Molecular solvation theory studies of liquid oleyl alcohol and molecular partitioning in water-oleyl alcohol mixture

Roy, Dipankar; Kovalenko, Andriy

Publisher's version / Version de l'éditeur:
https://doi.org/10.1016/j.cplett.2021.138726
Chemical Physics Letters, 777, 16 August 2021, pp. 1-5, 2021-05-14

Molecular Solvation Theory Studies of Liquid Oleyl Alcohol and Molecular Partitioning in Water-Oleyl Alcohol Mixture

Dipankar Roy[1], Andriy Kovalenko[1,2,3,*]

[1]Department of Mechanical Engineering, University of Alberta, 10-203 Donadeo Innovation Centre for Engineering, 9211-116 Street NW, Edmonton, Alberta T6G 1H9, Canada

[2]Department of Biological Sciences, University of Alberta, CW 405, Biological Sciences Bldg., Edmonton, Alberta T6G 2E9, Canada

[3]Nanotechnology Research Centre, National Research Council of Canada, 11421 Saskatchewan Drive, Edmonton, Alberta T6G 2M9, Canada

e-Mail: andriy.kovalenko@ualberta.ca

Keywords

3D-Reference Interaction Site Model, liquid oleyl alcohol, partial distribution function. Molecular simulation. partition function, quantitative structure activity relation.

Abstract

The 3D-RISM-KH molecular solvation theory is applied to liquid long chain unsaturated oleyl alcohol to explore the structure of the liquid and to calculate alcohol/water partition coefficients. The united atom TraPPE force field parameters resulted in comparable structural information between RISM and MD simulations. The RISM theory and the MD simulation point to a compact liquid structure of oleyl alcohol. The molecular partition coefficients were computed with a root mean square deviation of 0.48 unit via a predictive quantitative structure activity relationship.

Introduction

Oleyl alcohol (*cis*-9-Octadecen-1-ol, CAS no. 143-28-2) is an unsaturated long chain alcohol (C-18 chain) and a clear liquid mostly used in cosmetic applications. Oleyl alcohol is used in liquid-liquid extraction processes, and to make liquid membranes [1,2]. A detailed structure of pure liquid form of this long chain unsaturated alcohol is not reported till date, except phase behavior in mixed solvent environments and in lipid membrane simulations [3,4]. We have applied the 3-dimensional reference interaction site model (3D-RISM) molecular solvation theory with the Kovalenko-Hirata (KH) closure relation to understand the structure of pure oleyl alcohol. In the absence of experimental structural data on the liquid state, we have compared our findings with those obtained from molecular dynamics (MD) simulations. Further, we have used excess chemical potentials of 14 solutes in oleyl alcohol and in water to build a quantitative model for predicting molecular partitioning between oleyl alcohol and water media.

The 3D-RISM theory was founded by Chandler and coworkers [5,6]. The mathematical derivations underlying the theory were reported elsewhere [7-9]. Briefly, the total correlation function h_γ of solvent sites in 3D is related to the 3D direct correlation function c_γ and site-site bulk susceptibility function χ for solvent site γ around a solute of arbitrary shape at position **r** as:

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

A mathematical function, a closure relation, is required to integrate the infinite interaction diagrams generated from the 3D-RISM equation. One of the most used closure relations is the Kovalenko-Hirata closure, with successful applications in different domains of molecular

simulation techniques [10]. The excess chemical potential μ, and thus the solvation free energy, is obtained from the analytical form of the KH closure via the form:

$$\mu_{\text{solv}} = \sum_{\gamma} \int_{V} d\boldsymbol{r}\, \Phi_{\gamma}(\boldsymbol{r}) \text{ and } \Phi_{\gamma}(\boldsymbol{r}) = \rho_{\gamma} k_B T \left[\frac{1}{2} h_{\gamma}^2(\boldsymbol{r}) \Theta(-h_{\gamma}(\boldsymbol{r})) - c_{\gamma}(\boldsymbol{r}) - \frac{1}{2} h_{\gamma}(\boldsymbol{r}) c_{\gamma}(\boldsymbol{r})\right]$$

The $\Theta_{\gamma}(r)$ is the Heaviside step function. Important thermodynamic parameters are derived from the excess chemical potential for solute sites *u* and solvent sites *v*, at a temperature *T*, as:

$$\Delta\mu = \Delta\varepsilon_{uv} + \Delta\varepsilon_{vv} - T\Delta S_{\text{V}}$$

One of the drawbacks of the 3D-RISM theory is that the calculated internal pressures are wrong, and hence the absolute number of excess chemical potential showed significant deviations from experimental solvation free energies, although the trends are reproduced correctly [11]. Several correction schemes were developed to make 3D-RISM computed solvation free energies close to absolute solvation free energy. However, in this manuscript, we have used the Gaussian fluctuation (GF) excess chemical potentials (μ^{GF}) to calculate partition coefficients, without any correction schemes.

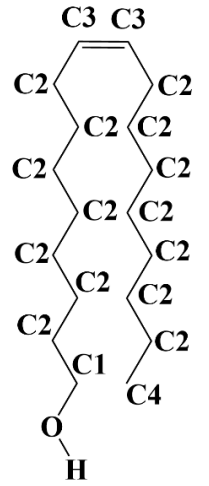
Computational Methods and Database

The oleyl alcohol geometry was optimized at the B3LYP/cc-pVTZ level of theory using the Gaussian16 quantum chemical software [12-15]. The optimized structure was verified as true minimum by using the vibrational analysis. The geometry and atomic charges computed at this level is used for RISM calculations. The solvent susceptibility function of pure oleyl alcohol is computed using the extended-RISM (X-RISM) formalism at temperature 298.15 K. The van der Walls parameters for atomic sites of oleyl alcohols were represented using the united atom TraPPE (Transferable Potentials for Phase Equilibria) force field [16-18]. The choice of a united atom

model is made to reduce the number of solvent sites in order to achieve a better computational performance and computing algorithm convergence. The hydroxyl hydrogen atom is provided with non-zero σ and parameters [19]. To reduce computational demand, sites with similar charges and van der Walls parameters were counted as equivalent sites with multiplicity of *n* (*n* being the number of such equivalent cites). The force field parameters are provided in Table 1. The oleyl alcohol/water partition coefficient ($LogP_{Ol\text{-}W}$) data was collected from the work of Collander [20]. The lowest energy conformation of the solutes, obtained using the MMFF94 force field, was used for solute parametrization [21,22]. The solutes are parameterized using the GAFF force field with AM1-BCC charges, and also with the UFF parameters with AM1 charges, for comparison purposes [23-26]. For the excess chemical potential in water, a modified SPC water model was used for bulk susceptibility calculations [27]. The 3D-RISM-KH calculations were performed on a uniform cubic 3D-grid of 128 × 128 × 128 points in the box of size 64 × 64 × 64 $Å^3$ representing a solute with a few solvation layers with convergence accuracy set to 10^{-5} of the modified direct inversion in the iterative subspace (MDIIS) solver. All the RISM calculations were performed using our in-house code, a working version of which is available in the AMBERTOOLS package. All the MD simulations were performed using the GROMACS package [28]. The modified GROMOS96 united atom topology of oleyl alcohol was used for liquid simulation [29,30]. The solvent box of pure oleyl alcohol was composed of 512 molecules, and was subjected to 1 *ns* NVT and NPT equilibration without any constraints under periodic boundary conditions. The target temperature was set to 298.15 K and target pressure to 1 bar using Berendsen thermostat. The temperature and density profiles were used to judge the adequacy of the equilibration steps. The final production runs were 10 *ns* long. All the radial distribution functions were calculated using the built-in GROMACS utilities. The 2D-molecular descriptor calculations were generated by

using the PaDEL descriptor package [31]. The performance of predictive modeling was assessed using the mean absolute deviation (MAD) and root mean square deviation (RMSD) values. The RMSE is calculated as: $\mathrm{RMSD} = \sqrt{\frac{\sum_1^N (x_{\mathrm{experiment}} - x_{\mathrm{theory}})}{N}}$, where N is the number of non-missing data points, $x_{\mathrm{experiment}}$ is the experimental partition coefficient, and x_{theory} is the computed partition coefficient.

Table 1. United atom TRAPPE force field parameters employed for oleyl alcohol.

	Solvent Site	Charge	σ (Å)	ε (kcal/mol)
	H	0.435	0.700	0.046
	O	-0.700	3.02	0.185
	C1	0.265	3.95	0.0914
	C2	0.000	3.95	0.0914
	C3	0.000	3.73	0.0934
	C4	0.000	3.75	0.195

Results and Discussion

The NVT and NPT equilibration of the oleyl alcohol solvent box resulted in a stable system at 298 K with a density of ~0.84 g/cm^3 (experimental density: 0.8489 g/cm^3 at 20 °C). The radial distributions of different sites yielded some interesting observations for intermolecular separations. The hydroxyl hydrogens had first maximum at 2.8 Å and a broad peak around 10 Å. The first maxima for the oxygen atoms was shifted to ~3.2 Å with shoulders, and a broad peak

was observed after 8 Å. The intermolecular separation of the hydroxyl oxygen and hydrogen atoms has maxima at 3.7 Å and 8.5 Å. Evidently, the polar hydroxyl part of the oleyl alcohol is arranged in an ordered fashion. The radial distribution functions for the carbon sites also point to an ordered arrangement, although less tightly packed than the polar ends. For example, the intermolecular separations for the unsaturated centers were located at 3.5 Å and ~8 Å. The solvent susceptibility calculations for the liquid oleyl alcohol were attempted with various combination of united atom force field parameters and quantum chemical charges. However, the AMBER and OPLS united atom parameters with the B3LYP/cc-pVTZ charges resulted in negative compressibility of the liquid phase, which is indicative of unphysical nature of the conserved susceptibility function. The TraPPE parameters with modified parameters for hydroxyl hydrogen provided a physical solution for susceptibility function. Hence, for all the 3D-RISM-KH calculations in the oleyl alcohol medium, solvent susceptibility obtained with the TraPPE parameters was used. The partial distribution functions calculated using the X-RISM formalism with the TraPPE force field parameters are in a qualitatively good agreement with those computed via the MD simulation with the modified gromos96 force field. The X-RISM-KH calculations also provide an ordered structure of the oleyl alcohol. These distribution functions are summarized in Table 2.

Table 2. First and second maxima (in Å) of the radial distribution functions from the MD and partial distribution function from the X-RISM-KH calculations.[a]

G	MD	X-RISM	g	MD	X-RISM
H···H	2.8	2.2, 4.98	H···O	3.74, 8.5(b)	4.02, 8.42
O···O	3.2(s), 8.1(b)	3.18, 7.53	O···C1	4.04, ~8(b)	4.18, 8.78
C1···C1	4.7, 9.3(b)	4.13, 7-10(m)	C1···C2	3-5(m)	3-6(m), ~9(b)

C2···C2	3.5, ~8	3.6, 7.88	C2···C3	2.5, 4-6(m)	4-6(m),
C3···C3	5.16	1.55, 3.92	C3···C4	5, ~10(m)	4.05, ~8.5(b)
C4···C4	4.04, 8.9(b)	4.5, 9.2	O···C3	1.64, 9.54	3.75, 8.45

[a]Please refer to Table 1 for solvent sites. b: Broad peak, m: multiple peaks, s: sharp peak.

The experimental molecular partitioning coefficients between oleyl alcohol and water of fourteen compounds were used to build a predictive QSAR model using excess chemical potentials of the solutes computed in the oleyl alcohol and water media. The 2D-molecular descriptors were incorporated in the QSAR model building process. The atomic polarity indices (*bpol*) have a decent correlation with the target molecular permeability as well as the topological polar surface area (TopoPSA) [32]. The 3D-RISM-KH computed excess chemical potentials in water and oleyl alcohol were used as other descriptors in model building. The excess chemical potentials in water have a weak correlation with the $LogP_{Ol\text{-}W}$. The correlations of different descriptors with target partition function is provided in supplementary materials. The multiple linear regression (MLR) model for QSAR prediction has the functional form:

$$LogP_{Ol\text{-}W} = a_0 + a_1 \times \text{bpol} + a_2 \times \mu^{GF}{}_{W} + a_3 \times \mu^{GF}{}_{Ol} + a_4 \times \text{TopoPSA}$$

The performance of the model against fourteen experimental data points, with the GAFF parameters of the solutes, are: $a_0 =$ -0.2926, $a_1 =$ -0.21, $a_2 = 0.25$, $a_3 =$ -0.15, $a_4 =$ -0.06, $R^2 = 0.56$. Here, a_i are the coefficients obtained from MLR analysis. Replacing the GAFF solute parameters with the UFF parameters yielded a similar performance. The computed and experimental oleyl alcohol-water partition coefficients and molecular descriptors used in the calculations are provided in Table 3. Thus, the choice of force field parameters for solutes did not show any effect on the

overall predictive performance of the QSAR model, as evident from the RMSD values of prediction (~0.48 units for both the GAFF and UFF parameters).

Table 3. The QSAR model computed Oleic acid/Water partition functions of 14 solutes computed using the X-RISM-KH theory with the UFF and GAFF force field parameters for the solutes.

Name	$LogP_{Ol\_w}$ (Experimental)	$LogP_{Ol\_w}$ (RISM/UFF)	$LogP_{Ol\_w}$ (RISM/GAFF)
Acetic acid	0.22	0.48	0.45
Butyric acid	2.9	1.66	1.65
Dimethylmalonic acid	0.49	0.40	0.43
Formic acid	0.12	-0.05	-0.03
Glutaric acid	0.11	0.85	0.80
Glycolic acid	0.02	-0.11	-0.07
Lactic acid	0.062	0.32	0.35
Maleic acid	0.11	0.02	0.02
Malic acid	0.018	-0.42	-0.41
Malonic acid	0.049	-0.28	-0.34
2-hydroxyisobutyric acid	0.14	0.68	0.73
Propionic acid	0.81	1.03	0.99
Pyruvic acid	0.12	0.67	0.70
Tricarbalylic acid	0.03	-0.05	-0.06
MAD		0.37	0.37
RMSD		0.48	0.48

Conclusion

In this report, we have used the 3D-RISM-KH molecular solvation theory to understand the liquid state of oleyl alcohol. A united atom model of oleyl alcohol is used for this purpose. The RISM calculated partial distribution functions agree well will the radial distribution functions computed using MD simulations. The solvent susceptibility function of oleyl alcohol obtained from the X-RISM-KH calculations with the united atom TraPPE parameters for heavy atoms and modified hydroxyl hydrogen parameters were used to calculate GF-excess chemical potentials of fourteen solutes to build a QSAR model of oleyl alcohol/water molecular partition coefficients. The model for partition function calculation performed reasonably well irrespective of the choice of force field parameters for the solute molecules. The molecular partitioning predictions using the new QSAR model reported here are calibrated against a relatively small data set of organic acids. Extension of this model to compounds with other functionalities should be done cautiously via prior benchmarking with additional experimental data, if available.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the NSERC Discovery Grant (RES0029477), and the Alberta Prion Research Institute Explorations VII Research Grant (RES0039402).

Acknowledgements

Generous computing time provided by WestGrid (www.westgrid.ca) and Compute Canada/Calcul Canada (www.computecanada.ca) is acknowledged.

Appendix A. Supplementary data

Supplementary data to this article contains partial distribution function and radial distribution function plots, density and temperature profile from MD equilibration runs, molecular descriptors used for partition function predictions.

References

1. T. M. Boudreau and G. A. Hill, *Process Biochem.*, 2006, **41**, 980-983.

2. M. Matsumura, H. Kataoka, M. Sueki and K. Araki, *Bioprocess Eng*., 1988, **3**, 93-100.

3. Y. Kono, N. Naito, A. Yoshimura and K. Aramaki, *J. Oleo. Sci*., 2010, **59**, 581-587.

4. S. Han, *Chem. Phys. Lipid.*, 2013, **175-176**, 79-83.

5. H. C. Anderson and D. Chandler, *J. Chem. Phys.*, 1972, **57**, 1918-1929.

6. D. Chandler and H. C. Anderson, *J. Chem. Phys.*, 1972, **57**, 1930-1937.

7. J. Hansen and I. McDonald, Theory of Simple Liquids, 2nd ed., Academic Press: London, 1986.

8. A. Kovalenko, Three-dimensional Rism Theory for Molecular Liquids and Solid-Liquid Interfaces, in: Hirata F (Ed.), Molecular Theory of Solvation, Kluwer Academic Publishers; Norwell, MA, USA, 2003, pp. 169–276.

9. A. Kovalenko, *Pure Appl. Chem.*, 2013, **85**, 159-199.

10. A. Kovalenko and F. Hirata, *J. Chem. Phys.*, 2000, **112**, 10391.

11. V. Sergiievskyi, G. Jeanmairet, M. Levesque and D. Borgis, *J. Chem. Phys.*, 2015, **143**, 184116.

12. A. D. Becke, *J. Chem. Phys*., 1993, **98**, 5648-5652.

13. C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.

14. T. H. Dunning Jr., *J. Chem. Phys.*, 1989, **90**, 1007-1023.

15. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian 16, Revision B.01, Gaussian, Inc., Wallingford CT, 2016.

16. B. Chen, J. J. Potoff and J. I. Siepmann, *J. Phys. Chem. B*, 2001, **105**, 3093-3104.

17. C. D. Wick, M. G. Martin and J. I. Siepmann, *J. Phys. Chem. B*, 2000, **104**, 8008-8016.

18. M. G. Martin and J. I. Siepmann, *J. Phys. Chem. B*, 1998, **102**, 2569-2577.

19. D. Roy, N. Blinov and A. Kovalenko, *J. Phys. Chem. B*, 2017, **121**, 9268-9273.

20. R. Collander, *Acta Chim. Scandinav.*, 1951, **5**, 774-780.

21. T. A. Halgren and R. B. Nachbar, *J. Comput. Chem.*, 1996, **17**, 587-615.

22. N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminfo.*, 2011, **3**, 33.

23. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.

24. A. Jakalian, B. L. Bush, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2000, **21**, 132-146.

25. A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024-10035.

26. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 13, 3902-3909.

27. F. Hirata and R. M. Levy, *Chem. Phys. Lett.*, 1987, **136**, 267-273

28. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, *SoftwareX*, 2015, **1-2**, 19-25.

29. C. Oostenbrink, A. Villa, A. E. Mark and W. F. Van Gunsteren, *J. Comput. Chem.*, 2004, **25**, 1656-1676.

30. A. K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink and A. E. Mark, *J. Chem. Theory Comput.*, 2011, **7**, 4026-4037.

31. C. W. Yap, *J. Comput. Chem.*, 2011, **32**, 1466-74.

32. P. Ertl, *J. Med. Chem.* 2000, **43**, 3714-7.