Molecular Modeling of Hydrophobic Effects in Complex Biomolecular Systems: From Simple Mixtures to Protein-Interface Aggregation

Von der Fakultät Energie-, Verfahrens- und Biotechnik
der Universität Stuttgart zur Erlangung der Würde eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

Vorgelegt von

Sven Benson

geboren in Stuttgart

Hauptberichter: Prof. Dr. Jürgen Pleiss
Mitberichter: Prof. Dr. Joachim Groß
Prof. Jens Abildskov

Tag der mündlichen Prüfung: 03.12.2014

Institut für Technische Biochemie der Universität Stuttgart

2014

Contents of this dissertation have already been published:

Benson S.P., Pleiss J. (2013) "Incomplete Mixing versus Clathrate-like Structures: A Molecular View on Hydrophobicity in Methanol-Water Mixtures." *Journal of Molecular Modeling*, Volume 19, Issue 8, pp 3427-3436.

Benson S.P., Pleiss J. (2014) "Molecular Dynamics Simulations of Self-Emulsifying Drug Delivery Systems (SEDDS): Influence of Excipients on Droplet Nanostructure and Drug Localization." *Langmuir*, Volume 30, Issue 28, pp 8471-8480.

Benson S.P., Pleiss J. (2014) "The Solvent Flux Method (SFM): A Case Study of Water Access to Candida Antarctica Lipase B." *Journal of Chemical Theory and Computation*, Volume 10, Issue 11, pp 5206-5214.

Table of Contents

Zusammenfassung

Hydrophobizität ist ein Begriff, der die Bildung molekularer Strukturen im wässrigen Milieu beschreibt und findet daher in nahezu jedem biomolekularen Kontext Anwendung. Hydrophobizität lässt sich nicht durch einen präzisen Formalismus begründen; es handelt sich dabei um ein abstraktes Konzept, welches das strukturbildende "Verhalten" von Molekülen in Wechselwirkung mit Wasser beschreibt. Die Terminologie ist dabei irreführend, da eine Abstoßung zwischen unpolaren Molekülgruppen und Wasser impliziert wird, wobei in der Tat anziehende Kräfte vorliegen, die sich auf atomare Dipole zwischen Elektronen und Protonen zurückführen lassen. Obwohl es längst als gesichert gilt, dass sich die zentrale Triebkraft von Hydrophobizität nicht auf die „Scheu vor Wasser" sondern eher auf eine Art „Narzissmus des Wassers", d.h. auf die präferierte Interaktion von Wassermolekülen mit ihresgleichen, zurückführen lässt, ist der Begriff Hydrophobizität im wissenschaftlichen Sprachgebrauch fest etabliert. Dies lässt sich auf die Arbeiten von Kautzmann zurückführen, in denen Proteinstabilität mit hydrophoben Wechselwirkungen korreliert wurde. Gegenstand dieser Dissertation sind im weitesten Sinne die Arbeiten des Autors bezüglich struktureller und dynamischer Charakterisierung hydrophober Effekte in biomolekularen Systemen, wobei Systeme in drei verschiedenen Größenordnungen untersucht wurden: Binäre Methanol-Wasser Mischungen, die Aggregation von Triglyzeridtröpfchen in wässriger Lösung, sowie die Interaktion von Enzymen mit großflächigen Triglyzerid-Wasser-Grenzflächen. In der Studie "Incomplete Mixing versus Clathrate-Like Structures: A Molecular View on Hydrophobicity in Methanol-Water Mixtures" (Chapter 4.1) wurden unterschiedliche Eigenschaften von Methanol-Wasser Mischungen analysiert und hinsichtlich der molekularen Grundlagen von Exzessgrößen untersucht, insbesondere in Bezug auf Entropie und der Diskrepanz zwischen realem und idealem Mischungsverhalten. Hierbei wurde die Hypothese von Frank und Evans untersucht, welche besagt, dass sich diese Abweichungen auf strukturierte Wassermoleküle bei der Solvatisierung unpolarer Moietäten in Lösung zurückführen lassen. Darüber hinaus wurde die Hypothese *incomplete mixing* analysiert, welche besagt, dass sich das Phänomen lediglich durch molekulare Entmischung erklären lässt. In der zugrundeliegenden Studie wurde *incomplete mixing* durch einen neu eingeführten Parameter, den *molecular association bias*, bestätigt. Dieser Parameter wurde

aus einer statistischen Evaluierung atomarer Nachbarschaftsbeziehungen abgeleitet. Unterstützende Indizien für die Hypothese von Frank und Evans konnten bei der Untersuchung geeigneter dynamischer und struktureller Kenngrößen der Mischungssysteme nicht erhoben werden. Ein intuitiv zugängliches Modell wurde daraufhin vorgeschlagen, welches die variable molekulare Selbstdiffusionskonstante in Abhängigkeit des Mischungsverhaltens beschreibt. Die Selbstdiffusionskonstante wurde dabei mit der Formation bzw. dem Zusammenbruch von systemweiten Wasserstoffbrückennetzwerken in Verbindung gebracht. In einer Folgestudie "Molecular Dynamics Simulations of Self-Emulsifying Drug Delivery Systems (SEDDS): Influence of Excipients on Droplet Nanostructure and Drug Localization." (Chapter 4.2) wurde die molekulare Strukturbildung im Kontext selbstemulsifizierender Systeme untersucht. Hierbei konnte die Bildung von Lipidtröpfchen bestehend aus Wasser, Triglyzerid, Tensid und hydrophoben Wirkstoffmolekülen auf Grundlage zufällig verteilter Startstrukturen modelliert und dabei eine Hierarchie von Assoziationspräferenzen zwischen molekularen Moietäten aufgezeigt werden. Während die Tröpfchenbildung auf der präferentiellen Selbstassoziation von Wassermolekülen beruht, konnte die Formation von Nanostrukturen im Inneren der Tröpfchen auf die Aggregation von hydrophilen Triglyzerid-Kopfgruppen zurückgeführt werden. Lamellare Substrukturen, induziert durch selbstassoziierte hydrophile Triglyzerid-Kopfgruppen, wurden als dominantes strukturbildendes Motiv im Inneren von SEDDS-Tröpfchen identifiziert. Es konnte gezeigt werden, dass sich durch Exzipientenvariation, wie z.B. der Kettenlänge von Triglyzerid, dem Zusatz von Di-, Monoglyzerid oder Polyethylenglycol, eine spezifische Nanoumgebung für die Stabilisierung von Wirkstoffmolekülen einstellen lässt. Durch die gezielten Modulationsmöglichkeiten der Nanostruktur von Tröpfchen ergeben sich Designpotentiale für SEDDS Formulierungen bezüglich der Stabilisierung von Wirkstoffen, der Kontrolle von Wirkstoff-Abgabe-Raten und dem Schutz vor abbauenden Enzymen im Verdauungstrakt. Zu diesen Enzymen zählen Hydrolasen, welche Esterverbindungen von Triglyceriden in wässrigem Milieu spalten, wodurch die Oberfläche von SEDDS-Tröpfchen sukzessive abgebaut wird. In der Studie "The Solvent Flux Method (SFM): A Case Study of Water Access to Candida Antarctica Lipase B" (Chapter 4.3) wurde die hydrolytische Aktivität der Hydrolase *Candida antarctica* lipase B (CALB) in ihrer katalytisch aktiven Konformation modelliert, d.h. adsorbiert an einer phasenseparierten Grenzschicht bestehend aus Triglyzerid und Wasser. Hierbei war der Substratzugangskanal

zum aktiven Zentrum in Richtung der Grenzschicht orientiert und darin eingebettet. Die eingesetzten Triglyzeridschichtkoordinaten stellten ein Modell der Oberfläche eines großflächigen Triglyzeridaggregats dar, unter der Voraussetzung, dass die Oberflächenkrümmung des Aggregats in der Größenordnung von Proteinen vernachlässigbar ist. Es konnte gezeigt werden, dass Triglyzerid und Wasser nicht vollständig phasenseparіert vorliegen, d.h. Wasser drang in großen Mengen in die Triglyzeridschicht ein, wodurch sich eine wasserreiche Übergangsphase in der Größenordnung von einigen Angstrom ausbildete. Im Anschluss an eine Analyse der Dynamik und der Assoziationspräferenz von Wassermolekülen wurde vorgeschlagen, dass der Influx von Wasser in die Grenzschicht durch die lamellaren Assoziationsstrukturen von polaren Triglyzeridmoietäten induziert wird. Des Weiteren konnte gezeigt werden, dass CALB während der Simulation nicht lediglich an der Schicht adsorbiert, sondern darin leicht eintaucht, was auf die wasserreiche Übergangsphase zurückgeführt wurde, in der CALB effektiv „schwimmen“ kann. Die *solvent flux method* (SFM) wurde als allgemeingültige molekulardynamische Methode für synthetisches Hydrolase-Engineering konzipiert und wurde eingesetzt, um den Einfluss von Wasser in das aktive Zentrum von CALB in aktiver Konformation zu modellieren. Signifikanter Einfluss von Wasser durch den Substratkanal konnte hierbei identifiziert werden, obwohl der Substratkanal vollständig in der Triglyzeridphase eingebettet vorlag. Zwei polare Aminosäurereste wurde an der Oberfläche von CALB identifiziert, die repräsentativ für präferierten Einfluss von Wassermolekülen via der Triglyzerid-Wasser Grenzschicht sein könnten. Es wurde vorgeschlagen, dass die kooperative Interaktion zwischen polaren Aminosäureresten und selbstassoziierten polaren Triglyzeridmoitäten zur Diffusion von Wasser in der Grenzschicht beitragen. Wassermoleküle könnten daher zwischen Protein und Grenzschicht „hindurchsickern“. Darüber hinaus konnte ein potentieller sekundärer Wasserkanal identifiziert und die Existenz eines bekannten primären Wasserkanals bestätigt werden. Daraus wurde geschlossen, dass SFM in Verbindung mit akkuraten Grenzflächenmodellen einen nützlichen Beitrag zum Design substratspezifischer Enzymvarianten mit hydrophoben Aminosäuresubstitutionen leisten kann, wobei die Konzentration von Wassermolekülen im aktiven Zentrum verringert und der Einfluss von präferierten Nukleophilen in synthetischen Applikationen erhöht wird.

Summary

Hydrophobicity is a term commonly used to discuss the formation of molecular structures in aqueous solution, and since water is ubiquitous in cellular systems, it may be applied in virtually every biomolecular context. Hydrophobicity is not a first-principle parameter but an abstract concept to describe the "behavior" of molecules in aqueous environments. The terminology of hydrophobicity is misleading, because it implies repulsion or a lack of attraction between nonpolar groups and water, when in fact attractive interactions persist due to atom dipoles. Although it has long been recognized that the driving force of structure formation in aqueous environments is founded in water's "narcissism", i.e., water self-preference, rather than in a general "fear of water", the term hydrophobicity has established itself ever since Kautzmann related protein stability to hydrophobic interactions. Due to its false implications, hydrophobicity can be a cause of confusion and the culprit of misleading deductions. Presented in this dissertation is the author's work on the structural and dynamical characterization of hydrophobic effects in biomolecular systems in the broadest sense, whereby molecular systems on three different size scales are covered: binary mixtures of methanol and water, aggregation of triglyceride droplets in aqueous solution and enzymes that interact with triglyceride-water interfaces of large-scale triglyceride aggregates. The study on "Incomplete Mixing versus Clathrate-Like Structures: A Molecular View on Hydrophobicity in Methanol-Water Mixtures" (Chapter 4.1) models methanol-water mixture properties to investigate the molecular mechanisms of excess effects, particularly relating to the origins of entropy deviations from ideal mixture behavior. Thereby, the hypothesis of Frank and Evans is examined, which relates entropy deviations from ideal mixture behavior to the existence of structured water molecules involved in the solvation of nonpolar moieties. The hypothesis of incomplete mixing is also examined, which states that a simple demixing of molecules is sufficient to explain the entropy deviation. In the aforementioned study, incomplete mixing was supported by evidence derived from a proposed parameter, the molecular association bias, which was conceived from a statistical evaluation of atomistic nearest neighbor relationships. Conversely, no evidence was found to support the Frank and Evans hypothesis in methanol-water mixtures, after investigating dynamic properties of water molecules in the first solvation shell of the methyl moiety of methanol. Furthermore, a pictorial model was

proposed to explain the variable molecular mobility in different ratios of methanol-water mixtures, which is quantifiable by the molecular self-diffusion coefficient. The self-diffusion minimum was thereby linked to the formation or the breakdown of the system-wide hydrogen bonding network of water. In a follow-up study "Molecular Dynamics Simulations of Self-Emulsifying Drug Delivery Systems (SEDDS): Influence of Excipients on Droplet Nanostructure and Drug Localization." (Chapter 4.2), the structural implications of hydrophobicity in the context of self-emulsifying lipidic droplets were explored. Droplets were shown to aggregate from a random distribution of mixtures consisting of water, triglyceride, surfactants and hydrophobic drug molecules. In a hierarchy of molecular association preferences, triglyceride droplets formed as a consequence of water self-preference, while the nanostructure within droplets was found to be affected by the tendency of polar triglyceride moieties to self-associate. It was shown that through excipient variations, such as triglyceride fatty acid chain length, the addition of di- and monoglyceride molecules or surfactant poly(ethylene glycol), it is possible to facilitate different nanoenvironments for the solubilization of drug molecules. Lamellar-like patterns due to self-associated polar triglyceride moieties were observed to be a dominant structural feature in SEDDS droplets. It was shown that SEDDS nanostructure could be altered by the presence of other excipient molecules, especially by the surfactant monoglyceride. The ability to rationally engineer the nanoenvironments of SEDDS droplets by controlled excipient variations could prove beneficial to SEDDS formulation design and could potentially allow for a tuning of drug stabilization, drug release-rates and for protecting drugs from degradative enzymes of the gastrointestinal tract. Among those enzymes are hydrolases, which cleave ester bonds of triglyceride molecules in the presence of water and thus successively degrade the surface of SEDDS droplets during digestion. In the study "The Solvent Flux Method (SFM): A Case Study of Water Access to Candida Antarctica Lipase B" (Chapter 4.3), hydrolase activity of *Candida antarctica* lipase B (CALB) was modeled in its catalytically active conformation, i.e., attached to a triglyceride layer with its active site substrate channel oriented towards a triglyceride-water interface. The applied planar interface layer model was thereby representative of large-scale triglyceride aggregate surfaces, on the premise that the surface curvature is negligible in relation to the protein size. It was found that the triglyceride-water interface is wet, i.e., the aqueous solvent environment and the triglyceride phase were not strictly separated in the model. Water was shown to penetrate into the layer in significant amounts, effectively

creating a water-rich transitional phase with a thickness of several angstroms. After analyzing water molecule dynamics and association preferences, it was proposed that the influx of water into the triglyceride layer is mediated by the lamellar structures induced by the self-association of polar triglyceride moieties. Moreover, CALB did not simply adsorb onto the layer in simulation, instead it slightly immersed into the triglyceride layer, which was perceived to be a consequence of the water-rich transitional phase, in which CALB may effectively "swim". The solvent flux method (SFM), which was developed as a general-purpose MD method for hydrolase engineering, was applied to model the influx of water molecules to the active site of CALB in its attached conformation. Significant water influx was thereby detected through the substrate channel, despite it being fully enveloped by the triglyceride phase. Two polar residues were identified at the surface of CALB that could be representative for preferred entryways for water molecules via the triglyceride-water interface. It was proposed that the cooperative interaction between the polar amino acid residues and the self-associated polar triglyceride moieties may facilitate water diffusion. Water molecules were thus considered to "leak" in between the protein and the water-rich transitional phase of the triglyceride layer. Beyond the identification of these water entryways, SFM also revealed the presence of a potential secondary water channel and confirmed a known water channel in CALB. It was thus concluded that in conjunction with accurate interface models, SFM could prove beneficial for the engineering of substrate-specific hydrophobic enzyme variants that exclude water molecules from the active site and increase the influx of preferred nucleophiles in synthetic applications of hydrolases.

1 Introduction

In this chapter, fundamental methods, concepts and theory of molecular dynamics simulations (MD) are outlined (Chapter 1.1 and 1.2), followed by a brief presentation of the models that are applied in the core studies of this dissertation (Chapter 1.3).

1.1 Molecular Dynamics Simulations

1.1.1 Method Context

MD simulations provide atomistic detail on the molecular structures and dynamics of biomolecular systems and can thus provide insights that are otherwise inaccessible or difficult to obtain by experimental means. X-ray crystallography offers only limited information on dynamic properties, which were shown to systematically underestimate the flexibility of protein B-factors (Kuzmanic et al. 2014), and it relies on the ability to generate crystals of a sample, which is not always feasible. NMR is currently restricted by an extensive data evaluation process and size limitations (Marion 2013), which can be particularly problematic when resolving the dynamics of large proteins or protein complexes. MD is not the only available computational method to assess the molecular structure of biomolecular systems, but offers some key advantages over other methods. For the characterization of dynamic properties, for example, it is better suited than Monte Carlo methods (MC), which rely on the sampling of ensemble conformations by performing random displacements and thereby evaluating relative energies. While the characterization of energy barriers is thus more deliberate with MC and often more efficient, MC does not rigorously resolve the dynamic evolution of molecular systems as well as MD. Both MD and MC suffer from the fact that bond-breaking cannot be modeled and also that its parameters are kept constant during simulation in conventional implementations of the methods. Quantum mechanical (QM) approaches are in principle better suited for this task, but suffer from the drawback that they are computationally extremely demanding and therefore only viable for the characterization of small systems. Hybrid QM/MM approaches are typically coupled to MD and have been successfully

applied in exploring catalytic functions of proteins (van der Kamp et al. 2013). In large biomolecular systems, these hybrid approaches are a means to implement a QM level of detail to specific structural elements as required, while utilizing the computational efficiency of MD to model the dynamics of the remaining interactions. Thus, MD distinguishes itself in the efficient computational characterization of the time-propagation of molecular systems, provided the phenomena of interest may be resolved within computationally accessible timescales.

1.1.2 Outline of Formalism

Methodically, MD solves the equations of motions of classical mechanics for systems consisting of N interacting particles (eq.1).

$$m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2} = m_i \frac{\partial \boldsymbol{v}_i}{\partial t} = m_i \boldsymbol{a}_i = \boldsymbol{F}_i \qquad i = 1,2,3,\dots,N \tag{eq.1}$$

$$\boldsymbol{F}_i = -\frac{\partial \boldsymbol{V}_i}{\partial \boldsymbol{r}_i} \tag{eq.2}$$

The forces $\boldsymbol{F}_i$ acting between atoms are negative derivatives of the potential functions that define the interactions (eq.2). Technically, it is not feasible to solve these equations in a continuous manner; instead, discrete time-steps Δt need to be defined, typically in the timescale of a femtosecond (10^{-15}s = 1fs) to resolve the fastest motions within the system, which in the biomolecular context are usually hydrogen bond vibrations. By constraining bond-stretching vibrations or bond-angle vibrations with algorithms such as SHAKE (Ryckaert et al. 1977) or LINCS (Hess et al. 1997), it is possible to increase the time-step and therefore enable simulation to either finish faster or to increase the overall timeframe that is achievable with available computational resources. Time-steps of 2-5fs are thus made viable for systems containing organic compounds. In MD, the equations of motions are solved by an integrator, e.g., with the leap-frog algorithm (Hockney et al. 1974).

$$\boldsymbol{v_i}\left(t+\tfrac{1}{2}\Delta t\right)=\boldsymbol{v_i}\left(t-\tfrac{1}{2}\Delta t\right)+\frac{\Delta t}{m}\boldsymbol{F_i}(t) \quad \text{(eq.3)}$$

$$\boldsymbol{r_i}(t+\Delta t)=\boldsymbol{r_i}(t)+\Delta t\boldsymbol{v_i}\left(t+\tfrac{1}{2}\Delta t\right) \quad \text{(eq.4)}$$

A trajectory can thus be generated that consists of a succession of system conformations at time-steps Δt, whereby each conformation depends on the coordinates $\boldsymbol{r}_i$ and the velocities $\boldsymbol{v}_i$ of the preceding conformation, as well as the potential functions $\boldsymbol{V}_i$ that define the interactions between the N atoms in the system. The potential functions $\boldsymbol{V}_i$ determine the time-dependent propagation of a simulated system in accordance with the conditions applied at the beginning of a simulation, which involves a set of atom coordinates $\boldsymbol{r}_0$ and a distribution of velocities $\boldsymbol{v}_0$ over all atoms. The potential functions $\boldsymbol{V}_i$ can be divided into two subgroups, the bonded and the non-bonded interactions and can furthermore be supplemented by additional potential terms, such as restraints or other external forces. A large variety of viable combinations of potential functions exist for protein simulations alone (Adcock et al. 2006) and the choice of which combination to apply for a given model is by no means trivial. In the following, the most common implementation in the GROMACS MD simulation engine (van der Spoel et al. 2005) is briefly outlined to introduce fundamental concepts.

1.1.3 Potential Functions

Bonded potential functions (eq.5-8) describe intramolecular interactions, i.e., define the geometry of molecules and are often modeled with a harmonic oscillator. The modeled interactions include bond-stretching between atoms *i* and *j* (eq.5), bond-angle vibrations between an atom triplet i-j-k (eq.6), proper dihedrals or torsion angles between the ijk and the jkl planes of consecutive atoms i-j-k-l (eq.7) and improper dihedrals, a potential function that is applied to specific groups of atoms to maintain chirality of atom groups or the planarity of planar groups (eq.8).

Bonded-potentials

Harmonic Bonding Potential $V_b(r_{ij}) = \frac{1}{2} k_{ij}^b (r_{ij} - b_{ij})^2$ (eq.5)

Harmonic Angle Potential $V_a(\theta_{ijk}) = \frac{1}{2} k_{ijk}^\theta (\theta_{ijk} - \theta_{ijk}^0)^2$ (eq.6)

Proper Dihedral Potential $V_d(\phi_{ijkl}) = k_\phi (1 + \cos(n\phi_{ijkl} - \phi_s))^2$ (eq.7)

Improper Dihedral Potential $V_{id}(\xi_{ijkl}) = \frac{1}{2} k_\xi (\xi_{ijkl} - \xi_0)^2$ (eq.8)

$k_{ij}^b, k_{ijk}^\theta, k_\phi, k_\xi$: force constants || $b_{ij}, \theta_{ijk}^0, \phi_s, \xi_0$: reference values || $r_{ij}, \theta_{ijk}, \phi_{ijkl}, \xi_{ijkl}$: variables

Non-bonded potential functions (eq.9-10) define intermolecular interactions and are most commonly split into two parts, the Coulomb potential (eq.9), which models the long-range attractive or repulsive forces (1/r) that occur due to atom charges, and the Lennard-Jones potential (eq.10), which is a unified representation of the short-range repulsive ($1/r^{12}$) and the medium-range attractive ($1/r^6$) forces that occur due to overlapping electron orbitals or atomic dipoles respectively.

Non-bonded potentials

Coulomb Potential $V_C(r_{ij}) = f \frac{q_i q_j}{\varepsilon_r r_{ij}} \qquad f = \frac{1}{4\pi\varepsilon_0}$ (eq.9)

Lennard-Jones Potential $V_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^6}$ (eq.10)

$q_i q_j$: charges || $C_{ij}^{(12)}, C_{ij}^{(6)}$: atom specific LJ-parameters || $\varepsilon_0, \varepsilon_r$: vacuum and relative permittivity

1.1.4 Statistical Mechanics

Macroscopic thermodynamic properties can be extracted from the information contained within simulation trajectories, but not without applying fundamental concepts of statistical mechanics, which are a necessary condition for relating the microscopic trajectory information to the thermodynamic state. The *thermodynamic state* of a system can be defined by a minimal set of state variables that is sufficient to uniquely determine all other thermodynamic properties of a system, while the *microscopic state* of a system can be defined by the positions $\boldsymbol{r}_i$ and the momenta $\boldsymbol{p}_i=\boldsymbol{m}_i\boldsymbol{v}_i$ of all N atoms in the system, or in statistical mechanics terms: a single point in the 6N dimensional phase space, whereby the factor 6 refers to the 3 positions and 3 momenta coordinates of each atom. In statistical mechanics, a set of points in phase space that satisfies the conditions of the same thermodynamic system state is referred to as an *ensemble*. MD simulation trajectories contain sets of positions $\boldsymbol{r}_i$ and momenta $\boldsymbol{p}_i$, which can belong to the same ensemble if all points in phase space represent the same thermodynamic state. To ensure that this is the case, a minimal set of thermodynamic state variables is kept constant throughout a simulation. One such set of variables is composed of the number of atoms N, the pressure P and the temperature T, which is referred to as the isobaric-isothermal or abbreviated the NPT ensemble. Strictly speaking, MD generates time averages, not ensemble averages and the latter is the prerequisite for the aforementioned statistical mechanics approach to relate the microscopic to the thermodynamic state in a statistically significant manner. However, the *ergodic hypothesis* states that given a sufficiently long sampling period, time-averaging and ensemble averaging are equal. Thus it becomes feasible to extract macroscopic properties from MD trajectories, provided the simulation is conducted under ensemble conditions.

1.1.5 Ensembles and Coupling Algorithms

The NPT ensemble is particularly useful for modeling common standard experimental conditions (P=1bar and T=298.15K). However, different ensembles exist that could otherwise be suitable, such as the microcanonical ensemble (NVE), with constant number of atoms N, volume V and energy E, or the canonical ensemble (NVT) with constant number of atoms N, volume V and temperature T. Both of the latter ensembles could, for

example, be utilized to extract the pressure from a trajectory after appropriate MD simulations within the ensemble constraints. An NPT ensemble can be generated by introducing coupling algorithms to periodically rescale the temperature and the pressure during a simulation.

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau_T} \qquad \text{(eq.11)}$$

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_P} \qquad \text{(eq.12)}$$

The Berendsen thermostat (eq.11) and barostat (eq.12) thereby weakly couple their respective quantities to a predefined constant value of an external bath (Berendsen et al. 1984). The Berendsen coupling scheme is very stable and efficient in converging systems that are far from equilibrium back to equilibrium values. By variation of the coupling constants τ_T and τ_p, it is also possible to modify the strength of the exponential decay when rescaling a respective value to its predefined constant value, which is an integral part of the solvent flux method (Chapter 4.3). However, the Berendsen coupling scheme does not rigorously model the canonical ensemble. Thus, for equilibrium simulations, where a more accurate ensemble representation is desired, it is considered more appropriate to choose other coupling schemes, such as Nosé-Hoover (Nose 1984; Hoover 1985) for temperature coupling and Parinello-Rahman (Parrinello et al. 1981) for pressure coupling. It is common practice in MD to perform energy minimization (Zimmermann 1991) and equilibration steps before extracting data from a simulation, to ensure that the system has adjusted to the model environment, and hereby the different coupling schemes each hold their own merit. While minimization does not require thermo- or barostats, equilibration is best conducted with the Berendsen coupling scheme until appropriate observable properties converge to their expected or constant values.

1.2 Model Simplifications and Validity

The practical application of MD generally constitutes a compromise between accuracy and efficiency. This requires careful consideration and integration of all the following concepts to attain reliable results coupled with high statistical significance, the latter of which is

mostly limited by the available computational resources but also by some method-inherent restrictions.

1.2.1 Empirical Force Fields

For reasons of computational efficiency, non-bonded interactions are commonly described in a pairwise additive manner, i.e., interactions between a single atom and the atoms of the surrounding system are calculated on an atom-to-atom basis and then summed up for a representation of the full interaction energy. This excludes polarization effects that stem from interactions of three or more atoms, which can lead to deviations between properties calculated by MD and experimental data. Applying polarizable models can improve the accuracy of MD, but at significant computational cost; however, development in this regard has intensified and the near future could potentially see current pairwise approaches fully replaced by polarizable models (Cieplak et al. 2009). A justification for the pairwise approach is that polarizability is modeled in an effective manner by fitting the potential function parameters to experimental data under specific environmental conditions. The sets of parameters and the respective potential functions that are thus obtained are referred to as force fields. During the last decades, a range of such force fields have been developed and are frequently refined in cycles of experimental validation and empirical adjustment to reproduce mainly condensed-state properties in the biomolecular context. Since the creation of robust models is a very laborious task, it is not efficient or economical to repeat this process separately for every individual system, and thus force fields are designed to be transferable within the scope of similar systems. Among the most prominent biomolecular force fields that were attained in this manner are the all atom AMBER (Cornell et al. 1995), OPLS-AA (Jorgensen et al. 1996) and the united atom GROMOS 53A5/53A6 force fields (Oostenbrink et al. 2004), the latter of which introduces further simplifications to increase computational efficiency by unifying certain groups of atoms to be represented by a single moiety, such as treating an aliphatic carbon atom and its hydrogen atoms as a single entity for calculations. A necessary feature of force fields of pairwise additive potentials is that often only a single or in the best case a limited number of descriptions exist for each atom, which is defined by an atom type and differs depending on its molecular context. An oxygen atom, for example, will be represented by a different charge,

dependent on whether it is involved in a hydroxyl group, in an ester bond or in a water molecule. Once an atom type is assigned to an atom in a simulation, it cannot be changed during the course of a simulation to a more appropriate description when it appears in a different molecular context. Consistency of the models applied in a simulation is therefore of utmost importance when utilizing transferable force fields. The choice of force field and the range of applicability will most probably never be a trivial matter, regardless of whether polarizability is added to the potential functions or not, and will likely always entail a residual degree of chemical intuition when evaluating the results of a simulation. Therefore, the validation of a MD model should always be a priority and if possible be correlated to experimental data, although paradoxically enough, MD as a method seems to be at its best in instances where experimental approaches fail.

1.2.2 Non-Bonded Interactions

A further factor of uncertainty in MD stems from the computational requirements in the calculation of non-bonded interactions. Typical protein simulation systems consist of 20 000 to 50 000 atoms, when explicitly including water solvent molecules, but more complex systems can quickly exceed the 100 000 atoms mark. Calculating the pairwise interactions between an atom and all other atoms, for every atom and at every time-step of a simulation, is computationally not feasible for systems of the respective sizes. This problem is enhanced by the fact that periodic boundary conditions are usually applied in MD simulation systems, to avoid the effects of finite borders and to mimic infinitely large systems. In practical terms, this implies that a simulation system is considered to be a cell that is bordered on all sides by copies of itself. To make the calculation of long-range interactions feasible, cutoffs can be introduced beyond which interactions are either ignored entirely or no longer explicitly calculated, e.g., by a smooth convergence of the potential to zero. The Coulomb potential (eq.9) is particularly problematic in this regard, due to its long range (potential decay of 1/r) and the Lennard-Jones potential (eq.10) to a lesser degree, due to its relatively short range (potential decay of $1/r^6$ and $1/r^{12}$). While introducing simple cutoffs to the Lennard-Jones interactions is thus commonly accepted, the approximation of the long-range Coloumb interactions is usually more refined. The most frequently applied methods stem from the Ewald summation, which greatly increases

the computational efficiency in systems with periodic boundary conditions. Ewald summation involves splitting the Coulomb potential into two terms, a short-range term that is calculated in real space and a long-range term that is calculated by Fourier transform in reciprocal space. Compared to the direct summation of pairwise interactions, the sum of the short-range and the sum of the long-range interactions both quickly converge, which significantly increases the overall efficiency when calculating interactions between charged atoms. This approach is further improved by methods such as the particle mesh Ewald algorithm (Darden et al. 1993; Essmann et al. 1995), which additionally improves the performance of the reciprocal sum.

1.3 Applied Molecular Models

When evaluating the concept of hydrophobicity with MD, special attention should be given to the water model. Numerous water models have emerged in literature in the past decades, many of which were crafted for general-purpose use in biomolecular systems. Development was thereby guided by the principle of reduced computational demand while forfeiting as little accuracy as possible. When simulating large solutes such as proteins in aqueous solution, water molecules often tend to make up more than 80% of the coordinates in a system, since ample space between the protein and the periodic boundaries needs to be provided, to avoid interaction artifacts between the protein and its self-images. While it is possible to implicitly treat solvation by dielectric continuum models, which is computationally very efficient, studies on the widely-used generalized Born continuum solvent model have shown that conformational ensembles for proteins greatly deviate from experimental data (Zhou 2003), and therefore simulations in full atomistic detail are commonly preferred if accuracy is valued above extended simulation times in biomolecular systems. A variety of rigid and flexible explicit water models exist for use in MD. Flexible water models should, in principle, offer a better representation of reality and were shown to be more accurate in reproducing a range of properties, such as viscosities (Medina et al. 2011) or dielectric constants (Mizan et al. 1996). However, for other properties, such as molecular structure or self-diffusion, rigid models were shown to perform amiably (Wu et al. 2006; Mizan et al 1996; Medina et al. 2011). The compromise between rigid water models and increased computational efficiency is thus the more

common model application in biomolecular systems. Another common simplification is constraining hydrogen bonds and water angle-vibrations. While this introduces a further potential to the calculations in MD and therefore increases the computational burden, it also allows for simulations to be conducted with longer time-steps, thus making longer simulation times more accessible overall. Moreover, the negative impact of maintaining water molecule rigidity can be significantly decreased by applying efficient constraint algorithms, e.g., SETTLE (Miyamoto et al. 1992), which was specifically developed to increase the computational efficiency when implementing water constraints. Another factor to consider is the consistency with commonly applied transferable force fields, which partially persists simply due to historical reasons. When choosing a water model, the choice should always incorporate the conditions under which solutes or co-solvents have been parameterized, since the water-solute interactions will have significant impact on the accuracy of solvation properties. The widely-used transferable force field OPLS-AA (Jorgensen et al. 1996), for example, has originally been crafted with the intention of modeling organic liquids in combination with rigid TIP3P and TIP4P water models (Jorgensen et al. 1983). These water molecules differ in the number of interaction sites that are included in the model. By increasing the number of interaction sites from three (TIP3P) to four (TIP4P), three for representing the atoms and one for representing the dipole between the oxygen and the lone pairs of the oxygen atom, the second peak in the oxygen-oxygen radial distribution function was much improved in comparison to experimental data (Jorgensen et al. 1983). This implies an overall improved model representation of the molecular structure of water. The TIP4P water model in combination with the OPLS-AA model for methanol were thus chosen for analyzing the mixture structure in respect to excess properties in the study "Incomplete Mixing versus Clathrate-Like Structures: A Molecular View on Hydrophobicity in Methanol-Water Mixtures" (Chapter 4.1) and also because high correlation with experimental data on relevant properties had previously been obtained for this force field combination by van der Spoel et al. (van der Spoel et al. 2006). Adding a higher level of detail to a molecular model does not, however, imply an improvement in terms of model accuracy. For example, TIP4P is capable of modeling the isothermal compressibility κ much better than TIP3P can, while in the case of the thermal expansion coefficient α, the roles are reversed (Table 1). It is unfortunately often not possible to adjust a model's parameters to represent all thermodynamic properties equally well.

Table 1: Calculated and experimental properties for liquid water at 25°C and 1atm. (reconstructed from Jorgensen et al. (Jorgensen et al. 1983), experimental data from Jorgensen (Jorgensen 1982) and Dorsey (Dorsey et al. 1940).

	SPC	TIP3P	BF	TIPS2	TIP4P	ST2	Expt.
d(g/cm³)	0.971	0.982	1.181	0.927	0.999	0.925	0.997
$-E_f$ (kcal/mol)	10.18	9.86	10.49	9.88	10.07	10.37	9.92
ΔH_{vap} (kcal/mol)	10.77	10.45	11.08	10.47	10.66	10.96	10.51
C_p (cal/mol deg)	23.4	16.8	16	18.9	19.3	22.2	17.99
10^5 α (deg^{-1})	58	41	59	88	94	-69	25.7
10^6 κ (atm^{-1})	27	18	18	56	35	63	45.8

Since it is difficult to predict potential artifacts when combining separately parameterized and optimized models, SPC/E (Berendsen et al. 1981), another common 3-site water model was chosen due to the consistency with lipid parameters of Berger et al. (Berger et al. 1997) in the study "Molecular Dynamics Simulations of Self-Emulsifying Drug Delivery Systems (SEDDS): Influence of Excipients on Droplet Nanostructure and Drug Localization." (Chapter 4.2). Berger lipids were originally designed for united atom bilayer simulations of the lipid dipalmitoylphosphatidylcholine (DPPC) in an NPT ensemble, whereby adjustments were mainly introduced to the Lennard-Jones parameters of the aliphatic CH_n groups, to improve the volume per lipid and lipid density in the layer phase, which resulted in a model that is able to reproduce aliphatic side chain aggregation of lipids in excellent agreement with experimental data (Berger et al. 1997). The integration of the CH_n group Lennard-Jones parameters into the united atom GROMOS 53A6 force field (Oostenbrink et al. 2004) is seamless, considering that the Berger parameterization was conducted using the bonded parameters of GROMOS. The Berger lipid model assumes charge neutrality in the CH_n groups in the aliphatic side chain (Figure 1, DPPC atoms C37-C50 and C18-C31). Although the GROMOS 53A6 force field has been developed for simulations of biomolecules in explicit water and in this context was shown to display improved hydration and solvation properties over previous iterations of the force field (Oostenbrink et al. 2004), a further improvement of the non-bonded interaction parameters for oxygen compounds, including ester bonds, was recently introduced by Horta et al. (Horta et al. 2011) that allows for a modeling of solvation properties with improved accuracy. Thereby, the 53A6 bond-stretching and the bond-angle vibration parameters around the oxygen compounds were modified according to optimized molecular geometries by quantum mechanics, as well as an extensive refinement of all

non-bonded parameter sets for each individual oxygen compound. The parameter adjustment was tested by the authors of the study against the observables liquid density, the enthalpy of vaporization and the solvation free energy of the oxygen compounds in water as well as in cyclohexane (Horta et al. 2011). This parameterization is referred to as $53A6_{OXY}$ and was used in conjunction with the remaining unmodified 53A6 standard parameters to model the polar moieties of mono-, di-, triglyceride molecules (Figure 1).

DPPC

TRIG

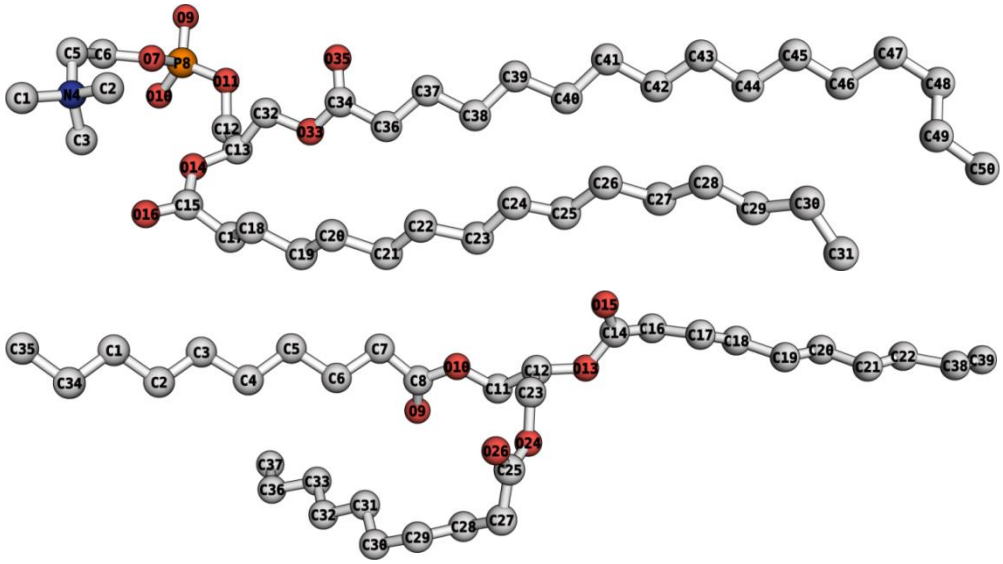


Figure 1: Comparison of united atom molecular structures of dipalmitoylphosphatidylcholine (DPPC) that was parameterized by Berger et al. and triglyceride (TRIG), for which the Berger lipid parameters were applied

Winger et al. reported that the behavior of polyethylene glycol (PEG), which is also applied as an excipient in the SEDDS study, significantly deviates from experimental observations when parameterized with the 53A6 force field, which was reportedly resolved by the 53A6_OE adaptation (Winger et al. 2009). Hereby, the Lennard-Jones parameters of the ether oxygen atoms were changed and the polarity between ether oxygen atoms and adjacent carbon atoms was increased (Figure 2). Since this parameter optimization by Winger et al. (Winger et al. 2009) was specific for PEG, it was preferred over the optimized $53A6_{OXY}$ parameter set.

PEG-6

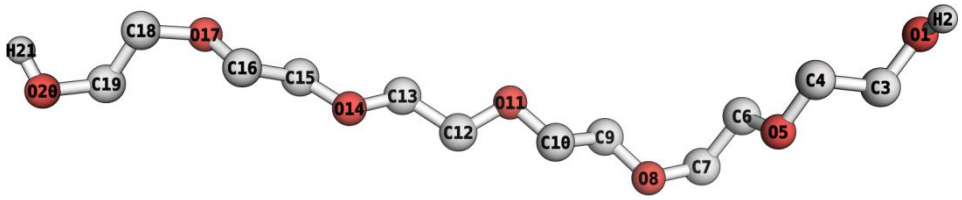


Figure 2: Polyethylene glycol (PEG-6) was parameterized in accordance to the 53A6_OE adaptation of the GROMOS 53A3 force field (Oostenbrink et al. 2004).

Cyclosporine A (CyA) was used as a model drug molecule in conjunction with varied excipient compositions consisting of mono-, di-, triglycerides and PEG-6. CyA is a circular peptide consisting of 11 amino acids (Figure 3), including the nonstandard amino acids D-alanine (Dal-1), N-methyl-L-leucine (Mel-2, Mel-3, Mel-4, Mel-8, Mel-10), N-methyl-L-valine (Mev-4), (4R)-4-[(E)-2-butenyl]-4,N-dimethyl-L-threonine (Meb-5), L-2-aminobutyric acid (Abu-6) and N-methylglycine (Sar-7). The parameters of CyA were parameterized according to the standard GROMOS 53A6 force field parameters.

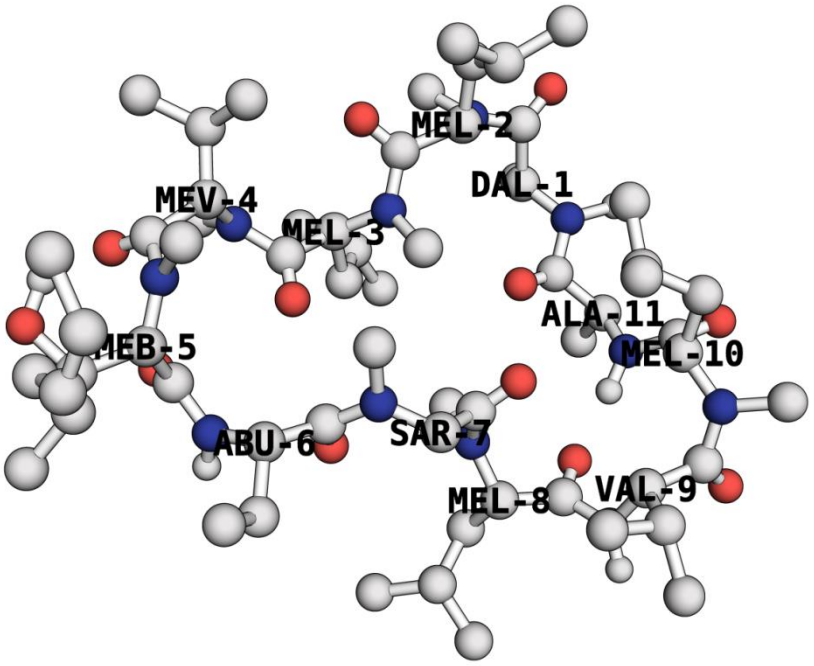


Figure 3: Cyclosporine A (CyA), a circular peptide containing the standard amino acids valine (VAL-9), L-alanine (ALA-11) and the nonstandard residues D-alanine (DAL-1), N-methyl-L-leucine (MEL-2, MEL-3, MEL-4, MEL-8, MEL-10), N-methyl-L-valine (MEV-4), (4R)-4-[(E)-2-butenyl]-4,N-dimethyl-L-threonine (MEB-5), L-2-aminobutyric acid (ABU-6), N-methylglycine (SAR-7).

For the study "The Solvent Flux Method (SFM): A Case Study of Water Access to Candida Antarctica Lipase B." (Chapter 4.3), the Berger lipid model was adapted to the OPLS-AA force field (Jorgensen et al. 1996) according to the reparameterization of Neale an Pomès (Neale and Pomès. 2008), who introduced changes to the dihedral parameters and thereby optimized the hydration free energies for both aqueous and hydrophobic environments. To model the interactions of the protein *Candida antarctica* lipase B (CALB) with the interface of a large triglyceride aggregation structure, a planar triglyceride layer was created by Gruber et al. (Gruber et al. 2012) (Figure 4). The protein CALB was parameterized according to the standard parameters from the OPLS-AA force field.

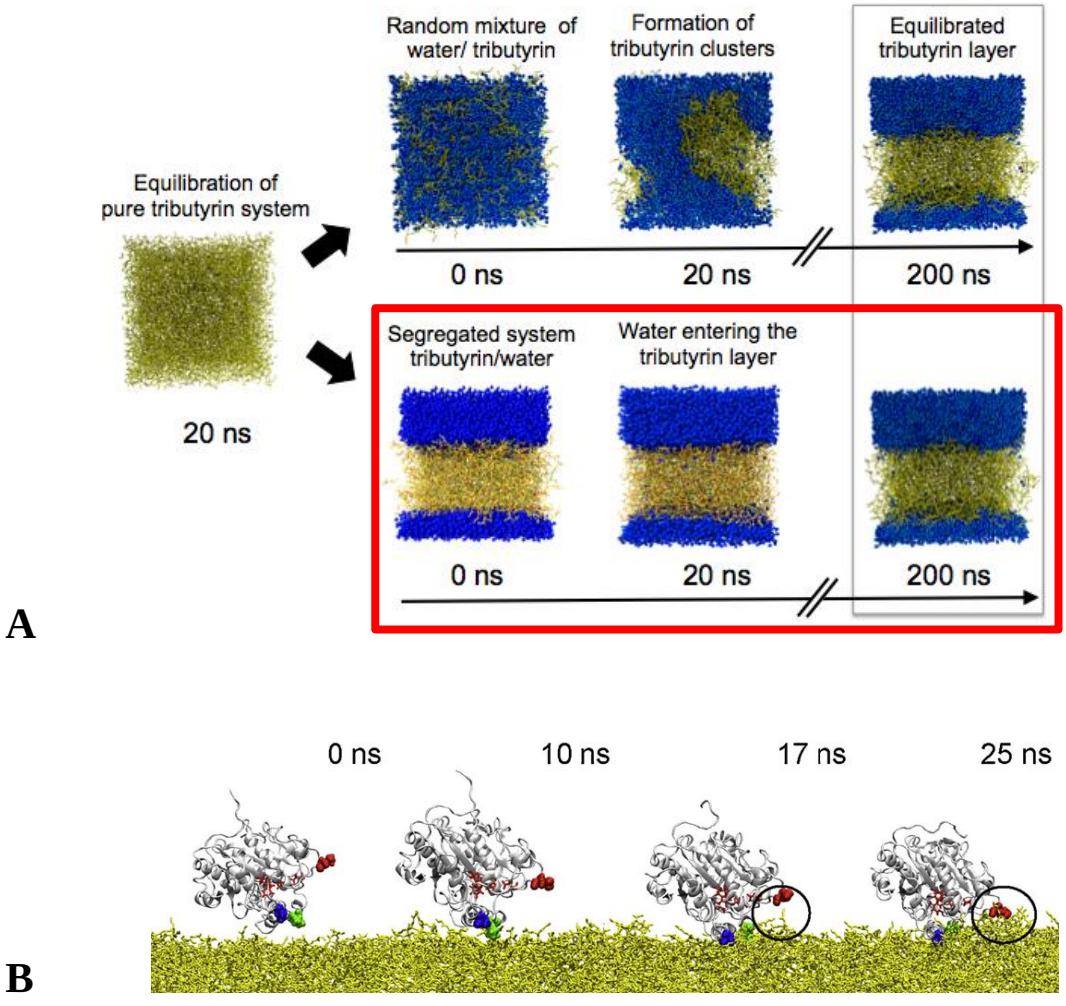


Figure 4: A) Triglyceride molecules were equilibrated without water for ~20ns in a cubic cell to form a triglyceride layer, a new system was created containing the triglyceride layer and water (red outline). This system was thoroughly equilibrated for 200ns+ in an NPT ensemble with anisotropic pressure coupling, thus separately rescaling the pressure in the x/y layer-plane and orthogonal to the plane in z-direction / B) The protein *Candida antarctica* lipase B was adsorbed to the equilibrated triglyceride-water interface during 25ns of simulation, a thorough equilibration of 500ns+ was performed hereafter (not shown) – pictures from Gruber et al. (Gruber et al. 2012).

2 Results

This section should be regarded as an extension of the studies presented in Chapter 4, which were conducted in full by the author of this dissertation. Therefore, only the major results that support the follow-up discussion (Chapter 3) on different hydrophobicity scales are briefly outlined herein.

2.1 Mixtures of Methanol and Water

In the study "Incomplete Mixing versus Clathrate-Like Structures: A Molecular View on Hydrophobicity in Methanol-Water Mixtures" (Chapter 4.1), the molecular origins of excess properties in methanol-water mixtures were explored. In this context, MD was applied to examine a range of structural (radial distribution functions, clustering) and dynamic properties (hydrogen bond lifetimes, molecular self-diffusion and rotational autocorrelation) for a range of methanol-water mixtures that represent the full extent of the mixture spectrum. Two hypotheses that seek to relate observable entropy deviations from ideal mixture behavior (excess entropy) to molecular mechanisms were thereby examined. The Frank and Evans hypothesis (Frank et al. 1945) proposes the existence of structured water molecules surrounding nonpolar solutes, while the more recent hypothesis of incomplete mixing (Dixit et al. 2002; Guo et al. 2003) proposes demixing of water and nonpolar moieties as a source for entropy deviations. To evaluate the Frank and Evans hypothesis in methanol-water mixtures, hydrogen bond lifetimes (Figure 5A), molecular self-diffusion (Figure 6A) and rotational mobility (Figure 5B) of water molecules adjacent to the methyl moieties of methanol, i.e., clathrate-like water, were investigated and compared to the respective properties of bulk water molecules. No deviations between clathrate-like water dynamics versus bulk water dynamics were observed and thus the Frank and Evans model was not supported by the results of the study. To examine incomplete mixing in methanol-water mixtures, the molecular association bias parameter M_{AB} was conceived that relies on the statistical evaluation of nearest neighbor relationships. Hereby, the number of molecules of type B that are found as nearest neighbors to molecules of type A are related to the respective number of neighbors that

would be expected of an ideal molecular distribution, whereby A and B can be either water or methanol. For example, M_{wm} denotes the number of nearest methanol molecule neighbors to water molecules, whereas M_{ww} denotes the number of nearest water molecule neighbors to other water molecules (Figure 6B). M_{AB}=1 is expected of an ideal distribution, while M_{AB}>1 is referred to as a positive association bias and M_{AB}<1 as a negative association bias.

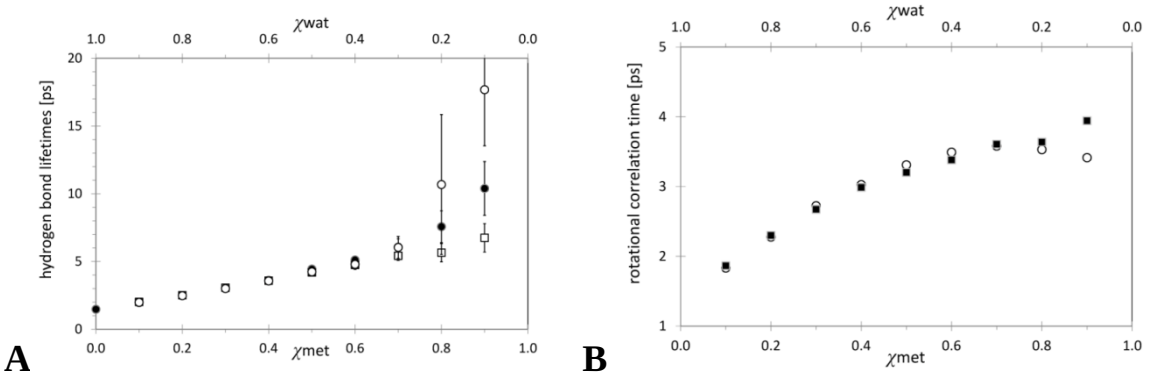


Figure 5 A: Hydrogen bond lifetime τww between water molecules - clathrate-like water (□), bulk water (○) and all water molecules (●); deviations were only observed for mixture systems where the low water content no longer allowed for significant statistics, which is marked by high standard deviations that were averaged over multiple trajectory intervals. / B: Deviations between rotational autocorrelation for the clathrate-like water group (○) and the bulk water group (■) were minor. Standard deviations were averaged over multiple trajectory intervals and ranged from 0.01-0.1ps.

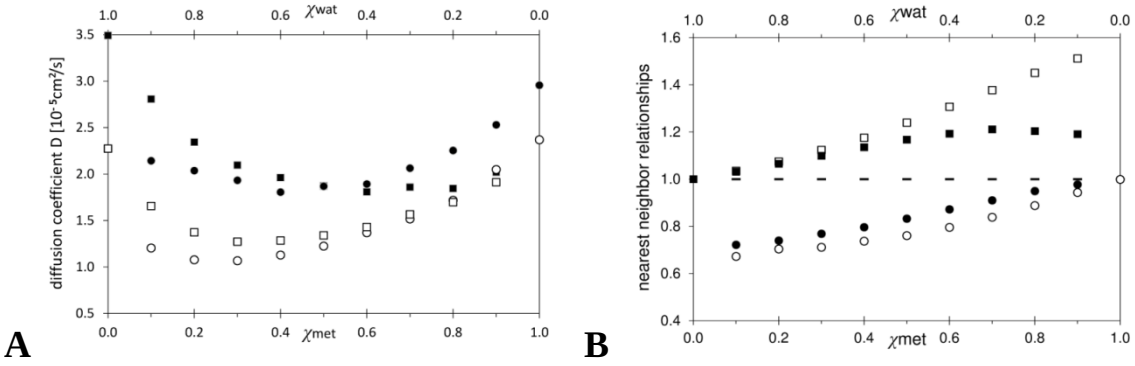


Figure 6: A) Molecular self-diffusion coefficients D for simulated water (TIP4P: ■), experimental water (□), simulated methanol (OPLS-AA: ●) and experimental methanol (○). Standard deviation and detailed statistics are provided in the supporting information of Chapter 4.1. Experimental data was obtained from Derlacki et al. (Derlacki et al. 1985). Self-diffusion of clathrate-like water and bulk water was in very close agreement with the overall diffusion (not shown). / B) Molecular association bias between different molecules - water-water M_{ww}: ■ / methanol-water M_{mw}: □ / water-methanol M_{wm}: ● / methanol-methanol M_{mm}: ○. Values differ significantly from ideal mixture expectations (dashed line). Standard deviations were averaged over multiple trajectory intervals and ranged from 0.15-3.45.

Significant deviations in structural association from ideal mixture behavior were observed in the molecular association biases and thus the hypothesis of incomplete mixing was supported by the results of the study. Furthermore, the molecular basis for the minimum in molecular self-diffusion was investigated (Figure 6A), which was found to coincide with the formation or breakdown of the hydrogen bonding network of water in that both are observable at approximately equimolar ratio (Figure 7). It was possible to contextualize this observation with the aforementioned structural and dynamic properties to formulate a holistic perspective on the molecular mobility in methanol-water mixtures (Chapter 4.1). The mixture system was thereby perceived to be separated into two hydrogen bonding regimes: hydrogen-bonded water clusters embedded in methanol for mixtures with low water content and methanol molecules within a system-wide hydrogen-bonded water network for mixtures with high water content.

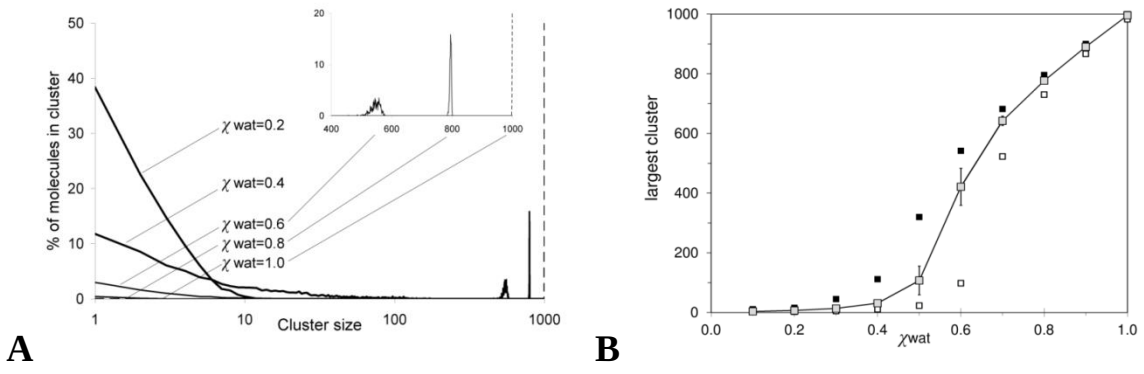


Figure 7: Clustering with cutoff criterion 3.5Å in methanol-water mixture systems containing 1000 molecules – A) Distribution (percentage) of water molecules that are involved in clusters of different size N; for molar fraction χwat<0.5 water is contained in small clusters of size N<100, while for χwat>0.5 all water molecules are interconnected in a system-wide hydrogen bonding network. Continuous lines are applied for the purpose of visual distinction. / B) Largest water molecules clusters in methanol-water mixtures (biggest overall cluster: ■ / largest cluster on average: ■ / smallest overall cluster: □); at χwat~0.5, the system-wide hydrogen bonding network forms or collapses upon changing the mixture composition.

2.2 Lipid Droplet Formation: Self-Emulsifying Drug Delivery Systems

The molecular association bias analysis was extended to systems of lipid emulsions in the MD study "Molecular Dynamics Simulations of Self-Emulsifying Drug Delivery Systems (SEDDS): Influence of Excipients on Droplet Nanostructure and Drug Localization." (Chapter 4.2). Herein, the formation of SEDDS droplets was simulated to ascertain whether insights into the droplet nanostructure may be utilized for the design of drug delivery formulations. Common SEDDS formulation excipients were thereby systematically varied, including the fatty acid chain length, PEG-6 concentration, mono- and diglyceride content as well as droplet loading. The analyzed properties included the molecular association bias, to contextualize the implications of excipient changes with changes in droplet nanostructure, as well as the immersion depth of the model drug molecule cyclosporine A (CyA), to evaluate whether changes in droplet nanostructure induce different forms of drug localization within SEDDS droplets. In the following, all results associated with the nanostructure and drug localization are presented in form of depictions of droplets. A full quantitative description in form of molecular association biases and CyA immersion depths may be referenced in Chapter 4.2. A standard system consisting of capric (C10) triglyceride molecules, a single drug molecule and water served as a reference to evaluate the implications of excipient variations and is outlined in green in the following figures. Drug loading was investigated by increasing the amount of drug molecules contained in the system. A preferred localization of the circular peptide CyA in the outer periphery of the droplet and in close proximity to the aqueous solvation shell was determined, as well as a negligible impact of CyA on the nanostructure of the droplet (Figure 8). In the system containing 10 CyA molecules, CyA displayed an increased molecular association bias towards the aqueous environment, which implies drug precipitation due to droplet overloading.

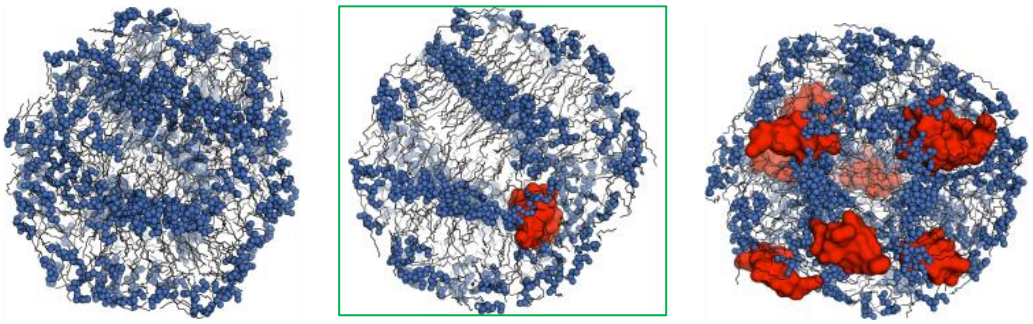


Figure 8: Drug loading variation in SEDDS droplets consisting of drug molecule CyA (red surface) in capric triglyceride molecules with polar (blue spheres) and nonpolar (grey lines) moieties – left: 0x drug molecules, middle: 1x drug molecule, right: 10x drug molecules. The standard reference system is outlined in green.

The fatty acid chain lengths of triglyceride excipients were varied to reveal distinctively different nanostructures, which were classified as "unstructured" for caproic (C6), "lamellar-like" for capric (C10) (standard system) and "vesicle-like" for myristic (C14) triglyceride systems (Figure 9). The propensity of CyA to accumulate in the interfacial region of the droplet and the aqueous environment was maintained in all systems. However, CyA was shown to intermittently immerse more deeply into the droplet nanostructure in the capric (C10) system, which was attributed to the highly ordered lamellar-like self-association structure of polar triglyceride moieties and accredited to the fact that a stabilization of the polar moieties of CyA is thereby possible on two sides of CyA, instead of a single side as in vesicle-like systems.

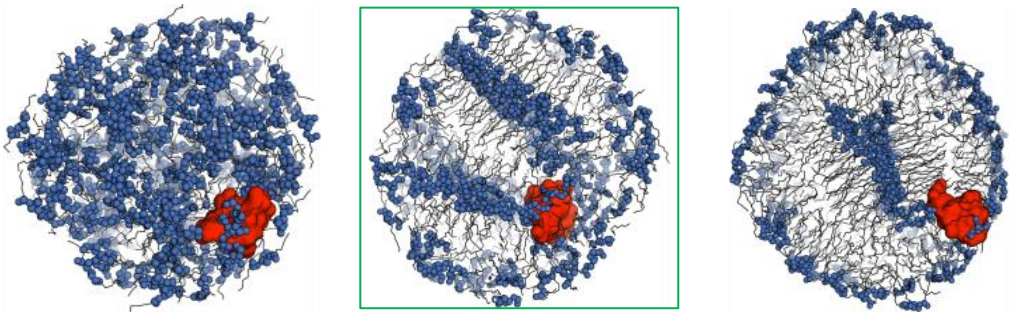


Figure 9: Fatty acid chain length variation in SEDDS droplets consisting of drug molecule CyA (red surface) in capric triglyceride molecules with polar (blue spheres) and nonpolar (grey lines) moieties – left: caproic (C6) triglycerides, middle: capric (C10) triglycerides, right: myristic (C14) triglycerides. The standard reference system is outlined in green.

Poly(ethylene glycol) (PEG-6) was added to the capric (C10) standard system in 1:4 and 2:3 ratios respectively. The addition of PEG-6 enhanced the preexisting molecular association bias and accumulated as an encasing shell on the surface of the droplet, thereby partially shielding CyA from the aqueous environment (Figure 10). CyA was therefore slightly more immersed in the droplet, particularly in the C10 3:2 PEG-6 system. The nanostructure of triglyceride was preserved. The positive molecular association bias of CyA towards triglyceride molecules implied that triglyceride molecules facilitate the solubilization of CyA and not PEG-6.

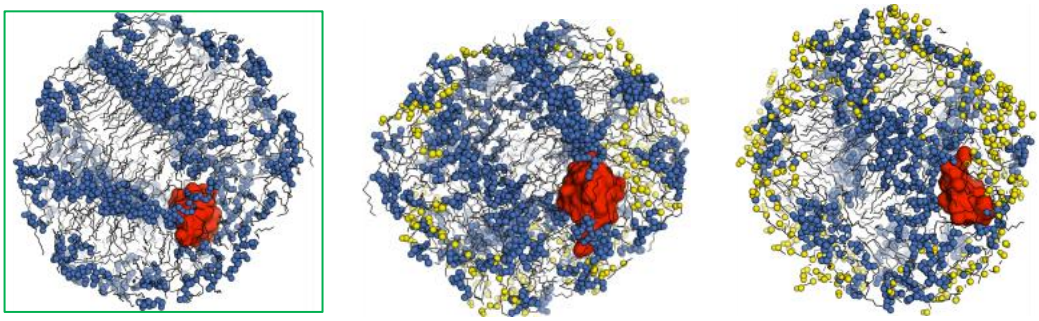


Figure 10: Addition of surfactant poly(ethylene glycol) (PEG-6) (polar moieties: yellow) to SEDDS droplets consisting of drug molecule CyA (red surface) in capric triglyceride molecules with polar (blue spheres) and nonpolar (grey lines) moieties – left: no PEG-6 added, middle: ratio of triglyceride 4:1 PEG-6, right: ratio of triglyceride 3:2 PEG-6. The standard reference system is outlined in green.

The addition of capric (C10) di- and monoglyceride to the capric (C10) triglyceride standard system in 1:1 ratios yielded significant changes in the overall droplet nanostructure and the localization of CyA therein. The polar moieties of di- and, to a higher degree, monoglyceride molecules preferentially associated at the aqueous interface, while their respective nonpolar moieties protruded into the droplet nanostructure. This was perceived as the major cause for the observable changes in respect to the standard system. While the preferential localization of CyA close to the droplet surface was still maintained for diglyceride, the impact of monoglyceride nanostructure perturbance in the droplet interior induced the formation of a core structure of polar triglyceride moieties, which proved sufficient to facilitate a nanoenvironment that was preferential to CyA solubilization (Figure 11).

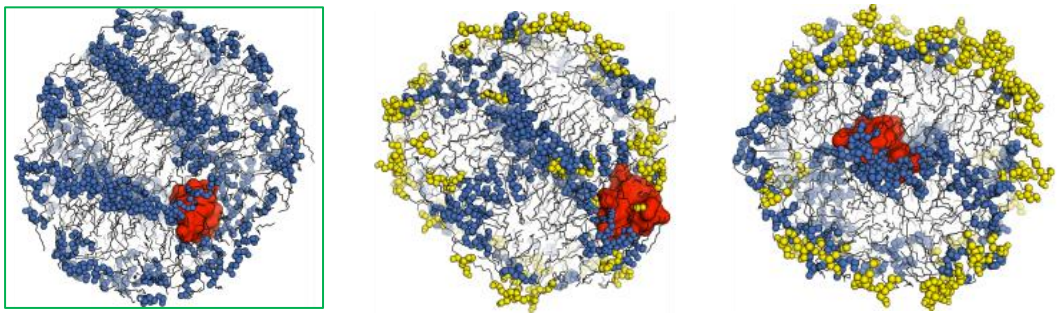


Figure 11: Addition of di- and monoglycerides (polar moieties: yellow) to SEDDS droplets consisting of drug molecule CyA (red surface) in capric triglyceride molecules with polar (blue spheres) and nonpolar (grey lines) moieties – left: only triglyceride, middle: ratio of triglyceride 1:1 diglyceride, right: ratio of triglyceride 1:1 monoglyceride. The standard reference system is outlined in green.

The standard system was furthermore decreased and increased in size by variation of the number of triglyceride molecules contained in the system (Chapter 4.2). The nanostructure was thereby found to be size-invariant within the computed size range, which was restricted to droplet diameters of no more than 10nm due to the full atomistic model description of the systems. Although essentially no water molecules were found within the droplet interior in all systems, the aqueous and the triglyceride phases were not found to be clearly separated, i.e., the triglyceride-water interface was not structured in a manner that minimizes the water-triglyceride surface area. Instead, water was found to locally penetrate the droplet for several angstroms (Figure 12). This observation gave cause to study water accessibility to planar triglyceride-water interfaces, which is outlined in the following section.

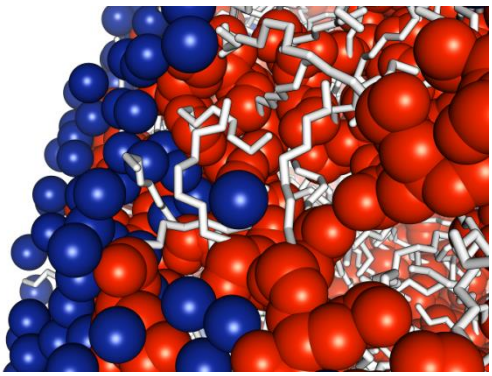


Figure 12: Interface of SEDDS droplet with aqueous environment, water (blue) is shown to slightly penetrate into the the interior of the droplet at self-associated polar triglyceride moieties (red).

2.3 Triglyceride-Water Interfaces: Large-Scale Aggregates in Water

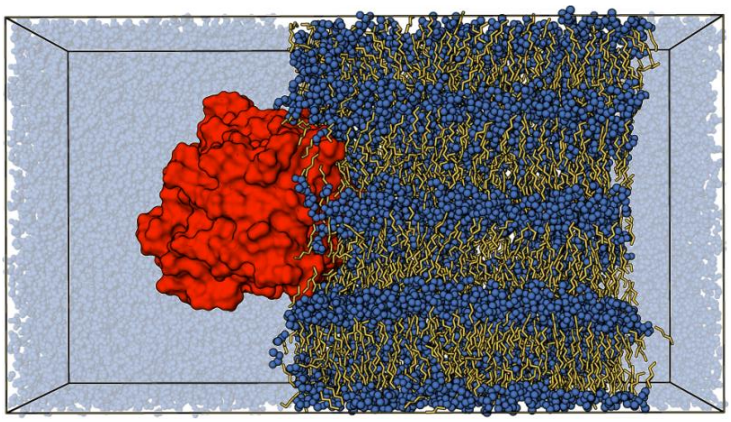


Figure 13: Simulation system consisting of *Candida antarctica* lipase B attached to an interface of triglyceride (here: tricaproin (C6)) and water (red surface: protein / dark blue spheres: polar triglyceride moieties / orange sticks: nonpolar triglyceride molecules / water: light blue spheres in the background).

The data presented in this section is part of an ongoing study, which has not yet been submitted for review. Three planar triglyceride layer systems were created in accordance with the preliminary work conducted by Gruber et al. (Gruber et al. 2012) (Figure 4), to model the interaction of the protein *Candida antarctica* lipase B (CALB) with large-scale triglyceride-water interfaces in an oil-in-water emulsion (Figure 13). The systems consisted of water, the protein CALB and three triglycerides of different fatty acid chain lengths, tributyrin (C4), tricaproin (C6) and tricaprylin (C8). To examine the structure of these interfaces, atomistic profiles of the system components were created along the coordinate z-axis orthogonal to the triglyceride layer plane (Figure 14). The protein was found to immerse into the triglyceride layer rather than to adsorb onto its surface, while water was found to enter the layer in significant amounts, effectively creating a transitional water-rich interphase between the water bulk and the relatively dry core region of the triglyceride layer. The water-rich interphase thereby coincides with the protein immersion depth of approximately 5Å.

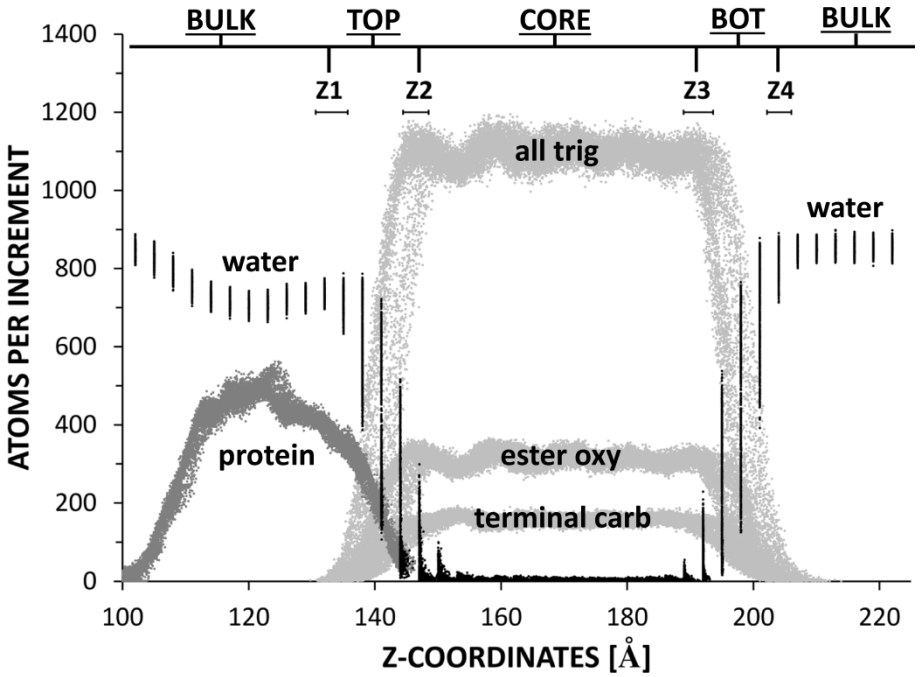


Figure 14: Molecular structure of the interface of protein triglyceride (tributyrin) and water during 100ns simulation; the coordinate profile depicts the number of atoms in 3Å increments along the axis orthogonal to the layer plane (z-axis), triglyceride atoms are additionally split into groups containing all atoms (all trig) as well as the polar oxygen atoms of triglyceride ester bonds (ester oxy) and the terminal carbon atoms of the fatty acid side chains (terminal carb), to illustrate that triglyceride molecules are not oriented with their polar moieties towards the bulk water phase. Four phases were defined in the system in respect to water content, BULK water, TOP and BOT as the water-rich interphases of the triglyceride layer and the CORE region of the layer. Error bars for Z1-Z4 mark the fluctuations of z-coordinates during simulation.

Four different phases were differentiated for analysis, the water-rich triglyceride-water interphase including the immersed protein (TOP: Z1-Z2), the relatively dry triglyceride layer core region (CORE: Z2-Z3), the water-rich triglyceride-water interphase without protein (BOT: Z3-Z4) and the bulk water phase (BULK) on either side of the layer. The phase boundaries of Z2 and Z3 were thereby defined as the z-coordinates where the measured atoms per increment reached 90% of the maximal value for the first time in respect to the bulk water phase from either side of the layer, while Z1 and Z4 marked the z-coordinates of the first observable triglyceride atom in respect to the bulk water phase. The positions of Z1-Z4 varied during a 100ns MD simulation despite a periodic removal of the center of mass movement of the triglyceride layer. The variations were attributed to triglyceride layer fluctuations of approximately 8Å along the z-axis (Figure 15A). The protein immersion depth was obtained by calculating the fluctuations in the z-component of the center of mass distance between the protein and the triglyceride layer and was shown not to correlate with the triglyceride layer fluctuations, despite being similar in scale

(Figure 15B). It was therefore concluded that the protein is loosely embedded in the water-rich triglyceride-water interphase, which may facilitate a high degree of protein mobility or "swimming".

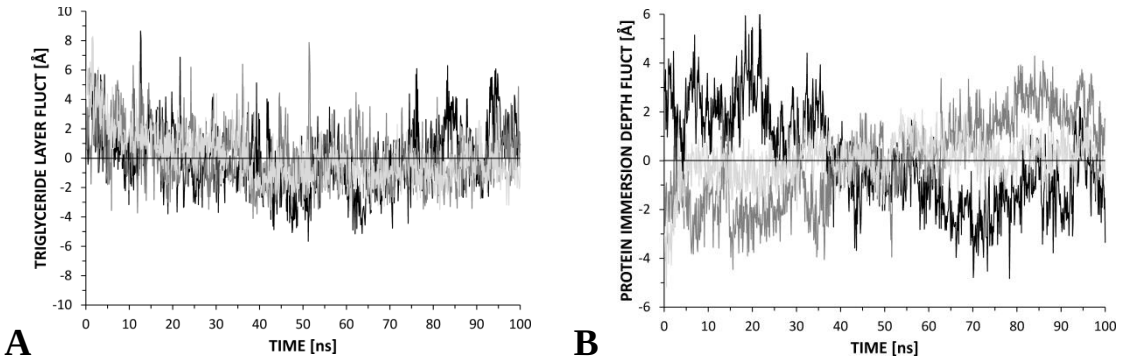


Figure 15: A) Triglyceride layer fluctuations, calculated as thickness variations in z-direction of the layer – black: tributyrin (C4), dark grey: tricaproin (C6), light grey: tricaprylin (C8) / B) Protein immersion depth fluctuations in the triglyceride-water interface, calculated as the deviations in the z-component of the center of mass distances between the protein and the triglyceride layer.

Table 2: Density of water in the different phases BULK, TOP, CORE and BOT in triglyceride layers of tributyrin (C4), tricaproin (C6) and tricaprylin (C8).

	density [g/l]			
water density in ...	BULK	TOP	CORE	BOT
tributyrin (C4)	1003.87 ± 11.00	673.80 ± 72.70	8.35 ± 2.04	440.75 ± 54.47
tricaproin (C6)	1021.32 ± 12.25	647.69 ± 82.76	5.19 ± 1.78	211.99 ± 38.62
tricaprylin (C8)	965.76 ± 20.59	571.51 ± 75.74	5.31 ± 2.33	199.41 ± 35.85

The density of water in the different triglyceride layer phases BULK, TOP, CORE and BOT was calculated to quantify the observed water distribution (Table 2). Within the defined boundaries, the density in the TOP phase was consistently higher than in the BOT phase, which implies that the presence of the protein at the triglyceride-water interface may induce an increased water penetration into the layer. For tributyrin (C4), water densities in both TOP and BOT phases retained about half of the water density of the BULK phase. The BULK density correlated well with the experimental value ρwat=997g/l (Mikhail et al. 1961). The relatively "dry" CORE phase still displayed water densities in the range of 5-8% in respect to the BULK. Elongating the fatty acid chain length of triglyceride with tributyrin (C4), tricaproin (C6) and tricaprylin (C8) caused a sharp decline in water

density, particularly for the BOT phase, which implies that a higher density of polar triglyceride moieties might be favorable to water influx.

The mobility of water molecules was characterized by calculating the molecular self-diffusion coefficients in all respective phases (Figure 16). For the tributyrin (C4) system, the water self-diffusion coefficient $D_{BULK}=1.88\cdot10^{-5}$ [cm²/s] of the BULK phase was slightly lower than the experimental value of systems consisting of pure water $D_{EXP}=2.16\cdot10^{-5}$ [cm²/s] (Derlacki et al. 1985), which was attributed to the presence of two continuous triglyceride surfaces that restrict water mobility along the z-axis. In the water-rich interphases TOP and BOT, water maintained more than half of the mobility of bulk water, with $D_{BOT}=1.34\cdot10^{-5}$ [cm²/s] and $D_{TOP}=1.10\cdot10^{-5}$ [cm²/s]. The presence of the protein was found to have a detrimental effect on water mobility (D_{TOP}), which may be related to protein hydration or simply to the additional surface area, but might also be influenced by factors pertaining to the interaction of the protein with the triglyceride-water interface. Even in the relatively dry triglyceride layer core, residual water diffusion was maintained with $D_{CORE}=1.10\cdot10^{-5}$ [cm²/s]. Analysis of water self-diffusion for tricaproin (C6) and tricaprylin (C8) yielded similar diffusion constants. The molecular origin of this overall high residual mobility was perceived to lie in the lamellar structural order of the triglyceride layer, in which the nonpolar triglyceride fatty acid side chains were arranged in tangential orientation to the triglyceride layer plane (Figure 13). The nonpolar triglyceride moieties were thereby significantly exposed to the water bulk, while polar triglyceride moiety association induced semi crystalline - structures and polar nanoenvironments that expanded throughout the triglyeride layer (Figure 16). The polar nanoenvironments of the triglyceride layer were perceived to be potential mediators for water mobility, which may facilitate hydrogen bonds to diffusing water molecules. Since evidence for this claim should be reflected in the molecular structure of the triglyceride layer, the molecular association bias in hydrogen bonding distance (cutoff 3.5Å between water oxygen atoms) was analyzed (Table 3).

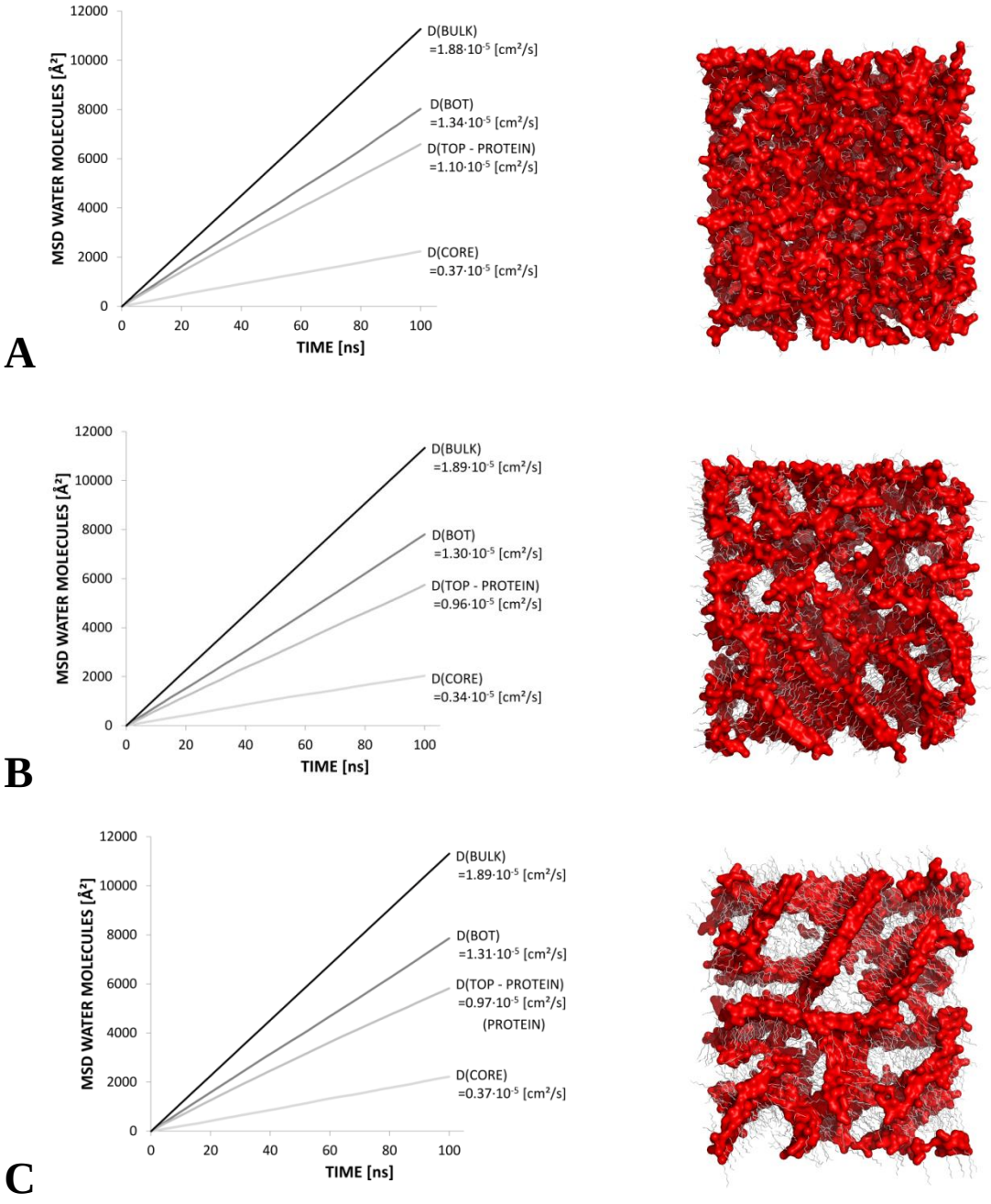


Figure 16: Mobility of water in triglyceride-water interfaces of A: tributyrin (C4), B: tricaproin (C6), C: tricaprylin (C8) – Left Column: Mean square displacement (MSD) and corresponding molecular self-diffusion coefficients of water in the phases BULK, TOP, CORE and BOT as defined by atomistic profiles (Figure 14). / Right Column: Depiction of polar triglyceride moieties (red surface), which may contribute to the water mobility in the triglyceride layer.

Table 3: Molecular association bias (mol. assoc. bias) of water molecules in the triglyceride layer phases TOP, BOT and CORE for tributyrin (C4), tricaproin (C6) and tricaprylin (C8) with cutoff criterion 3.5Å, whereby Nwe denotes the bias of water oxygen towards polar ester ogyxgen atoms of triglyceride, Nwa the bias of water oxygen towards the aliphatic carbon atoms of triglyceride and Nww the bias of water oxygen to other water oxygen atoms. The reference value for Nwe, Nwa and Nww in a random distribution is 1.

		tributyrin (C4) mol. assoc. bias			tricaproin (C6) mol. assoc. bias			tricaprylin (C8) mol. assoc. bias		
TOP	M_{we}	0.54	±	0.07	0.62	±	0.09	0.74	±	0.14
	M_{wa}	0.19	±	0.02	0.16	±	0.02	0.15	±	0.03
	M_{ww}	2.50	±	0.41	4.23	±	0.62	4.63	±	0.65
CORE	M_{we}	1.83	±	1.88	2.04	±	2.37	2.13	±	3.11
	M_{wa}	0.60	±	0.58	0.46	±	0.51	0.41	±	0.48
	M_{ww}	11.99	±	12.90	26.26	±	18.09	32.77	±	16.34
BOT	M_{we}	0.58	±	0.08	0.77	±	0.11	0.64	±	0.07
	M_{wa}	0.21	±	0.03	0.19	±	0.03	0.14	±	0.02
	M_{ww}	1.49	±	0.18	1.58	±	0.22	1.67	±	0.21

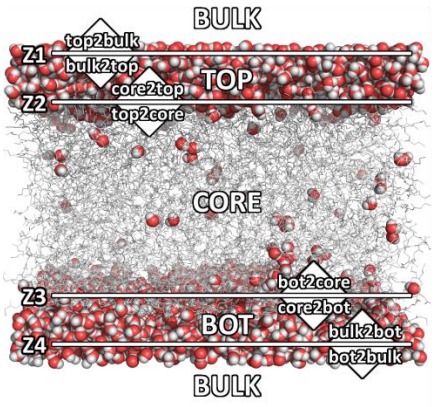
The preferential self-association of water (M_{ww}) was apparent in all phases of the triglyceride layer. In the TOP phase, M_{ww} was consistently higher than in the BOT phase, which implies that the presence of the protein has a direct influence on the water structure at the triglyceride-water interface. In the CORE phase, M_{ww} reached the highest values, which suggests that the residual water contained therein has a high propensity to cluster. M_{we}, the association bias of water towards the polar ester oxygen atoms of triglyceride was consistently higher in TOP, CORE and BOT phases than M_{wa}, the respective association bias towards the nonpolar triglyceride atoms. Elongating the nonpolar moiety of the triglyceride fatty acid side chains, from tributyrin (C4), tricaproin (C6) and tricaprylin (C8), enhances the respective association biases; M_{ww} and M_{we} both significantly increased, while M_{wa} decreased. The observed molecular association biases support the interpretation of polar nanoenvironments as mediators for water diffusion; as polar atoms become scarce in relation to nonpolar atoms for systems of longer triglyceride fatty acid side chains, the aggregated polar triglyceride substructures become more segregated and thus water molecules associate more strongly to this polar nanoenvironment while still retaining their high degree of molecular mobility (Figure 16).

To obtain a holistic perspective of water mobility in the triglyceride layer, the exchange of water molecules between the BULK, TOP, MID and BOT phases was evaluated by flux

calculations (Table 4). Hereby, the number of water molecules that moved out of a particular phase and into another phase after discrete time-steps of 30ps was summed up and averaged over a 30ns simulation. The resulting value was normalized to a surface area parallel to the layer surface plane (x-y plane) and thus the respective flux was determined in molecules/(Å²µs). As expected, the highest water exchange occurred between the BULK phase and the water-rich transitional TOP and BOT phases, while the water exchange between the TOP and BOT phases and the CORE phase was significantly lower. However, the flux rates for water molecules from BOT phase to CORE phase were found to be a factor 10 higher than the flux rates from TOP phase to CORE phase, which implies an involvement of the protein in water retention. The more frequent exchange of water molecules between the BULK and TOP phases compared to BULK and BOT phases may be ascribed to the protein as well, which effectively constitutes a barrier between bulk water and the triglyceride surface and thus minimizes the solvent accessible surface area of the triglyceride layer.

Table 4: Measurement of the water flux between the BULK, TOP, CORE and BOT phases in the triglyceride layer, for tributyrin (C4), tricaproin (C6) and tricaprylin (C8).

	tributyrin (C4) molecules/[Å²µs]	tricaproin (C6) molecules/[Å²µs]	tricaprylin (C8) molecules/[Å²µs]
top2bulk	519.38 ± 166.74	407.94 ± 154.49	376.31 ± 126.42
bulk2top	519.68 ± 166.27	407.66 ± 155.11	374.75 ± 124.23
core2top	3.29 ± 3.85	2.66 ± 3.54	0.76 ± 5.47
top2core	3.30 ± 3.89	2.67 ± 3.53	0.76 ± 1.64
core2bulk	1.11 ± 1.89	2.65 ± 3.65	3.42 ± 3.84
bulk2core	1.18 ± 22.40	2.66 ± 16.51	3.52 ± 15.88
bot2core	23.71 ± 20.67	14.28 ± 13.33	15.88 ± 12.56
core2bot	23.78 ± 20.95	14.29 ± 13.00	15.97 ± 12.77
bulk2bot	688.50 ± 205.49	578.50 ± 170.62	512.09 ± 108.04
bot2bulk	689.52 ± 206.92	578.33 ± 169.91	512.28 ± 109.43



To clarify the involvement of the protein in potential water retention as well as the increased self-association of water molecules and the heightened water density in the TOP phase, the solvent flux method (SFM) was developed as an accelerated MD approach to model the influx of water molecules into the protein CALB.

2.4 Water Influx in Proteins: The Solvent Flux Method (SFM)

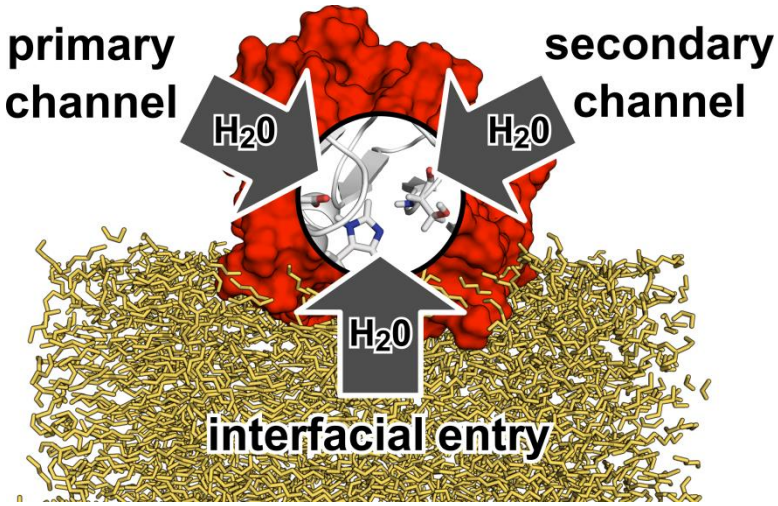


Figure 17: *Candida antarctica* lipase B (red) attached to a triglyceride layer (orange) in water (not shown) - a primary water channel was validated by the solvent flux method SFM (Chapter 4.3), a potential secondary channel was proposed and significant water influx to the active site (sticks) via the substrate channel was observed, which is fully enveloped by triglyceride molecules of the layer phase.

In the study "The Solvent Flux Method (SFM): A Case Study of Water Access to Candida Antarctica Lipase B.", the influx of water molecules to the active site cavity of the enzyme CALB was modeled for the identification of potential water channels and positions relevant to the solvent-related engineering of hydrolase applications (Chapter 4.3). Hereby, external forces were periodically introduced after Δt_{ITER} in the form of a radial velocity pulse onto all water molecules contained within a sphere enclosing the protein and its first solvation shell, which allowed for an accelerated passage of water molecules over energy barriers that obstruct the entry to the active site cavity. Simultaneously, a single water molecule was removed in close proximity to a reference atom in the active site cavity with the same periodicity Δt_{ITER} as the introduced velocity pulse, which induced a water concentration gradient and provided additional space for water to occupy after influx. Moreover, the periodic removal of water constitutes a model for enzymatic hydrolase activity, which consumes water molecules during the cleavage of ester bonds of triglyceride. Furthermore, after a given number of SFM iterations n_{EVAC}, it was ensured that the active site cavity of CALB had been evacuated of any previously present water molecules and therefore any water molecules removed for iterations $n_i > n_{EVAC}$ necessarily

had to have passed through a potential entryway in the protein structure to reach the active site. Hence, for n_i it was possible to reconstruct the pathways of water molecules entering the active site cavity in form of pseudo-trajectories with the time-step of Δt_{ITER}. This allowed for a highly accelerated characterization of the inherently slow water influx process and the identification of potential avenues of water entry (Figure 18 and Table 5).

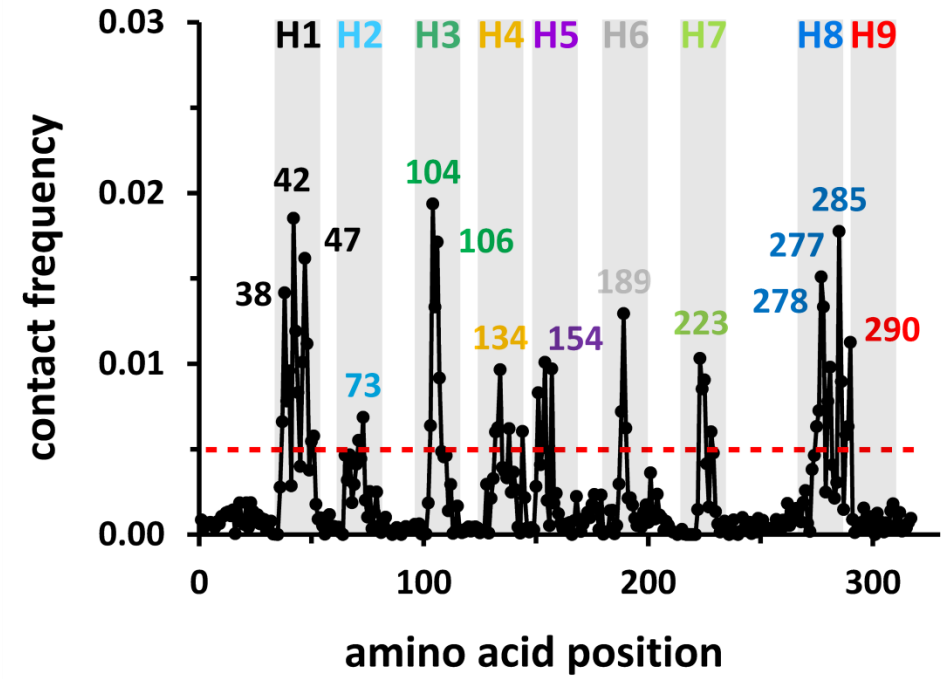


Figure 18: Contact frequencies between water molecules and amino acid positions during water influx to the active site of *Candida antarctica* lipase B (details in Chapter 4.3) define hotspot areas H1-H9 and hotspot positions that were defined as peaks above a threshold (red dotted line).

All observed hotspot areas and hotspot positions were evaluated on the basis of the relative contact frequency (Table 5), the positioning in the protein structure (Figure 19A) and the superimposition of influx pathways (Figure 19B).

Table 5: Evaluation of hotspot areas H1-H9 and hotspot positions according to relative contact frequencies (rel. freq.) and localization in the protein structure (details in Chapter 4.3).

	pos.	AA	property	rel. freq.	d_{act} [Å]	positioning	Assessment	chan
	38	**PRO**	hphob	0.71	5.7	periphery	adjacent to primary water channel	-
H1	**42**	**THR**	**hphil**	**1.00**	**7.8**	**periphery**	**primary water channel**	**PWC**
	47	**SER**	**hphil**	**0.85**	**8.5**	**periphery**	**primary water channel**	**PWC**
H2	**73**	**LEU**	hphob	0.27	10.8	surface	probable barrier, adjacent to Lys290	-
H3	**104**	**TRP**	hphob	0.96	4.8	active site	active site	-
	106	**GLN**	**hphil**	**0.98**	**4.8**	**active site**	**primary water channel**	**PWC**
H4	**134**	**ASP**	**neg**	**0.54**	**11.5**	**periphery**	**potential secondary water channel**	**SWC**
H5	**154**	**VAL**	hphob	0.56	12.0	periphery	entrance substrate channel	-
H6	189	ILE	hphob	0.96	9.4	surface	entrance substrate channel	-
H7	**223**	**ASP**	**neg**	**0.65**	**12.7**	**surface**	**potential interfacial water entrance**	**IWE1**
	277	**LEU**	hphob	0.77	8.7	periphery	probable barrier, adj. to Asp223 and Ser47	
H8	**278**	**LEU**	hphob	0.88	6.9	periphery	entrance substrate channel	
	285	**ILE**	hphob	0.94	11.3	periphery	entrance substrate channel	
H9	**290**	**LYS**	**pos**	**0.52**	**17.7**	**surface**	**potential interfacial water entrance**	**IWE2**

Four plausible avenues for water entry to CALB were thereby identified, a previously characterized (Larsen et al. 2000) water channel (PWC), a potential secondary water channel (SWC) and two potential interfacial water entrances (IWE1 and IWE2) that were marked by polar residues in direct contact with the triglyceride-water interface. IWE1 and IWE2 were hereby regarded as positions that may interact with the highly self-associated polar triglyceride moieties that were observed in the triglyceride layer (Chapter 2.3) and may thus serve as potential mediators for the influx of water through the substrate channel (Figure 19 and 20).

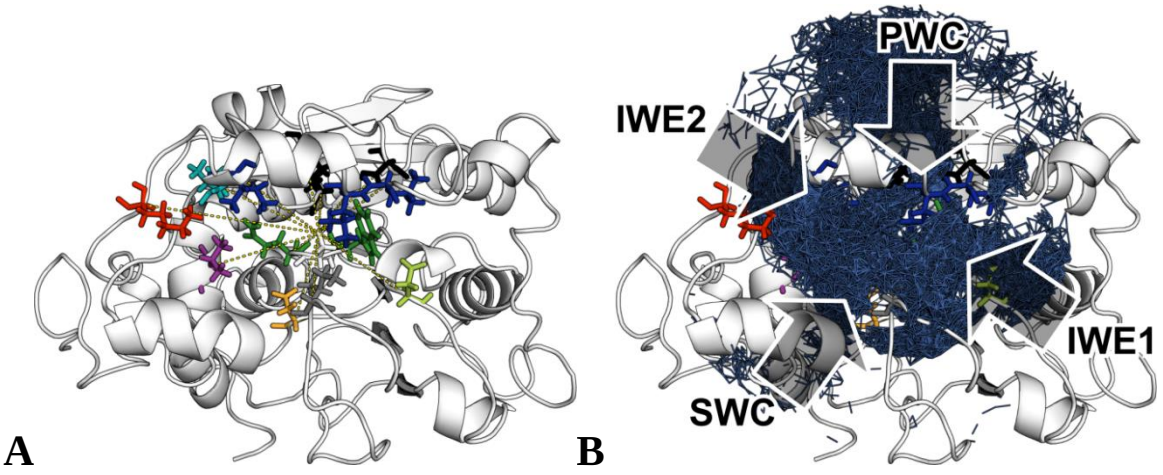


Figure 19: A) Hotspot positions in the CALB protein structure with colors corresponding to H1-H9 of Figure 18, dotted lines point towards the active site Ser105, protein is represented in the bottom view, i.e. from the perspective of the triglyceride-water interface that CALB is attached to (see Figure 17) / B) Water influx pathways from five full SFM runs are superimposed onto the same depiction as Figure 4A to illustrate potential water entryways (arrows). PWC is a known water channel (Larsen et al. 2000), SWC a potential secondary channel, whereas IWE1 and IWE2 are hydrophilic amino acids in direct contact with the triglyceride-water interface and may contribute to water influx via the substrate channel.

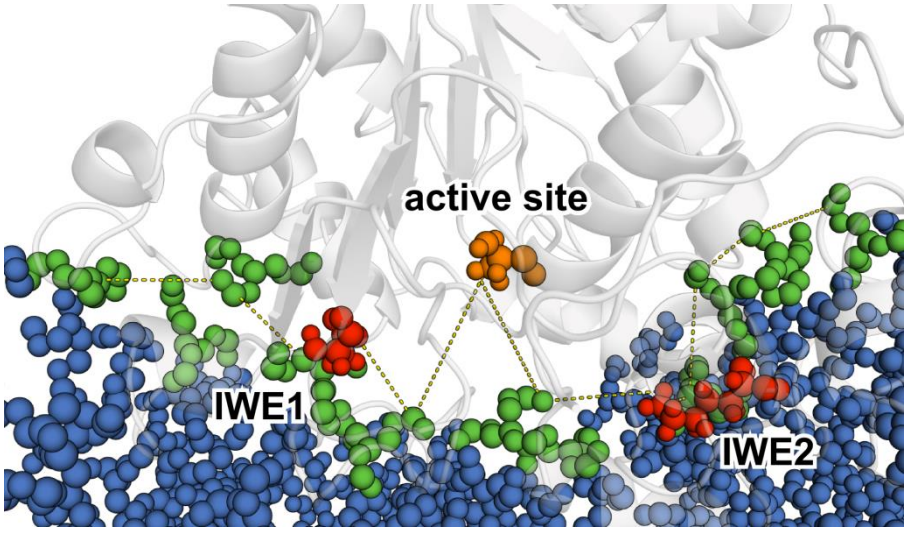


Figure 20: Hypothetical pathways (dotted line) of water molecules to the active site (orange spheres) along the potential interfacial water entrances IWE1/2 (red spheres) in between the protein (light grey) and the triglyceride-water interface, where lamellar-like association structures of polar triglyceride moieties (blue spheres) are in close contact (green spheres) to the protein and may thus facilitate hydrogen bonds for water diffusion through the substrate channel.

3 Discussion

In the framework of this dissertation, molecular mobility and aggregation were investigated by molecular dynamics simulations (MD) in condensed-phase biomolecular systems, which ranged from methanol-water mixtures, lipidic droplets and large-scale interfaces interacting with proteins. The studies covered in this context can be loosely linked by concepts relating to hydrophobicity, the discussion of which is the intended focus of this section. More detailed discussions related to the individual studies are provided in the relevant sections of Chapter 4.

Hydrophobicity does not, as the terminology suggests, imply a water "phobia" of nonpolar solutes in aqueous environments. The contrary is the case, attractive van der Waals forces persist between nonpolar atoms and water molecules, which stem from permanent and induced dipoles within atoms and originate from the spatial distribution of negatively charged electrons surrounding positively charged protons in atom nuclei. However, the driving force for structural assembly of nonpolar moieties in condensed-phase aqueous environments, the hydrophobic interaction, is ascribed to the strong attractive Coulomb force between water molecules, which supersedes the relatively weak attractive van der Waals forces (Galli 2007) between water and nonpolar solutes. The bipolar water molecule spans continuous networks of hydrogen bonds, each water molecule thereby associating with up to four binding partners at a time. All nonpolar solute molecules (or moieties) that are introduced into this system will subsequently associate and arrange in a manner that perturbs this hydrogen bonding network as little as possible. In fact, the solvation free energy of simple nonpolar spheres in aqueous environments is commonly regarded to be mainly composed of energy contributions that stem from either water hydrogen bond reorientation (entropic) or water hydrogen bond breaking (entphalpic) (Chandler 2005). To attain, from this biologist's perspective, an intuitively accessible parameter to evaluate the implications of hydrophobicity on structural association in biomolecular systems, the molecular association bias was conceived in the study "Incomplete Mixing versus Clathrate-Like Structures: A Molecular View on Hydrophobicity in Methanol-Water Mixtures" (Chapter 4.1). The molecular association bias is based on a simple statistical relation between observable nearest atom neighbors and the corresponding expected value for a random molecular distribution. It was applied to describe the segregation of

molecules in methanol-water mixtures in terms of a descriptive hierarchy of molecular association preferences. Water self-preference was thereby observed, but also the preference of methanol towards water, the disfavor of water towards methanol and the disfavor of methanol towards itself as a molecular partner for association (Figure 6B). Local segregation of methanol and water, also known as incomplete mixing (Dixit et al. 2002; Guo et al. 2003), was previously shown to be a sufficient criterion to quantify the methanol-water mixture entropy deviations from ideal mixture behavior (Soper et al. 2006), which contradicts the Frank and Evans (Frank et al. 1945) hypothesis of structured water surrounding solutes as the source of excess entropy. Analysis of the dynamic properties of hydrogen bond lifetimes (Figure 5A), rotational autocorrelation (Figure 5B) and molecular self-diffusion (Figure 6A) of water molecules within the first solvation shell surrounding the nonpolar methyl moieties of methanol provided no support for structured water in the methanol-water mixture. Although Dzugutov (Dzugutov 1996) states that excess entropy, as a measure of inaccessible configurations of a system, can in principle be directly linked to excess self-diffusion, a pictorial description that may serve to explain the changing molecular mobility in methanol-water mixtures was proposed. Clustering analysis revealed that the hydrogen bonding network breaks down or forms at an approximately equimolar ratio (Figure 7), which coincides with the minimum in the self-diffusion coefficients for both water and methanol (Figure 6A). Chandler states (Chandler 2002) that when the size of a solute cavity in water extends beyond a threshold of approximately 1nm, the nature of hydrophobicity changes as the number of potential hydrogen bonds between water molecules is depleted. This concept is usually related to large solutes that disrupt hydrogen bonding instead of small (<1nm) solutes, such as a single methyl group of methanol, where it is proposed that existing hydrogen bonds can instead be maintained merely by bond reorientation. However, with increasing methanol content in methanol-water mixtures, a disruption of the system-wide hydrogen bonding network could induce similar effects as ascribed to large solutes, which have also been likened to cavities (Chandler 2005) in the aqueous phase. The disintegration of cooperative water hydrogen bonding was thus considered to force water molecules into clusters within methanol, which was also consistent with the observation that hydrogen bond lifetimes between water molecules increased exponentially after the breakdown of the hydrogen bonding network (Figure 5A). When further decreasing water content in methanol-rich mixtures, the renewed increase in water diffusion was thus rationalized by the diffusion of

water clusters, which, according to the Einstein-Stokes relation, increases as cluster size decreases. When increasing water content in water-rich mixtures, after a system-wide hydrogen bonding network is formed, the increase in water mobility was accredited to the fact that more potential hydrogen bonding partners are made available. Methanol was considered to follow water behavior similar to a small cluster of solute molecules in water-rich mixtures until the breakdown of the hydrogen bonding network, after which methanol gradually attains its bulk mobility. This pictorial model of molecular structure and mobility in methanol-water mixtures is consistent with the concept of incomplete mixing. It is also consistent with the implications of the relation between diffusion and entropy by Dzugutov (Dzugutov 1996), namely that molecular mobility can be ascribed to the configurations that are accessible to molecules in simple liquid systems.

The molecular association bias as a quantity to describe the molecular structure in condensed-phase aqueous mixtures was extended to lipid droplets, specifically to self-emulsifying drug delivery systems (SEDDS), which are lipidic formulations that are routinely applied in the oral delivery of poorly water-soluble drug molecules (Singh et al. 2009). The study "Molecular Dynamics Simulations of Self-Emulsifying Drug Delivery Systems (SEDDS): Influence of Excipients on Droplet Nanostructure and Drug Localization." (Chapter 4.2) focused on the investigation of the SEDDS nanostructure to gather insight as to drug molecule solubilization, which may prove beneficial to formulation design. The hydrophobic effects in regard to structure formation of amphiphilic assemblies, such as SEDDS droplets, are commonly described in similar terms as the hydrophobic effects regarding purely nonpolar solutes, with the inclusion of the interaction between hydrophilic solvent moieties and water. Thereby, the hydrophilic "headgroup" is considered to always be oriented towards the aqueous environment (Chandler 2005). This can explain the molecular structure observed for a range of microscopic assemblies, such as micelles or lipid bilayers (Larson 1999). The implications of these considerations are that the nanostructure within amphiphilic lipid droplets is strictly determined by water self-preference, while the interactions between the solute molecules themselves play a subordinate role in nanostructure formation. The results of the SEDDS study illustrate that the nanostructure within a hydrophobic lipid droplet can significantly change upon varying the composition of SEDDS excipients. Molecular association biases suggested that the self-association preference of polar amphiphilic

triglycerides moieties in particular has a significant impact on nanostructure formation within droplets (size <10nm in simulation). From this perspective, it is tempting to infer a hierarchy of self-preferences to explain SEDDS nanostructure. Hereby, water self-preference induces the formation of microscopic triglyceride aggregates in form of droplets, while the self-preference of polar moieties could induce the observed unstructured, lamellar-like or vesicle-like nanostructures within triglyceride aggregates and finally, hydrophobic or amphiphilic drug molecules, like CyA, embed themselves into the droplet nanostructure according to their association preferences. In this regard, drug molecules in SEDDS droplets can be perceived as solvated within the solvent environment of the lipid droplet, which in turn is solvated in water. Discussions on hydrophobic phenomena commonly focus on the characterization of the interface between hydrophobic substances and the aqueous environment, likely because nanostructures within the size-range of lipidic droplets are very difficult to characterize experimentally (Roke et al. 2012). While interfacial characterization may be sufficient to attain a full structural description of a micelle or a lipid bilayer, which do not exceed the length scale of two solute molecules, triglyceride molecules are known to form emulsions of large-scale aggregates, ranging from several hundred to several thousand nm (Chung et al. 2001; Jurado et al. 2006). In these instances, water-induced structural formation at the surface of an aggregate can be expected to gradually transition into the structure associated with the condensed state of bulk triglyceride. A major focus of SEDDS formulation design, however, is to reduce the size of emulsion droplets diameters as far as possible by introducing surfactants to the formulations. Thereby, current self-nanoemulsifying drug delivery systems (SNEDDS) manage to reduce droplet sizes to approximately 10nm (Singh et al. 2009; Zidan et al. 2007), where the transition of surface-induced to bulk-like structure may not yet have been completed. In this regard, nanostructure formation in droplets could be influenced by the interaction between lipids and in the case of triglyceride molecules the self-association of its polar moieties. In the SEDDS study presented in this dissertation, droplets were modeled with diameters of 10nm, and results obtained therein suggest that within these size scales, interior droplet nanostructures vary greatly due to factors that cannot exclusively be linked to the interfacial interactions between triglyceride and water. Elongation of the fatty acid chain length of triglyceride, for example, induced significantly different polar nanoenvironments within the droplet interior, which may be observed in the heightened molecular association bias between

polar triglyceride moieties. Polar headgroups of triglyceride molecules were thereby shown to not necessarily orient towards the aqueous environment. Subsequently, lamellar-like patterns formed within the SEDDS droplets due to the presence of triglyceride, and hydrophobic moieties of triglyceride were exposed at the aqueous interface to a significant degree. Unlike micelles that are restricted in size due to the orientation of polar headgroups towards the aqueous environment and the subsequent packing of hydrophobic fatty acid side chains in micelle interiors, large-scale aggregates triglycerides do not seem to follow that pattern. Hence, the organization of the interior nanostructure must, at least in part, be facilitated by triglyceride interactions. The nanostructure in SEDDS droplets was thus perceived to be influenced by both the hydrophobic effects induced at the aqueous interface as well as the interactions between polar triglyceride headgroups. Conversely, di- and monoglyceride molecules did display the familiar reorientation of polar headgroups towards the aqueous surrounding as well as the protrusion of nonpolar tailgroups into the interior, which significantly altered the preexisting (in respect to the standard system) droplet nanostructures dominated by triglyceride. This marked difference in hydrophobic behavior may explain why mono- and diglycerides may be applied as surfactants, while triglycerides cannot. Moreover, considering poly(ethylene glycol) (PEG) in this context illustrates how other amphiphilic molecules do not adhere to the binary hydrophobic behavior of polar moieties oriented towards the aqueous phase and nonpolar moieties oriented towards the oil phase, simply due to the chemical structure. Strictly speaking, PEG is an amphiphile, but is composed of a succession of alternating polar and nonpolar moieties (Figure 2). In SEDDS droplet simulations, it was found that PEG is localized at the droplet surface, thereby forming encasing shells around an interior lipid droplet, with significant exposure of nonpolar moieties towards the aqueous environment. Light scattering studies have revealed that PEG displays extended rather than compact conformations and does not necessarily aggregate in in aqueous environments (Devanand et al. 1990). Evaluating the nanoscale structuring of molecules simply on the basis of generalized hydrophobic concepts may thus not be sufficient to fully capture the extent of observable nanostructure phenomena involving triglyceride aggregates. Analyzing molecular association biases may help paint a clearer picture.

With the explicit modeling of water, the SEDDS system size, in terms of number of atoms, were computationally already very costly. A simple upscaling of the droplet systems

(10nm) to model hydrophobic effects on large-scale aggregates (1000nm+) was therefore not feasible. Thus, the planar model by Gruber et al. (Gruber et al. 2012) was applied that mimics the negligible interfacial surface curvature of large-scale triglyceride aggregates. Hereby, it was found that after extensive equilibration, similar lamellae as those previously observed in the structure of SEDDS systems aggregated in orthogonal orientation in respect to the planar triglyceride-water interface in a highly ordered, almost crystalline manner (Figure 13). Furthermore, water was shown to penetrate the initially dry triglyceride system in significant amounts. This observation was extensively characterized in terms of triglyceride layer structure, molecular association biases and water mobility, as part of an ongoing study by the author to quantify protein interactions on triglyceride-water interfaces. Hereby, a major focus will be to correlate the obtained results to experimentally observable properties. Direct experimental data on the nanostructure of triglyceride interfaces in aqueous environments is unfortunately lacking, since few experimental methods are able to resolve droplets or micellar systems in the condensed-state and thus few studies have been conducted in this regard (Roke et al. 2012). However, data on solid state crystalline triglyceride aggregates is available, wherein it has been suggested that triglycerides stack in lamellae substructures (Marangoni et al. 2012; Acevedoet al. 2010), which supports the observations on the aggregation structure obtained from simulation. Water uptake into the triglyceride nanostructure is also supported by experimental data, for example by HPLC analysis studies, which have correlated water content in tricaprylin systems with monoglyceride content (Rane et al. 2008) as well as ester concentration for various lipids (Cao et al. 2004). Particularly the latter study corroborates the influx of water molecules into the triglyceride-water interface at the self-associated polar triglyceride moieties, which was found in the simulation studies by the author. Surprisingly, few computational studies were found in literature that characterize triglyceride systems, only one of which applied MD simulations in the condensed-phase with modern force fields (Sum et al. 2003). Therein, it was illustrated that the OPLS force field, which was used in the study in conjunction with the Berger lipid model (Chapter 1.3), performs well in predicting triglyceride viscosities; it also outlined the strong tendency of polar triglyceride moieties to self-associate.

In the study “The Solvent Flux Method (SFM): A Case Study of Water Access to Candida Antarctica Lipase B.”, a method was devised with the intention of creating a general-

purpose MD approach to quantify the solvent flux in enzymes, in particular for hydrolase related engineering. Thereby, it was possible to identify potential avenues of water entry to the enzyme *Candida antarctica* lipase B (CALB), which has evolved to include water as a co-substrate in its catalytic function of cleaving triglycerides. As was shown in an experimental study of Larsen et al. on CALB (Larsen et al. 2000), specific mutations to hinder water influx can be productively utilized for the engineering of hydrolase variants that facilitate the reverse synthetic reaction instead of hydrolysis. Thus designed hydrolases are highly beneficial for synthetic chemical processes, since water is excluded as a competing nucleophile to the acyl-enzyme intermediate. The molecular model of the triglyceride-water interface, which served as a benchmark for SFM, suggests that water influx does not only occur by means of a single water channel, but also that the self-associated polar nanoenvironments of triglyceride may play an important role in facilitating a sufficient water resupply to the catalytic machinery at the active site of CALB. The promising results obtained thus far motivate a continued development of such methods in combination with refining the accuracy of models such as the triglyceride-water interfaces to attain predictive capabilities. In the context of biomolecular systems, enzyme engineering offers the added benefit of a binary validation process in form of functional enzyme variants to support insights gathered from molecular models. This may reduce the necessity to conduct time-consuming and cost-intensive experimental studies or the computational reproduction of macroscopic observable properties.

4 Publications

4.1 Incomplete Mixing Versus Clathrate-Like Structures: A Molecular View on Hydrophobicity in Methanol-Water Mixtures

4.1.1 Abstract

The underlying molecular mechanisms of macroscopic excess properties were studied by molecular dynamics simulations for different compositions of methanol-water mixtures. Structural data (nearest neighbor relationships, clustering analysis) and dynamic data (hydrogen bond lifetimes, rotational autocorrelation, translational diffusion) were evaluated. Nearest neighbor relationships provide quantitative evidence and a pictorial description of incomplete mixing at the molecular level as a source for mixture anomalies, while a comparative study of water surrounding methyl moieties versus water in the bulk-like environment provides evidence against the hydrophobicity model of clathrate-like hydration. Furthermore, the formation or breakdown of the system-wide hydrogen bonding network at a critical threshold of approximately equimolar mixture is perceived to separate the mixture system into two hydrogen bonding regimes: hydrogen-bonded water clusters embedded in methanol for mixtures with low water content and methanol molecules within a system-wide hydrogen-bonded water network for mixtures with high water content.

4.1.2 Introduction

Deviations from ideal mixture behavior (excess effects) in aqueous solutions are a widespread phenomenon that is quantified mostly by macroscopically observable properties such as entropy[1]. However, measuring macroscopic observables does not explain the molecular origin of the anomalies in mixtures, which results from the nature of the intermolecular interactions of the different molecule types. Solution chemical reaction rates, ion transport in solution or membranes, protein folding and enzymatic activity are but a few instances where an in-depth understanding of molecular mechanisms is of paramount importance. Excess effects in aqueous solutions are generally ascribed to

hydrophobicity. While hydrophobicity is a useful chemical concept, even after decades of detailed investigation the underlying molecular mechanisms are still controversial[2,3]. As prevalent and intuitive as “hydrophobicity” may seem, it is important to note that hydrophobicity is not a first-principle parameter, but an abstract concept. The terminology of “hydrophobicity” in itself is misleading, because it implies a lack of attraction between polar water molecules and nonpolar groups, when in fact attractive interactions persist due to induced dipoles[4]. The cause of hydrophobicity is widely attributed to the comparatively strong electrostatic attraction between water molecules via hydrogen bonding[2,3]. In this connotation, "hydrophobicity" is best described as the “hydrophilicity” of water in preference to molecules with weaker dipoles.

One of the most seminal works on molecular association behavior in aqueous solutions is that of Frank and Evans[5] published in 1945, which introduced the model of “ice-like” clathrate structures of immobilized water molecules surrounding nonpolar moieties. Kauzmann[6] later correlated the hydrophobic interaction with protein folding and stability and introduced it as a driving force for structural assembly and phase separation in solution. While these ideas undoubtedly had profound influence on the perception of how water interacts with nonpolar solutes on a molecular level, the scientific evidence is inconclusive and thus the molecular mechanisms of hydrophobicity are still a topic of high controversy. Evidence for immobilized water molecules has been found, e.g. in early neutron diffraction experiments[7-9] and Monte Carlo simulations[10]. However, more recent molecular dynamics simulations[11], ab initio calculations[12,13] and diffraction experiments[14,15] have contested these findings.

Incomplete mixing has been proposed as an alternative concept to explain the molecular source of excess properties in aqueous mixtures[14,16,1]. In contrast to the localized immobilization of water surrounding nonpolar solutes of the clathrate model, incomplete mixing explains excess entropy by a system-wide molecular segregation.

Hydrophobicity is widely perceived to be a multifaceted problem that manifests itself differently dependent on the interface between the hydrophobic compound and the aqueous environment, such as fully miscible liquids, small micelles, proteins in solution, or fully separated phases[2,3]. In aqueous solutions, hydrogen bonds are perturbed when small nonpolar solutes are introduced to the mixture, which leads to a structural rearrangement of

water molecules (entropic effect). In contrast, large hydrophobic interfaces are perceived to reduce the number of possible hydrogen bonds in the system (enthalpic effect). This differentiation however seems arbitrary, considering that structural rearrangement of water molecules would also take place in the presence of large hydrophobic interfaces, while likewise the number of possible hydrogen bonds in solution would eventually decrease as the concentration of small nonpolar solutes increases.

To study hydrophobicity and the molecular mechanisms of excess effects with molecular dynamics simulations, we chose water-methanol mixtures as simple model systems.

4.1.3 Methods

4.1.3.1 Simulation Details

MD simulations were performed at 298.15 K and 1 bar under periodic boundary conditions in an NPT ensemble. The Berendsen coupling scheme[17] was used, with coupling constants of 0.4 ps for temperature and 1.2 ps for pressure coupling. The leap-frog algorithm[18] was used for all simulations with a time step of 2 fs. C-H and O-H bond lengths were constrained with the LINCS algorithm[19]. Long-range electrostatics was treated by the particle-mesh Ewald algorithm (PME)[20,21]. Lennard-Jones interactions were capped at 1.4 nm. The transferable all atom OPLS (optimized potential for liquid simulation) force field[22] was used for methanol; water was parameterized by the TIP4P model[10]. During an equilibration phase of 20 ns, energy, density, and radial distribution functions were monitored, followed by a production phase of 10 ns for analysis. The GROMACS 4.0.7 software was used for simulations and analysis[23,24].

4.1.3.2 Molecular Self-Diffusion

The molecular self-diffusion coefficient $\boldsymbol{D}_A$ is a transport property that describes the translational mobility of particles A and can be regarded as a meaningful quantity to link molecular mobility to the macroscopic deviation from ideal mixture behavior witnessed in

excess properties[25]. It is calculated from the mean-square displacement and the Einstein relation by averaging over all particles[26].

$$\lim_{t\to\infty}\left\langle(\mathbf{r}_i(t)-\mathbf{r}_i(0))^2\right\rangle_A = 6D_A t \qquad \text{(eq. 1)}$$

4.1.3.3 Nearest Neighbor Relationships

To characterize changes in local structure for the different compositions of a binary methanol-water mixture, the deviation of the nearest neighbor relationships in the simulated systems from their expected values in an ideal mixture was calculated. Nearest neighbor relationships were defined by a radial cutoff criterion of 0.35 nm (Online Resource Fig.S4). $\boldsymbol{N}_{ix}^{\text{total}}$ denotes the total number x of molecules (methanol and water) that are neighbors to molecule centers of type i (either methanol or water). $\boldsymbol{N}_{ii}^{\text{ideal}}$ denotes the number of water-water or methanol-methanol neighbor relationships as expected in an ideal mixture. In a fully randomized molecule distribution, $\boldsymbol{N}_{ii}^{\text{ideal}}$ may be calculated from $\boldsymbol{N}_{ix}^{\text{total}}$ and the molar fraction $\boldsymbol{\chi}_i$ by

$$N_{ii}^{ideal} = p_{ii}^{ideal} \cdot N_{ix}^{total} = \chi_i \cdot N_{ix}^{total} \qquad \text{(eq.2)}$$

where $\boldsymbol{p}_{ii}^{ideal}$ denotes the probability of finding a neighbor of molecule type i to a molecule center of type i for an ideally randomized molecule distribution, assuming that the difference in molecule size of both molecule types in the system is negligible (ideal mixture criterion). In an ideal mixture, $\boldsymbol{p}_{ii}^{ideal}$ should be equal to the molar fraction $\boldsymbol{\chi}_i$ of molecule type i. In an ideal binary mixture, the number of nearest neighbors between two molecules of different types $\boldsymbol{N}_{ij}^{ideal}$ (the number of neighbors of molecule type j to centers of molecule type i) is calculated by

$$N_{ij}^{ideal} = N_{ix}^{total} - N_{ii}^{ideal} = (1-\chi_i)\cdot N_{ix}^{total} \quad (i \neq j) \qquad \text{(eq. 3)}$$

The deviation between simulated (real) and ideal mixtures can thus be quantified by the ratios $\boldsymbol{M}_{ii}$ and $\boldsymbol{M}_{ij}$:

$$M_{ii} = \frac{N_{ii}^{real}}{N_{ii}^{ideal}} = \frac{N_{ii}^{real}}{\chi_i \cdot N_{ix}^{total}} \quad \text{and} \quad M_{ij} = \frac{N_{ij}^{real}}{N_{ij}^{ideal}} = \frac{N_{ij}^{real}}{(1-\chi_i)\cdot N_{ix}^{total}} \qquad \text{(eq. 4)}$$

with $\boldsymbol{N}_{ii}^{real}$, $\boldsymbol{N}_{ij}^{real}$, and $\boldsymbol{N}_{ix}^{total}$ being derived from the simulation. $\boldsymbol{M}_{ii}$ is therefore a quantity to describe the association bias of molecules *i* within a binary mixture and the resulting structural arrangement of the two molecule types. Likewise, $\boldsymbol{M}_{ij}$ characterizes the interaction between the different molecule types. In $\boldsymbol{M}_{ij}$, the first index *i* signifies the central molecule type and *j* the neighboring molecule type ($\boldsymbol{M}_{ij} \neq \boldsymbol{M}_{ji}$). For ideal binary mixtures we expect $\boldsymbol{M}_{ii}=1$ and $\boldsymbol{M}_{ij}=1$, while for incomplete mixing we expect $\boldsymbol{M}_{ii}>1$ if molecules of type *i* favor the proximity to molecules of the same type and $\boldsymbol{M}_{ij}>1$, if molecules of type *i* favor the proximity to molecules of the other type *j*.

4.1.3.4 Hydrogen Bond Lifetime

Upon separating water molecules into a clathrate-like water and a bulk water group, it is necessary to take heed of water molecules quickly leaving the first water shell around the methyl moieties due to fast translational diffusion. To circumvent this difficulty, the evaluation of correlation times provides a simple and descriptive property to compare the duration of the existence of a signal. Therefore, the hydrogen bond (HB) correlation function $\boldsymbol{c}_h(t)$[27] was analyzed,

$$c_h(t) = \frac{\langle h(0)h(t)\rangle}{\langle h\rangle} \qquad \text{(eq. 5)}$$

where $\boldsymbol{h}$ denotes the existence function and $\langle h \rangle$ the total population of HBs. The existence function equals unity if a HB exists and is zero otherwise, under the condition that $\boldsymbol{h}(0)=1$. This HB lifetime description has been called the "intermittent hydrogen bond correlation function"[28], alluding to the fact that by construction it is time-independent of bond-breaking events. From the correlation of the existence function it is possible to define the reactive flux correlation function $\boldsymbol{K}$(t)[27].

$$K(t) = \frac{dc_h(t)}{dt} = kc_h(t) - k'n(t) \qquad \text{(eq. 6)}$$

The term $k' \cdot \boldsymbol{n}$(t) describes bond-breaking and the term $k \cdot \boldsymbol{c}_h$(t) bond-forming. From here on it is possible to define the HB lifetime $\boldsymbol{\tau}_h$ as the inverse of the forward rate constant k.

$$\tau_h = \frac{1}{k} \qquad \text{(eq. 7)}$$

4.1.3.5 Rotational Autocorrelation

The rotational autocorrelation time can be calculated by integration or fitting of the rotational autocorrelation function C_h(t).

$$C_h(t) = \left\langle P_n\left[u(0) \cdot u(t)\right]\right\rangle \qquad \text{(eq. 8)}$$

P_n is the rank n Legendre polynomial and $\boldsymbol{u}$(t) a vector associated with the geometry of the molecule to be analysed. Specifically the normal vector to the water molecule plain $\boldsymbol{u}(t)=\boldsymbol{r}_{OH1} \times \boldsymbol{r}_{OH2}$ was used[29]. Applying the rank 2 Legendre polynomial makes calculated data comparable to experimental NMR data[30,31]. Instead of integration, the trapezoidal rule was used to acquire the Riemann sum between intervals of discrete time steps of 2 fs.

4.1.3.6 Differentiation of Clathrate-Like and Bulk Water Molecules

To determine whether there is a difference in the dynamics of water molecules surrounding nonpolar methyl moieties (clathrate-like water) and free water molecules (bulk water) within methanol-water mixtures, the two groups were analyzed separately. Clathrate-like water molecules were assigned by applying a distance cutoff criterion between the carbon atom of methanol and the oxygen atom of water, where 0.45 nm corresponds to the first minimum of the radial distribution function g(r)[Cmet-Owat] (Online Resource Fig.S4). The full simulation trajectories were segmented into equally long parts of 500 ps and were analyzed after redefining clathrate-like and bulk water for every segment. Since the long tail of autocorrelation functions significantly influences results, a combined scheme of explicit integration until 5 ps and fitting after 5 ps was applied, which has previously been benchmarked for similar systems[29]. For mixtures below χ_{wat}<0.6, water molecules could no longer be separated into clathrate-like water and bulk water groups when applying the cutoff criterion of 0.45 nm, because all water molecules were assigned to the clathrate-like water group from thereon. Therefore, the cutoff criterion was reduced to 0.35nm, which corresponds to the first peak maximum of g(r)[Cmet-Owat]. Reducing the cutoff criterion to 0.35nm resulted in a shift of the exponential increase of HB lifetimes τ_{ww} from χ_{wat}~0.5 to χ_{wat}~0.2 (Online Resource Fig.S6), which suggests that the water molecules in the outer half (0.35nm-0.45nm) of the first coordination sphere statistically mask the effect of exponentially rising HB lifetimes from 0.5>χ_{wat}>0.2. The large standard deviations at low water content are a consequence of the small number of bulk water molecules.

4.1.4 Results

Multiple molecular dynamics simulations were performed on condensed phase systems containing pure water, pure methanol and nine intermittent methanol-water mixtures (χwat=0.0, 0.1, 0.2, … , 1.0). The molecular systems were equilibrated thoroughly (20ns), while monitoring potential energy, density, and radial distribution functions. A system size of 1000 molecules with periodic boundary conditions proved sufficient for analysis, since simulations of larger systems led to comparable results (Online Resource Fig.S3). To explore the molecular basis for deviations from ideal mixture behavior, all simulations were analyzed for structural properties such as density (Online Resource Fig.S1), radial

distribution functions (Online Resource Fig.S2), nearest neighbor relationships, clustering analysis and dynamic properties such as molecular self-diffusion, rotational autocorrelation and hydrogen bond (HB) lifetimes. Molecules surrounding the nonpolar methyl moiety (clathrate-like water) and "free" water molecules (bulk water) in the methanol-water mixtures were separated into two groups and analyzed individually when appropriate, in order to implicitly assess increased order of clathrate-like water as a possible source for the lower than expected entropy in the mixture.

4.1.4.1 Diffusion Coefficient

Molecular self-diffusion coefficients for methanol and water molecules were analyzed to capture excess behavior on a molecular level and to compare simulation results to experimental data (Fig.1).

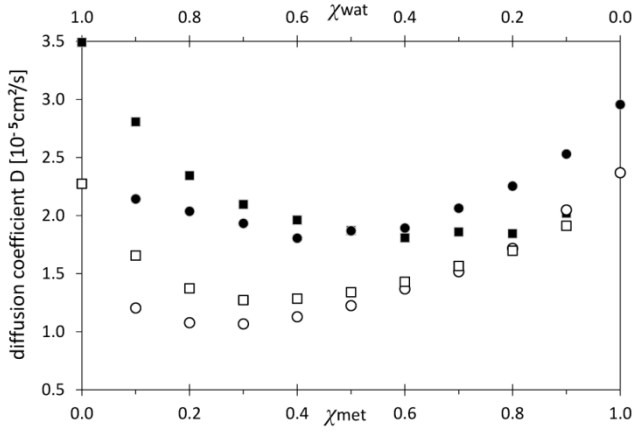


Fig.1: Molecular self-diffusion coefficients D in methanol-water mixtures for simulated water (TIP4P: ■), experimental water[32] (□), simulated methanol (OPLS/AA: ●) and experimental methanol[32] (○). Simulation data was obtained from mean square displacement and the Einstein relation. Standard deviation and detailed statistics of the calculations are provided in the Online Resource (Fig.S5).

For low molar fractions of water, the diffusion coefficient of water molecules was found to be high ($\boldsymbol{D}_{\chi wat=0.1}$ = 2.02 $\cdot 10^{-5}$ cm²/s) and decreased with increasing water content to a minimum at $\boldsymbol{\chi}_{wat} \approx 0.5$ ($\boldsymbol{D}_{\chi wat=0.5}$ = 1.81 $\cdot 10^{-5}$ cm²/s). For $\boldsymbol{\chi}_{wat} > 0.5$, the coefficient increased

again until the maximum value ($\boldsymbol{D}_{\chi wat=1.0}$ = 3.59 ·10^{-5} cm²/s) for pure water was reached (experimental: $\boldsymbol{D}_{\chi wat=1.0}$ = 2.16 ·10^{-5} cm²/s[32]). Similarly, the diffusion coefficient of methanol was found to be high for low molar fractions of methanol ($\boldsymbol{D}_{\chi met=0.1}$=2.14 ·$10^{-5}$ cm²/s). It decreased with increasing methanol content until a minimum at $\boldsymbol{\chi}_{met}\approx$0.4 ($\boldsymbol{D}_{\chi met=0.4}$=1.81 ·$10^{-5}$ cm²/s) and increased again until the maximum of $\boldsymbol{D}_{\chi met=1.0}$=2.96 ·$10^{-5}$ cm²/s for pure methanol (experimental: $\boldsymbol{D}_{\chi met=1.0}$=2.5 ·$10^{-5}$ cm²/s[32]). Thus, the minima of the diffusion coefficients for water and for methanol were both found in the vicinity of $\boldsymbol{\chi}_{met}$=0.4-0.6 ($\boldsymbol{\chi}_{wat}$=0.4-0.6) of the mixture, while the highest diffusion coefficients were found for the pure substances.

Diffusion coefficients are also a meaningful quantity to probe for variations in translational mobility between clathrate-like water and bulk water, which is a possible source for the lower than expected entropy observed for methanol-water mixtures. However, water molecules in both groups displayed no discernible difference in diffusion coefficients and were in close agreement with the diffusion coefficient presented for all molecules (Fig.1). This suggests that there is no significant difference in translational mobility between clathrate-like water and bulk water. Compared to experimental values[33], the simulated diffusion coefficients were consistently overestimated, but the relative trend was adequately represented. These findings are in agreement with previously published data[34,29].

4.1.4.2 Rotational Autocorrelation

To compare the dynamics of clathrate-like water and bulk water, the rotational autocorrelation time $\boldsymbol{\tau}_{w\perp}$ for a vector orthogonal to the water molecule plane was separately analyzed for clathrate-like water and bulk water. Water molecules in the clathrate-like water and bulk water groups displayed similar rotational correlation times $\boldsymbol{\tau}_{w\perp}$ for all mixture compositions, ranging from $\boldsymbol{\tau}_{w\perp}(\chi met=0.1)$=1.83 ps to $\boldsymbol{\tau}_{w\perp}(\chi met=0.9)$=3.94 ps (Fig.2), which is in agreement to experimental values of $\boldsymbol{\tau}_{w\perp}(\chi met=0.0)$~2.0 ps [30,31] and simulation data of TIP4P water model literature of $\boldsymbol{\tau}_{w\perp}(\chi met=0.0)$= 0.8-1.7 ps [29]. A linear increase for 0.1<$\boldsymbol{\chi}_{wat}$<0.5 and a convergence to maximal values for 0.5<$\boldsymbol{\chi}_{wat}$<0.9 was observed.

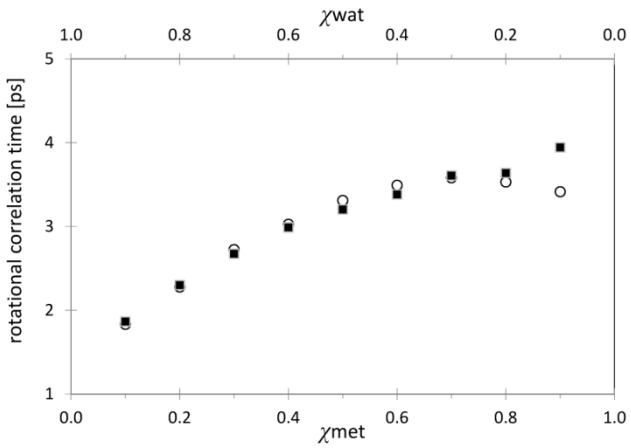


Fig.2: Rotational correlation time τw⊥ of the normal vector to the water molecule plain in methanol-water mixtures calculated via the integral of the rotational autocorrelation function (rank 2 Legendre polynomial) of the clathrate-like water group (○) and the bulk water group (■), after applying a spherical differentiation cutoff criterion for clathrate-like water of 0.35nm around the methyl group. Standard deviations averaged over multiple trajectory intervals range from 0.01-0.1ps.

4.1.4.3 Hydrogen Bond Lifetimes

The hydrogen bond (HB) lifetime $\boldsymbol{\tau}$ can be regarded as a characteristic quantity for the dynamics of hydrogen-bonded clusters (Fig.3a). For pure water, HBs between water molecules had a short lifetime ($\boldsymbol{\tau}_{ww}(\chi\text{wat}=1.0)$=1.47 ps)**.** Upon decreasing water content in the methanol-water mixture, the lifetime of water-water HBs increased linearly up to an equimolar ratio ($\boldsymbol{\tau}_{ww}(\chi\text{wat}=0.5)$ = 4.4 ps). When decreasing the water content further ($\boldsymbol{\chi}_{\text{wat}}$<0.5), $\boldsymbol{\tau}_{ww}$ increased almost exponentially until a maximum at $\boldsymbol{\chi}_{\text{wat}}$=0.1 was reached ($\boldsymbol{\tau}_{ww}(\chi\text{wat}=0.1)$=10.39 ps). In contrast, the lifetime of HBs between methanol molecules $\boldsymbol{\tau}_{mm}$ changed only slightly for the different mixtures. HB lifetimes were short for low methanol content ($\boldsymbol{\tau}_{mm}(\chi\text{met}=0.1)$=1.17 ps) and slightly increased to $\boldsymbol{\tau}_{mm}(\chi\text{met}=0.6)$=2.6 ps just above equimolar composition. For higher molar fractions of methanol, the lifetime remained at almost constant levels between $\boldsymbol{\tau}_{mm}$=2.6-3.06 ps. The lifetimes of HBs between methanol and water $\boldsymbol{\tau}_{mw}$ were found at levels between $\boldsymbol{\tau}_{ww}$ and $\boldsymbol{\tau}_{mm}$, with a linear increase from $\boldsymbol{\tau}_{mw}(\chi\text{met}=0.1)$=1.8 ps to $\boldsymbol{\tau}_{mw}(\chi\text{met}=0.1)$=5.72 ps. When comparing water molecules in clathrate-like water and bulk water groups, we found lifetimes $\boldsymbol{\tau}_{ww}$ of bulk water to be significantly higher than $\boldsymbol{\tau}_{ww}$ of clathrate-like water for $\boldsymbol{\chi}_{\text{wat}}$<0.5 when applying a

differentiation cutoff criterion of 0.45nm (Online Resource Fig.S6) and χ_{wat}<0.2 when applying a cutoff criterion of 0.35nm (Fig.3b).

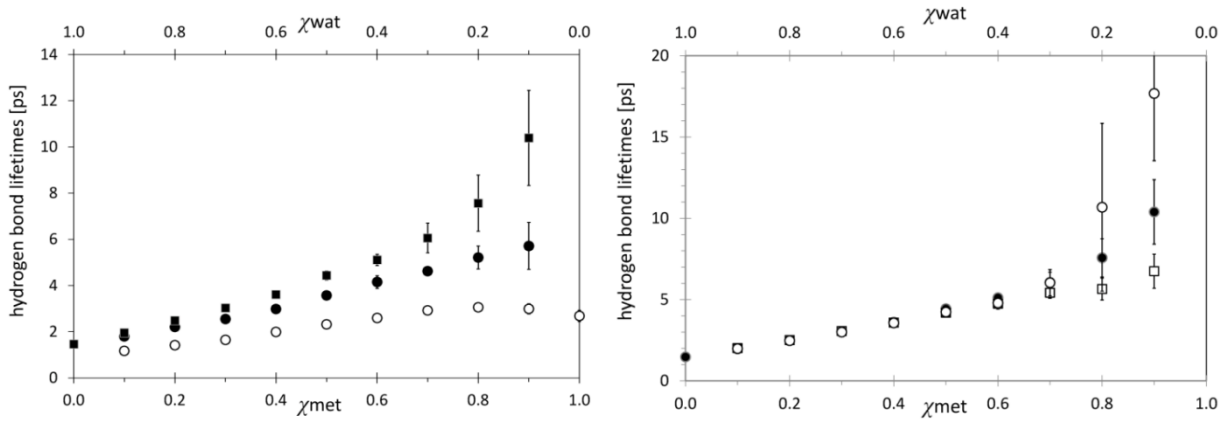


Fig.3: Hydrogen bond lifetime τ in methanol-water mixtures as the inverse of the forward lifetime of the autocorrelation function (ACF) for (a) interactions water-water τww (■), methanol-water τmw (●), methanol-methanol τmm (○) and (b) interactions water-water τww of the clathrate-like water group (□), the bulk water group (○) and all water molecules (●), after applying a spherical differentiation cutoff criterion for clathrate-like water of 0.35nm surrounding the methyl group. Error bars display standard deviations averaged over multiple trajectory intervals.

4.1.4.4 Nearest Neighbor Relationships

To find evidence for incomplete mixing at the molecular level, nearest neighbor relationship analysis between molecules within the methanol-water mixtures was performed by examining the ratios $\boldsymbol{M}_{ii}=\boldsymbol{N}_{ii}^{real}/\boldsymbol{N}_{ii}^{ideal}$ and $\boldsymbol{M}_{ij}=\boldsymbol{N}_{ij}^{real}/\boldsymbol{N}_{ij}^{ideal}$ between the number of neighbors in the simulated ($\boldsymbol{N}_{ii}^{real}$) mixtures and the calculated values for ideal mixtures ($\boldsymbol{N}_{ii}^{ideal}$) (Fig.4).

The neighbor relationships $\boldsymbol{M}_{ww}$ (central water molecule to water neighbors), $\boldsymbol{M}_{wm}$ (central water molecule to methanol neighbors), $\boldsymbol{M}_{mm}$ (central methanol molecule to methanol neighbors) and $\boldsymbol{M}_{mw}$ (central methanol molecule to water neighbors) were analyzed for all methanol-water mixtures. Starting from $\boldsymbol{M}_{ww}$=1.00 at χ_{wat}=1.0 (pure water), $\boldsymbol{M}_{ww}$ increased upon decreasing the water content, until $\boldsymbol{M}_{ww}$=1.21 was reached at χ_{wat}=0.3 and then remained constant. Conversely, $\boldsymbol{M}_{wm}$ (water-methanol neighbors) was consistently lower than ideal levels at the same mixture compositions as $\boldsymbol{M}_{ww}$, with a minimum of $\boldsymbol{M}_{wm}$=0.72 at χ_{wat}=0.9. Thus, in all methanol-water mixtures the water molecules preferred other water molecules as neighbors rather than methanol molecules (positive bias). Unlike the water-

water nearest neighbor relation, $\boldsymbol{M}_{mm}$ (methanol-methanol neighbors) decreased upon decreasing the methanol content, starting from $\boldsymbol{M}_{mm}$=1.00 at χ_{met}=1.0 (pure methanol) until $\boldsymbol{M}_{mm}$=0.67 at χ_{met}=0.1. $\boldsymbol{M}_{mw}$ (methanol-water neighbors) was consistently higher than ideal levels at the same mixture compositions as $\boldsymbol{M}_{mm}$, with a maximum of $\boldsymbol{M}_{mw}$=1.51 at χ_{met}=0.9. Thus, methanol molecules display the tendency to disfavor other methanol molecules as neighbors (negative bias), but prefer water molecules. We therefore witness demixing of methanol and water molecules within the molecular system, with water being the preferred neighbor to both water and methanol molecules.

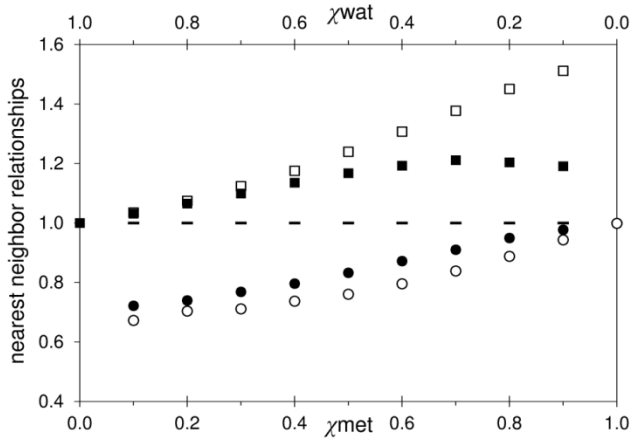


Fig.4: Nearest neighbor relationships between water and methanol molecules within a cutoff distance representative of hydrogen bonding (0.35nm) are presented for methanol-water mixtures by the relation $M_{ii}=N_{ii}^{real}/N_{ii}^{ideal}$ of neighbors N_{ii}^{real} found for simulated mixture systems to neighbors N_{ii}^{ideal} expected for ideal mixtures of the same composition. Bias in the mixture systems due to molecular interactions is perceived relative to ideal mixture expectations (dashed line) – interactions: water-water M_{ww}: ■ / methanol-water M_{mw}: □ / water-methanol M_{wm}: ● / methanol-methanol M_{mm}: ○. Standard deviations averaged over multiple trajectory intervals range from 0.15-3.45.

4.1.4.5 Cluster Formation

An analysis of largest clusters was conducted to assess hydrogen bond percolation in the methanol-water mixture system, by applying a single-linkage algorithm with a spherical distance cutoff criterion of 0.35nm (hydrogen bonding cutoff criterion) between oxygen atoms. Since the size of largest clusters vary during simulation, the largest cluster N_{max} and the smallest cluster N_{min} found throughout the entire simulation for each molar fraction was

determined, as well as the average largest cluster size $N_{aver.}$ The methanol cluster sizes (Fig.5b) decreased from N_{aver}=25 (χ_{met}=1.0) to N_{aver}=2 (χ_{met}=0.1).

The overall largest clusters were significantly larger for the entire range of mixture compositions, spanning from N_{max}=89 (χ_{met}=1.0) to N_{max}=8 (χ_{met}=0.1), whereas the overall smallest cluster spanned from N_{min}=12 (χ_{met}=1.0) to N_{min}=1 (χ_{met}=0.1). In general, an exponential decrease of largest cluster size for decreasing methanol content was observed. In contrast, the size of water largest clusters (Fig.5a) hyperbolically decreased for high water content from N_{aver}=996 at χ_{wat}=1.0 to N_{aver}=421 at χ_{wat}=0.6, until an inflection point was reached close to equimolar mixture (χ_{wat}=0.5) with N_{aver}=107. From thereon, an exponential decay was observed for χ_{wat}<0.5 until the minimum at N_{aver}=4 (χ_{wat}=0.1). The deviations of the overall largest and smallest clusters found throughout the analysis did not vary as dramatically for water as they did for methanol, except in the vicinity of the inflection point, where N_{max}=542 (χ_{wat}=0.6) and N_{min}=98 (χ_{wat}=0.6).

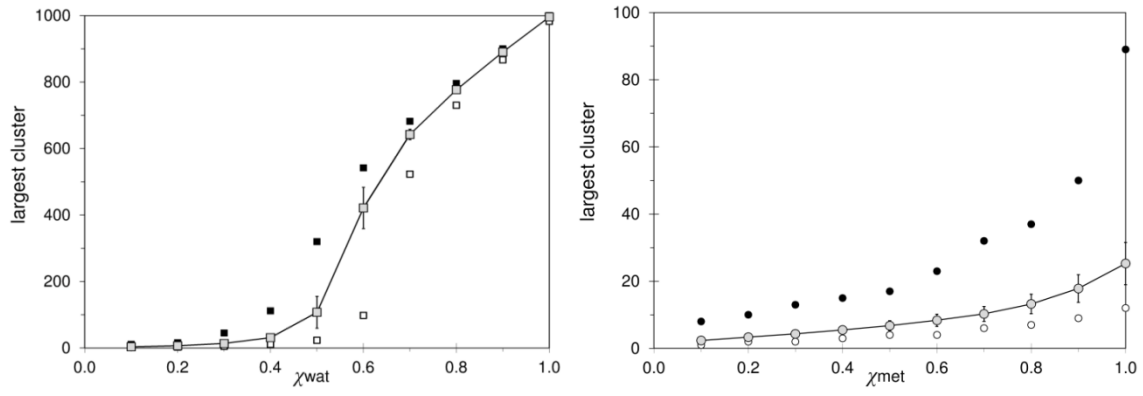


Fig.5: Clustering analysis of molecules in methanol-water mixtures using a spherical cutoff criterion of 0.35nm for (a) largest clusters of water molecules (biggest overall cluster: ■ / largest cluster on average: ■ / smallest overall cluster: □) and (b) for largest clusters of methanol (biggest overall cluster: ● / largest cluster on average: ● / smallest overall cluster: ○). Error bars display standard deviations averaged over multiple trajectory intervals.

4.1.5 Discussion

4.1.5.1 Incomplete Mixing versus Clathrate-Like Structure

The Frank and Evans model postulates increased structural order of water molecules surrounding nonpolar moieties in aqueous solutions, which, if it held true, could clarify the molecular basis of hydrophobicity in general and the lower than expected entropy of methanol-water mixtures in particular. Clathrate-like hydration is a highly controversial topic. While some investigations of aqueous solutions confirm the existence of clathrate-like hydration[35-38], others have found no evidence for structural rearrangement of water molecules surrounding nonpolar solutes in aqueous solutions[39-45]. Indeed, excess entropy in methanol-water mixtures was recently quantified solely based on radial distribution functions reflecting molecular-scale segregation within methanol-water mixtures, "without the need to invoke icebergs"[1]. It should be possible to determine if clathrate-like water structures contribute to excess entropy by comparing dynamic properties of clathrate-like water and bulk water molecules. Indeed, Dzugutov et al. proposed a formalism that directly correlates diffusion and entropy[46,25]. Particularly, reorientational correlation of water molecules has been widely used to investigate the existence of immobilized hydrate water, but immobilization could also affect translational diffusion. Therefore, differences in the dynamics of clathrate-like and bulk water are a necessary condition of the Frank and Evans model. Mere structural observations such as the existence of water clathrates around methanol molecules are insufficient to support the model. Pratt's statement that "clathrate is in the eye of the beholder"[47] frames this consideration nicely, by suggesting that while conceiving clathrate-like structures may not be necessary for a quantitative description of thermodynamic properties in aqueous solutions, if you look for them then you are likely to find them. Despite this premise, we did not find clathrate-like structures when comparing translational, rotational dynamics and HB lifetimes of molecules in the clathrate-like water group with molecules in the bulk water group. Our results are therefore in agreement with findings that refute clathrate-like structures as a source of excess entropy in methanol-water mixtures[39-45].

In spite of the sizable evidence against clathrate-like structures and the fact that quantitative theories in support of clathrate-like hydration are scarce[48], the Frank and Evans model has remained popular in the general perception of hydrophobicity and is still

the focus of widespread investigation. Human intuition plays an important role in science as a catalyst for the development of new ideas and in this regard a pictorial model is certainly more inspiring than less intuitive models like an abstract comprehension of the superimposition of proximal radial information that contradicts clathrate-hydration[1]. The conception of an equally simple and pictorial alternative model that accounts for the molecular origin of hydrophobicity without incorporating clathrate-hydration would be desirable. Thereby it is certainly not sufficient to merely disprove existing clathrate models, which appears particularly challenging due to the abundance of inconclusive and contradictory findings.

The concept of incomplete mixing[14,16,1] lays its focus on molecular-scale segregation throughout the mixture rather than localized structural arrangements. It is widely recognized and also confirmed by our results that molecular segregation in aqueous solutions is driven by hydrogen bonding preference. For methanol-water mixtures, the observed positive bias of the water-water association and the negative bias of the methanol-methanol association, as well as the preference of methanol to associate with water as a neighbor were demonstrated by nearest neighbor relationships. It indicates that the distinctive preference of water as a hydrogen bonding partner leads to a system-wide change in structural order within methanol-water mixtures. Furthermore, if hydrogen bonding is the predominant intermolecular interaction in aqueous solutions, HB lifetimes are expected to reflect the observed association bias. Indeed, the water-water HB lifetimes $\boldsymbol{\tau}_{ww}$ for all water molecules were found to increase substantially when increasing the methanol content in the mixture, whereas the methanol-methanol HB lifetimes $\boldsymbol{\tau}_{mm}$ were shown to be consistently lower than the HB lifetimes of both the water-water ($\boldsymbol{\tau}_{ww}$) and the mixed methanol-water ($\boldsymbol{\tau}_{mw}$) interactions. These results suggest incomplete mixing at the molecular level as a source for the excess effects observed for methanol-water mixtures. Thereby nearest neighbor relationships $\boldsymbol{M}_{ii}$ quantify the degree of molecular segregation and the bias witnessed for the different molecule types in simple, pictorial terms, based on a single parameter, nearest neighbor relationships, which is derived from first-principles on the molecular scale.

4.1.5.2 Hydrophobicity in the Context of Percolation

Neither of the above mentioned models offers any explanation for the localization and the molecular causes for the minimum in excess properties of methanol-water mixtures, such as entropy or molecular self-diffusion. Pure water and aqueous solutions with high water content are known to percolate[49-51], which implies that a continuous system-wide water cluster exists, in which all water molecules are interconnected via hydrogen bonds. For every point in time a lattice of water molecules can be constructed, in which a hydrogen bond between lattice points may be defined as "occupied" or "empty"[52]. Bond percolation persists as long as a single continuous graph can be constructed that connects all lattice points within the system. A percolation threshold $\boldsymbol{p}_{c=}\ \boldsymbol{\chi}^{c}_{wat}$ close to equimolar mixture has been reported for methanol-water mixtures[53,54], which was confirmed by our clustering analysis. The discrepancy between the number of maximal and minimal observable clusters in our largest cluster analysis ($\boldsymbol{N}_{\max}$ and $\boldsymbol{N}_{\min}$, respectively) suggests that the transition from one continuous water cluster (percolation) to isolated water clusters is not precisely localized. For $\boldsymbol{\chi}_{wat}$=0.6, system-wide water molecule clusters (e.g. cluster size $\boldsymbol{N}_{\max}$=542) exist temporarily, while at other times only isolated largest clusters can be found in the mixtures (e.g. cluster size $\boldsymbol{N}_{\max}$=98). It should also be noted that the presented data does not consider how methanol is involved in the hydrogen network, which might slightly lower the percolation threshold with the formation of bi-percolating networks[53].

A relation between percolation and anomalies in thermodynamic quantities has previously been suggested[53], and hydrogen bonding has been linked to macroscopic excess properties at critical mixture compositions[55]. A crossover point for different hydrogen bonding environments from water-rich to methanol-rich environment was previously proposed at $\boldsymbol{\chi}_{wat}$~0.4[56]. Since the percolation threshold for water molecules at $\boldsymbol{\chi}_{wat}$~0.5 coincides with the minimum in molecular self-diffusion for both methanol and water, it is assumed that the mobility of molecules, and thus entropy, is directly affected by percolation. The HB lifetimes analysis underscores this hypothesis, particularly $\boldsymbol{\tau}_{ww}$, which depends linearly on $\boldsymbol{\chi}_{wat}$ for $\boldsymbol{\chi}_{wat}$<0.5 and exponentially for $\boldsymbol{\chi}_{wat}$>0.5, indicating a fundamental change in hydrogen bonding dynamics at the percolation threshold. The difference in increased $\boldsymbol{\tau}_{ww}$ for bulk water in comparison to clathrate-like water suggests that this effect arises due to the significant enhancement of water-water hydrogen bonding in the bulk, which is exactly the opposite effect of what would be expected of clathrate-like

hydration and thus is in stark contradiction to the Frank and Evans model. It is however in agreement with the view that the loss of hydrogen bonding drives the segregation of nonpolar moieties from water[2] and offers a fresh perspective on how enthalpic and entropic contributions to the free energy of solvation may be interpreted for methanol-water mixtures.

In general, hydrophobicity in aqueous solutions is considered to be a multifaceted problem[2,3], where the enthalpic breaking of hydrogen bonds dominates for large solutes, whereas for small solutes in aqueous solution, like methanol-water mixtures, the free energy is perceived to be mainly affected by entropic reordering of hydrogen bonds. In methanol-water mixtures, entropic reordering of water molecules to maintain bulk-like tetrahedral hydrogen bond coordination may obviously only take place above the percolation threshold (χ_{wat}>0.5), where a sufficient number of water molecules are available as bonding partners. Therefore, the formation or the breakdown of the system-wide hydrogen bonding network have a major effect on hydrogen bond dynamics. On the basis of our findings, we suggest that the percolation threshold is not only a critical threshold for excess properties such as molecular self-diffusion, but is also a divisive threshold for two distinctively different hydrophobicity regimes. A possible pictorial model is described in the following subsections.

4.1.5.3 Molecular Mechanism of Water Diffusion

The observed minimum of molecular self-diffusion of water at $\chi_{wat}\approx$0.5 could thus be explained: For high water content (χ_{wat}>0.5) in the methanol-water mixture, the percolation in the system-wide hydrogen bonding network persists as the association bias between water molecules becomes more prevalent and individual hydrogen bonds are maintained for longer durations. Decreasing the water content in the mixture coincides with an increasing nearest neighbor bias $\boldsymbol{M}_{\mathrm{ww}}$ and rising HB lifetimes τ_{ww}, suggesting an enhanced water-water interaction, thus reducing the overall mobility of water molecules as fewer binding partners are available. Just above the percolation threshold (χ^{c}_{wat}>0.5), the network resembles a tight mesh of water molecules and its restricted mobility explains the minimal diffusion coefficient and the high degree of structural order. Below the percolation threshold (χ_{wat}<0.5) the hydrogen bonding network breaks down and we find isolated water

clusters, the size of which decrease with decreasing water content. While at χ_{wat}=0.4 the majority of clusters consist of 10-50 water molecules, at χ_{wat}=0.1 the majority of clusters found are below size 10 (Online Resource Fig.S7b). For clusters of decreasing diameter, the observed increase in diffusion could be explained by the Einstein-Stokes relation. This would also explain the significant increase of the HB lifetimes $\boldsymbol{\tau}_{ww}$ for χ_{wat}<0.5 between water molecules in the bulk water group, while $\boldsymbol{\tau}_{ww}$ in the clathrate-like water group remain at relatively low levels.

4.1.5.4 Molecular Mechanism of Methanol Diffusion

Nearest neighbor relationships ($\boldsymbol{M}_{\mathrm{mm}}$ and $\boldsymbol{M}_{\mathrm{mw}}$) suggest that methanol molecules preferably associate with water molecules in the mixture rather than with other methanol molecules. This conclusion is supported by the fact that HB lifetimes $\boldsymbol{\tau}_{mw}$ between methanol and water molecules are consistently higher than HB lifetimes $\boldsymbol{\tau}_{mm}$ between methanol molecules. In contrast to water, pure methanol does not form a system-wide network, but clusters into small rings or chains[57,58]. At low water content, water molecules do not perturb this local structure, which is in agreement with the constant HB lifetime $\boldsymbol{\tau}_{mm}$ between methanol molecules at χ_{met}>0.5. As the water content increases, the size of the water clusters increases and methanol molecules increasingly interact with the water clusters, which could explain an overall decrease in diffusion of methanol adapting to decreasing diffusion of water. At the percolation threshold $\chi^{c}_{wat}\approx$0.5, the formation of a system-wide HB network of water significantly perturbs the existing methanol structures by spatial constraints. The decreasing methanol-water HB lifetimes could be explained by competing water-water HB in the system-wide network, which would effectively result in increased mobility of methanol molecules.

4.1.6 Summary

Clathrate-like hydration as a source for excess entropy in methanol-water mixtures could not be observed. At high water content, the dynamic properties of clathrate-like water and of bulk water were similar. At low water content, the hydrogen bond lifetimes of bulk water were significantly enhanced in comparison to clathrate-like water, contradicting

clathrate-like hydration. Instead, incomplete mixing was observed and quantified by nearest neighbor relationships, which offers an intuitive and pictorial way to describe hydrophobicity. Furthermore, the hydrogen bond percolation threshold ($\chi^{c}_{wat} \approx 0.5$) in the mixture was shown to coincide with the minimum in molecular self-diffusion. Two distinctively different hydrophobicity regimes were conceived: At low water content, isolated water clusters are embedded in loosely coupled methanol molecule structures, while at high water content methanol molecules are embedded in a system-wide hydrogen-bonded water network.

4.1.7 Acknowledgements

We thank Joachim Groß (University of Stuttgart) for helpful discussions. This work was funded by SimTech Cluster of Excellence at the University of Stuttgart.

4.1.8 Supporting Information

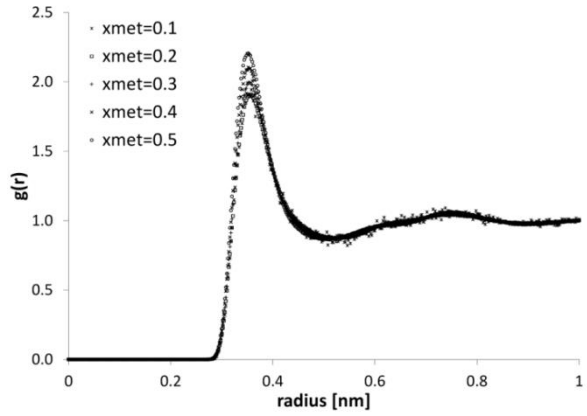


Fig. S1: Radial distribution functions g(r) C1-Ow for methanol-water mixture compositions χ_{wat}=0.1-0.5 used to distinguish the first coordination sphere of possible "clathrate" water molecules (CLATH group) surrounding the aliphatic methyl moiety of methanol molecules.

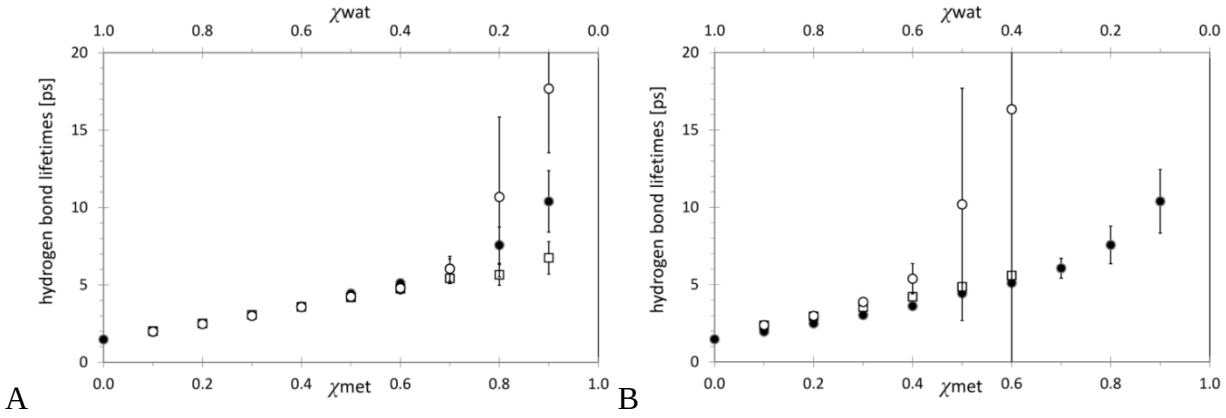


Fig. S2: Hydrogen bond lifetime τ in methanol-water mixtures as the inverse of the forward lifetime of the autocorrelation function (ACF) for interactions water-water τ_{ww} of groups CLATH (□), group REST(○), all water molecules (●), after applying a spherical differentiation cutoff criterion for CLATH of (A) 0.35nm and (B) 0.45nm surrounding the methyl group.

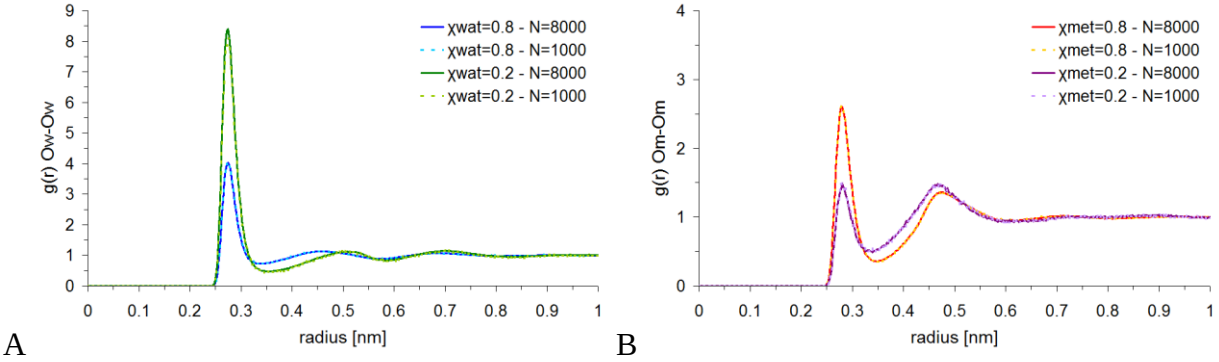


Fig. S3: Comparison of radial distribution functions (A) g(r) Ow-Ow and (B) g(r) Om-Om for methanol-water mixture compositions χ_{wat}=0.2 and χ_{wat}=0.8 between systems of different absolute size (number of molecules N). Small systems (N=1000) equilibrated to the same structural arrangement of molecules as large systems (N=8000), thus small systems proved adequate for the structural analysis conducted in this work.

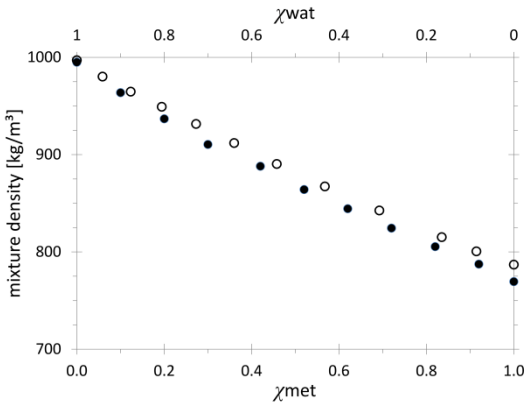


Fig. S4: Density $\rho(\chi_i)$ of the methanol-water mixtures: In contrast to the self-diffusion coefficient, the density dependence on molar fractions $\rho(\chi_i)$ of methanol-water mixtures was almost linear (black dots), corresponding to the behavior expected of an ideal mixture and ranges from $\rho_{\chi wat=1.0}$=0.995g/cm³ for pure water to $\rho_{\chi met=1.0}$=0.770g/cm³ for pure methanol. The standard deviation ranged from 0.003g/cm³ to 0.004g/cm³. Experimental data (Mikhail & Kimel 1961) was added for comparison (white dots).

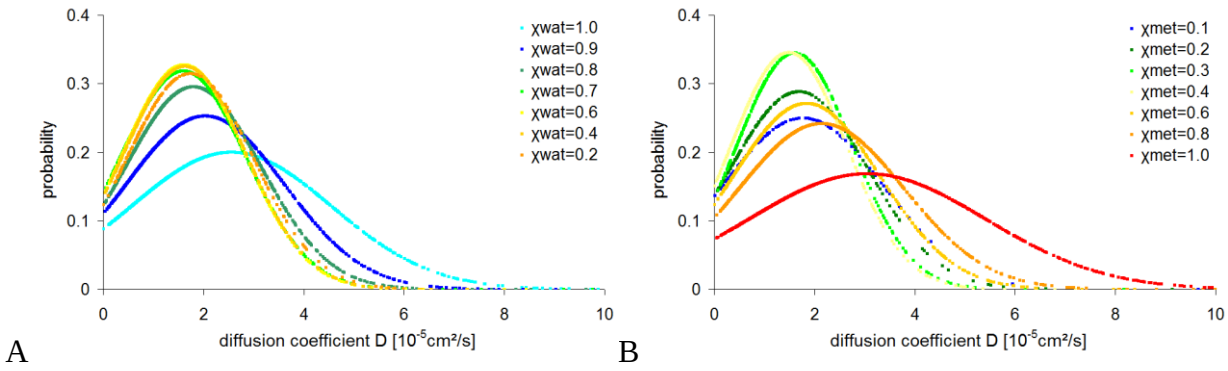


Fig. S5: Probability density functions of the self-diffusion coefficients of (A) all water molecules and (B) all methanol molecules in the simulated methanol-water mixture systems. Standard deviations vary from max. 1.901 ·10^{-5} cm²/s at χ_{wat}=1.0 to min. 1.145 ·10^{-5} cm²/s at χ_{wat}=0.4 for water and from max. 2.361 ·10^{-5} cm²/s at χ_{met}=1.0 to min. 1.120 ·10^{-5} cm²/s at χ_{met}=0.3 for methanol.

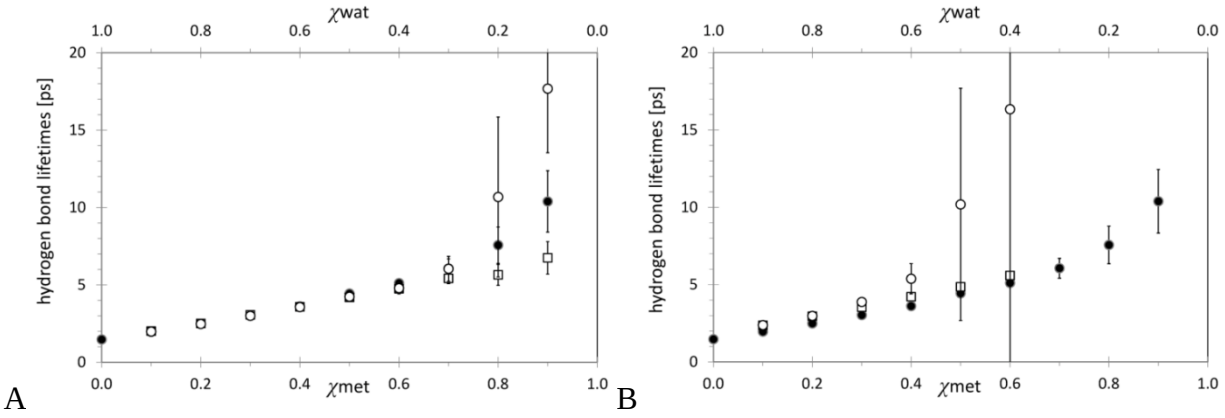


Fig. S6: Hydrogen bond lifetime τ in methanol-water mixtures as the inverse of the forward lifetime of the autocorrelation function (ACF) for interactions water-water τ_{ww} of groups CLATH (□), group REST(○), all water molecules (●), after applying a spherical differentiation cutoff criterion for CLATH of (A) 0.35nm and (B) 0.45nm surrounding the methyl group.

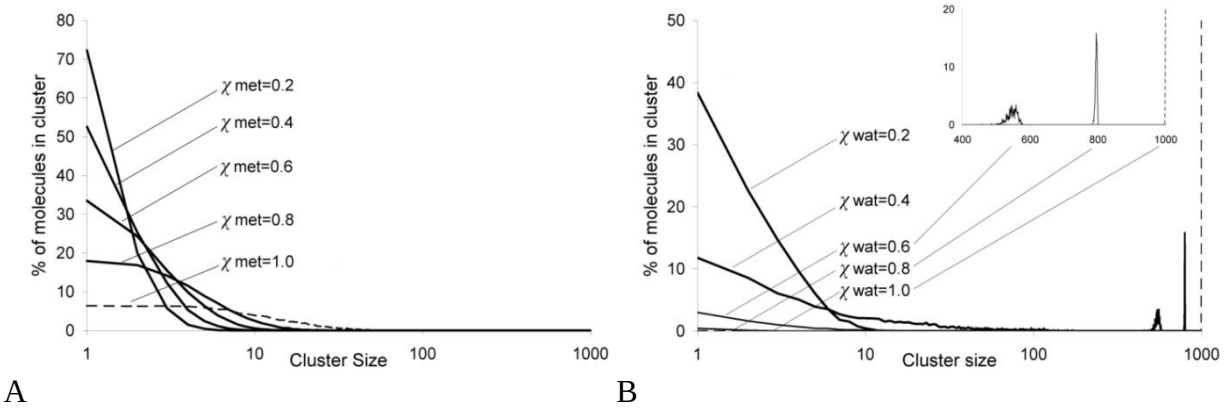


Fig. S7: Clustering behavior of methanol (A) and water (B) molecules in methanol-water mixtures at molar fractions χ_{met} and χ_{wat}, respectively. Data was gathered from a single-linkage algorithm between oxygen atoms using a spherical cutoff radius (0.35nm). Continuous lines are applied for the purpose of visual distinction between the clustering behavior at the different molar fractions and representative of discrete integer values for the different cluster sizes.

4.1.9 References

1. Soper AK, Dougan L, Crain J, Finney JL (2006) Excess Entropy in Alcohol-Water Solutions: a Simple Clustering Explanation. J Phys Chem B 110 (8):3472-3476

2. Chandler D (2002) Hydrophobicity: Two Faces of Water. Nature 417 (6888):491-491

3. Galli G (2007) Dissecting Hydrophobicity. Proceedings of the National Academy of Sciences of the United States of America 104 (8):2557-2558

4. Jones JE (1924) On the Determination of Molecular Fields. II. From the Equation of State of a Gas. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character 106 (738):463-477

5. Frank HS, Evans MW (1945) Free Volume and Entropy in Condensed Systems .3. Entropy in Binary Liquid Mixtures - Partial Molal Entropy in Dilute Solutions - Structure and Thermodynamics in Aqueous Electrolytes. J Chem Phys 13 (11):507-532

6. Kauzmann W (1959) Some Factors in the Interpretation of Protein Denaturation. Advances in Protein Chemistry 14:1-63

7. Dixit S, Poon WCK, Crain J (2000) Hydration of Methanol in Aqueous Solutions: a Raman Spectroscopic Study. J Phys: Condens Matter 12 (21):323-328

8. Soper AK, Finney JL (1993) Hydration of Methanol in Aqueous-Solution. Physical Review Letters 71 (26):4346-4349

9. Wakisaka A, Komatsu S, Usui Y (2001) Solute-Solvent and Solvent-Solvent Interactions Evaluated Through Clusters Isolated from Solutions: Preferential Solvation in Water-Alcohol Mixtures. Journal of Molecular Liquids 90 (1-3):175-184

10. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys 79 (2):926-935

11. Meng EC, Kollman PA (1996) Molecular Dynamics Studies of the Properties of Water Around Simple Organic Solutes. Journal of Physical Chemistry 100 (27):11460-11470

12. Morrone JA, Hasllinger KE, Tuckerman ME (2006) Ab Initio Molecular Dynamics Simulation of the Structure and Proton Transport Dynamics of Methanol-Water Solutions. J Phys Chem B 110 (8):3712-3720. doi:10.1021/jp0554036

13. van Erp TS, Meijer EJ (2001) Hydration of Methanol in Water. A DFT-Based Molecular Dynamics Study. Chemical Physics Letters 333 (3-4):290-296

14. Dixit S, Crain J, Poon WCK, Finney JL, Soper AK (2002) Molecular Segregation Observed in a Concentrated Alcohol-Water Solution. Nature 416 (6883):829-832

15. Finney JL, Bowron DT, Daniel RM, Timmins P, Roberts MA (2003) Molecular and Mesoscale Structures in Hydrophobically Driven Aqueous Solutions. Biophysical Chemistry 105 (2-3):391-409. doi:10.1016/s0301-4622(03)00104-2

16. Guo JH, Luo Y, Augustsson A, Kashtanov S, Rubensson JE, Shuh DK, Agren H, Nordgren J (2003) Molecular Structure of Alcohol-Water Mixtures. Physical Review Letters 91 (15)

17. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR (1984) Molecular-Dynamics with Coupling to an External Bath. J Chem Phys 81 (8):3684-3690

18. Hockney RW, Goel SP, Eastwood JW (1974) Quiet High-Resolution Computer Models of a Plasma. Journal of Computational Physics 14 (2):148-158

19. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical-integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of n-Alkanes. Journal of Computational Physics 23 (3):327-341

20. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A Smooth Particle Mesh Ewald method. J Chem Phys 103 (19):8577-8593

21. York DM, Darden TA, Pedersen LG (1993) The Effect of Long-Range Electrostatic Interactions in Simulations of Macromolecular Crystals - a Comparison of the Ewald and Truncated List Methods. J Chem Phys 99 (10):8345-8348

22. Jorgensen WL, Maxwell DS, TiradoRives J (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. J Am Chem Soc 118 (45):11225-11236

23. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. Journal of Chemical Theory and Computation 4 (3):435-447

24. van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) GROMACS: Fast, Flexible, and Free. Journal of Computational Chemistry 26 (16):1701-1718

25. Dzugutov M (2001) A Universal Scaling Law for Atomic Diffusion in Condensed Matter. Nature 411 (6838):720-720

26. Allen M.P. TDJ (1987) Computer Simulations of Liquids Oxford University Press Inc, Oxford,

27. Luzar A (2000) Resolving the Hydrogen Bond Dynamics Conundrum. J Chem Phys 113 (23):10663-10675

28. Rapaport DC (1983) Hydrogen-Bonds in Water Network Organization and Lifetimes. Molecular Physics 50 (5):1151-1162

29. van der Spoel D, van Maaren PJ, Berendsen HJC (1998) A Systematic Study of Water Models for Molecular Simulation: Derivation of Water Models Optimized for Use with a Reaction Field. J Chem Phys 108 (24):10220-10230

30. Halle B, Wennerstrom H (1981) Interpretation of Magnetic-Resonance Data from Water Nuclei in Heterogeneous Systems. J Chem Phys 75 (4):1928-1943

31. Struis R, Debleijser J, Leyte JC (1987) Dynamic Behavior and Some of the Molecular Properties of Water Molecules in Pure Water and in MgCl2 Solutions. Journal of Physical Chemistry 91 (6):1639-1645

32. Derlacki ZJ, Easteal AJ, Edge AVJ, Woolf LA, Roksandic Z (1985) Diffusion-Coefficients of Methanol and Water and the Mutual Diffusion-Coefficient in Methanol Water Solutions at 278-K and 278-K. Journal of Physical Chemistry 89 (24):5318-5322

33. Mikhail SZ, Kimel WR (1961) Densities and Viscosities of Methanol-Water Mixtures. J Chem Eng Data 6 (4):533-537. doi:10.1021/je60011a015

34. Wensink EJW, Hoffmann AC, van Maaren PJ, van der Spoel D (2003) Dynamic Properties of Water/Alcohol Mixtures Studied by Computer Simulation. J Chem Phys 119 (14):7308-7317. doi:10.1063/1.1607918

35. Lankau T, Konrad O (2007) Hydrophobic Solvation: Aqueous Methane Solutions. Journal of Chemical Education 84 (5):864

36. Bowron DT, Filipponi A, Roberts MA, Finney JL (1998) Hydrophobic Hydration and the Formation of a Clathrate Hydrate. Physical Review Letters 81 (19):4164-4167

37. Titantah JT, Karttunen M (2012) Long-Time Correlations and Hydrophobe-Modified Hydrogen-Bonding Dynamics in Hydrophobic Hydration. Journal of the American Chemical Society (22):9362-9368

38. Rezus YLA, Bakker HJ (2007) Observation of Immobilized Water Molecules Around Hydrophobic Groups. Physical Review Letters 99 (14):148301

39. Jorgensen WL, Buckner JK (1987) Use of Statistical Perturbation-Theory for Computing Solvent Effects on Molecular-Conformation - Butane in Water. Journal of Physical Chemistry 91 (24):6083-6085

40. Kincaid RH, Scheraga HA (1982) Acceleration of Convergence in Monte-Carlo Simulations of Aqueous-Solutions Using the Metropolis Algorithm - Hydrophobic Hydration of Methane. Journal of Computational Chemistry 3 (4):525-547

41. Rosenberg RO, Mikkilineni R, Berne BJ (1982) Hydrophobic Effect on Chain Folding - the Trans to Gauche Isomerization of Normal-Butane in Water. Journal of the American Chemical Society 104 (26):7647-7649

42. Rossky PJ, Zichi DA (1982) Molecular Librations and Solvent Orientational Correlations in Hydrophobic Phenomena. Faraday Symposia of the Chemical Society (17):69-78

43. Rossky PJ, Friedman HL (1980) Benzene-Benzene Interaction in Aqueous-Solution. Journal of Physical Chemistry 84 (6):587-589

44. Laage D, Stirnemann G, Hynes JT (2009) Why Water Reorientation Slows without Iceberg Formation Around Hydrophobic Solutes. The Journal of Physical Chemistry B 113 (8):2428-2435

45. Qvist J, Halle B (2008) Thermal Signature of Hydrophobic Hydration Dynamics. Journal of the American Chemical Society 130 (31):10345-10353

46. Dzugutov M (1996) A Universal Scaling Law for Atomic Diffusion in Condensed Matter. Nature 381 (6578):137-139

47. Pratt LR (2002) Molecular Theory of Hydrophobic Effects: "She Is Too Mean To Have Her Name Repeated." Annual Review of Physical Chemistry 53:409-436

48. Hummer G, Garde S, Garcia AE, Pratt LR (2000) New Perspectives on Hydrophobic Effects. Chemical Physics 258 (2-3):349-370

49. Blumberg RL, Stanley HE, Geiger A, Mausbach P (1984) Connectivity of Hydrogen-Bonds in Liquid Water. J Chem Phys 80 (10):5230-5241

50. Geiger A, Stillinger FH, Rahman A (1979) Aspects of the Percolation Process for Hydrogen-Bond Networks in Water. J Chem Phys 70 (9):4185-4193

51. Stanley HE, Teixeira J (1980) Interpretation of the Unusual Behavior of H2O and D2O at Low-Temperatures - Tests of a Percolation Model. J Chem Phys 73 (7):3404-3422

52. Shante VKS, Kirkpatrick S (1971) Introduction to Percolation Theory. Advances in Physics 20 (85):325

53. Dougan L, Bates SP, Hargreaves R, Fox JP, Crain J, Finney JL, Reat V, Soper AK (2004) Methanol-Water Solutions: a Bi-Percolating Liquid Mixture. J Chem Phys 121 (13):6456-6462. doi:10.1063/1.1789951

54. Bako I, Megyes T, Balint S, Grosz T, Chihaia V (2008) Water-Methanol Mixtures: Topology of Hydrogen Bonded Network. Phys Chem Chem Phys 10 (32):5004-5011. doi:10.1039/b808326f

55. Sato T, Chiba A, Nozaki R (2000) Hydrophobic Hydration and Molecular Association in Methanol-Water Mixtures Studied by Microwave Dielectric Analysis. J Chem Phys 112 (6):2924-2932

56. Pascal TA, Goddard WA Hydrophobic Segregation, Phase Transitions and the Anomalous Thermodynamics of Water/Methanol Mixtures. J Phys Chem B 116 (47):13905-13912

57. Haughney M, Ferrario M, McDonald IR (1987) Molecular-Dynamics Simulations of Liquid Methanol. Journal of Physical Chemistry 91 (19):4934-4940

58. Yamaguchi T, Hidaka K, Soper AK (1999) The Structure of Liquid Methanol Revisited: a Neutron Diffraction Experiment at -80 Degrees C and +25 Degrees C. Molecular Physics 97 (4):603-605

4.2 Molecular Dynamics Simulations of Self-Emulsifying Drug Delivery Systems (SEDDS): Influence of Excipients on Droplet Nanostructure and Drug Localization

4.2.1 Abstract

In this study, molecular dynamics (MD) simulations were applied to model the lipidic nanoscale droplets that form when self-emulsifying drug delivery systems (SEDDS) disperse into microemulsions in the gastrointestinal (GI) tract. The major focus lay on investigating the influence of SEDDS excipient composition on the nanostructure of droplets and the localization of drug molecules within them, by monitoring the drug immersion depth and the molecular association bias between hydrophilic and hydrophobic moieties. A SEDDS standard system consisting of capric (C10) fatty acid chain length triglycerides and drug molecule cyclosporin A (CyA) was compared to systematic excipient variations. Investigating the drug loading capacity of droplets yielded a negligible influence of drug molecules on the droplet nanostructure; increasing the drug load merely resulted in increased drug exposure to the aqueous environment. The variation of triglyceride fatty acid chain lengths yielded clearly distinguishable droplet association patterns (unstructured, lamellar-like and vesicle-like), which could prove beneficial for predicting and engineering drug solubilization in SEDDS. The addition of surfactant poly (ethylene glycol) (PEG-6) revealed the formation of an encapsulating surfactant shell with a negligible impact on the droplet interior triglyceride nanostructure, which could potentially be utilized to protect drug molecules from digestion. Mono- and diglyceride molecules displayed an increased tendency to accumulate at the droplet surface as well, in accordance with their capacity to act as surfactants, while also significantly interfering with the interior droplet nanostructure. The addition of monoglyceride molecules in particular caused the CyA molecule to be solubilized in a hydrophilic droplet core region consisting of polar triglyceride moieties. This mode of drug localization was in stark contrast to that of other systems, where CyA was predominantly found in the interfacial region of the aqueous environment.

4.2.2 Introduction

The oral delivery of hydrophobic drugs represents a significant technical challenge for the development of pharmaceutical products, with approximately 40% of newly launched drugs displaying poor solubility in aqueous media. Self-emulsifying drug-delivery systems (SEDDS) are lipidic formulations that increase the bioavailability and stabilize absorption rates of orally administered hydrophobic drugs. Drug bioavailability is generally considered to be a function of permeation and dissolution.[1] Herein SEDDS distinguish themselves from other lipidic formulations by the capacity to spontaneously form stable isotropic oil-in-water (o/w) microemulsions of fine lipidic droplets upon mild agitation in aqueous media. Drug molecules are thereby homogeneously distributed throughout the gastrointestinal tract (GI) while being embedded in a dissolved state within a precipitation-suppressing nanoenvironment. This makes SEDDS particularly promising vehicles for hydrophobic drug compounds that exhibit dissolution-rate-limited absorption; however, the benefits are considered to be transferable to other drug categories as well.[2, 3] SEDDS designs are generally classified according to their excipient composition of oils, surfactants/emulsifiers and cosolvents/coemulsifiers.[4, 5] The most basic form of SEDDS (type I) consists of drug and oil molecules, most commonly triglycerides with medium-chain-length fatty acids.[3] The preference for simple triglyceride systems stems from the fact that they are readily digested by pancreatic lipases, which entails health and safety benefits. However, type I formulations are known to exhibit relatively poor loading capacity. To counter this limitation, type II+ formulations are designed with the addition of surfactants (and coemulsifiers), which reduces the size of stable dispersion droplets to the micrometer (SMEDDS) and nanometer (SNEDDS) scales. While significant progress has been made toward controlling the microscopic dispersion properties of SEDDS, i.e., size and stability of droplets, few methods exist to evaluate the droplet nanostructure, which is considered to be intimately linked to drug solubility, loading capacity, release rates, and localization.[6-15] As a recent physiology-based pharmacokinetics solubility study demonstrates, effective methods of accurately predicting drug solubility are lacking.[16] Although phase diagram studies, simplified mathematical models, and computational studies have shown promise in predicting solubility for practical applications,[17-19] most of these methods are restricted to the microscale. However, to fully exploit the potential of SEDDS formulation design, additional methods are desirable that can predictively assess

and contextualize the nanostructure, drug solubilization, and SEDDS composition. In fact, the estimation of drug solubility in the context of the complex nanostructures of lipid aggregates has been identified as one of the major challenges in drug-delivery formulation design.[20] It is no trivial task to predetermine a phase structure of lipidic formulations, even when its excipients are clearly defined. Besides the thermodynamics of self-assembly, the forces between the amphiphilic lipid molecules within the aggregates as well as the solution conditions need to be accounted for.[21] Indeed, ideal solubility and regular solubility theories were shown to fail in adequately predicting the solubility of drugs in lipidic solutions, due to the complex and disparate interfaces that form upon mixing different excipients.[22] Under GI tract solution conditions in particular, bile salts and hydrolytic enzymes represent additional factors of uncertainty; during the digestive phase, they can influence the composition and nanostructure of lipid aggregates in fundamental ways.[3] Due to its rigorous first-principles methodology, molecular dynamics (MD) shows great potential in elucidating both the SEDDS droplet nanostructure and the localization of drug molecules therein, to the extent that MD was even advocated as the most powerful route for predicting solubility on the basis of chemical structures[20]. It is therefore surprising that although the literature on the MD characterization of liquid-phase molecular self-assembly is broad, studies mostly cover single systems or a limited range of systems, phase transitions or the characterization of distinct phases, e.g., lipid bilayer formation.[23] To our knowledge, only a few MD studies directly focus on lipidic formulations for pharmaceutical applications[24-26] and only a single study specifically covers SEDDS, but with a focus on the formation of o/w microemulsions and not the nanostructure of SEDDS droplets in particular.[27] This study, on the other hand, sets out to explore the potential of MD in characterizing single SEDDS droplets, with the aim of evaluating the implications of systematic excipient variations on droplet nanostructure (molecular association bias) and drug localization (immersion depth).

4.2.3 Experimental Section

4.2.3.1 Simulation Details

Molecular dynamics simulations were performed with GROMACS 4.5.3[28, 29] software at T=310 K and 1 bar, using the leapfrog integrator[30] with a time step of 4 fs. Simulations were conducted in an NPT ensemble, applying Nose-Hoover[31, 32] temperature coupling in 0.4 ps intervals and isotropic Parrinello-Rahman[33] pressure coupling in 4.0 ps intervals. The center-of-mass movement was removed every 0.4 ps independently for the drug, triglycerides/surfactants, and water molecules. Periodic boundary conditions were applied in all dimensions. All bonds were constrained with the LINCS algorithm.[34] Long-range electrostatics were calculated with the particle mesh Ewald (PME) method.[35, 36] Lennard-Jones interactions were treated with a cutoff and capped at 1.4 nm. For cyclosporin A (CyA) parametrization the united atom GROMOS 53a6 force field[37] was used, which already incorporates the necessary nonstandard amino acid residues of CyA. Bonds, angles, and dihedrals for cyclization were added to the force field in accordance with standard peptide bonds with appropriate reparameterization of the termini. To parametrize tri-, di- and monoglycerides, the Berger lipids model[38] was added to the GROMOS 53a6 force field in conjunction with $53A6_{OXY}$ parameters[39]. Surfactant poly(ethylene glycol) (PEG) was parametrized with 53A6_OE charges and Lennard-Jones parameters.[40] The SPCE[41] water model was applied.

4.2.3.2 Simulation Systems

Since the phase structure of triglyceride/water or triglyceride/surfactant/water mixtures is known to depend on solute/solvent ratios as well as packing parameters and is susceptible to periodic boundary condition artifacts,[42] preliminary trial runs with varied size and atom density were conducted until a single droplet of our model system was formed without artifacts. Subsequently, the initial conformation of all systems under consideration was set up with a constant atom number density, $\mathbf{N}_{ATOM}/[nm^3]$(Table 1) to ensure comparable starting conditions. To reduce computational demand, the droplet sizes were set to the low limit (~7.5 nm) of SEDDS droplets found in commercial applications.[3] All systems were generated as unbiased random molecular arrangements.

Table 1. Summary of droplets under investigation[a]

no.	CyA	excipients	initial density $N_{ATOMS}/[nm^3]$	no. molecules $N_{SUBSTRATE}$	no. molecules $N_{SURFACTANT}$	no. molecules N_{WATER}	composition % w_{SOLUTE}/w_{WATER}	sim. time [ns]	final diam. [nm]	final box volume $[nm^3]$
1	0	C10	4.04	200	0	57059	10.80	200	7.5	1912
2	1	C10	4.04	200	0	57424	10.84	200	7.5	1925
3	10	C10	4.04	200	0	63564	10.74	200	7.5	2126
4	1	C6	4.04	200	0	40104	10.78	200	7.3	1339
5	1	C14	4.04	200	0	75416	10.73	200	8.6	2530
6	1	C10 3 : 2 PEG-6	4.04	120	80	47636	10.53	200	7.5	1587
7	1	C10 4 : 1 PEG-6	4.04	160	40	52522	10.70	200	7.4	1756
8	1	C10 - TRI 1 : 1 DI	4.04	200	0	50532	10.63	200	7.6	1689
9	1	C10 - TRI 1 : 1 MONO	4.04	200	0	41492	10.88	200	7.1	1388

[a]Molecules were randomly placed into cubic systems at an equal atom number density of 4.04 atoms / [nm³] and simulated for a total of 200ns to form droplets of diameter ~7.5nm.

Systems were energy minimized with the steepest-descent algorithm, and equilibration was performed for at least 100 ns in total. The convergence of the potential energy, the atom density and the molecular association bias to constant values was monitored to ensure equilibration. An additional 100 ns simulation time was used for analysis. The entire process was repeated three times with varied random molecular arrangements and velocity distributions to improve the statistical significance.

4.2.3.3 Analysis

The GROMACS 4.5.3 software package was applied to monitor the clustering behavior of SEDDS droplets. Hereby, nonwater atoms within d<0.5 nm were defined as belonging to a cluster. Custom-made scripts were applied to analyze the drug immersion depth and the molecular association bias within droplets. The drug immersion depth was defined as the distance of the center of mass (COM) of a drug molecule to the closest water molecules of the droplet's first water solvation shell ($\mathbf{d_{com}}$), as well as the minimal distance of any drug molecule atom to the water solvation shell ($\mathbf{d_{min}}$) (Figure 1A).

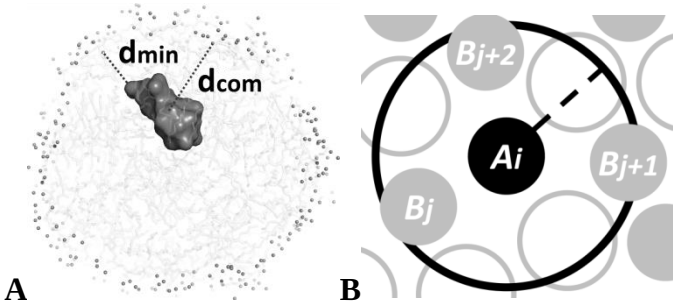


Figure 1. SEDDS simulation analysis. (A) Method of analyzing the immersion depth: d_{com}, center-of-mass distance to the water solvation shell; d_{min}, minimal distance of a drug atom to the water solvation shell. (B) Method of analyzing the molecular association bias: atoms of molecule type B were counted as neighbors when located within a spherical cutoff distance from central atoms of type A.

The molecular association bias[43] was quantified between several groups of atoms: CyA drug atoms, PEG-6 molecules, water molecules, and the polar (hydrophilic) or nonpolar (hydrophobic) moieties of triglycerides. Neighboring atoms of group B with respect to central atoms of group A within a radial cutoff of d<0.35 nm were defined as nearest neighbors (Figure 1B). The molecular association bias $\boldsymbol{M_{AB}}$ is defined as the probability relation of the nearest neighbors derived from simulation $\boldsymbol{p_{AB}^{real}}$ to the expected values of ideal mixtures $\boldsymbol{p_{AB}^{ideal}}$:

$$M_{AB} = \frac{p_{AB}^{real}}{p_{AB}^{ideal}} \qquad p_{AB}^{ideal} = \frac{n_B}{n_{total}} \qquad p_{AB}^{real} = \frac{N_{AB}}{\sum_i N_{Ai}} \qquad i = A,B,C,...$$

Hereby, $\boldsymbol{n_B}$ denotes the number of atoms belonging to group *B*, and $\boldsymbol{n_{total}}$ denotes the total number of atoms in the analyzed system. $\boldsymbol{N_{AB}}$ denotes the total number of neighbors of group *B* to all atoms of central group *A*, which is divided by the total sum of all neighbors to atoms of group *A*. For ideal binary mixtures, we therefore expect $\boldsymbol{M_{AB}}$=1, while $\boldsymbol{M_{AB}}$>1 is expected for a positive bias if molecules of group *A* favor the proximity to molecules of group *B* and $\boldsymbol{M_{AB}}$<1 for a negative molecular association bias.

4.2.4 Results

Single SEDDS droplets of oil-in-water (o/w) microemulsions containing the drug cyclosporin A (CyA) were modeled by molecular dynamics (MD) simulations, with an emphasis on droplet nanostructure and drug localization. The droplet nanostructure was analyzed by quantifying the molecular association bias between hydrophobic (PHOB) and hydrophilic (PHIL) moieties of SEDDS excipients, whereas drug localization was quantified by the immersion depth of drug molecules within the droplet nanostructure. To systematically evaluate the potential of this MD approach to SEDDS design, we defined a standard system consisting of a single CyA drug molecule in conjunction with capric (C10) triglyceride molecules and introduced or altered excipients that are commonly applied in commercial SEDDS formulations (Figure 2).

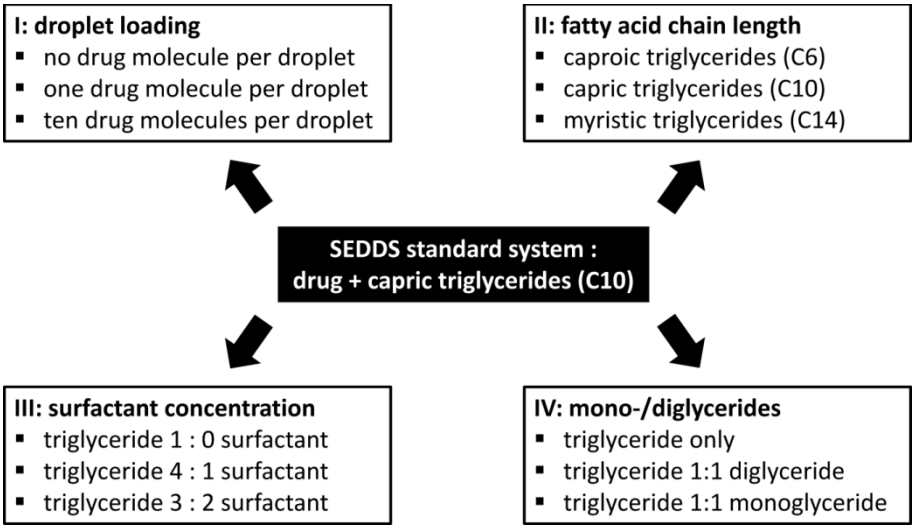


Figure 2: A type I SEDDS system consisting of cyclosporin A and capric triglycerides in aqueous solution was varied in its excipient composition. (I) Droplet loading: variation of drug content. (II) Fatty acid chain length: caproic (C6), capric (C10), and myristic (C14) triglycerides. (III) Surfactant concentration: addition of poly (ethylene glycol) (PEG-6). (IV) Degree of glycerol esterification: variation of tri-, di-, and monoglyceride composition

Droplet loading with different amounts of drug molecules, the influence of the triglyceride fatty acid chain length, the addition of the surfactant poly(ethylene glycol) (PEG-6), and the addition of mono- and diglycerides were evaluated. Starting from random mixtures of all components, all systems spontaneously formed droplets within the equilibration phase

of 100 ns (Table S2). In their equilibrated state, all triglyceride droplets were essentially dry, with only a few residual water molecules remaining within the droplet interior. Snapshots of the equilibrated systems are depicted in Figures 3-6.

Table 2: Drug molecule immersion depth dmin and dcom in SEDDS droplets of varying excipient composition – dmin represents the minimal distance between the drug molecule and the droplet's first water solvation shell, whereas dcom represents the distance between the center of mass of the drug molecule and the closest water molecule of the first water solvation shell. All depicted data is taken from single 100ns MD simulation runs after initial 100ns of equilibration. dcom of the "droplet loading" system is averaged over its 10 CyA drug molecules contained within in the system. More detailed data is available in the supporting information (Table S2).

	standard		*dropl. loading*	*fatty acid chain length*				*surfactant concentration*				*glycerol esterification*			
	C10		10xCyA/C10	C6		C14		trig 4:1 surf		trig 3:2 surf		tri 1:1 mono		tri. 1:1 di	
time [ns]	dmin/dcom [Å]		dcom/s.d. [Å]	dmin/dcom [Å]		dmin/dcom [Å]		dmin/dcom [Å]		dmin/dcom [Å]		dmin/dcom [Å]		dmin/dcom [Å]	
0	12.0	22.1	14.1 ± 3.8	3.6	10.8	3.6	12.2	10.8	18.0	8.6	17.1	17.8	24.5	6.0	13.3
5	14.8	22.5	13.3 ± 3.3	3.8	10.2	3.6	10.3	13.9	20.5	8.8	15.8	17.2	25.7	3.5	10.6
10	13.2	19.7	13.7 ± 3.3	5.2	12.0	3.5	11.6	11.8	19.6	7.4	15.2	16.5	25.2	3.5	10.4
15	13.0	20.9	14.2 ± 2.7	4.1	13.4	3.8	13.1	11.2	17.7	8.1	16.7	17.9	28.0	4.8	12.5
20	12.5	23.0	14.1 ± 3.0	3.6	11.7	3.7	11.8	14.4	20.0	11.8	18.9	20.0	26.1	3.7	10.3
25	11.2	19.4	14.0 ± 3.8	3.6	11.2	3.5	11.3	10.5	18.0	7.5	15.0	16.7	26.7	3.8	9.6
30	12.1	18.1	15.1 ± 3.7	3.5	9.9	3.7	10.5	10.8	18.2	4.6	13.1	18.1	25.1	3.5	11.9
35	13.1	21.8	14.6 ± 2.8	3.5	10.2	3.7	10.8	14.3	21.0	8.1	15.5	14.6	21.5	4.1	10.3
40	11.9	19.3	14.5 ± 3.4	3.6	10.3	3.6	11.4	11.2	23.1	7.2	14.7	17.0	26.3	3.6	10.2
45	11.3	19.3	13.7 ± 3.5	3.5	10.2	3.5	12.2	12.1	17.6	9.8	17.3	17.4	24.9	3.6	9.7
50	12.4	21.1	13.9 ± 2.5	3.5	10.8	3.5	10.5	11.5	19.1	8.9	16.2	17.7	24.5	3.6	10.8
55	11.9	22.5	13.6 ± 2.1	3.7	11.3	3.6	11.8	10.4	16.6	7.3	15.2	16.9	23.7	3.5	9.7
60	11.5	19.0	15.3 ± 3.6	3.5	11.2	3.5	10.6	10.6	18.2	6.3	13.6	16.6	24.1	3.7	10.5
65	10.9	17.9	13.9 ± 2.9	3.7	11.0	3.6	11.8	11.7	21.9	7.5	14.7	18.2	25.0	3.5	9.6
70	10.5	19.3	14.6 ± 3.3	3.6	12.1	3.7	12.5	11.9	19.2	8.1	15.3	21.1	27.4	3.6	10.3
75	10.7	21.7	13.7 ± 2.3	3.7	12.3	4.0	12.7	11.3	20.4	7.0	13.7	19.0	26.1	3.6	10.7
80	5.0	13.5	14.0 ± 2.6	4.4	13.9	3.6	10.7	15.2	25.9	6.3	14.5	18.0	24.5	3.6	12.8
85	3.5	15.2	13.1 ± 3.7	3.6	11.6	3.7	13.1	11.4	21.8	6.6	14.9	18.3	27.6	3.5	11.9
90	6.2	14.1	13.1 ± 3.5	7.1	14.1	3.5	12.0	12.5	21.2	7.3	16.0	20.7	27.2	3.7	11.1
95	7.4	17.8	13.9 ± 3.2	3.7	10.6	3.6	10.4	11.8	21.7	6.9	14.0	17.5	23.6	3.5	11.9
100	7.3	15.6	13.0 ± 3.1	3.6	10.1	3.7	11.0	10.7	22.1	7.6	15.5	20.1	29.5	3.6	10.4

4.2.4.1 Standard System: Nanostructure, Drug Localization and Size Variation

Common solution-phase structures involving amphiphilic molecules, e.g., micelles, vesicle, or lipid bilayers, are known to display a bias of hydrophilic moieties toward water molecules at the aqueous interface. For the triglyceride-water phase structure of the type I

SEDDS standard system however, the appropriate molecular association bias was not pronounced (Table 3: TRIGPHIL-WAT: 0.81). Instead, lamellar-like substructures within the droplet interior were observed (Figure 3E), which is reflected in the high association bias of hydrophilic triglyceride moieties (TRIGPHIL-TRIGPHIL: 2.13). Considering that the bias between hydrophobic triglyceride moieties was merely slightly increased (TRIGPHOB-TRIGPHOB: 1.11) suggests that the tendency of hydrophilic triglyceride moieties to associate is indeed what drives the formation of SEDDS droplet nanostructures. Furthermore, drug molecule CyA displayed a strong association bias toward hydrophilic triglyceride moieties (DRUG-TRIGPHIL: 2.76) and to a lesser degree to the water solvation shell of the droplet (DRUG-WAT: 1.23). A potential explanation for this preference may be inferred when monitoring the drug molecule immersion depth (Table 2) after equilibration (t>100 ns). While the drug molecule could indeed come into direct contact with the droplet's aqueous interface (t=185 ns: dmin~3.5 Å), a complete submergence in the outer periphery of the droplet nanostructure (~10 Å) was the more common form of drug localization. The capacity of the drug molecule to diffuse into the droplet's interior during the simulation may be accredited to the lamellar-like droplet nanostructure and the subsequent potential of hydrophilic stabilization on two sides of amphiphilic circular peptide CyA. The effect of droplet size was investigated by increasing the number of capric triglyceride (C10) molecules from 200 to 500 (droplet diameters of 75 and 100 Å, respectively) and by decreasing the number of triglyceride molecules to 100 (droplet diameter of 60 Å). The presence of lamellar-like nanostructures (Figure S7), the molecular association biases (Figure S8), and the immersion depth of CyA in the droplet (Table S2) were independent of droplet size.

4.2.4.2 I: Droplet Loading

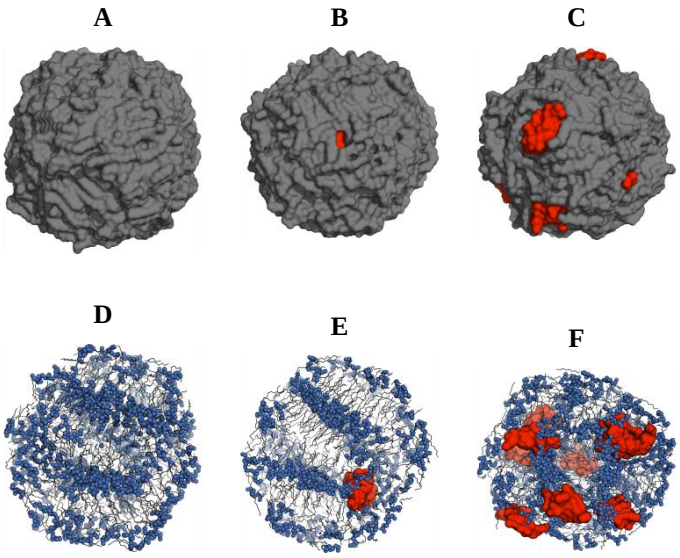


Figure 3: Depiction of drug localization and molecular association in SEDDS droplets (for larger image versions see Figure S2) – upper row: drug (red surface) immersion in droplet (grey surface), lower row: molecular association of hydrophilic triglyceride moieties (blue), left column: capric triglycerides (C10) without drug molecule, middle column: capric triglycerides (C10) with single drug molecule, right column: capric triglycerides (C10) with ten drug molecules

Table 3: Molecular association bias between hydrophilic ($_{PHIL}$) and hydrophobic ($_{PHOB}$) triglyceride (TRIG), water (WAT) and drug (DRUG) moieties – Variation of droplet loading (0x/1x/10x drug molecules), data taken from final 25ns of a 200ns simulation

atom groups	*0xCySA*	*1xCySA*	*10xCySA*
$TRIG_{PHOB}$ – *WAT*	0.28 ± 0.02	0.27 ± 0.01	0.27 ± 0.01
$TRIG_{PHIL}$ – *WAT*	0.80 ± 0.02	0.81 ± 0.03	0.75 ± 0.02
DRUG – *WAT*	n.a. ± n.a.	1.23 ± 0.12	2.68 ± 0.77
$TRIG_{PHOB}$ – $TRIG_{PHOB}$	1.10 ± 0.00	1.11 ± 0.00	1.11 ± 0.00
$TRIG_{PHOB}$ – $TRIG_{PHIL}$	0.96 ± 0.00	0.98 ± 0.00	0.97 ± 0.00
$TRIG_{PHIL}$ – $TRIG_{PHOB}$	0.52 ± 0.00	0.53 ± 0.00	0.53 ± 0.00
$TRIG_{PHIL}$ – $TRIG_{PHIL}$	2.13 ± 0.01	2.13 ± 0.01	2.14 ± 0.00
DRUG – $TRIG_{PHOB}$	n.a. ± n.a.	0.20 ± 0.04	0.30 ± 0.02
DRUG – $TRIG_{PHIL}$	n.a. ± n.a.	2.76 ± 0.12	2.42 ± 0.23

The removal of the drug molecule or the addition of further drug molecules did not significantly alter the droplet nanostructure, even when loading up to 10 drug molecules (10xCyA) into a single droplet with diameter <10 nm (Figure 3F). Merely the positive molecular association bias of drug molecules toward the aqueous environment ($TRIG_{DRUG}$-$TRIG_{WAT}$) increased from 1.23 in standard system 1xCyA to 2.68 in the 10xCyA system (Table 3), which is likely a consequence of droplet overloading. With immersion depths of d_{com}≈10 Å (Table 2), the CyA molecules were preferentially localized at the droplet surface and in contact with the aqueous environment (Figure 3B,C). Few exceptions were observed in the overloaded 10xCyA system, where the aqueous interface seems to be too small to accommodate all CyA molecules and a submerged localization of CyA seems to be favored over CyA aggregation.

4.2.4.3 II: Fatty Acid Chain Length Variation

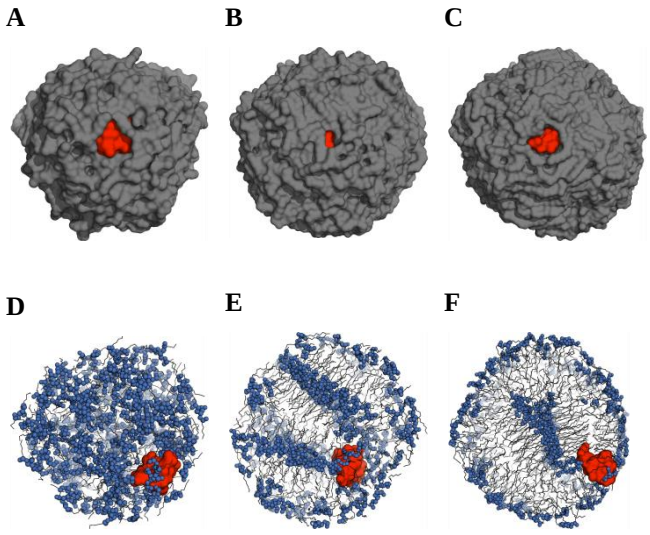


Figure 4: Depiction of drug localization and molecular association in SEDDS droplets (for larger image versions see Figure S3) – upper row: drug (red surface) immersion in droplet (grey surface), lower row: molecular association of hydrophilic triglyceride moieties (blue), left column: caproic triglycerides (C6), middle column: capric triglycerides (C10), right column: myristic triglycerides (C14)

Table 4: Molecular association bias between hydrophilic ($_{PHIL}$) and hydrophobic ($_{PHOB}$) triglyceride (TRIG), water (WAT) and drug (DRUG) moieties – Variation of fatty acid chain length (C6/C10/C14), data taken from final 25ns of a 200ns simulation

atom groups	*CAPROIC (C6)*	*CAPRIC (C10)*	*MYRISTIC (C14)*
$TRIG_{PHOB}$ – *WAT*	0.32 ± 0.02	0.27 ± 0.01	0.25 ± 0.00
$TRIG_{PHIL}$ – *WAT*	0.59 ± 0.02	0.81 ± 0.03	1.08 ± 0.02
DRUG – *WAT*	1.35 ± 0.51	1.23 ± 0.12	1.95 ± 0.71
$TRIG_{PHOB}$ – $TRIG_{PHOB}$	1.05 ± 0.01	1.11 ± 0.00	1.11 ± 0.00
$TRIG_{PHOB}$ – $TRIG_{PHIL}$	1.10 ± 0.01	0.98 ± 0.00	0.93 ± 0.01
$TRIG_{PHIL}$ – $TRIG_{PHOB}$	0.66 ± 0.00	0.53 ± 0.00	0.47 ± 0.00
$TRIG_{PHIL}$ – $TRIG_{PHIL}$	1.52 ± 0.01	2.13 ± 0.01	2.68 ± 0.01
DRUG – $TRIG_{PHOB}$	0.29 ± 0.13	0.20 ± 0.04	0.26 ± 0.09
DRUG – $TRIG_{PHIL}$	1.83 ± 0.22	2.76 ± 0.12	3.08 ± 0.39

When the fatty acid chain length is varied, a clear discrepancy in the molecular aggregation behavior of hydrophilic and hydrophobic moieties within the droplet interior was observed. For triglycerides of shorter (C6) fatty acids chain length, no segregation patterns were visually distinguishable (Figure 4D), like the lamellar-like patterns of the C10 systems (Figure 4E). Elongating the fatty acid chain lengths (C14) leads to the formation of a hydrophilic core and a hydrophilic droplet surface, reminiscent of a vesicular phase structure (Figure 4F). This prominent change in nanostructure is particularly apparent in the strong molecular association bias of hydrophilic triglyceride moieties ($TRIG_{PHIL}$-$TRIG_{PHIL}$), which increased from C6 (1.52), C10 (2.13), and C14 (2.68) (Table 4), but also in the increasing bias of hydrophilic triglyceride moieties toward the aqueous environment ($TRIG_{PHIL}$-WAT), which varied from C6: 0.59 to C10: 0.81 to C14: 1.08. The positive hydrophilic association bias is also reflected in the direct environment of the CyA molecule, which is increasingly hydrophilic for longer fatty acid side chains, in terms of both triglyceride association (DRUG-$TRIG_{PHIL}$) and the water environment (DRUG-WAT). Consequently, for the vesicle-like C14 system the drug was not immersed deeply in the droplet, but remained localized at the surface within a hydrogen bonding distance of d_{min}≈3.5 Å with respect to the solvent environment (Table 2). In the unstructured C6 system, the drug molecule was mostly found at the surface as well but occasionally was immersed up to d_{min}≈7.1 Å, which is well below the immersion depths of up to d_{min}≈14.8 Å for the lamellar-like C10 system, with its potential of stabilizing the hydrophilic moieties on two sides of circular peptide CyA.

4.2.4.4 III: Addition of Surfactant PEG-6

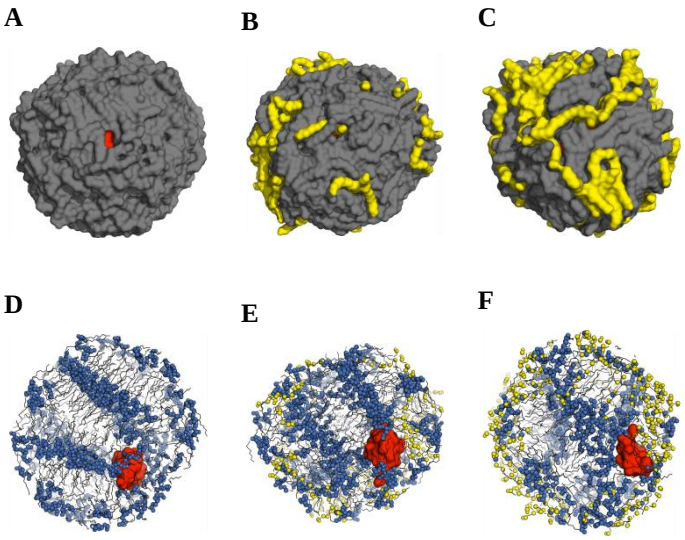


Figure 5: Depiction of drug localization and molecular association in SEDDS droplets (for larger image versions see Figure S4) – upper row: drug (red spheres) immersion in droplet (grey surface), lower row: drug surface accessibility of hydrophilic triglyceride moieties (blue) and surfactant (yellow), left column: capric triglycerides (C10) without surfactant, middle column: capric triglycerides (C10) in 4:1 ratio with polyethylene glycol (PEG-6), right column: capric triglycerides (C10) in 3:2 ratio with polyethylene glycol (PEG-6)

Table 5: Molecular association bias between hydrophilic ($_{PHIL}$) and hydrophobic ($_{PHOB}$) triglyceride (TRIG), water (WAT), surfactant (SURF) and drug (DRUG) moieties – Variation of surfactant PEG-6 content (TRIG only/TRIG 4:1 PEG/TRIG 3:2 PEG), data taken from final 25ns of a 200ns simulation

atom groups	*TRIG*	*TRIG 4 : 1 PEG*	*TRIG 3 : 2 PEG*
$TRIG_{PHOB}$ – *WAT*	0.27 ± 0.01	0.23 ± 0.01	0.17 ± 0.01
$TRIG_{PHIL}$ – *WAT*	0.81 ± 0.03	0.64 ± 0.02	0.51 ± 0.03
DRUG – *WAT*	1.23 ± 0.12	2.05 ± 0.84	0.47 ± 0.53
$TRIG_{PHOB}$ – $TRIG_{PHOB}$	1.11 ± 0.00	1.28 ± 0.01	1.54 ± 0.01
$TRIG_{PHOB}$ – $TRIG_{PHIL}$	0.98 ± 0.00	1.11 ± 0.01	1.34 ± 0.01
$TRIG_{PHOB}$ – *SURF*	n.a. ± n.a.	0.05 ± 0.01	0.06 ± 0.01
$TRIG_{PHIL}$ – $TRIG_{PHOB}$	0.53 ± 0.00	0.60 ± 0.00	0.71 ± 0.00
$TRIG_{PHIL}$ – $TRIG_{PHIL}$	2.13 ± 0.01	2.43 ± 0.01	2.94 ± 0.02
$TRIG_{PHIL}$ – *SURF*	n.a. ± n.a.	0.13 ± 0.02	0.13 ± 0.01
DRUG – $TRIG_{PHOB}$	0.20 ± 0.04	0.45 ± 0.07	0.49 ± 0.16
DRUG – $TRIG_{PHIL}$	2.76 ± 0.12	2.23 ± 0.30	3.22 ± 0.29
DRUG – *SURF*	n.a. ± n.a.	0.15 ± 0.24	0.39 ± 0.23

The standard system was modified to contain capric (C10) triglyceride molecules in 4:1 and 3:2 ratios with poly(ethylene glycol) (PEG) molecules. The addition of PEG-6 molecules only slightly enhanced the molecular association bias of hydrophilic and hydrophobic triglyceride moieties (Table 5). This can be accredited to a general segregation of triglyceride and PEG-6 molecules and can be deduced from the negative association bias of triglyceride molecules toward surfactant molecules ($TRIG_{PHIL}$-SURF and $TRIG_{PHOB}$-SURF). PEG-6 molecules predominantly covered the surface of the droplet, encasing the drug molecule partially (Figure 5B, TRIG 4:1 PEG-6) or fully (Figure 5C, TRIG 3:2 PEG-6). Hence, the CyA molecule was immersed more deeply (Table 2) within the droplet than in systems without PEG-6, depending on the thickness of the encapsulating shell. Judging from the negative drug molecule association bias toward surfactants and the positive drug association bias toward hydrophilic triglyceride moieties, CyA molecules were preferentially localized at the triglyceride-surfactant interface, partially or fully shielded from the aqueous environment. The positive association bias toward polar triglyceride molecules (DRUG-$TRIG_{PHIL}$) implies that CyA solubilization might be facilitated by triglyceride molecules rather than by surfactant molecules.

4.2.4.5 IV: Addition of Mono- and Diglycerides

Mono- and diglyceride molecules were added to the standard system in an equimolar ratio. Due to the additional hydrophilic hydroxyl groups, the polar glycerol moieties of both mono- and diglycerides were predominantly located at the droplet surface, in direct contact with the aqueous surrounding. While diglycerides were also found within the droplet in sizable amounts (Figure 6E), monoglycerides were almost exclusively localized at the droplet surface (Figure 6F). Consequently, triglyceride and mono/diglyceride molecules were segregated similarly to the surfactant systems (Table 6). However, two disparate drug-stabilization schemes could be distinguished. For the tri-/diglyceride system, the CyA molecule was accessible to the water solvation shell at the droplet surface (Figure 6B), similarly to the previously discussed systems, while in the tri/monoglyceride system the CyA molecule was stabilized at the droplet core, without any exposure to the aqueous environment (Figure 6C).

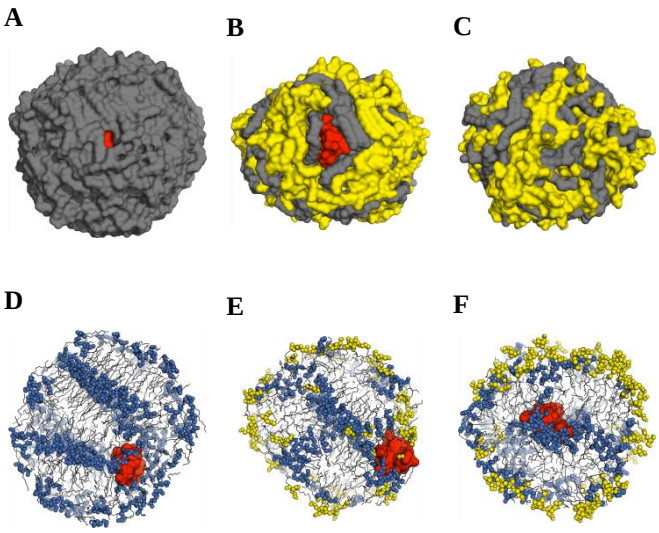


Figure 6: Depiction of drug localization and molecular association in SEDDS droplets (for larger image versions see Figure S5) – upper row: drug (red surface) immersion in droplet (grey surface), lower row: drug surface accessibility of hydrophilic triglyceride (blue), diglyceride (yellow), monoglyceride (yellow) moieties, left column: capric triglycerides (C10), middle column: capric tri- and diglycerides in equimolar ratio, right column: capric tri- and monoglycerides in equimolar ratio

Table 6: Molecular association bias between hydrophilic (PHIL) and hydrophobic (PHOB) triglyceride (TRIG), water (WAT) and drug (DRUG) moieties – variation of degree of glycerol esterification (TRI only/TRI 1:1 DI/TRI 1:1 MONO), data taken from final 25ns of a 200ns simulation

atom groups	*TRI*	*TRI 1 : 1 DI*	*TRI 1 : 1 MONO*
$TRIG_{PHOB}$ – *WAT*	0.27 ± 0.01	0.16 ± 0.00	0.15 ± 0.01
$TRIG_{PHIL}$ – *WAT*	0.81 ± 0.03	0.48 ± 0.03	0.45 ± 0.02
DRUG – *WAT*	1.23 ± 0.12	2.10 ± 0.25	0.00 ± 0.00
$TRIG_{PHOB}$ – $TRIG_{PHOB}$	1.11 ± 0.00	2.00 ± 0.01	1.75 ± 0.01
$TRIG_{PHOB}$ – $TRIG_{PHIL}$	0.98 ± 0.00	1.73 ± 0.02	1.51 ± 0.01
$TRIG_{PHOB}$ – $nGLYC_{PHOB}$	n.a. ± n.a.	0.01 ± 0.00	0.03 ± 0.00
$TRIG_{PHOB}$ – $nGLYC_{PHIL}$	n.a. ± n.a.	0.06 ± 0.01	0.06 ± 0.01
$TRIG_{PHIL}$ – $TRIG_{PHOB}$	0.53 ± 0.00	0.91 ± 0.01	0.81 ± 0.01
$TRIG_{PHIL}$ – $TRIG_{PHIL}$	2.13 ± 0.01	3.71 ± 0.02	3.26 ± 0.02
$TRIG_{PHIL}$ – $nGLYC_{PHOB}$	n.a. ± n.a.	0.04 ± 0.01	0.06 ± 0.02
$TRIG_{PHIL}$ – $nGLYC_{PHIL}$	n.a. ± n.a.	0.28 ± 0.00	0.21 ± 0.02
DRUG – $TRIG_{PHOB}$	0.20 ± 0.04	0.26 ± 0.08	0.42 ± 0.14
DRUG – $TRIG_{PHIL}$	2.76 ± 0.12	3.11 ± 0.25	4.76 ± 0.30
DRUG – $nGLYC_{PHOB}$	n.a. ± n.a.	0.36 ± 0.09	0.00 ± 0.00
DRUG – $nGLYC_{PHIL}$	n.a. ± n.a.	0.77 ± 0.22	0.00 ± 0.00

In the latter case, the droplet nanostructure consisted of an outer shell of monoglyceride molecules with its hydrophilic headgroups oriented toward the aqueous phase and a central hydrophilic core of polar ester groups of triglyceride molecules. While monoglyceride molecules formed a micellar-phase structure, triglyceride molecules formed an inverted micellar-phase structure. The CyA molecule was hereby deeply immersed within the droplet at d_{min}≈25-30 Å (Table 2), where it was stabilized via polar interactions at the surface of the hydrophilic triglyceride core.

4.2.5 Discussion

4.2.5.1 Implications of Droplet Loading

The potential of MD simulations for determining the nanostructure and the localization of drug molecules within SEDDS droplets was evaluated by modeling systems of varied lipid excipients that are common to SEDDS design (Figure 2), in conjunction with the immunosuppressant CyA drug molecule, which is applied in a range of commercial applications.[44-47] The solubilization capacity of drug molecules in lipidic formulations is of relevance to SEDDS not only to avoid drug precipitation but also to define the minimal volume per unit dose. We have therefore analyzed the effect of increasing drug content in SEDDS droplets. No significant alterations of the molecular association behavior between triglyceride moieties were observed, even in the case of overloading, which implies that the influence of the drug-triglyceride interaction on the nanostructure of the droplet may indeed be negligible. We did, however, observe a significant increase in the molecular association bias of drug molecules toward the aqueous environment, which points toward increased drug precipitation. While the solubilization of drug molecules is of primary concern for improving bioavailability,[20] controlling drug release profiles is also of significance,[48-50] and indeed, recent studies on supersaturated SEDDS have shown that high drug concentration in droplets can be beneficial for bioavailability.[51, 52] In instances where poorly water-soluble drug molecules are stabilized too strongly within the hydrophobic droplet nanostructure, increasing the drug load is possibly a remedy. In this regard, the molecular association bias of drug molecules toward the aqueous environment may be an interesting parameter not only to quantify the propensity of drug molecules to

precipitate but also potentially tune release rates of drug molecules from the nanostructure. The fact that no significant aggregation of CyA was observed during the simulation signifies that the preferred equilibrium environment of amphiphilic peptides is indeed the triglyceride droplet nanostructure, which further encourages the possibility of engineering the nanostructure of droplets toward accommodating drug-specific requirements in SEDDS formulation.

4.2.5.2 Implications of Fatty Acid Chain Length Variation

The simulated systems of varied fatty acid chain lengths aggregated to distinctly different nanostructures, ranging from seemingly unstructured (C6) to lamellar-like (C10) to vesicle-like (C14) molecular arrangements. The unifying feature of the observed aggregation patterns was the positive molecular association bias of hydrophilic and hydrophobic triglyceride moieties, which is substantiated by a previous study of Sum et al., where similar bilayer-like substructures within the interior of larger triglyceride aggregates were observed.[53] Current SEDDS formulations focus primarily on the droplet microstructure. While the droplet size distribution can be engineered by the addition of surfactants,[3] the nanostructure of the droplets and the location of the drug molecules are currently not addressed by SEDDS design. Controlling the nanostructure and thus the localization of drug molecules in droplets would greatly contribute to the efficacy of SEDDS formulations. It seems feasible when considering that the preferred equilibrium environment of any molecule is defined by its chemical structure[13] and that drug localization and solubility are clearly correlated.[20] For a predetermined droplet size, triglyceride excipients offer the highest potential for nanostructure engineering, since the choice of triglycerides is independent of the droplet size. Variability in drug precipitation for formulations of different triglyceride chain lengths has been reported,[54, 55] which could be caused by the discrepancies in nanostructures that enable different forms of drug solubilization. For the vesicle-like nanostructure (C14), we observed the formation of a hydrophilic core of polar triglyceride moieties, a highly hydrophobic interphase and a hydrophilic surface environment. Thus, particularly polar and amphiphilic drug molecules could also be accommodated in a vesicle-like nanostructure in the core region of the droplet as an alternative to the commonly observed stabilization at the droplet surface in

the case of CyA. The potential for core stabilization of CyA was in fact observed for the mono/triglyceride systems, which featured a vesicle-like nanostructure similar to that of the C14 system (see below). Surely, the preferred localization will depend on and vary with the unique chemical structure of a drug molecule. For shorter chain systems C6 and C10, no discernible nanostructure could be identified that would suggest the core stabilization of CyA molecules, which were exclusively found at the interface between the droplet and the aqueous environment. While this localization might be considered to be beneficial in some instances,[20] it would be detrimental to peptides such as CyA, which are susceptible to digestion. Furthermore, in C10 systems a partial submergence of CyA within the droplet was observed, which could be attributed to the lamellar-like aggregation type that allows for hydrophilic contacts on either side of the circular peptide. If systematically reproducible, this would constitute a compromise between accessibility and protection from the GI tract and could therefore be of relevance to SEDDS design considerations. From these findings we gather that for SEDDS applications it should be feasible to affect the droplet nanostructure through variations of triglyceride fatty acid chain lengths and thus improve the solubilization of drug molecules within droplets. However, the influence of surfactant molecules on the nanostructure would also need to be accounted for.

4.2.5.3 Implication of Adding Surfactants

Surfactants are amphiphilic molecules that are commonly applied in the stabilization of emulsions/dispersions by acting as mediators between hydrophobic and hydrophilic moieties at oil-water interfaces. The choice of surfactant in SEDDS formulations varies greatly, [3] and due to the capacity to engineer stable dispersions of defined droplet size, surfactants have to be regarded as a necessity for any rational SEDDS design approach. While the impact of surfactants on the microscopic dispersion characteristics of microemulsions is well established in experiments and literature, the effect of surfactants on the nanostructure within droplets is not as straightforward. The PEG-6 surfactant chosen in this study consists of chains of alternating strongly polar and weakly polar moieties. The observed aggregation of PEG-6 at the droplet surface is therefore not surprising. Next to a slight pronunciation of the existing molecular association bias in the triglyceride phase of the droplet, the formation of prominent surfactant shells was

observed, which partially covered CyA molecules to a degree that varied with surfactant concentration. Although CyA molecules were localized at the interphase between triglyceride and PEG-6 in all instances, the molecular association bias suggests that drug stabilization is mainly facilitated by the triglyceride phase. Effectively, PEG-6 could therefore be utilized not only to define the dispersion droplet size but also to facilitate a protective layer, encasing a preserved triglyceride phase, shielding drug molecules from the GI tract environment, and interfering with the rate at which triglyceride molecules are digested. The observed encapsulation would, however, be specific to PEG or surfactants of a similar chemical structure. Contrary to PEG, it is to be expected that polar surfactants which feature a marked hydrophobic moiety should interact more strongly with the hydrophobic nanostructure of the triglyceride phase. In this study this was modeled in mono- and diglycerides.

4.2.5.4 Implications of Adding Mono- and Diglycerides

Mono- and diglycerides differ from surfactants such as PEG in the long aliphatic fatty acid side chains and the polar glycerol moiety, which gives them an overall molecular polarity. Therefore, they act as surfactants for condensed-phase triglyceride aggregates; consequently, the hydrophilic moieties of monoglyceride and to a lesser degree diglyceride were abundantly localized at the droplet surface. In simulation, the aliphatic side chains protruded into the hydrophobic droplet interior and directly interfered with the droplet nanostructure, which enhanced the existing molecular association bias between triglyceride moieties in comparison to that of the capric (C10) type I standard system. While lamellar-like nanostructures were still present in the diglyceride system, in the monoglyceride system a vesicle-like nanostructure with a distinguishable core of hydrophilic triglyceride moieties was formed, where drug molecule CyA was solubilized. This form of drug localization should be highly desirable and feasible for many amphiphilic drugs, particularly for strongly polar molecules; the hydrophilic moiety could be accommodated at the core, and the hydrophobic moiety, in the hydrophobic droplet interphase. However, the same considerations apply to drug localization at the aqueous interface as well, which was indeed observed in simulations of the myristic (C14) type I SEDDS systems that featured a similar hydrophilic droplet core (see above). Predicting drug localization with

statistically significant MD studies of single systems with defined excipients could be a feasible approach to predetermining the preferred locus of drug solubilization in the formulation design phase.

4.2.5.5 Molecular Modeling Approach for SEDDS Design

This study takes a more general approach of analyzing different SEDDS excipient variations and assessing the potential of MD on SEDDS design. One should bear in mind that the preferred locus of drug solubilization is never a binary problem but always a matter of probability and that the MD method accommodates this by converging thermodynamic systems to local potential energy minima. It should be computationally feasible to extend the approach applied in this study to single systems with defined excipients and thereby generate statistics that significantly represent the preferred drug localization in a specific SEDDS formulation. Due to the computational demand of this study, we excluded the influence of bile salts and digestion-induced changes to the droplet composition[3]. The force field combination applied in this work constitutes a well-established parameter set, which was specifically refined for modeling comparable condensed-phase lipidic systems and should therefore adequately reproduce the interactions that govern the molecular association of SEDDS droplets. The general united atom GROMOS 53a6 force field[37] used as a basis for this study has been parametrized to reproduce thermodynamic properties of liquids, liquid mixtures and polar amino acid analogues in particular. The $53a6_{OXY}$ parameter set[39] represents a specific redesign for improved solvation properties for a range of oxygen moieties, including esters, while the 53A6_OE parameter set has been optimized for poly(ethylene glycol)[40]. The Berger lipids model[38] has been optimized to reproduce lipid properties accurately.

4.2.5.6 Influence of the Aqueous Interface and the Droplet Size

The association bias between hydrophilic triglyceride moieties ($TRIG_{PHIL}$-$TRIG_{PHIL}$) was consistently higher than the association bias of polar triglyceride moieties with water ($TRIG_{PHIL}$-WAT), which implies that the influence of the triglyceride-water interface plays a subordinate role in droplet nanostructure formation. This can be directly observed in the

structural features of the droplets (Figure 3-6). Instead of associating in a manner that minimizes the water-accessible hydrophobic surface area of triglycerides, lamellar-like patterns of the polar triglyceride moieties formed and the hydrophobic fatty acid chains were significantly exposed to the bulk water. While these observations might appear to be counterintuitive, they are in agreement with the notion that in aqueous solution the dominant driving force for molecular structuring is the self-association of water into hydrogen bonding networks.[43] In this regard, the aqueous surroundings merely define the spherical, minimal volume of the droplet in order to maximize hydrogen bonding within the bulk water. Likewise, the nanostructure of the droplet is driven by the self-association of polar triglyceride moieties, indicated by their high association bias. Therefore, it is expected to be a result of excipient variations and droplet size, e.g., upon formation of lamellar structures. As a consequence, larger droplets are expected to feature multiple layers as observed for the model systems of this study, which is in agreement with prior reports on bulky triglyceride aggregates.[53] Generalized correlations between droplet size and drug molecule localization are, however, difficult to deduce on the basis of the presented data, and in light of the variable chemical composition of drug molecules should be evaluated on a case-by-case basis. While variations of the number of triglycerides between 100 and 500 molecules revealed no significant changes in the droplet nanostructure or in the localization of the drug molecules, these observations should not be extrapolated to significantly larger aggregates. Modeling larger systems in atomic detail is computationally demanding. However, since the droplets were essentially dry, in future studies an implicit treatment of water would allow for increasing the droplet size by a factor of 10 and thus studying aggregates on the micrometer scale.

4.2.6 Conclusion

The bioavailability and controlled release rates of orally administered drugs are significantly affected by how drugs are solubilized in lipidic drug-delivery systems. Being able to rationally predetermine the optimal environment and the localization of drug molecules in the nanostructure of SEDDS droplets would allow for a more holistic strategy in formulation design, which is currently focused on engineering the droplet size of microemulsions. In this study, the dependence of the droplet nanostructure and drug

localization on the SEDDS excipient composition was systematically demonstrated by molecular dynamics simulations. Hereby, the molecular association bias of hydrophilic and hydrophobic excipient moieties was introduced as a simple parameter to distinguish specific droplet nanostructures. Fatty acid chain length variations were shown to induce clearly distinguishable aggregation patterns of hydrophilic triglyceride moieties from seemingly unstructured (C6) to lamellar-like (C10) to vesicle-like (C14). Mono- and diglyceride molecules were shown to interfere with the droplet nanostructure and to feature the surfactant-specific aggregation of polar moieties at the lipid-water interface. PEG-6 surfactants were found to accumulate at the droplet surface, without perturbing the encapsulated triglyceride droplet. Increasing the amount of drug molecules loaded into a droplet did not display any alteration of its nanostructure; however, increased exposure of the drug molecules to the aqueous environment was observed. The findings encourage the transfer of the approach taken in this study to more specific SEDDS applications, where it could be more rigorously developed and refined.

4.2.7 Acknowledgements

We thank the German Research Foundation (DFG) for the financial support of this project within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart. We thank the high performance computing center Stuttgart (HLRS) for their support and for supplying the computational resources.

4.2.8 Supporting Information

Molecular association bias during simulation for selected atom groups of the standard system consisting of drug molecule CyA with 200 capric (C10) triglyceride molecules in aqueous solution. Detailed data on the cluster formation of SEDDS droplets and CyA drug immersion depths. Larger versions of Figures 3-6, depicting the droplet nanostructures of the systems covered in this study. Material is available free of charge via the Internet at http://pubs.acs.org/.

Table S1: Molecular association bias of the standard system (1x CyA / 200x capric (C10) triglyceride) in aqueous solution - During 200ns of NPT simulation (310K/1bar) systems equilibrated into a single SEDDS droplet of diameter ~7.5nm, starting from a random molecular arrangement. Fluctuations in drug molecule values reflect the variability of CyA immersion in the standard system (Table S2). Data for t<100ns may contain artifacts due to periodic boundary conditions, which are not accounted for in the calculation.

	$TRIG_{PHIL}$			$TRIG_{PHOB}$			DRUG		
[ns]	WAT	$TRIG_{PHOB}$	$TRIG_{PHIL}$	WAT	$TRIG_{PHOB}$	$TRIG_{PHIL}$	WAT	$TRIG_{PHOB}$	$TRIG_{PHIL}$
0	0.67	0.60	2.47	0.47	1.38	1.07	3.43	0.19	1.56
5	0.75	0.56	2.12	0.32	1.15	1.01	3.68	0.21	1.59
10	0.77	0.55	2.12	0.29	1.15	0.99	3.33	0.26	1.66
15	0.77	0.55	2.11	0.29	1.14	1.01	3.40	0.52	1.30
20	0.86	0.53	2.10	0.27	1.12	0.98	2.81	0.33	1.92
25	0.84	0.52	2.13	0.26	1.13	0.97	2.78	0.05	2.57
30	0.80	0.53	2.11	0.28	1.12	0.98	2.36	0.34	2.07
35	0.81	0.53	2.11	0.27	1.12	0.98	1.93	0.52	1.81
40	0.80	0.53	2.12	0.27	1.12	0.98	2.88	0.41	1.75
45	0.78	0.53	2.12	0.28	1.11	0.98	3.34	0.28	1.87
50	0.84	0.53	2.12	0.29	1.12	0.97	2.04	0.22	2.42
55	0.77	0.53	2.14	0.32	1.12	0.97	1.93	0.24	2.44
60	0.83	0.53	2.12	0.29	1.11	0.98	1.75	0.44	2.07
65	0.77	0.54	2.12	0.29	1.11	0.98	1.20	0.24	2.69
70	0.77	0.53	2.13	0.29	1.11	0.97	1.52	0.31	2.43
75	0.77	0.53	2.12	0.27	1.11	0.98	1.86	0.25	2.48
80	0.78	0.53	2.13	0.28	1.11	0.97	1.92	0.25	2.44
85	0.79	0.52	2.15	0.29	1.11	0.96	0.84	0.33	2.60
90	0.83	0.52	2.12	0.26	1.11	0.97	1.38	0.19	2.74
95	0.81	0.53	2.12	0.28	1.11	0.98	0.43	0.31	2.75
100	0.79	0.52	2.14	0.29	1.11	0.96	0.43	0.31	2.75
105	0.81	0.52	2.12	0.27	1.10	0.98	2.59	0.10	2.62
110	0.80	0.53	2.13	0.28	1.11	0.97	2.30	0.17	2.52
115	0.81	0.52	2.13	0.30	1.11	0.96	1.96	0.15	2.66
120	0.80	0.53	2.13	0.27	1.11	0.97	2.62	0.17	2.42
125	0.80	0.53	2.12	0.27	1.10	0.97	2.07	0.20	2.53
130	0.83	0.52	2.13	0.28	1.11	0.97	1.67	0.27	2.49
135	0.84	0.52	2.13	0.28	1.11	0.97	0.88	0.06	3.19
140	0.80	0.53	2.12	0.28	1.11	0.97	0.98	0.13	3.00
145	0.81	0.52	2.14	0.28	1.11	0.97	1.67	0.10	2.87
150	0.78	0.52	2.14	0.29	1.10	0.97	1.27	0.30	2.53
155	0.79	0.52	2.13	0.28	1.11	0.97	0.65	0.17	3.01
160	0.85	0.53	2.11	0.25	1.11	0.99	1.78	0.24	2.52
165	0.82	0.52	2.13	0.28	1.11	0.97	2.15	0.23	2.40
170	0.80	0.53	2.12	0.25	1.11	0.98	1.35	0.24	2.64
175	0.86	0.52	2.12	0.26	1.12	0.98	1.18	0.21	2.75
180	0.84	0.52	2.12	0.27	1.11	0.98	1.09	0.22	2.76
185	0.80	0.53	2.11	0.27	1.11	0.98	1.40	0.26	2.59
190	0.77	0.53	2.14	0.29	1.11	0.98	1.07	0.15	2.93
195	0.79	0.53	2.13	0.27	1.11	0.97	1.28	0.14	2.89
200	0.79	0.53	2.13	0.28	1.11	0.97	1.35	0.25	2.63

Table S2: Clustering of SEDSS systems (clust.) and immersion depth of drug molecule CySA over time. The 500xC10 system required a further 100ns to cluster into a single droplet; the first 100ns and the last 100ns are shown.

	system size variation									*fatty acid chain length*						*surfactant concentration*						*glycerol esterification*						*droplet loading*																	
	100xC10			200xC10			500xC10			C6			C14			trigl.4:1 surf.			trigl.3:2 surf.			tri 1:1 mono			tri. 1:1 di			0xCySA		10xCySA															
time [ns]	no. clust.	d_{min} [Å]	d_{com}	no. clust.	d_{min} [Å]	d_{com}	no. clust.	d_{min} [Å]	d_{com}	no. clust.	d_{min} [Å]	d_{com}	no. clust.	d_{min} [Å]	d_{com}	no. clust.	d_{min} [Å]	d_{com}	no. clust.	d_{min} [Å]	d_{com}	no. clust.	d_{min} [Å]	d_{com}	no. clust.	d_{min} [Å]	d_{com}	no. clust.		no. clust.	d_{com} [Å]													
0	36	3.5	9.8	79	3.5	12.4	169	3.5	9.6	79	3.5	8.6	71	3.5	12.4	74	3.5	9.8	66	3.5	9.1	55	3.5	11.5	58	3.5	10.3	79	-	101	11.5	10.8	11.5	11.1	13.3	9.0	9.5	11.3	12.5	11.0
5	6	3.8	12.8	9	3.6	11.0	24	3.5	9.8	9	3.6	11.9	8	3.5	8.9	7	3.5	10.7	11	3.5	11.6	12	3.6	11.7	8	3.5	10.7	10	-	9	8.9	10.2	9.8	9.4	11.2	12.3	12.6	12.0	9.2	12.0
10	4	3.5	11.9	4	3.7	10.6	13	3.6	11.6	4	3.9	11.4	7	3.5	10.4	6	3.5	9.3	7	3.6	11.4	10	7.6	16.0	5	3.5	10.4	7	-	8	13.0	10.7	14.3	12.6	9.5	11.1	9.9	12.0	10.6	11.1
15	3	4.5	13.9	3	3.6	10.3	11	3.6	10.8	3	3.6	11.2	5	3.5	9.3	5	3.5	10.7	7	3.5	10.0	5	8.0	16.3	4	3.7	10.7	5	-	7	10.6	11.6	11.0	11.2	10.9	9.9	12.9	12.7	13.3	10.3
20	2	3.7	12.2	1	4.1	13.4	11	3.6	12.8	1	3.6	11.6	5	3.5	11.1	3	3.8	11.0	5	3.6	12.9	3	9.9	17.7	2	3.5	12.2	3	-	5	11.2	16.5	13.4	10.4	9.9	10.2	12.5	13.5	10.2	9.6
25	1	3.8	11.4	1	3.9	13.8	9	3.5	10.4	1	3.5	10.6	4	3.7	11.9	3	3.5	12.5	4	3.6	14.0	2	7.8	16.5	2	5.4	12.9	3	-	5	10.1	14.1	14.5	12.2	11.8	10.9	10.4	13.9	8.8	10.8
30	1	5.7	12.3	1	9.3	20.4	6	3.5	13.2	1	3.5	12.1	4	3.5	12.2	3	6.1	13.8	3	3.6	10.9	2	14.9	21.0	2	3.6	13.2	3	-	4	11.8	15.7	10.1	11.1	12.1	11.3	10.7	12.8	10.4	11.1
35	1	7.7	13.9	1	11.2	21.4	5	3.8	13.0	1	3.7	12.5	4	3.7	10.5	2	3.6	12.3	3	3.7	11.9	1	15.9	23.6	1	3.6	10.6	3	-	3	9.0	18.8	13.1	11.9	11.2	12.8	10.6	14.4	12.1	10.7
40	1	5.9	12.9	1	10.9	19.4	5	3.5	13.4	1	4.6	12.4	2	3.6	10.6	2	3.5	13.9	3	6.3	14.7	1	16.9	24.8	1	3.8	10.3	3	-	3	13.3	14.9	10.9	10.5	15.6	12.1	9.7	16.2	11.3	10.3
45	1	6.7	14.6	1	7.2	17.8	5	3.7	14.0	1	3.6	11.7	2	3.6	12.4	1	8.2	16.3	2	6.7	14.9	1	15.6	25.4	1	11.0	18.2	2	-	3	9.9	18.0	12.0	11.7	12.9	11.5	9.6	15.6	13.3	9.8
50	1	8.0	15.0	1	8.3	17.3	4	6.2	12.7	1	4.6	11.2	2	3.5	11.0	1	11.3	18.0	2	4.3	12.8	1	14.8	23.0	1	8.3	16.6	2	-	3	12.5	16.6	9.4	11.6	16.5	11.2	8.8	13.9	14.3	11.0
55	1	6.3	12.8	1	6.0	13.5	3	3.7	10.8	1	4.9	12.1	2	3.6	10.8	1	8.2	14.6	1	3.6	13.5	1	13.2	21.7	1	7.4	14.4	2	-	2	12.4	17.6	11.9	11.5	16.0	10.9	11.0	15.9	10.7	10.0
60	1	7.4	14.1	1	10.1	17.9	3	3.7	12.6	1	6.5	16.2	1	3.7	11.7	1	9.1	17.4	1	4.7	12.2	1	15.5	24.3	1	6.2	14.4	2	-	2	10.4	17.3	16.3	12.4	14.5	11.9	8.5	15.7	9.9	10.8
65	1	7.2	17.3	1	11.1	18.6	3	3.8	14.6	1	5.4	13.1	1	3.5	11.3	1	11.2	20.5	1	5.3	13.3	1	15.8	21.9	1	3.5	11.9	2	-	2	11.4	18.9	15.2	11.4	17.1	11.2	11.7	16.8	11.7	9.9
70	1	8.2	18.0	1	11.8	19.1	3	3.9	13.2	1	5.5	14.1	1	3.6	11.4	1	8.7	15.9	1	3.9	13.3	1	20.0	26.4	1	3.5	11.0	2	-	2	12.4	20.6	16.4	11.5	15.1	11.0	11.5	18.7	10.1	9.2
75	1	11.0	19.9	1	10.5	18.6	3	3.5	12.7	1	3.5	11.0	1	3.5	10.9	1	3.8	15.1	1	8.8	15.7	1	18.1	25.8	1	3.5	11.6	2	-	2	8.9	19.2	12.9	12.0	15.2	13.7	12.5	15.8	9.6	10.1
80	1	11.0	20.1	1	13.6	22.7	3	3.5	11.9	1	4.9	13.9	1	3.5	11.3	1	6.0	16.1	1	6.7	14.1	1	17.9	25.1	1	3.5	11.2	2	-	2	12.4	18.4	16.4	10.5	17.1	10.5	10.5	16.6	9.1	9.0
85	1	7.1	17.1	1	13.3	19.4	3	3.5	13.6	1	5.0	13.3	1	3.5	12.1	1	10.4	18.8	1	7.6	15.8	1	16.5	26.0	1	3.7	11.1	1	-	2	11.4	17.6	14.2	11.9	14.6	9.0	11.1	17.7	10.7	9.4
90	1	7.9	17.9	1	13.9	18.9	2	3.6	12.4	1	3.8	12.4	1	3.5	12.2	1	8.2	19.0	1	8.5	18.3	1	15.3	24.0	1	3.7	13.1	1	-	2	10.8	20.6	15.1	11.6	15.4	9.7	8.8	17.0	11.2	10.1
95	1	9.9	19.6	1	15.9	23.0	2	6.7	16.9	1	7.0	13.0	1	3.7	11.7	1	8.6	15.8	1	11.2	19.2	1	18.8	26.4	1	3.6	10.6	1	-	1	10.9	19.2	16.0	12.0	16.5	11.2	12.1	19.7	10.6	11.4
100	1	7.9	18.1	1	12.0	22.1	2	5.2	12.0	1	3.6	10.8	1	3.6	12.2	1	10.8	18.0	1	8.6	17.1	1	17.8	24.5	1	6.0	13.3	1	-	1	11.2	15.8	19.6	9.6	17.3	11.0	17.7	17.6	13.5	8.2
105	1	9.1	19.1	1	14.8	22.5	2	3.5	11.4	1	3.8	10.2	1	3.6	10.3	1	13.9	20.5	1	8.8	15.8	1	17.2	25.7	1	3.5	10.6	1	-	1	10.0	16.4	18.6	11.3	15.7	9.4	15.2	16.3	10.4	9.6
110	1	10.3	20.2	1	13.2	19.7	2	3.7	11.1	1	5.2	12.0	1	3.5	11.6	1	11.8	19.6	1	7.4	15.2	1	16.5	25.2	1	3.5	10.4	1	-	1	10.8	16.5	19.3	11.1	16.1	9.2	16.3	15.9	12.2	9.6
115	1	10.7	17.7	1	13.0	20.9	2	3.5	11.9	1	4.1	13.4	1	3.8	13.1	1	11.2	17.7	1	8.1	16.7	1	17.9	28.0	1	4.8	12.5	1	-	1	14.0	17.6	16.7	12.1	15.5	9.5	18.6	13.7	11.3	13.4
120	1	10.3	19.9	1	12.5	23.0	2	3.7	11.8	1	3.6	11.7	1	3.7	11.8	1	14.4	20.0	1	11.8	18.9	1	20.0	26.1	1	3.7	10.3	1	-	1	14.8	18.6	16.6	12.5	14.7	9.5	16.0	17.2	10.8	10.0
125	1	8.0	18.2	1	11.2	19.4	2	3.6	12.6	1	3.6	11.2	1	3.5	11.3	1	10.5	18.0	1	7.5	15.0	1	16.7	26.7	1	3.8	9.6	1	-	1	15.2	19.7	18.3	11.7	13.8	8.4	17.1	16.6	10.7	8.8
130	1	9.3	20.1	1	12.1	18.1	2	4.7	12.2	1	3.5	9.9	1	3.7	10.5	1	10.8	18.2	1	4.6	13.1	1	18.1	25.1	1	3.5	11.9	1	-	1	14.6	21.4	18.1	12.1	17.3	11.7	19.0	15.9	11.2	9.8
135	1	9.2	20.2	1	13.1	21.8	2	4.0	11.9	1	3.5	10.2	1	3.7	10.8	1	14.3	21.0	1	8.1	15.5	1	14.6	21.5	1	4.1	10.3	1	-	1	13.5	15.6	18.4	13.6	17.8	12.2	16.5	17.4	10.5	10.8
140	1	8.6	18.4	1	11.9	19.3	2	6.1	13.0	1	3.6	10.3	1	3.6	11.4	1	11.2	23.1	1	7.2	14.7	1	17.0	26.3	1	3.6	10.2	1	-	1	14.3	15.1	19.9	12.7	17.2	10.7	16.3	18.7	9.4	11.0
145	1	8.5	19.1	1	11.3	19.3	1	6.9	14.6	1	3.5	10.2	1	3.5	12.2	1	12.1	17.6	1	9.8	17.3	1	17.4	24.9	1	3.6	9.7	1	-	1	12.6	16.8	18.5	10.7	14.1	9.3	15.8	19.1	10.0	9.9
150	1	11.5	20.6	1	12.4	21.1	1	7.7	14.1	1	3.5	10.8	1	3.5	10.5	1	11.5	19.1	1	8.9	16.2	1	17.7	24.5	1	3.6	10.8	1	-	1	11.1	14.0	18.9	13.1	15.0	11.6	16.3	16.1	11.5	11.3
155	1	12.0	19.7	1	11.9	22.5	1	6.6	14.1	1	3.7	11.3	1	3.6	11.8	1	10.4	16.6	1	7.3	15.2	1	16.9	23.7	1	3.5	9.7	1	-	1	13.7	13.6	16.8	11.9	16.1	10.1	14.3	15.4	13.2	11.0
160	1	10.4	16.5	1	11.5	19.0	1	6.9	14.5	1	3.5	11.2	1	3.5	10.6	1	10.6	18.2	1	6.3	13.6	1	16.6	24.1	1	3.7	10.5	1	-	1	15.5	17.9	21.0	13.0	17.9	9.6	17.9	17.2	11.5	11.1
165	1	9.4	16.1	1	10.9	17.9	1	7.0	14.1	1	3.7	11.0	1	3.6	11.8	1	11.7	21.9	1	7.5	14.7	1	18.2	25.0	1	3.5	9.6	1	-	1	14.1	14.6	16.1	12.4	16.3	9.1	17.7	15.8	8.7	14.3
170	1	9.1	17.8	1	10.5	19.3	1	6.5	14.0	1	3.6	12.1	1	3.7	12.5	1	11.9	19.2	1	8.1	15.3	1	21.1	27.4	1	3.6	10.3	1	-	1	14.0	13.6	17.7	14.0	18.6	9.8	17.1	17.8	8.2	15.4
175	1	10.0	18.6	1	10.7	21.7	1	8.2	15.7	1	3.7	12.3	1	4.0	12.7	1	11.3	20.4	1	7.0	13.7	1	19.0	26.1	1	3.6	10.7	1	-	1	13.0	14.2	14.8	10.5	17.9	12.0	15.2	15.4	9.8	13.8
180	1	10.6	17.8	1	5.0	13.5	1	8.2	16.5	1	4.4	13.9	1	3.6	10.7	1	15.2	25.9	1	6.3	14.5	1	18.0	24.5	1	3.6	12.8	1	-	1	14.6	15.4	15.9	13.9	14.7	8.9	17.3	16.9	11.8	10.3
185	1	6.9	13.4	1	3.5	15.2	1	8.1	16.1	1	3.6	11.6	1	3.7	13.1	1	11.4	21.8	1	6.6	14.9	1	18.3	27.6	1	3.5	11.9	1	-	1	13.0	12.1	20.3	9.7	14.7	9.6	15.5	17.5	7.7	10.5
190	1	7.7	14.7	1	6.2	14.1	1	8.8	16.7	1	7.1	14.1	1	3.5	12.0	1	12.5	21.2	1	7.3	16.0	1	20.7	27.2	1	3.7	11.1	1	-	1	9.9	15.4	18.5	9.5	16.6	8.5	13.8	17.6	10.4	10.7
195	1	7.8	16.3	1	7.4	17.8	1	8.3	16.0	1	3.7	10.6	1	3.6	10.4	1	11.8	21.7	1	6.9	14.0	1	17.5	23.6	1	3.5	11.9	1	-	1	9.7	15.5	19.9	14.1	16.5	11.2	12.4	17.4	9.4	13.0
200	1	9.6	16.9	1	7.3	15.6	1	8.7	16.2	1	3.6	10.1	1	3.7	11.0	1	10.7	22.1	1	7.6	15.5	1	20.1	29.5	1	3.6	10.4	1	-	1	11.4	11.9	19.4	10.7	16.8	9.6	11.8	16.1	10.5	11.7

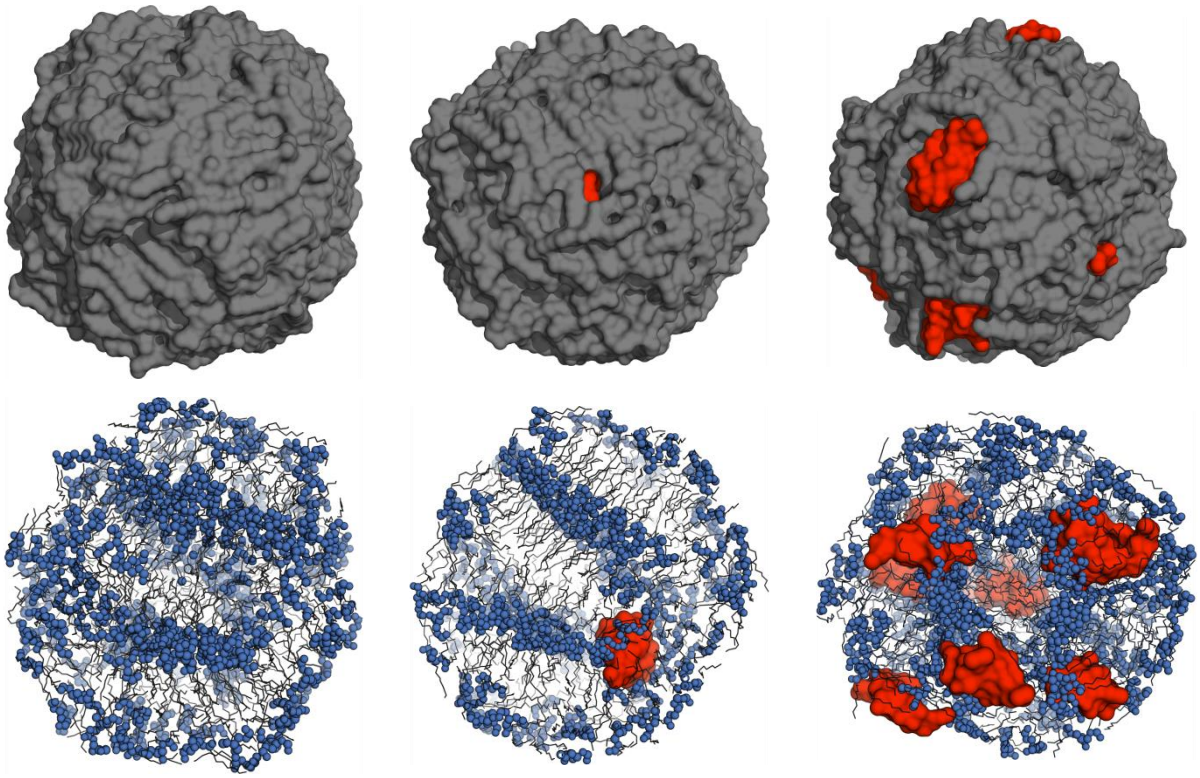


Figure S3: Larger size depiction of manuscript Figure 3: Drug CyA localization and molecular association in SEDDS droplets – upper row: drug (red surface) immersion in droplet (grey surface), lower row: molecular association of hydrophilic triglyceride moieties (blue), left column: caproic triglycerides (C6), middle column: capric triglycerides (C10), right column: myristic triglycerides (C14)

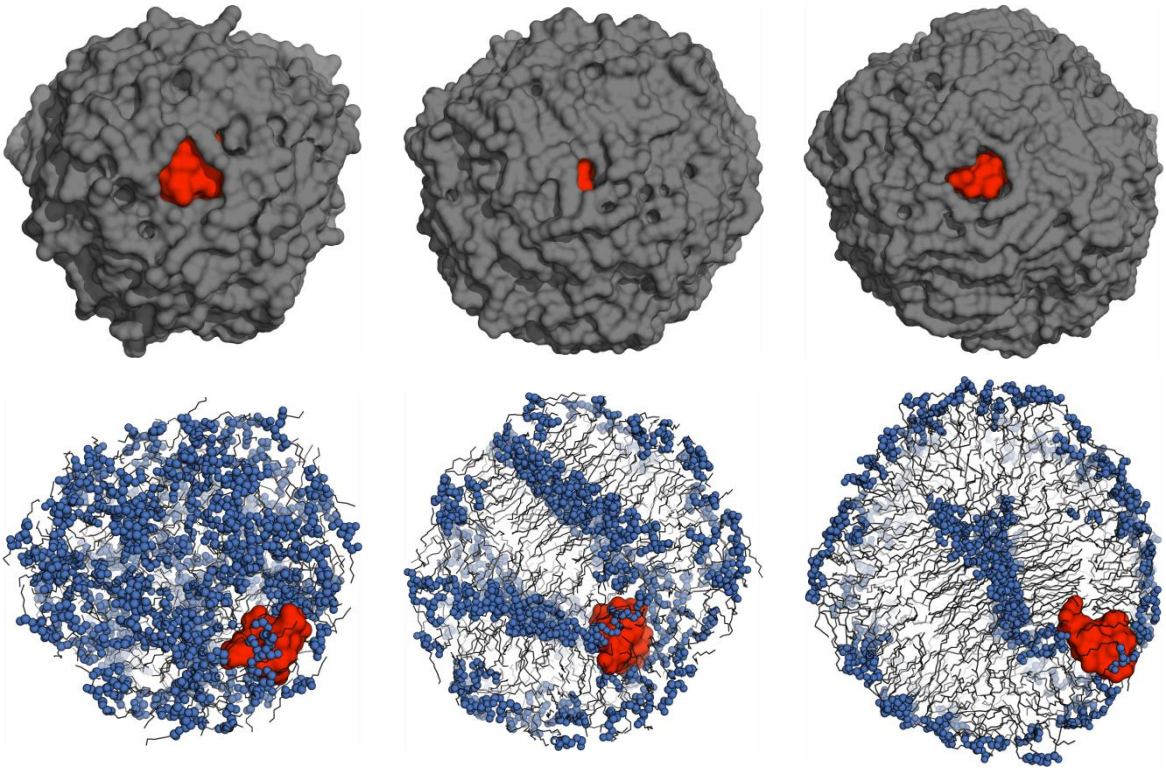


Figure S4: Larger size depiction of manuscript Figure 4: Drug CyA localization and molecular association in SEDDS droplets – upper row: drug (red surface) immersion in droplet (grey surface), lower row: molecular association of hydrophilic triglyceride moieties (blue), left column: caproic triglycerides (C6), middle column: capric triglycerides (C10), right column: myristic triglycerides (C14)

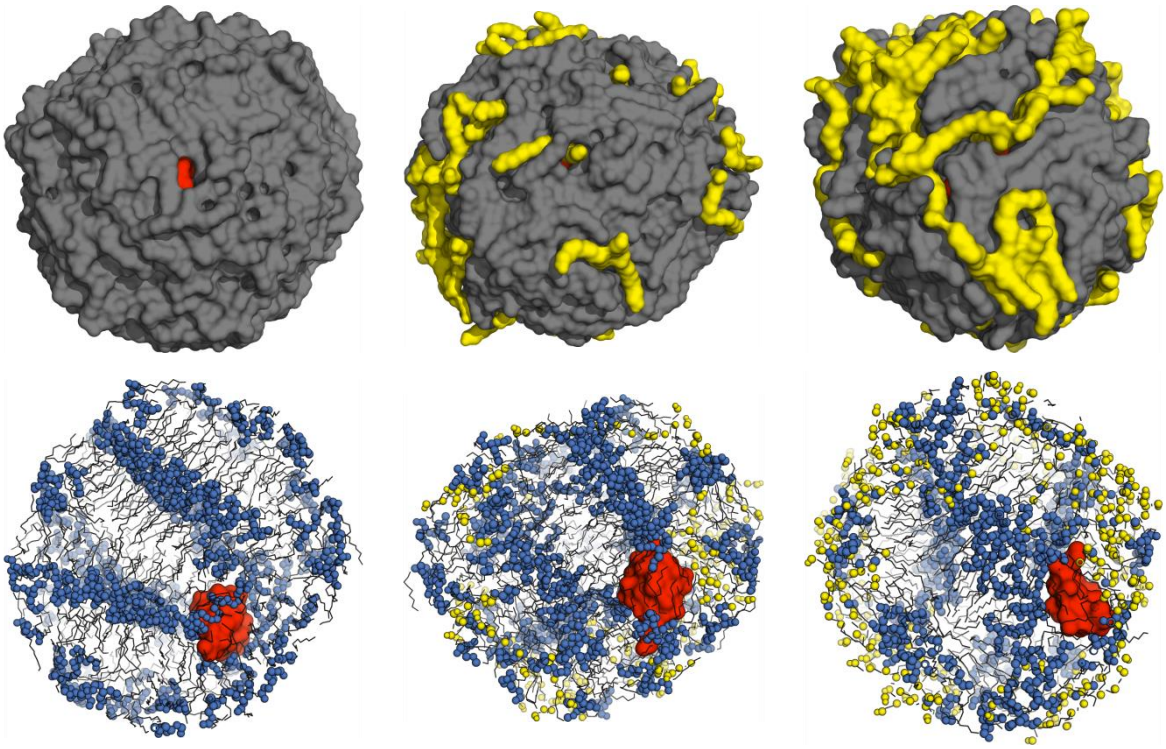


Figure S5: Larger size depiction of manuscript Figure 5: Drug CyA localization and molecular association in SEDDS droplets – upper row: drug (red spheres) immersion in droplet (grey surface), lower row: drug surface accessibility of hydrophilic triglyceride moieties (blue) and surfactant (yellow), left column: capric triglycerides (C10) without surfactant, middle column: capric triglycerides (C10) in 4:1 ratio with polyethylene glycol (PEG-6), right column: capric triglycerides (C10) in 3:2 ratio with polyethylene glycol (PEG-6)

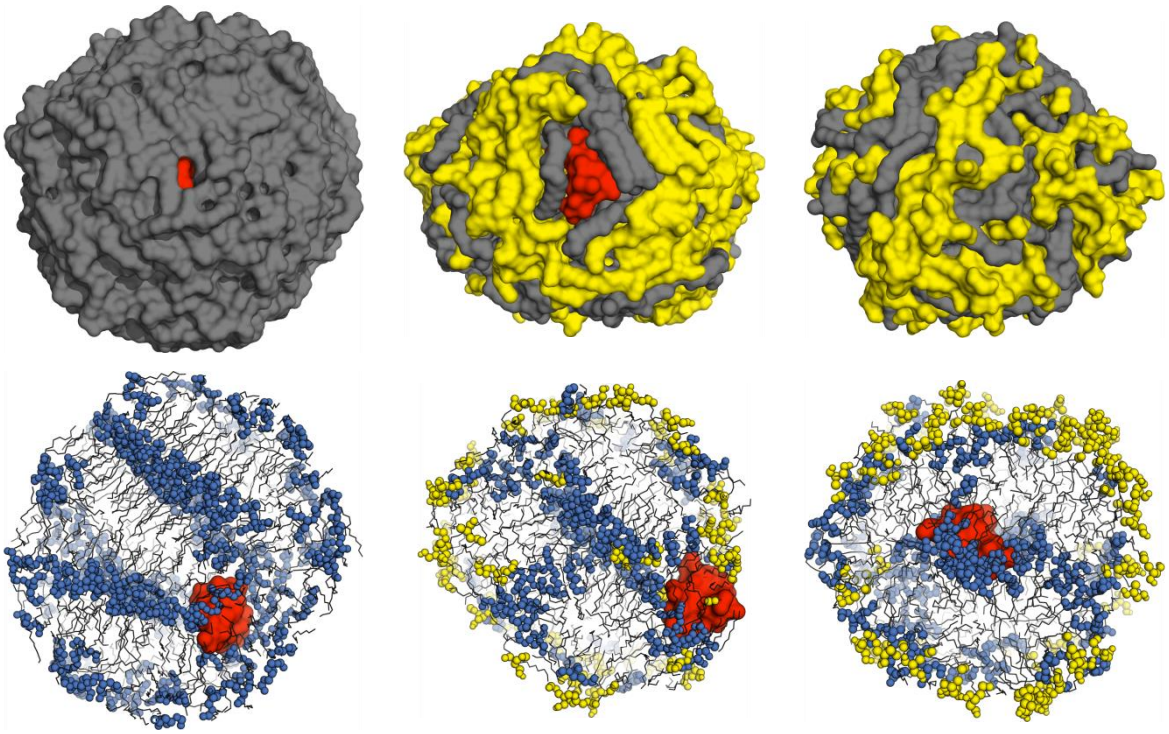


Figure S6: Larger size depiction of manuscript Figure 6: Drug CyA localization and molecular association in SEDDS droplets – upper row: drug (red surface) immersion in droplet (grey surface), lower row: drug surface accessibility of hydrophilic triglyceride (blue), diglyceride (yellow), monoglyceride (yellow) moieties, left column: capric triglycerides (C10), middle column: capric tri- and diglycerides in equimolar ratio, right column: capric tri- and monoglycerides in equimolar ratio

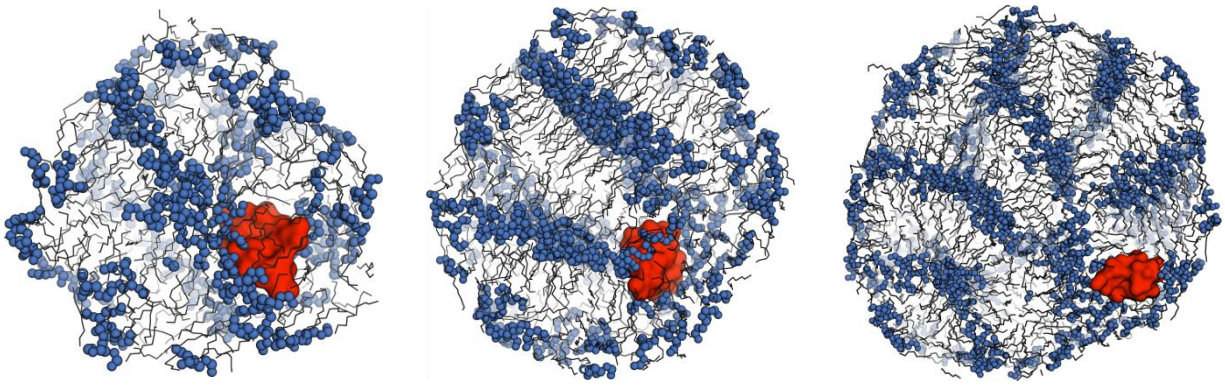


Figure S7: Implications of system size variation on the molecular association in a SEDDS droplet nanostructure of capric (C10) triglyceride excipients (blue spheres: hydrophilic moieties / grey lines: hydrophobic moieties) and the localization of drug CyA (red surface) in a SEDDS droplet nanostructure – depicted from left to right: systems of 100, 200 and 500 triglyceride molecules of diameters 60, 75 and 100Å respectively

Figure S8: Molecular association bias of size variations of the capric (C10) triglyceride SEDDS standard system, ranging from 100, 200 to 500 triglyceride molecules

	100C10	*200C10*	*500C10*
$TRIG_{PHOB}$ – *WAT*	0.28 ± 0.02	0.27 ± 0.01	0.27 ± 0.01
$TRIG_{PHIL}$ – *WAT*	0.79 ± 0.03	0.81 ± 0.03	0.81 ± 0.02
DRUG – *WAT*	1.98 ± 0.30	1.23 ± 0.12	0.27 ± 0.62
$TRIG_{PHOB}$ – $TRIG_{PHOB}$	1.13 ± 0.01	1.11 ± 0.00	1.09 ± 0.00
$TRIG_{PHOB}$ – $TRIG_{PHIL}$	0.98 ± 0.01	0.98 ± 0.00	0.96 ± 0.00
$TRIG_{PHIL}$ – $TRIG_{PHOB}$	0.53 ± 0.00	0.53 ± 0.00	0.51 ± 0.00
$TRIG_{PHIL}$ – $TRIG_{PHIL}$	2.13 ± 0.01	2.13 ± 0.01	2.13 ± 0.00
DRUG – $TRIG_{PHOB}$	0.30 ± 0.09	0.20 ± 0.04	0.47 ± 0.06
DRUG – $TRIG_{PHIL}$	2.29 ± 0.21	2.76 ± 0.12	2.43 ± 0.14

4.2.9 References

(1) Amidon, G.; Lennernäs, H.; Shah, V.; Crison, J. A Theoretical Basis for a BiopharmaceuticDrug Classification: The Correlation of in Vitro Drug Product Dissolution and in Vivo Bioavailability. Pharm. Res. 1995, 12, 413-420.

(2) Kohli, K.; Chopra, S.; Dhar, D.; Arora, S.; Khar, R. K. Self-Emulsifying Drug DeliverySystems: An Approach to Enhance Oral Bioavailability. Drug Discovery Today. 2010, 15, 958-965.

(3) Singh, B.; Bandopadhyay, S.; Kapil, R.; Singh, R.; Katare, O. P. Self-Emulsifying Drug Delivery Systems (SEDDS): Formulation Development, Characterization, and Applications. Crit. Rev. Ther. Drug Carrier Syst. 2009, 26, 427-451.

(4) Pouton, C. W. Lipid Formulations for Oral Administration of Drugs: Non-Emulsifying, Self-Emulsifying and 'Self-Microemulsifying' Drug Delivery Systems. Eur. J. Pharm. Sci. 2000, 11, S93-S98.

(5) Pouton, C. W. Formulation of Poorly Water-Soluble Drugs for Oral Administration: Physicochemical and Physiological Issues and the Lipid Formulation Classification System. Eur. J. Pharm. Sci. 2006, 29, 278-287.

(6) Aboofazeli, R.; Barlow, D. J.; Lawrence, M. J. Particle Size Analysis of Concentrated Phospholipid Microemulsions: II. Photon Correlation Spectroscopy. AAPS PharmSci. 2000, 2, 1–10.

(7) Amar, I.; Aserin, A.; Garti, N. Solubilization Patterns of Lutein and Lutein Esters in Food Grade Nonionic Microemulsions. J. Agric. Food Chem. 2003, 51, 4775-4781.

(8) Garti, N.; Amar, I.; Yaghmur, A.; Spernath, A.; Aserin, A. Interfacial Modification and Structural Transitions Induced by Guest Molecules Solubilized in U-Type Nonionic Microemulsions. J. Dispersion Sci. Technol. 2003, 24, 397-410.

(9) Garti, N.; Avrahami, M.; Aserin, A. Improved Solubilization of Celecoxib in U-Type Nonionic Microemulsions and their Structural Transitions with Progressive Aqueous Dilution. J. Colloid Sci. 2006, 299, 352-365.

(10) Malcolmson, C.; Barlow, D. J.; Lawrence, M. J. Light-Scattering Studies of Testosterone Enanthate Containing Soybean Oil/C18 : 1E10/Water Oil-in-Water Microemulsions. J. Pharm. Sci. 2002, 91, 2317-2331.

(11) Malcolmson, C.; Satra, C.; Kantaria, S.; Sidhu, A.; Lawrence, M. J. Effect of Oil on the Level of Solubilization of Testosterone Propionate into Nonionic Oil-in-Water Microemulsions. J. Pharm. Sci. 1998, 87, 109-116.

(12) Spernath, A.; Yaghmur, A.; Aserin, A.; Hoffman, R. E.; Garti, N. Food-Grade Microemulsions Based on Nonionic Emulsifiers: Media To Enhance Lycopene Solubilization. J. Agric. Food Chem. 2002, 50, 6917-6922.

(13) Spernath, A.; Yaghmur, A.; Aserin, A.; Hoffman, R. E.; Garti, N. Self-Diffusion Nuclear Magnetic Resonance, Microstructure Transitions, and Solubilization Capacity of Phytosterols and Cholesterol in Winsor IV Food-Grade Microemulsions. J. Agric. Food Chem. 2003, 51, 2359-2364.

(14) von Corswant, C.; Thorén, P. E. G. Solubilization of Sparingly Soluble Active Compounds in Lecithin-Based Microemulsions: Influence on Phase Behavior and Microstructure. Langmuir 1999, 15, 3710-3717.

(15) Warisnoicharoen, W.; Lansley, A. B.; Lawrence, M. J. Light-Scattering Investigations on Dilute Nonionic Oil-in-Water Microemulsions. AAPS PharmSci. 2000, 2, 16–26.

(16) Poulin, P.; Jones, R. D. O.; Jones, H. M.; Gibson, C. R.; Rowland, M.; Chien, J. Y.; Ring, B. J.; Adkison, K. K.; Ku, M. S.; He, H.; Vuppugalla, R.; Marathe, P.; Fischer, V.; Dutta, S.; Sinha, V. K.; Björnsson, T.; Lavé, T.; Yates, J. W. T. PHRMA CPCDC Initiative on Predictive Models of Human Pharmacokinetics, Part 5: Prediction of Plasma Concentration–Time Profiles in Human by Using the Physiologically-Based Pharmacokinetic Modeling Approach. J. Pharm. Sci. 2011, 100, 4127-4157.

(17) Prajapati, H. N.; Dalrymple, D. M.; Serajuddin, A. T. M. A Comparative Evaluation of Mono-, Di- and Triglyceride of Medium Chain Fatty Acids by Lipid/Surfactant/Water Phase Diagram, Solubility Determination and Dispersion Testing for Application in Pharmaceutical Dosage Form Development. Pharm. Res. 2012, 29, 285-305.

(18) Sacchetti, M.; Nejati, E. Prediction of Drug Solubility in Lipid Mixtures from the Individual Ingredients. AAPS PharmSciTech. 2012, 13, 1103-1109.

(19) Persson, L. C.; Porter, C. J.; Charman, W. N.; Bergström, C. A. Computational Prediction of Drug Solubility in Lipid Based Formulation Excipients. Pharm. Res. 2013, 30, 3225-3237.

(20) Rane, S. S.; Anderson, B. D. What Determines Drug Solubility in Lipid Vehicles: Is it Predictable? Adv. Drug Delivery Rev. 2008, 60, 638-656.

(21) Israelachvili, J. N. Intermolecular and Surface Forces, 3rd ed.; Academic Press: Burlington, MA, 2011.

(22) Anderson, B. D.; Rytting, J. H.; Higuchi, T. Solubility of Polar Organic Solutes in Nonaqueous Systems: Role of Specific Interactions. Pharm. Sci. 1980, 69, 676-680.

(23) King, D. T.; Warren, D. B.; Pouton, C. W.; Chalmers, D. K. Using Molecular Dynamics to Study Liquid Phase Behavior: Simulations of the Ternary Sodium Laurate/Sodium Oleate/Water System. Langmuir. 2011, 27, 11381-11393.

(24) Kasimova, A. O.; Pavan, G. M.; Danani, A.; Mondon, K.; Cristiani, A.; Scapozza, L.; Gurny, R.; Müller, M. Validation of a Novel Molecular Dynamics Simulation Approach for Lipophilic Drug Incorporation into Polymer Micelles. J. Phys. Chem. B. 2012, 116, 4338-4345.

(25) Marrink, S.-J.; Tieleman, D. P. Molecular Dynamics Simulation of a Lipid Diamond Cubic Phase. J. Am. Chem. Soc. 2001, 123, 12383-12391.

(26) Rane, S. S.; Anderson, B. D. Molecular Dynamics Simulations of Functional Group Effects on Solvation Thermodynamics of Model Solutes in Decane and Tricaprylin. Mol. Pharmaceutics. 2008, 5, 1023-1036.

(27) Warren, D.; King, D.; Benameur, H.; Pouton, C.; Chalmers, D. Glyceride Lipid Formulations: Molecular Dynamics Modeling of Phase Behavior During Dispersion and Molecular Interactions Between Drugs and Excipients. Pharm Res. 2013, 30, 3238-3253.

(28) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J. Chem. Theory Comput. 2008, 4, 435-447.

(29) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. J. Comput. Chem. 2005, 26, 1701-1718.

(30) Wensink, E. J. W.; Hoffmann, A. C.; van Maaren, P. J.; van der Spoel, D. Dynamic Properties of Water/Alcohol Mixtures Studied by Computer Simulation. J. Chem. Phys. 2003, 119, 7308-7317.

(31) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. Phys. Rev. A 1985, 31, 1695-1697.

(32) Nose, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. J. Chem. Phys. 1984, 81, 511-519.

(33) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. J. Appl. Phys. 1981, 52, 7182-7190.

(34) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of n-Alkanes. J. Comput. Phys. 1977, 23, 327-341.

(35) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. J. Chem. Phys. 1995, 103, 8577-8593.

(36) York, D. M.; Darden, T. A.; Pedersen, L. G. The Effect of Long-Range Electrostatic Interactions in Simulations of Macromolecular Crystals - a Comparison of the Ewald and Truncated List Methods. J. Chem. Phys. 1993, 99, 8345-8348.

(37) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. J. Comput. Phys. 2004, 25, 1656-1676.

(38) Berger, O.; Edholm, O.; Jähnig, F. Molecular Dynamics Simulations of a Fluid Bilayer of Dipalmitoylphosphatidylcholine at Full Hydration, Constant Pressure, and Constant Temperature. Biophys. J. 1997, 72, 2002-2013.

(39) Horta, B. A. C.; Fuchs, P. F. J.; van Gunsteren, W. F.; Hünenberger, P. H. New Interaction Parameters for Oxygen Compounds in the GROMOS Force Field: Improved Pure-Liquid and Solvation Properties for Alcohols, Ethers, Aldehydes, Ketones, Carboxylic Acids, and Esters. J. Chem. Theory Comput. 2011, 7, 1016-1031.

(40) Winger, M.; de Vries, A. H.; van Gunsteren, W. F. Force-field Dependence of the Conformational Properties of α,ω-Dimethoxypolyethylene Glycol. Mol. Phys. 2009, 107, 1313-1321.

(41) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. J. Phys. Chem. 1987, 91, 6269-6271.

(42) Mitchell, D. J.; Tiddy, G. J. T.; Waring, L.; Bostock, T.; McDonald, M. P. Phase Behaviour of Polyoxyethylene Surfactants with Water. Mesophase Structures and Partial Miscibility (Cloud Points). J. Chem. Soc., Faraday Trans. 1 1983, 79, 975-1000.

(43) Benson, S. P.; Pleiss, J. Incomplete Mixing Versus Clathrate-Like Structures: A Molecular View on Hydrophobicity in Methanol-Water Mixtures. J. Mol. Model. 2013, 19, 3427-3436.

(44) Gershanik, T.; Benzeno, S.; Benita, S. Interaction of a Self-Emulsifying Lipid Drug Delivery System with the Everted Rat Intestinal Mucosa as a Function of Droplet Size and Surface Charge. Pharm. Res. 1998, 15, 863-869.

(45) Sander, C.; Holm, P. Porous Magnesium Aluminometasilicate Tablets as Carrier of a Cyclosporine Self-Emulsifying Formulation. AAPS PharmSciTech. 2009, 10, 1388-1395.

(46) Zidan, A. S.; Sammour, O. A.; Hammad, M. A.; Megrab, N. A.; Habib, M. J.; Khan, M. A. Quality by Design: Understanding the Formulation Variables of a Cyclosporine A Self-Nanoemulsified Drug Delivery Systems by Box-Behnken Design and Desirability Function. Int. J. Appl. Pharm. 2007, 332, 55-63.

(47) Odeberg, J. M.; Kaufmann, P.; Kroon, K.-G.; Höglund, P. Lipid Drug Delivery and Rational Formulation Design for Lipophilic Drugs with Low Oral Bioavailability, Applied to Cyclosporine. Eur. J. Pharm. Biopharm. 2003, 20, 375-382.

(48) Abdel-Mottaleb, M. M. A.; Neumann, D.; Lamprecht, A. In Vitro Drug Release Mechanism from Lipid Nanocapsules (LNC). Int. J. Pharm. 2010, 390, 208-213.

(49) Attama, A. A.; Reichl, S.; Muller-Goymann, C. C. Sustained Release and Permeation of Timolol from Surface-Modified Solid Lipid Nanoparticles through Bioengineered Human Cornea. Curr. Eye Res. 2009, 34, 698-705.

(50) zur Mühlen, A.; Schwarz, C.; Mehnert, W. Solid Lipid Nanoparticles (SLN) for Controlled Drug Delivery - Drug Release and Release Mechanism. Eur. J. Pharm. Biopharm. 1998, 45, 149-155.

(51) Thomas, N.; Holm, R.; Garmer, M.; Karlsson, J.; Müllertz, A.; Rades, T. Supersaturated Self-Nanoemulsifying Drug Delivery Systems (Super-SNEDDS) Enhance the Bioavailability of the Poorly Water-Soluble Drug Simvastatin in Dogs. AAPS J. 2013, 15, 219-227.

(52) Thomas, N.; Holm, R.; Müllertz, A.; Rades, T. In Vitro and In Vivo Performance of Novel Supersaturated Self-Nanoemulsifying Drug Delivery Systems (Super-SNEDDS). J. Controlled Release 2012, 160, 25-32.

(53) Lei, Z. G.; Chen, B. H.; Li, C. Y.; Liu, H. Predictive Molecular Thermodynamic Models for Liquid Solvents, Solid Salts, Polymers, and Ionic Liquids. Chem. Rev. 2008, 108, 1419-1455.

(54) Deckelbaum, R. J.; Hamilton, J. A.; Moser, A.; Bengtssonolivecrona, G.; Butbul, E.; Carpentier, Y. A.; Gutman, A.; Olivecrona, T. Medium-Chain Versus Long-Chain Triacylglycerol Emulsion Hydrolysis by Lipoprotein-Lipase and Hepatic Lipase - Implications for the Mechanisms of Lipase Action. Biochemistry 1990, 29, 1136-1142.

(55) Prasad, D.; Chauhan, H.; Atef, E. Studying the Effect of Lipid Chain Length on the Precipitation of a Poorly Water Soluble Drug from Self-Emulsifying Drug Delivery System on Dispersion into Aqueous Medium. J. Pharm. Pharmacol. 2013, 65, 1134-1144.

4.3 The Solvent Flux Method (SFM): A Case Study of Water Access to *Candida Antarctica* Lipase B

4.3.1 Abstract

The solvent flux method (SFM) was developed to comprehensively characterize the influx of solvent molecules into the active site of a protein and to thereby detect possible entryways in the protein structure as well as amino acid positions relevant to solvent-related enzyme engineering, particularly to improve reverse hydrolysis reaction rates in non-aqueous media. SFM introduces a solvent concentration gradient and adds artificial forces to accelerate solvent influx to the active site in molecular dynamics simulations and thus allows for an evaluation of solvent entryways with high statistical significance in short timescales. *Candida antarctica* lipase B (CALB), with its active site entrance channel oriented towards a triglyceride-water interface, i.e., its functionally active conformation, served as a model system for SFM to evaluate the influx of water molecules. A known water channel in CALB was thereby identified as a proof of principle for SFM, a secondary water channel and two avenues for water entry that contribute to the general propensity of water molecules to "leak" in between the protein and the triglyceride-water interface.

4.3.2 Introduction

The diversity of proteins has evolved under aqueous solution conditions of the cytosol in living organisms, creating an inseparable link between form and function of proteins and the presence of water. Aqueous environments determine protein folding, allow for educts and products to diffuse in and out of enzymes or between enzymes in cascade reactions; multiple enzymes may interact or assemble into complexes and work in unison within elaborate metabolic networks. Water may be directly involved in enzyme reactions; the enzyme class (EC) that hereby stands out in regards to commercial interest are hydrolases (3.-.-.-), specifically those cleaving ester bonds (3.1.-.-), where water acts as a nucleophile on an acyl-enzyme intermediate and thus facilitates the release of alcohol and acid constituents. While this natural degradative enzyme function is of substantial interest to a

range of industrial applications[1], such as food processing, breaking down of natural materials or waste treatment, significant strides have been made during the last decades to shift the chemical equilibrium towards the reverse synthesis reaction by transferring hydrolases into non-natural organic solvent environments[2-4] or more recently ionic liquids[5]. A broad range of substrates may thus be introduced to the reaction media as competing nucleophiles to the acyl-enzyme intermediate, which enables a diverse set of synthetic possibilities, such as transesterification, aminolysis or thiolysis[2, 4-6]. Owing to the evolutionary adaptation of hydrolases to aqueous environments, transferring enzymes to anhydrous conditions raises new challenges like maintaining enzyme stability and optimizing the activity of the desired synthetic reaction. This is addressed by diverse avenues of biotechnological engineering, such as chemical modifications of enzyme surfaces[7-9], crosslinking[10], lyophilization[11] that exploits the phenomenon of pH memory[3], enzyme immobilization techniques[12], engineering solvent conditions[13] with appropriate log P values[14] or developing solvent tolerant variants by directed evolution[15, 16]. Despite these efforts, the presence of residual hydration water seems to be a prerequisite for enzyme activity in non-aqueous environments, which raises the fundamental problem of competing native hydrolysis reaction and the desired reverse synthetic reaction[2]. Moreover, defining and maintaining "dry" conditions is difficult as well as costly to achieve on the industrial scale, which motivated Larsen *et al.*[17] to propose an approach to suppress water as a nucleophile in *Candida antarctica* lipase B (CALB) by rationally devising variants that diminish the influx of water to the active site cavity and thus dramatically increase transacylation rates for vinyl butyrate over hydrolysis rates in butanol-water mixtures. CALB catalyzes hydrolysis of triglycerides in its native physiological setting; the natural substrate triglyceride being poorly water-soluble forms phase-separated interfaces in aqueous environments and thus CALB needs to attach to these interfaces in its active conformation, which is with its substrate channel oriented towards the triglyceride phase. The most obvious entryway for water molecules is thus blocked by a hydrophobic barrier. However, a constant influx of water to the active site cavity is a prerequisite for hydrolysis to occur and therefore, Larsen et al.[17] proposed the existence of a water channel, which was targeted for mutation and implicitly validated by experiment. The aforementioned study supports the notion that a molecular understanding of the role of residual water in non-aqueous reaction media is an integral part for furthering the development of hydrolase applications[4]. Due to the difficulty of examining individual water molecules by

experimental means, molecular dynamics (MD) studies have been conducted extensively to elucidate enzyme-solvent interactions under various conditions, which are comprehensively summarized in a recent review by Lousa et al.[18]. Herein it is outlined how MD analysis on overall protein stability[19-21] and flexibility[22-24] in organic solvents or ionic liquids[25, 26] may offer deductions on mechanistic principles such as pH memory or the required minimal water content[27], or even how significant parameters such as water activity may be modeled in a manner[28, 29] that is consistent with experimental data[30]. By analyzing hydration water and its influence on protein secondary structure flexibility, it may also be possible to indirectly identify water entrance channels[31]. In the present study, we set out to contribute to this vibrant area of research by proposing an MD-based approach, the solvent flux method (SFM), which allows for a holistic characterization of the influx of solvent molecules through the protein structure and into the active site cavity. We hereby detect structural features relevant to the solvent influx, as well as positions that may prove beneficial to the rational design of enzyme variants for hydrolase applications. In this context, SFM was benchmarked for the influx of water molecules to the active site of CALB by comparing predictions to the experimental results by Larsen et al.[17]

4.3.3 Experimental Section

4.3.3.1 Simulation Details

Molecular dynamics simulations were performed with the GROMACS 4.5.3[32, 33] software in an NPT ensemble at 298.15K and 1bar, using the leap-frog integrator[34] with a time step of 2fs. The Berendsen thermo- and barostats were applied for reasons of robustness and efficiency when converging systems far from equilibrium to the equilibrated state. This was deemed necessary since the method developed in this study introduces significant periodic rescaling and reorientation of water molecule velocity vectors close to the protein (see results section for SFM details). Water molecules with adjusted velocity vectors were relatively loosely coupled in intervals of 50ps to prolong the velocity rescaling effect under Berendsen temperature coupling (Figure S9 and S10), while the remaining bulk water molecules, the triglyceride molecules and the protein were coupled relatively tightly in intervals of 0.1ps. To ensure that the structural integrity of the system remained intact

during simulation, i.e. to avoid protein denaturation or distortions in the protein-triglyceride interface, position restraints with force constants of 1000 kJ mol^{-1} nm^{-2} were applied to the protein backbone and the triglyceride molecules in all spatial dimensions (Figure S7). To avoid artifacts due to position restraints under NPT conditions, the reference coordinates of each molecule group's center of mass were rescaled periodically. Semi-isotropic pressure coupling was applied independently to the x-y plane and in z-dimension for a more accurate pressure representation of the planar triglyceride layer[35] in the model system. Pressure coupling intervals were set to 100ps for the adjusted water molecules and to 5ps for the protein, triglyceride and bulk water molecules. The center of mass movements were removed every 100 simulation steps independently for all system components. Periodic boundary conditions were applied in all three dimensions. Hydrogen bonds were constrained with the LINCS algorithm[36]. Long-range electrostatics were calculated with the Particle-Mesh Ewald (PME) method[37, 38]. Lennard-Jones interactions were treated with a cutoff and capped at 1.2nm. The OPLS/AA all-atom force field[39] was used to parameterize the protein lipase B from *Candida antarctica* (CALB). Structural information of CALB was obtained from the Protein Data Bank (PDB: 1TCA[40]). Triglyceride molecules were parameterized with the Berger lipid model[41] and the 1-4 interactions adapted with the half-ε double-pairlist method of Neale and Pomès[42] for consistency with the OPLS-AA force field. Water molecules were parameterized with the TIP3P[43] model.

4.3.3.2 System Creation and Equilibration

A model system was created to mimic the conditions of the functionally active conformation of CALB at the triglyceride-water interface. As a prerequisite step, a random molecular arrangement of caproic (C6) triglyceride molecules and water molecules had to be equilibrated into a phase-separated triglyceride layer, according to the approach of Gruber et al.[44]. The resulting triglyceride-water interface hereby represents a model for a droplet surface found in triglyceride in water emulsions (Figure 1). This model was deemed appropriate considering that the diameter of CALB (~5nm) is three orders of magnitude smaller than experimentally determined droplet diameters (~2.0µm)[45]; hence the droplet surface curvature is expected to be negligible on the model scale. Furthermore,

when considering that the Coulomb potential decays with 1/r, a layer thickness of ~7nm should be sizeable enough to ensure that the most significant Coulomb contributions to the non-bonded potential between the protein and the triglyceride surface are accounted for. Moreover, the lamellar-like nanostructures observed for the triglyceride layer model (Figure 1) are in agreement with previously published results on large-scale triglyceride aggregates[46]. The potential implication of these structural features on the mobility of water molecules are subject of a detailed follow-up study and will therefore not be elaborated on in this context. CALB was added to the triglyceride-water system and was attached to the layer with the active site entrance pointing towards the triglyceride phase, thus representing a model for the active binding configuration for triglyceride hydrolysis. The system was equilibrated for a further 500ns, while monitoring protein adhesion and immersion depth (unpublished data).

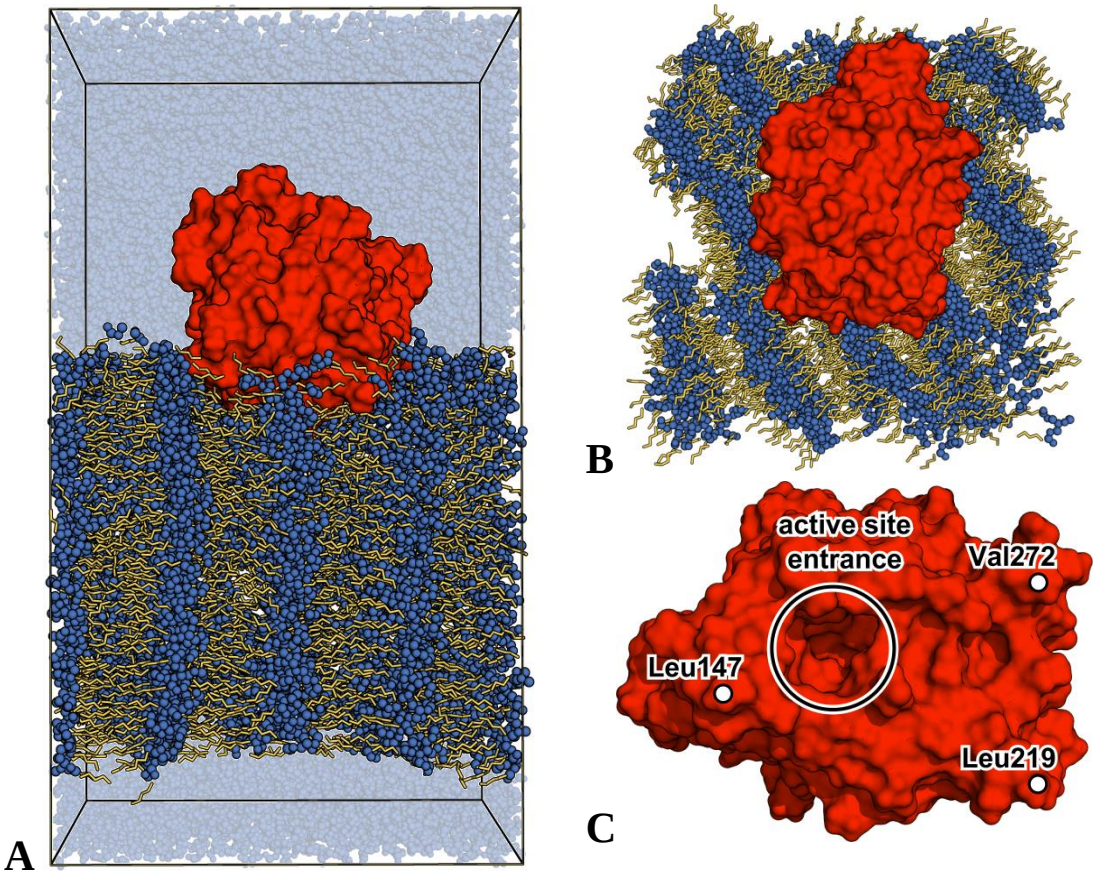


Figure 1: Simulation system consisting of *Candida antarctica* lipase B attached to a triglyceride-water interface (red surface: protein / dark blue spheres: polar triglyceride moieties / orange sticks: nonpolar triglyceride moieties) – A) Side view of full system including water (light blue spheres in the background) / B) Top view of protein and triglyceride layer / C) Bottom view of protein, white circle depicting the entrance to the active site, Leu147, Leu219 and Val272 are anchor residues by which CALB attaches to the triglyceride layer[44].

4.3.3.3 The Solvent Flux Method (SFM)

The design and benchmarking of SFM was tightly coupled to the test case system of CALB at the triglyceride-water interface, with water as the influx solvent. Two major technical considerations were formative for SFM during the design phase. Firstly, it was necessary to differentiate solvent molecules that pass through an entryway and into the active site cavity from other bulk solvent molecules. This was achieved in a recursive manner, by exploiting the fact that solvent molecules present within the active site cavity necessarily had to have passed through an entryway of the protein at a preceding point in time. It was thus possible to reconstruct a solvent molecule pathway through the protein and therefore to detect general avenues of entry on a statistical basis. To identify such water molecules and to ensure that they had indeed passed into the active site cavity from the exterior, the water molecule closest to the reference atom at the intended influx site (here: oxygen atom of Ser105 of the catalytic triad of CALB) and within a spherical water removal cutoff of 0.35nm was periodically removed from the simulation system after Δt_{ITER}=10ps (Figure 2). Not only did this ensure that after a given number of iterations all solvent molecules previously present within the active site cavity were evacuated, it also introduced an increased influx of solvent molecules due to the solvent concentration gradient between the active site of the protein and the exterior solvent bulk. Moreover, the removal of water served as a mechanistic model for water consumption during the native CALB hydrolysis reaction. The induced solvent influx ties in with the second major technical consideration, which is the acceleration of conventional MD to rapidly overcome rate-limiting energy barriers specific for water influx and thus realize a significant data depth in a feasible duration of real-time. This was achieved by introducing a reorientation and rescaling of velocity vectors of water molecules that were located within a water velocity rescaling cutoff of 3nm surrounding the active site reference, which was adjusted to incorporate the entirety of the protein and its first solvation shell. At the beginning of every iteration, velocity magnitudes of water molecules present within this cutoff were rescaled with a constant factor v_{MULT}, which effectively defines the strength of a periodic velocity pulse that is introduced to overcome the influx-rate-limiting energy barriers in a timely manner. The orientation of the new vectors was defined by assigning a fraction v_{OLD} of the respectively rescaled vector magnitudes to the unit vectors in the original direction, while a larger fraction v_{CENT}=100%-v_{OLD} was assigned to unit vectors pointing towards the active site reference (Figure 2). Summation of these vector pairs yielded the new velocity

vectors v_{NEW}, which were used to update the initial conditions at the beginning of an iteration. Thereafter, temperatures were allowed to converge back towards equilibrium during Δt_{ITER}. It proved necessary to additionally fine-tune the duration of this effect by modulating the temperature coupling parameter τ_{ITER}, thus coupling of the velocity-adjusted water molecules more loosely than the other system elements. The process of single water molecule removal, velocity readjustment and simulation continuation was repeated for a predefined number of iteration steps n_{ITER}. With ΔN_{WAT} denoting the sum of all water molecules removed during a full SFM run, we can express the total water removal rate at the active site reference as $\Delta N_{WAT}/n_{ITER}$, with $\Delta N_{WAT}/n_{ITER}=1$ signifying water removal at every iteration.

4.3.4 Results

In this study, we present the computational solvent flux method (SFM) that is devised on the basis of non-equilibrium molecular dynamics (MD) simulations and comprehensively characterizes the access of solvent molecules from the bulk medium to the active site cavity of a protein. The intended purpose of SFM is to identify solvent entrance pathways and amino acid positions for solvent-related enzyme engineering, particularly in the context of improving reverse hydrolysis reaction rates. In the following, SFM and its parameters are thoroughly benchmarked in a case study of the enzyme *Candida antarctica* lipase B (CALB) with the solvent water (Figure 1). The feasibility of SFM analysis is then assessed by comparing results with the successful experimental study of Larsen et al. on the same system[17], which was conducted under the premise of the existence of a specialized water channel.

4.3.4.1 SFM Parameter Optimization in the CALB and Water System

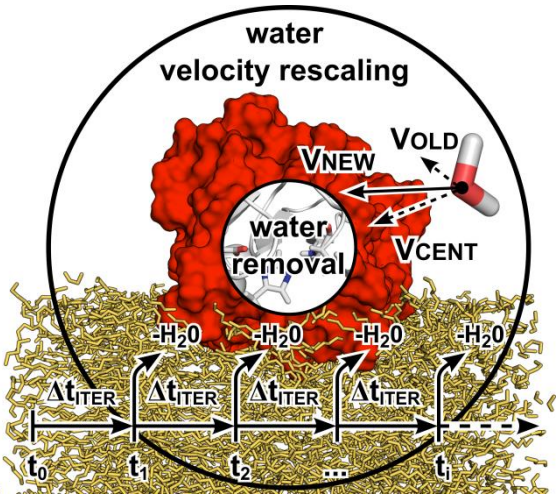


Figure 2: Velocity vector magnitudes of water molecules found within the water velocity rescaling cutoff around an active site reference atom are rescaled by a factor of v_{MULT} and then transformed to v_{NEW} after every iteration Δt_{ITER}. After every iteration Δt_{ITER} a single water molecule is removed within the water removal cutoff radius. The direction of v_{NEW} is defined by X% of the rescaled magnitudes that are allocated to the original vector directions (v_{OLD}) and by Y% of the rescaled magnitudes that are allocated to the direction of the active site reference atom (V_{CENT}), whereby X+Y=100 and Y>>X.

The major parameters to configure SFM include the factor v_{MULT} by which velocity vectors of solvent molecules are rescaled, the temperature coupling constants τ_{ITER} of velocity-adjusted water molecules, the strength of position restraints (pos.res.) that are applied to the protein backbone and triglyceride molecules and the percentage v_{CENT} that defines the distribution of a rescaled velocity magnitude onto the vectors v_{OLD} and v_{CENT} (Figure 2 – details see methods section). These parameters were extensively tuned to attain a setup that balances the need for high water removal rates $\Delta N_{WAT}/n_{ITER}$ at the active site, while maintaining the structural integrity of the system despite the artificial forces introduced by SFM (Table 1). During parameter tuning, the RMSD was monitored to assess the structural integrity of the system. The RMSD was calculated by fitting the protein structure obtained after n_{ITER}=100 iterations onto the initial protein structure, while $\Delta T_{ITER}=|T_{100}-T_{10}|$ was monitored to observe potential heat-buildup in the systems and is defined by the temperature difference of the velocity-adjusted water molecules between iterations n_{ITER}=100 and n_{ITER}=10, averaged over the predefined iteration duration of Δt_{ITER}=10ps

(Table 1). No significant heat build-up was observed in any of the analyzed systems (Figure S4-S10). A system setup that corresponds to conventional MD simulation conditions without artificial external forces (Table 1 - system no.1) served as a reference to compare systems tested during parameter tuning. As a first step, the velocity scaling factor v_{MULT} was varied from 1.5 to 4.0 (systems no. 2-7), while the remaining parameters were initially set to position restraints of 100 kJ mol^{-1} nm^{-2}, v_{CENT}=75% and τ_{ITER}=0.1ps. Increasing v_{MULT} introduces the necessary energy to overcome barriers that water molecules would otherwise be unlikely to pass within Δt_{ITER}=10ps, as the water removal rate $\Delta N_{WAT}/n_{ITER}$=0.02 (system no.1) demonstrates. While $\Delta N_{WAT}/\Delta t_{ITER}$ increased only slightly from initially 0.02 to 0.06, v_{MULT}=3.5 (system no. 6) was maintained from thereon as a relatively high velocity rescaling factor in foresight of increasing $\Delta N_{WAT}/\Delta t_{ITER}$ via adjustment of the remaining parameters. The RMSD=0.079nm (system no. 6) was slightly increased in spite of position restraints 100 kJ mol^{-1} nm^{-2}, which is why the position restraints (pos.res.) were varied as a second step, from 0 to 5000 kJ mol^{-1} nm^{-2} (system no. 8-11) to achieve a compromise between stable structures and conformational freedom of the protein backbone. Position restraints of 1000 kJ mol^{-1} nm^{-2} proved adequate in maintaining the RMSD=0.059nm (system no. 10) at standard levels of RMSD=0.069nm (system no. 1). As a third step, the amount of the rescaled water velocity magnitudes that is assigned to v_{CENT} was varied from 0 to 100% (system no. 12-15), which caused an increase of $\Delta N_{WAT}/\Delta t_{ITER}$ from 0.06 to 0.16 for v_{CENT}=90%. The RMSD=0.069nm (system no. 14) was thereby not significantly altered. A residual v_{OLD}=10%=100%-v_{CENT} was maintained in the original vector directions to ensure residual sideways velocity components to laterally bypass potential barriers. The most significant increase in $\Delta N_{WAT}/\Delta t_{ITER}$ from 0.16 to 0.99 was obtained during the fourth and final parameter tuning step (system no. 16-22) by variation of the temperature coupling constant τ_{ITER} of the velocity-adjusted water molecules from 0.5ps to 100ps. The resulting final parameter set (system no. 21) of v_{MULT}=3.5, pos.res.=1000 kJ mol^{-1} nm^{-2}, v_{CENT} =90% and τ_{ITER}=50ps ensured the removal of water after nearly every iteration with $\Delta N_{WAT}/\Delta t_{ITER}$=0.99, while maintaining the structural integrity of the system with RMSD=0.059nm and without heating artifacts due to velocity adjustment.

Table 1: A systematic benchmarking of SFM parameters v_{MULT}, v_{CENT}, τ_{ITER} and position restraints (pos.res.) was conducted for n_{ITER}=100 iterations with time intervals of Δt_{ITER}=10ps, to establish a parameter set that maximizes the water removal rate $\Delta N_{WAT}/n_{ITER}$ (0.00: no water molecule removed, 1.00: water molecule removed after every iteration). Moreover, the structural integrity of the protein was monitored (RMSD), as well as the difference in temperatures between n_{ITER}=10 and n_{ITER}=100 iterations (ΔT_{ITER}), to ensure that the accelerated water molecules do not accumulate residual heat during SFM.

		variation of SFM parameters				wat. rem. rate at active site	structural integrity of protein and heat build-up in accelerated water					
	no.	v_{MULT}	pos. res. [kJ mol^{-1} nm^{-2}]	v_{CENT} [%]	τ_{ITER} [ps]	$\Delta N_{WAT}/n_{ITER}$ 0.00 - 1.00	RMSD† [nm]			ΔT_{ITER}†† [K]		
velocity magnitude magnitude scaling	1	1.0	0	0	0.1	0.02	0.069	±	0.002	0.03	±	4.78
	2	1.5	100	75	0.1	0.01	0.064	±	0.003	0.04	±	3.51
	3	2.0	100	75	0.1	0.03	0.068	±	0.003	0.77	±	3.90
	4	2.5	100	75	0.1	0.10	0.060	±	0.006	0.31	±	4.27
	5	3.0	100	75	0.1	0.05	0.081	±	0.007	0.63	±	3.69
	6	**3.5**	100	75	0.1	0.06	0.079	±	0.008	0.77	±	4.54
	7	4.0	100	75	0.1	0.11	0.078	±	0.009	1.16	±	4.06
position restraints	8	3.5	0	75	0.1	0.08	0.112	±	0.012	0.00	±	3.88
	9	3.5	500	75	0.1	0.09	0.076	±	0.003	0.71	±	3.60
	10	3.5	**1000**	75	0.1	0.14	0.059	±	0.004	1.76	±	4.13
	11	3.5	5000	75	0.1	0.18	0.050	±	0.001	0.99	±	3.43
velocity vect. reorientation	12	3.5	1000	0	0.1	0.07	0.060	±	0.001	0.45	±	3.71
	13	3.5	1000	50	0.1	0.09	0.065	±	0.002	0.84	±	4.57
	14	3.5	1000	**90**	0.1	0.16	0.069	±	0.004	0.75	±	4.44
	15	3.5	1000	100	0.1	0.24	0.065	±	0.005	0.23	±	3.98
temperature coupling constant	16	3.5	1000	90	0.5	0.40	0.067	±	0.005	1.85	±	4.88
	17	3.5	1000	90	1	0.65	0.069	±	0.005	4.32	±	5.58
	18	3.5	1000	90	5	0.55	0.081	±	0.005	6.89	±	8.57
	19	3.5	1000	90	10	0.57	0.074	±	0.006	4.10	±	10.61
	20	3.5	1000	90	25	0.73	0.070	±	0.005	0.09	±	10.83
	21	**3.5**	**1000**	**90**	**50**	**0.99**	**0.068**	**±**	**0.004**	**0.60**	**±**	**9.37**
	22	3.5	1000	90	100	1.00	0.067	±	0.004	10.21	±	11.38

† RMSD after nITER=100 iterations, averaged over ΔtITER=10ps after backbone fitting to starting structure of nITER=1

†† $\Delta T_{ITER}=T_{100}-T_{10}$ of temperatures at n_{ITER}=100 and n_{ITER}=10, averaged over Δt_{ITER}=10ps (largest s.d. shown)

4.3.4.2 Identifying Sources of Influx with SFM in the CALB and Water Model System

SFM was applied to test for water influx in the system of CALB attached to a triglyceride-water interface (Figure 1), using the parameter set as optimized in the previous section. Influx occurred towards the active site reference oxygen atom of Ser105 of CALB, due to the water concentration gradient between the protein exterior and the active site, induced by the periodic removal of water molecules and the periodic introduction of a radial velocity pulse to the water molecules in close proximity of the protein. After an iteration n_{EVAC}=100, it was ensured that all water molecules that were originally present in the active site cavity had been removed and thus any water molecules that were consecutively removed in $n_i > n_{EVAC}$ necessarily had to have passed through a water entryway of the protein to reach the active site. The pathway of any water molecule that was thus removed

during an iteration n_i, with $n_{EVAC}<n_i<n_{ITER}=500$, was reconstructed by superimposing the system coordinates in a least square fit of protein backbone atoms of all prior iterations $n_j<n_i$. Moreover, the contact frequency between the removed water molecules and the position-specific amino acids of the protein was evaluated on the basis of the reconstructed water pathways by counting the number of iterations during which a water molecule was found within the contact distance $r_{HSPOT}=0.4$nm to a specific amino acid. The entire procedure was repeated 50 times for three different initial system configurations for a total of 150 independent SFM runs, each consisting of 500 iterations of $\Delta t_{ITER}=10$ps, amounting to a total of 750ns of accelerated SFM analysis. The three system configurations corresponded to the minimal, the maximal and an average RMSD value, derived from a prior 100ns MD simulation under conventional conditions without external forces or water removal (Figure S1). The contact frequencies from all SFM runs were normalized and merged for a holistic and statistically significant representation of probable water molecule entryways during influx to the active site cavity of CALB; amino acid positions corresponding to increased contact frequencies were defined as hotspot areas H1-H9 (Figure 3).

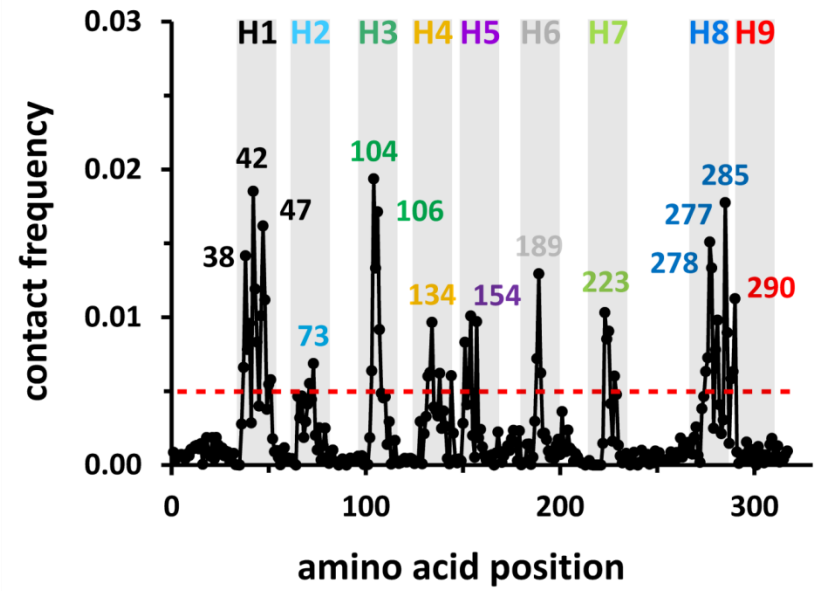


Figure 3: Contact frequency between influx water and position-specific amino acid are presented as the number of observed water contacts relative to all evaluated system conformations of the SFM analysis (all potential contacts). Hotspot areas H1-H9 were defined as peaks greater than a contact frequency cutoff value (red dotted line) and were arbitrarily colored for reasons of easy distinction and reference in Table 2 and Figure 4. Single positions that correspond to the highest values of the hotspot areas were labeled and defined as hotspot positions.

The individual hotspot positions were evaluated (Table 2) on the basis of their relative position in the CALB structure (Figure 4A) and by analyzing reconstructed water influx pathways (Figure 4B). Hotspot positions were predominantly found in close proximity to the triglyceride-water interface (Figure 1), with significant influx occurring through the substrate channel (Figure 4B), which corresponds to positions Val154 of H5, Ile189 of H6 and Leu278/Ile285 of H8. Positions Thr42/Ser47 of H1 and Gln106 of H3 coincide with the mutational strategy on CALB of Larsen et al.[17], who proposed and implicitly validated a water channel at these exact positions. Due to the highest relative contact frequency (Table 2), we refer to this channel as the primary water channel (PWC) in the following. Three new potential water entryways were identified, which display high relative contact frequencies and feature polar residues on the protein surface, at Val134 of H4, which we refer to as the secondary water channel (SWC) and at two positions Asp223 of H7 and Lys290 of H9 that are in immediate contact with the triglyceride interface, which we refer to as IWE1 and IWE2 respectively. Other potential hotspot positions, such as Pro38 of H1, Leu73 of H2 and Leu277 of H8 were estimated to be potential barriers, which obtain their heightened contact frequency due to the close proximity to one or more probable entryways, or as in the case of Trp104 of H3 their location next to the influx reference residue Ser105.

Table 2: Compiled hotspot areas H1-H9 obtained from 150 SFM runs analyzing water influx into CALB attached to a triglyceride-water interface depicted in Figure 1 – Data shown includes amino acid positions (pos.), amino acids (AA), amino acid properties in regards to charge and hydrophobicity/-philicity (property), the relative contact frequency calculated by SFM (rel. freq.), the distance (d_{act}) of residue Cα carbon atoms relative to the oxygen reference atom of Ser105 at the active site, the spatial positioning of the amino acid in the protein structure (positioning), a general assessment of the individual hotspot positions (assessment) and a channel definition (chan), whereby PWC is the water channel reported by Larsen et al.[17], SWC a potential secondary water channel and IWE1/IWE2 potential water entrances located at the protein-triglyceride interface that may contribute to the water influx via the substrate channel.

	pos.	AA	property	rel. freq.	d_{act} [Å]	positioning	assessment	chan
	38	**PRO**	hphob	0.71	5.7	periphery	adjacent to primary water channel	-
H1	**42**	**THR**	**hphil**	**1.00**	**7.8**	**periphery**	**primary water channel**	**PWC**
	47	**SER**	**hphil**	**0.85**	**8.5**	**periphery**	**primary water channel**	**PWC**
H2	**73**	**LEU**	hphob	0.27	10.8	surface	probable barrier, adjacent to Lys290	-
H3	**104**	**TRP**	hphob	0.96	4.8	active site	active site	-
	106	**GLN**	**hphil**	**0.98**	**4.8**	**active site**	**primary water channel**	**PWC**
H4	**134**	**ASP**	**neg**	**0.54**	**11.5**	**periphery**	**potential secondary water channel**	**SWC**
H5	**154**	**VAL**	hphob	0.56	12.0	periphery	entrance substrate channel	-
H6	**189**	**ILE**	hphob	0.96	9.4	surface	entrance substrate channel	-
H7	**223**	**ASP**	**neg**	**0.65**	**12.7**	**surface**	**potential interfacial water entrance**	**IWE1**
	277	**LEU**	hphob	0.77	8.7	periphery	probable barrier, adj. to Asp223 and Ser47	
H8	**278**	**LEU**	hphob	0.88	6.9	periphery	entrance substrate channel	
	285	**ILE**	hphob	0.94	11.3	periphery	entrance substrate channel	
H9	**290**	**LYS**	**pos**	**0.52**	**17.7**	**surface**	**potential interfacial water entrance**	**IWE2**

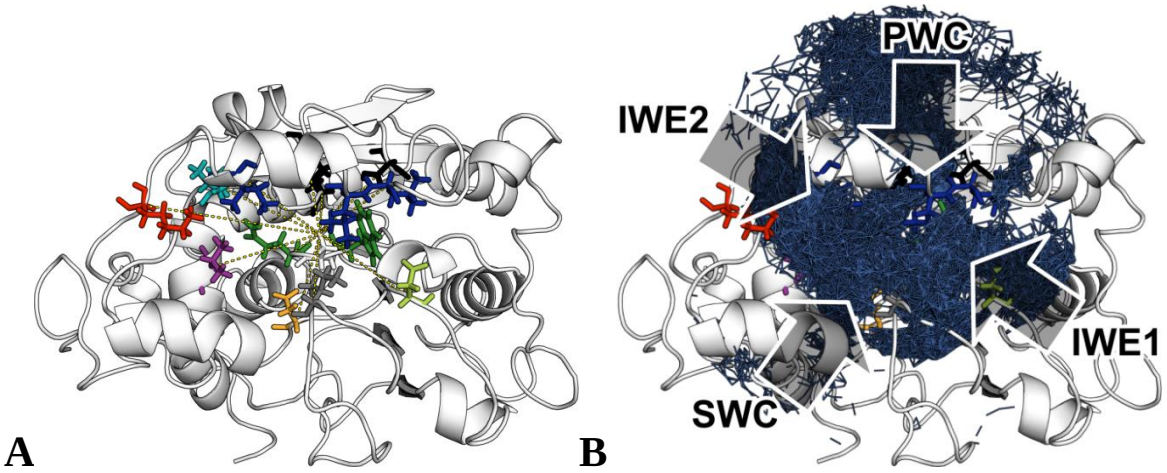


Figure 4: (A) Hotspot positions in the CALB protein structure with colors corresponding to H1-H9 of Figure 3, dotted lines point towards the active site Ser105, protein is represented in the bottom view, i.e. from the perspective of the triglyceride-water interface that CALB is attached to (see Figure 1) (B) Water influx pathways from five full SFM runs of 500 iterations length are superimposed onto the same depiction as Figure 4A to illustrate potential water entryways (arrows). Helices between the PWC and SWC channels and the triglyceride layer stop water "leakage" in between the protein and the interface, whereas IWE1 and IWE2 significantly contribute to the influx via the substrate channel, which is situated in the central part of the protein in both depictions.

PWC was shown to be the most frequented water entryway (Figure 5A and Table 2); even residue Gln46, defined as the outer entrance in the study of Larsen et al.[17], was resolved by SFM, albeit not as one of the highest contact frequency peaks (Figure 3). Based on the nomenclature of Uppenberg et al.[40], water influx via PWC occurred in the loop region between α5 and β2 (Figure 5A). Access to channel SWC (Figure 5B) occurred in the loop region between α5 and β5, which is located in relatively close proximity to the catalytically active Ser105 and is structurally slightly submerged relative to the protein surface, yet fully solvent accessible between α7 and the loop region between β6 and β7. Access to entrance IWE1 (Figure 5C), which is located directly at the boundaries between CALB and the triglyceride layer, occurred at the loop preceding α9, which was slightly denatured at its N-terminus due to the direct interaction with the triglyceride layer. IWE1 may thus contribute to the influx of water via the triglyceride layer and through the substrate channel, as water molecules were observed to "leak" in between the protein and the triglyceride layer (Figure 4B). Access to entrance IWE2 (Figure 5D) occurred in the loop region between β8 and β9 close to the C-terminus of CALB, which also directly borders the triglyceride-water interface and can thus potentially also contribute to the interfacial leakage.

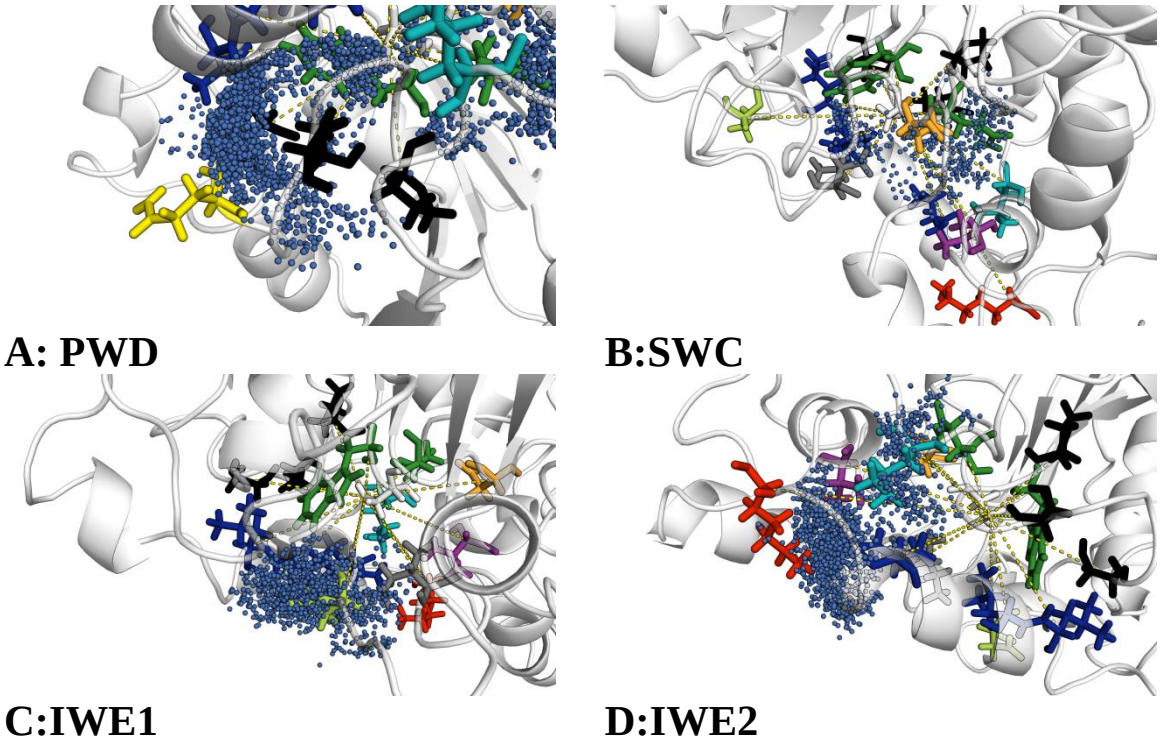


Figure 5: Water influx to CALB attached to a triglyceride-water interface (Figure 1), dotted lines point towards the catalytically active Ser105 (A) The outer PWC channel entrance is defined by Gln46 (yellow sticks) and the residues Thr42/Ser47 (middle black sticks) (B) Channel SWC is defined by Asp134 (orange) as a potential water entrance in close proximity to Ser105 (C) Entrance IWE1 (light green) is defined by Asp223 and is located at the triglyceride-water interface and may thus contribute to interfacial leakage of water (D) Entrance IWE2 (light green) is defined by Lys290 and is located at the triglyceride-water interface and may thus also contribute to interfacial leakage of water.

4.3.5 Discussion

4.3.5.1 Water Influx to the Active Site of CALB Modeled by SFM

Results from SFM suggest that the influx of water molecules into the active site of CALB is a more complex phenomenon than originally anticipated. It appears that the triglyceride layer, which CALB is attached to (Figure 1), is by no means a hydrophobic barrier that excludes water molecules from entering the enzyme via the substrate channel. On the contrary, significant interfacial leakage in between the protein and the triglyceride interface was observed (Figure 4B) that is comparable in significance to the water influx via the primary water channel PWC[17] and higher than the influx via the potential secondary channel SWC (Table 2). This was not expected, because during equilibration of the protein-triglyceride system, the protein immersed gradually into the layer (Figure S1B), which implies that interfacial leakage is unlikely to be caused by the model of protein

adsorption but due to the structure of triglyceride association. Interfacial leakage may be accredited to the fact that triglyceride molecules in the layer do not appear to associate in a manner that minimizes the interactions between the nonpolar triglyceride phase and the polar aqueous surrounding[47]. If the triglyceride-water interface model is accurate (Figure 1), the lamellar-like self-association structure of polar triglyceride moieties might indeed serve as a polar microenvironment that can facilitate an influx of water molecules through the interface and into the active site via the substrate channel. If the interfacial model proves to be correct, IWE1 and IWE2 are potentially such avenues for water entry at the protein-triglyceride-water interface. These considerations are subject to an ongoing study to clarify the involvement of triglyceride molecule association on interfacial water.

4.3.5.2 SFM Modeling Quality and Applicability

It was demonstrated that the solvent flux method (SFM) presented in this study is able to successfully identify solvent entryways in proteins, by revealing the water channel and key positions in CALB that may afford synthesis versus hydrolysis in non-aqueous media when appropriately mutated, as demonstrated by Larsen et al.[17]. Although SFM introduces artificial external forces to MD simulations by the periodic velocity adjustment of water molecules and may thus raise concerns about misleading biases, one must consider the limitations of conventional MD in sampling rare events such as the passage of a molecule through a channel, both in real and computational time. Despite the optimistic predictions that may be drawn from Moore's law[48] that the increase in computational power will eventually allow for the rigorous thermodynamic ensemble modeling of even the most complex molecular systems, present day reality lies in computational millisecond timescales at best. Exploring complex problems such as a comprehensive modeling of the solvent flux through an enzyme with significant statistics does therefore not seem feasible by conventional MD approaches, at least not in the foreseeable future. Increasing the sampling proficiency of molecular dynamics by introducing biases to the potential function to overcome energy barriers has therefore become common practice in the application of MD and has been applied in many variations, such as accelerated molecular dynamics (aMD)[49], replica-exchange molecular dynamics (REMD)[50], steered molecular dynamics (SMD)[51] or random acceleration molecular dynamics (RAMD)[52]. The multitude of

successful applications and insights attained by these methods is testament to their usefulness in resolving questions of a biophysical nature that would otherwise be inaccessible. SFM offers the means to rapidly and comprehensively sample solvent flux behavior throughout a protein and thus allows for an atomistic resolution of structural features and positions that are relevant to enzyme-solvent interactions. It is thereby unique in its approach of simultaneously modeling the accelerated influx of multiple solvent molecules that randomly come into contact with the protein surface from the solvent bulk, and to extract meaningful data in form of contact frequencies and compiled pathways of individual water molecules during influx. In this regard, other comparable methods are either less deliberate in overcoming specific energy barriers[49, 50], are restricted to single molecules and predefined geometric reaction coordinates[51], or are more suited to resolving the efflux of single substrates[52]. SFM holds the potential of exploring solvent-enzyme interactions beyond the hydration-related context[20, 21, 23, 25, 26, 53], which seems to be required, considering that identifying the role of residual water in facilitating enzyme activity in non-aqueous synthetic hydrolase applications has proven to be nontrivial[54] and studies of "dry" systems have shown that residual water is in fact not restricted to hydration of the protein, but instead partitions into the solvent under equilibrium conditions[55]. Understanding the behavior of this free water may thus prove particularly useful for hydrolase applications in non-aqueous environments[4], specifically to eliminate water or to improve solvent/substrate entryways for synthesis reactions by rational enzyme engineering. SFM may hereby aid in resolving the competition between synthesis and degradation, which remains a fundamental problem that is difficult to overcome without a deeper mechanistic understanding of a particular enzyme application.

4.3.6 Acknowledgements

We would like to thank the German Research Foundation (DFG) for financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart. We thank the high performance computing center Stuttgart (HLRS) for their support and for supplying the computational resources.

4.3.7 Supporting Information

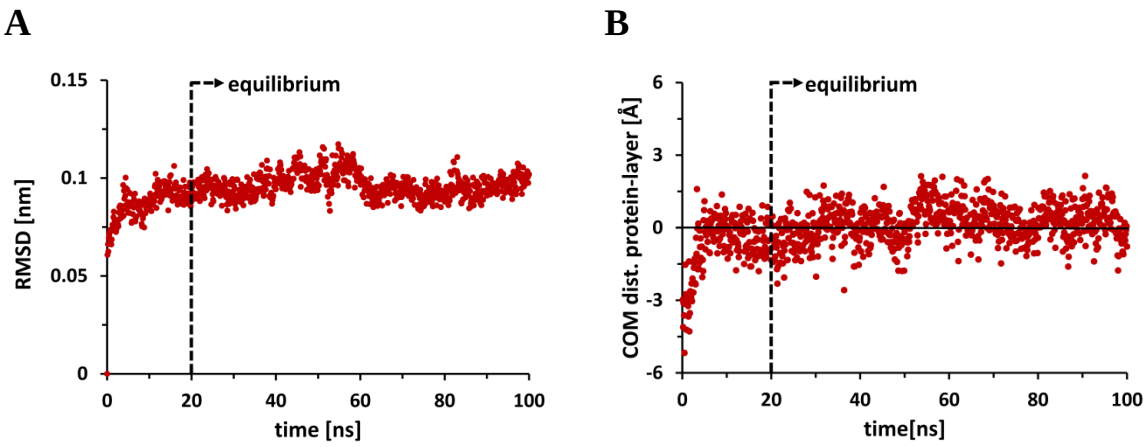


Figure S1: 100ns of MD simulation was performed on a pre-equilibrated system consisting of CALB attached to a triglyceride-water interface (see manuscript Figure 1) - A) The root mean square deviation (RMSD) of the protein was monitored while performing a least square fit of backbone atoms onto the starting structure. System conformations corresponding to the maximal, the minimal and an average RMSD value for t>50ns were selected as starting points for three different SFM runs, each of which were repeated 50 times for 500 iterations of Δt_{ITER}=10ps for a combined simulation time of 750ns / B) The equilibration of the aqueous protein-substrate interface was monitored by the deviations in the vertical component of the distance between the centers of mass (COM) of the protein and the triglyceride layer. The system equilibrated within 20ns.

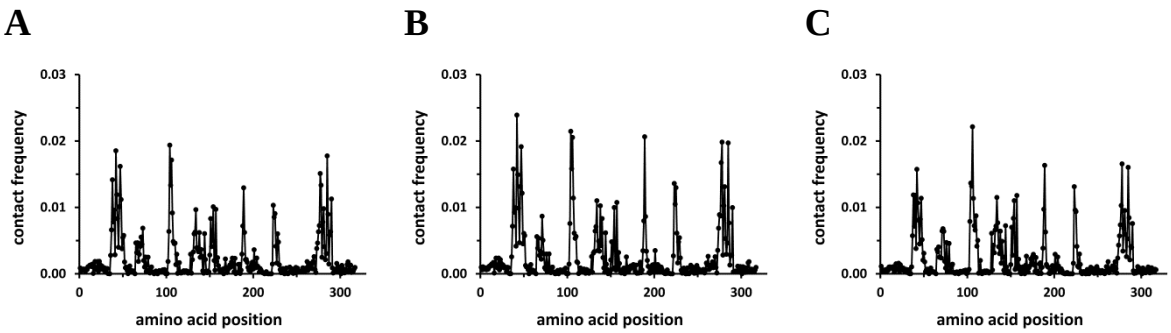


Figure S2: Compiled contact frequency plots of data from 50 independent SFM runs with different starting velocity distributions and starting structures corresponding to A) minimal, B) average, C) maximal RMSD values as derived from a 100ns equilibration simulation (see Figure S1) – largest s.d. between peaks was 0.0047, all data was combined to Figure 3A of manuscript.

Table S3: Definition of the major SFM parameters v_{MULT}, v_{CENT}, τ_{ITER} and position restraints that were varied to maximize the water removal rate $\Delta N_{WAT}/\Delta t_{ITER}$ at the active site while maintaining the structural integrity of the protein and the triglyceride layer.

param	unit	description
v_{MULT}	-	scalar factor by which the velocity vector magnitudes of all water molecules are multiplied that are found within a cutoff distance surrounding the active site reference atom
pos. res.	[kJ mol^{-1} nm^{-2}]	position restraints that are applied to the polypeptide backbone atoms to protect from denaturation and to the triglyceride molecules of the layer for maintaining its structural integrity
v_{CENT}	[%]	percentage of vector magnitude after rescaling with VMULT that is assigned to a new unit vector, which points from the respective water molecule towards the center reference atom in the active site, whereas 100% - VCENT is assigned to the inital vector's direction
τ_{ITER}	[ps]	temperature coupling constant of the Berendsen thermostat, which rescales the temperature back to a predefined value - the higher the coupling constant, the slower the exponential decay

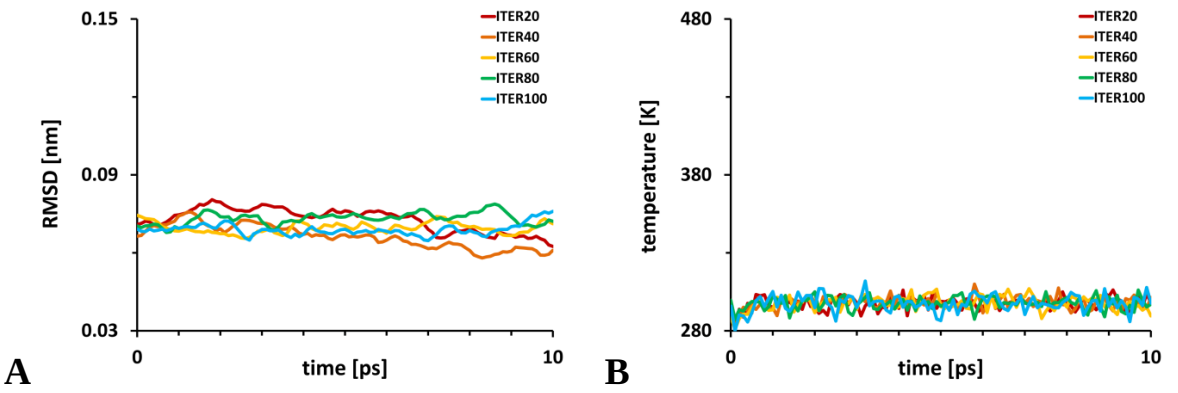


Figure S4: Analysis at NPT standard conditions with solvent flux parameters v_{MULT}=1.0, v_{CENT},=0, Δt_{ITER},=10ps, τ_{ITER}=0.1ps for iterations 20, 40, …, 100 - A) The root mean square deviation (RMSD) of the protein backbone after a given iteration, fitted onto the protein backbone of the initial equilibrated reference structure / B) After seeding fresh velocities, the temperature rapidly converged to the reference temperature of 298.15K under Berendsen coupling conditions.

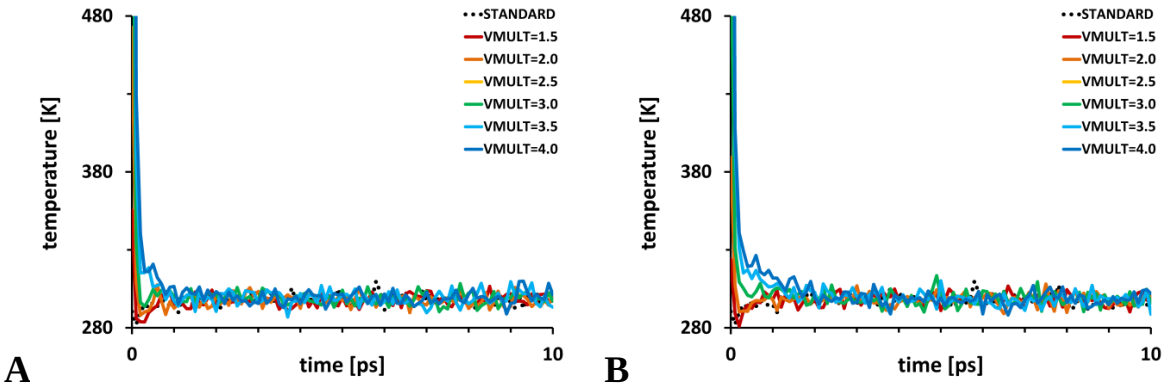


Figure S5: Temperature of the VEL group water molecules under NPT standard conditions, varying the factor v_{MULT}=1.5, 2.0, …, 4.0 by which the VEL group velocity vectors are multiplied, while keeping v_{CENT}=75%, τ_{ITER}=0.1ps and posre=100 kJ mol^{-1} nm^{-2} constant. Values are depicted for A) 10 iterations and B) 100 iterations of Δt_{ITER}=10ps. After velocity adjustment, temperatures rapidly converge to the 298.15K reference temperature under Berendsen coupling conditions. No consecutive heat build-up over the iterations is observed in any of the systems.

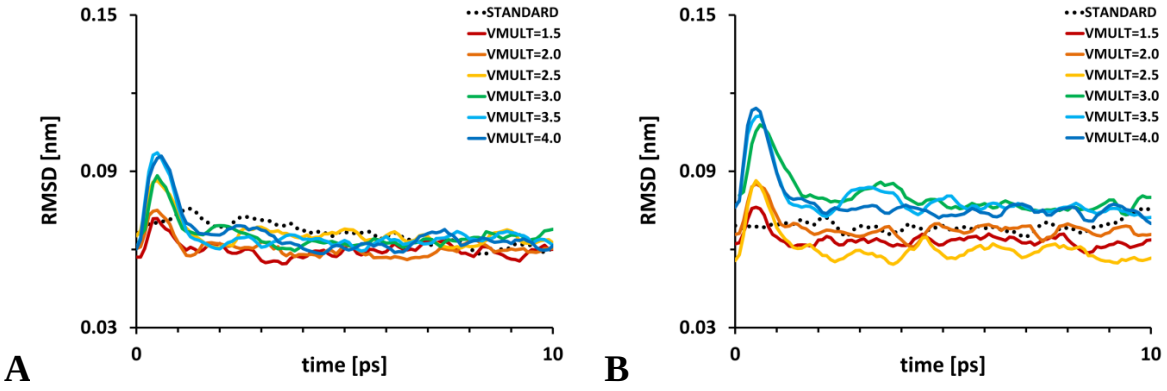


Figure S6: Root mean square deviation of the protein backbone under NPT standard conditions, varying the factor v_{MULT}=1.5, 2.0, …, 4.0 by which the VEL group velocity vectors are multiplied, while keeping v_{CENT}=75%, τ_{ITER}=0.1ps and posre=100 kJ mol^{-1} nm^{-2} constant. Values are depicted for A) 10 iterations and B) 100 iterations of Δt_{ITER}=10ps. After 100 iterations, deviations from the standard conditions (dotted line) are observed for v_{MULT}>2.5.

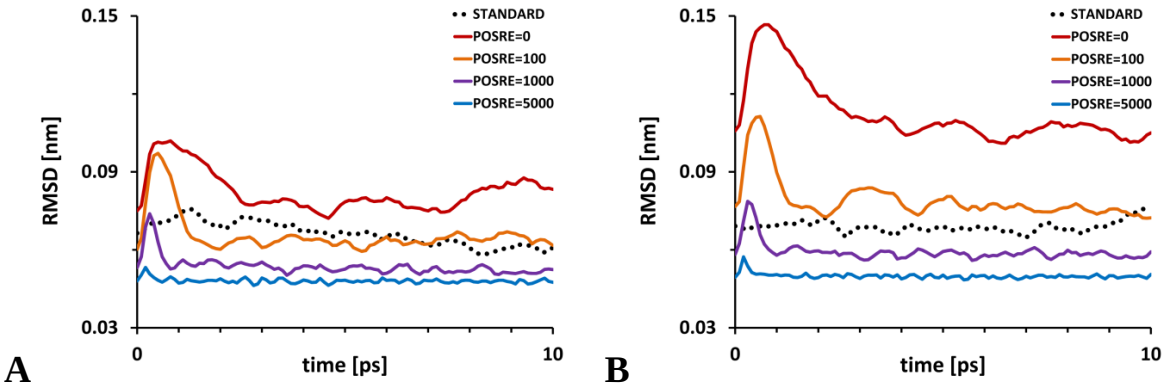


Figure S7: Root mean square deviation of the protein backbone under NPT standard conditions, varying backbone position restraints posre=0, 100, 1000, 5000 kJ mol^{-1} nm^{-2}, while keeping v_{MULT}=3.5, v_{CENT}=75% and τ_{ITER}=0.1ps constant. Values are depicted for A) 10 iterations and B) 100 iterations of Δt_{ITER}=10ps. After 100 iterations, significant structural distortions are observed relative to the standard conditions (dotted line) for position restraints <1000 kJ mol^{-1} nm^{-2}.

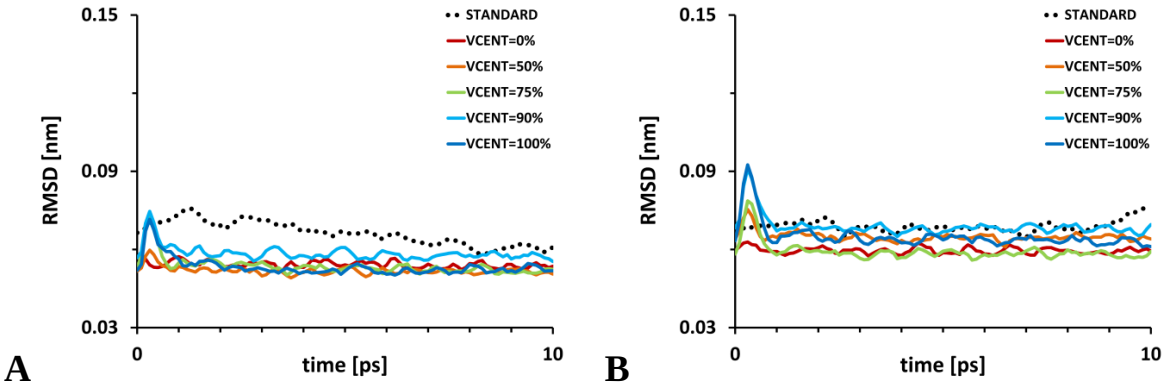


Figure S8: Root mean square deviation of the protein backbone under NPT standard conditions, varying the percentage v_{CENT}=0, 50, 75, 90, 100% by which the velocity vectors are reoriented towards the active site reference atom, while keeping v_{MULT}=3.5, τ_{ITER}=0.1ps and posre=1000 kJ mol^{-1} nm^{-2} constant. Values are depicted for A) 10 iterations and B) 100 iterations of Δt_{ITER}=10ps. After 100 iterations, only slight structural distortions relative to standard conditions (dotted line) are briefly observed at the onset of the velocity vector adjustment.

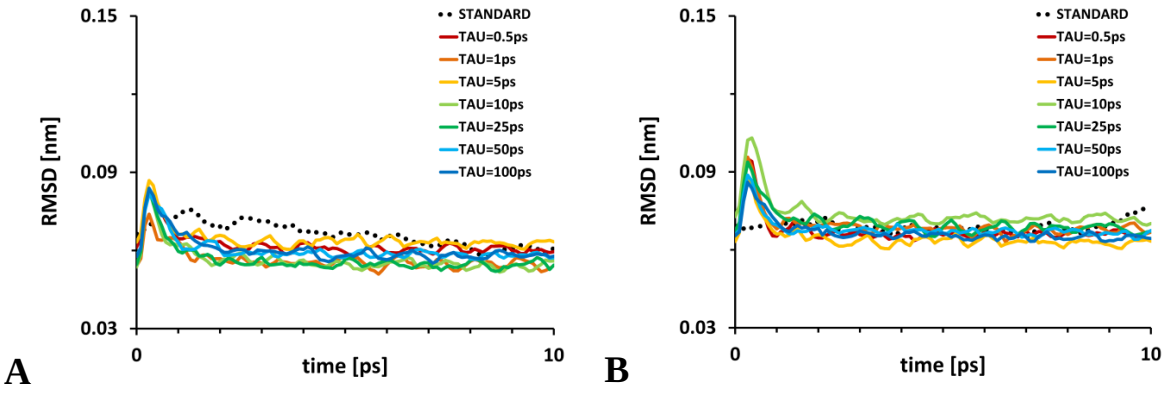


Figure S9: Root mean square deviation of the protein backbone under NPT standard conditions, varying the temperature coupling constant τ_{ITER}=0.5, 1, 5, ..., 100ps applied to the VEL group, while keeping v_{MULT}=3.5, v_{CENT}=90% and posre=1000 kJ mol^{-1} nm^{-2} constant. Values are depicted for A) 10 iterations and B) 100 iterations of Δt_{ITER}=10ps. After 100 iterations, only slight structural distortions relative to standard conditions (dotted line) are briefly observed at the onset of the velocity vector adjustment.

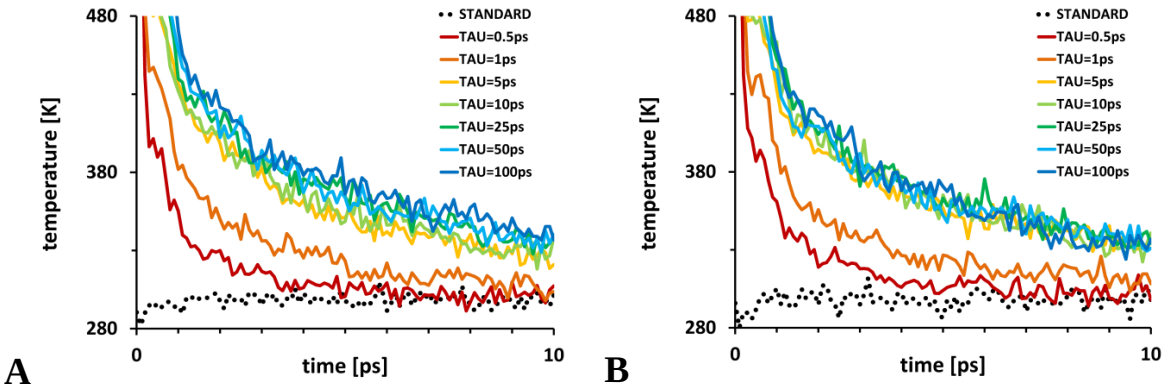


Figure S10: Temperature of the VEL group water molecules under NPT standard conditions, varying the temperature coupling constant τ_{ITER}=0.5, 1, 5, ..., 100ps applied to the VEL group, while keeping v_{MULT}=3.5, v_{CENT}=90% and posre=1000 kJ mol^{-1} nm^{-2} constant.. Values are depicted for A) 10 iterations and B) 100 iterations of Δt_{ITER}=10ps. After velocity adjustment, temperatures as expected converge more slowly for larger τ_{ITER} to the 298.15K reference temperature under Berendsen coupling conditions. However, no consecutive heat build-up over the iterations is observed in any of the systems.

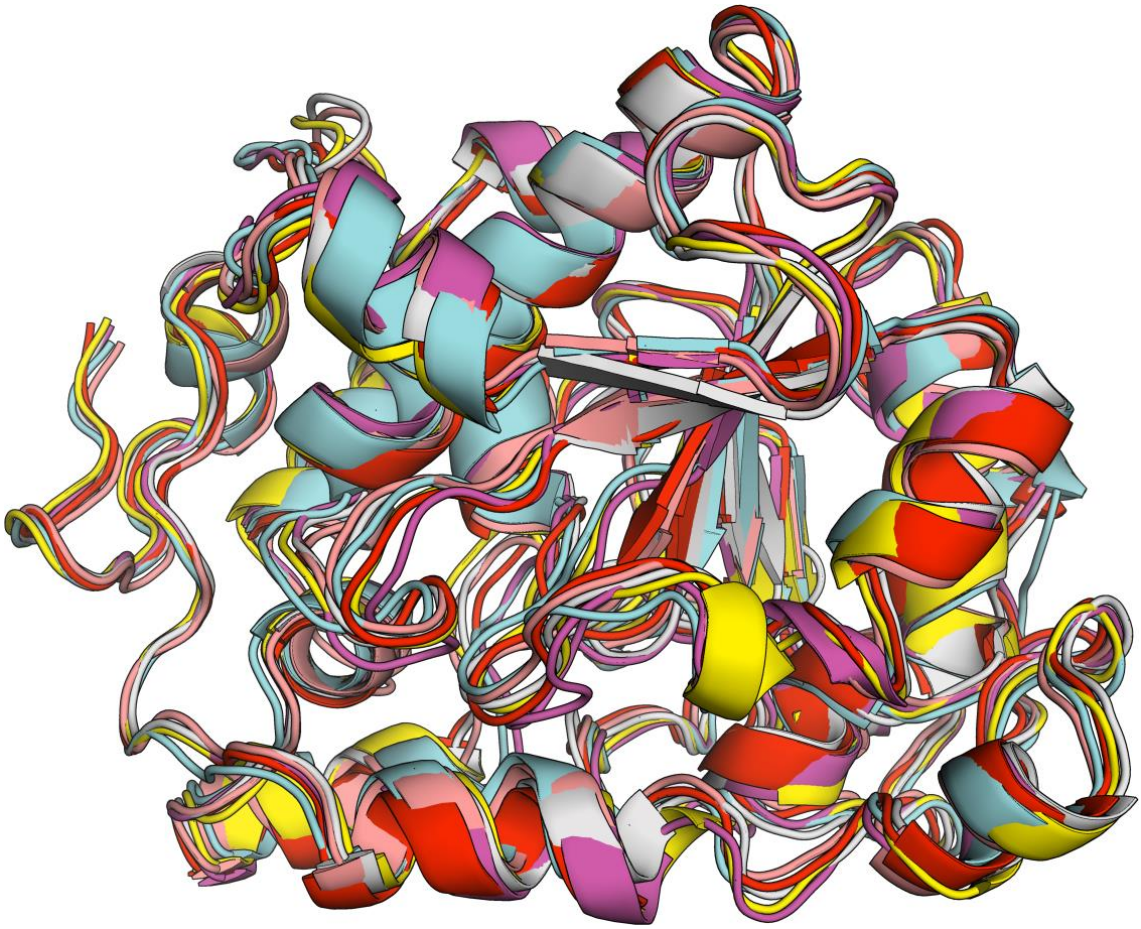


Figure S11: CALB after 100 SFM optimization iterations of v_{MULT} (yellow), v_{CENT} (pink), τ_{ITER} (blue) and position restraints (grey), superimposed with the starting structure (red).

4.3.8 References

(1) Panda, T.; Gowrishankar, B. S. Production and Applications of Esterases. Appl. Microbiol. Biotechnol. 2005, 67, 160-169.

(2) Gupta, M. N.; Roy, I. Enzymes in Organic Media. Eur. J. Biochem. 2004, 271, 2575-2583.

(3) Zaks, A.; Klibanov, A. M. Enzymatic Catalysis in Nonaqueous Solvents. J. Biol. Chem. 1988, 263, 3194-3201.

(4) Hudson, E. P.; Eppler, R. K.; Clark, D. S. Biocatalysis In Semi-Aqueous and Nearly Anhydrous Conditions. Curr. Opin. Biotechnol. 2005, 16, 637-643.

(5) van Rantwijk, F.; Sheldon, R. A. Biocatalysis in Ionic Liquids. Chem. Rev. 2007, 107, 2757-2785.

(6) Schmid, A.; Dordick, J. S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. Industrial Biocatalysis Today and Tomorrow. Nature 2001, 409, 258-268.

(7) Distel, K. A.; Zhu, G. Y.; Wang, P. Biocatalysis Using an Organic-Soluble Enzyme for the Preparation of Poly(Lactic Acid) in Organic Solvents. Bioresour. Technol. 2005, 96, 617-623.

(8) Mine, Y.; Zhang, L.; Fukunaga, K.; Sugimura, Y. Enhancement of Enzyme Activity and Enantioselectivity by Cyclopentyl Methyl Ether in the Transesterification Catalyzed by Pseudomonas Cepacia Lipase Co-Lyophilized with Cyclodextrins. Biotechnol. Lett. 2005, 27, 383-388.

(9) Jiang, Y.; Dalton, H. Chemical Modification of the Hydroxylase of Soluble Methane Monooxygenase Gives One Form of the Protein with Significantly Increased Thermostability and Another That Functions Well in Organic-Solvents. Biochim. Biophys. Acta-Gen. Subj. 1994, 1201, 76-84.

(10) Govardhan, C. P. Crosslinking of Enzymes for Improved Stability and Performance. Curr. Opin. Biotechnol. 1999, 10, 331-335.

(11) Lindsay, J. P.; Clark, D. S.; Dordick, J. S. Combinatorial Formulation of Biocatalyst Preparations for Increased Activity in Organic Solvents: Salt Activation of Penicillin Amidase. Biotechnol. Bioeng. 2004, 85, 553-560.

(12) Persson, M.; Wehtje, E.; Adlercreutz, P. Factors Governing the Activity of Lyophilised and Immobilised Lipase Preparations in Organic Solvents. Chembiochem 2002, 3, 566-571.

(13) Wescott, C. R.; Klibanov, A. M. The Solvent Dependence of Enzyme Specificity. Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol. 1994, 1206, 1-9.

(14) Laane, C. Medium-Engineering For Bio-Organic Synthesis. Biocatal. Biotransform. 1987, 1, 17-22.

(15) Bornscheuer, U. T.; Pohl, M. Improved Biocatalysts by Directed Evolution and Rational Protein Design. Curr. Opin. Chem. Biol. 2001, 5, 137-143.

(16) Cirino, P. C.; Arnold, F. H. Protein Engineering of Oxygenases for Biocatalysis. Curr. Opin. Chem. Biol. 2002, 6, 130-135.

(17) Wittrup Larsen , M.; Zielinska, D. F.; Martinelle, M.; Hidalgo, A.; Jensen, L. J.; Bornscheuer, U. T.; Hult, K. Suppression of Water as a Nucleophile in Candida antarctica Lipase B Catalysis. ChemBioChem 2000, 11, 796-801.

(18) Lousa, D.; Baptista, A. M.; Soares, C. M. A Molecular Perspective on Nonaqueous Biocatalysis: Contributions from Simulation Studies. Phys. Chem. Chem. Phys. 2013, 15, 13723-13736.

(19) Tejo, B.; Salleh, A.; Pleiss, J. Structure and Dynamics of Candida Rugosa Lipase: the Role of Organic Solvent. J. Mol. Model. 2004, 10, 358-366.

(20) Soares, C. M.; Teixeira, V. H.; Baptista, A. M. Protein Structure and Dynamics in Nonaqueous Solvents: Insights from Molecular Dynamics Simulation Studies. Biophys. J. 2003, 84, 1628-1641.

(21) Norin, M.; Haeffner, F.; Hult, K.; Edholm, O. Molecular Dynamics Simulations of an Enzyme Surrounded by Vacuum, Water, or a Hydrophobic Solvent. Biophys. J. 1994, 67, 548-559.

(22) Rehm, S.; Trodler, P.; Pleiss, J. Solvent-Induced Lid Opening in Lipases: A Molecular Dynamics Study. Protein Sci. 2010, 19, 2122-2130.

(23) Micaelo, N. M.; Soares, C. M. Modeling Hydration Mechanisms of Enzymes in Nonpolar and Polar Organic Solvents. FEBS J. 2007, 274, 2424-2436.

(24) Trodler, P.; Schmid, R. D.; Pleiss, J. Modeling of Solvent-Dependent Conformational Transitions in Burkholderia Cepacia Lipase. BMC Struct. Biol. 2009, 9.

(25) Klähn, M.; Lim, G. S.; Seduraman, A.; Wu, P. On the Different Roles of Anions and Cations in the Solvation of Enzymes in Ionic Liquids. Phys. Chem. Chem. Phys. 2011, 13, 1649-1662.

(26) Micaelo, N. M.; Baptista, A. n. M.; Soares, C. u. M. Parametrization of 1-Butyl-3-methylimidazolium Hexafluorophosphate/Nitrate Ionic Liquid for the GROMOS Force Field. J. Phys. Chem. B 2006, 110, 14444-14451.

(27) Affleck, R.; Xu, Z. F.; Suzawa, V.; Focht, K.; Clark, D. S.; Dordick, J. S. Enzymatic Catalysis and Dynamics in Low-Water Environments. PNAS 1992, 89, 1100-1104.

(28) Branco, R. J. F.; Graber, M.; Denis, V.; Pleiss, J. Molecular Mechanism of the Hydration of Candida antarctica Lipase B in the Gas Phase: Water Adsorption Isotherms and Molecular Dynamics Simulations. ChemBioChem 2009, 10, 2913-2919.

(29) Wedberg, R.; Abildskov, J.; Peters, G. H. Protein Dynamics in Organic Media at Varying Water Activity Studied by Molecular Dynamics Simulation. J. Phys. Chem. B 2012, 116, 2575-2585.

(30) Valivety, R. H.; J. Halling, P.; Macrae, A. R. Reaction Rate with Suspended Lipase Catalyst Shows Similar Dependence on Water Activity in Different Organic Solvents. Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol. 1992, 1118, 218-222.

(31) Bös, F.; Pleiss, J. Multiple Molecular Dynamics Simulations of TEM beta-Lactamase: Dynamics and Water Binding of the omega-Loop. Biophys. J. 2009, 97, 2550-2558.

(32) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J. Chem. Theory Comput. 2008, 4, 435-447.

(33) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. J. Comput. Chem. 2005, 26, 1701-1718.

(34) Wensink, E. J. W.; Hoffmann, A. C.; van Maaren, P. J.; van der Spoel, D. Dynamic Properties of Water/Alcohol Mixtures Studied by Computer Simulation. J. Chem. Phys. 2003, 119, 7308-7317.

(35) Marrink, S. J.; Mark, A. E. Effect of Undulations on Surface Tension in Simulated Bilayers. J. Phys. Chem. B 2001, 105, 6122-6127.

(36) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of n-Alkanes. J. Comput. Phys. 1977, 23, 327-341.

(37) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. J. Chem. Phys. 1995, 103, 8577-8593.

(38) York, D. M.; Darden, T. A.; Pedersen, L. G. The Effect of Long-Range Electrostatic Interactions in Simulations of Macromolecular Crystals - a Comparison of the Ewald and Truncated List Methods. J. Chem. Phys. 1993, 99, 8345-8348.

(39) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. Development And Testing Of The OPLS All-Atom Force Field On Conformational Eenergetics And Properties Of Organic Liquids. J. Am. Chem. Soc. 1996, 118, 11225-11236.

(40) Uppenberg, J.; Hansen, M. T.; Patkar, S.; Jones, T. A. The Sequence, Crystal Structure Determination and Refinement of Two Crystal Forms of Lipase B from Candida Antarctica. Structure 1994, 2, 293-308.

(41) Berger, O.; Edholm, O.; Jähnig, F. Molecular Dynamics Simulations of a Fluid Bilayer of Dipalmitoylphosphatidylcholine at Full Hydration, Constant Pressure, and Constant Temperature. Biophys. J. 1997, 72, 2002-2013.

(42) Neale, C.; Pomès, R. Combination rules for united-atom lipids and OPLS-AA proteins, this unpublished document is available online from the Pomès lab at http://www.pomeslab.com/files/lipidCombinationRules.pdf.

(43) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. J. Phys. Chem. 1987, 91, 6269-6271.

(44) Gruber, C. C.; Pleiss, J. Lipase B from Candida Antarctica Binds to Hydrophobic Substrate-Water Interfaces via Hydrophobic Anchors Surrounding the Active Site Entrance. J. Mol. Catal. B-Enzym. 2012, 84, 48-54.

(45) Jurado, E.; Camacho, F.; Luzón, G.; Fernández-Serrano, M.; García-Román, M. Kinetic Model for the Enzymatic Hydrolysis of Tributyrin in O/W Emulsions. Chem. Eng. Sci. 2006, 61, 5010-5020.

(46) Sum, A. K.; Biddy, M. J.; de Pablo, J. J.; Tupy, M. J. Predictive Molecular Model for the Thermodynamic and Transport Properties of Triacylglycerols. J. Phys. Chem. B 2003, 107, 14443-14451.

(47) Moore, G. E. Cramming More Components onto Integrated Circuits (Reprinted from Electronics, pg 114-117, April 19, 1965). Proc. IEEE 1998, 86, 82-85.

(48) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. J. Chem. Phys. 2004, 120, 11919-11929.

(49) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. Chem. Phys. Lett. 1999, 314, 141-151.

(50) Lu, H.; Isralewitz, B.; Krammer, A.; Vogel, V.; Schulten, K. Unfolding of Titin Immunoglobulin Domains by Steered Molecular Dynamics Simulation. Biophys. J. 1998, 75, 662-671.

(51) Lüdemann, S. K.; Lounnas, V. r.; Wade, R. C. How Do Substrates Enter and Products Exit the Buried Active Site of Cytochrome P450cam? 1. Random Expulsion Molecular Dynamics Investigation of Ligand Access Channels and Mechanisms. J. Mol. Biol. 2000, 303, 797-811.

(52) Trodler, P.; Pleiss, J. Modeling Structure and Flexibility of Candida Antarctica Lipase B in Organic Solvents. BMC Struct. Biol. 2008, 8.

(53) Torres, S.; Castro, G. R. Non-Aqueous Biocatalysis in Homogeneous Solvent Systems. Food Technol. Biotechnol. 2004, 42, 271-277.

(54) Halling, P. J. What Can We Learn by Studying Enzymes in Non-Aqueous Media? Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci. 2004, 359, 1287-1297.

5 Bibliography

Abdel-Mottaleb, MMA *et al.*, 2010 In Vitro Drug Release Mechanism from Lipid Nanocapsules (LNC). *Int. J. Pharm., 390,* 208-213.

Aboofazeli, R *et al.*, 2000 Particle Size Analysis of Concentrated Phospholipid Microemulsions: II. Photon Correlation Spectroscopy. *AAPS Pharmsci, 2,* 1–10.

Acevedo, NC; Marangoni, AG, 2010 Characterization of the Nanoscale in Triacylglycerol Crystal Networks. *Cryst. Growth Des., 10,* 3327-3333.

Adcock, SA; McCammon, JA, 2006 Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev., 106,* 1589-1615.

Affleck, R *et al.*, 1992 Enzymatic Catalysis and Dynamics in Low-Water Environments. *PNAS, 89,* 1100-1104.

Allen M.P., TDJ, 1987 Computer Simulations of Liquids *Oxford University Press Inc*.

Amar, I *et al.*, 2003 Solubilization Patterns of Lutein and Lutein Esters in Food Grade Nonionic Microemulsions. *J. Agric. Food Chem., 51,* 4775-4781.

Amidon, G *et al.*, 1995 A Theoretical Basis for a Biopharmaceutic Drug Classification: The Correlation of in Vitro Drug Product Dissolution and in Vivo Bioavailability. *Pharm. Res., 12,* 413-420.

Anderson, BD *et al.*, 1980 Solubility of Polar Organic Solutes in Nonaqueous Systems: Role of Specific Interactions. *Pharm. Sci., 69,* 676-680.

Attama, AA *et al.*, 2009 Sustained Release and Permeation of Timolol from Surface-Modified Solid Lipid Nanoparticles through Bioengineered Human Cornea. *Curr. Eye Res., 34,* 698-705.

Bako, I *et al.*, 2008 Water-Methanol Mixtures: Topology of Hydrogen Bonded Network. *Phys. Chem. Chem. Phys., 10,* 5004-5011.

Benson, SP; Pleiss, J, 2013 Incomplete Mixing Versus Clathrate-Like Structures: A Molecular View on Hydrophobicity in Methanol-Water Mixtures. *J. Mol. Model., 19,* 3427-3436.

Benson, SP; Pleiss, J, 2014 Molecular Dynamics Simulations of Self-Emulsifying Drug Delivery Systems (SEDDS): Influence of Excipients on Droplet Nanostructure and Drug Localization. *Langmuir (in press).*

Berendsen, HJC *et al.*, 1987 The Missing Term in Effective Pair Potentials. *J. Phys. Chem., 91,* 6269-6271.

Berendsen, HJC *et al.*, 1984 Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.*, *81*, 3684-3690.

Berendsen, HJC *et al.*, 1981 Intermolecular Forces (Jerusalem Symposia) *Springer*.

Berger, O *et al.*, 1997 Molecular Dynamics Simulations of a Fluid Bilayer of Dipalmitoylphosphatidylcholine at Full Hydration, Constant Pressure, and Constant Temperature. *Biophys. J.*, *72*, 2002-2013.

Blumberg, RL *et al.*, 1984 Connectivity of Hydrogen-Bonds in Liquid Water. *J Chem Phys*, *80*, 5230-5241.

Bornscheuer, UT; Pohl, M, **2001** Improved Biocatalysts by Directed Evolution and Rational Protein Design. *Curr. Opin. Chem. Biol.*, *5*, 137-143.

Bös, F; Pleiss, J, **2009** Multiple Molecular Dynamics Simulations of TEM beta-Lactamase: Dynamics and Water Binding of the omega-Loop. *Biophys. J.*, *97*, 2550-2558.

Bowron, DT *et al.*, 1998 Hydrophobic Hydration and the Formation of a Clathrate Hydrate. *Phys. Rev. Lett.*, *81*, 4164-4167.

Branco, RJF *et al.*, 2009 Molecular Mechanism of the Hydration of Candida antarctica Lipase B in the Gas Phase: Water Adsorption Isotherms and Molecular Dynamics Simulations. *ChemBioChem*, *10*, 2913-2919.

Cao, Y *et al.*, 2004 Predictive Relationships for the Effects of Triglyceride Ester Concentration and Water Uptake on Solubility and Partitioning of Small Molecules into Lipid Vehicles. *J. Pharm. Sci.*, *93*, 2768-2779.

Chandler, D, **2002** Hydrophobicity: Two Faces of Water. *Nature*, *417*, 491-491.

Chandler, D, **2005** Interfaces and the Driving Force of Hydrophobic Assembly. *Nature*, *437*, 640-647.

Chung, H *et al.*, 2001 Stability of the Oil-in-Water Type Triacylglycerol Emulsions. *Biotechnol. Bioprocess Eng.*, *6*, 284-288.

Cieplak, P *et al.*, 2009 Polarization Effects in Molecular Mechanical Force Fields. *J Phys Condens Matter*, *21*, 333102.

Cirino, PC; Arnold, FH, **2002** Protein Engineering of Oxygenases for Biocatalysis. *Curr. Opin. Chem. Biol.*, *6*, 130-135.

Cornell, WD *et al.*, 1995 A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.*, *117*, 5179-5197.

Darden, T *et al.*, 1993 Particle Mesh Ewald: An Nlog(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.*, *98*, 10089-10092.

Deckelbaum, RJ *et al.*, 1990 Medium-Chain Versus Long-Chain Triacylglycerol Emulsion Hydrolysis by Lipoprotein-Lipase and Hepatic Lipase - Implications for the Mechanisms of Lipase Action. *Biochemistry, 29,* 1136-1142.

Derlacki, ZJ *et al.*, 1985 Diffusion Coefficients of Methanol and Water and the Mutual Diffusion Coefficient in Methanol-Water Solutions at 278 and 298 K. *J. Phys. Chem., 89,* 5318-5322.

Devanand, K; Selser, JC, **1990** Polyethylene Oxide Does not Necessarily Aggregate in Water. *Nature, 343,* 739-741.

Distel, KA *et al.*, 2005 Biocatalysis Using an Organic-Soluble Enzyme for the Preparation of Poly(Lactic Acid) in Organic Solvents. *Bioresour. Technol., 96,* 617-623.

Dixit, S *et al.*, 2002 Molecular Segregation Observed in a Concentrated Alcohol-Water Solution. *Nature, 416,* 829-832.

Dixit, S *et al.*, 2000 Hydration of Methanol in Aqueous Solutions: a Raman Spectroscopic Study. *J. Phys. Condens. Matter, 12,* 323-328.

Dixit, S *et al.*, 2002 Water Structure and Solute Association In Dilute Aqueous Methanol. *Europhys. Lett., 59,* 377-383.

Dorsey, EEc, **1940** Properties of Ordinary Water-Substances *Reinhold.*

Dougan, L *et al.*, 2004 Methanol-Water Solutions: a Bi-Percolating Liquid Mixture. *J. Chem. Phys., 121,* 6456-6462.

Dzugutov, M, **1996** A Universal Scaling Law for Atomic Diffusion in Condensed Matter. *Nature, 381,* 137-139.

Dzugutov, M, **2001** A Universal Scaling Law for Atomic Diffusion in Condensed Matter. *Nature, 411,* 720-720.

Essmann, U *et al.*, 1995 A Smooth Particle Mesh Ewald Method. *J. Chem. Phys., 103,* 8577-8593.

Finney, JL *et al.*, 2003 Molecular and Mesoscale Structures in Hydrophobically Driven Aqueous Solutions. *Biophys. Chem., 105,* 391-409.

Frank, HS; Evans, MW, **1945** Free Volume and Entropy in Condensed Systems III. Entropy in Binary Liquid Mixtures; Partial Molal Entropy in Dilute Solutions; Structure and Thermodynamics in Aqueous Electrolytes. *J. Chem. Phys., 13,* 507-532.

Galli, G, **2007** Dissecting Hydrophobicity. *Proc. Natl. Acad. Sci. USA, 104,* 2557-2558.

Garti, N *et al.*, 2003 Interfacial Modification and Structural Transitions Induced by Guest Molecules Solubilized in U-Type Nonionic Microemulsions. *J. Dispersion Sci. Technol., 24,* 397-410.

Garti, N *et al.*, 2006 Improved Solubilization of Celecoxib in U-Type Nonionic Microemulsions and their Structural Transitions with Progressive Aqueous Dilution. *J. Colloid Sci.*, *299*, 352-365.

Geiger, A *et al.*, 1979 Aspects of the Percolation Process for Hydrogen-Bond Networks in Water. *J. Chem. Phys.*, *70*, 4185-4193.

Gershanik, T *et al.*, 1998 Interaction of a Self-Emulsifying Lipid Drug Delivery System with the Everted Rat Intestinal Mucosa as a Function of Droplet Size and Surface Charge. *Pharm. Res.*, *15*, 863-869.

Govardhan, CP, 1999 Crosslinking of Enzymes for Improved Stability and Performance. *Curr. Opin. Biotechnol.*, *10*, 331-335.

Gruber, CC; Pleiss, J, 2012 Lipase B from Candida Antarctica Binds to Hydrophobic Substrate-Water Interfaces via Hydrophobic Anchors Surrounding the Active Site Entrance. *J. Mol. Catal. B-Enzym.*, *84*, 48-54.

Guo, JH *et al.*, 2003 Molecular Structure of Alcohol-Water Mixtures. *Phys. Rev. Lett.*, *91*.

Gupta, MN; Roy, I, 2004 Enzymes in Organic Media. *Eur. J. Biochem.*, *271*, 2575-2583.

Halle, B; Wennerstrom, H, 1981 Interpretation of Magnetic-Resonance Data from Water Nuclei in Heterogeneous Systems. *J. Chem. Phys.*, *75*, 1928-1943.

Halling, PJ, 2004 What Can We Learn by Studying Enzymes in Non-Aqueous Media? *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci.*, *359*, 1287-1297.

Hamelberg, D *et al.*, 2004 Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.*, *120*, 11919-11929.

Haughney, M *et al.*, 1987 Molecular-Dynamics Simulations of Liquid Methanol. *J. Phys. Chem.*, *91*, 4934-4940.

Hess, B *et al.*, 1997 LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.*, *18*, 1463-1472.

Hess, B *et al.*, 2008 GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.*, *4*, 435-447.

Hockney, RW *et al.*, 1974 Quiet High-Resolution Computer Models of a Plasma. *J. Comput. Chem.*, *14*, 148-158.

Hoover, WG, 1985 Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A.*, *31*, 1695-1697.

Horta, BAC *et al.*, 2011 New Interaction Parameters for Oxygen Compounds in the GROMOS Force Field: Improved Pure-Liquid and Solvation Properties for Alcohols, Ethers, Aldehydes, Ketones, Carboxylic Acids, and Esters. *J. Chem. Theory Comput.*, *7*, 1016-1031.

Hudson, EP *et al.*, 2005 Biocatalysis In Semi-Aqueous and Nearly Anhydrous Conditions. *Curr. Opin. Biotechnol.*, *16*, 637-643.

Hummer, G *et al.*, 2000 New Perspectives on Hydrophobic Effects. *Chem. Phys.*, *258*, 349-370.

Israelachvili, JN, Intermolecular and Surface Forces, 3rd ed.; *Elsevier: USA*.

Jiang, Y; Dalton, H, **1994** Chemical Modification of the Hydroxylase of Soluble Methane Monooxygenase Gives One Form of the Protein with Significantly Increased Thermostability and Another That Functions Well in Organic-Solvents. *Biochim. Biophys. Acta-Gen. Subj.*, *1201*, 76-84.

Jones, JE, **1924** On the Determination of Molecular Fields. II. From the Equation of State of a Gas. *Proc. R. Soc. London*, *106*, 463-477.

Jorgensen, WL, **1982** Revised TIPS for Simulations of Liquid Water and Aqueous Solutions. *J. Chem. Phys.*, *77*, 4156-4163.

Jorgensen, WL; Buckner, JK, **1987** Use of Statistical Perturbation-Theory for Computing Solvent Effects on Molecular-Conformation - Butane in Water. *J. Phys. Chem.*, *91*, 6083-6085.

Jorgensen, WL *et al.*, 1983 Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.*, *79*, 926-935.

Jorgensen, WL; Madura, JD, **1983** Quantum and Statistical Studies of Liquids .25. Solvation and Conformation of Methanol in Water. *J. Am. Chem. Soc.*, *105*, 1407-1413.

Jorgensen, WL *et al.*, 1996 Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.*, *118*, 11225-11236.

Jurado, E *et al.*, 2006 Kinetic Model for the Enzymatic Hydrolysis of Tributyrin in O/W Emulsions. *Chem. Eng. Sci.*, *61*, 5010-5020.

Kasimova, AO *et al.*, 2012 Validation of a Novel Molecular Dynamics Simulation Approach for Lipophilic Drug Incorporation into Polymer Micelles. *116*, 4338-4345.

Kauzmann, W, **1959** Some Factors in the Interpretation of Protein Denaturation. *Adv. Protein Chem.*, *14*, 1-63.

Kincaid, RH; Scheraga, HA, **1982** Acceleration of Convergence in Monte-Carlo Simulations of Aqueous-Solutions Using the Metropolis Algorithm - Hydrophobic Hydration of Methane. *J. Comput. Chem.*, *3*, 525-547.

King, DT *et al.*, **2011** Using Molecular Dynamics to Study Liquid Phase Behavior: Simulations of the Ternary Sodium Laurate/Sodium Oleate/Water System. *Langmuir*, *27*, 11381-11393.

Klähn, M *et al.*, **2011** On the Different Roles of Anions and Cations in the Solvation of Enzymes in Ionic Liquids. *Phys. Chem. Chem. Phys.*, *13*, 1649-1662.

Kohli, K *et al.*, **2010** Self-Emulsifying Drug Delivery Systems: An Approach to Enhance Oral Bioavailability. *Drug Discov. Today.*, *15*, 958-965.

Kuzmanic, A *et al.*, **2014** X-ray Refinement Significantly Underestimates the Level of Microscopic Heterogeneity in Biomolecular Crystals. *Nat. Commun.*, *5*.

Laage, D *et al.*, **2009** Why Water Reorientation Slows without Iceberg Formation Around Hydrophobic Solutes. *J. Phys. Chem. B*, *113*, 2428-2435.

Laane, C, **1987** Medium-Engineering For Bio-Organic Synthesis. *Biocatal. Biotransform.*, *1*, 17-22.

Lankau, T; Konrad, O, **2007** Hydrophobic Solvation: Aqueous Methane Solutions. *J. Chem. Educ.*, *84*, 864.

Larson, RG, **1999** The Structure and Rheology of Complex Fluids *OUP US.*

Lei, ZG *et al.*, **2008** Predictive Molecular Thermodynamic Models for Liquid Solvents, Solid Salts, Polymers, and Ionic Liquids. *Chem. Rev.*, *108*, 1419-1455.

Lindsay, JP *et al.*, **2004** Combinatorial Formulation of Biocatalyst Preparations for Increased Activity in Organic Solvents: Salt Activation of Penicillin Amidase. *Biotechnol. Bioeng.*, *85*, 553-560.

Lousa, D *et al.*, **2013** A Molecular Perspective on Nonaqueous Biocatalysis: Contributions from Simulation Studies. *Phys. Chem. Chem. Phys.*, *15*, 13723-13736.

Lu, H *et al.*, **1998** Unfolding of Titin Immunoglobulin Domains by Steered Molecular Dynamics Simulation. *Biophys. J.*, *75*, 662-671.

Lüdemann, SK *et al.*, **2000** How Do Substrates Enter and Products Exit the Buried Active Site of Cytochrome P450cam? 1. Random Expulsion Molecular Dynamics Investigation of Ligand Access Channels and Mechanisms. *J. Mol. Biol.*, *303*, 797-811.

Luzar, A, **2000** Resolving the Hydrogen Bond Dynamics Conundrum. *J. Chem. Phys.*, *113*, 10663-10675.

Malcolmson, C *et al.*, 2002 Light-Scattering Studies of Testosterone Enanthate Containing Soybean Oil/C18 : 1E10/Water Oil-in-Water Microemulsions. *J. Pharm. Sci., 91,* 2317-2331.

Malcolmson, C *et al.*, 1998 Effect of Oil on the Level of Solubilization of Testosterone Propionate into Nonionic Oil-in-Water Microemulsions. *J. Pharm. Sci., 87,* 109-116.

Marangoni, AG *et al.*, 2012 Structure and Functionality of Edible Fats. *Soft Matter, 8,* 1275-1300.

Marion, D, 2013 An Introduction to Biological NMR Spectroscopy. *Mol. Cell. Proteomics, 11,* 3006-3025.

Marrink, S-J; Tieleman, DP, 2001 Molecular Dynamics Simulation of a Lipid Diamond Cubic Phase. *J Am Chem Soc, 123,* 12383-12391.

Marrink, SJ; Mark, AE, 2001 Effect of Undulations on Surface Tension in Simulated Bilayers. *J. Phys. Chem. B, 105,* 6122-6127.

Medina, JS *et al.*, 2011 Molecular Dynamics Simulations of Rigid and Flexible Water Models: Temperature Dependence of Viscosity. *Chem. Phys., 388,* 9-18.

Meng, EC; Kollman, PA, 1996 Molecular Dynamics Studies of the Properties of Water Around Simple Organic Solutes. *J. Phys. Chem., 100,* 11460-11470.

Micaelo, NM *et al.*, 2006 Parametrization of 1-Butyl-3-methylimidazolium Hexafluorophosphate/Nitrate Ionic Liquid for the GROMOS Force Field. *J. Phys. Chem. B, 110,* 14444-14451.

Micaelo, NM; Soares, CM, 2007 Modeling Hydration Mechanisms of Enzymes in Nonpolar and Polar Organic Solvents. *FEBS J., 274,* 2424-2436.

Mikhail, SZ; Kimel, WR, 1961 Densities and Viscosities of Methanol-Water Mixtures. *J. Chem. Eng. Data, 6,* 533-537.

Mine, Y *et al.*, 2005 Enhancement of Enzyme Activity and Enantioselectivity by Cyclopentyl Methyl Ether in the Transesterification Catalyzed by Pseudomonas Cepacia Lipase Co-Lyophilized with Cyclodextrins. *Biotechnol. Lett., 27,* 383-388.

Mitchell, DJ *et al.*, 1983 Phase Behaviour of Polyoxyethylene Surfactants with Water. Mesophase Structures and Partial Miscibility (Cloud Points). *J. Chem. Soc., Faraday Trans., 79,* 975-1000.

Miyamoto, S; Kollman, PA, 1992 Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem., 13,* 952-962.

Mizan, TI *et al.*, 1996 Comparison of Rigid and Flexible Simple Point Charge Water Models at Supercritical Conditions. *J. Comput. Chem., 17,* 1757-1770.

Moore, GE, **1998** Cramming More Components onto Integrated Circuits (Reprinted from Electronics, pg 114-117, April 19, 1965). *Proc. IEEE, 86,* 82-85.

Morrone, JA *et al.*, **2006** Ab Initio Molecular Dynamics Simulation of the Structure and Proton Transport Dynamics of Methanol-Water Solutions. *J. Phys. Chem. B, 110,* 3712-3720.

Neale and Pomès 2008 Combination rules for united-atom lipids and OPLS-AA proteins. http://www.pomeslab.com/files/lipidCombinationRules.pdf.

Norin, M *et al.*, **1994** Molecular Dynamics Simulations of an Enzyme Surrounded by Vacuum, Water, or a Hydrophobic Solvent. *Biophys. J., 67,* 548-559.

Nose, S, **1984** A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys., 81,* 511-519.

Odeberg, JM *et al.*, **2003** Lipid Drug Delivery and Rational Formulation Design for Lipophilic Drugs with Low Oral Bioavailability, Applied to Cyclosporine. *Eur. J. Pharm. Biopharm., 20,* 375-382.

Oostenbrink, C *et al.*, **2004** A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Phys., 25,* 1656-1676.

Panda, T; Gowrishankar, BS, **2005** Production and Applications of Esterases. *Appl. Microbiol. Biotechnol., 67,* 160-169.

Parrinello, M; Rahman, A, **1981** Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys., 52,* 7182-7190.

Pascal, TA; Goddard, WA, **2012** Hydrophobic Segregation, Phase Transitions and the Anomalous Thermodynamics of Water/Methanol Mixtures. *J. Phys. Chem. B, 116,* 13905-13912.

Persson, LC *et al.*, **2013** Computational Prediction of Drug Solubility in Lipid Based Formulation Excipients. *Pharm. Res., 30,* 3225-3237.

Persson, M *et al.*, **2002** Factors Governing the Activity of Lyophilised and Immobilised Lipase Preparations in Organic Solvents. *Chembiochem, 3,* 566-571.

Poulin, P *et al.*, **2011** PHRMA CPCDC Initiative on Predictive Models of Human Pharmacokinetics, Part 5: Prediction of Plasma Concentration–Time Profiles in Human by Using the Physiologically-Based Pharmacokinetic Modeling Approach. *J. Pharm. Sci., 100,* 4127-4157.

Pouton, CW, **2000** Lipid Formulations for Oral Administration of Drugs: Non-Emulsifying, Self-Emulsifying and 'Self-Microemulsifying' Drug Delivery Systems. *Eur. J. Pharm. Sci., 11,* S93-S98.

Pouton, CW, **2006** Formulation of Poorly Water-Soluble Drugs for Oral Administration: Physicochemical and Physiological Issues and the Lipid Formulation Classification System. *Eur. J. Pharm. Sci.*, *29*, 278-287.

Prajapati, HN *et al.*, **2012** A Comparative Evaluation of Mono-, Di- and Triglyceride of Medium Chain Fatty Acids by Lipid/Surfactant/Water Phase Diagram, Solubility Determination and Dispersion Testing for Application in Pharmaceutical Dosage Form Development. *Pharm. Res.*, *29*, 285-305.

Prasad, D *et al.*, **2013** Studying the Effect of Lipid Chain Length on the Precipitation of a Poorly Water Soluble Drug from Self-Emulsifying Drug Delivery System on Dispersion into Aqueous Medium. *J. Pharm. Pharmacol.*, *65*, 1134-1144.

Pratt, LR, **2002** Molecular Theory of Hydrophobic Effects: "She Is Too Mean To Have Her Name Repeated." *Annu. Rev. Phys. Chem.*, *53*, 409-436.

Qvist, J; Halle, B, **2008** Thermal Signature of Hydrophobic Hydration Dynamics. *J. Am. Chem. Soc.*, *130*, 10345-10353.

Rane, S *et al.*, **2008** Quantitative Solubility Relationships and the Effect of Water Uptake in Triglyceride/Monoglyceride Microemulsions. *Pharm. Res.*, *25*, 1158-1174.

Rane, SS; Anderson, BD, **2008** Molecular Dynamics Simulations of Functional Group Effects on Solvation Thermodynamics of Model Solutes in Decane and Tricaprylin. *Mol. Pharmaceutics*, *5*, 1023-1036.

Rane, SS; Anderson, BD, **2008** What Determines Drug Solubility in Lipid Vehicles: Is it Predictable? *Adv. Drug Delivery Rev.*, *60*, 638-656.

Rapaport, DC, **1983** Hydrogen-Bonds in Water Network Organization and Lifetimes. *Mol. Phys.*, *50*, 1151-1162.

Rehm, S *et al.*, **2010** Solvent-Induced Lid Opening in Lipases: A Molecular Dynamics Study. *Protein Sci.*, *19*, 2122-2130.

Rezus, YLA; Bakker, HJ, **2007** Observation of Immobilized Water Molecules Around Hydrophobic Groups. *Phys. Rev. Lett.*, *99*, 148301.

Roke, S; Gonella, G, **2012** Nonlinear Light Scattering and Spectroscopy of Particles and Droplets in Liquids. *Annu. Rev. Phys. Chem.*, *63*, 353-378.

Rosenberg, RO *et al.*, **1982** Hydrophobic Effect on Chain Folding - the Trans to Gauche Isomerization of Normal-Butane in Water. *J. Am. Chem. Soc.*, *104*, 7647-7649.

Rossky, PJ; Friedman, HL, **1980** Benzene-Benzene Interaction in Aqueous-Solution. *J. Phys. Chem.*, *84*, 587-589.

Rossky, PJ; Zichi, DA, **1982** Molecular Librations and Solvent Orientational Correlations in Hydrophobic Phenomena. *Faraday. Symp. Chem. Soc.*, 69-78.

Ryckaert, JP *et al.*, 1977 Numerical-integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of n-Alkanes. *J. Comput. Phys.*, *23*, 327-341.

Sacchetti, M; Nejati, E, 2012 Prediction of Drug Solubility in Lipid Mixtures from the Individual Ingredients. *AAPS PharmSciTech*, *13*, 1103-1109.

Sander, C; Holm, P, 2009 Porous Magnesium Aluminometasilicate Tablets as Carrier of a Cyclosporine Self-Emulsifying Formulation. *AAPS PharmSciTech*, *10*, 1388-1395.

Sato, T *et al.*, 2000 Hydrophobic Hydration and Molecular Association in Methanol-Water Mixtures Studied by Microwave Dielectric Analysis. *J. Chem. Phys.*, *112*, 2924-2932.

Schmid, A *et al.*, 2001 Industrial Biocatalysis Today and Tomorrow. *Nature*, *409*, 258-268.

Shante, VKS; Kirkpatrick, S, 1971 Introduction to Percolation Theory. *Adv. Phys.*, *20*, 325.

Singh, B *et al.*, 2009 Self-Emulsifying Drug Delivery Systems (SEDDS): Formulation Development, Characterization, and Applications. *Crit. Rev. Ther. Drug Carrier Syst.*, *26*, 427-451.

Soares, CM *et al.*, 2003 Protein Structure and Dynamics in Nonaqueous Solvents: Insights from Molecular Dynamics Simulation Studies. *Biophys. J.*, *84*, 1628-1641.

Soper, AK *et al.*, 2006 Excess Entropy in Alcohol-Water Solutions: a Simple Clustering Explanation. *J. Phys. Chem. B*, *110*, 3472-3476.

Soper, AK; Finney, JL, 1993 Hydration of Methanol in Aqueous-Solution. *Phys. Rev. Lett.*, *71*, 4346-4349.

Spernath, A *et al.*, 2002 Food-Grade Microemulsions Based on Nonionic Emulsifiers: Media To Enhance Lycopene Solubilization. *J. Agric. Food Chem.*, *50*, 6917-6922.

Spernath, A *et al.*, 2003 Self-Diffusion Nuclear Magnetic Resonance, Microstructure Transitions, and Solubilization Capacity of Phytosterols and Cholesterol in Winsor IV Food-Grade Microemulsions. *J. Agric. Food Chem.*, *51*, 2359-2364.

Stanley, HE; Teixeira, J, 1980 Interpretation of the Unusual Behavior of H2O and D2O at Low-Temperatures - Tests of a Percolation Model. *J. Chem. Phys.*, *73*, 3404-3422.

Struis, R *et al.*, 1987 Dynamic Behavior and Some of the Molecular Properties of Water Molecules in Pure Water and in MgCl2 Solutions. *J. Phys. Chem. B*, *91*, 1639-1645.

Sugita, Y; Okamoto, Y, 1999 Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.*, *314*, 141-151.

Sum, AK *et al.*, 2003 Predictive Molecular Model for the Thermodynamic and Transport Properties of Triacylglycerols. *J. Phys. Chem. B, 107*, 14443-14451.

Tejo, B *et al.*, 2004 Structure and Dynamics of Candida Rugosa Lipase: the Role of Organic Solvent. *J. Mol. Model., 10*, 358-366.

Thomas, N *et al.*, 2013 Supersaturated Self-Nanoemulsifying Drug Delivery Systems (Super-SNEDDS) Enhance the Bioavailability of the Poorly Water-Soluble Drug Simvastatin in Dogs. *AAPS J., 15*, 219-227.

Thomas, N *et al.*, 2012 In Vitro and In Vivo Performance of Novel Supersaturated Self-Nanoemulsifying Drug Delivery Systems (Super-SNEDDS). *J. Control. Release., 160*, 25-32.

Titantah, JT; Karttunen, M, 2012 Long-Time Correlations and Hydrophobe-Modified Hydrogen-Bonding Dynamics in Hydrophobic Hydration. *J. Am. Chem. Soc., 22*, 9362-9368

Torres, S; Castro, GR, 2004 Non-Aqueous Biocatalysis in Homogeneous Solvent Systems. *Food Technol. Biotechnol., 42*, 271-277.

Trodler, P; Pleiss, J, 2008 Modeling Structure and Flexibility of Candida Antarctica Lipase B in Organic Solvents. *BMC Struct. Biol., 8*.

Trodler, P *et al.*, 2009 Modeling of Solvent-Dependent Conformational Transitions in Burkholderia Cepacia Lipase. *BMC Struct. Biol., 9*.

Uppenberg, J *et al.*, 1994 The Sequence, Crystal Structure Determination and Refinement of Two Crystal Forms of Lipase B from Candida Antarctica. *Structure, 2*, 293-308.

Valivety, RH *et al.*, 1992 Reaction Rate with Suspended Lipase Catalyst Shows Similar Dependence on Water Activity in Different Organic Solvents. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol., 1118*, 218-222.

van der Kamp, MW; Mulholland, AJ, 2013 Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry, 52*, 2708-2728.

van der Spoel, D *et al.*, 2005 GROMACS: Fast, Flexible, and Free. *J. Comput. Chem., 26*, 1701-1718.

van der Spoel, D *et al.*, 1998 A Systematic Study of Water Models for Molecular Simulation: Derivation of Water Models Optimized for Use with a Reaction Field. *J. Chem. Phys., 108*, 10220-10230.

van der Spoel, D *et al.*, 2006 Thermodynamics of Hydrogen Bonding in Hydrophilic and Hydrophobic Media. *J. Phys. Chem. B, 110*, 4393-4398.

van Erp, TS; Meijer, EJ, **2001** Hydration of Methanol in Water. A DFT-Based Molecular Dynamics Study. *Chem. Phys. Lett., 333*, 290-296.

van Rantwijk, F; Sheldon, RA, **2007** Biocatalysis in Ionic Liquids. *Chem. Rev., 107*, 2757-2785.

von Corswant, C; Thorén, PEG, **1999** Solubilization of Sparingly Soluble Active Compounds in Lecithin-Based Microemulsions: Influence on Phase Behavior and Microstructure. *Langmuir, 15*, 3710-3717.

Wakisaka, A *et al.*, **2001** Solute-Solvent and Solvent-Solvent Interactions Evaluated Through Clusters Isolated from Solutions: Preferential Solvation in Water-Alcohol Mixtures. *J. Mol. Liq., 90*, 175-184.

Warisnoicharoen, W *et al.*, **2000** Light-Scattering Investigations on Dilute Nonionic Oil-in-Water Microemulsions. *AAPS Pharmsci, 2*, 16–26.

Warren, D *et al.*, **2013** Glyceride Lipid Formulations: Molecular Dynamics Modeling of Phase Behavior During Dispersion and Molecular Interactions Between Drugs and Excipients. *Pharm. Res., 30*, 3238-3253.

Wedberg, R *et al.*, **2012** Protein Dynamics in Organic Media at Varying Water Activity Studied by Molecular Dynamics Simulation. *J. Phys. Chem. B, 116*, 2575-2585.

Wensink, EJW *et al.*, **2003** Dynamic Properties of Water/Alcohol Mixtures Studied by Computer Simulation. *J. Chem. Phys., 119*, 7308-7317.

Wescott, CR; Klibanov, AM, **1994** The Solvent Dependence of Enzyme Specificity. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol., 1206*, 1-9.

Winger, M *et al.*, **2009** Force-field Dependence of the Conformational Properties of α,ω-Dimethoxypolyethylene Glycol. *Mol. Phys., 107*, 1313-1321.

Wittrup Larsen , M *et al.*, **2000** Suppression of Water as a Nucleophile in Candida antarctica Lipase B Catalysis. *ChemBioChem, 11*, 796-801.

Wu, Y *et al.*, **2006** Flexible Simple Point-Charge Water Model with Improved Liquid-State Properties. *J. Chem. Phys., 124*.

Yamaguchi, T *et al.*, **1999** The Structure of Liquid Methanol Revisited: a Neutron Diffraction Experiment at -80 Degrees C and +25 Degrees C. *Mol. Phys., 97*, 603-605.

York, DM *et al.*, **1993** The Effect of Long-Range Electrostatic Interactions in Simulations of Macromolecular Crystals - a Comparison of the Ewald and Truncated List Methods. *J. Chem. Phys., 99*, 8345-8348.

Zaks, A; Klibanov, AM, **1988** Enzymatic Catalysis in Nonaqueous Solvents. *J. Biol. Chem., 263*, 3194-3201.

Zhou, R, **2003** Free Energy Landscape of Protein Folding in Water: Explicit vs. Implicit Solvent. *Proteins: Struct., Funct., Bioinf., 53,* 148-161.

Zidan, AS *et al.*, **2007** Quality by Design: Understanding the Formulation Variables of a Cyclosporine A Self-Nanoemulsified Drug Delivery Systems by Box-Behnken Design and Desirability Function. *Int. J. Appl. Pharm., 332,* 55-63.

Zimmermann, K, **1991** ORAL: All Purpose Molecular Mechanics Simulator and Energy Minimizer. *J. Comput. Chem., 12,* 310-319.

zur Mühlen, A *et al.*, **1998** Solid Lipid Nanoparticles (SLN) for Controlled Drug Delivery - Drug Release and Release Mechanism. *Eur. J. Pharm. Biopharm., 45,* 149-155.

Acknowledgements

I would like to thank Prof. Dr. Bernard Hauer for the opportunity to conduct my research at the Institute for Technical Biochemistry.

I am especially grateful to my supervisor Prof. Dr. Jürgen Pleiss for the confidence in allowing me to conduct my research and for his constant support and motivation. I also thank my co-supervisor Prof. Dr. Joachim Groß for his efforts and insightful discussions.

The German Research Foundation (DFG) I would like to thank for financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart.

The High Performance Computing Center Stuttgart (HLRS), the Karlsruhe Institute for Technology (KIT) and the BW-Grid project I would like thank for access to the required computational resources and for the technical support.

All members of the Institute for Technical Biochemistry, and the Bioinformatics group in particular, I thank for very formative years.

A special thanks to my partner and my family for their unwavering support, love and for standing by me no matter what. It will be returned in kind.

Declaration

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature and acknowledgement of collaborative research and discussions.

Stuttgart, July 8th 2014

_____

Sven Benson