Isabela Quintela Matos

Partition Coefficient Calculations of Molecules Mimicking Asphaltenes Through Molecular Simulation Using The Coarse-Grained SAFT-γ Mie Force Field

Rio de Janeiro

2018

Isabela Quintela Matos

Partition Coefficient Calculations of Molecules Mimicking Asphaltenes Through Molecular Simulation Using The Coarse-Grained SAFT-γ Mie Force Field

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Tecnologia de Processos Químicos e Bioquímicos, Escola de Química, Universidade Federal do Rio de Janeiro, como requisitos parcial à obtenção do título de Mestre em Engenharia Química.

Universidade Federal do Rio de Janeiro

Escola de Química

Programa de Pós-Graduação em Tecnologia de Processos Químicos e Bioquímicos

Supervisor: Charlles Rubber de Almeida Abreu

Co-supervisor: Papa Matar Ndiaye

Rio de Janeiro

2018

Isabela Quintela Matos
Partition Coefficient Calculations of Molecules Mimicking Asphaltenes Through Molecular Simulation Using The Coarse-Grained SAFT-γ Mie Force Field/ Isabela Quintela Matos. – Rio de Janeiro, 2018-
46 p. : il. (algumas color.) ; 30 cm.

Supervisor: Charlles Rubber de Almeida Abreu

Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro
Escola de Química
Programa de Pós-Graduação em Tecnologia de Processos Químicos e Bioquímicos, 2018.

1. Palavra-chave1. 2. Palavra-chave2. 2. Palavra-chave3. I. Orientador. II. Universidade xxx. III. Faculdade de xxx. IV. Título

Errata sheet

Elemento opcional da ABNT (2011, 4.2.1.2). Exemplo:

FERRIGNO, C. R. A. **Tratamento de neoplasias ósseas apendiculares com reimplantação de enxerto ósseo autólogo autoclavado associado ao plasma rico em plaquetas**: estudo crítico na cirurgia de preservação de membro em cães. 2011. 128 f. Tese (Livre-Docência) - Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, 2011.

Folha	**Linha**	**Onde se lê**	**Leia-se**
1	10	auto-conclavo	autoconclavo

Isabela Quintela Matos

Partition Coefficient Calculations of Molecules Mimicking Asphaltenes Through Molecular Simulation Using The Coarse-Grained SAFT-γ Mie Force Field

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Tecnologia de Processos Químicos e Bioquímicos, Escola de Química, Universidade Federal do Rio de Janeiro, como requisitos parcial à obtenção do título de Mestre em Engenharia Química.

Trabalho aprovado. Rio de Janeiro, 24 de novembro de 2012:

Charlles Rubber de Almeida Abreu
Orientador

Professor
Convidado 1

Professor
Convidado 2

Rio de Janeiro
2018

Este trabalho é dedicado às crianças adultas que,
quando pequenas, sonharam em se tornar cientistas.

Acknowledgements

Os agradecimentos principais são direcionados à Gerald Weber, Miguel Frasson, Leslie H. Watter, Bruno Parente Lima, Flávio de Vasconcellos Corrêa, Otavio Real Salvador, Renato Machnievscz[1] e todos aqueles que contribuíram para que a produção de trabalhos acadêmicos conforme as normas ABNT com LaTeX fosse possível.

Agradecimentos especiais são direcionados ao Centro de Pesquisa em Arquitetura da Informação[2] da Universidade de Brasília (CPAI), ao grupo de usuários *latex-br*[3] e aos novos voluntários do grupo *abnTeX2*[4] que contribuíram e que ainda contribuirão para a evolução do abnTeX2.

[1] Os nomes dos integrantes do primeiro projeto abnTeX foram extraídos de <http://codigolivre.org.br/projects/abntex/>

[2] <http://www.cpai.unb.br/>

[3] <http://groups.google.com/group/latex-br>

[4] <http://groups.google.com/group/abntex2> e <http://www.abntex.net.br/>

"Não vos amoldeis às estruturas deste mundo,
mas transformai-vos pela renovação da mente,
a fim de distinguir qual é a vontade de Deus:
o que é bom, o que Lhe é agradável, o que é perfeito.
(Bíblia Sagrada, Romanos 12, 2)

Abstract

Segundo a ABNT (2003, 3.1-3.2), o resumo deve ressaltar o objetivo, o método, os resultados e as conclusões do documento. A ordem e a extensão destes itens dependem do tipo de resumo (informativo ou indicativo) e do tratamento que cada item recebe no documento original. O resumo deve ser precedido da referência do documento, com exceção do resumo inserido no próprio documento. (...) As palavras-chave devem figurar logo abaixo do resumo, antecedidas da expressão Palavras-chave:, separadas entre si por ponto e finalizadas também por ponto.

Palavras-chave: latex. abntex. editoração de texto.

Abstract

This is the english abstract.

Keywords: latex. abntex. text editoration.

List of Figures

List of Tables

List of symbols

Γ	Letra grega Gama
Λ	Lambda
ζ	Letra grega minúscula zeta
$\in$	Pertence

Contents

1 Introduction

2 Literature Review

2.1 Coarse Grained Force Fields

Molecular simulations can be carried out at different levels of descriptions. The detailed atomistic level or *ab initio*level is described by the laws of quantum mechanics. The system consists of a set of subatomic particulars in which Schrodinger's equation is solved for all of them. The next level is the atomistic description. It considers that the system is made up of atoms following the laws of statistical mechanics. Force fields at this level are based on pair potentials with Coulombic charged sites, which account for the molecular interactions. The contributions due to to intramolecular interactions like bond-stretching, angle-bending and torsion are also usually accounted by these kind of force fields. When the scale of the simulations needs to be increased and the atomistic simulations become too computationally expensive, the coarse-grained (CG) description is more suited. It considers that the system is made up of pseudo atoms or beads that contain multiple atoms.

There is a obvious loss of information in grouping atoms, hence it is necessary to assure that the process of eliminating unnecessary or unimportant information ('coarse graining') doesn't affect the system's physical behavior. The coarse grained force fields are developed by mapping the atomistic model to define the pseudo atoms with the intetion of assuring that the model has accuracy, transferability, robustness, and computational efficiency. This mapping is normally done by grouping similar funcional groups. The level of coarse-graining also needs to be defined, up to 6 heavy atoms (non-hydrogen atoms) per bead in order to not loose much detail and maintain isotropic representations of the beads (SHINODA; DEVANE; KLEIN, 2007; MARRINK *et al.*, 2007; HADLEY; MCCABE, 2012). The CG force field can be parametrized following two different approaches: bottoms up and top down. The bottoms up approach uses information of a more detailed scale such as the *ab initio* description or the atomistic description to obtain the information necessary to the parametrization. This method depends highly of the quality of the detailed model to succeed. Meanwhile, the top down methodology obtains the parameters from one larger scales. This information at larger scales could be experimentally observed data like thermodynamic properties or native-structure based properties.

One of the first applications of coarse grained models is the study of protein folding (LEVITT; WARSHE, 1975; LEVITT, 1976). These earlier protein CG models were based on the structure of the molecule and they contributed for the knowledge of the physicochemical forces associated with protein folding and protein interactions

(KOGA; TAKADA, 2001). More recent models focused on retaining the protein's chemical specificity. The Bereau and Deresmo model (BEREAU; DESERNO, 2009) has a up to four-bead representation and was used in studies of protein folding and aggregation. However, this model still needs tuning to improve stability of proteins (BEREAU; BACHMANN; DESERNO, 2010). The OPEP (Optimized Potential for Efficient Protein Structure Prediction) model (STERPONE *et al.*, 2014; STERPONE; DERREUMAUX; MELCHIONNA, 2015) has up to six-bead representation. It was used to investigate a variety of phenomena, ranging from protein folding to *ab initio* peptide structure prediction (BARDUCCI; BONOMI; DERREUMAUX, 2011; CHEBARO *et al.*, 2009b; CHEBARO *et al.*, 2009a). Other CG protein models used in the literature are the Scorpion (solvated coarse-grained protein interaction) (BASDEVANT; BORGIS; HA-DUONG, 2013), the UNRES (united residue) (ADAM *et al.*, 2014) and the MARTINI model (LARS *et al.*, 2013). The later one is the most popular model for the CG modeling of membrane proteins (MARRINK; TIELEMAN, 2013). The MARTINI model is also extensively used as CG model for water. This model represents four water molecules as one bead using a shifted Lennard Jones potential for the non bonded interactions. Though its extensive use, the MARTINI water model doesn't properly represent properties as interfacial tension and compressibility (HE *et al.*, 2010) and can freeze at room temperature (WINGER *et al.*, 2009; MARRINK *et al.*, 2007), what makes necessary the use of anti-freeze agents during the simulations. This behavior can be explained by the high level of coarse graining (4:1), the lack of explicit charges and the use of the 12-6 potential. With the idea of improving the MARTINI model, Chiu, Scott e Jakobsson (2010) used the Morse Potential, which is softer than the LJ potential. Meanwhile, Shinoda, Devane e Klein (2007) used different forms of the Mie potential. They concluded that a 12-4 Mie Potential was ideal for the all water cross interactions and a 9-6 Mie Potential was suited for all the solute-solute interactions.

Outside of the Martini framework, He *et al.* (2010) studied different levels of coarse-graining for water ranging for one to 4 molecules per bead using different Mie and Morse potentials. Other works also assessed the use of Soft-core potentials to study aqueous solutions of surfactants (SHINODA; DEVANE; KLEIN, 2007), ionic liquids (BHARGAVA; KLEIN, 2009), lipids (SHINODA; DEVANE; KLEIN, 2010) and membranes (PANTANO; KLEIN, 2009). Other CG force field for water based on the Mie Potential is the SAFT-γ Mie (LOBANOVA *et al.*, 2015). In this model, the water molecule can be represented by two different one isotropic bead interacting via a 8-6 Mie Potential models. The CGW1-vle model was parametrized using saturated-liquid density and vapor pressure data, and should be used for simulations of aqueous systems' fluid-phase equilibria at high temperatures and pressures. This model still suffers from premature freezing with a triple point at 343 K. The other model, CGW1-ift, was parametrized using saturated-liquid density and vapor-liquid interfacial tension, hence it is best

suited for interfacial properties calculations. Both models have temperature-dependent size and energy parameters and performed well for these properties over the entire temperature range of the liquid. The SAFT-γ Mie force field have also been applied to other compounds with satisfactory results. Müller e Mejía (2017) parametrized the force field for aromatic compounds and tested it with simulations of fluid phase equilibrium. Herdes, Totton e Müller (2015) carried out simulations of alkanes and light gases with this force fields. Binary and ternary mixtures of water, carbon-dioxide and water (LOBANOVA *et al.*, 2016), thermodynamic and transport properties of carbon dioxide and methane (AIMOLI; MAGINN; ABREU, 2014a; AIMOLI; MAGINN; ABREU, 2014b) and water/oil interfacial tension (HERDES *et al.*, 2017) were also studied with this force field.

2.2 Solvation Free Energies Based on Molecular Dynamics

Free energies can be expressed as averages over ensembles of atomic configurations generated using Monte Carlo or molecular dynamics techniques. In the canonical ensemble, the free energy is given by:

$$F(N,V,T) = -\kappa_b T Q(N,V,T) \tag{2.1}$$

where $Q(N,V,T))$ is the partition function of the canonical ensemble. Meanwhile, the average over the isothermal-isobaric ensemble gives the Gibbs free energy:

$$G(N,P,T) = -\kappa_b T \frac{1}{V_0} \int_0^\infty dV \exp(-\kappa_b T P V) Q(N,V,T) \tag{2.2}$$

Since it is only possible to obtain free energy differences, Solvation free energy calculations based on molecular dynamics estimate the difference between the Gibbs free energies of end states, more specifically the difference between the solute alone in the gas phase and the solute interacting with the solvent. In order to these energy differences be accurate, the sates' phase integral must have sufficient overlap (KLIMOVICH; SHIRTS; MOBLEY, 2015). This can be achieved by calculating the free energy difference between a series of intermediates states. The result of these differences are independent of the path chosen since the free energy is a state function. That's why the states used typically don't have a physical sense, they are alchemical states which are only linking the physical states of interest.

The solvation free energy calculations follow a thermodynamic cycle to gradually insert the solute molecule into the solvent as illustrated in the Figure 2.2.1. According to

this cycle, the free energy of solvation can be expressed as:

$$\Delta G_{solv} = \Delta G_{1\to4} = \Delta G_{1\to2} + \Delta G_{2\to3} + \Delta G_{3\to4} \tag{2.3}$$

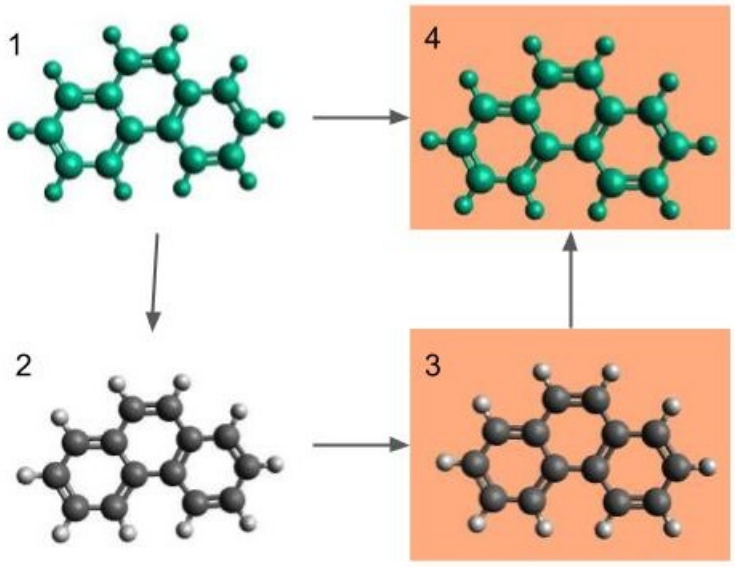

Figure 2.2.1 – Thermodynamic cycle for solvation free energy calculations with molecular dynamics (Adapted from Klimovich, Shirts e Mobley (2015))

The solvation free energy between sates 1 and 2 in the cycle is the one associated with turning off the molecule's non bonded interactions in the gas phase. The following transformation, $\Delta G_{2\to3}$, is the free energy of moving the non-interacting molecule in the gas phase to the solvent and is equal to zero since the transformation of a non interacting molecule doesn't depend on the environment. Lastly, $\Delta G_{3\to4}$ is the free energy required to the the non-interaction molecule in the aqueous phase regain its non-bonded interactions. The solvation free energy calculation can be classified according to the types of the non bonded interactions that are turned of in the $1 \to 2$ and $3 \to 4$ parts of the cycle. If both the non-bonded interactions with the environment and the internal interactions are turned of, this is the annihilation free energy. Meanwhile, if only the non-bonded interactions with the environment are turned off,this is the decoupling free energy. In this later case, $\Delta G_{1\to2} = 0$ and the $\Delta G_{solv} = \Delta G_{3\to4}$. The methods used to carry out theses transformations scale the solute charges to zero and then turn of the interactions corresponding to the Lennard Jones potential. In order to carry out the later process, a modified potential with a coupling parameter λ is used. Each λ represent an alchemical state and, when $\lambda = 0$, there is no interaction with the solvent and, when $\lambda = 1$, the interactions are fully activated. The coupling of the λ parameter could be linear, but it could generate numerical problems related to the exponential part of the Potential. That's why the non-linear soft-core scheme (BEUTLER *et al.*, 1994) is usually

used, the so called soft core Lennard-Jones potential is given by:

$$U_{LJ}^{sc}(r) = 4\lambda\epsilon \left\{ \frac{1}{[\alpha(1-\lambda)^2 + (r/\sigma)^6]^2} - \frac{1}{\alpha(1-\lambda)^2 + (r/\sigma)^6} \right\} \tag{2.4}$$

where α is a constant in which the value of 0.5 is normally assumed to it. The $\Delta G_{3\to4}$ can be then obtained by doing independent simulations in different values of λ or by doing expanded ensemble simulations (LYUBARTSEV *et al.*, 1992) which samples all state in a single simulation. This method allows a faster sampling across the alchemical states considering that the kinetic barriers are not substantial.

2.3 Post simulation methods

The data obtained with molecular dynamics simulations method explained in the section above contain the potential energies correspondent to each λ. These potential energies obtained then needs to be post processed and analyzed in order to calculate the solvation free energies. Some of the widely used method for these calculations are going to be briefly describe below.

2.3.1 Thermodynamic integration

The thermodynamic integration method (KIRKWOOD, 1935) uses equilibrium averages to evaluate the derivative of the potential energy with respect to the coupling parameter. Then, the free energies are obtained as the integration of the derivatives of the initial and final state:

$$\Delta G_{solv} = \int_0^1 \frac{\partial U}{\partial \lambda} d\lambda \tag{2.5}$$

The integration in Eq. (2.5) is obtained by interpolating the output data form the simulations in different ways. Some examples of methods for the interpolations are the trapezoidal rule or natural cubic spline (PALIWAL; SHIRTS, 2011). There are also other more complex schemes that are usually system specific as the works of Jorge *et al.* (2010) and Shyu e Ytreberg (2010) and that use fitting functions to interpolate the data.

2.3.2 Free energy of Pertubation (FEP)

The free energy of perturbation method (ZWANZIG, 1954) is the oldest and the of the most general purpose strategy to calculate free energy differences. In this method, the difference between two thermodynamic states A and B is given by:

$$\Delta G_{AB} = -\frac{1}{\beta} ln \langle e^{-\beta(U_B - U_A)} \rangle_A \tag{2.6}$$

According to the equation above, the free energy difference is calculated by doing an average over the potential energies of state A and B obtained during the simulation of state A. This method requires a great overlap between states, the state B needs to represent a small perturbation in state A, in order to obtain a rapid convergence of the free energy difference. To assure the overlap, it is possible to carry out simulations in N intermediate sates between A and B, so Eq. (2.6) becomes:

$$\Delta G_{AB} = -\frac{1}{\beta} ln \left(\frac{1}{N} \sum_{i=0}^{N+1} e^{-\beta(U_{i+1}-U_i)} \right)_i \tag{2.7}$$

The way of calculating ΔG of Eq. (2.7) is called Exponential Averaging (EXP) (ZWANZIG, 1955; PALIWAL; SHIRTS, 2011). The direction of the transformation is also important in this method. If the direction is of decreasing entropy, the step is of insertion (ΔG_{AB}) and method is called insertion exponential averaging (IEXP). The direction of increasing entropy is called a deletion step (ΔG_{BA}) and the method is labeled as deletion exponential averaging (DEXP). These directions can yield different values of free energy differences due the under sampling in the tail regions of the ΔG_{AB} distribution (KLIMOVICH; SHIRTS; MOBLEY, 2015; POHORILLE; JARZYNSKI; CHIPOT, 2010). These problems makes the EXP methods not suited to calculate free energy differences when the system hasn't a sufficient a overlap. For these cases, the Bennet Acceptance Ratio or the Multi-State Beneet Acceptance Ratio is more indicated.

2.3.3 Bennet Acceptance Ratio (BAR)

The BAR method (BENNETT, 1976) was developed with the intent of eliminating the bias in the free energy estimation. It uses the uncorrelated samples of the potential energy in both directions ($A \rightarrow B$ and $B \rightarrow A$) to obtain the free energy differences using the information in a statically optimal way. The free energy difference between two intermediate states (i and j) is calculated by the self-consistence solution of the following equations:

$$\Delta G_{ij} = \frac{1}{\beta} ln \left(\frac{\sum_{k=1}^{N_j} \frac{1}{1+\exp[-\beta(\Delta U_k^j - C)]}}{\sum_{l=1}^{N_i} \frac{1}{1+\exp[-\beta(\Delta U_l^i - C)]}} \right) + C - \frac{1}{\beta} ln \left(\frac{N_j}{N_i} \right) \tag{2.8}$$

$$C = \Delta G_{ij} + \frac{1}{\beta} ln \left(\frac{N_j}{N_i} \right) \tag{2.9}$$

The total free energy difference between end states is then given by the sum over the differences of consecutive intermediate states. This method also provides a function

to obtain the minimum variance for the free energy difference. The variance equation for any value of C is given by:

$$s_{ij}^2 = \frac{1}{\beta^2 N_i}\left[\frac{\langle f^2(x)\rangle_i}{\langle f(x)\rangle_i^2} - 1\right] + \frac{1}{\beta^2 N_j}\left[\frac{\langle f^2(x)\rangle_j}{\langle f(x)\rangle_j^2} - 1\right] \tag{2.10}$$

where $f(x) = 1/(1+x)$ is the Fermi equation and $x = \exp[\beta(\Delta U - C)]$. The variance of the free energy difference between end sates can be calculated by assuming independent errors and summing over the variance of consecutive intermediate states. However, this assumption is not correct and there is no general formula to obtain a statistically unbiased estimate of an entire transformation with the BAR method (PALIWAL; SHIRTS, 2011).

There are tow other methods related to the BAR method that don't solve the Eqs. (2.8) and (2.9) self consistently. By doing that, the free energy difference will not have minimum variance, but significant space and disk memory can be saved since the averages of Eqs. (2.8) - (2.10) are accumulated.The two methods are the Unoptimized Bennett Acceptance Ratio (UBAR) and Range-Based Bennett Acceptance Ratio (RBAR). The first one avoids the self consistently resolution of the BAR equations by defining $C = \beta^{-1} ln(N_j/N_i)$. The UBAR method requires that the intermediate free energies differences are approximately equal to zero to obtain optimal estimations. Meanwhile, the RBAR method selects a range initial guesses of the constant C in order to calculet a range of ΔG_{ij}. The value of free energy difference correspondent to the minimum variance is then used as input in Eq. (2.9) to calculate the value of C. Hence, this method requires a good estimation of the initial range of the values of C. IN therms of accuracy, the UBAR method can be as accurate as the BAR method, but it may end up being as computational costly (PALIWAL; SHIRTS, 2011).

2.3.4 Multistate Bennet Acceptance Ratio (MBAR)

The MBAR method (SHIRTS; CHODERA, 2008) is a further development of the BAR method. The method proposes a estimator that computes free energies and their uncertainties of all the K states by minimizing the KxK matrix of variances simultaneously. The estimator solves self consistently for each G_i the following equation:

$$G_i = \frac{1}{\beta} ln \sum_{k=1}^{K} \sum_{n=1}^{N_k} \frac{\exp[-\beta U_i(x_{kn})]}{\sum_{l=1}^{K} N_l \exp[\beta(G_l - U_l(x_{kn}))]} \tag{2.11}$$

The equation above requires that the potential energy of uncorrelated configuration n to be evaluated for all K states ($U_i(x_{kn})$ and for all the uncorrelated configuration

snapshots (N_k) from state k. The free energy change between the states is given then by $\Delta G_{ij} = G_j - G_i$. The statistical variance of $S_{ij}^2 \Delta G_{ij}$ is given by the matrix covariances:

$$s_{ij}^2 \Delta G_{ij} \equiv cov(-ln\hat{Z}_j/\hat{Z}_i, -ln\hat{Z}_j/\hat{Z}_i) \tag{2.12}$$

where $\hat{Z}_j$ and $\hat{Z}_i)$ are the partition functions of states i and j. The MBAR method can be considered as limit case of the weighted histogram analysis method (WHAM) (KUMAR *et al.*, 1992) for computing free energies. The Eq. (2.11) becomes equal to the WHAM equations if the histogram width tend to zero. Despite this, the MBAR is still more suited because it doesn't have the bias associated with the discretization and it allows to calculate an error estimate.

3 Fundamentals of the Computational Methods

3.1 SAFT-γ Mie Force Field

3.1.1 SAFT-VR Mie EoS

The SAFT-VR Mie equation of state (LAFITTE *et al.*, 2013) is the basis for the SAFT-γ Mie coarse grained force field (AVENDAÑO *et al.*, 2011). This EoS was initially developed to describe chain molecule formed from fused Mie segments using the Mie attractive and repulsive potential. The Mie potential is a type of generalized Lennard-Jones potential that can be used to describe explicitly repulsive interactions of different hardness/softness and attractive interactions of different ranges, and is given by:

$$U_{Mie}(r) = \epsilon \frac{\lambda_r}{\lambda_r - \lambda_a} \left(\frac{\lambda_r}{\lambda_a}\right)^{\left(\frac{\lambda_a}{\lambda_r - \lambda_a}\right)} \left[\left(\frac{\sigma}{r}\right)^{\lambda_r} - \left(\frac{\sigma}{r}\right)^{\lambda_a}\right] \tag{3.1}$$

where ϵ is the potential well depth, σ is the segment diameter, r is the distance between the spherical segments, λ_r is the repulsive exponent and λ_a is the attractive exponent. This equation uses the Barker e Henderson (1976) high perturbation expansion of the Helmholtz free energy up to third order and an improved expression for the radial distribution function (RDF) of Mie monomers at contact to obtain a equation able to give an accurate theoretical description of the vapor-liquid equilibria and second derivative properties (LAFITTE *et al.*, 2013). For a non-associating fluid, the Helmholtz free energy is:

$$\frac{A}{N\kappa_b T} = a = a^{IDEAL} + a^{MONO} + a^{CHAIN} \tag{3.2}$$

3.1.1.1 Ideal Contribution

The ideal contribution for a mixture is given by:

$$a^{IDEAL} = \sum_{i=1}^{N_c} x_i \ln\left(\rho_i {\Lambda_i}^3\right) - 1 \tag{3.3}$$

where $x_i = N_i/N$ is the molar fraction of component i, $\rho_i = N_i/V$ is the number density, N_i is the number of molecules of each component and Λ_i^3 is de Broglie wavelength.

3.1.1.2 Monomer Contribution

The monomer contribution describes the interactions between Mie segments and can be expressed for a mixture as:

$$a^{MONO} = \left(\sum_{i=1}^{N_c} x_i m_{s,i} \right) a^M \tag{3.4}$$

In the equation above, $m_{s,i}$ is the number of spherical segments making up the molecule i and a^M is the monomer dimensionless Helmholtz free energy and it is expressed as a third order perturbation expansion in the inverse temperature (BARKER; HENDERSON, 1976):

$$a^M = a^{HS} + \beta a_1 + \beta {a_2}^2 + \beta {a_3}^3 \tag{3.5}$$

where $\beta = \kappa_b T$ and a^{HS} is the hard-sphere dimensionless Helmholtz free energy for a mixture :

$$a^{HS} = \frac{6}{\pi \rho_s} \left[\left(\frac{\zeta_2^3}{\zeta_3^2} - \zeta_0 \right) \ln(1-\zeta_3) + \frac{3\zeta_1\zeta_2}{1-\zeta_3} + \frac{\zeta_2^3}{\zeta_3(1-\zeta_3)^2} \right] \tag{3.6}$$

The variable $\rho_s = \rho \sum_i^{N_c} x_i ms, i$ is the total number density of spherical segments and ζ_l are the moments of the number density:

$$\zeta_l = \frac{\pi \rho_s}{6} \left(\sum_{i=1}^{N_c} x_{s,i} d_{ii}^l \right), l = 0, 1, 2, 3 \tag{3.7}$$

where $x_{s,i}$ is the mole fraction of the segments and is related through the mole fraction of component i (x_i) by:

$$x_{s,i} = \frac{m_{s,i} x_i}{\sum_{k=1}^{N_c} m_{s,k} x_k} \tag{3.8}$$

The effective hard-sphere diameter d_{ii} for the segments is:

$$d_{ii} = \int_0^{\sigma_{ii}} (1 - \exp(-\beta U_{ii}^{Mie}(r))) dr \tag{3.9}$$

The integral in Eq. (3.9) is normally obtained by means of Gauss-Legendre with a 5-point quadrature (PAPAIOANNOU *et al.*, 2014). The detailing of the terms of Eq. (3.4) can be found in Lafitte *et al.* (2013).

3.1.1.3 Chain Contribution

The chain formation of m_s tangentially bonded Mie segments contribution is based on the first-order pertubation theory (TPT1) (PAPAIOANNOU *et al.*, 2014) and can be expressed as:

$$a^{CHAIN} = -\sum_{i=1}^{N_c} x_i (m_{s,i} - 1) \ln(g_{ii}^{Mie}(\sigma_{ii})) \tag{3.10}$$

The $g_{ij}^{Mie}(\sigma_{ij})$ term correspond to the value of the radial distribution function (RDF) of the hypothetical Mie system evaluated at the effective diameter and can be obtained with the perturbation expansion:

$$g_{ij}^{Mie}(\sigma_{ij}) = g_{d,ij}^{HS}(\sigma_{ij}) \exp[\beta\epsilon g_{1,ij}(\sigma_{ij})/g_{d,ij}^{HS}(\sigma_{ij}) + (\beta\epsilon)^2 g_{2,ij}(\sigma_{ij})/g_{d,ij}^{HS}(\sigma_{ij})] \tag{3.11}$$

The other terms in the equations above are explicitly exposed in the original article (LAFITTE *et al.*, 2013).

3.1.1.4 Ring Contribution

There are two forms for the Helmholtz free energy for rings formed from m_s tangentially bonded segments in the literature. The first one (LAFITTE *et al.*, 2012) considered that the difference between a chain and a ring molecule is that the latter one has one more bond that is connecting the first segment to the last. With this assumption, the Eq. (3.10) can be adapted to rings by:

$$a^{RING} = -\sum_{i=1}^{N_c} x_i m_{s,i} \ln(g_{ii}^{Mie}(\sigma_{ii})) \tag{3.12}$$

According to Lafitte *et al.* (2012), Eq. (3.12) needs an additional parametrization with molecular simulation data so the EoS can be used in molecular simulations, but this procedure is not the necessary for chain molecules. Recently Müller e Mejía (2017) tried to correct this inconsistency by means of developing the ring free energy based on the work of Müller e Gubbins (1993) who obtained rigorous expressions for hard fluids with molecular geometries of rings of $m_s = 3$. The final expression developed for the ring dimensionless Helmholtz free energy is:

$$a^{RING} = -\sum_{i=1}^{N_c} x_i (m_{s,i} - 1 + \chi_i \eta_i) \ln(g_{ii}^{Mie}(\sigma_{ii})) \tag{3.13}$$

$\eta_i = m_{s,i}\rho_i\sigma_{ii}^3/6$ is the packing fraction and χ_i is a parameter which depends on $m_{s,i}$ and on the geometry of the ring of each component i. For a value of $\chi = 0$ Eq. (3.13) is equal to Eq. (3.10). Meanwhile, the system corresponds to a hard sphere system of triangles when $\chi = 1.3827$. Müller e Mejía (2017) also calculated values of ζ for $m_s = 3, m_s = 4, m_s = 5, m_s = 7$ with pseudo-experimental data from molecular dynamics (MD) for a defined pure fluid. The values of χ for each geometry estimated can be seen in the Figure 3.1.1.

3.1.1.5 Combining rules for the intermolecular potential parameters

Lafitte *et al.* (2013) also suggested mixing rules for the potential parameters based on Lorentz-Berthelot combining rules (ROWLINSON; SWINTON, 1982):

$$\sigma_{ij} = \frac{\sigma ii + \sigma jj}{2} \tag{3.14}$$

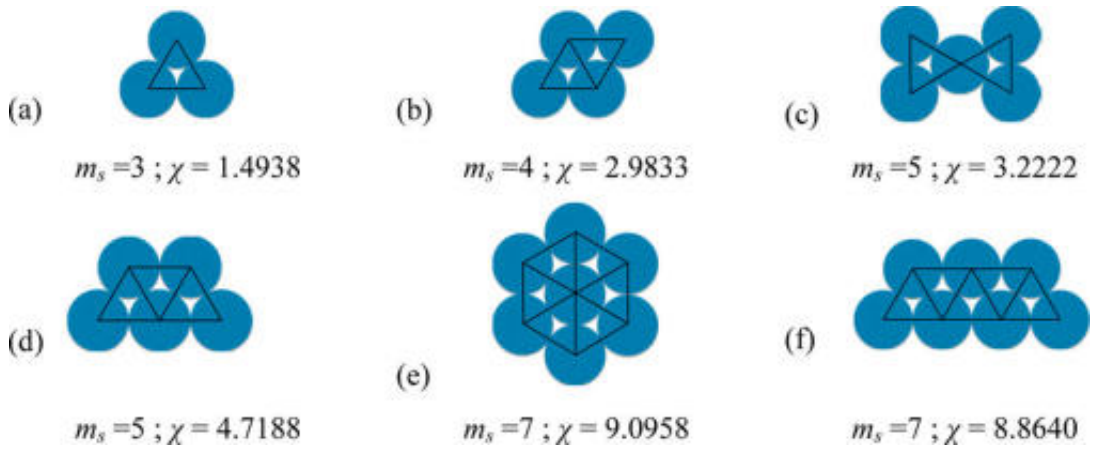

Figure 3.1.1 – Values for parameter χ according to the ring geometry (MÜLLER; MEJÍA, 2017)

$$\lambda_{k,ij} - 3 = \sqrt{(\lambda_{k,ii} - 3)(\lambda_{k,jj} - 3)}, k = r, a \tag{3.15}$$

$$\epsilon_{ij} = (1 - k_{ij}) \frac{\sqrt{\sigma_{ii}^3 \sigma_{jj}^3}}{\sigma_{ij}^3} \sqrt{\epsilon_{ii}\epsilon_{jj}} \tag{3.16}$$

The k_{ij} is a binary interaction parameter to correct the deviations of the Lorentz-Berthelot rule for chemically distinct compounds. This parameter can also be fitted to experimental data or pseudo experimental data.

3.1.2 Parameter Estimation for the SAFT-γ Mie Force Field

The SAFT-γ Mie Force Field uses a coarse graining top down methodology in its parameterization. This methodology aims to obtain the intermolecular parameters from macroscopic experimental data like fluid-phase equilibrium or superficial tension data. The idea is that the force field's parameters estimated with the the SAFT-VR Mie EoS can be used on molecular simulations since both the equation of state and the force field use the same explicit intermolecular potential model (Mie potential). This correspondence between models has already been seem for a variety of fluids in which this force field was parameterized and this success in the representation of the properties of real fluids can be imputed to the degrees of freedom of Mie Potential (HERDES; TOTTON; MÜLLER, 2015). This flexibility also provides an exploration of a very large parameter space without using a iterative simulation scheme (AVENDAÑO *et al.*, 2011).

Each substance has initially five parameters to be estimated (m_s,σ,ϵ,λ_r and λ_a) according to Eq. (3.1). The number of segments are usually fixed in an integer value so it can be used in the coarse grained simulations. The attractive parameter can also be

fixed since there is a high correlation between the attractive and repulsilve parameter. Usually, the parameter is fixed in the London value of 6, which is expected to be a good representation of the dispersion scale of most simple fluids that don't have strong polar interactions (RAMRATTAN *et al.*, 2015; HERDES; TOTTON; MÜLLER, 2015). There are two strategies to obtain the parameters of each substance: one is by fitting the Saft-Vr Mie EoS to experimental data as vapor pressure and liquid density and the other is using correspondent state parametrization. The first one, generally, minimizes the following unweighted least-squares objective function:

$$\min_{\sigma,\epsilon,\lambda_r} F_{obj}(\sigma,\epsilon,\lambda_r) = \sum_{i=1}^{N_p}\left(\frac{P_v^{SAFT}(T_i,\sigma,\epsilon,\lambda_r)-P_v^{exp}(T_i)}{P_v^{exp}(T_i)}\right)^2 + \sum_{i=1}^{N_p}\left(\frac{\rho_l^{SAFT}(T_i,\sigma,\epsilon,\lambda_r)-\rho_l^{exp}(T_i)}{\rho_l^{exp}(T_i)}\right)^2 \tag{3.17}$$

where N_p is the number of experimental points, P_v is the vapor pressure and ρ_l is the saturated liquid density. The minimized properties can also change and other possible properties as superficial tension and speed of sound can also be taken into account. These multiple parameters make it necessary the use of a wide range of experimental data since multiple solutions can be found for the fit. So one need to be careful in deciding the level of coarse graining (i.e. the parameter m_s) and subsequent parameter space that will not result in some physical inconsistencies like a fluid with premature freezing.

Lafitte *et al.* (2012) suggested that the two corrections factors (c_σ and c_ϵ) should be estimated with simulation data when using Eq. (3.12) for the ring contribution. They are related to the EoS parameters by scaled parameters:

$$\sigma^{scaled} = c_\sigma \sigma^{SAFT} \tag{3.18}$$

$$\epsilon^{scaled} = c_\epsilon \epsilon^{SAFT} \tag{3.19}$$

According to Lafitte *et al.* (2012), these corrections are necessary because the approximations employed in the EoS theory generate discrepancies between molecular simulations and the EoS results for ring molecules modeled with Eq. (3.12). The objective function for this second estimation is given by:

$$\min_{c_\sigma,c_\epsilon} F_{obj}(c_\sigma,c_\epsilon) = \sum_{i=1}^{N_p}\left(\frac{P_v^{sim}(T_i,\sigma^{SAFT},\epsilon^{SAFT})-P_v^{SAFT}(T_i,\sigma^{scaled},\epsilon^{scaled})}{P_v^{sim}(T_i,\sigma^{SAFT},\epsilon^{SAFT})}\right)^2 + \sum_{i=1}^{N_p}\left(\frac{\rho_{liq}^{sim}(T_i,\sigma^{SAFT},\epsilon^{SAFT})-\rho_{liq}^{SAFT}(T_i,\sigma^{scaled},\epsilon^{scaled})}{\rho_{liq}^{sim}(T_i,\sigma^{SAFT},\epsilon^{SAFT})}\right)^2 \tag{3.20}$$

The repulsive parameter is maintained in the value found on the minimization of Eq. (3.17), so the refined values for the force field are:

$$\sigma^{sim} = \sigma^{SAFT}/c_\sigma \tag{3.21}$$

$$\epsilon^{scaled} = \epsilon^{SAFT}/c_\epsilon \tag{3.22}$$

It is interesting to point out that this new parametrization is not necessary when using Eq. (3.13) as the ring contribution. The other method to obtain the force field parameters is the correspondent state parametrization for the EoS SAFT-VR Mie (MEJÍA; HERDES; MÜLLER, 2014). This method considers that the unweighted volume average of the attractive contribution to the Mie intermolecular potential,a_1, can be given a mean field approximation:

$$a_1 = 2\pi\rho\sigma^3\epsilon\alpha \tag{3.23}$$

The van der Waals constant, α, considering $\lambda_a = 6$ is related by the Mie exponents by:

$$\alpha = \frac{1}{\epsilon\sigma^3}\int_\sigma^\infty \phi(r)r^2dr = \frac{\lambda_r}{3(\lambda_r-3)}\left(\frac{\lambda_r}{6}\right)^{6/(\lambda_r-6)} \tag{3.24}$$

The parametrization in this method starts by using the experimental acentric factor, ω, for each molecule with a fixed value of m_s to obtain the value of the repulsive exponent with the following Padé series:

$$\lambda_r = \frac{\sum_{i=0} a_i\omega^i}{1+\sum_{i=1} b_i\omega^i} \tag{3.25}$$

a_i and b_i are dependent parameters of the number of segments and a table with its values is presented in the original paper (MEJÍA; HERDES; MÜLLER, 2014). Substituting λ_r into Eq. (3.24), the van der Waals constant can be found. The reduced critical potential T_c^* can also be related to α by a Padé series:

$$T_c^* = \frac{\sum_{i=0} c_i\alpha^i}{1+\sum_{i=1} d_i\alpha^i} \tag{3.26}$$

The values of c_i and d_i are also available in the original paper. The reduced temperature of the equation above is used in conjunction with the experimental critical

temperature, T_c, to find the energy parameter with the relation below:

$$T_c^* = \frac{\kappa_b T_c}{\epsilon} \tag{3.27}$$

The diameter parameter, however, is not obtained with the critical properties, but with the reduced liquid density,$\rho_{T_r=0.7}$, at the reduced temperature ,T_r, of 0.7. This density is also obtained with a Padé series using parameters obtained by Mejía, Herdes e Müller (2014):

$$\rho^*_{T_r=0.7} = \frac{\sum_{i=0} j_i \alpha^i}{1 + \sum_{i=1} k_i \alpha^i} \tag{3.28}$$

The relation among the equation above, σ and the experimental density is given by:

$$\rho^*_{T_r=0.7} = \rho_{T_r=0.7} \sigma^3 N_{av} \tag{3.29}$$

where N_{av} is The Avogadro number. This correspondent state method has the advantage of only requiring critical data, that it is available for a great range of fluids, and one liquid density point. In addition to that, there is an available online parameter database obtained with this strategy (ERVIK; MEJÍA; MÜLLER, 2016).

The binary interaction parameter k_{ij} of Eq. (3.16) is necessary to adjust the mixture behaviour of chemically distinct components. Normally, it is fitted to experimental binary vapor liquid equilibrium or superficial tension data with the SAFT-VR Mie EoS (MÜLLER; MEJÍA, 2017; LOBANOVA *et al.*, 2016). However, Ervik *et al.* (2016) used molecular simulation results to fit the parameter to the superficial tension data of the mixture water-toluene. The strategy followed by them was to do simulations in three values of k_{ij} and then refine the parameter value until a value in good agreement with the experimental data was found.

3.2 Expanded Ensemble Method

Instead of doing various simulations in different values of λ, expanded ensemble simulations (LYUBARTSEV *et al.*, 1992) were developed to allow a non-Boltzmann sampling scheme of different states in only one simulation. The statistical expanded ensemble, Z^{EE} can be defined as a sum of sub ensembles Z_i in different values of λ:

$$Z^{EE} = \sum_{i=1}^{N} Z_i(\lambda_i) exp(\eta_i) \tag{3.30}$$

where N is the number of alchemical states and η_i is the arbitrary weight of the subs ensemble Z_i at each state. This weights allow the sampling of the states during the simulation with equal frequency. A sufficient number of states needs to be sampled in order to obtain the reduced free energies differences between the states $(\hat{f}_i - \hat{f}_j)$ from the simulations' results. The weights are calculated doing trials simulations. For the first simulation, the values of η are chosen or set to zero and the histogram of the states visited is obtained. With this histogram is possible to estimate the free energy differences and, since the weights are related to the free energies by Eq. (3.31)), the next values of η can be calculated from the previous result. This iteration goes on until a uniform distribution is secured.

equilibrium properties are no longer estimated via arithmetic averages, but via properly weighted averages.

$$(\eta_i - \eta_j)_{k+1} = \beta(\hat{f}_i - \hat{f}_j)_k \tag{3.31}$$

These acquired weights can then be used in a longer simulation to obtain the data necessary for the free energy differences calculations.

3.3 Gibbs Ensemble Monte Carlo

4 Methodology

5 Results and Discussion

6 Conclusions

Bibliography

ADAM, L.; MACIEJ, B.; CEZARY, C.; EWA, G.; YI, H.; DAWID, J.; PAWEL, K.; MACIEJ, M.; MARIUSZ, M.; A, M. M.; ANDREI, N.; STANISLAW, O.; A, S. H.; K, S. A.; RAFAL, S.; TOMASZ, W.; YANPING, Y.; BARTLOMIEJ, Z. A unified coarse-grained model of biological macromolecules based on mean-field multipole-multipole interactions. *Journal of molecular modeling*, v. 20, p. 2306, 2014.

AIMOLI, C. G.; MAGINN, E. J.; ABREU, C. Force field comparison and thermodynamic property calculation of supercritical co_2 and ch_4 using molecular dynamics simulations. *Fluid Phase Equilib.*, v. 368, p. 80–90, 2014.

AIMOLI, C. G.; MAGINN, E. J.; ABREU, C. R. Transport properties of carbon dioxide and methane from molecular dynamics simulations. *The Journal of Chemical Physics*, v. 141, p. 134101, 2014.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 6028*: Resumo - apresentação. Rio de Janeiro, 2003. 2 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 14724*: Informação e documentação — trabalhos acadêmicos — apresentação. Rio de Janeiro, 2005. 9 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 14724*: Informação e documentação — trabalhos acadêmicos — apresentação. Rio de Janeiro, 2011. 15 p. Substitui a Ref. ABNT (2005).

AVENDAÑO, C.; LAFITTE, T.; GALINDO, A.; ADJIMAN, C. S.; JACKSON, G.; MULLER, E. A. Saft-γ force field for the simulation of molecular fluids.1. a single-site coarse grained model of carbon dioxide. *The Journal of Physical Chemistry B*, v. 115, p. 11154–11169, 2011.

BARDUCCI, A.; BONOMI, M.; DERREUMAUX, P. Assessing the quality of the opep coarse-grained force field. *Journal of Chemical Theory and Computation*, v. 7, p. 1928– 1934, 2011.

BARKER, J. A.; HENDERSON, D. What is "liquid"? understanding the states of matter. *Review of Modern Physics*, v. 48, p. 587–671, 1976.

BASDEVANT, N.; BORGIS, D.; HA-DUONG, T. Modeling protein-protein recognition in solution using the coarse-grained force field scorpion. *Journal of Chemical Theory and Computation*, v. 9, p. 803–813, 2013.

BENNETT, C. Efficient estimation of free energy differences from monte carlo data. *Journal of Computational Physics*, v. 22, p. 245–268, 1976.

BEREAU, T.; BACHMANN, M.; DESERNO, M. Interplay between secondary and tertiary structure formation in protein folding cooperativity. *Journal of the American Chemical Society*, v. 132, p. 13129– 13131, 2010.

BEREAU, T.; DESERNO, M. Generic coarse-grained model for protein folding and aggregation. *Journal of Chemical Physics*, v. 130, p. 235106, 2009.

BEUTLER, T.; MARK, A.; SCHAIK, R. van; GERBER, P.; GUNSTEREN, W. van. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chemical Physics Letters.*, v. 222, p. 529–539, 1994.

BHARGAVA, B.; KLEIN, M. L. Formation of micelles in aqueous solutions of a room temperature ionic liquid: a study using coarse grained molecular dynamicss. *Molecular Simulation*, v. 107, p. 393–401, 2009.

CHEBARO, Y.; DONG, X.; LAGHAEI, R.; DERREUMAUX, P.; MOUSSEAU, N. Pep-fold: an online resource for de novo peptide structure prediction. *Nucleic Acids Research*, v. 37, p. 498– 503, 2009.

CHEBARO, Y.; DONG, X.; LAGHAEI, R.; DERREUMAUX, P.; MOUSSEAU, N. Replica exchange molecular dynamics simulations of coarse-grained proteins in implicit solvent. *Journal of Physical Chemistry B*, v. 113, p. 267– 274, 2009.

CHIU, S.; SCOTT, H.; JAKOBSSON, E. A coarse-grained model based on morse potential for water and n-alkanes. *Journal of Chemical Theory and Computation*, v. 6, p. 851–863, 2010.

ERVIK, A.; LYSGAARD, M. O.; HERDES, C.; JIMÉNEZ-SERRATOS, G.; MLLER, E. A.; MUNKEJORD, S. T.; MÜLLER, B. A multiscale method for simulating fluid interfaces covered with large molecules such as asphaltenes. *Journal of Computational Physics*, v. 327, p. 576–611, 2016.

ERVIK, A.; MEJÍA, A.; MÜLLER, E. A. Bottled saft: A web app providing saft-γ mie force field parameters for thousands of molecular fluids. *Journal of Chemical Information and Modeling*, v. 56, p. 1609–1614, 2016.

HADLEY, K. R.; MCCABE, C. Coarse-grained molecular models of water: a review. *Molecular Simulation*, v. 38, p. 671–681, 2012.

HE, X.; SHINODA, W.; DEVANE, R.; KLEIN, M. L. Exploring the utility of coarse-grained water models for computational studies of interfacial systems. *Molecular Simulation*, v. 108, p. 2007–2020, 2010.

HERDES, C.; ERVIK, A.; MEJÍA, A.; MÜLLER, E. A. Prediction of the water/oil interfacial tension from molecular simulations using the coarse-grained saft-γ mie force field. *Fluid Phase Equilib.*, 2017.

HERDES, C.; TOTTON, T. S.; MÜLLER, E. A. Coarse grained force field for the molecular simulation of natural gases and condensates. *Fluid Phase Equilibria*, v. 406, p. 91–100, 2015.

JORGE, M.; GARRIDO, N.; QUEIMADA, A.; ECONOMOU, I.; MACEDO, E. Effect of the integration method on the accuracy and computational efficiency of free energy calculations using thermodynamic integration. *Journal of Chemical Theory and Computation*, v. 6, p. 1018–1027, 2010.

KIRKWOOD, J. Statistical mechanics of fluid mixtures. *Journal of Chemical Physics*, v. 3, p. 300–313, 1935.

KLIMOVICH, P. V.; SHIRTS, M. R.; MOBLEY, D. L. Guidelines for the analysis of free energy calculations. *Journal of Computer-Aided Molecular Design*, v. 29, p. 397–411, 2015.

KOGA, N.; TAKADA, S. Roles of native topology and chain-length scaling in protein folding: a simulation study with a go-like model. *Journal of Molecular Biology*, v. 331, p. 171–180, 2001.

KUMAR, S.; BOUZIDA, D.; SWENDSEN, R.; KOLLMAN, P.; ROSENBERG, J. The weighted histogram analysis method for free-energy calculations on biomolecules. 1. the method. *Journal of Computational Chemistry*, v. 13, p. 1011–1021, 1992.

LAFITTE, T.; APOSTOLAKOU, A.; AVENDANO, C.; GALINDO, A.; ADJIMAN, C. S.; MULLER, E. A.; JACKSON, G. Accurate statistical associating fluid theory for chain molecules formed from mie segments. *The Journal of Chemical Physics*, v. 139, p. 154504, 2013.

LAFITTE, T.; AVENDAÑO, C.; PAPAIOANNOU, V.; GALINDO, A.; ADJIMAN, C. S.; JACKSON, G.; MÜLLER, E. A. Saft-γ force field for the simulation of molecular fluids: 3. coarse-grained models of benzene and hetero-group models of n-decylbenzene. *Molecular Physics*, v. 110, p. 1189–1203, 2012.

LARS, V.; PERIOLE, X.; TIELEMAN, D. P.; MARRINK, S. J. Improved parameters for the martini coarse-grained protein force field. *Journal of Chemical Theory and Computation*, v. 9, p. 687–697, 2013.

LEVITT, M. A simplified representation of protein conformations for rapid simulation of protein folding. *Journal of Molecular Biology*, v. 104, p. 59–107, 1976.

LEVITT, M.; WARSHE, A. Computer-simulation of protein folding. *Nature*, v. 253, p. 694–698, 1975.

LOBANOVA, O.; AVENDAÑO, C.; LAFITTE, T.; MÜLLER, E. A.; JACKSON, G. Saft-γ force field for the simulation of molecular fluids: 4. a single-site coarse-grained model of water applicable over a wide temperature range. *Mol. Phys.*, v. 113, p. 1228–1249, 2015.

LOBANOVA, O.; MEJÍA, A.; JACKSON, G.; MLLER, E. A. Saft-γ force field for the simulation of molecular fluids 6: Binaryand ternary mixtures comprising water, carbon dioxide, and n-alkanes. *The Journal of Chemical Thermodynamics*, v. 93, p. 320–336, 2016.

LYUBARTSEV, A. P.; MARTSINOVSKI, A. A.; SHEVKUNOV, S. V.; VORONTSOV-VELYAMINOV, P. N. New approach to monte carlo calculation of the free energy: Method of expanded ensembles. *Journal of Chemical Physics*, v. 96, p. 1776–1783, 1992.

MARRINK, S. J.; RISSELADA, H. J.; YEFIMOV, S.; TIELEMAN, D. P.; VRIES, A. H. de. The martini force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B*, v. 111, p. 7812–7824, 2007.

MARRINK, S. J.; TIELEMAN, D. P. Perspective on the martini model. *Chemical Society Reviews*, v. 42, p. 6801– 6822, 2013.

MEJÍA, A.; HERDES, C.; MÜLLER, E. A. Force fields for coarse-grained molecular simulations from a corresponding states correlation. *Industrial and Chemical Engineering Research*, v. 53, p. 4131–4141, 2014.

MÜLLER, E. A.; GUBBINS, K. E. Simulation of hard triatomic and tetratomic molecules. a test of associating fluid theory. *Molecular Physics*, v. 80, p. 957–976, 1993.

MÜLLER, E. A.; MEJÍA, A. Extension of the saft-vr mie eos to model homonuclear rings and its parametrization based on the principle of corresponding states. *Langmuir*, -, p. A–L, 2017.

PALIWAL, H.; SHIRTS, M. R. A benchmark test set for alchemical free energy transformations and its use to quantify error in common free energy methods. *Journal of Chemical Theory and Computation*, v. 7, p. 4115–4134, 2011.

PANTANO, D.; KLEIN, M. L. Characterization of membrane-protein interactions for the leucine transporter from aquifex aeolicus by molecular dynamics calculations. *Journal of Physical Chemistry B*, v. 113, p. 13715–13722, 2009.

PAPAIOANNOU, V.; LAFITTE, T.; AVENDAÑO, C.; ADJIMAN, C. S.; JACKSON, G.; MÜLLER, E. A.; GALINDO, A. Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from mie segments. *The Journal of Chemical Physics*, v. 140, p. 054107, 2014.

POHORILLE, A.; JARZYNSKI, C.; CHIPOT, C. Good practices in freeenergy calculations. *Journal of Physical Chemistry B*, v. 114, p. 10235–10253, 2010.

RAMRATTAN, N.; AVENDAÑO, C.; MÜLLER, E.; GALINDO, A. A corresponding-states framework for the description of the mie family of intermolecular potentials. *Molecular Physics*, v. 113, p. 1–16, 2015.

ROWLINSON, J. S.; SWINTON, F. L. *Liquid and Liquid Mixtures*. 3. ed. London: Butterworth Scientific, 1982.

SHINODA, W.; DEVANE, R.; KLEIN, M. L. Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants. *Molecular Simulation*, v. 33, p. 27–36, 2007.

SHINODA, W.; DEVANE, R.; KLEIN, M. L. Zwitterionic lipid assemblies: molecular dynamics studies of monolayers, bilayers, and vesicles using a new coarse grain force field. *Journal of Physical Chemistry B*, v. 114, p. 6836–6849, 2010.

SHIRTS, M. R.; CHODERA, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *Journal of Chemical Physics*, v. 129, p. 124105, 2008.

SHYU, C.; YTREBERG, F. M. Reducing the bias and uncertainty of free energy estimates by using regression to fit thermodynamic integration data. *Journal of Computational Chemistry*, v. 6, p. 1018–1027, 2010.

STERPONE, F.; DERREUMAUX, P.; MELCHIONNA, S. Protein simulations in fluids: Coupling the opep coarse-grained force field with hydrodynamics. *Journal of Chemical Theory and Computation*, v. 11, p. 1843– 1853, 2015.

STERPONE, F.; MELCHIONNA, S.; TUFFERY, P.; PASQUALI, S.; MOUSSEAU, N.; CRAGNOLINI, T.; CHEBARO, Y.; ST-PIERRE, J.-F.; KALIMERI, M.; BARDUCCI, A.; LAURIN, Y.; TEK, A.; BAADEN, M.; NGUYEN, P. H.; DERREUMAUX, P. The opep protein model: from single molecules, amyloid formation, crowding and hydrodynamics to dna/rna systems. *Chemical Society Reviews*, v. 43, p. 4871– 4893, 2014.

WINGER, M.; TRZESNIAK, D.; BARON, R.; GUNSTEREN, W. van. On using a too large integration time step in molecular dynamics simulations of coarse-grained molecular models. *Physical Chemistry Chemical Physics*, v. 11, p. 1934, 2009.

ZWANZIG, R. W. High-temperature equation of state by a perturbation method. i. nonpolar gases. *The Journal of Chemical Physics*, v. 22, p. 1420, 1954.

ZWANZIG, R. W. High-temperature equation of state by a perturbation method. ii. polar gases. *The Journal of Chemical Physics*, v. 23, p. 1915, 1955.

Appendix

APPENDIX A – Quisque libero justo

Quisque facilisis auctor sapien. Pellentesque gravida hendrerit lectus. Mauris rutrum sodales sapien. Fusce hendrerit sem vel lorem. Integer pellentesque massa vel augue. Integer elit tortor, feugiat quis, sagittis et, ornare non, lacus. Vestibulum posuere pellentesque eros. Quisque venenatis ipsum dictum nulla. Aliquam quis quam non metus eleifend interdum. Nam eget sapien ac mauris malesuada adipiscing. Etiam eleifend neque sed quam. Nulla facilisi. Proin a ligula. Sed id dui eu nibh egestas tincidunt. Suspendisse arcu.

APPENDIX B – Nullam elementum urna vel imperdiet sodales elit ipsum pharetra ligula ac pretium ante justo a nulla curabitur tristique arcu eu metus

Nunc velit. Nullam elit sapien, eleifend eu, commodo nec, semper sit amet, elit. Nulla lectus risus, condimentum ut, laoreet eget, viverra nec, odio. Proin lobortis. Curabitur dictum arcu vel wisi. Cras id nulla venenatis tortor congue ultrices. Pellentesque eget pede. Sed eleifend sagittis elit. Nam sed tellus sit amet lectus ullamcorper tristique. Mauris enim sem, tristique eu, accumsan at, scelerisque vulputate, neque. Quisque lacus. Donec et ipsum sit amet elit nonummy aliquet. Sed viverra nisl at sem. Nam diam. Mauris ut dolor. Curabitur ornare tortor cursus velit.

Morbi tincidunt posuere arcu. Cras venenatis est vitae dolor. Vivamus scelerisque semper mi. Donec ipsum arcu, consequat scelerisque, viverra id, dictum at, metus. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut pede sem, tempus ut, porttitor bibendum, molestie eu, elit. Suspendisse potenti. Sed id lectus sit amet purus faucibus vehicula. Praesent sed sem non dui pharetra interdum. Nam viverra ultrices magna.

Aenean laoreet aliquam orci. Nunc interdum elementum urna. Quisque erat. Nullam tempor neque. Maecenas velit nibh, scelerisque a, consequat ut, viverra in, enim. Duis magna. Donec odio neque, tristique et, tincidunt eu, rhoncus ac, nunc. Mauris malesuada malesuada elit. Etiam lacus mauris, pretium vel, blandit in, ultricies id, libero. Phasellus bibendum erat ut diam. In congue imperdiet lectus.

Annex

ANNEX A – Morbi ultrices rutrum lorem.

Sed mattis, erat sit amet gravida malesuada, elit augue egestas diam, tempus scelerisque nunc nisl vitae libero. Sed consequat feugiat massa. Nunc porta, eros in eleifend varius, erat leo rutrum dui, non convallis lectus orci ut nibh. Sed lorem massa, nonummy quis, egestas id, condimentum at, nisl. Maecenas at nibh. Aliquam et augue at nunc pellentesque ullamcorper. Duis nisl nibh, laoreet suscipit, convallis ut, rutrum id, enim. Phasellus odio. Nulla nulla elit, molestie non, scelerisque at, vestibulum eu, nulla. Ut odio nisl, facilisis id, mollis et, scelerisque nec, enim. Aenean sem leo, pellentesque sit amet, scelerisque sit amet, vehicula pellentesque, sapien.

ANNEX B – Cras non urna sed feugiat cum sociis natoque penatibus et magnis dis parturient montes nascetur ridiculus mus

Sed consequat tellus et tortor. Ut tempor laoreet quam. Nullam id wisi a libero tristique semper. Nullam nisl massa, rutrum ut, egestas semper, mollis id, leo. Nulla ac massa eu risus blandit mattis. Mauris ut nunc. In hac habitasse platea dictumst. Aliquam eget tortor. Quisque dapibus pede in erat. Nunc enim. In dui nulla, commodo at, consectetuer nec, malesuada nec, elit. Aliquam ornare tellus eu urna. Sed nec metus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas.

ANNEX C – Fusce facilisis lacinia dui

Phasellus id magna. Duis malesuada interdum arcu. Integer metus. Morbi pulvinar pellentesque mi. Suspendisse sed est eu magna molestie egestas. Quisque mi lorem, pulvinar eget, egestas quis, luctus at, ante. Proin auctor vehicula purus. Fusce ac nisl aliquam ante hendrerit pellentesque. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Morbi wisi. Etiam arcu mauris, facilisis sed, eleifend non, nonummy ut, pede. Cras ut lacus tempor metus mollis placerat. Vivamus eu tortor vel metus interdum malesuada.