Expanding the scope of integral equation-based solvation theory

Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)

Die Dissertation wurde angefertigt im Zeitraum vom 01.11.11 bis zum 14.12.15 und der Fakultät Chemie und Chemische Biologie der Technischen Universität Dortmund vorgelegt

von
Roland Frach.

Dortmund, 2015

Erstgutachter: Prof. Dr. Stefan M. Kast
Zweitgutachter: Prof. Dr. Roland Winter

Danksagungen

Mein ganz besonderer Dank geht an Prof. Dr. Stefan M. Kast, der mir die Arbeit an diesem interessanten Thema erst ermöglicht hat. Durch seine vielen Anregungen, jedoch auch durch die Möglichkeit zur freien Arbeitsweise konnte der Grundstein für diese Dissertationsschrift gelegt werden.

Daneben danke ich natürlich Prof. Dr. Roland Winter für die Erstellung des Korreferats.

Beiden langjährigen Bürokollegen und Lektoren der Arbeit Leonhard Henkes und Florian Mrugalla, die mir immer mit Hilfe, guter Laune sowie Motivationen und (meistens) guten Ideen zur Seite standen, möchte ich nochmals ganz herzlich danken. Ebenso bedanke ich mich bei Franziska Hoffgaard, Jochen Heil, Daniel Tomazic, Patrick Kibies, Martin Urban und bei allen anderen Kollegen für die Zusammenarbeit sowie die nette Zeit hier im Arbeitskreis. Nicht zu vergessen sind natürlich Prof. Dr. Alfons Geiger, dessen Tür und Ohr für geplagte Promovierende immer offen standen, und Anneliese Ahlke für Rat und Tat bei allen bürokratischen Schwierigkeiten. Bei allen „alten“ Schulfreunden, Universitätskommilitonen, Fußballkollegen, Teammitgliedern der TTF Lünen und jedem Freund außerhalb meines akademischen „Werdegangs“ und meiner „Sportlaufbahn“ danke ich natürlich auch vielmals.

An meine Eltern, meinen Bruder und meine Familie möchte ich meinen herzlichsten Dank richten für die Zeit, die in mich investiert werden musste und für alles was sie aus mir gemacht haben. Mein größter Dank geht abschließend an meine Frau Sandra, die mir während der ganzen Promotion Kraft, Trost und Geborgenheit geschenkt hat sowie viel Zeit zum Lesen meiner Dissertationsschrift investierte.

Weitere Anmerkungen

Teile des zweiten Kapitels sind bereits veröffentlicht in der Publikation:

R. Frach, S. M. Kast, „Solvation Effects on Chemical Shifts by Embedded Cluster Integral Equation Theory“, *J. Phys. Chem. A* **2014**, *118*, 11620.

1D-RISM-Ergebnisse für reines C_6H_6 und C_6F_6 und 3D-RISM-Verteilungen von C_6H_6 und C_6F_6 in C_6H_6 und C_6F_6 mit Originalkraftfeldern wurden bereits in Vorarbeiten[*] präsentiert und in dieser Arbeit in einem breiteren Kontext in Kapitel 3 wiederverwendet.

[*] R. Frach, „Quantenchemische Modelle für nichtwässrige Dienaminlösungen“, Masterarbeit, TU Dortmund **2011**.

ZUSAMMENFASSUNG

Implizite Lösungsmittelmodelle werden verwendet, um die Eigenschaften von Lösungen mit realisierbarem Rechenaufwand vorauszusagen und zu verstehen. Viele dieser impliziten Lösungsmittelmodelle ignorieren jedoch die Granularität des Lösungsmittels und scheitern an der Beschreibung wichtiger gerichteter Wechselwirkungen, wie z. B. Wasserstoffbrücken. Einen Weg zur Beschreibung verschiedenster Lösungsmitteleigenschaften, der die Merkmale der Lösungsmittelmoleküle nicht vernachlässigt, ermöglicht das *Reference Interaction Site Model* (RISM). Hierbei wird das Lösungsmittel über seine Paarverteilungsfunktion charakterisiert. Diese stellt ein Maß für die Nahordnung der Lösungsmittelteilchen dar und ist im Rahmen des RISM-Ansatzes verbunden mit der freien Solvatationsenergie. Eine Kombination des RISM-Modells mit quantenmechanischer Präzision für die gelöste Substanz führt zu dem so genannten *Embedded Cluster Reference Interaction Site Model*, EC-RISM. Diese Kombination wird durch die Anwendung eines selbstkonsistenten Ansatzes, bei dem die Wechselwirkung zwischen dem Solvens und der gelösten Substanz durch eingebettete Punktladungen beschrieben wird, erreicht.

Bisher wurde die EC-RISM-Methodik nur für Studien von freien Energien und verwandten Größen von dipolaren, hauptsächlich wässrigen Lösungen, verwendet. Die Anwendbarkeit von EC-RISM auf andere Eigenschaften von Lösungen sowie für nicht dipolare Lösungsmittel wurde bisher nicht sichergestellt. Daher stehen im Folgenden zwei Aspekte im Fokus der Arbeit. Einerseits werden chemische NMR-Verschiebungen berechnet, mit denen getestet wird, ob der Einfluss des Solvens auf die Wellenfunktion des gelösten Teilchens geeignet beschrieben wird und um die Wellenfunktion selbst zu überprüfen. Andererseits werden Benzol- und Hexafluorbenzolmodelle entwickelt um zu erforschen, ob diese komplexen, aber ähnlichen Lösungsmittel mit Hilfe von Integralgleichungsmethoden wie RISM unterscheidbar sind. Anschließend werden diese Modelle verwendet, um eine chemische Reaktion zu untersuchen, bei der eine Steigerung der Stereoselektivität durch den Einsatz von Hexafluorbenzol im Vergleich mit Benzol ausgelöst wird.

Die Ergebnisse der Untersuchungen zeigen, dass die Anwendung von EC-RISM die chemische Verschiebung in wässriger Lösung systematisch verbessert und somit auch eine plausible Wellenfunktion wiedergibt. Weiterhin wird bestätigt, dass sich Lösungen in Benzol und Hexafluorbenzol adäquat beschreiben lassen und sogar Stereoselektivitätsvoraussagen in Übereinstimmung mit Experimenten möglich sind. Hierbei können verschiedenste

Lösungsmitteleinflüsse separiert und analysiert sowie betrachtet und diskutiert werden, wobei die Bedeutung sowohl dispersiver als auch multipolarer Wechselwirkungsanteile verdeutlicht wird.

ABSTRACT

Implicit solvation models are used to predict and to understand various properties of solutions within feasible time scales. But common implicit approaches ignore the granularity of the solvent and fail to describe substantial directional interactions like hydrogen bonds. A different way to calculate various properties of solvents without that retains this features is presented in this work, the reference interaction site model (RISM). Thereby the solvent is characterized by its pair distribution function, which is a measure for the near-order of the solvent particles. Furthermore the pair distribution functions are connected within the RISM approach with the free energy of solvation. Combining the RISM approach with quantum-mechanical precision for the solute leads to the so-called embedded cluster reference interaction site model, EC-RISM. This is achieved by application of a self-consistent approach for both parts that are connected with a cluster of embedded point charges, which describes the solvent-solute interaction.

By now the EC-RISM methodology was only applied for studies of free energies and related subjects of dipolar, mostly aqueous solutions. The applicability of EC-RISM for different properties of solutions as well as for nondipolar solvents was so far not ensured in the past. Here, two important aspects are in the focus; on the one hand, NMR chemical shifts are calculated to test whether the EC-RISM approach properly describes the solvent influence on the wave function of the solute and to test the wave function itself. On the other hand benzene and hexafluorobenzene models are developed to research if these complex and very similar solvents are distinguishable with integral equation theories like RISM. Furthermore these models are applied for the investigation of a chemical reaction that shows a stereoselectivity enhancement that is caused by the application of hexafluorobenzene compared to benzene.

These investigations show that EC-RISM systematically improves the chemical shifts in aqueous solutions and therefore displays the adequacy of the corresponding wave function. Additionally it is confirmed that benzene and hexafluorobenzene are properly described and that even the stereoselectivity of the previously mentioned chemical reaction is correctly predicted. Thereby the different solvent influences are separated, analyzed, considered and discussed, whereas the relevance of dispersive and multipolar parts is elucidated.

LIST OF CONTENTS

1 INTRODUCTION

1.1 Motivation

A vast amount of chemical reactions occur in solution. Either in artificially or natural environment, solutions are frequently the only surrounding for a wide range of chemical reactions to occur. In contrast to solids, where particles exhibit a low mobility, and gases, for which the particle density at moderate temperatures is very low, reactions in the liquid state are in general relatively fast and in many cases controllable. Therefore solids are the right choice for substances that should not react, for e.g. by degradation. In comparison gas phase reactions are appropriate for simple processes or explosives, while solutions are suited for complex and highly specific reactions. Nonetheless there are exceptions to this statement and it is for sure an oversimplification, but proper for very many cases, reactions in aqueous solutions of living organisms being the most remarkable example for complex mechanisms.

Because of their capability to largely alter the reactivity of solutes, many different solvents are applied routinely in everyday laboratory work. An appropriate solvent choice is still difficult, but crucial for the desired reaction outcome. A specific solvent is often picked by its polarity, due to desired or sometimes known favored or disfavored dipolar interactions. Furthermore various solvents are often chosen because they directly intervene in the reaction mechanisms. In many cases the consequences of the solvent choice is not easily predictable and many solvents have to be screened before the wanted yields are achieved.

In the last decades computational solvation models improved the understanding of processes in the liquid state and even quantitatively predicted several properties of solutions.[1] By these findings issues in the solvent selection process, i. e. for organic reactions or properties of solutions[2,3] were recognized and bypassed. However up to now there is no solvent model that is applicable and accurate enough for each problem in the laboratory flask.

Thus plenty different models have been developed for specific subjects. These approaches are commonly differentiated between explicit and implicit solvent models. One of the most prominent explicit model is the molecular dynamics (MD) simulation[4] methodology, where molecular forces are numerically integrated over time to approximate the phase space for a given system and ensemble. In force field MD (ffMD) simulations forces are evaluated only for nuclei by using derivatives of parametrized empirical force fields. Electrons are not explicitly treated and their influence on the systems has to be included in these force fields. On high performance computation environments ffMD simulations techniques are capable to describe systems with a million of nuclei over some milliseconds. Due to fast improvements on hard- and software and relatively good parallelizability of ffMD techniques the treatable particle number and simulation length are rapidly increasing. A detailed quantum-mechanical description of particle dynamics is achievable with Car-Parrinello[5]- or other *ab initio*-MD (aiMD) simulation[6] techniques. These methods allow for the investigation of reactivity and electronic properties like excited states, electron transfer and so on. Nevertheless due to much higher complexity and inferior scalability of these techniques, the number of particles and simulation timescales are by far smaller compared to classical MD. In general these techniques are powerful, but often much too time consuming and comprehensive to sufficiently sample the phase space of many systems like highly diluted compounds. Additionally essential thermodynamic properties of moderately complicated systems are not directly accessible and must be sophisticatedly sampled by perturbative methods like free energy perturbation (FEP)[7,8] or importance sampling methodologies i. e. umbrella sampling.[9]

Instead implicit methods focus on the properties of solutes. They are commonly used when the properties of the solvent are not of primary importance, but its influence on the solute has to be represented properly.[10] A lot of these methodologies are based on the Poisson-Boltzmann equation[11] or on further simplifications like the generalized-born model (GB)[12] that are often, but not necessarily, combined with force field descriptions of the solute interactions. Another established class of methods are the polarizable continuum methods (PCM), which are routinely applied to supplement quantum-chemical calculations of solutes.[13] Paying the price with a loss in chemical detail and granularity, continuum methods have proven to be computationally fast and adequately accurate for many use cases.

An alternative to pure explicit or implicit models are the fluid phase integral equations[14] that are historically based the Ornstein-Zernike (OZ) equation.[15] In the context of these models the solvent is described by its local solvent density. Consequently not each individual solvent molecule has to be considered, but at the same time the solvent retains its granularity

and its directional interactions with a solute. The OZ equation itself represents a theory for simple spherical monoatomic liquids. For the treatment of complex molecular liquids the Molecular Ornstein-Zernike (MOZ)[16] equation or the site-site based reference interaction site model (RISM)[17,18] can be applied. Various types of approximations exist for the latter, which address different types of problems, for example polymers (PRISM)[19], spherical solvent distributions (1D RISM)[17,18] or more complex three dimensional surroundings (3D RISM).[20,21] Within approaches like RISM-SCF[22] or the embedded cluster RISM (EC-RSIM)[23] integral equation theory has also been coupled with quantum-mechanical electronic structure theory for solutes. In the past the coupling between RISM and quantum-mechanics (RISM/QM), among others especially EC-RISM proved to be extremely useful for calculations of relative Gibbs free energy in aqueous solutions.[23,24] But neither its relevance for free energies or kinetic properties in nondipolar solutions, nor any spectroscopic properties have ever been investigated. Furthermore suitable methodologies for these quantities are untested or not developed for the EC-RISM framework. Additionally also the quality of the electronic wave function was not ensured with any quantity before. Focused in the context of this work are exactly these issues to investigate the opportunities that are achievable by the EC-RISM approach.

1.2 Aims and objectives

In the past EC-RISM was almost only applied to issues concerning Gibbs free energies of solutes in aqueous dilutions and hereby deduced properties like tautomer ratios[24] and acidity constants.[23] In the context of this work studies are performed to scrutinize the physical properness of EC-RISM and to expand its space of applications, in particular to capture the adequacy of the electronic structure by estimation of chemical shieldings and chemical shifts because of their close correspondence to the wave function. Furthermore the range of investigated solvents was extended by the development and testing of nondipolar solvent models. To this end even the computation of kinetics properties for a complicated organic reaction was possible. The following chapter shall outline these projects.

After briefly introducing the basic theory in the next part of the work, the second chapter focusses on a combination of a chemical shielding evaluation technique with the EC-RISM approach. Chemical shieldings and even more chemical shifts, can be precisely measured in solution and connected to the electron density. Therefore chemical shifts are frequently used to test the electronic density and wave function of different theoretical models, without the need of high level quantum-chemical reference values.[25] For this reason the calculation of chemical shifts was chosen as test quantity for the wave functions achieved by EC-RISM. This is accomplished by coupling the gauge invariant atomic orbitals (GIAO) methodology[26] for chemical shieldings with EC-RISM solvation theory. Because of its low molecular size, experimentally known chemical shift values and its importance as peptide backbone model system, this coupled approach is applied to *N*-methylacetamide. Furthermore the influence of the level of theory for the quantum-chemical part as well as for the 3D RISM approach on the chemical shifts is investigated.

The third and fourth chapters addresses the topic of nonaqueos fluids and focusses on nonaqueos and nondipolar solvation. This part provides new insights into the capability of EC-RISM to describe such solvents, which was so far only investigated for dimethylsulfoxide solutions that are certainly nonaqueos but dipolar.[27] In contrast to the sparsely investigated RISM/QM solvent models a vast amount of different liquids was already studied by RISM approaches uncoupled to quantum chemistry. For only a small amount of dipolar solvents, like ionic liquids[28] or acrolein[29], results for RISM/QM approaches are established. Here, benzene and hexafluorobenzene are selected. Both solvent molecules have no permanent dipole moment and from the structural point of view are very similar. Additionally their

macroscopic properties resemble each other, making the distinction between both challenging for theory and precisely because of that an appropriate choice for testing EC-RISM. Different solvent models of both liquids are created by 1D RISM calculations based on classical force fields. To validate the solvent structure the resulting one dimensional solvent distributions are compared to the MD simulation results. Afterwards regular 3D RISM and EC-RISM calculations are performed for small solutes and compared to literature results. The free energies of these solutes is analyzed and divided into different contributions to give an estimate of dispersive and electrostatic influences of benzene and its perfuorinated counterpart. Additionally the influence of specific solute-solvent electrostatics is investigated and critically discussed.

The fourth chapter concentrates on a description of a stereoselective solvent dependent nitro-Michael addition.[30] Other RISM/QM calculations for example by Naka et al.[31] or Ida et al.[32] lay the foundation for reaction mechanism investigation by integral equation methodologies. Nevertheless only simple reactions in aqueous solutions were focused so far. In the here presented work a theoretical investigation of a remarkable stereoselective Michael addition is accomplished, for which the enantiomeric ratio changes significantly after solvent substitution of benzene with hexafluorobenzene. Therefore this part couples the essence of the previous chapter with a complex and chemically important issue. Furthermore it is shown how well the previously introduced models perform for these challenging problems. Moreover a new component to the interpretation of the solvent effect for this specific reaction mechanism is provided and an attempt to characterize chemical reaction kinetics with EC-RISM is established.

1.3 Theoretical basics

1.3.1 Classical density functional integral equation theory

In analogy to the Hohenberg-Kohn-theorems[33] that connects the minimum electronic energy of the ground state with a unique electronic density in the context of the fermionic or quantum-mechanical density functional theory (DFT), a similar approach is provable for the classical density functional theory[34,35] that relates the free energy of a fluid system $A[\rho]$ to its particle density ρ by

$$\begin{aligned} A[\rho] &= A^{\mathrm{id}} + A^{\mathrm{ex}}[\rho] \\ &= A^{\mathrm{id}} + A^{\mathrm{ex}}[\rho_0] - \int \Delta\rho(\mathbf{r}) \frac{\delta A^{\mathrm{ex}}[\rho_0]}{\delta\rho(\mathbf{r})} \mathrm{d}\mathbf{r} \\ &\quad - k_{\mathrm{B}}T \int_0^1 \mathrm{d}\lambda (1-\lambda) \int\int \Delta\rho(\mathbf{r}) \Delta\rho(\mathbf{r}') c(\mathbf{r},\mathbf{r}';\lambda) \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'. \end{aligned} \tag{1}$$

for a simple fluid.[36]

A central property in equation (1) and also in this work is the local particle density $\rho(\mathbf{r})$. Additionally equation (1) implies that, beside the easy to calculate ideal part of the free energy A^{id}, the excess part A^{ex} is needed. But it is only accessible by knowledge of the thermodynamics of a reference system, precisely its excess free energy $A^{\mathrm{ex}}[\rho_0]$, which is in general also unknown. Further the second order direct correlation function

$$c(\mathbf{r},\mathbf{r}') = -(k_{\mathrm{B}}T)^{-1} \frac{\delta^2 A^{\mathrm{ex}}[\rho]}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} \tag{2}$$

is required. This direct correlation function is the second functional derivative of the excess free energy w. r. t. differences of the local particle density $\rho(\mathbf{r})$ of particles at position $\mathbf{r}$ and on another particle with position $\mathbf{r}'$, which accordingly quantifies the system's free energy susceptibility as response to density fluctuations. At the same time the free energy depends on the difference in the local solvent density $\Delta\rho(\mathbf{r})$ between both systems. Furthermore a thermodynamic integration[37] has to be performed that couples the target and the reference system by a parameter λ for a given temperature T multiplied with the Boltzmann constant k_{B}. Relation (1) is also extendable to molecular fluids, but the following subchapter refrains from deducing a more sophisticated theory and introduces into the field of classical DFT in a manageable way. Comments on the theory of molecular fluids and the corresponding consequences are given in the next subchapters.

Percus presented an idea to gain access to a solvable approach for a classical DFT of pure liquids and mixtures.[38] He assumed the reference state with $\lambda = 0$ to be an undisturbed homogenous fluid with one particle α located at position $\mathbf{r}_\alpha = 0$. In this reference state this particle does not interact with all other particles. In the final state with $\lambda = 1$ this particle interacts with all solvent particles γ by pairwise additive interaction potential terms.

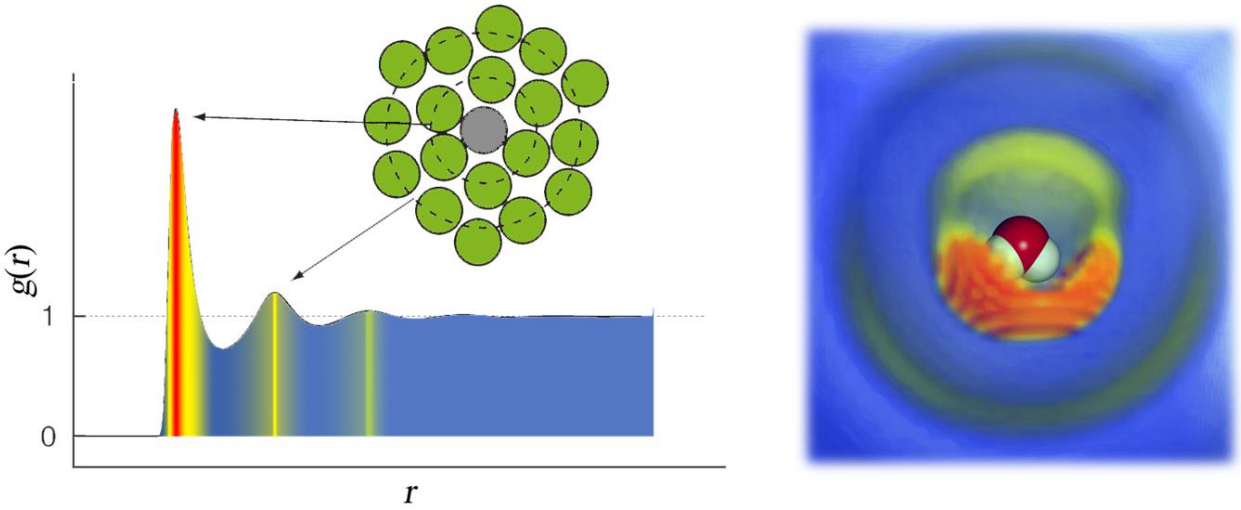


Figure 1.1 Illustration of a radial pair distribution function $g(r)$ (left) and a three dimensional solvent distribution (right).[39] Both distributions are color-coded in the same way, following from very high (red), through medium high (yellow) up to relatively low values (blue).

Then it can be shown that the local density of the final state ($\lambda = 1$) is connected to the pair distribution function of an isotropic and homogeneous system

$$g(\mathbf{r}_\alpha, \mathbf{r}_\gamma) = \langle N(N-1) \rangle \frac{p^{(2)}(\mathbf{r}_\alpha, \mathbf{r}_\gamma)}{\rho^2} \tag{3}$$

in the grand canonical ensemble by

$$\rho(\mathbf{r}_\gamma; \lambda = 1) = \rho g(\mathbf{r}_\alpha = 0, \mathbf{r}_\gamma; \lambda = 1)\,. \tag{4}$$

The pair distribution function connects the probability density $p^{(2)}(\mathbf{r}_\alpha, \mathbf{r}_\gamma)$ of finding a solvent particle at position $\mathbf{r}_\gamma$ simultaneously with a solute particle α at position $\mathbf{r}_\alpha$. Furthermore the probability density (3) is connected to the particle number N and the macroscopic particle density of the system ρ. Additionally

$$\Delta\rho(\mathbf{r}_\gamma) = \rho(\mathbf{r}_\gamma; \lambda = 1) - \rho(\mathbf{r}_\gamma; \lambda = 0) = \rho g(\mathbf{r}_\alpha, \mathbf{r}_\gamma) - \rho \overset{\text{def}}{=} \rho h(\mathbf{r}_\alpha, \mathbf{r}_\gamma)\,. \tag{5}$$

follows for the density difference and defines the total correlation function h. As illustrated in Figure 1.1 the pair distribution function fluctuates at short distances, indicating a distinct

short-range order of the particles that is directly proportional to the solvation shells. The fluctuations vanish at longer distances from the particle of interest (α).

By assuming just a small perturbation of the homogeneous fluid reference state by the particle α, it is possible to approximate $c(\mathbf{r},\mathbf{r}';\lambda) \approx c_0(\mathbf{r},\mathbf{r}')$ for all values of λ, where c_0 is the direct correlation function of the reference. Then it can be shown that

$$A[\rho] \approx A^{\mathrm{id}} + A^{\mathrm{ex}}[\rho_0] - \mu_0^{\mathrm{ex}} \int \Delta\rho(\mathbf{r})\mathrm{d}\mathbf{r} - \frac{1}{2} k_{\mathrm{B}} T \int\int \Delta\rho(\mathbf{r})\Delta\rho(\mathbf{r}')c_0(\mathbf{r},\mathbf{r}')\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \tag{6}$$

follows for the free energy (1) with μ_0^{ex} being the excess chemical potential of the reference state.[36] After insertion of equation (6) into the grand potential of the system

$$\Omega[\rho] = A[\rho] + \int U(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} - \mu \int \rho(\mathbf{r})\mathrm{d}\mathbf{r} \tag{7}$$

and minimization of (7) w. r. t. the local particle density it is possible to show that the density

$$\rho(\mathbf{r}) = \rho_0 \exp\left[-\frac{U(\mathbf{r})}{k_{\mathrm{B}} T} + \int \Delta\rho(\mathbf{r}')c_0(\mathbf{r},\mathbf{r}')\mathrm{d}\mathbf{r}' \right] \tag{8}$$

minimizes the grand potential. Here, $U(\mathbf{r})$ is the total interaction energy, μ is the total excess chemical potential and ρ_0 is the density of the reference system. By following Percus' idea further and assuming that particle α is one specific solvent particle γ itself, the particle number between both states is expected to remain constant and the particles interact via pairwise additive interaction potentials $u_{\alpha\gamma}(r_{\alpha\gamma})$, it is possible to deduce the so called hypernetted-chain (HNC)[40,41,42] closure relation

$$g(r_{\alpha\gamma}) = \exp\left[-\frac{u_{\alpha\gamma}(r_{\alpha\gamma})}{k_{\mathrm{B}} T} + \rho \int h(r_{\alpha\gamma'})c(r_{\gamma'\gamma})\mathrm{d}\mathbf{r}_{\gamma'} \right] \tag{9}$$

for a uniform fluid. Under these circumstances the resulting pair distribution function only depends on the distance $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$ between two particles and the macroscopic particle density simplifies to $\rho = \rho_0 = \rho_\gamma$. The HNC closure is correct up to second order in the density change. An exact closure relation has the form

$$g(r_{\alpha\gamma}) = \exp\left[-\frac{u_{\alpha\gamma}(r_{\alpha\gamma})}{k_{\mathrm{B}} T} + \rho \int c(r_{\gamma'\gamma})h(r_{\alpha\gamma'})\mathrm{d}\mathbf{r}_{\gamma'} + B \right], \tag{10}$$

where the higher order correlations are part of the so called bridge function B. Unfortunately a closed analytically and numerically computable form of the bridge function is not accessible.

The HNC closure is not solvable even when the total interaction potential is given in the form of a known pair potential, because both, the direct correlation function and the pair distribution function have to be determined. This can be achieved by application of the Ornstein-Zernike equation[15] (OZ)

$$h(r_{\alpha\gamma}) = c(r_{\alpha\gamma}) + \rho\int c(r_{\alpha\gamma'})h(r_{\gamma'\gamma})\mathrm{d}\mathbf{r}_{\gamma'}\,, \tag{11}$$

that was accidently found by Leonard Ornstein and Frederik Zernike and later proven for different systems by e. g. Zernike and Prins[43], Goldstein[44], Klein et al.[45]. One manageable way to deduce the OZ equation is to expand the functional definition of the Delta function by

$$\delta(\mathbf{r}-\mathbf{r}') = \frac{\delta\rho^{(1)}(\mathbf{r})}{\delta\rho^{(1)}(\mathbf{r}')} = \int H(\mathbf{r},\mathbf{r}'')H^{-1}(\mathbf{r}'',\mathbf{r}')\mathrm{d}\mathbf{r}'\,, \tag{12}$$

and to insert the density-density correlation function

$$H(\mathbf{r},\mathbf{r}') = \rho(\mathbf{r})\rho(\mathbf{r}')h(\mathbf{r},\mathbf{r}') + \rho(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')\,, \tag{13}$$

and its inverse $H^{-1}(\mathbf{r},\mathbf{r}')$. The OZ equation contributes with a second relation that connects *g*, respectively *h* to *c*. Therefore it is possible to solve the nonlinear system of equation numerically and for some cases even analytically.[46,47] The nonlinear system of equations consisting of the closure and Ornstein-Zernike relation is separable into three subsystems for a pure solvent (*vv*), an infinitely diluted solution (*uv*) and a mixture of solutes in an infinitely diluted solution (*uu*). A possible approach to solve these separated equations is elucidated in the next subchapter.

1.3.2 The reference interaction site model

The system of equations consisting of the closure relation (10) and the Ornstein-Zernike equation (11) that was presented in the previous subchapter is only applicable to simple spherical particles. A generalization of the Ornstein-Zernike equation to complex molecular fluids is the molecular Ornstein Zernike relation[48]

$$h(\mathbf{r}_{\alpha\gamma},\mathbf{\Omega}_{\alpha},\mathbf{\Omega}_{\gamma}) = c(\mathbf{r}_{\alpha\gamma},\mathbf{\Omega}_{\alpha},\mathbf{\Omega}_{\gamma}) + \frac{\rho}{8\pi^2}\int\int c(\mathbf{r}_{\alpha\gamma'},\mathbf{\Omega}_{\alpha},\mathbf{\Omega}_{\gamma'})h(\mathbf{r}_{\gamma'\gamma},\mathbf{\Omega}_{\gamma'},\mathbf{\Omega}_{\gamma})\mathrm{d}\mathbf{r}_{\gamma'}\mathrm{d}\mathbf{\Omega}_{\gamma'} \tag{14}$$

that incorporates not only the distance between two molecules, but also the relative orientation, described by the Euler angles $\mathbf{\Omega}_i$. Because of its high dimensionality and difficult solvability, Chandler and Anderson introduced a new approach for molecular fluids,

circumventing the full Molecular Ornstein-Zernike equation, the so called site-site OZ or 1D reference interaction site model (1D RISM).[17,18] Here, a molecule consists of explicit sites with intramolecular distances l_{ij}. In consequence a new intramolecular distribution function, the so-called intramolecular correlation ω can be introduced that is defined as

$$\omega(r_{\alpha\gamma}) = \frac{\delta(|\, r_{\alpha\gamma} - l_{\alpha\gamma}\,|)}{4\pi l_{\alpha\gamma}^2} \tag{15}$$

for rigid molecules with $\delta(x)$ as Dirac delta function. One proper way to treat molecules is elucidated by Chandler et al. using the so called 'rational closure'. [49] It passes over the united atom limit, defines an intramolecular correlation function beyond the rigid molecule ansatz and circumvents inconsistencies in the closure/OZ system of equations by directly formulating a self-consistent approach. Furthermore Cortis, Rossky and Friesner showed that these rational closures are connected to Boltzmann weighted averages of particle distances over molecular orientations.[50] The intramolecular correlation restricts the total correlation h and in consequence the pair distribution function g, to reproduce the given intramolecular structure, by

$$\mathbf{h} = \boldsymbol{\omega} * \mathbf{c} * \boldsymbol{\omega} + \boldsymbol{\omega} * \mathbf{c} * \boldsymbol{\rho}\mathbf{h}\,. \tag{16}$$

Therefore, on one hand, it is not necessary to deal with Euler angles between molecules, but on the other hand it is still possible to treat complex molecules. Equation (16) consists of the matrix indices $\mathbf{h} = (h_{\alpha\gamma}(r_{\alpha\gamma}))^{uv}$, $\mathbf{c} = (c_{\alpha\gamma}(r_{\alpha\gamma}))^{uv}$ and $\boldsymbol{\rho} = ((\rho_\alpha)^u,(\rho_\gamma)^v)$ with α being one distinct site related to the molecules of type u and γ as one distinct site related to molecules of type v. In contrast to simple spherical particles, all molecules can have a variety of different sites α or γ, which are commonly equal to the molecules' nuclei. A closed solution of the 1D RISM approach is accessible by application of Percus' idea and redefine the molecule u to be part of the v molecules or, in a more physical picture, to define the sol**u**te (u) to be part of the sol**v**ent (v). Afterwards a sub-equation of (16), the pure solvent or 1D-vv RISM

$$\mathbf{h}^{vv} = \boldsymbol{\omega}^v * \mathbf{c}^{vv} * \boldsymbol{\omega}^v + \boldsymbol{\omega}^v * \mathbf{c}^{vv} * \boldsymbol{\rho}^v\mathbf{h}^{vv} \tag{17}$$

is solvable after choosing a closure, an intermolecular pair potential U and the macroscopic solvent density of each site $\boldsymbol{\rho}^v = (\rho_\gamma)^v$. The solvent-solvent total correlation function $\mathbf{h}^{vv}$ is then applicable to estimate the solvent surroundings around a solute in an infinitely diluted solution with $\boldsymbol{\rho}^u \rightarrow 0$ by

$$\mathbf{h}^{uv} = \boldsymbol{\omega}^u * \mathbf{c}^{uv} * \boldsymbol{\omega}^v + \boldsymbol{\omega}^u * \mathbf{c}^{uv} * \boldsymbol{\rho}^v\mathbf{h}^{vv}, \tag{18}$$

which is also known as 1D-*uv* RISM equation. After definition of a new correlation function, the solvent susceptibility function

$$\boldsymbol{\chi} = \boldsymbol{\rho}^{v}\boldsymbol{\omega}^{v} + \boldsymbol{\rho}^{v}\mathbf{h}^{vv}\boldsymbol{\rho}^{v}, \tag{19}$$

the 1D-*uv* RISM equation can be reduced to

$$\mathbf{h}^{uv} = \boldsymbol{\omega}^{u} * \mathbf{c}^{uv} * (\boldsymbol{\rho}^{v})^{-1}\boldsymbol{\chi}. \tag{20}$$

The study of charged and polar solutes and solvents with RISM was invented first by Hirata and Rossky with their extended RISM methodology, XRISM[51] and further developed some years later by Perkyns and Pettitt to a dielectrically consistent approach, called DRISM.[52,53] Applying the DRISM formalism, for which the pure solvent dielectric constant and the molecular dipole moment of the solvent molecule must be given, especially the pair distribution functions of solutes in finite-concentration salt solutions could be enhanced without empirically fitted parameters.[52]

1.3.3 3D RISM

A three dimensional expansion of the RISM formalism, 3D RISM, was not developed before the mid-90th by Ikeguchi and Doi[54] as well as Beglov et al.[55,56] and Kovalenko and Hirata.[21] In contrast to the 1D RISM approach solution of 3D RISM results in an anisotropic solvent total correlation function

$$h_{\gamma}(\mathbf{r}) = \sum_{\gamma'}(\rho_{\gamma'})^{-1}\int c_{\gamma'}(\mathbf{r}-\mathbf{r}')\chi_{\gamma\gamma'}(|\mathbf{r}'|)\mathrm{d}\mathbf{r}' \tag{21}$$

around an infinitely dissolved solute. According to the 1D-*uv* RISM approach, here the 3D RISM equation (21), in combination with an appropriate closure, is solvable with a proper precomputed solvent susceptibility function. Thus commonly the 1D-*vv* RISM result is selected. Furthermore the three dimensional HNC closure[57]

$$h_{\gamma}(\mathbf{r}) = \exp(-\beta U_{\gamma}(\mathbf{r}) + h_{\gamma}(\mathbf{r}) - c_{\gamma}(\mathbf{r})) - 1 \tag{22}$$

or the numerically more stable partial series expansion of order *k* (PSE-*k*)[58] set of closures

$$h_{\gamma}(\mathbf{r}) = \begin{cases} \sum_{i=0}^{k}\left(-\beta U_{\gamma}(\mathbf{r}) + h_{\gamma}(\mathbf{r}) - c_{\gamma}(\mathbf{r})\right)^{i}/i! - 1 & \Leftrightarrow \quad -\beta u_{\gamma}(\mathbf{r}) + h_{\gamma}(\mathbf{r}) - c_{\gamma}(\mathbf{r}) > 0 \\ \exp[-\beta U_{\gamma}(\mathbf{r}) + h_{\gamma}(\mathbf{r}) - c_{\gamma}(\mathbf{r})] - 1 & \Leftrightarrow \quad -\beta u_{\gamma}(\mathbf{r}) + h_{\gamma}(\mathbf{r}) - c_{\gamma}(\mathbf{r}) \le 0 \end{cases} \tag{23}$$

that gradually approach the HNC solution, are usually applied. The inversed temperature is indicated here as $\beta = 1 / k_B T$. Furthermore $U_\gamma(\mathbf{r})$ describes the total interaction potential between the solute and the solvent site γ on the spatial point $\mathbf{r}$. Therefore an interaction potential of the form

$$U_\gamma(\mathbf{r}) = \sum_\alpha \left(\frac{1}{4\pi\varepsilon_0} \frac{q_\alpha q_\gamma}{|\mathbf{r}-\mathbf{R}_\alpha|} + 4\varepsilon_{\alpha\gamma} \left(\frac{\sigma_{\alpha\gamma}^{12}}{|\mathbf{r}-\mathbf{R}_\alpha|^{12}} - \frac{\sigma_{\alpha\gamma}^{6}}{|\mathbf{r}-\mathbf{R}_\alpha|^{6}} \right) \right) \tag{24}$$

is commonly applied with ε_0 as the dielectric permittivity of the vacuum and q_γ as well as q_α as partial charge of the corresponding solvent and solute site, the latter which is in general localized on the solute site coordinates $\mathbf{R}_\alpha$. In (24) the van der Waals interactions are described by a Lennard-Jones potential with $\varepsilon_{\alpha\gamma}$ interpretable as potential depth and $\sigma_{\alpha\gamma}$ as contact distance, both usually taken from molecular mechanics force fields. The electrostatic point charge part U_γ^{elec} of the potential is in general divided into

$$U_\gamma^{\text{elec}}(\mathbf{r}) = U_\gamma^{\text{elec,S}}(\mathbf{r}) + U_\gamma^{\text{elec,L}}(\mathbf{r}) \tag{25}$$

with a short range part

$$U_\gamma^{\text{elec,S}}(\mathbf{r}) = \sum_\alpha \frac{1}{4\pi\varepsilon_0} \frac{q_\alpha q_\gamma}{|\mathbf{r}-\mathbf{R}_\alpha|} \operatorname{erfc}(\kappa |\mathbf{r}-\mathbf{R}_\alpha|) \tag{26}$$

and a long range part

$$U_\gamma^{\text{elec,L}}(\mathbf{r}) = \sum_\alpha \frac{1}{4\pi\varepsilon_0} \frac{q_\alpha q_\gamma}{|\mathbf{r}-\mathbf{R}_\alpha|} \operatorname{erf}(\kappa |\mathbf{r}-\mathbf{R}_\alpha|) \tag{27}$$

following the Ewald summation idea[59] to properly treat long range effects without truncation errors. The latter is evaluated in the reciprocal space after Fourier transformation for which it gets short ranged by a properly chosen smearing factor κ.

As mentioned in chapter 1.3.1 the classical density theory approaches connect the solvent density with the free energy of the system. In principle, the free energy is also accessible for various approximations of the molecular Ornstein-Zernike equation, as well as for the 1D RISM[60] and 3D RISM[61,62] theory by

$$\begin{aligned} \left(\frac{\partial A}{\partial N} \right)_{V,T} &= \mu^{\text{id}} + \mu^{\text{ex,HNC}} + \mu^{\text{ex,Bridge}} \\ &= \beta^{-1} \ln \rho\Lambda^3 + \beta^{-1}\rho \int \mathrm{d}\mathbf{r} \left[\left(\frac{1}{2}h^2 - c - \frac{1}{2}hc \right) + \int_0^1 \mathrm{d}\lambda (h(\mathbf{r};\lambda)+1) \frac{\partial B(\mathbf{r};\lambda)}{\partial \lambda} \right]. \end{aligned} \tag{28}$$

In equation (28) the chemical potential μ is separated into its ideal part μ^{id} that depends on the temperature, the solvent density and the thermal wavelength Λ, thus easily computable, and the excess chemical potential μ^{ex}. The excess chemical potential inherits the problematic part, because a solution is related to the chosen bridge function and sometimes not numerically computable or not path independent.[61] In case of the HNC closure the bridge function related part is simply zero and for the PSE-*k* closures[58]

$$\mu^{\text{ex,Bridge}} = \mu^{\text{ex,PSE-}k} = -\beta^{-1}\rho\int \mathrm{d}\mathbf{r}\,\Theta(h)(-\beta U + h - c)/(k+1)! \tag{29}$$

with $\Theta(x)$ as Heaviside step function.

1.3.4 Embedded cluster RISM approach for quantum chemistry in solution

The 1D-*uv* RISM and 3D RISM approaches couple the solute-solvent interactions with the solvent distribution, which implicitly incorporates the solvent-solvent interactions through its solvent susceptibility function. If we for example consider the transfer of a rigid molecule with a hypothetically unpolarizable electron density from vacuum to solution the excess chemical potential calculated with RISM would be equivalent to the solvation free energy ΔG_{solv}. But what is missing for realistic free energy estimation with RISM is the intramolecular interaction of the solute and its change due to the solvent surrounding. A coupling of different approaches for inter- and intramolecular interactions was done, for example with classical molecular mechanics force fields[63,64] or with polarizable force fields like “atomic multipole optimized energetics for biomolecular applications” approach AMOEBA.[65,66] But more detailed studies require a quantum-mechanical picture of the solute of interest. Therefore, based on the variational principle for both, the electronic energy of the solute and the free energy of the whole solution, a self-consistent method has been introduced for 1D RISM by Ten-no et al. (RISM-SCF)[67] and was extended by explicit treatment of the spatial electron density distribution (SEDD) by Sato et el. (RISM-SCF-SEDD).[68] A combination of 3D RISM with quantum-chemical DFT[69] and with *ab initio* molecular orbital theory (3D RISM-SCF) was used in the workgroup of Hirata.[70] A more elaborate combination is to couple the MOZ equation with quantum-mechanics, which was done by Yoshida and Kato for very simple cases[71] and extended by the multi-center MOZ[72,73] approach by Kido et al. (MC-MOZ-SCF).[74]

An alternative approach that has not to be variational, but which is computationally cheap and easily implemented with every kind of quantum-chemical codes, is the embedded cluster RISM methodology, EC-RISM. The EC-RISM strategy was first applied for partial charge optimization in solid-state environment[75,76] and then further developed for infinite dilutions by Kloss et al.[23] The first step of EC-RISM is the calculation of the electronic wave function or electron density and the corresponding quantum-mechanical energy E_{sol} of the solute in vacuum. Afterwards, taking the vacuum result, the molecular electrostatic potential $\varphi(\mathbf{r})$ is generated and used to estimate partial charges q_α on the solute sites localized on the solute site coordinates $\mathbf{R}_\alpha$. These partial charges are then subsequently used to calculate the solute-solvent interaction potential of a first solvent density approximation with 3D RISM. Therefore a classical force field interaction potential approach of the form (24)-(27) is usable.

Alternatively the exact quantum-mechanical electrostatic potential can be implemented by estimating

$$U_\gamma^\varphi(\mathbf{r}) = q_\gamma \varphi(\mathbf{r}) \tag{30}$$

for the electrostatic solvent-solute interaction. In this case the classical Coulomb potential with Ewald splitting is evaluated to account for long range electrostatics. The total solute-solvent interaction is approximated by

$$\begin{aligned} U(\mathbf{r}) &= U_\gamma^\varphi(\mathbf{r}) + U_\gamma^{\mathrm{LJ}}(\mathbf{r}) \\ &= U_\gamma^{\mathrm{elec,S}}(\mathbf{r}) + \Delta U_\gamma^{\varphi,\mathrm{elec}}(\mathbf{r}) + U_\gamma^{\mathrm{elec,L}}(\mathbf{r}) + U_\gamma^{\mathrm{LJ}}(\mathbf{r}), \end{aligned} \tag{31}$$

for which $U_\gamma^{\mathrm{LJ}}(\mathbf{r})$ is the Lennard-Jones interaction energy and $\Delta U_\gamma^{\varphi,\mathrm{elec}}(\mathbf{r})$ is the difference between the exact quantum-mechanical electrostatic potential and the point charge representation.[66]

After convergence of the 3D RISM/closure system of equations the sum

$$A_{\mathrm{sol}}^*(V,T) = E_{\mathrm{sol}} + \mu^{\mathrm{ex}} \tag{32}$$

between the quantum-chemical energy of the solute in solution E_{sol} and the excess chemical potential μ^{ex} defines a first estimation of a per particle free energy of the solute A^*_{sol} in solution. It can be connected to the solvation free energy of one solute particle

$$\Delta A_{\mathrm{solv}}^*(V,T) = A_{\mathrm{sol}}^*(V,T) - A_{\mathrm{gas}}^*(V,T) = \mu_{\mathrm{solv}}^*(V,T) - \mu_{\mathrm{gas}}^*(V,T) \tag{33}$$

in the Ben-Naim standard state,[77,78] which is indicated by the star and corresponds to the hypothetical transfer process of a molecule from a 1 M ideal gas state to a hypothetical 1 M

solution at infinite dilution. In eq. (33) $A^*_{gas}(V,T)$ is the free energy of the solute in the gas phase, $\mu^*_{solv}(V,T)$ is the chemical potential in solution and $\mu^*_{gas}(V,T)$ in gas phase. At thermodynamic equilibrium conditions, due to an equal difference of chemical potentials, eq. (33) is equivalent to the Gibbs free energy of solvation of a solute molecule

$$\Delta A^*_{solv}(V,T) = \Delta G^*_{solv}(p,T) = \mu^*_{solv}(p,T) - \mu^*_{gas}(p,T) \tag{34}$$

in the Ben-Naim standard state.[79]

After the first EC-RISM step the solute has polarized the solvent, but a self-consistent polarization between the solute and the solvent is not achieved, because the solvent has no influence on the solute so far. Therefore a background charge density that represents the solvent electrostatics is calculated from the solvent distribution function by

$$\rho_q(\mathbf{r}) = \sum_\gamma q_\gamma \rho_\gamma g_\gamma(\mathbf{r}) . \tag{35}$$

and can be discretized on an explicit grid, giving a set of background charges

$$q(\mathbf{r}_i) = \int_{V_i} \rho_q(\mathbf{r})\, \mathrm{d}\mathbf{r} \approx \rho_q(\mathbf{r}_i) V_i , \tag{36}$$

with V_i as volume of one grid cell i. These are used to calculate the Coulomb potential between $\rho_q(\mathbf{r})$ and the solute nuclei and electrons, described by the Hamiltonian $\hat{H}_{uv}$, which is then added to the common quantum-mechanical Hamiltonian of the solute $\hat{H}_{sol}$. This a a new Hamiltonian of the whole system is defined by

$$\hat{H}_{tot} = \hat{H}_{sol} + \hat{H}_{uv} . \tag{37}$$

Solving the Schrödinger equation with this Hamiltonian provides an approximation of the total wave function ψ_{tot} and the total energy E_{tot} of the system. Here the focus remains on the change of the electronic energy of the solute by solvent polarization and on that account the interaction between the solvent charge density and the molecule of interest has to be evaluated by

$$E_{uv} = \int \mathrm{d}\mathbf{r} \rho_q(\mathbf{r}) \varphi(\mathbf{r}) . \tag{38}$$

In the following it is subtracted from E_{tot}, giving us a new solute electronic energy

$$E_{sol} = E_{tot} - E_{uv} . \tag{39}$$

Afterwards a new estimation of the solvent surroundings is carried out and the previous steps are repeated until self-consistency for the free energy is reached providing us with a

final per particle free energy of the solute that is again estimated by eq. (32). In the following, if not further specified the designator (*) for the standard state will be dropped and the Ben-Naim standard state is applied. Furthermore in analogy to eq. (34) a per particle free energy of the solute molecule in infinite dilution is defined as

$$G_{\mathrm{sol}} = E_{\mathrm{sol}} + \mu^{\mathrm{ex}} . \tag{40}$$

2 SOLVATION EFFECTS ON CHEMICAL SHIFTS BY EMBEDDED CLUSTER INTEGRAL EQUATION THEORY[i]

2.1 Introduction

Chemical shifts are key quantities obtained from nuclear magnetic resonance (NMR) spectroscopy to determine chemical structure. With increasing molecular size, conformational freedom, and chemical complexity the task to assign spectral lines to nuclei becomes even more difficult, which gives rise to increasing demand for accurate computational prediction. Empirical group contribution assignment is not sufficient if, for instance, detailed conformational information has to be deduced from spectra, even if more elaborate NMR techniques are employed. Quantum-chemical calculations are in principle capable of filling the gap since NMR parameters are observables that can be inferred from a compound's wave function. Since the first practical application of quantum-chemically computed chemical shifts by Schindler and Kutzelnigg[80] *ab initio* and density functional theory (DFT) calculations have provided meaningful support for the interpretation of complex spectra for example of small molecules[81,82] or even of peptides.[83] While certainly much more time consuming than common empirical NMR spectra simulation programs,[84] quantum chemistry allows for the

[i] Reproduced in part with permissions from R. Frach, S. M. Kast, *J. Phys. Chem. A* **2014**, *118*, 11620. Copyright 2014 American Chemical Society.

investigation of molecular conformations and their surroundings on NMR parameters under controlled conditions and without parametrization.

However, accurate quantum-chemical predictions of measured chemical shifts (and of course other properties such as coupling constants) suffer from several sources of error.[85] Besides limitations of basic quantum-mechanical factors such as basis sets and levels of theory, these inaccuracies result from uncertainties about conformational ensemble averages that have to be considered[86,87] as well as from the necessity to describe solute-solvent interactions in a condensed phase environment appropriately. In particular, the presence and properties of a solvent directly influences chemical shielding constants by perturbing the vacuum electron density of a solvated molecule,[88] and indirectly by modulating the conformational distribution of the molecule of interest.[89] Therefore, quantum-chemical predictions of chemical shifts require not only an adequate description of the solute alone, but also of the surrounding solvent and its interaction with the solute. Continuum solvation models based on the dielectric response such as the polarizable continuum model (PCM)[13] are commonly used for the description of polarization effects and typically improve chemical shift calculations when directional interactions are of minor importance.[90] Beyond chemical shift predictions, combinations of conformational analysis with PCM-based NMR coupling constant calculations can guide the interpretation of complex experiments.[91] However, PCM (and other self-consistent reaction field approaches) ignore directional interactions, higher-order electric multipoles, or dynamic solvation effects and can therefore not fully account for observable solvent contributions.[92] As an alternative, explicit solvent molecules can be added, for instance within quantum-mechanical/classical-mechanical (QM/MM) models,[93] by fragmentation schemes, like the adjustable density matrix assembler (ADMA),[94,95] or by combining both, as in the automatic fragmentation QM/MM[96] model that can be combined with classical or *ab initio* molecular dynamics (MD) simulations[87] to deal with the conformational sampling problem. Explicit solvent calculations dramatically enhance chemical shift predictions,[87] but are not routinely used because of their computational demand particularly when coupled with simulations.

As alternative to continuum models, integral equation based theories can be used for describing solvent distributions around solute molecules on a molecular level, since the granular solvent properties are preserved in this case. One computationally manageable way is to employ the reference interaction site model (RISM)[17,18] integral equation theory, more specifically in the form of the three-dimensional (3D RISM)[97,98] approximation, that yields

solvent-site distribution functions around arbitrarily shaped solutes. One way to couple this representation of the solvent structure with quantum-chemical calculations is to map the resulting solvent-charge distribution onto discrete, embedding point charges that represent an additional component of the solute's electronic Hamiltonian. In this way, the solvent distribution polarizes the solutes, which in turn exhibits a solvent-modulated electrostatic potential on the solvent sites giving rise to a change of the solvent distribution by solving the 3D RISM equations for the modified solute-solvent potential. Repeating these steps until self-consistency between solute's electronic and solvent structure is obtained establishes the "embedded cluster reference interaction site model" (EC-RISM).[23] This approach was developed for predicting thermodynamic solvation properties since RISM theory yields analytical expressions for the excess chemical potential that, by adding the electronic energy, allows for the determination of the solute's free energy G in solution. EC-RISM utilizes common molecular solvent models such that the granular and directional nature of solute-solvent interactions is preserved. The approach was successfully applied to conformational equilibria and pK_a shifts in water,[23] to tautomeric equilibria within the SAMPL2 predictions contest,[24] and to pK_a predictions in dimethyl sulfoxide (DMSO).[27] In all cases examined so far, a close agreement between theory and experiment is found.

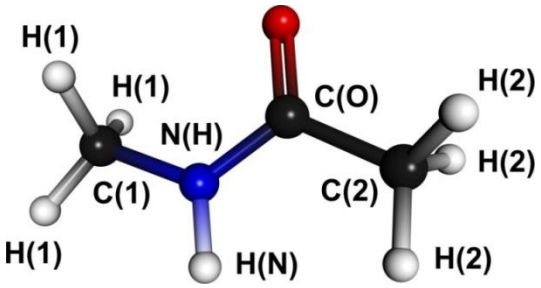


Figure 2.1 Structure of N-methylacetamid (NMA) with nuclear designators used throughout.

Since in the past the EC-RISM methodology was only tested for the sum of electronic energy and chemical potential, there was so far no possibility to determine whether the promising results were the consequence of a fortuitous cancellation of error. As a stringent test of the quality of the electronic structure component, the approach is extended to magnetic shielding calculations. This facilitates the computation of isotropic chemical shifts

$$\delta_i = \sigma_{\text{ref}} - \sigma_i \tag{41}$$

by subtracting the isotropic nuclear magnetic shielding σ_i of the nucleus of interest i from its reference value σ_{ref}. To this end, the EC-RISM procedure is easily augmented by adding a

GIAO[26] (gauge invariant atomic orbital) nuclear magnetic shielding determination step to the final electronic structure calculation after self-consistency has been reached. As a proof of principle, the procedure is applied to the prediction of the chemical shifts of all nuclei (except for oxygen for which the presently applied level of theory does not yield satisfactory results with nearly an order of magnitude larger deviations as shown by Auer[99]) of the small bioorganic building block *N*-methylacetamide (NMA) in aqueous solution (for nuclear designators, see Figure 2.1) and examine the impact of various approximations to the 3D RISM theory as well as the effect of varying the basis set size and the accuracy of treating solute-solvent electrostatic interactions.

2.2 Computational details

GIAO calculations were applied to both optimized vacuum and solution phase geometries, in the latter case by using PCM calculation results also for EC-RISM postprocessing, as was shown earlier to be a reasonable approximation.[24] For comparison, GIAO results were also obtained for PCM wave functions. Following on one hand standard procedures for the quantum-chemical part, the B3LYP hybrid functional[100,101,102] and the 6-31+G(d) basis set[103] within the Gaussian 03[104] program package (Rev. D.02) were throughout employed, using also its implementation of the CHELPG procedure[105] for the determination of point charges and for the 3D RISM part the PSE closures with orders 1 to 4 and the HNC closure. On the other hand, the performance of the larger aug-cc-pVTZ[106,107] basis set in conjunction with B3LYP was tested. Additionally, beside a point charge representation for the solute-solvent interaction, EC-RISM solvent structure calculations were performed with the solute's exact electrostatic potential to examine the solute-solvent interaction. Furthermore the performance of the basis set size was investigated using a series of the Pople basis sets (6-31G(d), 6-31+G(d), 6-31+G(d,p), 6-311++G(d,p)) in combination with only PSE-1 and HNC closure relations and point charge electrostatics. Except for the nuclear magnetic shielding calculation *in vacuo* the minimum energy structure studied was generated by geometry optimization with the "integral equation formalism" (not to be confused with integral equation theory) IEF-PCM model[13] in water using default atomic parameters. The PCM free energy in solution was calculated as the sum of electrostatic and nonelectrostatic terms. Only the *trans* conformation depicted in Figure 2.1 was taken into account since the *cis* form obtained from rotating around the N(H)-C(O) bond was found to be energetically disfavored by 1.9 kcal mol^{-1} by

B3LYP/6-31+G(d)/PCM optimization, therefore essentially not contributing to a thermal ensemble at ambient temperature.

The chemical shielding estimation was done by the secondary standard methodology

$$\delta_i = \sigma_{\text{sec,ref}} - \sigma_i + \delta_{\text{sec,ref}}\,, \tag{42}$$

where the calculated nuclear magnetic shielding reference in water, $\sigma_{\text{sec,ref}}$, is shifted by the chemical shift $\delta_{\text{sec,ref}}$ of the reference compound to common primary standards. Therefore we use the nuclear magnetic shielding values of TMS (tetramethylsilane) for ^{1}H, NH_3 for ^{15}N and 1,4-dioxane for ^{13}C in aqueous solution, thus avoiding discrepancies due to inconsistent solvent models. The nuclear magnetic shielding of TMS, NH_3 and 1,4-dioxane in aqueous solution were used as secondary standards for ^{1}H, ^{15}N and ^{13}C with chemical shifts of -0.094 ppm[108], -19.4 ppm[109] and 66.6 ppm[110] to TMS in $CDCl_3$ and neat NH_3. The solvation model for the reference calculations were used throughout corresponding to the NMA solvation model. The proton chemical shifts of both methyl groups were estimated by averaging over the corresponding ^{1}H chemical shifts. To describe repulsion-dispersion interactions within the EC-RISM solvation model between the solvent and the solute we used the 12-6 Lennard-Jones potential with the "General Amber Force Field" GAFF (version 1.4, March 2010)[111,112] parameters, and parameters from Makrodimitri et al.[113] for the silicon of TMS (Table 2.1).

Table 2.1 Lennard-Jones parameters of all solute species.

	σ / Å	ε / cal mol^{-1}
C(1)	3.400	109.4
H(1)	2.471	15.7
N(H)	3.250	170.0
H(N)	1.069	15.7
C(O)	3.400	86.0
O(C)	2.960	210.0
C(2)	3.400	109.4
H(2)	2.650	15.7
TMS	σ / Å	ε / cal mol^{-1}
Si	3.385	585.0
C	3.400	109.4
H	3.650	15.7
NH_3	σ / Å	ε / cal mol^{-1}
N	3.250	170.0
H	1.069	15.7
1,4-dioxane	σ / Å	ε / cal mol^{-1}
C	3.400	109.4
O	3.000	170.0
H	2.471	15.7

Procedural parameters for EC-RISM calculations at 298.15 K follow closely earlier studies[24] in particular, we used a grid of 104^3 points with 0.3 Å spacing for the 3D RISM solvent structure estimation and a convergence threshold of 10^{-5} kcal mol^{-1} in the Gibbs free energy of the EC-RISM procedure. As in earlier work,[24] water is described by a modified variant of the SPC/E[114] model. All computational predictions were compared with experimental data taken from Exner et al.[115]

2.3 Results and discussion

2.3.1 Statistics

Based on the calculated raw shielding data summarized in Tables S1-S3 in the appendix isotropic chemical shifts were computed, listed in Table 2.2-Table 2.4 and displayed in Figure 2.2 and Figure 2.3. Note that the point charge approach is indicated by the superscript "*q*", whereas the exact electrostatics approach is denoted by superscript "φ" throughout.

As expected (see Table 2.2 and Figure 2.2), the presence and adequate treatment of the solvent has the most substantial impact on the amide proton H(N), the amide carbon C(O), and the nitrogen nucleus N(H), whereas only modest influence on the methyl groups is found. Looking first at a comparison of all solvation methods under investigation for the 6-31G+(d) basis set, we find that the chemical shifts of both methyl group carbons deviate only by around -1.54 to -1.99 ppm for C(1) and -3.58 to -2.22 ppm for C(2). For this basis set the ^{1}H nuclei of the methyl groups show a similar trend with a deviation of around -0.06 to 0.20 ppm for H(1) and -0.34 to 0.16 ppm for H(2). Regarding the aug-cc-pVTZ results (see Figure 2.2 and Table 2.3), the methyl group chemical shifts can be considered quite insensitive, whereas the deviation range for the C(1) and C(2) nuclei is slightly greater. This insensitivity to the presence of a solvent was also captured in other calculations[115] for NMA, suggesting that interaction with solvent molecules is not significant for these nuclei. In contrast, chemical shifts of the amide nuclei depend heavily on the solvation modeling methodology employed. The EC-RISM model cannot account for partial electron transfer between solute and solvent. Technically, this would require sampling of solvent configurations, implying dramatically larger effort. However, based on conclusions drawn by Exner and co-workers[87], electron transfer does not significantly influence chemical shifts.

Table 2.2 Chemical shifts (in ppm) calculated by GIAO/B3LYP/6-31+G(d) in the gas phase (with vacuum-optimized structures) and using the solvation models PCM and EC-RISM (PSE-[1-4] and HNC) with optimized PCM structures compared with experimental values (deviations are shown in parenthesis). Superscript "*q*" indicates the point charge model for electrostatic solute-solvent interactions while superscript "φ" refers to the exact quantum-mechanical electrostatic potential. Statistical quality is measured by the root mean square deviation (RMSD) and the mean signed error (MSE).

	C(1)	H(1)	N(H)	H(N)	C(O)	C(2)	H(2)	RMSD	MSE
Literature	28.65	2.63	114.24	7.73	177.24	24.32	1.90		
Vacuum	27.06 (-1.59)	2.56 (-0.07)	89.82 (-24.42)	4.20 (-3.53)	158.73 (-18.51)	20.74 (-3.58)	1.56 (-0.34)	11.75	-7.43
PCM	26.81 (-1.84)	2.57 (-0.06)	103.16 (-11.09)	6.18 (-1.55)	164.25 (-12.99)	21.00 (-3.32)	1.70 (-0.20)	6.64	-4.43
PSE-1^q	27.09 (-1.56)	2.81 (0.18)	118.15 (3.91)	6.55 (-1.18)	168.17 (-9.07)	21.93 (-2.39)	2.02 (0.12)	3.91	-1.43
PSE-1$^\varphi$	26.76 (-1.89)	2.80 (0.17)	114.36 (0.12)	6.54 (-1.19)	167.99 (-9.25)	22.04 (-2.28)	2.01 (0.11)	3.70	-2.03
PSE-2^q	27.10 (-1.55)	2.82 (0.19)	121.16 (6.92)	6.74 (-0.99)	168.78 (-8.46)	22.00 (-2.32)	2.05 (0.15)	4.28	-0.87
PSE-2$^\varphi$	26.71 (-1.94)	2.82 (0.19)	116.18 (1.94)	6.74 (-0.99)	168.57 (-8.67)	22.09 (-2.23)	2.04 (0.14)	3.56	-1.65
PSE-3^q	27.10 (-1.55)	2.83 (0.20)	122.16 (7.92)	6.82 (-0.91)	168.99 (-8.25)	22.02 (-2.30)	2.06 (0.16)	4.46	-0.68
PSE-3$^\varphi$	26.69 (-1.96)	2.82 (0.19)	116.69 (2.45)	6.82 (-0.91)	168.76 (-8.48)	22.11 (-2.21)	2.05 (0.15)	3.54	-1.54
PSE-4^q	27.11 (-1.54)	2.83 (0.20)	122.54 (8.30)	6.84 (-0.89)	169.06 (-8.18)	22.03 (-2.29)	2.06 (0.16)	4.54	-0.61
PSE-4$^\varphi$	26.68 (-1.97)	2.82 (0.19)	116.83 (2.59)	6.86 (-0.87)	168.82 (-8.42)	22.11 (-2.21)	2.05 (0.15)	3.53	-1.51
HNC^q	27.11 (-1.54)	2.83 (0.20)	122.75 (8.51)	6.86 (-0.87)	169.10 (-8.14)	22.03 (-2.29)	2.06 (0.16)	4.58	-0.57
HNC$^\varphi$	26.66 (-1.99)	2.82 (0.19)	116.84 (2.60)	6.89 (-0.84)	168.85 (-8.39)	22.10 (-2.22)	2.05 (0.15)	3.52	-1.50

The largest discrepancy to experimental values is observed for vacuum calculations. The PCM implicit solvent approach improves the results, but, for the smaller 6-31+G(d) basis set, only to an extent where the deviations to experimental values are still off at least by -1.55 ppm for H(N), -11.09 ppm for N(H) and -12.30 ppm for C(O). For all these nuclei the EC-RISM solvation model combined with B3LYP/6-31+G(d) level of theory further improves the chemical shift predictions, leading to deviations of 8.51 ppm for N(H) and -9.25 ppm for C(O) in the worst case. Furthermore the EC-RISM approach with the HNC closure for this basis set reduces the amide proton shift error down to -0.84 ppm, which is nearly half the deviation of the best PCM result.

Table 2.3 Chemical shifts (in ppm) calculated by GIAO/B3LYP/aug-cc-pVTZ in the gas phase (with vacuum-optimized structures) and using the solvation models PCM and EC-RISM (PSE-[1-4] and HNC) with optimized PCM structures compared with experimental values (deviations are shown in parenthesis). Data for EC-RISM closure approximations PSE-4$^{\varphi}$ and HNC$^{\varphi}$ and the large basis set have been omitted due to divergence of the corresponding 3D RISM solutions.

	C(1)	H(1)	N(H)	H(N)	C(O)	C(2)	H(2)	RMSD	MSE
literature	28.65	2.63	114.24	7.73	177.24	24.32	1.90		
vacuum	24.08 (-4.57)	2.65 (0.02)	97.84 (-16.40)	4.97 (-2.76)	168.40 (-8.84)	17.81 (-6.51)	1.70 (-0.20)	7.73	-5.61
PCM	23.57 (-5.08)	2.67 (0.04)	113.34 (-0.90)	6.84 (-0.89)	174.29 (-2.95)	18.41 (-5.91)	1.83 (-0.07)	3.19	-2.25
PSE-1^{q}	23.89 (-4.76)	2.73 (0.10)	125.48 (11.24)	7.00 (-0.73)	178.91 (1.67)	19.49 (-4.83)	1.99 (0.09)	5.01	0.40
PSE-1$^{\varphi}$	25.06 (-3.59)	2.90 (0.27)	124.37 (10.13)	7.23 (-0.50)	181.04 (3.80)	20.76 (-3.56)	2.16 (0.26)	4.52	0.97
PSE-2^{q}	23.91 (-4.74)	2.75 (0.12)	128.14 (13.90)	7.18 (-0.55)	179.72 (2.48)	19.60 (-4.72)	2.01 (0.11)	5.91	0.94
PSE-2$^{\varphi}$	25.00 (-3.65)	2.91 (0.28)	123.10 (8.86)	7.42 (-0.31)	181.74 (4.50)	20.85 (-3.47)	2.18 (0.28)	4.21	0.93
PSE-3^{q}	23.91 (-4.74)	2.75 (0.12)	129.04 (14.80)	7.24 (-0.49)	180.00 (2.76)	19.63 (-4.69)	2.02 (0.12)	6.22	1.13
PSE-3$^{\varphi}$	24.97 (-3.68)	2.91 (0.28)	119.72 (5.48)	7.50 (-0.23)	181.96 (4.72)	20.87 (-3.45)	2.19 (0.29)	3.34	0.49
PSE-4^{q}	23.92 (-4.73)	2.75 (0.12)	129.38 (15.14)	7.27 (-0.46)	180.10 (2.86)	19.65 (-4.67)	2.02 (0.12)	6.34	1.20
PSE-4$^{\varphi}$	24.95 (-3.70)	2.92 (0.29)	-	7.54 (-0.19)	182.02 (4.78)	20.87 (-3.45)	2.19 (0.29)	-	-
HNC^{q}	23.92 (-4.73)	2.75 (0.12)	129.57 (15.33)	7.28 (-0.45)	180.16 (2.92)	19.65 (-4.67)	2.02 (0.12)	6.41	1.24
HNC$^{\varphi}$	24.94 (-3.71)	2.92 (0.29)	-	7.57 (-0.16)	182.05 (4.81)	20.88 (-3.44)	2.19 (0.29)	-	-

This is a clear indication that directional solvent interactions that are properly resolved by EC-RISM theory are immediately responsible for the experimentally observed data. As also pointed out by Bader[116] and Exner et al.,[115] this solvent directionality is crucial for capturing nuclear magnetic shieldings of the amide groups. As illustrated a continuum description is insufficient when a small basis sets is applied (< 6-311++G(d,p)).

The results with the examined series of Pople basis sets (Table 2.4 and Figure 2.3) reveal that an increasing basis set size systematically increases the PCM and vacuum results for the N(H) and C(O) nuclei, whereas the H(N) chemical shift is improved up to the 6-31+G(d,p) basis and worsens for the largest employed Pople basis (6-311++G(d,p)). This indicates that the error compensation between the reference and the H(N) shielding, which is in general important[117] for chemical shift estimations with DFT methods, is not as effective for the H(N) nucleus with B3LYP/6-311++G(d,p) level of theory as for the other nuclei.

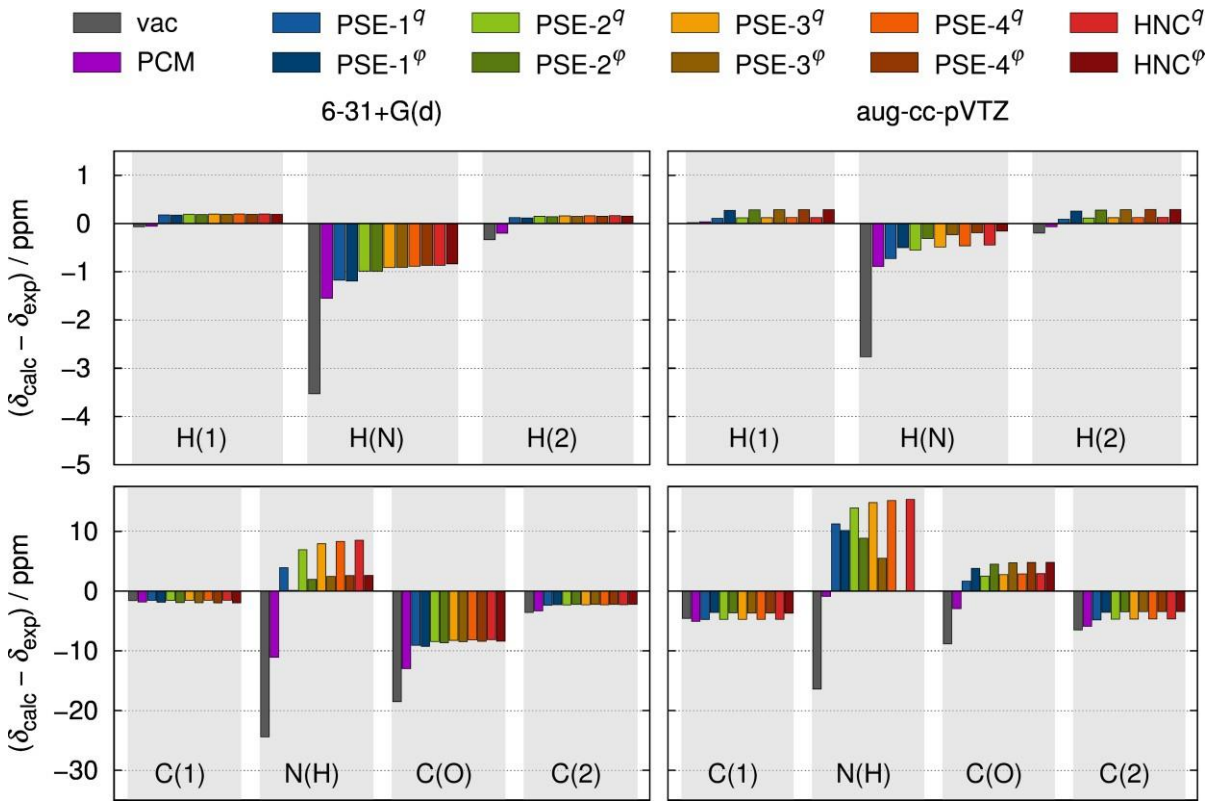


Figure 2.2 Nucleus-specific deviations between calculated chemical shifts δ_{calc} of NMA and experimental values[115] δ_{exp} from vacuum and various solution-phase calculations, employing GIAO/B3LYP with the basis sets 6-31+G(d) (left) and aug-cc-pVTZ (right). Superscript "*q*" indicates the point charge model for electrostatic solute-solvent interactions while superscript "φ" refers to the exact quantum-mechanical electrostatic potential.

A different behavior is observed for the EC-RISM results in combination with different Pople basis sets, for which the C(O) chemical shift improves using a larger basis, the N(H) drastically worsens and the H(N) shift shows a similar trend as for the PCM results. Overall the chemical shifts of the C(O) and N(H) increase with the amount and diffusity of the orbitals used, but the EC-RISM results, starting from a negative deviation to experiments with the 6-31+G(d) basis overshoot the experimental values, ending up at a much too high chemical shift for the C(O) and N(H) nuclei using the 6-311++G(d,p) and both examined closures. Therefore error cancellation is more effective in the EC-RISM^q case for smaller basis sets than for a larger 6-311++G(d,p).

Table 2.4 Chemical shifts (in ppm) calculated by GIAO/B3LYP and a series of Pople basis sets (6-31G(d), 6-31+G(d), 6-31+G(d,p), 6-311++G(d,p)) in the gas phase (with vacuum-optimized structures) and using the solvation models PCM and EC-RISM (PSE-1 and HNC) with optimized PCM structures compared with experimental values (deviations are shown in parenthesis).

	C(1)	H(1)	N(H)	H(N)	C(O)	C(2)	H(2)	RMSD	MSE
Literature	28.65	2.63	114.24	7.73	177.24	24.32	1.90		
Vacuum									
6-31G(d)	26.22 (-2.43)	2.62 (-0.01)	85.39 (-28.85)	3.81 (-3.92)	156.34 (-20.90)	24.11 (-0.21)	1.45 (-0.45)	13.58	-8.11
6-31+G(d)	27.06 (-1.59)	2.56 (-0.07)	89.82 (-24.42)	4.20 (-3.53)	158.73 (-18.51)	20.74 (-3.58)	1.56 (-0.34)	11.75	-7.43
6-31+G(d,p)	23.85 (-4.80)	2.83 (0.20)	96.35 (-17.89)	4.44 (-3.29)	159.15 (-18.09)	21.42 (-2.90)	1.59 (-0.31)	9.93	-6.73
6-311++G(d,p)	23.40 (-5.25)	2.55 (-0.08)	99.71 (-14.53)	4.46 (-3.27)	167.59 (-9.65)	17.64 (-6.68)	1.57 (-0.33)	7.44	-5.69
PCM									
6-31G(d)	29.46 (0.81)	2.48 (-0.15)	93.34 (-20.90)	5.72 (-2.01)	161.92 (-15.32)	23.48 (-0.84)	1.55 (-0.35)	9.83	-5.54
6-31+G(d)	26.81 (-1.84)	2.57 (-0.06)	103.16 (-11.09)	6.18 (-1.55)	164.25 (-12.99)	21.00 (-3.32)	1.70 (-0.20)	6.64	-4.43
6-31+G(d,p)	26.33 (-2.32)	2.61 (-0.02)	106.77 (-7.47)	6.70 (-1.03)	165.52 (-11.72)	20.35 (-3.97)	1.74 (-0.16)	5.55	-3.81
6-311++G(d,p)	24.89 (-3.76)	2.61 (-0.02)	108.95 (-5.29)	6.00 (-1.73)	177.22 (-0.02)	18.78 (-5.54)	1.95 (0.05)	3.29	-2.33
PSE-1[q]									
6-31G(d)	29.53 (0.88)	2.56 (-0.07)	106.47 (-7.77)	5.87 (-1.86)	165.58 (-11.66)	24.14 (-0.18)	1.70 (-0.20)	5.35	-2.98
6-31+G(d)	27.09 (-1.56)	2.81 (0.18)	118.15 (3.91)	6.55 (-1.18)	168.17 (-9.07)	21.93 (-2.39)	2.02 (0.12)	3.91	-1.43
6-31+G(d,p)	26.59 (-2.06)	2.68 (0.05)	119.91 (5.67)	6.90 (-0.83)	169.48 (-7.76)	21.31 (-3.01)	1.92 (0.02)	3.90	-1.13
6-311++G(d,p)	24.50 (-4.15)	2.79 (0.16)	130.14 (15.90)	6.86 (-0.87)	180.82 (3.58)	20.12 (-4.20)	2.05 (0.15)	6.56	1.51
HNC[q]									
6-31G(d)	29.52 (0.87)	2.57 (-0.06)	109.99 (-4.25)	6.11 (-1.62)	166.37 (-10.87)	24.22 (-0.10)	1.73 (-0.17)	4.47	-2.32
6-31+G(d)	27.11 (-1.54)	2.83 (0.20)	122.75 (8.51)	6.86 (-0.87)	169.10 (-8.14)	22.03 (-2.29)	2.06 (0.16)	4.58	-0.57
6-31+G(d,p)	26.61 (-2.04)	2.70 (0.07)	124.11 (9.87)	7.21 (-0.52)	170.43 (-6.81)	21.41 (-2.91)	1.96 (0.06)	4.73	-0.33
6-311++G(d,p)	24.51 (-4.14)	2.81 (0.18)	134.69 (20.45)	7.15 (-0.58)	182.02 (4.78)	20.28 (-4.04)	2.09 (0.19)	8.24	2.41

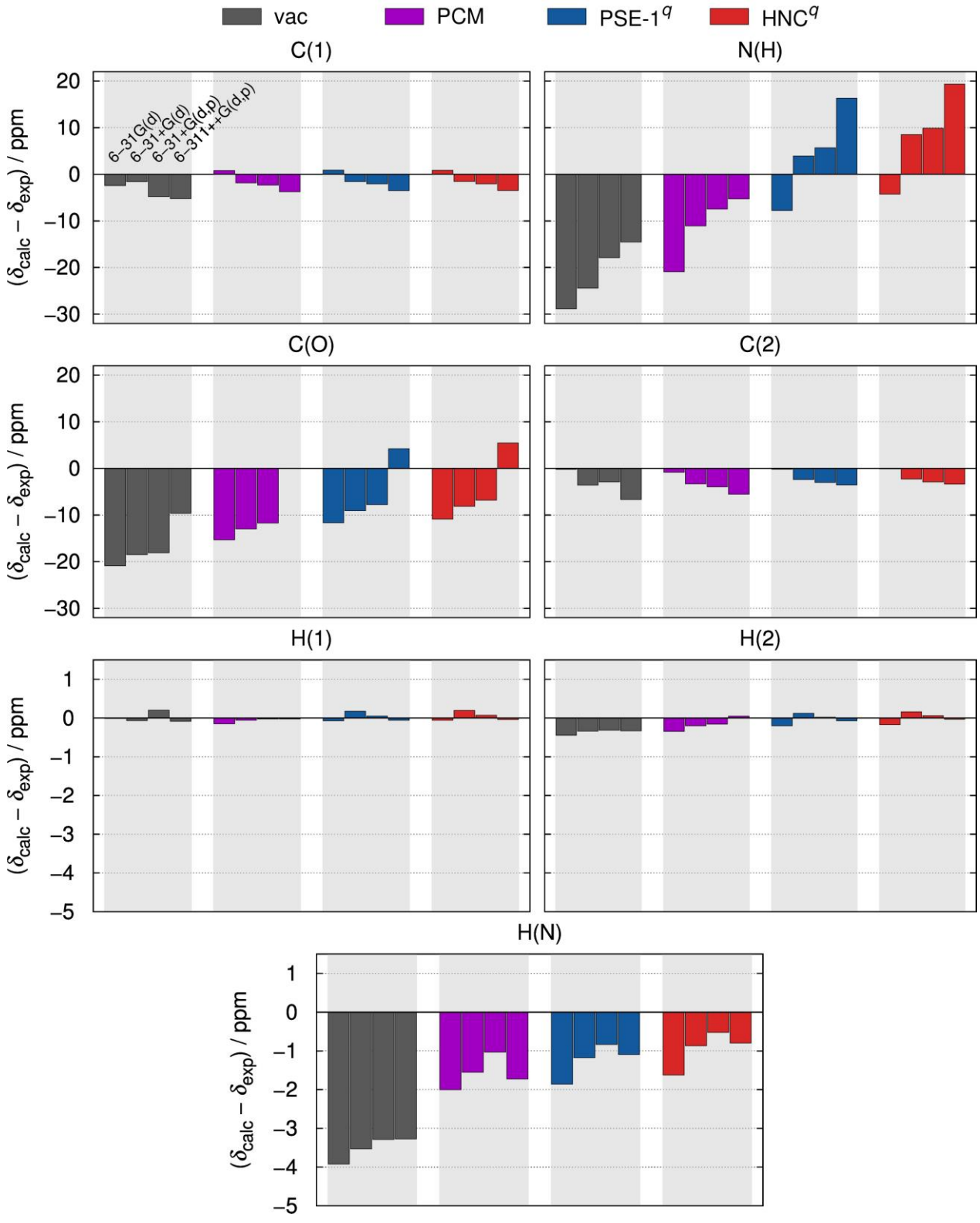


Figure 2.3 Nucleus specific results for the Pople basis set series with deviations between calculated shifts δ_{calc} and experimental values[115] δ_{exp} from vacuum and various solution-phase calculations, employing GIAO/B3LYP

Changing the basis set to the much more expensive aug-cc-pVTZ basis (see Figure 2.2) dramatically improves the amide nuclei chemical shifts for the PCM solvation model. The chemical shift errors for the carbonyl carbon within the EC-RISM solvation model also

decrease with the larger aug-cc-pVTZ basis, but this choice apparently worsens, similar to the 6-311++G(d,p) results, the chemical shifts for the amide nitrogen N(H) in comparison with smaller Pople basis sets and to a lesser extent also for non-amide carbons and hydrogens. On the other hand, in agreement with findings by others,[117] a triple-ζ basis is required to describe the amide carbonyl C(O) chemical shift properly. Similarly, also the amide proton H(N) chemical shift prediction is improved down to a deviation of only -0.16 ppm for EC-RISM(HNC) calculations when using exact electrostatics. Exact electrostatics (EC-RISM$^{\varphi}$) appears to be most essential for the amide nitrogen N(H) and to a lesser extent for the amide hydrogen predictions, for both in this manner examined basis sets (we will rationalize the origin of this effect below). Even the tendency of the EC-RISM chemical shifts with point charge representation to increase with higher PSE orders and basis set sizes is inverted for the nitrogen nucleus applying exact electrostatics with the aug-cc-pVTZ basis.

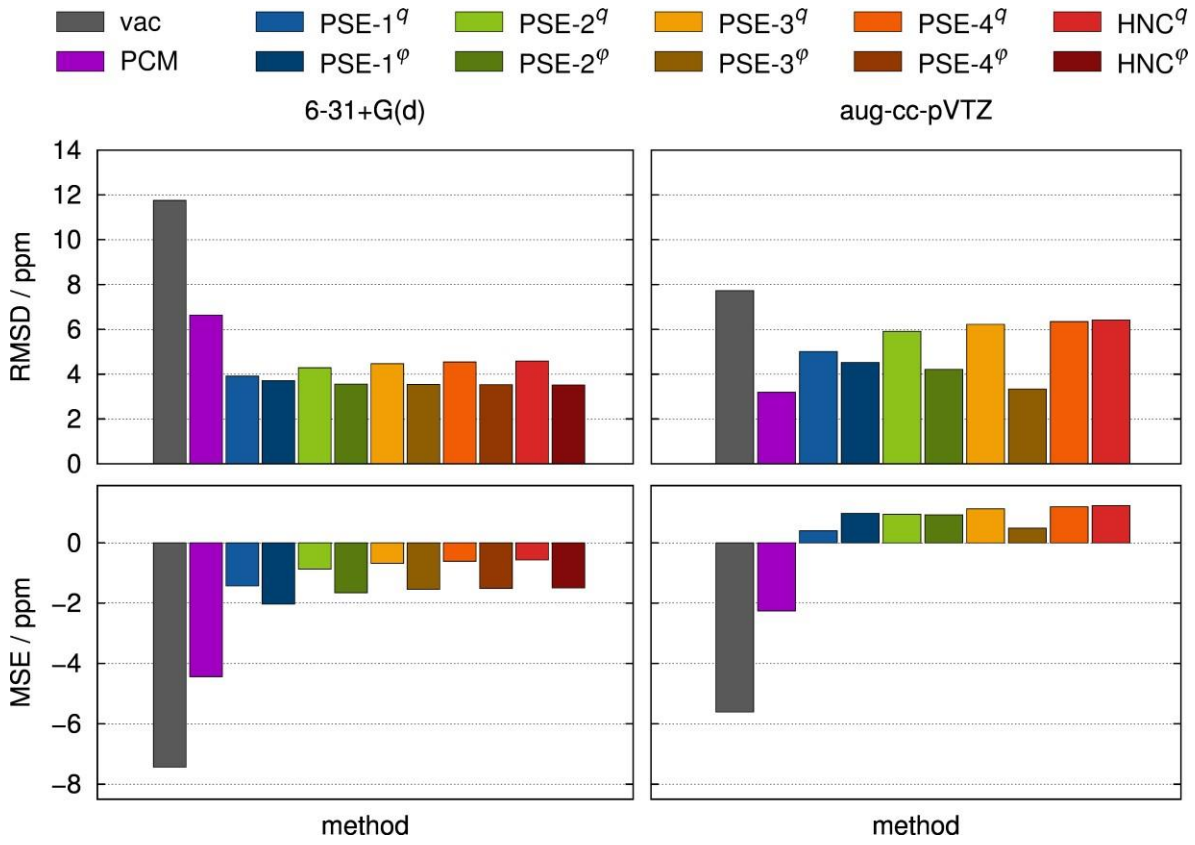


Figure 2.4 Root mean square deviation (RMSD) and mean signed error (MSE) of the calculated chemical shifts over all nuclei investigated with respect to experimental values.[115] Methodology and color codes correspond to Figure 2.2.

This trend is also a general pattern as reflected by two statistical descriptors, the root mean square deviation (RMSD) and the mean signed error (MSE). Figure 2.4 and Figure 2.5 reveals that the deviations to experimental chemical shifts, estimated with the EC-RISM solvation

model in combination with all closure relations, nearly all basis sets (except 6-311++G(d,p)) and both solute-solvent interaction schemes, are smaller than every PCM result with two exception, the RMSD of the B3LYP/aug-cc-pVTZ/PCM and B3LYP/6-311++G(d,p)/PCM. Its values of 3.19 and 3.29 ppm represent the overall smallest RMSDs, closely followed by the B3LYP/aug-cc-pVTZ/EC-RISM(PSE-3$^{\varphi}$) result with an RMSD that is only 0.15 ppm, respectively 0.05 ppm larger. Note that, due to convergence problems for the reference compound ammonia, the EC-RISM(PSE-4$^{\varphi}$) and EC-RISM(HNC$^{\varphi}$) chemical shift statistics are not available, but we can expect a continuation of the trend towards smaller RMSD for exact electrostatics.

Regarding the preformed basis set series evaluation, the MSE of all methods increase with the basis set size and the RMSDs significantly decrease for vacuum and PCM calculations. When using the Pople basis with point charge representation the RMSD for the B3LYP/EC-RISM/PSE-1 results has a minimum at the 6-31+G(d,p), whereas the RMSD results of the HNC closure is constantly worsening.

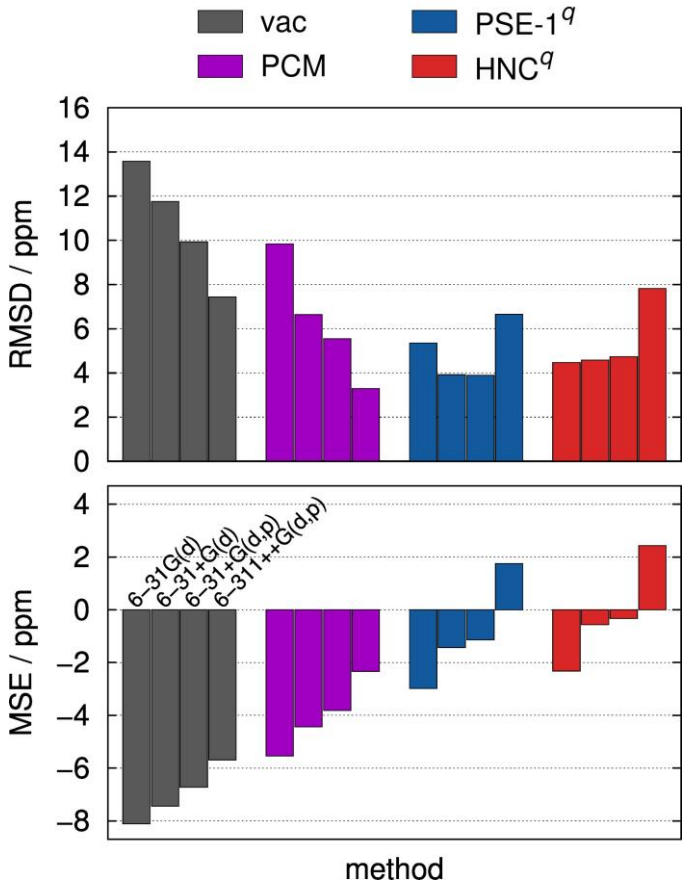


Figure 2.5 Root mean square deviation (RMSD) and mean signed error (MSE) of the calculated chemical shifts for the Pople basis set series over all nuclei investigated with respect to experimental values.[115] Methodology and color codes correspond to Figure 2.3.

However, RMSD data should be interpreted with caution. The RMSD is only statistically meaningful if the data set is large enough to obey symmetric deviations from the mean. For our very small sample size comprising only a single compound with a few nuclei, the RMSD has to be viewed in conjunction with the MSE since we are then in a better position to reveal systematic in contrast to random deviations. Taking both RMSD and MSE metrics together shows that the largest applied basis set with the exact electrostatics approach, B3LYP/aug-cc-pVTZ/EC-RISM(PSE-3^{q}), performs best for the integral equation model. The much larger MSE of the PCM model compared with EC-RISM (-2.25 vs. 0.49) indicates more pronounced systematic error whereas the EC-RISM approach appears to be more balanced. Similarly, data for the larger aug-cc-pVTZ basis set in comparison with the smaller Pople basis sets indicate less random scatter for the former. Enhancing the level of theory for both, the quantum-chemical and the solvation part of the model allows for a systematic improvement, except for the EC-RISM/6-311++G(d,p) results. In this case the EC-RISM^{q} approach results in too high chemical shifts for the C(O) and N(H) nuclei. In consequence the RMSD and MSE are larger. It is questionable whether the PCM results benefit from fortunate error cancellation for the larger basis sets (6-311++G(d,p), aug-cc-pVTZ) that is by far not as fortunate for EC-RISM with point charge electrostatics or if these trend is systematic. Furthermore the aug-cc-pVTZ results indicate that a large basis set should be combined with exact electrostatics.

In summary, within the sequence of EC-RISM closure approximations, we can reaffirm earlier observations that PSE-3 represents the best compromise between numerical stability and accuracy; PSE-1 (KH) is clearly inferior. While this statement holds most notably for exact electrostatics, it also holds for the point charge approximation at least for the small basis set, with the exception of the amide nitrogen that always benefits from an exact treatment. In order to adequately describe the shielding of this critical nucleus, a large basis sets appears to be necessary, however at the prize to also compute exact electrostatic interactions since the point charge model is clearly insufficient in this case.

2.3.2 Physical rationalization

The impact of the treatment of electrostatics particularly for N(H) has to be translated into solvation patterns since the structure of a polarizing environment determines the chemical shift. Its deviation from experimental values vastly increases for the point charge EC-RISM approximation with higher order closure relation, resulting in a large overall RMSD. As already pointed out by Bader,[116] the amide nitrogen chemical shift of NMA is very sensitive to surrounding charges, in particular to those located in the amide bond plane.

In order to rationalize the effect of different treatments of solute-solvent electrostatics, we computed the difference between the solvent charge densities of the point charge approach and the exact electrostatics approach,

$$\Delta\rho_q(\mathbf{r}) = \rho_q(\mathbf{r}, \mathrm{HNC}^{\varphi}) - \rho_q(\mathbf{r}, \mathrm{HNC}^{q}). \tag{43}$$

Figure 2.6 illustrates this difference in the amide bond plane for the converged EC-RISM/B3LYP/6-31+G(d) results. Largest discrepancies are found along the N-H bond axis and around the carbonyl oxygen. Bader[116] showed that a localized extra charge moved along the N-H bond axis barely influences the ^{15}N chemical shift, whereas the impact of charges moved along the N-C(O) and N-C(2) bond is large. Corresponding charge density differences to the point charge approach are indeed found around the carbonyl oxygen, indicating the relevance of lone electron pairs that are only captured by the exact electrostatics approach.

Higher PSE orders induce more pronounced solvent structuring.[58] Regarding deviation from experimental values, the chemical shifts tend to move to lower fields with increasing PSE-*n* order, indicating that NMA is polarized more strongly by enhancing solvent structure. Figure 2.7 shows that the chemical shifts of core NMA amide nuclei depend almost linearly on its dipole moment, which increases in parallel with the PSE order. Among the amide atoms, the chemical shift of the C(O) shows less dependence on the dipole moment with a slope of 2.57 ppm/D to 3.62 ppm/D, but particularly chemical shifts of H(N) (0.81-1.03 ppm/D) as well as of N(H) (7.52-14.06 ppm/D) change dramatically (relative to their experimental value) in response to enhanced solvent structure. The dipole moments of the NMA molecules are en bloc higher when using a point charge representation in comparison with the exact electrostatic environment. Further the EC-RISM dipole moments for the 6-31+G(d) and aug-cc-pVTZ basis sets are at least 1 or more Debye higher than the vacuum and PCM dipole moments (see Table S6-S8).

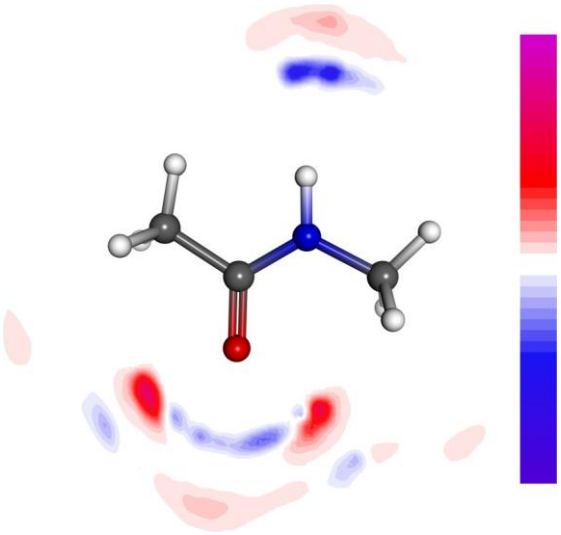


Figure 2.6 Solvent charge density difference $\Delta\rho_q(\mathbf{r})$ between exact quantum-mechanical electrostatics and the point charge solute-solvent interaction approximation within the EC-RISM/HNC solvent model for the B3LYP/6-31+G(d) level of theory. The color code covers ±0.429 e/Å^3; for NMA the extreme values are -0.429 and 0.345 e/Å^3, respectively.

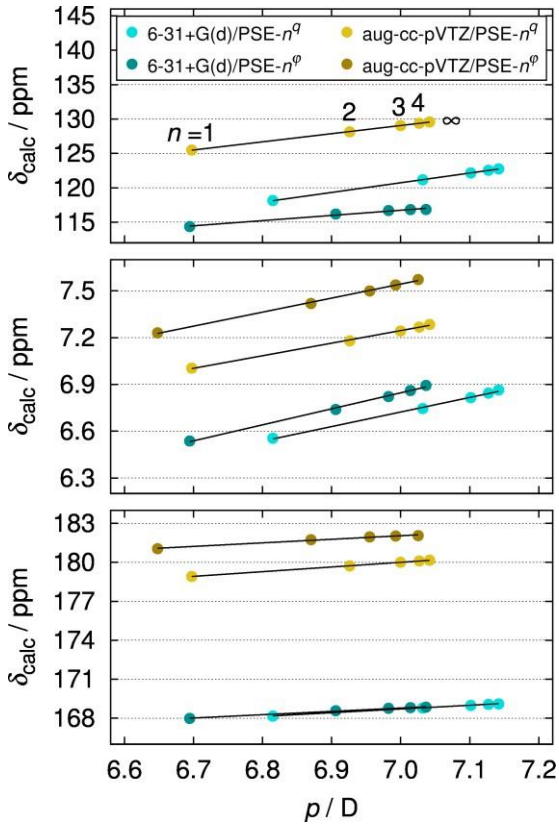


Figure 2.7 Correlation between NMA dipole moment and calculated chemical shifts for increasing PSE order (analogous trend for all lines) along with linear regression curves (data is found in supporting Tables S6-S8), from top to bottom: N(H), H(N), C(O).

Due to the close correspondence of the chemical shift and the dipole moment it is questionable if the overshooting of EC-RISM in combination with a high PSE-*n* closure order and a highly polarizable basis set is in fact the consequence of an overpolarization of the molecule of interest, here NMA or of the reference molecules. Furthermore error cancellation for the quantum-chemical level of theory should not be that effective, when the reference molecule and NMA are polarized to a different degree.

In comparison with the dipole moments of 5.79 D achieved by Du and Wei[118] with their HF/6-31G(d) QM/3d-RISM-HNC methodology our results (6.70 – 7.14 D) are by far higher. But they are in very good agreement with the Car-Parinello molecular dynamics results (6.96 D)[119], although our PSE-4 and HNC results are slightly higher. With values of 5.83 D for 6-31+G(d) and 4.07 D for aug-cc-pVTZ the PCM dipole moments are consistently not as large as the EC-RISM results. The description of the solute wave function by the aug-cc-pVTZ basis set and exact electrostatics representation has also always a lower dipole moment that counteracts overpolarization effects.

2.4 Concluding remarks

In summary, this proof-of-principle study has succeeded to demonstrate the basic applicability of the EC-RISM methodology to NMR chemical shift determination and its predictive advantages compared with the PCM model. We have shown that EC-RISM with sufficient quality of the closure approximation is capable of capturing the essential physical effects of a polarizing environment on electronic structure. This is particularly true when the full quantum-mechanical electrostatic potential is applied for the solute-solvent interaction description. Further the presented results indicate that the polarization of the wave function by the EC-RISM solvent surroundings is reasonable and comparative to the PCM wave function, although the electron density is polarized more strongly. Compared to the PCM model EC-RISM does not gain such an effective error cancellation for larger basis sets using only a point charge representation of the electrostatic solute-solvent interactions. As a general suggestion, the combination of the small 6-31+G(d) basis with the PSE-3 closure and exact electrostatics appears to be a reasonable compromise between computational speed and predictive power, especially for small molecules. For the treatment of electron-rich compounds like anions the 6-31+G(d,p) basis set with a small PSE order is recommended, but has to be tested in combination with higher PSE orders in future work. It performs nearly as good as the 6-

31+G(d)/PSE-3 combination and the slightly larger basis set is more appropriate to capture polarization of small anions.

This motivates us to pursue the necessary steps for further developments, such as the examination of the quantum-chemical level of theory, impact of dispersion/repulsion models, consistent EC-RISM-optimal geometries, conformational ensembles, prediction of coupling constants, and so forth, which remain untreated in this work. From a practical point of view, a regression analysis for a large data set of compounds could be advantageous since this would eliminate the need for a sophisticated choice and adequate treatment of reference compounds. An interesting and straight-forward extension would be the combination with ADMA-type fragment-based quantum-chemical calculations[94,95] in order to scale up the methodology to large systems such as proteins. But even in the present form, EC-RISM has the potential to be a useful complement of the computational chemistry arsenal to tackle solution-phase NMR problems. This is especially the case for issues, where common continuum solvation methodologies are not applicable like investigations of chemical shifts under high hydrostatic pressure.

3 SOLVATION STRUCTURE AND THERMODYNAMICS OF NONDIPOLAR LIQUIDS[ii]

3.1 Introduction

The influence of pure solvents on solutes is predominantly explained by dipolar interactions between solvent and solute or through collective terms like "hydrophobic effect" or "hydrophobic interaction" that are commonly used for aqueous solutions when dipolar solute-solvent interactions are missing.[120,121] Additionally, packing effects are particularly important for weak up to non-interacting solvents.[122] Nevertheless, the question of how to distinguish between solvents with similar molecular packing, but without dipolar interactions arises. A good example for such a delicate solvent pair is benzene (C_6H_6)/hexa-fluorobenzene (C_6F_6). Both molecules are highly symmetric (D_{6h}) and lack dipole moments. Both solvents are not easily polarizable, as indicated by their low relative dielectric permittivities ($\varepsilon_r = 2.27$ for C_6H_6 and $\varepsilon_r = 2.02$ for C_6F_6 [123]). In condensed phase and in the crystal structure both exhibit a characteristic so-called "herringbone structure"[124,125] that is preferred due to beneficial quadrupolar electrostatic interactions and contains T-shape like aromatic ring dimers (see Figure 3.1).[126,127,128] Furthermore the molecular quadrupole moments of C_6H_6 ($\Theta = -33.3 \pm 2.1$ Cm^2)[129] and C_6F_6 ($\Theta = 31.7 \pm 1.7$ Cm^2)[129] have nearly the same magnitude, but opposite signs, augmenting a π-π stacking between C_6H_6 and C_6F_6 molecules.[124]

[ii] The C_6H_6 and C_6F_6 solvent models that are applied and extended in this chapter, and the corresponding 1D and 3D RISM calculations are partially based on: R. Frach, *Quantenchemische Modelle für nichtwässrige Dienaminlösungen*, master thesis, TU Dortmund, **2011**.'

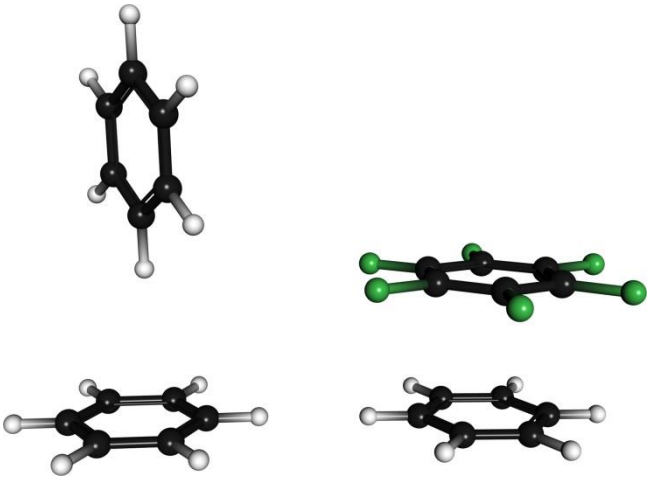


Figure 3.1 Illustration of a T-shaped benzene dimer and a π-stacked dimer of benzene and hexafluorobenzene.

But there are still fundamental differences between both solvents regarding thermodynamic and kinetic properties, like solvation free energies ΔG_{solv}[130] or rate constants of stereoselective reactions.[131,30] For detailed rationalizations of these phenomena, methods are required that reach beyond dielectric effects and incorporate subtle structural solvent reorganization differences. Classical force field molecular dynamics (ffMD)[4] simulations treat solvation in a detailed manner, but lack a quantum-chemical description of the solute, which is needed, for example, for the computational investigation of chemical reactions. This deficiency is tackled by *ab initio* MD[6] simulations, however at much higher computational expense. Unlike *ab initio* MD calculation, dielectric continuum solvation methods like the polarizable continuum model (PCM)[13] or the conductor-like screening model (COSMO)[132] are comparatively computationally cheap. Indeed, such methods usually fail to resolve subtle differences in the solvation structure. A couple of quadrupolar extensions for PCM are accessible as described by Jeon et al.,[133,134] However these extensions suffer from a high amount of necessary experimental parameters and a sophisticated extensibility to higher multipoles.

One way to fill the gap between a computationally feasible solvent model, which retains the solvent structure and a quantum-mechanical solute description, is to use integral equation theory with a coupled quantum-chemical solute description. A computationally balanced and accurate way of implementing integral equation theory is the reference interaction site model (RISM)[17,18] approach. Starting from the one-dimensional level of theory, the so called 1D RISM method, a solvent model, characterized by its solvent distribution functions and its site-site susceptibility, can be generated. Afterwards, by using the precomputed solvent susceptibility, it is possible to calculate spatial solute-solvent distribution function for

arbitrary shaped solutes by solving the three dimensional form of the reference interaction site model (3D RISM).[56,57] To describe the interaction between the RISM solvent model and the solute, the solvent-charge distribution, which results from the 3D RISM solvent distribution, is mapped onto a discrete point charge grid, representing the solvent influence on the electronic Hamiltonian of the solute. The solute is polarized and in turn polarizes the solvent, consequently modifying the solvent distribution in a subsequent 3D RISM calculation. In this way, by reiterating the previous steps, a self-consistent procedure is established, a procedure that is known as "embedded cluster reference interaction site model" (EC-RISM).[23] Up to now this approach was mostly and successfully validated for polar solvents like water[23] or DMSO.[27]

Here, a diverse investigation of C_6H_6 and C_6F_6 solutions by the EC-RISM methodology is presented. Thus it is determined how far the thermodynamics depend on the quantum-mechanical level of theory, the integral equation approach and the solvent model. Therefore 1D RISM solvent susceptibility functions, that we will simply call 'solvent models', are generated and evaluated by comparing integral equation results to MD simulations. Afterwards these solvent models are used to evaluate Gibbs free energy differences of hexafluorobenzene and benzene solution of small molecules in the Ben-Naim standard state[77], the so-called transfer Gibbs free energy

$$\Delta G^{*}_{\mathrm{trans}} = G^{*}_{\mathrm{sol}}(\mathrm{C_6F_6}) - G^{*}_{\mathrm{sol}}(\mathrm{C_6H_6})\,. \tag{44}$$

The transfer Gibbs free energy is connected to the ratio of Henry constants k_{H} of a small gas molecule by

$$\frac{k_{\mathrm{H}}(\mathrm{C_6H_6})}{k_{\mathrm{H}}(\mathrm{C_6F_6})} = \frac{\rho(\mathrm{C_6H_6})}{\rho(\mathrm{C_6F_6})}\exp[-\beta(\Delta G^{*}_{\mathrm{trans}})] \tag{45}$$

and therefore a significant observable to estimate and understand the solvability of gases.[135] Note that, for computation of absolute Henry law constants

$$p^{0}k_{\mathrm{H}}^{-1} = \exp[-\beta\Delta G^{0}_{\mathrm{solv}}], \tag{46}$$

one would apply the usual standard state concentration and the corresponding Gibbs free energy of solvation with

$$\Delta G^{0}_{\mathrm{solv}} = \Delta G^{*}_{\mathrm{solv}} + \Delta G^{*\to 0} = G^{*}_{\mathrm{sol}} - G^{*}_{\mathrm{vac}} + \Delta G^{*\to 0}, \tag{47}$$

for which $\Delta G^{*\rightarrow 0}$ corresponds to the free energy change from the 1 M infinitely diluted solution to the usual standard state[136] with p^0 as standard pressure and G^*_{vac} is the vacuum Gibbs free energy of the solute. In the following this chapter concentrates on transfer Gibbs free energies and Henry constants will not be estimated, therefore the standard state designators will be dropped.

In contrast to earlier integral equation theoretical attempts to describe C_6H_6 and C_6F_6 by Lowden et al.[137] and by Steinhauser et al.[138], applying 3D RISM, this work steps beyond radial distribution functions and show that significant structural details are lost by radial averaging, which have substantial influence on the thermodynamics. Additionally structural differences for the three dimensional solvent distribution by either changing the representation of solute-solvent electrostatics or even more by interchanging the quadrupole moments of the solvent molecules are observed. The connection between RISM solvation patterns and electrostatic quadrupole moments is rationalized by examination of the “like” pairs (C_6H_6 in C_6H_6 and C_6F_6 in C_6F_6) and “unlike” pairs (C_6H_6 in C_6F_6 and C_6F_6 in C_6H_6) with 1D, 3D and EC-RISM. By further deactivation of all electrostatic interactions for both solvents and employment of pure dispersive models, the solvent influence is separated into parts, investigating the importance of van der Waals and quadrupole interactions.

3.2 Computational details

3.2.1 Solvent susceptibilities and 1D RISM calculations

The solvent susceptibility function χ is the key component for RISM calculations, as it defines the most important part of the solvent model. To gain access to this quantity for benzene and hexafluorobenzene, pure solvent 1D RISM calculations were performed. This results in solvent susceptibilities, which were utilized as starting steps for further higher level calculations. Therefore classical force fields were employed for the solvent site-site interactions in the form of a Coulomb potential for electrostatic interactions and a Lennard-Jones 12-6-potential with Lorentz-Berthelot mixing rules for all dispersion interactions. Force field parameters for C_6H_6 and C_6F_6, namely the solvent site charge q, the Lennard-Jones potential depth ε and the corresponding contact distance σ were taken from a hybrid Amber/OPLS force field.[139,140] For comparison a second force field model were chosen for C_6H_6 developed by Cornell et al.[141,39] To model C_6F_6 the corresponding force field parameters of the mixed Amber/OPLS approach[140,39] were applied. Geometric mixing rules that are common for the OPLS force field were not used to stay consistent with the otherwise applied parametrization. Ideal highly symmetric (D_{6h}) C_6H_6 and C_6F_6 structures were employed. For an overview of all solvent parameters and bond distances see Table 3.1.

Table 3.1 Bond length r_{ij}, macroscopic solvent densities and classical force field parameters for benzene[139,141] and hexafluorobenzene.[140]

	C_6H_6 Amber/OPLS[139]	C_6H_6 Cornell[141]	C_6F_6 Amber/OPLS[140]
q_C / e	-0.115	-0.103	0.130
q_H / e	0.115	0.103	-
q_F / e	-	-	-0.130
σ_C / Å	3.55	3.40	3.55
σ_H / Å	2.42	2.60	-
σ_F / Å	-	-	2.85
ε_C / kcal mol^{-1}	0.070	0.086	0.070
ε_H / kcal mol^{-1}	0.030	0.015	-
ε_F / kcal mol^{-1}	-	-	0.061
r_{CC} / Å	1.400	1.410	1.400
r_{CH} / Å	1.080	1.090	-
r_{CF} / Å	-	-	1.347
ρ / nm^{-3}	6.785	6.785	5.218

The density was set to 873.4 kg/m^3 for benzene, to 1606.3 kg/m^3 for hexafluorobenzene (corresponding to a particle density of 6.785 nm^{-3} and 5.218 nm^{-3})[142] and a value of 298.15 K was applied for the temperature. The 1D RISM equations were solved with HNC or PSE-*k* closures, *k* ranging from 1 up to 4 on a logarithmically spaced grid with 512 points analogously to earlier work[143]. The grid ranged from 0.0059 Å to a maximum distance of 164.02 Å. Solutions were converged to a maximum residuum norm of $1 \cdot 10^{-6}$ for the direct correlation function. The dielectrically consistent RISM approach (DRISM)[52,53,] that corrects the solvent structure of dipolar solutions by the use of the molecular dipoles was not employed due to the missing dipole moments of both solvent molecules.

Furthermore two additional sets of solvent models were generated, one set for which the partial charges of C_6H_6 and C_6F_6 are assessed to zero and one set for which the partial charges between C_6H_6 and C_6F_6 were interchanged, thus interchanging the quadrupole moments of both molecules. The Lennard-Jones parameters and the molecular structures of the Amber/OPLS models have been used for both sets. Subsequently the uncharged solvent models will be referred to as 'q_0-C_6H_6' and 'q_0-C_6F_6' and the models with reversed partial charges as 'q_{rev}-C_6H_6' and 'q_{rev}-C_6F_6'. MD simulations were performed to validate the 1D RISM pair distribution functions of C_6H_6 and C_6F_6 and to get a grasp of the influence of the bridge function on the solvation structure. The force field models for all MD simulations were chosen corresponding to the 1D RISM calculations. Therefore cubic simulation cells were generated with a scale of 5.3^3 nm^3 containing 1010 benzene molecules and 5.8^3 nm^3 containing 1018 hexafluorobenzene molecules. The packmol software (version 15.133)[144] was used to construct simulation starting points. Gromacs (version 4.6.3)[145,146] was used for MD equilibration and production runs. Cutoff radii of 1.2 nm for real space interactions were selected for all calculations. Long ranged electrostatics were treated with the smooth particle mesh Ewald technique.[147,148] Furthermore a time step of 1 fs was chosen. After equilibration within either the NVT ensemble (T = 298.15 K) on the one hand and the NpT ensemble (p = 1 bar, T = 298.15 K) on the other hand production runs were performed in both ensembles with a simulation time scale of 10 ns. The isotherm-isobar simulations were conducted to review the force field specific solvent density. Apart from this simulations in the canonical ensemble were performed in order to achieve pair distribution functions with the experimental density. Temperature coupling was realized through the Nosé-Hoover thermostat[149,150] and the desired pressure was adjusted with the Parrinello-Rahman coupling scheme[151,152], both with relaxation times of 1 ps and 5 ps respectively.

3.2.2 Spatial solvent distributions and transfer Gibbs free energies

The determination of excess free energies and three dimensional solvent structures was performed by 3D RISM calculations on a cubic grid with 120^3 grid points and a grid spacing of 0.3 Å. The 3D RISM solutions were converged to a maximum residuum norm of $5 \cdot 10^{-6}$ for the direct correlation function. The 3D RISM closure was always selected accordingly to the closure of the solvent susceptibility function. EC-RISM calculations were performed with a convergence criterion of 10^{-3} kcal/mol for the total Gibbs free energy of the solute. The transfer Gibbs free energy of a set of nine small molecules, namely He, Ne, Ar, N_2, O_2, CO, CO_2, CH_4 and CF_4 for which transfer Gibbs free energies have been measured by Wilhelm and Battino,[130] were estimated in accordance to equation (44). A range of different quantum-chemical approaches, namely the Hartree-Fock (HF) approach[153], Møller-Plesset perturbation theory with second order correlation correction (MP2)[154,155], PBE0[156] density functional theory and B3LYP hybrid exchange functional[100,101,102] were tested to determine their influence on ΔG_{trans}. The solute structures were estimated by the integral equation formalism PCM (IEF-PCM, $\varepsilon_r(C_6H_6) = 2.2706$, $\varepsilon_r(C_6F_6) = 2.029$) geometry optimization with three different Pople basis sets (6-31G(d), 6-311G(d,p), 6-311+G(d,p))[103] as well as three different Dunning basis sets (aug-cc-pVDZ, aug-cc-pVTZ, aug-cc-pVQZ).[106,107] Afterwards, using these precomputed structures with corresponding levels of theory, transfer Gibbs free energies were calculated with EC-RISM, additionally employing classical force field parameters from literature ((He, Ne, Ar, N_2, O_2)[157], CO[158], CO_2[159], CH_4[160], CF_4[161]) for dispersion interactions (see Table 3.2). A polarization of the solute was not expected for the uncharged solvent models. For this reason the examined levels of theory were limited to B3LYP with 6-311G(d,p) and aug-cc-pVTZ basis sets for the q_0-models. The solute-solvent interaction was either described by point charges, indicated by EC-RISM^q or by the exact electrostatic potential (ESP), abbreviated by EC-RISM$^\varphi$. The quantum-chemical part of EC-RISM was performed by the Gaussian 03 program package (Rev. D.02)[104], as well as the implemented fit to the ESP using the ChelpG partial charge determination procedure[162]. Beside the last EC-RISM step, all precedent steps were evaluated with HF level of theory.

The results were compared to IEF-PCM[13], to experimental values of Evans et al.[130] and furthermore to 1D and 3D RISM results utilizing unpolarized vacuum point charges. 3D RISM solvent distribution function for the “like” and “unlike” pairs of C_6H_6 and C_6F_6 were obtained with the same procedural parameters as for the small molecules. The solute structures, point charges and Lennard-Jones parameter were chosen in accordance to the 1D

RISM solvent models for pure 3D RISM calculations. To obtain a qualitative picture of the quantum-chemical influence on the like and unlike solute-solvent pairs just a small 6-31G(d) basis set was chosen. The exact ESP was used for the solute-solvent part here.

Table 3.2 Employed Lennard-Jones potential parameters for EC-RISM calculations. Parameters for the noble gases and oxygen as well as nitrogen were taken from Makrodimitri et al.[157] Parameters for carbon monoxide, carbon dioxide, methane and perfluoromethane were applied according to further literature force fields.[158,159,160,161] To distinguish between interaction sites of molecules with two different types of atoms the intended interaction sites are highlighted by underlining.

	σ / Å	ε / kcal/mol
He	2.570	0.020
Ne	2.760	0.067
Ar	3.410	0.236
N_2	3.310	0.074
O_2	3.090	0.089
$\underline{C}O$	3.636	0.032
$C\underline{O}$	2.979	0.196
$\underline{C}O_2$	2.757	0.056
$C\underline{O}_2$	3.033	0.160
$\underline{C}H_4$	2.001	0.095
$C\underline{H}_4$	0.835	0.017
$\underline{C}F_4$	3.500	0.097
$C\underline{F}_4$	2.950	0.053

3.3 Results and discussion

3.3.1 Pure solvent radial pair distribution functions

All solvent susceptibility functions of C_6H_6 and C_6F_6 were constructed by pure solvent 1D RISM calculations in combination with each earlier mentioned closure.[39] The sole exception is the combination of the Amber/OPLS force field with reversed partial charges for C_6F_6. Here a converged total correlation function could not be achieved using the HNC closure. The solvent site-site correlation functions for the whole set of possible force field-closure combinations are shown in Figures S1-S9 of the appendix. In contrast to earlier results for water[58,163] the pure solvent pair correlation functions of C_6H_6 and C_6F_6 are almost closure-insensitive.

In comparison with MD results the 1D RISM pair distribution functions predominantly exhibit shorter site-site contact distances as illustrated in Figure 3.2. This is consistent with other 1D RISM calculations and is the motivation for the so-called 'repulsive bridge corrections'[164], which broaden exactly these site-site distance at least for water. Beyond the contact distances the structural elements of C_6H_6 are well reproduced by 1D RISM theory, but sometimes shifted to shorter distances, even in the case of larger reaching structural elements like the second and third solvation shell. Examples are the two maxima of the carbon-carbon pair distribution function that are reproduced, but slightly shifted to shorter distances. A comparison of the 1D RISM/AMBER/OPLS results of benzene with the Cornell force field model distribution functions reveals that both force fields yield nearly identical results with slightly different heights of the first two maxima of the hydrogen-hydrogen g_{HH} and carbon-carbon pair distribution function g_{CC}. This is in principle also true for the simulated results.

The hexafluorobenzene structure has, focusing on the simulated g_{CC} and carbon-fluorine distribution functions g_{CF}, similar structural features, compared to the corresponding benzene distribution functions. Nevertheless these radial distribution functions are not reproduced that well by the 1D RISM as for the benzene case. It is reasonable to infer that the major difference between the radial solvation structures is exhibited by the fluorine-fluorine g_{FF} and the hydrogen-hydrogen pair distributions g_{HH}. Particularly for the g_{FF} function, where a maximum around 3 Å is observable, there is none for the g_{HH} function.

An estimation of the simulated solvent density by isotherm-isobar simulation shows that each force field model underestimates the solvent density by 6.9-12.4 % as tabulated in Table 3.3. Here the Amber/OPLS force field performs slightly better for the densities than the Cornell force field. It can be assumed that the pair distribution functions related to the NpT simulation results could be slightly higher with a force field model that reflects the densities more appropriate, as it is the case for the NVT simulation results.

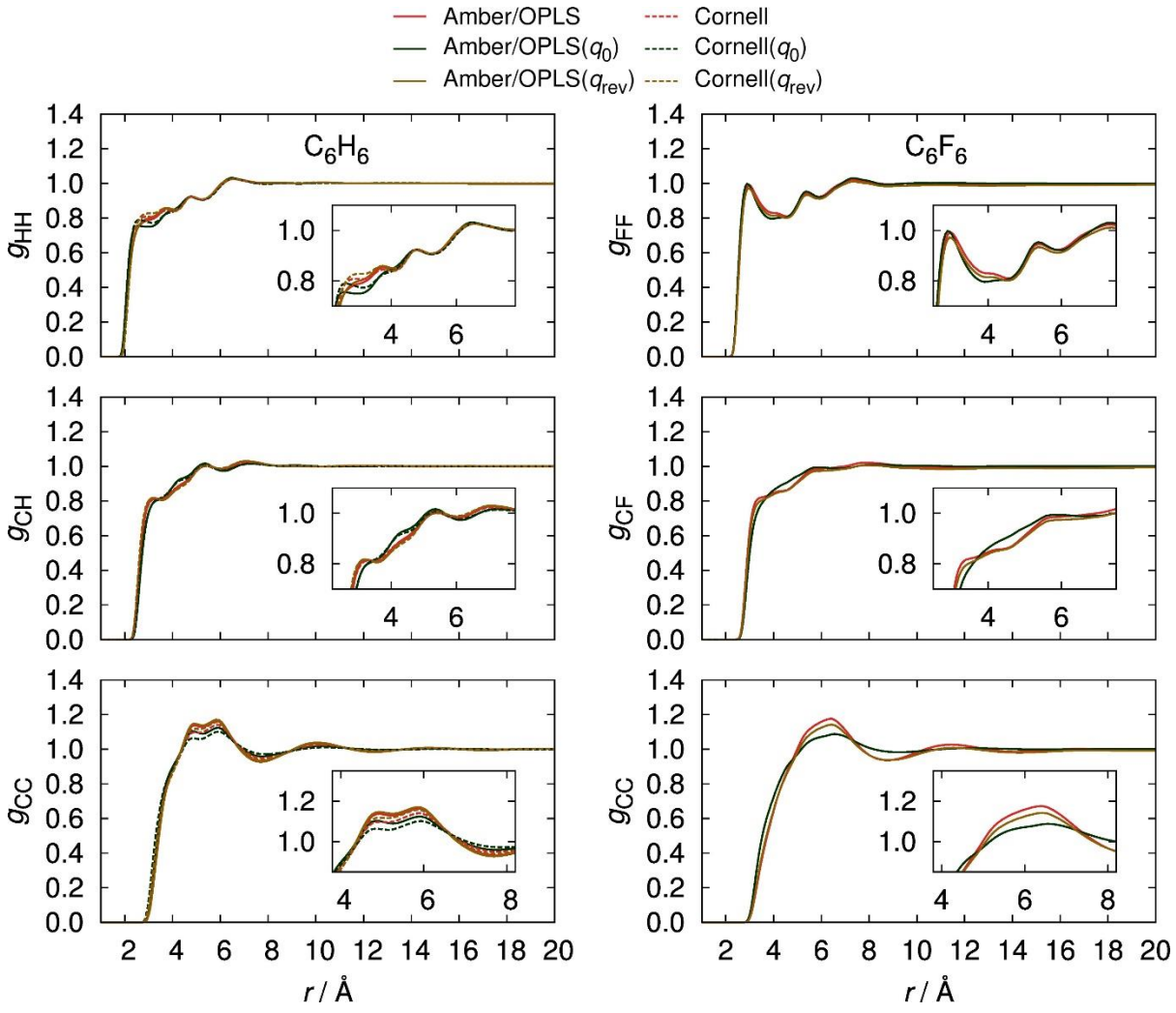


Figure 3.2 Radial site-site distribution function of C_6H_6 (left) and C_6F_6 (right) calculated with 1D RISM/HNC closure in comparison with MD results. The results have been obtained with either the Amber/OPLS or the Cornell force fields. The underlying 1D RISM data is based on the master thesis of R. Frach.[39]

Table 3.3 Comparison of the experimental densities[142] of benzene and hexafluorobenzene with the isothermal-isobaric simulated densities. Additionally the relative difference $\Delta\rho_{rel}$ to the experimental values is tabulated.

	C_6H_6		C_6F_6	
	ρ / kg m^{-3}	$\Delta\rho_{rel}$ / %	ρ / kg m^{-3}	$\Delta\rho_{rel}$ / %
exp	873.4	-	1606.3	-
Amber/OPLS	813.3	6.9	1489.5	7.3
Cornell	765.0	12.4	-	-

The underlying cause behind the difference of the g_{HH} and g_{FF} functions seems not to be the different quadrupolar moments, but the difference in the dispersion interactions. As shown in Figure 3.3, where the results without electrostatic moments (q_0) and reversed quadrupole moments (q_{rev}) are illustrated, the drastic change in the electrostatics of both solvents leads to nearly the same 1D RISM results for radial solvation. It has to be noted that especially the exchange of partial charges has only a minor impact on the pair distribution functions, while the results with the uncharged model are notable slightly flatter.

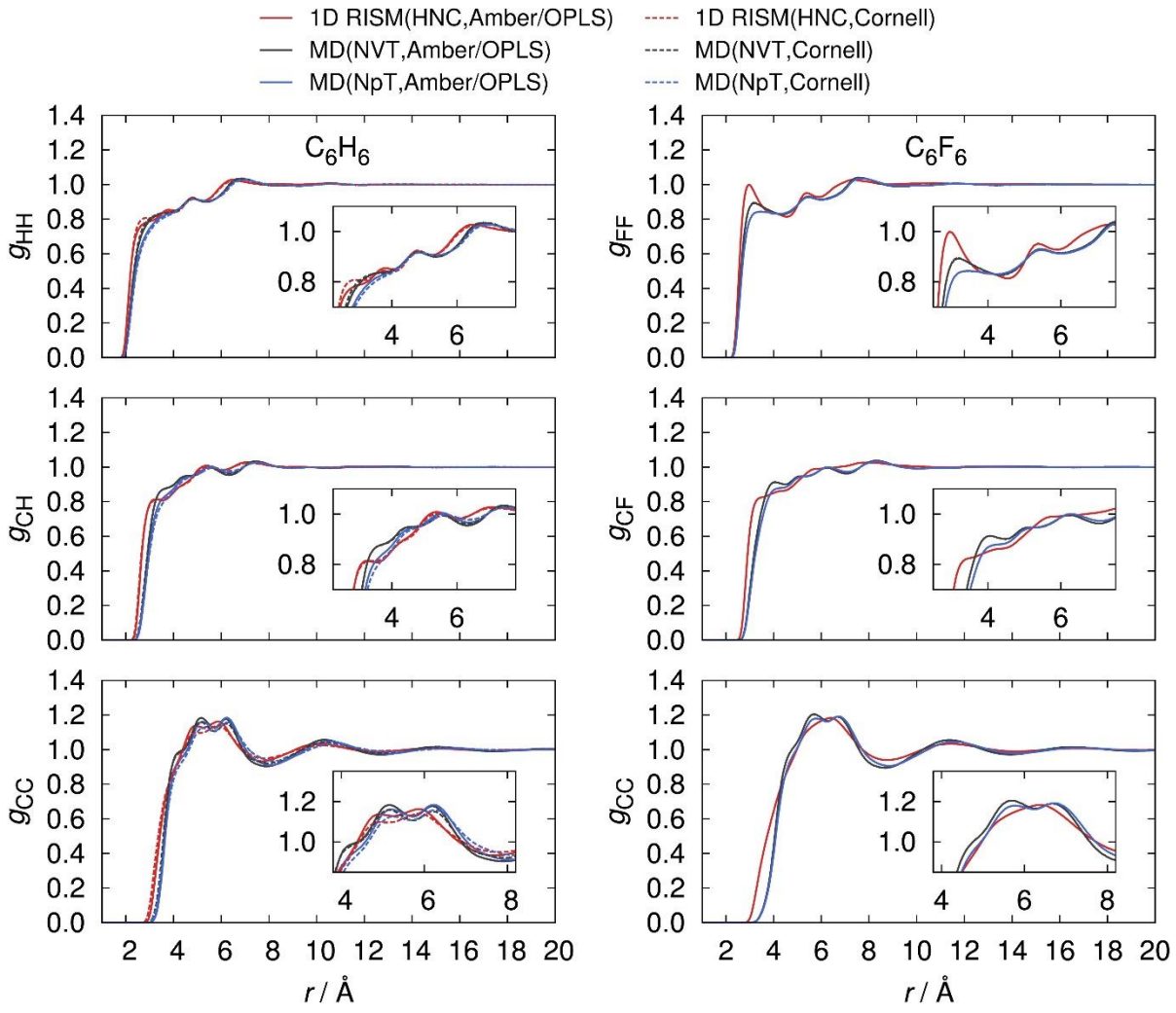


Figure 3.3 1D RISM radial site-site distribution function of benzene (left) and hexafluorobenzene (right) in comparison with the reversed charge model (q_{rev}-C_6F_6, q_{rev}-C_6H_6) and the uncharged model results (q_0-C_6F_6 and q_0-C_6H_6). The 1D RISM results have been obtained with the HNC closure for the q_0-C_6H_6 and q_0-C_6F_6 solvent models and, due to convergence difficulties, with the PSE-4 closure for the q_{rev}-C_6F_6 and q_{rev}-C_6H_6 models. The underlying 1D RISM data is based on the master thesis of R. Frach.[39]

These distributions functions can be compared to the radial pair distributions of 3D RISM calculations of the solute-solvent pairs C_6H_6/C_6H_6 and C_6F_6/C_6F_6 after averaging over the angular coordinates using Lebedev-Laikov grids.[165] The resulting radial distribution functions are illustrated in Figure 3.4. These radial distribution functions are in much better agreement with the simulated results for short distances than the initial 1D RISM pair distribution functions. Furthermore the shifts of the maxima that are observed for the 1D RISM results are corrected by 3D RISM and consequently almost always at the right distance. But the heights of the radially averaged 3D RISM distribution functions are much lower, probably due to the use of the PSE-1 closure as 3D RISM closure. In conclusion, the radial solvation structure is appropriately represented by 3D RISM, but the peak height is underestimated. Nonetheless, particularly the better behavior for small distances that is not expressed in the radially averaged 1D RISM theory has a great impact on the solvation free energies.[164]

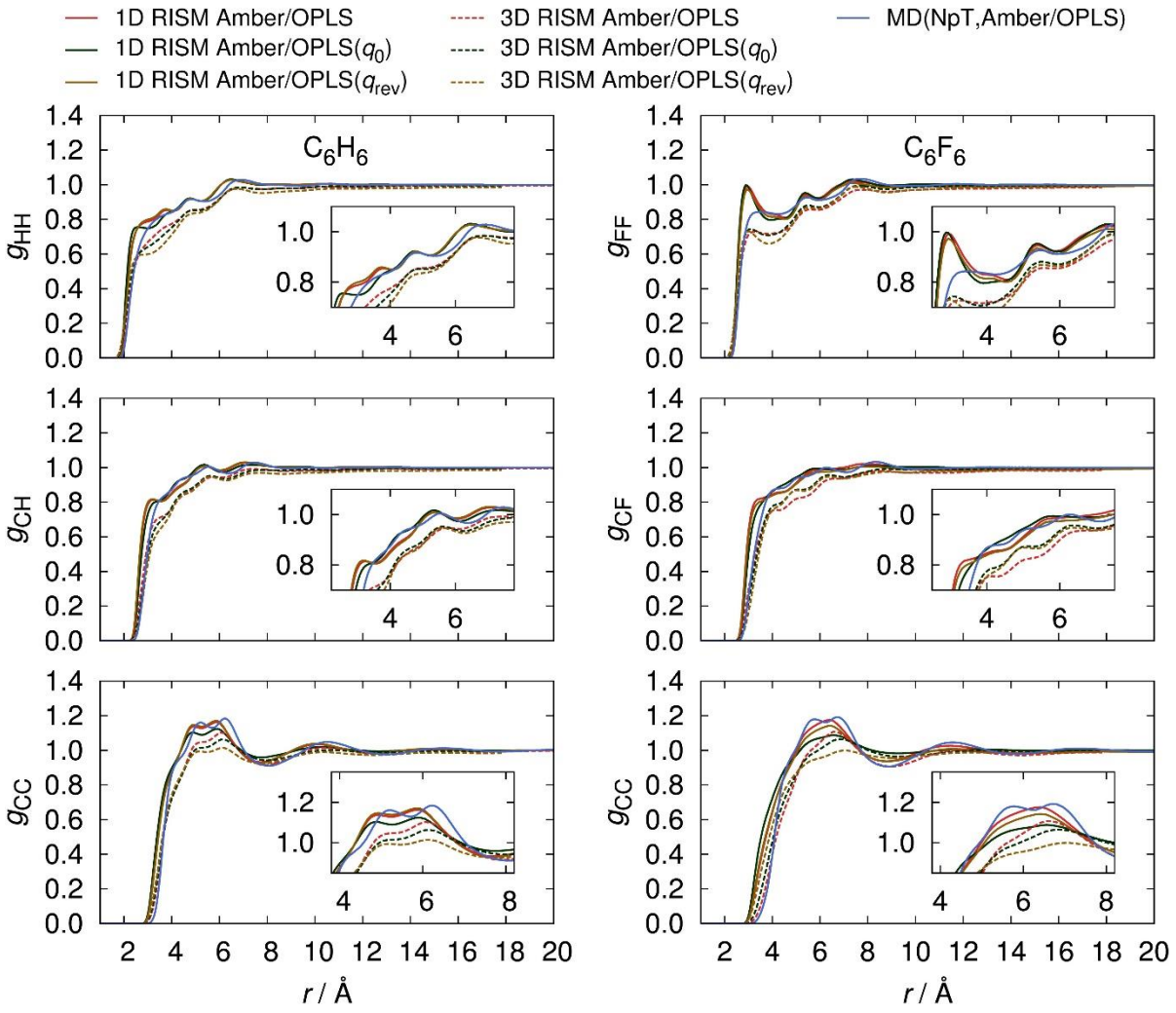


Figure 3.4 Radial distribution functions of pure C_6H_6 and C_6F_6 calculated by either 1D RISM theory or radially averaged 3D RISM theory in comparison with simulated MD results. The underlying 1D RISM data is based on the master thesis of R. Frach.[39]

Regarding the three dimensional solvent distributions of the like and unlike pairs, as illustrated by Figure 3.5 and Figure 3.6, a fundamental difference in the solvation pattern that is not reproduced by 1D RISM theory is obtained. The regular 3D RISM results with the AMBER/OPLS benzene and hexafluorobenzene models are reasonable in the context of matching quadrupole interactions between both ring systems and indicate π-stacking for the unlike C_6H_6/C_6F_6 and C_6H_6/C_6F_6 pairs and T-shaped solvation structures for the like C_6H_6/C_6H_6 and C_6F_6/C_6F_6 pairs. Our results exhibit similarities with the crystal structure packing of C_6H_6 and C_6F_6 [124] and are in very good agreement with the 'reconstructed three dimensional solvent structures estimated by the empirical potential structural refinement' analysis technique[166,167] carried out by Headen et al.[168] An inversion of the solvent quadrupole moment yields an inversed three dimensional solvation pattern, as presented in the first and third column of Figure 3.6. Further turning off the partial charges of the solvent molecules totally disrupts the T-shape like solvation patterns. After reduction of the spatial distribution functions of the q_0- and q_{rev}-model results to radial solvent distribution functions, as also illustrated in Figure 3.4, a leveling effect on the radial averaging to the solvation distribution is observed. This means that, not as strong as for the pure 1D RISM results, but still notable, the 3D RISM radial distribution functions of the reversed q_{rev} and the regular solvent models are very similar to each other, despite being nearly the total opposite in three dimensions. Therefore the C_6H_6/C_6H_6 solvent distribution of the regular solvent model matches the distributions of the q_{rev}-results for the pair C_6F_6/C_6H_6 and so forth.

As illustrated in Figure 3.5 there are just minor differences in the solvation pattern between the pure 3D RISM and the EC-RISM$^{\varphi}$ results, although the solute's exact electrostatic potential is applied for the EC-RISM$^{\varphi}$ calculation and the empirical Amber/OPLS force field partial charges are used for the pure 3D RISM evaluations. The largest difference between the generic 3D RISM and EC-RISM results is exhibited in the hexafluorobenzene in hexafluorobenzene case. Nonetheless the overall solvation patterns are very similar here. The consequences of the radial averaging and the solvent model on Gibbs free energies of small diluted molecules are presented and discussed in the next chapter.

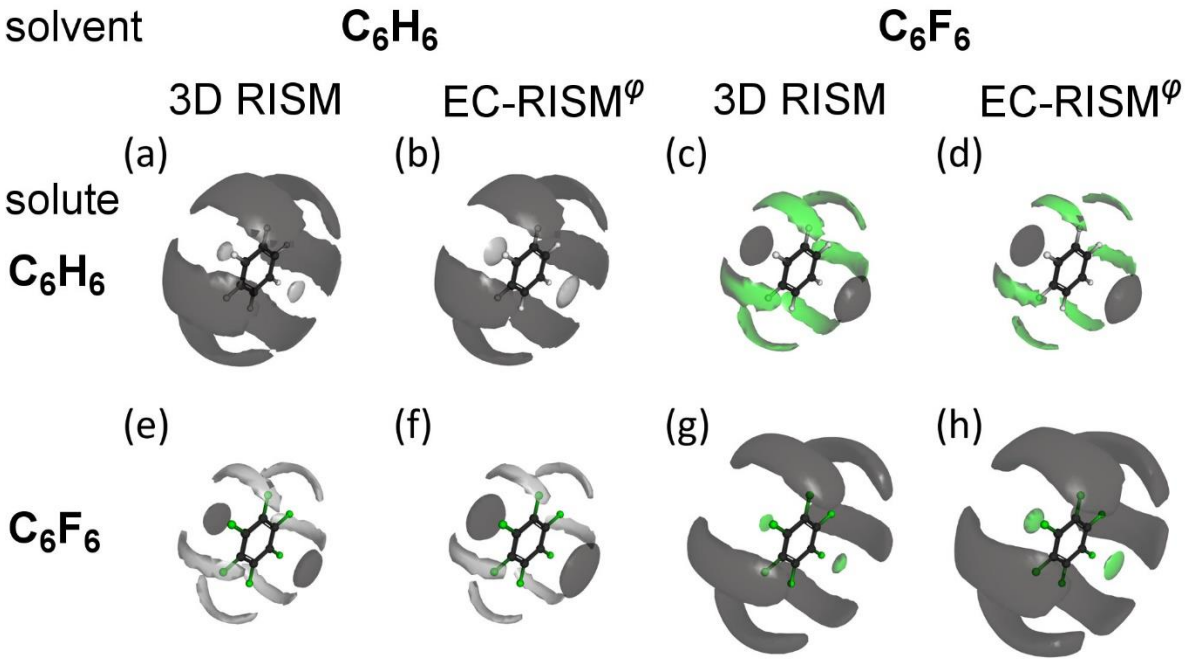


Figure 3.5 Isosurfaces of the solvent distribution functions of C_6H_6 on the left half ((a), (b), (e), (f)) and C_6F_6 on the right half ((c), (d), (g), (h)) around dissolved C_6H_6 (bottom) and C_6F_6 (top) calculated with 3D RISM in the first and third column ((a), (e), (c), (g)) or EC-RISM in the second and forth column ((b), (f), (d), (h)). The underlying data is based on the master thesis of R. Frach.[39] The solvent density is estimated with solvent susceptibilities based on 1D RISM calculation with the Amber/OPLS force field model in combination with the PSE-1 closure. The isosorface value is 12% smaller than the corresponding solvent site density maximum. B3LYP level of theory with 6-31(d) basis set is used for the EC-RISM$^{\varphi}$ iteration. Solvent densities of carbon atoms are illustrated in dark gray, the hydrogen and fluorine solvent densities are set to light gray and green.

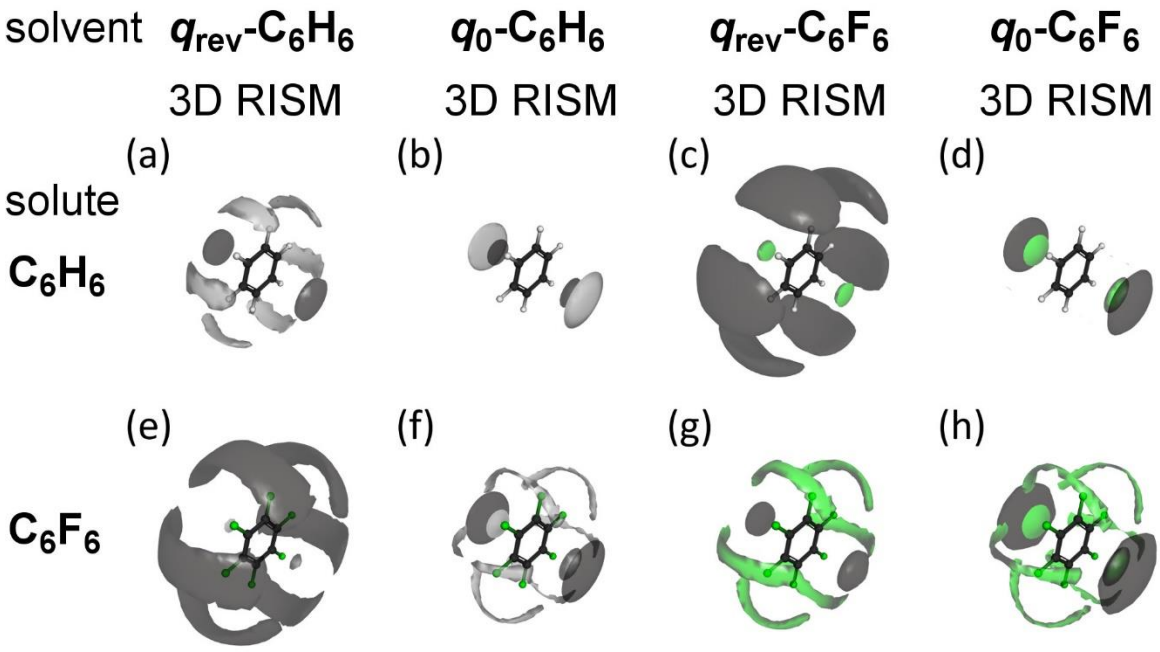


Figure 3.6 Isosurfaces of the spatial solvent distribution functions of the interchanged quadrupole moment solvent models q_{rev}-C_6H_6 in the first column ((a), (e)) and q_{rev}-C_6F_6 in the third column ((c), (g)) around benzene (bottom) and hexafluorobenzene (top) calculated with 3D RISM. Additionally the solvent distributions of the uncharged models q_0-C_6H_6 in the second column ((b), (f)) and q_0-C_6F_6 in the fourth column ((d), (h)) are depicted. Isosurfaces are arranged correspondingly to Figure 3.5.

3.3.2 Gibbs free energies and solvation structures of small molecules

Transfer Gibbs free energies of small molecules for each C_6H_6 and C_6F_6 solvent model were estimated by EC-RISM calculations. Therefore the exact electrostatic potential, which is addressed as EC-RISM$^{\varphi}$ or self-consistent ChelpG partial charges (EC-RISM^{q}) were established in order to describe the solute-solvent potential. EC-RISM results are compared to regular 1D RISM and 3D RISM outcomes for which the electrostatics are modeled using ChelpG partial charges obtained from vacuum quantum-chemical calculations. Due to convergence difficulties, especially for C_6F_6, only results obtained with the PSE-1 closure for both, the solvent susceptibility function and the 3D RISM part are presented. The results are summed up in Tables S8-S22 of the appendix. Figure 3.7 distinguishes between different solvent models and solute-solvent interaction approaches averaged over all quantum-chemical levels of theories. The variability of the averaged results that quantifies the overall influence of the quantum-mechanical level of theory is indicated by the standard deviations listed in Table 3.4. The influence of the quantum-chemical part will be discussed later on.

All integral equation methodologies clearly differentiate between C_6H_6 and C_6F_6 solvation for the process

Molecule solved in C_6H_6 → Molecule solved in C_6F_6

such that the RISM transfer Gibbs free energies show the same trend as the experimental values and exhibit negative values. Note that this is not the case for the PCM results (see Table 3.4) that misrepresent the directionality of the transfer process and have small positive transfer Gibbs free energies for every level of theory and molecule. Therefore we will skip the PCM model results in our figures to focus more on integral equation approaches. The EC-RISM results with the ordinary solvent models exhibit, except for the CO_2 results, a difference to the experiments that is far below 0.3 kcal/mol. Especially the results for argon, oxygen and carbon monoxide are very close to experimental results. In the case of the uncharged (q_0) as well as the reversed partial charge model (q_{rev}), nearly each ΔG_{trans} is by far not negative enough and therefore underestimated the free energy of the transfer process. The sole exception is CO_2, for which the ΔG_{trans} prediction with the uncharged solvent model results has the smallest deviation from experiments. All realistic solvent approaches predict ΔG_{trans} values lower than -0.931 kcal/mol for CO_2, which is off by 0.4 or even more kcal/mol. 1D RISM calculations perform quite well for the three noble gases as well as for nitrogen, but compared to the 3D and EC-RISM results, quite bad for the other compounds.

Table 3.4 Estimated averaged transfer Gibbs free energies for the 9 investigated molecules in kcal/mol. The geometries are estimated through PCM geometry optimization in benzene and hexafluorobenzene solution with HF, PBE0, B3LYP and MP2 combined with different Pople and Dunning basis sets. The listed results are averaged over the quantum-chemical levels of theory. Vacuum partial charges for the different solute geometries are calculated and used for the pure 3D RISM calculations. 1D RISM calculations are only performed for aug-cc-pVTZ geometries and the related vacuum partial charges. Here, the plus-minus sign indicates not the standard error of the averaged results, but the standard deviation giving an estimation of the spread of the underlying unaveraged quantum-mechanical results.

ΔG_{trans}	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
Exp[130]	-0.600	-0.670	-0.600	-0.830	-0.640	-0.690	-0.490	-0.360	-1.230
PCM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.014 ± 0.001	0.004 ± 0.000	0.022 ± 0.002	0.088 ± 0.004	0.007 ± 0.000	0.023 ± 0.002
3D RISM^{qvac}									
Amber/OPLS	-0.370 ± 0.000	-0.419 ± 0.000	-0.541 ± 0.000	-0.666 ± 0.001	-0.598 ± 0.001	-0.675 ± 0.001	-0.965 ± 0.003	-0.125 ± 0.000	-1.060 ± 0.004
Cornell	-0.335 ± 0.000	-0.400 ± 0.000	-0.573 ± 0.000	-0.702 ± 0.001	-0.642 ± 0.001	-0.728 ± 0.002	-1.150 ± 0.004	-0.219 ± 0.000	-1.218 ± 0.004
EC-RISM^{q}									
Amber/OPLS	-0.370 ± 0.000	-0.420 ± 0.000	-0.542 ± 0.000	-0.670 ± 0.002	-0.600 ± 0.002	-0.659 ± 0.002	-0.931 ± 0.004	-0.127 ± 0.003	-1.039 ± 0.003
Cornell	-0.335 ± 0.000	-0.400 ± 0.000	-0.573 ± 0.000	-0.705 ± 0.002	-0.643 ± 0.001	-0.702 ± 0.003	-1.110 ± 0.007	-0.220 ± 0.003	-1.196 ± 0.003
EC-RISM$^{\varphi}$									
Amber/OPLS	-0.372 ± 0.000	-0.423 ± 0.001	-0.555 ± 0.001	-0.721 ± 0.002	-0.631 ± 0.002	-0.723 ± 0.002	-0.935 ± 0.005	-0.334 ± 0.010	-1.031 ± 0.004
Cornell	-0.337 ± 0.000	-0.403 ± 0.001	-0.585 ± 0.001	-0.767 ± 0.002	-0.678 ± 0.002	-0.779 ± 0.003	-1.107 ± 0.007	-0.405 ± 0.009	-1.195 ± 0.004
EC-RISM^{q} Amber/OPLS									
q_0	-0.349 ± 0.000	-0.397 ± 0.000	-0.507 ± 0.000	-0.618 ± 0.000	-0.561 ± 0.000	-0.611 ± 0.000	-0.595 ± 0.000	-0.200 ± 0.000	-0.846 ± 0.000
q_{rev}	-0.323 ± 0.000	-0.364 ± 0.000	-0.467 ± 0.000	-0.564 ± 0.001	-0.514 ± 0.001	-0.583 ± 0.002	-0.100 ± 0.014	-0.227 ± 0.003	-0.607 ± 0.006
1D RISM^{qvac}									
Amber/OPLS	-0.367	-0.417	-0.540	-0.826	-0.727	-0.847	-1.142	-0.280	-1.691

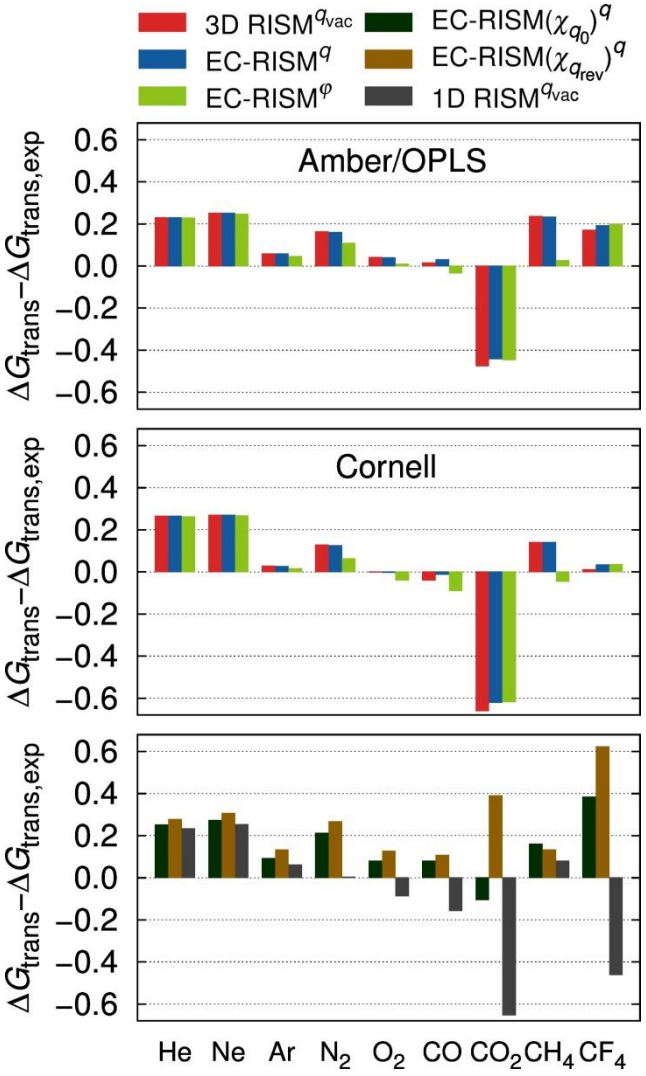


Figure 3.7 Deviation of the averaged transfer Gibbs free energies in kcal/mol for the 9 gas molecules from experimental values of pure 1D RISM, pure 3D RISM, EC-RISM^{q} and EC-RISM$^{\varphi}$ calculations. Therefore Amber/OPLS, Cornell, Amber/OPLS(q_0) and Amber/OPLS(q_{rev}) solvent models in quantum-chemical level of theories consistent with the precomputed PCM solute structures have been used.

Figure 3.8 (as well as Figure S10 and Table S24) illustrates the root mean squared deviation (RMSD) of each combination between a quantum-chemical level of theory and EC-RISM as well as 3D RISM. As depicted in this figure, the basis set and the level of theory have no systematic influence on the pure 3D RISM results. In contrast the EC-RISM^{q} and even more the EC-RISM$^{\varphi}$ results improve with enlargement of the basis. The results for the q_{rev} solvent models also improve, but not as stringent as for the more physically plausible models. The smallest deviation is achieved by MP2/aug-cc-pVTZ in combination with EC-RISM$^{\varphi}$, which is just slightly better than MP2/aug-cc-pVQZ/EC-RISM$^{\varphi}$. Furthermore the influence of the basis set is more pronounced than the influence of the *ab initio* or DFT level of theory.

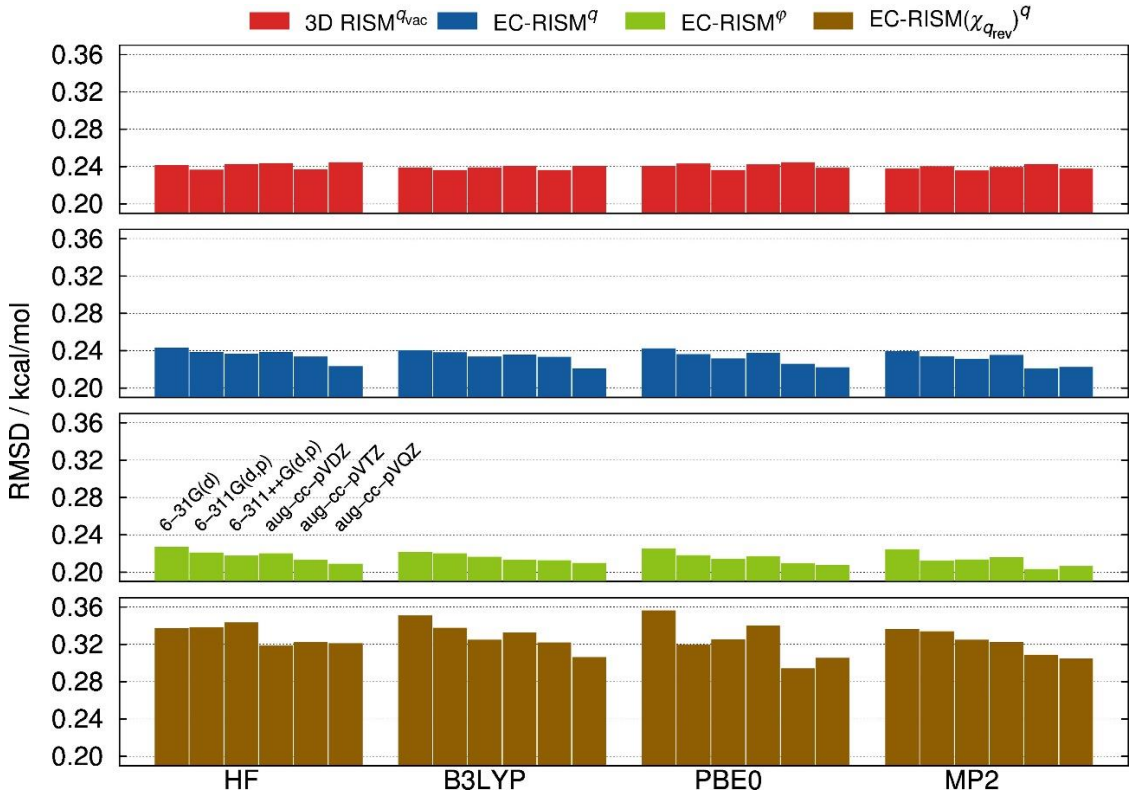


Figure 3.8 Root mean square deviations from experimental values of pure 3D RISM, EC-RISM^{q} and EC-RISM$^{\varphi}$ calculations with the AMBER/OPLS and Amber/OPLS(q_{rev}) solvent models averaged over all molecules and differentiated by the quantum-chemical level of theory.

Figure 3.8 illustrates that the results change more with the solvation model than with the quantum-chemical level of theory. With help of the comparison of the overall mean signed errors (MSE) and root mean square deviations (RMSD) from experimental values, shown in Table 3.5 and Figure 3.9, it is recognizable that the combination of the Amber/OPLS force field with the EC-RISM methodology and exact electrostatic potential is the best choice for transfer Gibbs free energies calculations. The Cornell solvent model has the smallest MSE, but a much larger RMSD for each methodology. Overall the results are improved by EC-RISM in comparison with generic 3D RISM and are again enhanced by exact electrostatics in comparison with ChelpG partial charges. Surprisingly the second best RMSD results is provided by the uncharged (q_0) EC-RISM^{q} Amber/OPLS model calculations. However, here the ΔG_{trans} values are always slightly higher than experimental values, resulting in a higher MSE and therefore indicating an unbalanced, systematically shifted methodology. The 1D RISM results deviate nearly as much as the unphysical q_{rev}-model, but with a much better MSE. The unphysical q_{rev}-model results exhibit the highest difference to experimental values and the largest MSE between all RISM approaches.

Table 3.5 Calculated root mean squared deviations (RMSD) and mean signed (MSE) in kcal/mol. The geometries are estimated through PCM geometry optimization in C_6H_6 and C_6F_6 solution with HF, PBE0, B3LYP and MP2 theory combined with different basis sets. The results are averaged over the quantum-chemical levels of theory and all molecules. Vacuum partial charges for the solutes are applied for pure 3D RISM calculations. 1D RISM calculations are only performed for aug-cc-pVTZ geometries.

	RMSD	MSE
PCM	0.778	0.783
3D RISM$^{q\text{vac}}$		
Amber/OPLS	0.239	0.086
Cornell	0.278	0.018
EC-RISM^{q}		
Amber/OPLS	0.233	0.094
Cornell	0.266	0.028
EC-RISM$^{\varphi}$		
Amber/OPLS	0.214	0.048
Cornell	0.259	-0.018
EC-RISM^{q}(Amber/OPLS)		
q_0	0.220	0.178
q_{rev}	0.325	0.295
1D RISM$^{q\text{vac}}$(Amber/OPLS)	0.316	-0.091

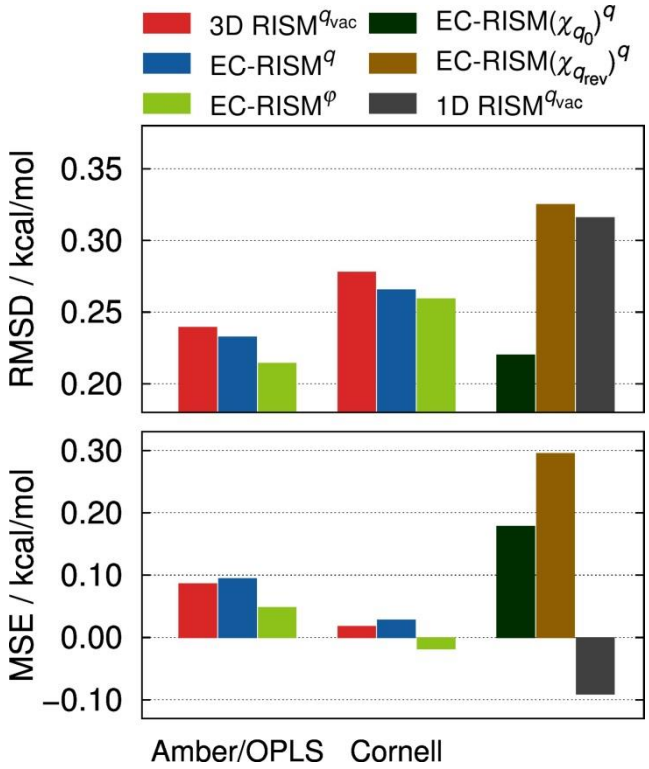


Figure 3.9 Overall root mean square deviations (RMSD) and mean signed errors (MSE) of 1D RISM, 3D RISM and EC-RISM results for transfer Gibbs free energies estimated with Amber/OPLS, Cornell, Amber/OPLS(q_0) and Amber/OPLS(q_{rev}) solvent models in combination with geometries consistent to the quantum-chemical level of theory of the precomputed PCM solute structures.

The q_0-model results can be understood as properties of solutes diluted in a pure dispersive solvent. Therefore the results in Figure 3.9 and Table 3.5 already illustrate that dispersive effects are important for the transfer Gibbs free energy of small molecules, but that a realistic representation of the electrostatics improves these results that are in general not negative enough. In contrast, as in the q_{rev}-case, where the solvent partial charges are reversed, the addition of electrostatic solute-solvent interactions worsens the results by further increasing the ΔG_{trans} values. To quantify the influence of dispersive solvent-solute interactions the so-called fraction of the purely dispersive interaction to the transfer Gibbs free energy can be estimated by

$$x_{disp} = \frac{\Delta G_{trans}(M) - \Delta G_{trans}(\text{EC-RISM}(\chi_{q_0}))}{\Delta G_{trans}(M)} \tag{48}$$

with M as any solvation method other than EC-RISM^q with the q_0-solvation model. Due to small, but fundamental differences in the solvation structures, as illustrated by the solvent distributions in Figure 3.2, a consistent x_{disp} value of around 94% for all molecules, even for the noble gases, is achieved. Lower values clearly indicate the influence of electrostatics. In consequence a juxtaposition of the q_0-model results with the ΔG_{trans} values, obtained with AMBER/OPLS force fields, reveals that approximately 87% - 89% of the transfer free energy is due to pure dispersion (see Table 3.6). The fraction x_{disp} in Table 3.6 is therefore evaluated and averaged over the same level of theories that are used in combination with the q_0-solvent model to get a fair comparison. The transfer Gibbs free energies of CO_2 and CF_4 have the highest difference to the q_0-model results and thereby, with 63 – 64 % and 79 – 82 % the lowest amount of dispersive influence. Surprisingly there is nearly no difference between the 3D RISM and EC-RISM^q results with B3LYP/6-311G(d,p) and B3LYP/aug-cc-pVTZ level of theory. In addition the EC-RISM results with exact electrostatics and with ChelpG point charges are very similar, except for N_2, O_2 and CO, where x_{disp} is lower for the EC-RISM$^\varphi$ results.

Table 3.6 Fraction of the dispersive influence on the Gibbs free energies of transfer x_{disp} estimated by equation (48). Furthermore the averaged fractions $\langle x_{disp} \rangle$ are listed. The given values are only calculated for the level of theory consistent to the examined q_0-model calculations.

x_{disp}	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4	$\langle x_{disp} \rangle$
3D RISM	94%	95%	94%	93%	93%	90%	63%	99%	79%	89%
EC-RISM^q	94%	95%	94%	93%	93%	92%	64%	99%	81%	89%
EC-RISM$^\varphi$	93%	93%	91%	86%	88%	85%	64%	98%	82%	87%

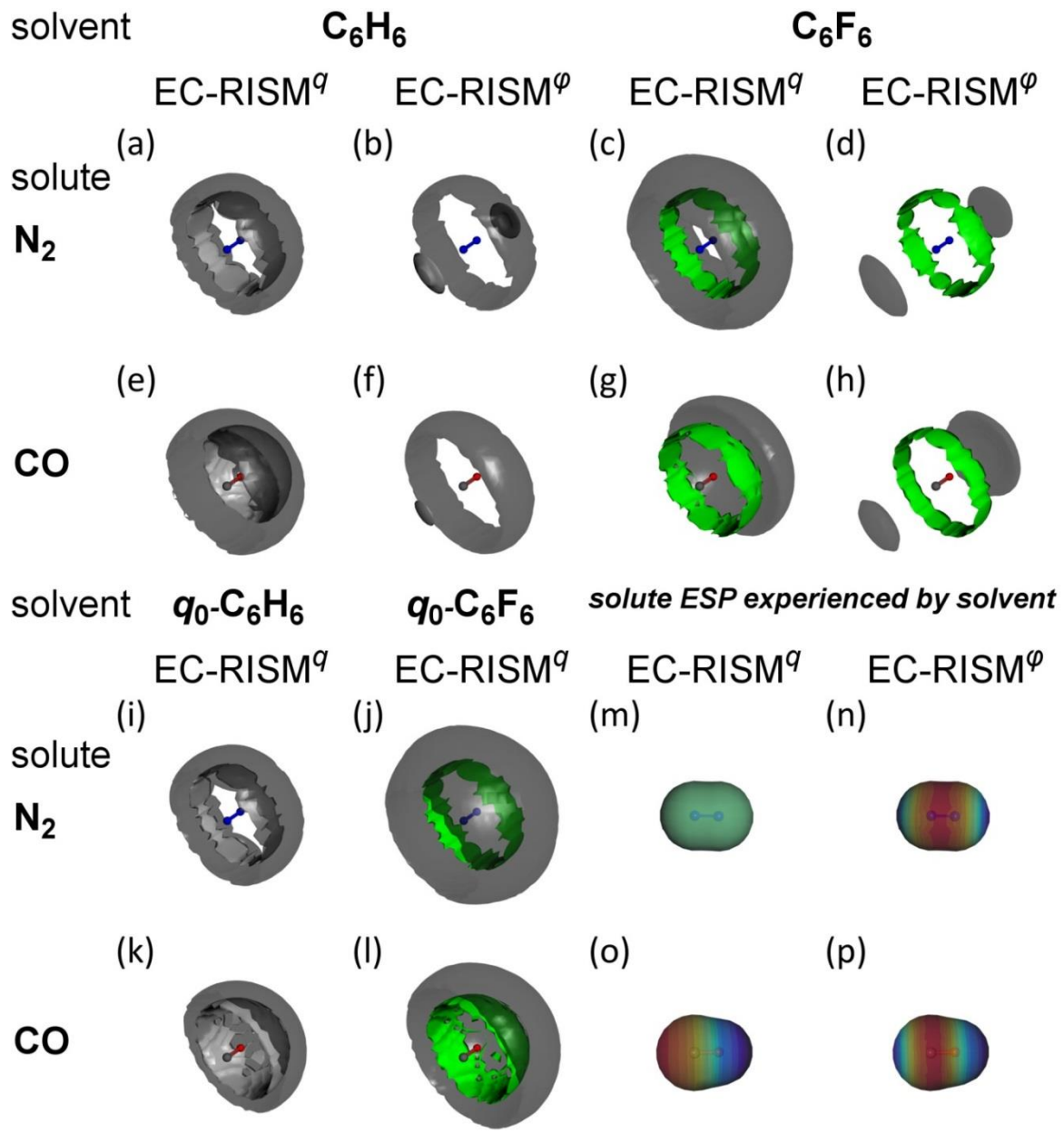


Figure 3.10 Solvent density of ordinary AMBER/OPLS C_6F_6 (c, d, g, h) and C_6H_6 (top, a, b, e, f), q_0-C_6H_6 (i, k) and q_0-C_6F_6 (j, l) around N_2 (a, b, c, d, i, j) and CO (e, f, g, h, k, l), calculated by EC-RISM with B3LYP/aug-cc-pVTZ level of theory. Electrostatic solute-solvent interactions are described by ChelpG partial charges (a, c, e, g, i-l) or by exact electrostatic interactions (b, d, g, h). Isosurfaces of carbon (dark gray, transparent), hydrogen (light gray) and fluorine (green) are set to 96% of the corresponding maximum solvent density. The solute electrostatic potential (ESP), here mapped onto isosurfaces, experienced by the solvent of the corresponding EC-RISM model, are illustrated (m-p) with blue as negative and red as positive ESP regions.

To explain the influence of the electrostatic representation on the transfer Gibbs free energies and in particular on the differences between quadrupolar and dispersive solvation, it is helpful to focus on the N_2 and CO results. Figure 3.10 depicts that the EC-RISM^q solvation structure around N_2 is dominated by two solvent density "rings". A comparison of these densities with the solvent structure around N_2 that is achieved with the q_0-solvation model shows that these rings are already gained by pure dispersive solvation. Figure 3.10 (m)-(n)

illustrates the electrostatic potential of the solute that is experienced by the solvent in the corresponding EC-RISM calculation. In the EC-RISM^q calculations of N_2 the solvent does not experience any electrostatic solute-solvent interaction at all, for which in contrast the solute in the EC-RISM$^\varphi$ calculation interacts by higher electrostatic moments with the solvent. Consequently the solvent structure of the EC-RISM$^\varphi$ counterpart drastically differs in comparison with the EC-RISM^q and EC-RISM(q_0)^q results. Thus N_2 and O_2 can be polarized by EC-RISM$^\varphi$ despite both molecules have partial charges of zero. This higher multipole interaction can be regarded as the cause for the influenceability of the N_2, as well as the O_2 results by electrostatic interactions and consecutive polarization, which is not accessible by regular 3D RISM and EC-RISM^q models. These findings are supported by calculations findings of Mohan et al.[169] for the C_6F_6/N_2 dimer. Another example for which the x_{disp} values are not similar is CO, although it is a dipolar molecule and multipolar influences appear not to be substantial. But again, as for N_2, the EC-RISM$^\varphi$ densities around CO are clearly different to the EC-RISM^q results, which is in fact again explainable by the complex molecular electrostatic potential of CO (see Figure 3.10 (o)-(p)).

To further differentiate between energetic contributions and to investigate polarization effects we split the transfer Gibbs free energy into

$$\Delta G_{\text{trans}} = \Delta E_{0,\text{trans}} + \Delta E_{\text{pol,trans}} + \Delta\mu^{\text{ex}}_{0,\text{trans}} + \Delta\mu^{\text{ex}}_{\text{pol,trans}}, \tag{49}$$

where

$$\Delta E_{0,\text{trans}} = E_{0,\text{sol}}(\text{C}_6\text{F}_6) - E_{0,\text{sol}}(\text{C}_6\text{H}_6) \tag{50}$$

and

$$\Delta\mu_{0,\text{trans}} = \mu^{\text{ex}}_{0,\text{sol}}(\text{C}_6\text{F}_6) - \mu^{\text{ex}}_{0,\text{sol}}(\text{C}_6\text{H}_6) \tag{51}$$

are the unpolarized contributions of the electronic energy (eq. (50)) and the excess chemical potential (eq.(51)) of the solution that correspond to the 3D RISM results with vacuum partial charges. The other two contributions that describe the influence of self-consistent polarization are described by

$$\begin{aligned}\Delta E_{\text{pol,trans}} &= E_{\text{sol}}(\text{C}_6\text{F}_6) - E_{\text{sol}}(\text{C}_6\text{H}_6) - \Delta E_{0,\text{trans}} \\ \Delta\mu^{\text{ex}}_{\text{pol,trans}} &= \mu^{\text{ex}}(\text{C}_6\text{F}_6) - \mu^{\text{ex}}(\text{C}_6\text{H}_6) - \Delta\mu_{0,\text{trans}}.\end{aligned} \tag{52}$$

The estimated contributions, averaged over all quantum-mechanical levels of theory, are illustrated in Figure 3.11 for the EC-RISM^q results with the AMBER/OPLS solvent model. The calculated transfer Gibbs free energy is dominated by the unpolarized excess chemical

potential of the solution. The changes of ΔE or $\Delta\mu^{ex}$ that are caused by polarization are timid and play a minor role for the ΔG_{trans} value predictions here.

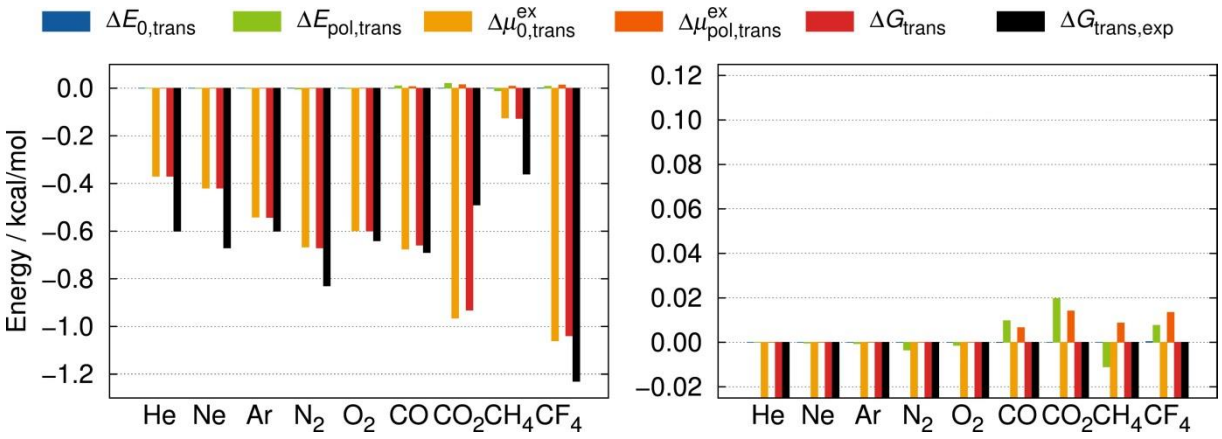


Figure 3.11 Averaged energetic contributions to the calculated transfer Gibbs free energy for the EC-RISM^q model with AMBER/OPLS solvent susceptibility functions in comparison with experimental values. The left panel displays the whole energetic scale and the right part of the figure focuses on smaller contributions.

3.4 Concluding remarks

In summary a computationally inexpensive, but nevertheless realistic solvent model for benzene and hexafluorobenzene was developed. Both solvent models yield reasonable radial pair distribution functions and spatial solvent distributions. Furthermore, with transfer Gibbs free energies with a total RMSD of just about 0.2 to 0.3 kcal/mol our models predict thermodynamic observables in good agreement with experiments and with a comparable accuracy to methods like scaled particle theory.[170,171,130] Beside 3D RISM and EC-RISM provide insight on the dominating structural differences between benzene and hexafluorobenzene solvents. Furthermore it is shown that 1D RISM loses crucial solvent structure information by radial averaging. Thereby the shortcomings of the 1D RISM/HNC combination for short distances, which are often amended by a repulsive bridge correction, are improved by 3D RISM. The applied non-additive approach that utilizes totally uncharged solvent models is able separate the dispersion and electrostatic influences and quantifies their importance for the calculated transfer Gibbs free energies. In this context it is observed that a proper representation of the dispersion interaction already captures a great amount of the Gibbs free energy difference of C_6H_6 and C_6F_6 solutions and therefore of the Henry constants. Based on the q_{rev}-model studies it is revealed that changes in the electronic energy or excess

chemical potential, due to electrostatic interactions and the associated polarization, are nevertheless meaningful and systematically improve the results, when a realistic representation is used and worsens the results, when an unphysical mixture of electrostatics and dispersion (q_{rev}) is applied. Thus the electrostatics can be viewed as essential for a differential solvation effect, whereas the dispersion defines a baseline for the difference between benzene and hexafluorobenzene.

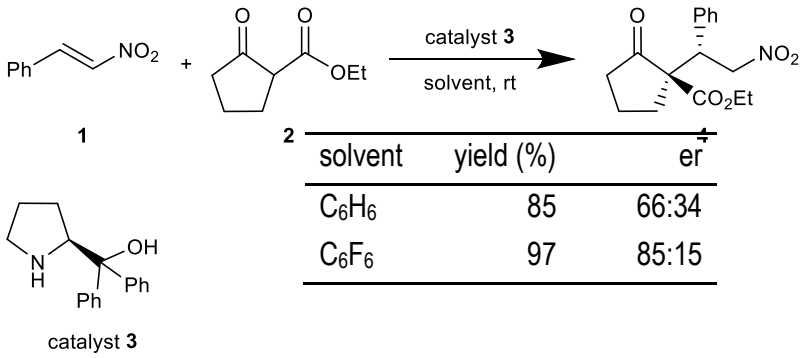
In consequence the use of exact solute-solvent electrostatics, which yields a more realistic solvent description and slightly, but systematical, better thermodynamics is emphasized and supported by more reasonable solvation patterns that are not properly depicted by a point charge representation.

The combination of an exact electrostatic representation for the solute-solvent interactions yields the possibility to incorporate higher multipole moments in a simple way and to treat phenomena like sigma hole effects[172] between solute and solvent molecules without specially parametrized force fields. Furthermore the applied models, combined with quantum-mechanical solute descriptions in the context of the EC-RISM framework, are useful starting points for further investigations of kinetics and thermodynamics of more complex processes. The next chapter will exactly focus on one of these more complex processes and focus on an unsymmetrically catalyzed organic reaction in both solvents.

4 SOLVENT CONTROLLED STEREOSELECTIVITY OF NONDIPOLAR LIQUIDS

4.1 Introduction

It is well known that potential energy surfaces of chemical reactions are dramatically influenced by the solvent surrounding. Due to this influence, a drastic change of the solvent's polarity is often connected to a change in the yield of a desired product. This is often explained with a stabilization of specific intermediate states or reaction products due to favored or unfavored dipolar interactions. Nevertheless yields can already by largely altered by e. g. much smaller perturbations originating from nondipolar solvents and additives. Remarkable examples of such behavior are summarized by Sugiishi et al.[173], where stereoselectivity-control is achieved by replacement of common solvents with perfluorinated organic compounds. In fact the modification of stereoselectivity by fluorinated additives, solvents or functional groups is just one aspect of the not fully understood 'fluorine effect'.[174,175,176,177] One extraordinary reaction of this kind was highlighted by Cavallo and coworkers were perfluorobenzene (C_6F_6) improves the stereoselectivity of a nitro-Michael addition (see scheme 1) in comparison with various other solvents, including also benzene (C_6H_6).[30] Both solvents, C_6H_6 and C_6F_6, are very similar in their molecular structure and their macroscopic physical properties. Still they exhibit different influences on the presented Michael addition, which are not explainable by simple dipolar models. In their work Lattanzi et al.[30] suggested that C_6F_6 enlarges the energetic gap between the transition states (TSs) that correspond to the two resulting enantiomers.



solvent	yield (%)	er
C_6H_6	85	66:34
C_6F_6	97	85:15

Scheme 1. Asymmetric catalysed nitro-Michael addition with results for the yield and the enantiomeric ratios (er).[30]

They underpinned their hypothesis with quantum-chemical calculations of these transition states (see Figure 4.1), postulating a concerted mechanism and modeled the solvent influence by placing one or respectively two explicit C_6F_6 molecules in the TSs. These DFT calculations revealed that one C_6F_6 molecule is energetically favored if it is stacked on the enolate π-orbitals of the transition states forming a steric hindrance for the reaction system.

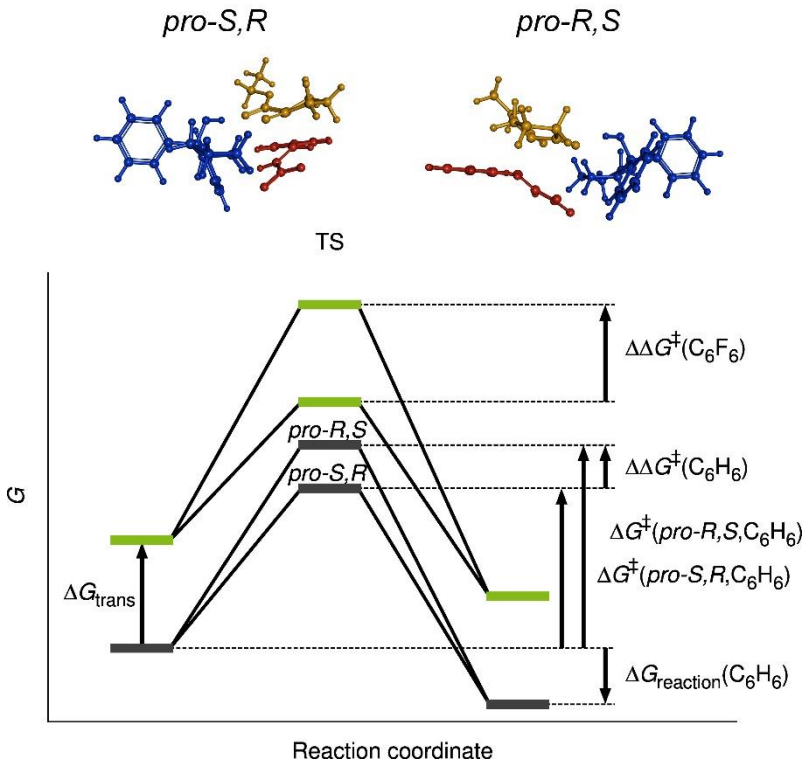


Figure 4.1 Transition states (top, abbrev. TS) of the nitro-Michael addition illustrated in scheme 1 as proposed by Lattanzi et al.[30] estimated with BP98[100,178]/SVP[179] level of theory. The TSs consist of the catalyst (blue), the enolate (yellow) and the nitrostyrene (red). Furthermore a schematic illustration of the reaction kinetics (bottom) in C_6H_6 (grey) and C_6F_6 (green) solvation with various and throughout the chapter important Gibbs free energies are implied.

Here the energetics of the reactants, products and TSs of the presented nitro-Michael addition in C_6H_6 and C_6F_6 solution are investigated with the embedded cluster reference interaction site model (EC-RISM).[23] EC-RISM represents a thermodynamically consistent methodology for infinite dilution conditions and is capable to estimate granular solvent distributions without the need of manual molecule placement as it was performed in the preliminary literature work.[30] This is achieved by a combination of a full quantum-chemical treatment for the molecules of interest and a classical density functional theory approach for the solvent. For this purpose the three dimensional reference interaction site model (3D RISM) is applied.[20,21] By solution of the nonlinear system of equations that consists of the 3D RISM equation itself and a so called closure relation, it is possible to predict the solvation structure and the excess chemical potential μ^{ex}. By calculating the solvent distribution function we gain insight into specific, directional solvent-solute interactions and influences in thermodynamic equilibrium conditions. Additionally EC-RISM gets access to the per particle Gibbs free energy of the solute by

$$G_{\mathrm{sol}} = E_{\mathrm{sol}} + \mu^{\mathrm{ex}}. \tag{53}$$

The electronic energy of the solute E_{sol} is estimated by a self-consistent solution of the electrostatic solute-solvent and solvent-solute interaction.

Based on the different solute free energies we scrutinize the reaction kinetics by calculating substantial free energy differences as illustrated in Figure 4.1 and elaborated by the thermodynamic cycle in Figure 4.2. Their proper definition is given by

$$\begin{aligned}
&G_{\mathrm{educts}}(v) = G_{\mathrm{sol}}(\mathbf{1}, v) + G_{\mathrm{sol}}(\mathbf{2}, v) + G_{\mathrm{sol}}(\mathbf{3}, v)\\
&G_{\mathrm{prod}}(v) = G_{\mathrm{sol}}(\mathbf{4}, v) + G_{\mathrm{sol}}(\mathbf{3}, v)\\
&\Delta G_{\mathrm{reaction}}(v) = G_{\mathrm{prod}}(v) - G_{\mathrm{educts}}(v)\\
&\Delta G^{\ddagger}(TS, v) = G_{\mathrm{sol}}(TS, v) - G_{\mathrm{educts}}(v)\\
&\Delta G_{\mathrm{trans}}(u) = G_{\mathrm{sol}}(u, \mathrm{C_6F_6}) - G_{\mathrm{sol}}(u, \mathrm{C_6H_6}),
\end{aligned} \tag{54}$$

with *v* being any of the two solvents (C_6H_6 and C_6F_6), *u* as an arbitrary solute and *TS*, which stands for one of the two transition states, called *pro-R,S* or *pro-S,R*. The *pro-R,S* transition state yields the *S,R*-product and the *pro-S,R* transition states reacts further to the *R,S*-product. The bold letters in eq. (54) represent the compounds according to scheme 1. The upper half of thermodynamic cycle in Figure 4.2 gives an overview of the reaction pathway in benzene (grey) and hexafluorobenzene (green) and depicts that the Gibbs free energy difference of the

TSs, the reaction barrier $\Delta G^{\ddagger}(TS,v)$ is normalized to the per particle Gibbs free energy of the educts $G_{sol}(TS,v)$.

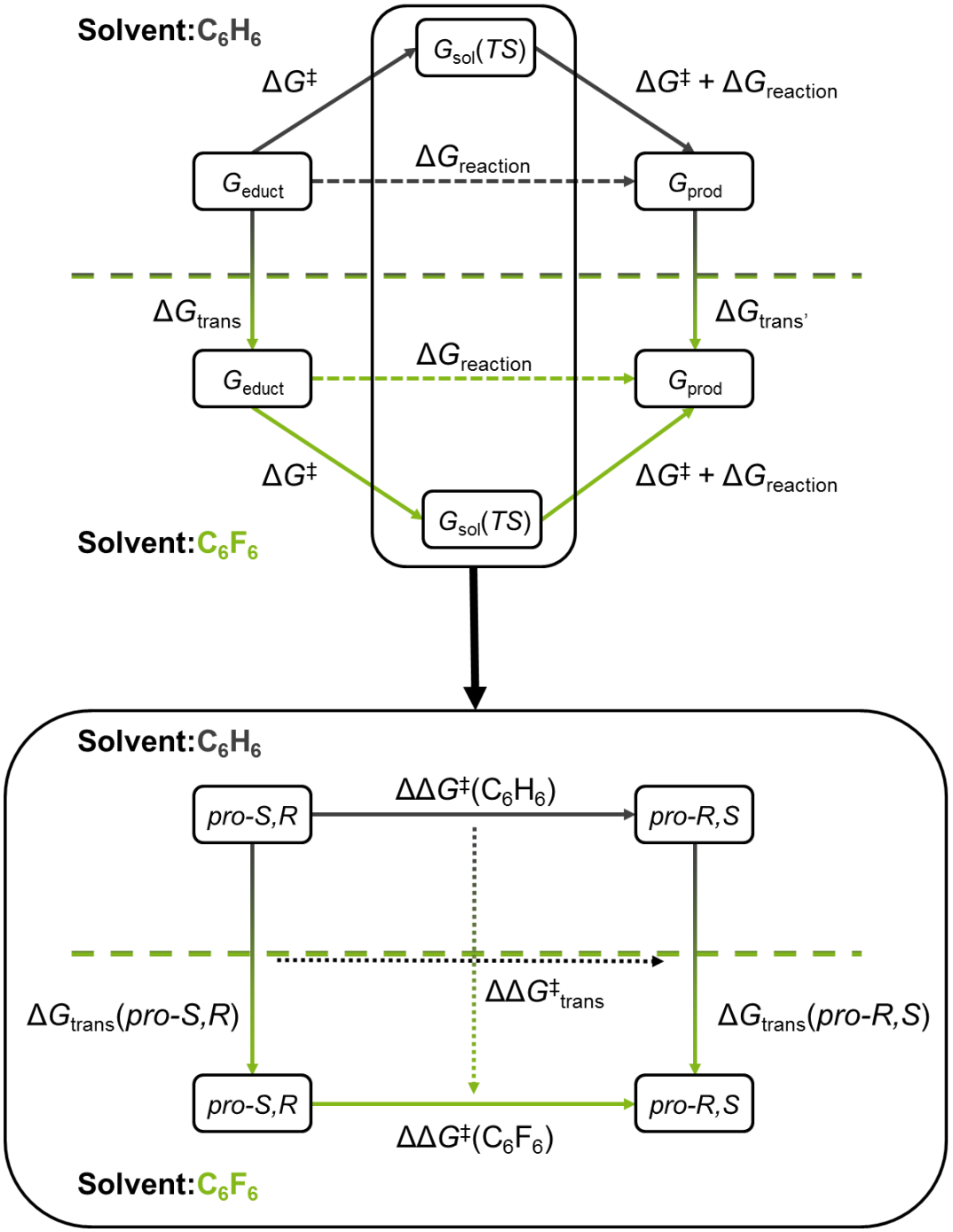


Figure 4.2 Thermodynamic cycle illustrating important Gibbs free energies. The upper half illustrates the whole reaction pathway. The lower half focusses on the energetics between the transition states (*pro-R,S* and *pro-S,R*). Reactions or processes in benzene are indicated with grey arrows, reactions in hexafluorobenzene with green ones. The dashed arrows in the upper half indicate that a direct reaction is not possible and has to cross a transition state.

The properties in eq. (54) are mainly useful to estimate two key values in order to characterize the stereoselectivity of the nitro-Michael addition. One of these values is the difference between the normalized Gibbs free energies of the transition states

$$\Delta\Delta G^{\ddagger}(v) = \Delta G^{\ddagger}(pro\text{-}R,S,v) - \Delta G^{\ddagger}(pro\text{-}S,R,v)\,. \tag{55}$$

As indicated by the lower half of Figure 4.2, which concentrates on the transitions states and thereby related properties, the $\Delta\Delta G^{\ddagger}$ values are solvent specific. Here they will be applied to approximate the enantiomeric ratio (er) and therefore the stereoselectivity of the reaction by

$$\text{er} \approx (1+\exp[-\beta\Delta\Delta G^{\ddagger}])^{-1}/(1+\exp[\beta\Delta\Delta G^{\ddagger}])^{-1}, \tag{56}$$

in one of the solvents. This relationship will be properly deduced in the next chapter. The other key value is the difference between the $\Delta\Delta G^{\ddagger}$ values in both solvents

$$\begin{aligned}\Delta\Delta G^{\ddagger}_{\text{trans}} &= \Delta\Delta G^{\ddagger}(\text{C}_6\text{F}_6) - \Delta\Delta G^{\ddagger}(\text{C}_6\text{H}_6)\\ &= \Delta G^{\ddagger}(pro\text{-}R,S,\text{C}_6\text{F}_6) - \Delta G^{\ddagger}(pro\text{-}S,R,\text{C}_6\text{F}_6)\\ &\quad -\left[\Delta G^{\ddagger}(pro\text{-}R,S,\text{C}_6\text{H}_6) - \Delta G^{\ddagger}(pro\text{-}S,R,\text{C}_6\text{H}_6)\right]\\ &= \Delta G_{\text{trans}}(pro\text{-}R,S) - \Delta G_{\text{trans}}(pro\text{-}S,R).\end{aligned} \tag{57}$$

It quantifies the change of stereoselectivity after solvent transfer in terms of a free energy difference. In the following an increase in stereoselectivity after exchange of benzene with hexafluorobenzene will be referred to as stereoselectivity enhancement. A positive $\Delta\Delta G^{\ddagger}_{\text{trans}}$ value corresponds to the experimentally observed stereoselectivity enhancement from an er of 66:34 in C_6H_6 to an er of 85:15 in C_6F_6. Consequently a negative value shows a decreasing discrimination, which is not consistent with the experimental results.

Note that in this chapter on one hand the discrimination between specific solutes in both solvents, the ΔG_{trans} values (see lower half of Figure 4.2, arrows from top to bottom) and on the other hand the discrimination between two different solutes, respectively the transition states by $\Delta\Delta G^{\ddagger}$ in one of the solvents (see lower half of Figure 4.2, arrows from left to right) will be investigated. A combination of both, the $\Delta\Delta G^{\ddagger}_{\text{trans}}$ describes the surprising effect of hexafluorobenzene on the considered organic reaction. As deduced by eq. (54) and the lower half of Figure 4.2 the difference of the transfer Gibbs free energies of both TSs $\Delta G_{\text{trans}}(pro\text{-}R,S) - \Delta G_{\text{trans}}(pro\text{-}S,R)$ is equivalent to $\Delta\Delta G^{\ddagger}(\text{C}_6\text{F}_6) - \Delta\Delta G^{\ddagger}(\text{C}_6\text{H}_6)$ and consequential the motivation to use only two Δ-operators.

This chapter will demonstrate that there is a much greater correspondence between EC-RISM results and experimental enantiomeric ratios than between polarizable continuum model (PCM) results[13] and experiments. We emphasize the importance of the quadrupole electrostatic interactions for the stereoselectivity enhancement by application of an uncharged, pure dispersive solvent model. Supportively, an investigation of the electronic structure calculation method is performed that considers the influence of various quantum-chemical

levels of theory and different transition state structure geometries. Further the robustness of EC-RISM is tested by modification of the solute parametrization, applying either physically motivated or unphysical force fields. Additionally the Gibbs free energies are broken down into the solutes electronic and the solvent components to determine their separate impact on the stereoselectivity.

4.2 Methodology

For simplification and feasibility some assumptions to model the stereoselectivity of the nitro-Michael reaction are used. Firstly, the most aspects of solvent influence are investigated for the two TSs (see Figure 4.1) that have been estimated by Lattanzi et al[30], estimated by transition state searches with BP86/SVP level of theory. Therefore, as illustrated by Figure 4.1, only free energy differences of these transition states, without a prior exhaustive search for other reaction pathways, are calculated. Furthermore it is assumed that the resulting stereoselectivity is not controlled by the product energetics, but by reaction kinetics, thus by the energy of the two TSs. This assumption is applicable when both products are energetically equivalent, leading to common kinetically controlled stereoselective reactions.[180] Both assumptions allow for the deduction of a simple formula to estimate the enantiomeric ratio (er) of the reaction. Starting from the approximation that the probability p_i to receive one of the two enantiomers i is directed by

$$p_i \approx p_{j\to i} = \frac{\exp[-\beta G_{\mathrm{sol}}(j \to i)]}{Z} \tag{58}$$

of the corresponding transition state $j\to i$, the enantiomeric ratio for the here considered nitro-Michael addition can be defined as

$$\begin{aligned} \mathrm{er} &= [R,S]/[S,R] \approx p_{R,S} \,/\, p_{S,R} \\ &\approx p_{pro-S,R\to R,S} \,/\, p_{pro-R,S\to S,R} \\ &= \frac{\exp[-\beta G_{\mathrm{sol}}(pro\text{-}S, R \to R, S)]}{Z} \Big/ \frac{\exp[-\beta G_{\mathrm{sol}}(pro\text{-}R, S \to S, R)]}{Z}. \end{aligned} \tag{59}$$

The $j\to i$-notation implies that transition state j reacts to the product i and $G_{\mathrm{sol}}(j\to i)$ is the corresponding Gibbs free energy of the transition state. This assumption is applicable when reverse reactions are prohibited and each product i is achieved by just one reaction pathway

with the corresponding transition state $j \rightarrow i$. In equation eq. (59) the partition function is defined as

$$Z = \exp[-\beta G_{\mathrm{sol}}(pro\text{-}S, R \rightarrow R, S)] + \exp[-\beta G_{\mathrm{sol}}(pro\text{-}R, S \rightarrow S, R)] \tag{60}$$

for this reaction. After normalization of (59) to the Gibbs free energy of the educts G_{educts} the enantiomeric ratios can be calculated by

$$\begin{aligned} \mathrm{er} &\approx \frac{\exp[-\beta(G_{\mathbf{sol}}(pro\text{-}S, R) - G_{\mathrm{educts}})]}{Z\exp[\beta G_{\mathrm{educts}}]} \Big/ \frac{\exp[-\beta G_{\mathbf{sol}}(pro\text{-}R, S) - G_{\mathrm{educts}}]}{Z\exp[\beta G_{\mathrm{educts}}]} \\ &= \frac{\exp[-\beta \Delta G^{\ddagger}(pro\text{-}S, R)]}{Z\exp[\beta G_{\mathrm{educts}}]} \Big/ \frac{\exp[-\beta \Delta G^{\ddagger}(pro\text{-}R, S)]}{Z\exp[\beta G_{\mathrm{educts}}]} \end{aligned} \tag{61}$$

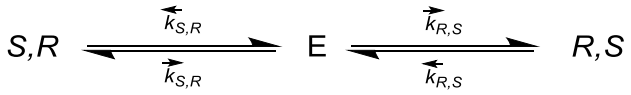
using a shorter notation for $j \rightarrow i$ that corresponds to the definition at the beginning of this chapter. This redefinition is applicable in this case, because the products and their transition states are assumed to be precisely assignable. By introducing $\Delta\Delta G^{\ddagger}$ and insertion of eq. (60) into eq. (61) the er can be further reduced to

$$\mathrm{er} \approx \frac{\exp[-\beta \Delta G^{\ddagger}(pro\text{-}S, R)]\exp[\beta \Delta G^{\ddagger}(pro\text{-}S, R)]}{Z\exp[\beta G_{\mathrm{educts}}]\exp[\beta \Delta G^{\ddagger}(pro\text{-}S, R)]} \Big/ \frac{\exp[-\beta \Delta G^{\ddagger}(pro\text{-}R, S)]\exp[\beta \Delta G^{\ddagger}(pro\text{-}R, S)]}{Z\exp[\beta G_{\mathrm{educts}}]\exp[\beta \Delta G^{\ddagger}(pro\text{-}R, S)]}$$

$$= (1 + \exp[-\beta \Delta\Delta G^{\ddagger}])^{-1} / (1 + \exp[\beta \Delta\Delta G^{\ddagger}])^{-1}. \tag{62}$$

The resulting equation (62) is the theoretically accessible maximum er for the investigated reaction, defining the major thermodynamic influence by application of a simple two-state model without reverse reaction. Consequentially it is assumed that the enantiomeric ratio of the products is the same as the concentration ratio of the transition states. This approach ignores the effects of reaction kinetics that, for example by back-and-forth reactions, lower the er over time until the thermodynamic equilibrium, the racemic mixture with a ratio of 50:50 is reached.

Within a follow up approach a more sophisticated model is obtained by consideration of simple textbook reaction kinetics. Therefore the compound concentrations are described as illustrated by scheme 2. This approach is applied to check whether or not eq. (61) is valid or if kinetic effects are important. In consequence relative reaction times are accessible.



Scheme 2. Ansatz for the kinetic model. Therefore it is assumed that all educts are in equimolar concentrations. Additionally all educt concentrations are combined to one educt concentration [E]. Furthermore back-and-forth reaction is allowed.

The concentrations of all educt compounds are condensed to one educt concentration [E] implying equimolarity. Here, a back-and-forth reaction of both products to the educts is allowed and therefore a system of differential equations of the form

$$\begin{pmatrix} -(\vec{k}_{R,S}+\vec{k}_{S,R}) & \overleftarrow{k}_{R,S} & \overleftarrow{k}_{S,R} \\ \vec{k}_{R,S} & -\overleftarrow{k}_{R,S} & 0 \\ \vec{k}_{S,R} & 0 & -\overleftarrow{k}_{S,R} \end{pmatrix} \begin{pmatrix} [\mathrm{E}](t) \\ [R,S](t) \\ [S,R](t) \end{pmatrix} = \frac{\partial}{\partial t} \begin{pmatrix} [\mathrm{E}](t) \\ [R,S](t) \\ [S,R](t) \end{pmatrix} \tag{63}$$

is obtained. Solution of eq. (63) results in the concentration profiles of [*R*,*S*], [*S*,*R*] and [E] over time *t*. The system of equations (63) is analytically solvable by the method of variation of constants, involving the determination of the roots of a third order polynomial. The solution of the third order polynomial was performed with the Mathematica 10.02 software.[181] Additionally initial values of $[\mathrm{E}](t=0) = [\mathrm{E}]_0$ for the starting educt concentration and $[R,S](t=0) = 0$ and $[S,R](t=0) = 0$ for the products were applied. Solving eq. (63) results yields a solutions for the integrated rate laws with the form

$$\begin{aligned}
[\mathrm{E}](t) &= \frac{[\mathrm{E}]_0}{C}\left\{2\overleftarrow{k}_{R,S}\overleftarrow{k}_{S,R}k' + \exp\left[-\frac{1}{2}\left(k_{\mathrm{sum}}+k'\right)t\right]\cdot\right. \\
&\left\{\left(\vec{k}_{R,S}\vec{k}_{S,R}(\vec{k}_{R,S}-\overleftarrow{k}_{R,S}+\vec{k}_{S,R}) + (\vec{k}_{R,S}+\overleftarrow{k}_{R,S})(\vec{k}_{R,S}+\vec{k}_{S,R})\overleftarrow{k}_{S,R} - \vec{k}_{R,S}\overleftarrow{k}_{S,R}^2\right)\left(1-\exp[k't]\right)\right. \\
&\quad + \vec{k}_{R,S}\overleftarrow{k}_{S,R}k'\exp[k't] + \vec{k}_{S,R}\overleftarrow{k}_{R,S}k'\exp[k't] \\
&\quad \left.\left. + \vec{k}_{R,S}\overleftarrow{k}_{S,R}k' + \vec{k}_{S,R}\overleftarrow{k}_{R,S}k'\right\}\right\} \\
[R,S](t) &= \frac{[\mathrm{E}]_0}{C}\vec{k}_{R,S}\left\{2\overleftarrow{k}_{S,R}k' + \exp\left[-\frac{1}{2}\left(k_{\mathrm{sum}}+k'\right)t\right]\cdot\right. \\
&\left\{\left(\vec{k}_{R,S}\overleftarrow{k}_{S,R} - \overleftarrow{k}_{S,R}(\vec{k}_{S,R}+\overleftarrow{k}_{S,R}) + \overleftarrow{k}_{R,S}(2\vec{k}_{S,R}+\overleftarrow{k}_{S,R})\right)\left(\exp[k't]-1\right)\right. \\
&\quad \left.\left. - \overleftarrow{k}_{S,R}k'\exp[k't] - \overleftarrow{k}_{S,R}k'\right\}\right\} \\
[S,R](t) &= \frac{[\mathrm{E}]_0}{C}\vec{k}_{S,R}\left\{2\overleftarrow{k}_{R,S}k' + \exp\left[-\frac{1}{2}\left(k_{\mathrm{sum}}+k'\right)t\right]\cdot\right. \\
&\left\{\left(\vec{k}_{S,R}\overleftarrow{k}_{R,S} - \overleftarrow{k}_{R,S}(\vec{k}_{R,S}+\overleftarrow{k}_{R,S}) + \overleftarrow{k}_{S,R}(2\vec{k}_{R,S}+\overleftarrow{k}_{R,S})\right)\left(\exp[k't]-1\right)\right. \\
&\quad \left.\left. - \overleftarrow{k}_{R,S}k'\exp[k't] - \overleftarrow{k}_{R,S}k'\right\}\right\}
\end{aligned} \tag{64}$$

with

$$
\begin{aligned}
k_{\text{sum}} &= \vec{k}_{R,S} + \vec{k}_{S,R} + \overleftarrow{k}_{R,S} + \overleftarrow{k}_{S,R} \\
k' &= \sqrt{k_{\text{sum}}^2 - 4(\vec{k}_{R,S}\overleftarrow{k}_{S,R} + \vec{k}_{S,R}\overleftarrow{k}_{R,S} + \overleftarrow{k}_{R,S}\overleftarrow{k}_{S,R})} \\
C &= 2(\vec{k}_{R,S}\overleftarrow{k}_{S,R} + \vec{k}_{S,R}\overleftarrow{k}_{R,S} + \overleftarrow{k}_{R,S}\overleftarrow{k}_{S,R})k'.
\end{aligned} \tag{65}
$$

Further the reaction rate constants are approximated by

$$
\begin{aligned}
\vec{k}_{R,S} &= A\exp[-\beta\Delta G^{\ddagger}(pro\text{-}S,R)] \\
\vec{k}_{S,R} &= A\exp[-\beta\Delta G^{\ddagger}(pro\text{-}R,S)] \\
\overleftarrow{k}_{R,S} &= A\exp[-\beta(\Delta G^{\ddagger}(pro\text{-}S,R) - \Delta G_{\text{reaction}})] \\
\overleftarrow{k}_{S,R} &= A\exp[-\beta(\Delta G^{\ddagger}(pro\text{-}R,S) - \Delta G_{\text{reaction}})].
\end{aligned} \tag{66}
$$

Since the Arrhenius pre-exponential factor A are not estimated directly, A will be considered as equal for each reaction rate constant. Using approach (66) for the rate constants often occurring multiplication of rate constants are transformable to

$$
\begin{aligned}
\vec{k}_{R,S}\overleftarrow{k}_{S,R} &= A^2\exp[-\beta\Delta G^{\ddagger}(pro\text{-}S,R)]\exp[-\beta(\Delta G^{\ddagger}(pro\text{-}R,S) - \Delta G_{\text{reaction}})] \\
&= A^2\exp[-\beta(\Delta G^{\ddagger}(pro\text{-}S,R) - \Delta G_{\text{reaction}})]\exp[-\beta\Delta G^{\ddagger}(pro\text{-}R,S)] \\
&= \vec{k}_{S,R}\overleftarrow{k}_{R,S}.
\end{aligned} \tag{67}
$$

With (67) the reaction rates (65) can be rearranged to

$$
\begin{aligned}
[\mathrm{E}](t) &= 2\frac{[\mathrm{E}]_0}{C}\Big\{\overleftarrow{k}_{R,S}\overleftarrow{k}_{S,R}k' + \vec{k}_{R,S}\overleftarrow{k}_{S,R}(\vec{k}_{R,S} + \vec{k}_{S,R} + k')\exp\big[-(k_{\text{sum}} + k')t/2\big] \\
&\quad + \vec{k}_{R,S}\overleftarrow{k}_{S,R}(k' - \vec{k}_{R,S} - \vec{k}_{S,R})\exp\big[-(k_{\text{sum}} - k')t/2\big]\Big\} \\
[R,S](t) &= \frac{[\mathrm{E}]_0}{C}\vec{k}_{R,S}\overleftarrow{k}_{S,R}\Big\{2k' + (3\vec{k}_{R,S} + \overleftarrow{k}_{R,S} - \vec{k}_{S,R} - \overleftarrow{k}_{S,R} - k')\exp\big[-(k_{\text{sum}} + k')t/2\big] \\
&\quad + (\vec{k}_{S,R} + \overleftarrow{k}_{S,R} - 3\vec{k}_{R,S} - \overleftarrow{k}_{R,S} - k')\exp\big[-(k_{\text{sum}} - k')t/2\big]\Big\} \\
[S,R](t) &= \frac{[\mathrm{E}]_0}{C}\vec{k}_{R,S}\overleftarrow{k}_{S,R}\Big\{2k' + (3\vec{k}_{S,R} + \overleftarrow{k}_{S,R} - \vec{k}_{R,S} - \overleftarrow{k}_{R,S} - k')\exp\big[-(k_{\text{sum}} + k')t/2\big] \\
&\quad + (\vec{k}_{R,S} + \overleftarrow{k}_{R,S} - 3\vec{k}_{S,R} - \overleftarrow{k}_{S,R} - k')\exp\big[-(k_{\text{sum}} - k')t/2\big]\Big\}.
\end{aligned} \tag{68}
$$

Further applying (67) to the definition of constant C a beneficial new definition

$$
\begin{aligned}
C &= 2(\vec{k}_{R,S}\overleftarrow{k}_{S,R} + \vec{k}_{S,R}\overleftarrow{k}_{R,S} + \overleftarrow{k}_{R,S}\overleftarrow{k}_{S,R})k' = 2(2\vec{k}_{R,S}\overleftarrow{k}_{S,R} + \overleftarrow{k}_{R,S}\overleftarrow{k}_{S,R})k' \\
&= 2(2\vec{k}_{R,S}\overleftarrow{k}_{S,R} + \vec{k}_{R,S}\exp[\beta\Delta G_{\text{reaction}}]\overleftarrow{k}_{S,R})k' \\
&= 2\vec{k}_{R,S}\overleftarrow{k}_{S,R}k'(\exp[\beta\Delta G_{\text{reaction}}] + 2)
\end{aligned} \tag{69}
$$

is gained and the reaction rates can be further modified to

$$
\begin{aligned}
[\mathrm{E}](t) &= [\mathrm{E}]_0 \left\{ \frac{1}{2\exp[-\beta\Delta G_{\mathrm{reaction}}]+1} + \frac{k'+\vec{k}_{R,S}+\vec{k}_{S,R}}{k'(\exp[\beta\Delta G_{\mathrm{reaction}}]+2)} \exp\left[-k_{\mathrm{eff},1}t\right] \right. \\
&\left. + \frac{k'-\vec{k}_{R,S}-\vec{k}_{S,R}}{k'(\exp[\beta\Delta G_{\mathrm{reaction}}]+2)} \exp\left[-k_{\mathrm{eff},2}t\right] \right\} \\
[R,S](t) &= [\mathrm{E}]_0 \left\{ \frac{1}{\exp[\beta\Delta G_{\mathrm{reaction}}]+2} + \frac{\vec{k}_{S,R}+\overleftarrow{k}_{S,R}-3\vec{k}_{R,S}-\overleftarrow{k}_{R,S}-k'}{2k'(\exp[\beta\Delta G_{\mathrm{reaction}}]+2)} \exp\left[-k_{\mathrm{eff},1}t\right] \right. \\
&\left. + \frac{3\vec{k}_{R,S}+\overleftarrow{k}_{R,S}-\vec{k}_{S,R}-\overleftarrow{k}_{S,R}-k'}{2k'(\exp[\beta\Delta G_{\mathrm{reaction}}]+2)} \exp\left[-k_{\mathrm{eff},2}t\right] \right\} \\
[S,R](t) &= [\mathrm{E}]_0 \left\{ \frac{1}{\exp[\beta\Delta G_{\mathrm{reaction}}]+2} + \frac{\vec{k}_{R,S}+\overleftarrow{k}_{R,S}-3\vec{k}_{S,R}-\overleftarrow{k}_{S,R}-k'}{2k'(\exp[\beta\Delta G_{\mathrm{reaction}}]+2)} \exp\left[-k_{\mathrm{eff},1}t\right] \right. \\
&\left. + \frac{3\vec{k}_{S,R}+\overleftarrow{k}_{S,R}-\vec{k}_{R,S}-\overleftarrow{k}_{R,S}-k'}{2k'(\exp[\beta\Delta G_{\mathrm{reaction}}]+2)} \exp\left[-k_{\mathrm{eff},2}t\right] \right\}
\end{aligned}
\tag{70}
$$

with $k_{\mathrm{eff},1} = (k_{\mathrm{sum}} + k')/2$ and $k_{\mathrm{eff},2} = (k_{\mathrm{sum}} - k')/2$. From eq. (70) an effective form

$$
[i](t) = [\mathrm{E}]_0 \left(x_{t\to\infty}(i) + x_1(i)\exp[-k_{\mathrm{eff},1}t] + x_2(i)\exp[-k_{\mathrm{eff},2}t] \right) \tag{71}
$$

is obtained, for which $x_{t\to\infty}$ is recognizable as the molar ratio in thermodynamic equilibrium of compound i. The assignment of the compound specific time-independent constants $x_1(i)$ and $x_2(i)$ is straightforward.

4.3 Computational details

As quantum-mechanical level of theory, the B3LYP[100,101,102] hybrid functional with 6-311+G(d,p) basis was applied as general approach for all combinations of solute force fields, solvent models and TS structures.[103] All product and educt structures were generated by IEF-PCM[13] geometry optimization with the Gaussian09 program package.[182] The original transition state structures of Lattanzi et al.[30] were used for every calculation where no other information is specified. These transition state structures were estimated with the BP98[100,178] hybrid functional in combination with a SVP basis set.[179] Subsequently Lattanzi et al performed singly point calculations with the M06 functional and the larger TZVP basis set.[179] However, the influence of the transition state structure on the reaction pathway was investigated by re-optimization of the given original transition state structures to give new TSs corresponding to the B3LYP/6-311+G(d,p) level of theory. Furthermore, as purely wave

function based methodology, MP2/6-311+G(d,p) level of theory was chosen. Additionally, using the ORCA(version 3.0)[183] program package, the M06 hybrid functional[184] was applied. The computationally very demanding transition state search was not conducted for these electronic structure methods. The consequences are discussed later in the results part of this chapter. An empirical dispersion correction like the common D3 correction of Grimme et al.[185] was not taken into account, because it cancels out when differences between both solvents are estimated. The D3 correction only depends on the molecular structure, not on the wave function and in consequence has no influence on the stereoselectivity enhancement, due to the use of identical TSs in both solvents.

Benzene and hexafluorobenzene solvent susceptibilities based on the AMBER/OPLS force field[139,140,141] (see chapter 3.2.1) were applied for 3D RISM calculations. Furthermore, the uncharged solvent models (q_0) were used to investigate the influence of missing electrostatic solvent-solute interactions closely. Due to its indulgent convergence behavior and the minor influence, at least for 1D RISM calculation, as shown earlier, the PSE-1 closure was chosen. 3D RISM was solved on a cubic grid with 150x150x150 grid points and a grid spacing of 0.2 Å. A modified direct inversion of iterative subspace method (MDIIS)[186] was used, for which a maximum residuum norm of 10^{-5} for the direct correlation function was selected, to give a converged solution. The treated compounds were parametrized with the generalized amber force field (GAFF, version 1.4, March 2010)[187,112] to model solute-solvent dispersive interaction by the Lennard-Jones interaction potential. The corresponding parameters are given in the appendix. To analyze to which degree the force field influences the results, additional calculations were conducted either with the OPLS all-atom force field or an approach, for which all Lennard-Jones parameters were equal (for parameters see appendix).[188] A convergence threshold of $2 \cdot 10^{-4}$ kJ mol^{-1} for the free energy was applied. In consequence of the findings of the previous chapter exact electrostatics were computed (EC-RISM$^{\varphi}$) to account for the coupling between the solute and the solvent. Additionally calculations with the point charge representation were performed (EC-RISM^{q}) to test the influence of the electrostatic representation on molecules that are more complex than the small molecules treated in chapter 3. To achieve EC-RISM convergence in a feasible manner only the last EC-RISM step was evaluated with higher level of theory and the Hartree-Fock procedure was applied before.

4.4 Results

This section is segmented into five parts.

1. Pure **IEF-PCM results are compared to EC-RISM findings** with **exact electrostatics** (**EC-RISM$^{\varphi}$**), **point charges** (**EC-RISM^{q}**) or the **uncharged solvent models** (**EC-RISM(χ_{q0})^{q}**). The level of theory is uniformly set to B3LYP/6-311+G(d,p) in this chapter and the original TS structures (BP98/SVP) are applied.
2. The **influence of the quantum-mechanical level of theory** is investigated as well as the applied solute parametrization.
3. The consequences of **TS re-optimization** are demonstrated.
4. **The excess chemical potential of the solution** and **the solute's electronic energy** are **separately analyzed** to quantify which parts determine the stereoselectivity enhancement.
5. The findings of the ansatz for the **reaction kinetics** are demonstrated.

4.4.1 Solvent model comparison

The results for the reaction barriers ($\Delta G^{\ddagger}$), the transfer Gibbs free energies (ΔG_{trans}), the Gibbs free energies of reaction ($\Delta G_{reaction}$) and the $\Delta\Delta G^{\ddagger}$ values are summarized in Figure 4.3. The achieved enantiomeric ratios, the $\Delta\Delta G^{\ddagger}$ and $\Delta\Delta G^{\ddagger}_{trans}$ values are additionally listed in Table 4.1.

Table 4.1 Comparison of calculated $\Delta\Delta G^{\ddagger}$ and er values with experimental results.[30] The geometries of the transition states of Lattanzi et al.[30] (BP98/SVP) are used and treated with a B3LYP/6-311+G(d,p) level of theory. IEF-PCM or EC-RISM/PSE-1 solvation models are applied. Two different solvent susceptibilities are used for EC-RISM, either a regular force field parametrized C_6H_6 and C_6F_6 or uncharged C_6H_6 and C_6F_6 models (q_0-models). Electrostatics are computed with the exact electrostatic potential of the solute, indicated by EC-RISM$^{\varphi}$ or with partial charges (EC-RISM^{q}). Equation (4) is used to evaluate the er.

Model	$\Delta\Delta G^{\ddagger}(C_6H_6)$ / kJ mol^{-1}	er(C_6H_6)	$\Delta\Delta G^{\ddagger}(C_6F_6)$ / kJ mol^{-1}	er(C_6F_6)	$\Delta\Delta G^{\ddagger}_{trans}$ / kJ mol^{-1}
Literature	1.64[a]	66:34	4.39[a]	85:15	2.75
PCM	4.93	88:12	4.75	87:13	-0.18
EC-RISM$^{\varphi}$	3.65	81:19	5.78	91:9	2.13
EC-RISM^{q}	3.62	81:19	5.65	91:9	2.03
EC-RISM^{q}(q_0-models)	1.88	68:32	1.82	68:32	-0.06

[a] Literature $\Delta\Delta G^{\ddagger}$ are calculated by inversion of equation (4) with the throughout applied assumptions.

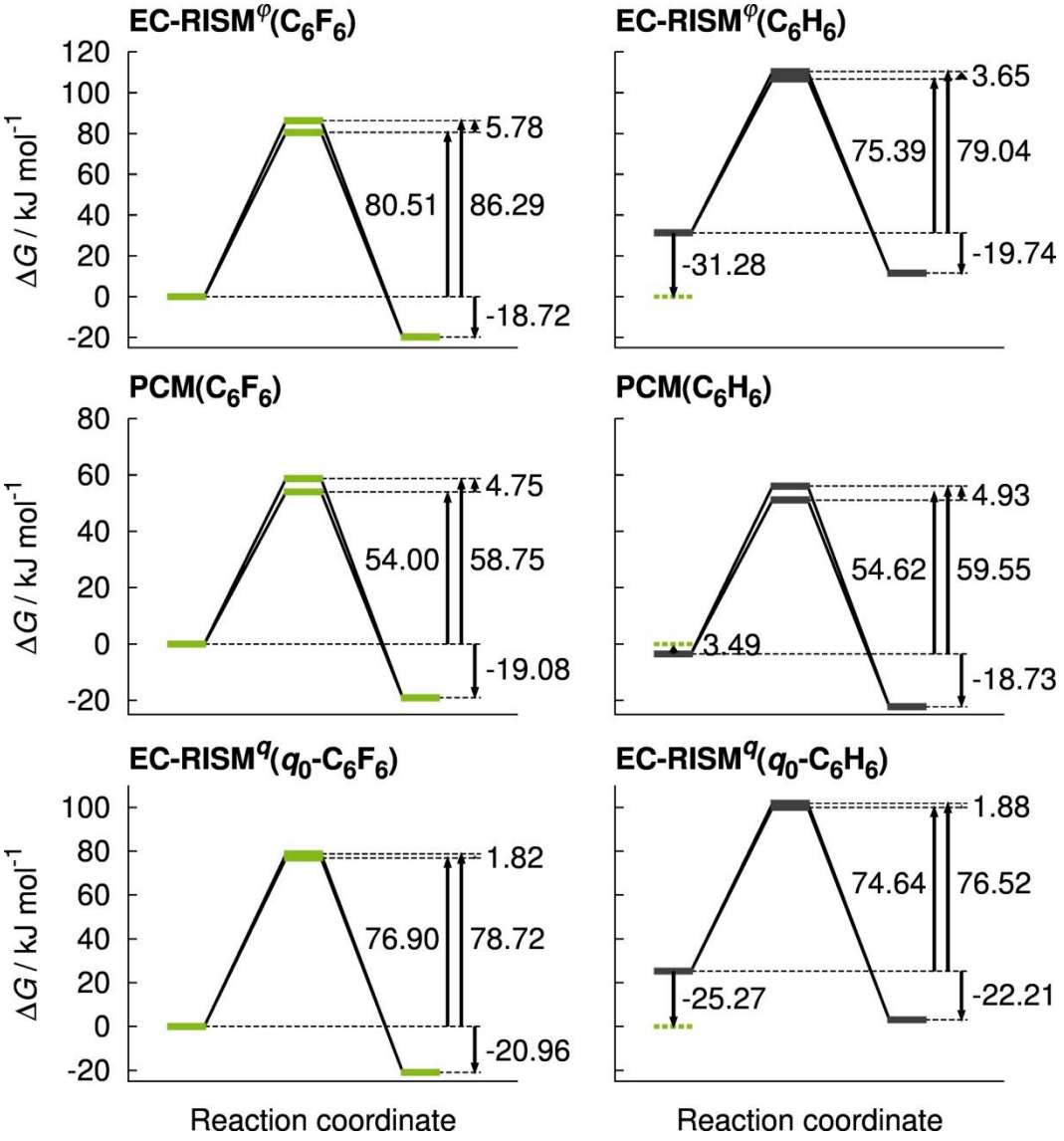


Figure 4.3 Reaction profile results for the considered nitro-Michael addition in hexafluorobenzene (left) and benzene (right) for regular EC-RISM$^{\varphi}$ (top), PCM (mid) and the EC-RISM^{q} approaches with uncharged solvents q_0-C_6H_6 and q_0-C_6F_6 (bottom). The results are normalized to the Gibbs free energy of the educts G_{educt} in hexafluorobenzene of the applied solvation approach.

The enantiomeric ratios in Table 4.1 with values of 88:12 in C_6H_6 and 87:13 in C_6F_6 for PCM and 81:19 in C_6H_6 and 91:9 in C_6F_6 for EC-RISM$^{\varphi}$ are slightly overestimated by both solvation methodologies, but consistent with the experimental stereoselectivity (66:34 in C_6H_6, 85:15 in C_6F_6), favoring the *R,S*-product. Nevertheless EC-RISM$^{\varphi}$ is capable to distinguish between the solvent specific $\Delta\Delta G^{\ddagger}$ values in C_6H_6 (3.65 kJ mol^{-1}) and C_6F_6 (5.78 kJ mol^{-1}), whereas PCM solvation with 4.93 kJ mol^{-1} and 4.75 kJ mol^{-1} displays nearly identical $\Delta\Delta G^{\ddagger}$s. This results, as indicated by the negative $\Delta\Delta G^{\ddagger}_{trans}$ value of -0.18 kJ mol^{-1} for PCM, yield the wrong solvent effect on the stereoselectivity by hexafluorobenzene. Otherwise the stereoselectivity enhancement is well captured by EC-RISM$^{\varphi}$, for which the er improves from 66:34 to 85:15 by solvent exchange. A comparison of the EC-RISM$^{\varphi}$ results

with the EC-RISM^{q} model calculations, as listed in Table 4.1, indicates that exact electrostatics slightly improve the $\Delta\Delta G^{\ddagger}_{trans}$ value, but that the point charge representation also captures the experimentally verified effect of hexafluorobenzene.

The right stereoisomer is also determined by the uncharged, purely dispersive solvent models q_0-C_6H_6 and q_0-C_6F_6, but the estimated ers are, with 68:32 in C_6H_6 and 68:32 in C_6F_6, as consequence of the narrower difference between the transition state barriers (as depicted by Figure 4.3), much lower. Furthermore enhancement of the stereoselectivity by hexafluorobenzene is not reproduced by these solvent models. Therefore solvent-solute electrostatic interactions are essential to express the experimental stereoselectivity enhancement by C_6F_6, but not to predict the dominating enantiomer.

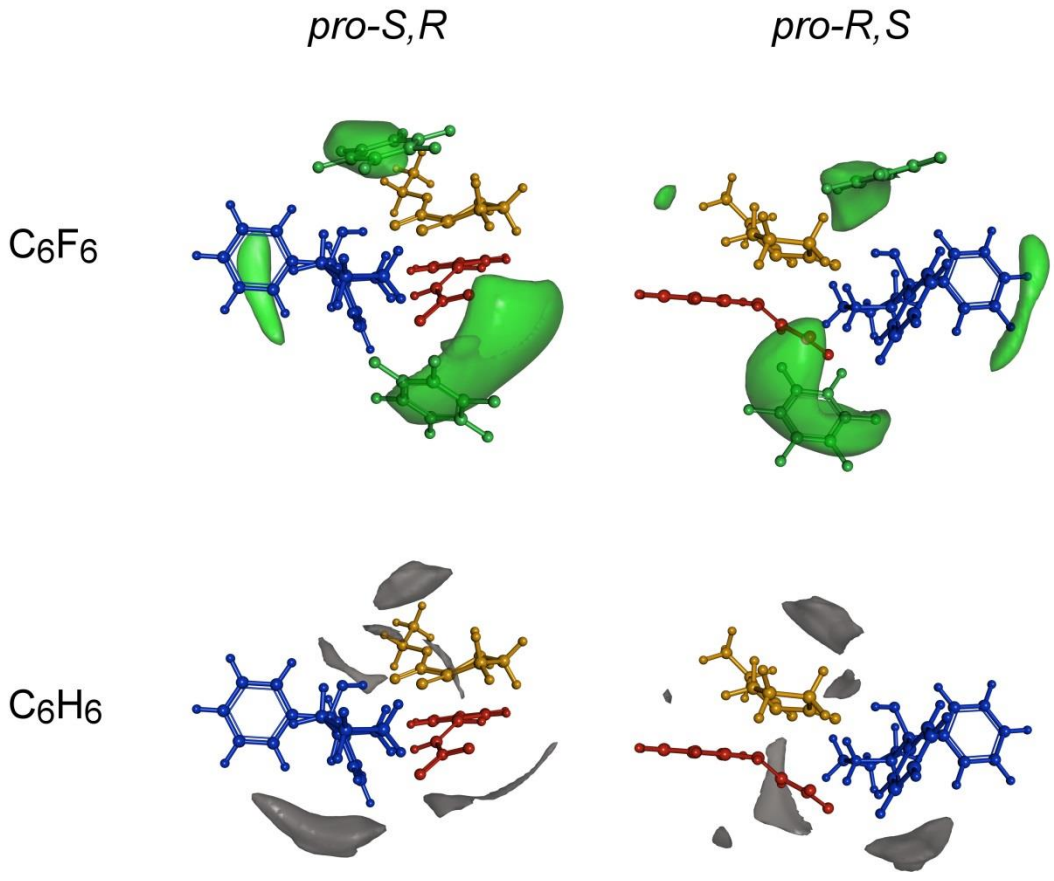


Figure 4.4 Carbon solvent distribution functions of C_6F_6 (top) and C_6H_6 (bottom) around the *pro-S,R* and *pro-R,S* TSs. The explicit C_6F_6 molecules in the illustrations on top were not included in the calculations and are superimposed from the original DFT setting of Lattanzi et al.[30] B3LYP/6-311+G(d,p) is chosen as level of theory with regular EC-RISM$^{\varphi}$(C_6H_6) and EC-RISM$^{\varphi}$ (C_6F_6) solvent models. Isosurface values are set to 2.0.

The er that is calculated by eq. (62) represents the theoretical upper limit for the stereoselectivity. Accordingly, due to a missing full description of kinetic effects, unknown experimental uncertainty and laboratory compound loss, the applied EC-RISM approach is suited to describe the solvent influence on the nitro-Michael reaction. This is corroborated by solvent density results illustrated in Figure 4.4. There the EC-RISM results correlate with the explicitly placed and minimized C_6F_6 molecules that were not included in our TSs and originate from the work of Lattanzi et al.[30] This is also the case for the crucial solvent density near the phenol group of the enolate that was mentioned to be very important for the stereoselective discrimination.[30] The solvent density of benzene exhibits a different spatial distribution around the TSs that is not in the region of the reactive center. This could as well be one important reason for the inferior influence on the er.

4.4.2 Impact of quantum-mechanics and force field approaches

As summed in table Table 4.2 the solute's force field has a considerable influence on the transition state energetics. Still the applied GAFF force field result possesses the best correspondence to the experimental values. The OPLS force field performs slightly worse, but has still a qualitative agreement to the experimental findings. Both physically motivated force field approaches predict the experimental stereoselectivity enhancement of hexafluorobenzene, showing that the EC-RISM approach is quite robust against small force field changes. The use of an unphysical force field like the 'all uniform σ & ε dispersion parameter' approach, reveals that an unreasonable choice for the solute-solvent dispersion yields energetics that drastically deviate from the desired results.

Table 4.2 Comparison between the calculated $\Delta\Delta G^{\ddagger}$ and er values, coupled with EC-RISM^q in combination with different force field approaches for the solute parametrization. Two common force fields (GAFF and OPLS) are applied. Further an approach, for which all solute-solvent dispersion parameters are set to the same σ (2.00 Å) and ε (0.3011 kJ mol^{-1}) values, is tested.

force field	GAFF	OPLS	equal σ & ε
$\Delta\Delta G^{\ddagger}(C_6H_6)$ / kJ mol^{-1}	3.62	4.06	42.97
$\Delta\Delta G^{\ddagger}(C_6F_6)$ / kJ mol^{-1}	5.65	5.32	9.93
$\Delta\Delta G^{\ddagger}_{trans}$ / kJ mol^{-1}	2.03	1.26	-33.04
er(C_6H_6)	81:19	84:16	100:0
er(C_6F_6)	91:9	90:10	98:2

Table 4.3 presents the results for different quantum-mechanical levels of theory. As before, the polarizable continuum methodology is not able to distinguish between C_6H_6 and C_6F_6. The $\Delta\Delta G^{\ddagger}$ values for the M06 density functional theory and MP2 perturbation treatment are much higher than the B3LYP/6-311+G(d,p) results. In consequence, because of the Boltzmann weighting (see eq. (62)), the ers are significantly higher than the experimental values. But the BP98/SVP 'transition states' structures are not maxima on the potential energy surface for the MP2/6-311+G(d,p)/EC-RISM and M06/6-311G(d,p)/EC-RISM levels of theory as demonstrated by the results in Table 4.4. This is probably due to the inconsistency between the optimized *TS* structures that were estimated with BP86/SVP and the level of theory applied to optimize the educt structures. In consequence the deduced relations for the reaction thermodynamics and the associated ers are not applicable for the MP2 and M06 results.

Table 4.3 Comparison between the calculated $\Delta\Delta G^{\ddagger}$ (in kJ mol^{-1}) and er values estimated with EC-RISM$^{\varphi}$ in combination with different quantum-chemical approaches for the solute wave function.

QM method(TS)	$\Delta\Delta G^{\ddagger}(C_6H_6)$ / kJ mol^{-1}	$\Delta\Delta G^{\ddagger}(C_6F_6)$ / kJ mol^{-1}	$\Delta\Delta G^{\ddagger}_{trans}$ / kJ mol^{-1}	er(C_6H_6)	er(C_6F_6)
B3LYP/6-311+G(d,p)// BP98/SVP	3.65	5.78	2.13	81:19	91:9
MP2/6-311+G(d,p)// BP98/SVP	16.21	18.55	2.34	100:0	100:0
M06/6-311G(d,p)// BP98/SVP	10.49	12.19	1.70	98:2	99:1

Table 4.4 Calculated $\Delta G^{\ddagger}$ and $\Delta G_{reaction}$ values (in kJ mol^{-1}) estimated with EC-RISM$^{\varphi}$ in combination with quantum-chemical approaches for the solute wave function corresponding to Table 4.3.

QM method(TS)	$\Delta G^{\ddagger}$(*pro-S,R*)		$\Delta G_{reaction}$	
Solvent	C_6H_6	C_6F_6	C_6H_6	C_6F_6
B3LYP/6-311+G(d,p)// BP98/SVP	75.39	80.51	-19.72	-19.74
MP2/6-311+G(d,p)// BP98/SVP	-113.82	-110.10	-145.29	-144.56
M06/6-311G(d,p)// BP98/SVP	-19.80	-15.86	-90.98	-88.98

4.4.3 Excess chemical potential and electronic energy

To segregate the influence that determines the enantiomeric ratio, the Gibbs free energies of the transition states are separated into

$$\Delta G^{\ddagger}(TS,v) = \Delta E^{\ddagger}(TS,v) + \Delta\mu^{\ddagger,\mathbf{ex}}(TS,v) \tag{72}$$

with $\Delta E^{\ddagger}$ as the solutes electronic energy and $\Delta\mu^{\ddagger,\mathrm{ex}}$ as excess chemical potential of the solution, both normalized to the corresponding energy of the educts. Thus the Gibbs free energy difference between the transition states is representable as

$$\begin{aligned}\Delta\Delta G^{\ddagger}(v) &= \Delta E^{\ddagger}(pro\text{-}R,S,v) - \Delta E^{\ddagger}(pro\text{-}R,S,v)\\ &+ \Delta\mu^{\ddagger,\mathbf{ex}}(pro\text{-}R,S,v) - \Delta\mu^{\ddagger,\mathbf{ex}}(pro\text{-}S,R,v)\\ &= \Delta\Delta E^{\ddagger}(v) + \Delta\Delta\mu^{\ddagger,\mathbf{ex}}(v),\end{aligned} \tag{73}$$

defining an electronic ($\Delta\Delta E^{\ddagger}$) and an excess chemical potential difference ($\Delta\Delta\mu^{\ddagger,\mathrm{ex}}$). The split-up is presented in Figure 4.5 and the corresponding data is summed in Tables S33-S38 of the appendix.

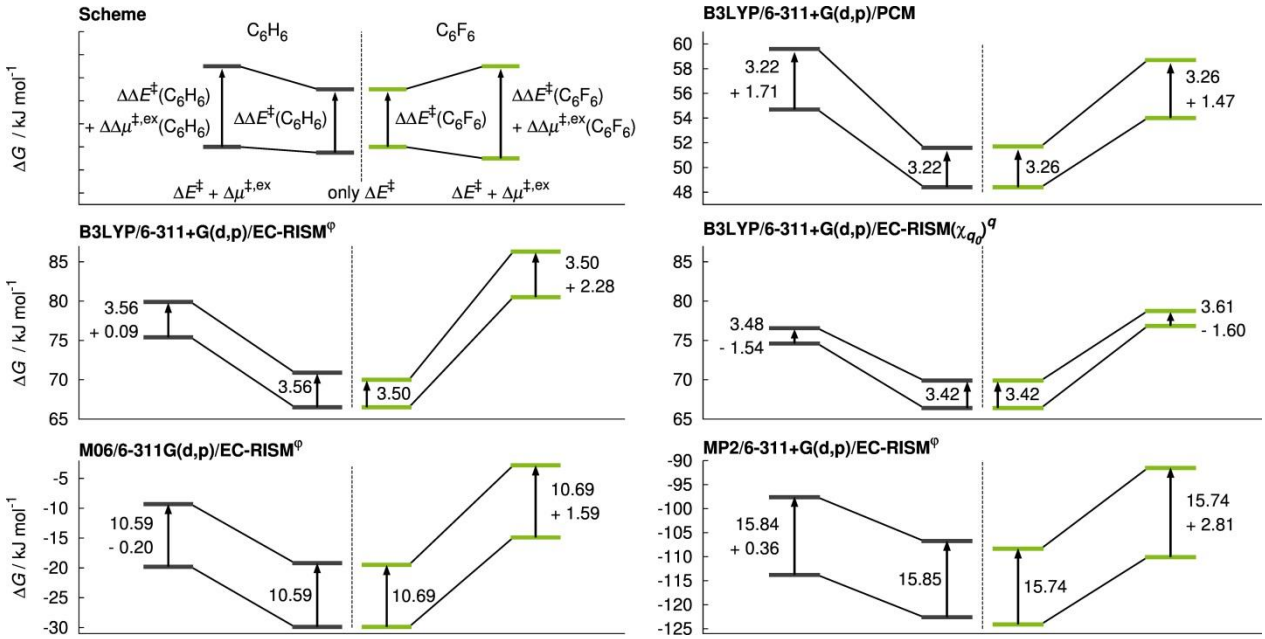


Figure 4.5 Separation of the Gibbs free energies of the transition states into the electronic energy and the excess chemical potential. The colouring is consistent with previously presented figures. The original transition state structures (BP86/SVP) are applied and post-processed with the depicted levels of theory. The Gibbs free energies are normalized to the Gibbs free energy of the educts.

Accordingly to Figure 4.5 the $\Delta\Delta E^{\ddagger}$ is dominated by the applied quantum-mechanical level of theory and nearly unimpeded by the solvent. The *R*,*S*-product is therefore always the preferred enantiomer as indicated by uniform positive values between 3.22 kJ mol^{-1} and 15.85 kJ mol^{-1}, independent of the applied solvation model. Addition of the excess chemical potential part consistently increases the Gibbs free energy of every solvation model. In the case of PCM the excess chemical potential is increased nearly for the same amount in C_6H_6 and C_6F_6, indicated by $\Delta\Delta\mu^{\ddagger,ex}$ values of 1.71 kJ mol^{-1} for C_6H_6 and 1.47 kJ mol^{-1} for C_6F_6. In consequence PCM does not distinguish between both solvents and the $\Delta\Delta G^{\ddagger}$ values are almost equal as shown previously. Otherwise EC-RISM predicts $\Delta\Delta\mu^{\ddagger,ex}$ values of about -0.20 - 0.34 kJ mol^{-1} for the regular C_6H_6 model, accordingly an equal excess chemical potential increase of both TSs in benzene. Contrary the corresponding $\Delta\Delta\mu^{\ddagger,ex}$ values in C_6F_6 are, with 1.58 - 2.81 kJ mol^{-1}, computed to be significantly higher. Therefore the improvement of stereoselectivity caused by C_6F_6 ensues from the solvent specific difference of the $\Delta\Delta\mu^{\ddagger,ex}$ results between both solvents. The results with the q_0-models exhibit values for $\Delta\Delta E^{\ddagger}$ that are comparable with the results for B3LYP/6-311+G(d,p)/EC-RISM$^{\varphi}$. The addition of $\Delta\Delta\mu^{\ddagger,ex}$ for the purely dispersive solvent models narrows the initial $\Delta\Delta E^{\ddagger}$ gap by -1.54 kJ mol^{-1} in C_6H_6 and by -1.60 kJ mol^{-1} in C_6F_6. Indeed, as for PCM, the electronic energies of the transition states are shifted by the same amount in both solvents and the purely dispersive solvent model is not capable to reproduce the experimental stereoselectivity enhancement. This shows that the electrostatic solvent-solute interaction is important, especially for the excess chemical potential part, because it moderately modifies the electronic energy gap in C_6H_6, but substantially increases it in C_6F_6.

4.4.4 Consequences of consistent transition state structures

Starting from the original BP98/SVP structures a sophisticated transition state search was performed for the B3LYP/6-311+G(d,p) level of theory,. Afterwards new Gibbs free energies are computed with the IEF-PCM and EC-RISM$^{\varphi}$ approach using B3LYP/6-311+G(d,p) quantum-mechanics. The results are expressed in Table 4.5 and Figure 4.6. Surprisingly the $\Delta\Delta G^{\ddagger}$ results and consequently the enantiomeric ratios are nearly unaltered. The new EC-RISM$^{\varphi}$ results still distinguish between both solvents expressed by the negative $\Delta\Delta G^{\ddagger}_{trans}$ and er values. Thus at a first glance it seems that the TSs search has no impact. Indeed, changes of about 26 kJ mol^{-1} for the reaction barriers are consistently achieved for the EC-RISM$^{\varphi}$ and IEF-PCM solvation models. But, this drastic decrease has no impact on the enantiomeric

ratios obtained by the applied two-state model that neglects the reaction kinetics. The cause of the unchanged ers is a nearly complete cancellation of the changes in the reaction barriers for $\Delta\Delta G^{\ddagger}$ values that determine the predicted ers.

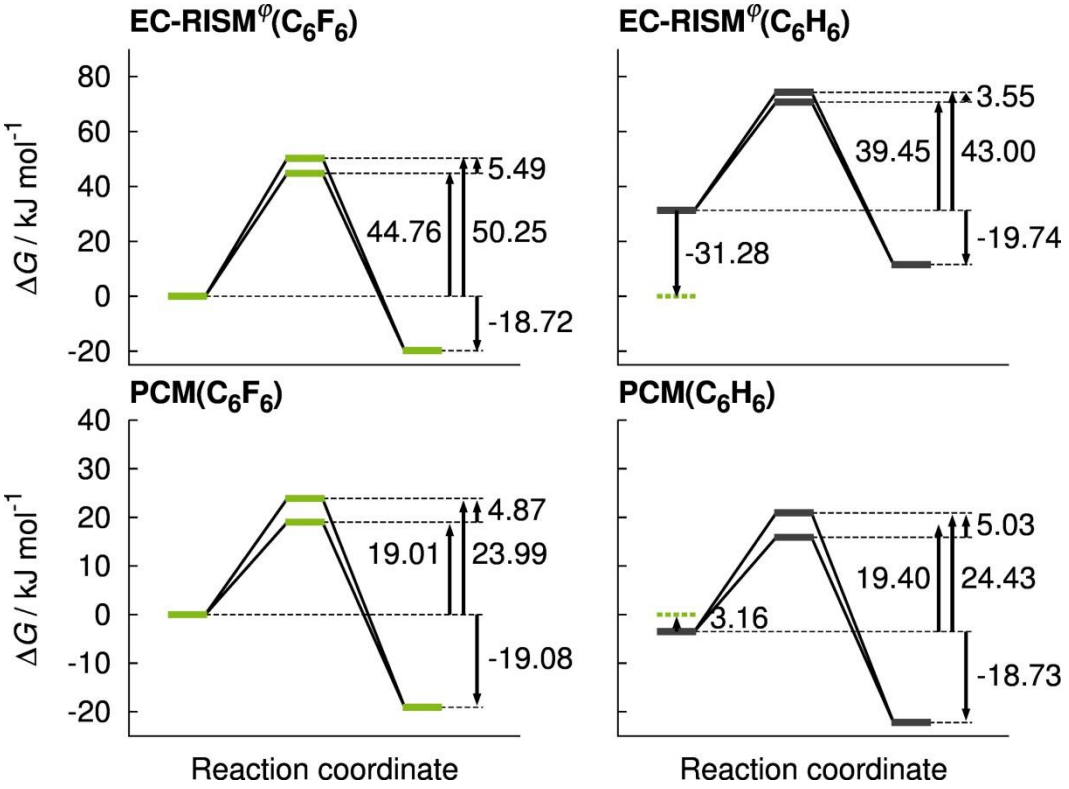


Figure 4.6 Reaction profile plot in the same style as Figure 4.3 with results for IEF-PCM and EC-RISM$^{\varphi}$ solvation for a set of structures that are consistently optimized with B3LYP/6-311+G(d,p).

Table 4.5 Comparison between $\Delta\Delta G^{\ddagger}$ and er values estimated on one hand with the original TSs of Lattanzi et al.[30] (BP98/SVP) and on the other hand with transition state structures corresponding to B3LYP/6-311+G(d,p)/EC-RISM$^{\varphi}$//B3LYP/6-311+G(d,p). All Gibbs free energies are given in kJ mol^{-1}.

	EC-RISM$^{\varphi}$		IEF-PCM	
	B3LYP/ 6-311+G(d,p)// BP98/SVP	B3LYP/ 6-311+G(d,p)// B3LYP/ 6-311+G(d,p)	B3LYP/ 6-311+G(d,p)// BP98/SVP	B3LYP/ 6-311+G(d,p)// B3LYP/ 6-311+G(d,p)
$\Delta G^{\ddagger}$(*pro-R,S*,C_6H_6)	79.04	43.00	59.55	24.43
$\Delta G^{\ddagger}$(*pro-S,R*,C_6H_6)	75.39	39.45	54.62	19.40
$\Delta G^{\ddagger}$(*pro-R,S*,C_6F_6)	86.29	50.25	58.75	23.88
$\Delta G^{\ddagger}$(*pro-S,R*,C_6F_6)	80.51	44.76	54.00	19.01
$\Delta\Delta G^{\ddagger}$($C_6H_6$)	3.65	3.55	4.93	5.03
$\Delta\Delta G^{\ddagger}$($C_6F_6$)	5.78	5.49	4.75	4.87
$\Delta\Delta G^{\ddagger}_{trans}$	2.13	1.94	-0.18	-0.16
er(C_6H_6)	81:19	81:19	88:12	88:12
er(C_6F_6)	91:9	90:10	87:13	88:12

4.4.5 Reaction kinetics

As shown earlier the applied transition state structures have no impact on the enantiomeric ratios for the two-state models without reverse reaction. Here, we will point out that the change in energetics, due to the chosen TSs, indeed has effects on the compound concentrations over time. The resulting concentration profiles are illustrated in Figure 4.7. The corresponding coefficients are summed in Tables S41 and S42 of the appendix.

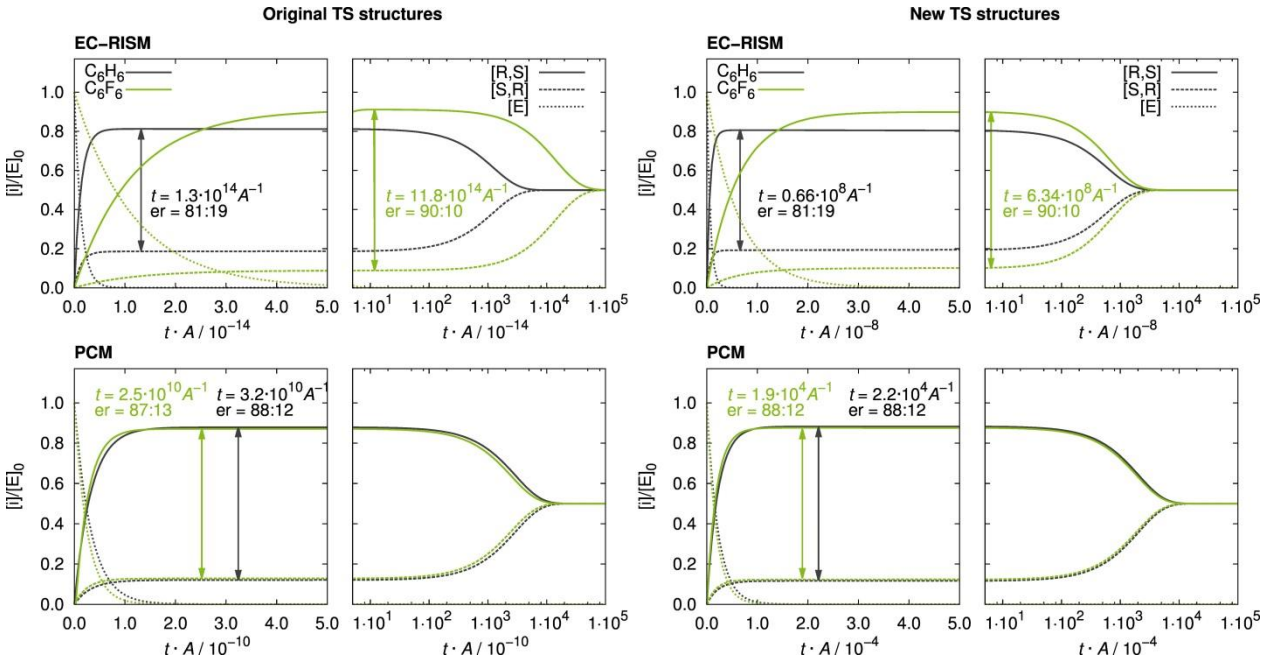


Figure 4.7 Concentration profiles normalized to the educts starting concentration $[E]_0$ in dimensions of the pre-exponential factor *A*. Illustrated are the results of the kinetic model based on the solution of equation (62) with EC-RISM$^{\varphi}$ (top) and PCM (bottom) solvation in combination with the original BP988/SVP transition state geometries (left) and the newly optimized (B3LYP/6-311+G(d,p)) TS structures (right). Due to the wide range of reaction constants, the plotted time scales (*x*-axes) vary to reproduce clearly reasonable illustrations.

The EC-RISM model predicts that the er maximum is achieved approximately 9-10 times faster in benzene than in hexafluorobenzene, although the er is higher in the latter, whereas PCM predicts a reaction that is assumed to be 70% slower in benzene. The highest possible enantiomeric ratio is congruent to the previously established results. By this means the reaction kinetics have no influence on the achievable maximum er, but the newly optimized transition state structures substantially influence the reaction times. The relative Gibbs free energies of the new TSs geometries are approximately 25 kJ mol^{-1} lower than the free energies of the BP98/SVP structures, which in consequence shortens the reaction times by 5-6 orders of magnitude. Nevertheless all concentration profiles exhibit relatively long periods of constantly high ers that are in general desired in experiments.

4.5 Concluding remarks

In summary a starting point for further EC-RISM supported investigations of complex solvent mediated organic reactions is provided here. It is demonstrated that the EC-RISM model of C_6F_6 reproduces calculations with explicit solvent molecules and furthermore allows for studies of surrounding influences by the application of hypothetical models like the applied uncharged q_0-C_6H_6 and q_0-C_6F_6 solvents. To this end the favored enantiomer is achieved with every applied approach, but the intricate solvent discrimination is a consequence of the solvent model and cannot be reproduced by PCM. Furthermore the dispersive solvent models are not capable to predict the experimentally observed stereoselectivity increase by hexafluorobenzene compared to benzene, emphasizing the importance of quadrupolar solvent-solute electrostatic interactions. The applied chemical kinetics approach is fairly simple, but has the advantage to be analytically solvable with the capability to estimate relative reaction durations. Absolute reaction times are not explicitly calculated due to lacking pre-exponential factors *A*. These could be obtained by calculations with a set of temperatures or by transition state theory[189,190] for which a high amount of additional computational effort is needed for vibrational frequency analysis. Furthermore a wider range of quantum-chemical levels of theory should be applied to corroborate the presented findings that were achieved by only one of the three used quantum-mechanical approaches.

Nevertheless EC-RISM proves to be a feasible and useful solvent model for each reaction step, including the products, transition states and educts. EC-RISM is easy to apply, works without the need of complex solvent particle positioning sampling that is essential for example in cluster approaches. Furthermore compared with empirical models for yield predictions like PC-SAFT.[191,192] Altogether here a solvent specific approach that gives the observer opportunity for insight into the solvation influence on each step of the reaction is established.

5 SUMMARY AND CONCLUSION

In summary new applications of EC-RISM theory were developed, implemented and evaluated in order to extend the scope and to assess the reliability of the solvation model. On one hand a framework for nuclear magnetic resonance chemical shifts prediction was established; on the other hand an approach for reaction thermodynamics and kinetics was realized and tested.

The wave function of infinitely diluted solutions of molecules obtained by EC-RISM was tested and examined by chemical shift evaluations under the influence of various flavors of the solvation methodology. In consequence a physically plausible polarization model of the solute's electronic structure in aqueous environment that is consistent with literature results was obtained. In comparison with chemical shifts of common implicit solvation models, the EC-RISM solvent polarizes the dissolved molecule more strongly, which results in higher unshielded nuclei and better experimental agreement for small basis sets in general and also for larger basis sets in the specific case, when solute-solvent interactions are treated exactly, without a partial charge approximation. Compared to the simpler point charge procedure, exact electrostatics incorporate crucial multipolar electrostatic effects on the solvent distribution.

These findings are in good agreement with the Gibbs free energy studies of infinitely diluted solutions of small molecules in benzene and hexafluorobenzene for which the more accurate exact electrostatic description systematically improves the results. Additionally, as for the results in aqueous solution, the obtained solvent distributions with exact electrostatics are significantly different from the solvation structures calculated by a point charge representation. The application of a solvent model with reversed point charges supports this hypothesis by worsening the agreement with experimental findings. Furthermore the use of the uncharged solvent models quantifies the fact that a general discrimination of benzene and

hexafluorobenzene is already achieved by dispersion interactions. But, as shown for the transition states of the investigated Michael addition, pure dispersion is insufficient to explain the differential solvation effect of hexafluorobenzene, and electrostatic, quadrupolar solvent-solute interactions are essential to describe the stereoselectivity enhancement shown experimentally. This observation and rationalization for various systems represents a major conclusion of this thesis.

In future work the applied decomposition of purely dispersive and electrostatic interactions should be analyzed for other fluorinated solvents and additives in order to scrutinize whether or not electrostatics interaction are in general crucial for the fluorine effect. It is therefore reasonable to start with solvents that are similar to hexafluorobenzene such as other polyfluorinated benzene compounds. Alternatively, investigations of different perhalogenated benzenes should be performed in order to identify unique properties of fluorinated solvents. Furthermore it is well known that fluorine-containing solvents cause notable changes in the chemical shifts of many solutes.[193] An investigation of these solvation effects would require a combination of the EC-RISM chemical shift estimation technique with the nondipolar solvent models developed in this work.

6 REFERENCES

[1] R. E. Skyner, J. L. McDonagh, C. R. Groom, T. van Mourik, J. B. O. Mitchell, *Phys. Chem. Chem. Phys.* **2015**, *17*, 6174.

[2] R. Gani, C. Jiménez-González, D. J. C. Constable, *Computers and Chemical Engineering* **2005**, *29*, 1661.

[3] T. Zhou, Z. Lyu, Z. Qi, K. Sundmacher, *Chemical Engineering Science* **2015**, *137*, 613.

[4] D. Frenkel, B. Smit, *Understanding Molecular Simulations: From Algorithms to Applications*, Academic Press, London, **2001**.

[5] R. Car, M. Parinello, *Phys. Rev. Lett.* **1985**, *55*, 2471.

[6] D. Marx, J. Hutter, *Ab Initio Molecular Dynamics*, Cambridge University Press, New York, **2009**.

[7] R. W. Zwanzig, *J. Chem. Phys.* **1954**, *22*, 1420.

[8] D. Wu, D. A. Kofke, *J. Chem. Phys.* **2005**, *123*, 054103.

[9] G. M. Torrie, J. P. Valleau, *J. Comput. Phys.* **1977**, *23*, 187.

[10] C. J. Cramer, D. Truhlar, *Chem. Rev*. **1999**, *99*, 2161.

[11] B. Honig, A. Nicholls, *Science* **1995**, *268*, 1144.

[12] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc*. **1990**, *112*, 6127.

[13] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105*, 2999.

[14] D. A. McQuarrie, *Statistical mechanics*, HarperCollins, New York, **1976**.

[15] L. S. Ornstein, F. Zernike, *Proc. Acad. Sci. Amsterdam* **1914**, *17*, 793.

[16] C. G. Gray, C. G. Gubbins, *Theory of Molecular Fluids*, Oxford University Press, Belfast, **1984**.

[17] D. Chandler, H. C. Andersen, *J. Chem. Phys.* **1972**, *57*, 1918.

[18] D. Chandler, H. C. Andersen, *J. Chem. Phys.* **1972**, *57*, 1930.

[19] K. S. Schweizer, J. G. Curro, *Adv. Polym. Sci.* **1994**, *116*, 321.

[20] D. Beglov, B. Roux, *J. Chem. Phys. B* **1997**, *101*, 7821.

[21] A. Kovalenko, F. Hirata, *Chem. Phys. Lett.* **1998**, *290*, 237.

[22] A. Kovalenko, F. Hirata, *J. Mol. Liq.* **2001**, *90*, 215.

[23] T. Kloss, J. Heil, S. M. Kast, *J. Phys. Chem. B*, **2008**, *112*, 4337.

[24] S. M. Kast, J. Heil, S. Güssregen, K. F. Schmidt, *J. Comput. Aided Mol. Des.* **2010**, *24*, 343.

[25] R. F. Ribeiro, A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Chem. Theory Comput.* **2009**, *5*, 2284.

[26] R. Ditchfield, *Mol. Phys.* **1972**, *27*, 789.

[27] J. Heil, D. Tomazic, S. Egbers, S. M. Kast, *J. Mol. Model.* **2014**, *20*, 2161.

[28] S.-H. Chong, S.-i. Miura, G. Basu, F. Hirata, *J. Phys. Chem.* **1995**, *99*, 10526.

[29] M. Malvadi, S. Bruzzone, C. Chiappe, S. Gusarov, A. Kovalenko, *J. Phys. Chem. B* **2009**, *113*, 3536.

[30] A. Lattanzi, C. De Fusco, A. Rosso, A. Poater, L. Cavallo, *Chem. Commun.* **2012**, *48*, 1650.

[31] K. Naka, H. Sato, A. Morita, F. Hirata, S. Kato, *Theor. Chem. Acc.* **1999**, *102*, 165.

[32] K. Ida, D. Yokogawa, A. Ikeda, H. Sato, S. Sakaki, *Phys. Chem. Chem. Phys.* **2009**, *11*, 8556.

[33] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864.

[34] T. Morita, K. Hiroike, *Prog. theort. Phys.* **1961**, *25*, 537.

[35] C. De Dominicis, *J. Math. Phys.* **1962**, *3*, 983.

[36] J.-P. Hansen, I. R. McDonald, *Theory of Simple Liquids*, Academic Press, London, **2006**.

[37] J. G. Kirkwood, *J. Chem. Phys.* **1935**, *3*, 300.

[38] J. K. Percus, *Phys. Rev. Lett.* **1962**, *8*, 462.

[39] R. Frach, *Quantenchemische Modelle für nichtwässrige Dienaminlösungen*, master thesis, TU Dortmund, **2011**.

[40] E. Meeron, *J. Chem. Phys.* **1958**, *28*, 630.

[41] K. Hiroike, *J. Phys. Soc. Japan* **1958**, *13*, 1497.

[42] K. Hiroike, *Prog. Theor. Phys.* **1960**, *23*, 1003.

[43] F. Zernike, J. Prins, *Zeits. f. Physik* **1927**, *41*, 184.

[44] L. Goldstein, *Phys. Rev.* **1951**, *83*, 289-298.

[45] M. J. Klein, L. Tisza, *Phys. Rev.* **1949**, *76*, 1861.

[46] J. L. Lebowitz, *Phys. Rev.* **1964**, *133*, A895.

[47] L. Blum, F. Vericat, J. N. Herrera-Pacheco, *J. Stat. Phys.* **1992**, *66*, 249.

[48] L. Blum, A. J. Torruella, *J. Chem. Phys.* **1972**, *56*, 303.

[49] D. Chandler, J. D. McCoy, S. J. Singer, *J. Chem. Phys.* **1986**, *85*, 5977.

[50] C. M. Cortis, P. J. Rossky, R. A. Friesner, *J. Chem. Phys.* **1997**, *107*, 6400.

[51] F. Hirata, P. J. Rossky, *Chem. Phys. Lett.* **1981**, *83*, 329.

[52] J. S. Perkyns, B. M. Pettitt, *Chem. Phys. Lett.* **1992**, *190*, 626.

[53] J. S. Perkyns, B. M. Pettitt, *J. Chem. Phys.* **1992**, *97*, 7656.

[54] M. Ikeguchi, J. Doi, *J. Chem. Phys.* **1995**, *103*, 5011.

[55] D. Beglov, B. Roux, *J. Chem. Phys.* **1995**, *103*, 360.

[56] D. Beglov, B. Roux, *J. Chem. Phys.* **1996**, *104*, 8678.

[57] A. Kovalenko, F. Hirata, *J. Chem. Phys.* **2000**, *112*, 10391.

[58] S. M. Kast, T. Kloss, *J. Chem. Phys.* **2008**, *129*, 236101.

[59] P. Ewald, *Ann. Phys.* **1921**, *369*, 253.

[60] L. Lue, D. Blankschtein, K. *J. Phys. Chem.* **1992**, *96*, 8582.

[61] S. M. Kast, *Phys. Rev. E* **2003**, *67*, 041203.

[62] L. Lue, D. Blankschtein, *J. Phys. Chem.* **1992**, *96*, 8582.

[63] T. Miyata, F. Hirata, *J. Comput. Chem.* **2008**, *29*, 871.

[64] T. Luchko, S. Gusarov, D. R. Roe, C. Simmerling, D. A. Case, J. Tuszynski, A. Kovalenko, *J. Chem. Theory Comput.* **2010**, *6*, 607.

[65] J. W. Ponder, C. Wu, R. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, L. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio Jr. , M. Head-Gordon, G. N. L. Clark, M. E. Johnson, T. Head-Gordon, *J. Phys. Chem. B* **2010**, *114*, 3143.

[66] F. Hoffgaard, J. Heil, S. M. Kast, *J. Chem. Theory Comput.* **2013**, *9*, 4718.

[67] S. Ten-no, F. Hirata, S. Kato, *Chem. Phys. Lett.* **1993**, *214*, 391.

[68] H. Sato, F. Hirata, S. Kato *J. Chem. Phys.* **1996**, *105*, 1546.

[69] A. Kovalenko, F. Hirata, *J. Chem. Phys.* **1999**, *110*, 10095.

[70] H. Sato, A. Kovalenko, F. Hirata, *J. Chem. Phys.* **2000**, *112*, 9463.

[71] N. Yoshida, S. Kato, *J. Chem. Phys.* **2000**, *113*, 4974.

[72] D. Yokogawa, H. Sato, T. Imai, S. Sakaki, *J. Chem. Phys.* **2009**, *130*, 064111.

[73] D. Yokogawa, H. Sato, S. Sakaki, *J. Mol. Liq.* **2009**, *147*, 112.

[74] K. Kido, K. Kasahara, D. Yokogawa, H. Sato *J. Chem. Phys.* **2015**, *143*, 014103.

[75] J. Sauer, *Chem Rev.* **1989**, *89*, 199.

[76] S. Hauptmann, H. Dufner, J. Brickmann, S. M. Kast, R. S. Berry, *Phys. Chem. Chem. Phys.* **2003**, *5*, 635.

[77] A. Ben Naim, *J. Phys. Chem.* **1978**, *82*, 792.

[78] A. Ben-Naim, Y. Marcus, *Chem. Phys.* **1984**, *81*, 2016.

[79] S.-H. Chong, S. Ham, *J. Chem. Theory Comput.* **2015**, *11*, 378.

[80] M. Schindler, W. Kutzelnigg, *J. Chem. Phys.* **1982**, *76*, 1919.

[81] J. C. Facelli, *J. Phys. Chem. B* **1998**, *102*, 2111.

[82] S. G. Smith,J. M. Goodman, *J. Am. Chem. Soc.* **2010**, *132*, 12946.

[83] C. V. Sumowski, M. Hanni, S. Schweizer, C. Ochsenfeld, *J. Chem. Theory Comput.* **2014**, *10*, 122.

[84] M. Pérez, T, M. Peakman, A. Alex, P. D. Higginson, J. C. Mitchell, M. J. Snowden, I. Marao, *J. Org. Chem.* **2006**, *71*, 3103.

[85] M. W. Lodewyk, M. R. Siebert, D. J. Tantillo, *Chem. Rev.* **2012**, *112*, 1839.

[86] F. A. A. Mulder, M. Filatov, *Chem. Soc. Rev.* **2010**, *39*, 578.

[87] M. Dračinsky, H. M. Möller, T. E. Exner, *J. Chem. Theory Comput.* **2013**, *9*, 3806.

[88] M. Dračinsky, P. Bouř, *J. Chem. Theory Comput.* **2010**, *6*, 288.

[89] D. A. Zichi, P. J. Rossky, *J. Chem. Phys.* **1986**, *84*, 1712.

[90] R. Cammi, B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1999**, *110*, 7627.

[91] S. Nozinovic, P. Gupta, B. Fürtig, C. Richter, S. Tüllmann, E. Duchardt-Ferner, M. C. Holthausen, H. Schwalbe, *Angew. Chem. Int. Ed.* **2011**, *50*, 5397.

[92] M. N. Manalo, A. C. de Dios, *Magn. Reson. Chem.* **2002**, *40*, 781.

[93] Q. Cui, M. Karplus, *J. Phys. Chem. B* **2000**, *104*, 3721.

[94] T. E. Exner, P. G. Mezey, *J. Comput. Chem.* **2003**, *24*, 1980.

[95] A. Frank, L. Onila, H. M. Möller, T. E. Exner, *Proteins* **2011**, *79*, 2189.

[96] T. Zhu, X. He, J. Z. H. Zhang, *Phys. Chem. Chem. Phys.* **2012**, *14*, 7837.

[97] D. Beglov, B. Roux, *J. Chem. Phys.* **1998**, *104*, 8678.

[98] A. Kovalenko, F. Hirata, *J. Chem. Phys.* **1998**, *112*, 4337.

[99] A. A. Auer, *J. Chem. Phys.* **2009**, *131*, 024116.

[100] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098.

[101] C. Lee, W. Yang, G. Parr, *Phys. Rev. B* **1988**, *37*, 785.

[102] S. H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*, 1200.

[103] R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 724.

[104] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, et al., *Gaussian 03,* Gaussian, Inc., Wallingford CT, **2004**.

[105] C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **1990**, *11*, 361.

[106] T. H. Dunning Jr., *J. Chem. Phys.* **1989**, *90*, 1007.

[107] R. A. Kendall, T. H. Dunning Jr., R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796.

[108] R. E. Hoffman, *Magn. Reson. Chem.* **2006**, *44*, 606.

[109] W. M. Litchman, A. Mohammed Jr., A. E. Florin, *J. Am. Chem. Soc.* **1969**, *91*, 6574.

[110] V. Thanabal, D. O. Omencinsky, M. D. Reily, W. L. Cody, *J. Biomol. NMR* **1994**, *4*, 47.

[111] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157.

[112] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. Model.* **2006**, *25*, 247260.

[113] Z. A. Makrodimitri, V. E. Raptis, I. G. Economou, *J. Phys. Chem. B* **2006**, *110*, 16047.

[114] S. Maw, H. Sato, S. Ten-no, F. Hirata, *Chem. Phys. Lett.* **1997**, *276*, 20.

[115] T. E. Exner, A. Frank, I. Onila, H. M. Möller, *J. Chem. Theory Comput.* **2012**, *8*, 4818.

[116] R. Bader, *J. Phys. Chem. B* **2009**, *113*, 347.

[117] T. Helgaker, M. Jaszuński, K. Ruud, *Chem. Rev.* **1999**, *99*, 293.

[118] Q. Du, D. Wei, *J. Phys. Chem. B* **2003**, *107*, 13463.

[119] M. P. Gaigeot, R. Vuilleumier, M. Sprik D. Borgis, *J. Chem. Theory Comput.* **2005**, *1*, 772.

[120] D. Chandler, *Nature* **2005,** *437*, 640.

[121] D. V. Matyushov, R. Schmid, *J. Chem. Phys.* **1996**, *105*, 4729.

[122] D. Chandler, J. D. Weeks, H. C. Andersen, *Science* **1983**, *220*, 787.

[123] G. R. Freeman, N. Gee, *J. Chem. Thermodynamics* **1993**, *25*, 549.

[124] D. Ruivo, An. B. Pereiro, J. M. S. S. Esperança, J. N. Canongia Lopes, L. P. N. Rebelo, *J. Phys. Chem. B* **2010**, *114*, 12589.

[125] S. Lorenzo, G. R. Lewsi, I. Dance, *New J. Chem.* **2000**, *24*, 295.

[126] R. Laatikainen, J. Ratilainen, R. Sebastian, H. Santa, *J. Am. Chem. Soc.* **1995**, *117*, 11006.

[127] G. Milano, F. Müller-Plathe, *J. Phys. Chem. B* **2004**, *108*, 7415.

[128] D. B. Ninković, J. M. Andrić, S. N. Malkov, S. D. Zarić, *Phys. Chem. Chem. Phys.* **2014**, *16*, 11173.

[129] J. Vrbancich, G. L. D. Ritchie, *J. Chem Soc. Faraday II* **1980**, *76*, 648.

[130] E. Wilhelm, R. Battino, *J. Chem. Thermydynamics* **1971**, *3*, 761.

[131] C. Samojłowicz, M. Bieniek, A. Pazio, A. Makal, K. Woźniak, A. Poater, L. Cavallo, J. Wójcik, K. Zdanomowski, K. Grela, *Chem. Eur. J.* **2011**, *17*, 12981.

[132] A. Klamt, G. Schüürmann, *J. Chem. Soc. Perkin Trans.* **1993**, *2*, 799.

[133] J. Jeon, H. Kim, *J. Solution Chem.* **2001**, *30*, 849.

[134] J. Jeon, H. Kim, *J. Chem. Phys.* **2008**, *129*, 174505.

[135] D. Kerlé, R. Ludwig, A. Geiger, D. Paschek, *J. Phys. Chem. B* **2009**, *113*, 12727.

[136] V. S. Bryantsev, M. S. Diallo, W. A. Goddard III, *J. Phys. Chem. B* **2008**, *112*, 9709.

[137] L. J. Lowden, D.Chandler, *J. Chem. Phys.* **1974**, *61*, 5228.

[138] O. Steinhauser, I. Hausleithner, H. Bertagnolli, *Chem. Phys.* **1987**, *111*, 371.

[139] H. A. Carlson, T. B. Nguyen, M. Orozco, W. L. Jorgensen, *J. Comp. Chem.* **1993**, *14*, 1240.

[140] W. L. Jorgensen, J. P. Ulmschneider, J. Tirado-Rives, *J. Chem. Phys. B* **2004**, *108*, 16264.

[141] W. D. Cornell, P. Cieplak, C. I. Bayly, I. Gould, K. M. Merz Jr., D. M. Ferguson, D. C. Spellmeyer, T. J. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179.

[142] J. H. Dymaond, J. Robertson, J. Isdale, *Int. J. Thermodyn.* **1981**, *2*, 223.

[143] S. M. Kast, *Phys. Chem. Chem. Phys.* **2001**, *3*, 5087.

[144] L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, *J. Comput. Chem.* **2009**, *30*, 2157.

[145] D. van der Spoel, F. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comp. Chem.* **2005**, *26*, 1701.

[146] B. Hess, C. Kutzner, D. van der Spoel, F. Lindahl, *J. Chem. Theory Comp.* **2008**, *4*, 435.

[147] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089.

[148] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577.

[149] S. Nosé, *Mol. Phys.* **1984**, *52*, 255.

[150] W. G. Hoover, *Phys. Rev. A* **1985**, *31*, 1695.

[151] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182.

[152] S. Nosé, M. L. Klein, *Mol. Phys.* **1983**, *50*, 1055.

[153] C. C. J. Roothan, *Rev. Mod. Phys.* **1951**, *23*, 69.

[154] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 0618.

[155] Head-Gordon, J. A. Pople, M. J. Frisch, *Chem. Phys. Lett.* **1988**, *153*, 503.

[156] C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158.

[157] Z. A. Makrodimitri, R. Dohrn, I. G. Economou, *Macromolecules* **2007**, *40*, 1720.

[158] A. Martín-Calvo, D. Lahoz-Martín, S. Calero, *J. Phys. Chem. C* **2011**, *116*, 6655.

[159] A. J. Huth, J. M. Stueve, V. V. Guliants, *Membrane Sci*. **2012**, *403*, 236.

[160] Q. Cui, V. H. Smith Jr., *J. Chem. Phys.* **2000**, *113*, 10240.

[161] E. K. Watkins, W. L. Jorgensen, *J. Phys. Chem. A* **2001**, *105*, 4118.

[162] C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **1990**, *11*, 361.

[163] I. S. Joung, T. Luchko, D. A. Case, *J. Chem. Phys.* **2013**, *138*, 044103.

[164] A. Kovalenko, F. Hirata, *J. Chem. Phys.* **2000**, *113*, 2793.

[165] V. I. Lebedev, D. N. Laikov, *Dokl. Math.* **1999**, *59*, 477.

[166] A. K. Soper, *Chem. Phys.* **1996**, *202*, 295.

[167] A. K. Soper, *Phys. Rev. B* **2005**, *72*, 104204.

[168] T. F. Headen, C. A. Howard, N. T. Skipper, M. A. Wilkinson, D. T. Bowron, A. K. Soper, *J. Am. Chem. Soc.* **2010**, *132*, 5735.

[169] N. Mohan, C. H. Suresh, A. Kumar, S. R. Gadre, *Phys. Chem. Chem. Phys.* **2013**, *15*, 18401.

[170] H. Reiss, H. L. Frisch, J. L. Lebowitz, *J. Chem. Phys.* **1959**, *31*, 369.

[171] H. Reiss, H. L. Frisch, J. L. Lebowitz, *J. Chem. Phys.* **1960**, *32*, 119.

[172] P. Auffinger, F. A. Hays, E. Westhof, P. S. Ho, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16789.

[173] T. Sugiishi, M. Matsugi, H. Hamamoto, H. Arnii, *RSC Adv.* **2015**, *5*, 17269.

[174] K. Reichenbächer, H. I. Süss, J. Hullinger, *Chem. Soc. Rev.* **2004**, *34*, 22.

[175] K. Müller, C. Faeh, F. Diederich, *Science* **2007**, *317*, 1881.

[176] D. 'Hagan, *Chem. Soc. Rev.* **2008**, *37*, 308.

[177] T. Böttcher, D. Duda, N. Kalinovich, O. Kazakova, M. Ponomarenko, K. Vlasov, M. Winter, G.-V. Rpscenthaler, *Prog. in Solid State Chem.* **2014**, *42*, 202.

[178] P. Perdew, *Phys. Rev. B* **1986**, *33*, 8822.

[179] A. Schaefer, H. Horn, R. Ahlrichs, *J. Chem. Phys.* **1992**, *97*, 2571.

[180] G. Cainelly, P. Galetti, D. Giacomini, *Chem. Soc. Rev.* **2009**, *38*, 990.

[181] Wolfram Research, Inc., Mathematica, Version 10.2, Champaign, IL, **2015**.

[182] Gaussian 09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson et al. and D. J. Fox, Gaussian, Inc., Wallingford CT, **2009**.

[183] F. Neese, *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2012**, *2*, 73.

[184] Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, **2008**, *120*, 215.

[185] S, Grimme, J. Antony, S. Ehrlich, S. Krieg, *J. Chem. Phys.* **2010**, *132*, 152104.

[186] A. Kovalenko, S. Ten-no, F. Hirata, *J. Comput. Chem.* **1999**, *20*, 928.

[187] J. Wang; R. M. Wolf; J. W. Caldwell; P. A. Kollman; D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157.

[188] W. L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.* **1988**, *110*, 1657.

[189] K. J. Laidler, M. C. King, *J. Phys. Chem.* **1983**, *87*, 2557.

[190] D. G. Truhlar, B. C. Garrett, S. J. Klippenstein, *J. Phys. Chem.* **1996**, *100*, 12771.

[191] J. Gross, G. Sadowski, *Ind. Eng. Chem. Res.* **2001**, *40*, 1244.

[192] F. S. Emami, A. Vahid, E. K. Szymkuć, P. Dittwald, K. Molga, B. A. Grybowski, *Angew. Chem. Int. Ed.* **2015**, *127*, 10947.

[193] M. R. Bacon, G. E. Maciel, *J. Am. Chem. Soc.* **1973**, *95*, 2413.

A1 APPENDIX

A1.1 Solvation effects on chemical shifts by embedded cluster integral equation theory

Table S1. Calculated isotropic chemical shielding constants (GIAO/B3LYP/6-31+G(d)) for optimized NMA structures from all models studied. For all EC-RISM calculations, the optimal PCM structure was used.

	σ / ppm						
	C(1)	H(1)	N(H)	H(N)	C(O)	C(2)	H(2)
vacuum	162.12	29.41	153.22	27.77	30.46	168.45	30.41
PCM	162.21	29.38	143.79	25.78	24.77	168.02	30.26
PSE-1^{q}	162.01	29.31	133.62	25.57	20.93	167.16	30.10
PSE-1$^{\varphi}$	162.40	29.31	135.65	25.58	21.18	167.13	30.10
PSE-2^{q}	162.02	29.30	131.83	25.38	20.33	167.11	30.07
PSE-2$^{\varphi}$	162.47	29.30	134.23	25.38	20.60	167.09	30.08
PSE-3^{q}	162.02	29.30	131.28	25.31	20.13	167.10	30.07
PSE-3$^{\varphi}$	162.49	29.30	133.81	25.30	20.42	167.07	30.07
PSE-4^{q}	162.02	29.30	131.09	25.28	20.06	167.09	30.06
PSE-4$^{\varphi}$	162.50	29.30	133.68	25.26	20.36	167.07	30.07
HNC^{q}	162.02	29.30	130.99	25.26	20.02	167.09	30.06
HNC$^{\varphi}$	162.51	29.30	133.63	25.22	20.33	167.08	30.07

Table S2. Calculated isotropic chemical shielding constants for optimized NMA structures with a series of Pople basis sets (GIAO/B3LYP/(6-31G(d),6-31+G(d,p),6-311++G(d,p))). For all EC-RISM calculations, the optimal PCM structure was used.

	σ / ppm						
	C(1)	H(1)	N(H)	H(N)	C(O)	C(2)	H(2)
vacuum							
6-31G(d)	165.02	29.47	150.25	28.28	34.91	167.14	30.64
6-31+G(d,p)	166.97	28.72	152.53	27.11	31.67	169.40	29.97
6-311++G(d,p)	155.53	29.33	140.31	27.42	11.34	161.28	30.31
PCM							
6-31G(d)	161.97	29.60	146.60	26.36	29.51	167.95	30.53
6-31+G(d,p)	163.77	28.93	145.70	24.83	24.59	169.76	29.79
6-311++G(d,p)	154.18	29.14	126.30	25.19	2.44	159.59	30.00
PSE-1^q							
6-31G(d)	161.95	29.53	139.42	26.22	25.90	167.34	30.39
6-31+G(d,p)	163.57	28.86	135.52	24.65	20.68	168.85	29.63
6-311++G(d,p)	154.05	29.08	115.17	25.02	-2.27	158.43	29.83
HNC^q							
6-31G(d)	161.95	29.53	139.42	26.22	25.90	167.34	30.39
6-31+G(d,p)	163.57	28.84	132.84	24.34	19.75	168.77	29.59
6-311++G(d,p)	154.03	29.06	112.14	24.72	-3.48	158.26	29.78

Table S3. Calculated isotropic chemical shielding constants (GIAO/B3LYP/aug-cc-pVTZ) for optimized NMA structures from all models studied. For all EC-RISM calculations, the optimal PCM structure was used.

	σ / ppm						
	C(1)	H(1)	N(H)	H(N)	C(O)	C(2)	H(2)
Vacuum	155.31	29.00	141.27	26.68	10.98	161.58	29.94
PCM	155.51	28.97	130.94	24.80	4.80	160.68	29.80
PSE-1^q	155.31	28.92	120.24	24.65	0.29	159.71	29.67
PSE-1$^\varphi$	154.23	28.74	119.05	24.42	-1.76	158.53	29.49
PSE-2^q	155.31	28.91	118.22	24.48	-0.51	159.62	29.64
PSE-2$^\varphi$	154.28	28.74	117.59	24.23	-2.46	158.44	29.47
PSE-3^q	155.31	28.91	117.59	24.41	-0.78	159.59	29.64
PSE-3$^\varphi$	154.31	28.73	117.19	24.15	-2.68	158.41	29.46
PSE-4^q	155.30	28.90	117.36	24.39	-0.88	159.58	29.63
PSE-4$^\varphi$	154.32	28.73	117.07	24.11	-2.75	158.40	29.46
HNC^q	155.30	28.90	117.24	24.37	-0.94	159.57	29.63
HNC$^\varphi$	154.33	28.73	117.06	24.08	-2.78	158.39	29.46

Table S4**.** Calculated isotropic reference chemical shielding constants (GIAO/B3LYP with two basis sets on optimized B3LYP geometries, *in vacuo* or using PCM for solution calculations). The ^{15}N chemical shielding of ammonia in aqueous solution was used as secondary NMR reference with a known chemical shift of -19.4 ppm.[a1] The hydrogen atoms of TMS and the carbon atoms of 1,4-dioxane were selected for secondary ^{1}H and ^{13}C reference chemical shielding constants with known chemical shifts in water of -0.094 for ^{1}H[a2] and 66.6 ppm for ^{13}C.[a3] Shielding constants for ammonia with EC-RISM closure approximations PSE-4$^{\varphi}$ and HNC$^{\varphi}$ and the large basis set could not be obtained due to divergence of the corresponding 3D RISM solutions.

	σ_{ref} / ppm					
	6-31+G(d)			aug-cc-pVTZ		
	^{13}C	^{1}H	^{15}N	^{13}C	^{1}H	^{15}N
vacuum	189.19	31.97	243.04	179.39	31.65	239.11
PCM	189.02	31.95	246.95	179.09	31.64	244.28
PSE-1^{q}	189.09	32.12	251.77	179.20	31.65	245.71
PSE-1$^{\varphi}$	189.17	32.12	250.01	179.29	31.65	243.42
PSE-2^{q}	189.11	32.12	252.99	179.22	31.66	246.36
PSE-2$^{\varphi}$	189.18	32.12	250.41	179.28	31.65	240.69
PSE-3^{q}	189.12	32.12	253.44	179.22	31.66	246.62
PSE-3$^{\varphi}$	189.18	32.12	250.49	179.28	31.65	236.90
PSE-4^{q}	189.12	32.12	253.63	179.22	31.66	246.74
PSE-4$^{\varphi}$	189.18	32.12	250.50	179.27	31.65	-
HNC^{q}	189.12	32.12	253.74	179.22	31.66	246.81
HNC$^{\varphi}$	189.18	32.12	250.48	179.27	31.65	-

Table S5. Calculated isotropic reference chemical shielding constants (GIAO/B3LYP with two basis sets on optimized B3LYP geometries, *in vacuo* or using PCM for solution calculations). The ^{15}N chemical shielding of ammonia in aqueous solution was used as secondary NMR reference with a known chemical shift of -19.4 ppm.[a1] The hydrogen atoms of TMS and the carbon atoms of 1,4-dioxane were selected for secondary ^{1}H and ^{13}C reference chemical shielding constants with known chemical shifts in water of -0.094 for ^{1}H[a2] and 66.6 ppm for ^{13}C.[a3] Shielding constants for ammonia with EC-RISM closure approximations PSE-4$^{\varphi}$ and HNC$^{\varphi}$ and the large basis set could not be obtained due to divergence of the corresponding 3D RISM solutions.

	σ_{ref} / ppm		
	^{13}C	^{1}H	^{15}N
vacuum			
6-31G(d)	191.25	32.09	235.64
6-31+G(d,p)	190.82	31.56	248.87
6-311++G(d,p)	178.93	31.88	240.02
PCM			
6-31G(d)	191.43	32.08	239.95
6-31+G(d,p)	190.11	31.53	252.47
6-311++G(d,p)	178.53	31.86	242.64
PSE-1^{q}			
6-31G(d)	191.48	32.09	245.89
6-31+G(d,p)	190.16	31.55	255.42
6-311++G(d,p)	178.55	31.87	245.31
HNC^{q}			
6-31G(d)	191.50	32.09	247.89
6-31+G(d,p)	190.18	31.55	256.95
6-311++G(d,p)	178.54	31.87	246.83

Table S6. Dipole moments p of NMA after EC-RISM convergence for two basis sets with various closure approximations.

	6-31+G(d)	aug-cc-pVTZ
closure	p / D	p / D
vacuum	4.07	3.90
PCM	5.83	4.07
PSE-1^q	6.81	6.70
PSE-1$^\varphi$	6.69	6.65
PSE-2^q	7.03	6.93
PSE-2$^\varphi$	6.91	6.87
PSE-3^q	7.10	7.00
PSE-3$^\varphi$	6.98	6.96
PSE-4^q	7.13	7.03
PSE-4$^\varphi$	7.01	6.99
HNC^q	7.14	7.04
HNC$^\varphi$	7.04	7.03

Table S7. Linear regression results (slope, intercept, and coefficient of determination R^2) between calculated chemical shifts δ_{calc} and dipole moments p for calculations with the EC-RISM solvation model based on the point charge approach (PSE-n^q) for the electrostatic solute-solvent interaction.

	6-31+G(d)			aug-cc-pVTZ		
	R^2	slope / ppm D^{-1}	intercept / ppm	R^2	slope / ppm D^{-1}	intercept / ppm
N(H)	1.000	14.06	22.31	1.000	11.85	46.07
H(N)	0.997	0.94	0.18	0.998	0.81	1.51
C(O)	1.000	2.87	148.62	1.000	3.62	154.69

Table S8. Linear regression results (slope, intercept, and coefficient of determination R^2) between calculated chemical shifts δ_{calc} and dipole moments p for calculations with the EC-RISM solvation model based on the exact quantum-mechanical potential (PSE-n^φ) for the electrostatic solute-solvent interaction.

	6-31+G(d)			aug-cc-pVTZ		
	R^2	slope / ppm D^{-1}	intercept / ppm	R^2	slope / ppm D^{-1}	intercept / ppm
N(H)	0.980	7.52	64.10	-	-	-
H(N)	0.996	1.03	-0.34	0.998	0.91	1.25
C(O)	0.993	2.57	150.81	0.981	2.75	162.79

A1.2 Solvation structure and thermodynamics of nondipolar liquids

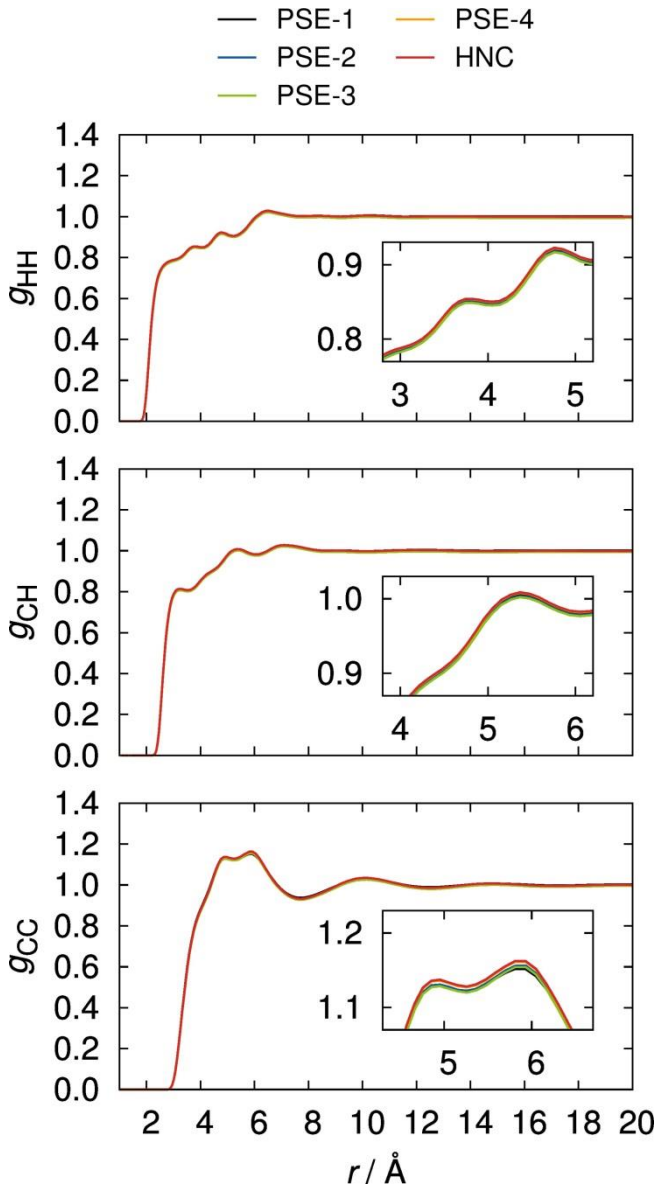


Figure. S1. 1D RISM radial site-site distribution function of benzene, calculated with the hypernetted chain closure and four different orders of the partial series expansion closure. Furthermore the Amber/OPLS force field parameters were taken for the illustrated results.

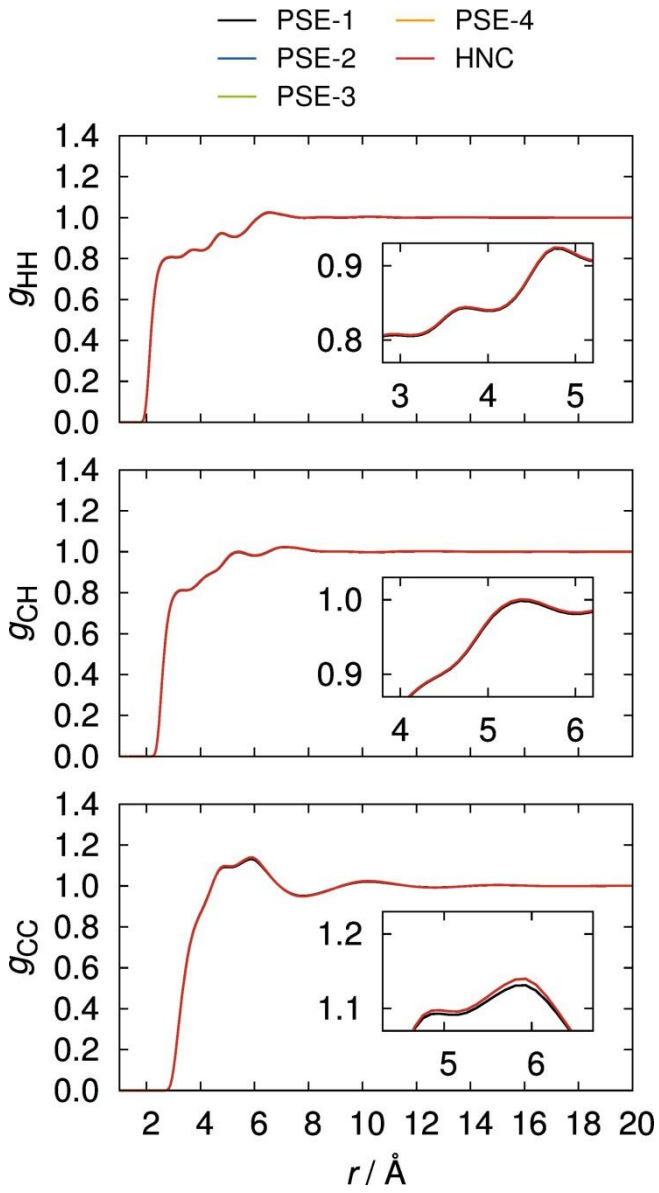


Figure. S2. 1D RISM radial site-site distribution function of benzene, calculated with the hypernetted chain closure and four different orders of the partial series expansion closure. Furthermore the Cornell force field parameters were taken for the illustrated results.

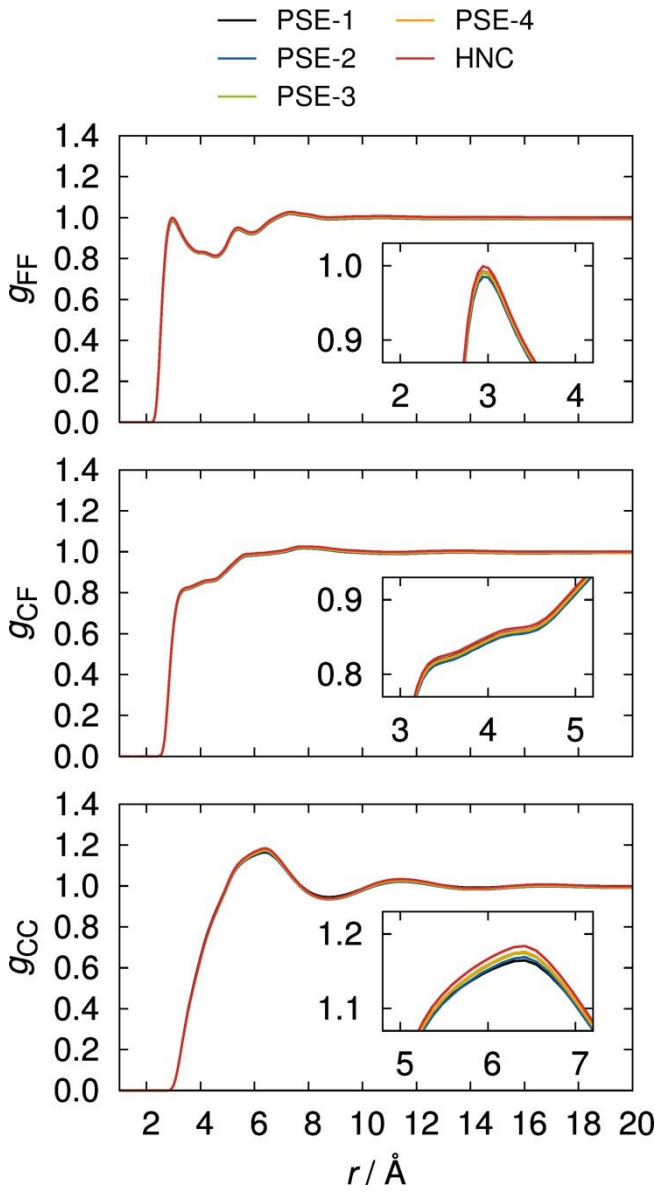


Figure. S3. 1D RISM radial site-site distribution function of hexafluorobenzene, calculated with the hypernetted chain closure and four different orders of the partial series expansion closure. Furthermore the Amber/OPLS force field parameters were taken for the illustrated results.

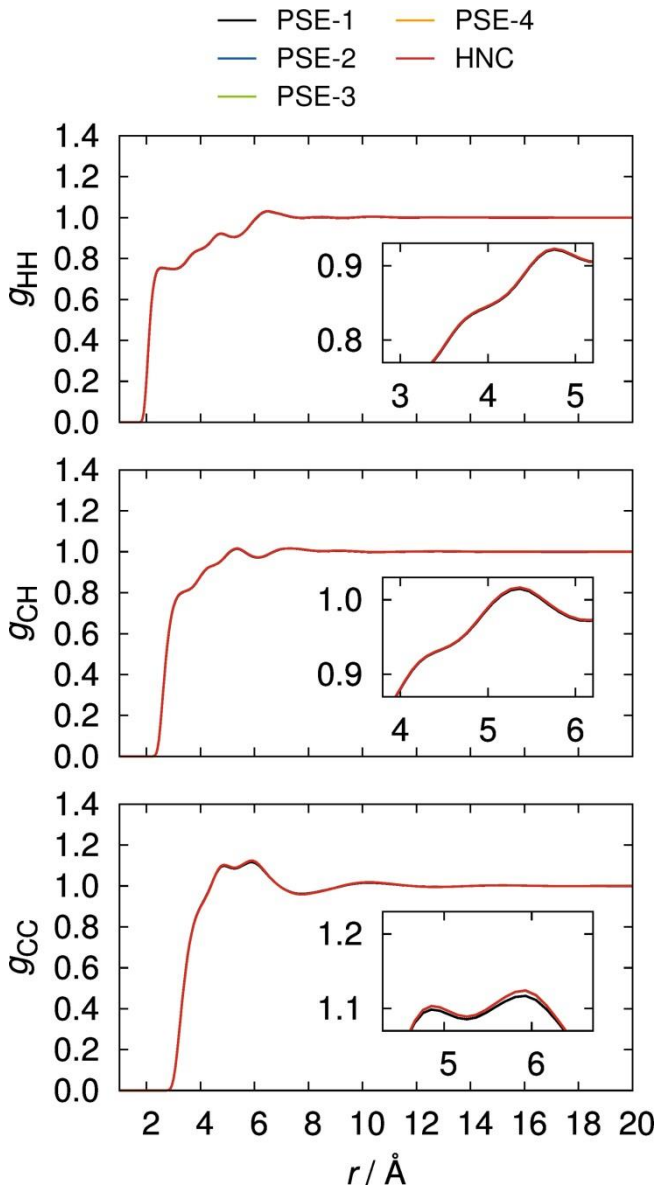


Figure. S4. 1D RISM radial site-site distribution function of benzene without partial charges, calculated with the hypernetted chain closure and four different orders of the partial series expansion closure. Furthermore the Amber/OPLS force field parameters were taken for the illustrated results, but the partial charges were omitted and set to zero.

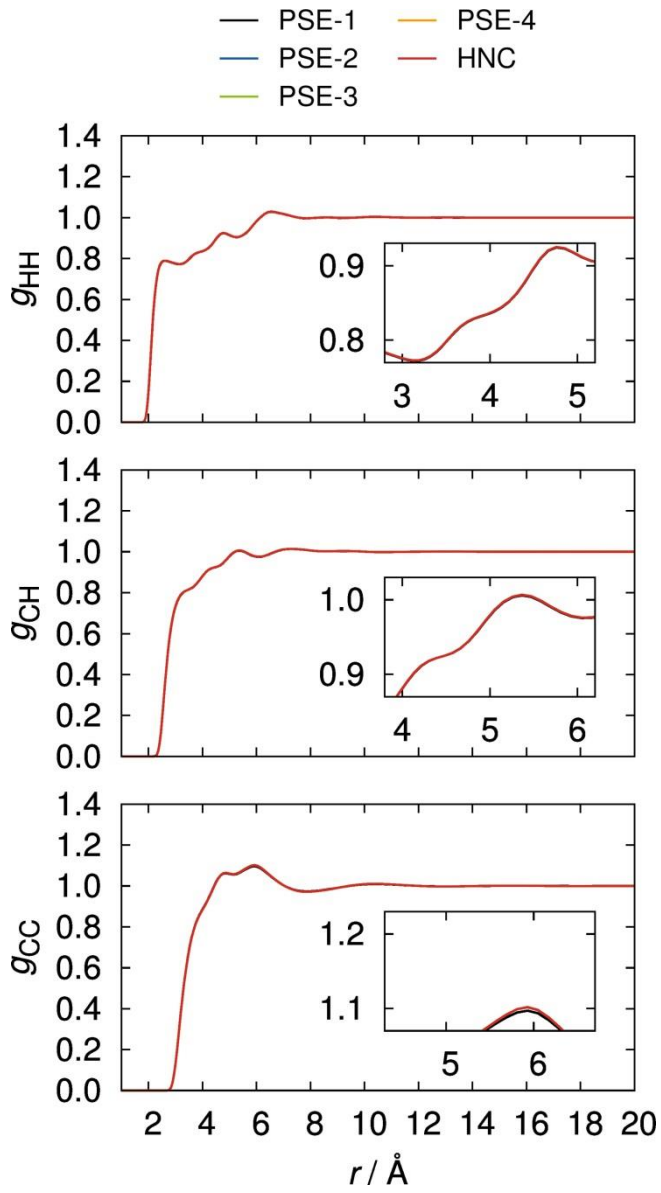


Figure. S5. 1D RISM radial site-site distribution function of benzene without partial charges, calculated with the hypernetted chain closure and four different orders of the partial series expansion closure. Furthermore the Cornell force field parameters were taken for the illustrated results, but the partial charges were omitted and set to zero.

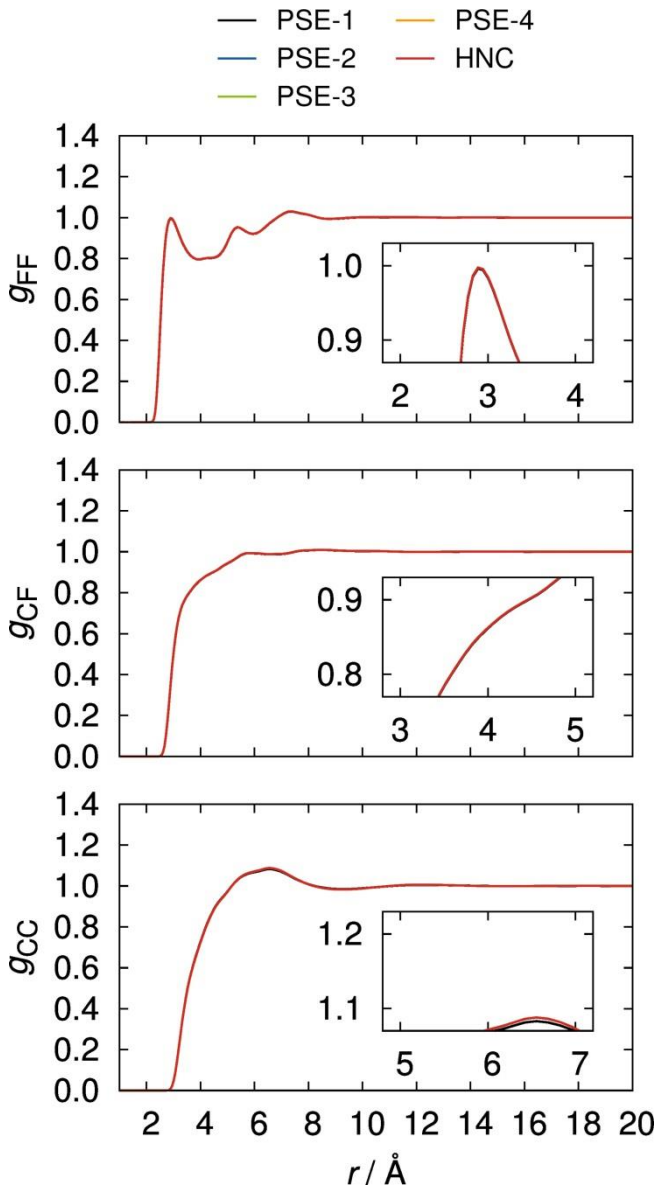


Figure. S6. 1D RISM radial site-site distribution function of hexafluorobenzene without partial charges, calculated with the hypernetted chain closure and four different orders of the partial series expansion closure. Furthermore the Amber/OPLS force field parameters were taken for the illustrated results, but the partial charges were omitted and set to zero.

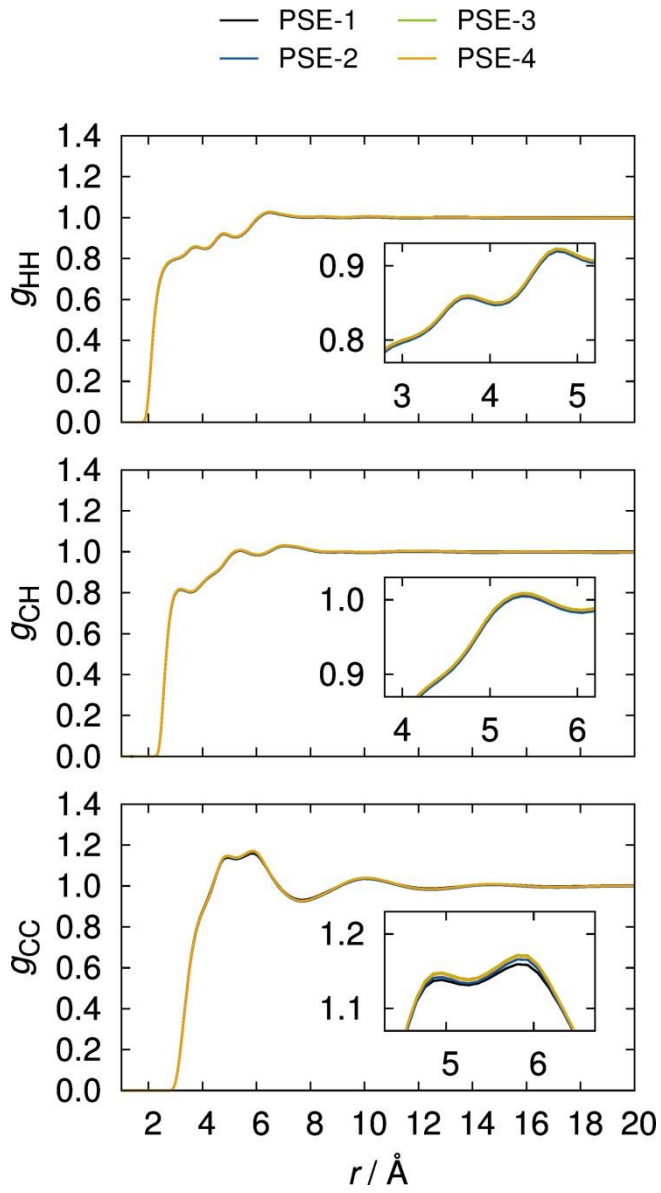


Figure. S7. 1D RISM radial site-site distribution function of benzene with the partial charges of hexafluorobenzene, calculated with four different orders of the partial series expansion closure. Results with the hypernetted chain closure could not been achieved. The Amber/OPLS force field parameters for benzene have been used for dispersive interactions and the Amber/OPLS force field parameters of hexafluorobenzene were used for the partial charges, defining our inversed quadrupole moment model abbreviated as q_{rev}-C_6H_6.

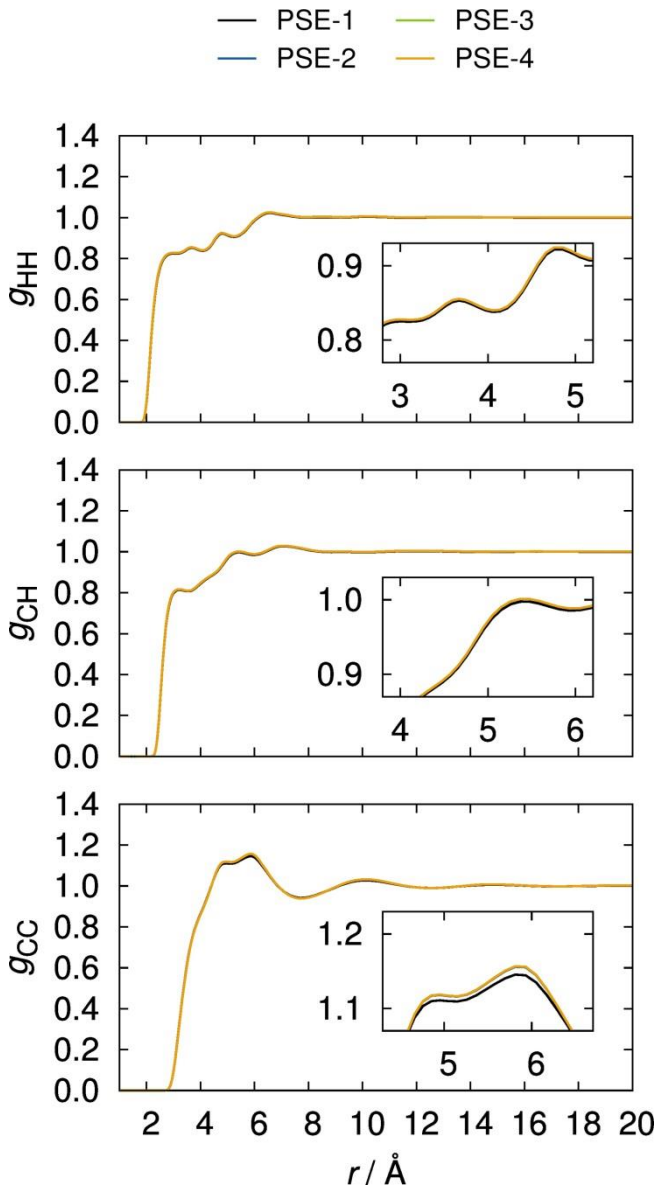


Figure. S8. 1D RISM radial site-site distribution function of benzene with the partial charges of hexafluorobenzene, calculated with four different orders of the partial series expansion closure. Results with the hypernetted chain closure could not been achieved. The Cornell force field parameters for benzene have been used for dispersive interactions and the Amber/OPLS force field parameters of hexafluorobenzene were used for the partial charges.

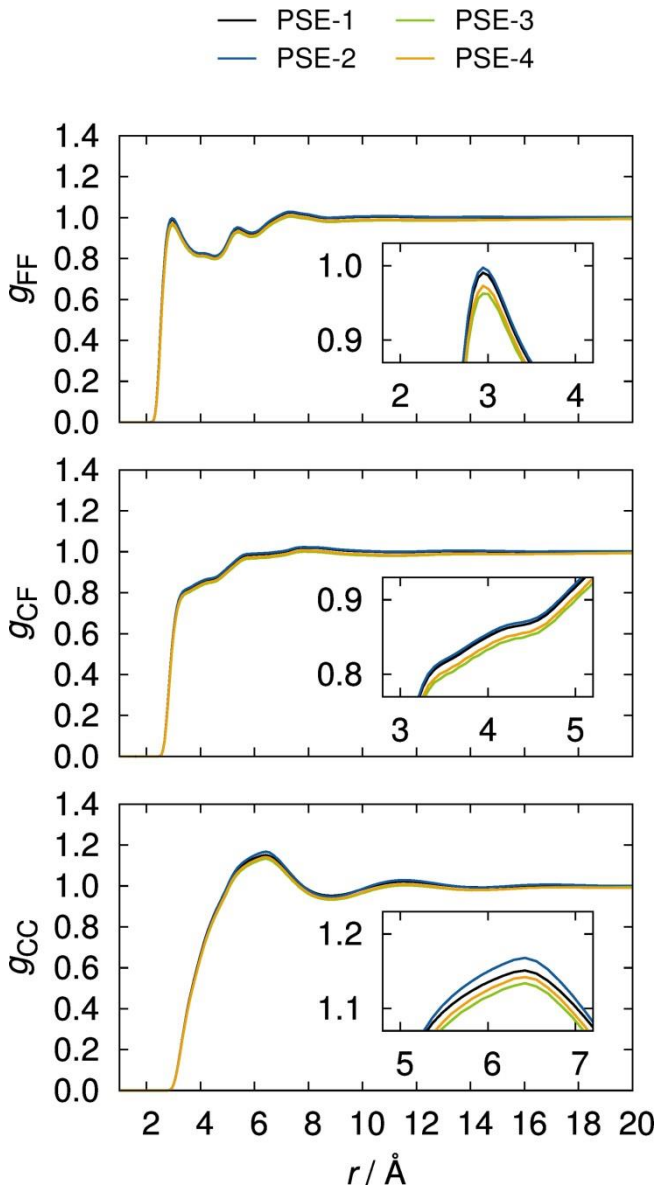


Figure. S9. 1D RISM radial site-site distribution function of hexafluorobenzene with the partial charges of benzene, calculated with four different orders of the partial series expansion closure. Results with the hypernetted chain closure could not been achieved. The Amber/OPLS force field parameters for hexafluorobenzene have been used for dispersive interactions and the Amber/OPLS force field parameters of benzene were used for the partial charges, defining our inversed quadrupole moment model abbreviated as q_{rev}-C_6F_6.

TABLE S8. IEFPCM results for the Gibbs free energies of all 9 gas molecules in benzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	-1791.641	-80618.900	-330555.488	-68363.457	-93886.676	-70744.364	-117743.261	-25222.894	-273371.665
6-311G**	-1794.613	-80649.113	-330576.122	-68381.052	-93913.104	-70764.176	-117776.791	-25231.584	-273446.430
6-311+G**	-1794.613	-80651.672	-330576.310	-68382.390	-93915.714	-70765.154	-117779.321	-25231.632	-273450.515
aug-cc-pVDZ	-1791.984	-80632.670	-330572.574	-68374.712	-93904.489	-70755.388	-117762.135	-25225.702	-273403.363
aug-cc-pVTZ	-1795.422	-80655.840	-330580.343	-68390.957	-93927.184	-70772.607	-117792.073	-25234.531	-273483.673
aug-cc-pVQZ	-1795.634	-80662.418	-330582.509	-68395.437	-93933.838	-70777.541	-117800.279	-25236.189	-273503.081
B3LYP									
6-31G*	-1824.202	-80882.425	-331021.977	-68727.535	-94327.267	-71102.945	-118336.995	-25425.717	-274520.589
6-311G**	-1827.955	-80917.905	-331044.609	-68747.521	-94355.372	-71126.029	-118374.822	-25435.360	-274606.852
6-311+G**	-1827.955	-80923.867	-331045.034	-68749.876	-94358.911	-71127.809	-118378.518	-25435.478	-274614.924
aug-cc-pVDZ	-1825.489	-80903.531	-331039.703	-68739.256	-94345.901	-71116.013	-118357.911	-25427.125	-274563.348
aug-cc-pVTZ	-1829.002	-80927.560	-331048.873	-68756.681	-94367.794	-71133.924	-118388.753	-25438.297	-274641.781
aug-cc-pVQZ	-1829.232	-80934.372	-331051.179	-68761.352	-94374.748	-71139.176	-118397.593	-25440.302	-274662.716
PBE0									
6-31G*	-1812.055	-80816.893	-330911.143	-68649.814	-94232.116	-71022.689	-118212.533	-25387.128	-274266.755
6-311G**	-1815.449	-80850.293	-330932.642	-68667.951	-94258.100	-71043.531	-118247.262	-25395.706	-274347.367
6-311+G**	-1815.449	-80855.390	-330932.973	-68669.937	-94261.190	-71045.024	-118250.466	-25395.849	-274354.188
aug-cc-pVDZ	-1813.104	-80836.101	-330928.413	-68661.239	-94249.748	-71034.690	-118232.342	-25389.166	-274306.391
aug-cc-pVTZ	-1816.446	-80859.131	-330936.949	-68676.784	-94270.866	-71051.211	-118260.853	-25398.249	-274381.567
aug-cc-pVQZ	-1816.658	-80865.904	-330939.154	-68681.415	-94277.799	-71056.404	-118269.600	-25400.140	-274402.272
MP2									
6-31G*	-1798.670	-80713.225	-330641.650	-68558.847	-94094.887	-70921.944	-118039.960	-25309.107	-273875.062
6-311G**	-1810.102	-80780.274	-330668.567	-68585.059	-94141.439	-70955.550	-118097.322	-25338.401	-274037.758
6-311+G**	-1810.102	-80785.039	-330669.022	-68587.906	-94145.208	-70957.699	-118102.053	-25338.586	-274048.710
aug-cc-pVDZ	-1808.903	-80762.485	-330669.416	-68574.734	-94129.118	-70943.129	-118078.848	-25331.186	-273991.860
aug-cc-pVTZ	-1816.520	-80826.848	-330712.704	-68627.548	-94202.322	-70998.002	-118174.157	-25360.515	-274221.512
aug-cc-pVQZ	-1818.052	-80848.941	-330728.184	-68645.681	-94227.106	-71017.154	-118206.907	-25368.640	-274297.534

TABLE S9. IEFPCM results for the Gibbs free energies of all 9 gas molecules in hexafluorobenzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	-1791.641	-80618.900	-330555.488	-68363.445	-93886.672	-70744.331	-117743.139	-25222.889	-273371.640
6-311G**	-1794.613	-80649.113	-330576.122	-68381.037	-93913.098	-70764.145	-117776.668	-25231.578	-273446.399
6-311+G**	-1794.613	-80651.672	-330576.310	-68382.375	-93915.707	-70765.120	-117779.193	-25231.626	-273450.473
aug-cc-pVDZ	-1791.983	-80632.670	-330572.574	-68374.703	-93904.487	-70755.357	-117762.020	-25225.696	-273403.337
aug-cc-pVTZ	-1795.422	-80655.840	-330580.343	-68390.946	-93927.181	-70772.578	-117791.963	-25234.525	-273483.646
aug-cc-pVQZ	-1795.634	-80662.418	-330582.509	-68395.427	-93933.835	-70777.512	-117800.170	-25236.183	-273503.053
B3LYP									
6-31G*	-1824.202	-80882.425	-331021.977	-68727.521	-94327.264	-71102.922	-118336.916	-25425.711	-274520.596
6-311G**	-1827.955	-80917.905	-331044.609	-68747.504	-94355.366	-71126.006	-118374.737	-25435.353	-274606.818
6-311+G**	-1827.955	-80923.866	-331045.034	-68749.860	-94358.904	-71127.784	-118378.424	-25435.472	-274614.877
aug-cc-pVDZ	-1825.488	-80903.531	-331039.703	-68739.244	-94345.898	-71115.990	-118357.827	-25427.119	-274563.335
aug-cc-pVTZ	-1829.001	-80927.560	-331048.873	-68756.668	-94367.791	-71133.901	-118388.671	-25438.289	-274641.761
aug-cc-pVQZ	-1829.231	-80934.372	-331051.179	-68761.339	-94374.744	-71139.154	-118397.511	-25440.296	-274662.689
PBE0									
6-31G*	-1812.055	-80816.893	-330911.143	-68649.801	-94232.113	-71022.667	-118212.455	-25387.120	-274266.743
6-311G**	-1815.449	-80850.293	-330932.642	-68667.935	-94258.094	-71043.508	-118247.179	-25395.698	-274347.354
6-311+G**	-1815.449	-80855.390	-330932.973	-68669.922	-94261.183	-71045.000	-118250.376	-25395.841	-274354.165
aug-cc-pVDZ	-1813.103	-80836.101	-330928.413	-68661.228	-94249.745	-71034.667	-118232.262	-25389.159	-274306.361
aug-cc-pVTZ	-1816.445	-80859.131	-330936.948	-68676.772	-94270.863	-71051.189	-118260.775	-25398.242	-274381.549
aug-cc-pVQZ	-1816.657	-80865.904	-330939.153	-68681.403	-94277.795	-71056.382	-118269.523	-25400.132	-274402.250
MP2									
6-31G*	-1798.669	-80713.225	-330641.650	-68558.831	-94094.885	-70921.935	-118039.893	-25309.101	-273875.046
6-311G**	-1810.102	-80780.274	-330668.567	-68585.040	-94141.434	-70955.541	-118097.253	-25338.395	-274037.731
6-311+G**	-1810.102	-80785.039	-330669.022	-68587.888	-94145.202	-70957.689	-118101.975	-25338.580	-274048.685
aug-cc-pVDZ	-1808.903	-80762.485	-330669.415	-68574.722	-94129.117	-70943.120	-118078.779	-25331.180	-273991.842
aug-cc-pVTZ	-1816.519	-80826.848	-330712.704	-68627.535	-94202.319	-70997.993	-118174.093	-25360.508	-274221.490
aug-cc-pVQZ	-1818.051	-80848.941	-330728.184	-68645.668	-94227.103	-71017.144	-118206.842	-25368.632	-274297.515

TABLE S10. EC-RISM[q] results for the Gibbs free energies of all 9 gas molecules in benzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Amber/OPLS force field was applied for dispersion interaction. ChelpG partial charges were evaluated for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	-1788.939	-80616.155	-330552.844	-68359.922	-93883.642	-70740.806	-117740.434	-25222.027	-273367.856
6-311G**	-1791.910	-80646.368	-330573.478	-68377.503	-93910.060	-70760.639	-117773.989	-25230.708	-273442.586
6-311+G**	-1791.910	-80648.927	-330573.665	-68378.846	-93912.667	-70761.605	-117776.510	-25230.755	-273446.646
aug-cc-pVDZ	-1789.280	-80629.925	-330569.930	-68371.202	-93901.475	-70751.853	-117759.346	-25224.813	-273399.523
aug-cc-pVTZ	-1792.718	-80653.094	-330577.698	-68387.453	-93924.167	-70769.093	-117789.321	-25233.663	-273479.877
aug-cc-pVQZ	-1792.931	-80659.673	-330579.865	-68391.934	-93930.822	-70774.032	-117797.536	-25235.322	-273499.295
B3LYP									
6-31G*	-1821.499	-80879.680	-331019.333	-68723.949	-94324.169	-71099.423	-118334.466	-25424.831	-274516.714
6-311G**	-1825.251	-80915.160	-331041.965	-68743.919	-94352.263	-71122.510	-118372.280	-25434.470	-274602.931
6-311+G**	-1825.251	-80921.121	-331042.389	-68746.281	-94355.793	-71124.284	-118375.944	-25434.585	-274610.940
aug-cc-pVDZ	-1822.784	-80900.785	-331037.059	-68735.683	-94342.812	-71112.497	-118355.325	-25426.226	-274559.388
aug-cc-pVTZ	-1826.296	-80924.814	-331046.228	-68753.122	-94364.705	-71130.422	-118386.169	-25437.413	-274637.864
aug-cc-pVQZ	-1826.526	-80931.627	-331048.534	-68757.792	-94371.659	-71135.676	-118395.012	-25439.420	-274658.805
PBE0									
6-31G*	-1809.352	-80814.148	-330908.500	-68646.237	-94229.036	-71019.175	-118210.019	-25386.224	-274262.922
6-311G**	-1812.745	-80847.548	-330929.998	-68664.360	-94255.009	-71040.016	-118244.743	-25394.804	-274343.505
6-311+G**	-1812.745	-80852.645	-330930.328	-68666.350	-94258.091	-71041.505	-118247.923	-25394.943	-274350.278
aug-cc-pVDZ	-1810.399	-80833.356	-330925.768	-68657.678	-94246.680	-71031.180	-118229.793	-25388.256	-274302.484
aug-cc-pVTZ	-1813.741	-80856.386	-330934.303	-68673.233	-94267.795	-71047.714	-118258.308	-25397.352	-274377.717
aug-cc-pVQZ	-1813.953	-80863.159	-330936.509	-68677.863	-94274.728	-71052.909	-118267.060	-25399.243	-274398.422
MP2									
6-31G*	-1795.967	-80710.480	-330639.006	-68555.198	-94091.748	-70918.453	-118037.217	-25308.228	-273871.152
6-311G**	-1807.398	-80777.529	-330665.923	-68581.402	-94138.309	-70952.070	-118094.606	-25337.502	-274033.839
6-311+G**	-1807.398	-80782.293	-330666.375	-68584.251	-94142.070	-70954.201	-118099.316	-25337.681	-274044.747
aug-cc-pVDZ	-1806.199	-80759.737	-330666.768	-68571.112	-94125.999	-70939.635	-118076.107	-25330.259	-273987.874
aug-cc-pVTZ	-1813.815	-80824.101	-330710.058	-68623.947	-94199.212	-70994.532	-118171.460	-25359.623	-274217.622
aug-cc-pVQZ	-1815.348	-80846.194	-330725.538	-68642.083	-94224.002	-71013.688	-118204.222	-25367.754	-274293.626

TABLE S11. EC-RISM[q] results for the Gibbs free energies of all 9 gas molecules in hexafluorobenzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Amber/OPLS force field was applied for dispersion interaction. ChelpG partial charges were evaluated for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	-1789.308	-80616.574	-330553.386	-68360.584	-93884.234	-70741.476	-117741.398	-25222.151	-273368.882
6-311G**	-1792.279	-80646.787	-330574.019	-68378.164	-93910.650	-70761.303	-117774.952	-25230.834	-273443.636
6-311+G**	-1792.279	-80649.346	-330574.207	-68379.507	-93913.257	-70762.272	-117777.483	-25230.881	-273447.705
aug-cc-pVDZ	-1789.650	-80630.344	-330570.471	-68371.864	-93902.066	-70752.521	-117760.301	-25224.938	-273400.560
aug-cc-pVTZ	-1793.088	-80653.514	-330578.240	-68388.113	-93924.757	-70769.758	-117790.265	-25233.783	-273480.909
aug-cc-pVQZ	-1793.300	-80660.092	-330580.406	-68392.593	-93931.412	-70774.695	-117798.477	-25235.443	-273500.325
B3LYP									
6-31G*	-1821.869	-80880.099	-331019.874	-68724.617	-94324.771	-71100.087	-118335.398	-25424.950	-274517.731
6-311G**	-1825.621	-80915.579	-331042.506	-68744.586	-94352.863	-71123.168	-118373.216	-25434.595	-274603.981
6-311+G**	-1825.621	-80921.540	-331042.931	-68746.947	-94356.393	-71124.947	-118376.899	-25434.710	-274612.009
aug-cc-pVDZ	-1823.153	-80901.205	-331037.600	-68736.351	-94343.413	-71113.162	-118356.263	-25426.356	-274560.435
aug-cc-pVTZ	-1826.666	-80925.233	-331046.769	-68753.788	-94365.305	-71131.083	-118387.096	-25437.542	-274638.908
aug-cc-pVQZ	-1826.896	-80932.046	-331049.075	-68758.457	-94372.258	-71136.336	-118395.937	-25439.549	-274659.847
PBE0									
6-31G*	-1809.722	-80814.567	-330909.041	-68646.905	-94229.636	-71019.838	-118210.946	-25386.330	-274263.929
6-311G**	-1813.115	-80847.967	-330930.539	-68665.025	-94255.606	-71040.673	-118245.674	-25394.916	-274344.542
6-311+G**	-1813.115	-80853.064	-330930.869	-68667.016	-94258.688	-71042.166	-118248.868	-25395.054	-274351.330
aug-cc-pVDZ	-1810.769	-80833.775	-330926.310	-68658.346	-94247.278	-71031.842	-118230.724	-25388.371	-274303.515
aug-cc-pVTZ	-1814.110	-80856.805	-330934.845	-68673.898	-94268.393	-71048.373	-118259.228	-25397.466	-274378.745
aug-cc-pVQZ	-1814.323	-80863.578	-330937.050	-68678.528	-94275.325	-71053.567	-118267.977	-25399.357	-274399.449
MP2									
6-31G*	-1796.336	-80710.899	-330639.548	-68555.883	-94092.359	-70919.098	-118038.114	-25308.353	-273872.175
6-311G**	-1807.768	-80777.948	-330666.465	-68582.083	-94138.916	-70952.709	-118095.502	-25337.648	-274034.885
6-311+G**	-1807.768	-80782.714	-330666.919	-68584.938	-94142.680	-70954.850	-118100.239	-25337.837	-274045.813
aug-cc-pVDZ	-1806.569	-80760.160	-330667.314	-68571.807	-94126.614	-70940.288	-118077.020	-25330.421	-273988.922
aug-cc-pVTZ	-1814.185	-80824.522	-330710.602	-68624.632	-94199.822	-70995.181	-118172.358	-25359.767	-274218.654
aug-cc-pVQZ	-1815.717	-80846.615	-330726.081	-68642.766	-94224.612	-71014.337	-118205.117	-25367.894	-274294.656

TABLE S12. EC-RISM[q] results for the Gibbs free energies of all 9 gas molecules in benzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Cornell force field was applied for dispersion interaction. ChelpG partial charges were evaluated for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	-1788.973	-80616.174	-330552.813	-68359.886	-93883.598	-70740.754	-117740.247	-25221.932	-273367.699
6-311G**	-1791.944	-80646.387	-330573.446	-68377.468	-93910.016	-70760.589	-117773.799	-25230.614	-273442.426
6-311+G**	-1791.944	-80648.946	-330573.634	-68378.810	-93912.623	-70761.554	-117776.317	-25230.661	-273446.486
aug-cc-pVDZ	-1789.315	-80629.944	-330569.898	-68371.166	-93901.431	-70751.802	-117759.162	-25224.719	-273399.366
aug-cc-pVTZ	-1792.752	-80653.113	-330577.667	-68387.417	-93924.124	-70769.043	-117789.139	-25233.568	-273479.719
aug-cc-pVQZ	-1792.965	-80659.692	-330579.833	-68391.898	-93930.779	-70773.982	-117797.355	-25235.227	-273499.137
B3LYP									
6-31G*	-1821.534	-80879.699	-331019.302	-68723.913	-94324.126	-71099.378	-118334.279	-25424.736	-274516.557
6-311G**	-1825.285	-80915.179	-331041.933	-68743.884	-94352.220	-71122.467	-118372.091	-25434.376	-274602.772
6-311+G**	-1825.285	-80921.140	-331042.358	-68746.246	-94355.750	-71124.238	-118375.750	-25434.491	-274610.779
aug-cc-pVDZ	-1822.818	-80900.804	-331037.027	-68735.648	-94342.769	-71112.451	-118355.140	-25426.131	-274559.231
aug-cc-pVTZ	-1826.331	-80924.833	-331046.197	-68753.087	-94364.662	-71130.377	-118385.987	-25437.319	-274637.706
aug-cc-pVQZ	-1826.561	-80931.646	-331048.502	-68757.757	-94371.616	-71135.632	-118394.831	-25439.326	-274658.648
PBE0									
6-31G*	-1809.387	-80814.167	-330908.468	-68646.202	-94228.993	-71019.130	-118209.831	-25386.131	-274262.766
6-311G**	-1812.780	-80847.567	-330929.966	-68664.324	-94254.965	-71039.973	-118244.554	-25394.711	-274343.345
6-311+G**	-1812.780	-80852.664	-330930.297	-68666.315	-94258.048	-71041.460	-118247.730	-25394.850	-274350.117
aug-cc-pVDZ	-1810.433	-80833.375	-330925.737	-68657.643	-94246.637	-71031.135	-118229.609	-25388.162	-274302.327
aug-cc-pVTZ	-1813.775	-80856.405	-330934.272	-68673.198	-94267.752	-71047.670	-118258.127	-25397.258	-274377.560
aug-cc-pVQZ	-1813.987	-80863.178	-330936.477	-68677.828	-94274.685	-71052.865	-118266.879	-25399.149	-274398.265
MP2									
6-31G*	-1796.001	-80710.499	-330638.975	-68555.164	-94091.706	-70918.420	-118037.063	-25308.133	-273870.999
6-311G**	-1807.433	-80777.548	-330665.892	-68581.368	-94138.267	-70952.039	-118094.449	-25337.411	-274033.683
6-311+G**	-1807.433	-80782.312	-330666.344	-68584.218	-94142.028	-70954.170	-118099.154	-25337.590	-274044.591
aug-cc-pVDZ	-1806.234	-80759.757	-330666.737	-68571.079	-94125.957	-70939.602	-118075.951	-25330.169	-273987.721
aug-cc-pVTZ	-1813.850	-80824.120	-330710.027	-68623.913	-94199.170	-70994.498	-118171.306	-25359.530	-274217.469
aug-cc-pVQZ	-1815.382	-80846.214	-330725.507	-68642.049	-94223.960	-71013.654	-118204.068	-25367.660	-274293.473

TABLE S13. EC-RISM$^{\varphi}$ results for the Gibbs free energies of all 9 gas molecules in benzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Amber/OPLS force field was applied for dispersion interaction. The exact molecular electronic potential was used for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	-1788.938	-80616.155	-330552.842	-68359.962	-93883.639	-70740.934	-117740.376	-25221.973	-273367.834
6-311G**	-1791.909	-80646.368	-330573.472	-68377.556	-93910.065	-70760.766	-117773.929	-25230.641	-273442.559
6-311+G**	-1791.909	-80648.926	-330573.657	-68378.892	-93912.667	-70761.728	-117776.434	-25230.684	-273446.607
aug-cc-pVDZ	-1789.279	-80629.922	-330569.922	-68371.219	-93901.459	-70751.964	-117759.266	-25224.736	-273399.487
aug-cc-pVTZ	-1792.717	-80653.093	-330577.690	-68387.479	-93924.156	-70769.200	-117789.242	-25233.579	-273479.843
aug-cc-pVQZ	-1792.930	-80659.672	-330579.854	-68391.959	-93930.811	-70774.137	-117797.457	-25235.238	-273499.262
B3LYP									
6-31G*	-1821.496	-80879.680	-331019.330	-68723.981	-94324.161	-71099.548	-118334.412	-25424.781	-274516.698
6-311G**	-1825.249	-80915.160	-331041.958	-68743.964	-94352.260	-71122.637	-118372.228	-25434.378	-274602.909
6-311+G**	-1825.249	-80921.119	-331042.376	-68746.313	-94355.781	-71124.400	-118375.865	-25434.477	-274610.898
aug-cc-pVDZ	-1822.781	-80900.777	-331037.040	-68735.679	-94342.779	-71112.599	-118355.236	-25426.099	-274559.348
aug-cc-pVTZ	-1826.294	-80924.809	-331046.214	-68753.132	-94364.680	-71130.522	-118386.090	-25437.297	-274637.827
aug-cc-pVQZ	-1826.524	-80931.623	-331048.518	-68757.801	-94371.635	-71135.774	-118394.935	-25439.304	-274658.770
PBE0									
6-31G*	-1809.350	-80814.148	-330908.497	-68646.270	-94229.030	-71019.301	-118209.965	-25386.187	-274262.907
6-311G**	-1812.744	-80847.548	-330929.992	-68664.405	-94255.008	-71040.144	-118244.691	-25394.734	-274343.483
6-311+G**	-1812.744	-80852.643	-330930.316	-68666.385	-94258.082	-71041.624	-118247.852	-25394.857	-274350.238
aug-cc-pVDZ	-1810.396	-80833.349	-330925.755	-68657.681	-94246.653	-71031.286	-118229.708	-25388.157	-274302.447
aug-cc-pVTZ	-1813.739	-80856.381	-330934.291	-68673.246	-94267.774	-71047.818	-118258.233	-25397.257	-274377.683
aug-cc-pVQZ	-1813.951	-80863.155	-330936.494	-68677.875	-94274.707	-71053.010	-118266.985	-25399.148	-274398.390
MP2									
6-31G*	-1795.966	-80710.480	-330639.004	-68555.257	-94091.742	-70918.553	-118037.159	-25308.176	-273871.136
6-311G**	-1807.398	-80777.529	-330665.917	-68581.473	-94138.308	-70952.177	-118094.545	-25337.435	-274033.817
6-311+G**	-1807.397	-80782.292	-330666.366	-68584.314	-94142.068	-70954.308	-118099.237	-25337.610	-274044.713
aug-cc-pVDZ	-1806.198	-80759.734	-330666.760	-68571.143	-94125.983	-70939.727	-118076.022	-25330.185	-273987.845
aug-cc-pVTZ	-1813.815	-80824.100	-330710.049	-68623.225	-94199.200	-70994.619	-118171.379	-25359.550	-274217.595
aug-cc-pVQZ	-1815.347	-80846.193	-330725.527	-68642.122	-94223.991	-71013.773	-118204.141	-25367.682	-274293.600

TABLE S14. EC-RISM$^{\varphi}$ results for the Gibbs free energies of all 9 gas molecules in hexafluorobenzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Amber/OPLS force field was applied for dispersion interaction. The exact molecular electronic potential was used for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	-1789.309	-80616.574	-330553.386	-68360.667	-93884.254	-70741.656	-117741.332	-25222.230	-273368.844
6-311G**	-1792.280	-80646.787	-330574.021	-68378.264	-93910.686	-70761.486	-117774.884	-25230.942	-273443.593
6-311+G**	-1792.280	-80649.347	-330574.210	-68379.606	-93913.293	-70762.464	-117777.414	-25230.999	-273447.661
aug-cc-pVDZ	-1789.650	-80630.345	-330570.475	-68371.931	-93902.079	-70752.697	-117760.236	-25225.075	-273400.518
aug-cc-pVTZ	-1793.088	-80653.514	-330578.243	-68388.185	-93924.774	-70769.929	-117790.193	-25233.908	-273480.868
aug-cc-pVQZ	-1793.300	-80660.092	-330580.410	-68392.666	-93931.429	-70774.866	-117798.403	-25235.566	-273500.283
B3LYP									
6-31G*	-1821.870	-80880.099	-331019.875	-68724.694	-94324.785	-71100.260	-118335.331	-25425.034	-274517.698
6-311G**	-1825.621	-80915.579	-331042.508	-68744.680	-94352.890	-71123.349	-118373.153	-25434.719	-274603.941
6-311+G**	-1825.621	-80921.541	-331042.936	-68747.042	-94356.422	-71125.140	-118376.829	-25434.853	-274611.969
aug-cc-pVDZ	-1823.155	-80901.209	-331037.610	-68736.416	-94343.428	-71113.338	-118356.193	-25426.516	-274560.398
aug-cc-pVTZ	-1826.667	-80925.236	-331046.776	-68753.856	-94365.320	-71131.254	-118387.033	-25437.690	-274638.870
aug-cc-pVQZ	-1826.897	-80932.048	-331049.082	-68758.525	-94372.273	-71136.506	-118395.872	-25439.696	-274659.808
PBE0									
6-31G*	-1809.723	-80814.567	-330909.042	-68646.982	-94229.651	-71020.012	-118210.880	-25386.421	-274263.897
6-311G**	-1813.116	-80847.967	-330930.541	-68665.120	-94255.635	-71040.854	-118245.612	-25395.039	-274344.504
6-311+G**	-1813.116	-80853.065	-330930.874	-68667.111	-94258.719	-71042.359	-118248.808	-25395.198	-274351.293
aug-cc-pVDZ	-1810.770	-80833.778	-330926.317	-68658.410	-94247.293	-71032.018	-118230.656	-25388.532	-274303.479
aug-cc-pVTZ	-1814.111	-80856.807	-330934.850	-68673.966	-94268.408	-71048.544	-118259.165	-25397.616	-274378.713
aug-cc-pVQZ	-1814.323	-80863.579	-330937.057	-68678.596	-94275.341	-71053.737	-118267.913	-25399.506	-274399.412
MP2									
6-31G*	-1796.337	-80710.899	-330639.548	-68555.983	-94092.370	-70919.251	-118038.050	-25308.435	-273872.142
6-311G**	-1807.768	-80777.948	-330666.466	-68582.202	-94138.941	-70952.876	-118095.434	-25337.758	-274034.846
6-311+G**	-1807.768	-80782.714	-330666.922	-68585.055	-94142.708	-70955.030	-118100.168	-25337.958	-274045.774
aug-cc-pVDZ	-1806.569	-80760.162	-330667.318	-68571.886	-94126.622	-70940.449	-118076.953	-25330.563	-273988.889
aug-cc-pVTZ	-1814.185	-80824.523	-330710.605	-68623.951	-94199.833	-70995.335	-118172.286	-25359.900	-274218.619
aug-cc-pVQZ	-1815.718	-80846.615	-330726.086	-68642.851	-94224.623	-71014.489	-118205.042	-25368.026	-274294.620

TABLE S15. EC-RISM$^{\varphi}$ results for the Gibbs free energies of all 9 gas molecules in benzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Cornell force field was applied for dispersion interaction. The exact molecular electronic potential was used for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	-1788.972	-80616.174	-330552.811	-68359.915	-93883.590	-70740.866	-117740.194	-25221.896	-273367.670
6-311G**	-1791.944	-80646.387	-330573.441	-68377.508	-93910.015	-70760.701	-117773.745	-25230.567	-273442.392
6-311+G**	-1791.944	-80648.946	-330573.627	-68378.844	-93912.618	-70761.662	-117776.249	-25230.611	-273446.441
aug-cc-pVDZ	-1789.314	-80629.941	-330569.891	-68371.175	-93901.412	-70751.899	-117759.090	-25224.665	-273399.324
aug-cc-pVTZ	-1792.752	-80653.112	-330577.659	-68387.433	-93924.108	-70769.137	-117789.068	-25233.507	-273479.679
aug-cc-pVQZ	-1792.964	-80659.691	-330579.824	-68391.914	-93930.763	-70774.075	-117797.283	-25235.166	-273499.098
B3LYP									
6-31G*	-1821.531	-80879.699	-331019.300	-68723.935	-94324.113	-71099.489	-118334.230	-25424.704	-274516.533
6-311G**	-1825.284	-80915.179	-331041.928	-68743.916	-94352.211	-71122.580	-118372.044	-25434.309	-274602.742
6-311+G**	-1825.284	-80921.138	-331042.347	-68746.267	-94355.733	-71124.342	-118375.679	-25434.411	-274610.732
aug-cc-pVDZ	-1822.815	-80900.797	-331037.011	-68735.637	-94342.734	-71112.542	-118355.061	-25426.038	-274559.186
aug-cc-pVTZ	-1826.329	-80924.829	-331046.184	-68753.089	-94364.634	-71130.466	-118385.917	-25437.232	-274637.663
aug-cc-pVQZ	-1826.559	-80931.642	-331048.488	-68757.758	-94371.589	-71135.719	-118394.762	-25439.239	-274658.607
PBE0									
6-31G*	-1809.385	-80814.167	-330908.466	-68646.224	-94228.981	-71019.242	-118209.783	-25386.108	-274262.742
6-311G**	-1812.778	-80847.567	-330929.961	-68664.357	-94254.959	-71040.087	-118244.508	-25394.662	-274343.316
6-311+G**	-1812.778	-80852.662	-330930.286	-68666.339	-94258.034	-71041.566	-118247.667	-25394.788	-274350.072
aug-cc-pVDZ	-1810.431	-80833.369	-330925.725	-68657.638	-94246.607	-71031.229	-118229.533	-25388.090	-274302.285
aug-cc-pVTZ	-1813.773	-80856.401	-330934.261	-68673.202	-94267.727	-71047.762	-118258.060	-25397.188	-274377.520
aug-cc-pVQZ	-1813.986	-80863.175	-330936.464	-68677.832	-94274.660	-71052.955	-118266.812	-25399.079	-274398.226
MP2									
6-31G*	-1796.000	-80710.499	-330638.973	-68555.209	-94091.694	-70918.508	-118037.010	-25308.099	-273870.975
6-311G**	-1807.432	-80777.548	-330665.886	-68581.424	-94138.260	-70952.133	-118094.394	-25337.364	-274033.653
6-311+G**	-1807.432	-80782.312	-330666.336	-68584.266	-94142.019	-70954.264	-118099.082	-25337.541	-274044.551
aug-cc-pVDZ	-1806.233	-80759.754	-330666.730	-68571.099	-94125.937	-70939.683	-118075.874	-25330.117	-273987.685
aug-cc-pVTZ	-1813.849	-80824.119	-330710.019	-68623.178	-94199.153	-70994.575	-118171.232	-25359.478	-274217.435
aug-cc-pVQZ	-1815.381	-80846.213	-330725.497	-68642.076	-94223.944	-71013.729	-118203.994	-25367.609	-274293.440

TABLE S16. 3D RISM results for the Gibbs free energies (excess chemical potentials) of all 9 gas molecules in benzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Amber/OPLS force field was applied for dispersion interaction. Vacuum ChelpG partial charges were evaluated for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	2.702	2.763	2.717	3.448	3.026	3.291	1.815	0.830	3.677
6-311G**	2.702	2.763	2.717	3.436	3.010	3.286	1.779	0.833	3.632
6-311+G**	2.702	2.763	2.717	3.436	3.012	3.282	1.750	0.832	3.629
aug-cc-pVDZ	2.702	2.763	2.717	3.448	3.018	3.290	1.844	0.842	3.685
aug-cc-pVTZ	2.702	2.763	2.717	3.431	3.014	3.283	1.852	0.821	3.637
aug-cc-pVQZ	2.702	2.763	2.717	3.429	3.012	3.282	1.851	0.821	3.631
B3LYP									
6-31G*	2.702	2.763	2.717	3.490	3.094	3.310	1.865	0.844	3.832
6-311G**	2.702	2.763	2.717	3.474	3.081	3.306	1.833	0.842	3.799
6-311+G**	2.702	2.763	2.717	3.474	3.081	3.300	1.796	0.841	3.802
aug-cc-pVDZ	2.702	2.763	2.717	3.488	3.085	3.310	1.898	0.853	3.853
aug-cc-pVTZ	2.702	2.763	2.717	3.468	3.081	3.303	1.913	0.836	3.811
aug-cc-pVQZ	2.702	2.763	2.717	3.466	3.078	3.302	1.912	0.836	3.805
PBE0									
6-31G*	2.702	2.763	2.717	3.485	3.077	3.308	1.857	0.843	3.789
6-311G**	2.702	2.763	2.717	3.471	3.063	3.304	1.825	0.841	3.753
6-311+G**	2.702	2.763	2.717	3.472	3.063	3.298	1.796	0.840	3.754
aug-cc-pVDZ	2.702	2.763	2.717	3.484	3.068	3.308	1.889	0.850	3.807
aug-cc-pVTZ	2.702	2.763	2.717	3.466	3.064	3.301	1.905	0.836	3.766
aug-cc-pVQZ	2.702	2.763	2.717	3.464	3.061	3.300	1.904	0.836	3.760
MP2									
6-31G*	2.702	2.763	2.717	3.528	3.141	3.319	1.888	0.838	3.839
6-311G**	2.702	2.763	2.717	3.511	3.108	3.315	1.841	0.840	3.767
6-311+G**	2.702	2.763	2.717	3.512	3.107	3.307	1.812	0.839	3.770
aug-cc-pVDZ	2.702	2.763	2.717	3.530	3.122	3.320	1.930	0.853	3.865
aug-cc-pVTZ	2.702	2.763	2.717	3.503	3.109	3.312	1.926	0.832	3.779
aug-cc-pVQZ	2.702	2.763	2.717	3.497	3.101	3.310	1.920	0.830	3.780

TABLE S17. 3D RISM results for the Gibbs free energies (excess chemical potentials) of all 9 gas molecules in hexafluorobenzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Amber/OPLS force field was applied for dispersion interaction. Vacuum ChelpG partial charges were evaluated for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	2.332	2.344	2.175	2.786	2.434	2.622	0.859	0.704	2.651
6-311G**	2.332	2.344	2.175	2.775	2.420	2.622	0.823	0.706	2.582
6-311+G**	2.332	2.344	2.175	2.776	2.422	2.616	0.785	0.705	2.571
aug-cc-pVDZ	2.332	2.344	2.175	2.786	2.427	2.623	0.898	0.715	2.649
aug-cc-pVTZ	2.332	2.344	2.175	2.771	2.424	2.619	0.917	0.699	2.605
aug-cc-pVQZ	2.332	2.344	2.175	2.769	2.422	2.619	0.918	0.698	2.602
B3LYP									
6-31G*	2.332	2.344	2.175	2.821	2.492	2.629	0.886	0.720	2.775
6-311G**	2.332	2.344	2.175	2.808	2.481	2.632	0.857	0.717	2.714
6-311+G**	2.332	2.344	2.175	2.808	2.481	2.622	0.809	0.717	2.708
aug-cc-pVDZ	2.332	2.344	2.175	2.820	2.485	2.631	0.932	0.728	2.786
aug-cc-pVTZ	2.332	2.344	2.175	2.802	2.481	2.628	0.961	0.712	2.746
aug-cc-pVQZ	2.332	2.344	2.175	2.801	2.479	2.628	0.963	0.711	2.743
PBE0									
6-31G*	2.332	2.344	2.175	2.817	2.477	2.628	0.881	0.719	2.741
6-311G**	2.332	2.344	2.175	2.806	2.465	2.631	0.852	0.716	2.678
6-311+G**	2.332	2.344	2.175	2.806	2.465	2.621	0.814	0.716	2.670
aug-cc-pVDZ	2.332	2.344	2.175	2.816	2.470	2.631	0.925	0.726	2.748
aug-cc-pVTZ	2.332	2.344	2.175	2.801	2.467	2.628	0.955	0.713	2.710
aug-cc-pVQZ	2.332	2.344	2.175	2.799	2.464	2.628	0.956	0.712	2.707
MP2									
6-31G*	2.332	2.344	2.175	2.854	2.532	2.630	0.902	0.714	2.781
6-311G**	2.332	2.344	2.175	2.839	2.504	2.635	0.856	0.715	2.689
6-311+G**	2.332	2.344	2.175	2.840	2.503	2.623	0.816	0.715	2.684
aug-cc-pVDZ	2.332	2.344	2.175	2.855	2.516	2.633	0.956	0.729	2.797
aug-cc-pVTZ	2.332	2.344	2.175	2.832	2.504	2.631	0.964	0.708	2.721
aug-cc-pVQZ	2.332	2.344	2.175	2.828	2.498	2.632	0.962	0.705	2.723

TABLE S18. 3D RISM results for the Gibbs free energies (excess chemical potentials) of all 9 gas molecules in benzene solution. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Cornell force field was applied for dispersion interaction. Vacuum ChelpG partial charges were evaluated for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	2.667	2.744	2.748	3.484	3.069	3.342	2.001	0.924	3.834
6-311G**	2.667	2.744	2.748	3.471	3.054	3.335	1.966	0.926	3.792
6-311+G**	2.667	2.744	2.748	3.472	3.056	3.332	1.941	0.926	3.790
aug-cc-pVDZ	2.668	2.744	2.748	3.483	3.062	3.341	2.026	0.936	3.842
aug-cc-pVTZ	2.668	2.744	2.748	3.466	3.058	3.333	2.032	0.916	3.795
aug-cc-pVQZ	2.668	2.744	2.748	3.464	3.055	3.331	2.030	0.915	3.789
B3LYP									
6-31G*	2.667	2.744	2.748	3.525	3.137	3.365	2.053	0.939	3.988
6-311G**	2.667	2.744	2.748	3.509	3.124	3.358	2.021	0.936	3.959
6-311+G**	2.667	2.744	2.748	3.510	3.124	3.354	1.989	0.936	3.962
aug-cc-pVDZ	2.668	2.744	2.748	3.523	3.128	3.364	2.081	0.947	4.010
aug-cc-pVTZ	2.668	2.744	2.748	3.503	3.124	3.355	2.092	0.931	3.968
aug-cc-pVQZ	2.668	2.744	2.748	3.501	3.121	3.354	2.092	0.930	3.963
PBE0									
6-31G*	2.667	2.744	2.748	3.520	3.120	3.363	2.044	0.938	3.946
6-311G**	2.667	2.744	2.748	3.507	3.106	3.357	2.013	0.935	3.913
6-311+G**	2.667	2.744	2.748	3.507	3.106	3.352	1.988	0.935	3.914
aug-cc-pVDZ	2.668	2.744	2.748	3.519	3.111	3.362	2.072	0.944	3.965
aug-cc-pVTZ	2.668	2.744	2.748	3.501	3.107	3.354	2.085	0.931	3.923
aug-cc-pVQZ	2.668	2.744	2.748	3.499	3.105	3.352	2.084	0.930	3.918
MP2									
6-31G*	2.667	2.744	2.748	3.563	3.184	3.376	2.076	0.932	3.995
6-311G**	2.667	2.744	2.748	3.546	3.151	3.369	2.031	0.934	3.927
6-311+G**	2.667	2.744	2.748	3.547	3.150	3.363	2.005	0.934	3.931
aug-cc-pVDZ	2.668	2.744	2.748	3.564	3.165	3.377	2.114	0.948	4.022
aug-cc-pVTZ	2.668	2.744	2.748	3.538	3.151	3.367	2.107	0.926	3.937
aug-cc-pVQZ	2.668	2.744	2.748	3.532	3.143	3.364	2.101	0.924	3.937

TABLE S19. EC-RISM^q results for the Gibbs free energies of all 9 gas molecules in a model for benzene solution that was parametrized without partial charges on benzene. We refer to this model as q_0-C_6H_6. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Amber/OPLS force field was applied for dispersion interaction.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
B3LYP									
6-311G**	-1825.236	-80915.106	-331041.816	-68743.762	-94352.121	-71122.352	-118371.421	-25434.369	-274602.643
aug-cc-pVTZ	-1826.267	-80924.761	-331046.079	-68752.965	-94364.563	-71130.259	-118385.381	-25437.311	-274637.592

TABLE S20. EC-RISM^q results for the Gibbs free energies of all 9 gas molecules in a model for hexafluorobenzene solution that was parametrized without partial charges on benzene. We refer to this model as q_0-C_6F_6. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Amber/OPLS force field was applied for dispersion interaction.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
B3LYP									
6-311G**	-1825.581	-80915.503	-331042.323	-68744.381	-94352.682	-71122.964	-118372.016	-25434.569	-274603.489
aug-cc-pVTZ	-1826.618	-80925.158	-331046.587	-68753.583	-94365.124	-71130.870	-118385.976	-25437.511	-274638.438

TABLE S21. EC-RISM[q] results for the Gibbs free energies of all 9 gas molecules in a model for benzene solution that has been parametrized with the partial charges of hexafluorobenzene. We refer to this model as q_{rev}-C_6H_6. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Amber/OPLS force field was applied for dispersion interaction. ChelpG partial charges were evaluated for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	-1788.995	-80616.205	-330552.844	-68359.958	-93883.665	-70740.852	-117740.790	-25221.960	-273367.989
6-311G**	-1791.966	-80646.417	-330573.477	-68377.538	-93910.082	-70760.682	-117774.347	-25230.641	-273442.750
6-311+G**	-1791.966	-80648.977	-330573.665	-68378.881	-93912.689	-70761.650	-117776.881	-25230.687	-273446.819
aug-cc-pVDZ	-1789.337	-80629.974	-330569.929	-68371.237	-93901.497	-70751.898	-117759.690	-25224.743	-273399.666
aug-cc-pVTZ	-1792.775	-80653.144	-330577.697	-68387.488	-93924.189	-70769.137	-117789.653	-25233.591	-273480.022
aug-cc-pVQZ	-1792.987	-80659.722	-330579.864	-68391.968	-93930.844	-70774.075	-117797.866	-25235.250	-273499.438
B3LYP									
6-31G*	-1821.556	-80879.730	-331019.333	-68723.987	-94324.195	-71099.462	-118334.816	-25424.758	-274516.828
6-311G**	-1825.308	-80915.209	-331041.964	-68743.956	-94352.288	-71122.547	-118372.634	-25434.401	-274603.083
6-311+G**	-1825.308	-80921.171	-331042.388	-68746.317	-94355.818	-71124.324	-118376.319	-25434.514	-274611.109
aug-cc-pVDZ	-1822.840	-80900.835	-331037.057	-68735.720	-94342.836	-71112.537	-118355.671	-25426.161	-274559.527
aug-cc-pVTZ	-1826.353	-80924.864	-331046.227	-68753.158	-94364.729	-71130.461	-118386.499	-25437.348	-274638.005
aug-cc-pVQZ	-1826.583	-80931.676	-331048.533	-68757.828	-94371.683	-71135.714	-118395.341	-25439.355	-274658.945
PBE0									
6-31G*	-1809.409	-80814.198	-330908.499	-68646.274	-94229.061	-71019.214	-118210.365	-25386.133	-274263.031
6-311G**	-1812.802	-80847.598	-330929.997	-68664.396	-94255.033	-71040.052	-118245.093	-25394.718	-274343.649
6-311+G**	-1812.802	-80852.695	-330930.327	-68666.387	-94258.115	-71041.544	-118248.291	-25394.854	-274350.436
aug-cc-pVDZ	-1810.455	-80833.405	-330925.767	-68657.715	-94246.704	-71031.219	-118230.135	-25388.171	-274302.613
aug-cc-pVTZ	-1813.797	-80856.435	-330934.302	-68673.269	-94267.819	-71047.752	-118258.635	-25397.268	-274377.848
aug-cc-pVQZ	-1814.009	-80863.208	-330936.507	-68677.899	-94274.751	-71052.947	-118267.384	-25399.159	-274398.552
MP2									
6-31G*	-1796.023	-80710.530	-330639.006	-68555.248	-94091.779	-70918.474	-118037.437	-25308.168	-273871.259
6-311G**	-1807.455	-80777.579	-330665.923	-68581.450	-94138.338	-70952.091	-118094.831	-25337.464	-274033.975
6-311+G**	-1807.455	-80782.345	-330666.379	-68584.306	-94142.103	-70954.233	-118099.576	-25337.656	-274044.907
aug-cc-pVDZ	-1806.256	-80759.792	-330666.774	-68571.174	-94126.035	-70939.668	-118076.355	-25330.242	-273988.009
aug-cc-pVTZ	-1813.872	-80824.154	-330710.060	-68624.001	-94199.245	-70994.563	-118171.694	-25359.588	-274217.747
aug-cc-pVQZ	-1815.404	-80846.246	-330725.540	-68642.135	-94224.035	-71013.719	-118204.454	-25367.715	-274293.748

TABLE S22. EC-RISM[q] results for the Gibbs free energies of all 9 gas molecules in a model for benzene solution that has been parametrized with the partial charges of hexafluorobenzene. We refer to this model as q_{rev}-C_6H_6. The corresponding geometries have been estimated through IEFPCM geometry optimization with the given level of theory. The Amber/OPLS force field was applied for dispersion interaction. ChelpG partial charges were evaluated for the solute-solvent electrostatic interactions.

	He	Ne	Ar	N_2	O_2	CO	CO_2	CH_4	CF_4
HF									
6-31G*	-1789.318	-80616.569	-330553.310	-68360.521	-93884.174	-70741.420	-117740.832	-25222.185	-273368.589
6-311G**	-1792.289	-80646.782	-330573.944	-68378.100	-93910.590	-70761.253	-117774.381	-25230.867	-273443.297
6-311+G**	-1792.289	-80649.342	-330574.132	-68379.443	-93913.197	-70762.218	-117776.897	-25230.915	-273447.356
aug-cc-pVDZ	-1789.660	-80630.339	-330570.396	-68371.800	-93902.006	-70752.468	-117759.748	-25224.974	-273400.256
aug-cc-pVTZ	-1793.098	-80653.509	-330578.165	-68388.049	-93924.697	-70769.707	-117789.724	-25233.820	-273480.596
aug-cc-pVQZ	-1793.310	-80660.087	-330580.331	-68392.529	-93931.352	-70774.645	-117797.939	-25235.479	-273500.013
B3LYP									
6-31G*	-1821.879	-80880.095	-331019.799	-68724.554	-94324.712	-71100.044	-118334.882	-25424.990	-274517.478
6-311G**	-1825.630	-80915.574	-331042.431	-68744.523	-94352.804	-71123.130	-118372.689	-25434.631	-274603.681
6-311+G**	-1825.630	-80921.536	-331042.856	-68746.885	-94356.334	-71124.903	-118376.346	-25434.747	-274611.686
aug-cc-pVDZ	-1823.163	-80901.200	-331037.525	-68736.289	-94343.354	-71113.117	-118355.742	-25426.389	-274560.152
aug-cc-pVTZ	-1826.676	-80925.229	-331046.694	-68753.725	-94365.246	-71131.041	-118386.587	-25437.574	-274638.619
aug-cc-pVQZ	-1826.906	-80932.041	-331049.000	-68758.394	-94372.199	-71136.295	-118395.430	-25439.581	-274659.560
PBE0									
6-31G*	-1809.732	-80814.562	-330908.966	-68646.842	-94229.576	-71019.796	-118210.433	-25386.385	-274263.681
6-311G**	-1813.125	-80847.962	-330930.464	-68664.963	-94255.547	-71040.636	-118245.151	-25394.965	-274344.248
6-311+G**	-1813.125	-80853.059	-330930.794	-68666.953	-94258.629	-71042.123	-118248.326	-25395.105	-274351.018
aug-cc-pVDZ	-1810.778	-80833.770	-330926.235	-68658.283	-94247.219	-71031.799	-118230.210	-25388.418	-274303.242
aug-cc-pVTZ	-1814.120	-80856.800	-330934.770	-68673.835	-94268.333	-71048.333	-118258.727	-25397.513	-274378.467
aug-cc-pVQZ	-1814.332	-80863.573	-330936.975	-68678.465	-94275.266	-71053.528	-118267.478	-25399.404	-274399.171
MP2									
6-31G*	-1796.346	-80710.894	-330639.473	-68555.811	-94092.299	-70919.082	-118037.672	-25308.387	-273871.918
6-311G**	-1807.778	-80777.943	-330666.389	-68582.012	-94138.855	-70952.698	-118095.053	-25337.670	-274034.585
6-311+G**	-1807.778	-80782.708	-330666.842	-68584.864	-94142.617	-70954.831	-118099.758	-25337.851	-274045.495
aug-cc-pVDZ	-1806.579	-80760.153	-330667.236	-68571.728	-94126.547	-70940.266	-118076.563	-25330.434	-273988.648
aug-cc-pVTZ	-1814.195	-80824.516	-330710.525	-68624.557	-94199.758	-70995.160	-118171.914	-25359.788	-274218.379
aug-cc-pVQZ	-1815.727	-80846.609	-330726.005	-68642.692	-94224.547	-71014.316	-118204.674	-25367.917	-274294.383

TABLE S23. Maxima of the spatial solvent site (γ) distribution functions $g_\gamma(\mathbf{r})$ around benzene and hexafluorobenzene. Listed below are the results of 3D RISM calculations with the Amber/OPLS force field model used for the solvent susceptibility functions and for the solute-solvent interaction. The PSE-1 closure was picked to solve the 3D RISM equations. Additionally a B3LYP level of theory with 6-31G(d) pople basis and solute-solvent electrostatic evaluations via the exact electrostatic potential was chosen for the EC-RISM calculations.

		C_6H_6	C_6F_6
Solvent model	γ	max($g_\gamma(\mathbf{r})$)	
C_6H_6	C	1.59	2.54
	H	2.30	1.58
C_6F_6	C	2.28	1.29
	F	1.84	2.83
C_6H_6 (EC-RISM)$^\varphi$	C	1.62	2.10
	H	2.04	1.57
C_6F_6 (EC-RISM)$^\varphi$	C	2.21	1.70
	F	1.89	2.12
q_{rev}-C_6H_6	C	2.27	1.60
	H	1.55	2.74
q_{rev}-C_6F_6	C	1.40	2.65
	F	2.50	1.89
q_0-C_6H_6	C	1.77	1.81
	H	1.61	1.55
q_0-C_6F_6	C	1.52	1.57
	F	1.96	1.92

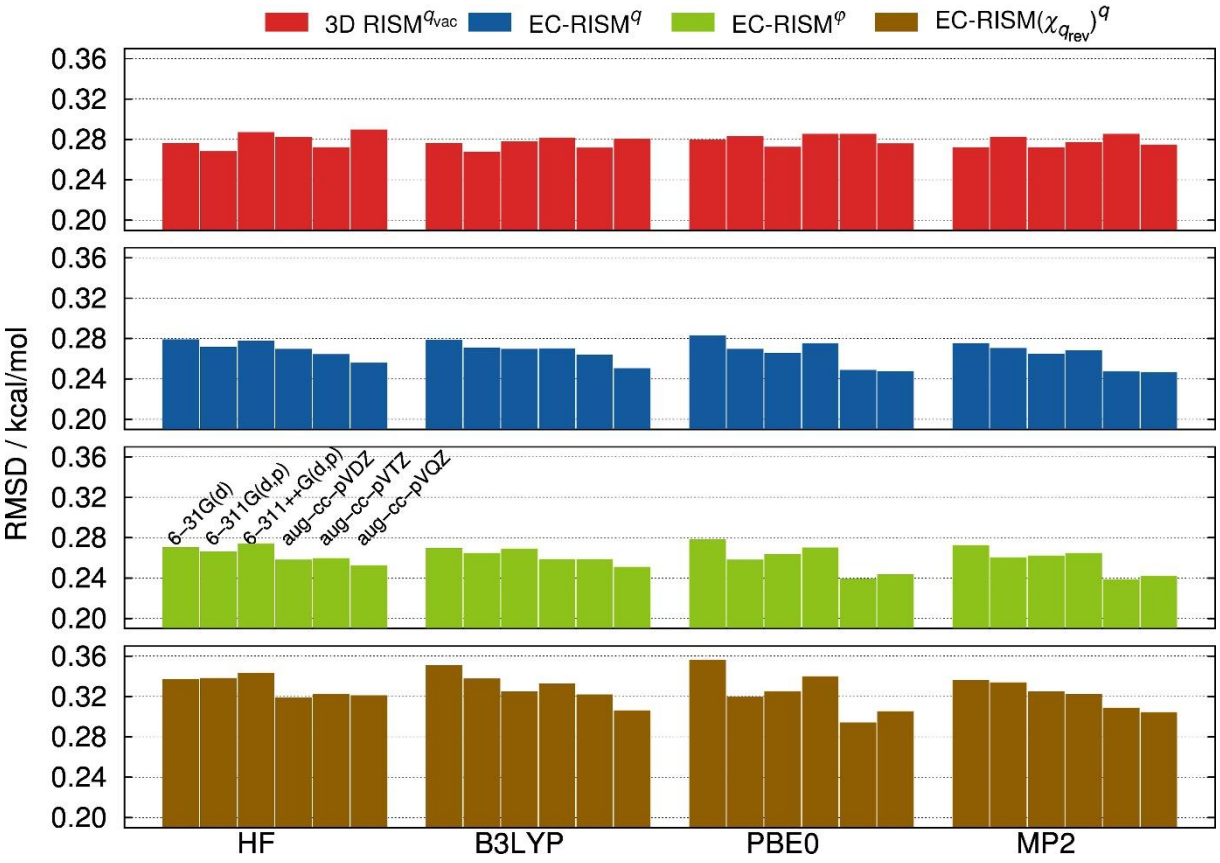


Figure. S10. Root mean square deviations from experimental values of pure 3D RISM, EC-RISM^{q} and EC-RISM$^{\varphi}$ calculations with the Cornell and Amber/OPLS(q_{rev}) solvent models averaged over all molecules and differentiated by the quantum-chemical level of theory.

TABLE S24. Root mean square deviations (RMSD) in kcal/mol to experimental values averaged over all gas molecules and divided into the different quantum-chemical levels of theory and solvation models.

RMSD	3D RISM^{qvac} AMBER/OPLS	EC_RISM^{q} AMBER/OPLS	EC_RISM$^{\varphi}$ AMBER/OPLS	3D RISM^{qvac} Cornell	EC_RISM^{q} Cornell	EC_RISM$^{\varphi}$ Cornell	EC_RISM^{q} (χ_{qrev})
HF							
6-31G*	0.241	0.243	0.227	0.276	0.279	0.270	0.337
6-311G**	0.236	0.238	0.221	0.268	0.271	0.266	0.338
6-311+G**	0.242	0.236	0.217	0.287	0.277	0.274	0.343
aug-cc-pVDZ	0.243	0.238	0.220	0.282	0.269	0.258	0.318
aug-cc-pVTZ	0.236	0.233	0.213	0.272	0.264	0.259	0.322
aug-cc-pVQZ	0.244	0.223	0.209	0.289	0.256	0.252	0.321
B3LYP							
6-31G*	0.238	0.240	0.221	0.276	0.279	0.269	0.351
6-311G**	0.236	0.238	0.220	0.267	0.271	0.264	0.337
6-311+G**	0.239	0.233	0.216	0.278	0.269	0.269	0.324
aug-cc-pVDZ	0.240	0.235	0.213	0.281	0.270	0.258	0.332
aug-cc-pVTZ	0.236	0.233	0.213	0.271	0.264	0.258	0.322
aug-cc-pVQZ	0.240	0.220	0.210	0.280	0.250	0.251	0.306
PBE0							
6-31G*	0.240	0.242	0.225	0.279	0.283	0.278	0.356
6-311G**	0.243	0.236	0.218	0.283	0.269	0.258	0.319
6-311+G**	0.236	0.231	0.214	0.272	0.265	0.263	0.325
aug-cc-pVDZ	0.242	0.237	0.217	0.285	0.275	0.270	0.340
aug-cc-pVTZ	0.244	0.226	0.209	0.285	0.248	0.239	0.294
aug-cc-pVQZ	0.238	0.222	0.207	0.276	0.247	0.243	0.305
MP2							
6-31G*	0.237	0.239	0.224	0.271	0.275	0.272	0.336
6-311G**	0.240	0.233	0.212	0.282	0.270	0.260	0.334
6-311+G**	0.235	0.231	0.213	0.272	0.265	0.262	0.324
aug-cc-pVDZ	0.239	0.235	0.216	0.277	0.268	0.265	0.322
aug-cc-pVTZ	0.242	0.221	0.203	0.285	0.247	0.239	0.309
aug-cc-pVQZ	0.237	0.222	0.206	0.274	0.247	0.242	0.304

A1.3 Solvent controlled stereoselectivity of nondipolar liquids

Table S25**.** Gibbs free energies in solution of the considered compounds calculated with EC-RISM$^{\varphi}$, the regular benzene and hexafluorobenzene solvent models and B3LYP/6-311+G(d,p) level of theory with original transition state structures.

	$G_{sol}(C_6H_6)$ / kJ mol^{-1}	$G_{sol}(C_6F_6)$ / kJ mol^{-1}
nitrostyrene(**1**)	-1350287.66	-1350298.47
β-ketoester(enole)(**2**)	-1412317.36	-1412325.48
catalyst(**3**)	-2072630.77	-2072643.12
E(**1**+**2**+**3**)	-4835235.79	-4835267.07
pro-S,R	-4835160.40	-4835186.56
pro-R,S	-4835156.75	-4835180.78
R,S-product(**4**)	-2762624.72	-2762642.73
S,R-product	-2762624.79	-2762642.81
4+**3**	-4835255.49	-4835285.85
S,R-product+**3**	-4835255.56	-4835285.93

Table S26. Gibbs free energies in solution of the considered compounds calculated with IEF-PCM solvation and B3LYP/6-311+G(d,p) level of theory with original transition state structures.

	$G_{sol}(C_6H_6)$ / kJ mol^{-1}	$G_{sol}(C_6F_6)$ / kJ mol^{-1}
nitrostyrene(**1**)	-1350300.93	-1350299.38
β-ketoester(enole)(**2**)	-1412332.93	-1412332.08
catalyst(**3**)	-2072633.78	-2072632.69
E(**1**+**2**+**3**)	-4835267.64	-4835264.15
pro-S,R	-4835213.02	-4835210.15
pro-R,S	-4835208.09	-4835205.40
R,S-product(**4**)	-2762652.61	-2762650.54
S,R-product	-2762652.56	-2762650.54
4+**3**	-4835286.39	-4835283.23
S,R-product+**3**	-4835286.34	-4835283.23

Table S27**.** Gibbs free energies in solution of the considered compounds calculated with EC-RISM^q, the uncharged q_0-C_6H_6 and q_0-C_6F_6 solvent models and B3LYP/6-311+G(d,p) level of theory with original transition state structures.

	$G_{sol}(C_6H_6)$ / kJ mol^{-1}	$G_{sol}(C_6F_6)$ / kJ mol^{-1}
nitrostyrene(**1**)	-1350278.05	-1350284.49
β-ketoester(enole)(**2**)	-1412311.14	-1412318.64
catalyst(**3**)	-2072620.43	-2072631.76
E(**1**+**2**+**3**)	-4835209.62	-4835234.89
pro-S,R	-4835133.10	-4835156.17
pro-R,S	-4835134.98	-4835157.99
R,S-product(**4**)	-2762611.40	-2762624.09
S,R-product	-2762611.40	-2762624.09
4+**3**	-4835231.83	-4835255.85
S,R-product+**3**	-4835231.83	-4835255.85

Table S28. Gibbs free energies in solution of the considered compounds calculated with EC-RISM^q, and B3LYP/6-311+G(d,p) level of theory with original transition state structures.

	$G_{sol}(C_6H_6)$ / kJ mol^{-1}	$G_{sol}(C_6F_6)$ / kJ mol^{-1}
nitrostyrene(**1**)	-1350287.28	-1350298.05
β-ketoester(enole)(**2**)	-1412317.01	-1412325.51
catalyst(**3**)	-2072628.15	-2072640.05
E(**1**+**2**+**3**)	-4835232.44	-4835263.61
pro-S,R	-4835153.76	-4835182.98
pro-R,S	-4835150.14	-4835177.33
R,S-product(**4**)	-2762623.75	-2762642.38
S,R-product	-2762623.65	-2762642.29
4+**3**	-4835251.90	-4835282.43
S,R-product+**3**	-4835251.80	-4835282.34

Table S29. Gibbs free energies in solution of the transition states calculated with EC-RISM^q, and B3LYP/6-311+G(d,p) level of theory with original transition state structures that were parametrized with the OPLS force field.

	$G_{sol}(C_6H_6)$ / kJ mol^{-1}	$G_{sol}(C_6F_6)$ / kJ mol^{-1}
pro-S,R	-4835157.91	-4835182.93
pro-R,S	-4835153.85	-4835177.61

Table S30. Gibbs free energies in solution of the transition states calculated with EC-RISM^q, and B3LYP/6-311+G(d,p) level of theory with original transition state structures that were parametrized with the 'all equal LJ-parameters'.

	$G_{sol}(C_6H_6)$ / kJ mol^{-1}	$G_{sol}(C_6F_6)$ / kJ mol^{-1}
pro-S,R	-4835155.60	-4835141.32
pro-R,S	-4835112.63	-4835131.39

Table S31. Gibbs free energies in solution of the considered compounds calculated with EC-RISM$^{\varphi}$, the regular benzene and hexafluorobenzene solvent models and M06/6-311G(d,p) level of theory with original transition state structures.

	$G_{sol}(C_6H_6)$ / kJ mol^{-1}	$G_{sol}(C_6F_6)$ / kJ mol^{-1}
nitrostyrene(**1**)	-1349321.73	-1349331.32
β-ketoester(enole)(**2**)	-1411366.82	-1411374.33
catalyst(**3**)	-2071054.58	-2071064.66
E(**1**+**2**+**3**)	-4831743.12	-4831770.30
pro-S,R	-4831762.92	-4831786.17
pro-R,S	-4831752.43	-4831773.98
R,S-product(**4**)	-2760779.53	-2760795.00
S,R-product	-2760779.72	-2760794.63
4+**3**	-4831834.10	-4831859.66
S,R-product+**3**	-4831834.30	-4831859.29

Table S32. Gibbs free energies in solution of the considered compounds calculated with EC-RISM$^{\varphi}$, the regular benzene and hexafluorobenzene solvent models and MP2/6-311+G(d,p) level of theory with original transition state structures.

	$G_{sol}(C_6H_6)$ / kJ mol^{-1}	$G_{sol}(C_6F_6)$ / kJ mol^{-1}
nitrostyrene(**1**)	-1346532.35	-1346542.26
β-ketoester(enole)(**2**)	-1408376.59	-1408384.51
catalyst(**3**)	-2066520.45	-2066532.62
E(**1**+**2**+**3**)	-4821429.39	-4821429.39
pro-S,R	-4821543.21	-4821569.49
pro-R,S	-4821527.00	-4821550.94
R,S-product(**4**)	-2755054.23	-2755071.26
S,R-product	-2755054.52	-2755071.33
4+**3**	-4821574.68	-4821603.88
S,R-product+**3**	-4821574.98	-4821603.95

Table S33. Solute electronic energy and excess chemical potential of the considered compounds calculated with EC-RISM$^{\varphi}$, the regular benzene and hexafluorobenzene solvent models and B3LYP/6-311+G(d,p) level of theory with original transition state structures.

	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$\mu^{ex}(C_6F_6)$ / kJ mol^{-1}	μ^{ex} (C_6F_6) / kJ mol^{-1}
nitrostyrene(**1**)	-1350288.99	-1350288.83	1.33	-9.63
β-ketoester(enole)(**2**)	-1412326.30	-1412326.28	8.94	0.79
catalyst(**3**)	-2072644.00	-2072644.12	13.23	1.00
E(**1**+**2**+**3**)	-4835259.29	-4835259.23	23.50	-7.84
pro-S,R	-4835192.76	-4835192.76	32.36	6.19
pro-R,S	-4835189.20	-4835189.26	32.45	8.47
R,S-product(**4**)	-2762636.54	-2762636.63	11.82	-6.10
S,R-product	-2762636.53	-2762636.62	11.74	-6.19
4+**3**	-4835280.54	-4835280.75	25.05	-5.10
S,R-product+**3**	-4835280.53	-4835280.74	24.97	-5.19

Table S34. Solute electronic energy and excess chemical potential of the considered compounds calculated with EC-RISM^q, the regular benzene and hexafluorobenzene solvent models and B3LYP/6-311+G(d,p) level of theory with original transition state structures.

	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$\mu^{ex}(C_6F_6)$ / kJ mol^{-1}	μ^{ex} (C_6F_6) / kJ mol^{-1}
nitrostyrene(**1**)	-1350289.00	-1350288.86	1.72	-9.20
β-ketoester(enole)(**2**)	-1412326.37	-1412326.37	9.36	0.85
catalyst(**3**)	-2072643.91	-2072644.03	15.77	3.99
E(**1**+**2**+**3**)	-4835259.28	-4835259.26	26.85	-4.36
pro-S,R	-4835191.91	-4835192.84	38.15	9.87
pro-R,S	-4835188.43	-4835189.23	38.29	11.90
R,S-product(**4**)	-2762636.65	-2762636.83	12.89	-5.55
S,R-product	-2762636.65	-2762636.83	13.00	-5.46
4+**3**	-4835280.56	-4835280.86	28.66	-1.56
S,R-product+**3**	-4835280.56	-4835280.86	28.77	-1.47

Table S35. Solute electronic energy and excess chemical potential of the considered compounds calculated with EC-RISM^q, the uncharged q_0-benzene and hexafluorobenzene solvent models and B3LYP/6-311+G(d,p) level of theory with original transition state structures.

	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$\mu^{ex}(C_6F_6)$ / kJ mol^{-1}	μ^{ex} (C_6F_6) / kJ mol^{-1}
nitrostyrene(**1**)	-1350287.87	-1350287.87	9.82	3.38
β-ketoester(enole)(**2**)	-1412325.91	-1412325.91	14.77	7.28
catalyst(**3**)	-2072643.15	-2072643.15	22.72	11.39
E(**1**+**2**+**3**)	-4835256.93	-4835256.93	47.31	22.05
pro-S,R	-4835190.49	-4835190.49	55.51	32.50
pro-R,S	-4835187.07	-4835187.07	53.97	30.90
R,S-product(**4**)	-2762635.13	-2762635.13	23.73	11.04
S,R-product	-2762635.13	-2762635.13	23.73	11.04
4+**3**	-4835278.28	-4835278.28	46.45	22.43
S,R-product+**3**	-4835278.28	-4835278.28	46.45	22.43

Table S36**.** Solute electronic energy and excess chemical potential of the considered compounds calculated with EC-RISM^q, the regular benzene and hexafluorobenzene solvent models and IEF-PCM/6-311+G(d,p) level of theory with original transition state structures.

	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$\mu^{ex}(C_6F_6)$ / kJ mol^{-1}	μ^{ex} (C_6F_6) / kJ mol^{-1}
nitrostyrene(**1**)	-1350286.32	-1350286.61	-14.60	-12.76
β-ketoester(enole)(**2**)	-1412325.16	-1412325.26	-7.78	-6.82
catalyst(**3**)	-2072624.63	-2072624.74	-9.16	-7.95
E(**1**+**2**+**3**)	-4835236.11	-4835236.61	-31.54	-27.53
pro-S,R	-4835187.75	-4835188.17	-25.27	-21.97
pro-R,S	-4835184.53	-4835184.91	-23.56	-20.50
R,S-product(**4**)	-2762633.41	-2762633.70	-19.20	-16.82
S,R-product	-2762633.41	-2762633.70	-19.20	-16.82
4+**3**	-4835258.04	-4835258.44	-28.36	-24.77
S,R-product+**3**	-4835258.04	-4835258.44	-28.36	-24.77

Table S37. Solute electronic energy and excess chemical potential of the considered compounds calculated with EC-RISM$^{\varphi}$, the regular benzene and hexafluorobenzene solvent models and M06/6-311G(d,p) level of theory with original transition state structures.

	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$\mu^{ex}(C_6F_6)$ / kJ mol^{-1}	μ^{ex} (C_6F_6) / kJ mol^{-1}
nitrostyrene(**1**)	-1349323.43	-1349323.26	1.70	-8.06
β-ketoester(enole)(**2**)	-1411375.90	-1411375.88	9.09	1.56
catalyst(**3**)	-2071066.78	-2071066.60	12.21	1.10
E(**1**+**2**+**3**)	-4831766.12	-4831765.74	22.99	-5.40
pro-S,R	-4831796.03	-4831795.87	33.12	9.70
pro-R,S	-4831785.34	-4831785.27	32.91	11.29
R,S-product(**4**)	-2760787.11	-2760787.10	7.58	-7.90
S,R-product	-2760787.25	-2760786.69	7.53	-7.94
4+**3**	-4831853.89	-4831853.70	19.79	-7.90
S,R-product+**3**	-4831854.04	-4831853.29	19.74	-7.90

Table S38. Solute electronic energy and excess chemical potential of the considered compounds calculated with EC-RISM$^{\varphi}$, the regular benzene and hexafluorobenzene solvent models and MP2/6-311+G(d,p) level of theory with original transition state structures.

	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$E_{sol}(C_6H_6)$ / kJ mol^{-1}	$\mu^{ex}(C_6F_6)$ / kJ mol^{-1}	μ^{ex} (C_6F_6) / kJ mol^{-1}
nitrostyrene(**1**)	-1346532,00	-1346530,70	-0,35	-11,56
β-ketoester(enole)(**2**)	-1408384,82	-1408384,64	8,23	0,13
catalyst(**3**)	-2066532,05	-2066532,18	11,60	-0,44
E(**1**+**2**+**3**)	-4821448,87	-4821447,52	19,48	-11,87
pro-S,R	-4821571,48	-4821571,62	28,27	2,13
pro-R,S	-4821555,63	-4821555,87	28,64	4,94
R,S-product(**4**)	-2755063,87	-2755062,65	9,64	-8,61
S,R-product	-2755062,65	-2755062,62	8,14	-8,71
4+**3**	-4821595,92	-4821594,83	21,23	-8,61
S,R-product+**3**	-4821594,70	-4821594,79	19,72	-8,61

Table S39. Gibbs free energies in solution of the transition states calculated with IEF-PCM, after re-optimization with B3LYP/6-311+G(d,p) level of theory.

	$G_{sol}(C_6H_6)$ / kJ mol^{-1}	$G_{sol}(C_6F_6)$ / kJ mol^{-1}
pro-S,R	-4835248.24	-4835245.14
pro-R,S	-4835243.21	-4835240.27

Table S40**.** Gibbs free energies in solution of the transition states calculated with EC-RISM$^{\varphi}$, after re-optimization with B3LYP/6-311+G(d,p) level of theory.

	$G_{sol}(C_6H_6)$ / kJ mol^{-1}	$G_{sol}(C_6F_6)$ / kJ mol^{-1}
pro-S,R	-4835196.34	-4835216.82
pro-R,S	-4835192.79	-4835222.31

Table S41. Coefficients of the effective integrated rate law calculated with EC-RISM and IEF-PCM solvation approaches in regular benzene and hexafluorobenzene solvation models with BP98/TZVP geometry optimized transition state structures.

Method	solvent	[*i*]	$x_{t\to\infty}(i)$	$x_1(i)$	$x_2(i)$	$k_{1,\mathrm{eff}}$ / (10^8 *A*)	$k_{1,\mathrm{eff}}$ / (10^8 *A*)
IEFPCM	C_6H_6	E	0.00026	0.99959	0.00015		
		R,S	0.49987	-0.87929	0.37942	0.0306814	$3.40458\cdot10^{-6}$
		S,R		-0.12030	-0.37957		
IEFPCM	C_6F_6	E	0.00023	0.99965	0.00013		
		R,S	0.49989	-0.87144	0.37155	0.0397547	$4.03751\cdot10^{-6}$
		S,R		-0.12811	-0.37168		
EC-RISM$^{\varphi}$	C_6H_6	E	0.00017	0.99976	0.00007		
		R,S	0.49913	-0.81326	0.31334	$7.62128\cdot10^{-6}$	$8.06634\cdot10^{-10}$
		S,R		-0.18650	-0.31341		
EC-RISM$^{\varphi}$	C_6F_6	E	0.00025	0.99958	0.00017		
		R,S	0.49987	-0.91111	0.41124	$8.62377\cdot10^{-7}$	$7.61797\cdot10^{-11}$
		S,R		-0.08847	-0.41141		

Table S42. Coefficients of the effective integrated rate law calculated with EC-RISM and IEF-PCM solvation approaches in regular benzene and hexafluorobenzene solvation models with re- optimized transition state structures. Therefore B3LYP/6-311+G(d,p) level of theory was applied.

Method	solvent	[*i*]	$x_{t\to\infty}(i)$	$x_1(i)$	$x_2(i)$	$k_{1,\mathrm{eff}}$ / (10^8 *A*)	$k_{1,\mathrm{eff}}$ / (10^8 *A*)
IEFPCM	C_6H_6	E	0.00026	0.99958	0.00015		
		R,S	0.49987	-0.88349	0.38362	45194.4	4.86287
		S,R		-0.11609	-0.38378		
IEFPCM	C_6F_6	E	0.00023	0.99964	0.00013		
		R,S	0.49989	-0.87675	0.37686	53300.9	5.22054
		S,R		-0.12290	-0.37699		
EC-RISM$^{\varphi}$	C_6H_6	E	0.00017	0.99976	0.00007		
		R,S	0.49913	-0.80706	0.30715	15.1973	0.00164927
		S,R		-0.19270	-0.30721		
EC-RISM$^{\varphi}$	C_6F_6	E	0.00025	0.99959	0.00016		
		R,S	0.49987	-0.90122	0.40134	1.59798	0.000143023
		S,R		-0.09837	-0.40151		

Table S43. Original (BP86/SVP) structure of the *pro-R,S* TS of Lattanzi et al.[30] and the applied GAFF force field parameter.

Atom	x / Å	y / Å	z / Å	σ / Å	ε / 10^{-21} J
O	-0.064621	0.345609	0.903188	2.960	1.459
C	0.979616	0.908534	1.299364	3.400	0.598
C	0.979351	2.194645	2.122113	3.400	0.760
C	2.465051	2.583137	2.242120	3.400	0.760
C	3.226503	1.243020	2.090682	3.400	0.760
C	2.357955	0.426405	1.137511	3.400	0.598
H	0.531897	1.950542	3.113588	2.650	0.109
H	0.322336	2.954916	1.654334	2.650	0.109
H	2.738023	3.275043	1.416854	2.650	0.109
H	2.706821	3.105231	3.189745	2.650	0.109
H	3.301145	0.718353	3.070090	2.650	0.109
H	4.270058	1.365289	1.733455	2.650	0.109
C	2.527127	-1.036502	1.000351	3.400	0.598
O	1.815753	-1.783932	0.330390	2.960	1.459
O	3.598741	-1.489943	1.715553	3.000	1.181
C	3.850559	-2.910662	1.663081	3.400	0.760
H	4.085922	-3.197084	0.614651	2.471	0.109
H	2.922522	-3.454204	1.944191	2.471	0.109
N	0.824458	1.361674	-2.004729	3.250	1.181
O	0.673338	2.495947	-1.499862	2.960	1.459
O	-0.077782	0.855992	-2.797236	2.960	1.459
C	1.947199	0.597745	-1.758082	3.400	0.598
H	1.994115	-0.342622	-2.316655	2.511	0.104
C	2.913523	1.036161	-0.820940	3.400	0.598
H	2.815417	2.106009	-0.576497	2.600	0.104
C	4.315780	0.551142	-0.930918	3.400	0.598
C	7.020579	-0.269318	-1.217827	3.400	0.598
C	5.382153	1.442071	-0.656966	3.400	0.598
C	4.636086	-0.764362	-1.350480	3.400	0.598
C	5.971770	-1.167397	-1.493212	3.400	0.598
C	6.720140	1.037823	-0.799590	3.400	0.598
H	5.152578	2.476256	-0.353106	2.600	0.104
H	3.825318	-1.479292	-1.558320	2.600	0.104
H	6.198944	-2.193140	-1.825682	2.600	0.104
H	7.532073	1.751922	-0.588599	2.600	0.104
H	8.068623	-0.588162	-1.333812	2.600	0.104
N	-1.103084	-1.010190	-1.223524	3.250	1.181
C	-2.621053	-1.066905	-1.105759	3.400	0.760
C	-2.915359	-2.579107	-1.125615	3.400	0.760
C	-1.843660	-3.149278	-2.063419	3.400	0.760
C	-0.588923	-2.361167	-1.670333	3.400	0.760
H	-3.027005	-0.591892	-2.023068	1.960	0.109
H	-3.950227	-2.793549	-1.451205	2.650	0.109
H	-2.789676	-2.976680	-0.096016	2.650	0.109

H	-1.698778	-4.242395	-1.951886	2.650	0.109
H	-2.112785	-2.954471	-3.123973	2.650	0.109
H	-0.629639	-0.723305	-0.318947	1.069	0.109
H	-0.772351	-0.252879	-1.929102	1.069	0.109
C	5.004031	-3.216999	2.603603	3.400	0.760
H	5.918223	-2.666010	2.303415	2.650	0.109
H	5.230803	-4.302796	2.585681	2.650	0.109
H	4.755094	-2.933211	3.646543	2.650	0.109
C	-3.190303	-0.329061	0.161565	3.400	0.760
H	0.127293	-2.210173	-2.500434	1.960	0.109
H	-0.039415	-2.810324	-0.822689	1.960	0.109
O	-2.498810	-0.806886	1.311414	3.066	1.462
H	-1.678824	-0.256014	1.398952	0.000	0.000
C	-4.671351	-0.748617	0.283423	3.400	0.598
C	-5.613102	-0.332412	-0.683056	3.400	0.598
C	-5.102263	-1.560781	1.350422	3.400	0.598
C	-6.955327	-0.734781	-0.592842	3.400	0.598
H	-5.299661	0.334061	-1.503169	2.600	0.104
C	-6.449307	-1.956466	1.444670	3.400	0.598
H	-4.367093	-1.867086	2.108377	2.600	0.104
C	-7.378751	-1.550920	0.472486	3.400	0.598
H	-7.678433	-0.398731	-1.353299	2.600	0.104
H	-6.773623	-2.587069	2.288390	2.600	0.104
H	-8.433118	-1.861590	0.547480	2.600	0.104
C	-3.122294	1.213475	0.113870	3.400	0.598
C	-3.663643	1.906404	1.223979	3.400	0.598
C	-2.549996	1.962112	-0.930900	3.400	0.598
C	-3.639566	3.306092	1.281932	3.400	0.598
H	-4.108944	1.328966	2.049185	2.600	0.104
C	-2.517718	3.369319	-0.869096	3.400	0.598
H	-2.062205	1.487167	-1.795238	2.600	0.104
C	-3.067250	4.046387	0.228776	3.400	0.598
H	-4.071229	3.824620	2.153373	2.600	0.104
H	-2.028923	3.924023	-1.684231	2.600	0.104
H	-3.043203	5.147181	0.272340	2.600	0.104

Table S44. Original (BP86/SVP) structure of the *pro-S,R* TS of Lattanzi et al.[30] and the applied GAFF force field parameter.

Atom	x / Å	y / Å	z / Å	σ / Å	ε / 10^{-21} J
O	-0.469729	2.391392	1.032950	2.960	1.459
C	-1.696702	2.285659	1.183945	3.400	0.598
C	-2.618040	3.462913	1.523158	3.400	0.760
C	-4.044791	2.898625	1.402006	3.400	0.760
C	-3.886457	1.381767	1.675895	3.400	0.760
C	-2.500086	1.047168	1.135302	3.400	0.598
H	-2.382060	3.769170	2.568941	2.650	0.109
H	-2.397586	4.330726	0.871053	2.650	0.109
H	-4.422742	3.064854	0.370781	2.650	0.109
H	-4.768882	3.379882	2.090070	2.650	0.109
H	-3.924516	1.172464	2.769262	2.650	0.109
H	-4.690883	0.765167	1.223273	2.650	0.109
C	-1.831338	-0.209933	1.501476	3.400	0.598
O	-0.665433	-0.530315	1.236920	2.960	1.459
O	-2.660857	-1.051943	2.176272	3.000	1.181
C	-2.106216	-2.322915	2.583414	3.400	0.760
H	-1.791239	-2.883060	1.676900	2.471	0.109
H	-1.190792	-2.141212	3.187178	2.471	0.109
N	-1.062868	2.143242	-1.993058	3.250	1.181
O	-1.742142	3.174669	-1.792309	2.960	1.459
O	0.100618	2.221451	-2.574488	2.960	1.459
C	-1.507481	0.889808	-1.624406	3.400	0.598
H	-0.830032	0.067910	-1.879595	2.511	0.104
C	-2.748609	0.750253	-0.957019	3.400	0.598
H	-3.383899	1.645902	-1.047500	2.600	0.104
C	-3.484583	-0.540457	-1.026269	3.400	0.598
C	-4.958825	-2.960270	-1.240814	3.400	0.598
C	-4.899582	-0.531181	-1.091327	3.400	0.598
C	-2.821835	-1.792764	-1.074385	3.400	0.598
C	-3.552368	-2.985533	-1.181453	3.400	0.598
C	-5.629882	-1.726612	-1.197241	3.400	0.598
H	-5.430763	0.434424	-1.085571	2.600	0.104
H	-1.722912	-1.832845	-1.023721	2.600	0.104
H	-3.017012	-3.947768	-1.224753	2.600	0.104
H	-6.729640	-1.691980	-1.253582	2.600	0.104
H	-5.528038	-3.899420	-1.328785	2.600	0.104
N	1.685728	2.082296	-0.519627	3.250	1.181
C	1.994949	3.551986	-0.315670	3.400	0.760
C	3.510600	3.611954	-0.053871	3.400	0.760
C	3.850928	2.199879	0.452191	3.400	0.760
C	2.963634	1.294304	-0.420068	3.400	0.760
H	1.384963	3.890325	0.542271	1.960	0.109
H	1.656378	4.100900	-1.215061	1.960	0.109
H	4.061059	3.823945	-0.995317	2.650	0.109

H	3.775571	4.408693	0.669299	2.650	0.109
H	3.553669	2.070756	1.514773	2.650	0.109
H	4.923482	1.944131	0.362335	2.650	0.109
H	3.397581	1.232120	-1.439980	1.960	0.109
H	0.995586	1.807076	0.231489	1.069	0.109
C	2.717172	-0.143989	0.148532	3.400	0.760
C	4.106331	-0.747475	0.468210	3.400	0.598
C	6.664607	-1.802783	1.061512	3.400	0.598
C	5.018328	-1.045802	-0.567662	3.400	0.598
C	4.486248	-0.990667	1.802412	3.400	0.598
C	5.757592	-1.516813	2.095678	3.400	0.598
C	6.289527	-1.565654	-0.273664	3.400	0.598
H	4.726267	-0.891860	-1.619491	2.600	0.104
H	3.766006	-0.765383	2.602034	2.600	0.104
H	6.039955	-1.704994	3.144256	2.600	0.104
H	6.987964	-1.796857	-1.093930	2.600	0.104
H	7.659605	-2.215461	1.292673	2.600	0.104
C	1.998292	-1.094311	-0.834044	3.400	0.598
C	0.759197	-3.002840	-2.531589	3.400	0.598
C	1.764889	-0.805882	-2.194065	3.400	0.598
C	1.597675	-2.356641	-0.337184	3.400	0.598
C	0.986683	-3.300664	-1.174060	3.400	0.598
C	1.146717	-1.752131	-3.035473	3.400	0.598
H	2.034908	0.168796	-2.625769	2.600	0.104
H	1.774751	-2.592736	0.723071	2.600	0.104
H	0.691143	-4.280704	-0.766136	2.600	0.104
H	0.964988	-1.498050	-4.091638	2.600	0.104
H	0.281113	-3.743449	-3.192067	2.600	0.104
O	1.993575	-0.012991	1.360789	3.066	1.462
H	1.017957	-0.169373	1.221949	0.000	0.000
H	1.123343	1.983607	-1.453785	1.069	0.109
C	-3.173074	-3.067564	3.368122	3.400	0.760
H	-4.074123	-3.241845	2.746097	2.650	0.109
H	-2.782418	-4.052078	3.697977	2.650	0.109
H	-3.478025	-2.497237	4.269186	2.650	0.109

Table S45.OPLS force field parameter for the *pro-R,S* and *pro-S,R* structures.

pro-R,S						*pro-S,R*					
Atom	σ/Å	ε/10^{-21} J	Atom	σ/Å	ε/10^{-21} J	Atom	σ/Å	ε/10^{-21} J	Atom	σ/Å	ε/10^{-21} J
O	2.960	1.459	H	2.500	0.208	O	2.960	1.459	H	2.500	0.208
C	3.750	0.730	H	2.500	0.208	C	3.750	0.730	H	2.500	0.208
C	3.500	0.459	H	0.000	0.000	C	3.500	0.459	H	2.500	0.208
C	3.500	0.459	H	0.000	0.000	C	3.500	0.459	H	2.500	0.208
C	3.500	0.459	C	3.500	0.459	C	3.500	0.459	H	0.000	0.000
C	3.500	0.459	H	2.500	0.208	C	3.500	0.459	C	3.500	0.459
H	2.500	0.208	H	2.500	0.208	H	2.500	0.208	C	3.550	0.486
H	2.500	0.208	H	2.500	0.208	H	2.500	0.208	C	3.550	0.486
H	2.500	0.208	C	3.500	0.459	H	2.500	0.208	C	3.550	0.486
H	2.500	0.208	H	2.500	0.208	H	2.500	0.208	C	3.550	0.486
H	2.500	0.208	H	2.500	0.208	H	2.500	0.208	C	3.550	0.486
H	2.500	0.208	O	3.120	1.181	H	2.500	0.208	C	3.550	0.486
C	3.750	0.730	H	0.000	0.000	C	3.750	0.730	H	2.420	0.208
O	2.960	1.459	C	3.550	0.486	O	2.960	1.459	H	2.420	0.208
O	2.900	0.973	C	3.550	0.486	O	2.900	0.973	H	2.420	0.208
C	3.500	0.459	C	3.550	0.486	C	3.500	0.459	H	2.420	0.208
H	2.500	0.208	C	3.550	0.486	H	2.500	0.208	H	2.420	0.208
H	2.500	0.208	H	2.420	0.208	H	2.500	0.208	C	3.550	0.486
N	3.250	0.834	C	3.550	0.486	N	3.250	0.834	C	3.550	0.486
O	2.960	1.181	H	2.420	0.208	O	2.960	1.181	C	3.550	0.486
O	2.960	1.181	C	3.550	0.486	O	2.960	1.181	C	3.550	0.486
C	3.500	0.459	H	2.420	0.208	C	3.500	0.459	C	3.550	0.486
H	2.500	0.104	H	2.420	0.208	H	2.500	0.104	C	3.550	0.486
C	3.500	0.459	H	2.420	0.208	C	3.500	0.459	H	2.420	0.208
H	2.500	0.208	C	3.550	0.486	H	2.500	0.208	H	2.420	0.208
C	3.550	0.486	C	3.550	0.486	C	3.550	0.486	H	2.420	0.208
C	3.550	0.486	C	3.550	0.486	C	3.550	0.486	H	2.420	0.208
C	3.550	0.486	C	3.550	0.486	C	3.550	0.486	H	2.420	0.208
C	3.550	0.486	H	2.420	0.208	C	3.550	0.486	O	3.120	1.181
C	3.550	0.486	C	3.550	0.486	C	3.550	0.486	H	0.000	0.000
C	3.550	0.486	H	2.420	0.208	C	3.550	0.486	H	0.000	0.000
H	2.420	0.209	C	3.550	0.486	H	2.420	0.209	C	3.500	0.459
H	2.420	0.209	H	2.420	0.208	H	2.420	0.209	H	2.500	0.208
H	2.420	0.209	H	2.420	0.208	H	2.420	0.209	H	2.500	0.208
H	2.420	0.209	H	2.420	0.208	H	2.420	0.209	H	2.500	0.208
H	2.420	0.209				H	2.420	0.209			
N	3.250	1.181				N	3.250	1.181			
C	3.500	0.459				C	3.500	0.459			
C	3.500	0.459				C	3.500	0.459			
C	3.500	0.459				C	3.500	0.459			
C	3.500	0.459				C	3.500	0.459			
H	2.500	0.208				H	2.500	0.208			
H	2.500	0.208				H	2.500	0.208			
H	2.500	0.208				H	2.500	0.208			

A1.4 Additional references

[a1] W. M. Litchman, A. Mohammed Jr., A. E. Florin, *J. Am. Chem. Soc.* **1969**, *91*, 6574.

[a2] R. E. Hoffman, *Magn. Reson. Chem.* **2006**, *44*, 606.

[a3] V. Thanabal, D. O. Omencinsky, M. D. Reily, W. L. Cody, Solution. *J. Biomol. NMR* **1994**, *4*, 47.

Eidesstattliche Versicherung (Affidavit)

_____ _____

Name, Vorname
(Surname, first name)

Matrikel-Nr.
(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

_____ _____

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

_____

_____

_____

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.
Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

_____ _____

Ort, Datum
(Place, date)

Unterschrift
(Signature)