Modeling Trapping Mechanism for PCB Adsorption on Activated Carbon

Master of Science thesis in process technology

by

Bjørnar Jensen

Department of Physics and Technology

University of Bergen, Norway

November 2009

Abstract

This thesis aims to investigate the trapping mechanism of polychlorinated biphenyls (PCB) on activated carbon. In the pursuit of this goal, molecular dynamics (MD) simulations were the method chosen. To be able to model a system of activated carbon, PCB and a triacylglycerol, individual molecular models had to be constructed for each of the components. The background for doing this work is that the levels of dioxin-like PCB, along with polychlorinated dibenzo-p-dioxin and dibenzofuran (PCDD/F) in fishmeal and fish oil produced for use in feed for salmon is above present European legislation levels in some regions. To reduce these levels, experiments have been carried out where these species are adsorbed on activated carbon. The adsorption process seems to be efficient for PCDD/F but less efficient for dioxin-like PCB.

Two PCB molecules were chosen, namely 3,3',4,4'-Tetrachlorobiphenyl and 2,3',4,4',5-Pentachlorobiphenyl, PCB congener 77 and 118 respectively. In addition, a triacylglycerol in the form of triolein was chosen to represent the fish oil. These molecules were drawn up using Schrödinger's Maestro graphical user interface. A Brookhaven Protein Databank file was generated from this software, and this was used as a basis for a molecular structure and force field file for MDynaMix v5.1 (MD51), which was the MD software used in this thesis. The force field parameters characterizing these molecules were obtained from the force field known as OPLS – Optimized Potential for Liquid Simulations, for which the main author is Prof. William L. Jorgensen. Partial atomic charges were obtained from quantum chemical calculations using density functional theory, with B3LYP and STO-6-31+G** basis set, and the Solvation Model 6 in order to get the charges for water solvated molecules.

Activated carbon (AC) is a graphitic material. Thus, graphite was used as the basis for the model of AC. A computer software utility was written during this thesis to generate a model of AC. This was necessary in order to build the model since it contains no less than 31232 carbon atoms. Similarly, in order to assign partial atomic charges to the individual atoms in this model a new software tool was written to do the job.

The models for the PCB molecules were validated with respect to their molar enthalpies of vaporization. However, the models for AC and triolein were not validated due to the time constraint. From the validations of PCB models, it was found that the model for PCB 77 gives

good results for the molar enthalpy of vaporization; however, a slight deviation from the experimental values is observed. PCB congener 118 is able to replicate the experimental values for the molar enthalpy of vaporization. Nevertheless, the conclusion about these model is that the phenyl-phenyl ring torsion angles does not have the correct parameters. Consequently, even though the models can replicate thermodynamic properties, they need new torsion angle parameters for the phenyl-phenyl ring torsion angle in order to replicate the correct molecular structures.

Furthermore, the individual components were merged into a large simulation, using a modified version of the MD51, which allowed an external force to rapidly move the liquid phase down on the AC model. This external force was subsequently scaled down over time. Again, due to the short time available for this thesis, the main simulation is still an immature system, with simulation time less than 1ns. Currently the simulation time is ~0.6ns. The speed of the simulation is roughly 1ns per month. No definitive results were found as to the trapping mechanism of PCB on AC, due to both the short period of simulation time, and some discrepancies observed in the simulation. This latter problem was known when the main simulation was started, but it was concluded that some valuable information could still be gathered. In order to get more definitive results, a new simulation must be set up, where a nitrogen atmosphere is a viable option to avoid the discrepancies observed. Simulations with runtimes longer than the current one is also needed.

An extended abstract for an article based upon this work has already been submitted. However, the article is not completed in time to be attached to this thesis.

Acknowledgements

Contents

1   Introduction

Polychlorinated biphenyls (PCB) are classified as a persistent organic pollutant (POP). These are lipophilic environmental pollutants. As such, PCB biomagnifies in the food chain (Oterhals et al. 2007). Fish oil is produced from the fat of fishes, mainly herring, sprat, sand eel and blue whiting. Due to biomagnification, the concentration is increased when processed into fish oils. From previous research, fish oils are proven the main source of POP in high-fat fish feed formulations (Oterhals et al. 2007) (Lundebye et al. 2004) (Hites et al. 2004). In order to develop a cost-effective solution for decontamination of fish-oils and fishmeal, a study was carried out (Oterhals et al. 2007), in which activated carbon (AC) was used as an adsorbent.

Activated carbon has been widely used as an adsorbent in the food industry, as well as others that utilizes adsorption process with activated carbon for various separation purposes. As a result of his study, Oterhals suggested a hypothesis that efficient adsorption onto AC was dependent on a planar molecular conformation. If this is the case, the usability of the adsorption process onto AC would be limited to a small class of POP. In this thesis, the goal is to use molecular dynamics (MD) as a method to investigate the hypothesis proposed by Oterhals.

In order to apply MD molecular-level models must be created first. The adsorbent, the solvent and the adsorbate all have to be modeled and then combined in a simulation system. Thus, the first tasks consisted of investigation into the general molecular structure of AC, PCBs and triacylglycerols (the solvent). A minimum of two different PCBs had to be selected, one with a co-planar molecular structure, and one with a mono-ortho chlorine substitution. In addition, a triacylglycerol had to be chosen as the solvent suitable to represent the fish oil. The major challenge though, lay in constructing the activated carbon model. Several issues needed to be solved here, and will be explained in detail later on in this thesis. Because of the size of the tasks, this thesis will mainly detail the construction of models, their verification, and the task of assembling the system. Some initial simulations has been run, though without definitive results as time constraint became an issue.

2 Definition and goals

The primary objective in this thesis is to identify the trapping mechanism of PCBs and to explain why some PCB congeners show a much stronger affinity towards the activated carbon than others. In doing so, Oterhals' hypothesis should either be proven correct or dismissed.

In order to achieve the primary objective the following sub-goals must be accomplished:

i. Investigate the structure of activated carbon and construct a model for activated carbon.

ii. Select at least two different PCB congeners and model them.

iii. Choose a suitable triacylglycerol to represent the fish oil, and construct a model capable of representing this constituent.

iv. The individual models should be validated against experimental quantities in order to ensure they represent the real-life molecules.

v. The individual models must all be incorporated into a large scale simulation.

vi. Equilibration of the large scale simulation, and a subsequent extended run must be completed in order to acquire a sufficient dataset to reach the primary objective.

3 Scientific methods

As stated in the previous section, the main goal of this thesis was to investigate the PCBs selectivity on activated carbon, in order to test the validity of Oterhals' hypothesis for the trapping mechanism of activated carbon. Early on, it had been decided that the most feasible way of doing this would be by atomistic simulations of the activated carbon, the fish oil constituents and PCBs. This was due to that trapping of PCBs in the slit pores might not be solely due to geometric constraining effects. Attractive interactions between the activated carbon and the PCBs might also have a considerable effect on both the adsorption process, and the subsequent immobilization of PCB in the slit pores.

To use ab initio simulations was simply not an option due to the large size of the system. Ab initio methods could at best be used to characterize the individual models. That left Molecular Dynamics (MD) and Monte Carlo (MC) simulations. A disadvantage with MC is that kinetic information, which potentially could be of value for later analysis of the adsorption kinetics and relative sensitivities, is lacking. This is in contrast to MD, which readily offers this pinpoint information. Another drawback of MC is that a random change in configuration, which is the essence of MC, could be too large and consequently result in skipping important steps needed to the trapping mechanism. Thus, critical changes in the system, or states of the system, could be missed.

For the reasons discussed above, the decision was made to go with MD as our method of choice for modeling the system and investigating the trapping mechanism of PCB onto activated carbon.

4 Computational chemistry and molecular dynamics

This section details the theory of the methods used in this thesis. It is divided into two main sections. A general oversight of quantum chemistry is given in the first part, while the concept of molecular dynamic simulation is the second part.

4.1 Quantum chemistry

In quantum chemistry (QC) methods, the focus is on electrons and atomic nuclei as building blocks for atoms. This is in contrast to force field methods, where atoms, pseudo-atoms and even amino acids are the building blocks.

Quantum chemistry apply quantum mechanics (QM) to find energies, charges, reaction paths, transition states, solvation energies and so on. QC is all about particles. Moreover, properties of these particles can be found through solving the Schrödinger equation.

QC as a field has seen a tremendous rise in popularity and applicability in the last decade. It is, however, not a simple technique, and it demands that the user has some knowledge about its inner workings, in order to interpret the results. Different options will yield quite different results, some more correct than others. Computers might do the calculations, but humans must interpret the results. Thus, knowledge is key to QC methods, as well as many other fields.

4.1.1 The Schrödinger equation

In quantum mechanics, the wavefunction contains all the dynamical information about the system it describes (Atkins and de Paula 2002). It is a mathematical tool that can have both a real and a complex value. The wavefunction itself cannot be related to a physical measurement. However, the probability of finding the particle in a given volume at a given time, will be given by equation (4.1).

$$\text{P} = \iiint \Psi^*(x, y, z, t)\Psi(x, y, z, t)dxdydz$$

(4.1)

Here P is the probability and Ψ^*(x, y, z, t) is the complex conjugate of Ψ(x, y, z, t). This means that the square of the modulus of the wavefunction is related to a physical measurement (Hinchliffe 2003).

By solving the Schrödinger equation one can obtain the wavefunction. The Schrödinger equation has been named after the Austrian physicist Erwin Schrödinger. He introduced his now famous equation describing the wavefunction of any system in 1926 (Atkins and de Paula 2002). There are several ways to write the Schrödinger equation. For example, the time-dependent Schrödinger equation, and the time-independent Schrödinger equation. The focus here will be the time-independent Schrödinger equation (4.2).

$$\hat{H}\Psi = E\Psi \tag{4.2}$$

Where $\hat{H}$ is the Hamiltonian operator, Ψ is the wavefunction and E is the total energy. The Hamiltonian operator is given by equation (4.3).

$$\hat{H} = \left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V(x, y, z) \right] \tag{4.3}$$

Here $\hbar$ is $h/2\pi$, where h is Planck's constant. V(x, y, z) is a three dimensional potential field, and m is the mass of the particle. The Hamiltonian operator operates upon the wave function Ψ, to the right of the $\hat{H}$. The Schrödinger equation is an eigenvalue equation. That is, operating (by an operator such as $\hat{H}$) on a function returns the same function multiplied with a constant (Jensen 2007).

4.1.2 Basis sets

An atomic orbital is a one-electron function used to describe the electronic distribution around an atom (Lowe and Peterson 2006). An example would be a solution to the Schrödinger equation for the hydrogen atom; this would be a hydrogenic orbital. In the context used in this section, the atomic orbitals have a different name. When used in basis sets, atomic orbitals are known as *basis functions* (Hinchliffe 2003).

To apply the quantum theory to molecules, MO-LCAO are used. MO-LCAO is an acronym for Molecular Orbital – Linear Combination of Atomic Orbitals. That is, by taking linear combinations of atomic orbitals, or basis functions, it is possible to construct molecular orbitals. A molecular orbital, in the MO-LCAO theory, is thus a mathematical combination of atomic orbitals.

There have been proposed a number of basis functions. Two of the most used are the Slater type orbital (4.4) and Gaussian type orbital (4.5).

$$\chi_{\xi,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{n-1}e^{-\xi r} \tag{4.4}$$

$$\chi_{\xi,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{2n-2-l}e^{-\xi r^2} \tag{4.5}$$

In equation (4.4) and equation (4.5), χ is the orbital; ξ is an adaptable parameter related to the electric shielding; n, l and m are quantum numbers; r is the radial distance; θ is the azimuthal angle ; φ is the polar (zenith) angle; N is the normalization constant and $Y_{l,m}(\theta, \varphi)$ are spherical harmonic functions.

However, both of these two basis functions have their limitations. The Slater type orbital requires much more computational power to solve than the Gaussian type orbitals. On the other hand, the Gaussian type orbitals are inferior to the Slater type orbitals in terms of accuracy. The Gaussian type orbitals have a zero slope at the nucleus position, while the Slater type orbitals have a cusp. The discontinuous behavior of the Slater type orbitals at the nucleus is desired, and as such the Gaussian type orbitals fail to reproduce this behavior (Jensen 2007). Another problem with the Gaussian type orbitals is that they fall off too quickly far away from the nucleus, due to the r^2 dependence in the exponential function.

Despite their differences, both Slater type orbitals (STO) and Gaussian type orbitals (GTO) can be used to construct basis sets. However, due to the deficiencies in the GTOs, more GTOs have to be used to achieve the same accuracy as with a STO. As a general rule of thumb, three GTOs are needed to represent one STO (Jensen 2007). Still, it is computationally faster to use multiple GTOs per STO.

A minimal basis set is a basis set that contains the fewest basis functions possible. A more comprehensive basis set is denoted double zeta (DZ) type basis. This doubles all the basis functions. For a hydrogen atom, a minimal basis set will only have one s-function to describe the 1s orbital. A DZ basis set would use two s-functions to describe the 1s orbital. By using different exponents (often denoted by the same Greek letter ζ, zeta, in STOs) in the exponential functions, a better description of the electron distribution is possible (Jensen 2007). Another way is using split valence basis. In the split valence basis sets, additional orbitals are only considered for the valence electrons. Thus, a double zeta split valence basis set would only double the orbitals for the valence electrons, but not the core electrons. With large molecules composed of heavy atoms, the savings in computational time from excluding

multiple orbitals for core electrons can be quite substantial. A triple zeta would have three times as many basis functions as a minimal basis set.

To further increase the accuracy of the basis sets, polarization functions can be added as well. Polarization functions are higher angular momentum functions. For hydrogen atoms, this would mean the addition of a p-function to the basis set, thus allowing for polarization of the s-orbitals. Similar a d-function can be added to polarize p-orbitals for heavier atoms, and so on. A set of diffuse functions can also be included; these are functions with small exponents in the exponential functions, thus they extend over a larger distance from the nucleus, or the center set for the basis function. Diffuse functions are needed when dealing with diffuse electronic distributions, like in anions or excited states.

In this thesis, several basis sets were used during QC calculations. STO-3G minimal basis set was one. This is a Slater type orbital, approximated by three Gaussian type orbitals, and is a Pople style basis set (Jensen 2007). The number 3 in STO-3G denotes the number of primitive Gaussians that are used to represent the core electrons. In the Pople style basis sets, the exponents in the exponential functions are fitted to match the STO. In contracted basis sets, the linear coefficients in front of the individual primitive Gaussians are fixed. A single contracted Gaussian type orbital (CGTO) is given by a number of primitive GTOs (PGTO), where the linear coefficients in front of each PGTO are fixed. Equation (4.6) shows the relationship between CGTOs and PGTOs.

$$\chi(CGTO) = \sum_i^k a_i \chi_i(PGTO)$$

(4.6)

Here i is an index over all PGTOs, a_i is the linear coefficient in front of the PGTO χ_i. Another basis set used was the STO-3-21G, which is a split valence double zeta basis set of contracted GTOs. The number 3 denotes the number of PGTOs used for core electrons. The numbers that follow indicate that this is a split valence, double zeta basis. Number 2 shows that the inner part of the valence shell is represented by two PGTOs. The last number, 1, tells that the outer part of the valence shell is represented by only one PGTO. The G stands for Gaussian.

The third type of basis sets used in this thesis is the STO-6-31+G**. This is also a Pople style basis set, with split valence double zeta basis. In addition, this basis sets also has added diffuse functions, denoted by the "+". This indicates that a diffuse set of basis functions has been added for s- and p-functions for heavy atoms (i.e. all but hydrogen atoms). The final part

of the description is the ** which indicated that polarization functions are added, for both hydrogen atoms, and heavier atoms.

4.1.3 Density Functional Theory

"The density functional theory of Hohenberg and Kohn, is in principle an exact theory" (Hinchliffe 2003). The basis for Density functional theory (DFT) is that the ground state energy of any electronic system is determined by a functional of the electron density. Thus, if we know the electron density of a system, in three-dimensional space, it is possible to calculate the ground state energy of that system, or any other ground state property, without knowing the wave function.

However, the "correct" functional for transforming the electron density into the exact ground state energy is unknown. What we do have is approximations. The energy functional can be divided into three parts. $T[\rho]$ is the functional for the kinetic energy. The attraction between the nuclei and the electrons is given by $E_{ne}[\rho]$, and the electron-electron repulsion is given by $E_{ee}[\rho]$ (Jensen 2007). Furthermore, the electron-electron repulsion may be split into two terms; a Coulomb and an exchange term, $J[\rho]$ and $K[\rho]$ respectively.

In the Kohn-Sham theory, the DFT energy is given by equation (4.7). In this equation, the subscript S in the kinetic energy term indicates that the energy is calculated from a Slater determinant (determinant of a matrix of spinorbitals).

$$E_{DFT}[\rho] = T_S[\rho] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho]$$
(4.7)

The last term in equation (4.7) is the exchange-correlation term, and by equating the E_{DFT} to the exact energy for the system, we can find the expression for E_{XC} (4.8).

$$E_{xc} = (T[\rho] - T_S[\rho]) + (E_{ee}[\rho] - J[\rho])$$
(4.8)

Kohn and Sham realized that since the exact functional for the kinetic energy of the system was unknown, they should shift their focus to calculating as much of the energy as possible rigorously. Since the Hartree-Fock (HF) method using Slater determinants is able to give the exact energy for a non-interacting system, provided that it is non-degenerate, the wave function kinetic energy from this method was implemented (Koch and Holthausen 2001). This expression is given in eq. (4.9). Here i goes over all electrons, φ_i is molecular orbitals, and the subscript S in T_S indicates that the kinetic energy is calculated from a Slater determinant.

$$T_S = -\frac{1}{2}\sum_{i}^{N}\langle\varphi_i|\nabla^2|\varphi_i\rangle \tag{4.9}$$

As the kinetic energy expression will be rigorous only for non-interacting systems, kinetic energy calculated from equation (4.9) is not equal to the true kinetic energy of the system. The deviation from the true kinetic energy is included in the exchange-correlation part. More specifically, it is the first term in parenthesis on the right hand side of equation (4.8).

Since all but the $E_{XC}[\rho]$ is known, and are exact, this is where the DFT of Kohn-Sham becomes an approximation. The quality of the results from DFT thus depends on the quality of the $E_{XC}[\rho]$ term. In general, this means that if the exact form of $E_{XC}[\rho]$ had been known DFT would be exact. Several approximations for $E_{XC}[\rho]$ have been suggested by various researchers. In this thesis, the B3LYP hybrid functional was used for the $E_{XC}[\rho]$. B3LYP is acronym for Becke 3 parameter Lee-Yang-Parr, and it is given in equation (4.10) (Jensen 2007).

$$E_{XC}^{B3LYP} = (1-a)E_X^{LSDA} + aE_X^{exact} + b\Delta E_X^{B88} + (1-c)E_C^{LSDA} + cE_C^{LYP} \tag{4.10}$$

Where E_X is the energy from the exchange contribution; E_C is the energy from the electron correlation. a, b and c are the three parameters fitted to experimental data, typical values being $a \sim 0.2$, $b \sim 0.7$ and $c \sim 0.8$. LSDA stands for Local Spin Density Approximation; the E_X^{exact} is the exact exchange energy for the KS (Kohn-Sham) orbitals. However, since the KS orbitals are not equal to the HF orbitals this equation will not yield the correct exact exchange energy. Recall that the KS orbitals are for non-interacting systems. E_X^{B88} was proposed by Becke, and is a correction to the exchange energy from the LSDA method. E_C^{LSDA} is the energy contribution from electron correlation in from the LSDA method; while E_C^{LYP} is the energy contribution from electron correlation formulated by Lee, Yang and Parr.

4.1.4 Continuum Solvation Models and the Solvation Model 6

The modeling of liquids and solvation free energies has traditionally been done by classical approaches, like Molecular Dynamics or Monte Carlo simulations. These methods use solvent molecules to model the liquid explicitly. Depending on the solvent, the number of solvent molecules needed can be substantial, and will result in corresponding computational efforts. This is time consuming and requires extensive computational power. Continuum solvation models replace the explicit solvent molecules with an implicit solvent, in the form of a

continuum. The solute is then modeled in a solvent reaction field, as a cavity (with partial atomic charges assigned to the atoms) in a dielectric (Tomasi and Persico 1994).

Solvation Model 6 (SM6) utilizes this approach. Interlocking spheres, slightly larger than van der Waals radius for the corresponding atom, make up the cavity for the solute molecule. The interactions between the solute and the solvent are parameterized in a semi-empirical fashion, using specific parameters for each type of solvent (Tomasi and Persico 1994). Semi-empirical methods are also used for the quantum description of the quantum-continuum dielectric model. According to its authors, the SM6 model benefits from the explicit inclusion of a single solvent molecule in the system (Kelly et al. 2005). The partial atomic charges for SM6 are obtained from Charge Model 4 (CM4) (Olson et al. 2007). CM4 uses partial atomic charges from the Löwdin population analysis or redistributed Löwdin population analysis in combination with parameters obtained from a training set to estimate the partial atomic charges. Löwdin population analysis, the redistributed Löwdin population analysis, as well as CM4 and electrostatic potential (ESP) fitted charges are discussed in the next section.

Continuum solvation models are widely used to obtain theoretical predictions of properties depending on the specifics of solvent-solute interactions. This was the reason for the utilization of SM6 in this thesis. However, the free energy of solvation was not of interest. What was of interest was the possible effect on the partial atomic charges. Considering that the model system contains a liquid, charges from vacuum might not be applicable. Comparison of partial atomic charges obtained in vacuum and in a water solvated system was made to determine which charges should be used.

4.1.5 Partial atomic charges from QM

Partial atomic charges can be estimated through various methods. One of the simpler methods, and as such a very popular method, is the Mulliken population analysis (Mulliken 1955). This method divides the number of electrons among all the basis functions in use, and then sums the contributions N(A) over the basis functions that belong to atom A, included half the contribution from overlap integrals. The partial atomic charge, q_A, for atom A is then

$$q_A = Z_A - N(A) \qquad \text{(4.11)}$$

Where Z_A is atomic number.

Löwdin population analysis is another method of obtaining the partial atomic charges centered on each atom in a molecule. This method improves the Mulliken population analysis in that it transforms the atomic orbitals into orthonormal basis functions before carrying out the population analysis (Cramer 2004). This is done because the charges derived from the Mulliken population analysis were found lacking in precision as well as tending to become unphysical large in case of large basis sets.

Several other problems also afflicted the Mulliken population analysis (Jensen 2007). One of these was that it could assign individual basis functions with occupation number greater than 1, and lower than 0. This in turn implies that an atomic orbital could have more than 2 electrons, or less than 0 electrons, respectively. For an AO to have more than 2 electrons would be in violation of the Pauli principle. Having less than 0 electrons physically makes no sense. A third problem was that the contributions from the overlap integrals were evenly split over the two atoms. As such, it failed to account for the electronegativity of atoms, which may result in one atom having a greater part of this contribution.

Löwdin's method fixed the last two problems; however, it is not independent of the basis sets. Especially when diffuse functions are included into the basis sets, the results can become unstable (Cramer 2004). If the basis functions are not centered on atoms, but bonds instead, yet another issue arises, since the charges would be completely incorrect, as no electrons would be assigned to atoms (Jensen 2007). This would lead to atomic charges being equal to the atomic number.

The Löwdin charge from the population analysis is given by equation (4.12) (Thompson et al. 2002) (Lowdin 1950).

$$q_k(L\ddot{o}wdin) = Z_k - \sum_{i\epsilon k}(\boldsymbol{S}^{1/2}\boldsymbol{P}\boldsymbol{S}^{1/2})_{ii}$$

(4.12)

In this equation q_k(Löwdin) is the Löwdin charge on atom k, Z_k is the nuclear charge, $\boldsymbol{P}$ is the electronic density matrix, $\boldsymbol{S}^{1/2}$ is the square root of the overlap matrix. The summation index, i, runs over all basis functions centered on atom k.

To somewhat alleviate the problems with the LPA charges when using diffuse basis set, Thompson et al. proposed a new method they called RLPA, or Redistributed Löwdin Population Analysis (Thompson et al. 2002). In RLPA, the contribution from the second part of eq. (4.12) is partitioned into two contributions. One contribution from the tight (non-

diffuse) orthogonalized basis functions and one contribution from the diffuse orthogonalized basis functions. In LPA it is this latter contribution, i.e. the contribution from the diffuse basis functions, which can have excess charge that may not belong to the atom in question. RLPA seeks to distribute the diffuse contribution to other atoms based on the geometry of the molecule and the degree of diffuseness of the basis function in question. Further details can be found in the original paper by (Thompson et al. 2002).

CM4 is the charge model used in SM6. It uses partial atomic charges obtained from LPA, or RLPA in case of diffuse basis sets, to build a new set of partial atomic charges. The LPA or RLPA charges are used as a starting point. CM4 then adds a new contribution based upon the Mayer bond order and two parameters. These two parameters are optimized to minimize sum of the squares of deviation of dipole moments when compared to the known dipole moments of the training set for CM4. The equation for CM4 charges is (Olson et al. 2007),

$$q_k = q_k^0 + \sum_{k \neq k'} T_{kk'}(B_{kk'})$$

(4.13)

Where, q_k is the CM4 partial atomic charge for atom k; q_k^0 is partial atomic charge from LPA or RLPA for atom k and $T_{kk'}$ is a quadratic function of the Mayer bond order $B_{kk'}$, between atom k and k'. The quadratic function is given by,

$$T_{kk'}(B_{kk'}) = (D_{Z_k Z_{k'}} + C_{Z_k Z_{k'}} B_{kk'}) B_{kk'}$$

(4.14)

The two CM4 parameters are $D_{Z_k Z_{k'}}$ and $C_{Z_k Z_{k'}}$. They depend on the method used to find the initial charges q_k^0 in equation (4.13), as well as the density function of the basis set. According to Olson et al. the CM4 charges should be useful for parameterization of force fields (Olson et al. 2007).

The software package used in this thesis included the option of using Solvation Model 6, mentioned in the section above. When this option is utilized, then Redistributed Löwdin Population Analysis (RLPA) (Thompson et al. 2002) is automatically carried out for basis sets containing diffuse functions. This is done since SM6 requires this to complete its calculations. For basis sets that do not include diffuse functions, normal Löwdin population analysis is carried out instead.

In contrast to the Löwdin population analysis, the ESP charges are considered a better approach for obtaining partial atomic charges for force field method simulations. This is due

15

to the way these charges are calculated. Simply put, the molecule is wrapped in a grid at a distance slightly longer than the van der Waals radius. A single point charge is then positioned on the grid at a coordinate, and the force that this point charge experiences is proportional to the partial atomic charges on the molecule. This procedure is carried out a large number of times. In the end, a representation of the electrostatic potential surface as it appears from the outside is built. The individual atoms are then assigned partial atomic charges so that this electrostatic potential can be replicated in a best possible manner. As usual, there are some drawbacks. Atoms lying closer to the electrostatic potential surface get more or less correct charges. However, atoms deep inside the molecule or partially hidden behind another atom are not that well represented. This is due to the shielding effect, i.e. the outer atom will mask, to some extent, the true magnitude of the charge hidden behind it. Jaguar does not support ESP calculations with the SM6 model. As such, ESP charges from a water solvated system were not calculated for any of the models.

4.1.6 Maestro & Jaguar software package

Schrödinger, LLC, publishes the software package used for QM simulations in this thesis. The package consists of the graphical user interface (GUI) Maestro, and the ab initio software, Jaguar. The Maestro GUI allows one to construct molecular models from elements or functional groups, or a combination of both. The models constructed in Maestro are then used as input for the Jaguar program.

Simulations in Jaguar can either be run via the Maestro GUI or the command line by issuing commands or using a batch job script. The GUI can be used to generate these scripts, thus providing an easy starting point to Jaguar scripting. The Jaguar program supports a variety of basis sets, as well as calculation options.

In this work Maestro version 8.5, release 207 was utilized along with Jaguar version 7.5 release 207 (Maestro 2008), (Jaguar 2008).

4.2 Molecular dynamics

Molecular Dynamics (MD) is a technique that follows the time evolution of a collection of molecules, or atoms. By observing the system through time one wishes to observe the systems trajectory through phase space. In MD, phase space is a 6N-dimensional space, where N is the number of particles. 3N of these dimensions are due to the positions of the particles, and 3N is

due to the momentums. Phase space holds all possible positions and momentums. An ensemble can be viewed as a collection of points in phase space that for instance has constant energy. For each time step, MD generates new configurations that correspond to points in phase space. These points in phase space, and corresponding configurations, are connected through time and are time-reversible. This way, time-dependent properties can be calculated. The governing equation, Newton's second law (4.15), relates the acceleration to the force (Hinchliffe 2003).

$$\vec{F} = m\vec{a} = m\frac{d^2\vec{R}}{dt^2}$$

(4.15)

In equation (4.15) $\vec{F}$ is the force as a vector; m is the mass of the particle; $\vec{a}$ is the acceleration as a vector; $\vec{R}$ is the position vector and t is the time.

Contrary to what the term suggests, one does not get the forces directly from the various force field packages. What it does give is formulas and the needed parameters to calculate the potential energy. The forces acting on the system is obtained by taking the negative gradient of the potential energy, U, of the system (4.16). This is detailed in the following sections.

$$\vec{F} = -\nabla U$$

(4.16)

The potential energy of a system is a combination of different terms. The terms can be divided into two main classes. These two classes are the *bonded* contributions and *non-bonded* contributions. Thus, we have that the potential energy, U, for the system is given by equation (4.17).

$$U = U_{bonded} + U_{non-bonded}$$

(4.17)

Where the bonded contribution in MD51 is

$$U_{bonded} = U_{bond} + U_{angle} + U_{torsion} + U_{improper\ torsion}$$

(4.18)

Note that improper torsions are not utilized in any of the models in this thesis. Therefore, improper torsions will not be discussed any further. The non-bonded contributions are given by equation (4.19).

$$U_{non-bonded} = U_{Lennard-Jones} + U_{electrostatic}$$

(4.19)

17

Molecular dynamic simulations performed in this thesis used MDynaMix v5.1. The program is based on the McMOLDYN program by Aatto Laaksonen, but has later been substantially modified and extended to by Alexander Lyubartsev (Lyubartsev and Laaksonen 2000). Professor Tatiana Kuznetsova has also added to this program, by including support for output in the Protein Data Bank format (.PDB). MDynaMix (hereby referred to as MD51) is available for download from the University of Stockholm (Lyubartsev and Laaksonen).

The force field used in this thesis is the "Optimized Potentials for Liquid Simulations", or better known as OPLS, has been developed by Prof. William L. Jorgensen. He was kind enough to send the complete set of force field parameters. The references for OPLS are (Jorgensen et al. 1996; Damm et al. 1997; Jorgensen and McDonald 1998; McDonald and Jorgensen 1998; Rizzo and Jorgensen 1999; Price et al. 2001; Watkins and Jorgensen 2001; Jorgensen et al. 2004; Jensen and Jorgensen 2006).

4.2.1 Ensembles in molecular dynamics

An ensemble in statistical mechanics has two meanings. One of the meanings refers to the constraints, for example constant temperature, volume and number of particles, which are the constraints for a canonical ensemble (Dill and Bromberg 2002). In the other meaning, it refers to a collection of microstates, where all possible microstates within the constraints make up the ensemble. A microstate can be considered a snapshot of the configuration of the system.

The micro-canonical ensemble is the natural ensemble generated by MD (Jensen 2007). In the micro-canonical ensemble, the number of particles (N), volume (V) and the energy (E) is kept constant. That is, MD conserves energy. Energy may shift from potential energy to kinetic energy, but the total energy of the system will remain constant, except for the energy that is artificially lost or gained due to numerical round off errors or truncations in integration routines and computer precision.

The term micro-canonical ensemble stems from statistical mechanics. Thus, there are some constraints related to an integral over the micro-canonical partition function. This integral has a probability function that ensures that the system's Hamiltonian is equal to the unique total energy of the corresponding insulated micro-canonical ensemble. The probability function in the integral is known as Boltzmann's distribution law, and gives the probability of the system being in a given microstate. For an ideal MD simulation, the sampled energy states would have to correspond to the same probability function, as the statistical mechanics theoretical

integral would give. As such, it would generate the same entropy generation. However, different MD settings can alter the sampling, and result in significant differences in entropy. For instance, a too long time step would make the system skip what might be important microstates. A too short time step might lead to the same microstate being sampled more frequently than the specific probability function in the statistical mechanics. Since a simulation cannot be run indefinitely, not all possible configurations will be visited. Even if a simulation could run indefinitely, it would not necessarily help for the simple reason that if the temperature control is not accurately sampling the canonical ensemble, the temperature control will impose a shift in the energy states for the potential and kinetic parts. Thus, with too small time step, only a part of phase space may be explored, and the probability distribution will be skewed. Again, these two scenarios will lead to different entropies. This is in contrast to the integration in statistical mechanics where one specific Hamiltonian and one total energy can only lead to a single unique entropy. This difference between statistical mechanics and ensembles in MD sampling will also be reflected in the other ensembles, or systems, described below. In MD, the NVT ensemble is trying to reproduce the canonical ensemble in statistical mechanics through the appropriate exchange of heat with an infinitively large heat bath. For other ensembles, the exchange can be for mass or mechanical work.

Even though the NVE ensemble is MD's natural ensemble, it is possible to simulate other ensembles. For example, the canonical ensemble, where number of particles (N), volume (V) and the temperature (T) are constant, the Isothermal-Isobaric ensemble where N, pressure (P) and T are constant, and finally the grand canonical ensemble where V, E and the chemical potential (μ), are constant. To generate the ensembles that are "unnatural" to MD, some special techniques have to be used. This thesis will not discuss how NPT and VEμ can be generated. However, it will show how temperature can be controlled and on the average kept close to constant, since simulations carried out here were done in the NVT ensemble, which intends to mimic the canonical ensemble in statistical mechanics.

In order to sample a canonical ensemble distribution, NVT has to be kept constant. Statistical mechanics states that the temperature of the system is proportional to its kinetic energy (4.20) (Jensen 2007).

$$\frac{m\langle \vec{v}^2 \rangle}{2} = \frac{3}{2}k_B T \tag{4.20}$$

Here $\langle \vec{v}^2 \rangle$ is the mean velocity squared, k_B is Boltzmann's constant and T is temperature in Kelvin.

In order to keep the temperature constant, one may scale the velocities of particles. However, this is unphysical, as the canonical partition function will not be generated then. Whenever velocity is scaled, it creates a discontinuity in the systems trajectory through phase space. A better approach is to use the Nosé-Hoover thermostat discussed in detail in section 4.2.4. This approach does generate the NVT sampling and will correspond to the sampling involved in the canonical partition function if the temperature control is ergodic. The ergodicity hypothesis simply says that the ensemble average is equal to the time average. Basically, this means that if a system in MD is observed over a long enough period of time, it will visit all possible positions in phase space. The partition function could then easily be found by summing over $e^{-E/k_B T}$, where E is the energy of the microstate, and the result would be ergodic. However, this is usually not possible due to the immense volume in the 6N-dimensional phase space.

4.2.2 Periodic boundary conditions

Periodic boundary conditions (PBC), is an important concept in molecular dynamics. The idea is to surround the simulation system with its own replicas. This is usually done on all sides of the system; however, it can on occasions be done in only one or two dimensions. The reasoning behind PBC is to avoid two general problems (Hinchliffe 2003). The first problem is that if rigid walls surround the simulation system, then the molecules or atoms close to the wall will experience a different force than the molecules or atoms near the center of the system. If the goal is to simulate the behavior of a bulk liquid then clearly, the interactions with the surrounding walls are undesired. Also by imposing rigid walls in the simulation the ability to replicate larger systems are somewhat lost. Mainly due to the lack of processing power it is not possible at this point in time to simulate a real-life molecular system, like for instance one liter of liquid water. The overwhelming number of molecules, and thus atoms, would demand an impressive amount of computational power.

The second problem with not using PBC is that a decrease in density can occur. Consider a simulation box, without rigid walls. A number of molecules or atoms are evenly positioned in this box. At some point in time, a molecule or atom can cross over the edge of the box and into oblivion. This affects the density of the system since there is now one atom or molecule

less remaining in the simulation box. Since density is defined as mass divided by volume, and mass is number of atoms multiplied by the molecular weight, when the number of atoms in the simulation box decreases, the density decreases as well. With PBC, any atoms leaving one side of the simulation box instantaneously enters from the other side from an image of the true system. This is illustrated in Figure 4-1. Thus, the density of the simulation system remains constant as the number of atoms remains constant, given that the volume is constant.
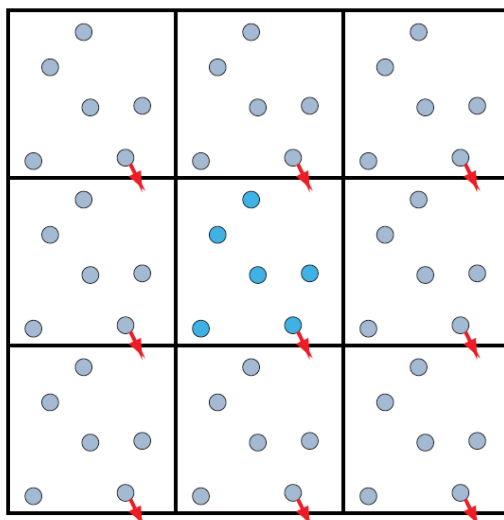
Figure 4-1: Periodic boundary condition.

In the two dimensional representation of PBC (Figure 4-1) the real system is surrounded by identical copies in all directions. Thus, in a 3D setup 26 copies of itself would surround the real system. Ideally for the system in this thesis PBC would only be required in x and y dimensions. This is due to the fact that in the z dimension the lower boundary is the AC model, and the top should be the liquid solution. As this was not an option in MD51, PBC were utilized in all 3 dimensions.

4.2.3 The Leapfrog algorithm

The Leapfrog algorithm is one of several algorithms used to solve evolution in MD, the equations hence determines the evolution of positions and momenta. The Verlét algorithm is probably the best known algorithm. However, Leapfrog algorithm is considered to be one of the most stable as well as accurate techniques in molecular dynamics. Leapfrog is used to solve Newton's equation of motion (4.15). This equation is a second order differential equation, but it can easily be written as two first order differential equations (4.21) and (4.22).

$$\vec{F} = m\frac{d\vec{v}}{dt} \qquad (4.21)$$

In equation (4.21) $\vec{F}$ is the force; m is the mass of the particle; $\vec{v}$ is velocity and t is time. In the equation below, $\vec{r}$ is the position vector.

$$\vec{v} = \frac{d\vec{r}}{dt} \qquad (4.22)$$

For a particle A, the velocity is $\vec{v}_A(t)$. By using two Taylor expansions for $\vec{v}_A(t)$, (4.23) and (4.24), and subtracting them we get an algorithm for numerical integration in small time steps (4.25). Δt is the time step, or change in time.

$$\vec{v}_A\left(t + \frac{\Delta t}{2}\right) = \vec{v}_A(t) + \left(\frac{d\vec{v}_A}{dt}\right)_t \frac{\Delta t}{2} + \frac{1}{2}\left(\frac{d^2\vec{v}_A}{dt^2}\right)_t \left(\frac{\Delta t}{2}\right)^2 + \cdots \qquad (4.23)$$

From the Taylor expansions the velocity (or more precise, an approximation of the velocity) at time t + Δt/2 can be found from the velocity at t, pluss the time-derivative of the velocity (acceleration) multiplied by Δt/2, pluss the time-derivative of the acceleration multiplied by $(\Delta t/2)^2$, and so on.

$$\vec{v}_A\left(t - \frac{\Delta t}{2}\right) = \vec{v}_A(t) - \left(\frac{d\vec{v}_A}{dt}\right)_t \frac{\Delta t}{2} + \frac{1}{2}\left(\frac{d^2\vec{v}_A}{dt^2}\right)_t \left(\frac{\Delta t}{2}\right)^2 + \cdots \qquad (4.24)$$

Similarly, the velocity at time t - Δt/2 can be found from equation (4.24).

$$\vec{v}_A\left(t + \frac{\Delta t}{2}\right) = \vec{v}_A\left(t - \frac{\Delta t}{2}\right) + \left(\frac{d\vec{v}_A}{dt}\right)_t \Delta t + \cdots \qquad (4.25)$$

Equation (4.25) gives the velocity at time t + Δt/2 from the velocity at t - Δt/2 and the acceleration times Δt. Usually, eq. (4.25) and (4.28) are truncated after second term on the right hand side, and this introduces a truncation error in the order of $(\Delta t)^3$. The acceleration is calculated from Newton's second law (4.15). Using the same procedure as above similar Taylor expansions are made for $\vec{r}_A(t)$, (4.26) and (4.27).

$$\vec{r}_A\left(t + \frac{\Delta t}{2}\right) = \vec{r}_A(t) + \left(\frac{d\vec{r}_A}{dt}\right)_t \frac{\Delta t}{2} + \frac{1}{2}\left(\frac{d^2\vec{r}_A}{dt^2}\right)_t \left(\frac{\Delta t}{2}\right)^2 + \cdots \qquad (4.26)$$

$$\vec{r}_A\left(t - \frac{\Delta t}{2}\right) = \vec{r}_A(t) - \left(\frac{d\vec{r}_A}{dt}\right)_t \frac{\Delta t}{2} + \frac{1}{2}\left(\frac{d^2\vec{r}_A}{dt^2}\right)_t \left(\frac{\Delta t}{2}\right)^2 + \cdots \tag{4.27}$$

Again, by subtracting equation (4.26) and (4.27) a new equation (4.28) gives the position at time t + Δt.

$$\vec{r}_A(t + \Delta t) = \vec{r}_A(t) + \vec{v}_A\left(t + \frac{\Delta t}{2}\right)\Delta t + \cdots \tag{4.28}$$

Here it is easy to see why it is called the Leapfrog algorithm. The velocity is calculated for t + Δt/2. However, the position is calculated for t + Δt. Thus, the velocity leapfrogs over the position, in time, and likewise the position leapfrogs over the velocity in time. If needed or desired, the velocity can be calculated by equation (4.29) in order to get the velocity and position at the same time.

$$\vec{v}_A(t) = \frac{\vec{v}_A\left(t + \frac{\Delta t}{2}\right) + \vec{v}_A\left(t - \frac{\Delta t}{2}\right)}{2} \tag{4.29}$$

4.2.4 Nosè-Hoover thermostat

As mentioned earlier, the temperature of the system is proportional to the (average) kinetic energy of the particles in the system (eq. (4.20)). Therefore, to keep the temperature constant, the average kinetic energy has to be kept constant as well. The simplest method for keeping the temperature constant is through velocity scaling. This method simply applies that the velocities are all scaled to correspond with the given temperature at regular times, or when the deviation of the average temperature from the given temperature exceeds a set tolerance. This is considered unphysical, as there is no guarantee that the sampling corresponds to the canonical ensemble for the system.

 A better solution was presented by S. Nosé (Nose 1984). He connected the simulation to a heat-bath, and allowed heat to be transported to and from the simulation system. This was done by extending the Hamiltonian with an additional degree of freedom, and reformulating the equations of motion. The extended Hamiltonian considered the simulation system as well as the heat-bath to make up the complete system. Through this the conservation law (of energy) holds for the complete system, but the total energy of the simulation system is allowed to fluctuate by exchanging energy with the heat-bath (Nose 1984). William G.

Hoover simplified Nosé's work in a paper presented in 1985 (Hoover 1985). An undesired side effect of the original proposal from Nosé was that the time itself in the simulation, which is used to evaluate kinetic properties and time-averages, was variable. This made the implementation of the method more difficult. Hoover simplified the equations presented by Nosé in such a way that time scaling was avoided, and as such made the task of implementation easier (Hoover 1985).

The following equations are the Nosé-Hoover thermostat. Equation (4.30) is the Nosé-Hoover Hamiltonian. Note that it is not a true Hamiltonian, since the equations of motion cannot be derived from it (Frenkel and Smit 2002). The two last terms on the right hand side corresponds to an added degree of freedom (Nose 1991).

$$\mathcal{H}_{NH} = \sum_i \frac{\vec{p}_i^2}{2m_i} + \Phi(\vec{q}) + \frac{\xi^2 Q}{2} + g k_B T \ln s$$

(4.30)

Here $\mathcal{H}_{NH}$ is the Nosé-Hoover Hamiltonian; m_i is the mass of particle i, $\vec{p}_i$ is the momentum for the particle; $\Phi(\vec{q})$ is a potential energy of a system; ξ is the friction coefficient; Q is a parameter that acts as a mass for the motion of s and the temperature of the heat-bath; N is the total number of particles; g is the number of degrees of freedom; k_B is the Boltzmann constant; T is the temperature and s is a scaling factor.

The Nosé-Hoover equations of motion are as follows:

$$\frac{d\vec{q}_i}{dt} = \frac{\vec{p}_i}{m_i}$$

(4.31)

In equation (4.31) $\vec{q}_i$ is the position of particle i; $\vec{p}_i$ is the momentum of particle i, and m_i is the mass of particle i. Equation (4.32) gives the change of momentum with time to be

$$\frac{d\vec{p}_i}{dt} = -\frac{\partial \phi}{\partial \vec{q}_i} - \xi \vec{p}_i$$

(4.32)

Where ϕ is the potential energy and ξ is the friction coefficient. It is important to note that the friction coefficient is a variable. This variable can have both positive and negative values. From the equation, it is clear that a positive value of ξ will lead to a reduction of the momentum, thus slowing the particle down. Similarly, a negative value of ξ will accelerate the particle. It is through this friction coefficient that the Nosé-Hoover thermostat controls the temperature.

Equations (4.31) and (4.32) are in effect the equations of motion with the inclusion of a friction force (Nose 1991).

$$\frac{d\xi}{dt} = \frac{\left[\sum_i \frac{\vec{p}_i^{\,2}}{m_i} - gk_BT\right]}{Q}$$

(4.33)

For this set of equations ((4.31)-(4.34)), which are utilizing real variables; g has the value of 3N. In the equation above, (4.33), one can see how the friction coefficient varies with time. It depends on the sum over the momentum for all the particles divided by the mass of each particle. This would be the total kinetic energy of the system. The average kinetic energy is then subtracted from this value. By dividing the difference between the actual kinetic energy and the desired kinetic energy by Q, the change in the friction coefficient with time is calculated. For a large Q (parameter acting as a mass) a slower change will occur. A small Q will lead to rapid changes. The actual value for Q must be carefully chosen in the implementation. Too small value for Q and the system is not guaranteed to equilibrate (Nose 1991). Too large value for Q would be inefficient as the system would spend too long to equilibrate. Furthermore, the value for Q affects the entropy generation in a much greater degree than the energy sampling. This implies that even if the energy sampling might not be significantly affected by the value of Q, the generation of entropy may be significantly affected. In turn, the resulting free energy of the system will be affected by the choice of Q. The value of Q can be estimated from the dependence of Q on the system size and temperature. In MD51 Q is indeed estimated in the software depending on the system, as such it is not an input parameter that the end-user has to calculate manually.

$$\frac{d \ln s}{dt} = \xi$$

(4.34)

ξ, the friction coefficient, can be found from two equations, (4.33) and (4.34). It is, however, usual to get the friction coefficient from equation (4.33). Equation (4.34) is not needed, as equation (4.31) to (4.33) form a closed set. But, it can be used as a diagnostic tool, since the Nosé-Hoover Hamiltonian (equation (4.30)) has to be conserved throughout the entire simulation (Frenkel and Smit 2002).

4.2.5 Intra-molecular interactions

In every molecule, there are contributions from intra-molecular interactions. These interactions include interactions such as bond stretching, angle bending, twisting of torsions

and some cross interactions. A cross interaction is an interaction that involves for instance both bond stretching and angle bending at the same time, which yields a different interaction energy than the sum of the individual contributions. Other intra-molecular interactions can be electrostatic or Lennard-Jones interactions between atoms separated by more than three bonds.

However, this section limits itself to the three main types of intra-molecular interactions. These are the bond stretching, angle bending and twisting of torsion angles.

4.2.5.1 Bond stretching

The covalent bonds in a molecule can be modeled as a spring-and-balls model. In this model the two atoms that are bonded together are represented as balls. The bond itself is represented by a spring. Below, Figure 4-2 shows a sketch of how the model envisions a covalent bond between two atoms.

Figure 4-2: A covalent bond modeled by the spring-and-ball model.

This spring has an equilibrium length; this would correspond to the (average) bond length that this bond has. At equilibrium length there is no contribution to the internal energy from bond stretching, as the equilibrium length equals the actual bond length. This results in that the parenthesis in equation (4.36) is zero, and there is no contribution to the potential energy from this bond. The strength of the bond is given by the strength of the spring. In MD51 Hook's law (4.35) is used to model the bond, which is consistent with OPLS.

$$\vec{F}_s(\vec{R}) = -k_s(\vec{R} - \vec{R}_e) \qquad (4.35)$$

Here $\vec{F}_s$ is the force as a function of the length of the spring $\vec{R}$ (or distance between the atoms); k_s is the force constant and $\vec{R}_e$ is the equilibrium length of the spring (or bond). The potential energy, U, is given by equation (4.36). It is now easy to see that equation (4.35) is related to equation (4.36) by the negative gradient of the latter to obtain the former.

$$U(\vec{R}) = \frac{1}{2}k_s(\vec{R} - \vec{R}_e)^2 \qquad (4.36)$$

This type of bond is called a *harmonic bond* due to the harmonic oscillation of the solution to the equations above. Note that both OPLS and MD51 combines the constant factor of ½ and the force constant into one force constant.

4.2.5.2 Angle bending

Similar to bond stretching there is also contributions from the bending of angles. The angle bending is also modeled by a spring model. An angle is defined between three atoms, which are bonded together. The spring is then visualized between the two outermost atoms. Hook's law models the stretching and compression of this spring.
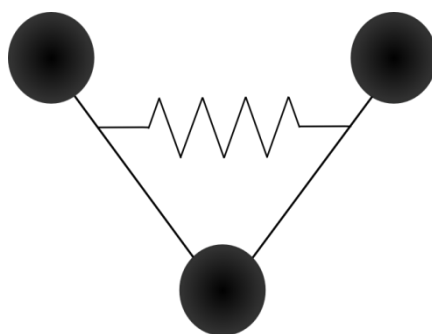
Figure 4-3: Balls and spring angle bend model.

Figure 4-3 visualizes how the spring is working with the angle bending. If the angle is increased or decreased, then the length of the spring will deviate from the equilibrium length and a force working in the opposite direction will occur.

$$U_{ang}(\theta) = \frac{1}{2}k_\theta(\theta - \theta_e)^2 \qquad \text{(4.37)}$$

Equation (4.37) gives the harmonic potential for a covalent angle in MD51. The angle is θ and the equilibrium angle is θ_e, the force constant for the angle is k_θ. When comparing eq. (4.37) with eq. (4.36), the similarities to the covalent bonds are obvious. Again, the force can be found by taking the negative gradient of equation (4.37). Note that both OPLS and MD51 combines the constant factor of ½ and the force constant into one force constant.

4.2.5.3 Torsions

Torsions are the angles between four atoms bonded together in a chain A-B-C-D. In the ball-and-spring model, one envisions two planes. One formed from A-B-C and one from D-C-B. The spring is then pulling or pushing these planes together or apart, depending on the current angle and the parameters for the torsion.

There are several different options for torsions in MD51. The equation that the OPLS force field is using is a Fourier series. The parameters given in the force field is the coefficients in front of the cosines. An option exists to also specify the phase angels, but in this system they are set to zero. This is due to the fact that the phase angles were not given in the force field parameter file, and thus assumed to be zero. Equation (4.38) shows how OPLS treats torsions. This equation reduces into equation (4.39) when the phase angles are set to zero, which is the MD51 torsion potential for the MM3 type torsions.

$$E_{torsion} = \sum_i \frac{V_1^i}{2}\left[1 + \cos\left(\phi_i + f_1^i\right)\right] + \frac{V_2^i}{2}\left[1 - \cos\left(2\phi_i + f_2^i\right)\right]$$
$$+ \frac{V_3^i}{2}\left[1 + \cos\left(3\phi_i + f_3^i\right)\right]$$

(4.38)

Here, V_1^i, V_2^i and V_3^i is the Fourier parameters for atom i, that are given in the force field. ϕ_i is the torsion angle, and $f_x^i \in [1,2,3]$ is the phase angle which should also be given by the force field.

$$E_{torsion} = \sum_i \frac{V_1^i}{2}[1 + \cos\phi_i] + \frac{V_2^i}{2}[1 - \cos 2\phi_i] + \frac{V_3^i}{2}[1 + \cos 3\phi_i]$$

(4.39)

The latest version of OPLS have removed the phase angles, but included a fourth term in the Fourier series above. Thus, $\frac{V_4^i}{2}[1 - \cos 4\phi_i]$ should ideally be added to equation (4.39). However, for all the torsions in the system in this thesis the fourth Fourier parameter was 0. Modifying the code to accommodate the extra term was then considered unnecessary, as there would be no contributions from this term when using the OPLS parameters.

4.2.6 Inter-molecular interactions

Inter-molecular interactions are non-bonded interactions. They occur between molecules or atoms, or a mixture of both. In MD51 there's two different interactions that make up the inter-molecular interactions. These are the van der Waals forces and the electrostatic forces. The van der Waals forces are short-range forces, while the electrostatic forces are long-range forces. The ranges of the forces are decided by how quickly they diminish.

4.2.6.1 Van der Waals type forces and Lennard-Jones 12-6 potentials

The short-ranged interactions or forces are generally known as van der Waals forces. There are three main contributions to the van der Waals forces. The three forces are the Keesom force, the Debye force and London dispersion forces.

Keesom force is the contribution from the interaction of two polar molecules. These are the forces from dipole-dipole interactions, or quadrupole-quadrupole interactions or octupole-octupole interactions. The moments in question are permanent moments and not induced ones. The abovementioned interactions can be described by a sum of a multipole expansion. The forces resulting from these interactions can be attractive or repulsive, depending on the orientation of the multipoles. The largest contribution will be from a dipole-dipole interaction.

The Debye force is from the interaction of a permanent dipole which sets up an electric field, which in turn can induce a dipole moment in another molecule. An induced dipole can come in addition to, or reduce the effect of a permanent dipole moment.

The third force, London dispersion force, stems from induced dipole – induced dipole interactions. These arise from the electrons moving in the molecules. This movement can and will make small instantaneous dipoles. As a dipole is created, an electric field is set up. The electric field then induces a dipole in a nearby molecule. These dipoles are short-lived, and often have constantly changing directions.

To model the van der Waals force there exist a number of different potentials. The Lennard-Jones 12-6 potential (4.40) is perhaps the most known of them all. This is also the potential used in MD51.

$$U_{LJ} = \sum_{non-bonded} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right]$$

(4.40)

This potential has three main elements. The well depth ϵ_{ij}, the repulsive term which is the first term inside the square bracket and the attractive term which is the second term in the square bracket. The well depth gives the strength of the interactions, or the maximum possible attractive force to be more precise. The parameter r_{ij} is the distance between atoms i and j. In MD51 the sum goes over all non-bonded atoms or molecules, as denoted by the indexes i and j. The practical use of this interaction model is as "effective" short-range pair interactions. Physically a third atom in contact with i and j will have an impact on the short-range

interactions between i and j. Similarly for the effects of more molecules surrounding i and j. All effects of multiple neighbors are thus considered to be "lumped" into the energies expressed through the model parameters in (4.40). The parameter σ_{ij} is a distance parameter (Prausnitz et al. 1999). This is where the potential energy is zero due to the repulsive and attractive forces balancing out. Figure 4-4 shows the Lennard-Jones potential for Argon. The well depth ϵ for Argon (in OPLS) is 0.979056 kJ/mole; by multiplying this parameter with -1 the minimum energy of the interaction is found. This corresponds to the equilibrium separation distance. A red dotted line in Figure 4-4 shows the minimum energy. Similar, σ is shown by a green line in the same figure. σ has the value of 3.401 Å for Argon in the OPLS force field.
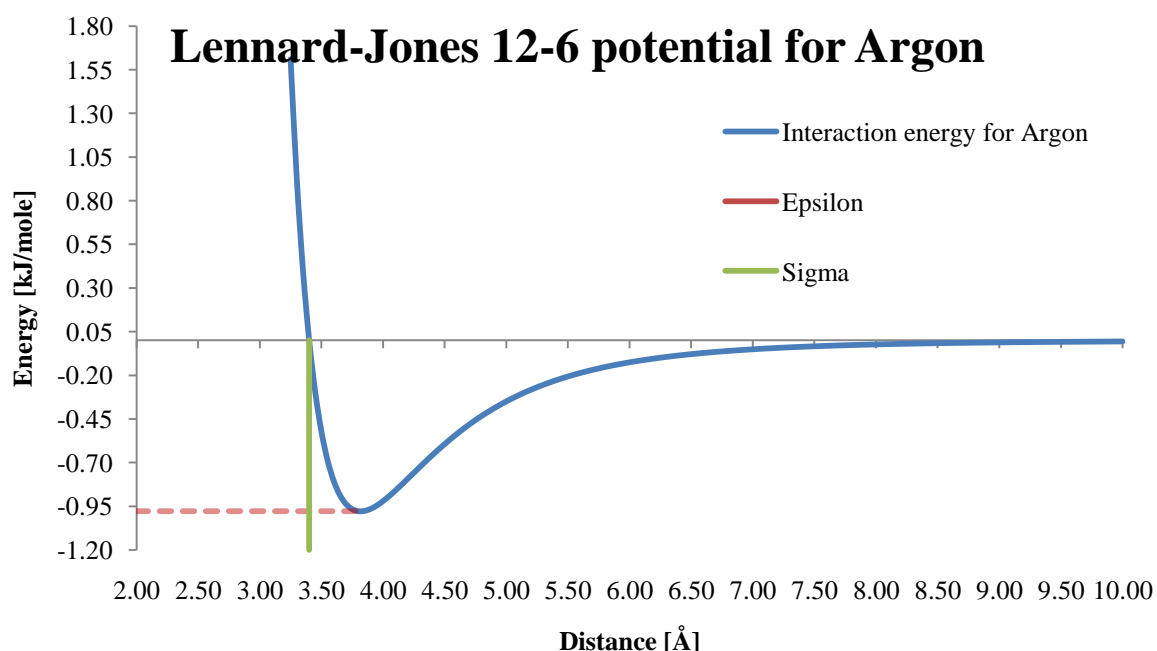
Figure 4-4: Lennard-Jones potential energy for Argon as a function of center-to-center distance.

The repulsive term would be better represented by an exponential function. However, the form r^{12} was selected due to the simplicity of computing this number as the square of r^6. This saves quite a bit of computational power, as the exponential function is demanding to compute. The repulsive force is due to that two electrons with the same quantum number cannot occupy the same space. This is known as the "Pauli exclusion principle", and it states that "No more than two electrons may occupy any given orbital, and if two do occupy one orbital, then their spins must be paired." (Atkins and de Paula 2002).

4.2.6.1.1 Lorentz-Berthelot rule

To calculate the Lennard-Jones potential energy, two parameters are needed. These are the well depth ϵ and the distance parameter σ. In literature, most of the parameters available are for pure substances. Different experimentally measured properties can be used for obtaining these parameters for unlike species. Examples include gas phase viscosity, energy, density and gas/liquid coexistence curve. But, as indicated above, the values for σ and ϵ are normally only available for the interaction between two atoms of the same type, for example Argon-Argon.

Considering the number of different atoms there are a lot more potential cross-interactions between two molecules of different species. To calculate these parameters by regression on experimental data is quite a task. As such, it is not normal to do so. Instead, mixing rules, or combination rules are used to estimate the values for the parameters. Lorentz-Berthelot rule (equation (4.41) and (4.42)) is one such way of estimating the parameters for unlike atoms. This is also the default way of estimating these unlike parameters in MD51.

$$\sigma_{ij} = \frac{1}{2}\left(\sigma_i + \sigma_j\right) \tag{4.41}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \tag{4.42}$$

In the Lorentz-Berthelot rule the unlike distance parameter is obtained by taking the arithmetic average of the distance parameter for each species. The energy parameter, or well depth, is obtained by taking a geometric average (Kuznesova 2001). Average excluded diameter according to equation (4.41) is considered to be reasonable, while (4.42) is based on experimental observations close to the critical region (Croxton 1975).

4.2.6.2 Electrostatic interactions and Ewald summation

An atom consists of nucleus and the electrons that surround the nucleus. This is most often visualized using the orbital-model. The atom in its normal state is electrostatic neutral. I.e. the number of electrons matches the number of protons in the nucleus. Electrons and protons have charges of equal magnitude but opposite polarity. Thus, they balance each other and the net charge is zero. In a molecule, two or more atoms are bonded together. Some of these atoms, like oxygen in H_2O, may have a greater electronegativity than other atoms, like the hydrogen atoms in H_2O. As a result, the electrons that belong to the two hydrogen atoms tend to spend more time, on average, on the oxygen-side of the hydrogen. This leads to a shift in electron

distribution where the oxygen atom appears to be slightly negative, and the hydrogen atoms appears to be slightly positive.

Because of this partial atomic charges were coined. Partial atomic charges are the charge that an atom appears to have when it is in a molecule. While the molecule itself is electrostatic neutral, parts of the molecule can have a (partial) charge, either positive or negative. Thus, an electrostatic neutral molecule might have a dipole moment, a quadrupole moment or higher moments. Between charged particles there are forces acting. Thus, the partial charge on an atom will give rise to a force between the atom, and other (charged) atoms in the vicinity. The same goes for a molecule with a dipole moment, or higher moments, there will be a force between this and any other charged particle in the vicinity. Examples of particles here would be another molecule, or an atom.

To calculate these forces Coulomb's law is used. Coulomb's law (4.43) gives the electrostatic force between two point charges. The magnitude of the force is dependent on the distance between the charges, the magnitude of the charges and the permittivity of the medium between the charges. A positive force implies that the charges are repelling each other, similar a negative force implies attraction.

$$\vec{F}_{AB} = \frac{1}{4\pi\epsilon_r\epsilon_0} \frac{Q_A Q_B}{R_{AB}^3} \vec{R}_{AB}$$

(4.43)

In Coulomb's law $\vec{F}_{AB}$ is the force vector; ϵ_0 and ϵ_r are the permittivity in vacuum and the relative permittivity, respectively; the Qs are the charges on the particles (A and B); R_{AB} is the distance between the particles and $\vec{R}_{AB}$ is the vector from A to B. Since both the permittivity in vacuum and the relative permittivity are positive quantities it is clear that if both A and B are of the same polarity the resulting force will be positive. Likewise if A and B have different polarities the force will be negative and the two point charges will be attracted.

In MD, the force field does not give the forces directly. However, it does give the potential energy. By taking the negative gradient of the potential energy the force is acquired, eq. (4.16). Thus, the equation used in MD is for the potential energy between two point charges, not the force. The equation used in MD51 for the electrostatic interactions between atom pairs is equation (4.44). Here the sum is over all possible non-bonded atoms, and r_{ij} is the distance between atom i and j.

$$U_{el} = \sum_{non-bonded} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

<div align="right">(4.44)</div>

This is however a truth with some modifications. In reality, it would be too costly to calculate the (electrostatic) potential energy between all possible atoms, and the neighboring images (from PBC). Thus, a cut-off distance is introduced. If the distance between two atoms is greater than the cut-off distance then equation (4.44) is not evaluated. Instead, the electrostatic interaction energy contribution is set equal to zero. This cut-off distance approach has some unphysical effects though, and the inaccuracy is quite severe (Frenkel and Smit 2002). It is possible to use Ewald summation instead, and this is done in MD51. Ewald summation method is a very common way of calculating electrostatic interaction in a system with periodic boundary conditions.

Ewald summation is able to calculate the electrostatic interaction exactly, to a numerical threshold (Jensen 2007). The Ewald summation method works by surrounding every charged particle by a diffuse charge distribution of the same magnitude as the charge, but of opposite sign (Frenkel and Smit 2002). This will then cancel the charge q, and turn the interaction into a short-range interaction as some of the charge is screened. Next, it is needed to compensate for the charge distribution that was introduced. In order to do so a new charge distribution, equal in magnitude to the first but of opposite signs, is added. This compensating charge distribution is both periodic and of a smooth nature. A Fourier series can then be used to represent this. This series should be rapidly converging.

The electrostatic potential is now made up of three contributions. Namely, one contribution due to the point charge, one contribution due to the screening charge distribution and one from the compensating charge distribution. The Coulomb self-interaction will also have to be considered. This interaction is due to images of the charge itself in the replica systems from the PBC.

$$U_{Ewald\ (Coloumb)} = U_{reciprocal} - U_{self} + U_{short-range}$$

<div align="right">(4.45)</div>

Here, $U_{Ewald\ (Coloumb)}$ is the total electrostatic contribution to the potential energy; $U_{reciprocal}$ is the contribution from Fourier space, or reciprocal space; U_{self} is the contribution from self-interaction and $U_{short-range}$ is the contribution from the screened point charges. Each of these contributions are given in equations (4.46), (4.48) and (4.49) (Frenkel and Smit 2002).

$$U_{reciprocal} = \frac{1}{2V} \sum_{k \neq 0} \frac{4\pi}{k^2} \left| \rho(\vec{k}) \right|^2 e^{-k^2/4\alpha}$$

(4.46)

In equation (4.24) V is the volume; $\vec{k} = \left(\frac{2\pi}{L} \right) \vec{l}$ where $\vec{l} = (l_x, l_y, l_z)$ being the lattice vector in Fourier space; $\rho(\vec{k})$ is the charge density given by equation (4.47) and α is parameter for the width of the Gaussian describing the compensating charge distribution.

$$\rho(\vec{k}) \equiv \sum_{i=1}^{N} q_i e^{(i\vec{k} \cdot \vec{r_i})}$$

(4.47)

For the charge density given above, equation (4.47), q_i is the charge of particle i; $\vec{r_i}$ is the position of particle i, and N is the total number of particles.

$$U_{self} = (\alpha/\pi)^{1/2} \sum_{i=1}^{N} q_i^2$$

(4.48)

The α-parameter in equation (4.48) is the same as in equation (4.46). This is due to the interaction between a continuous Gaussian charge distribution and the point charge q_i.

$$U_{short-range} = \frac{1}{2} \sum_{i \neq j}^{N} \frac{q_i q_j \, erfc(\sqrt{\alpha} r_{ij})}{r_{ij}}$$

(4.49)

The function erfc() in the equation for the short-range contributions (4.49) is the complementary error function.

By utilizing Ewald summation for the Coulombic interactions the computational time is reduced, while the accuracy can be kept at an acceptable level. Ewald summation scales as $N^{3/2}$, while Coulomb's law scales as N^2, where N is the number of particles.

4.2.7 The addition of a constant force

During the assembly of the system, it quickly became obvious that the normal schemes for positioning molecules into the simulation system were somewhat lacking. One issue was that the AC model had to be accurately positioned to get the correct PBC. Since it was desirable to have the PBC to replicate the AC model in X and Y dimension, without any gaps, it could not be randomly positioned. Thus, the program written to assign the charges to the AC model was modified. The new modification shifted the AC models center of mass (COM) to origin. This

was done, as MD51 requires that the COM is positioned at the origin. By doing so, the accurate positions of each atom in AC model were known.

The next move was to add the liquid solution of triolein and PCBs on top of, but a distance away from the AC model. When the simulation was started, the system was allowed to evolve through time. The idea was that the liquid solution would settle down on top of, and in the pores of the AC model. However, this did not happen. Instead, the liquid expanded and occupied all available space in the system. In order to move the solution down into the pores of the AC model, a constant force had to be added. This force was to operate only in the z dimension, such that the molecules would be pushed down into the pores.

MD51 had an option to apply an external electric field to work on the system. After a short discussion with Prof. Kuznetsova, the decision was made to try to utilize this external field in such a manner that the two parts of the system could be successfully merged. Upon reviewing the actual code in MD51 it soon became clear that the code, such as it was, would not do the job correctly.

As the code was written for an electric field, the force it applied to the atoms in the molecules was dependent on the partial atomic charges. To use the code as it was, would mean that molecules would rotate into favorable positions with respect to the electric field. The effect of such an ordering would be completely unphysical. Additionally, the equilibration time of the system would likely have had to been increased by a great deal, in order for the system to "forget" the starting position. The original code is displayed in the textbox below.

```
*   External electrostatic field
if ( IEXT.eq.1 ) then
        FULTIM = TIM + NSTEP * DT
        EFIELD = EXAMPL * COS( EXFREQ * FULTIM * 0.5 / PI )
        do I = NAB( TASKID ) , NAE( TASKID )
                IS = NSITE( I )
                GZ( I ) = GZ( I ) + CHARGE( IS ) * EFIELD
        end do
end if
```

This code then had to be modified in such a way that it would apply a force of the same magnitude to all atoms, in all molecules. The simple modification to the code is shown below.

```
*   External electrostatic field
if ( IEXT.eq.1 ) then
        EFIELD = EXAMPL
        do I = NAB( TASKID ), NAE( TASKID )
                IS = NSITE( I )
                GZ( I ) = GZ( I ) + EFIELD
        end do
end if
```

There are only a few modifications done here. First, the calculation of the variable FULTIM was removed, as it was no longer needed. This variable was included to make the field time dependent. Secondly, the expression for EFIELD was reduced to simply equal the EXAMPL (external amplitude, a parameter given in the md.input file). The EFIELD could have been completely removed and replaced by only EXAMPL, but the line was kept to simplify future modification or scaling of the external amplitude. The final modification was to remove the influence from the partial atomic charge on the force applied to the atom. This was simply done by removing the CHARGE(IS) variable that was multiplied with the EFIELD. In sum, these minor modifications made it possible to apply a constant external force on each molecule.

The solution that was implemented here is not as elegant as desired. Instead of a constant force, a constant acceleration might be a better option. For future modification, the goal should be to give the force in acceleration in the input file. This would be more intuitive, and the force acting on each atom can then be calculated from $F_z = ma_z$.

The constant force field is activated the same way as the time dependent electric field. That is, by the entry of "El_field *field_strength frequency*". The frequency is ignored in the modification done, so an arbitrary frequency should be included to avoid an error message.

5 Modeling the system

Current section deals with the actual construction of the individual models, as well as parameterization of the models. Some general theory for the molecules to be model has been included, along with information on the molecular structures. The first subsection describes the activated carbon, its properties and how the model was built, along with parameters and partial atomic charges. The remaining two subsections describe the model for the fish oil and the PCB models, respectively.

5.1 Activated carbon

Activated carbon (AC) is a widely used, low cost adsorbent. AC is not a single uniform substance, but rather a descriptive group of substances. All ACs consists for the most part of carbon atoms. In order to make a carbonous material into AC it has to go through a carbonization process, as well as an activation process. Heating the carbonous material to high temperatures, but below 800°C, in an inert atmosphere removes most of the oxygen, nitrogen and hydrogen atoms (Bansal and Goyal 2005). This process is called pyrolysis. The activation process can be thermal, physical or chemical. During the activation process, the porosity of the carbonous material is increased. The pore distribution is also affected by the increase in porosity.

Activated carbon has a large surface area to volume ratio. Due to the large surface area, AC has many possible adsorption sites. The pores present in AC are divided into three classes, micro-pores, meso-pores and macro-pores, where the pore diameter, d, is d < 2nm, 2nm ≤ d < 5nm and d ≥ 5nm respectively.

5.1.1 Molecular structure of activated carbon

Considering the number of different types of activated carbon on the market there is no easy way, if it is even possible, to define the molecular structure of AC in a clear and concise manner. AC has, by nature, an unordered structure. It is not a specific uniform molecule or a functional group. Thus, the normal ways for chemists to describe a substance has some limitations when it comes to AC.

Nevertheless, there are some key features common for AC. One is that AC is a carbonous material, as mentioned above. The name itself, activated carbon, implies as much. Hexagonal

rings make up most of the structure in AC, thus it can also be classified as a graphitic material. As further detailed in section 5.1.2, graphite consists of carbon atoms in layers of hexagonal ring structures. There is, however, a quite significant difference between AC and graphite. AC is highly unordered, while graphite is ordered in a crystal structure. The level of chaotic structure is a significant part of what gives AC its properties.

As stated above, AC consists mainly of carbon atoms. Oxygen, nitrogen and hydrogen are also often, if not always, present as well. These elements are referred to as impurities, or heteroatoms. Impurities themselves can play a significant role when it comes to the adsorption properties of AC. Heteroatoms in the AC can form different functional groups, which in turn can change the preference for different kinds of AC to adsorb different kinds of molecules (polar, non-polar) (Bansal and Goyal 2005). For instance, oxygen atoms in AC may easily form hydrogen bonds with the hydrogen atoms in Methane.

Figure 5-1: Electron microscope picture of granulated activated carbon.

Nevertheless, the main reason as to why ACs are such effective adsorbents are the pores and pore-distributions. The high porosity of AC leads to a very large surface area. Activated carbons can easily have surface areas of $1000 m^2 g^{-1}$ or more (Marsh and Rodríguez-Reinoso 2006). With such a large surface area, the possibilities for the adsorbate to be exposed to the adsorbent is large. The reason for the large surface area is mainly attributed to the micro-

38

pores, meso- and macro-pores contribute less to the total surface area. Some of the pore structure can be seen in Figure 5-1, where an electron microscope has been used to depict a sample of granulated activated carbon. When making ACs, carbonous material go through a carbonization process where the material is exposed to high temperatures in an inert atmosphere. This removes a large degree of the impurities and as such creates "holes" in the structure where those impurities were located. In the activation process these holes can be increased due to gasification of surface atoms (Marsh and Rodríguez-Reinoso 2006). Carbon atoms in the carbonous material may react with for instance oxygen atoms and turn into carbon monoxide or carbon dioxide. There are several other ways to increase the porosity and change the pore distribution.

5.1.2 Graphite as a model of activated carbon

Graphite in its natural form is a crystal consisting of carbon atoms with trigonal sp^2 bonding between the atoms (Pierson 1993). That is, each carbon atom is bonded to three other carbon atoms. The bonded carbon atoms form hexagonal rings.

There are two different crystal structures of graphite. Most common is the hexagonal graphite. Rhombohedral graphite is the other structure. This structure is thermodynamically unstable, and never found in pure form in the nature. Rhombohedral graphite is always found in combination with hexagonal graphite. Hexagonal graphite is stacked with a −ABABAB- stacking order. The A and B planes are identical, but offset such that the atoms does not reside directly above each other. Each plane is separated in z dimension by 3.348Å. Between two A-A or B-B planes, the distance in the z dimension is 6.696Å. The hexagonal graphite crystal structure was taken from "The American Mineralogists Crystal Structure Database" (Downs and Hall-Wallace 2003) and the entry in the database was from (Wyckoff 1963).
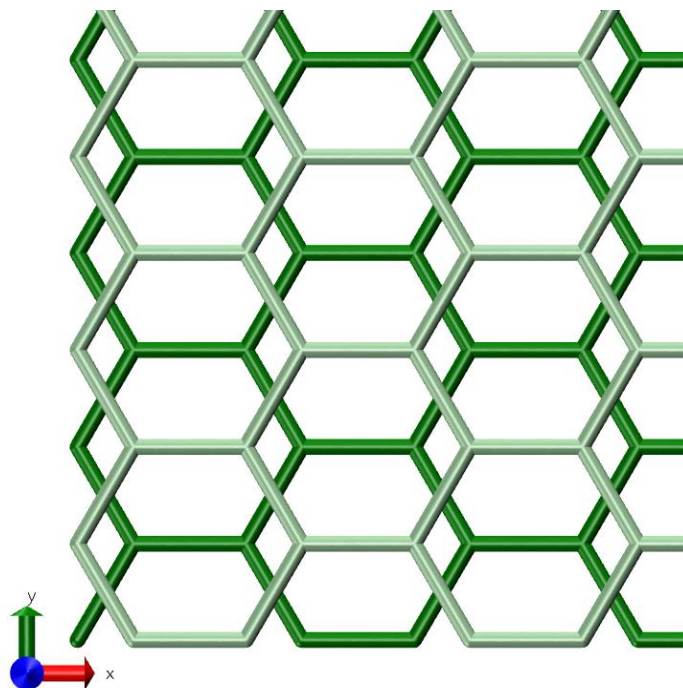
Figure 5-2: Hexagonal structure of graphite, A and B layers.

Figure 5-2 shows how the A and B planes are stacked on top of each other. The top most plane, the B-plane, is offset along the y-axis by 1.23Å compared to the A-plane (dark green). As mentioned above, the distance between two sheets is 3.348Å. A third sheet, which then would be an A-plane again, would be identically positioned in x and y coordinates as the first A-plane, however it would be 6.696Å away from the first one along the z-axis. Thus, there are no covalent bonds between the sheets of graphite. Covalent bonds in graphite is exclusively in the x and y dimensions in Figure 5-2. The force that keeps the sheets together in the z dimension is due to van der Waals forces.

In Figure 5-3, the structure of a graphite sheet has been decomposed into Cartesian coordinates, the bonds have been left out for clarity. Here, every other row (along the y-axis) has been given a different color, to make the symmetry more clear. It is easy to see that every other row is identical to each other. If we take a closer look at the x-axis coordinates for the first green row, the distance between the two first green atoms is 2.13Å − 0.71Å = 1.42Å, this is approximately the length of a carbon-carbon bond in graphite. However, between the second and third atoms, the distance is 4.97Å − 2.13Å = 2.84Å. This is twice the normal carbon-carbon bond and as such, there is no covalent bond between these two atoms.
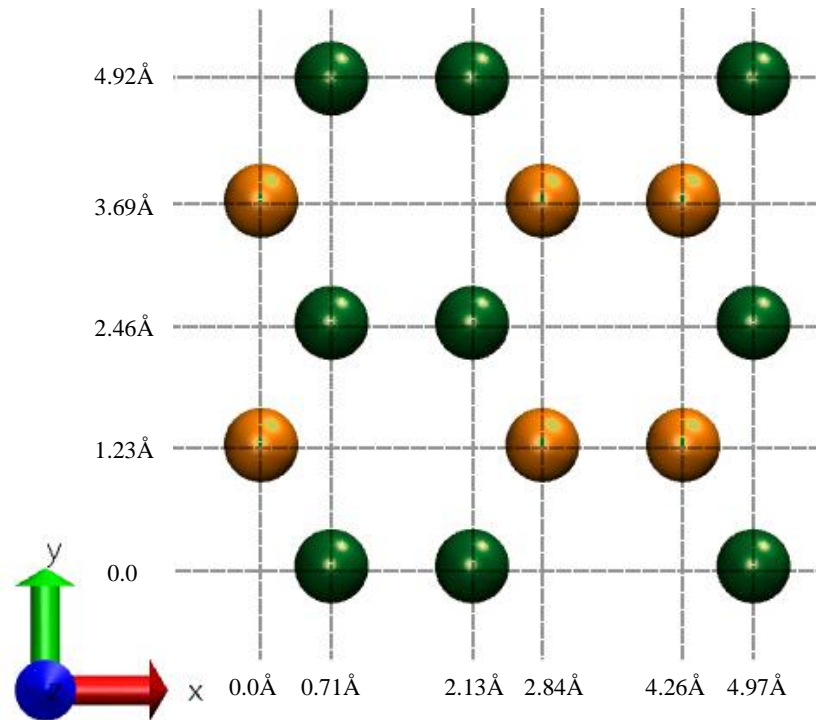
Figure 5-3: Graphite sheet in Cartesian coordinate system.

Similarly, for the first row of orange atoms, the distance between the first two atoms is 2.84Å. Between the second and third atoms, the distance is 4.26Å - 2.84Å = 1.42Å, again this corresponds to the carbon-carbon bond length in graphite. The distance between each of the green and orange rows is 1.23Å. The angle between any three bonded atoms in this structure is 120 degrees. To find the distance between the second green atom in the first row, and the second orange atom in the second row some minor calculation required. By subtracting the x-axis coordinate for the green atom from the x-axis coordinate from the orange atom the difference in x direction is found to be 2.84Å – 2.13Å = 0.71Å. Similar for the y-axis, the difference is 1.23Å. From the Pythagorean theorem the bond length is found to be

$$\sqrt{\left[(0.71\text{Å})^2 + \left(1.23\text{Å}\right)^2\right]} = 1.42\text{Å}.$$ Therefore, it is clear that all carbon-carbon bonds in graphite have a bond length of ~1.42Å. It is also clear that the green and orange rows are identical, but they are offset from each other along the x-axis by 2.13Å.

AC is a graphitic material. That is, it is not graphite, but it contains sections with graphite structure. However, there are some defects. These defects make AC chaotic and porous. Considering the fact that AC is a graphitic material, the idea of using graphite as model sprang to mind. Nevertheless, to model AC defects needed to be inserted in the graphite

41

structure. The decision made in this thesis was to model slit-pores. From studying the section on computer models of AC in Harry Marsh's book, Activated Carbon (Marsh and Rodríguez-Reinoso 2006), it soon became clear that computer models of AC is a field of knowledge all by itself. In the end, it was decided to make a very simplified model of AC, by simply removing parts of the planes above each other to get a slit-pore. However, by doing so quite a bit of the chaotic nature of AC would be lost. As such, in the section for further work there is an entry devoted to making this model more chaotic to get a better model of AC (9.1.3).

Several different models were considered. Initially the model had four pores, and utilized symmetry in both the x- and z-axis, Figure 5-4 (b). This would allow for a better utilization of the PBC. However, due to the sheer size of such a system, and the increased need for solvent molecules in order to avoid long-range electrostatic interactions between the AC model and itself this model was dismissed. A significant potential problem with this model would be the potential for long-range electrostatic interaction between the solution with itself across the graphite structure, thus making it more favorable or unfavorable for the solution to enter the pores. Another model that was considered was to simply use two sheets of graphite to make up a pore. This model was also cast aside due to that the effects of atoms further inside the AC model would not be represented, as these atoms would not be present in such a model. Furthermore, a pure sheet of graphite would have none of the defects of AC, other than the abrupt edges of the graphite sheets. Defects could have been introduced by using incomplete graphite sheets, and/or inclusion of impurities in the sheets. In the end, this model was considered lacking.

Finally, the model shown in Figure 5-4 (a) was selected. This model allows for interactions between the solution of triolein and PCB with the atoms in the pores, as well as long-range electrostatic interactions with atoms deeper inside the model. These interactions however, will be small compared to the interactions with the atom on the edge of the pore, due to the shielding effect these atoms have. Another feature of the model is that the pores can be considered to be defects in the graphite structure. This is because the graphite planes are not cut along any crystal faces. Also, note that in Figure 5-4 (a) there are no pores on the bottom part of the model, due to the aforementioned potential problems due to long-range electrostatic interactions. The final model ended up having 31232 carbon atoms. The Lennard-Jones parameters were taken from (Wongkoblap and Do 2008), where $\sigma = 3.4$Å and ε/k_BT was given to be 28K (0.2328 kJ/mol). QC simulations were used to obtain the partial

atomic charges for the atoms residing in close proximity to the pores. This is further detailed in section 5.1.3.
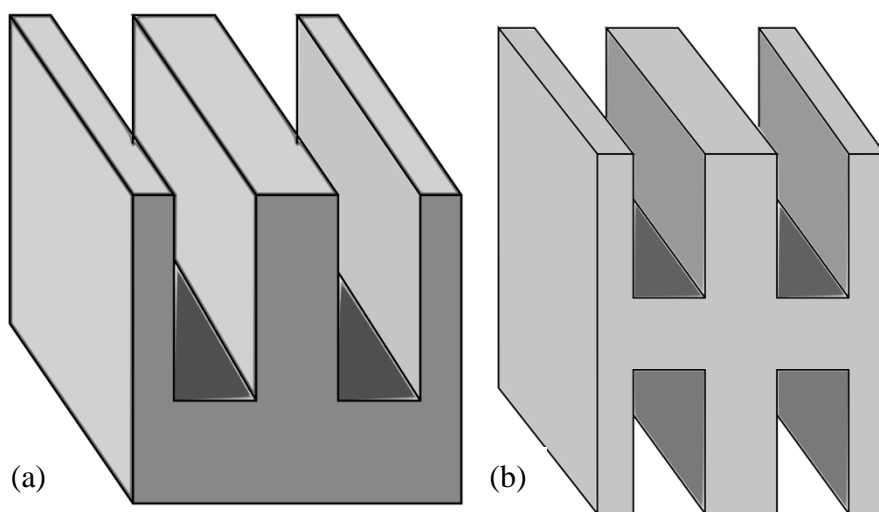
Figure 5-4: Sketch of suggested models. a) current model, b) rejected model.

5.1.3 Using Maestro/Jaguar to acquire the atomic charges

To obtain the charges for the graphite structure a small graphite crystal was used to perform some QM/QC calculations. The method utilized for the graphite was a single-point energy calculation, followed by an electrostatic potential (ESP) calculation. This was not done using SM6, so no charges for a water solvated system are available. Charges obtained using this method, are from vacuum calculations. Attempts were made to make Jaguar calculate charges from a water solvated model using SM6, unfortunately the model failed to converge after several attempts. Given enough time, this would likely have succeeded.
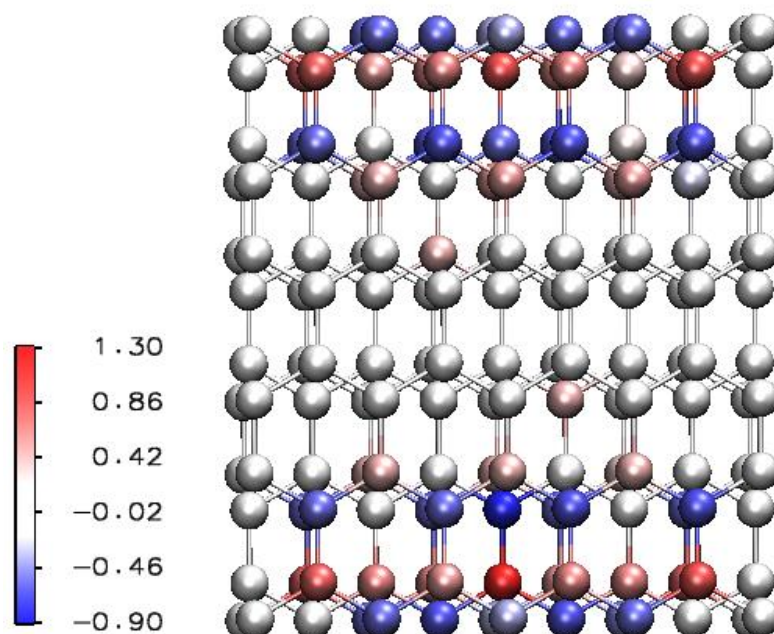
Figure 5-5: Graphical representation of charges on a small graphite structure.

Considering the size of the AC model is was impossible to use that model in the QM/QC calculations. Thus, a smaller graphite crystal was used; this crystal is shown in Figure 5-5. In that figure, the atoms are displayed as balls colored by the magnitude and polarity of the partial atomic charges assigned to them from the ESP calculation. The crystal has three sheets of graphite. In total, it has 162 atoms, and the partial atomic charges obtained from the ESP calculation ranges from +1.3e to -0.9e. These charges were analyzed with respect to the position of the atoms. It turned out that the charges aligns roughly in rows, with alternating polarity for each row. Figure 5-5 clearly shows this alternating distribution of charges. The distribution of the magnitudes of the charges within each row, is however wide.

Table 5-A: Charges from small graphite crystal.

Top Layer								
Row	1st	2nd	3rd	4th	1st	2nd	3rd	4th
Nr from left								
1	-0.273	0.989	-0.633	0.107	-0.258	0.987	-0.674	0.144
2	-0.658	0.710	-0.654	0.522	-0.653	0.748	-0.740	0.569
3	-0.412	0.632	-0.600	0.543	-0.461	0.771	-0.762	0.667
4	-0.575	0.916	-0.600	0.467	-0.675	0.998	-0.657	0.574
5	-0.246			0.119	-0.278			0.118
Avg*	-0.549	0.812	-0.622	0.510	-0.597	0.876	-0.708	0.603

Bottom Layer								
Row	1st	2nd	3rd	4th	1st	2nd	3rd	4th
Nr from left								
1	-0.273	0.991	-0.659	0.135	-0.258	0.943	-0.631	0.133
2	-0.665	0.768	-0.756	0.578	-0.615	0.697	-0.651	0.508
3	-0.447	0.750	-0.765	0.666	-0.458	0.764	-0.727	0.596
4	-0.668	1.003	-0.705	0.609	-0.690	1.009	-0.706	0.590
5	-0.265			0.151	-0.260			0.159
Avg*	-0.593	0.878	-0.721	0.618	-0.588	0.853	-0.678	0.565

Middle Layer								
Row	1st	2nd	3rd	4th	1st	2nd	3rd	4th
Nr from left								
1	-0.212	0.097	-0.087	-0.293	-0.139	0.050	-0.103	-0.008
2	-0.753	0.749	0.151	0.005	-0.699	0.674	-0.141	-0.154
3	-0.799	1.334	-0.953	-0.293	-0.661	1.071	-0.629	0.177
4	-0.165	0.706	0.188	-0.059	-0.138	0.488	0.358	-0.371
5		0.041	-0.129			0.067	-0.050	
Avg*	-0.482	0.930	-0.204	-0.160	-0.409	0.744	-0.137	-0.089

Table 5-A shows the charges from the four top and bottom rows, for each of the layers. The rows are given ordered numbers from first to fourth, while the atoms in each row are numbered from the left from one to five. The * notation on the average value means that for rows with 5 atoms, the average only runs over the three central atoms. This was done to avoid the large deviation that these edge charges had, when compared to the more central atoms. For the rows that only had four atoms the average goes over all four values. Next, the averages were taken over the rows, such that from each layer there was one average charge per row.

Table 5-B: Charges as row averages, and final charges.

Layer \ Row	1st	2nd	3rd	4th
Top	-0.573	0.844	-0.665	0.557
Middle	-0.446	0.837	-0.171	-0.124
Bottom	-0.590	0.866	-0.700	0.591
Average	-0.536	0.849	-0.512	0.341

In the end, the charges corresponding to each row was found by taking the average over the layers. These charges was then used in the utility to assign charges to the AC model.

5.1.4 Generating the structure and assigning the charges

To generate the atomistic model of AC it quickly became clear that this could not be done by hand, in any reasonable time or fashion. Due to that it was decided that the structure were to measure roughly 80Å by 80Å by 60Å, in x, y and z dimensions respectively. When dividing the length of the model in x dimension by an average distance between atoms in that dimension 80Å/2.125Å you get ~38 atoms. In the y dimension, the average distance between the rows of atoms is 1.226Å. Thus, 80Å/1.226Å gives ~65 rows of atoms. From this rudimentary estimation, the number of carbon atoms in each whole sheet of graphite was calculated to be 2470. With ca. 3.35Å between the sheets of graphite 60Å/3.35Å, give ~18 sheets. To fill up this volume by carbon atoms with graphite configuration a total of 44.460 atoms would be needed. The real number would be somewhat smaller due to the slit pores in the model, however this simple calculation justified spending some time on making a software tool to generate the structure. Figure 5-6 shows the final AC model, as it was generated by the software.

To make this software, knowledge of the molecular structure of graphite was a key element. The inter-atom spacing in x, y and z dimension was needed in order to correctly position the atoms. First the maximum number of elements, i.e. atoms, rows or graphite sheets, were calculated in each direction. For the x dimension, this was done by looping over the positions for the atoms. The first position was at the origin, the next position was at origin + 2.83558Å, the third at 2.83558Å + 1.41779Å, and the fourth at 4.25337Å + 2.83558Å, and so forth counting the number of positions reached before the total distance had passed 80Å. Y and z dimensions were much easier, considering that the spacing in those dimensions were constant. Thus, the relevant lengths only had to be divided by the corresponding spacing to get the

number of positions in those dimensions. A constant factor of one was added to the y and z dimension to get the dimension as close to the goal as possible.
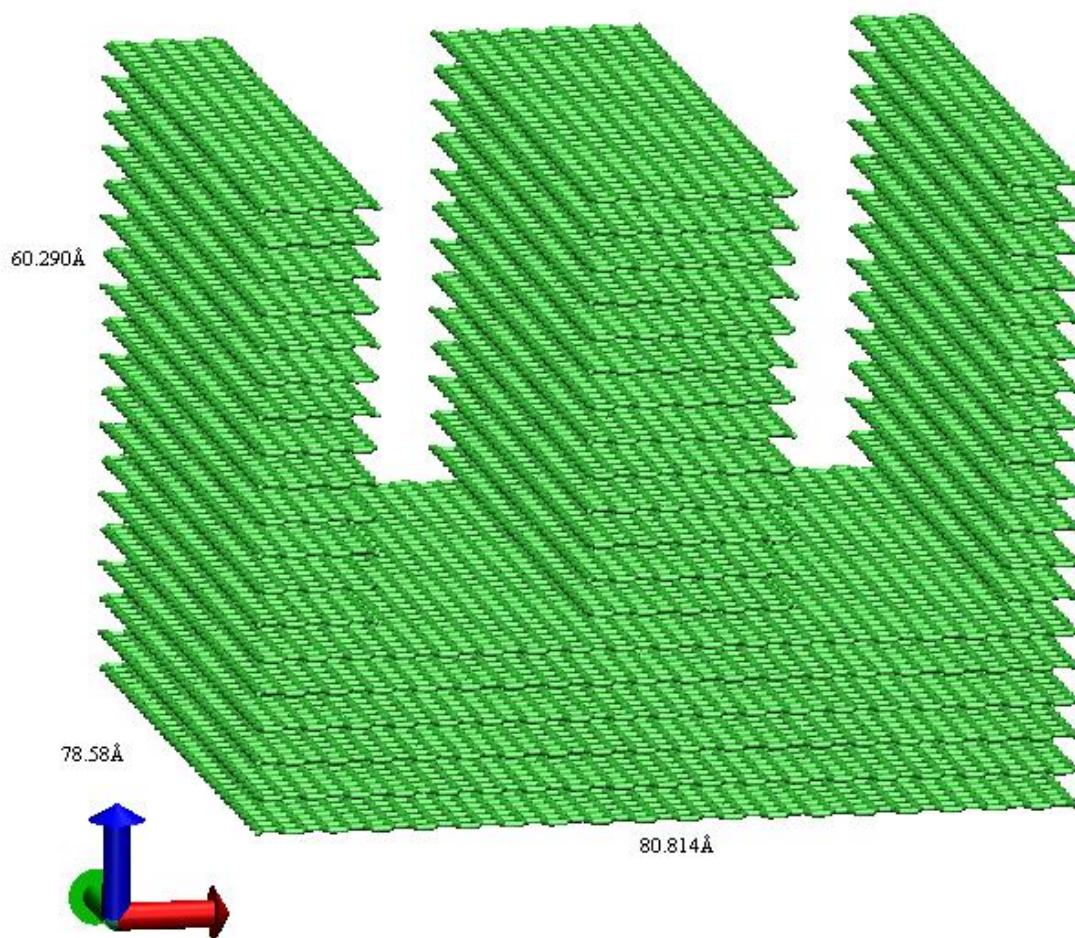
Figure 5-6: Final AC model. X, Y and Z-axis are red, green and blue respectively.

The number of elements in each dimension was then used to make a 3D array that could hold the required number of carbon atoms. Now that the initial setup was complete, the actual positioning of the carbon atom objects could start. Several checks had to be made each time a carbon object was to be positioned. First, it had to be established if the graphite sheet (z dimension) that the object belonged to was an even numbered sheet or an odd numbered sheet. This is due to the staggered configuration of the graphite sheet. If such a test were not to be carried out, all sheets would have been positioned with the atoms directly above each other, and would not fulfill the specifications for a graphite crystal. Secondly, it had to be determined if the z position was high enough that it would correspond to being in an area

where there might be a pore. If this were the case, then a special positioning rule would have to be executed. Else, normal positioning would carry on. Third, a check to find out if the y dimension row was odd or even had to be done as well. This had to be done to offset the positions of the carbon atom objects correctly. When all checks had been carried out a function initialized the carbon atom object and assigned the correct coordinates to it.

When reaching the height where there might be pores, as defined by parameters set in the program, the special positioning rule was carried out. This simply checked to see if the position where the carbon atom object was to be positioned would correspond to a position in a pore. If this was the case the object would still be initialized in order to prevent null references, but the position would be set to a specified coordinate, outside of the actual structure. Upon writing the output files, these atoms would be neglected and not written. In the final model in Figure 5-6, the widths of the pores are 21.41Å and 21.27Å for the left and right pore, respectively. Both pores had a depth of ~43.5Å.

Before writing the output files, the program searches through the structure and finds all the bonds by performing a distance search. After which the corresponding bond distances are stored in an array along with the corresponding atoms belonging to this bond. Similar operations are not carried out for angles and torsions, as this was to be a fixed molecule. The actual bonds are only there in the first place so that MD51 will count it as a molecule and not as a list of single atoms. Torsions and angles are not needed for this purpose, and as such intentionally omitted to conserve computational power.

Two output files are written from this software tool. One MD51 molecular input file, containing the relevant parameters such as element, position, mass, partial atomic charge, and Lennard-Jones parameters. It also contains the bonds section, with bond distance and a force constant. The second file is a Brookhaven Protein Databank (.pdb) file. This was used for visual representation and verification that the software building the AC model indeed did what it was supposed to do. The pdb file is also used as input to the ChargePlacer utility.

When the time came to assign the partial atomic charges on the atoms in the model, a new software utility had to be made. The task of manually identifying and assigning the corresponding charges to every atom would simply not be a viable solution. As mentioned above, the initial charges were obtained using ESP from QM calculations, but then averaged out. It is these average charges that are assigned.

This software opens up and reads in the AC model from the pdb-file generated by the first software tool. It then loops over all the atoms, comparing the coordinates to those of the first, second, third and fourth "row" of atoms neighboring to the pore. In Figure 5-6 the rows of atoms are aligned along the green axis, or y-axis to be more precise. The coordinates of these rows were manually entered into the software. Upon identifying which row the atom belonged to, if any at all, the corresponding charge was assigned. When the loop was finished and all atoms neighboring to a pore had been assigned their charges, the software then calculated the total charge of the AC model.

Since the AC model was supposed to be neutral, the bulk charges had to be such that they neutralized the (surplus) charge. However, simply assigning the same value to all bulk atoms would be considered unphysical. The solution that was implemented was to assign a random charge centered on the average charge needed to neutralize the model. This was done by the formula $Charge = \langle C \rangle + ([2 - 4 * n] * \langle C \rangle)$, where $\langle C \rangle$ is the average charge and $n \in [0.0, 1.0]$. The numbers 2 and 4 in relation with each other decides the range of variation allowed. Thus, the average charge is the starting point, and then a deviation from the average charge is assigned. When all the bulk atoms had been assigned their random charges, the total charge of the AC model was calculated again. If the total charge was deviating from zero by more than ±0.005, the random distribution was rejected, and a new distribution would be assigned. This process keeps going until an acceptable distribution has been found.
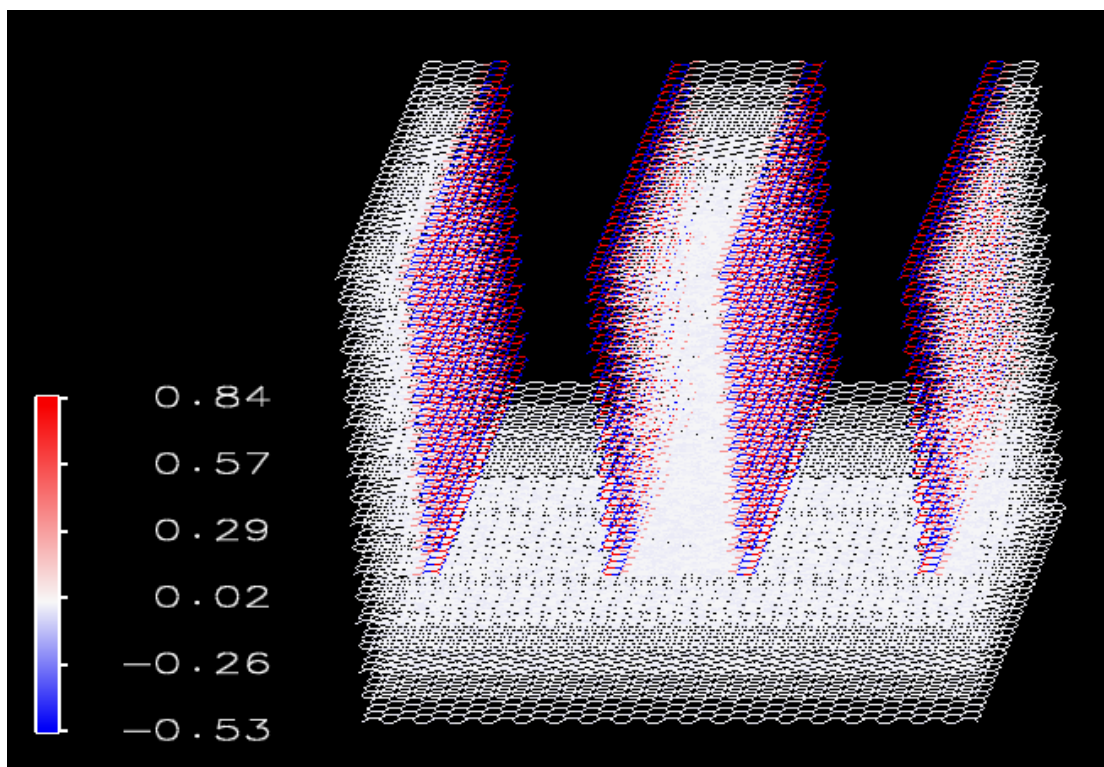
Figure 5-7: Visualization of charge distribution on AC model.

The final AC model, along with the charge distribution on the model is shown in Figure 5-2. Here the rows of charges assigned to the edges of the pores are easily visible. As the color scale bar to the left shows, the magnitude of the charges range from -0.53e to +0.84e. The bulk atoms have randomly assigned charges centered on the average value needed to make the model electrostatic neutral. Due to the large number of atoms in the bulk of this model, these charges are close to zero, and are colored white in this figure.

5.2 Fish oil constituents – Triolein

To carry out the MD simulation a representative model of fish oil had to be made. In order to make this model, the initial task was to identify the major constituents of fish oil. Four candidates were selected from the paper by (Laakso et al. 1990). In this paper, they had analyzed fish oils using High-Performance Liquid Chromatography (HPLC). Myristic acid (14:0), palmitic acid (16:0), oleic acid (18:1) and erucic acid (22:1) was the four major constituents present with the following concentrations in molepercentage, 11.1, 17.8, 11.1 and 14.5, respectively. Together these four constituents make up 54.5 molepercentage of the fish oil from Atlantic Herring used in the paper.

Ideally, these four molecules should all have been used, in appropriate relative concentrations as a basis for a model fish oil. However, it was decided to only go with one of them and then later add more if it was deemed necessary for the purpose of adsorption selectivity estimations. With PCB being lipophilic, the molecules with the longer hydrocarbon chains seemed more appropriate to use, even though the mole percentage were lower than for instance for the palmitic acid. Based upon this, and a discussion with Oterhals, the choice fell on using the oleic acid (18:1) in the form of triacylglycerol, named triolein.

5.2.1 Molecular structure of triolein

The triolein molecule has 167 atoms. It has three hydrocarbon chains with 18 carbons in each chain. Between the ninth and tenth carbon there is a double bond, on all three chains. The long hydrocarbon chains are hydrophobic. Each chain connects to an acyl group, thus giving the name triacyl. The connection, or backbone, between the three chains is a glycerol group, except for that the oxygen atoms are not connected to a hydrogen atom, but instead to a carbon atom. Triolein is an amphiphilic molecule, where the head is hydrophilic, but the hydrocarbon chains as mentioned are hydrophobic. The chemical formula for triolein is $C_{57}H_{104}O_6$. Figure 5-8 shows a graphical representation of the triolein model. This figure was made using VMD, unfortunately VMD does not show the double bonds that exists at certain places in the molecule.
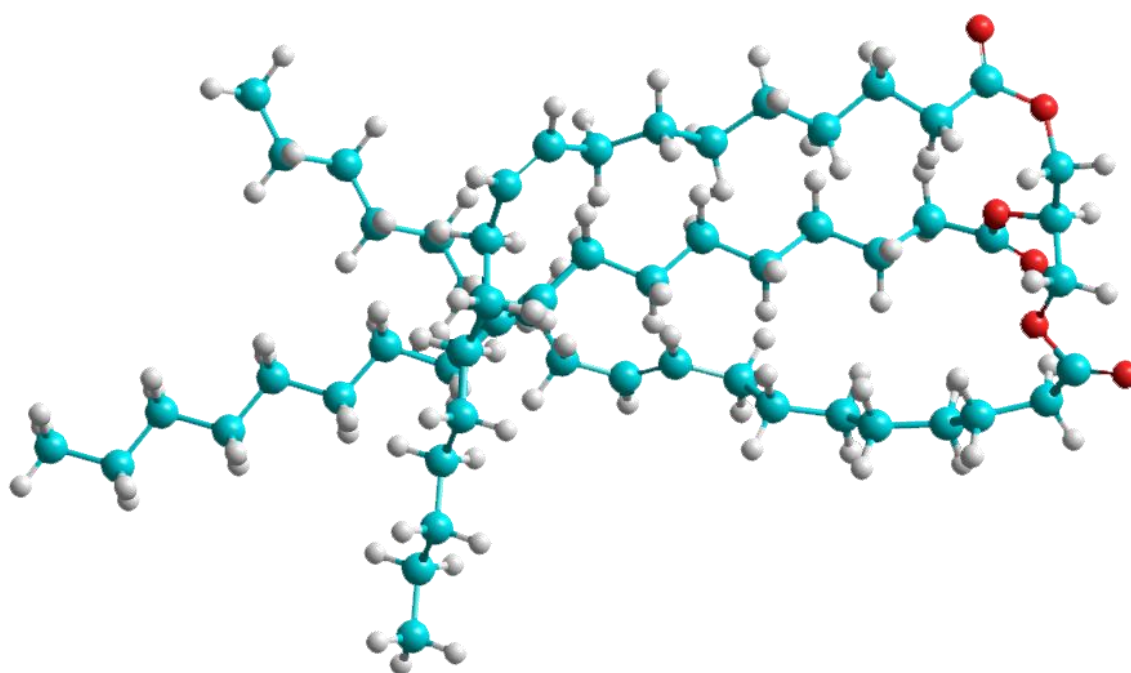
Figure 5-8: Graphical representation of the triolein model.

5.2.2 Characterization of triolein

Characterization of the triolein molecule was done over several steps. Initially the molecule had to be drawn up using Maestro. The drawing in Maestro was then used to make an input file for Jaguar. A Jaguar script was written to carry out the geometrical optimization with successively increasing basis sets. The basis sets started out with STO-3G, followed by STO-3-21G, followed by STO-6-31G** and finally STO-6-31+G**. After the final geometrical optimization, a PDB file was generated. This pdb file was in turn used as a basis for the MD51 molecular input file.

To acquire the partial atomic charges a Single Point Energy (SPE) calculation was done in Jaguar, using the STO-6-31+G** basis set and the SM6 in order to have a water solvated system. By doing this the charges for both a vacuum and water solvated system were found from CM4. In this work the partial atomic charges from the water solvated system was utilized. These charges replaced the charges that are included in the OPLS force field.

To model the stretching of bonds, bending of angles and torsions the OPLS force field parameters were used. As such, the atoms in the triolein molecule had to be classified with the correct atom types in order to find the corresponding parameters. Initially this may seem like an insurmountable task given that there are 167 atoms in this molecule. However, quite a lot of them are of the same atom type. Thus, only a few atom types are needed to describe the different atoms. The atom types used to describe triolein and the corresponding values for the Lennard-Jones parameters sigma (σ) and epsilon (ε) is given in Table 5-C.

Table 5-C: OPLS force field parameters for triolein

Atom Type	Description	Sigma [Å]	Epsilon [kJ/mole]
HC[1]	H in alkane	2.500	0.126
HC[2]	H in alkene	2.420	0.126
CT	C in alkane	3.500	0.276
CM	C in alkene	3.550	0.318
C	C= in ester	3.750	0.439
O	O= in ester	2.960	0.879
OS	-OR in ester	3.000	0.711

The superscripts in Table 5-C on the HC entries are put there to distinguish between hydrogen atoms in an Alkane[1], and hydrogen atoms in an Alkene[2]. These are two different atom types, as the force field parameters indicate, but have the same descriptive characters.

Stretching of covalent bonds also needs to be described. Two parameters are needed for each type of bond, specifically the equilibrium length of the bond and the force constant. The two atom types that make up the bond define the type of bond. Each covalent bond has its own characteristics.

Table 5-D: OPLS parameters for covalent bonds in triolein.

Bond	Bond Length [Å]	Force Constant [kJ/mole/Å^2]
CT-CT	1.529	1121.312
CT-HC[1]	1.090	1422.560
CT-CM	1.510	1326.328
CM-CM	1.340	2297.016
CM-HC[2]	1.080	1422.560
CT-C	1.522	1326.328
C-O	1.229	2384.880
C-OS	1.327	895.376
CT-OS	1.410	1338.880

Table 5-D lists the force field parameters for the bond stretching of the different types of bonds in triolein. The nine bond types listed make up all 166 covalent bonds in triolein.

Table 5-E: OPLS parameters for covalent angles in triolein.

Angle	Equilibrium Angle [deg]	Force Constant [kJ/mole/rad^2]	Angle	Equilibrium Angle [deg]	Force Constant [kJ/mole/rad^2]
CT-CT-CT	112.7	244.136	OS-C-CT	111.4	338.904
HC[1]-CT-HC[1]	107.8	138.072	OS-CT-HC[1] (*)	N/A	N/A
CT-CT-HC[1]	110.7	156.900	C-CT-CT	111.1	263.592
CT-CM-HC[2]	117.0	146.440	CT-CT-CM	111.1	263.592
CM-CM-CT	124.0	292.880	CM-CT-HC[1]	109.5	146.440
CM-CM-HC[2]	120.0	146.440	CT-OS-C	116.9	347.272
C-CT-HC[1]	109.5	146.440	O-C-OS	123.4	347.272
O-C-CT	120.4	334.720	CT-CT-OS	109.5	209.200

Table 5-E shows the covalent angles in triolein, along with their equilibrium angles and force constants. In triolein, there are 318 covalent angles. For five of the covalent angles, no force field parameters were available. These five were of the type OS-CT-HC and an asterisk marks this entry in Table 5-E. Considering that the OS-CT-CT angle is defined, as well as CT-CT-HC, all atoms concerned by the missing angle type still have some parameters. The five angles in question are because of this, not included in the MD51 molecular input file. This means that the angles are still there, but they are practically unconstrained, as they have no force constant assigned. The angles do of course exist and MD51 will sample them, but there will be no contribution to the energy from them. Thus, the file only contains 313 angles. All the angles are described using only the force field parameters given in Table 5-E. When comparing the force constants for bonds and angles it is evident that the force constants for bonds are on average more powerful. The force constants for the bonds also have a slightly wider distribution.

The final parameters needed to describe the triolein molecule is the parameters for the torsions. As mentioned in section 4.2.5.3, these are in OPLS based upon a Fourier series. Thus, there are three parameters pr torsion. There are no equilibrium torsion angles given, as it is for the normal angles. The equilibrium angle is given by the lowest energy for the Fourier series.

Table 5-F: OPLS parameters for torsions in triolein.

Torsion	V_1 [kJ/mole]	V_2 [kJ/mole]	V_3 [kJ/mole]	Torsion	V_1 [kJ/mole]	V_2 [kJ/mole]	V_3 [kJ/mole]
C-CT-CT-CT	-7.100	-1.908	2.448	CT-CT-CT-CT	5.439	-0.209	0.837
C-CT-CT-HC[1]	0.000	0.000	-0.318	CT-CT-CT-HC[1]	0.000	0.000	1.255
CM-CM-CT-CT	1.448	1.695	-3.782	CT-CT-CT-OS	N/A	N/A	N/A
CM-CM-CT-HC[1]	0.000	0.000	-1.556	CT-OS-C-O	0.000	21.439	0.000
CM-CT-CT-CT	N/A	N/A	N/A	CT-C-OS-CT	19.535	21.439	0.000
CM-CT-CT-HC[1]	0.000	0.000	1.531	CT-CT-C-OS	0.000	0.000	-2.314
HC[2]-CM-CT-HC[1]	0.000	0.000	1.331	C-OS-CT-CT	-5.104	-0.527	1.766
C-OS-CT-HC[1]	0.000	0.000	0.828	HC[1]-CT-C-O	0.000	0.000	0.000
CT-CM-CM-CT	0.000	58.576	0.000	HC[1]-CT-C-OS	0.000	0.000	0.552
HC[2]-CM-CM-CT	0.000	58.576	0.000	CT-CT-C-O	-1.159	5.138	-2.904
CT-CT-CM-HC[2]	N/A	N/A	N/A	HC[1]-CT-CT-OS	0.000	0.000	1.958
HC[1]-CT-CT-HC[1]	0.000	0.000	1.255	OS-CT-CT-OS	-2.301	0.000	0.000
HC[2]-CM-CM-HC[2]	0.000	58.576	0.000				

In triolein there is 450 torsions, all described by the 25 unique torsion types listed in Table 5-F. For three of these unique torsions no parameters were available in OPLS. These torsions are CM-CT-CT-CT, CT-CT-CM-HC[2] and CT-CT-CT-OS. Of them, there are 14 entries in total. All atoms in these three torsions are part of other torsions, and as such not without parameters. The impact, however, on the accuracy of the simulation due to this lack of parameters for these torsions are unknown. Without these 14 entries, there are only 436 entries in the MD51 molecular input file.

As it had been decided to obtain the partial atomic charges from QM calculations, by using the SM6 to make the system water solvated, as well as from ESP, a comparison of the charges and a subsequent decision on which set to choose had to be done. A selection of the charges obtained is displayed in Table 5-G. Considering that there are 167 atoms in the triolein molecule the list of charges are too extensive to show in its complete form here. Thus, only the charges on atoms of particular interest is shown in Table 5-G. A complete table of the charges obtained for the triolein molecule from all the methods is included in Appendix A.

Table 5-G: Comparison of charges from QM calculation for triolein.

Atom	Vacuum				Water solvated		
	LPA	RLPA	CM4	ESP	LPA	RLPA	CM4
O(=C)	-0.31270	-0.25419	-0.35098	-0.53058	-0.38029	-0.31856	-0.40404
O	-0.34491	-0.21574	-0.25516	-0.46201	-0.35771	-0.22968	-0.27191
C(=O)	0.22770	0.22708	0.34472	0.71992	0.23351	0.22937	0.33675
C(=C)	-0.23533	-0.10819	-0.06760	-0.35219	-0.23921	-0.11205	-0.07158
C(-O)	-0.09310	-0.01878	0.08041	0.15002	-0.08008	-0.00553	0.08954

The atoms which charges are displayed in the table above are from one chain in the triolein molecule. Within the molecule the charges for these atoms for all three chains were reasonable consistent. Some differences between LPA, RLPA and the CM4 and ESP are evident for the backbone carbon atom that is bonded to an oxygen atom with a single bond. The two former methods give negative charges, while ESP and CM4 show positive charges for the same atom. For the other atoms, it is clear that ESP consistently give charges of a greater magnitude than the other three methods. Considering that the complete simulation was not supposed to be a vacuum simulation, charges from a water solvated system seemed more appropriate. Due to the inconsistencies and known deficiencies of LPA, this charge distribution could readily be dismissed. That left the choice between charges from RLPA and

CM4. RLPA might be better than LPA. It does however give much the same charges as LPA, but of smaller magnitudes. Particularly interesting is it when comparing the charges on the carbon atom, which is connected by a double bond to an oxygen atom. Here LPA, CM4 and ESP all give charges of a greater magnitude. The final decision was to use the CM4 charge distribution. Charges from ESP calculation on a water solvated system would have been interesting for comparison reasons; unfortunately, Jaguar does not support this combination.

A sample molecular input file for MD51 for the PCB congener 77 molecule, with all the partial atomic charges, LJ parameters as well as force constants is included in appendix B. The molecular input file for the triolein molecule is included on the CD accompanying this thesis. Due to the large size of the file it was not included as an appendix here.

5.3 Polychlorinated biphenyls

Polychlorinated biphenyls (PCB) are classified as a persistent organic pollutant, as mentioned earlier. PCB consists of two phenyl ring structures, which are connected. The ring structures are made up of carbon atoms. Connected to each carbon atom, except the two atoms that are used to join the ring structures together, is either a hydrogen atom or a chlorine atom. This means that there are 12 carbon atoms, and 10 positions for hydrogen or chlorine atoms. The general chemical formula for PCB is $C_{12}H_{10-x}Cl_x$, where X is between 1 and 10. In total, there are 209 different congeners of PCB.

PCB are lipophilic molecules, and as such they biomagnifies in the food chain. To avoid further biomagnifications when producing fish oil and fishmeal from fish byproducts, herring, sprat, sand eel, capelin or blue whiting it is desirable to remove the PCB present. A previous study carried out by Otherhals showed that adsorption of POP onto activated carbon can be a viable method. However, there were some concerns voiced in that study. Oterhals' experimental work showed that the mono-ortho substituted DL-PCB included in his study showed a limited tendency to adsorb onto AC. While the non-ortho substituted DL-PCB adsorbed to a much higher degree.

It was this observation that made Oterhals to form and propose his hypothesis that efficient adsorption of PCB onto AC were dependent on a planar molecular conformation. If this is the case then the versatility of the adsorption process in order to reduce the levels of POP in fish oil will be highly limited.

5.3.1 Molecular structure and toxicity

In this thesis the numbering of the congeners are the BZ numbering where each congener has a unique number from 1 to 209 depending on the number of chlorine substitutions and positions (Ballschmiter et al. 1992). Out of the 209 different PCB congeners, 12 are classified as dioxin-like and given the name DL-PCB. These 12 DL-PCB congeners have four common traits, that all need to be present for it to be dioxin-like. The traits are as follows (EPA 1980):

- Co-planar congener with non-ortho substitution or mono-ortho substitution at either the 2, 2', 6 or 6' position (see Figure 5-9).
- Must have a total of four or more chlorine substitutions.
- Both para positions (4, 4') must be chlorinated.
- Two or more of the meta positions (3, 3', 5 and 5') must be chlorinated.
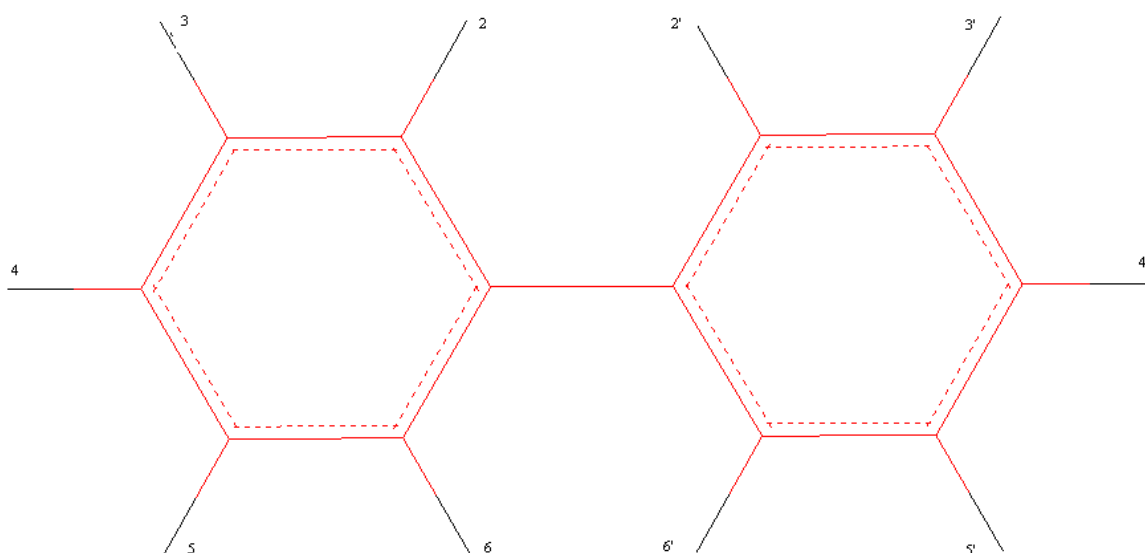
Figure 5-9: Structure of PCB and numbered positions.

The DL-PCB has a higher toxic equivalent (TEQ) rating than the normal PCB congeners. The revised TEQ values for PCB from 2005 by World Health Organization (WHO) are 0.0001 and 0.00003 for PCB congener 77 and 118, respectively (Van den Berg et al. 2006). The European Union's European Commission has set limits to the TEQ levels that legally may be present in foodstuff (Wiborg et al. 2008). For marine oils, i.e. fish body oil, fish liver oil and oils from marine organisms, the limit is 10pg/g fat for foodstuff intended for human consumption. Both the levels accepted in foodstuff by the EU, and the respective TEQ levels of the different PCBs are revised from time to time.

5.3.2 Conformers and adsorption

Based upon Oterhals' experimental study two DL-PCB congeners were selected. Both of these congeners exhibits dioxin-like traits. The congeners were selected due to the large difference in the amount adsorbed during the experiments. PCB congener 77, a co-planar non-ortho substituted DL-PCB had been adsorbed onto the AC with an average level of 74.7%, relative to the alkali-refined and bleached oil (Oterhals et al. 2007). PCB congener 118, a co-planar mono-ortho substituted DL-PCB had been adsorbed onto the AC to a much lower degree, only 10.1% relative to the alkali-refined and bleached oil. This large difference makes these two congeners ideal to compare.

What is clear from the experiment by Oterhals is that the molecular structure of the PCB congeners to large degree decides the level adsorbed. To determine the trapping mechanisms of PCB onto AC, and how the molecular structure affects the adsorption process is the main motivation for this study.

5.3.3 Characterization of PCB congeners 77 and 118

The chemical formulas for PCB congener 77 and 118 were taken from the table over different congeners (EPA 1980). Based upon these formulas the molecules were drawn up in Maestro, see Figure 5-10. Once this was done a QM simulation were set up to carry out the geometric optimization of the molecules. In this simulation DFT/B3LYP was used with the 6-31+G** basis set. A single point energy (SPE) calculation was also carried out, both in vacuum and in water by utilizing SM6. From the SPE calculation the partial atomic charges were obtained from CM4, for both vacuum and water solvated system.
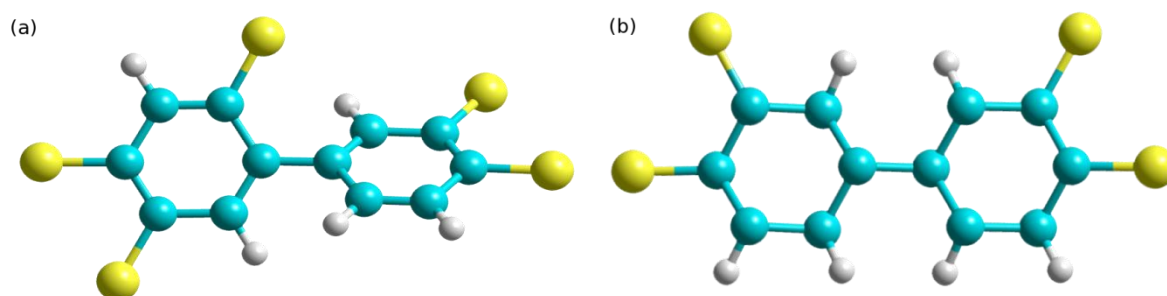
Figure 5-10: (a) PCB congener 118. (b) PCB congener 77.

In Figure 5-10 above, note that the phenyl ring of PCB congener 118 is rotated at an angle, while the phenyl rings of PCB congener 77 are aligned in the same plane. Also, note that the

phenyl rings are aromatic, but VMD does not show double bonds to indicate the aromatic nature of the rings.

The OPLS force field was used to characterize the bond stretching, angle bending, torsions and the Lennard-Jones parameters. However, the partial atomic charges from OPLS were replaced with the charges obtained from the QM calculations. In order to use the OPLS force field parameters the atoms in the models had to be assigned an atom type. In the case of PCB, there are four different atom types applicable. These are listed in Table 5-H. A fifth atom type could have been added for the carbon atoms bonded with the chlorine atoms.

Table 5-H: OPLS atom type and LJ parameters for PCB.

Atom Type	Description	Sigma [Å]	Epsilon [kJ/mole]
CA	C in benzene	3.550	0.293
HA	H in benzene	2.420	0.126
Cl	Cl in chlorobenzene	3.400	1.255
C!	C1 in biphenyl	3.550	0.293

However, the only difference between the CA atom type and the carbon atom type connected to a chlorine atom was the charges. The other parameters were the same. Considering that the charges from OPLS are not in use in this work, the inclusion of a fifth atom type was not necessary. When comparing the parameters in Table 5-H, note the large value for the well depth for the chlorine atom.

To describe the covalent bond interactions, force field parameters are also needed for the interactions between atoms of same or different atom types. For PCB there are five different types of bonds, specifically CA-HA, CA-CA, CA-Cl, CA-C! and C!-C!. The OPLS force field parameters for the bonds in PCB are given in Table 5-I.

Table 5-I: OPLS parameters for covalent bonds in PCB.

Bond	Bond Length [Å]	Force Constant [kJ/mole/Å^2]
C!-C!	1.460	1610.840
CA-C!	1.400	1962.296
CA-CA	1.400	1962.296
CA-Cl	1.725	1255.200
CA-HA	1.080	1535.528

For covalent angles, three atoms are involved. This leads quickly to many different possible combinations. The eight different angles used are listed in Table 5-J. Note that the entry marked by (*) is not necessarily correct. There was no information available in the OPLS force field for that combination of atom types. However, when comparing C!-CA-HA with CA-CA-HA the parameters are identical. Based upon this the assumption was made that it would be valid to use the parameters for the CA-CA-Cl for the C!-CA-Cl angle. All the equilibrium angles in Table 5-J are equal; it is only the force constants that separate the different angles.

Table 5-J: OPLS parameters for covalent angles in PCB.

Angle	Equilibrium Angle [deg]	Force Constant [kJ/mol/rad^2]
C!-CA-Cl (*)	120.0	313.800
C!-CA-HA	120.0	146.440
CA-C!-C!	120.0	263.592
CA-C!-CA	120.0	263.592
CA-CA-C!	120.0	263.592
CA-CA-CA	120.0	263.592
CA-CA-Cl	120.0	313.800
CA-CA-HA	120.0	146.440

The final part of the force field is the torsions. As mentioned in section 4.2.5.3, the torsions describe the angle between two planes made up from four atoms, i.e. the planes made from A-B-C and D-C-B. The parameters used for describing the torsions in the PCB are given in Table 5-K.

Table 5-K: OPLS parameters for torsions in PCB.

Torsion	V1 [kJ/mole]	V2 [kJ/mole]	V3 [kJ/mole]	Torsion	V1 [kJ/mole]	V2 [kJ/mole]	V3 [kJ/mole]
C!-CA-CA-HA	0.0	30.334	0.0	Cl-CA-CA-Cl	0.0	30.334	0.0
CA-C!-C!-CA	0.0	9.079	0.0	Cl-CA-CA-HA	0.0	30.334	0.0
CA-CA-C!-C!	0.0	30.334	0.0	HA-CA-C!-C!	0.0	30.334	0.0
CA-CA-CA-CA	0.0	30.334	0.0	HA-CA-C!-CA	0.0	30.334	0.0
CA-CA-CA-Cl	0.0	30.334	0.0	HA-CA-CA-HA	0.0	30.334	0.0
CA-CA-CA-HA	0.0	30.334	0.0				

All of the torsions appear to be equal in Table 5-K, with one exception which is the parameter deciding how the two phenyl rings are rotated in relation to each other. Note that this section has only listed the different possible bonds, angles and torsions in use. In a molecule, there can be several bonds, angles or torsions of the same type. The same goes for the molecular input file for MD51. The input file does indeed contain more bonds, angles and torsions than the tables in this section may indicate. However, many of the entries are identical but with regards to different atoms.

The partial atomic charges for both PCB congeners were obtained from QM calculations. Two different calculations were carried out. One used ESP to find the partial atomic charges in a vacuum; the other calculation used the SM6 model. From the SM6 model calculation six different set of charges were obtained. These were charges from Löwdin population analysis (LPA), Redistributed Löwdin population analysis (RLPA) and charges from Charge Model 4 (CM4). SM6 calculates charges from both vacuum and a water solvated system.

Table 5-L: Selected charges from QM calculations for PCB congener 77.

Atom	Vacuum				Water solvated		
	LPA	RLPA	CM4	ESP	LPA	RLPA	CM4
C3	-0.17780	-0.11763	-0.00205	0.01083	-0.18469	-0.12482	-0.01118
H8	0.21471	0.12813	0.08750	0.09150	0.21856	0.13220	0.09181
Cl8	0.14058	0.09719	-0.02367	-0.06629	0.13005	0.08679	-0.03128

In Table 5-A, C3 is the carbon atom connected to Cl8. The hydrogen atom was randomly selected. When comparing the charges, some questions arise. Chlorine is normally considered very electronegative. However, both LPA and RLPA give positive charges for the chlorine atom. This is also seen for the other chlorine atoms in both PCB congener 77 and 118. This strongly supports that LPA is not a good way to obtain partial atomic charges when using diffuse basis sets. RLPA, which is supposed to alleviate some of the problems with LPA, returns less positive charges. Nevertheless, they are still positive, which is not at all the expected polarity. LPA and RLPA are consistently positive in both vacuum and in a water solvated system. CM4 and ESP on the other hand give negative charges. Granted, they are not of any considerable magnitude when comparing to the OPLS parameter of -0.180e in chlorobenzene. Another interesting difference is that ESP gives a positive charge on C3 and

this is consistent with the charges from OPLS, however the magnitude is very different. All the other methods give negative charges for the same atom. Intuitively, it would not be unlikely for C3 to have a positive partial atomic charge, when bonded to a very electronegative atom such as chlorine. However, considering that ESP charges are not available for a water solvated system and influenced by the fact that CM4 charges from a water solvated system were used for triolein; the charges from CM4 with a water solvated system were used for the PCB models as well. A table of all charges for both PCB congeners can be found in appendix A.

The molecular input file in MD51 for PCB congener 77 is included in Appendix B. This file contains all the force field parameters needed for this model, as well as partial atomic charges and atomic weight. The molecular input file for PCB congener 118 is included on the CD accompanying this thesis.

6 Setting up the initial simulations

The task of setting up the simulation system was split into three parts. First, a system containing a mixture of PCBs and triolein were set up. Second, the model for the activated carbon was introduced. Third, the external force discussed in section 4.2.7 was applied, downscaled over time and eventually turned off. The details of these tasks will be further discussed in the subsections below.

6.1 Liquid phase - PCBs and triolein

To set up the liquid phase, two main tasks were carried out. These tasks are detailed in their respective subsections below.

6.1.1 Estimating the number of molecules – Density calculations

The size of the system was not arbitrary chosen. From density calculations where the molecular weight of triolein was known and the volume of the pores, along with what was judged an absolute minimum of liquid phase above the AC, the absolute minimum number of triolein molecules was estimated. The density of triolein is well known from literature, in this case it was taken from (Fong et al. 1988) to be 0.915 g/ml. Thus, the objective was to find the number of triolein molecules needed in order to make the simulation system have approximately the same density, with the given volume.

The volume of each of the two pores was calculated to be $20\text{Å} * 80\text{Å} * 60\text{Å} = 9.6 * 10^{-26}$ m^3. In discussion with Prof. Kuznetsova an it was decided that an absolute minimum height of liquid on top of the graphite was 20Å. This was to allow the triolein molecules some leeway as to how they want to move. A too thin layer would lead to unrealistic motions, and too thick layer would consist of too many molecules and thus demand too much computational time. An additional reason as to why a minimum thickness of the liquid layer on top was important was to shield some of the long-range electrostatic interactions from the AC model with the pores of the replica above it (due to PBC). Thus, the volume on top of the AC model was calculated to be $80\text{Å} * 80\text{Å} * 20\text{Å} = 1.28 * 10^{-25}$ m^3. Summing together the contributions from the pores and the top of the AC model the minimum volume needed to be filled was calculated to be $2 * 9.6 * 10^{-26}m^3 + 1.28 * 10^{-25}m^3 = 3.23 * 10^{-25}$ m^3.

From general physics we know that $\rho = m/V$, where ρ is density, m is mass, and V is volume. The relationship between mass and number of molecules is given by $m = N * M_w$, where N is the number of molecules in moles and M_w is the molecular weight. By combining these two expressions, the number of molecules in moles was found. Multiplying by Avogadro's number, N_A, converts the result from number of moles to number of molecules. The complete formula for the minimum number of triolein molecules needed is given by equation (6.1).

$$N = \frac{\rho * V}{M_w} * N_A$$

$$= \frac{915 \frac{kg}{m^3} * 25.6 * 10^{-26} \, m^3}{0.885 \frac{kg}{mole}} * 6.022 * 10^{23} \frac{molecules}{mole} \qquad \textbf{(6.1)}$$

$$\approx 200 \, molcules$$

Considering that this was set to be the absolute minimum of triolein molecules needed, some additional molecules were added. The final number for triolein ended up being 252 molecules.

An important aspect of this thesis was to study the trapping mechanisms for PCBs. In view of this, and also to gain some statistical significance of the numbers, a high concentration of PCBs were deemed necessary. The number of molecules for the two different species of PCB would also have to be equal for a better comparison. For the complete system in this thesis, there were 51 molecules of each of the PCB species used, PCB congener 77 and congener 118. Thus, in total 102 PCB molecules.

The final liquid part of the system contains 252 triolein molecules, and 51 molecules of each PCB congener, in total 354 molecules. This gives the following mole fractions $X_{triolein} = 0.712$, $X_{PCB} = 0.288$, where $X_{PCB} = X_{PCB77} + X_{PCB118}$. For each of the individual species of PCB the mole fraction is thus $X_{PCB77} = X_{PCB118} = 0.144$.

6.1.2 Compression of the system

When starting a molecular dynamics simulation the molecules have to be positioned into the simulation box. This can be done manually by assigning the centers of mass for each molecule a coordinate inside the simulation box. It is however difficult to do so. In MD51, there are several methods for positioning the atoms in the simulation box. The different

options for random positioning are FCC lattice, cubic lattice, cylindrical and spherical holes. The FCC lattice option was chosen in this instance. Random rotation of the molecules was left off.

To avoid overlapping molecules, or atoms packed too close to each other, which would lead to tremendous repulsive forces, a large box was used as the starting point. This box measured 80.814Å x 78.58Å x 1100.0Å. The x and y dimensions was set equal to those needed for PBC to correctly replicate the AC model in those dimensions. The height of the system was set sufficiently high as to avoid the overlapping problem.

As a side effect from the need to start in a large box, the density of the system was very low, and not at all close to the real density of triolein of 915 kg/m³. Even though the addition of PCBs will have some effect on the density of the liquid phase, this was not taken into account, as no experimental data of densities for a system of these concentrations were available. Thus, the density of triolein was the target density for the system. In order to achieve this density the only variable volume parameter, the height, was adjusted at intervals, such that the volume would eventually be sufficiently small as to correspond to the target density. To allow the molecules to move around and reorganize rather easily a high simulation temperature was used. Considering the size of the system in number of atoms, bonds, angles and torsions it was decided to reduce the number of molecules during this compression stage. By reducing the number of molecules, the computational time could be significantly reduced.

Figure 6-1: Initial system for liquid phase.

1100.0Å

78.58Å

80.814Å

The general idea behind this was that since PBC were already utilized, then when the target density was reached the system could "physically" be replicated in z dimension. This would not lead to overlapping atoms or molecules, as PBC ensures that there is a void, or space, on the opposite face for the molecules stretching over the boundaries of one face. In view of this,

the rather strange numbers of molecules becomes clear. The number of molecules for each species can be divided by 3 without any remainder. Initially, the number of molecules was set to 250 and 50, respectively for the triolein and each of the PCBs.

The initial simulation box in Figure 6-1 had, because of the decision above, 84 triolein molecules and 17 molecules of each of the species of PCB. In total, there were 118 molecules in the simulation box. System temperature was set high, to a temperature of 2000K. This would allow the molecules to have a considerable kinetic energy, and as such be able to move out of unfavorable positions that might occur each time the height of the system was reduced. Each time step in the compression phase of the simulation was set to be 0.1ps, and simple velocity scaling was utilized. The use of simple velocity scaling was primarily to avoid having Nosé-Hoover breaking down during the run, as the routine embedded in MD51 aborts with large deviations in temperature. Such deviations can readily arise due to molecules or atoms being forced too close to each other by the ongoing compression. The time between each compression phase was set to be 20.000 time steps, or simply 2ps.

For every run, or compression phase, the height of the box was reduced by roughly 10%. Thus, in the second run the system height was 1000Å, the third was reduced by a 100 more angstroms down to 900Å. For every run, the number of atoms or molecules in close proximity was monitored, such that in case of a too rapid compression the degree or compression could be reduced. At the end of the compression task, the system had the following dimensions: (80.814 x 78.58 x 24.26)Å, with a resulting volume of $1.54 * 10^{-25} m^3$ or $154.0 * 10^3$ Å^3. The corresponding density was 0.915 g/cm^3.

MD51 dumps, if the option is turned on, a file containing the final position of each atom, in each molecule, in the xmol format. This format simply contains a header with the number of atoms, the dimensions of the system along with the age of the system in picoseconds. Then there is a list consisting of the element of the atom in question along with its x, y and z coordinate in angstroms. A simulation can also be initiated with the positions from such a file. With that in mind, a copy of this file was opened in Excel. The x and y coordinates were left as they were. In addition, the original z coordinates were kept as they were, but all atoms with coordinates were duplicated in such a way that there were now three times the atoms listed. However, this left atoms in exactly the same position. In order to position these three smaller systems on top of each other in such a manner that they would be a seamless larger system the z coordinates were manipulated. This was done by simply adding the z dimension of the box,

namely 24.26Å, to each of the z coordinates for one of the copies. For the other copy, the same value was subtracted. The result was that the original small system was now in the middle between two replicas of itself, as it appeared to be in the simulation due to PBC.

The final stage was to multiply the z dimension in the header by three in order to assign the system its new size. As this was completed the liquid system now had 252 triolein molecules, and 51 molecules of each PCB species, all in a box with the dimensions of (80.814 x 78.58 x 72.78)Å. A new simulation was then started from this configuration in order to allow the molecules to move around some more and reorder themselves, to avoid any unintended side effects. This can also be considered as an equilibration period for the liquid system. Nevertheless, the complete system would still have to go through equilibration.

6.2 Complete system

When all the individual components of the system were ready, meaning the individual molecular models and the complete liquid phase a few minor tasks remained.

6.2.1 Inserting the Activated Carbon model

As has been mentioned from time to time earlier in this thesis the precise positioning of the AC model is an important detail. The positioning was important in order to allow the periodic boundary conditions to seamlessly replicate the AC model in x and y dimensions. In order to do so, the correct system size had to be calculated from the AC model in such a way that the distance between the real model and its images in x and y dimensions would correspond to the normal graphite spacing. By doing this, the system would not have any "gaps" or cracks between itself and the replicas due to PBC.

To obtain this an additional distance of 1.42Å was added to the dimensions of the system in the x-dimension, this would correspond to ½ of the carbon-carbon bond length in graphite on each side of the model. When the PBC then replicates the real model in x-dimension the distance between the images would be the correct bond length. Similar was done in the y dimension, however the distance added there was only 1.226Å, corresponding to the spacing between the rows of atoms in that dimension. Once this was done the model replicates in x and y dimensions perfectly, and the pores appears (visually) to be longer, and there are four extra pores, two on each side of the main model.

The positioning was done manually by making an xmol formatted file. The ChargePlacer utility had already placed the center of mass at origin when assigning the charges, so most of the work was done there; some modifications to the positioning along the z-axis were done though. Finally the AC model had to be merged with the liquid phase. Again, by manually manipulating the xmol formatted file, the liquid phase was positioned on top of the AC model, though some distance apart to avoid for the long chains of the triolein molecules to overlap with the AC model.

6.2.2 Applying and downscaling of external force

The final part of setting up the simulation was to actually start the simulation, as well as in a rapid manner let the liquid phase enter the pore structure of the AC model. It was for this reason that MD51 was modified to allow for the application of an external force working in the z direction. By activating the modification through the keyword, and testing different strengths of the force, a suitable value for the force was found. Mind that this was done by always resetting the system back to scratch, such that the effect of the force could be visually compared between the runs.

The starting value for the force was set to -1.000.000 V/cm (Volt pr cm). This was converted to internal units. In the internal units this force was a dimensionless number corresponding to -3.830 * 10^{-11}. The conversion to internal units is a function of the (total) weight of the system in kg and the length of the time steps used.

Since this was an external force, that was not supposed to be there and influence the behavior of the system, but only act as an accelerant, or a time saving function for bringing the two systems into close contact, it had to be turned off once the job was done. However, as mentioned earlier MD naturally conserves energy, since it is using the classical laws of motion. From Newton's third law it is known that for every force working in one direction there is an equal force of equal magnitude working in the opposite direction. Thus, by simply turning the force off the molecules rapidly started moving away from the AC model. The obvious solution was to downscale the force systematically, over several iterations while allowing the system to settle, much like what was done for the compression.

The procedure to downscale the force and eventually turn it off was done as described here. The system was allowed to evolve for 20.000 time steps between each reduction of the force as a main rule. On some occasions twice this. The time steps was 0.1ps, so 20.000 time steps

corresponds to 2ps. The temperature of the system was set to 330K; this is in the range where normal simulations using these models will be. Initially the force was reduced by as much as 50%, since the external force was still strong enough to stop the upwards motion during the 2ps. When the force had a value of 250.000 the reduction was reduced to roughly $1/3^{rd}$ each time. The final stages was 25.000, 15.000, 10.000, 5.000 and 0. By this time, the force was so small that hardly any effect was observed between the scaling, the force being dominated by LJ and long-range electrostatics.

The physical size of the simulation during this time was (80.814 x 78.58 x 220.0)Å. The large height of the system was to keep interactions from long-range electrostatics between the liquid phase and the bottom of the replica of the AC model due to PBC to a minimum.

6.3   Computational demands of final model

The sheer size of this system in atoms, bonds, angles and torsions imposes some requirements for the computational power of the hardware in order to run efficiently. Both during the preliminary and initial setup of the simulations several obstacles unveiled themselves, which led to considerable time and effort spent to overcome the obstacles. This section deals with some of these problems, and how they were solved, or circumvented.

6.3.1   Compilation issues and the memory models

In the start of this thesis, MD43 was the software that was chosen to run the simulations in. MD43 is an older version of MDynaMix. However, Prof. Tatiana Kuznetsova has done some extensive modifications to it, by adding support for the aforementioned pdb file format. Her main contribution was the addition of support for quarternions, thus restricting the molecules to only rotational and translational motion. MD43 did unfortunately not support the number of atoms needed. The restriction was due to that the memory model for the runtime image exceeded 2 GB. This severely narrowed down the list of compilers that could support the compilation of the program. All the arrays needed to hold information about the atoms were mapped out during compile time, i.e. it was a static memory model. If the software had been written differently, as to allocate the tables at runtime, then this would not have been an issue.

Three compilers were found which could support a memory model exceeding 2 GB, one of these was GNU gfortran v4.3.4, the other was an Intel compiler and the last one was PGI's fortran compiler. Upon trying out these compilers it was discovered that gfortran v4.3.4

supported the use of a keyword "mcmodel=large" which would allow the program to compile and run. For the Intel and PGI compilers, even though they also had support for a large memory model, there was no such luck. The software did compile, but upon execution the program return "Not a Number", NaN, for all the calculations. As such, these were not able to run any real simulations, even though they went through the motions. Nevertheless, there now was one working compiler that managed to handle the size of the system in question. It could compile the software, it ran and it returned numbers. The nail in the coffin for this compiler was the lack of a "Message Passing Interface" (MPI) wrapper that could compile the program to run at multiple CPUs. At the current time, no such wrapper existed, that we could find for this compiler. Therefore, the compiler could not be used as the size of the systems demands parallel processing to be able to do any simulations within a reasonable timeframe.

By trial and error, the limit of MD43 was established to be ~15.000 atoms before crossing into the domain of a large memory model. Reducing the size of the system was contemplated before the reality of the low number of atoms supported was discovered. The general idea was to reduce the base (bottom part) of the AC model to reduce the number of atoms; however, the size of the pores and the amount of molecules in the liquid phase could not be reduced by any considerable amount. The AC model has 31.232 atoms, and the total system has 75.560 atoms. Thus, the liquid phase of the system is actually responsible for slightly more than half the number of atoms. It was clear that reducing the size of the system was not a viable option.

At this point Prof. Tatiana Kuznetsova contacted Alexander Lyubartsev, one of the authors of the software. His recommendation was to try MD51 instead, as it had been reworked to some extent. Alexander Lyubartsev ran successfully a system of 75.000 water molecules, i.e. 225.000 atoms. This was taken into account, and the decision was made to try it. The source code for MD51 was downloaded from the MDynaMix homepage.

As MD51 proved to cope with the system size, it was decided to utilize this new version. However, this version lacks our in-house modifications. As such, the support for quarternions was no longer an option. Furthermore, it had no support for output in the pdb file format; Prof. Kuznetsova added this on request. Some additional modifications were also made in the part of the program reading the mmol (molecular input) files. This only supported a four-letter description of the atoms. The AC model on the other hand had six-letter descriptions, a character for the element and a unique number denoting which atom it was. MD51 also had

better support for parallel processing than the older MD43 version, with support for 128 and 96 CPUs respectively.

6.3.2 Computational requirements in terms of processing power

A downside by using MD51 is that the constrained dynamics option, where bond lengths are fixed, is apparently much less robust. The reason for this is yet unknown, but we attribute it to the fact that MD51 abandoned the updating of velocities as a part of the second half step. It is observed that constrained dynamics breaks down much more easily in MD51. This is unfortunate as constrained dynamics decreases the number of force-pairs that has to be calculated each time step, and thus speeds up the simulation. Another reason for using constrained dynamics is that a larger time step can often be used. This is because it is the high frequency vibrational motion of the bonds involving hydrogen atoms that requires such short steps in order to gain high enough resolution to model these correctly. Nevertheless, MD51 was working while MD43 was not.

Keeping in mind that there were 89.098 bonds, 82.548 angles and 115.176 torsions in this simulation, not to mention the LJ interactions and the long-range electrostatic interactions, the numbers clearly indicate that a lot of computational work has to be carried out during a simulation. Currently, when running on the Hexagon cluster (supercomputer at Bergen Center for Computational Science) using 128 CPUs the simulation is only able to progress with a speed of ~1250-1350 time steps each hour. While this might seem like a lot, the truth of it is that it is painfully slow. Especially when considering that the time step currently is set to 1fs, and that a simulation should run at the least run in excess of several nanoseconds, to accumulate statistical significant data. In order for the simulation to reach 1ns, it has to do $1 * 10^6$ time steps. At the current speed, a reasonable estimate with using 1300 time-steps pr hour would then be 32 days for each nanosecond. More computational power would have been appreciated. Nevertheless, while the computational resources are available, MD51 does not support more CPUs.

6.4 Validation runs

To carry out validation of the independent models each species has to have two separate runs. One simulation has to be carried out as a gas phase; another has to be carried out as a liquid phase. By utilizing the data gathered from these simulations, the molar enthalpy of vaporization can be estimated. This procedure was carried out for both of the PCB congeners.

For the triolein this was not an option as there was no experimental data on the enthalpy of vaporization available. This is due to that triolein actually decomposes before reaching its boiling point and as such experimental measurements of the enthalpy of vaporization is not possible under normal conditions.

6.4.1 Gas phase run

The gas phase runs were set up identically for both of the PCB congeners. It was a single molecule simulation with a temperature of 360.0K. The internal thermostat had a relaxation time of 50fs for both simulations. PBC was a cubic box with dimensions of 1000Å in x, y and z dimensions. The time step for the simulation was set to 0.1fs, and the duration of the run was set to 2ns, or 20.000.000 time steps. This was done as a single CPU run, as parallel processing would not have had any positive effect on such a small system. Additionally, this run was carried out using MD43, as MD51 does not support internal thermostats. When trying to use MD51 for this task, the internal temperature was 0K, while the translational and rotational temperatures were several thousand degrees Kelvin.

6.4.2 Liquid phase run

For the liquid phase runs, both simulations had 128 PCB molecules. The density of the systems was set to 1.5 g/cm^3. System temperature was set to 360.0K, and Nosé-Hoover thermostat was used with relaxation time of 15fs. The time steps was set to 1.0fs, and simulation ran for 1.000.000 time steps, which corresponds to 1ns. X, y and z dimensions were 35.90Å in each direction. The systems ran using 32 CPUs.

7 Results and discussion

In this chapter results from validations and simulations will be presented. The first part deals with the validations of the individual models. Followed by results and discussions around the results obtained from simulation run of the complete system. In the final part some discussion around the observed behavior of the PCB congeners in terms of the torsion angle parameters are presented.

7.1 Validation of the models

Validation of the individual models is an important aspect of running a realistic simulation. If the individual models fail to replicate the properties of the real molecules, then clearly the results from any simulations using these models will be dubious at best. In this thesis, thermodynamic properties were used to run validation tests on the PCB congeners. If the models could replicate, within a certain degree, the thermodynamic properties of PCB from experimental work, then the models would be approved. If on the other hand the models would fail to replicate the thermodynamic properties, further analysis and subsequent modification to the models would be necessary.

The thermodynamic property that was used in the validation of the PCB congeners was the molar enthalpy of vaporization. This property can be estimated by running two simulations, one as a liquid phase and one as a gas phase. Through the assumption of ideal gas for the gas phase, the molar enthalpy of vaporization can be estimated from equation (7.1).

$$\Delta h_{vap}(T, P) = \left[\langle u_{gas}(T) \rangle + RT \right] - \left[\langle u_{liq}(T, P) \rangle + Pv \right] \tag{7.1}$$

Here R is the universal gas constant; T is temperature in Kelvin; P is pressure in Pascal for the liquid phase; v is the molar volume for the liquid phase and $\langle u_{gas}(T) \rangle$ and $\langle u_{liq}(T, P) \rangle$ are the molar internal energies for the gas and liquid phase respectively.

Pressures obtained from the MD simulations for the liquid phase were unrealistically high, as shown in the respective sections in the results. The assumption made was that the pressures should in reality be close to 1 atm, at the current temperature and density. From the molar volumes calculated which were in the order of 10^{-4} m^3/mole, the contribution from Pv were

insignificant when added on to the value for $\langle u_{liq}(T,P)\rangle$. Thus, no error of significance is introduced by omitting the Pv contribution in the calculation of the molar enthalpy of vaporization. Due to the above reasoning, a simplified version of equation (7.1) was used for calculating the molar enthalpy of vaporization. This simplified expression is given in equation (7.2).

$$\Delta h_{vap}(T,P) = \big[\langle u_{gas}(T)\rangle + RT\big] - \langle u_{liq}(T,P)\rangle \qquad \textbf{(7.2)}$$

7.1.1 PCB congener 77

From the gas and liquid phase simulations for PCB congener 77, estimates of the molar internal energies were obtained. These were used along with the temperature of the simulation to calculate the molar enthalpy of PCB congener 77 as a gas. Similarly, the molar enthalpy of PCB congener 77 as a liquid was calculated. By taking the difference between these two enthalpies, the molar enthalpy of vaporization is found.

Table 7-A: Data from gas and liquid phase runs for PCB congener 77.

U_{gas} [kJ/mole]	U_{liq} [kJ/mole]	T [K]	P [MPa]	V [m^3]	V_m [m^3/mole]
214.019 ± 3.598	124.810 ± 0.033	360.0	274.87 ± 0.73	4.63E-26	3.32E-04

The data in Table 7-A is the computational values returned from MD51 and MD43, liquid and gas phase simulations respectively. It is immediately clear that the pressure obtained in the liquid phase simulation is not even close to the target pressure of approximately 1 atm. To use this pressure in the subsequent calculations of the thermodynamic properties would be to commit a severe error of judgment.

In Table 7-B the individual contributions for eq. (7.2) has been calculated. Since the Pv term has been intentionally neglected, the molar enthalpy for the liquid phase is equal to the molar internal energy of the liquid phase. The molar enthalpy for the gas phase contains the contribution from the RT term.

Table 7-B: Calculated molar enthalpies and quantities from sim. data for PCB congener 77.

$h_{gas}(T)$	h_{liq}	RT
[kJ/mole]	[kJ/mole]	[kJ/mole]
217.012 ± 3.598	124.810 ± 0.033	2.993

The final step in order to obtain an estimate of the molar enthalpy of vaporization is to subtract the molar enthalpy of the liquid phase, from that of the gas phase, as shown in eq. (7.2). This leads to the result in Table 7-C, where the estimated value is shown, along with the experimental value (Nakajoh et al. 2006).

Table 7-C: Molar enthalpy of vaporization for PCB congener 77.

Δh_{vap} estimate	Δh_{vap} experimental	Deviance
[kJ/mole]	[kJ/mole]	[kJ/mole]
92.202 ± 3.631	83.47 ± 1.85	8.73 ± 5.48

The estimated value for the molar enthalpy of vaporization does fall outside of the experimental value. However, this limited deviation for the model is acceptable.

7.1.2 PCB congener 118

The procedure here is identical to that for PCB congener 77. Again, the Pv term has intentionally been omitted for reasons stated above in section 7.1 and 7.1.1.

When comparing the values in Table 7-D with that of Table 7-A some differences are observed. The pressures obtained for PCB congener 118 is substantially lower than the pressures obtained for PCB congener 77. Some differences in the pressures were to be expected, although the pressures should have been approximately 1 atmosphere which corresponds to 101.325 kPa, however the difference in pressure between the two congeners of more than 2000 atmospheres are well outside acceptable deviations. Again, this only contributes to confirm that the pressures obtained are unreliable. There are also some differences when comparing the gas and liquid phase molar internal energies, though they are not large.

Table 7-D: Data from gas and liquid phase runs for PCB congener 118.

U_{gas} [kJ/mole]	U_{liq} [kJ/mole]	T [K]	P [MPa]	V [m^3]	V_m [m^3/mole]
221.848 $\pm$ 4.361	132.347 $\pm$ 0.048	360.0	71.37 $\pm$ 1.26	4.63E-26	3.32E-04

Again, the calculation of the molar enthalpies of gas and liquid phases were carried out. The molar enthalpy for both the gas phase and liquid phase are slightly higher for PCB congener 118 (Table 7-E), as to that of PCB congener 77 (Table 7-B). Contribution to the gas phase molar enthalpy from the RT term is the same for both systems, which should be obvious since the temperature is the same.

Table 7-E: Calculated molar enthalpies and quantities from sim. data for PCB congener 118.

$h_{gas}(T)$ [kJ/mole]	h_{liq} [kJ/mole]	RT [kJ/mole]
224.841 $\pm$ 4.361	132.347 $\pm$ 0.048	2.993

By subtracting the molar enthalpy of the liquid phase from that of the gas phase, the estimated value for the molar enthalpy of vaporization is obtained. This value is shown in Table 7-F. When comparing this to the experimental value (Nakajoh et al. 2006), and considering the uncertainties, the estimated molar enthalpy of vaporization agrees with the experimental molar enthalpy of vaporization. The result of this validation is promising, and better than expected.

Table 7-F: Molar enthalpy of vaporization for PCB congener 118.

Δh_{vap} estimate [kJ/mole]	Δh_{vap} experimental [kJ/mole]	Deviance [kJ/mole]
92.494 $\pm$ 4.409	87.03 $\pm$ 3.88	5.46 $\pm$ 8.29

This does not guarantee that the model will behave as the real molecule under all circumstances. Nevertheless, it does indicate that the model is capable of representing PCB congener 118, at least in some circumstances. A more detailed point will be made of this in section 7.3, where the OPLS parameters for the torsion angles for the two different congeners are compared to the energy profile for these torsions from QM simulations.

7.1.3 Triolein

For the PCB molecules the validation of the models were done by comparing the estimated molar enthalpies of vaporization to those obtained from experiments. For the triolein molecule this was not an option. Triolein has a high boiling point, and the molecule decomposes and breaks into smaller parts long before it actually reaches the boiling point. The search after experimental molar enthalpies for triolein was thus unsuccessful. A different approach to the validation of this molecule is needed.

Molar enthalpies of vaporization can also be estimated from expressions for functional groups, molecular weight and boiling points. However, the results from such methods are vague at best. To pass them on as an equal to experimental data would be completely out of place. As such, it can be discussed whether a validation towards data obtained in this manner is indeed valid.

A better option might be to validate the radial distribution functions (RDF) from liquid simulations to RDFs obtained from structural analysis by experimental X-Ray diffraction data. This is frequently done with larger biological molecules, which is what triolein is. Due to the limited time that was available, this has not been done. Nevertheless, it is suggested for further work in section 9.1.2.

7.1.4 AC model

Again, facing the task of validation of the model is difficult. AC does not decompose in the same manner that triolein does, however the melting point of graphite is very high, 4800 ± 200 K at ~100 MPa (Savvatimskiy 2005). At this point, it also ceases to be graphite, as it loses the hexagonal rigid structure and simply becomes liquid carbon, or sublimes into carbon gas. Given that the AC model is a solid, and that the characteristics of interest is its high efficiency as an adsorbent, the molar enthalpy of vaporization would have been the wrong way to validate this model anyway. The reason for this is that the molar enthalpy of vaporization does not give any true insight as to if it can replicate the adsorption characteristics of AC. As with triolein, a different approach to the validation of this model is required.

Activated carbons are used as adsorbents in wide areas, and a normal way to classify its properties as an adsorbent is by BET-isotherms of adsorption of N_2. Thus, a simulation run with N_2 would have been a viable method of validation. Ideally, the BET-isotherm should be

acquired for the Norit SA 4 PAH HF activated carbon that Oterhals used in his experimental work, but BET-isotherms from other AC could be used as well, for a more general approach. This has not been done in this thesis due to the short time available. It is unfortunate that the AC model has not been validated in any way. The AC model is a major part of finding the trapping mechanism for the PCBs. In the section 9.1.3, which are proposals for further work, the validation of the AC model through simulations with nitrogen, and subsequent comparisons with experimental data have been suggested.

7.2   Complete system simulation

The bulk of the time spent on this master thesis has been used building the individual models, finding the appropriate parameters for the models, and making a working simulation of the complete system. Just the task of putting the individual models together in a larger system, compiling and making this run took months of work. With this in mind, it should be no surprise that the final simulation of the complete system is an immature system. At the time this is written, it has barely managed to complete 0.5ns. Considering the immaturity of the system, the results displayed in this section can only be viewed as an indication of the behavior of the system, and not as de facto results.

7.2.1   Visual observations

Part of the analysis of the main simulation was done by visual inspection. Visual observations that employ human eyes that are quite powerful can note certain key features and deficiencies that might be obscured by graphs and charts. As mentioned in section 6.1.1, the number of molecules in the liquid phase was selected such that the density of triolein at one atmosphere would be replicated when the liquid fills the pores, and extends 20Å above the AC model. When studying the left hand side of Figure 7-1, it becomes clear that the liquid phase in the starting configuration extends much further than 20Å along the z-axis on top of the AC model. As such, the starting liquid density in the system is incorrect.
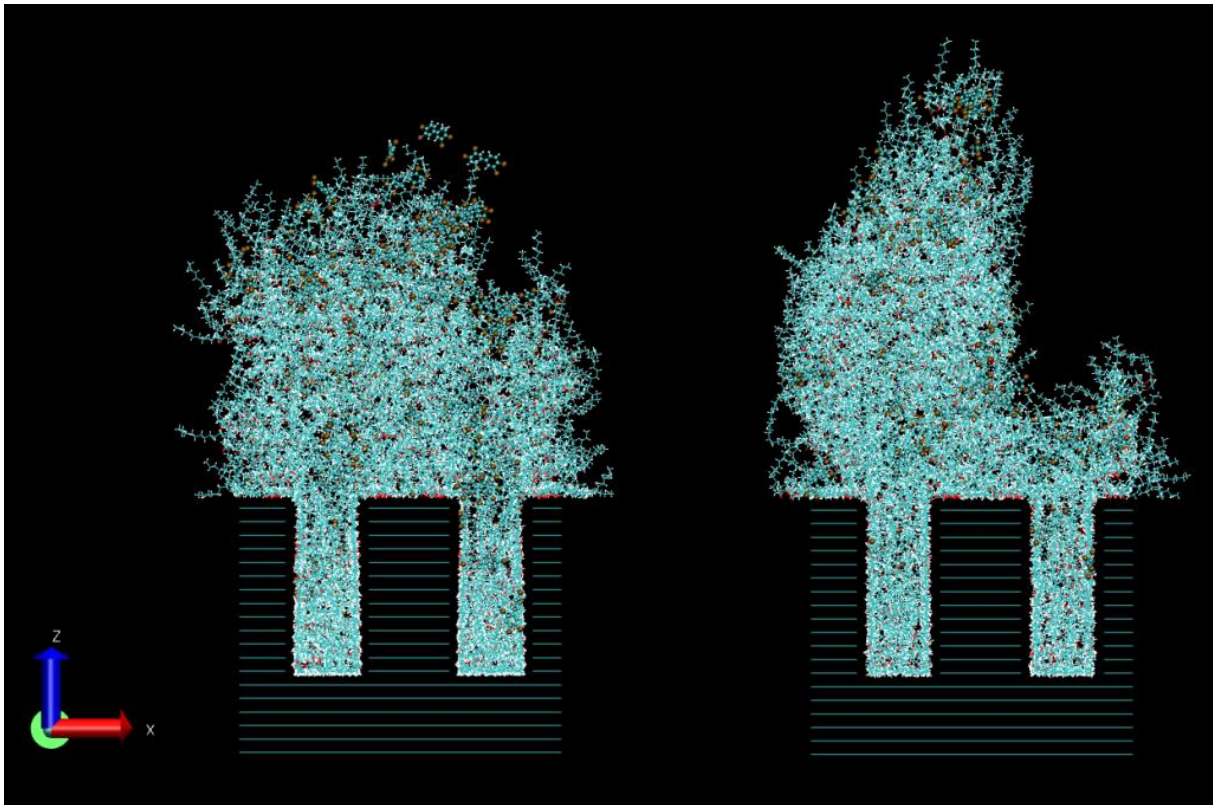
Figure 7-1: Comparison of starting configuration (left) and system after 0.45ns (right).

When the external force was on, and at its maximum, the density of the liquid phase was too high. During the necessary downscaling period, the liquid phase expanded, as expected. It was known that the starting configuration for this run did not have the bulk density of 915 kg/m^3, but one would expect the final liquid density distribution to be reasonably close to the bulk one. Our observation led us to believe that the LJ potentials apparently are not strong enough to keep the liquid in the system close to the AC model, in the current setup and other methods must be applied. One such method would be to be more restrictive with the PBC, and setup the system in such a way that the space between the top and bottom of the AC model would be 20Å. It can however be argued whether this would then truly represent the liquid phase that we want to study. In addition, due to the long-range electrostatic interactions from the pores in the graphite, there would be no guarantee that the entire systems behavior would not be dominated by these interactions. Prof. Kuznetsova suggested what could be a simple solution to this problem in a discussion about this shortcoming. Adding an atmosphere of nitrogen atoms, or other gasses, on top of the liquid phase could possibly alleviate some of these effects. This method will likely be the easiest way to accomplish a higher system pressure, while avoiding the undesired side effects discussed above. Addition of a gas on top will have

a cost in computational time, but it is likely that the benefits by adding the gas will outweigh the cost.

The right hand side of Figure 7-1 shows the system after 0.45ns, i.e. after roughly two weeks of run time. At this time, the simulation has deteriorated. The liquid phase now extends 129Å from the top of the AC model. It has also shifted to the left, and rearranged itself in such a manner that there is a void ~20Å above the right pore. No molecules are present in this volume. However, over the left pore the case is completely the opposite. Here, there is an abundance of molecules and no sign of such a void. Also, note that on the right hand side of the figure, the liquid phase peak is now closer to the base of the AC model than to the top of the AC model due to the PBC. There exists a possibility that this effect correctly depicts either the impact of the AC pores of the surface roughness of the PCB-triolein liquid. This extremely skewed distribution in the liquid phase profile might turn out to be a temporary occurrence as the system responds to the combined effect of the external force switching off and the unphysical influence of the AC models bottom due to the PBC. In such a case, a possible remedy would involve the use of annealing quenching, i.e. a drastic decrease in temperature and subsequent warm-up, to dampen dynamic memory effects. The unphysical effect may be mitigated through the introduction of a protective nitrogen blanket above the liquid phase, similar to quite a few experimental setups involving biological compounds.

Another interesting feature is that the triolein molecules are consistently oriented with the backbones containing oxygen atoms into the bulk liquid phase, while the hydrocarbon tails are pointing outwards. In Figure 7-2, the large blue and pink balls in the liquid phase is backbone oxygen and carbon atoms of the triolein molecule, respectively. Additionally, in this figure the atoms, which are colored by partial atomic charges, reveals the tendency of the oxygen atoms in the triolein molecules to arrange themselves in close proximity to the pores of the AC model.
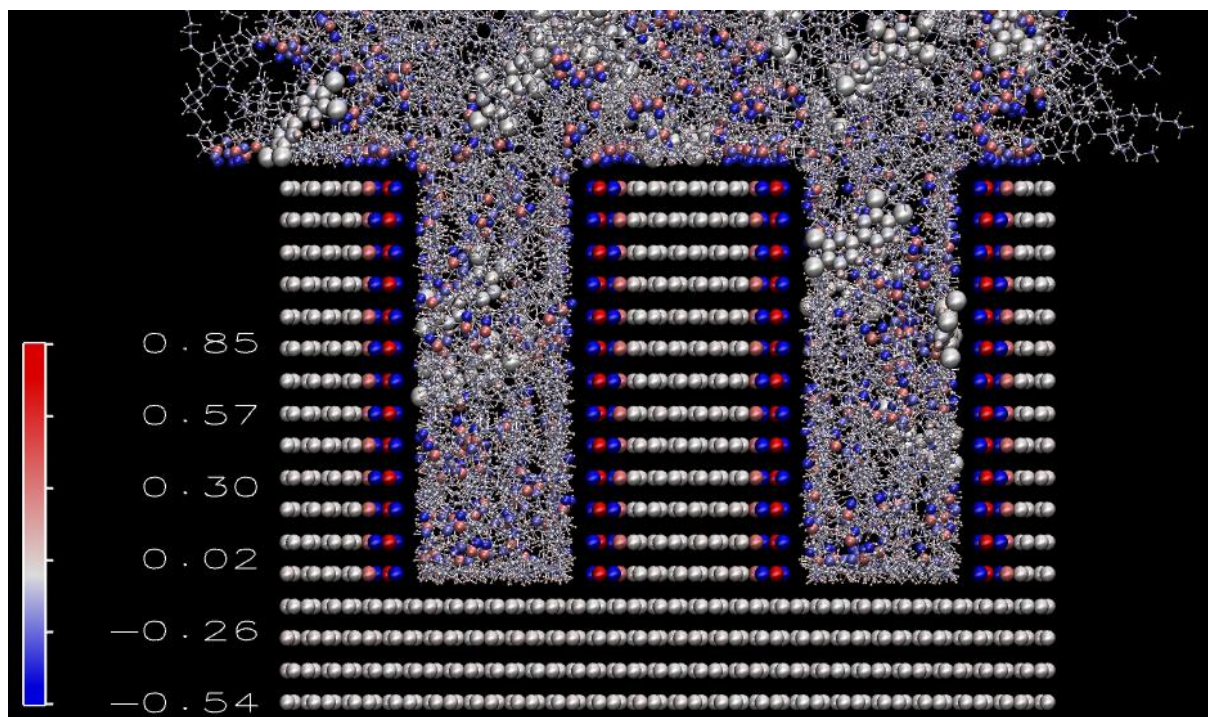
Figure 7-2: Orientation of triolein towards charges on AC model.

As discussed previously, and easily seen by studying the color bar scale, partial atomic charges of atoms closer to the pores are quite large. The electrostatic interactions between the pore walls, and the carbonyl functional group, will thus dominate. For example, a triolein molecule that has managed to orient itself above the charge distribution in the pore walls, has never been seen to leave. The hydrocarbon chains do move around, but the backbones, appear to have a restricted range of motion. Whether this truly replicates a real-life system, or if this is only an artifact of rather strong partial atomic charges in this simulation remains to be ascertained. It is also clear from Figure 7-2 that the partial atomic charges on the PCBs are quite insignificant compared to some of those in the triolein model and the AC model.
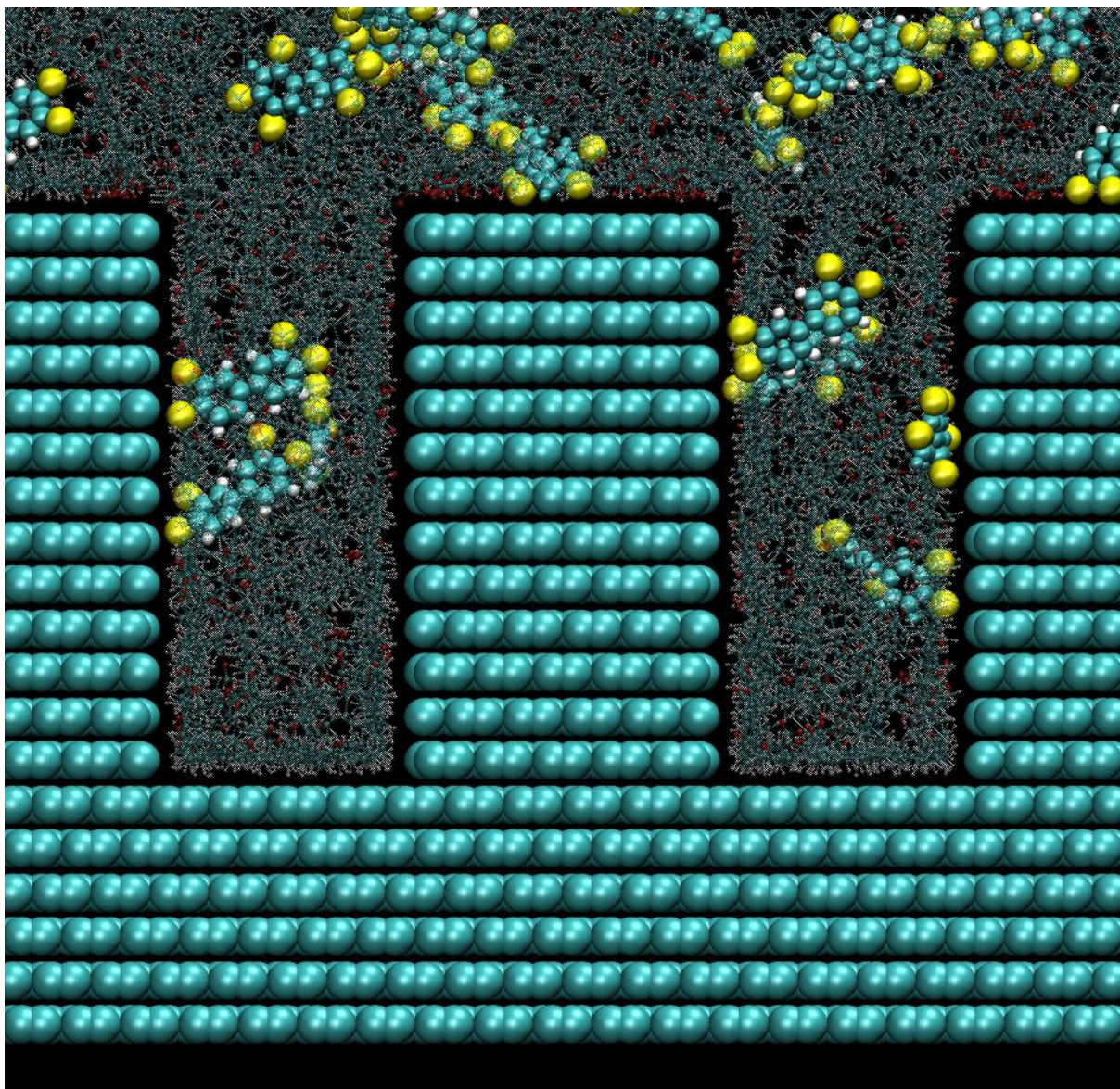
Figure 7-3: Orientation of PCBs trapped in pores.

Figure 7-3 shows the PCBs trapped in the pores of the AC model. From the orientation of the trapped molecules, the chlorine atoms on the PCBs seem to show a preference to the AC model. Five of the seven PCBs in the pores are oriented with two chlorine atoms in close proximity to the AC model, while the rest of the molecule extends into the center of the pores. Additionally two molecules in close proximity of the upper side of the AC model display similar behavior. Of the seven trapped molecules, only two is oriented flat against the AC model. The one in the right pore is clearly visible and positioned adjacent to the pore wall, the other PCB that is in the left pore has the same orientation, but is separated from the AC model by a ~4-5Å layer of triolein molecules. This is brought to attention due to that in the behavior displayed here whether the molecule has a planar conformation or not seems to be irrelevant. If positioning the chlorine atoms close to the pore walls is enough to trap the molecules, then

with the current pore diameter the trapping mechanism appears to be less sensitive to whether a molecule has a planar conformation or not. That is if the molecules that are only "attached" by their chlorine atoms can indeed be considered to have adsorbed. This can be discussed, but when compared to the PCB congener 118 that has a planar conformation and is placed flat and adjacent to the pore wall on the right hand side, the difference is very distinct. From this comparison it is clear that the "attached" molecules cannot be said to have undergone physical adsorption in same extent as the single PCB congener 118 molecule. However, if the model is capable to represent the real molecule, then this is in contradiction to the expectations based upon the hypothesis that is investigated. This behavior would be expected of PCB congener 77, but congener 118 should, according to the hypothesis not display this behavior. The rotational barrier and the increased energetic state of the molecule due to the unfavorable angle of the phenyl-phenyl torsion angle should have been deterring the molecule from this conformation. Insufficient statistical data is available to draw any conclusions based on this one molecule, not to mention that at this point the system has only been running for 0.5ns. Nevertheless, this observation, along with the general observation that the phenyl-phenyl rings torsion angle appears to be close to identical for both the congeners, suggests that a closer look at the OPLS parameters for these torsion angles in the PCB models is warranted.

7.2.2 Radial distribution functions

During the downscaling period of the external forces some radial distribution functions (RDF) were made using VMD. These RDFs were between each of the PCB congeners in the pores including a 5Å layer on the top of the AC model, and the AC model itself. Similarly, two new RDFs were made, but this time for the "production" run of the simulation. These new RDFs are made based on simulation data from ~0.5-0.6ns. All four RDFs are plotted in Figure 7-4. This figure shows that during this time there has been a new development in the system. In the RDFs from the downscaling period, the PCB congener 77 (dashed blue line) displayed a higher affinity for the AC model than the PCB congener 118 (dashed red line).
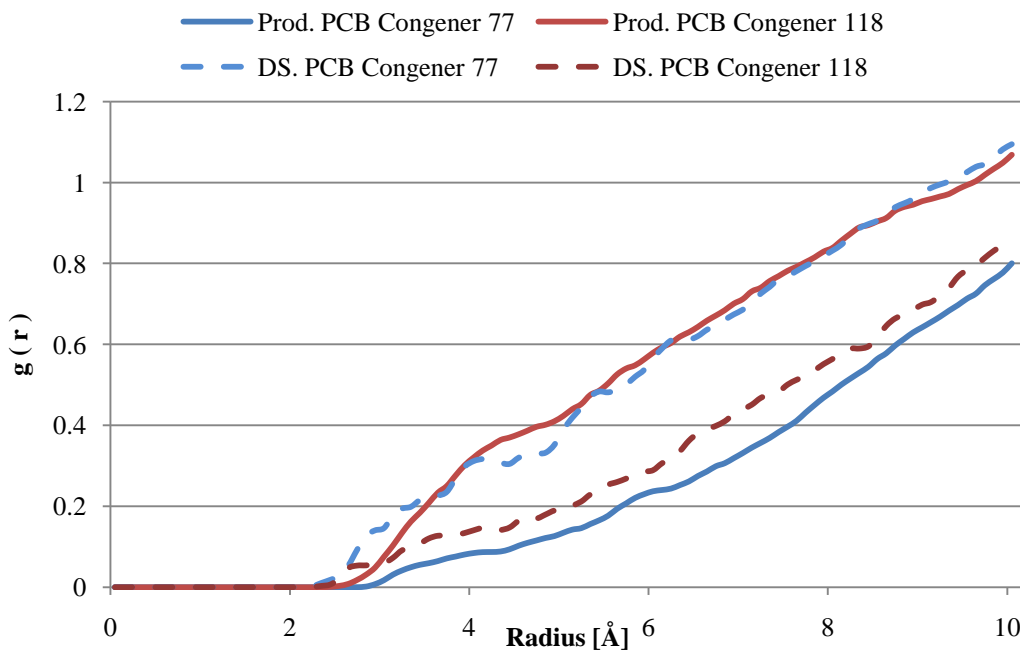
Figure 7-4: RDF comparison of PCB in pores between the downscaling (DS) and production runs.

However, when plotting the RDFs for the production run, there has been a significant change. Now PCB congener 118 displays the highest affinity towards the AC model. Actually, the RDFs swapped places so consistently that the data was triple checked in order to ascertain that it indeed was correct, and that no mislabeling had occurred. Furthermore, the integral values have changed significantly for both PCB congeners. These values are not displayed in the figure, but the values in the downscaling period were 0.93 for PCB congener 77 and 0.64 for congener 118. In the latest RDFs, the same integral values are 0.18 and 0.28, respectively.

The same trend is observed when plotting RDFs for the AC model with the chlorine atoms of the two PCB congeners, Figure 7-5. Again, the PCB congener 118 comes out with the highest probability for being close to the AC model. PCB congener 77 consistently shows less affinity towards the AC, until the radius is approximately 9Å. Interesting enough, from a distance of ~4Å to ~8Å the curves are very similar in form, but only of different magnitude.

84
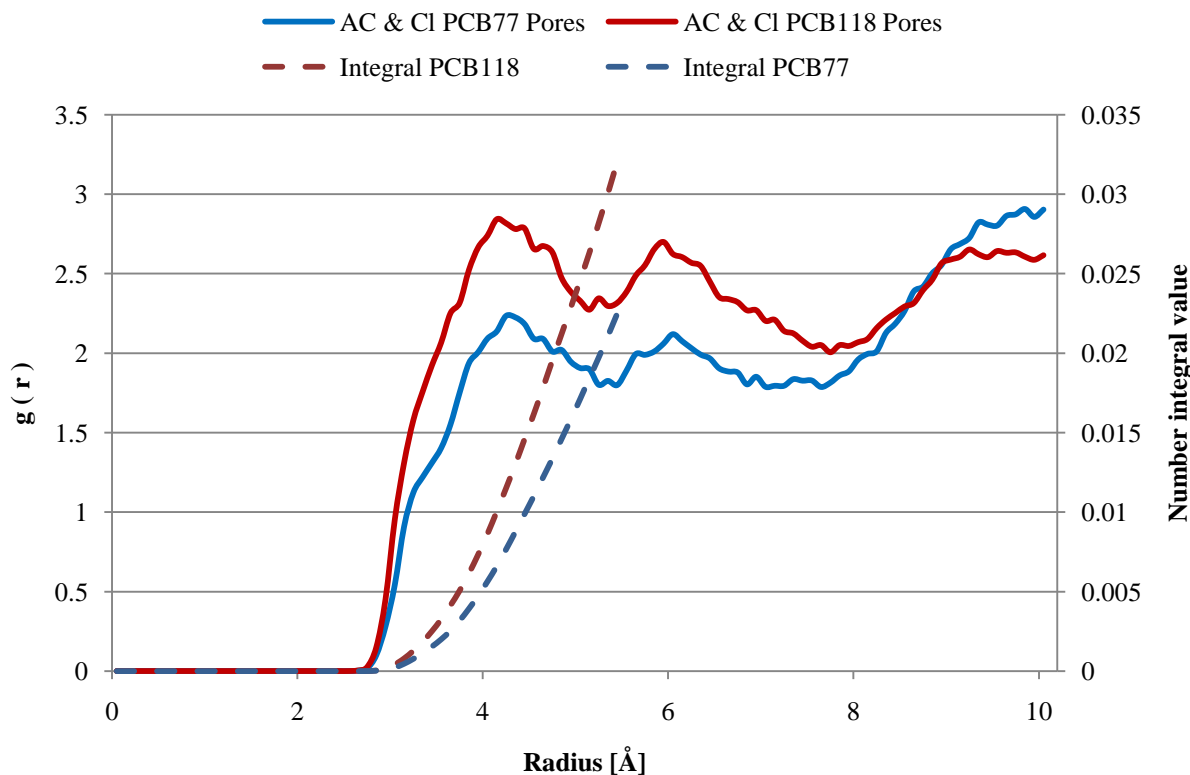
Figure 7-5: RDF of AC with chlorine of PCBs in pores.

The single molecule that has the planar conformation adjacent to the pore wall can explain some of the contribution to RDF of PCB congener 118. When the molecule is positioned in this way, all five chlorine atoms have approximately the same distance to the AC model, as such their contribution is significant. For the molecules that are only attached by two chlorine atoms and where the rest of the molecule extends into the pore, the contributions are averaged over a larger volume.
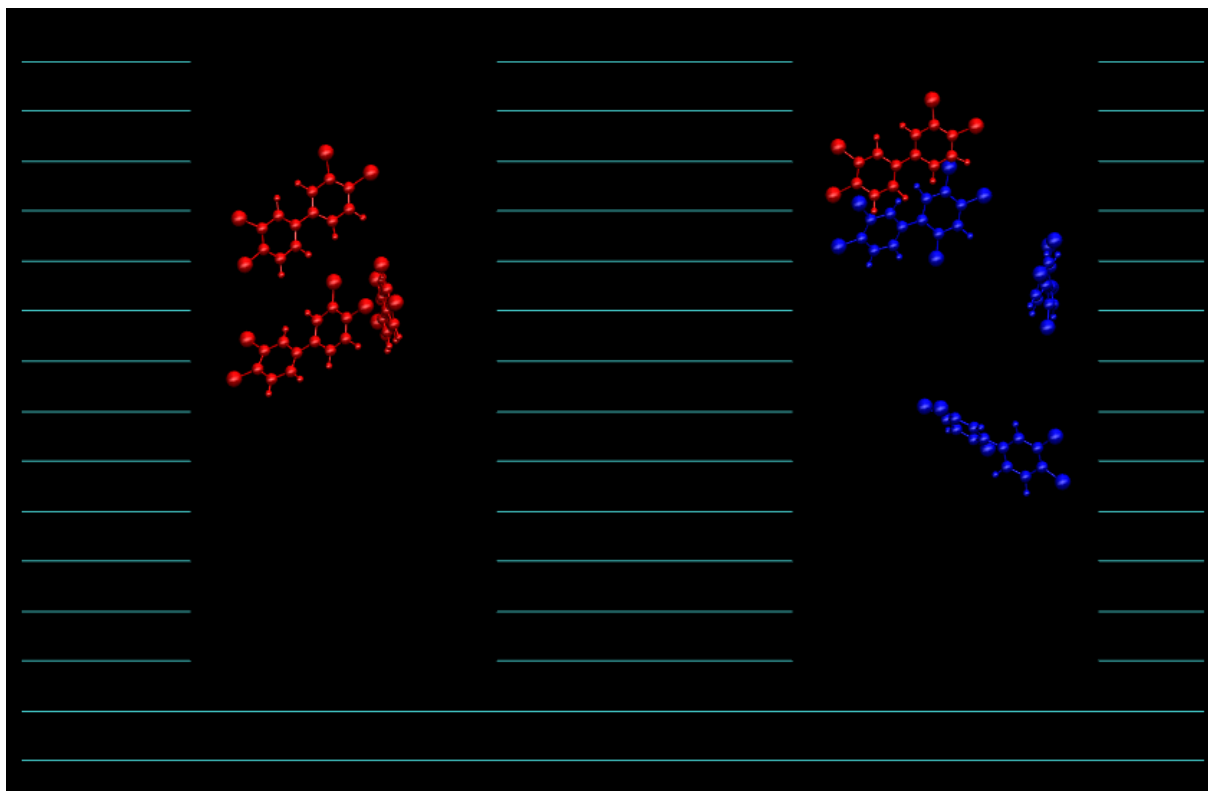
Figure 7-6: PCB congener 77 (red) and PCB congener 118 (blue) in pores at time of RDFs above.

When comparing Figure 7-6, to the data from the RDFs it is clear that the RDFs do not give the complete picture. For one, the RDFs indicate that there are larger amounts of PCB congener 118 in the pores, than PCB congener 77. By visual inspection it is clear that this is not the case. The number of PCB congener 77 is 4 versus 3 for PCB congener 118. However, the RDFs give contributions from individual atoms. Moreover, from the figure only one PCB molecule can truly be considered to have undergone physical adsorption. The molecule staying in the same relative orientation throughout several trajectory files supports this. The molecule in question is of congener 118. This molecule has a planar orientation relative to the pore walls. Five other molecules do show some interactions with the pore walls. In these molecules, the chlorine atoms show a preference to anchor to the pore walls. For the one adsorbed molecule, from visual inspection of the trajectory, it seems that the loss of a rotational degree of freedom might be compensated for by the favorable lower energy state from the LJ potential interaction with the pore wall. This is based on the observation that the molecule, when adsorbed to the pore wall in such a manner, was not seen to rotate around all three axis. However, no entropic energy calculations has been carried out, and MD51 does not support this for one specific molecule, so no confirmation in way of numerical estimates can

be given. In section 7.3, the torsion angle parameters for the phenyl-phenyl torsion angle is under extensive scrutiny, and it is not unlikely that better representations of this torsion angle will have significant impact on the molecules ability to adsorb to the pore wall in such a planar configuration. Another issue is that the amount of triolein molecules in the pores seems to restrict the freedom of movement for the PCB molecules. As such, the molecules that has already attached to the pore walls with the chlorine atoms, seems to experience great difficulties in obtaining a similar orientation to that of the one adsorbed molecule. In fact, the PCB molecules might have to tear loose from the pore walls first, before being able to orient in such a manner.
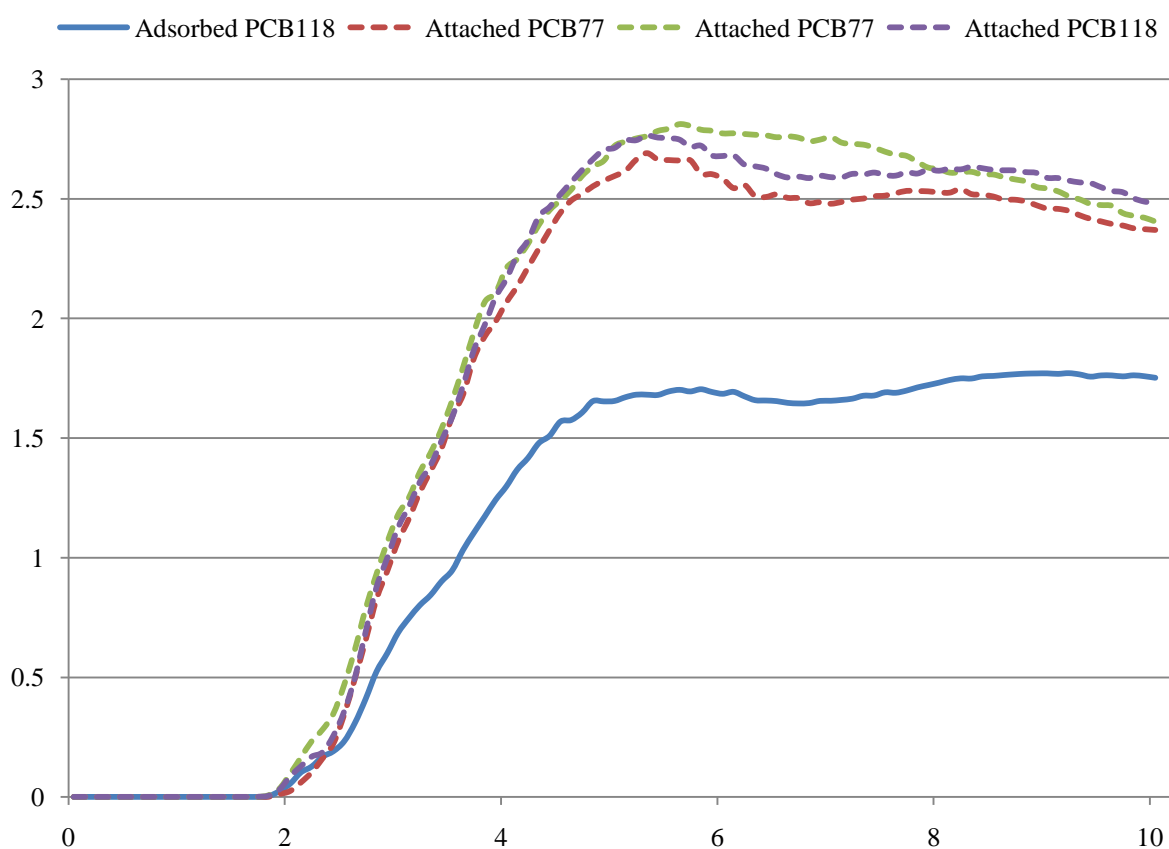
Figure 7-7: RDF for adsorbed and attached PCBs - a comparison.

Figure 7-7 clearly shows that the relative amount of triolein molecules around the PCBs that are not oriented in the planar configuration in relation to the pore walls is higher than for the single PCB that is discussed above, and that is considered to have adsorbed. This should be obvious as the PCB molecules that are only "attached" by their chlorine atoms have more of their surface available to interact with triolein molecules.

7.3   Evaluation of OPLS torsion angle parameters for PCB

During the simulation of the complete system, some discrepancies concerning the relative behavior of the two PCB molecules were observed. From the QM calculations, and the corresponding pictures of the optimized conformational geometry of the PCBs, Figure 5-10, it was seen that the torsion angles between the two phenyl-rings were different. The result from the QM calculations were, however, expected. The discrepancies that was observed in the simulation was that the models failed to replicate this behavior.

Considering that the two models both use the same force field parameters for those torsion angles, this was to some degree expected. Nevertheless, due to the fact that PCB congener 118 has a chlorine substitution in the 2-position, it was hypothesized that the more dominating LJ parameters might to some degree cause a repulsion between the hydrogen atom in the 2'-position and the chlorine atom. The hypothesis was that this repulsion would contribute to force the torsion angle in the PCB congener 118 more away from the equilibrium position, and as such replicate the difference between the two congeners. It was due to this chain of thought that the attention were focused upon the torsion angles for the phenyl rings, and that this discrepancy was observed.

The difference in this torsion angle might just be what enables PCB congener 77 to adsorb better than PCB congener 118. This has been one of the main theories throughout this thesis. As such, the torsion angles for the PCB congeners should have proper parameters that can replicate the equilibrium angles. Due to the observation that the torsion angles in question appeared to be virtually identical for both the congeners the decision was made to further investigate the parameters for the torsion angles between the phenyl rings. This was done in three steps. Initially, a profile of the energy contribution from the Fourier series used to model the torsion was generated by calculating the energy for every degree within a 360-degree rotation of the torsion angle. Furthermore, both the PCB models made in Maestro was used to do a "rigid coordinate scan" where the model was kept fixed, except for the torsion angles in question. This was set up such that a single-point-energy calculation was carried out one by one while keeping everything but the torsion angle fixed. The torsion angle was varied in 1-degree steps from 0 to 359 degrees. For this task DFT was used with B3LYP and the 6-31+G** basis set. Jaguar then created an output file, which held the respective molecular energies for the different geometries.

These energies were extracted from the output file, and put into Excel. In the QM calculations, the results were negative energies that varied with respect to the torsion angle. By subtracting the most negative value from all of the energies, positive quantities were obtained. This is in line with how MD treats bonds, angles and torsions, where deviation from equilibrium leads to a positive contribution to the internal energy of the molecule. At equilibrium, the contribution to the internal energy would be zero. From these data, graphs showing the energy profiles for the torsion angles between the phenyl rings were plotted, along with the energy profile from the torsion angle using the OPLS parameters.
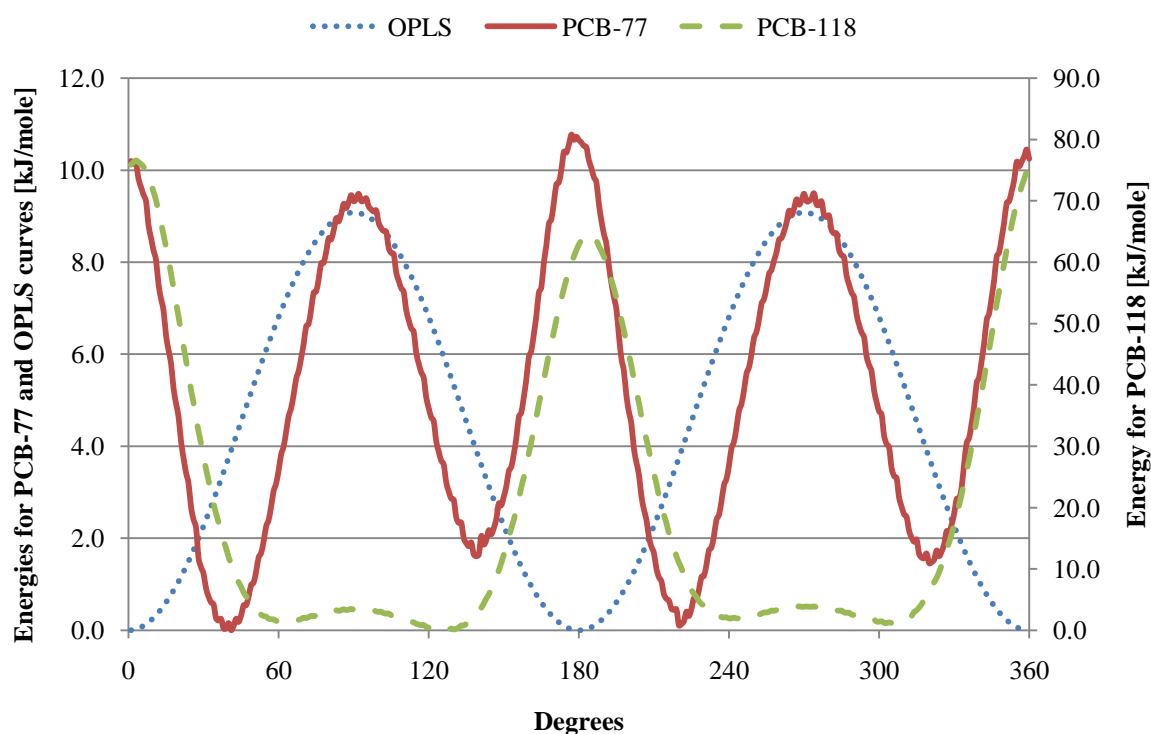
Figure 7-8: Comparison of energy contributions of torsion angle from OPLS and QM for PCB 77 and PCB 118.

When comparing the data obtained from QM for the individual congeners to that obtained from OPLS, which is identical for both congeners, in Figure 7-8 it is immediately clear that OPLS fails catastrophically to replicate the energy profiles. If one compares the magnitude of the energy contributions from the QM calculation, for the PCB congeners the difference is immense. The difference in energy between most favorable and least favorable geometry for PCB congener 77 is ~10kJ/mole. For PCB congener 118, the same difference is ~75kJ/mole. OPLS has a difference in energy of ~9kJ/mole, and as such can energetically represent PCB congener 77. However, OPLS fails to represent the global minimum for PCB congener 77, as it puts the global minimum at zero degrees. It also fails to represent the local minimums at

~140 degrees and at 320 degrees, in OPLS there is a minimum again at 180 degrees. This is a major shortcoming. When it comes to PCB congener 118, the deficiencies are far worse. As mentioned, the difference in energy between the most and least favorable geometry is much larger than what the OPLS torsion parameters describe. Furthermore, the global minimum is at 126 degrees, while the OPLS puts it at 0 and 180 degrees. For the PCB congener 118, there is a local minimum at ~61 degrees. This is in no way replicated by the OPLS torsion parameters, or event the functional form. The energetic difference between the local minimum (1.41kJ/mole) at ~61 degrees and the local maximum (3.49kJ/mole) at ~91 degrees, which separate the local minimum from the global minimum, is only ~2.08kJ/mole. Thus, if the PCB conformer 118 were at the local minimum it would not take all that much to twist it past the local maximum and into the global minimum. Therefore, when the OPLS parameters indicate that both the local minimum and the global minimum are almost the most unfavorable geometries it is clear that both the parameters and the functional form are not good enough.

Attempts were made to fit new parameters to the Fourier series in order to replicate the energy profiles from the QM calculations. This was done with MathLab and using the method of least squares. However, it was a futile attempt, which was more or less clear from the start given that the Fourier series used in the implementation in MD51 only supports three parameters, and none of these had the correct multiplier inside the cosine functions in order to replicate the four minima and maxima. Other potentials for the torsion angles, included in MD51, were examined. Unfortunately, none is capable for replicating, or even approximating the energy profiles obtained from the QM calculations.

A new potential has to be introduced in order to be able to approximate these profiles. This has not been done in this work, again due to time constraint. It has been listed in section 9, as a suggestion to further work.

As mentioned in section 7.1.2, validation of a model does not guarantee that it will behave exactly like a real molecule. The validation done in that section showed that the model was good, at least capable of approximating thermodynamic properties. However, the validation did not give any indications of the fact that the torsion angles between the phenyl rings were poorly represented. Without any expected behavior pattern of the molecules, this could have been missed, and it is very likely that such a mistake could have a significant impact on the results. From this, it is clear that some work still remains for the PCB models.

8 Conclusions

Based upon the results and discussions in section 7, some conclusions have been made. The validations of PCB congener 77 and 118 were successful, and the models are capable to replicate thermodynamic properties. Nevertheless, the models are still not accurate enough for the kind of simulations that they were intended. Considering that geometrical constraining of the torsion angles between the phenyl rings was our original working hypothesis, the models failure to replicate these torsions accurately might be considered a shortcoming. Thus, the models are not precise enough to give real insight into the trapping mechanism of PCBs, or to verify or invalidate Oterhals' hypothesis. Further work has to be carried out in order to replicate the functional form and the energy barriers for the torsion angles between the phenyl rings. In addition, since MD51 does not currently have any potential capable to replicate this functional form, modifications to the code is required to add support for such a potential. A Fourier series is likely the best option, but this time sinus functions has to be included as well.

For the triolein molecule, no validation was done. This is unfortunate and efforts should be made to validate this model in the future. As mentioned earlier, one possibility is to compare RDFs from liquid phase simulations of pure triolein molecules to RDFs from X-Ray diffraction studies. Similarly, the AC model should be verified. As previously discussed, this can be done by running a simulation with nitrogen gas and the AC model. Results from such a simulation can in turn be compared to experimental BET-isotherms.

For the complete system, it is clear that as it is now, the target density is not achieved. With a higher density, more molecules would have been in contact with the AC model. For a new run, actions should be taken to improve on this situation. As discussed above adding an atmosphere of nitrogen atoms might be one solution. One possible drawback is the solubility of nitrogen in triolein. However, this is a fairly common experimental setup, and attempting this might be worthwhile. Currently the number of PCB molecules in the pores is too low, for any statistical significant data to be gathered. It is also safe to conclude that simulations have to run for a much longer period of time, than what has been possible to do in these circumstances. From the data gathered from the complete system simulation, there can be no conclusions made as to the trapping mechanisms of PCB onto AC. Furthermore, no conclusions as to the selectivity of the adsorption process can be made. This is both due to the low number of PCB molecules in the pores, as well as the short period of time that the

simulation has been running, and as mentioned above, the torsion angles between the phenyl rings in the PCBs are not accurate enough. Even if these two main points had been fulfilled, there would be no guarantee that any conclusions drawn from this model would have been correct, as right now the AC model is not validated. Without the validation of the AC model, it is unknown if it displays similar properties as AC or not.

Under the current density, the charge distribution on the AC model appears to be dominating the behavior of the triolein molecules. The charged end of the molecule displays a strong preference for the top side of the pore walls. For the AC model, the charges were obtained from QM calculations with ESP as the method of obtaining the charges. These charges are from vacuum simulations, while all the other charges in the system are CM4 charges from a water solvated system. Again, due to the lack of time and difficulties getting the self-consistent field to converge, charges from a water solvated system has not yet been obtained.

It is the conclusion, that overall the models look promising and that this approach to finding the trapping mechanisms of PCB onto AC has some merits. Nevertheless, some work still remains with the models. A new master student continuing from where this work leaves this, with some of the suggestions for further work should be able to carry out most of the changes suggested here, as well as run a substantially longer simulation and acquire a better data set in order to evaluate Oterhals' hypothesis.

9 Further work

In this section, proposals for further work are listed. The different proposals have been divided into subsections. In the first section, there are yet another three subsections related to the individual models. The final two sections cover proposals for the complete simulation system, and modifications to MD51.

9.1 Future improvements to models

9.1.1 PCBs

Currently there are seven sets of charges available for the PCBs. One set is of charges obtained from ESP calculations in vacuum. The remaining six sets of charges are from Löwdin population analysis, Redistributed Löwdin population analysis and CM4 for both vacuum and water solvated system.

ESP charges are considered more correct for use in force field methods due to the way they are obtained. As such, ESP charges from a water solvated system would have been very interesting in order to compare these with the charges already obtained. An effort should be made to find software that could do this. Improved charges might contribute to make the models more accurate, with respect to the interaction with AC and triolein. In addition, a sensitivity study on the charges and the effect of the charges when it comes to interactions with triolein and the AC model should be carried out.

Another issue with the PCB models is the parameters for the torsions between the phenyl rings. The parameters from OPLS are in this regard, poor at best. They completely fail to give the minimum energy at the correct torsion angles. In addition, in OPLS the same torsional parameters are used for all PCBs. The QM calculation done over a 360-degree twist of the phenyl rings for PCB congener 77 and congener 118 show a large difference in the magnitude of the energy, and the energy distribution for the congeners. Special torsional parameters should be fitted to each of the PCB congeners. This will affect how the models behave. It may have a significant effect on PCB congener 118, when considering its ability to enter smaller pores.

9.1.2 Triolein

For triolein, the charges used in this model were obtained from CM4 for a water solvated system. However, ESP charges for a water solvated system would have been interesting for comparison reasons. A sensitivity study on the charges would help to determine the importance of correct partial atomic charges on this model. This is especially important considered that so far the triolein model has yet to be validated.

Furthermore, the verification of the triolein model would be beneficial for determining the validity of the simulation system as a whole. This was not performed since no experimental enthalpy of vaporization was found, or may even be available as triolein decomposes before reaching its boiling point. Different approaches to the validation of the triolein model should be investigated. One method could be to compare RDF for the triolein molecule in a pure liquid phase simulation, with RDFs from X-Ray diffraction. This is most likely the best way of validating this model.

9.1.3 Activated carbon and the GraphiteMaker utility

Validation of the AC model itself might not be viable through the normal methods. However, running simulations with AC and N_2 could be interesting in order to check how the adsorption isotherms turn out. Comparison of BET-isotherms for different types of activated carbons could be compared to a study of the adsorption of N_2 onto the AC model. The general idea is that the model can be checked against experimental data, and a measurement of how good or bad it is could be obtained. Right now, the workings of the AC model, good or bad, is unknown. It is however known that it lacks many of the features that make AC such an effective adsorbent.

Real activated carbon consists not only of carbon atoms. There are impurities or heteroatoms like oxygen, nitrogen, hydrogen and sulfur as well. Inclusion of heteroatoms in the pores of the model will not only contribute to make the model more realistic, it will most likely also affect the charge distribution in the pore to a large degree. Consequently, the behavior of molecules in close proximity to the pores, or in the pores will most likely be affected. This does not only affect the adsorption process through charges, but the different LJ parameters for different atoms can also have a significant effect.

As with the other models, the validity of the partial atomic charges can be questioned. In the simulation carried out in this thesis, the electrostatic interactions between the AC model and

the triolein completely dominates the distribution of the triolein molecules in the liquid phase. Obtaining water solvated ESP charges for the AC model would be very interesting for comparison reasons. Different software would have to be utilized in order to do so. At the current time, Jaguar does not support this combination. In addition, the trend when utilizing water solvation is that the magnitudes of the charges are reduced compared to the magnitudes of the vacuum charges. If this is also the case for the AC model, then this will help alleviate the dominating character of the electrostatic interactions. As such, CM4 charges would be at least interesting as a comparison, and likely a good option.

In order to build the AC model, the utility GraphiteMaker was a necessity. It was however only coded with one purpose in mind, and that was to create the model used in this thesis. Additional features should be added to this utility for future use. Currently the utility is hard coded when it comes to the dimensions of the AC model. The same applies to the shape of the model. This limits the models that the utility can generate to one. Modification to the program in order to allow the user to enter the dimensions of the system, as well as the number of pores and their respective sizes, depths and positions would greatly improve the utility. Also the ability to have variable pore diameters where the pore can narrow in as a function of depth would be interesting, and certainly relevant when trying to replicate the true structure of AC.

Another issue is the failure of the utility to replicate the chaotic, or unordered, nature of AC. By adding the option to have a random distribution of atoms, within a specified limit, in the pore would give a better representation of AC.

Along these lines, the corners of the top of the pores could have an impact on how easily molecules enters or leaves the pores. Having the option to use rounded corners in the pores may give some interesting insight into this mechanism. Prof. Kvamme originally proposed this idea, and it is included here as it obviously has merits, and should be implemented.

Furthermore, the utility that builds the model should be merged with the utility that positions the charges. In addition, the option to supply the random number generator used for the charge distribution with a seed should included. This would allow a specific charge distribution to be used multiple times. As it is now, each time the ChargePlacer utility is used a new, unique, distribution is generated.

This model may be used to model several other systems with AC, given that continuous work with it is done. In order to do so, several of the modifications here should be carried out, as well as consecutive validations to see the impact of the different contributions in the model.

9.2 Future simulation systems

The simulation in this thesis was carried out with pore diameters approximately equal to 20Å. For future systems, both smaller and larger pores should be investigated as well. By doing this comparisons on how the molecules enter the pores, behaves in the pores and leaves the pores can be compared. As a result a better understanding of the trapping mechanism could possibly be reached. This would also allow better to confirm or reject Otherhals' hypothesis, as well as give inspiration for new more precise hypotheses. Pores in the range below 2nm should be given more emphasis, as these will maximize the restrictions on the rotational motions for the PCBs.

The inclusion of several types of triacylglycerols might be beneficial. By including more triacylglycerols, the model fluid for fish oils will be more complete. Focus should be on palmitic acid (16:0), and metnoe acid (14:0). The inclusion of erucic acid (22:1) could also be considered beneficial. Selective adsorptions of dioxins, dibenzofurans and PCB from fish oil onto AC or silicates can also be investigated using the model fish oil.

As discussed earlier, the addition of a gas atmosphere on top of the liquid phase can allow the liquid phase to have a density that is more along the lines of the target density. This was proposed by Prof. Kuznetsova during a discussion of the problem of the ever expanding liquid phase.

9.3 Future modifications to MD51

MD51 has to be modified in order to allow for a new kind of torsion angle potential. The current potentials available in MD51 fail to replicate the energy profiles around the connection of the phenyl rings that was obtained from the QM simulations. Fourier series using both sinuses and cosines will likely be able to approximate the functional forms.

Currently MD51 only give the interactions summed over the different types of molecule. For example, it will give the PCB congener $77-$ triolein interaction energy. In order to calculate the fractional coverage of by the Langmuir equation, or the Langmuir adsorption constant it

would be beneficial to add the possibility to output the interaction energy for specific sites. Since MD51 is already calculating these values, it is just a matter of modifying the code to output the interaction energies for specific sites on demand.

Due to the failure of the complete system to stay at the target density, some modifications to MD51 might alleviate this. Other simulation tools allow for a vacuum box or a continuum to be positioned on top of the simulation system. In theory, this could allow the simulation to be carried out, without a large increase in computational time, and at the same time keep the density of the system more consistent. This might be a bit ambitious, considering that the simple solution to the problem is just to add a gas atmosphere on top instead.

Another beneficial modification to MD51 would be to allow for PBC in only X and Y direction. This would be a good solution, when there is no real need for it. PBC along the z-axis is only adding extra computational work in this system setup. Such a modification could also potentially reduce the height of the base of the AC model, as it now is 20 Å to avoid electrostatic interactions across the AC model itself (from the fluid). By reducing the base of the AC model to half of what it is now, i.e. to 10Å, a reduction of ~7500 carbon atoms would be achieved.

The modification carried out on MD51 to allow for a constant external force was rudimentary. A more elegant solution would be to add the external force as a keyword on its own, and also to specify the force as an acceleration in Å/fs^2 instead of V/cm. This would be much more intuitive. Furthermore, it should be possible to set a time interval of when to downscale the force, as well as by how much it should be downscaled. This would be a great time saver for anyone using this feature, as they would not have to restart the simulation every time one downscaling step has been performed. The same feature would also have been beneficial when compressing a system to a target density. It is not always possible to immediately start a new simulation once one run has completed. The time lost can be significant when a compression phase can take weeks from the start.

Currently MD51 only allows 128 CPUs to be used. With a system of this size, the speed of the simulation is simply too slow. Modification of MD51 to allow for the usage of more CPUs would allow for systems on this scale, as well as systems of larger scales to be carried out within reasonable timeframes. As the computer industry is evolving, the availability of the power is not an issue right now, however the simulations appear to become larger and more complex and support for more than 128 CPUs is sorely needed. An interesting project

would be to make MD51 able to run utilizing graphics card. AMDs latest card, which costs ~3000 NKr has the theoretical computational power of 2.7 TeraFLOPS, while the complete Hexagon supercomputer has the theoretical computational power of 51.7 TeraFLOPS. Considering that Hexagon has 5552 cores then one such graphics card is the equivalent of 290 cores. A single desktop computer can have four of these, and *assuming* a very poor scaling of 50% then this would equal 575 cores (counted 100% of the first card and 50% cores for the other 3 cards), at the cost of 12.000NKr. Considering the rapid development of GFX cards this might just be the way to go.

References

Atkins, P. and J. de Paula (2002). Physical Chemistry, Oxford University Press.

Ballschmiter, K., R. Bacher, et al. (1992). "The determination of chlorinated biphenyls, chlorinated dibenzodioxins, and chlorinated dibenzofurans by GC-MS." Journal of High Resolution Chromatography 15(4): 260-270.

Bansal, R. C. and M. Goyal (2005). Activated carbon adsorption. Boca Raton, La. ; London, Taylor & Francis.

Cramer, C. J. (2004). Essentials of computational chemistry : theories and models. Chichester, Wiley.

Croxton, C. A. (1975). Introduction to liquid state physics. London, Wiley.

Damm, W., A. Frontera, et al. (1997). "OPLS all-atom force field for carbohydrates." Journal of Computational Chemistry 18(16): 1955-1970.

Dill, K. A. and S. Bromberg (2002). Molecular driving forces : statistical thermodynamics in chemistry and biology. New York ; London, Garland.

Downs, R. T. and M. Hall-Wallace (2003). "The American mineralogist crystal structure database." American Mineralogist 88(1): 247-250.

EPA (1980). PCB Congener table. U. S. E. P. Agency.

Fong, B. S., J. P. Despres, et al. (1988). "Interactions of High-Density Lipoprotein Subclasses (Hdl2 and Hdlc) with Dog Adipocytes - Selective Effects of Cholesterol and Saturated Fat Feeding." Journal of Lipid Research 29(5): 553-561.

Frenkel, D. and B. Smit (2002). Understanding molecular simulation : from algorithms to applications. San Diego, Calif. ; London, Academic.

Hinchliffe, A. (2003). Molecular modelling for beginners. Chichester, Wiley.

Hites, R. A., J. A. Foran, et al. (2004). "Global Assessment of Organic Contaminants in Farmed Salmon." Science 303(5655).

Hoover, W. G. (1985). "Canonical dynamics: Equilibrium phase-space distributions." Physical Review A 31(Copyright (C) 2009 The American Physical Society): 1695.

Jaguar (2008). Schrödinger. New York, NY, Schrodinger LLC.

Jensen, F. (2007). Introduction to computational chemistry, John Wiley & Sons Ltd.

Jensen, K. P. and W. L. Jorgensen (2006). "Halide, ammonium, and alkali metal ion parameters for modeling aqueous solutions." Journal of Chemical Theory and Computation 2(6): 1499-1509.

Jorgensen, W. L., D. S. Maxwell, et al. (1996). "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids." Journal of the American Chemical Society 118(45): 11225-11236.

Jorgensen, W. L. and N. A. McDonald (1998). "Development of an all-atom force field for heterocycles. Properties of liquid pyridine and diazenes." Theochem-Journal of Molecular Structure 424(1-2): 145-155.

Jorgensen, W. L., J. P. Ulmschneider, et al. (2004). "Free energies of hydration from a generalized Born model and an ALL-atom force field." Journal of Physical Chemistry B 108(41): 16264-16270.

Kelly, C. P., C. J. Cramer, et al. (2005). "SM6: A density functional theory continuum solvation model for calculating aqueous solvation free energies of neutrals, ions, and solute-water clusters." Journal of Chemical Theory and Computation 1(6): 1133-1152.

Koch, W. and M. C. Holthausen (2001). A chemist's guide to density functional theory. Weinheim ; Chichester, Wiley-VCH.

Kuznesova, T. (2001). Molecular modeling for thermodynamic properties of bulk and interfacial systems. [Bergen], University of Bergen, Department of Physics**:** 1 b. (flere pag.).

Lowdin, P. O. (1950). "ON THE NON-ORTHOGONALITY PROBLEM CONNECTED WITH THE USE OF ATOMIC WAVE FUNCTIONS IN THE THEORY OF MOLECULES AND CRYSTALS." Journal of Chemical Physics **18**(3): 365-375.

Lowe, J. P. and K. A. Peterson (2006). Quantum chemistry. Oxford, Elsevier Academic.

Lundebye, A.-K., M. H. G. Berntssen, et al. (2004). "Dietary uptake of dioxins (PCDD/PCDFs) and dioxin-like PCBs in Atlantic salmon (Salmo salar)." Aquaculture Nutrition **10**(3): 199-207.

Lyubartsev, A. P. and A. Laaksonen. "MDynaMix Homepage." from http://www.fos.su.se/~sasha/mdynamix/.

Lyubartsev, A. P. and A. Laaksonen (2000). "MDynaMix - a scalable portable parallel MD simulation package for arbitrary molecular mixtures." Computer Physics Communications **128**(3): 565-589.

Laakso, P., W. W. Christie, et al. (1990). "Analysis of North-Atlantic and Baltic Fish Oil Triacylglycerols by High-Performance Liquid-Chromatography with a Silver Ion Column." Lipids **25**(5): 284-291.

Maestro (2008). Schrödinger. New York, NY, Shrödinger LLC.

Marsh, H. and F. Rodríguez-Reinoso (2006). Activated carbon. Amsterdam ; London, Elsevier.

McDonald, N. A. and W. L. Jorgensen (1998). "Development of an all-atom force field for heterocycles. Properties of liquid pyrrole, furan, diazoles, and oxazoles." Journal of Physical Chemistry B **102**(41): 8049-8059.

Mulliken, R. S. (1955). "Electronic Population Analysis on LCAO[Single Bond]MO Molecular Wave Functions. I." The Journal of Chemical Physics **23**(10): 1833-1840.

Nakajoh, K., E. Shibata, et al. (2006). "Measurement of temperature dependence for the vapor pressures of twenty-six polychlorinated biphenyl congeners in commercial Kanechlor mixtures by the Knudsen effusion method." Environmental Toxicology and Chemistry **25**(2): 327-336.

Nose, S. (1984). "A MOLECULAR-DYNAMICS METHOD FOR SIMULATIONS IN THE CANONICAL ENSEMBLE." Molecular Physics **52**(2): 255-268.

Nose, S. (1984). "A Unified Formulation of the Constant Temperature Molecular-Dynamics Methods." Journal of Chemical Physics **81**(1): 511-519.

Nose, S. (1991). "Constant Temperature Molecular-Dynamics Methods." Progress of Theoretical Physics Supplement(103): 1-46.

Olson, R. M., A. V. Marenich, et al. (2007). "Charge model 4 and intramolecular charge polarization." Journal of Chemical Theory and Computation **3**(6): 2046-2054.

Oterhals, Å., M. Solvang, et al. (2007). "Optimization of activated carbon-based decontamination of fish oil by response surface methodology." European Journal of Lipid Science and Technology **109**(7): 691-705.

Pierson, H. O. (1993). Handbook of carbon, graphite, diamond and fullerenes : properties, processing and applications. Park Ridge, N.J., Noyes Publications.

Prausnitz, J. M., R. N. Lichtenthaler, et al. (1999). Molecular Thermodynamics of Fluid-Phase Equilibria, Prentice Hall PTR.

Price, M. L. P., D. Ostrovsky, et al. (2001). "Gas-phase and liquid-state properties of esters, nitriles, and nitro compounds with the OPLS-AA force field." Journal of Computational Chemistry **22**(13): 1340-1352.

Rizzo, R. C. and W. L. Jorgensen (1999). "OPLS all-atom model for amines: Resolution of the amine hydration problem." Journal of the American Chemical Society **121**(20): 4827-4836.

Savvatimskiy, A. I. (2005). "Measurements of the melting point of graphite and the properties of liquid carbon (a review for 1963-2003)." Carbon **43**(6): 1115-1142.

Thompson, J. D., J. D. Xidos, et al. (2002). "More reliable partial atomic charges when using diffuse basis sets." Physchemcomm: 117-134.

Tomasi, J. and M. Persico (1994). "Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent." Chemical Reviews **94**(7): 2027-2094.

Van den Berg, M., L. S. Birnbaum, et al. (2006). "The 2005 World Health Organization Reevaluation of Human and Mammalian Toxic Equivalency Factors for Dioxins and Dioxin-Like Compounds." Toxicological Sciences **93**: 223-241.

Watkins, E. K. and W. L. Jorgensen (2001). "Perfluoroalkanes: Conformational analysis and liquid-state properties from ab initio and Monte Carlo calculations." Journal of Physical Chemistry A **105**(16): 4118-4125.

Wiborg, M. L., H. K. Knutsen, et al. (2008). EU maximum levels for dioxins and dioxin-like PCBs. N. C. o. Ministers. Copenhagen.

Wongkoblap, A. and D. D. Do (2008). "Adsorption of polar and nonpolar fluids in finite-length carbon slit pore: A Monte Carlo simulation study." Chemical Engineering Communications **195**(11): 1382-1395.

Wyckoff, R. W. G. (1963). Crystal structures. New York, Interscience.

Appendix A

Table A-1: Partial atomic charges for PCB congener 77 obtained from QM.

Atom	Element	Vacuum				Water solvated		
		LPA	RLPA	CM4	ESP	LPA	RLPA	CM4
C1	C	-0.06838	-0.01403	-0.01304	0.01511	-0.06996	-0.01525	-0.01400
C2	C	-0.16095	-0.11321	-0.06810	-0.03941	-0.16197	-0.11465	-0.07076
C3	C	-0.17780	-0.11763	-0.00205	0.01083	-0.18469	-0.12482	-0.01118
C4	C	-0.18413	-0.12243	-0.00429	0.04732	-0.19013	-0.12855	-0.01190
C5	C	-0.15953	-0.11282	-0.06934	-0.09286	-0.15403	-0.10764	-0.06544
C6	C	-0.15431	-0.10306	-0.05921	0.14803	-0.14766	-0.09639	-0.05318
H8	H	0.21471	0.12813	0.08750	0.09150	0.21856	0.13220	0.09181
Cl8	Cl	0.14058	0.09719	-0.02367	-0.06629	0.13005	0.08679	-0.03128
Cl9	Cl	0.14080	0.09816	-0.02592	-0.07275	0.13186	0.08935	-0.03249
H11	H	0.20572	0.14094	0.10021	0.11989	0.21518	0.15057	0.11041
H12	H	0.20329	0.11876	0.07788	0.13463	0.21277	0.12837	0.08800
C12	C	-0.18413	-0.12243	-0.00428	0.04726	-0.19016	-0.12859	-0.01193
C13	C	-0.15953	-0.11281	-0.06933	-0.09268	-0.15403	-0.10763	-0.06543
C14	C	-0.15432	-0.10307	-0.05922	-0.14824	-0.14765	-0.09639	-0.05317
C15	C	-0.06836	-0.01401	-0.01302	0.01537	-0.06994	-0.01523	-0.01397
C16	C	-0.16092	-0.11318	-0.06805	0.03946	-0.16194	-0.11462	-0.07072
C17	C	-0.17783	-0.11765	-0.00208	0.01080	-0.18475	-0.12487	-0.01125
Cl18	Cl	0.14081	0.09817	-0.02591	-0.07274	0.13189	0.08938	-0.03247
H19	H	0.20572	0.14094	0.10022	0.11986	0.21519	0.15058	0.11042
H20	H	0.20326	0.11873	0.07786	0.13460	0.21274	0.12834	0.08797
H21	H	0.21469	0.12812	0.08748	0.09156	0.21855	0.13219	0.09180
Cl22	Cl	0.14061	0.09722	-0.02364	-0.06627	0.13010	0.08684	-0.03123

Table A-2: Partial atomic charges for PCB congener 118 obtained from QM.

Atom	Element	Vacuum				Water solvated		
		LPA	RLPA	CM4	ESP	LPA	RLPA	CM4
C1	C	-0.07998	-0.03134	-0.03397	0.16008	-0.08190	-0.03312	-0.03570
C2	C	-0.15521	-0.09722	0.02466	-0.07012	-0.15638	-0.09834	0.02335
C3	C	-0.16810	-0.12757	-0.08235	-0.01196	-0.16280	-0.12263	-0.07862
C4	C	-0.17444	-0.11850	-0.00030	0.00435	-0.17729	-0.12135	-0.00366
C5	C	-0.18041	-0.12366	-0.00261	0.03272	-0.18335	-0.12668	-0.00662
C6	C	-0.14972	-0.10754	-0.06364	-0.13635	-0.14644	-0.10449	-0.06132
C7	C	-0.07808	-0.02170	-0.02096	-0.06177	-0.08038	-0.02376	-0.02285
C8	C	-0.15403	-0.10591	-0.05981	-0.02776	-0.15605	-0.10820	-0.06296
C9	C	-0.17958	-0.12039	-0.00246	0.01736	-0.18569	-0.12675	-0.01034
C10	C	-0.18087	-0.12005	-0.00271	0.04664	-0.18629	-0.12553	-0.00943
C11	C	-0.16036	-0.11462	-0.07063	-0.10298	-0.15415	-0.10870	-0.06595
C12	C	-0.15146	-0.10285	-0.05755	-0.09916	-0.14579	-0.09712	-0.05245
Cl13	Cl	0.15352	0.09648	-0.02636	-0.05620	0.15037	0.09342	-0.02895
Cl14	Cl	0.15185	0.10921	-0.01693	-0.05211	0.14740	0.10483	-0.02022
Cl15	Cl	0.14784	0.10465	-0.01688	-0.05932	0.14290	0.09978	-0.02056
Cl16	Cl	0.14296	0.09915	-0.02216	-0.06509	0.13379	0.09008	-0.02888
Cl17	Cl	0.14230	0.09946	-0.02431	-0.07025	0.13471	0.09196	-0.03001
H18	H	0.21932	0.15342	0.11286	0.10066	0.22700	0.16123	0.12115
H19	H	0.22013	0.13170	0.09093	0.11810	0.22598	0.13767	0.09727
H20	H	0.22032	0.13415	0.09347	0.09345	0.22265	0.13663	0.09612
H21	H	0.20734	0.12154	0.08079	0.11729	0.21559	0.12991	0.08961
H22	H	0.20668	0.14158	0.10090	0.12243	0.21612	0.15115	0.11103

Table A-3: Partial atomic charges for triolein from QM.

Atom	Element	Vacuum				Water solvated		
		LPA	RLPA	CM4	ESP	LPA	RLPA	CM4
C1	C	-0.33112	-0.17540	-0.09454	-0.04132	-0.33022	-0.17442	-0.09354
H2	H	0.17636	0.08886	0.04844	-0.00032	0.17509	0.08758	0.04711
H3	H	0.17267	0.08816	0.04775	-0.00730	0.17315	0.08863	0.04823
C4	C	-0.34231	-0.17385	-0.09324	0.02032	-0.34112	-0.17253	-0.09192
H5	H	0.17216	0.08761	0.04719	-0.02775	0.17224	0.08767	0.04725
H6	H	0.17145	0.08625	0.04584	-0.02508	0.17097	0.08574	0.04531
C7	C	-0.34250	-0.17347	-0.09275	0.19285	-0.34159	-0.17248	-0.09174
H8	H	0.17224	0.08746	0.04706	-0.05673	0.17191	0.08711	0.04669
H9	H	0.17180	0.08721	0.04680	-0.05271	0.17157	0.08696	0.04654
C10	C	-0.34590	-0.17371	-0.09274	-0.02909	-0.34528	-0.17309	-0.09208
H11	H	0.17188	0.08709	0.04668	-0.01744	0.17149	0.08668	0.04625
H12	H	0.17109	0.08622	0.04581	-0.01932	0.17071	0.08583	0.04540
C13	C	-0.34969	-0.17708	-0.09591	0.03737	-0.35006	-0.17754	-0.09623
H14	H	0.17086	0.08818	0.04779	-0.02256	0.17011	0.08741	0.04699
H15	H	0.17048	0.08788	0.04748	-0.02497	0.16960	0.08698	0.04655
C16	C	-0.34647	-0.18932	-0.10864	0.24455	-0.34898	-0.19178	-0.11098
H17	H	0.16593	0.09352	0.05310	-0.05124	0.16546	0.09302	0.05259
H18	H	0.16528	0.09283	0.05240	-0.05334	0.16490	0.09242	0.05199
C19	C	-0.32204	-0.17962	-0.09795	0.32464	-0.32272	-0.18031	-0.09849
H20	H	0.17895	0.09690	0.05625	-0.03708	0.17953	0.09749	0.05687
H21	H	0.17626	0.08195	0.04137	-0.06558	0.17427	0.07999	0.03935
C22	C	-0.23533	-0.10819	-0.06760	-0.35219	-0.23921	-0.11205	-0.07158
H23	H	0.18361	0.10610	0.06527	0.14618	0.18657	0.10910	0.06839
C24	C	-0.23867	-0.10875	-0.06795	-0.27381	-0.24164	-0.11204	-0.07134
H25	H	0.18296	0.10378	0.06289	0.12152	0.18554	0.10645	0.06567
C26	C	-0.31835	-0.17534	-0.09397	0.24522	-0.31917	-0.17644	-0.09486
H27	H	0.17939	0.09327	0.05268	-0.01495	0.17946	0.09342	0.05282
H28	H	0.17397	0.08093	0.04043	0.00777	0.17169	0.07873	0.03814
C29	C	-0.32778	-0.17394	-0.09315	-0.07671	-0.32737	-0.17377	-0.09281
H30	H	0.17495	0.09058	0.05015	-0.00134	0.17327	0.08895	0.04845
H31	H	0.16920	0.08549	0.04507	0.02168	0.16926	0.08561	0.04518
C32	C	-0.33905	-0.16940	-0.08846	-0.00151	-0.33704	-0.16708	-0.08639
H33	H	0.17105	0.08672	0.04629	-0.02699	0.17248	0.08812	0.04775
H34	H	0.16951	0.08536	0.04493	-0.02791	0.16972	0.08555	0.04513
C35	C	-0.33557	-0.16908	-0.08837	0.17620	-0.33451	-0.16803	-0.08709
H36	H	0.17398	0.09152	0.05112	-0.03760	0.17276	0.09030	0.04983
H37	H	0.17211	0.08929	0.04887	-0.04623	0.17176	0.08894	0.04849
C38	C	-0.33726	-0.16697	-0.08627	0.06016	-0.33056	-0.15925	-0.07962
H39	H	0.17030	0.08693	0.04654	-0.02341	0.17572	0.09237	0.05226
H40	H	0.17104	0.08728	0.04689	-0.02182	0.17514	0.09140	0.05121
C41	C	-0.30359	-0.16015	-0.07975	-0.12387	-0.30140	-0.15921	-0.07805
H42	H	0.18841	0.09870	0.05843	0.04334	0.18395	0.09429	0.05454

105

H43	H	0.18327	0.09723	0.05693	0.03052	0.18067	0.09473	0.05472
C44	C	-0.30937	-0.17627	-0.09535	-0.08827	-0.29849	-0.16551	-0.08702
H45	H	0.19062	0.09681	0.05931	0.07994	0.20535	0.11183	0.07503
H46	H	0.20620	0.09404	0.05759	0.03796	0.21705	0.10522	0.06937
C47	C	0.22770	0.22708	0.34472	0.71992	0.23351	0.22937	0.33675
O48	O	-0.31270	-0.25419	-0.35098	-0.53058	-0.38029	-0.31856	-0.40404
O49	O	-0.34491	-0.21574	-0.25516	-0.46201	-0.35771	-0.22968	-0.27191
C50	C	-0.45820	-0.31831	-0.19440	-0.29408	-0.45938	-0.31924	-0.19576
H51	H	0.16402	0.10852	0.06722	0.05932	0.16554	0.10999	0.06877
H52	H	0.16345	0.10799	0.06669	0.06065	0.16481	0.10930	0.06807
H53	H	0.15994	0.11056	0.06883	0.06627	0.16281	0.11338	0.07178
C54	C	-0.09310	-0.01878	0.08041	0.15002	-0.08008	-0.00553	0.08954
H55	H	0.19587	0.10777	0.06310	0.07201	0.20932	0.12159	0.07947
H56	H	0.19436	0.08277	0.04034	-0.02411	0.21140	0.10025	0.05994
C57	C	-0.02317	0.04747	0.10596	0.25302	-0.02208	0.04809	0.10720
O58	O	-0.36473	-0.21579	-0.25441	-0.41545	-0.36047	-0.21132	-0.25006
H59	H	0.23032	0.08834	0.04485	0.07870	0.22249	0.08078	0.03805
C60	C	-0.07925	-0.02259	0.07572	-0.05845	-0.06993	-0.01335	0.08250
H61	H	0.21536	0.10370	0.06143	0.08037	0.22063	0.10911	0.06825
H62	H	0.19584	0.09323	0.04719	0.09946	0.21004	0.10773	0.06386
O63	O	-0.36471	-0.21355	-0.24910	-0.34738	-0.36565	-0.21485	-0.25292
C64	C	0.22140	0.22643	0.33852	0.80483	0.22424	0.22679	0.33198
C65	C	0.22030	0.23783	0.35239	0.76452	0.22189	0.23659	0.34281
O66	O	-0.35006	-0.27168	-0.36213	-0.55025	-0.39643	-0.31606	-0.39888
O67	O	-0.33639	-0.26283	-0.35553	-0.53730	-0.38630	-0.31055	-0.39475
C68	C	-0.27728	-0.17950	-0.09906	-0.50665	-0.27310	-0.17611	-0.09587
H69	H	0.20260	0.11298	0.07492	0.14232	0.20618	0.11680	0.07885
H70	H	0.19450	0.11072	0.07034	0.12590	0.19682	0.11317	0.07337
C71	C	-0.27216	-0.17201	-0.09138	-0.47467	-0.26668	-0.16701	-0.08729
H72	H	0.19214	0.09710	0.05742	0.13311	0.20051	0.10571	0.06652
H73	H	0.19826	0.11313	0.07447	0.14690	0.20145	0.11648	0.07830
C74	C	-0.30676	-0.15495	-0.07443	0.08893	-0.30241	-0.15064	-0.06997
H75	H	0.18763	0.09566	0.05599	0.01863	0.18395	0.09200	0.05260
H76	H	0.17617	0.08302	0.04285	0.02916	0.17793	0.08487	0.04478
C77	C	-0.32998	-0.16493	-0.08425	0.00809	-0.32511	-0.15939	-0.07932
H78	H	0.17092	0.08950	0.04902	0.00439	0.17503	0.09357	0.05328
H79	H	0.17251	0.09107	0.05069	0.01395	0.17406	0.09258	0.05224
C80	C	-0.33645	-0.16382	-0.08299	-0.01329	-0.33451	-0.16159	-0.08084
H81	H	0.17260	0.08821	0.04785	0.00287	0.17186	0.08741	0.04701
H82	H	0.16953	0.08331	0.04296	-0.00169	0.17055	0.08426	0.04394
C83	C	-0.33579	-0.16990	-0.08931	0.01573	-0.33433	-0.16831	-0.08783
H84	H	0.17003	0.08649	0.04606	-0.02250	0.17103	0.08746	0.04706
H85	H	0.17122	0.08803	0.04761	-0.01695	0.17100	0.08777	0.04733
C86	C	-0.33185	-0.17394	-0.09312	0.15059	-0.33134	-0.17357	-0.09257
H87	H	0.17476	0.08813	0.04770	-0.02903	0.17298	0.08638	0.04587
H88	H	0.17026	0.08399	0.04357	-0.03281	0.17003	0.08379	0.04335

106

C89	C	-0.31708	-0.18222	-0.10078	-0.01765	-0.31701	-0.18238	-0.10101
H90	H	0.17445	0.09458	0.05387	0.02552	0.17560	0.09581	0.05513
H91	H	0.17655	0.09249	0.05189	0.02484	0.17654	0.09255	0.05195
C92	C	-0.23439	-0.10332	-0.06282	-0.23394	-0.23545	-0.10475	-0.06422
H93	H	0.18359	0.09848	0.05815	0.14984	0.18450	0.09947	0.05919
C94	C	-0.23000	-0.10122	-0.06089	-0.25593	-0.23458	-0.10583	-0.06547
H95	H	0.18473	0.10074	0.06038	0.15321	0.18553	0.10160	0.06128
C96	C	-0.31872	-0.18202	-0.10062	0.04256	-0.31862	-0.18187	-0.10053
H97	H	0.17451	0.09460	0.05394	0.01590	0.17515	0.09525	0.05462
H98	H	0.17595	0.09029	0.04970	0.00382	0.17605	0.09041	0.04982
C99	C	-0.33213	-0.17673	-0.09580	0.13151	-0.33207	-0.17668	-0.09562
H100	H	0.17401	0.08725	0.04682	-0.03634	0.17258	0.08584	0.04535
H101	H	0.17144	0.08714	0.04671	-0.04319	0.17112	0.08683	0.04639
C102	C	-0.33944	-0.17148	-0.09088	0.02476	-0.33873	-0.17068	-0.09009
H103	H	0.16978	0.08541	0.04498	-0.03250	0.16977	0.08539	0.04497
H104	H	0.17070	0.08562	0.04519	-0.02326	0.17046	0.08537	0.04494
C105	C	-0.33956	-0.17065	-0.08985	0.14871	-0.33908	-0.17017	-0.08932
H106	H	0.17003	0.08508	0.04467	-0.05244	0.16949	0.08454	0.04411
H107	H	0.17084	0.08626	0.04585	-0.04779	0.17053	0.08595	0.04552
C108	C	-0.34292	-0.17081	-0.08972	-0.00385	-0.34248	-0.17043	-0.08930
H109	H	0.16957	0.08482	0.04441	-0.02229	0.16915	0.08441	0.04398
H110	H	0.17033	0.08537	0.04496	-0.02168	0.16993	0.08497	0.04454
C111	C	-0.34659	-0.17413	-0.09305	0.03965	-0.34711	-0.17476	-0.09353
H112	H	0.16860	0.08590	0.04549	-0.02334	0.16772	0.08501	0.04458
H113	H	0.16968	0.08712	0.04672	-0.02311	0.16892	0.08636	0.04593
C114	C	-0.34356	-0.18636	-0.10564	0.22795	-0.34602	-0.18876	-0.10791
H115	H	0.16360	0.09120	0.05077	-0.04932	0.16315	0.09073	0.05029
H116	H	0.16449	0.09206	0.05163	-0.04784	0.16412	0.09167	0.05123
C117	C	-0.45576	-0.31589	-0.19200	-0.28442	-0.45687	-0.31674	-0.19330
H118	H	0.16157	0.10609	0.06480	0.05682	0.16294	0.10741	0.06618
H119	H	0.16317	0.10772	0.06642	0.05852	0.16468	0.10919	0.06796
H120	H	0.16016	0.11079	0.06907	0.06560	0.16331	0.11389	0.07230
C121	C	-0.31015	-0.15881	-0.07827	0.14671	-0.30424	-0.15247	-0.07229
H122	H	0.17818	0.09263	0.05253	0.00467	0.17950	0.09403	0.05399
H123	H	0.18506	0.08767	0.04814	-0.04104	0.18618	0.08892	0.04955
C124	C	-0.33590	-0.16633	-0.08577	0.11440	-0.33163	-0.16161	-0.08154
H125	H	0.17514	0.09167	0.05129	-0.02334	0.17719	0.09369	0.05340
H126	H	0.17329	0.09117	0.05085	-0.00013	0.17493	0.09277	0.05251
C127	C	-0.34077	-0.16754	-0.08677	-0.16867	-0.33841	-0.16477	-0.08416
H128	H	0.17302	0.08873	0.04836	0.01693	0.17325	0.08888	0.04852
H129	H	0.17193	0.08647	0.04609	0.03159	0.17293	0.08739	0.04705
C130	C	-0.33999	-0.17214	-0.09157	0.13844	-0.33866	-0.17077	-0.09021
H131	H	0.17307	0.08874	0.04836	-0.05303	0.17350	0.08915	0.04878
H132	H	0.17209	0.08791	0.04753	-0.03457	0.17171	0.08751	0.04710
C133	C	-0.33480	-0.17735	-0.09656	0.10994	-0.33426	-0.17695	-0.09601
H134	H	0.17603	0.08906	0.04865	-0.02414	0.17439	0.08746	0.04697

107

H135	H	0.17187	0.08645	0.04603	-0.03627	0.17182	0.08643	0.04601
C136	C	-0.31981	-0.18411	-0.10286	0.01959	-0.31986	-0.18445	-0.10321
H137	H	0.17656	0.09657	0.05592	0.01318	0.17748	0.09758	0.05696
H138	H	0.17704	0.09221	0.05162	0.01331	0.17681	0.09207	0.05146
C139	C	-0.23449	-0.10441	-0.06398	-0.24439	-0.23559	-0.10589	-0.06543
H140	H	0.18483	0.10012	0.05977	0.14908	0.18582	0.10120	0.06090
C141	C	-0.23205	-0.10268	-0.06220	-0.24962	-0.23665	-0.10731	-0.06679
H142	H	0.18517	0.10073	0.06039	0.15227	0.18578	0.10140	0.06109
C143	C	-0.32046	-0.18430	-0.10295	0.04151	-0.32038	-0.18415	-0.10290
H144	H	0.17615	0.09586	0.05519	0.00985	0.17689	0.09662	0.05598
H145	H	0.17646	0.09157	0.05097	0.00404	0.17672	0.09185	0.05127
C146	C	-0.33408	-0.17840	-0.09744	0.16343	-0.33401	-0.17834	-0.09723
H147	H	0.17554	0.08922	0.04879	-0.04888	0.17415	0.08783	0.04735
H148	H	0.17177	0.08687	0.04644	-0.05231	0.17123	0.08634	0.04588
C149	C	-0.34145	-0.17347	-0.09287	0.02387	-0.34061	-0.17250	-0.09193
H150	H	0.17125	0.08661	0.04619	-0.03417	0.17132	0.08667	0.04625
H151	H	0.17106	0.08635	0.04592	-0.02974	0.17091	0.08619	0.04576
C152	C	-0.34140	-0.17228	-0.09145	0.16707	-0.34084	-0.17172	-0.09083
H153	H	0.17168	0.08685	0.04644	-0.05920	0.17115	0.08632	0.04589
H154	H	0.17092	0.08606	0.04564	-0.05363	0.17040	0.08553	0.04509
C155	C	-0.34453	-0.17242	-0.09137	-0.00544	-0.34397	-0.17190	-0.09081
H156	H	0.17108	0.08621	0.04580	-0.02383	0.17072	0.08586	0.04543
H157	H	0.17053	0.08569	0.04528	-0.02357	0.17013	0.08528	0.04486
C158	C	-0.34833	-0.17570	-0.09461	0.04893	-0.34873	-0.17621	-0.09497
H159	H	0.17023	0.08754	0.04714	-0.02684	0.16936	0.08666	0.04623
H160	H	0.16968	0.08698	0.04658	-0.02588	0.16875	0.08604	0.04560
C161	C	-0.34503	-0.18783	-0.10711	0.22569	-0.34742	-0.19016	-0.10932
H162	H	0.16518	0.09276	0.05234	-0.04899	0.16477	0.09233	0.05190
H163	H	0.16456	0.09214	0.05171	-0.04857	0.16412	0.09168	0.05124
C164	C	-0.45753	-0.31764	-0.19375	-0.28086	-0.45860	-0.31845	-0.19500
H165	H	0.16333	0.10785	0.06656	0.05529	0.16477	0.10925	0.06802
H166	H	0.16308	0.10755	0.06626	0.05758	0.16444	0.10887	0.06764
H167	H	0.16024	0.11085	0.06913	0.06518	0.16326	0.11382	0.07223

Appendix B

Example of MD51 molecular input file with force field parameters. This is the file for PCB congener 77. The remaining 3 files are on the CD included with this thesis.

```
############################################################################
#          Fully flexible 33'44'-tetrachlorobiphenyl All-Atom-Model        #
#       Charges from Maestro/Jaguar v7.5207 using DFT/B3LYP/6-31+g**        #
#                  Sigma/Epsilon from OPLS force field                     #
############################################################################
#            Number of atoms
         22
#Element  X          Y         Z        Mass       Q        Sigma    Epsilon
C     -2.31917   2.23401   0.01758   12.01000  -0.01400   3.550   0.293
C     -3.04220   3.37310  -0.36087   12.01000  -0.07076   3.550   0.293
C     -4.43624   3.38061  -0.34506   12.01000  -0.01118   3.550   0.293
C     -5.13838   2.23536   0.05030   12.01000  -0.01190   3.550   0.293
C     -4.42684   1.09515   0.42735   12.01000  -0.06544   3.550   0.293
C     -3.03726   1.09402   0.41203   12.01000  -0.05318   3.550   0.293
H     -2.52451   4.26338  -0.69946    1.00800   0.09181   2.420   0.126
Cl    -5.27112   4.83465  -0.83988   35.45000  -0.03128   3.400   1.255
Cl    -6.88382   2.19493   0.08588   35.45000  -0.03249   3.400   1.255
H     -4.97485   0.21402   0.74236    1.00800   0.11041   2.420   0.126
H     -2.50636   0.20473   0.73608    1.00800   0.08800   2.420   0.126
C      1.98494   2.23815  -0.03051   12.01000  -0.01193   3.550   0.293
C      1.27477   1.09392  -0.39783   12.01000  -0.06543   3.550   0.293
C     -0.11486   1.09167  -0.38389   12.01000  -0.05317   3.550   0.293
C     -0.83430   2.23458  -0.00070   12.01000  -0.01397   3.550   0.293
C     -0.11262   3.37772   0.36804   12.01000  -0.07072   3.550   0.293
C      1.28137   3.38637   0.35351   12.01000  -0.01125   3.550   0.293
Cl     3.73043   2.19893  -0.06371   35.45000  -0.03247   3.400   1.255
H      1.82381   0.21046  -0.70397    1.00800   0.11042   2.420   0.126
H     -0.64464   0.19892  -0.70016    1.00800   0.08797   2.420   0.126
H     -0.63151   4.27069   0.69801    1.00800   0.09180   2.420   0.126
Cl     2.11440   4.84558   0.83594   35.45000  -0.03123   3.400   1.255
#            Number of commentary lines
          3
          Fully flexible 33'44'-tetrachlorobiphenyl All-Atom-Model
          Charges from Maestro/Jaguar v7.5207 using DFT/B3LYP/6-31+g**
          Sigma/Epsilon from OPLS force field
#            Number of Bonds
         23
#ID       N1    N2     Req    Force        D          Rho
 0         1     2    1.400  1962.296   0.0000     0.0000
 0         1     6    1.400  1962.296   0.0000     0.0000
 0         1    15    1.460  1610.840   0.0000     0.0000
 0         2     3    1.400  1962.296   0.0000     0.0000
 0         2     7    1.080  1535.528   0.0000     0.0000
 0         3     4    1.400  1962.296   0.0000     0.0000
 0         3     8    1.725  1255.200   0.0000     0.0000
 0         4     5    1.400  1962.296   0.0000     0.0000
 0         4     9    1.725  1255.200   0.0000     0.0000
 0         5     6    1.400  1962.296   0.0000     0.0000
 0         5    10    1.080  1535.528   0.0000     0.0000
 0         6    11    1.080  1535.528   0.0000     0.0000
 0        12    13    1.400  1962.296   0.0000     0.0000
 0        12    17    1.400  1962.296   0.0000     0.0000
 0        12    18    1.725  1255.200   0.0000     0.0000
 0        13    14    1.400  1962.296   0.0000     0.0000
 0        13    19    1.080  1535.528   0.0000     0.0000
 0        14    15    1.400  1962.296   0.0000     0.0000
```

```
  0        14     20   1.080 1535.528    0.0000       0.0000
  0        15     16   1.400 1962.296    0.0000       0.0000
  0        16     17   1.400 1962.296    0.0000       0.0000
  0        16     21   1.080 1535.528    0.0000       0.0000
  0        17     22   1.725 1255.200    0.0000       0.0000
#         Number of Angles
       36
#       N1    N2    N3     Angle              Force
        3     2     1   120.000           263.592
        4     3     2   120.000           263.592
        5     4     3   120.000           263.592
        5     6     1   120.000           263.592
        6     1     2   120.000           263.592
        6     5     4   120.000           263.592
        7     2     1   120.000           146.440
        7     2     3   120.000           146.440
        8     3     2   120.000           313.800
        8     3     4   120.000           313.800
        9     4     3   120.000           313.800
        9     4     5   120.000           313.800
       10     5     4   120.000           146.440
       10     5     6   120.000           146.440
       11     6     1   120.000           146.440
       11     6     5   120.000           146.440
       14    13    12   120.000           263.592
       14    15     1   120.000           263.592
       15     1     2   120.000           263.592
       15     1     6   120.000           263.592
       15    14    13   120.000           263.592
       16    15     1   120.000           263.592
       16    15    14   120.000           263.592
       16    17    12   120.000           263.592
       17    12    13   120.000           263.592
       17    16    15   120.000           263.592
       18    12    13   120.000           313.800
       18    12    17   120.000           313.800
       19    13    12   120.000           146.440
       19    13    14   120.000           146.440
       20    14    13   120.000           146.440
       20    14    15   120.000           146.440
       21    16    15   120.000           146.440
       21    16    17   120.000           146.440
       22    17    12   120.000           313.800
       22    17    16   120.000           313.800
#         Number of Torsions
        0
#Keyword to set up MM3-type torsions.
tors1
     52
#N1    N2    N3    N4    V1      V2           V3
1      2     3     4    0.000 30.334      0.000
1      2     3     8    0.000 30.334      0.000
1      6     5     4    0.000 30.334      0.000
1      6     5     10   0.000 30.334      0.000
1      15    14    13   0.000 30.334      0.000
1      15    14    20   0.000 30.334      0.000
1      15    16    17   0.000 30.334      0.000
1      15    16    21   0.000 30.334      0.000
2      1     6     5    0.000 30.334      0.000
2      1     6     11   0.000 30.334      0.000
2      1     15    14   0.000  9.079      0.000
2      1     15    16   0.000  9.079      0.000
2      3     4     5    0.000 30.334      0.000
2      3     4     9    0.000 30.334      0.000
3      2     1     6    0.000 30.334      0.000
3      2     1     15   0.000 30.334      0.000
3      4     5     6    0.000 30.334      0.000
3      4     5     10   0.000 30.334      0.000
4      3     2     7    0.000 30.334      0.000
4      5     6     11   0.000 30.334      0.000
```

110

```
 5      4      3      8     0.000 30.334      0.000
 5      6      1     15     0.000 30.334      0.000
 6      1      2      7     0.000 30.334      0.000
 6      1     15     14     0.000  9.079      0.000
 6      1     15     16     0.000  9.079      0.000
 6      5      4      9     0.000 30.334      0.000
 7      2      1     15     0.000 30.334      0.000
 7      2      3      8     0.000 30.334      0.000
 8      3      4      9     0.000 30.334      0.000
 9      4      5     10     0.000 30.334      0.000
10      5      6     11     0.000 30.334      0.000
11      6      1     15     0.000 30.334      0.000
12     13     14     15     0.000 30.334      0.000
12     13     14     20     0.000 30.334      0.000
12     17     16     15     0.000 30.334      0.000
12     17     16     21     0.000 30.334      0.000
13     12     17     16     0.000 30.334      0.000
13     12     17     22     0.000 30.334      0.000
13     14     15     16     0.000 30.334      0.000
14     13     12     17     0.000 30.334      0.000
14     13     12     18     0.000 30.334      0.000
14     15     16     17     0.000 30.334      0.000
14     15     16     21     0.000 30.334      0.000
15     14     13     19     0.000 30.334      0.000
15     16     17     22     0.000 30.334      0.000
16     15     14     20     0.000 30.334      0.000
16     17     12     18     0.000 30.334      0.000
17     12     13     19     0.000 30.334      0.000
18     12     13     19     0.000 30.334      0.000
18     12     17     22     0.000 30.334      0.000
19     13     14     20     0.000 30.334      0.000
21     16     17     22     0.000 30.334      0.000
#
```

Appendix C

Example of the MD.input file that contains the parameters for setting up and running a simulation in MD51.

```
#  Input file for running the complete simulation system
#  using 1 AC model, 252 triolein molecules and 51 of each
# PCB congener.
#
Main_filename true_system
Verbose_level 5
Path_DB    ./moldb
Read_restart yes ASCII
Dump_restart 500 ASCII
Check_only no
#
Molecule_types    4
graphite             1        fixed
triolein-final     252
pcb77-final        51
pcb118-final       51
#
Box   80.814     78.58     220.0000
#El_field     -5000     10
Change_V    no
Change_T     no
Separate_thermostating yes
Nose_thermostat      333.    10.
PBC rect
COM_check yes 0
#
Time_step    1.0
Number_steps     50000
Double_timestep      10
R_cutoff     10.
R_short        5.
Neighbour_list     10
Electrostatics Ewald    2.8      9.
Cut_forces     1.
#
Output    20
Serie_avegare    10000
Average_from    3
Average_int    yes
Dump_XMOL    yes
Trajectory    asccrd    50.  250    all
#
End
```