NRC·CNRC

# NRC Publications Archive
# Archives des publications du CNRC

## Understanding the liquid states of cyclic hydrocarbons containing N, O, and S atoms via the 3D-RISM-KH molecular solvation theory

Roy, Dipankar; Kovalenko, Andriy

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.
For the publisher's version, please access the DOI link below./ Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

**Publisher's version / Version de l'éditeur:**
https://doi.org/10.3390/molecules27196563
*Molecules, 27, 19, pp. 1-13, 2022-10-04*

**NRC Publications Archive Record / Notice des Archives des publications du CNRC :**
https://nrc-publications.canada.ca/eng/view/object/?id=99c77d59-3a6d-44a9-8099-d4180c517ce9
https://publications-cnrc.canada.ca/fra/voir/objet/?id=99c77d59-3a6d-44a9-8099-d4180c517ce9

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at
https://nrc-publications.canada.ca/eng/copyright
READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site
https://publications-cnrc.canada.ca/fra/droits
LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

**Questions?** Contact the NRC Publications Archive team at PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

**Vous avez des questions?** Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

National Research Council Canada | Conseil national de recherches Canada

Canada

molecules

MDPI

*Article*

# Understanding the Liquid States of Cyclic Hydrocarbons Containing N, O, and S Atoms via the 3D-RISM-KH Molecular Solvation Theory

**Dipankar Roy [1] and Andriy Kovalenko [1,2,*]**

[1] Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2E9, Canada
[2] Nanotechnology Research Centre, National Research Council of Canada, 11421 Saskatchewan Drive, Edmonton, AB T6G 2M9, Canada
* Correspondence: andriy.kovalenko@ualberta.ca

**Abstract:** The 3D-reference interaction site model (3D-RISM) molecular solvation theory in combination with the Kovalenko–Hirata (KH) closure is extended to seven heterocyclic liquids to understand their liquid states and to test the performance of the theory in solvation free energy (SFE) calculations of solutes in select solvents. The computed solvent site distribution profiles were compared with the all-atom molecular dynamics (MD) simulations, showing comparable performances. The computational results were compared against the structural parameters for liquids, whenever available, as well as against the experimental SFEs. The liquids are found to have local ordered structures held together via weak interactions in both the RISM and MD simulations. The 3D-RISM-KH computed SFEs are in good agreement with the benchmark values for the tetrahydrothiophene-S,S-dioxide, and showed comparatively larger deviations in the case of the SFEs in the tetrahydrofuran continuum.

**Keywords:** heterocyclic compounds; molecular solvation theory; 3D-RISM-KH; force field; partial atomic charges; solvation free energy

check for updates

**Citation:** Roy, D.; Kovalenko, A. Understanding the Liquid States of Cyclic Hydrocarbons Containing N, O, and S Atoms via the 3D-RISM-KH Molecular Solvation Theory. *Molecules* **2022**, *27*, 6563. https://doi.org/10.3390/molecules27196563

Academic Editor: Benedito José Costa Cabral

Received: 3 September 2022
Accepted: 27 September 2022
Published: 4 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

Molecular solvation is essential to life processes. While a qualitative description of the solvation process is reported in numerous reports, quantitative applications are small in number, in comparison to the former. The difficulty is in the theoretical formulations of how solvents behave around a solute, providing a dielectric medium via polarity, as well as interacting directly with the solute(s) [1–5]. Explicit solvation models capture the dielectric effect of the solvent around a solute and showed to be effective in predicting and explaining chemical (re)activity [4–8]. The quantum chemical solvation models based on atomic surface charge models, i.e., COSMO, PCM, and SMD are a good example of this approach [7–12]. The inclusion of explicit solvent molecules, in combination with a continuum solvation model, also known as the cluster–continuum model, was used to account for explicit interactions between solvent and solute molecules [13–16]. Such models are very useful but limited to smaller sizes due to the large computational cost associated with electronic structure calculations. The molecular dynamic simulations use explicit solvent molecules defined via standard force field parameters to account for solute–solvent interactions. The statistical mechanics-based molecular solvation theory via the reference interaction site model (RISM, 1D, and 3D) is another prominent solvation model with vast applications in chemical, biological, and material science applications. The RISM formalism is rooted in the works of Chandler, Anderson, and coworkers [17–22]. The RISM provides probability density $\rho_\gamma g_\gamma(r)$ of finding interaction site $\gamma$ of solvent molecules at position $r$ around a solute molecule of arbitrary shape. The $\rho_\gamma$ and $g_\gamma$ are average number density in the bulk and normalized density distribution, respectively. The 3D-total correlation

function, $h_\gamma(\boldsymbol{r})$, is connected to the 3D-direct correlation function, $c_\alpha(\boldsymbol{r})$, via the site–site bulk solvent susceptibility $\chi_{\alpha\gamma}(r)$, for solvent sites $\alpha$ around the solute as:

$$h_\gamma(\boldsymbol{r}) = \sum_\alpha \int d\boldsymbol{r}' c_\alpha(\boldsymbol{r}-\boldsymbol{r}')\chi_{\alpha\gamma}(r')$$

The susceptibility $\chi_{\alpha\gamma}$ is obtained from the site density and intramolecular correlation function, $\omega_{\alpha\gamma}(r)$, from the dielectrically consistent RISM (DRISM) equation [23]: $\chi_{\alpha\gamma}(r) = \omega_{\alpha\gamma}(r) + \omega_{\alpha\gamma}(r)\,\rho_\gamma h_{\alpha\gamma}(r)$.

A closure relation is required for integrating an infinite chain of interactions involving intra- and intermolecular sites in order to get the distribution functions. The closure form proposed by Kovalenko and Hirata (KH) has the following functional form [24]:

$$g_\gamma(\boldsymbol{r}) = \begin{cases} \exp(-u_\gamma(\boldsymbol{r})/(k_BT) + h_\gamma(\boldsymbol{r}) - c_\gamma(\boldsymbol{r})) & \text{for} \quad g_\gamma(\boldsymbol{r}) \le 1 \\ 1 - u_\gamma(\boldsymbol{r})/(k_BT) + h_\gamma(\boldsymbol{r}) - c_\gamma(\boldsymbol{r}) & \text{for} \quad g_\gamma(\boldsymbol{r}) > 1 \end{cases}$$

The KH closure is known for providing numerical stability and easier convergence in computing 3D distribution functions for liquids of different polarities. The excess chemical potential ($\mu$) is obtained from a close analytical form:

$$\mu_{\text{solv}} = \sum_\gamma \int_V d\boldsymbol{r}\, \Phi_\gamma(\boldsymbol{r})$$

and

$$\Phi_\gamma(\boldsymbol{r}) = \rho_\gamma k_B T \left[\frac{1}{2}h_\gamma^2(\boldsymbol{r})\Theta(-h_\gamma(\boldsymbol{r})) - c_\gamma(\boldsymbol{r}) - \frac{1}{2}h_\gamma(\boldsymbol{r})c_\gamma(\boldsymbol{r})\right]$$

where $\Theta(x)$ is the Heaviside step function. The solvation free energy functional is written in the 3D-RISM-KH formalism via Gaussian fluctuation (GF) free energy:

$$\Delta G_{\text{solv}}^{\text{GF}} = k_BT \sum_\gamma \rho_\gamma \int_{R^3} [-c_\gamma(r) - \frac{1}{2}c_\gamma(r)h_\gamma(r)]\, dr$$

Interested readers are referred to Refs. [18,25–27] for the detailed theoretical derivation of the RISM theory and correlation functions mentioned here. The errors in 3D-RISM computations are results of erroneous calculation of internal pressure as well as inadequacies of the force field parameters used. Several correction schemes were developed to scale RISM-computed $\Delta G^{GF}$ to correctly correlate with solvation free energy. The most used "universal correction" scheme uses a linear regression scheme using RISM-computed partial molar volume (PMV) for scaling and has the following form [25]: $\Delta G_{\text{corrected}} = \Delta G^{\text{GF}} + a \cdot \text{PMV} + b$, and the parameters $a$ and $b$ are obtained from regression analysis.

The 3D-RISM-KH has been used successfully to model solvents of varying polarities, e.g., water, hydrocarbons, octanol, amines, ketones, etc., to name a few. For a review of the applications of the RISM-KH theory, please see Refs. [26–28]. The benchmarking of this theory to a new set of solvents depends on the quality of the data available in the literature. For example, diffraction data on the liquid form are the most suitable to judge the accuracy of RISM calculations. Solvation free energy (SFE) of solutes provides an additional test for the RISM models. However, these benchmark data are not often available in the literature. In the present work, we tested the performance of the 3D-RISM-KH molecular solvation theory in modeling the liquid state of seven heterocyclic ring systems containing N, O, and/or S atoms. The molecules we chose were: Pyrrole, Pyrrolidine, Furan, Tetrahydrofuran (THF), Thiophene, Tetrahydothiophene (thiolane), and Tetrahydrothiophene-S,S-dioxide (sulfolane) (Figure 1). The modeling of ring structures, especially the ones containing heteroatoms is often problematic due to inadequate force field parameters. We tested the well-used generalized AMBER force field (GAFF) and the universal force field (UFF) parameters in this study [29–31]. Additionally, we have compared the 3D-RISM-KH computed SFEs of 14 solutes in THF and sulfolane solvents against those reported in the MNSol

database [32]. The 3D-RISM-KH computed solvent profiles are compared against those computed using the molecular dynamics (MD) simulations, as well as against experimental data, whenever available. The present work is part of an ongoing effort to develop reliable solvent models for applications using the 3D-RISM molecular solvation theory.

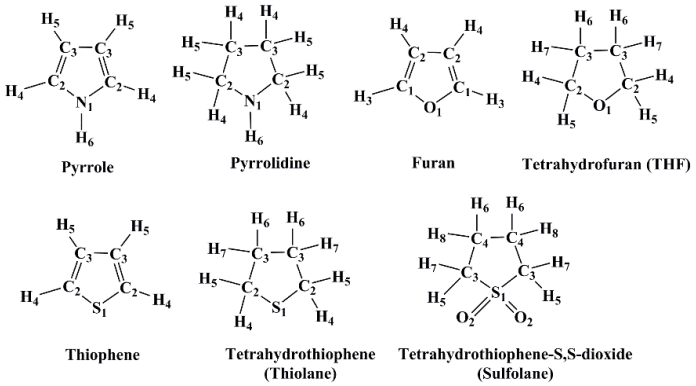



**Figure 1.** Cyclic molecules studied using the 3D-RISM-KH molecular solvation theory in this work. Equivalent solvent sites are marked with numbers. The highest number refers to the number of solvent sites present in each molecule.

## 2. Results

In this work, we compared and contrasted the 3D-RISM-KH and MD computed liquid state structures of seven heterocyclic liquids against available experimental data. For the RISM calculations, GAFF and UFF force field parameters were used. We experienced convergence issues for the solvent susceptibility calculations of the pure unsaturated rings (pyrrolidine, tetrahydrothiophene, tetrahydrofuran, and sulfolane) using the GAFF force field parameters with AM1-BCC charges. Such issues were not noticed for the aromatic rings. The UFF parameters in combination with the MP2/cc-pVTZ level computed atomic charges worked well for all seven systems. In this manuscript, all the reported pure liquid pairwise distribution functions (PDFs) are computed from the UFF-parameter-based solvent models. For the GAFF-computed parameters of aromatic heterocycles, please refer to the Supplementary Materials (Figures S1–S3). We observed noticeable differences between the GAFF- and UFF-computed distributions for the three liquid aromatic heterocycles. Further work is underway to compare the effect of force field parameters in aromatic systems via the 3D-RISM-KH theory. The GAFF parameters used for the all-atom MD simulation resulted in equilibrated systems. The physical parameters of the seven liquids and those obtained from the MD simulations are provided in Table 1. The structures of the individual liquids are discussed in the following sections. The solvation free energy (SFE) data in the two liquids are provided in the final section.

**Table 1.** Density (standard, and MD-computed) and dielectric constants used in this work for seven liquids.

| Molecule | Dielectric Constant | Density (Experimental, g·cm$^{-3}$) | Density (MD, g·cm$^{-3}$) [1] |
|---|---|---|---|
| Pyrrole | 8.00 | 0.968 [33] | 1.009 |
| Furan | 2.94 | 0.973 [34] | 0.964 |
| Thiophene | 2.73 | 1.065 [33] | 1.033 |
| Pyrrolidine | 8.30 | 0.852 [34] | 0.892 |
| Tetrahydrofuran (THF) | 7.42 | 0.883 [33] | 0.899 |
| Tetrahydrothiophene | 8.61 | 0.998 [35] | 0.977 |
| Sulfolane | 43.96 | 1.2606 [34] | 1.298 |

[1] Computed using the GAFF force field with AM1-BCC charges.

### *2.1. Liquid Structure of Pyrrole*

The aromatic pyrrole molecule is an important synthetic block in organic and medicinal chemistry. The molecule is known to form clusters from gas phase and solid phase studies [36–38]. The experimental results of the liquid structure of pyrrole are sparse [39,40]. The liquid clusters were reported to have [N]-H···π interactions forming T-shaped dimers from the molecular simulations as well as vibrational spectroscopy [41,42]. The N-N radial distribution function (rdf) showed multiple minima in the MD and RISM computations. The MD rdf has the first peak at ~3.4 Å and the second broad one at ~6.4 Å (Figure 2). The RISM-computed PDF has a broad peak with a shoulder in the region ~3.8–5.2 Å. The multiple peaks in the MD profile were reported previously [43,44]. The rdf of N-$C_2$ has one sharp peak at 2.2 Å and two nearby peaks at 3.4 Å and 5.5 Å. The first one is the intramolecular separation, which is not calculated in the RISM formalism. The RISM-computed PDFs for the same have a broad peak with multiple shoulders in the rebroad region of 3.9–6.2 Å. The N-H intermolecular distribution has maxima at 2.6 and 3.8 Å, and a broad region between ~5–7 Å. These findings are similar to those reported by Barone and coworkers [44] and were indicative of T-shaped H-π interactions. The RISM-computed PDFs for all the N-H intermolecular distributions have multiple peaks with a multishoulder broad peak between ~3–7 Å, indicative of the existence of different clusters in the liquid form.

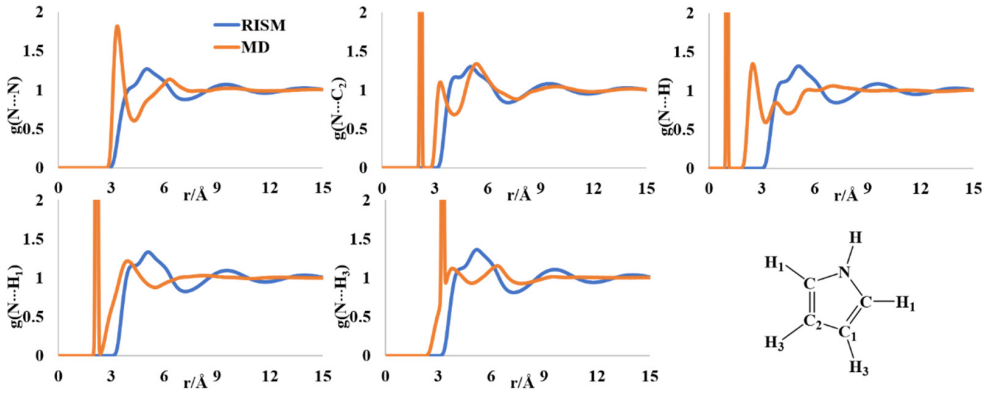



**Figure 2.** RISM- and MD-computed distribution functions for liquid pyrrole.

### *2.2. Liquid Structure of Furan*

Similar to pyrrole, experimental data on the structure of furan are not present in the literature, except for the solid phase and energy dispersive X-ray diffraction of the liquid [40,45]. The latter, however, provided structural factors processed with a sharpening factor, making it impossible to make a direct comparison. The authors in that work [45] pointed to the presence of the C-H···π, C-H···O, and H···H interaction in a complex long-range structure of the liquid.

The O-O and O-$C_2$ intermolecular distributions computed by the MD and RISM formalisms are comparable (Figure 3). For the O-O distribution function, the first peak with a shoulder appears at 4.6–5.8 Å region, followed by a broad peak at ~10 Å. The O-$C_2$ distribution has the first broad maxima in the region of 3.7–5.5 Å, for both the MD and RISM calculations, pointing to probable π-stacking interactions. The O-H (and $H_2$) distribution functions differ in shape between MD and RISM. The MD computed RDF for O-H intermolecular distribution has a sharp peak at 3.42 Å. This peak is replaced in the RISM-computed PDF by a broad peak in the region ~4–6.5 Å. This broad region is present in the PDF of the O-$H_2$ intermolecular distribution. The RISM computations are indicative of weak O···H($H_2$) interactions in the liquid state that could contribute to ordered structure.

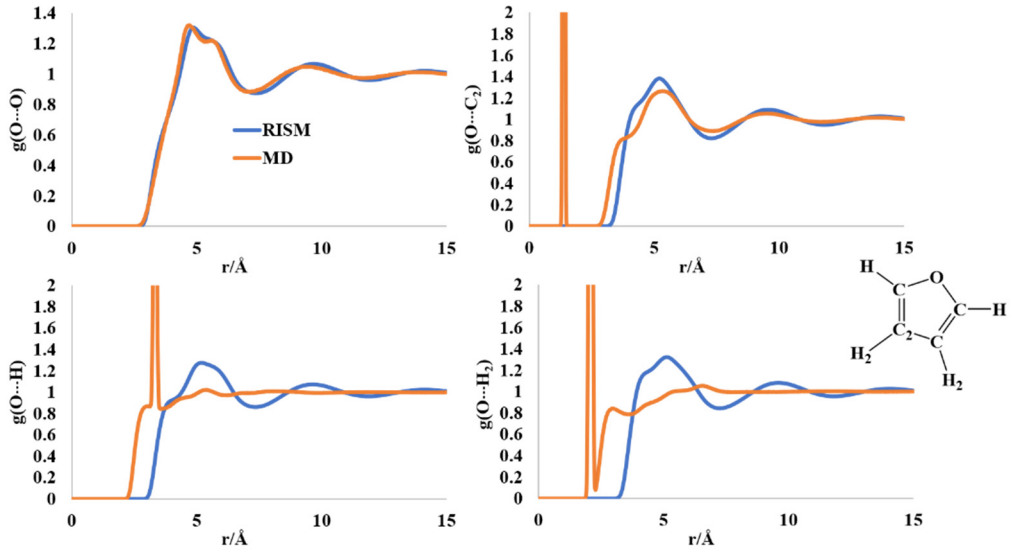



**Figure 3.** RISM- and MD-computed distribution functions for liquid furan.

### *2.3. Liquid Structure of Thiophene*

The S-S intermolecular PDF has peaks around 4.1 and 5.8 Å for thiophene. The MD computed rdf has a shoulder around 4 Å and a shouldered peak at ~5.5–6.3 Å (Figure 4). The two sharp peaks at ~2.1 Å and ~3.3 Å in the $C(C_2)$-$H_2(H)$ rdfs are intramolecular in nature, much as the two sharp peaks in the C-C2 rdf at 1.32 Å and ~2.3 Å. The C-H PDFs and rdfs did not provide sufficient information on the orderliness in the liquid thiophene structure. The first peak of the C-$C_2$ PDF was found at 3.5 Å, followed by a multishoulder hump in the region of 4.5–8 Å, which is present in the MD profile, too. The orderliness in the furan molecule seemed to be less than that of those in furran and pyrrole, although some ordered structures cannot be ruled out from the rdfs and PDFs, as reported in previous works [40,45].

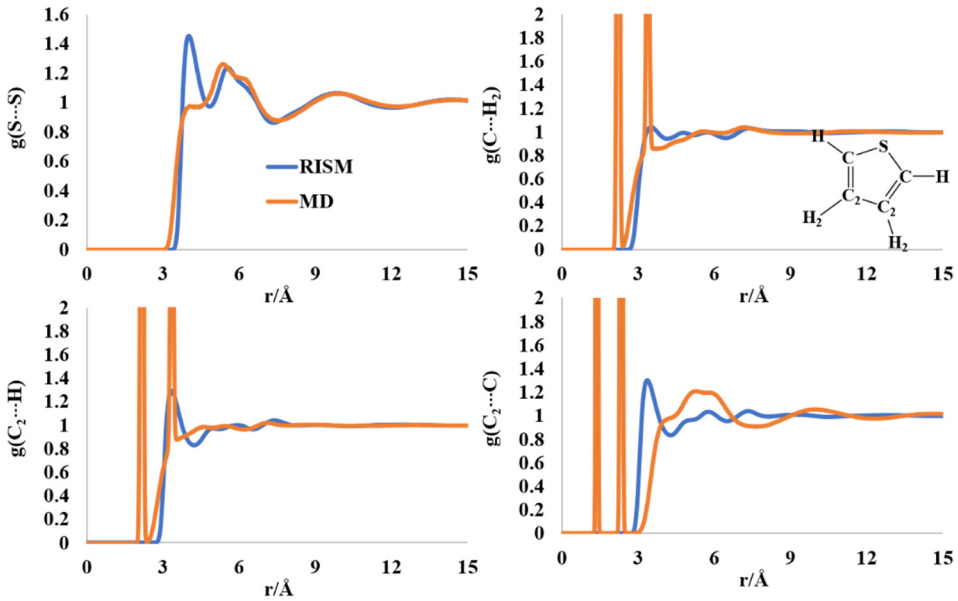



**Figure 4.** RISM- and MD-computed distribution functions for liquid thiophene.

### *2.4. Liquid Structure of Pyrrolidine*

The ring puckering of pyrrolidine received a lot of attention, and gas phase diffraction and theoretical modeling reports are present on the puckered ring of pyrrolidine [46–49],

although complete structural descriptions of the liquid state are absent in the chemical literature. The N-N rdf has a sharp peak at 3 Å and a broad peak in the region of 4.7–8 Å (Figure 5). The sharp peak is absent in the RISM PDF, though the broad peak is present in the aforementioned region. The N-C2 intermolecular distribution showed multiple minima in the region of 3.8–6 Å, with a broad region after 9 Å in both the MD and RISM computations. The intermolecular N-H distribution has two maxima at 2.8 and 3.9 Å in the MD profile. The first maximum in the RISM PDF appeared at ~3.4 Å. The position of the maximum in the intermolecular N-H(ring) distributions is comparable between the RISM and MD profiles. The presence of at least two dimeric forms was reported from spectroscopic analysis [50]. The intermolecular atomic distribution profiles computed by MD and RISM support this observation.

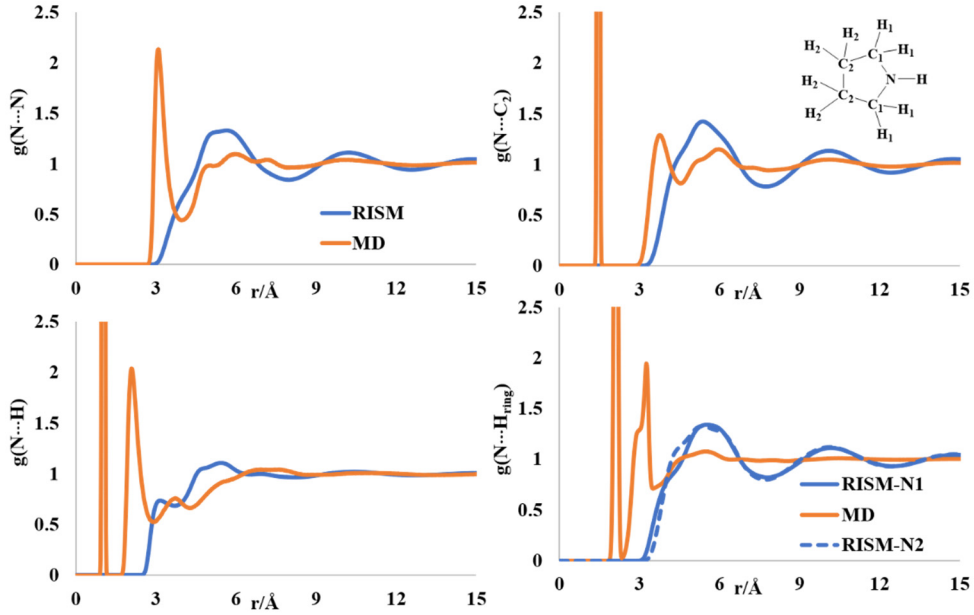



**Figure 5.** RISM- and MD-computed distribution functions for liquid pyrrolidine.

### *2.5. Liquid Structure of Tetrahydrofuran (THF)*

The structure of THF has received attention due to a flexible ring structure, like pyrrolidine, and discrepancies between the computed and experimental structure parameters in the liquid form [51–54]. The first maxima (two peaks) in the intermolecular O-O distribution are found at 4.6 Å (5.6 Å) and 5.31 Å (6.0 Å), for the MD and RISM calculations, respectively (Figure 6). The experimentally determined maximum for this distribution was reported at 5.3 Å (6.3 Å) [53]. The O-C intermolecular distribution has multipeak broad maxima in the region of ~4–6.4 Å from MD and RISM. The experimental distributions had two peaks at 4.2 Å and 6.8 Å. The O-C2 intermolecular peaks are at 3.8 Å and 6 Å for MD, while RISM had a broad peak with a shoulder at 5.6 Å. The experimentally reported maxima were at 4.2 Å and 7.9 Å [53]. Finally, the C-C2 intermolecular distribution peaks were obtained at 4.3 Å and 5.5 Å for the MD simulations, and at 3.4 Å and 5.5 Å for the RISM computations. The experimental maxima for this intermolecular separation were at 4.4 Å and 7.4 Å. The differences between the experimental rdfs and the theoretically computed ones were reported previously [54,55].

### *2.6. Liquid Structure of Tetrhydrothiophene (Thiolane)*

The RISM- and MD-computed intermolecular distribution profiles have a qualitative similar structure for liquid thiolane, although the positions of the maxima varied. The intermolecular S-S distribution has maxima at 4.2 Å and 5.7 Å in the RISM profile, and at 5.4 Å for the MD simulations (Figure 7). These peak positions are longer than the experimentally reported one at 2.65 Å [56]. The multiple maxima in the broad peak in

the region of 3.9 Å-7 Å for the S-C($C_2$) intermolecular distribution can be interpreted as possible ordered structures in the liquid form. The C-$C_2$ intermolecular distribution profile also points to possible cluster form(s).

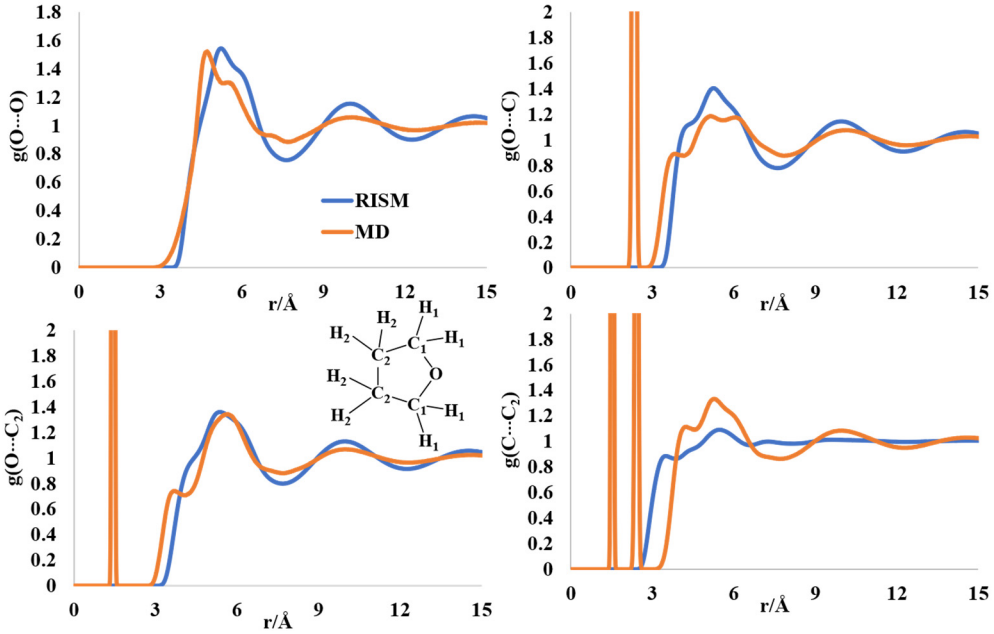



**Figure 6.** The RISM- and MD-computed distribution functions for liquid THF.

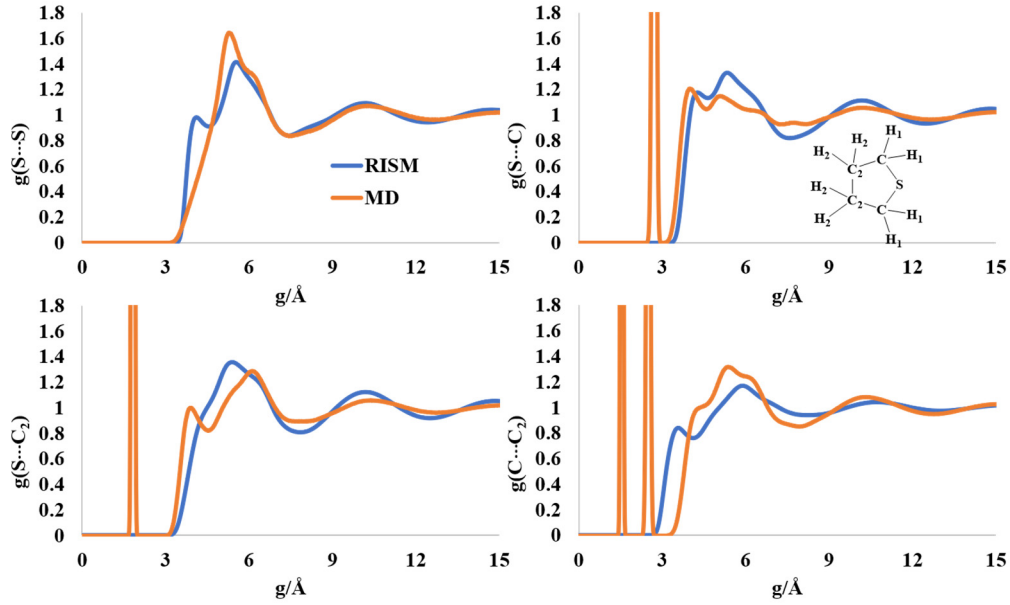



**Figure 7.** RISM- and MD-computed distribution functions for liquid thiolane.

### *2.7. Liquid Structure of Tetrahydrothiophene-S,S-dioxide*

The sulfolane molecule (and the class of compounds) have extensive applications in the field of electrochemistry and battery technology. The planarity of the ring and puckerings were investigated using diffraction and spectrochemical analysis in the solid and liquid phases [57–62]. Despite these reports, a full description of the liquid state as well as unambiguous ring puckering is absent in the literature. The interatomic S-S distribution has maxima at ~6 Å and 10.5 Å in both the RISM and MD profiles. The atomic charges computed at the MP2/cc-pVTZ level for the RISM calculations reflected the electronic

distribution differently than the AM1-BCC charges, especially for H-atoms. The AM1-BCC charges have two distinct sets of hydrogens, those attached to the carbon next to the S atom and those connected to the carbon atom(s) farthest from the S atom. The MP2/cc-pVTZ level charges divided the hydrogens into four different groups based on the envelope ring structure, where C-atoms adjacent to S are connected to two different types of hydrogens, and the two C-atoms away from S are also attached to two different types of H-atoms, solely based on atomic charges (Figure 8). The first two maxima in the intermolecular O···H($H_2$) distribution are found at 2.64 Å (2.9 Å in RISM) and 3.44 Å (3.5 Å in RISM) via the MD simulation. These findings are in qualitative agreement with the previous report on forcefield development for the sulfolane molecule [62]. Overall, the liquid sulfolane molecules are held together by O···H[C] weak hydrogen bonds.

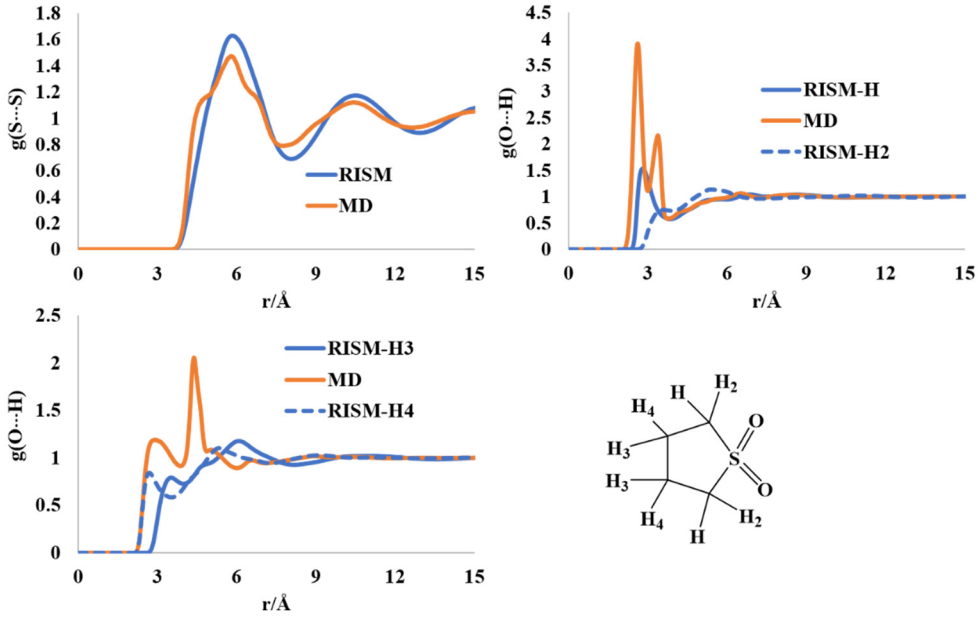



**Figure 8.** RISM- and MD-computed distribution functions for liquid sulfolane.

### *2.8. SFEs in Tetrahydrofuran (THF) and Tetrahydrothiophene-S,S-dioxide (Thiolane) Solvents*

The fourteen solutes (seven for each solvent) for which experimental solvation free energies were available in THF and sulfolane medium were computed using the 3D-RISM-KH theory. The solutes are parameterized using the UFF parameters with AM1 charges. The Gaussian fluctuation excess chemical potentials computed via the 3D-RISM-KH theory were corrected via the "universal correction" scheme. The fitting coefficients are provided in Table 2.

**Table 2.** Regression coefficients (a and b) used for the universal correction fitting of the solvation free energies obtained via the 3D-RISM-KH theory.

| Solvent | a (kcal/mol/Å$^3$) | b (kcal/mol) |
|---|---|---|
| THF | −0.1139 | −10.5505 |
| Sulfolane | −0.1715 | −7.1919 |

The corrected SFEs are provided in Table 3. The small number of experimental data is indeed a handicap in the calibration process. The computed SFEs in THF have a mean unsigned error (MAE) of 4.3 kcal/mol) and a relative mean squared error (RMSE) of 6.2 kcal/mol. The MAE and RMSE for the solvation free energies in sulfolane are 0.63 and 0.92 kcal/mol, respectively. The large RMSE values in the THF solvent point to further improvements required in the force field parameters for the solvent molecule, as well as a need for a larger benchmark dataset.

**Table 3.** Experimental SFEs ($\Delta G_{expt}$), 3D-RISM-KH computed excess chemical potential ($\mu^{GF}$), partial molar volume (PMV), and corrected SFEs ($\Delta G_{cor}$).

| Solute | Solvent | $\Delta G_{expt}$ (kcal/mol) | $\mu^{GF}$ (kcal/mol) | PMV (Å$^3$) | $\Delta G_{cor}$ (kcal/mol) |
|---|---|---|---|---|---|
| n-octane | THF | −5.39 | 44.13 | 352.55 | −6.57 |
| toluene | THF | −5.5 | 28.46 | 235.24 | −8.88 |
| ethanol | THF | −4.56 | 21.21 | 154.02 | −6.88 |
| tetrahydrofuran | THF | −4.25 | 24.21 | 192.77 | −8.29 |
| 1,4-dioxane | THF | −5.17 | 44.06 | 207.79 | 9.84 |
| 2-butanone | THF | −4.54 | 27.72 | 212.46 | −7.03 |
| nitromethane | THF | −5.09 | 21.21 | 152.51 | −6.71 |
| n-octane | sulfolane | −2.44 | 53.41 | 281.17 | −2.02 |
| toluene | sulfolane | −4.23 | 33.32 | 188.49 | −6.21 |
| ethanol | sulfolane | −4.3 | 22.90 | 114.90 | −4.00 |
| 1,4-dioxane | sulfolane | −4.9 | 30.16 | 163.54 | −5.09 |
| 2-butanone | sulfolane | −4.09 | 31.08 | 163.25 | −4.12 |
| butylamine | sulfolane | −4.25 | 35.56 | 182.73 | −2.98 |
| nitromethane | sulfolane | −5.28 | 20.98 | 110.01 | −5.09 |

## 3. Discussion

In this manuscript, we have explored the liquid states of seven heterocycles (aromatic and anon-aromatic) via molecular simulations. The 3D-RISM-KH theory and MD simulations are used for these all-atom simulations of the liquid state. The intermolecular radial distribution functions obtained from the MD simulations are qualitatively similar to the distributions computed via the RISM formalism, although some discrepancies are noted in the intermolecular distribution function involving hydrogen atoms. It is not possible to compare the accuracy of the two theoretical models, as there is not sufficient experimental data available to compare and contrast. However, one can summarize the local molecular cluster in these liquids held together by H···π, H···X(X=N/O) interactions for N- and O-containing systems. While the clustering in the thiophene and tetrahydrothiophene is less prominent than in the other liquids, one cannot rule out the effect of the dispersion effects, which are not treated explicitly in either of the approaches.

The 3D-RISM-KH computations of the SFEs in the two solvents provided encouraging results. The THF solvation model requires further improvements for use in the SFE applications, though the distribution profiles are in general in agreement with experimental findings.

The differences between the MD and RISM calculations of the solvent site distributions have some contributions from the differences in the force field parameters employed. Development of force field parameters for saturated ring systems is a non-trivial task. We have first-hand experience of difficulties in obtaining physically meaningful solvent susceptibility functions using available force field parameters for cyclic molecules [63,64]. The performances of the UFF parameters with the MP2/cc-pVTZ level atomic charges for RISM-KH provided reasonable results, in comparison to available experimental results. We noticed that the GAFF-computed RISM distributions have peak positions different than those from the UFF parameters. The PDFs for pyrrole from the RISM calculations are an interesting observation, as they have similar distribution profiles for all heavy atoms. The results may simply be that the force field in combination with RISM did not work for pyrrole. The N-containing heterocycles have been known to have issues with the GAFF parameters [65].

In this manuscript, solvent models of seven heterocyclic compounds were developed within the framework of the 3D-RISM-KH theory. These solvent models provided good structural and physical features when compared with the molecular dynamics simulations and available experimental results. These solvent models can be used for further applications in (bio)chemical research, and also as fragments for binding site mapping on biological receptor surfaces, using the molecular solvation theory.

## 4. Materials and Methods

The RISM-KH calculations were completed using our in-house code, a working version of which is implemented in the AMBERTOOLS software package [66]. The RISM-KH calculations were performed using the all-atom GAFF and UFF force field parameters. The former is combined with the AM1-BCC charges [29,30]. The UFF force field is combined with the atomic charges computed at the MP2/cc-pVTZ level on the B3LYP/cc-pVTZ level optimized geometries [31,66–71]. The electronic structure calculations were performed using the Gaussian16 suite of quantum chemical software [72]. For solvent susceptibility calculations, the extended-RISM (X-RISM) formalism was employed. The modified-DIIS convergence criterion (MDIIS) is set to 1E-10 for all the calculations. The reliability converged solvent susceptibility functions were verified using computed compressibility of the liquid. The equivalent solvent sites for pure liquids are numbered in Figure 1.

The benchmark solvation free energies in THF and sulfolane were obtained from the Minnesota solvation database (MNSol 2009) and consist of 14 datapoints [32]. For the 3D-RISM-KH calculations, we have used the geometry obtained from this database. The 3D-RISM-KH computed Gaussian fluctuation excess chemical potentials are corrected using multiple linear regressions. We have used the UFF force field parameters with semiempirical AM1 atomic charges for solute molecules [73]. The 3D-RISM-KH calculations on the solutes were performed on a uniform cubic 3D grid of 128 × 128 × 128 points in a box of size 64 × 64 × 64 $Å^3$. For a sample workflow for a RISM-KH-based molecular simulation, please see reference [27].

All the molecular dynamics (MD) simulations were performed using the GROMACS-2020.4 engine with the GAFF force field parameters and AM1-BCC charges for all the molecules in the liquid state [74]. A cubic simulation box consisting of 256 molecules was used for each system without any restraints. Energy minimized systems were subjected to 20 ns of NVT and NPT equilibration. The temperature and density profiles were checked from the resultant simulations to confirm equilibrated systems for further production simulations of 50 ns. All the trajectory analyses were performed using the modules provided in the GROMACS software package.

**Supplementary Materials:** The following are available online at https://www.mdpi.com/article/10.3390/molecules27196563/s1, **Table S1.** B3LYP/cc-pVTZ level optimized geometries of seven heterocycles studied: Pyrrole, Pyrrolidine, Furan, Tetrahydrofuran, Thiophene, Tetrahydrothiophene, Tetrahydrothiophene-S,S-dioxide, **Figure S1.** GAFF and UFF force field computed pairwise distribution functions of liquid pyrrole via the 3D-RISM-KH theory, Figure S2. GAFF and UFF force field computed pairwise distribution functions of liquid furan via the 3D-RISM-KH theory, **Figure S3.** GAFF and UFF force field computed pairwise distribution functions of liquid thiophene via the 3D-RISM-KH theory.

**Author Contributions:** Conceptualization, D.R. and A.K.; methodology, D.R. and A.K.; software, D.R. and A.K.; formal analysis, D.R.; writing, D.R. and A.K.; funding acquisition, A.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was financially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RES0029477), and the Canadian Consortium on Neurodegeneration in Aging (CCNA) grant (RES0051206).

**Data Availability Statement:** All pertinent data related to this paper are provided in the Supplementary Materials.

**Acknowledgments:** Generous computing time from the Digital Research Alliance of Canada (https://alliancecan.ca/en) is acknowledged.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References


1. Floris, F.; Tomasi, J. Evaluation of the dispersion contribution to the solvation energy. A simple computational model in the continuum approximation. *J. Comput. Chem.* **1989**, *10*, 616–627. [CrossRef]
2. Roux, B.; Simonson, T. Implicit solvent models. *Biophys. Chem.* **1999**, *78*, 1–20. [CrossRef]
3. Warshel, A.; Sharma, P.K.; Kato, M.; Parsonb, W.W. Modeling electrostatic effects in proteins. *Biochim. Biophys. Acta Prot. Proteom.* **2006**, *1764*, 1647–1676. [CrossRef] [PubMed]
4. Wagoner, J.A.; Baker, N.A. Assessing implicit models for nonpolar mean solvation forces: The importance of dispersion and volume terms. *Proc. Nat. Acad. Sci. USA* **2006**, *103*, 8331–8336. [CrossRef]
5. Schaefer, M.; Karplus, M. A Comprehensive Analytical Treatment of Continuum Electrostatics. *J. Phys. Chem.* **1996**, *100*, 1578–1599. [CrossRef]
6. Lamm, G. The Poisson-Boltzmann Equation. In *Reviews in Computational Chemistry*; Lipkowitz, K.B., Larter, R., Cundari, T.R., Boyd, D.B., Eds.; Willey: Hoboken, NJ, USA, 2003; Volume 19, pp. 147–366.
7. Tomasi, J. Selected features of the polarizable continuum model for the representation of solvation. *WIREs Comp. Mol. Sci.* **2011**, *1*, 855–867. [CrossRef]
8. Herbert, J.M. Dielectric continuum methods for quantum chemistry. *WIREs Comp. Mol. Sci.* **2021**, *11*, e1519. [CrossRef]
9. Klamt, A.; Schüürmann, G. COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans.* **1993**, *2*, 799–805. [CrossRef]
10. Klamt, A.; Moya, C.; Palomar, J. A Comprehensive Comparison of the IEFPCM and SS(V)PE Continuum Solvation Methods with the COSMO Approach. *J. Chem. Theory Comput.* **2015**, *11*, 4220–4225. [CrossRef]
11. Mennucci, B. Polarizable continuum model. *WIREs Comp. Mol. Sci.* **2012**, *2*, 386–404. [CrossRef]
12. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [CrossRef]
13. Pliego, J.R., Jr.; Riveros, J.M. Hybrid discrete-continuum solvation methods. *WIREs Comp. Mol. Sci.* **2020**, *10*, e1440. [CrossRef]
14. Pliego, J.R.; Riveros, J.M. The Cluster−Continuum Model for the Calculation of the Solvation Free Energy of Ionic Species. *J. Phys. Chem. A* **2001**, *105*, 7241–7247. [CrossRef]
15. Ho, J.; Coote, M.L. A universal approach for continuum solvent pKa calculations: Are we there yet? *Theor. Chem. Acc.* **2009**, *125*, 3–21. [CrossRef]
16. Pliego, J.R., Jr.; Riveros, J.M. Theoretical Calculation of p$K_a$ Using the Cluster-Continuum Model. *J. Phys. Chem. A* **2002**, *106*, 7434–7439. [CrossRef]
17. Chandler, D.; Andersen, H.C. Optimized cluster expansions for classical fluids. II. Theory of molecular liquids. *J. Chem. Phys.* **1972**, *57*, 1930. [CrossRef]
18. Chandler, D. Derivation of an integral equation for pair correlation functions in molecular fluids. *J. Chem. Phys.* **1973**, *59*, 2742. [CrossRef]
19. Chandler, D.; McCoy, J.D.; Singer, S.J. Singer, Density functional theory of nonuniform polyatomic systems. I. General formulation. *J. Chem. Phys.* **1985**, *85*, 5971. [CrossRef]
20. Chandler, D.; McCoy, J.D.; Singer, S.J. Density functional theory of nonuniform polyatomic systems. II. Rational closures for integral equations. *J. Chem. Phys.* **1986**, *85*, 5977. [CrossRef]
21. Chandler, D. Equilibrium structure and molecular motion in liquids. *Acc. Chem. Res.* **1974**, *7*, 246–251. [CrossRef]
22. Chandler, D.; Weeks, J.D.; Anderson, H.C. Van der Waals Picture of Liquids, Solids, and Phase Transformations. *Science* **1983**, *220*, 787–794. [CrossRef]
23. Perkyns, J.S.; Pettitt, B.M. A site-site theory for finite concentration saline solutions. *J. Chem. Phys.* **1992**, *97*, 7656–7666. [CrossRef]
24. Kovalenko, A.; Hirata, F. A molecular theory of liquid interfaces. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1785–1793. [CrossRef]
25. Palmer, D.S.; Frolov, A.; Ratkova, E.; Fedorov, M.V. Towards a universal method for calculating hydration free energies: A 3D reference interaction site model with partial molar volume correction. *J. Phys. Condens. Matt.* **2010**, *22*, 492101. [CrossRef]
26. Kovalenko, A. Multiscale Modeling of Solvation. In *Springer Handbook of Electro-Chemical Energy*; Breitkopf, C., Swider-Lyons, K., Eds.; Springer: Berlin/Heidelberg, Germany, 2017.
27. Roy, D.; Kovalenko, A. 3D-RISM-KH Molecular Solvation Theory. In *Multiscale Dynamics Simulations: Nano and Nano-Bio Systems in Complex Environments*; Salahub, D.R., Wei, D., Eds.; RSC Publishing: London, UK, 2021; pp. 254–286.
28. Roy, D.; Kovalenko, A. Extension of the approximate 3D-RISM-KH molecular solvation theory to liquid aniline and pyridines. *Res. Chem.* **2022**, *4*, 100365. [CrossRef]
29. Wang, J.; Wang, W.; Kollman, P.A.; Case, D.A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247260. [CrossRef]
30. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general AMBER force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [CrossRef] [PubMed]
31. Rappe, A.K.; Casewit, C.J.; Colwell, K.S.; Goddard, W.A.; Skiff, W.M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035. [CrossRef]
32. Marenich, A.V.; Kelly, C.P.; Thompson, J.D.; Hawkins, G.D.; Chambers, C.C.; Giesen, D.J.; Winget, P.; Cramer, C.J.; Truhlar, D.G. *Minnesota Solvation Database (MNSOL) Version 2012*; University of Minnesota: Minneapolis, MN, USA, 2020. [CrossRef]
33. Lide, D.R. (Ed.) *CRC Handbook of Chemistry and Physics*, 76th ed.; CRC Press Inc.: Boca Raton, FL, USA, 1995; pp. 3–474.

34. O'Neil, M.J. (Ed.) *The Merck Index—An Encyclopedia of Chemicals, Drugs, and Biologicals*, 13th ed.; Merck and Co., Inc.: Whitehouse Station, NJ, USA, 2006; pp. 738–1597.
35. Weast, R.C. (Ed.) *Handbook of Chemistry and Physics*, 60th ed.; CRC Press Inc.: Boca Raton, FL, USA, 1979; p. C-516.
36. Gomez-Zavaglia, A.; Fausto, R. Self-Aggregation in Pyrrole: Matrix Isolation, Solid State Infrared Spectroscopy, and DFT Study. *J. Phys. Chem. A* **2004**, *108*, 6953–6967. [CrossRef]
37. Goddard, R.; Heinemann, O.; Kruger, C. Pyrrole and a Co-crystal of 1H- and 2H-1,2,3-Triazole. *Acta Crystallogr.* **1997**, *53*, 1846–1850. [CrossRef]
38. Wilcox, W.S.; Goldstein, J.H. Evidence for a Completely Planar Structure of Pyrrole from Its Microwave Spectrum. *J. Chem. Phys.* **1952**, *20*, 1656. [CrossRef]
39. Gontrani, L.; Ramondo, F.; Caminiti, R. Energy dispersive X-ray diffraction and molecular dynamics meet: The structure of liquid pyrrole. *Chem. Phys. Lett.* **2006**, *417*, 200–205. [CrossRef]
40. Schomaker, V.; Pauling, L. The Electron Diffraction Investigation of the Structure of Benzene, Pyridine, Pyrazine, Butadiene-1,3, Cyclopentadiene, Fur an, Pyrrole, and Thiophene. *J. Am. Chem. Soc.* **1939**, *61*, 1769–1780. [CrossRef]
41. Gamba, Z.; Klein, M.L. Short-range structure of liquid pyrrole. *J. Chem. Phys.* **1990**, *92*, 697–698. [CrossRef]
42. Lautie, A.; Novak, A. Effect of Hydrogen Bonding on the NH Stretching Raman Band of Pyrrole. *J. Chem. Phys.* **1972**, *56*, 2479. [CrossRef]
43. McDonald, N.A.; Jorgensen, W.L. Development of an All-Atom Force Field for Heterocycles. Properties of Liquid Pyrrole, Furan, Diazoles, and Oxazoles. *J. Phys. Chem. B* **1998**, *102*, 8049–8059. [CrossRef]
44. Macchiagodena, M.; Mancini, G.; Pagliai, M.; Cardini, G.; Barone, V. New atomistic model of pyrrole with improved liquid state properties and structure. *Int. J. Quantum Chem.* **2018**, *118*, e25554. [CrossRef]
45. Gontrani, L.; Ramondo, F.; Caminiti, R. Furan and thiophene in liquid phase: An X-ray and molecular dynamics study. *Chem. Phys. Lett.* **2006**, *422*, 256–261. [CrossRef]
46. Pfafferott, G.; Oberhammer, H.; Boggs, J.E.; Caminati, W. Geometric Structure and Pseudorational Potential of Pyrrolidine. An ab Initio and Electron Diffraction Study. *J. Am. Chem. Soc.* **1985**, *107*, 2305–2309. [CrossRef]
47. Han, S.J.; Kang, Y.K. Pseudorotation in heterocyclic five-membered rings: Tetrahydrofuran and pyrrolidine. *J. Mol. Struct. THEOCHEM* **1996**, *369*, 157–165. [CrossRef]
48. Stortz, C.A.; Sarotti, A.M. Exhaustive exploration of the conformational landscape of mono- and disubstituted five-membered rings by DFT and MP2 calculations. *RSC Adv.* **2019**, *9*, 24134–24145. [CrossRef] [PubMed]
49. Denis, P.A. Coupled cluster investigation of the axial and equatorial isomers of pyrrolidine. *Mol. Phys.* **2012**, *111*, 9–17. [CrossRef]
50. Hesse, S.; Wassermann, T.N.; Suhm, M.A. Brightening and Locking a Weak and Floppy N−H Chromophore: The Case of Pyrrolidine. *J. Phys. Chem. A* **2010**, *114*, 10492–10499. [CrossRef] [PubMed]
51. Chandrasekhar, J.; Jorgensen, W.L. Monte Carlo simulations of liquid tetrahydrofuran including pseudorotation. *J. Chem. Phys.* **1982**, *77*, 5073. [CrossRef]
52. Girard, S.; Müller-Plathe, F. Molecular dynamics simulation of liquid tetrahydrofuran: On the uniqueness of force fields. *Mol. Phys.* **2003**, *106*, 779–787. [CrossRef]
53. Bowron, D.T.; Finney, J.L.; Soper, A.K. The Structure of Liquid Tetrahydrofuran. *J. Am. Chem. Soc.* **2006**, *128*, 5119–5126. [CrossRef]
54. Houteghem, M.V.; Ghysels, A.; Verstraelen, T.; Poelmans, W.; Waroquier, M.; Speybroeck, V.V. Critical Analysis of the Accuracy of Models Predicting or Extracting Liquid Structure Information. *J. Phys. Chem. B* **2014**, *118*, 2451–2470. [CrossRef]
55. Bedard-Hearn, M.J.; Larsen, R.E.; Schwartz, B.J. Understanding Nonequilibrium Solute and Solvent Motions through Molecular Projections: Computer Simulations of Solvation Dynamics in Liquid Tetrahydrofuran (THF). *J. Phys. Chem. B* **2003**, *107*, 14464–14475. [CrossRef]
56. Johnsson, M.; Persson, I. The structure of silver(I) iodide and bromide, and the silver(I) solvate in tetrahydrothiophene solution: An X-ray scattering and raman spectroscopic study. *Inorg. Chim. Acta* **1987**, *130*, 215–220. [CrossRef]
57. Naumov, V.A.; Semashko, V.N.; Shaidulin, S.A. Electron-diffraction investigation of the structure of the tetramethylene sulfone molecule. *J. Struct. Chem.* **1974**, *14*, 555–559. [CrossRef]
58. Combet, J.; Bee, M.; Djurado, D.; D'Assenza, G.; Commandeur, G. Structural and dynamical aspects of solid sulpholan in the temperature range 160–300 K I.—X-ray powder diffraction study. *Phys. B Condens. Matt.* **1997**, *233*, 95–101. [CrossRef]
59. Alonso, J.L. Microwave spectrum of sulfolane. *J. Mol. Struct.* **1981**, *73*, 71–78. [CrossRef]
60. Hargittai, I. On the ring planarity of sulpholane. *J. Mol. Struct.* **1981**, *77*, 313–314. [CrossRef]
61. Peemoeller, H.; Sharp, A.R.; Watton, A.; Kydon, D.W. Nuclear magnetic relaxation study of fully deuterated sulfolan. *J. Chem. Phys.* **1983**, *78*, 4337. [CrossRef]
62. Mukherji, S.; Avula, N.V.S.; Balasubramanian, S. Refined Force Field for Liquid Sulfolane with Particular Emphasis to Its Transport Characteristics. *ACS Omega* **2020**, *5*, 28285–28295. [CrossRef] [PubMed]
63. Hlushak, S.; Kovalenko, A. Effective Interactions and Adsorption of Heterocyclic Aromatic Hydrocarbons in Kaolinite Organic Solutions Studied by 3D-RISM-KH Molecular Theory of Solvation. *J. Phys. Chem. C* **2017**, *121*, 22092–22104. [CrossRef]
64. Luchko, T.; Blinov, N.; Limon, G.C.; Joyce, K.P.; Kovalenko, A. SAMPL5: 3D-RISM partition coefficient calculations with partial molar volume corrections and solute conformational sampling. *J. Comput. Aided Mol. Des.* **2016**, *30*, 1115–1127. [CrossRef]
65. Boyd, N.J.; Wilson, M.R. Optimization of the GAFF force field to describe liquid crystal molecules: The path to a dramatic improvement in transition temperature predictions. *Phys. Chem. Chem. Phys.* **2015**, *17*, 24851–24865. [CrossRef]

66. Case, D.A.; Aktulga, H.M.; Belfon, K.; Ben-Shalom, I.Y.; Berryman, J.T.; Brozell, S.R.; Cerutti, D.S.; Cheatham, T.E., III; Cisneros, G.A.; Cruzeiro, V.W.D.; et al. *Amber 2022*; University of California, San Francisco: San Francisco, CA, USA, 2022.
67. Frisch, M.J.; Head-Gordon, M.; Pople, J.A. Direct MP2 gradient method. *Chem. Phys. Lett.* **1990**, *166*, 275–280. [CrossRef]
68. Head-Gordon, M.; Head-Gordon, T. Analytic MP2 Frequencies Without Fifth Order Storage: Theory and Application to Bifurcated Hydrogen Bonds in the Water Hexamer. *Chem. Phys. Lett.* **1994**, *220*, 122–128. [CrossRef]
69. Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. [CrossRef]
70. Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789. [CrossRef]
71. Dunning, T.H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023. [CrossRef]
72. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 16, Revision B.01*, Gaussian, Inc.: Wallingford, CT, USA, 2016.
73. Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909. [CrossRef]
74. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19–25. [CrossRef]