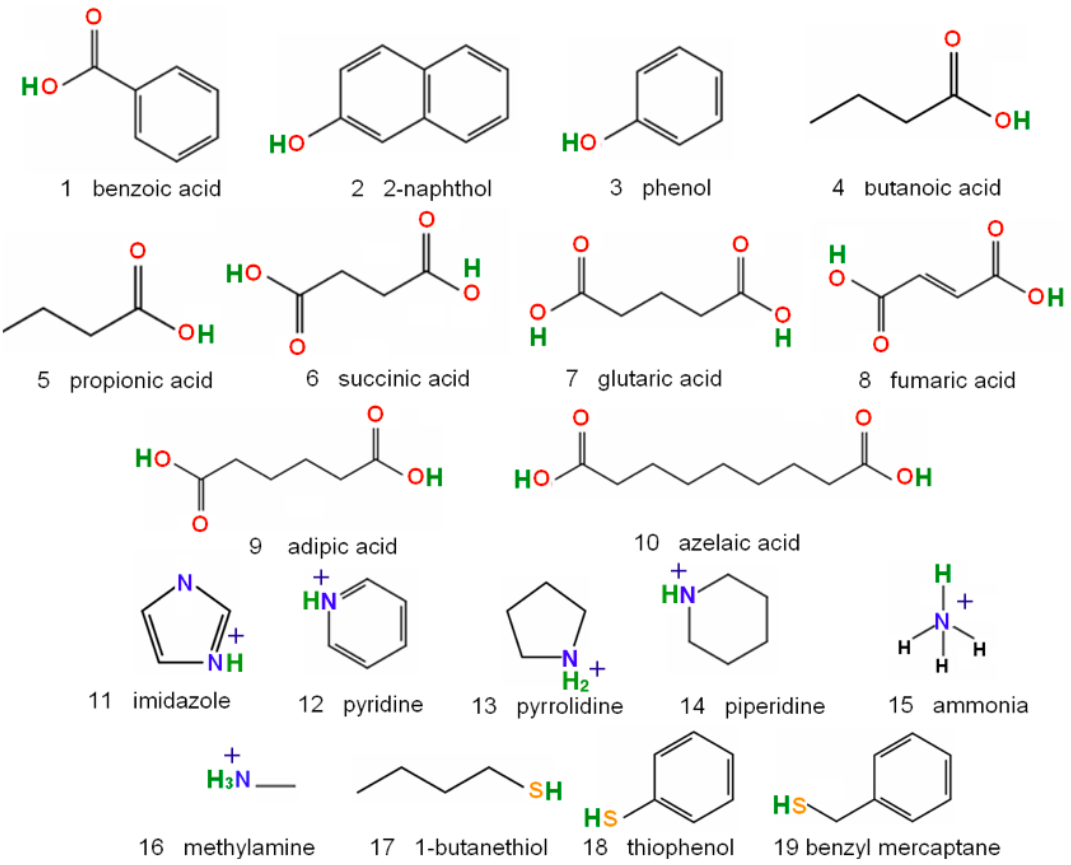

Figure 3.1: The 19 titratable molecules considered for pK$_A$ computations are displayed in the protonated state. The titratable oxygens, nitrogens, and sulfurs are highlighted in red, blue and orange, respectively.

The selection of these compounds is based on four specific criteria: (1) their pK$_A$ values are measured in more than two solvents; (2) they involve titratable oxygen, nitrogen, and sulfur atoms (3) they are small and rigid; (4) they can be flexible but with acidic groups on both chain ends to repel each other and then favor unique stretched conformations in gas and solvent phases. The publication related to this study is summarized in section 4.1

3.1.1 Quantum chemical computations and proton affinity

The gas-phase free energy (ΔG_{gas})) calculation, performed at the quantum chemical (QC) level of theory, should provide only a relatively small error source in computing pK_A values. Many different computational schemes are available to calculate this energy contribution. Among those, density functional theory (DFT) offers the best cost-efficiency approach [139]. The B3LYP functional [131, 132] and its derivatives are popular choices and their accuracy has been tested in numerous alternative studies. In 2003, Fu et al. have shown that the B3LYP/6-311++G(2df,p) method yielded gas-phase free energies within 2.3 kcal/mol of deviation from experimental values of organic acids reported in the NIST online database [140]. A few years later, Range et al. reported similar deviations applying the B3LYP functional to organic compounds [141]. Nevertheless, Schmidt am Busch et al., in 2004, obtained B3LYP average deviations in pK_A values of only 1.52 kcal/mol for a set of basic and acidic compounds, respectively [92]. Finally, in 2007 Bryantsev has published a much smaller absolute error of 0.78 kcal/mol by using B3LYP/6-311++G** [142]. The B3LYP results are therefore assumed to be sufficiently accurate. Therefore, this functional family has been used for all QC computations presented in this thesis. For geometry optimizations and calculation of the vibrational energies, the basis set 6-31G** of Jaguar 8.0 [127] was used. In contrast, the electronic ground state energies were computed with the quadruple zeta [cc-pVQZ(-g)] basis set. The electrostatic potentials (ESP) exploited to compute the atomic partial charges are computed for vacuum conditions using the 6-31G basis set of Jaguar 7.7 [143]. The ESP computations are performed in gas-phase in order to avoid the so-called electron leakage phenomenon that is observed when the electronic wave function is evaluated in the presence of a dielectric medium. This effect is particularly relevant for deprotonated states where the molecules carry excess negative charge. In nature, there is no dielectric medium, but the electron density of solute molecules is present that repels the solute electrons, resulting in an opposite effect. Hence, even under vacuum conditions such artifacts can occur. They are stronger if the applied basis-set is large. Therefore, in the present application the atomic partial charges were derived applying a relatively small basis set with respect to the one used for geometry optimizations. The accuracy of the quantum chemical computations was tested by comparing the energy differences between protonated and deprotonated molecular species in the gas-phase, as well as, the measured proton affinity values (see [1] for details).

3.1.2 Computing solvation energies

The computation of solvation energies of protonated and deprotonated molecular states were performed using an electrostatic approach which requires atomic partial charges. These are derived with the RESP (Restrained ElectroStatic Potential) fitting procedure detailed in section 2.2.1.1. This approach uses the electrostatic potential in the neighborhood of the considered molecule generated by the nuclear charges and the electronic wave function. By least squares fitting and application of restraints with a penalty function, the point charges at the nuclei are calculated to reproduce the molecular electrostatic potential ESP.
The computation of the solvation energies $\Delta G_{\text{solv}}\,(\text{AH})$ and $\Delta G_{\text{solv}}\,(A^-)$ is performed solving the Poisson equation with a finite difference method implemented in SOLVATE, from the energy program suite MEAD (macroscopic electrostatics with atomic detail) [86, 87]. The solvent description is operated by using a dielectric continuum with constants ε equal to 32.75, 46.7, 34.5 and 80 for methanol (MeOH), dimethyl sulfoxide (Me2SO), acetonitrile (MeCN), and water, respectively. Inside the volume of the solute molecules, where a detailed quantum chemical description is applied, the dielectric constant is set to unity. As outlined in section 2.2.1.1, the Poisson equations are solved using a two-step focusing scheme. Due to the size of the investigated molecule, cubic lattices of 189 points along the edge of the grid are used. The first coarse-grained simple cubic grid with a lattice constant of 0.4 Å is used to adopt the proper asymptotic of the ESP. The second high resolution grid with a lattice constant of $0.1A$ Å is embedded in the coarse-grained grid. Such a scheme has been tested to yield solvation energies within 0.1 pH units of accuracy [1]. The boundary between the solute and the solvent environment is defined by the solvent excluded molecular solute surface (SES). It is generated by rolling a specific solvent probe sphere over the volume, represented by the joint volumes of the atomic van der Waals spheres. The radii of the probe spheres used to generate the solute surfaces are 1.4, 2.05, 2.23, and 2.41 Å for water, MeOH, Me2SO, and MeCN, respectively. The solvent specific radii of the solute atoms also account for differences in solute solvent interactions. The hydrogen atoms of the solvent can approach the solute most closely. The electrostatic interactions of the polar solvent hydrogens with the solute are therefore dominant, whereas apolar hydrogens have a lower influence on the solvation energy. Such a difference is taken into account by the atomic radii used, smaller for protic and larger for aprotic solvents. Due to the nature of methanol, neither purely protic nor aprotic (see section 2.1.2 for more information), the corresponding atomic solute radii adopt values between the larger solute radii for Me_2SO and MeCN and the smaller radii for water. The hydrogen atoms are positively charged and are unlikely to come very close to each other; for this reason hydrogen solute radii are not changed between different solvents. In this thesis, a strategy of enhancement is used to enlarge the

radii from the reference water to the other solvents. The same approach has been used before, for instance in SM8 [144].
The values of the atomic radii used for water are close to the values used before to compute pK_A values [92], they are 1.4Å for nitrogen and oxygen atoms and 1.0Å for all the hydrogens, although radii of 1.2Å were applied in the previous study for all the non-titratable hydrogen atoms. For all carbon atoms 1.5Å is used. A larger radius of 2.0Å was used in [92] for aliphatic carbons, however its influence is extremely small due to the screening with the non-polar hydrogens and therefore not considered here. The radius of the sulfur atoms was optimized in the present application and is equal to 2.0Å. This value lies between the Coulomb radius of sulfur used for electrostatic energy computations in Jaguar (1.9Å) and the atomic radius used in SM8 (2.12Å). The radii of solute atoms used in the other solvents as well as their optimization methodology are detailed in section 4.1 and [1].

3.2 Investigation of di-oxo-manganese clusters

In this thesis, the *ab-initio* method has been applied in order to compute the pK_A values of two series of organometallic molecules inspired by the oxygen evolving complex of photosystem II. As introduced in section 2.2.2, these compounds are characterized by a di-oxo-manganese core and are solvated in acetonitrile (MeCN). The pK_A values are predicted by using the standard *ab-initio* strategy. The initial coordinates are defined according to the available 3,5-H-MnIV-MnIV-SOZMUP crystal structure from the Cambridge Crystallographic Data Center (CCDC) [145]. The structure optimization and the computation of vibrational energies are performed using the DFT with the B3LYP functional in combination with the Los Alamos National Laboratory LACVP** basis set (Hay and Wadt 1985a, Hay and Wadt 1985b) of Jaguar 8 [127]. The LACVP series of basis sets is a combination of the double-zeta 6-31G basis set with the LANL2DZ effective core basis set. Specifically, the atoms H - Ar are modeled by the 6-31G (or 6-31G** in the specific case) basis set while the heavier atoms are described using the LANL2DZ basis set. The electronic ground state energies are computed with the B3LYP* functional [146] in combination with 6-311G**++ and [cc-pVQZ(-g)], for the manganese ions and for all the other atoms, respectively. The B3LYP* functional differs from the original B3LYP regarding the weight of the exchange-correlation term, 0.15 for the first and 0.20 for the second functional. The functional B3LYP* shows a better energy quantification for atoms belonging to first row transition metals are present [146]. Moreover, the 6-311G**++ is the largest basis set available in Jaguar8.0 to describe the manganese atoms. The comparison of the performance of B3LYP and B3LYP* functionals applied here is shown in section 5.1. The electronic ground state energies

computations are performed at high-spin configurations for both the oxidation states of the di manganese complexes. However, the low-spin configuration of the atoms Mn1 (3/2) - Mn2 (3/2) is also investigated and obtained by including the effect of the antiferromagnetic coupling for the oxidized MnIV-MnIV complexes. The atomic partial charges are derived by using the standard RESP procedure (sections 1.3 and 3.1.2). However, these systems also include manganese (Mn) and chloride (Cl) atoms for which the atomic radii were not optimized. The van der Waals radius of manganese is equal to 1.48 Å, the value used for electrostatic energy computations in Jaguar. Because the manganese atoms are buried at the core of the organic scaffold, the radius does not change between different solvents. A radius equal to 1.50 Å was applied for the chloride atoms of the compounds solvated in MeCN, this value was optimized in [147]. In order to finally compute the pK_A values, also the free energy of proton solvation in MeCN are needed. In this thesis, a set of proton solvation energy values is applied and the details are shown in section 5.1 within the computational results.

3.3 Electrostatic transformation of pK_A values

The electrostatic transformation method shifts the pK_A value of a titratable residue from one solvent to another by computing the electrostatic solvation energies ESEs of the protonated and deprotonated molecular species. The effectiveness of this procedure is demonstrated by converting the pK_A values between different solvents for 30 compounds belonging to 10 different molecular families, as shown in figure 3.2. These compounds cover a wide range of organic molecules with oxygen, nitrogen, and sulfur as titratable atoms. Four different solvents are considered here: water, MeCN, Me2SO, and MeOH. However, the procedure can be applied to any other solvent for which the proton solvation energy is known.
The theoretical framework of this method is introduced in section 2.2.3 and the respective publication is summarized in section 4.2. All non-electrostatic contributions to the solvation energy are assumed to cancel in the difference of solvation energies of protonated and deprotonated species. For this reason, only the evaluation of the electrostatic contributions to the solvation energy need to be performed. All the molecular structures are optimized by energy minimization within the Jaguar 8.0 software using DFT with the B3LYP functional combined with the double-Z basis set 6-31G**. Atomic partial charges are required to evaluate the ESE and are determined by using the RESP procedure as detailed in section 2.2.1.1. The required electronic wave function is determined by QC computation in gas-phase using the methodology introduced in sections 3.1.1 and 3.1.2. The proton solvation energies that are essential to operate the transformation in the four considered

solvents correspond to the values obtained in the article [1] presented in section 4.2. The electrostatic transformation approach is valid for relatively rigid or small molecules. However, it can be generalized for flexible molecules if the appropriate structures are known.

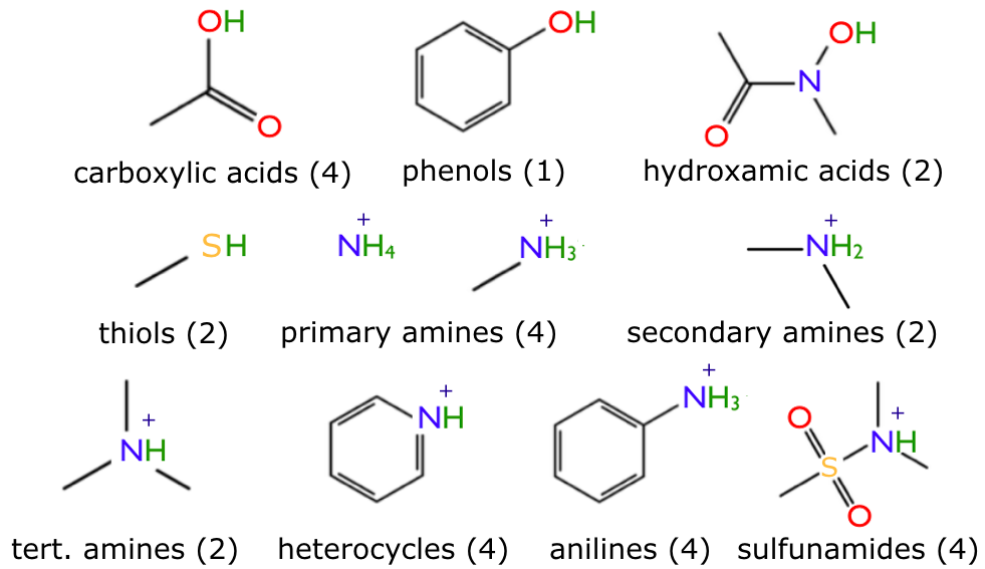

Figure 3.2: The ten different molecular families considered for pK$_A$ computations using the electrostatic transform method are shown. The families involve 30 different titratable molecules displayed in the protonated state. The number of considered compounds is given in brackets behind the family name. Titratable oxygens, nitrogens, and sulfurs are highlighted in red, blue, and yellow, respectively.

3.4 Empirical conversion of pK$_A$ values

In this section, the empirical conversion of pK$_A$ values from water to acetonitrile (MeCN), dimetyl sulfoxide (Me_2SO), and methanol (MeOH) is presented. As already introduced in section 2.2.4, the reference pK$_A$ values for water are either taken from Friesner et al. [58] or alternatively, obtained by using the pK$_A$ prediction module as implemented in Jaguar 8.0 [127]. The empirical conversion scheme is applied to 82 organic compounds belonging to twenty different molecular families shown in Figure 3.3.
The empirical parameters are optimized by fitting the predictions to benchmark pK$_A$ values in the three solvents. The pK$_A$ values of many organic compounds measured in pure Me_2SO have been determined [128]. Unfortunately, this is not

the case for non-dilute MeCN and MeOH. In these two solvents, the measurements are often performed in aqueous mixtures. Due to the bias by polar water clusters around the titratable atoms, the pK_A values of such compounds measured in acetonitrile and methanol cannot be used as a benchmark. For this reason, to enrich the database of benchmarks obtaining a similar amount of pK_A values in the solvents, the electrostatic transformation method [4] is applied. The accuracy of the electrostatic transformation method is sufficiently high and enables the application of the same set of empirical parameters to convert the pK_A values between the solvents.

The parameterization is performed over 18, 38, and 26 measured values, and 36, 30, and 35 computed values in MeCN, Me_2SO and MeOH, respectively. The empirical parameters for the twenty considered molecular families are shown in Table 3.1. For some families (alcohols, tetrazole, and indole and pyrroles) the empirical parameterization was not possible in the solvents MeCN and MeOH due to lack of measured data available for these solvents. As detailed in section 2.2.3, these molecules cannot be properly investigated by the electrostatic transformation approach due to technical limitations of the RESP procedure, so it is therefore not applied. The results and the set of benchmark data used are shown in section 5.2 and appendix (A1-20), respectively.

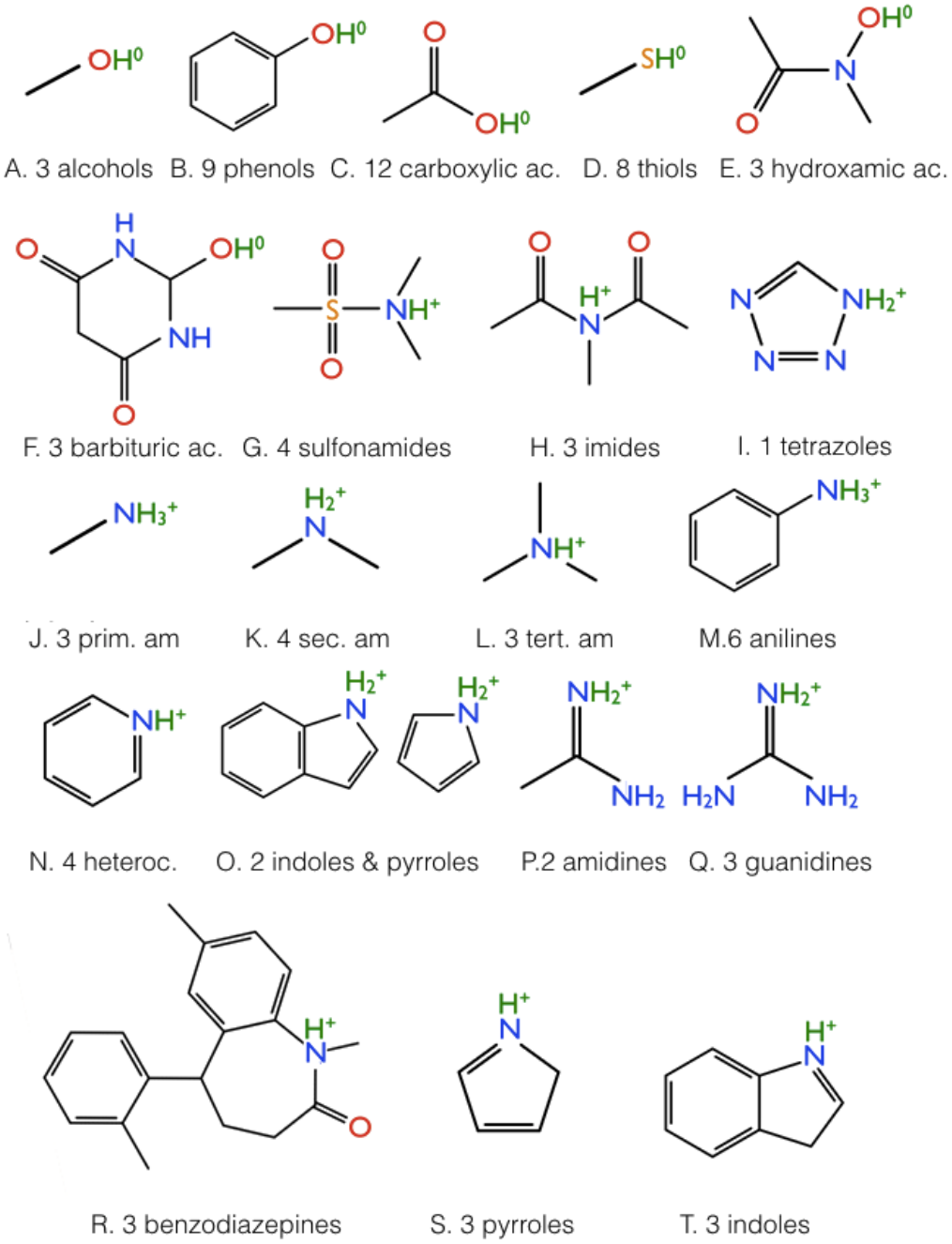

Figure 3.3: Twenty different molecular families involving a total of 82 different titratable molecules are displayed in the protonated state. The number of molecules of a specific family is given after the one-letter family name. Titratable oxygen, nitrogen and sulfur are highlighted in red, blue and yellow respectively. Titratable hydrogens are in green color and the corresponding total charge is highlighted with a + and 0 symbols.

Table 3.1: Average shifts of pK_A values (ΔpK_A) in MeCN, Me_2SO and MeOH relative to the values in water. The conversion is not applied for the alcohols, tetrazoles, and indoles and pirroles molecular families.

	family	ΔpK_A relative to water		
		ΔpK_A(MeCN)	ΔpK_A(Me_2SO)	ΔpK_A(MeOH)
A	alcohols		13.20	
B	phenols	16.30	7.90	4.20
C	carboxylic acids	15.40	7.00	5.00
D	thiols	12.70	3.70	2.70
E	sulfonamides	13.10	4.90	4.00
F	hydroxamic acids	15.00	6.60	5.00
G	imides	14.60	5.70	4.20
H	barbituric acids	15.50	6.50	4.70
I	tetrazoles		6.40	
J	primary amines	7.80	-0.50	0.30
K	secondary amines	7.90	0.20	0.00
L	tertiary amines	7.6	-0.4	0.00
M	anilines	6.60	-0.70	0.50
N	heterocycles	7.60	-1.00	0.10
O	indoles and pyrroles		5.80	
P	amidines	9.60	1.40	1.00
Q	guanidines	10.00	2.00	0.90
R	benzodiazepines	7.30	-0.60	-0.30
S	pyrroles (C-2 prot.)	6.60	-1.50	-0.30
T	indoles (C-3 prot.)	6.90	-0.90	-0.30

Chapter 4

Summary of publications

The two first author articles published within this thesis are summarized here. In the first section the article "Proton solvation in protic and aprotic solvents" [1], as well as a short erratum [2], are introduced along with important remarks about the van der Waals radii optimization performed in this study and used in all the other investigations. The second section is devoted to the presentation of the article "Computing pK_A values in different solvents by electrostatic transformation".

4.1 Proton solvation in protic and aprotic solvents

Proton concentration and transfer play a central role in both chemical and biological systems [148, 149, 150, 151, 152, 153]. To estimate the protonation equilibria in solution, the specific free energy of proton solvation needs to be known. However, the determination of this energy represents a challenging issue still debated in physical chemistry [1]. In this investigation, proton affinities, electrostatic energy of solvation, and pK_A values of organic molecules are computed in protic and aprotic solvents. Furthermore, the proton solvation energies in acetonitrile (MeCN), methanol (MeOH), water, and dimethyl sulfoxide (Me_2SO) are derived by matching computed to measured pK_A values for a set of 19 compounds (see figure 3.1 in section 3.1). The selection of these molecules is based on (i) the availability of experimental pK_A values in the different solvents, (ii) their structural rigidity, (iii) the wide range of pK_A values covered, and (iv) the involvement of titratable nitrogen, oxygen or sulfur atoms. As detailed in section 3.1, the pK_A values are computed by combining quantum chemistry and electrostatics in the ab-initio method. The computation of the proton affinity in vacuum is performed quantum chemically

with DFT, while the electrostatic solvation energy contributions are calculated by solving the Poisson equation. In this respect, the solute molecule is separated from the surrounding dielectric medium representing the dielectric boundary by the solute surface. As illustrated in section 1.2.3, this boundary surface is generated by rolling a solvent specific probe sphere all over the van der Waals volume of the solute. Except for sulfur, all the required atomic radii in water are taken from [92]. However, as these radii change in MeCN, MeOH, and Me_2SO, they are optimized to accurately predict the pK_A values. The van der Waals radii in water are 1.00, 1.50 Å, and 1.40 Å for hydrogen, carbon, and oxygen/nitrogen atoms, respectively, while the radius of sulfur optimized here is 2.0 Å. These solute radii are adjusted to non-aqueous conditions by multiplication with a solvent dependent enhancement factor α(solvent). All the optimized atom radii in non-aqueous solvents are listed in table 4.1 and are further applied in all the other investigations of this thesis. The positively charged hydrogen atoms repel each other. Hence, the mutual electrostatic interactions of those atoms do not contribute to the molecular solvation energy and the enhancement is not applied. On the other hand, the titratable atoms are particularly relevant and the magnitude of their multiplicative factors have a strong influence on the solvation energy. The optimization of the van der Waals radii is performed here by minimizing the pK_A-RMSD (root mean square deviation) in the different solvents. The enhancement factor used for the carbon, oxygen, and nitrogen atoms are 1.28, 1.12, and 1.25 in MeCN, MeOH, and Me_2SO, respectively. The multiplicative factor used to enlarge the sulfur atoms in Me_2SO is 1.09. Among all the titratable atoms, the radius of sulfur in water is noticeably the larger value. Consequently, its influence on the energy of solvation results is smaller than the one of oxygen and nitrogen and requires therefore a weaker enhancement. The van der Waals radius of sulfur was not optimized in MeCN, since no pK_A value of sulfur based compounds is computed in this solvent. A single sulfur titration is performed in MeOH. Since an optimization is not possible for such a solvent the value used in water was applied.

The dependencies of the RMSD in pK_A values are plotted in Figure 4.1 as a

Table 4.1: Optimized van der Waals atomic radii of solute atoms for acetonitrile (MeCN), methanol (MeOH), and dimethyl sulfoxide (Me_2SO).

Atom	MeCN	MeOH	Me_2SO
H	1.00	1.00	1.00
C	1.92	1.68	1.87
O/N	1.79	1.57	1.75
S	-	(2.00)	2.18
α	1.28	1.12	1.25

function of the enhancement factor α together with the resulting proton solvation free energy, which was used as an optimization parameter as well. Interestingly, the minimum in pK_A-RMSD values is flat for all the three non-aqueous solvents, whereas the proton solvation energies go through a plateau regime. Thus, it is possible to conclude that the proton solvation free energy value is robust and not sensitive to small variations of the enhancement factor. On the other hand, the computation of the pK_A values yields proton solvation energies with a high accuracy. The RMS deviations between calculated and available pK_A values are

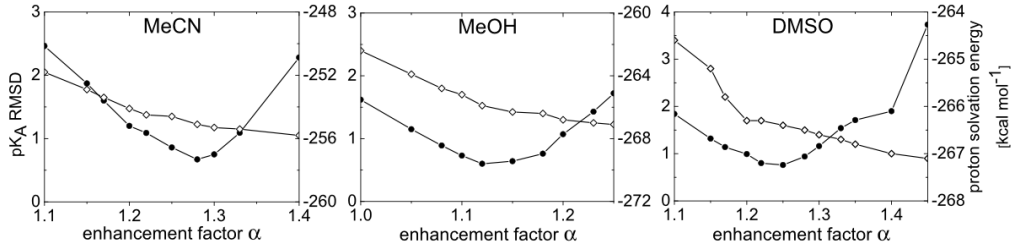

Figure 4.1: The pK_A-RMSD values ($\bullet$, left scale) and proton solvation energies ($\diamond$, right scale) for the solvents MeCN, MeOH, and Me_2SO are plotted as a function of the multiplicative enhancement factor α. The optimal factors α are 1.28, 1.12, 1.25 for MeCN, MeOH, and Me_2SO, respectively. The corresponding pK_A-RMSD values are 0.67, 0.60, 0.76 for MeCN, MeOH, and Me_2SO, respectively.

0.67, 0.65, 0.47, and 0.75 for MeCN, MeOH, water, and Me_2SO, considering 7, 11, 15, and 9 different titratable compounds, respectively, and excluding systematic errors (see [1] for more details). The proton solvation free energies derived here are -255.1, -265.9, -266.3, and -264.4 kcal/mol in MeCN, MeOH, water, and Me_2SO, respectively. The energy term in water is very close to the consensus value of -265.9 kcal/mol [93]. Additionally, the measured pK_A values of hydrated Me_2SO-H, hydronium ion, and MeOH-H are -1.04, -1.74, and -2.5 [154, 155, 156], respectively, whereas the one of hydrated MeCN-H is about -10 [157]. These values correlate qualitatively well with the proton solvation energies in the corresponding solvents. Hence, the lower the pK_A value of a solvent molecule in water, the easier it is to solvate the proton in that specific solvent. Such a correlation implies that the solvated proton is closely attached to a single solvent molecule. Consequently, the extent of proton solvation in MeOH and Me_2SO is similar, while in MeCN proton solvation is less negative. The correlation is also valid for water, in which the proton assumes an Eigen-cluster conformation [98]. In fact, the computation of the proton solvation energy in water performed by solvation of the Eigen cluster in a continuum dielectric medium yields -262.4 kcal/mol [99], a value close to the consensus value of -265.9 kcal/mol. The pK_A values computed in water for protonated Me_2SO, MeOH, and MeCN are -2.33, -4.96, and -12.81, respectively, in

good agreement with the measured values. Notably, the proton solvation energies in Me_2SO and water are similar. These two solvents have an opposite nature, the first is aprotic and the second is protic. Consequently, a consistent difference in solvation energies should be expected. However, such behavior is most probably due the sulfur atom, which is more polarizable than oxygen and nitrogen, the polar atoms of water and MeCN, respectively. As a result, sulfur can bind an excess proton stronger than can nitrogen in MeCN and so shift the energy of proton solvation toward the more negative value close to water.
The proton solvation energies computed here are suitable to improve the ab-initio prediction of pK_A values in different solvents. Moreover, these values are fundamental parameters of the electrostatic transformation and the empirical conversion methodologies designed in this thesis.

4.1.1 Erratum

In the original article a dielectric constant ε equal to 36.7 was erroneously used to represent the solvent Me_2SO. Here, the proper constant of 46.7 is applied and all the related computations are repeated [2]. Since relative energies are used to compute the pK_A value, the shift in dielectric constant $\Delta\varepsilon$=10 does not influence the calculated energy values. Although the total pK_A-RMSD in Me_2SO decreases here from 0.75 to 0.63 pH units, the large difference in dielectric environment does not affect the proton solvation free energy value, which is confirmed to be equal to -266.4 kcal/mol as in the original publication.

4.2 Computing pK_A values in different solvents by electrostatic transformation

The electrostatic transformation method requires moderate CPU power to titrate small molecules in different solvents with an accuracy of 0.7 pH units [4]. Taken optimized geometry, the electrostatic transformation requires only a few minutes to perform the predictions using a single CPU. A pK_A value known for a molecule in one solvent is transformed to the value in any other solvent if the proton solvation energy of both solvent environments is known. The electrostatic solvation energies of protonated and deprotonated molecular species are computed in the two environments of interest as detailed in section 3.3. Here, the methodology has been used to titrate 30 organic molecules belonging to 10 different molecular families in water, acetonitrile (MeCN), dimethyl sulfoxide (Me_2SO), methanol (meOH). A total of 77 pK_A measured values are considered in the four different solvents. The

methodology shows a high level of accuracy with pK_A-RMSD values obtained by matching measured with computed values are 0.64, 0.62, 0.66, and 0.58 in water, MeCN, Me_2SO, and MeOH, respectively. The methodology does not rely on any empirical factor and it can be applied to any other pairs of solvents if the respective proton solvation energies are known.

Chapter 5

Unpublished results and discussion

The unpublished results obtained in this thesis are presented in two sections. The first section is focused on the application of the ab-initio methodology to predict the pK_A values of two series of di-oxo-manganese organic compounds solvated in acetonitrile.
The second section is devoted to the parameterization of the empirical conversion methodology applied on 82 selected organic molecules belonging to 20 different molecular families. The empirical conversion method is designed in collaboration with the software company Schrödinger, LCC and aims to convert the pK_A values predicted by the Jaguar package in water [58] to the corresponding values in the three considered solvents of pharmaceutical interest: acetonitrile (MeCN), dimethyl sulfoxide (Me_2SO), and methanol (MeOH).

5.1 Protonation equilibria of di-oxo-manganese complexes

The ab-initio methodology is applied to investigate the protonation equilibria of a series of organometallic complexes, named SOZMPU [145]. As introduced in section 2.2.2 and shown in figure 5.1, such complexes are structurally characterized by a di-oxo-manganese central core analogue to the oxygen evolving complex (OEC) of Photosystem II. Here, the di-oxo-manganese cluster is coordinated by two 3,5-di(R)/5-R-bis(salicylidene)-1,3-propanediamine (salpn) residues with 3,5-R = H, Cl, or NO_2 and 5-R = Cl, or OCH_3. The ligand salpn has a net formal charge of -2 carried by the two oxygens. The compounds are unsable in

water and the measured pK_A values are available only in pure MeCN [123, 124].

The ab-initio computation of pK_A values requires high level quantum chemical

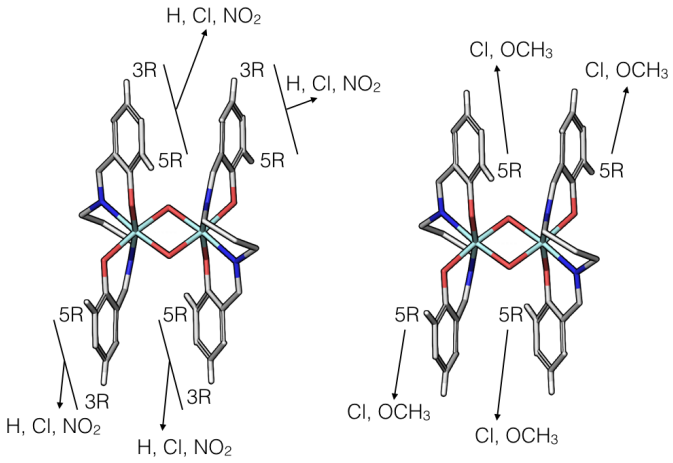

Figure 5.1: The SUZMUP series of di-oxo-manganese model compounds. The manganese, oxygen, and nitrogen atoms are colored in pale blue, red, and blue respectively. In SOZMPU are considered five derived molecules: 3,5R=H, Cl, or NO_2 and 5R=OCH3, or Cl.

and electrostatic methodologies as well as the free energy of proton solvation in the corresponding solvent. In a preliminary investigation in this thesis the ab-initio method [92] to perform computations for non-aqueous solvents was applied on the SUZMUP complexes. The first protonation of the μ-oxygen bridge was investigated in both the MnIV-MnIV and MnIV-MnIII oxidation states of SOZMUP. As detailed in section 2.2.1, the proton solvation free energy and the van der Waals radii of the solute are important values used to compute the free energies of solvation and pK_A values. The values of proton solvation free energy used was equal to -260.2 [71]. The van der Waals radii of hydrogen, carbon, nitrogen, oxygen, and chloride atoms [147] are listed in table 5.1, while the manganese ion radius is equal to 1.48 Å [127, 143]. The results obtained for all the derived molecules in both the oxidised and reduced states of SUZMUP are listed in tables 5.2. The obtained total root-mean-square deviations between the computed pK_A-values and the available experimental data (pK_A-RMSD) is 6.96 pH units.

In order to improve such poor results, the ab-initio method has been updated by determining more precise values of proton solvation free energies and by optimizing the van der Waals atomic radii as detailed in section 4.1 and [1]. The

Table 5.1: The van der Waals atomic radii of hydrogen, carbon, nitrogen, and oxygen are listed for MeCN. The radii are taken from [147].

atom	vdW radii (Å)
H	1.00
C	1.80
O/N	1.76
Cl	1.76

revised version of the ab-initio method has then been applied to the organometallic model complexes of interest. This investigation represents the first ab-initio pK_A computation of bridging oxygens in di-oxo-mangese complexes solvated in MeCN.

5.1.1 Updated ab-initio method

This section presents the results obtained by applying the updated ab-initio method on the SOZMUP complexes. Different strategies are tested here in order improve the computation of the pK_A values of such challenging systems. The B3LYP* functional is used to compute the ground state electronic energies of all the protonated and deprotonated di-oxo-manganese models. As explained in section 3.2, such a functional usually performs well when describing transition metal complexes. Nevertheless, the standard ab-initio method [1], based fully on DFT B3LYP, is also tested although limited to 3,5H-SOZMUP. The results of this study are therefore preliminary. In a further investigation, the B3LYP* and B3LYP performances may be compared including the effect of antiferromagnetic coupling between the

Table 5.2: The calculated and experimental pK_A values of the SOZMUP complex are shown for the different oxidation states MnIV-MnIV and MnIII-MnIV.

SOZMUP	exp	theory	dev
MnIV-MnIV 3,5-R=H	13.40	11.00	2.40
MnIII-MnIV 3,5-R=H	24.50	24.69	-0.19
MnIV-MnIV 3,5-R=Cl	10.80	3.05	7.75
MnIII-MnIV 3,5-R=Cl	20.20	19.27	0.93
MnIV-MnIV 3,5-R=NO_2	5.00	-6.62	11.62
MnIII-MnIV 3,5-R=NO_2	13.30	1.10	12.20
MnIV-MnIV 5-R=Cl	11.50	6.73	4.77
MnIV-MnIV 5-R=OCH_3	14.10	10.22	3.88

manganese ions as detailed in section 3.2. The influence of this effect on the pK_A computation is tested here, as detailed in section 3.2, on the first protonation equilibria in 3,5H-SOZMUP only.

5.1.1.1 Application of B3LYP*

The results obtained by investigating the protonation equilibria in SOZMUP are listed in table 5.3. Here, in contrast to the first preliminary investigation, the pK_A value is also computed for the second protonation of the 3,5-H-di-oxo-MnIV cluster. Although the total pK_A-RMSD between computed and measured pK_A values has significantly improved compared to the preceding preliminary computation, the pK_A-RMSD of 4.81 pH units is still very large. Nevertheless, such results can be used as a starting point to improve the strategy of ab-initio pK_A computations in double-core organometallic model systems.

Table 5.3: The calculated and measured pK_A values of the SOZMUP complexes are shown for the oxidation states MnIV-MnIV and MnIII-MnIV. In the fourth column the deviation between the measured and computed pK_A values is given as dev = exp - theory. *(s.p.) highlights the second protonation of the di-oxo-manganese cluster in the 3,5-H compound. All the other computations are performed for the first protonation of the oxo-bridges.

SOZMUP	exp	theory	dev
MnIV-MnIV 3,5-R=H (s.p.)*	6.50	4.45	2.05
MnIV-MnIV 3,5-R=H	13.40	13.59	0.19
MnIII-MnIV 3,5-R=H	24.50	28.01	-3.51
MnIV-MnIV 3,5-R=Cl	10.80	6.90	3.90
MnIII-MnIV 3,5-R=Cl	20.20	23.63	-3.43
MnIV-MnIV 3,5-R=NO_2	5.00	-3.49	8.49
MnIII-MnIV 3,5-R=NO_2	13.3	4.74	8.56
MnIV-MnIV 5-R=Cl	11.50	7.99	3.51
MnIV-MnIV 5-R=OCH_3	11.47	14.10	-2.63

5.1.1.1.1 Charge modeling to improve the pK_A prediction

Among the investigated compounds, the 3,5-NO_2 derived molecule shows the highest deviations from experiment with a variation of 8.49 and 8.56 pH units. Interestingly, the analysis of its atomic point charges in both protonation states highlights a charge asymmetry between the two di-nitro groups carried by the salpn ligands.

The nitrogen atoms placed on position 3 of the aromatic rings carry a partial charge of +0.92, while the ones on position 5 display a lower value of +0.82. Such a charge asymmetry between the groups may then affect the pK_A computation. This hypothesis is tested by applying a modeling strategy exploiting an isolated ring of the ligand. The atomic partial charges of deprotonated 1,3-dinitrobenzene are computed in gas-phase using the RESP procedure as detailed in section 2.2.1.1. The comparison between SOZMUP and model charges carried by the di-nitro groups, including their connecting atoms, is shown in figure 5.2. The atomic point

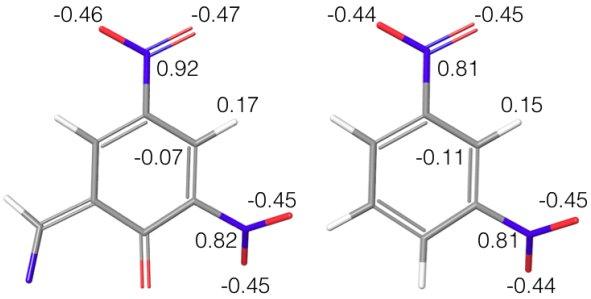

Figure 5.2: The comparison between the atomic partial charges of the 3,5-NO_2 aromatic ring of salpn (left) and 1,3-dinitrobenzene (right) is shown.

charge pattern of 1,3-dinitrobenzene shows that both the nitro substituents should carry the same charges. The partial charges of the nitro functional groups of 1,3-dinitrobenzene, as well as the ones of their connecting carbon atoms, are therefore employed to model the SOZMUP patterns. These charges are constant during the RESP procedure and are applied on protonated and deprotonated compounds in both their oxidation states, while all the other atomic charges can vary. The computed pK_A values are 0.55 and 6.96 for the oxidized and reduced molecular species, respectively. The corresponding deviations from experimental values are now 4.45 and 6.34 pH units in the two species. The modeled charges improve the results of about 50% and 25% with respect to the results of the preceding computations for the oxidized and reduced states, respectively. As a result, the total pK_A-RMSD for the SOZMUP series decreases from 4.81 to 3.69 pH units. Such a strategy may be applied in the future also to the other complexes.

5.1.1.2 Test of B3LYP

The standard ab-initio methodology is based on the application of the classical DFT/B3LYP approach. For this reason, the B3LYP functional is tested to compute the pK_A values of the first protonation for the two oxidation states of 3,5H-SOZMUP.

As a result the computed pK_A values are shifted from the original (what is original here is unclear) values of 13.59 and 28.01 to 14.42 and 29.26 for the oxidized and reduced states, respectively. The pK_A-RMSD of the two states rises from 2.48 for B3LYP* to 3.44 pH units for B3LYP.

5.1.1.3 Inclusion of the antiferromagnetic coupling effect

As described in section 3.2, the inclusion of the antiferromagnetic coupling effect has been tested on the system at B3LYP* and B3LYP levels of theory. Unfortunately, the ground state electronic energy does not converge when the quadruple zeta cc-pvqz(-g) basis set is used, as detailed in section 3.2. Therefore, the smaller lavc3p**++ basis set is applied to all the atoms to compute the energy contribution of both protonated and deprotonated molecules. Such a Mn1 (3/2) - Mn2 (3/2) spin investigation is however limited to the first protonation equilibrium of 3,5H-SOZMUP. While no convergence could be obtained using the B3LYP* functional, the predicted B3LYP-pK_A values are equal to 13.61 and 26.17 for the oxidized and reduced states, respectively, with deviations 0.21 and 1.67 pH units from the measured data. The pK_A-RMSD of 3,5H-SOZMUP decreases here from 2.48 and 3.44 obtained with high spin (S=3) using B3LYP* and B3LYP, respectively, compared to 1.19 pH units using B3LYP with low spin (S=0). In fact, these results agree better with experiment then the pK_A values obtained from the semi-empirical approach [136], where the corresponding deviation is 2.6 pH units (see section 2.3.3 for details). The inclusion of the antiferromagnetic coupling effect in the electronic ground state energy computation may considerably improve the accuracy of the ab-initio method. In that respect, the present investigation is preliminary and this issue should be addressed in more detail in future work.

5.1.2 Influence of the ligands on the pK_A values

An important question is how the pK_A of the di-oxo-manganese cluster is affected by its organic ligands. The salpn substituents of SOZMUP affect the pK_A values by influencing the electron density of the manganese ions [136]. Varying the substituents at the ligands shift the oxo-bridge pK_A value of the diMnIV-SOZMUP by a different amounts: the substituent 5-OCH_3, has a minor effect and decreases the pK_A value of the diMn-IV compound by only 1.93 units; 3,5-Cl shows a stronger influence and decreases the value by 2.60 units; the substituent 3,5-NO_2 has the strongest impact by shifting the pK_A value by 8.5 pH units. The 3,5-NO_2 substituent represents also the most computationally challenging ligand to investigate, giving rise to the highest deviation between experimental and ab-initio predicted pK_A values, which has been partially compensated by modeling in section 5.1.1.1.1.

5.1.3 Structural analysis of the optimized complexes

As previously detailed in section 3.2, the crystallographic structures of fully deprotonated diMn-IV 3,5-R-SOZMUP [145] is used as a starting point for all the geometry optimizations. Tables 5.4 lists the RMSD values of the two manganese and two μ-oxygens between the coordinates of the crystallographic and quantum chemically optimized di-oxo-manganese structures. This analysis shows that complexes with the same oxidation and protonation states show similar deviations from the crystal structure, apart from the protonated 3,5-NO_2-MnIII-MnIV compound. In fact, with a RMSD of 0.2651Å relative to the coordinates of the crystal structure, the core atoms of the di-nitro substituted complex show atom displacements that are 50% greater than the ones of the other molecules. This discrepancy in structural RMSD may influence the computed pK_A values. Figure 5.3 compares the crystallographic and optimized deprotonated structures of SOZMUP in the di-MnIV oxidation state. The optimized structure does not show relevant deviations from the crystallographic data and the two salpn ligands remain symmetric to each other, as expected.

Table 5.4: RMSD values are listed for geometry optimized di-oxo-manganese structures of the SOZMUP complex in different protonation and redox states relative to the corresponding crystal structure, which is deprotonated and in the redox state MnIV-MnIV. (d) refers to deprotonated, (p) to protonated, (ox) refers to MnIV-MnIV and (red) to the MnIII-MnIV.

subst./state	3,5-R=H	3,5-R=Cl	3,5-R=NO_2	5-R=Cl	5-R=OCH_3
d - ox	0.0133	0.0081	0.0080	0.0086	0.0138
d - red	0.0406	0.0397	0.0392		
p - ox	0.0896	0.0859	0.0884	0.0879	0.0860
p - red	0.1721	0.1742	0.2651		

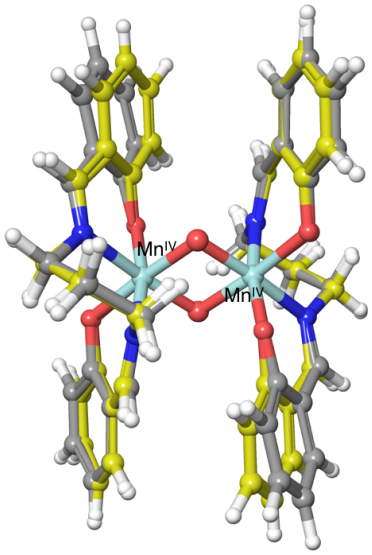

Figure 5.3: Crystallographic and optimized deprotonated MnIV-MnIV structure of SOZMUP. The crystallographic structure is highlighted in yellow, while the optimized structure is colored in gray. The manganese, oxygen, and nitrogen atoms are colored in pale blue, red, and blue, respectively. Hydrogen atoms are highlighted in white.

5.2 Empirical conversion method

In section 2.2.4, a fully empirical methodology is presented that converts macroscopic pK_A values of organic molecules from water to the corresponding values in MeCN, Me_2SO, and MeOH. The parameterization of the method is accomplished by matching the predicted pK_A values with the corresponding benchmark values for 82 different organic compounds involving titratable oxygen, nitrogen, or sulfur atoms. These compounds are gathered in 20 different molecular families of the solutes, which were previously defined in [58] and shown in figure 3.3 of section 3.4. The pK_A values in water are used as references and shifted to the values in the other solvents by applying a single additive parameter that is solvent and solute family dependent. The optimized parameters are shown in table 3.1 of section 3.4. Among the reference pK_A values, 53 are taken from Friesner et al. [58], while the remaining 31 values are calculated using the pK_A prediction module implemented in Jaguar 8.0 [127], which is theoretically based on [58]. The reference pK_A values are listed in tables A18-37 and their total root-mean-square deviation with respect to experimental values (pK_A-RMSD) is equal to 0.53 pH

units. The empirical parameterization is performed over 219 pK_A values computed and matched to the respective benchmark values in the three considered solvents. As previously introduced in section 3.4, measured pK_A values are not available for all the considered compounds solved in the three non-aqueous solvents. For this reason, in the case for which benchmark data are unavailable for a particular solvent, the electrostatic transformation method [4] is applied on the measured pK_A value available for the other solvents. The empirical parameterization is performed by matching the computed pK_A values to 18, 38, and 26 measured benchmark values, and 36, 30, and 35 electrostatically transformed values in MeCN, Me_2SO and MeOH, respectively (see figure 5.4, 5.5, 5.6, 5.7, 5.7 and tables A18-37 for details).
The total pK_A-RMSD between converted and benchmark values are, without three outliers, equal to 0.53, 0.55, and 0.44 for MeCN, Me_2SO, and MeOH, considering 55, 70, and 61 titratable compounds, respectively. Computed pK_A values with a deviation from experiment greater than 1.5 pH units are considered as outliers; namely 4-nitrophenol and succinic acid in MeCN with deviations from experiment equal to 2.30 and 1.90 pH units, respectively, and saccharin in MeCN and MeOH with deviations equal to 1.53 and 2.50. Nevertheless, even including those compounds, the pK_A-RMSD values increases only modestly to 0.68 and 0.54 in MeCN and MeOH, respectively. In this procedure, the outliers may be due to (i) solute-solvent conformations not properly covered by the empirical parameters, (ii) poor quality of the reference pK_A values in water, (iii) or systematic errors in the experimental titration of the selected benchmarks. A fourth reason could be related to the electrostatic transformation method used to enrich the benchmark database. The accuracy of this approach decreases when the atomic partial charges computed for a system in gas-phase are not appropriate in solution [4]. However, among the 101 benchmarks obtained by electrostatic transformation, only saccharin is identified as an outlier. Its poor prediction can be explained by the large deviation from experimental pK_A value taken from [58], which differs by 1.20 pH units from the measured value [158] (see table A22).
This method is optimized for the pK_A prediction module implemented in Jaguar 8.0 [127]. However, it can be generalized to be applied to the latest software version [66] as well as to directly convert measured pK_A values. Moreover, since the same set of parameters is valid to match the predictions to both experimentally measured and electrostatically transformed benchmark values, the robustness of the electrostatic transformation method [4] introduced by this thesis in section 2.2.3 is corroborated.

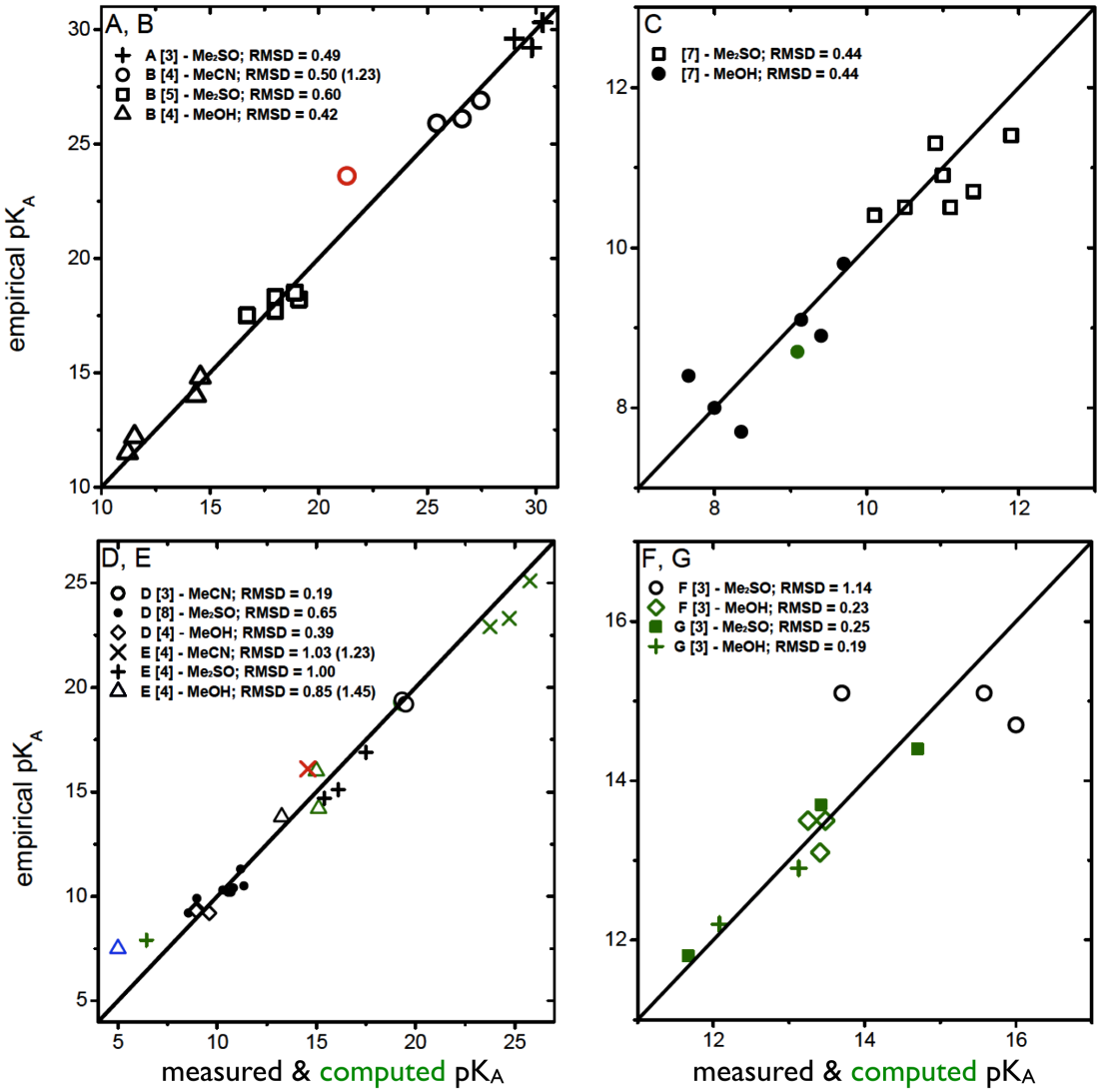

Figure 5.4: Correlation diagrams of measured and computed pK_A values are shown for [A] alcohols, [B] phenols, [C] carboxylic acids, [D] thiols, [E] sulfonamides, [F] hydroxamic acids, [G] barbituric acids. The matches with measured and electrostatic transformed benchmarks are shown in black and green, respectively. Deviations of more than 1.5 units between the computed and benchmark pK_A values are considered to be outliers and labeled in red or blue when the prediction is matched to measured values or electrostatic transformations, saccharin in MeCN and MeOH, respectively. The pKa-RMSD values including the outlier molecules are given in brackets.

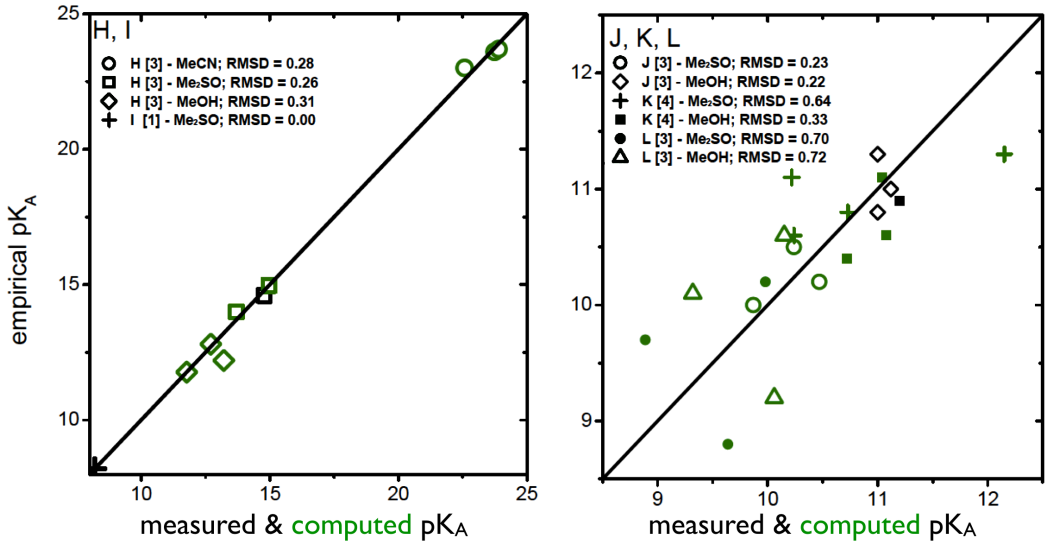

Figure 5.5: Correlation diagrams of measured and computed pK_A values are shown for [H] imides,[I] tetrazoles, [J] primary amines, [K] secondary amines, and [L] tertiary amines.

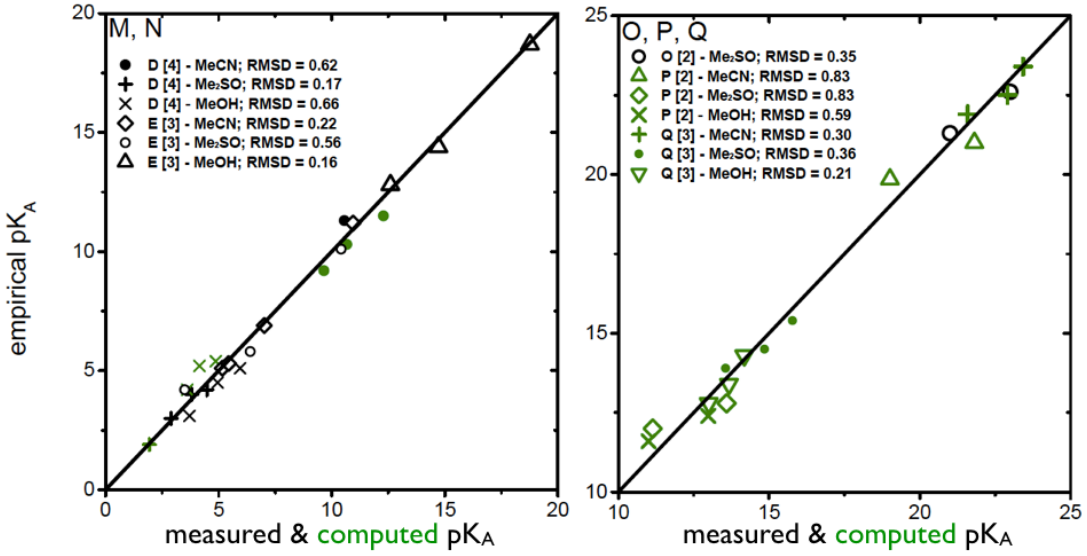

Figure 5.6: Correlation diagrams of the measured and computed pK_A values are shown for [M] anilines, [N] heterocycles, [O] indoles and pyrroles, [P] amidines, [Q] guanidines.

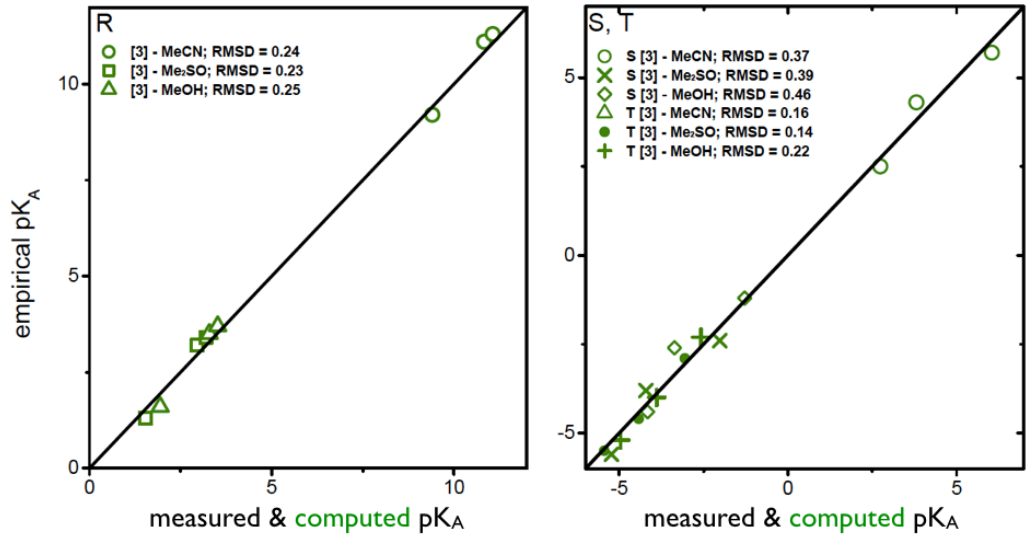

Figure 5.7: Correlation diagrams of the measured and computed pK_A values are shown for [R] benzodiazepines, [S] pyrroles C2-protonation, and [T] indoles C3-protonation.

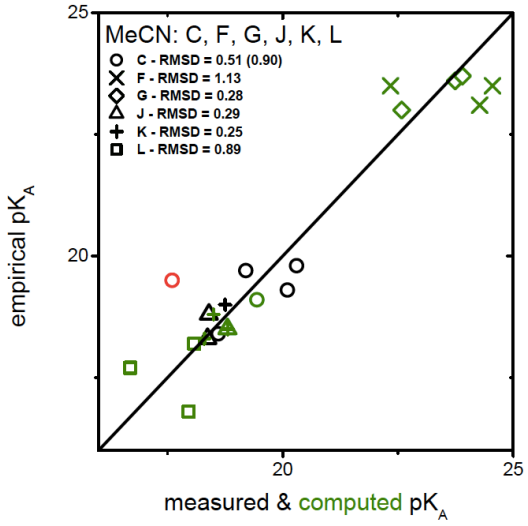

Figure 5.8: Correlation diagram of the measured and computed pK_A values in MeCN are shown for [C] carboxylic acids, [F] hydroxamic acids, [G] barbituric acids, [J] primary amines, [K] secondary amines, and [L] tertiary amines.

Chapter 6

Conclusions

The core of this thesis is comprised of four related topics concerning the prediction of pK_A values of organic and organometallic molecules solvated in different solvents. Section 4.1 describes the revision of the ab-initio methodology to accurately predict the pK_A values of molecules in acetonitrile (MeCN), methanol (MeOH), water, and dimethyl sulfoxide (Me_2SO) [1, 2]. Section 2.2.2 introduces the preliminary application of the revised ab-initio method to investigate the protonation equilibria of di-oxo-manganese model complexes solvated in MeCN. Section 4.2 introduces the newly established pK_A electrostatic transformation method [4] that can transform known pK_A values from one solvent to another regardless whether the involved solvents are protic or aprotic. Section 2.2.4 presents a new empirical methodology that converts known pK_A values of organic compounds in water to the corresponding values in acetonitrile (MeCN), methanol (MeOH), and dimethyl sulfoxide (Me_2SO). The application is demonstrated using the pK_A value predictor of Jaguar 8.0 [127].

6.1 Revision of the ab-initio method

In this thesis, the ab-initio method is updated to compute pK_A values of molecules in a series of protic and aprotic solvents. The experimental pK_A values of 19 organic compounds are reproduced with high accuracy using the proton solvation free energies of -255.1, -265.9, -266.3, and -266.4 kcal/mol in MeCN, MeOH, water, and Me_2SO, respectively [1]. The total deviations between computed and measured pK_A values in the different solvents (pK_A-RMSD) are equal to 0.67, 0.57, 0.47, and 0.63 in MeCN, MeOH, water, and Me_2SO, respectively.
The values of proton solvation energy obtained here agree semi-quantitatively with the ones predicted by Kelly et al. using the CPB approximation [71] as listed in

section 2.2.1.2. In contrast to previous works, the solute environment is modeled here as a dielectric medium. The high accuracy of the method is achieved without exploiting explicit solvent molecules that surround the solute. Such a purely electrostatic approach avoids any dependency on solute-solvent conformations as well as their averaging, an effect that is not trivial and solvent specific. On the other hand, this model does not account for specific solute-solvent conformations, which could be critical for the methodological accuracy. Some of the few resulting outliers could therefore result from the solvation model used.
The proton solvation free energy derived in water is mainly a parameter to match computed and measured pK_A values. As the value of proton solvation obtained in the present work is extremely close to the consensus value of -265.9 kcal/mol [93], it can be concluded that the optimized atomic radii for the present work are appropriate to compute electrostatic solvation energies. The radii of solute atoms for water are enlarged for the other solvents by a common enhancement factors, which are optimized by minimizing the pK_A-RMSD as detailed in section 4.1.
The proton solvation energy computed in MeCN agrees very well with the value of -255.2 kcal/mol experimentally obtained with the TATB approximation [102]. In Me_2SO, the value of the proton solvation free energy agrees with a measured value of -268.55 kcal/mol also based on absolute electrode potentials [104]. In comparison to the proton solvation energy obtained in other computational studies, the value obtained for MeCN agrees with values from other works apart from the one of [71] equal to -260.2 kcal/mol , whereas for Me_2SO it is close to two of four available values only. In MeOH, the proton solvation energy is very close to but systematically more negative than the previous measured and computed values (see section 2.2.1.2 for details).
The proton solvation free energies derived in this thesis correlate with the pK_A values of protonated MeCN, MeOH, and Me_2SO measured in water. This suggests that the proton solvated in MeCN, MeOH, and Me_2SO is attached to a single solvent molecule, while in water it assumes an Eigen-cluster conformation.
In this study, the atomic partial charges, as well as the vibrational and ground state electronic energies, are computed in gas-phase conditions and used for all the solvents. For this reason, the agreement between the experimental and computed proton solvation energy in water corroborates the results obtained in the other solvents as well. The free energies of proton solvation computed in this thesis are very accurate and suitable to improve the pK_A value prediction of molecules in non-aqueous solvents.

6.2 Application to di-oxo-manganese model complexes

In this thesis, an ab-initio method is applied to investigate the protonation equilibria of a series of di-oxo-manganese model complexes inspired by the oxygen evolving complex (OEC) of photosystem II. The series, referenced as SOZMUP, collects five different derived compounds in two oxidation states diMnIV (oxidized) and MnIII-MnIV (reduced) with substituents: 3,5-H, 3,5-Cl, 3,5-NO_2, 5-Cl, and 5-OCH_3. Different strategies are preliminarily explored to improve the accuracy of the pK_A prediction scheme. The DFT functionals B3LYP and B3LYP* are tested in combination with the cc-pvqz(-g) and lavc3p**++ basis sets to compute the ground state electronic energies. Such computations are performed with and without the effect of antiferromagnetic coupling. Moreover, a strategy of charge modeling improves the accuracy of the method in predicting the pK_A values of di-nitro derived compounds. The methodology provides predictions with a total pK_A-RMSD of 3.69 over a range of about 20 pH units using B3LYP*. The DFT functional B3LYP, preliminarily tested on 3,5-H-SOZMUP, yields a pK_A-RMSD of 3.44 pH units. Unfortunately, energy convergence is not achieved if the effect of the antiferromagnetic coupling is included using the standard procedure using quadruple zeta (see section 3.2 for details). Therefore, when including antiferromagnetic coupling the triple zeta lavc3p**++ basis set is applied to all the atoms of oxidized and reduced 3,5-H-SOZMUP. This improves the predictions by a factor of about three, with a final pK_A-RMSD of 1.19 units. Moreover, this deviation is improved by a factor of two with respect to the pK_A-RMSD of 2.47 pH units obtained with a semi-empirical approach in [136]. These preliminary single point energy calculations show poor convergence in the electronic ground state energy, and future investigations are required to understand the origin of the failure to converge. This study represents the first ab-initio investigation of the protonation equilibria in di-oxo-manganese model complexes and the results obtained justify the application of this method on the challenging OEC of photosystem II [3].

6.3 Electrostatic transformation method

In this thesis, the electrostatic transformation method has been introduced to compute the pK_A values of a large variety of titratable molecules in different solvents. Given an optimized geometry, the process requires only a few minutes to perform the predictions using a single CPU. However, the transformation can be accomplished only if the pK_A value is known in at least one of the investigated solvents. Contrary to the ab-initio method introduced in section 2.2.1, the ex-

pensive computations of vibrational and electronic ground state energies are not required. However, the geometry optimization of the molecule of interest is needed and performed using DFT in gas-phase conditions, as detailed in section 3.3. After geometry optimization, the atomic partial charges of the molecule are generated in gas-phase using the RESP procedure [159]. Only the electrostatic solvation energy contributions are used in the pK_A predictions, while all the non-electrostatic solvation energies cancel between the two protonation states of a solute molecule. A further energy contribution essential to transforming the pK_A values between two solvents are the proton solvation free energies in both solvents of interest. Here, the proton solvation energy derived in [1] for water, acetonitrile (MeCN), methanol (MeOH), and dimethyl sulfoxide (Me_2SO) are used.
The electrostatic transformation method is very accurate and it has been tested here for 30 different organic molecules by performing 77 predictions. The total pK_A-RMSD between computed and measured values is about 0.70 pH units including the outliers. This method shows therefore an accuracy comparable to alternative but computationally more expensive ab-initio approaches. The electrostatic transformation method is free from empirical factors and represents a robust and novel approach to compute pK_A values of organic molecules in protic and aprotic solvents.

6.4 Empirical conversion method

The third methodology introduced in this thesis is the empirical conversion method developed in collaboration with the software company Schrödinger, LCC. The method converts pK_A values predicted in water for instance by Jaguar 8.0 [127] to pK_A values in acetonitrile (MeCN), methanol (MeOH), and dimethyl sulfoxide (Me_2SO). The conversion is based on the application of a single additive empirical parameter, which depends on solvent and the molecular family of the solute. The empirical parameters are optimized for the three solvents by matching the computed pK_A values to benchmark pK_A values. The methodological optimization has then been performed over a total of 20 different molecular families gathering 82 different compounds. For cases in which the benchmark data are unavailable, these are derived by applying the electrostatic transformation method. The total pK_A-RMSD between converted and benchmark values, without the outliers, are equal to 0.53, 0.55, and 0.44 for MeCN, Me_2SO, and MeOH, respectively.
The empirical conversion has been optimized for the pK_A prediction module implemented in Jaguar 8.0. However, it can be alternatively parameterized to be applied to the latest software version [66] as well as to directly convert measured pK_A values from water to other solvents. Interestingly, in contrast to other previous

investigations [126], the empirical conversion method does not use multiplicative empirical factors to modify the slope of predicted values. Moreover, since the same set of empirical parameters is applied to match both measured and transformed benchmark pK_A values, it implicitly confirms the robustness of the electrostatic transformation methodology.

Bibliography

[1] E. Rossini and E. W. Knapp. "Proton solvation in protic and aprotic solvents". In: *J. Comput. Chem.* 37 (2016), pp. 1082–1091.

[2] E. Rossini and E. W. Knapp. "Erratum: Proton solvation in protic and aprotic solvents [J. Comput. Chem. 2015, 37, 1082-1091]". In: *J. Comput. Chem.* 37 (2016), pp. 1082–1091.

[3] A. Robertazzi, A Galstyan, and E. W. Knapp. "PSII manganese cluster: protonation of W2, O5, O4 and His337 in the S1 state explored by combined quantum chemical and electrostatic energy computations". In: *Biochim Biophys Acta.* 1837 (2014), pp. 1316–1321.

[4] E. Rossini, R. R Netz, and E. W. Knapp. "Computing pKa values in different solvents by electrostatic transformation". In: *J. Chem. Theory Comput.* 12 (2016), pp. 3360–3369.

[5] C. J. Cramer. *Essential of computational chemistry theories and models.* Wiley, 2004.

[6] P. Debye and E. Hückel. "Zur Theorie der Elektrolyte. I. Gefrierpunktserniedrigung und verwandte Erscheinungen". In: *Physik Z.* 24 (1923), pp. 185–206.

[7] J. Tomasi, B. Mennucci, and R. Cammi. "Quantum Mechanical Continuum Solvation Models". In: *Chem. Rev.* 105 (2005), pp. 2999–3094.

[8] I. Sakalli, J. Schöberl, and E.-W. Knapp. "mFES: A Robust Molecular Finite Element Solver for Electrostatic Energy Computations". In: *J. Chem. Theory Comput.* 10 (2014), pp. 5095–5112.

[9] J. Comput. Chem. "pKa in proteins solving the Poisson–Boltzmann equation with finite elements". In: *Sakalli, I. and Knapp, E.-W.* 36 (2015), pp. 2147–2157.

[10] A. Shrake and J. A. Rupley. "Environment and exposure to solvent of protein atoms. Lysozyme and insulin". In: *J Mol. Biol.* 79 (1973), pp. 351–371.

[11] D. T. Manallack. *The pKa distribution of drugs: application to drug discovery.* Dyes, Drugs: New Uses, and Implications, 2016.

[12] M. Meloun and S. Bordovská. "Benchmarking and validating algorithms that estimate pKa values of drugs based on their molecular structures". In: *Anal. Bioanal. Chem* 389 (2007), pp. 1267–1281.

[13] M. Rupp, R. Körner, and I. V. Tetko. "Estimation of acid dissociation constants using graph kernels". In: *Mol. Inf.* 29 (2010), pp. 731–741.

[14] A. Ščavničar, A. T. Balaban, and M. Pompe. "Application of variable anti-connectivity index to active sites. Modelling pKa values of aliphatic monocarboxylic acids". In: *SAR QSAR Environ. Res* 24 (2013), pp. 553–563.

[15] B. G. Tehan, E. J. Lloyd, M. G. Wong, W. R. Pitt, J. G. Montana, D. J. Manallack, and E. Gancia. "Estimation of pKa using semiempirical molecular orbital methods. Part 1: application to phenols and carboxylic acids". In: *Quant. Struct.-Act. Rel.* 21 (2002), pp. 457–472.

[16] B. G. Tehan, E. J. Lloyd, M. G. Wong, W. R. Pitt, E. Gancia, and D. J. Manallack. "Estimation of pKa using semiempirical molecular orbital methods. Part 2: application to amines, anilines and various nitrogen containing heterocyclic compounds". In: *Quant. Struct.-Act. Rel.* 21 (2002), pp. 473–485.

[17] E. Krieger, J. E. Nielsen, C. A. E. M. Spronk, and G. Vriend. "Fast empirical pKa prediction by Ewald summation". In: *J. Mol. Graphics Modell.* 25 (2006), pp. 481–486.

[18] J. Jover, R. Bosque, and J. Sales. "Neural network based QSPR study for predicting pKa of phenols in different solvents". In: *QSAR Comb. Sci.* 26 (2007), pp. 385–397.

[19] I. Ugur, A. Marion, S. Parant, J. H. Jensen, and G. Monard. "Rationalization of the pKa values of alcohols and thiols using atomic charge descriptors and its application to the prediction of amino acid pKa's". In: *J. Chem. Inf. Model.* 54 (2014), pp. 2200–2213.

[20] S. Geidl, R. Svobodová Vareková, V. Bendová, L. Petrusek, C.-M. Ionescu, Z. Jurka, R. Abagyan, and J. J. Koča. "How does the methodology of 3D structure preparation influence the quality of pKa prediction?" In: *Chem. Inf. Model.* 55 (2015), pp. 1088–1097.

[21] R. Parthasarathi, J. Padmanabhan, M. Elango, K. Chitra, V. Subramanian, and P. K. Chattaraj. "pKa prediction using group philicity". In: *J. Phys. Chem. A* 110 (2006), pp. 6540–6544.

[22] P. Kollman. "Free energy calculations: applications to chemical and biochemical phenomena". In: *Chem. Rev.* 93 (1993), pp. 2395–2417.

[23] J. Tomasi and M. Persico. "Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent". In: *Chem. Rev.* 94 (1994), pp. 2027–2094.

[24] C. J. Cramer and D. G. Truhlar. "Implicit solvation models: equilibria, structure, spectra, and dynamics". In: *Chem. Rev.* 99 (1999), pp. 2161–2200.

[25] M. Orozco and F. J. Luque. "Theoretical methods for the description of the solvent effect in biomolecular systems". In: *Chem. Rev.* 100 (2000), pp. 4187–4226.

[26] T. Simonson. "Electrostatics and dynamics of proteins". In: *Rep. Prog. Phys.* 66 (2003), pp. 737–787.

[27] C. J. Cramer and D. G. Truhlar. "A universal approach to solvation modeling". In: *Acc. Chem. Res.* 41 (2008), pp. 760–768.

[28] A. C. Lee and G. M. Crippen. "Predicting pKa". In: *J. Chem. Inf. Model.* 49 (2009), pp. 2013–2033.

[29] J. Ho and M. L. Coote. "A universal approach for continuum solvent pKa calculations are we there yet?" In: *Theor. Chem. Acc.* 125 (2010), pp. 3–21.

[30] K. S. Alongi, G. C. Shields, and R. A. Wheeler. *Theoretical calculations of acid dissociation constants: a review article. In Annual Reports in Computational Chemistry.* Vol. 6. Amsterdam: Elsevier, 2010, pp. 113–138.

[31] J. Ho and M. L. Coote. "First-principles prediction of acidities in the gas and solution phase". In: *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 1 (2011).

[32] J. E. Davies, N. L. Doltsinis, A. J. Kirby, C. D. Roussev, and M. Sprik. "Estimating pKa values for pentaoxyphosphoranes". In: *J. Am. Chem. Soc.* 124 (2002), pp. 6594–6599.

[33] D. Riccardi, P. Schaefer, and Q. Cui. "pKa calculations in solution and proteins with QM/MM free energy perturbation simulations: a quantitative test of QM/MM protocols". In: *J. Phys. Chem. B* 109 (2005), pp. 17715–17733.

[34] M. Sulpizi and M. Sprik. "Acidity constants from vertical energy gaps: density functional theory based molecular dynamics implementation". In: *Phys. Chem. Chem. Phys.* 10 (2008), pp. 5238–5249.

[35] N. Ghosh and Q. Cui. "pKa of residue 66 in Staphylococal nuclease. I. Insights from QM/MM simulations with conventional sampling". In: *J. Phys. Chem. B* 112 (2008), pp. 8387–8397.

[36] S. C. L. Kamerlin, M. Haranczyk, and A. Warshel. "Progress in ab initio QM/MM free-energy simulations of electrostatic energies in proteins: accelerated QM/MM studies of pKa, redox reactions and solvation free energies". In: *J. Phys. Chem. B* 113 (2009), pp. 1253–1272.

[37] M. Mangold, L. Rolland, F. Costanzo, M. Sprik, M. Sulpizi, and J. Blumberger. "Absolute pKa values and solvation structure of amino acids from density functional based molecular dynamics simulation". In: *J. Chem. Theory Comput.* 7 (2011), pp. 1951–1961.

[38] A. Warshel. “Dynamics of reactions in polar solvents. Semiclassical trajectory studies of electron-transfer and proton-transfer reactions”. In: *J. Phys. Chem.* 86 (1982), pp. 2218–2224.

[39] A. Warshel, F. Sussman, and G. King. “Free energy of charges in solvated proteins: microscopic calculations using a reversible charging process”. In: *Biochemistry* 25 (1986), pp. 8368–8372.

[40] R. Bürgi, P. A. Kollman, and W. F. Van Gunsteren. “Simulating proteins at constant pH: an approach combining molecular dynamics and Monte Carlo simulation”. In: *Proteins: Struct., Funct., Bioinf.* 47 (2002), pp. 469–480.

[41] A. M. Baptista, V. H. Teixeira, and C. M. Soares. “Constant-pH molecular dynamics using stochastic titration”. In: *J. Chem. Phys.* 117 (2002), pp. 4184–4200.

[42] T. Simonson, J. Carlsson, and D. A. Case. “Proton binding to proteins: pKa calculations with explicit and implicit solvent models”. In: *J. Am. Chem. Soc.* 126 (2004), pp. 4167–4180.

[43] E. J. Arthur, J. D. Yesselman, and C. L. Brooks. “Predicting extreme pKa shifts in staphylococcal nuclease mutants with constant pH molecular dynamics”. In: *Proteins: Struct., Funct., Bioinf.* 79 (2011), pp. 3276–3286.

[44] D. J. Bonthuis and R. R. Netz. “Beyond the continuum: how molecular solvent structure affects electrostatics and hydrodynamics at solid-electrolyte interfaces”. In: *J. Phys. Chem. B* 117 (2013), pp. 11397–11413.

[45] Y.-L. Lin, A. Aleksandrov, T. Simonson, and B. Roux. “An overview of electrostatic free energy computations for solutions and proteins”. In: *J. Chem. Theory Comput.* 10 (2014), pp. 2690–2709.

[46] S. T. Russell and A. Warshel. “Calculations of electrostatic energies in proteins. The energetics of ionized groups in bovine pancreatic trypsin inhibitor”. In: *J. Mol. Biol.* 185 (1985), pp. 389–404.

[47] A. Warshel and J. Aqvist. “Electrostatic energy and macromolecular function”. In: *Annu. Rev. Biophys. Biophys. Chem.* 20 (1991), pp. 267–298.

[48] V. Luzhkov and A. Warshel. “Microscopic models for quantum mechanical calculations of chemical processes in solutions: LD/ AMPAC and SCAAS/AMPAC calculations of solvation energies”. In: *J. Comput. Chem.* 13 (1992), pp. 199–213.

[49] J. Florián and A. Warshel. “Langevin dipoles model for ab initio calculations of chemical processes in solution: parameterization and application to hydration free energies of neutral and ionic solutes and conformational analysis in aqueous solution”. In: *J. Phys. Chem. B* 101 (1997), pp. 5583–5595.

[50] W. H. Orttung. “Direct solution of the Poisson equation for biomolecules of arbitrary shape, polarizability density and charge distribution”. In: *Ann. N.Y. Acad. Sci.* 303 (1977), pp. 22–37.

[51] M. K. Gilson, K. A. Sharp, and B. Honig. "Calculating the electrostatic potential of molecules in solution: method and error assessment". In: *J. Comput. Chem.* 9 (1988), pp. 327–335.

[52] D. Sitkoff, K. A. Sharp, and B. Honig. "Accurate calculation of hydration free energies using macroscopic solvent models". In: *J. Phys. Chem.* 98 (1994), pp. 1978–1988.

[53] B. Jayaram, Y. Liu, and D. L. Beveridge. "A modification of the generalized Born theory for improved estimates of solvation energies and pK shifts". In: *J. Chem. Phys.* 109 (1998), pp. 1465–1471.

[54] S. A. Hassan, F. Guarnieri, and E. L. Mehler. "A general treatment of solvent effects based on screened coulomb potentials". In: *J. Phys. Chem. B* 104 (2000), pp. 6478–6489.

[55] I. Sharma and G. A. Kaminski. "Calculating pKa values for substituted phenols and hydration energies for other compounds with the first-order fuzzy-border continuum solvation model". In: *J. Comput. Chem.* 33 (2012), pp. 2388–2399.

[56] A. Pomogaeva and D. M. Chipman. "Hydration energy from a composite method for implicit representation of solvent". In: *J. Chem. Theory Comput.* 10 (2014), pp. 211–219.

[57] A. Pomogaeva and D. M. Chipman. "Composite method for implicit representation of solvent in dimethyl sulfoxide and acetonitrile". In: *J. Phys. Chem. A* 119 (2015), pp. 5173–5180.

[58] J. J. Klicić, R. A. Friesner, S.-Y. Liu, and W. C. Guida. "Accurate prediction of acidity constants in aqueous solution via density functional theory and self-consistent reaction field methods". In: *J. Phys. Chem. A* 106 (2002), pp. 1327–1335.

[59] D. M. Chipman. "Computation of pKa from dielectric continuum theory". In: *J. Phys. Chem. A* 106 (2002), pp. 7413–7422.

[60] A. Klamt, F. Eckert, M. Diedenhofen, and M. E. Beck. "First principles calculations of aqueous pKa values for organic and inorganic acids using COSMO-RS reveal an inconsistency in the slope of the pKa scale". In: *J. Phys. Chem. A* 107 (2003), pp. 9380–9386.

[61] F. Eckert and A. Klamt. "Accurate prediction of basicity in aqueous solution with COSMO-RS". In: *J. Comput. Chem.* 27 (2006), pp. 11–19.

[62] H. Lu, X. Chen, and C.-G. Zhan. "First-principles calculation of pKa for cocaine, nicotine, neurotransmitters, and anilines in aqueous solution". In: *J. Phys. Chem. B* 111 (2007), pp. 10599–10605.

[63] S. Zhang, J. Baker, and P. Pulay. "A reliable and efficient first principles-based method for predicting pKa values. 1. Methodology." In: *J. Phys. Chem. A* 114 (2010), pp. 425–431.

[64] S. Zhang. "A reliable and efficient first principles-based method for predicting pKa values. 4. Organic bases". In: *J. Comput. Chem.* 33 (2012), pp. 2469–2482.

[65] T. Matsui, T. Baba, K. Kamiya, and Y. Shigeta. "An accurate density functional theory based estimation of pKa values of polar residues combined with experimental data: from amino acids to minimal proteins". In: *Phys. Chem. Chem. Phys.* 14 (2012), pp. 4181–4187.

[66] A. D. Bochevarov, E. Harder, T. F. Hughes, J. R. Greenwood, D. A. Braden, D. M. Philipp, D. Rinaldo, M. D. Halls, J. Zhang, and R. A. Friesner. "Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences". In: *Int. J. Quantum Chem.* 113 (2013), pp. 2110–2142.

[67] H. Farrokhpour and M. Manassir. "Approach for predicting the standard free energy solvation of H+ and acidity constant in nonaqueous organic solvents". In: *J. Chem. Eng. Data* 59 (2014), pp. 3555–3564.

[68] G. I. Almerindo, D. W. Tondo, and J. R. Pliego Jr. "Ionization of organic acids in dimethyl sulfoxide solution: a theoretical ab initio calculation of the pKa using a new parametrization of the polarizable continuum model". In: *J. Phys. Chem. A* 108 (2004), pp. 166–171.

[69] E. L. M. Miguel, P. L. Silva, and J. R. Pliego Jr. "Theoretical prediction of pKa in methanol: testing SM8 and SMD models for carboxylic acids, phenols, and amines". In: *J. Phys. Chem. B* 118 (2014), pp. 5730–5739.

[70] J.-N. Li, Y. Fu, L. Liu, and Q.-X. Guo. "First-principle predictions of basicity of organic amines and phosphines in acetonitrile". In: *Tetrahedron* 62 (2006), pp. 11801–11813.

[71] C. P. Kelly, C. J. Cramer, and D. G. Truhlar. "Single-ion solvation free energies and the normal hydrogen electrode potential in methanol, acetonitrile, and dimethyl sulfoxide". In: *J. Phys. Chem. B* 111 (2007), pp. 408–422.

[72] P. Poliak. "The DFT calculations of pKa values of the cationic acids of aniline and pyridine derivatives in common solvents". In: *Acta Chim. Slovaca* 7 (2014), pp. 25–30.

[73] Jr Montgomery JA, Frisch MJ, Ochterski JW, and Petersson GA. "A complete basis set model chemistry. VI. Use of density functional geometries and frequencies". In: *J. Chem. Phys.* 110 (1999), pp. 2822–2827.

[74] Jr Montgomery JA, MJ Frisch, Ochterski JW, and Petersson GA. "A complete basis set model chemistry. VII. Use of the minimum population localization method". In: *J. Chem. Phys.* 112 (2000), pp. 6532–6542.

[75] J. W. Ochterski, G. A. Petersson, and J. A. Montgomery. "A complete basis set model chemistry. V. Extensions to six or more heavy atoms". In: *J. Chem. Phys.* 104 (1996), pp. 2598–2619.

[76] F.C. Pickard, D.R. Griffith, S.J. Ferrara, M.D. Liptak, K.N. Kirschner, and G.C. Shields. "CCSD(T), W1, and other model chemistry predictions for gas-phase deprotonation reactions". In: *Int. J. Quantum Chem.* 106 (2006), pp. 3122–3128.

[77] E. K. Pokon, M. D. Liptak, S. Feldgus, and G. C. Shields. "Comparison of CBS-QB3, CBS- APNO, and G3 predictions of gas phase deprotonation data". In: *J. Phys. Chem. A* 105 (2001), pp. 10483–10487.

[78] M. D. Liptak and G. C. Shields. "Comparison of density functional theory predictions of gas-phase deprotonation data". In: *Int. J. Quantum Chem.* 105 (2005), pp. 580–587.

[79] K. M. Ervin and V. F. DeTuri. "Anchoring the gas-phase acidity scale". In: *J. Phys. Chem. A* 106 (2002), pp. 9947–9956.

[80] W. Danikiewicz. "How reliable are gas-phase proton affinity values of small carbanions? A comparison of experimental data with values calculated using Gaussian-3 and CBS compound methods". In: *Int. J. Mass Spectrom.* 285.86-94 (2009).

[81] O. Sackur. "Die Anwendung der kinetischen Theorie der Gase auf chemische Probleme". In: *Annalen der Physik* 36 (1911), pp. 958–980.

[82] O. Sackur. *Die Bedeutung des elementaren Wirkungsquantums für die Gastheorie und die Berechnung der chemischen Konstanten. Festschrift W. Nernst zu seinem 25 jährigen Doktorjubiläum gewidmet von seinen Schülern.* Ed. by Wilhelm Knapp. 405-423. Germany: Halle an der Salle, 1912.

[83] O. Sackur. "Die universelle Bedeutung des sog. elementaren Wirkungsquantums". In: *Annalen der Physik* 40 (1913), pp. 67–86.

[84] H. Tetrode. "Berichtigung zu meiner Arbeit: "Die chemische Konstante der Gase und das elementare Wirkungsquantum". In: *Annalen der Physik* 38 (1912), pp. 434–442.

[85] P. J. Mohr, B. N. Taylor, and D. B. Newell. "CODATA Recommended values of the fundamental physical constants: 2010". In: *J. Phys. Chem. Ref. Data* 41 (2012), pp. 043109/1–043109/84.

[86] D. Bashford and K. Gerwert. "Electrostatic Calculations of the pKa Values of Ionizable Groups in Bacteriorhodopsin". In: *J. Mol. Biol.* 224.473-486 (1992).

[87] D. Bashford, D. A. Case, C. Dalvit, L. Tennant, and P. E. Wright. "Electrostatic calculations of side-chain pKa values in myoglobin and comparison with NMR data for histidines". In: *Biochemistry* 32 (1993), pp. 8045–8056.

[88] G. C. Shields. "Chapter 8 – Theoretical Calculations of Acid Dissociation Constants: A Review Article". In: *Annu. Rep. Comput. Chem.* 6 (2010), pp. 113–138.

[89] D. M. Chipman. “Computation of pKa from Dielectric Continuum Theory”. In: *J. Phys. Chem. A* 106 (2002), pp. 7413–7422.

[90] J. Ho and M. L. Coote. “A universal approach for continuum solvent pKa calculations: are we there yet?” In: *Theor. Chem. Acc.* 125 (2010), pp. 3–21.

[91] P. G. Seybold and G. C. Shields. “Computational estimation of pKa values”. In: *WIREs Comput. Mol. Sci.* 5 (2015), pp. 290–297.

[92] M. Schmidt am Busch and E. W. Knapp. “Accurate pKa Determination for a Heterogeneous Group of Organic Molecules”. In: *ChemPhysChem* 5 (2004), pp. 1513–1522.

[93] M. D. Tissandier, K. A. Cowen, W. Y. Feng, E. Gundlach, M. H. Cohen, A. D. Earhart, J. V. Coe, and T. R. Tuttle. “The Proton’s Absolute Aqueous Enthalpy and Gibbs Free Energy of Solvation from Cluster-Ion Solvation Data”. In: *J. Phys. Chem. A* 102 (1998), pp. 7787–7794.

[94] D. M. Camaioni and C. A. Schwerdtfeger. “Comment on “Accurate Experimental Values for the Free Energies of Hydration of H+, OH-, and H3O+””. In: *J. Phys. Chem. A* 109 (2005), p. 10795.

[95] W. A. Donald and E. R. Williams. “An Improved Cluster Pair Correlation Method for Obtaining the Absolute Proton Hydration Energy and Enthalpy Evaluated with an Expanded Data Set”. In: *J. Phys. Chem. B* 14 (2010), pp. 13189–13200.

[96] M. D. Liptak and G. C. Shields. “Accurate pKa Calculations for Carboxylic Acids Using Complete Basis Set and Gaussian-n Models Combined with CPCM Continuum Solvation Methods”. In: *J. Am. Chem. Soc.* 123 (2001), pp. 7314–7319.

[97] M. D. Liptak, K. C. Gross, P. G. Seybold, S. Feldgus, and G. C. Shields. “Absolute pKa Determinations for Substituted Phenols”. In: *J. Am. Chem. Soc.* 124 (2002), pp. 6421–6427.

[98] M. Eigen. “Proton Transfer, Acid-Base Catalysis, and Enzymatic Hydrolysis. Part I: ELEMENTARY PROCESSES”. In: *Angew. Chem. Int.* 3 (1964), pp. 1–19.

[99] C. G. Zhan and D. A. Dixon. “Absolute Hydration Free Energy of the Proton from First-Principles Electronic Structure Calculations”. In: *J. Phys. Chem. A* 105 (2001), pp. 11534–11540.

[100] V. S. Bryantsev, M. S. Diallo, and W. A. Goddard III. “Calculation of Solvation Free Energies of Charged Solutes Using Mixed Cluster/Continuum Models”. In: *J. Phys. Chem. B* 112 (2008), pp. 9709–9719.

[101] N. F. Carvalho and J. R. Pliego Jr. “Cluster-continuum quasichemical theory calculation of the lithium ion solvation in water, acetonitrile and dimethyl sulfoxide: an absolute single-ion solvation free energy scale”. In: *Phys. Chem. Chem. Phys.* 17 (2015), pp. 26745–26755.

[102] G. H. Kalidas, Y. Marcus, and G. Hefter. "Gibbs Energies of Transfer of Cations from Water to Mixed Aqueous Organic Solvents". In: *Chem. Rev.* 100 (2000), pp. 819–852.

[103] Y. Marcus, M. J. Kamlet, and R. W. Taft. "Linear solvation energy relationships: standard molar Gibbs free energies and enthalpies of transfer of ions from water into nonaqueous solvents". In: *J. Phys. Chem.* 92 (1988), pp. 3613–3622.

[104] W. R. Fawcett. "The ionic work function and its role in estimating absolute electrode potentials". In: *Langmuir* 24 (2008), pp. 9868–9875.

[105] E. Westphal and J. R. Pliego Jr. "Absolute solvation free energy of Li+ and Na+ ions in dimethyl sulfoxide solution: A theoretical ab initio and cluster-continuum model study". In: *J. Chem. Phys.* 123 (2005), p. 074508.

[106] S. Hwang and D. S. Chung. "Calculation of the solvation free energy of the proton in methanol". In: *Bull. Korean Chem. Soc.* 26 (2005), pp. 589–593.

[107] Y. Marcus. "Thermodynamics of ion hydration and its interpretation in terms of a common model". In: *Pure Appl. Chem.* 59 (1987), pp. 1093–1101.

[108] J. R. Pliego Jr. and J. M. Riveros Jr. "Gibbs energy of solvation of organic ions in aqueous and dimethyl sulfoxide solutions". In: *Phys. Chem. Chem. Phys.* 4 (2002), pp. 1622–1627.

[109] Y. Fu, L. Liu, R. Q. Li, R. Liu, and Q. X. Guo. "First-Principle Predictions of Absolute pKa's of Organic Acids in Dimethyl Sulfoxide Solution". In: *J. Am. Chem. Soc.* 126 (2004), pp. 814–822.

[110] D. Himmel, S. K. Goll, I. Leito, and I. Krossing. "Anchor Points for the Unified Brønsted Acidity Scale: The rCCC Model for the Calculation of Standard Gibbs Energies of Proton Solvation in Eleven Representative Liquid Media". In: *Chem. Eur. J.* 17 (2011), pp. 5808–5826.

[111] S. Hwang and D. S. Chung. "Gas Phase Proton Affinity, Basicity, and pKa Values for Nitrogen Containing Heterocyclic Aromatic Compounds". In: *Bull. Kor. Chem. Soc.* 26 (2005), pp. 585–588.

[112] J. R. Pliego Jr. and E. L. M. Miguel. "Absolute Single-Ion Solvation Free Energy Scale in Methanol Determined by the Lithium Cluster-Continuum Approach". In: *J. Phys. Chem. B* 117 (2013), pp. 5129–5135.

[113] S. S. David and E. Meggers. "Inorganic chemical biology: from small metal complexes in biological systems to metalloproteins". In: *Curr Opin Chem Biol.* 12 (2008), pp. 194–196.

[114] S. S. David and E. Meggers. "Inorganic Chemical Biology: from Small Metal Complexes in Biological Systems to Metalloproteins". In: *Curr. Opin. Chem. Biol.* 12 (2008), pp. 194–196.

[115] U. Lindh. "Metal Biology: Aspects of Beneficial Effects". In: *Ambio* 36 (2007), pp. 107–110.

[116] R. H. Holm, P. Kennepohl, and E. I. Solomon. "Structural and Functional Aspects of Metal Sites in Biology". In: *Chem Rev.* 96 (1996), pp. 2239–2314.
[117] J. M. Mayer. "Proton-Coupled Electron Transfer: A Reaction Chemist's View". In: *Annu Rev Phys Chem.* 55 (2004), pp. 363–390.
[118] P. C. A. Bruijnincx and P. J. Sadler. "New Trends for Metal Complexes with Anticancer Activity". In: *Curr. Opin. Chem. Biol.* 12 (2008), pp. 197–206.
[119] J. Suh and W. S. Chei. "Metal complexes as artificial proteases: toward catalytic drugs". In: *Curr. Opin. Chem. Biol.* 12.207-213 (2008).
[120] K. J. Barnham and I. B. Ashley. "Metals in Alzheimer's and Parkinson's Diseases". In: *Curr. Opin. Chem. Biol.* 12 (2008), pp. 222–228.
[121] A. I. Bush. "The metal theory of Alzheimer's disease." In: *J. Alzheimers. Dis.* 33.Suppl1 (2013), S277–281.
[122] P. Chernev, I. Zaharieva, E. Rossini, A. Galstyan, H. Dau, and E.-W. Knapp. "Merging Structural Information from X-ray Crystallography, Quantum Chemistry, and EXAFS Spectra: The Oxygen-Evolving Complex in PSII". In: *J. Phys. Chem. B* 120 (2016), pp. 10899–10922.
[123] M. J. Baldwin and V. L. Pecoraro. "Energetics of proton-coupled electron transfer in high-valent Mn2 (µ-O) 2 systems: models for water oxidation by the oxygen-evolving complex of Photosystem II". In: *J. Am. Chem. Soc.* 118.11325-11326 (1996).
[124] M. J. Baldwin, A. Gelasco, and V. L. Pecoraro. "The effect of protonation on [Mn(IV)(µ2-O)]2 complexes". In: *Photosynth. Res.* 38 (1993), pp. 303–308.
[125] M. Rupp, R. Ramakrishnan, and O. A. von Lilienfeld. "Machine learning for quantum mechanical properties of atoms in molecules". In: *J. Phys. Chem. Lett.* 6 (2015), pp. 3309–3313.
[126] C. R. Nicoleti, V. G. Marini, L. M. Zimmermann, and V. G. Machado. "Anionic chromogenic chemosensors highly selective for fluoride or cyanide based on 4-(4-Nitrobenzylideneamine)phenol". In: *J. Braz. Chem. Soc.* 23 (2012), pp. 1488–1500.
[127] *Jaguar, Version 8.3*. Schrödinger, LLC. New York, NY, 2014.
[128] URL: http://www.chem.wisc.edu/areas/reich/pkatable/.
[129] C. P. Kelly, C. J. Cramer, and D. G. Truhlar. "SM6: A density functional theory continuum solvation model for calculating aqueous solvation free energies of neutrals, ions, and solute-water clusters". In: *J. Chem. Theory Comput.* 1 (2005), pp. 1133–1152.
[130] T. Zhu, J. Li, G. D. Hawkins, C. J. Cramer, and D. G. Truhlar. "Density functional solvation model based on CM2 atomic charges". In: *J. Chem. Phys.* 109 (1998), pp. 9117–9133.

[131] C. Lee, W. Yang, and R. G. Parr. "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density". In: *Phys. Rev. B* 37 (1988), pp. 785–789.

[132] A. D. Becke. "Density-functional thermochemistry. III. The role of exact exchange". In: *J. Chem. Phys.* 98 (1993), pp. 5646–5652.

[133] Jr. Dunning T. H. "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen". In: *J. Chem. Phys.* 90 (1989), pp. 1007–1023.

[134] C. M. Cortis and R. A. Friesner. "An automatic three-dimensional finite element mesh generation system for the Poisson-Boltzmann equation". In: *J. Comput. Chem.* 18 (1997), pp. 1570–1590.

[135] C. M. Cortis and R. A. Friesner. "Numerical solution of the Poisson-Boltzmann equation using tetrahedral finite-element meshes". In: *J. Comput. Chem.* 18 (1997), pp. 1591–1608.

[136] M. Amin, L. Vogt, S. Vassiliev, I. Rivalta, M. Sultan, D. Bruce, G. W. Brudvig, V. S. Batista, and M. R. Gunner. "Electrostatic effects on proton-coupled electron transfer in oxomanganese complexes inspired by the oxygen-evolving complex of photosystem II". In: *J Phys Chem B.* 117 (2013), pp. 6217–6226.

[137] Y. Song, J. Mao, and M. R. Gunner. "MCCE2: Improving protein pKa calculations with extensive side chain rotamer sampling". In: *J. Comp. Chem.* 30 (2009), pp. 2231–2247.

[138] W. Rocchia, E Alexov, and B. Honig. "Extending the applicability of the nonlinear poisson-boltzmann equation: multiple dielectric constants and multivalent ions". In: *J Phys Chem B.* 105 (2001), pp. 6507–6514.

[139] G. Frison and G. Ohanessian. "A comparative study of semiempirical, ab initio, and DFT methods in evaluating metal-ligand bond strength, proton affinity, and interactions between first and second shell ligands in Zn-biomimetic complexes." In: *J Comput. Chem.* 29 (2008), pp. 416–433.

[140] Y. Fu, L. Liu, R. Q. Li, R. Liu, and Q. X. Guo. "First-Principle Predictions of Absolute pKa's of Organic Acids in Dimethyl Sulfoxide Solution". In: *J. Am. Chem. Soc.* 126 (2004), pp. 814–822.

[141] K. Range, C. S. López, A. Moser, and D. M. York. "Multilevel and Density Functional Electronic Structure Calculations of Proton Affinities and Gas-Phase Basicities Involved in Biological Phosphoryl Transfer". In: *J. Phys. Chem. A* 110 (2006), pp. 791–797.

[142] V. S. Bryantsev, M. S. Diallo, and W. A. Goddard III. "pKa calculations of aliphatic amines, diamines, and aminoamides via density functional theory with a Poisson-Boltzmann continuum solvent model". In: *J. Phys. Chem. A* 111 (2007), pp. 4422–4430.

[143] *Jaguar, Version 7.7*. Schrödinger, LLC. New York, NY, 2010.

[144] A. V. Marenich, R. M. Olson, C. P. Kelly, Cramer. C. J., and D. J. Truhlar. "Self-Consistent reaction field model for aqueous and nonaqueous solutions based on accurate polarized partial charges". In: *J. Chem. Theory Comput.* 3 (2007), pp. 2011–2033.

[145] J. W. Gohdes and W. H. Armstrong. "Synthesis, structure, and properties of [Mn(salpn)(EtOH)2](ClO4) and its aerobic oxidation product [Mn(salpn)O]2". In: *Inorg. Chem.* 31 (1992), pp. 368–373.

[146] O. Salomon, M. Reiher, and B. A. Hess. "Assertion and validation of the performance of the B3LYP* functional for the first transition metal row and the G2 test set". In: *J. Chem. Phys.* 117 (2002), pp. 4729–4737.

[147] A. Galstyan and E.-W. Knapp. "Accurate redox potentials of mononuclear iron, manganese, and nickel model complexes". In: *J. Comput. Chem.* 2 (2009), pp. 203–211.

[148] E. Wicke, M. Eigen, and Th. Ackermann. "Über den Zustand des Protons (Hydroniumions) in wäßriger Lösung". In: *Z. Phys. Chem.* 1 (1954), p. 340.

[149] A. Warshel and S. T. Russell. "Calculations of electrostatic interactions in biological systems and in solutions". In: *Q. Rev. Biophys.* 17 (1984), pp. 283–422.

[150] R. G. Pearson. "Ionization potentials and electron affinities in aqueous solution". In: *J. Am. Chem. Soc.* 108 (1986), pp. 6109–6114.

[151] C. Lim, D. Bashford, and M. Karplus. "Absolute pKa calculations with continuum dielectric methods". In: *J. Phys. Chem.* 95 (1991), pp. 5610–5620.

[152] J. Aaqvist and A. Warshel. "Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches". In: *Chem. Rev.* 93 (1993), pp. 2523–2544.

[153] M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello. "On the Quantum Nature of the Shared Proton in Hydrogen Bonds". In: *Science* 275 (1997), pp. 817–820.

[154] G. Wada. "Basicity of Dimethyl Sulfoxide in an Aqueous Solution as Measured by the Spectrophotometric Indicator Method". In: *Bull. Chem. Soc. Jpn.* 42 (1969), pp. 890–893.

[155] P. M. Dewick. *Essentials of Organic Chemistry: For Students of Pharmacy, Medicinal Chemistry and Biological Chemistry*. Ed. by Wiley. Chichester, 2006.

[156] B. S. Chauchan. *Principles of Biochemistry and Biophysics*. New Dehli: University Science Press, 2008.

[157] R. F. Daley. *Organic Chemistry*. 2013. URL: http://www.ochem4-free.info/.

[158] R. P. Bell and W. C. E. Higginson. “The Catalyzed Dehydration of Acetaldehyde Hydrate, and the Effect of Structure on the Velocity of Protolytic Reactions”. In: *Proc. R. Soc. A.* 197 (1949), pp. 141–159.

[159] B. H. Besler, K. M. Merz, and P. A. Kollman. “Atomic charges derived from semiempirical methods”. In: *J. Comput. Chem.* 11 (1990), pp. 431–439.

List of abbreviations

QM - quantum mechanics;
MD - molecular dynamics;
MeCH - acetonitrile;
MeOH - methanol;
Me_2SO - dimethyl sulfoxide;
PBE - poisson boltzmann equation;
lPBE - linerized poisson boltzmann equation;
SASA - solvent accessible surface area;
SES - solvent exluded surface;
IEL - ion exclusion layer;
ESP - electrostatic potential;
RESP - restrained electrostatic potential method;
DFT - density functional theory;
PCM - polarizable continuum model;
QSAR - quantitative structure–activity relationship;
MEAD - macroscopic electrostatics with atomic detail;
CPA - cluster-pair approximation;
TATB - tetraphenylarsonium tetraphenylborate;
OEC - oxygen evolving complex;
ESE - electrostatic solvation energies;
FDM - finite difference method;
vdW - van der Waals;
COSMO - conductor-like screening model;
SMx - minnesota solvation models;
CM4 - charge model 4;
SCRF - self-consistent reaction field;
MCCE - multi-conformer continuum electrostatics;
RMSD - root-mean-square deviation;

List of Figures

List of Tables

Chapter 7

Appendix

7.1 Protonation equilibria of di-oxo-manganese complexes

Table A1: list of the optimized coordinates and atomic partial charges of the deprotonated 3,5-H-SOZMUP in the oxidation state MnIV-MnIV

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.5030074647	1.1934046532	8.0726547811	1.481296
O2	-2.7847988751	2.5221178971	7.5359586407	-0.733946
O3	-1.5776881965	0.4983696744	6.3085090832	-0.645222
O4	-2.7457135433	0.0842444680	8.7369514696	-0.802544
N5	-1.3664559141	2.2302773790	9.8162911510	-0.060594
N6	0.1255291794	2.3912034163	7.4775059484	-0.368141
C7	-3.7682624939	2.9271317733	8.2959089701	0.676724
C8	-4.9917448067	3.3236177626	7.7007438863	-0.358458
C9	-6.0448143446	3.7749056674	8.4758761021	-0.018483
C10	-5.9301966343	3.8810053285	9.8757093991	-0.241575
C11	-4.7355901042	3.5337882335	10.4776835323	-0.027158
C12	-3.6487226594	3.0484708270	9.7176057052	-0.277560
C13	-2.3832191556	2.8201025541	10.3558044274	0.040378
C14	-0.0526595248	2.2588562474	10.4519669248	-0.032859
C15	0.8909180680	3.2411935180	9.7062546929	-0.054543
C16	0.4470347102	3.5754916776	8.2710355754	0.165006
C17	0.8695799670	2.0960255875	6.4655557584	0.151598
C18	0.6360971128	1.0489428949	5.5073245713	-0.174359

Continued on next page

Table A1 – continued from previous page

Atom name	X	Y	Z	charge
C19	1.6228206661	0.8141301278	4.5250498778	-0.105165
C20	1.4255854357	-0.0911382658	3.4993733577	-0.175846
C21	0.1921481801	-0.7622593651	3.4208986562	-0.090405
C22	-0.7993788197	0.5494878899	4.3627884201	-0.224183
C23	-0.6039559170	0.3390446917	5.4532793528	0.500183
H24	-5.0732892831	3.2428655047	6.6213169023	0.161877
H25	-6.9783279823	4.0535034126	7.9922232651	0.110124
H26	-6.7659484761	4.2375937524	10.4694450149	0.129084
H27	-4.6190325752	3.6217540360	11.5558354836	0.104950
H28	-2.2616423480	3.2202972714	11.3681843751	0.105839
H29	-0.1526951005	2.5557248100	11.5030867712	0.050810
H30	0.3385600510	1.2425832730	10.4099196941	0.085525
H31	0.9525305526	4.1922396095	10.2487924572	0.015673
H32	1.8999689071	2.8161018795	9.6891645323	0.039563
H33	1.2334771513	4.1631457535	7.7778060511	0.007919
H34	-0.4568931521	4.1929152236	8.2995578790	0.005453
H35	1.7529355856	2.7191021657	6.2755329130	0.099153
H36	2.5561042202	1.3705398105	4.5872613456	0.111750
H37	2.1988509220	-0.2667981213	2.7584373936	0.121280
H38	0.0089671604	-1.4573329713	2.6045597862	0.116395
H39	-1.7602429029	-1.0490810731	4.2906968669	0.111976
Mn40	-1.5881218039	-1.1942142408	9.3352192600	1.471047
O41	-0.3455934539	-0.0844002406	8.6725770797	-0.801596
O42	-0.3034324055	-2.5225524682	9.8684441440	-0.733117
O43	-1.5148655163	-0.5006284623	11.1002087681	-0.641374
N44	-1.7258801935	-2.2294399236	7.5912177954	-0.059742
N45	-3.2142243873	-2.3959560331	9.9313719787	-0.362651
C46	0.6794948098	-2.9243827207	9.1060456936	0.677147
C47	-2.4893495975	-0.3420559736	11.9549589091	0.495523
C48	-0.7091871413	-2.8166190539	7.0486606612	0.041386
C49	-3.0415614397	-2.2607938091	6.9595195366	-0.032975
C50	-3.5310364321	-3.5826693411	9.1393234075	0.161274
C51	-3.9583210254	-2.1034303229	10.9440811181	0.148718
C52	1.9048911709	-3.3194839873	9.6981649519	-0.358398
C53	0.5578056364	-3.0435946009	7.6842844991	-0.277937
C54	-3.7275414238	-1.0552155439	11.9017582642	-0.170552
C55	-2.2966131803	0.5490291743	13.0438743852	-0.222497

Continued on next page

Table A1 – continued from previous page

Atom name	X	Y	Z	charge
H56	-0.8320413703	-3.2154378252	6.0358805316	0.105655
C57	-3.9767787200	-3.2527731298	7.7034677974	-0.053791
H58	-2.9437910525	-2.5508100703	5.9060415908	0.050559
H59	-3.4379795721	-1.2465554682	7.0092053284	0.085504
H60	-4.3144795762	-4.1734782006	9.6336226617	0.008606
H61	-2.6241388060	-4.1958158804	9.1112437162	0.006438
H62	-4.8392067692	-2.7297041454	11.1351578406	0.099285
C63	2.9574878047	-3.7672322227	8.9202906680	-0.018515
H64	1.9883478757	-3.2403815288	10.7775600481	0.161776
C65	1.6442407899	-3.5254739878	6.9213770339	-0.027033
C66	-4.7146982873	-0.8222159734	12.8840441126	-0.106150
C67	-3.2886279542	0.7602244510	13.9855628959	-0.090816
H68	-1.3375446141	1.0521603496	13.1148306427	0.111268
H69	-4.0279151837	-4.2048985583	7.1607936982	0.015672
H70	-4.9905845232	-2.8376439296	7.7185727654	0.039363
C71	2.8406660512	-3.8713109528	7.5204724181	-0.241684
H72	3.8924132862	-4.0446831930	9.4018648187	0.110093
H73	1.5257739442	-3.6116323287	5.8432950230	0.104840
C74	-4.5198799777	0.0850199441	13.9084458888	-0.175652
H75	-5.6463361788	-1.3814590959	12.8227497568	0.111894
H76	-3.1075908842	1.4573298919	14.8006762615	0.116497
H77	3.6760709969	-4.2252229477	6.9246667809	0.129124
H78	-5.2934288776	0.2593653375	14.6493721478	0.121301

Table A2: list of the optimized coordinates and atomic partial charges of the deprotonated 3,5-H-SOZMPU in the oxidation state MnIV-MnIII

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.3321630982	1.0648162316	7.9943337187	1.019831
O2	-2.6180622116	2.3680423055	7.1789735588	-0.637530
O3	-1.2674384531	0.2218204485	6.2572425224	-0.455141
O4	-2.6193733517	0.0348907696	8.6111973217	-0.682803
N5	-1.3601490221	2.2557308570	9.6435050200	-0.140211
N6	0.2626464810	2.3796997912	7.3672069680	-0.266640
C7	-3.6172325306	2.8831305014	7.8173992647	0.579791
C8	-4.7896842486	3.2807954753	7.1090107143	-0.324080
C9	-5.8694117713	3.8471317178	7.7591530679	-0.047528
C10	-5.8486729676	4.0833505955	9.1491641686	-0.245171
C11	-4.7123266463	3.7406696015	9.8591532790	-0.072603
C12	-3.5995502477	3.1378945098	9.2324498368	-0.196076
C13	-2.4080565926	2.9189768726	10.0053477352	0.048331
C14	-0.1745685826	2.2601091944	10.5059237350	0.148192
C15	1.1219522380	2.4726281024	9.6898107635	-0.107980
C16	0.8975298791	3.2323180757	8.3742294769	0.161868
C17	0.6809420434	2.4555928813	6.1525084196	0.102462
C18	0.1987208521	1.7051073136	5.0201514581	-0.132954
C19	0.7220455868	2.0354327269	3.7500837532	-0.123591
C20	0.3433044763	1.3685929041	2.5996644574	-0.196428
C21	-0.5864572687	0.3165849672	2.7148381502	-0.093168
C22	-1.1109887152	-0.0427680070	3.9405755757	-0.246738
C23	-0.7481104249	0.6330597845	5.1444340476	0.432703
H24	-4.8047675579	3.0995229968	6.0381260029	0.133750
H25	-6.7536936707	4.1166514011	7.1836528312	0.095042
H26	-6.7039746798	4.5288982780	9.6485540227	0.112284
H27	-4.6628901193	3.9260185424	10.9311931654	0.104323
H28	-2.3823423772	3.4033810899	10.9891279799	0.093730
H29	-0.2901363886	3.0662274291	11.2456909848	-0.013302
H30	-0.1491303136	1.3030589407	11.0326873028	0.046560
H31	1.8371973031	3.0347803816	10.3036234317	0.004091
H32	1.5594113785	1.5003319671	9.4539218080	0.074208
H33	1.8551991270	3.6150728710	7.9904921977	-0.000968
H34	0.2455091447	4.1001534820	8.5503208243	-0.008115
H35	1.4935758758	3.1620341897	5.9248136588	0.071796

Continued on next page

Table A2 – continued from previous page

Atom name	X	Y	Z	charge
H36	1.4443511377	2.8489990202	3.6914126627	0.098831
H37	0.7532384248	1.6460169472	1.6331223236	0.104299
H38	-0.8990281982	-0.2261920859	1.8243985817	0.102486
H39	-1.8258676877	-0.8551449147	4.0289797130	0.117100
Mn40	-1.5083911965	-1.1911336019	9.5018367804	1.280598
O41	-0.1996184903	-0.0462024909	8.7773822143	-0.733248
O42	-0.2165672247	-2.4714196359	10.1969151356	-0.742300
O43	-1.6584566408	-0.3629263482	11.4792046295	-0.609597
N44	-1.5465108133	-2.6755746716	7.7289447206	-0.370134
N45	-3.2934021033	-2.3373013492	10.0850092502	-0.390585
C46	0.7698578039	-3.0322544702	9.5600112020	0.682250
C47	-2.5604316130	-0.4373353296	12.3819892160	0.522615
C48	-0.5025617846	-3.3385355014	7.3934248299	0.211491
C49	-2.7149584227	-2.6654763118	6.8660575844	0.160823
C50	-3.8360201411	-3.2940494969	9.1135029284	0.198970
C51	-3.9867205050	-2.1520638342	11.1615754656	0.147253
C52	1.9836699322	-3.2873053198	10.2579009836	-0.343377
C53	0.7024235569	-3.4682698588	8.1921373105	-0.311696
C54	-3.7252453866	-1.2887065864	12.2840448176	-0.193872
C55	-2.4528851770	0.3476473001	13.5804393098	-0.287534
H56	-0.4896373416	-3.8826786492	6.4331814318	0.041200
C57	-4.0108367922	-2.6707396516	7.7144380528	-0.066050
H58	-2.6951908009	-3.5275252972	6.1769695126	-0.020310
H59	-2.6779659225	-1.7462310006	6.2728689065	0.024574
H60	-4.8069239522	-3.6710068825	9.4714588377	-0.022271
H61	-3.1519190794	-4.1507763577	9.0570885357	0.000250
H62	-4.9193445899	-2.7300633732	11.2627436159	0.063982
C63	3.0600301716	-3.9070020747	9.6479794504	-0.062617
H64	2.0302823503	-2.9611877852	11.2930734224	0.138708
C65	1.8128113258	-4.1090366921	7.6091053012	-0.048869
C66	-4.6770832597	-1.3005367770	13.3304828942	-0.100866
C67	-3.4057904389	0.3045678474	14.5763698821	-0.078883
H68	-1.5754554078	0.9832412694	13.6694965519	0.100746
H69	-4.7899801937	-3.2245748518	7.1748084972	-0.014345
H70	-4.3584027099	-1.6426262836	7.8463690737	0.062223
C71	2.9894965218	-4.3263775342	8.3081696783	-0.246078
H72	3.9741133143	-4.0698373886	10.2169313546	0.095301

Continued on next page

Table A2 – continued from previous page

Atom name	X	Y	Z	charge
H73	1.7307531364	-4.4340550146	6.5722352105	0.087155
C74	-4.5424744636	-0.5249363833	14.4668893822	-0.231167
H75	-5.5452586397	-1.9512931345	13.2231237840	0.089201
H76	-3.2767550330	0.9228324580	15.4641143772	0.095303
H77	3.8372162984	-4.8112164689	7.8319517415	0.106905
H78	-5.2896286634	-0.5531515036	15.2548558957	0.103605

Table A3: list of the optimized coordinates and atomic partial charges of the deprotonated 3,5-H-SOZMPU in the oxidation state MnIV-MnIV

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.2423627533	1.0812452501	8.1106806134	1.257564
O2	-2.4636323407	2.2182496045	7.3622006389	-0.597074
O3	-1.0093466169	0.1554167133	6.5111530436	-0.576044
O4	-2.4868488801	-0.0251016550	8.7859848260	-0.708325
N5	-1.2559264908	2.1902250459	9.8007871329	-0.248732
N6	0.3159938181	2.3682902744	7.5181283352	-0.301236
C7	-3.5395190420	2.6683897207	7.9945753063	0.525985
C8	-4.7056201628	2.9246523683	7.2504764285	-0.264973
C9	-5.8265928587	3.4560936829	7.8720357507	-0.045401
C10	-5.8207537275	3.7606087502	9.2439494298	-0.168634
C11	-4.6734107457	3.5334603484	9.9836908320	-0.070681
C12	-3.5237202053	2.9758597031	9.3843763629	-0.191336
C13	-2.3198492365	2.8338453128	10.1596442906	0.138379
C14	-0.0525480959	2.2287113771	10.6463610982	0.051969
C15	1.2303557053	2.4826959633	9.8234756008	-0.072690
C16	0.9731928797	3.2320424686	8.5144472031	0.065882
C17	0.6194896571	2.5429439262	6.2730111074	0.114394
C18	0.0722798629	1.8290559653	5.1504856089	-0.131915
C19	0.3884660663	2.2716121374	3.8478768656	-0.090303
C20	-0.0963251845	1.6140354309	2.7311296727	-0.163479
C21	-0.9032941726	0.4782225961	2.9047908038	-0.040657
C22	-1.2178958465	0.0113643973	4.1728870908	-0.248264
C23	-0.7397160458	0.6727067112	5.3214201754	0.425537
H24	-4.6969115666	2.6993415593	6.1892021909	0.158896
H25	-6.7211399941	3.6423836826	7.2841242569	0.132889
H26	-6.7041894910	4.1788126023	9.7148553329	0.139309
H27	-4.6479091998	3.7765374833	11.0430925252	0.120071
H28	-2.3048055725	3.3433667945	11.1278201803	0.106036
H29	-0.1860311077	3.0204697058	11.3937183444	0.055746
H30	0.0210737305	1.2671989457	11.1598948589	0.061488
H31	1.9256400129	3.0679993933	10.4339681781	0.050995
H32	1.7182551773	1.5316608665	9.5962184888	0.053060
H33	1.9183547802	3.6075979623	8.1048814683	0.056216
H34	0.3251873630	4.1002301804	8.6946089331	0.051707
H35	1.3600138032	3.3126465310	6.0274572421	0.117956

Continued on next page

Table A3 – continued from previous page

Atom name	X	Y	Z	charge
H36	1.0180595421	3.1507425759	3.7324665893	0.125162
H37	0.1439472273	1.9698248712	1.7349851528	0.138795
H38	-1.2898462206	-0.0452428069	2.0345394582	0.128974
H39	-1.8402784929	-0.8667558699	4.3090389677	0.129102
Mn40	-1.5388801084	-1.3791696313	9.5399211045	1.079171
O41	-0.0353751838	-0.2396161696	8.9925768196	-0.682552
O42	-0.3750073668	-2.7277930130	10.1748510873	-0.652789
O43	-1.5311753942	-0.6198718689	11.2382351938	-0.523082
N44	-1.5514868361	-2.4082309107	7.7876167478	0.025088
N45	-3.2214733207	-2.3633503548	9.9770724517	-0.199759
C46	0.6568276561	-3.2125554896	9.5197743181	0.581072
C47	-2.5658378832	-0.2870346835	11.9777263926	0.508552
C48	-0.5255674673	-3.0635692166	7.3442356897	-0.058142
C49	-2.8099804283	-2.3868390942	7.0431639263	-0.083193
C50	-3.5377843219	-3.5812606617	9.2167431604	0.026113
C51	-4.0563615797	-1.9417406745	10.8763653157	0.085203
C52	1.7851691210	-3.6631036332	10.2407107835	-0.322641
C53	0.6657405006	-3.3421018107	8.0939221883	-0.120896
C54	-3.8548993501	-0.8710310002	11.8028358620	-0.137263
C55	-2.3882719752	0.6318662480	13.0373991515	-0.269218
H56	-0.5817355655	-3.4713652690	6.3296268025	0.142105
C57	-3.8496750515	-3.3061966510	7.7365979447	-0.018460
H58	-2.6384106291	-2.7135479867	6.0112423527	0.071561
H59	-3.1551014368	-1.3530598111	7.0328648480	0.104250
H60	-4.3897123089	-4.0887382622	9.6832668301	0.057277
H61	-2.6724467527	-4.2439135608	9.3068917121	0.054998
H62	-4.9994062570	-2.4877373527	10.9740499678	0.140126
C63	2.8810748461	-4.1835956945	9.5740238715	-0.029054
H64	1.7636954463	-3.5763116030	11.3219717253	0.176452
C65	1.7930655177	-3.9014616768	7.4457712581	-0.122196
C66	-4.9195345710	-0.4940994776	12.6542387245	-0.109371
C67	-3.4541558516	0.9878586445	13.8481047529	-0.053166
H68	-1.3954808529	1.0348411415	13.2065888632	0.141091
H69	-3.8922143258	-4.2793844906	7.2346467378	0.040185
H70	-4.8419366113	-2.8528188040	7.6428345319	0.040211
C71	2.8978756634	-4.3025511946	8.1698358986	-0.173534
H72	3.7444844884	-4.5088783269	10.1479691757	0.133912

Continued on next page

Table A3 – continued from previous page

Atom name	X	Y	Z	charge
H73	1.7761130689	-4.0105596746	6.3638608395	0.136341
C74	-4.7353299731	0.4386207237	13.6555852843	-0.159009
H75	-5.8925856839	-0.9576977303	12.5110179526	0.131528
H76	-3.2916768906	1.6988796254	14.6540042696	0.131227
H77	3.7650351470	-4.7175572242	7.6673085411	0.142829
H78	-5.5587011791	0.7270608227	14.3001261422	0.139524
H79	0.5302629308	-0.6607437091	8.3273637577	0.339146

Table A4: list of the optimized coordinates and atomic partial charges of the double protonated 3,5-H-SOZMPU in the oxidation state MnIV-MnIV

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.2687884809	1.2621526190	7.8704814969	0.895526
O2	-2.3568803234	2.3985482107	6.9748707049	-0.500416
O3	-0.9844493713	0.2329653757	6.3783103440	-0.449469
O4	-2.7633956021	0.1222611807	8.4511166481	-0.608458
N5	-1.4281179958	2.4102312863	9.5278443035	-0.138753
N6	0.3400072758	2.3539322350	7.4153609206	-0.152614
C7	-3.4618740632	2.9644831964	7.4640007479	0.424447
C8	-4.5155377195	3.2540881416	6.5828932634	-0.241083
C9	-5.6559288831	3.8927067742	7.0552998495	-0.038466
C10	-5.7748258223	4.2677274153	8.4044589973	-0.116135
C11	-4.7346963031	4.0050716567	9.2788035992	-0.143343
C12	-3.5703617708	3.3378562529	8.8348589050	-0.042831
C13	-2.4695586503	3.1529711348	9.7414164395	0.020073
C14	-0.3193159901	2.4093343477	10.5004601114	-0.062594
C15	1.0560413642	2.5009812475	9.8059715627	-0.000154
C16	0.9964441058	3.1933514446	8.4462728940	-0.032218
C17	0.7073433368	2.5135394135	6.1764535304	0.052361
C18	0.1883937257	1.8334480976	5.0310522856	-0.076624
C19	0.5782655051	2.2539478219	3.7391190364	-0.095561
C20	0.1214858550	1.5902333073	2.6149359995	-0.123227
C21	-0.7155437115	0.4715739019	2.7666897278	-0.039112
C22	-1.0966707512	0.0220071769	4.0250920539	-0.207711
C23	-0.6594373516	0.6995683707	5.1742862414	0.363138
H24	-4.4105280450	2.9779393111	5.5392444937	0.170195
H25	-6.4662725241	4.1084855244	6.3651667797	0.151555
H26	-6.6706419099	4.7694431663	8.7540497544	0.154171
H27	-4.8057463732	4.3071128230	10.3204946409	0.152411
H28	-2.5180485788	3.7060671609	10.6840239583	0.146532
H29	-0.4607925344	3.2588770341	11.1781592737	0.094687
H30	-0.3870895588	1.4835601300	11.0745844672	0.093382
H31	1.7364165845	3.0656810037	10.4506911295	0.065643
H32	1.4832164173	1.5030499322	9.6826266543	0.031615
H33	2.0094721982	3.4267314608	8.1035755343	0.089976
H34	0.4422914262	4.1380026347	8.5171589722	0.087372
H35	1.4975951724	3.2442780036	5.9840941428	0.153223

Continued on next page

Table A4 – continued from previous page

Atom name	X	Y	Z	charge
H36	1.2340913586	3.1145997135	3.6381921194	0.144982
H37	0.4094463714	1.9256034175	1.6245930853	0.152858
H38	-1.0723598021	-0.0541149884	1.8856886062	0.148495
H39	-1.7356362364	-0.8469283682	4.1380009073	0.134151
Mn40	-1.8268694656	-1.2597074301	9.5376578704	0.887347
O41	-0.3317820683	-0.1193057540	8.9567778998	-0.606972
O42	-0.7396975622	-2.3991753757	10.4301137461	-0.499141
O43	-2.1075840318	-0.2312139308	11.0304739003	-0.446722
N44	-1.6699978420	-2.4055400768	7.8785488250	-0.133298
N45	-3.4361142004	-2.3483446085	9.9941907185	-0.151595
C46	0.3658198702	-2.9626063498	9.9392291977	0.424150
C47	-2.4354996203	-0.6950284784	12.2346772584	0.364111
C48	-0.6272262803	-3.1455715122	7.6618726221	0.017961
C49	-2.7796035089	-2.4038969000	6.9069868223	-0.061727
C50	-4.0928192894	-3.1884433217	8.9637736078	-0.030469
C51	-3.8035598539	-2.5078046802	11.2330741168	0.052119
C52	1.4193783708	-3.2546422921	10.8195910163	-0.240928
C53	0.4744914068	-3.3313685323	8.5671890103	-0.042375
C54	-3.2850007522	-1.8276436375	12.3783088798	-0.077427
C55	-1.9991436940	-0.0162044497	13.3835221520	-0.208805
H56	-0.5787862316	-3.6964022980	6.7179492401	0.146577
C57	-4.1541407485	-2.4982530604	7.6029356789	-0.001623
H58	-2.6378738159	-3.2521554216	6.2278758226	0.094192
H59	-2.7135751016	-1.4771202034	6.3342553733	0.093277
H60	-5.1058590977	-3.4212301757	9.3069372845	0.089538
H61	-3.5388752802	-4.1332320348	8.8936484838	0.086565
H62	-4.5941538982	-3.2381907252	11.4251569552	0.153281
C63	2.5600202078	-3.8913244662	10.3452958711	-0.038420
H64	1.3140585446	-2.9819974453	11.8641140437	0.170241
C65	1.6390918670	-3.9967915542	8.1212933062	-0.142972
C66	-3.6778798489	-2.2453490746	13.6702232832	-0.094987
C67	-2.3828330316	-0.4634319601	14.6420320070	-0.038765
H68	-1.3582070789	0.8512588495	13.2704294098	0.134548
H69	-4.8336723724	-3.0660706777	6.9599319178	0.065806
H70	-4.5841613824	-1.5013408029	7.7246383332	0.031948
C71	2.6791635501	-4.2619521243	8.9949606507	-0.116422
H72	3.3703076678	-4.1091924141	11.0348225698	0.151574

Continued on next page

Table A4 – continued from previous page

Atom name	X	Y	Z	charge
H73	1.7103231762	-4.2956036026	7.0786802025	0.152382
C74	-3.2219749902	-1.5804500668	14.7940765643	-0.123506
H75	-4.3353477324	-3.1046866146	13.7715078027	0.144941
H76	-2.0263906389	0.0628349482	15.5228582291	0.148522
H77	3.5751038697	-4.7623999113	8.6438765865	0.154217
H78	-3.5123568264	-1.9136082508	15.7844610178	0.152884
H79	0.3374049725	-0.5853311549	8.4324685221	0.340665
H80	-3.4342230692	0.5897221439	8.9719231107	0.341283

Table A5: list of the optimized coordinates and atomic partial charges of the protonated 3,5-H-SOZMPU in the oxidation state MnIV-MnIII

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.2253478206	1.1945257704	7.8966854501	1.289582
O2	-2.4783540914	2.3306537124	7.0651032167	-0.636084
O3	-0.9060480515	0.2582889116	6.2770746924	-0.565024
O4	-2.4644287127	0.1076207464	8.4826642063	-0.695592
N5	-1.3500329180	2.3330834263	9.5539308570	-0.216183
N6	0.2843066940	2.6073568049	7.3107144229	-0.304911
C7	-3.5765923386	2.7383466955	7.6580173593	0.535598
C8	-4.7292152879	2.9959728106	6.8811556297	-0.282452
C9	-5.8903862843	3.4657645833	7.4718303287	-0.059530
C10	-5.9538345445	3.7212530335	8.8536601647	-0.187888
C11	-4.8281926363	3.5053566924	9.6273978471	-0.090197
C12	-3.6364198567	3.0059399245	9.0599881232	-0.169677
C13	-2.4593650655	2.9078252336	9.8806929587	0.082237
C14	-0.1797625079	2.4340179317	10.4316988856	0.088725
C15	1.0974531244	2.8289477688	9.6486182347	-0.073247
C16	0.8129681933	3.5357168480	8.3168286539	0.098648
C17	0.6631739562	2.7455328949	6.0875009747	0.117407
C18	0.2366718075	1.9469104082	4.9636047854	-0.162722
C19	0.6381118948	2.3562120930	3.6736101748	-0.093926
C20	0.2893317185	1.6403944575	2.5426042963	-0.183221
C21	-0.4697212061	0.4678065097	2.6931923847	-0.071715
C22	-0.8678995583	0.0319889108	3.9457432732	-0.240621
C23	-0.5340217690	0.7542435960	5.1168967731	0.446581
H24	-4.6721887523	2.8036443775	5.8145027923	0.147862
H25	-6.7671215443	3.6414433487	6.8531070004	0.116130
H26	-6.8701281835	4.0904180012	9.3031285499	0.122803
H27	-4.8501123301	3.7133272746	10.6950317921	0.117953
H28	-2.5066414684	3.4042496837	10.8556352041	0.110813
H29	-0.3972580071	3.1808084367	11.2067010139	0.028882
H30	-0.0473271170	1.4626884167	10.9124457774	0.050303
H31	1.7013018698	3.4971827163	10.2731845312	0.027313
H32	1.6882206797	1.9330791179	9.4471229257	0.044632
H33	1.7299510567	4.0098205747	7.9411184406	0.029232
H34	0.0708717491	4.3320642812	8.4666370840	0.029639
H35	1.3762470705	3.5460908890	5.8484865856	0.094847

Continued on next page

Table A5 – continued from previous page

Atom name	X	Y	Z	charge
H36	1.2310620545	3.2641154765	3.5805876212	0.109669
H37	0.5978403004	1.9766279224	1.5578997975	0.122751
H38	-0.7513419444	-0.1087204233	1.8153024897	0.118760
H39	-1.4513887379	-0.8753268959	4.0655218104	0.119619
Mn40	-1.6233793215	-1.2093580140	9.5241916098	1.021869
O41	0.0805445473	0.1112031888	8.7620889425	-0.735537
O42	-0.4467175341	-2.4358616971	10.3951768949	-0.648283
O43	-1.8441948949	-0.2444917675	11.1575269794	-0.511101
N44	-1.6267447575	-2.4717545040	7.8784265161	-0.239095
N45	-3.5489761883	-2.3047682772	10.0346464851	-0.298100
C46	0.5084453384	-3.1316359367	9.8373995874	0.535216
C47	-2.5034015501	-0.5333643758	12.2508595506	0.459581
C48	-0.6626181894	-3.2950101529	7.6285169162	0.083453
C49	-2.7577450701	-2.3944714071	6.9443101813	0.008177
C50	-4.1871600443	-3.1092837517	8.9963012654	0.101022
C51	-4.0902706073	-2.2245554529	11.1973783664	0.126503
C52	1.6226326810	-3.5265065657	10.6213349268	-0.291287
C53	0.4767400088	-3.5565821338	8.4699664652	-0.135245
C54	-3.5977340881	-1.4584042247	12.3261551109	-0.157396
C55	-2.1449442909	0.1497229592	13.4426925772	-0.283589
H56	-0.7070807178	-3.8731763183	6.6969071658	0.097012
C57	-4.1357311584	-2.3680181115	7.6501672172	-0.006904
H58	-2.7020940230	-3.2533792448	6.2602038256	0.039970
H59	-2.6347759410	-1.4755513465	6.3648677466	0.081043
H60	-5.2305449859	-3.3365315769	9.2628864032	0.014161
H61	-3.6546653968	-4.0683804114	8.9192176929	0.029751
H62	-5.0195286691	-2.7829043831	11.3926719042	0.075679
C63	2.6527876272	-4.2705211233	10.0755945623	-0.068911
H64	1.6376950061	-3.2111490919	11.6599359638	0.150635
C65	1.5352711397	-4.3367156086	7.9514084948	-0.138590
C66	-4.2704226256	-1.6310624931	13.5553133101	-0.104620
C67	-2.8207516990	-0.0556986590	14.6326941469	-0.061634
H68	-1.3102359079	0.8420534872	13.3841711982	0.124575
H69	-4.8713447520	-2.8098625199	6.9674938378	0.004211
H70	-4.4271803260	-1.3292659589	7.8256143737	0.052187
C71	2.6230250287	-4.6829693190	8.7293472760	-0.179679
H72	3.4998290238	-4.5425741212	10.7009565900	0.117352

Continued on next page

Table A5 – continued from previous page

Atom name	X	Y	Z	charge
H73	1.4820586045	-4.6591300598	6.9130636995	0.119610
C74	-3.8992782747	-0.9525917738	14.7026291621	-0.195965
H75	-5.1048802717	-2.3295201808	13.5882644156	0.106246
H76	-2.5096558009	0.4869052961	15.5226754539	0.109967
H77	3.4366113737	-5.2692122985	8.3140874183	0.120621
H78	-4.4300673879	-1.1104992473	15.6362433444	0.118749
H79	0.5851471539	-0.3285489041	8.0634425326	0.341352

Table A6: list of the optimized coordinates and atomic partial charges of the deprotonated 3,5-Cl-SOZMPU in the oxidation state MnIV-MnIV

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.4715757873	1.1793254342	8.0396570148	1.505377
O2	-2.7629113310	2.5103092353	7.5066457104	-0.702481
O3	-1.5242497113	0.4830505963	6.2495059336	-0.663992
O4	-2.7365843568	0.1048536560	8.7355185774	-0.788791
N5	-1.3085930299	2.2307108455	9.7577304237	-0.133437
N6	0.1464822978	2.3806910580	7.4286246474	-0.392630
C7	-3.7335305715	2.9043864097	8.2765696984	0.688747
C8	-4.9856488024	3.3000045820	7.7224484421	-0.110819
C9	-6.0318317666	3.7407867918	8.5129661645	-0.053956
C10	-5.8663228502	3.8348664961	9.9040246936	0.009871
C11	-4.6642608915	3.5022274359	10.4925923201	-0.020037
C12	-3.6006100588	3.0294591760	9.6962833609	-0.412687
C13	-2.3154637245	2.8210946087	10.3103235218	0.163130
C14	0.0287145930	2.3059033628	10.3502579915	0.006074
C15	0.8849925728	3.3602869287	9.6076511355	-0.040528
C16	0.4288131046	3.6160082047	8.1625903619	0.097337
C17	0.9302455505	2.0487610151	6.4614747589	0.249049
C18	0.7181455039	0.9909803599	5.5078248981	-0.314221
C19	1.7490135921	0.7550326338	4.5751293976	-0.081353
C20	1.5423010885	-0.0906394033	3.5063598384	0.038584
C21	0.2797451237	-0.6633963354	3.3017449618	-0.070304
C22	-0.7398175726	-0.4282400573	4.2078303318	-0.078528
C23	-0.5559569296	0.3474125593	5.3950832811	0.588689
Cl24	-5.1987493640	3.2047680754	5.9914399134	-0.096501
H25	-6.9757442246	4.0140735000	8.0554774140	0.143723
Cl26	-7.2089094853	4.3997213874	10.8866777965	-0.134831
H27	-4.5345782473	3.5967811764	11.5660704881	0.126063
H28	-2.1712247071	3.2444082503	11.3095546987	0.092592
H29	-0.0492687405	2.5540601977	11.4111498194	0.058388
H30	0.4680684298	1.3126784964	10.2590418210	0.070957
H31	0.8471001680	4.3213971528	10.1344400601	0.018689
H32	1.9296183972	3.0302917862	9.6170573780	0.043429
H33	1.1874102533	4.2074192166	7.6341063098	0.029569
H34	-0.4991820655	4.1963138300	8.1614918092	0.035315
H35	1.8206186412	2.6633589037	6.2839322928	0.096716

Continued on next page

Table A6 – continued from previous page

Atom name	X	Y	Z	charge
H36	2.7049613148	1.2544967682	4.6981713751	0.131127
Cl37	2.8328969999	-0.4169235278	2.3622416673	-0.130598
H38	0.0922383890	-1.2775122519	2.4283646802	0.137255
Cl39	-2.3278367334	-1.0759410385	3.8466465024	-0.104076
Mn40	-1.6174466340	-1.1809039147	9.3694281495	1.501875
O41	-0.3517435687	-0.1060312164	8.6725720806	-0.789743
O42	-0.3277981473	-2.5130192985	9.9014364436	-0.701662
O43	-1.5649134527	-0.4847751199	11.1589596406	-0.664381
N44	-1.7809632842	-2.2320301343	7.6502840712	-0.135406
N45	-3.2365414658	-2.3799689110	9.9789930069	-0.391733
C46	0.6431618877	-2.9073376991	9.1317190957	0.689951
C47	-2.5348517577	-0.3456766735	12.0110010078	0.590386
C48	-0.7749395436	-2.8238140259	7.0981199459	0.163329
C49	-3.1182678702	-2.3069109536	7.0579508974	0.007909
C50	-3.5188986122	-3.6157586146	9.2459390050	0.098887
C51	-4.0218218200	-2.0457821813	10.9440736071	0.248847
C52	1.8949949725	-3.3037862204	9.6859816172	-0.111731
C53	0.5101828209	-3.0326808932	7.7120845263	-0.411896
C54	-3.8101495971	-0.9865794304	11.8960622949	-0.316021
C55	-2.3520438550	0.4313302357	13.1977087397	-0.079275
H56	-0.9197885005	-3.2479368237	6.0993571563	0.092679
C57	-3.9746750495	-3.3609267194	7.8007994348	-0.039965
H58	-3.0404646811	-2.5551992027	5.9971557534	0.058108
H59	-3.5574200676	-1.3135884435	7.1489653669	0.070537
H60	-4.2781610089	-4.2062908834	9.7744414718	0.029074
H61	-2.5912386429	-4.1965959141	9.2479046296	0.034726
H62	-4.9130446491	-2.6592852198	11.1210517747	0.096940
C63	2.9407058587	-3.7459736854	8.8953678364	-0.054526
Cl64	2.1081911705	-3.2088253944	11.4171093102	-0.096358
C65	1.5732445301	-3.5065202366	6.9157285784	-0.020851
C66	-4.8429150133	-0.7468102512	12.8256973940	-0.079687
C67	-3.3734330991	0.6699595812	14.1008599788	-0.069689
Cl68	-0.7632991224	1.0757334137	13.5619323209	-0.103887
H69	-3.9366524214	-4.3223977342	7.2746553925	0.018472
H70	-5.0193291318	-3.0309886668	7.7910013546	0.042908
C71	2.7749857738	-3.8403414128	7.5043357374	0.010501
H72	3.8843668475	-4.0201276935	9.3528356489	0.144014

Continued on next page

Table A6 – continued from previous page

Atom name	X	Y	Z	charge
H73	1.4433712418	-3.6010606485	5.8422685337	0.126295
C74	-4.6368679947	0.0998825347	13.8937437952	0.037617
H75	-5.7995876654	-1.2445430320	12.7012508743	0.130886
H76	-3.1866545492	1.2846306684	14.9739844163	0.137208
Cl77	4.1167455741	-4.4070510793	6.5216354400	-0.134858
Cl78	-5.9299751826	0.4312452833	15.0336441249	-0.130394

Table A7: list of the optimized coordinates and atomic partial charges of the deprotonated 3,5-Cl-SOZMPU in the oxidation state MnIV-MnIII

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.6740718840	1.4331757648	7.7595194488	1.417606
O2	-3.0283922321	2.7183480842	7.1628274502	-0.735952
O3	-1.3787175313	0.5880820403	5.7690201391	-0.656522
O4	-2.9873863980	0.3437027313	8.5303972377	-0.761641
N5	-1.6142664522	2.8895054548	9.5349428255	-0.408360
N6	0.0512700960	2.6451785254	7.1809941176	-0.477550
C7	-4.0354411189	3.1002707272	7.8864332365	0.705327
C8	-5.3243660180	3.2736025372	7.2943692957	-0.093617
C9	-6.4363530756	3.7001336481	8.0022893415	-0.116080
C10	-6.3085195254	3.9936463425	9.3630985026	0.044874
C11	-5.0857725300	3.8662564297	9.9964094014	-0.071859
C12	-3.9557081457	3.4241825519	9.2869485825	-0.419890
C13	-2.6916984719	3.3808856134	10.0131233988	0.306656
C14	-0.3499316246	2.9451255363	10.2358885775	0.157411
C15	0.6774078705	3.7352277613	9.3808241201	-0.028275
C16	0.2858746298	3.9002961495	7.9005507434	0.215140
C17	0.9403198367	2.3020687196	6.3098322000	0.292331
C18	0.9101179653	1.2029530008	5.3765846360	-0.324491
C19	2.0909804345	0.9887233350	4.6343614214	-0.125427
C20	2.1444667148	0.0259580655	3.6503342996	0.089046
C21	1.0058577099	-0.7366102132	3.3486247521	-0.170957
C22	-0.1597370090	-0.5244717348	4.0550212290	0.009048
C23	-0.2894236068	0.4308254714	5.1315400975	0.537416
Cl24	-5.5071643388	2.9159719407	5.5847205053	-0.143648
H25	-7.3924003474	3.8010082705	7.5017151237	0.141170
Cl26	-7.7285788669	4.5380616092	10.2668027094	-0.191388
H27	-4.9906557004	4.1057302082	11.0515823829	0.128801
H28	-2.7004847169	3.8259802066	11.0208425966	0.023721
H29	-0.4488447013	3.3948283316	11.2344130604	0.015654
H30	-0.0201664455	1.9102324327	10.3590227663	0.039596
H31	0.7926297825	4.7511504983	9.7818135347	-0.014203
H32	1.6591765331	3.2526871350	9.4504522835	0.015706
H33	1.0741637263	4.4795343540	7.3965650012	-0.029111
H34	-0.6412279773	4.4785137644	7.8368099515	0.015158
H35	1.8364029416	2.9339170771	6.2210814625	0.060794

Continued on next page

Table A7 – continued from previous page

Atom name	X	Y	Z	charge
H36	2.9675299285	1.5943797895	4.8470923382	0.123207
Cl37	3.6343046160	-0.2491606098	2.7386869281	-0.192801
H38	1.0367939601	-1.4811957222	2.5610687571	0.145067
Cl39	-1.5771627735	-1.4738063947	3.6112506612	-0.177590
Mn40	-1.8807615686	-0.8205789060	9.2657078802	1.166174
O41	-0.5757115672	0.1605696589	8.6080202330	-0.725488
O42	-0.6596284444	-2.2002738638	10.0422606403	-0.599100
O43	-1.8393062434	-0.0087502538	11.0378544117	-0.494350
N44	-1.9826308323	-1.9937331301	7.6195688688	-0.204336
N45	-3.5290277220	-2.0519473997	9.9263932032	-0.363368
C46	0.2588048965	-2.8140334795	9.3866303116	0.555489
C47	-2.2645968684	-0.4851469666	12.1578868695	0.458898
C48	-1.0118119818	-2.7534959261	7.2371351023	0.183387
C49	-3.1939206561	-1.9129417728	6.7896148960	0.187828
C50	-4.2803237070	-2.8271603452	8.9328751504	0.198237
C51	-3.8858499553	-2.1467168854	11.1562988186	0.250130
C52	1.4146919795	-3.3395479555	10.0505081224	-0.053505
C53	0.1820548622	-3.0767305501	7.9772702123	-0.315648
C54	-3.2727557903	-1.5004555694	12.2941222836	-0.273156
C55	-1.7590843344	0.0288161621	13.3996755366	-0.000860
H56	-1.1108702858	-3.2419408861	6.2623750609	0.067003
C57	-4.4805245132	-2.0359517430	7.6358556577	-0.202206
H58	-3.1506855786	-2.7194685770	6.0469772606	0.015042
H59	-3.1690924491	-0.9597588520	6.2591839998	0.044286
H60	-5.2528390266	-3.1244121481	9.3504939805	0.009409
H61	-3.7178047513	-3.7493085952	8.7272978972	-0.013850
H62	-4.7372545727	-2.7960862180	11.4040978049	0.056965
C63	2.4138362402	-4.0211616736	9.3838744639	-0.106871
Cl64	1.5578245598	-3.0881896776	11.7793907560	-0.137125
C65	1.1971436398	-3.7909124819	7.3078077119	-0.089536
C66	-3.7362922331	-1.9116085686	13.5603342635	-0.130195
C67	-2.2096329336	-0.3927357053	14.6365165841	-0.122778
Cl68	-0.4945888469	1.2426367905	13.3588193745	-0.149603
H69	-5.2468892590	-2.5388154545	7.0334832341	0.032269
H70	-4.8441581165	-1.0387285616	7.8897083042	0.103929
C71	2.2995873157	-4.2408862133	8.0021831530	0.050849
H72	3.2781753106	-4.3852271177	9.9275050703	0.137781

Continued on next page

Table A7 – continued from previous page

Atom name	X	Y	Z	charge
H73	1.1063840679	-3.9796879819	6.2424781291	0.138779
C74	-3.2110221451	-1.3683888138	14.7121499023	0.078446
H75	-4.5097667688	-2.6714916983	13.6212509489	0.129418
H76	-1.7843276107	0.0314536631	15.5386731718	0.137047
Cl77	3.5814252320	-5.1237338694	7.1676412452	-0.180187
Cl78	-3.7891479516	-1.8956358178	16.2952432244	-0.183572

Table A8: list of the optimized coordinates and atomic partial charges of the protonated 3,5-Cl-SOZMPU in the oxidation state MnIV-MnIV

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.0607257633	0.9362429415	8.0780705063	1.255356
O2	-2.2248328495	2.1356518068	7.3364110778	-0.555397
O3	-0.8730379575	0.0691023576	6.4197903116	-0.532536
O4	-2.3483537033	-0.1261378751	8.7665305011	-0.696026
N5	-1.0335520231	2.0579873103	9.7674971502	-0.281519
N6	0.5466640161	2.1869330490	7.5041025004	-0.366800
C7	-3.2853356573	2.6145317174	7.9573937695	0.545755
C8	-4.4519516259	2.9117881637	7.2157589098	-0.086182
C9	-5.5656377654	3.4699037094	7.8265538364	-0.030653
C10	-5.5358444115	3.7603950117	9.1992318918	0.016873
C11	-4.4016638551	3.5097130892	9.9523774716	-0.033442
C12	-3.2742963644	2.9249802259	9.3439837889	-0.302777
C13	-2.0666117813	2.7499322807	10.1186682129	0.185329
C14	0.1730289182	2.0564284722	10.6169419123	0.254572
C15	1.4745629096	2.2119353474	9.8066978773	-0.145353
C16	1.2864425783	2.9823135376	8.5039166098	0.095205
C17	0.8016147771	2.4214669995	6.2615633391	0.244813
C18	0.1445046169	1.8281491302	5.1239272955	-0.261717
C19	0.3804736024	2.4211527238	3.8692088369	-0.054191
C20	-0.2448434408	1.9259623997	2.7391138365	0.013072
C21	-1.0998015032	0.8239150583	2.8433036749	-0.012197
C22	-1.3173779501	0.2227978275	4.0756230322	-0.011368
C23	-0.7047864891	0.6929588853	5.2649579330	0.400074
Cl24	-4.4635413859	2.5711400145	5.5071787224	-0.062857
H25	-6.4536569924	3.6830771750	7.2424233588	0.149083
Cl26	-6.9490928524	4.4580258593	9.9523036175	-0.084129
H27	-4.3834934187	3.7570809553	11.0088654496	0.126102
H28	-2.0263504717	3.2781669116	11.0759226177	0.112394
H29	0.0758812732	2.8728430078	11.3411138703	-0.000364
H30	0.1867646500	1.1136340984	11.1643027046	0.008829
H31	2.2039135867	2.7416833975	10.4277469208	0.066902
H32	1.8947468825	1.2285992597	9.5830459112	0.062828
H33	2.2593751901	3.2704356056	8.0897366451	0.060133
H34	0.7200272441	3.9056976672	8.6856122318	0.043499
H35	1.5693534884	3.1656179994	6.0238444424	0.102678

Continued on next page

Table A8 – continued from previous page

Atom name	X	Y	Z	charge
H36	1.0408286823	3.2791101512	3.7935739051	0.136031
Cl37	0.0210135222	2.6677741960	1.1824519839	-0.079962
H38	-1.5975605577	0.4339213706	1.9628788441	0.140831
Cl39	-2.3981472867	-1.1462975100	4.1272887216	-0.114890
Mn40	-1.4613336334	-1.4704045186	9.6037208046	1.185204
O41	0.0747676338	-0.3995717435	9.0044042307	-0.688671
O42	-0.3058114047	-2.7985854025	10.3158621558	-0.640593
O43	-1.5267542520	-0.6709305109	11.3093784799	-0.551105
N44	-1.4480276625	-2.5721547988	7.9098978087	-0.281777
N45	-3.1493167182	-2.4431578283	10.0390634691	-0.233340
C46	0.7335612812	-3.2985907796	9.6993286733	0.617614
C47	-2.5754013517	-0.3532354293	12.0226467903	0.568796
C48	-0.4172696147	-3.2435646963	7.5076042601	0.173283
C49	-2.7110031174	-2.6271503084	7.1686125648	-0.002436
C50	-3.4140247876	-3.7330181392	9.3790348011	0.004358
C51	-4.0404463353	-1.9627601024	10.8486486793	0.168400
C52	1.8725068707	-3.7254762578	10.4374932337	-0.124051
C53	0.7727863961	-3.4818846970	8.2810163953	-0.254602
C54	-3.8807918564	-0.8646777574	11.7526037925	-0.266118
C55	-2.4543146159	0.4665064731	13.1821432097	-0.119655
H56	-0.4663245257	-3.7049293512	6.5168890367	0.107227
C57	-3.6958937712	-3.5863943821	7.8778295630	0.036669
H58	-2.5208787000	-2.9575958260	6.1453420216	0.105311
H59	-3.1065110369	-1.6124988045	7.1356232883	0.078926
H60	-4.2569804390	-4.2263765669	9.8743470509	0.058854
H61	-2.5296556032	-4.3536770011	9.5465522012	0.080523
H62	-4.9920187754	-2.4972883851	10.9152485315	0.143121
C63	2.9822802049	-4.2592544284	9.8044166325	-0.027216
Cl64	1.8603265532	-3.5507067109	12.1673950139	-0.049980
C65	1.9012160104	-4.0513265950	7.6493801347	-0.099950
C66	-5.0013091171	-0.4726065173	12.5181959651	-0.073565
C67	-3.5604394003	0.8553239988	13.9198826153	-0.020033
Cl68	-0.8731035973	0.9639541511	13.7331752276	-0.061339
H69	-3.6398523809	-4.5878326180	7.4368822831	0.027059
H70	-4.7158846412	-3.2213938837	7.7186374475	0.027664
C71	2.9959331058	-4.4148040691	8.4063943101	0.019099
H72	3.8429463444	-4.5588806996	10.3916780765	0.152504

Continued on next page

Table A8 – continued from previous page

Atom name	X	Y	Z	charge
H73	1.9039429050	-4.1995727974	6.5741923876	0.150207
C74	-4.8443986204	0.4085989001	13.5665880583	0.020920
H75	-5.9808140164	-0.8785592664	12.2870355607	0.139158
H76	-3.4273658870	1.4935050767	14.7863755945	0.141175
Cl77	4.4130366393	-5.0889122867	7.6397127498	-0.082442
Cl78	-6.2216223900	0.9345835098	14.5017206092	-0.081415
H79	0.6676207817	-0.8678384551	8.3971968385	0.344214

Table A9: list of the optimized coordinates and atomic partial charges of the protonated 3,5-Cl-SOZMPU in the oxidation state MnIV-MnIII

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.0210533537	1.0097379190	8.0045957802	1.426086
O2	-2.2452754453	2.2477379052	7.2848009594	-0.622146
O3	-0.8533206880	0.1590683863	6.2981732932	-0.595604
O4	-2.2987237265	-0.0128753160	8.6544676213	-0.747800
N5	-0.9732620981	2.0989730679	9.6945317831	-0.334379
N6	0.5368477464	2.3616787357	7.3837870423	-0.374932
C7	-3.2807295709	2.6818532072	7.9525581828	0.558448
C8	-4.4878387766	2.9993566306	7.2758255782	-0.053219
C9	-5.5885038805	3.5066036098	7.9461412997	-0.083292
C10	-5.5171538590	3.7339432699	9.3285204159	0.053208
C11	-4.3582347586	3.4632542810	10.0293587372	-0.069727
C12	-3.2418291829	2.9273198515	9.3572938715	-0.343478
C13	-2.0147123659	2.7470532870	10.0963412052	0.277787
C14	0.2536536938	2.1007045440	10.5067396887	0.209270
C15	1.5143896265	2.4011827523	9.6653117642	-0.198162
C16	1.2209112405	3.1904376951	8.3876497389	0.135531
C17	0.8139135585	2.5685643474	6.1452198080	0.244487
C18	0.2250121106	1.8990454533	5.0075473349	-0.289716
C19	0.5252084447	2.4327485148	3.7384265139	-0.082533
C20	0.0083435214	1.8550409053	2.5958873263	0.043148
C21	-0.8070121187	0.7221469060	2.6915667525	-0.074224
C22	-1.0936375758	0.1844541549	3.9368633781	-0.004561
C23	-0.5974674191	0.7392772842	5.1518888411	0.461833
Cl24	-4.5714126497	2.7239485543	5.5525407667	-0.096398
H25	-6.4996071026	3.7263455154	7.4016947566	0.144062
Cl26	-6.9215473535	4.3797648426	10.1608336355	-0.132961
H27	-4.3073405296	3.6511412425	11.0970216939	0.135025
H28	-1.9642503844	3.2300015817	11.0762588963	0.068105
H29	0.1334336556	2.8519477726	11.2946452789	0.034892
H30	0.3356118381	1.1219879299	10.9823597884	0.032075
H31	2.2121294916	2.9802319903	10.2802066175	0.057592
H32	2.0058013454	1.4645921812	9.3955316896	0.082324
H33	2.1530092267	3.5883849638	7.9662194906	0.040045
H34	0.5724450265	4.0478739024	8.6145777562	0.020801
H35	1.5577127376	3.3362134853	5.8973777253	0.080077

Continued on next page

Table A9 – continued from previous page

Atom name	X	Y	Z	charge
H36	1.1601882902	3.3101740996	3.6641538301	0.131986
Cl37	0.3690517379	2.5305004754	1.0166778189	-0.131338
H38	-1.2177367843	0.2632346758	1.7999366735	0.144545
Cl39	-2.1155021355	-1.2334054600	3.9990425093	-0.139524
Mn40	-1.5076406903	-1.3868812854	9.6372063891	1.292383
O41	0.2366508790	-0.1782395790	8.7666933590	-0.805312
O42	-0.2822986165	-2.6647019987	10.3917905177	-0.703711
O43	-1.7588721402	-0.5162200994	11.3499727889	-0.599659
N44	-1.5430991866	-2.6104115993	7.9622898624	-0.335228
N45	-3.4247254496	-2.5184629417	10.1061227799	-0.375125
C46	0.7063890156	-3.2128831207	9.7483425686	0.634937
C47	-2.7957493448	-0.3627822206	12.1189629806	0.510353
C48	-0.5338812743	-3.3135153218	7.5793147470	0.240995
C49	-2.7871969960	-2.5993191454	7.1920302070	-0.056452
C50	-3.7156224408	-3.7321903015	9.3523641185	0.156921
C51	-4.2319609311	-2.1590294099	11.0369731930	0.209076
C52	1.8989322982	-3.5796775236	10.4410939861	-0.094330
C53	0.6731586600	-3.5197702160	8.3481525594	-0.315324
C54	-4.0368249446	-1.0752283928	11.9832743401	-0.270867
C55	-2.7383303973	0.5365065773	13.2301814480	-0.035376
H56	-0.5878512886	-3.8231358562	6.6102284393	0.069354
C57	-3.8785583881	-3.4900604631	7.8422592184	0.055311
H58	-2.5941460499	-2.9329475721	6.1677775966	0.117678
H59	-3.1146630422	-1.5595801868	7.1539011382	0.079965
H60	-4.6163199271	-4.2338263984	9.7367784707	-0.010542
H61	-2.8700476623	-4.4114571846	9.5157257727	0.029370
H62	-5.1570623725	-2.7340899409	11.1899187439	0.087950
C63	2.9777712428	-4.1622517582	9.7999815275	-0.081773
Cl64	1.9952200030	-3.2554946949	12.1546201685	-0.095743
C65	1.7703976961	-4.1345620164	7.7081124829	-0.108338
C66	-5.1161782556	-0.8290529509	12.8557102591	-0.107433
C67	-3.8067948354	0.7742026672	14.0782600502	-0.105529
Cl68	-1.2301946993	1.3711305636	13.5710677349	-0.134585
H69	-3.8898925309	-4.4761505245	7.3620811106	-0.009892
H70	-4.8531840546	-3.0302676053	7.6442356420	0.014925
C71	2.9097036459	-4.4340638714	8.4248828774	0.041738
H72	3.8719673377	-4.4061200790	10.3619996652	0.147795

Continued on next page

Table A9 – continued from previous page

Atom name	X	Y	Z	charge
H73	1.7136779419	-4.3662808525	6.6488954974	0.140253
C74	-5.0123184004	0.0954509782	13.8753160617	0.056175
H75	-6.0390981375	-1.3856258559	12.7242177863	0.129767
H76	-3.6997343483	1.4746993572	14.8983975324	0.142954
Cl77	4.2915587334	-5.1715473865	7.6271745213	-0.139403
Cl78	-6.3727512622	0.4052376102	14.9438907179	-0.140805
H79	0.7038136556	-0.6251892933	8.0470696698	0.360191

Table A10: list of the optimized coordinates and atomic partial charges of the deprotonated 3,5-NO_2-SOZMPU in the oxidation state MnIV-MnIV before the charge modeling

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.4052465826	1.1837077984	8.0620037532	1.416976
O2	-2.6491666599	2.5523042869	7.5020551793	-0.622787
O3	-1.4612234424	0.4478003482	6.2809408665	-0.598327
O4	-2.7282653690	0.1394121275	8.7060811400	-0.803109
N5	-1.2444969682	2.2319356376	9.7736892228	-0.085949
N6	0.1951521462	2.3986908191	7.4187812432	-0.337573
C7	-3.6241492480	2.9701339027	8.2333413047	0.606768
C8	-4.8660674144	3.3825125049	7.6603518951	-0.275808
C9	-5.9418800317	3.7756320044	8.4291580134	-0.078829
C10	-5.7971210098	3.8558959473	9.8163944120	-0.091858
C11	-4.5994826262	3.5306328597	10.4341659640	-0.072782
C12	-3.5232025447	3.0685138395	9.6676158667	-0.183347
C13	-2.2491485907	2.8425051078	10.3084885906	-0.001616
C14	0.0927686123	2.2912358325	10.3802375355	0.065936
C15	0.9819770124	3.3046266007	9.6145638251	-0.095061
C16	0.4972119323	3.6098628357	8.1914744341	0.080306
C17	0.8985779930	2.1643500904	6.3670150949	0.156593
C18	0.6753264349	1.1228600974	5.3938512490	-0.147084
C19	1.6429685358	0.9514014934	4.3974644223	-0.116160
C20	1.4568413878	0.0223827282	3.3864978661	-0.053485
C21	0.2835555230	-0.7288606882	3.3158661589	-0.117082
C22	-0.6996007766	-0.5191017977	4.2620274322	-0.213902
C23	-0.5484891329	0.3695460444	5.3770769384	0.518874
H24	-6.8853762657	4.0265536587	7.9603526162	0.177494
H25	-4.5146693300	3.6356453794	11.5106055313	0.138738
H26	-2.1168765800	3.2605923221	11.3113340814	0.162964
H27	0.0146036445	2.5850120722	11.4294150872	0.019357
H28	0.5082432492	1.2857416228	10.3258937221	0.090876
H29	1.0136414132	4.2582108467	10.1535504927	0.058672
H30	2.0063984429	2.9192809040	9.5849443446	0.056698
H31	1.2488670397	4.2123567795	7.6671665909	0.047181
H32	-0.4265751231	4.1960856603	8.2232262715	0.045311
H33	1.7354608746	2.8346623759	6.1429492290	0.129333
H34	2.5543560054	1.5398406787	4.4034143522	0.150526

Continued on next page

Table A10 – continued from previous page

Atom name	X	Y	Z	charge
H35	0.1429193054	-1.4523431656	2.5216686039	0.181595
Mn36	-1.6785285336	-1.1890454040	9.3478972346	1.421788
O37	-0.3553260680	-0.1445761473	8.7044942795	-0.802904
O38	-0.4382467048	-2.5598699310	9.9104534640	-0.624608
O39	-1.6269321225	-0.4534297308	11.1294072079	-0.598688
N40	-1.8351521222	-2.2363182684	7.6356837840	-0.086229
N41	-3.2830276888	-2.4000395059	9.9861260654	-0.338219
C42	0.5383541526	-2.9793094877	9.1822868722	0.606349
C43	-2.5464224265	-0.3667784319	12.0254955735	0.518360
C44	-0.8277882728	-2.8436522610	7.1022434990	-0.002672
C45	-3.1709073993	-2.2976396230	7.0261093621	0.068219
C46	-3.5818053755	-3.6132537719	9.2155942779	0.079923
C47	-3.9932648696	-2.1595642832	11.0317839120	0.159038
C48	1.7757509519	-3.4002122460	9.7594155808	-0.275146
C49	0.4443150224	-3.0700057709	7.7469088478	-0.181844
C50	-3.7737023842	-1.1141291495	12.0017259049	-0.147896
C51	-2.3999626160	0.5263607013	13.1375136018	-0.212748
H52	-0.9562865016	-3.2593521743	6.0978834664	0.162586
C53	-4.0608588475	-3.3116284940	7.7897765022	-0.095504
H54	-3.0898995035	-2.5920094916	5.9773725799	0.018540
H55	-3.5879508963	-1.2927230594	7.0787099844	0.090330
H56	-4.3355784061	-4.2144390145	9.7383183299	0.047510
H57	-2.6577322311	-4.1992688549	9.1891943373	0.045235
H58	-4.8335484872	-2.8265896481	11.2528096904	0.128554
C59	2.8552190867	-3.7889213467	8.9934589842	-0.079120
C60	1.5243546085	-3.5277221415	6.9828623654	-0.073377
C61	-4.7491269606	-0.9326808171	12.9884724589	-0.116610
C62	-3.3902886832	0.7456533120	14.0740891305	-0.118109
H63	-4.0881945594	-4.2661989857	7.2523124831	0.058872
H64	-5.0861795066	-2.9283278884	7.8141682764	0.056532
C65	2.7185497467	-3.8576468845	7.6047386908	-0.091590
H66	3.7955300353	-4.0452992551	9.4657003729	0.177522
H67	1.4449366487	-3.6259182710	5.9053617333	0.139039
C68	-4.5669816869	0.0005443531	13.9964363192	-0.053612
H69	-5.6635335911	-1.5163022792	12.9772925496	0.150969
H70	-3.2527074992	1.4721627326	14.8660256817	0.181801
N71	2.4900507718	-0.1597411986	2.3679379352	0.825333

Continued on next page

Table A10 – continued from previous page

Atom name	X	Y	Z	charge
O72	2.2756245220	-0.9909751535	1.4870189892	-0.449650
O73	3.5039830972	0.5322455116	2.4615167020	-0.457272
N74	3.8554538114	-4.3048772755	6.8007425802	0.829474
O75	3.7059326285	-4.3401799616	5.5790170256	-0.458526
O76	4.8853412208	-4.6152889299	7.3977267150	-0.452700
N77	-6.9294126986	4.3090898465	10.6233210563	0.829517
O78	-6.7738806164	4.3515194809	11.8440410006	-0.458306
O79	-7.9620010323	4.6171240296	10.0297863818	-0.452802
N80	-5.6084633647	0.1936973034	15.0045026377	0.825996
O81	-5.3975882158	1.0284606135	15.8829241583	-0.449626
O82	-6.6252480318	-0.4932751327	14.9052010246	-0.457520
N83	1.9494758593	-3.4042707216	11.2188648153	0.922736
O84	3.0684115864	-3.1157539491	11.6442792188	-0.472546
O85	0.9845673294	-3.7242970027	11.9048388355	-0.466018
N86	-1.9366584425	-1.2837961724	4.1206871509	0.877626
O87	-2.9997082456	-0.6950013452	4.2660821561	-0.424049
O88	-1.8238457325	-2.4929627336	3.8700458240	-0.485494
N89	-5.0485491657	3.3724259296	6.2020788433	0.923502
O90	-6.1677825556	3.0709094718	5.7864900202	-0.472708
O91	-4.0903570371	3.6943276396	5.5076460497	-0.466243
N92	-1.1605023298	1.2861483054	13.2851282355	0.877745
O93	-1.2698484380	2.4961100636	13.5338141853	-0.485968
O94	-0.0989631362	0.6931416036	13.1461889951	-0.424229

Table A11: list of the optimized coordinates and atomic partial charges of the deprotonated 3,5-NO_2-SOZMPU in the oxidation state MnIV-MnIII before the charge modeling

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.2001252466	0.9183583345	8.2526524173	1.286237
O2	-2.4078193684	2.3006446765	7.4569972244	-0.572206
O3	-1.2571600548	0.0938651721	6.4937579122	-0.578124
O4	-2.5174724400	-0.0645992145	8.8922682731	-0.762133
N5	-1.1565074349	2.0801099537	9.8918859314	-0.149675
N6	0.4263360897	2.1436055506	7.5755799216	-0.332837
C7	-3.3910716220	2.8215151529	8.0760594060	0.527157
C8	-4.5750837820	3.2731300315	7.3943511891	-0.240152
C9	-5.6773832159	3.7607146078	8.0642260235	-0.102344
C10	-5.6264238757	3.9145225056	9.4517447093	-0.097642
C11	-4.4844189379	3.5686766253	10.1677311563	-0.079870
C12	-3.3861803877	3.0114011572	9.5119189857	-0.142107
C13	-2.1794194953	2.7673816048	10.2749569573	0.006357
C14	0.0588972420	2.0502074552	10.7230664896	0.103798
C15	1.3354595859	2.2670245188	9.8774693915	-0.096211
C16	1.0886799248	3.0099230758	8.5587831212	0.128489
C17	0.8690918648	2.1481863720	6.3706093135	0.126995
C18	0.3511272326	1.3914229241	5.2491896362	-0.110368
C19	0.9540084312	1.6020825146	4.0074587559	-0.122512
C20	0.5314764081	0.9111292561	2.8783629883	-0.077092
C21	-0.5257455788	-0.0012788373	2.9553834876	-0.144451
C22	-1.1512605108	-0.1883417297	4.1667575419	-0.180559
C23	-0.7269127539	0.4424517279	5.3874399600	0.510656
H24	-6.5693103634	4.0336223015	7.5150771848	0.166649
H25	-4.4650130437	3.7334670259	11.2399679952	0.138049
H26	-2.1347002691	3.2390425145	11.2619159498	0.148348
H27	-0.0323098649	2.8239967023	11.4936030689	0.016763
H28	0.0960597707	1.0774614794	11.2148409163	0.083165
H29	2.0521471497	2.8428352289	10.4733429682	0.034713
H30	1.7900533997	1.3004141492	9.6491076497	0.073874
H31	2.0385210510	3.3782867419	8.1483859100	0.025126
H32	0.4421310663	3.8817175512	8.7264275363	0.024159
H33	1.7251923281	2.7907698768	6.1285174114	0.106434
H34	1.7710884401	2.3081994723	3.9059058935	0.140504

Continued on next page

Table A11 – continued from previous page

Atom name	X	Y	Z	charge
H35	-0.8535463228	-0.5385920921	2.0742887989	0.184312
Mn36	-1.4386208520	-1.3986026144	9.6389717023	1.308340
O37	-0.1066088460	-0.2684684315	8.9610090760	-0.776657
O38	-0.1026655524	-2.7569901789	10.1378312706	-0.666924
O39	-1.6054930788	-0.6043805566	11.6577661022	-0.578623
N40	-1.6225526408	-2.7820732735	7.8373089741	-0.127136
N41	-3.1595376341	-2.5855194454	10.2852771525	-0.396310
C42	0.8474503933	-3.1745648014	9.3850388724	0.643980
C43	-2.6527094297	-0.3877709104	12.3244309316	0.484649
C44	-0.5897485325	-3.2982989364	7.2917662629	0.012909
C45	-2.9327222925	-2.8030634600	7.2149412889	0.032457
C46	-3.4405666404	-3.8298782929	9.5569968830	0.173825
C47	-3.9769148752	-2.2614189310	11.2263001148	0.180909
C48	2.1521289972	-3.4486994471	9.9231927277	-0.291323
C49	0.7063734349	-3.4054349021	7.9587060878	-0.222283
C50	-3.8842082848	-1.1447453012	12.1477431697	-0.143581
C51	-2.7406482940	0.6682422806	13.3173307561	-0.189476
H52	-0.6447221443	-3.7195116937	6.2765936330	0.123359
C53	-3.8994391509	-3.6365920228	8.0987223056	-0.068141
H54	-2.9027138859	-3.2012334642	6.1914231521	0.022834
H55	-3.2714728387	-1.7656592378	7.1680062483	0.087413
H56	-4.2074528071	-4.4096377533	10.0900247357	-0.002204
H57	-2.5152245847	-4.4137784765	9.5720394301	0.018971
H58	-4.8607290086	-2.8949937614	11.3836788383	0.093173
C59	3.2323660647	-3.8131093925	9.1445840863	-0.092528
C60	1.7915317270	-3.8294107062	7.1974222116	-0.059878
C61	-5.0151791269	-0.8650728712	12.9116455484	-0.127592
C62	-3.8916056475	0.9792385652	14.0077544354	-0.140325
H63	-4.0085453182	-4.6450227826	7.6798661658	0.025634
H64	-4.8934034594	-3.1760587004	8.0806426040	0.038688
C65	3.0459913762	-4.0131334481	7.7772138197	-0.124811
H66	4.2073801055	-3.9493004822	9.5942130763	0.164802
H67	1.6753022256	-4.0202532775	6.1356731853	0.130990
C68	-5.0288367512	0.1885379538	13.8235893670	-0.079541
H69	-5.9165294427	-1.4588575841	12.7983491707	0.138487
H70	-3.9130034833	1.8104241164	14.7017125444	0.173878
N71	1.1888711688	1.1387093434	1.6041487733	0.838677

Continued on next page

Table A11 – continued from previous page

Atom name	X	Y	Z	charge
O72	0.7901537816	0.4974224638	0.6278667567	-0.481547
O73	2.1115512154	1.9595486757	1.5690560785	-0.487277
N74	4.1666868068	-4.4331018785	6.9566531917	0.841673
O75	3.9643143610	-4.6155165484	5.7511868443	-0.490821
O76	5.2621787307	-4.5864212949	7.5064007191	-0.486471
N77	-6.7740898058	4.4778645543	10.1437881985	0.827496
O78	-6.6941174761	4.6230266633	11.3664699599	-0.473603
O79	-7.7613138001	4.7837074307	9.4687715925	-0.478014
N80	-6.2211893423	0.4626180524	14.5978579270	0.834336
O81	-6.1837725628	1.3934359824	15.4093642796	-0.485231
O82	-7.2108825208	-0.2537656387	14.4082897735	-0.491391
N83	2.3961238200	-3.3201446971	11.3629952552	0.928991
O84	3.4985989034	-2.8817921500	11.7086631920	-0.502325
O85	1.5146125887	-3.6890536501	12.1323462079	-0.475581
N86	-2.3038821721	-1.0907810006	4.1996943673	0.860157
O87	-3.3039529144	-0.7217476783	4.8055607371	-0.431721
O88	-2.1999698149	-2.1630021298	3.5940102631	-0.494972
N89	-4.6634721854	3.2097146408	5.9326439266	0.913030
O90	-5.7829220942	3.0320699806	5.4437866783	-0.487312
O91	-3.6334730384	3.3762021886	5.2849752897	-0.487072
N92	-1.5697948203	1.4918631049	13.5867405729	0.857018
O93	-1.7428559796	2.7197277921	13.6765915472	-0.508470
O94	-0.4827381097	0.9376493978	13.7121614541	-0.451041

Table A12: list of the optimized coordinates and atomic partial charges of the protonated 3,5-NO_2-SOZMPU in the oxidation state MnIV-MnIV before the charge modeling

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.3556240163	1.2830176288	8.0697658602	1.377900
O2	-2.5901311747	2.5212184393	7.5196371643	-0.561532
O3	-1.3063867072	0.4537597126	6.3662396998	-0.617843
O4	-2.6039735905	0.1647595919	8.7403983685	-0.774834
N5	-1.2210080996	2.3016897341	9.7897237779	-0.165618
N6	0.2073714338	2.5350256981	7.4302792905	-0.365404
C7	-3.6217603042	2.8876833258	8.2398663686	0.536539
C8	-4.8645765510	3.1853637482	7.6277019267	-0.217597
C9	-5.9700409055	3.5633246466	8.3668373118	-0.091932
C10	-5.8403136962	3.7145427551	9.7478634077	-0.041333
C11	-4.6409561747	3.4781019071	10.3969970282	-0.100685
C12	-3.5308877777	3.0393634944	9.6594237594	-0.138721
C13	-2.2517660495	2.8815797557	10.3124121098	0.040036
C14	0.1182528544	2.4304483732	10.3936718483	0.097489
C15	0.9542184831	3.4875484185	9.6300059127	-0.103244
C16	0.4573251927	3.7623748802	8.2069414134	0.027940
C17	0.8891508174	2.3684822743	6.3459901966	0.174777
C18	0.6884605847	1.3501256360	5.3471837815	-0.130193
C19	1.6029704959	1.2769901253	4.2830028819	-0.133101
C20	1.4057861907	0.3713028110	3.2578955109	-0.031568
C21	0.2760948563	-0.4484911579	3.2246524236	-0.081197
C22	-0.6499962472	-0.3351955279	4.2425140396	-0.284600
C23	-0.4672179776	0.5108793066	5.3740118280	0.526331
H24	-6.9212686410	3.7427343569	7.8792012449	0.192250
H25	-4.5813350150	3.6323544496	11.4691583620	0.145946
H26	-2.1391193651	3.3213321387	11.3078318885	0.174764
H27	0.0190898063	2.7193100810	11.4425331742	0.033038
H28	0.5845285514	1.4476497185	10.3472722880	0.056064
H29	0.9324279386	4.4418405833	10.1668909326	0.079900
H30	1.9980220027	3.1586755950	9.6084160473	0.071011
H31	1.1796626728	4.3949291179	7.6794571816	0.077306
H32	-0.4908625020	4.3088599787	8.2323460412	0.078094
H33	1.6782797120	3.0918002060	6.1185680901	0.143257
H34	2.4759328527	1.9208490258	4.2510888218	0.162333

Continued on next page

Table A12 – continued from previous page

Atom name	X	Y	Z	charge
H35	0.1301544876	-1.1455189927	2.4069556821	0.191586
Mn36	-1.6844405491	-1.2536285141	9.3971515295	1.188120
O37	-0.1774968997	-0.1244049058	8.8153579162	-0.693156
O38	-0.5116404139	-2.6410855564	9.9402480926	-0.575584
O39	-1.5806096894	-0.5217339092	11.1243822978	-0.557871
N40	-1.8251195310	-2.2344731896	7.6448692505	-0.036845
N41	-3.2985244666	-2.3100107857	9.9573988287	-0.215901
C42	0.4608655740	-3.1332453158	9.2336269772	0.497046
C43	-2.4973612317	-0.3537040124	12.0238389985	0.544232
C44	-0.8405921164	-2.8962542187	7.1263959770	-0.128193
C45	-3.1533475522	-2.2300597425	7.0101985540	0.065473
C46	-3.6358136526	-3.5235837745	9.1856227060	-0.044792
C47	-4.0192727246	-2.0249995587	10.9941403240	0.123117
C48	1.6445163729	-3.6297552691	9.8489991481	-0.224278
C49	0.3987218361	-3.2032538471	7.7988054110	-0.020830
C50	-3.7647236081	-1.0152140451	11.9799386285	-0.130858
C51	-2.2881566395	0.5256349157	13.1254106923	-0.243721
H52	-0.9665534122	-3.2829362921	6.1106552603	0.215691
C53	-4.0850400923	-3.2195266045	7.7527930075	-0.045063
H54	-3.0650401327	-2.5133531736	5.9595075408	0.015495
H55	-3.5318084214	-1.2106023652	7.0698275866	0.085617
H56	-4.4192047112	-4.0761696962	9.7141108750	0.090760
H57	-2.7397284565	-4.1504856554	9.1897619283	0.087955
H58	-4.8999205322	-2.6466813255	11.1744161363	0.161178
C59	2.7181008169	-4.0883464543	9.1119184288	-0.097499
C60	1.4729642593	-3.7293223972	7.0600566239	-0.153575
C61	-4.7338893291	-0.7854096775	12.9703992022	-0.124370
C62	-3.2634884405	0.7944104471	14.0664962202	-0.091878
H63	-4.1219187686	-4.1755288918	7.2196891367	0.074129
H64	-5.1010579984	-2.8129600180	7.7531402949	0.053428
C65	2.6164401832	-4.1385861126	7.7180077612	-0.038420
H66	3.6275419618	-4.4070790891	9.6085731515	0.195406
H67	1.4244754797	-3.8177036370	5.9796269533	0.154539
C68	-4.4836381918	0.1209442745	13.9830079142	-0.039129
H69	-5.6865325561	-1.3047033829	12.9555057487	0.167272
H70	-3.0862920131	1.5065103421	14.8648053392	0.191305
N71	2.3963267923	0.2742916481	2.1742890824	0.812297

Continued on next page

Table A12 – continued from previous page

Atom name	X	Y	Z	charge
O72	2.1764746302	-0.5499056830	1.2923076343	-0.425831
O73	3.3685366259	1.0215800274	2.2411154154	-0.430149
N74	3.7540660065	-4.6460340417	6.9343169676	0.816980
O75	3.6311669325	-4.6452247134	5.7118675017	-0.434885
O76	4.7376568376	-5.0250761174	7.5601802793	-0.422865
N77	-7.0132998270	4.1402555348	10.5307038231	0.812460
O78	-6.8672995116	4.2237345181	11.7475353239	-0.433274
O79	-8.0439987306	4.3732886465	9.9090788048	-0.419819
N80	-5.5174050831	0.3779058738	14.9984170373	0.818467
O81	-5.2481519727	1.1980054500	15.8701748406	-0.424643
O82	-6.5687102436	-0.2470800050	14.8903378861	-0.432276
N83	1.7664637580	-3.6360500702	11.3182612734	0.905507
O84	2.8814488835	-3.3993744421	11.7747523261	-0.447697
O85	0.7575664715	-3.8991330677	11.9622605804	-0.445194
N86	-1.8635465711	-1.1506212127	4.1709479316	0.929153
O87	-2.9374259915	-0.5894213244	4.3350812268	-0.414101
O88	-1.7139213887	-2.3618257309	3.9670093820	-0.485967
N89	-5.0200498497	3.0764705281	6.1657265107	0.902432
O90	-6.1127727437	2.6871860574	5.7633589234	-0.445867
O91	-4.0625402782	3.4007876154	5.4727800549	-0.446051
N92	-1.0070947184	1.2238895355	13.2402178626	0.895218
O93	-1.0590017146	2.4365111858	13.4860982482	-0.476223
O94	0.0179495551	0.5825755155	13.0599743267	-0.413407
H95	0.4294707718	-0.5370637511	8.1832511784	0.309774

Table A13: list of the optimized coordinates and atomic partial charges of the protonated 3,5-NO_2-SOZMPU in the oxidation state MnIV-MnIII before the charge modeling

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-0.8962713272	0.9533549308	8.1163744625	1.533677
O2	-2.1064358869	2.1943477620	7.3400652949	-0.578666
O3	-0.6775599412	0.0647621009	6.4202201695	-0.689605
O4	-2.2262190139	-0.0353699228	8.7128691210	-0.830293
N5	-0.8532792623	2.0675094892	9.7696117174	-0.196090
N6	0.6333070676	2.3083175328	7.4688700959	-0.360933
C7	-3.1169567836	2.7087123884	7.9601438115	0.456366
C8	-4.3191154343	3.0622820191	7.2765551679	-0.162807
C9	-5.4114881567	3.5981529823	7.9338972829	-0.123720
C10	-5.3238739169	3.8615295547	9.3003864034	-0.062920
C11	-4.1706787143	3.5756280275	10.0155433002	-0.083728
C12	-3.0800821381	2.9832402222	9.3710003211	-0.129402
C13	-1.8694136362	2.7746487447	10.1329683055	0.021585
C14	0.3737213585	2.0563687693	10.5891979338	0.084873
C15	1.6345714840	2.3544290328	9.7442826937	-0.069441
C16	1.3467535465	3.1353048801	8.4606477343	0.056471
C17	0.8074756032	2.5743031073	6.2226946650	0.155435
C18	0.1917293775	1.8891389027	5.1049841542	-0.208242
C19	0.3275822726	2.4655078868	3.8368211923	-0.040000
C20	-0.2051897762	1.8423284542	2.7184756246	-0.068956
C21	-0.8340857589	0.6024316059	2.8235209466	-0.089895
C22	-0.9122111346	0.0006840312	4.0644017123	-0.245682
C23	-0.4746339265	0.6294638838	5.2707701935	0.573030
H24	-6.3218412222	3.8147299784	7.3887549328	0.188637
H25	-4.1319846240	3.8103161880	11.0740828330	0.136496
H26	-1.8174710654	3.2789430060	11.1015256718	0.181899
H27	0.2520150816	2.7969101673	11.3862988409	0.027827
H28	0.4542223976	1.0721102754	11.0513750231	0.076588
H29	2.3304704725	2.9373075150	10.3563813337	0.056106
H30	2.1309529021	1.4173460574	9.4839696832	0.050992
H31	2.2836333865	3.5024415295	8.0249020027	0.058496
H32	0.7220015861	4.0116169320	8.6803696582	0.056077
H33	1.4857193090	3.3908261958	5.9483800829	0.123883
H34	0.8402818319	3.4131210894	3.7105236772	0.135708

Continued on next page

Table A13 – continued from previous page

Atom name	X	Y	Z	charge
H35	-1.2378532427	0.1132264386	1.9453316434	0.176783
Mn36	-1.5205304862	-1.4850988830	9.6439230705	1.268590
O37	0.3024734489	-0.2590226651	8.8871822770	-0.829435
O38	-0.3346349206	-2.8146987516	10.3737470175	-0.645480
O39	-1.7530995670	-0.6143648417	11.3628902264	-0.586343
N40	-1.6085461481	-2.6494152989	7.9673851320	-0.131514
N41	-3.4474242996	-2.5913301080	10.1390186045	-0.351655
C42	0.6309642644	-3.3765319239	9.7291755155	0.509749
C43	-2.7855642132	-0.4033765303	12.1019939586	0.502510
C44	-0.6266743390	-3.3725002588	7.5547726661	-0.015592
C45	-2.8778881273	-2.6050254488	7.2315592143	0.080186
C46	-3.7498828134	-3.8018210955	9.3793892778	0.126794
C47	-4.2301036703	-2.2456118766	11.0929784019	0.143391
C48	1.8275068662	-3.7886560161	10.3954511901	-0.233742
C49	0.5882516964	-3.6178900676	8.3068251063	-0.059391
C50	-4.0300234736	-1.1277878083	12.0031805570	-0.136303
C51	-2.7621159768	0.6285375269	13.1004641068	-0.221726
H52	-0.7081105618	-3.8603378391	6.5806162986	0.169785
C53	-3.9367197965	-3.5384136039	7.8742724189	-0.103961
H54	-2.7112133728	-2.8798699526	6.1862142875	0.030573
H55	-3.2145036048	-1.5674013810	7.2674085600	0.064820
H56	-4.6422676144	-4.3068774837	9.7753754271	0.026790
H57	-2.9002213785	-4.4792296287	9.5253301915	0.036033
H58	-5.1417647126	-2.8292071489	11.2819728420	0.111485
C59	2.9169156702	-4.3026554486	9.7218835817	-0.107456
C60	1.6748769345	-4.1981763980	7.6437037743	-0.174081
C61	-5.1036432048	-0.8119911089	12.8390031867	-0.116138
C62	-3.8508232053	0.9730716941	13.8790500832	-0.115657
H63	-3.9254726791	-4.5141238353	7.3748525427	0.045750
H64	-4.9248991641	-3.1017106566	7.6931352468	0.059123
C65	2.8298021824	-4.5119480224	8.3436064645	-0.070406
H66	3.8263781368	-4.5432442171	10.2585464225	0.181153
H67	1.6329939122	-4.3972937707	6.5781677656	0.180367
C68	-5.0220449874	0.2299157758	13.7523295887	-0.069204
H69	-6.0292295580	-1.3748291243	12.7800126913	0.145856
H70	-3.7930151841	1.7926844683	14.5850942689	0.182970
N71	-0.0811689059	2.4761394053	1.4063402439	0.816744

Continued on next page

Table A13 – continued from previous page

Atom name	X	Y	Z	charge
O72	-0.5491440929	1.8748279859	0.4408837217	-0.445758
O73	0.4865555068	3.5675987583	1.3567507350	-0.457581
N74	3.9734846580	-5.0762717913	7.6306973100	0.839959
O75	3.8590117196	-5.2431948053	6.4155781269	-0.471699
O76	4.9776835699	-5.3451486219	8.2885453094	-0.458756
N77	-6.4675613328	4.4632664009	9.9883863057	0.821157
O78	-6.3518827286	4.6763886875	11.1947551494	-0.448757
O79	-7.4647349251	4.7156664565	9.3150785254	-0.449582
N80	-6.1734570517	0.5445256552	14.5927636263	0.832770
O81	-6.0561864287	1.4722996606	15.3925040750	-0.458541
O82	-7.1875015164	-0.1396470618	14.4460349536	-0.463841
N83	1.9491858185	-3.6382314752	11.8530318300	0.914052
O84	3.0591679558	-3.3246087541	12.2862156621	-0.478883
O85	0.9546900615	-3.8644173568	12.5326074669	-0.452892
N86	-1.4768870774	-1.3462227504	4.1172551706	0.910689
O87	-2.4821455985	-1.5800496193	3.4528212677	-0.445295
O88	-0.8778819715	-2.1749512826	4.8077692953	-0.493593
N89	-4.4531616706	2.8344265929	5.8305253845	0.849230
O90	-5.5878918526	2.6135699291	5.4076187888	-0.457150
O91	-3.4390885669	2.8971822017	5.1402538429	-0.442251
N92	-1.5367272158	1.4009648561	13.2963642165	0.884348
O93	-1.6524937472	2.6347590538	13.3678504705	-0.502170
O94	-0.4746032624	0.7966898382	13.3784991280	-0.432665
H95	0.8081853363	-0.6769341973	8.1749165658	0.382074

Table A14: list of the optimized coordinates and atomic partial charges of the deprotonated 5-Cl-SOZMPU in the oxidation state MnIV-MnIV

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.4778542865	1.1897924232	8.0291044499	1.315010
O2	-2.7789454945	2.5158976914	7.5009395270	-0.634137
O3	-1.5292959191	0.4936166547	6.2361464888	-0.580022
O4	-2.7363193276	0.1055970439	8.7214746510	-0.695032
N5	-1.3243381637	2.2392537675	9.7475720616	-0.119509
N6	0.1387492758	2.3876167160	7.4169679238	-0.337424
C7	-3.7539601428	2.9040278509	8.2689710388	0.659402
C8	-5.0045486312	3.3000924964	7.7140612408	-0.159662
C9	-6.0528245910	3.7317606300	8.5080611326	0.007356
C10	-5.9090634053	3.8230488996	9.9034698438	-0.252411
C11	-4.6981018916	3.4871753493	10.4796324018	-0.017294
C12	-3.6241108366	3.0228044531	9.6910219385	-0.338075
C13	-2.3388917595	2.8165617044	10.3025741556	0.111097
C14	0.0084658516	2.3127281632	10.3483053676	0.034812
C15	0.8805824134	3.3522921343	9.6018442497	-0.058650
C16	0.4268710080	3.6164133649	8.1576220444	0.123050
C17	0.9224599421	2.0532719615	6.4491001096	0.170781
C18	0.7108015581	1.0011025111	5.4914228068	-0.229632
C19	1.7353073561	0.7586379989	4.5524928110	-0.081835
C20	1.5400963194	-0.0835071937	3.4753441628	-0.202087
C21	0.2695640299	-0.6470305675	3.2845529757	-0.039379
C22	-0.7522076840	-0.4130692319	4.1885052018	-0.089979
C23	-0.5653581435	0.3589880662	5.3773020717	0.522834
Cl24	-5.2092412814	3.2171242927	5.9770474372	-0.090814
H25	-6.9909049466	4.0025980954	8.0336045762	0.105567
H26	-6.7401830904	4.1650930414	10.5111649482	0.138588
H27	-4.5589320230	3.5702240123	11.5547879022	0.110531
H28	-2.1986353983	3.2286368765	11.3072731999	0.088349
H29	-0.0762996803	2.5728087594	11.4059045938	0.045883
H30	0.4421601901	1.3157255421	10.2725589084	0.062395
H31	0.8606794071	4.3144233230	10.1282657939	0.017639
H32	1.9198179224	3.0050723703	9.6080667267	0.038500
H33	1.1906179185	4.2061935362	7.6337808212	0.014722
H34	-0.4971558764	4.2033255445	8.1593971186	0.015625
H35	1.8172198951	2.6640066645	6.2774596479	0.097537

Continued on next page

Table A14 – continued from previous page

Atom name	X	Y	Z	charge
H36	2.6897392502	1.2625940298	4.6861157273	0.112281
H37	2.3357873552	-0.2787592978	2.7642215324	0.133435
H38	0.0656840399	-1.2625944951	2.4139814179	0.106068
Cl39	-2.3466957318	-1.0546471280	3.8259816816	-0.101060
Mn40	-1.6085999482	-1.1701478610	9.3604811134	1.311698
O41	-0.3496819630	-0.0891154552	8.6656561911	-0.698646
O42	-0.3202399064	-2.4983186468	9.9101157363	-0.628948
O43	-1.5560823664	-0.4686384202	11.1512708085	-0.582569
N44	-1.7556914110	-2.2273536466	7.6467130064	-0.135516
N45	-3.2286683305	-2.3629390938	9.9715455107	-0.336891
C46	0.6506808772	-2.9189816691	9.1555472483	0.653359
C47	-2.5130561232	-0.3450377144	12.0192698380	0.529148
C48	-0.7442143062	-2.8252945094	7.1075652344	0.112202
C49	-3.0831446436	-2.2911383285	7.0327118824	0.049083
C50	-3.5254341598	-3.5862292571	9.2245317130	0.128586
C51	-4.0058810280	-2.0352224760	10.9466814021	0.166944
C52	1.8857722152	-3.3385755176	9.7281000535	-0.158694
C53	0.5300370276	-3.0514784071	7.7340228319	-0.328239
C54	-3.7872830931	-0.9919371199	11.9121257597	-0.230176
C55	-2.3193373199	0.4206677386	13.2111821311	-0.092467
H56	-0.8804689051	-3.2428813508	6.1044647498	0.088342
C57	-3.9737090065	-3.3146522654	7.7802240806	-0.058619
H58	-2.9899218311	-2.5627152831	5.9786734822	0.042570
H59	-3.5069668268	-1.2891443692	7.0941923667	0.059026
H60	-4.2954985574	-4.1714855883	9.7442032907	0.012924
H61	-2.6068130449	-4.1816680361	9.2227639462	0.013974
H62	-4.9007185418	-2.6458231754	11.1182076256	0.098668
C63	2.9295968568	-3.8066584895	8.9499321177	0.006873
Cl64	2.0748266367	-3.2392708071	11.4658044575	-0.090796
C65	1.5995810861	-3.5522962531	6.9618659452	-0.022671
C66	-4.8038342366	-0.7618144434	12.8626357590	-0.080841
C67	-3.3332902750	0.6419623181	14.1270467045	-0.037972
Cl68	-0.7253392852	1.0702881598	13.5612045503	-0.101204
H69	-3.9695847967	-4.2775735059	7.2547939781	0.017036
H70	-5.0069871890	-2.9503776168	7.7733827090	0.037389
C71	2.7950326127	-3.9116765948	7.5547278422	-0.251411
H72	3.8563443790	-4.0959174849	9.4355983667	0.105819

Continued on next page

Table A14 – continued from previous page

Atom name	X	Y	Z	charge
H73	1.4691853597	-3.6464574744	5.8865035670	0.111765
C74	-4.6024669395	0.0724914110	13.9446433947	-0.205079
H75	-5.7569827733	-1.2694386742	12.7339139096	0.112208
H76	-3.1241059554	1.2521713317	15.0000971345	0.106138
H77	3.6227049712	-4.2828711011	6.9592430009	0.138298
H78	-5.3920260124	0.2572382699	14.6653503021	0.134232

Table A15: list of the optimized coordinates and atomic partial charges of the protonated 5-Cl-SOZMPU in the oxidation state MnIV-MnIV

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.0996256329	0.9723031449	8.0995048052	1.027232
O2	-2.2436493537	2.1001272154	7.2309655297	-0.479552
O3	-0.7876502872	0.0609648538	6.4774349868	-0.493926
O4	-2.3930618153	-0.0767066653	8.7960136406	-0.592444
N5	-1.2127971082	2.1518361574	9.7362148655	-0.246740
N6	0.4986740231	2.2555214296	7.5911940654	-0.280945
C7	-3.3413656494	2.6088670406	7.7529061387	0.517222
C8	-4.4461096862	2.8737744238	6.9125000703	-0.126778
C9	-5.5961985803	3.4643704168	7.4170813602	-0.001487
C10	-5.6765529312	3.8216272663	8.7718530698	-0.186608
C11	-4.5983125144	3.5943586081	9.6110062373	-0.065218
C12	-3.4280451093	2.9786964716	9.1231594595	-0.213650
C13	-2.2802814813	2.8367940023	9.9869421620	0.110021
C14	-0.0714625718	2.1928766878	10.6667381775	0.233890
C15	1.2792191849	2.3622023671	9.9429323042	-0.069782
C16	1.1579617592	3.0944524973	8.6106712865	0.020738
C17	0.8000201670	2.4805004797	6.3567712519	0.172562
C18	0.2342376183	1.8359287823	5.2013934763	-0.179686
C19	0.5093331453	2.4077716526	3.9439611801	-0.075262
C20	-0.0072964230	1.8586810032	2.7846291853	-0.200216
C21	-0.7871984208	0.7016776445	2.8741430736	0.011361
C22	-1.0454100859	0.1109571091	4.1037120473	-0.018238
C23	-0.5596402219	0.6566700438	5.3177369520	0.366276
Cl24	-4.3197172170	2.4476415719	5.2241987230	-0.051611
H25	-6.4310276753	3.6506772315	6.7492501668	0.123694
H26	-6.5808099524	4.2853169605	9.1513624983	0.149305
H27	-4.6466609255	3.8865021032	10.6566765081	0.124235
H28	-2.3164669485	3.3917931913	10.9294394802	0.125103
H29	-0.2360597519	3.0184871067	11.3687834221	-0.008829
H30	-0.0711557491	1.2587861907	11.2275128802	-0.008402
H31	1.9532191750	2.9242543396	10.5972054000	0.053162
H32	1.7337787802	1.3832551068	9.7724363814	0.041807
H33	2.1489003703	3.3920801342	8.2488890291	0.065500
H34	0.5640710407	4.0105435046	8.7330067802	0.052648
H35	1.5452913104	3.2544593514	6.1424889326	0.100801

Continued on next page

Table A15 – continued from previous page

Atom name	X	Y	Z	charge
H36	1.1248036163	3.3023614629	3.8987062908	0.129940
H37	0.1871333884	2.3106133765	1.8182529957	0.153700
H38	-1.2006328270	0.2471463041	1.9798440773	0.120485
Cl39	-2.0053126694	-1.3509099988	4.1184924065	-0.129779
Mn40	-1.5085451295	-1.4286760472	9.6320394640	0.942369
O41	0.0252210585	-0.3400830344	9.0600077021	-0.567472
O42	-0.3499495598	-2.7891596295	10.2677338234	-0.566573
O43	-1.4757715943	-0.6495082894	11.3476325627	-0.457770
N44	-1.5868127340	-2.4910172902	7.9207541674	0.009700
N45	-3.1886377329	-2.3847188975	10.1378534666	-0.158289
C46	0.6595890299	-3.2859630405	9.5985798706	0.547736
C47	-2.4408399369	-0.4158339581	12.1959514664	0.504884
C48	-0.5749082554	-3.1518719915	7.4562009179	0.045010
C49	-2.8720481088	-2.4921524652	7.2159221228	-0.280332
C50	-3.5319312373	-3.6275375042	9.4265012222	0.030276
C51	-4.0020103848	-1.9529429944	11.0530365309	0.079278
C52	1.8170160069	-3.7464984488	10.2819416419	-0.149975
C53	0.6398797976	-3.4282178348	8.1738294648	-0.165115
C54	-3.7612762497	-0.9311787407	12.0221679638	-0.174256
C55	-2.2009790497	0.3134238948	13.3960710264	-0.144669
H56	-0.6630265693	-3.5692984989	6.4491795878	0.101738
C57	-3.8861218054	-3.3996900225	7.9532230044	0.029329
H58	-2.7243835354	-2.8383730039	6.1921559579	0.197846
H59	-3.2213980609	-1.4606764494	7.1850893881	0.123818
H60	-4.3662986627	-4.1177122707	9.9397760718	0.045396
H61	-2.6636856188	-4.2867088348	9.5149634682	0.052579
H62	-4.9531680738	-2.4827676847	11.1566667220	0.141244
C63	2.8926416748	-4.2781993461	9.5887466408	0.005044
Cl64	1.8690551893	-3.6181937477	12.0182458155	-0.041857
C65	1.7402395750	-3.9961260970	7.4919203064	-0.104099
C66	-4.7915473290	-0.6176528121	12.9384323083	-0.092588
C67	-3.2279048803	0.6169659216	14.2751038365	-0.002285
Cl68	-0.5699112028	0.7953766944	13.8096447521	-0.045037
H69	-3.9336260932	-4.3850674637	7.4762965322	0.029214
H70	-4.8832384155	-2.9547561245	7.8706719799	0.040870
C71	2.8617797045	-4.4013953708	8.1885745600	-0.195197
H72	3.7650742643	-4.6030108560	10.1465053618	0.125186

Continued on next page

Table A15 – continued from previous page

Atom name	X	Y	Z	charge
H73	1.6902690045	-4.1087096606	6.4120359479	0.138286
C74	-4.5403283913	0.1763504306	14.0377715397	-0.175368
H75	-5.7836679970	-1.0288886855	12.7722928733	0.130002
H76	-2.9961717318	1.1889701388	15.1679611092	0.120292
H77	3.7138259391	-4.8244547099	7.6676826509	0.153423
H78	-5.3291568413	0.4249603354	14.7390969688	0.147924
H79	0.6313501067	-0.7971464549	8.4565508577	0.308904

Table A16: list of the optimized coordinates and atomic partial charges of the deprotonated 5-OCH_3-SOZMPU in the oxidation state MnIV-MnIV

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.4150279237	1.1606072436	8.0416326959	1.450285
O2	-2.6270619034	2.5307401970	7.4462934011	-0.598524
O3	-1.5054518045	0.4335066028	6.2736287130	-0.562263
O4	-2.7072851246	0.1011623006	8.6861777245	-0.837128
N5	-1.2854651078	2.2121913236	9.7611764914	-0.024769
N6	0.2230520842	2.3358142024	7.4341734044	-0.411171
C7	-3.6297359656	2.9574659125	8.1727217741	0.382701
C8	-4.8375336096	3.3627201413	7.5205152188	0.256219
C9	-5.8964182317	3.8835134995	8.2637866715	-0.200026
C10	-5.7391703874	4.0103545200	9.6362585301	-0.075991
C11	-4.6223654201	3.6581118379	10.3331892268	-0.143149
C12	-3.5425440991	3.1033460741	9.5913178394	-0.178153
C13	-2.2962322486	2.8453799906	10.2597573927	-0.013336
C14	0.0094736617	2.2008713900	10.4416391947	-0.065143
C15	1.0372830198	3.0834416069	9.6811783378	0.000251
C16	0.6044293684	3.4795677224	8.2622991792	0.117835
C17	0.8936706374	2.1131266016	6.3550916453	0.199684
C18	0.5980660964	1.1523245051	5.3275663419	-0.136413
C19	1.5174658825	1.0645439770	4.2390659189	-0.208822
C20	1.1745693395	0.2718047310	3.1925216333	-0.028413
C21	-0.0186152042	-0.4329577385	3.0789080233	-0.309842
C22	-0.9261162516	-0.3478549571	4.1312598643	0.381835
C23	-0.6226221662	0.4190971739	5.3149780587	0.263421
H25	-6.8038459368	4.1729096039	7.7410994158	0.159803
H27	-4.5344401254	3.7811058039	11.4099034017	0.133889
H28	-2.1842799847	3.2597671855	11.2669746635	0.136315
H29	-0.1182008437	2.5484747467	11.4699555451	0.065965
H30	0.3374273147	1.1629391479	10.4711784666	0.081641
H31	1.2114352726	4.0170767238	10.2303774058	-0.000662
H32	1.9961989533	2.5558078050	9.6351461164	0.036601
H33	1.4160495823	4.0412926888	7.7797851494	0.017016
H34	-0.2678041002	4.1405817519	8.3122025187	0.022837
H35	1.7646113129	2.7490836912	6.1532865429	0.093307
H36	2.4432405838	1.6336373965	4.2699787832	0.143460
H38	-0.2378841550	-1.0244750556	2.1963838195	0.165818

Continued on next page

Table A16 – continued from previous page

Atom name	X	Y	Z	charge
Mn40	-1.6021172129	-1.1856709610	9.3718630456	1.683326
O41	-0.3120769279	-0.1266342883	8.7131576828	-0.809372
O42	-0.3653354622	-2.5198055025	9.9995641726	-0.709717
O43	-1.6010569440	-0.4365259896	11.1214786553	-0.678892
N44	-1.6742463884	-2.2479305361	7.6551948978	-0.072221
N45	-3.2333673220	-2.3925530390	9.9314538589	-0.379356
C46	0.6469963804	-2.9405890268	9.2915104627	0.377144
C47	-2.5482799654	-0.3782981350	12.0132576954	0.352640
C48	-0.6350748880	-2.8492778642	7.1762753316	0.010331
C49	-2.9590719565	-2.2849076493	6.9604611974	-0.076948
C50	-3.5338712487	-3.5790572831	9.1301481901	0.099967
C51	-3.9670937181	-2.1481326865	10.9643960261	0.171475
C52	1.8478987320	-3.3416227463	9.9748602672	0.355535
C53	0.5994355153	-3.0778682927	7.8741250333	-0.162696
C54	-3.7534928150	-1.1365310427	11.9618550050	-0.165773
C55	-2.3374586262	0.4596275802	13.1685441498	0.335397
H56	-0.7115454309	-3.2648887416	6.1655720269	0.121151
C57	-3.9346301747	-3.2514360691	7.6828713301	-0.029315
H58	-2.8072763368	-2.5852868108	5.9197754740	0.097322
H59	-3.3454270800	-1.2666663924	6.9665671154	0.057366
H60	-4.3335975915	-4.1644069266	9.6041285586	0.016216
H61	-2.6301629528	-4.1980155406	9.1301347673	0.040014
H62	-4.8293156318	-2.8018135252	11.1461713428	0.102009
C63	2.9427740507	-3.8174305264	9.2559149641	-0.307450
C65	1.7186026999	-3.5861488493	7.1489215786	-0.195259
C66	-4.7386567350	-1.0183411483	12.9879955723	-0.194999
C67	-3.3085232006	0.5727309021	14.1595802024	-0.301056
H69	-3.9901680000	-4.2062262637	7.1450082210	0.004129
H70	-4.9404782678	-2.8171992570	7.6684034577	0.046781
C71	2.8204870709	-3.9114812739	7.8724735195	-0.042118
H72	3.8611239993	-4.1041465978	9.7569328197	0.168071
H73	1.6649185080	-3.6898026441	6.0681638380	0.148623
C74	-4.4761436495	-0.1679074435	14.0120666927	-0.034877
H75	-5.6495269408	-1.6090837228	12.9295258967	0.141290
H77	-3.1572190800	1.2145558653	15.0207343712	0.166475
O78	-4.9690016603	3.3137855349	6.1611150441	-0.469058
C79	-4.8246841842	2.0271426850	5.5386520530	0.201335

Continued on next page

Table A16 – continued from previous page

Atom name	X	Y	Z	charge
H80	-5.5779487835	1.3301897521	5.9289227764	0.012681
H81	-5.0106977899	2.1941816183	4.4748146558	0.029679
H82	-3.8252175438	1.6140828634	5.6848714803	0.039533
O83	-1.1352464784	1.1065214078	13.2010758330	-0.350750
C84	-0.7782733519	1.7762803130	14.3984163138	0.034239
H85	0.2426232175	2.1345243158	14.2522582371	0.078194
H86	-1.4342297327	2.6343656460	14.6012673880	0.041560
H87	-0.8042728785	1.0977000948	15.2610061062	0.058623
O88	1.8073918345	-3.2027737736	11.3264422816	-0.374420
C89	2.9604778963	-3.5654212892	12.0636509782	0.090914
H90	2.7190570994	-3.3802541005	13.1117088031	0.072534
H91	3.2099235285	-4.6274405639	11.9318831604	0.029815
H92	3.8308954280	-2.9576765902	11.7801423653	0.038604
O93	-2.1489595190	-0.9574663024	4.1353056536	-0.411987
C94	-2.5784101660	-1.5774644603	2.9346239057	0.061617
H95	-2.5798307345	-0.8701277077	2.0949616133	0.054539
H96	-1.9508799383	-2.4403938721	2.6712248296	0.042404
H97	-3.5983664753	-1.9208244510	3.1179947050	0.069652

Table A17: list of the optimized coordinates and atomic partial charges of the protonated 5-OCH_3-SOZMPU in the oxidation state MnIV-MnIV

Atom name	**X**	**Y**	**Z**	**charge**
Mn1	-1.1996031393	1.0731700518	8.0412889424	1.282226
O2	-2.4046688857	2.2594313844	7.3335363846	-0.513673
O3	-1.0281966211	0.1943996427	6.3972181226	-0.556083
O4	-2.4398127995	-0.0319889085	8.7147499219	-0.728532
N5	-1.1289029255	2.1694111228	9.7474608791	-0.204941
N6	0.3893601834	2.3369065458	7.4413511230	-0.346610
C7	-3.4396562028	2.7572423097	7.9966133553	0.285820
C8	-4.6212528892	3.0778653710	7.2713834193	0.317053
C9	-5.6955026154	3.6889130231	7.9243962055	-0.204982
C10	-5.5656089251	3.9641071789	9.2732708504	-0.021300
C11	-4.4637326548	3.6961280733	10.0319773271	-0.152430
C12	-3.3695224557	3.0653209102	9.3828113731	-0.134697
C13	-2.1505510845	2.8646599944	10.1244725100	0.079860
C14	0.0947400586	2.1694295047	10.5693594067	0.195702
C15	1.3686064302	2.3723720418	9.7235804361	-0.144205
C16	1.1196827211	3.1492258258	8.4310123240	0.051326
C17	0.6497769476	2.5398391138	6.1938234052	0.165760
C18	0.0043905242	1.9129666267	5.0674137319	-0.082914
C19	0.2481707194	2.4675103312	3.7818840019	-0.198449
C20	-0.3979140896	1.8872702913	2.7336201808	0.003295
C21	-1.2593765968	0.8071640149	2.8099244737	-0.241465
C22	-1.4916005184	0.2479626582	4.0706848880	0.317497
C23	-0.8482978649	0.7908701921	5.2273230035	0.214764
H25	-6.5854079919	3.9326845967	7.3511064056	0.183084
H27	-4.4048622381	3.9415469223	11.0892330890	0.149941
H28	-2.0872646581	3.3810832017	11.0870879593	0.128209
H29	0.0015390986	2.9745084195	11.3077693682	0.020044
H30	0.1352237691	1.2203578906	11.1054379190	-0.007326
H31	2.1007794210	2.9199893507	10.3264666791	0.061221
H32	1.8099608650	1.4038386842	9.4772332614	0.069652
H33	2.0707689017	3.4829302582	7.9992736986	0.067759
H34	0.5206813908	4.0464674888	8.6381373138	0.052255
H35	1.4149292681	3.2821154013	5.9398137362	0.109788
H36	0.9048978340	3.3258870761	3.6698161448	0.161360
H38	-1.7398047662	0.4121201812	1.9216619229	0.170790

Continued on next page

Table A17 – continued from previous page

Atom name	X	Y	Z	charge
Mn40	-1.4834627828	-1.3738472239	9.4961940042	1.249179
O41	0.0114348904	-0.2454600726	8.9153198983	-0.676388
O42	-0.3083253989	-2.6842352192	10.1829019002	-0.640196
O43	-1.5466196529	-0.5972651633	11.1883559469	-0.542898
N44	-1.4220929396	-2.4252614725	7.7657027198	-0.056821
N45	-3.1462278970	-2.3967122508	9.8949181817	-0.210111
C46	0.7447711485	-3.1637676984	9.5692236195	0.302039
C47	-2.6054863678	-0.3028074502	11.9039168987	0.347044
C48	-0.3769799112	-3.0786752474	7.3677393780	0.019908
C49	-2.6635542886	-2.4517234040	6.9920489057	-0.123991
C50	-3.4056871872	-3.6424265636	9.1577702091	-0.044421
C51	-4.0213645671	-1.9804815923	10.7569745375	0.108032
C52	1.8676901804	-3.5789790999	10.3652295357	0.365548
C53	0.8017361686	-3.3260951250	8.1526298126	-0.057343
C54	-3.8777008451	-0.9015124161	11.6840109598	-0.115321
C55	-2.4536958252	0.5897810747	13.0196676738	0.230128
H56	-0.4011945349	-3.5127220224	6.3633133017	0.131643
C57	-3.6764677365	-3.4137755030	7.6619830969	0.012952
H58	-2.4521783256	-2.7525600100	5.9624765373	0.107683
H59	-3.0456298580	-1.4324604543	6.9746388744	0.091310
H60	-4.2558846090	-4.1629735515	9.6126583150	0.066358
H61	-2.5245073448	-4.2761226088	9.2917550001	0.090374
H62	-4.9546832625	-2.5462775048	10.8309807474	0.139466
C63	3.0008133248	-4.1017404434	9.7405999302	-0.300801
C65	1.9566850155	-3.8865389908	7.5300466451	-0.263927
C66	-4.9951582535	-0.5715438132	12.5072655697	-0.200268
C67	-3.5553345679	0.9133030214	13.8111231413	-0.234910
H69	-3.6480539432	-4.3970479343	7.1782636308	0.031112
H70	-4.6874967765	-3.0161456490	7.5222053978	0.043395
C71	2.9818280855	-4.2238855534	8.3559762028	0.022250
H72	3.8694416298	-4.4019118734	10.3165972251	0.185104
H73	1.9895925848	-4.0258442287	6.4528397613	0.178729
C74	-4.7751758316	0.3251621315	13.5019114370	0.000147
H75	-5.9599105567	-1.0404544882	12.3342168896	0.161286
H77	-3.4612587309	1.5969996692	14.6479806215	0.170963
O78	-4.7002741576	2.8731233729	5.9254676112	-0.460517
C79	-4.8130336326	1.5070820890	5.5055159846	0.115378

Continued on next page

Table A17 – continued from previous page

Atom name	X	Y	Z	charge
H80	-5.7220923672	1.0523624333	5.9201337662	0.038696
H81	-4.8907367973	1.5355344815	4.4170196349	0.080026
H82	-3.9367988241	0.9223501110	5.7969455252	0.057931
O83	-1.1970943599	1.0664705039	13.2179888036	-0.281423
C84	-0.9408640267	1.7932244945	14.4169180783	-0.081316
H85	0.1307802312	1.9947329376	14.4222791225	0.112074
H86	-1.4885472987	2.7442017849	14.4391518371	0.083710
H87	-1.2054218322	1.2033947614	15.3023177268	0.104709
O88	1.7186432810	-3.4022631094	11.6951791978	-0.374975
C89	2.7878773236	-3.7950417779	12.5499708513	0.047958
H90	2.4519042904	-3.5788002211	13.5642277514	0.096635
H91	3.0006876428	-4.8669659444	12.4566691584	0.064734
H92	3.6991934875	-3.2220371236	12.3377565788	0.057387
O93	-2.3172339813	-0.8107509298	4.2953373361	-0.366531
C94	-3.0358205064	-1.3390315779	3.1837717096	-0.039602
H95	-3.6802459935	-0.5788313979	2.7258462782	0.080004
H96	-2.3587083804	-1.7456654444	2.4228594958	0.092940
H97	-3.6560349789	-2.1439416244	3.5808844961	0.098885
H98	0.5786470714	-0.6827310854	8.2623577709	0.336250

7.2 Empirical conversion method

Table A18: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of alcohol in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The predicted pK_A values for water were obtained with the prediction scheme of Friesner et al. [58] as implemented in the Jaguar pK_A prediction module [66]. The predicted pK_A values for the other solvents were obtained by applying the empirical transform to the pK_A values in water predicted with Friesner's scheme [58]. The pK_A values calculated using the Jaguar pK_A prediction module are highlighted in blue. The pK_A values calculated using the electrostatic transformation are highlighted in green.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	alcohol	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1A	methanol	15.50	16.40	-0.90				29.00	29.60	0.60			
2A	ethanol	16.40	16.00	-0.10				29.80	29.20	-0.60			
3A	methanol	17.10	17.10	0.00				30.30	30.30	0.00			
	pK_A-RMSD			0.52						0.49			

Table A19: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of phenol in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	phenol (p)	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1B	phenol	10.00	19.80	0.20	26.60	26.10	0.50	18.00	17.70	0.30	14.33	14.00	0.33
2B	4-amino-p	10.46	9.30	1.16		25.60			17.20			13.50	
3B	4-chloro-p	9.90	9.60	0.30	25.44	25.90	-0.46	16.70	17.50	-0.80		13.80	
4B	4-fluoro-p	10.20	10.40	-0.20		26.70		18.00	18.30	-0.30		14.60	
5B	4-methoxy-p	10.40	10.30	0.100		26.60		19.10	18.20	0.90		14.50	
6B	4-methyl-p	10.50	10.60	-0.10	21.45	26.90	0.55	18.90	18.50	0.40	14.55	14.80	-0.25
7B	4-nitro-p	7.20	7.30	-0.10	21.30	**23.60**	**−2.30**		15.20		11.20	11.50	-0.30
8B	4-formyl-p	7.60	7.60	0.00		24.30			15.50			11.80	
9B	2-nitro-p	7.17	8.00	-0.83		24.80			15.90		11.53	12.20	0.67
	pK_A-RMSD			0.50			1.23			0.60			0.42

Table A20: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of carboxylic acids in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	carbox ac	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1C	benzoic ac	4.22	3.90	0.32	20.10	19.30	0.80	11.00	10.90	0.10	9.40	8.90	0.50
2C	3,4-Me_2-benz	4.44	3.70	0.74	19.44	19.10	0.34	11.40	10.70	0.70	9.09	8.70	0.39
3C	3-OH-benz	4.08	3.50	0.58		18.90		11.10	10.50	0.60		8.50	
4C	4-Br-benz	3.96	3.50	0.46	19.20	19.70	-0.50	10.90	11.30	-0.40		9.30	
5C	4-Cl-benz	4.00	3.40	0.60		18.80		10.10	10.40	-0.30		8.40	
6C	glutaric ac	4.32	4.30	0.02	19.20	19.70	-0.50	10.90	11.30	-0.40		9.30	
7C	fumartic ac	3.03	3.00	0.03	18.60	18.40	0.20		10.00		8.00	8.00	0.00
8C	proprionic ac	4.85	4.80	0.05		20.20			11.80		9.70	9.80	-0.10
9C	adipic ac	4.41	4.40	0.01	20.30	19.80	0.50	11.90	11.40	0.50		9.40	
10C	succinic ac	4.20	4.10	0.10	17.60	19.50	-1.90		10.10		9.14	9.10	0.04
11C	malic ac	3.50	2.70	0.80		18.10			9.70		8.35	7.70	0.65
12C	malonic ac	2.90	3.40	-0.50		18.80			10.40		7.66	8.40	-0.74
	pK_A-RMSD			0.45			0.90			0.44			0.44

Table A21: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of thiols in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	thiols (t)	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1D	thiophenol	6.62	6.60	0.02	19.33	19.30	0.03	10.28	10.30	-0.02	8.95	9.30	-0.35
2D	4-Me_3-t	6.82	6.70	0.12	19.32	19.40	-0.08	10.82	10.40	0.42	8.96	9.40	-0.44
3D	3-Me_3-t	6.66	6.50	0.16	19.51	19.20	0.31	10.55	10.20	0.35	9.60	9.20	0.40
4D	M-toluidine	6.78	6.60	0.18		18.10		11.19	10.30	-0.40	8.95	9.30	-0.35
5D	4-Cl-benz	5.90	5.50	0.40		17.00		8.55	9.20	-0.30		8.00	
6D	glutaric ac	6.89	6.80	0.09		18.30		11.35	10.50	-0.40		9.30	
7D	fumartic ac	6.99	6.50	0.49		18.00		10.70	10.20	0.50		9.00	
8D	proprionic ac	6.02	6.20	-0.18		17.70		8.98	9.90	-0.92		8.70	
	pK_A-RMSD			0.12			0.19			0.65			0.39

Table A22: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of sulfonamides in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	sulfonamides (SA)	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1E	saccharin	1.80	3.00	-1.20	14.57	**16.10**	**−1.53**	6.43	7.90	-1.47	5.00	**7.50**	**−2.50**
2E	Piloty's acid	9.29	9.80	-0.51	23.75	22.90	-0.85	15.40	14.70	0.70	13.25	13.80	-0.55
3E	phenil-SA	10.10	10.20	-0.10	24.72	23.30	1.42	16.10	15.10	1.00	15.12	14.20	0.92
4E	methyl-SA	10.87	12.00	-1.13	25.76	25.10	0.66	17.50	16.90	0.60	15.00	16.00	-1.00
	pK_A-RMSD			0.86			1.17			1.00			1.45

Table A23: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of hydroxylic acids in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	hyd-ac	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1F	benz-hyd-ac	8.80	8.50	0.30	22.34	23.50	-1.16	13.70	15.10	-1.40	13.25	13.50	-0.25
2F	methyl-hyd-ac	8.50	8.10	0.40	24.57	23.10	1.17	16.00	14.70	1.30	13.41	13.10	0.31
3F	capryl-hyd-ac	9.00	8.50	0.50	24.55	23.50	1.05	15.58	15.10	0.48	13.48	13.50	-0.02
	pK_A-RMSD			0.41			1.13			1.14			0.23

Table A24: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of imides in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	imides	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1G	succinimide	9.60	8.70	0.90	23.46	23.30	0.16	14.70	14.40	0.30	13.13	12.90	0.23
2G	dimethadione	7.60	6.10	1.50	20.67	20.70	0.03	11.67	11.80	-0.13	10.09	10.30	-0.21
3G	phenytoin	8.30	8.00	0.10	22.38	22.60	-0.22	13.42	13.70	-0.28	12.08	12.20	-0.12
	pK_A-RMSD			1.02			0.16			0.25			0.19

Table A25: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of barbituric acids in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	bar-ac	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1H	5,5-dimeth-bar-ac	8.00	8.10	-0.10	23.74	23.60	0.14	14.79	14.60	0.19	12.71	12.80	-0.09
2H	5,5-meth-phen-bar-ac	7.40	7.50	-0.10	22.58	23.00	-0.42	13.70	14.00	-0.30	11.77	12.20	-0.43
3H	hexobarbital	8.20	8.20	0.00	23.90	23.70	0.20	14.97	14.70	0.27	13.21	12.90	0.31
	pK_A-RMSD			0.08			0.28			0.26			0.31

Table A26: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of tetrazoles in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	tetrazoles	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1I	tetrazole	4.90	4.80	0.10				8.20	8.20	0.00			
	pK_A-RMSD			0.10						0.00			

Table A27: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of primary amines in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	prim-am	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1J	methylamin	10.62	10.50	0.12	18.37	18.30	0.07	9.87	10.00	-0.13	11.00	10.80	0.20
2J	ethylamin	10.70	11.00	-0.30	18.40	18.80	-0.40	10.24	10.50	-0.26	11.00	11.30	-0.30
3J	butylamin	10.80	10.70	0.10	18.80	18.50	0.30	10.47	10.20	0.27	11.12	11.00	0.12
	pK_A-RMSD			0.20			0.29			0.23			0.22

Table A28: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of secondary amines in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	seco-am	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1K	dimethamin	10.64	10.90	-0.26	18.50	18.80	-0.30	10.22	11.10	-0.88	11.00	10.80	0.20
2K	diethamin	11.10	11.10	0.00	18.75	19.00	-0.25	12.15	11.30	0.85	11.04	11.10	-0.06
3K	dipropylamin	11.00	10.60	0.40	18.81	18.50	0.31	10.73	10.80	-0.07	11.08	10.60	0.48
4K	ethylpropyl.	10.80	10.40	-0.20	18.30	18.30	0.00	10.24	10.60	-0.36	10.72	10.40	0.32
	pK_A-RMSD			0.36			0.25			0.64			0.33

Table A29: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of tertiary amines in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	tert-am	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1L	trimethylamin	9.80	10.10	0.30	16.69	17.10	-0.41	8.89	9.70	-0.81	9.32	10.10	-0.78
2L	triethylamin	10.78	10.60	0.18	18.08	18.20	-0.12	9.88	10.20	-0.32	10.15	10.60	-0.45
3L	tripropylamin	10.70	9.20	1.50	17.96	16.80	1.16	9.46	8.80	0.84	10.60	9.20	0.86
	pK_A-RMSD			0.89			0.89			0.70			0.72

Table A30: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of anilines in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	anilines	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1M	aniline	4.60	4.70	-0.10	10.56	11.30	-0.74	3.82	4.00	-0.18	4.15	5.20	-1.05
2M	2-chloro-a	2.66	2.60	0.06	9.66	9.20	0.46	1.94	1.90	0.04	3.71	3.10	0.61
3M	M-toluidine	4.69	4.60	0.09		11.20			3.90		5.95	5.10	0.85
4M	4-chloro-a	4.00	4.00	0.00		10.60			3.30		4.95	4.50	0.45
5M	3-chloro-a	3.52	3.70	-0.18	10.67	10.30	0.37	2.90	3.00	-0.10	3.61	4.20	-0.59
6M	4-methyl-a	3.52	4.90	-0.32	12.29	11.50	0.79	4.48	4.20	0.28	4.87	5.40	-0.53
	pK_A-RMSD			0.17			0.62			0.17			0.66

Table A31: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of heterocycles in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	hetero	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1N	imidazole	7.05	6.80	1.45	14.72	14.40	0.32	6.40	5.80	0.60	7.02	6.90	0.12
2N	piperidine	11.10	11.10	0.00	18.76	18.70	0.06	10.42	10.10	0.32	10.94	11.20	-0.26
3N	quinoline	4.80	5.00	-0.20		12.60			4.00		5.16	5.10	0.06
4N	pyridine	5.30	5.20	0.25	12.60	12.80	-0.20	3.50	4.20	-0.70	5.44	5.30	0.14
	pK_A-RMSD			0.82			0.22			0.56			0.10

Table A32: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of indoles and pyrroles in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	ind and pyrr	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1O	indole	16.10	15.50	0.60				21.00	21.30	-0.30			
2O	pyrrole	17.50	16.80	0.70				23.00	22.60	0.40			
	pK_A-RMSD			0.65						0.35			

Table A33: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of amidines in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	amidines	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1P	benzamidine	11.53	11.40	0.13	21.81	21.00	0.81	13.59	12.80	0.79	12.98	12.40	0.58
2P	acetamidine	10.25	10.60	-0.35	19.01	19.85	-0.84	11.14	12.00	-0.86	11.00	11.60	-0.60
	pK_A-RMSD			0.26			0.83			0.83			0.70

Table A34: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of guanidines in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	guanidine (g)	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1Q	guanidine	13.80	12.50	1.30	22.91	22.50	0.41	14.85	14.50	0.35	13.67	13.40	0.27
2Q	methylg-g	13.40	13.40	0.00	23.43	23.40	0.03	15.78	15.40	0.38	14.18	14.30	-0.12
3Q	debrisoquine	13.00	11.90	1.10	21.59	21.90	-0.31	13.56	13.90	-0.34	13.00	12.80	0.20
	pK_A-RMSD			0.98			0.30			0.36			0.21

Table A35: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of benzodiazepines in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	benzodiaz (g)	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1R	1,3-dihydro-1-methyl-5-phenyl-1-4-bzodiaz-2-one	3.30	3.80	-0.50	10.83	11.10	-0.27	2.95	3.20	-0.25	3.29	3.50	-0.21
2R	1,3-dihydro-3-bzodiaz-2-one	1.70	1.90	-0.20	9.41	9.20	0.21	1.54	1.30	0.24	1.94	1.60	0.34
3R	1-3-dihydro-5-phenyl-1,4-bzodiaz-2-one	3.50	4.00	-0.50	11.07	11.30	-0.23	3.20	3.40	-0.20	3.52	3.70	-0.18
	pK_A-RMSD			0.42			0.24			0.23			0.25

Table A36: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of pyrroles (C-2 protonation) in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	pyrroles (C-2 p.)	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1S	pyrrole	-3.80	-4.10	0.30	2.71	2.50	0.24	-5.23	-5.60	0.37	-4.16	-4.40	0.24
2S	1-methylpyrrole	-2.90	-2.30	0.60	3.81	4.30	-0.49	-4.21	-3.80	-0.41	-3.36	-2.60	-0.76
3S	3-methylpyrrole	-1.00	-0.90	-0.10	6.05	5.70	0.35	-2.02	-2.40	-0.38	-1.28	-1.20	-0.08
	pK_A-RMSD			0.39			0.37			0.39			0.46

Table A37: Comparison of experimental (exp) and predicted (pred) pK_A values for the family of indoles C-3 protonation in water, acetonitrile (MeCN), dimethyl - sulfoxide (Me_2SO) and methanol (MeOH)[58, 128]. The same style of A18 is used here.

	family	pK_A in water			pK_A in MeCN			pK_A in Me_2SO			pK_A in MeOH		
No	indoles (C-3 p.)	exp	pred	dev	exp	pred	dev	exp	pred	dev	exp	pred	dev
1S	indole	-3.60	-3.70	0.10	3.42	3.20	0.22	-4.42	-4.60	-0.18	-3.89	-4.00	0.11
2S	1-methylindole	-2.30	-2.00	-0.30	4.80	4.90	-0.10	-3.05	-2.90	-0.15	-2.58	-2.30	-0.28
3S	3-methylindole	-4.60	-4.60	0.00	2.42	2.30	0.12	-5.44	-5.50	-0.06	-4.96	-5.20	-0.24
	pK_A-RMSD			0.18			0.16			0.14			0.22