The effect of calcite mineral on hydrate stability and CO_2 adsorption

Master of Science thesis in process technology

by

Øyvind Berge Sunnarvik

Department of Physics and Technology

University of Bergen, Norway

June 2011

What we know is a mere droplet;
what we don't know, an ocean.
 - Isaac Newton

Abstract

The background for conducting this study are the huge unused possibilities for storage of large quantities of the greenhouse gas carbon dioxide in old or depleted oil reservoirs or deep brackish aquifers. The method of capture and sequestration of CO_2 (CCS) emitted from large point polluters, such as coal power plants, has been presented as a way to alleviate the human effect on the environment. However, it is yet unknown how large quantities of CO_2 will affect complex geological formations such as reservoirs. The complexity stems from the presence of different minerals, varying pore distribution and the chemical makeup of the reservoir fluid. The impact that injection of large quantities of CO_2 in aforementioned systems poses must be further investigated and understood. This thesis is a first step in such an investigation. The calcite mineral is the most common mineral in the Earth's crust and a major component in most oil reservoirs that are being considered for storage and was considered a prime starting point. Temperature and pressure conditions of many reservoirs could initiate hydrate growth when CO_2 is introduced into a water-filled system. The inclusion of hydrate in the system was meant to study the minerals effect on the hydrate stability and growth.

In this work, molecular dynamic (MD) simulations have been applied to studies of the interactions and behavior of a model calcite mineral on a composite hydrate/water/CO_2 system. In order to study the effects of the calcite on the hydrate stability, the calcite was free to float in the water phase and the CO_2 was free to dissolve and interact with both calcite and hydrate. The effect of a calcite crystal on the stability of the hydrate was evaluated to estimate the behavior of hydrate and calcite mineral surrounded by relevant fluids in realistic reservoir conditions.

Three versions of the same model system were investigated. The difference between them was temperature and the force fields used to model water. This was done in order to investigate the differences between well-established water models, and how they handled CO_2 dissolved in water in the vicinity of a mineral surface.

The structuring effects around the incomplete hydrate cages were studied both visually and by radial distribution functions, with the conclusions fully supported by the density profiles. All three systems showed a large degree of water structuring that significantly differed depending on the environment. The adsorption of CO_2 in the incomplete half cages of hydrate was found to aid in the hydrate-like ordering of free water on both sides of the hydrate. The CO_2 adsorbed in greatest numbers when the width of the water layer was short, the distance required

in order to cross a broader water layer and the solubility of CO_2 in water reduced the number of adsorbed CO_2, yet the CO_2 had a strong preference for adsorption to hydrate compared to staying dissolved in water.

Water was found to form a layer of molecules around the calcite crystal. This layer has proven itself to be quite stable and have an ordering completely determined by the charge and geometry of the calcite. The potential energies also showed that calcite favored interactions with water rather than CO_2.

While the properties of the two water models were generally in good agreement with each other in all three systems, the main difference was observed near the calcite crystal. The SPC/E water model of at 313 K exhibited a much tighter packing of water molecules that resulted in attenuating the expected orientation of CO_2 molecules toward the calcite. The extent of the calcite's structuring effect on water was shown to vary depending on the cleaving plane. It was found that the $\{10\bar{1}4\}$ surface, predominantly found in nature, had exhibited the least structuring effect. This was due to the surface structure of this side consists of a layered structure where the alternating calcium atoms and carbonate ions. These components has a different structuring effect on water, and water molecules are oriented differently depending on which group they are closest to; the adsorbed water layer displayed a large degree of disorder. Due to this initial disorder the following water layers are not as widely affected by the initial water layer as seen on other surfaces. The size of the crystal together with inherently strong network of hydrogen bonds in water resulted in limited translational and rotational motion for the crystal.

Acknowledgement

I would like to take this opportunity to thank my supervisor, Professor Tatyana Kuznetsova, for valuable input and expertise. Her knowledge about molecular dynamics and the source code has proven an invaluable resource throughout my thesis. Another important person whom deserves both recognition and my gratitude is my co-supervisor, Professor Bjørn Kvamme, whose insight into the field of hydrates has been an inexhaustible source of inspiration and ideas. They have both been a great source of guidance and help in writing this thesis.

I would also like to thank Bjørnar Jensen for his encouragement and help throughout this master thesis. The intellectual, and less intellectual, discussions have been many and fruitful. The process behind this thesis would have been a lot harder without his ideas, tricks and his help.

Finally, my acknowledgement goes to my family for their understanding and support, and for being my link to a normal world when things have gotten hectic in my little corner of it.

University of Bergen
Department of Physics and Technology

Øyvind Berge Sunnarvik
1.6.2011

Nomenclature

Acronyms:

AMCSD	American mineralogist crystal database
CCS	Carbon capture and storage
DCD	Coordinate and trajectory file format
EPM2	Elementary physical version 2 model
F3C	Fully flexible charge model
I/O	Input/output interface
IR	Infrared radiation
LJ	Lennard Jones potential
MC	Monte Carlo
MD	Molecular Dynamics
MD43	Version of MDynaMix used in this thesis
MPI	Message passing interface
PBC	Periodic boundary conditions
PDB	Protein data bank
RDF	Radial distribution function
sI	Hydrate structure I
sII	Hydrate structure II
sH	Hydrate structure H
SPC/E	Simple point charge extended model
TCL	Tool command language
VMD	Visual molecular dynamics

Symbols

A	Helmholtz free energy
N	Number of atoms
V	Volume
E	Energy
T	Temperature
P	Pressure
M	Energy levels
Q	Partition function for the canonical ensemble
U	Internal energy
H	Enthalpy
G	Gibbs free energy
S	Entropy
Å	Angstrom, 10^{-10} meters
Q_m	A parameter for mass of motion in the Nosé-Hoover thermostat
M_w	Molecular weight
N_a	Avogadro's number
$\vec{F}$	Force vector
$\vec{E}_P$	Potential energy
$\vec{R}$	Position vector
$\vec{Q}$	Quaternion

q	Partial charge
k	Force constant
s	Scaling factor in the Nosé-Hoover thermostat
t	Time
m	Mass
p	Number of stable phases
c	Number of system components
g	Number of degrees of freedom
(x_x^i, x_y^i, x_z^i)	Position coordinates for particle i in phase space
(p_x^i, p_y^i, p_z^i)	Momentum coordinates for particle i in phase space
k_B	Boltzmann's constant
u_{AB}	Pair potential
r_{AB}	Distance between particles A and B
q_i	Scalar quantities in Quaternions, i=1,3
$\vec{a}$	Acceleration
$\vec{v}_i$	Velocity of particle i
$\vec{l}$	Lattice vector in Fourier space
erfc()	Complimentary error function in
$g(r)$	Pair correlation function
$\mathcal{H}$	Hamiltonian operator
α	Width of the Gaussian distribution
ϵ_0	Permittivity in vacuum
δ^-, δ^+	Partial atomic charge
μ	Chemical potential
ε	Distance parameter in LJ
σ	Energy parameter in LJ
ξ_i	Friction coefficient
ω	Angular velocity
φ, θ, ψ	Euler angles
τ	Torque
$\mathcal{O}()$	Truncation error
α, β, γ	Angels describing crystal periodicity
$\vec{a}_1, \vec{a}_2, \vec{a}_3$	Vectors describing crystal periodicity
$\rho(\vec{K})$	Charge density

Contents

1 Introduction

Carbon dioxide (CO_2) has been labeled a greenhouse gas due to its effect when present in the atmosphere where it absorbs infrared radiation (IR) emitted by the earth whilst still being transparent for incoming solar radiation, much like the sheets of glass in a greenhouse. The absorbed IR increases the molecular vibrations and motion of the CO_2 molecule, causing a warming of the atmosphere. The burning of fossil fuels and deforestation has led to an ever increasing concentration of CO_2 in the atmosphere since the industrial revolution. The idea that human made emissions might lead to a global warming was first introduced by Svante Arrhenius in 1896. The idea was dismissed by his fellow scientists, but in today's scientific community there is a great agreement that the extent of human prosperity is affecting the world we live in.

The two primary sources of CO_2 emission are from petroleum used for motor transport and from burning coal used for electricity. One promising way of reducing CO_2 emissions is carbon capture and sequestration (CCS) where the CO_2 is removed from the flue gas of large single point polluters, for instance coal power plants. The removed CO_2 is then injected into subsurface reservoirs. CO_2 injected into old oil reservoirs could help increase the production and lifetime of that reservoir, helping the economic situation of CCS. Studies have also shown that CO_2 can be injected into natural gas hydrate reservoirs to produce the natural gas without the dissolution of the hydrate structure (Kvamme et al. 2007). The amount of carbon bound in gas hydrates are conservatively estimated to be twice that found in all known fossil fuels on earth so the potential for a new energy source is vast, if it can be produced efficiently. What we need to know in order to understand and predict the effects of large scale sequestration is how CO_2 interacts with the surroundings in the reservoirs.

The likely injections or storage sites for large quantities of carbon dioxide are oil reservoirs or deep water aquifers. The oil reservoirs can either be depleted or carbon dioxide can be injected as a method of increasing oil production, while the aquifers are for storage of large quantities of captured carbon dioxide. What both have in common is that the available storage volume will consist of microscopic pores and connecting networks, this means that any constituent in the pores will have a large surface area in contact with the mineral in the reservoirs and the interactions between mineral and guest will have a great effect on the stability and efficiency of the storage capacity. In all storage sites there will be other components present

when the carbon dioxide is injected, one of the components will, nearly always, be water. In water aquifers this is evident, but in oil reservoirs the presence of large quantities of water is less explicit. In oil reservoirs that have been in production for a while, water has most likely been injected to maintain the production pressure of the well, and the injection well is therefore surrounded by large quantities of water. But water is also present in oil reservoirs where it has not been injected artificially. This water is a residue from when the oil reservoir was formed. Oil is formed from the remnants of living things that has died and collected millions of years ago. As time went by these remnants was covered by sediments that formed a large dome of impermeable rock, also called cap rock. As the temperature and pressure around the remnants increased they transformed into liquid hydrocarbons. Since hydrocarbons are lighter than water it rises until it reaches the cap rock. This dome traps the oil and keeps it from escaping. As the oil or gas rises it displaces the water already trapped in the dome. This process takes millions of years and before the water is displaced the rock has become permanently water wet and the intrusion of oil will not be able to displace all the water. There will always be residual water saturation in form of a thin water layer on the surface of the reservoir mineral.

Since limestone, which main component is calcium carbonate ($CaCO_3$), is ubiquitous, its precipitation and dissolution affects the conditions of many aqueous environments, including groundwater, rivers, lakes and the oceans. Furthermore, because of the reactivity of carbonate minerals being much greater than most other minerals, they have a great effect on their environment even if only present in relatively small amounts. The ability of CO_2 to react with the source rock when injected into reservoirs, so called mineral carbonation, is considered as a secure long term storage method for large quantities of CO_2. This is a process that occurs naturally as part of the geological carbon cycle, and ways of speeding up the process is being investigated (Xu et al. 2002; Oelkers et al. 2008). However the precipitation of carbonate minerals can unintentionally clog the pore network in the reservoir reducing permeability. The CO_2 will be injected either as a supercritical liquid or dissolved in water, either way the mineral environment will be aqueous and the investigation of the surface qualities of calcite has shown that it changes with the surrounding environment (wetted or non-wetted). In this thesis the calcite mineral will be suspended in a water phase and free to move; the water phase will be next to a CO_2 gas phase and CO_2 are expected to dissolve into water. A hydrate phase will also be present, adjacent to the water/mineral and CO_2 are expected to transverse the water phase and enter the half cavities of the hydrate. This hydrate and CO_2 is assumed to have a

structuring effect on the water extending the hydrate cage structure into the water, the calcite mineral will have a structuring effect on the water surrounding it. We will try to evaluate to what degree the water structure surrounding the mineral inhibits the hydrate structure.

2 Goals and choice of methods

The main goals of this M.Sc. project have been the following

- Achieve new insight into the mineral/fluid/hydrate interactions and implications for structures and adsorption pattern on surfaces of Calcite and hydrate.
- Seek new knowledge about possible destabilization of hydrate due to mineral - hydrate interactions.
- Investigate the possibilities for further hydrate growth in the composite hydrate/carbon dioxide/calcite/water system.

Choice of scientific method

Experimental facilities were out of reach and not relevant. The complexity of the goals would also impose extensive experimental challenges.

The use of atomistic simulation studies of model systems was therefore chosen.

The use of complete ab initio simulations was not an option due to the large size of the system. Ab initio methods could at best be used to characterize the individual models as even hybrid techniques with *ab initio* for closest surroundings and classical molecular dynamics for outside surroundings (Carr Parinelle type of simulations) would be extensive for the necessary system size and available computational resources in the project. That left Molecular Dynamics (MD) and Monte Carlo (MC) simulations. A disadvantage with MC is that it does not give any information about the kinetic energy in the system. It does not give us the trajectories and continuous movement of the molecules. Another disadvantage of MC is that an arbitrary change in configuration, which is the essence of MC, could be exceedingly large and subsequently result in skipping steps needed to describe the important dynamic development of the system. Therefore, critical changes in the system, or states of the system, could go undetected. On the other side small step sizes could trap the system in local energy minimum states and give rise to unrealistic progress of the system

It was decided that two systems with different water models were to be simulated in addition to one system at a lower temperature in order to investigate the effect of different water characterizations and to obtain some insight on the sensitivity of the results on the system conditions.

3 Calcite

Calcite is the most common of the carbonate minerals and the most stable polymorph of calcium carbonate. It is present in both geological and biological systems, and is often the primary constituent of shells of marine organisms. It is a common constituent of sedimentary rocks, limestone in particular, which is primarily formed of dead marine organism. In petroleum reservoirs calcite is present in both shallow seas; as limestone, and in deep waters; as chalk. Calcite can form mineral coating that acts as cement in sandstones and shale, filling fractures and decreasing permeability. And inversely the dissolution of calcite can greatly increase porosity and permeability.

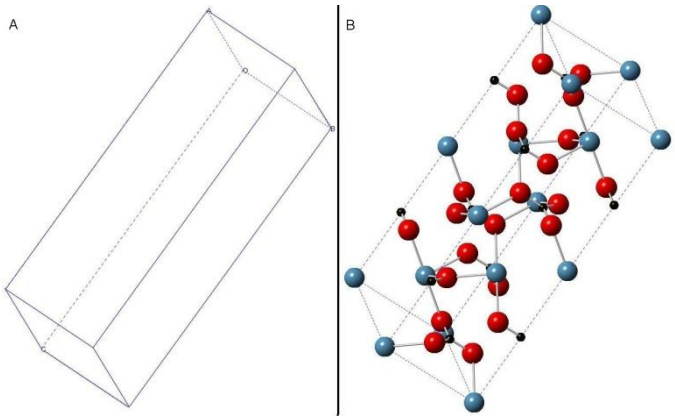

Figure 3-1.The rhombohedral unit cell of calcite crystal. B: The unit cell of calcite crystal. Red is oxygen, blue is calcite and black is carbon

Perfect calcite crystals are trigonal-rhombohedral in shape as shown in Figure 3-1 A, however, in nature the crystal is rarely perfect and over 300 crystal forms have been identified, in addition the occurrence of foreign metals will give the crystal a wide range of color.

When dealing with periodic systems, like a crystal, the periodicity can be described as a fundamental unit cell being replicated to form a system of the desired size. For a crystal the periodicity is repeated in 3 dimensions, the periodic cell can then be characterized by three vectors, often denoted $\vec{a}_1, \vec{a}_2$ and $\vec{a}_3$, spanning the physical space, with the length and angles between them defining the shape. The unit cell used in the thesis was found on the webpage of the American mineralogist crystal structure database (Downs and Hall-Wallace 2003) and was identified and described by Donald Graf (Graf 1961). The periodic cell has the dimensions $\vec{a}_1 = \vec{a}_2 = 4{,}991\text{Å}$ and $\vec{a}_3 = 17{,}062\text{Å}$ the angles are $\alpha = \beta = 90°, \gamma = 120°$. The program CrystalMaker® (Palmer)(2010) was used to create one large crystal, this large crystal was then cut to size using Visual Molecular Dynamics (VMD). The crystal was cut to expose

the $\{10\bar{1}4\}$ surface, this surface has been shown to be stable (Parker et al. 1993; Hwang et al. 2001) and is the surface commonly found exposed in nature.

The surface of calcite has been shown to have different properties when wet as opposed to dry. A molecular modeling study reported by Kvamme et.al (2007) has demonstrated how the excess surface energy of a wetted surface is significantly lower than that of a dry surface based on samplings from the model systems studied. In addition to wet and dry conditions, this study has shown that the presence of dissolved CO_2 raises the excess surface energy somewhat but it is still lower than for dry surface. This affects the chemical interactions and reactions on the surface of calcite.

CO_2 dissolves in water with a solubility that depends on pressure and temperature, and forms carbonic acid accordingly:

$$H_2O(liq) + CO_2(aq) \leftrightarrow H_2CO_3(aq) \quad (3\text{-}1)$$

Carbonic acid ionizes stepwise to produce bicarbonate and carbonate ions and affect the pH of the water.

$$H_2CO_3(aq) \leftrightarrow H^+(aq) + HCO_3^-(aq) \quad (3\text{-}2)$$

$$HCO_3^-(aq) \leftrightarrow H^+(aq) + CO_3^{2-}(aq) \quad (3\text{-}3)$$

When CO_2 in an aqueous phase comes in contact with calcite the following reactions take place:

$$CaCO_3(s) + H_2O(liq) + CO_2(aq) \leftrightarrow Ca^{2+} + 2 \cdot HCO_3^-(aq) \quad (3\text{-}4)$$

This is observed primarily when pH< 5-6, at higher pH than this precipitation is observed (Shiraki and Brantley 1995). Dissolution rates are proportional to the pH difference with respect to the equilibrium condition, and aqueous species are produced proportionally to the CO_2 pressure increase. (Espinoza and Santamarina 2010). In the system modeled in this thesis the addition of the supplementary ions $H^+(aq)$, $HCO_3^-(aq)$ and $CO_3^{2-}(aq)$ would further increase the accuracy of the model. But at this time the parameters needed for the implementation were not developed for all the ions. This is however suggested as part of further work.

4 Hydrate

Hydrate is an ice-like material that is formed under high pressure and low temperature conditions. Hydrate consists of water molecules connected through hydrogen bonds forming a cage-like structure around a guest molecule. Short range repulsive interactions between water and guest molecule and the hydrogen bonded network between the water molecules constituting the lattice ensures that the cage does not collapse. Short range non-polar attractions between water and guest molecules add to the stability of hydrate. Slightly polar guest molecules might lead to additional stabilization from electrostatic interactions between water and guest (example H_2S) (Kvamme and Forrisdahl 1993). The guest molecule is usually a smaller hydrocarbon from natural gas or liquid hydrocarbons and is trapped in the hydrate lattice. The guest molecule contributes to the stability of the hydrate; if it were removed the hydrate would collapse into regular ice or water if the hydrate where above the freezing point of water.

4.1 Hydrate structure

There are three main structures of hydrate. These are structure I (sI), structure II (sII) and structure H (sH) shown in Figure 4-1.

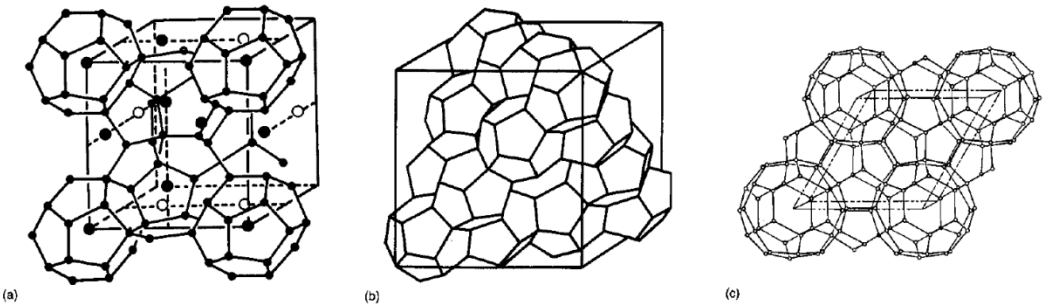

Figure 4-1. Unit cells of hydrate structure I (a), II (b) and H(c). (Sloan 1998)

The most common is sI and sII. sH is rare in an industrial context but are found in some natural hydrate sources, this is due to the complexity of formation involving at least three different molecules in three separate phases thus requiring mass transport across phase boundaries (Susilo et al. 2008). More recently discovered structures are structure T (Udachin et al. 2001) and the high pressure hydrates MH-II and MH-III (Loveday et al. 2001), but these are rarely observed. Some structural properties are listed in Table 4-I.

Table 4-I. Main geometrical properties of hydrate structures. (Sloan 1998)

Hydrate Crystal structure	I		II		H		
Cavity	Small	Large	Small	Large	Small	Medium	Large
Description	5^{12}	$5^{12}6^2$	5^{12}	$5^{12}6^4$	5^{12}	$4^3 5^6 6^3$	$5^{12}6^8$
Number of cavities/Unit cell	2	6	16	8	3	2	1
Average cavity radius, (Å)	3,95	4,33	3,91	4,73	3,91	4,06	5,71
Coordination number	20	24	20	28	20	20	36

sI is the simplest type of hydrate formed and the one used in this thesis. It consists of 46 water molecules forming two types of cavities:

- Dodecahedra (5^{12}). With twelve pentagonal faces.
- Tetradecahedron ($5^{12}6^2$). With twelve pentagonal and two hexagonal faces.

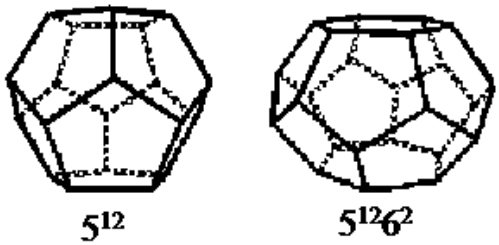

Figure 4-2. Structure I cavities

The unit cell of sI consists of 2 small and 6 large cages (Figure 4-1a) and encapsulates guest molecules with diameter between 4.2-6.0 Å. The 5^{12} and $5^{12}6^2$ hydrate structures are not stable on their own due to the expanded size relative to ice, the presence of the guest molecule in either the cell itself or in a large percentage of neighboring cavities are required.

There are four essential prerequisite that must be present if hydrate is to form and be stable. These are: a supply of hydrate forming guest molecules, access to a supply of water and conditions of low temperature and high pressure. Figure 4-3 shows a hydrate formation diagram in the pressure-temperature plane. In the hydrate free region the temperature and pressure conditions makes hydrates thermodynamically unstable and they will therefore not form. In the hydrate risk region hydrate can exist but they do not form and in the hydrate zone the temperature and pressure conditions facilitate hydrate growth. In the hydrate risk area the formation and presence of hydrate is hard to predict (Pickering et al.) and might be governed by the rate at which water and guest molecules can order themselves into a lattice structure

and by the rate at which the hydrate formers can be transported to the hydration formation zone.

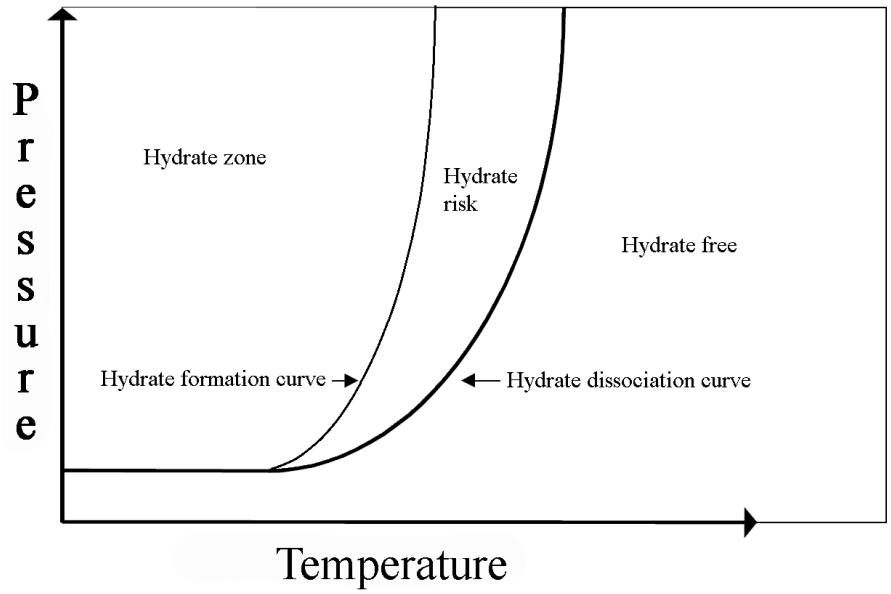

Figure 4-3. Hydrate formation diagram.

A phase is a distinct region of material that is chemically uniform and physically distinct. The Gibb's phase rule is based on the fundamental principles of thermodynamics. It gives us a theoretical foundation that allows us to describe the state of a system and makes it possible to estimate equilibrium conditions in a system in terms of intensive variable such as temperature and pressure. The Gibb's phase rule as given by equation (4-1) describes the possible number of degrees of freedom (g), i.e. the number of intensive variables, in terms of the maximum number of stable phases (p) and the number of system components (c).

$$g = c - p + 2 \qquad (4\text{-}1)$$

An example can be helpful in understanding how the rule woks. If we have a system of methane in a gas phase and water in a liquid phase and we want to consider the potential for hydrate formation. There are two components, water and methane, so c = 2 and three phases, liquid, gas and hydrate so c = 3 giving g = 1, with one intensive variable that can be defined in order for the system to be able to reach a unique solution. A unique analytic solution entails that the system could reach equilibrium (Sloan 1998). For the system used in this thesis there are three components: water, carbon dioxide and calcite so c = 3, there are one hydrate phase, two water phases, one gas phase and one solid calcite phase. The total number of phases are 5, such that the degree of freedom, g, is equal to 0. This means that the system simulated in this

thesis is a non-equilibrium system. Energy is allowed to flow and the interactions between molecules are allowed to form and shape the phases in the simulation. The number of particles and volume are set constant. The temperature is not allowed to reach an equilibrium temperature and is therefore a driving force for change in the system. The propulsion in the system is a function of minimum free energy as a function of conservation of mass and energy. Although the system might reach an energy minimum it will not be in equilibrium at this point and we do not know if it is a local or global minimum.

5 Computer simulation

Computer simulations can be seen as theoretical experiments. Simulations deal with models and not the "real thing" and are as such classified as theoretical. The verification process, in which the simulation set-up is established, a simulation is run and the resulting data are collected and analyzed are in the same fashion as an experimental scientist would do. In that sense we might consider Molecular Dynamics simulations as theoretical experiments which subsequently have to be related to experimental data for corresponding real systems in order to be used for interpretation of experiments and interpolations/extrapolations of experimental data.

Traditionally complex systems have been dealt with by simplifying the system until the subsystems are simple enough to be represented by solvable models which in turn have been verified by computer simulations. The progress in computer technology, especially when it comes to parallel computing, and algorithms have allowed for the complexity of the solvable system to increase and results from computer simulations have become directly comparable with experimental results. This allows for computer simulations to be used instead of experiments in regions where experimental data are hard to obtain, for instance in high pressure systems or where experiments become too expensive. Molecular simulation is also useful for interpreting observed data and functions as a link between theory and experiment. Adsorption phenomena are another area which is hard to investigate experimentally on an atomistic scale - at least when it comes to dynamic phenomena (combined thermodynamic and transport processes). Measuring contact angles for solid/liquid is a possible macroscopic technique for obtaining surface tension but interfacial free energy, on the other hand, is not possible to measure within present level of technology. And even though these interface properties are related, it is the interface free energy which is needed in studies of phase transitions.

There are two main approaches to performing computer simulations; the stochastic approach called Monte Carlo where a random number generator moves the process forward and a set of criteria is implemented to decide if the move is acceptable. And the deterministic approach called Molecular Dynamics (MD) where the forces acting on each molecule are calculated and used to move the system forward. The MD simulation technique has the advantage of giving information on the time properties of the system, i.e. how the system evolves with time. The MD simulation technique is used in this thesis and MD's main components along

with the mathematical concepts used to describe the physics behind the model will be outlined in the following sections.

5.1 Molecular dynamics

In a MD simulation one wishes to follow the time evolution of a systems particles through phase space. Phase space holds all the possible positions and momentum a system can have. A particle in phase space has 6 dimensions, 3 describing the position (x_x^i, x_y^i, x_z^i) and 3 describing the momentum (p_x^i, p_y^i, p_z^i), where the subscript is the dimensional direction and the superscript *i*, is the particle. A system of particles therefore has 6N-dimensions, where N is the total number of particles. An ensemble can be a selection of points in phase space limited by, for instance, all having the same energy. At any instance in time the system occupies one point in phase space, with time this point changes and creates a continuous curved line through phase space, this line maps the trajectory of the system. The governing equation in molecular dynamics, Newton's second law (5-1), relates the acceleration, $\vec{a}$, to the force, $\vec{F}$.

$$\vec{F} = m\vec{a} = m\frac{d^2\vec{R}}{dt^2} \tag{5-1}$$

In equation (5-1) m is the mass of the particle; $\vec{R}$ is the position vector and t is the time. In MD the force is evaluated from the potential energy $\vec{E}_P$ by equation (5-2). The potential energy is obtained from the force field package, which describes the formulas and necessary parameters for calculating the potential energy. Molecular simulation can also be derived from the Hamiltonian equations of motion using statistical mechanics. This approach is outlined in for instance "Understanding Molecular Simulation" by Frenkel and Smith (Frenkel and Smit 2002) or in the doctoral thesis by Ole Kr. Førrisdahl (Førrisdahl 2002).

$$\vec{F} = -\nabla\vec{E}_P \tag{5-2}$$

For an MD simulation the trajectory through phase space depends sensitively on the initial condition. This means that two trajectories that initially start very close will diverge exponentially as time progress. This divergence is caused by numerical errors that have two main sources. The discretization error that arises when a partial differential equation or integral is discretized into the corresponding difference equation, and a round-off error that arises when a computer rounds each number to some significant figure. An MD simulation will not give us the one and only possible outcome or path taken by a system, instead the simulation will

use statistical predictions to calculate the average development of a system where we know some of the initial conditions and the forces involved.

The trajectory obtained numerically does however need to be close to some true trajectory, this has only been proved for uninterestingly small systems. There exists considerable numerical evidence suggesting the existence of so-called shadow trajectories. Shadow trajectories are true (error free) trajectories of a dynamical system that remains close to a noisy numerical trajectory for a long time. If such a shadow exists the numerical trajectory can be considered reliable (Quinlan and Tremaine 1992). The main contributions to the precision of a MD simulation is predominantly limited by the accuracy of the bonded- and nonbonded- molecular potentials applied (Rowley and Pakkanen 1999), these will be discussed in chapters 5.8 and 5.9.

In this thesis the MDynaMix v4.3 (MD43) has been used. The program is a general purpose molecular simulation package that is designed for simulations of arbitrary mixtures of both rigid and flexible molecules, it does not include a specific force field package but can work with most conventional force fields (Lyubartsev and Laaksonen 2000). Professor Tatiana Kuznetsova has also added to this program, by included support for quaternion treatment for molecules and output in the protein database format (.PDB) and coordinate/velocity trajectory file (.DCD). The program is freely available under terms of GNU public license and is downloadable from their web page (MDynaMix).

5.2 Ensembles

Ensemble is a word coined by the American mathematical physicist JW Gibbs in 1878. Ensemble has, in statistical mechanics, two meanings. First it refers to a collection of all the possible microstates, or configurations of the arrangement of the system. Secondly it describes a set of constraints, for example constant temperature, volume and number of particles which is known as the canonical ensemble. Other thermodynamic quantities must be determined by ensemble averaging, which will be subjected to fluctuations.

The micro-canonical ensemble is the natural ensemble generated in an MD simulation. In the micro-canonical ensemble the number of particles (N), volume (V) and energy (E) is kept constant. In this ensemble the conservation of energy are assured, energy is free to shift from kinetic energy to potential energy but energy may not enter or leave the system. The results

from a micro-canonical simulation are however seldom directly comparable with experimental data.

This is because in a macroscopic system the number of microstates might be 10^{23}. Usually experiments cannot distinguish whether a system is in one microstate or another and in the micro canonical ensemble the entropy is connected to the microstate of the system through the well-known Boltzmann's law. Instead of looking at the microstates of a micro-canonical ensemble it can be useful to switch attention to the macro states of a canonical ensemble.

Therefore most experiments are not conducted under NVE conditions, instead of the energy being kept constant; it is the temperature (T) of the experimental system that is constant. An ensemble where NVT is constant is called a canonical ensemble. It is possible to run a simulation using the canonical ensemble but because the micro-canonical ensemble always will be the "native" ensemble any conversion must be handled with care so that the Boltzmann probability distribution (equation (5-3)) is conserved. There are also other ensembles like the Isothermal-Isobaric ensemble where N, pressure (P) and T are constant, and the grand canonical ensemble where V, E and the chemical potential (μ) are constant. Here only the NVT ensemble used in this thesis will be discussed.

In quantum mechanics the fundamental function that defines a system is called the wave function; the equivalent fundamental equation in statistical mechanics is called the partition function. For the canonical ensemble the partition function becomes:

$$Q_{N,V,T} = \sum_{i=1}^{M\to\infty} e^{-E_i(N,V)/k_B T} \qquad (5\text{-}3)$$

Where i is all the energy states having energy E_i, M is the energy levels, and $k_B$is the Boltzmann's constant. Equation (5-3) specify how particles are partitioned throughout the states effectively available to the system. In its original state it takes the form of an integral but in (5-3) it has been discretized for computational use. This partition function acts as a bridge between the quantum mechanical energy states of a macroscopic system and the thermodynamic properties of that system. The thermodynamic quantities are related to equation (5-3) as follows;

$$A = -k_B T \, ln \, Q \qquad (5\text{-}4)$$

$$U = k_B T^2 \left(\frac{\partial \ln \mathcal{Q}}{\partial T}\right)_{N,V} \tag{5-5}$$

$$H = U + pV \tag{5-6}$$

$$S = k_B \ln \mathcal{Q} + k_B T \left(\frac{\partial \ln \mathcal{Q}}{\partial T}\right)_{N,V} \tag{5-7}$$

$$G = H - TS \tag{5-8}$$

Where A is Helmholtz free energy, U is internal energy, H is enthalpy, S is entropy and G is Gibbs free energy.

When the volume and number of particles are specified and the temperature is held constant, the physical interpretation is that system of interest, system 1, is in thermal contact with a much larger system, system 2, in the case of a NVT ensemble called a heat bath. System 2 is isolated from the surrounding so no energy can leave. System 2 is presumed to be of infinite size so that the energy of the combined systems does not change. Meaning the energy of system 2 is so much greater than the energy of system 1 that any transfer of energy between systems 1 and 2 will not affect the energy of system 2. The entire system would then be a micro-canonical ensemble where V, N and E are constant. This will allow the energy of system 1 to fluctuate while the temperature, i.e. the systems average kinetic energy, remains constant. Equation (5-9) shows how the temperature of the system is related to kinetic energy; this example is for translational energy. The temperature control is enforced by means of a thermostat. The thermostat will add or remove energy from the boundaries of a MD system (system 1) trying to approximate the canonical ensemble. A number of thermostat methods are available; the simplest is a velocity scaling technique where the velocities of each particle are scaled up or down to keep the temperature constant. In practice this is undesirable since the adjustments of the velocities causes the trajectories to no longer be Newtonian.

$$\frac{1}{2} m_i \vec{v}_i^2 = \frac{3}{2} k_B T \tag{5-9}$$

Here m_i is the mass, $\vec{v}_i$ is the velocity and T is the temperature. A better option would be the Nose-Hoover thermostat described in section 5.7. This method does, if configured properly, generate NVT sampling and will correspond to the sampling involved in the canonical partition function if the temperature control is ergodic. The ergodicity hypothesis states that the ensemble average is equal to the time average, which means that if a MD system is allowed to evolve in time indefinitely it will eventually visit all possible positions in phase space with the

probability for each state as given by the canonical probability (exponent in (5-3)). The partition function (5-3) could then be found and the result would be ergodic. The ergodicity hypothesis makes the assumption that the average obtained by following a small number of particles over a long period of time is equivalent to averaging over a large number of particles over a short time (Jensen 2007).

5.3 PBC

Periodic boundary conditions (PBC) deals with the problem of surface effects on a bulk system and allows us to simulate large systems. Molecular simulations aim to provide information about a macroscopic system (of the order of 10^{23} molecules), but the computational restraints limits the system to a microscopic size of up to a few tens of thousands molecules. As the system size decreases the fraction of molecules in a surface position increases, for a cubical system of 1000 molecules almost half the molecules would be in a surface position (Frenkel and Smit 2002). In such a system the effects of surface conditions would dominate the system and cannot be neglected when trying to simulate bulk properties.

PBC mimics the effects of an infinite surrounding bulk by adding images of our primitive cell in an infinite periodic lattice around the primitive cell. The movement of a molecule in the primitive cell is exactly replicated in all other lattice cells. When a molecule drops out of the primitive cell the same molecule will enter on the opposite side, this maintains the number density in the primitive cell. This is illustrated in Figure 5-1, there are no walls at the boundaries and the system has no surface, the central box only forms a coordinate system for locating the N molecules. PBC preserve mass, particle number, total energy and linear momentum (Cramer 2004).

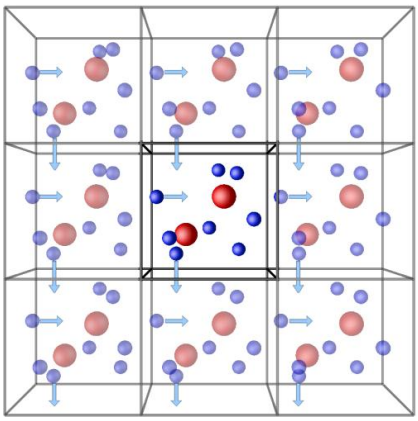

Figure 5-1. Periodic boundary condition. From (ISAACS)

To calculate the force on a given particle, one needs to include its interactions with the rest of N-1other particles in the simulation. Because of the PBC, we must also include all the particle images in the surrounding boxes. This is an infinite number of terms and in practice one would need to restrict the summation. The minimum image convention can be applied in the case of short-range forces. In the minimum image convention, the particle of interest lies at the center of a cell with the same size and shape as the primitive cell. The particle then interacts with the other particles in this box only as shown in Figure 5-2. A further approximation to reduce calculations is to implement a spherical cut-off, assuming that only the closest neighbors contribute to the force acting on a particle.

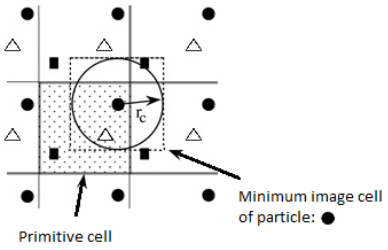

Figure 5-2 PBC with primitive cell, Minimum image cell and cut-off (r_c)

The minimum size of a periodic cell is determined by the largest cut-off length employed in the simulations. If the cut-off distances overlap the particle would "sense" its image and a nonexistent symmetry would be imposed on the system. It is often desired that the size be at least twice that of the minimum requirement so that one molecule do not sense another molecule twice, the one in the primitive cell and an image.

The cut off distance is the maximum distance that a force is evaluated, after this distance the force is considered negligible, particularly from a gradient standpoint. To prevent there from being a discontinuity in energy a smoothing function is often employed. This smoothing function gradually reduces the force to zero. The cut off method is only suited for short range forces, to handle long range forces the cut off distance would have to be quite large to avoid structural artifacts (Cramer 2004). The long range forces are instead handled by Ewald summation which will be discussed in chapter 5.9.2.

5.4 Verlet algorithm

The Verlet algorithm (Verlet 1967) is a method of integrating the equations of motion, allowing us to progress the simulated system in time. It is a compact and easy to use algorithm at the same time as it yields good results when compared with more complex methods. Like

Newton's equations of motion the Verlet algorithm is time reversible, meaning that if one were to reverse the momenta of all particles the system will backtrack its own trajectory. The Verlet algorithm has good energy conservation, although the short term (a few time steps) is only fair, the more important long time energy conservation shows little drift (Frenkel and Smit 2002)

The algorithm can be derived using Taylor expansion of the coordinate $\vec{x}_i$ of a particle i at time t, t+Δt and t-Δt:

$$\vec{x}_i(t+\Delta t) = \vec{x}_i(t) + \left(\frac{d\vec{x}_i}{dt}\right)\Delta t + \frac{1}{2}\left(\frac{d^2\vec{x}_i}{dt^2}\right)(\Delta t)^2 + \cdots \qquad (5\text{-}10)$$

$$\vec{x}_i(t-\Delta t) = \vec{x}_i(t) - \left(\frac{d\vec{x}_i}{dt}\right)\Delta t + \frac{1}{2}\left(\frac{d^2\vec{x}_i}{dt^2}\right)(\Delta t)^2 - \cdots \qquad (5\text{-}11)$$

Combining equation (5-10) and (5-11) and assuming that terms higher than second order is negligible we get the new position at time t+Δt expressed in terms of the position at previous time steps.

$$\vec{x}_i(t+\Delta t) = 2\vec{x}_i(t) - \vec{x}_i(t-\Delta t) + \left(\frac{d^2\vec{x}_i}{dt^2}\right)(\Delta t)^2 \qquad (5\text{-}12)$$

Equation (5-12) is known as the Verlet algorithm. We can see that there arises a round-off error when a small term is added to a difference of two large terms to generate the trajectory. There is also the problem that equation (5-12) does not explicitly contain the velocities $\vec{v}_l(t)$; they are instead obtained from the finite difference formula (5-13) in an additional step.

$$\vec{v}_l(t) = \frac{\vec{x}_i(t+\Delta t) - \vec{x}_i(t-\Delta t)}{2\Delta t} \qquad (5\text{-}13)$$

Equation (5-13) requires the position at both the previous and next time step. This handling of the velocities is rather awkward and while equation (5-12) is correct except for truncation errors in order of $\mathcal{O}(\Delta t^4)$ equation (5-13) is subject to truncation errors in order of $\mathcal{O}(\Delta t^2)$ decreasing the overall accuracy of the algorithm. Several schemes have been suggested to improve on the basic Verlet algorithm, one of these is the velocity Verlet algorithm.

5.4.1 The velocity Verlet algorithm

The velocity Verlet algorithm is a Verlet equivalent integration algorithm meaning that it yields trajectories similar to equation (5-12). It stores position, velocity and acceleration at the same time step and reduces the round-off error from equation (5-12), therefore limiting loss of precision on a computer.

$$\vec{x}_i(t+\Delta t) = \vec{x}_i(t) - \Delta t\, \vec{v}_i(t) + \frac{1}{2}(\Delta t)^2 \vec{a}_i \tag{5-14}$$

$$\vec{v}_i(t+\Delta t) = \vec{v}_i(t) + \frac{1}{2}\Delta t\, [\vec{a}_i(t) + \vec{a}_i(t+\Delta t)] \tag{5-15}$$

Where $\vec{a}_i(t)$ is the acceleration of molecule i at time t. The acceleration is related to the force through Newton's second law (5-1). Solving the velocity Verlet algorithm is done in two stages with a force calculation in between. First the new position at time $(t+\Delta t)$ is calculated using equation (5-14), and the velocities at $(t+\frac{1}{2}\Delta t)$ is calculated using equation (5-16)

$$\vec{v}_i\left(t+\frac{1}{2}\Delta t\right) = \vec{v}_i(t) + \frac{1}{2}\Delta t\, \vec{a}_i(t) \tag{5-16}$$

The forces and acceleration at time $(t+\Delta t)$ are then computed, and the velocity move is completed by equation (5-17)

$$\vec{v}_i(t+\Delta t) = \vec{v}_i\left(t+\frac{1}{2}\Delta t\right) + \frac{1}{2}\Delta t\, \vec{a}_i(t+\Delta t) \tag{5-17}$$

5.5 Constraint dynamics

In polyatomic molecules it becomes necessary to consider the stretching of interatomic bonds, bending motion and twisting motion. The twisting motion is especially important in long–chain molecules, and this effect has a low frequency in energy. Bond stretching on the other hand has a high frequency, but does not affect most systems in any significant way. At the same time when deciding the length of the time step, an important control parameter in MD, the maximum time step possible, is decided by the highest frequency in the system. With a time step of the order of femtoseconds, one nanosecond of simulation time requires one million steps. To achieve a simulation time of one microsecond, the simulation would have to run for one billion time steps. If the fastest vibration in the system is insignificant to the progres-

sion of the system, the ability to freeze these bond vibrations will yield a great reduction of number of time steps due to the ability to run the simulation with larger time steps. At normal temperatures the amplitude of vibration is small compared with molecular dimensions so this assumption is reasonable (Allen and Tildesley 1989). Assuming total rigidity on the other hand may be unrealistic. Constraint dynamics is a technique that allows selected degrees of freedom to be constrained while others are free to evolve under the influence of the forces acting upon them. The most commonly used method for applying constraints in molecular simulations is the SHAKE algorithm (Ryckaert et al. 1977).

In the SHAKE algorithm the atoms are first allowed to move under the influence of the forces, and subsequently the constraints are enforced. Each constraint is considered in turn and an iterative loop is employed to satisfy the constraint to within some tolerance. This tolerance is necessary as one constraint might be in conflict with another when molecules move around each other. The SHAKE algorithm is easily applied to the Verlet algorithm where the velocity term does not appear explicitly. If the same principles are applied to a velocity Verlet algorithm, the constriction algorithm is called RATTLE (Andersen 1983).

5.6 Quaternions

When looking at a molecules movement through space it is common to divide the motion into translation of the center off mass, and rotation about the center of mass. The former is easily handled with Newton's second law (5-1), where the force is the vector sum of all forces acting on the molecules center of mass. The rotational motion is governed by a twisting force, or torque (τ), around the center of mass. The torque acts in the rotational equation of motion in the same way that the force act in the translation equations, but due to the nature of orientation space the equations of orientation space are more complicated (Allen and Tildesley 1989).

The orientation of a body is usually specified as the relation between two systems with different reference axes, one fixed in space and one fixed with respect to the body rotating. To describe rotational motion a minimum of three angels are needed, these are usually some form of the Euler angles φ, θ and ψ.

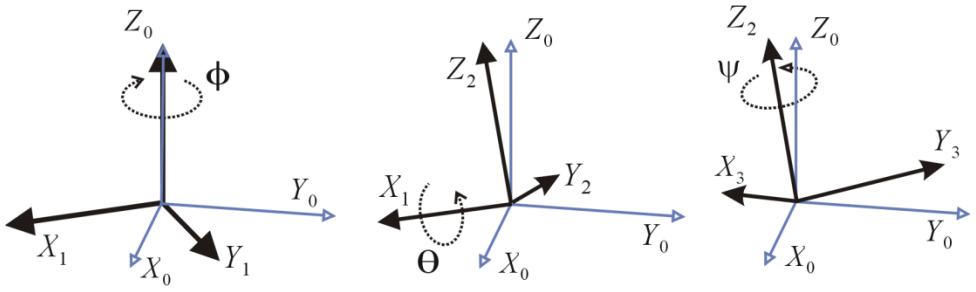

Figure 5-3. The Euler angels of rotation. Body-fixed axis in black and system-fixed in blue.

A rotation in Euler angels is decomposed into three different rotations. First, we rotate by φ about the system z-axis. This changes the x and y-axes to x_1 and y_1. Then we rotate by θ about the new x-axis, x_1, which changes the y and z-axes. Finally we rotate by ψ about the new z-axis.

There are some problems when computing Euler angels. First, rotation about each axis is stored in matrices; these matrices are then added together using matrix multiplication meaning that the order of multiplication is very important. If the order of rotation is Z, Y, X and the rotation in Y is 90 degrees the rotation in the X-axis is mapped onto the Y-axis, i.e. both the Y and X-axes rotate in the same plane compared to the system-axis. This is called gimbal lock and can be seen in the translational equation of motion for φ (4-1).

$$\dot{\varphi} = -\omega_x^s \frac{\sin\varphi\cos\theta}{\sin\theta} + \omega_y^s \frac{\sin\varphi\cos\theta}{\sin\theta} + \omega_z^s \qquad (5\text{-}18)$$

Where ω_x^s is the angular velocity in x direction viewed from the system-fixed axis; $\dot{\varphi}$ is the rate of change in the Euler angle. If θ approaches 0 or π, the presence of sin θ will make equation (4-1) diverge, the mathematical equivalent of a gimbal lock. The Euler approach also includes trigonometric functions that are computationally expensive to evaluate.

A better result is achieved by using quaternions. A quaternion $\vec{Q}$ is a set of four generalized parameters and have a well-behaved equations of motion. The four scalar quantities in a quaternion are $\vec{Q} = (q_0, q_1, q_2, q_3)$, where the last three can be thought off as a vector in imaginary space, i.e. $q_1^2 = q_2^2 = q_3^2 = -1$. q_0 is the real part, analog to the real part in representing complex numbers with a sum of real and imaginary parts.

Quaternions have been implemented in the MD version used in this thesis by Professor Tatyana Kuznetsova.

5.7 Nosé-Hover thermostat

The temperature in a MD simulation is related to kinetic energy according to equation (5-9). The simplest way of controlling the temperature is to scale the velocities of the particles, in effect introducing a non-Newtonian acceleration or deceleration to keep the kinetic energy constant, as mentioned in chapter 5.2. This method has its uses when the system is far from equilibrium as it is fast and computationally inexpensive, and is commonly used in the initial stages of a simulation.

A different approach is to endeavor to reproduce the canonical phase-space distribution; this will permit the kinetic energy to fluctuate. The implementation of such a technique would allow the results to be correlated with that of many-body simulations with Gibb's and Jaynes's statistical mechanics (Hoover 1985). One such method is the Nosé-Hoover thermostat first introduced by Shuichi Nosé in 1984 (Nosè 1984), and further developed by William Hoover in 1985 (Hoover 1985). In the Nosé-Hoover thermostat a heat bath is considered an integral part of the system. The bath is free to exchange energy with the system but not particles or volume. The Nosé-Hoover thermostat is the function that regulates the kinetic energy in order to keep the temperature in the main system constant. The Hamiltonian ($\mathcal{H}$) is the total energy of the system, to model the combined system Nosé implemented an extended Hamiltonian;

$$\mathcal{H}_N = \sum_i^N \frac{\vec{p}_i{}^2}{2m_i} + \mathrm{E}_P(\vec{q}_i) + \frac{Q_m \xi^2}{2} + g k_B T \ln s \qquad (5\text{-}19)$$

Here m_i is the mass, $\vec{p}_i^2$ is the momentum, $\mathrm{E}_P(\vec{q}_i)$ is a potential energy of a system, ξ_i is a friction coefficient, Q_m is a parameters that acts as a mass for the motion of s and the energy of the heath bath, 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 magnitude of Q_m determines the coupling between the system and the reservoir. In the original thermostat suggested by Nosé the time was scaled as dt`=dt/s where t` is real time and t is virtual time, the virtual variables are introduced for control of the temperature. The system is here sampled at virtual time steps leading to an uneven sampling in real time. This uneven sampling is impractical for the investigation of dynamical properties of the system. Hoover instead suggested that if a new variable ξ (4-1) is introduced the Nosé equations of motion could be simplified.

$$\xi = \left(\frac{1}{s}\right)\frac{ds}{dt`} = s\frac{p_s`}{Q_m} \tag{5-20}$$

The basic equations of motion for the Nosé-Hoover thermostat are (Nosè 1991):

$$\frac{d\vec{q}_i}{dt} = \frac{\vec{p}_i}{m_i} \tag{5-21}$$

$$\frac{d\vec{p}_i}{dt} = \frac{\partial \mathrm{E}_P}{\partial \vec{q}_i} - \xi\vec{p}_i \tag{5-22}$$

$$\frac{d\vec{\xi}_i}{dt} = \frac{\left(\sum_i^N \frac{\vec{p}_i^{\,2}}{m_i} - gk_BT\right)}{Q_m} \tag{5-23}$$

$$\frac{d\ln s}{dt} = \xi \tag{5-24}$$

Equation (5-21) and (5-22) gives the change of position and momentum respectively, these are effectively the same as normal but modified with a frictional force ξ. A positive ξ slows the particle down whereas there is acceleration in the negative ξ region. In equation (5-23) we see how the constant Q_m determines how quickly the system reacts to changes in temperature, i.e. how quickly energy flows across from the heat bath to the system. If Q_m is too large energy will not flow across the boundary, set Q_m too small and the system will overcorrect and fluctuate wildly, in both cases the system will fail to reach equilibrium. The sum is the total kinetic energy for the system, while gk_BT is the average kinetic energy. Here T is the temperature set for the system therefore the average kinetic energy is the desired energy for the system. From equation (5-23) we see that the thermostat only scales the temperature of the system when the momentaneous temperature of the system deviates from the desired temperature, if there is no deviation equation (5-19) reduces to the normal Hamiltonian.

Due to the fact that the introduction of real and virtual variables is not canonical, equation (5-19) is no longer a proper Hamiltonian. It can however be shown (Nosè 1991) that when equation (5-24) is added to equations (5-21) to (5-23), which in themselves form a closed system of equations, that equation (5-19) is conserved as a pseudo-Hamiltonian for the Nosé-Hoover thermostat. And given proper values of the thermostat masses it might get close to canonical for very simple systems but impossible for systems with many degrees of freedom since the thermostat masses will be given "universal masses" applied to all different species.

5.8 Bonded interaction

There are intra-molecular interactions in every molecule. The stability of these interactions is affected by both the movement in the individual atoms and by surrounding molecules. When the stability of these interactions fluctuates so does the magnitude of the force generated by these, changing both the shape of the atom and how the atom affects its surroundings. Intra-molecular interactions include interactions such as bond stretching, angle bending, torsion and cross interactions. Cross interactions are correcting terms used where for instance the sum of energies from bond stretching and angle bending does not equal the true change in energy from these interactions. Cross interactions therefore comes as an addition or reduction to the total interaction energies.

5.8.1 Bond stretching

Bond stretching is the oscillatory motion of two molecules sharing a covalent bond. This motion can be modeled as a spring-and-balls model as in Figure 5-4 below where the bond itself is modeled as a spring. When the molecules, or balls, are moved past their equilibrium point to an extended position the spring contracts back to the equilibrium position. The momentum then pulls the balls closer before the spring pushes the balls back to equilibrium.

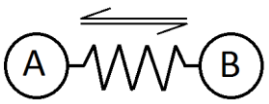

Figure 5-4. Bond stretching

The equilibrium bond length is the average bond length that the bond would have in a diatomic molecule, corresponding to the minimum energy structure of the molecule. In a polyatomic molecule the terms in the force field energy will produce a bond length slightly longer than that of the diatomic molecule. In MD43 Hook's law (5-25) is used to model the bond length.

$$\vec{F}_{str}(\vec{R}) = -k_{str}(\vec{R} - \vec{R}_{eq}) \qquad (5\text{-}25)$$

Here $\vec{F}_{str}$ is the force as a function of the length of the spring $\vec{R}$, k_{str} is the force constant. The negative sign indicates that the force is opposite the direction of displacement, restoring the system to its equilibrium state. $(\vec{R} - \vec{R}_{eq})$ is the distance the spring has been displaced. The potential energy is given by (5-26).This type of bond is called harmonic bond due to the harmonic oscillation of the solution of these equations.

$$U_{str}(\vec{R}) = \frac{1}{2}k_{str}(\vec{R} - \vec{R}_{eq})^2 \quad (5\text{-}26)$$

5.8.2 Angle bending

In a molecule consisting of three or more atoms there will be angle vibration. Angle bending is modeled in a similar way as bond stretching. An angle is defined by the three atoms and a spring is visualized between the "legs" as shown in Figure 5-5.

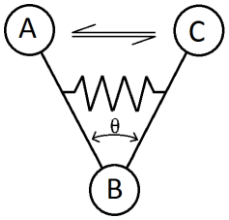

Figure 5-5. Angle bending

Hook's law (5-27) is used to model the stretching and contracting of the spring. If the angle θ is changed, the length of the spring changes and a force working in the opposite direction will occur.

$$U_{ang}(\theta) = \frac{1}{2}k_{ang}(\theta - \theta_{eq})^2 \quad (5\text{-}27)$$

The harmonic potential for a covalent angle is found using equation (5-27). k_{ang} is the force constant, θ is the angle made by A-B-C and the equilibrium angle is θ_{eq}.

5.8.3 Torsion

When four or more molecules, A-B-C-D, are bonded together where A-B, B-C and C-D are bonded together there will be a torsion force around the B-C bond due to the torque applied by the A-B and C-D arms.

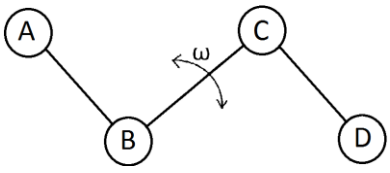

Figure 5-6. Torsion bending

Torsion is a three dimensional phenomena, in Figure 5-6 it is drawn in a two dimensional plane. To generate torsion around angle ω, molecule A and D would have to rise out of or into the plane, torsion would also appear if the molecule spins around the B-C bond and molecule A and D have different velocities.

The torsion energy is fundamentally different from the stretching and bending energies in three aspects (Jensen 2007):

a) In addition to the torsion energy, the energy from the non-bonded forces contributes against the torsion effect. The torsion parameters are therefore intimately linked to the non-bonded parameters.
b) If the bond is rotated 360˚ the torsion energy should return to the same value. The torsion energy function must therefore be periodic.
c) The energy cost of rotating around a bond is often low, therefore large deviations from the minimum energy structure may occur and a Taylor expansion is inadequate.

A Fourier series like (5-28) would incorporate the periodicity and give the desired accuracy.

$$E_{tors}(\omega) = \sum_{n=1} V_n \cos(n\omega) \qquad (5\text{-}28)$$

Where n describe the periodicity of the rotation, V_n is a constant describing the combined forces from a). Molecules composed of atoms with a maximum of four valence electrons (most organic molecules) are found to have three energy minima corresponding to three terms in the Fourier series (5-28) (n=1,2,3). It is customary to shift the zero point of the potential energy by adding a factor of one to each term (Jensen 2007). Equation (5-29) is the expression used for torsional energy in MD43.

$$E_{tors}(\omega) = \frac{1}{2}V_1[1 + \cos(\omega)] + \frac{1}{2}V_2[1 - \cos(2\omega)] + \frac{1}{2}V_3[1 + \cos(3\omega)] \qquad (5\text{-}29)$$

5.9 Non-bonded interaction

Inter-molecular forces are non-bonded forces that occur between non-bonded molecules. Inter-molecular forces are responsible for the non-ideal behavior and bulk properties in matter and they are generally weaker than the intra-molecular forces. The inter-molecular forces consist of the van der Waals forces and electrostatic forces. The van der Waals forces are short range, while the electrostatic is long range. The range is decided by how fast the force diminishes when moving away from the originating point. The computational evaluation of non-bonded energy is by far the most time-consuming process. This is because the bonded contributions grow linearly with the system size, while the non-bonded grow as the square of the

system size as the majority of atom pairs are not bonded (Jensen 2007). Through statistical mechanics it is straightforward to argue that inter-molecular interactions between two atoms will depend on the existence of surrounding atoms since they will affect the electrons of the two atoms in consideration. As such in the discussion below pairwise interactions are considered as "effective" interactions between pairs of atoms.

5.9.1 Van der Waals forces

The van der Waals forces are short range inter-molecular forces. The van der Waals force is comprised of the dipole-dipole, dipole-induced dipole and dispersion forces. Dipole-dipole forces are a result of the interaction between permanently polarized molecules, they are electrostatic of origin and the larger the dipole moment the larger the force. Dipole-dipole forces are also known as the Keesom force. The dipole-induced dipole, or Debye, force is the product of a permanent dipole and a nonpolar molecule. The permanent dipole set up an electric field around itself that reorients the electrons in the nonpolar molecule, accordingly creating a polarity to that molecule. The dispersion force is exhibited between two nonpolar molecules. Polarity in a molecule is caused by one of its atoms exerting a greater pull on the shared electrons exposing the positive nucleus of the less electronegative atom. A nonpolar molecule lacks this misalignment of electrons and on a time average the electrons “spend equal time” around all the atoms in the molecule. The instantaneous location of electrons however, can due to the correlated movement of electrons, be concentrated on one side of the molecule and induce a dipole. These dipoles are short-lived and fluctuate in direction. The dispersion force is also known as the London force and is the weakest of the three, while the Keesom force is the strongest.

The model of the van der Waals forces needs to be highly repulsive at short ranges to account for the negative electron clouds repelling each other. At intermediate distances the London, Debye and Keesom interactions create an attractive force. At longer distances the force approaches zero. A popular function that obeys these requirements is the Lennard-Jones potential.

$$E_{ij}^{LJ} = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{5-30}$$

It relates the potential energy of two molecules (i and j) to their distance of separation in terms of the parameters ϵ and σ. ϵ is an energy parameter that relates to the minimum energy corre-

sponding to the equilibrium separation. σ is a distance parameter that is equal to the intermolecular separation when the potential energy is zero (Prausnitz et al. 1999). The $\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}$ is the repulsive part and $\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}$ is the attractive term. A typical Lennard-Jones curve is shown in Figure 5-7.

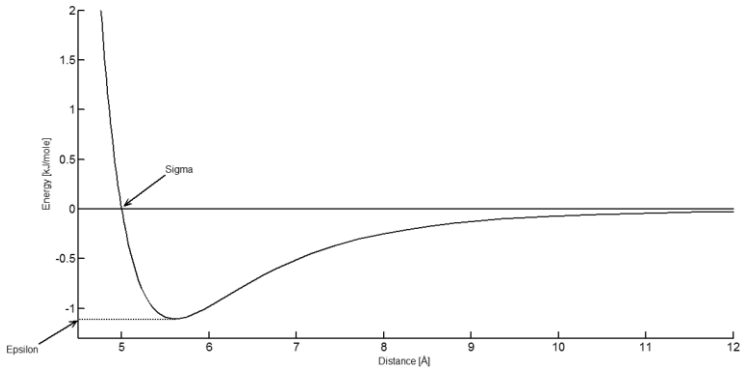

Figure 5-7. Lennard- Jones curve

Another more accurate model is the Buckingham potential, which has replaced the repulsive term with an exponential; more closely resembling the way the electron density falls off exponentially with distance from the nucleus. The Buckingham potential has a third parameter that acts as a fitting parameter, this allows for some flexibility in the shape of the repulsive interactions. Normally the repulsive part of the Buckingham potential is softer than that of LJ, giving a softer contact. Despite the increased accuracy of the Buckingham potential the Lennard-Jones model is most often employed in force fields. This is due to the reduced computational cost of not having to compute exponentials, which take approximately 5 times longer to calculate than simple adding and subtracting. The calculation of non-bonded interactions is the most time consuming part of the potential energy calculations and reducing it will have large consequences on the overall performance of the simulation. The van der Waals energy is calculated for atom pair interactions, but because the pair parameters are fitted to experimental data, the parameters have therefore incorporated many-body properties. This incorporation of many-body properties seems to reduce the disadvantage of the Lennard-Jones potential when applied to large systems (Kuznetsova 2001). This is especially true for pure components, when dealing with a mixture the simplified nature of the potentials makes the correlation less displayed due to the possible poorer packing of the two molecule types represented by the averaged sigma.

The parameters in equation (5-30) are for interaction between two atoms of the same type, when looking at two different atoms the atomic parameters of each need to be combined. There are several ways of doing this; in MD43 the Lorentz-Berthelot mixing rules (5-31) are applied. Here ε^{AB} is given as geometric mean and σ^{AB} is represented by arithmetic mean of the atomic parameters of each molecule.

$$\sigma^{AB} = \frac{1}{2}(\sigma^{A} + \sigma^{B})$$
$$\varepsilon^{AB} = \sqrt{\varepsilon^{A}\varepsilon^{B}} \qquad (5\text{-}31)$$

Where A and B are the atom types. This and other mixing rules are only estimations that reduce the number of independent parameters for which known values are needed. There are doubts about how accurate the results of these estimations are (Rowley and Pakkanen 1999) and ideally each pairwise potential should be parameterized using *ab initio* simulations (rigorous quantum mechanical simulations) but this is a very difficult task to achieve, and would demand a large amount of parameters in each MD simulation.

5.9.2 Electrostatic forces

A molecule in its normal state is electrostatic neutral. I.e. the number of electrons and protons are balanced. If two such electrostatically neutral atoms are covalently bonded and one atom is more electronegative than the other, this atom will pull more on the shared electrons and it will be a polar molecule. The molecule on a whole is still electrostatically neutral, but the charges on each of its atoms are slightly different. An example of this is water, H_2O. In H_2O the oxygen is more electronegative than the hydrogen and the molecule is dipolar. The oxygen will be slightly negative and the hydrogen will be slightly positive when compared to their unbound charge. In Figure 5-8 we see a water molecule and how the electrons (black dots) are envisioned shared, the hydrogen has one electron itself and one from the oxygen atom, the negative electrons will lay slightly closer to the electropositive hydrogen.

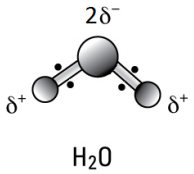

Figure 5-8. Water molecule with partial atomic charges.

The charges that atoms have when they take part in an electrostatic neutral polar molecule are called partial atomic charges. A positive partial atomic charge is represented with δ^+ and a negative charge by δ^-.

The electrostatic energy of interaction between polar molecules is in accordance with Coulomb's Law:

$$u_{AB} = \frac{1}{4\pi\epsilon_0}\frac{q_A q_B}{r_{AB}} \tag{5-32}$$

Where u_{AB} is the pair potential, ϵ_0 is the permittivity in vacuum, q is the charge of particle A or B, r_{AB} is the particle distance, $\frac{1}{4\pi\epsilon_0}$ is a proportionality constant. If the charges are of the same polarity u_{AB} is positive and if the polarity differ the force will be negative and A and B will attract each other.

If it is a polar molecule the electrostatic energy will dominate the potential energy and it is therefore important that the model gives an accurate representation of the involved components. If one were to calculate the Columbic forces between all possible atoms and their images (PBC) in an MD simulation, the simulation would be too costly, this is partly because the Columbic forces decay slowly. One option is to implement a cut-off distance as shown in Figure 5-2, and then the computationally expensive tail is simply ignored. This approach introduces some unphysical effects and the accuracy is quite poor since the remaining force are small but numerous (Frenkel and Smit 2002). Instead the Ewald summation technique is a very common method to calculate electrostatic interaction in a system with periodic boundaries.

In Ewald summation the forces are split into a "near"- and "far"-field contribution. In the "near"-field the point charge q_i is surrounded by a diffuse charge distribution, or cloud, with a total charge equal, but of opposite sign, to q_i. The electrostatic potential from particle i is now short range, and can be calculated easily. The screening distribution is usually in the form of a Gaussian distribution and the width decides how rapidly the charge converges. Next it is

needed to compensate for the screening distribution that was introduced, as we do not want to calculate the electrostatic interaction between two screened particles. To cancel out the original distribution a new distribution of equal magnitude but of opposite sign is introduced. The interaction between two Gaussian charge distributions are long range but because it is both smooth and periodic it can be evaluated in reciprocal space by Fourier methods (Jensen 2007). In addition a term excluding the self-interaction is needed; this is a result of the interaction between the charge cloud and the point charge at its center.

The electrostatic potential U_{Ewald} is now comprised of the Fourier-space contribution $U_{reciprocal}$, the contribution from self interaction U_{self} and the short range potential $U_{short\ range}$

$$U_{Ewald} = U_{reciprocal} - U_{self} + U_{short\ range} \tag{5-33}$$

$$U_{reciprocal} = \frac{1}{2V}\sum_{\vec{K}\neq 0}\frac{4\pi}{k^2}\left|\rho(\vec{K})\right|^2 e^{-k^2/4\alpha} \tag{5-34}$$

Where V is the volume of a cubic box with length L, $\vec{K} = \left({}^{2\pi}/_{L}\right)\vec{l}$ where $\vec{l} = (l_x, l_y, l_z)$ the lattice vectors in Fourier space. $\rho(\vec{K})$ is the charge density (5-35) and α is the width of the Gaussian distribution.

$$\rho(\vec{K}) \equiv \sum_{i=1}^{N} q_i e^{-i\vec{K}\cdot\vec{r}_i} \tag{5-35}$$

In equation (5-35) N is total number of particles, q_i is the charge of particle i, i is the imaginary unit and $\vec{r}_i$ the position of particle i.

$$U_{self} = ({}^{\alpha}/_{\pi})^{1/2}\sum_{i=1}^{N} q_i^2 \tag{5-36}$$

Where α is the same as in (5-34) due to the interaction between the point charge and the Gaussian.

$$U_{short\ range} = \frac{1}{2}\sum_{i\neq j}^{N}\frac{q_i q_j \mathrm{erfc}\left(\sqrt{\alpha} r_{ij}\right)}{r_{ij}} \tag{5-37}$$

The erfc() is the complementary error function. The calculations of Ewald sum method scales as $N^{3/4}$, and requires only slightly more computer time than the cutoff-based method at the same time as giving much better results.

5.10 RDF

Radial distribution functions (RDF), also called pair correlation function $g(r)$, is a way to characterize the distribution of particles in a homogeneously distributed system via the probability of finding a particle at a given distance. RDFs are the ratio of the actual density of particles at radius r to the average density ρ (Dill and Bromberg 2003). The radial distance around a test particle is discretized and the number of particles in each shell described by the true density $\rho g(r)V$, where V is the spherical volume of the shell. When the true density in a shell is the same as the average density, $g(r) = 1.0$. If $g(r) > 1.0$ the particle density in the shell is higher than ρ.

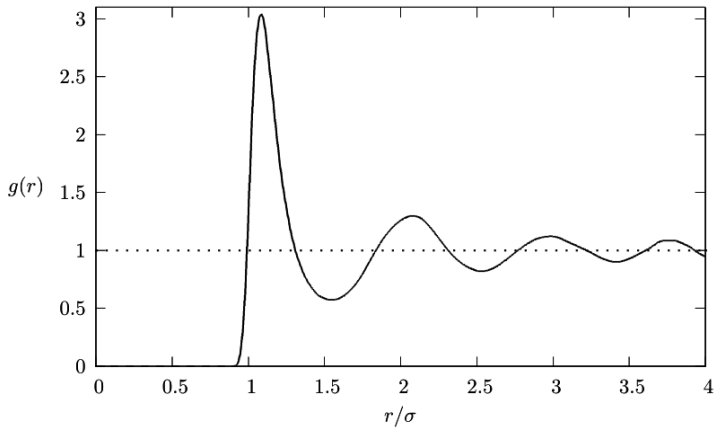

Figure 5-9. RDF for a Lennard Jones liquid

Figure 5-9 shows a Lennard-Jones liquid RDF. Since particles cannot overlap we see that $g(r) = 0$ for $r < 0{,}8$ then there is a peak indicating an excess number of particles comprising the first neighboring shell. After the second shell there is little correlation and the RDF goes toward 1, uniform distribution.

It is possible to calculate RDFs directly in the MDynaMix package. This requires the manual input of each atom pair to be evaluated, for instant O in H_2O and the first O in CO_2 and another for the second O in CO_2. The implementation does not allow specifying a volume, region or part of the system, like for instance bulk volume or interface. An effect of this is that MD43 will include contributions from selected atoms regardless of their position in the system, thus obscuring a possible local structure. The visualization program Visual Molecular

Dynamics (VMD) (Humphrey et al. 1996) that is used in this thesis to visually evaluate the trajectories calculated by MDynaMix also has a built inn function that calculates RDF. VMD has a very powerful atom selection language that allows the selections for the RDF to be very accurate. However, because VMD only have the data output from MD43, and MD43 have been set to generate output at 200 step intervals, the resolution of the RDFs are lower than if MD43 calculates them continuously while the simulation is running.

5.11 Molecular models

In order to produce accurate results the force fields requires models, or potentials, for each atom. The accuracy of these models will decide the accuracy of the whole simulation. Often there are many different models to choose from. In such cases choosing an appropriate model involves weighing the strengths of each model against the requirements of the simulation in mind. For other molecules there might be an issue trying to find models that are parameterized correctly for the temperature and pressure conditions in question. Here the choice of one model can often govern the choice of others; as the potential models must be of the same functional form. Especially the function describing the van der Waals forces (Chapter 5.9.1) might be a problem as the more accurate Buckingham potential is not as widely parameterized for all substances as the more popular Lennard-Jones potentials are. The models used to represent the water, carbon dioxide and calcite molecules are listed below.

For the water two models are used, the extended simple point charge (SPC/E) and a fully flexible 3 point model with good energy conservation (F3C). In both models the water molecule is represented by three point charges bound together by bonds. The assumption of point charges leads to an incorrect value for the permanent dipole of water. To correct for this the H-O-H angle are set slightly wider than the experimentally found angle (109.42 instead of 104.45). The SPC/E model is a reparameterization of the single point charge model (SPC); the oxygen charge has been changed resulting in better density and an improved diffusion constant. The F3C model, developed by (Levitt et al. 1997), is a simple water model that allows for efficient computation. It includes full flexibility, good energy conservation and a smooth atom-based truncation. It was developed in order to simulate macromolecules in large-scale simulations; it is therefore a simple model that nevertheless has been show to correspond well with experimental data. The main difference between the models is that the F3C has a slight van der Waals interaction on the hydrogen atom offsetting the otherwise unshielded Coulombic attraction between positively charged water hydrogens and calcite anions. Both these

models are widely used and have been compared for effectiveness f. inst. (Levitt et al. 1997; Mark and Nilsson 2001; Lin et al. 2010).

The parameters for SPC/E water has been taken from (Berendsen et al. 1987). Parameters for F3C are taken from (Levitt et al. 1997), but here the SPC/E charges have been kept due to better accordance with RDF-s (Kvamme et al. 2009).

Table 5-I. Parameters for the water models.

		SPC/E	F3C
R(OH), [Å]		1,0	1,0
HOH, degree		109,47	109,47
Charge, q [Q]	O	-0,8476	-0,8476
	H	+0,4238	+0,4238
σ [Å]	O	3,166	3,1660
	H	0,0	0,8021
ε [kJ/mol]	O	0,6507	0,8021
	H	0,0	0,04184
A [kJ Å^6/mol]	O	0,0	2889,7
	H	0,0	106,20
B [kJ/mol]	O	0,0	293206,3
	H	0,0	11537,16
C [1/Å]	O	0,0	3,659
	H	0,0	3,875

The carbon dioxide in the system was modeled using the elementary physical model (EPM2) developed by Harris and Yung (Harris and Yung 1995). It is a rigid three-site model with a Lennard-Jones charge at each atom. It reproduces a quadrupole moment close to that shown by experiments (Harris and Yung 1995). The quadrupole moment describe the effective shape of the ellipsoid of nuclear charge distribution. This model is widely used and has shown to produce good results when paired with the water models used here (Demurov et al. 2002; Zhang and Singer 2011).

Table 5-II. Parameter for Carbon Dioxide

		EPM2
R(CO), [Å]		1,149
HOH, degree		180
Charge q [Q]	C	0,6512
	O	-0,3256
σ [Å]	C	2,7570
	O	3,0330
ε [kJ/mol]	C	0,2339
	O	0,6694

For the calcite crystal the potential is made up of one model for the Ca^{+2} ion and one for the carbonate ion CO_3^{-2}. The location of the ions is decided by the crystal structure shown in Figure 3-1 and the input to MDynaMix is in the form of XYZ coordinates for each atom in the crystal. The charges are balanced in such a manner that the crystal as a whole is neutral. The calcium is represented by a 1-site model and the carbonate by a 4-site model. This model has been used previously in (Kvamme et al. 2009)

Table 5-III. Parameters for the Calcite crystal.

	Ca^{+2}	C	O
Charge, q [Q]	2,0	1,2046	-1,0682
σ [Å]	0,899	3,742	2,851
ε [kJ/mol]	113,819	0,5021	0,6657
A [kJ Å^6/mol]	55686,7	2432,71	1123,56
B [kJ/mol]	82942,86	369822,7	230230,1
C [1/Å]	2,198	3,6019	3,9602

5.13 Running in parallel

In parallel programming, single tasks are split into a number of subtasks that can be computed relatively independently and then aggregated to form a single coherent solution. In the following part a brief description of the hardware and software used in this thesis will be given.

5.13.1 Hardware

The numerical analysis in this thesis has been simulated on Hexagon, a Cray XT4 supercomputer associated with the University of Bergen. Hexagon consists of 1388 quad-core processors, a total of 5552 processor cores, and it has a computational capacity of 51.7 teraflops at peak performance. The processors are AMD 64-bit Opteron cores that communicate internally on a node true an integrated high speed direct point-to-point link, relinquishing the need for a peripheral controller which would compete for main memory access and reduce the latency experienced by traditional networking technologies. The nodes are in turn connected to a seastar2 chip, a dedicated communication chip that divests the processors of message handling routines. These seastar2 chips interconnect in a 3D torus topology (Figure 5-10) providing a rigorous high-bandwidth backbone structure to the computer system. In a torus network the nodes on the ends are connected to the nodes on the other side as seen in Figure 5-10B. Information passing through the network will pass through each node, the torus configuration halves the longest path, i.e. reduces the diameter of the system by half. As a torus the physical system is implemented as a folded torus to reduce the cable length between nodes. The 7-ported seastar2 provides one processor port and six network ports corresponding to the positive or negative x, y and z direction of the torus. The Cray XT system is not a single large scale system but can be viewed as an aggregation of small systems, creating a highly adaptive system and facilitates easy up scaling. The seastar2 technology allows for up to 30 000 nodes to be interconnected.

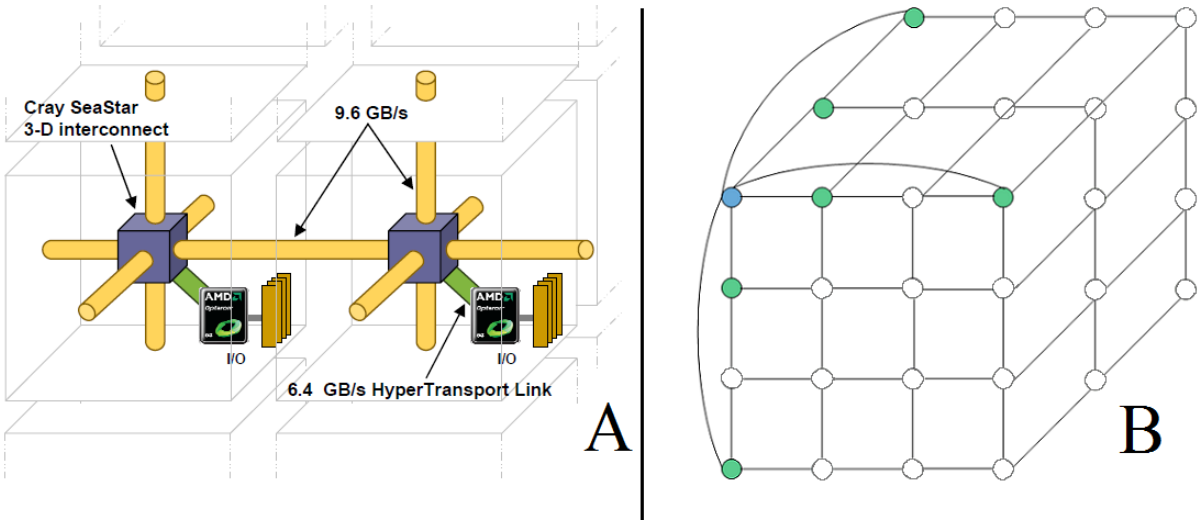

Figure 5-10. A; Seastar2 configuration of nodes with processor link (green). Picture from (Seastar2). B; Torus network, Blue node is connected to six green nodes (All connections are not shown for the sake of simplicity).

5.13.2 Software

For an operating system Cray XT4 runs a stripped down Linux version that provides the user with the standard look and feel of a Linux command system and shells. This enables the users to write their own program code as if they were sitting on a desktop Linux machine, it also ensures that programs written for use on Hexagon are transferable to other platforms. This Linux environment is designed to run large complex applications and it can be configured to match different system configurations and workloads. Its light weight kernel ensures that the scalability of the users programs is not hindered by the presence of the operating system.

When a program is run on Hexagon it is run in parallel across several nodes. That is the program or problem we want to calculate is divided into smaller processes and each processor calculates one part before the result is transported back to a central processor for merging or to an I/O node and to be stored in a memory bank. The method responsible for transporting the packages of information to and across nodes is the Message-Passing Interface (MPI). MPI is a message-passing library specification meaning that is it is a standard, not an implementation. It contains a specific set of rules and specification that software programs can follow to communicate with each other when run on several processors. MPI is not a sanctioned standard yet it has become a widely used mean for communication when running in parallel. On Hexagon the implementation is MPICH2 which is an extensively used implementation of MPI.

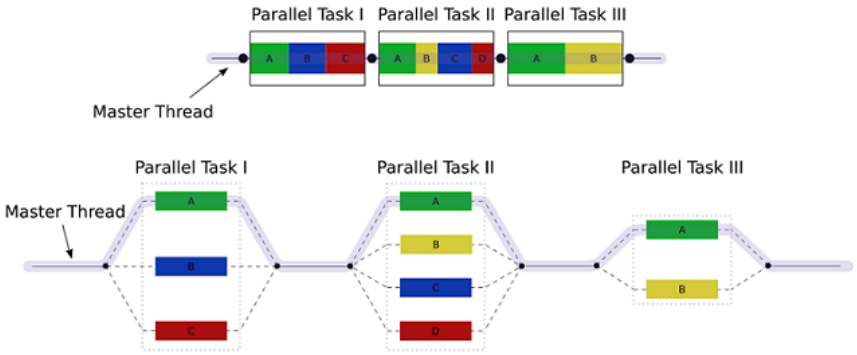

Figure 5-11. Parallelization of three tasks that run in sequence, after each parallel task the results are gathered and a new parallel task is started.

The molecular simulation program used in this thesis, M.DynaMix, is well suited to run on both single desktop computers as well as in parallel on massive parallel computers. When used on a supercomputer the parallelization strategy implemented is called the "replicated data" technique. In this method all the nodes keep the position of all the particles in the system, while the calculation of different contributions to the forces is divided between the nodes and done in parallel. After the individual processors are done calculating the progression of the forces of the particles they are responsible for the updated trajectories or coordinates of the particles are sent back to a central processor where the main position list are updated and sent back to each processor. One advantage of this method is the simple and efficient distribution of the force calculations over the processors for nearly all types of force fields and molecular structure. An obvious weakness of the technique is the relative high communication cost from collecting the contributions to the total forces and distributing the new particle positions again to all the nodes at the end of each time step. However these communication operations are unavoidable in most general parallel MD scheme.

5.13.3 Number of processors

When deciding how many processors to run a simulation on one needs to consider how many molecules or atoms there are in the simulation. As the processors will each calculate the forces interacting between the atoms assign to them, if too few processors are chosen the speed will be slower than optimal and if there are more processors than molecules; the force calculations between two molecules will be split on several processors, and time will be lost on transferring data between individual processors. How efficient the simulation program itself is coded for parallelization matters when deciding the number of processors. In MDynaMix there are two global operations, i.e. an operation that transfers data to a central processor.

These are the final summation of the forces and the broadcasting of the new coordinates to all nodes. Lyubartsev and Laaksonen (Lyubartsev and Laaksonen 2000) report that this transport of data between processors becomes noticeable when passing 64 processors on the parallel computer available to them. The implementation of the parallel algorithms in MDynaMix is directed at the atoms and atom coordinates, giving a relatively simple code that is suitable for simulating a variety of different molecular systems. The general code is in a manner that allows the parallelization to be achieved with a minimum of additional data structures and special operations allowing for an excellent scalability with the number of atoms and attached processors.

6 Simulation

In the design of the system it was important that it would depict a physically realistic system. The system that we ended up using had a water-wetted calcite crystal in an aqueous phase, the hydrate and crystal was separated by water as they would be in real life. In addition a gas phase was added to act as a reservoir of CO_2 that could dissolve and interact with the components of the system. This CO_2 phase could either represent injected CO_2 in context with CCS or CO_2 present in a reservoir. The effects of PBC are also important when considering the order of building blocks. The details of the construction process are outlined in the following parts.

6.1 Building the system

The system was constructed using a building block approach, each part being built individually before the parts were combined to form the whole system. The input required for MD43 is number of molecules and the size of the box. It was decided that three systems were to be simulated, two at different temperatures and a third with a different water model. In the following subchapters the details of building the systems will be discussed.

6.1.1 Estimating the number of molecules - Density calculations

In order to achieve the proper density of each block the number of molecules was calculated using the density (ρ) of the phase described. The density is in turn a function of temperature (T) and pressure (P), the same parameters that describes the simulation environment, in setting the pressure at this stage it is ensured that the simulation process will advance at the desired pressure. Equation (6.1) was used to calculate number of molecules and the results are shown in Table 6-I.

$$N = \frac{\rho * V}{M_w} * N_a \qquad \textbf{(6.1)}$$

Equation (6.1) was derived from the density formula ρ=m/V, where m is mass and V is volume. The number of molecules is given by the definition of mass $m=N*M_w$ where N is number of molecules in moles and M_w is the molecular weight. Multiplying by Avogadro's number (N_a) converts from number of moles to number of molecules.

In deciding the size of the system the number of molecules also plays a significant role. The larger the system the more molecules and the more computationally expensive the simulation becomes. The x and y directions is decided by the dimensions of the hydrate structure thus leaving the z direction to be adjusted according to our requirements for the system. As mentioned a larger system will require more computational power, especially the K values in the Ewald summation (Chapter 5.9.2) is limiting the length as it sums over the lattice vector in Fourier space and number of molecules.

6.1.2 The calcite crystal

The calcite crystal was constructed from its unit cell using CrystalMaker. This crystal was reduced in VMD to an appropriate size and to expose the desired surfaces while retaining the correct molecular formula of calcite. The replication process was performed by running a script, a text file containing a recipe of computer commands written in the TCL language. A TCL language interpreter is built into VMD to allow its users to create their own add-ons and scripts in order to automate repetitive and time-consuming tasks. Because the $\{10\bar{1}4\}$ surface was to be exposed, a much larger crystal than the finished one was needed. After the large crystal was created and the $\{10\bar{1}4\}$ surface identified, layers of the crystal was removed until the size and shape was satisfactory. The final crystal used in the simulations can be seen in Figure 6-1, in figure A the $\{10\bar{1}4\}$ surface is in the right and left side, figure B is a rotation of A by 90 around the z axis, figure C is a view onto the $\{10\bar{1}4\}$ surface. Due to the geometry of the crystal the carbonate groups appear to be connected, especially in figure B and C, there are no connections between the different CO_3 groups; only the overlapping of Carbon (in B) and Oxygen (in C) in depth. After the resizing the crystal it no longer has a unit cell structure, meaning that it can no longer automatically be copied in all directions. A .mol file contains the structural information and force-field parameters md43 uses to construct the system and evaluate the interaction forces. When constructing the .mol file for the calcite crystal the placement of each atom was specified by its z, y and z coordinates.

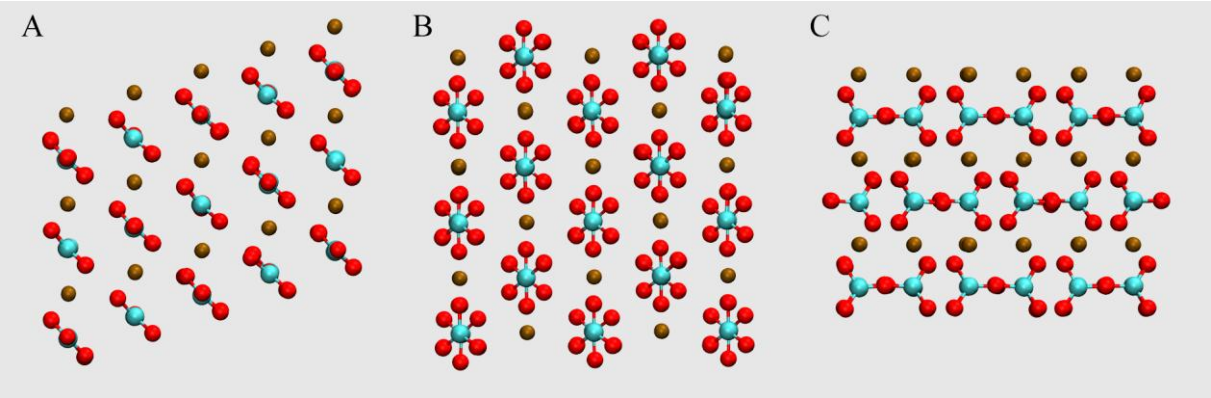

Figure 6-1. The calcite crystal viewed from three different angels, showing the highly ordered geometry of the crystal. Brown is calcium, blue is carbon and red is oxygen.

In order to investigate how the calcite crystal was affected by interaction with the other constituents of the system the crystal needed to be free floating in a body of water. This allowed the crystal to react and move as the forces from the hydrate, CO_2 and water affected the crystal until it reached an equilibrium position. This also decided the size of the water itself as enough water was needed for the calcite crystal to rotate freely, and not to be constrained by surface effects. This makes the diagonal of the crystal important as it is the longest part of the crystal and in the worst case scenario the diagonal could align itself with the z axis requiring the largest water phase. Due to the rhombohedral shape of the crystal any increase in size of the $\{10\bar{1}4\}$ surface would lead to a large increase in length of the crystals diagonal (see Figure 6-1 A).

6.1.3 The gas and water phases

The size of the water phase is dictated in x and y by the size of the hydrate phase, in z direction the concerns for the calcite crystal as mentioned in the previous chapter together with restrictions in overall system size are the main criteria. Similar concerns was the basis for deciding the size of the gas segment, it needed to be large enough for the main water body to be shielded from the hydrate, it needed to be large enough to embody bulk properties and it needed to have enough molecules to allow gas to diffuse into the water layers to interact with the hydrate and calcite. To investigate if the hydrate structure facilitated further structuralization of water molecules a thin water layer was placed between the gas segment and the hydrate structure. This water layer was not thick enough to display bulk properties and would at such facilitate sizable and rapid transferal of CO_2 molecules up to the hydrate surface at the same time as it contained building material for further hydrate structuring. The system now

has two gas/water/hydrate boundaries, one thin and one thick that will show if the gas diffuses through the main water body past the calcite crystal and onto the hydrate surface.

6.1.4 The hydrate phase

The hydrate phase was obtained from earlier work done in the research group. It comprises of four large cages in the x and y directions as shown in Figure 6-2 (A). We can also see that the hydrate phase has unit cell properties in the x and y direction, meaning that it can be expanded by placing images of the main cell in the x and y directions. This is useful when periodic boundary conditions are applied in x and y directions as this would create one large slab of hydrate. In the z direction the hydrate phase is two full large cages deep, in addition there are incomplete cages at each side as seen in Figure 6-2 (A) and (B). These incomplete cages lead to the loss of unit cell properties in the z direction, but the advantage of this is that the rigorous structure of the hydrate cages are exposed to free water and free carbon dioxide, or other constituents for other systems. This allows us to study how free components interact with and are affected by the hydrate. One other advantage is that the smaller hydrate cages are exposed. Without being completed these smaller cages have the possibility to admit larger guest molecules than is possible in the finished structure, this is interesting when looking at the formation of hydrate as guest molecules of different sizes might exchange place as the chemical and thermodynamic conditions change. In Figure 6-2 (A) and (B) the hydrate phase has been rotated 90° about the z axis, by comparing the two pictures we can see that the large ages at each side are in different states of completion this allows us to evaluate the degree of affluence afflicted by the hydrate structure on its neighbors.

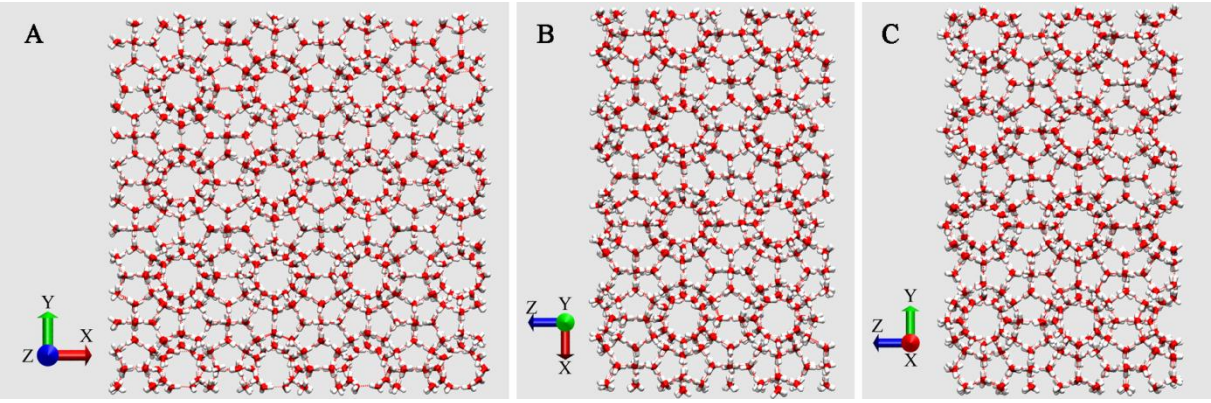

Figure 6-2. The hydrate phase at different rotations. (A): the xy plane. (B): the xz plane. (C): the yz plane

6.1.5 Constructing the main system

The different segments can be seen in Figure 6-3 where the x axis represents the depth of the system and is of the same magnitude as the y axis that is along the vertical side.

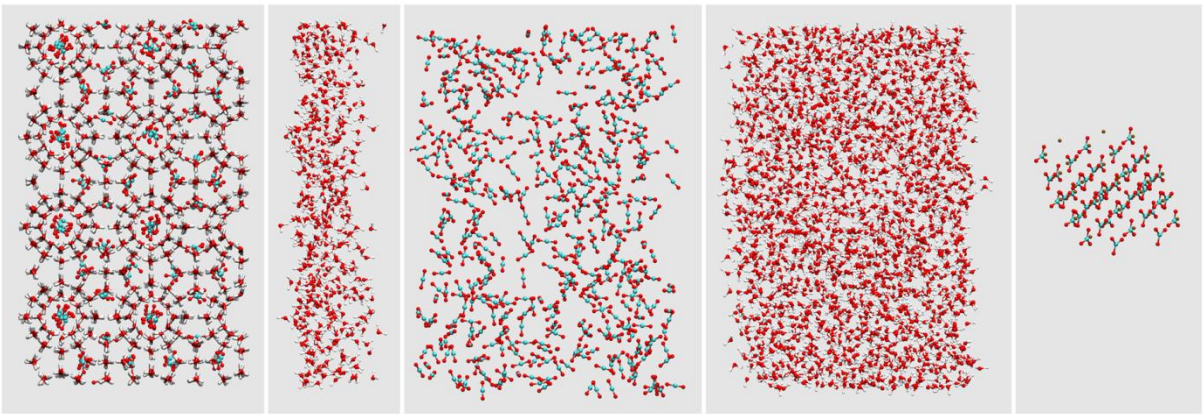

Figure 6-3. The individual parts of the system in the order they are assembled. Here The Z axis is horizontal and the Y axis is vertical, The x axis is the depth and is equal in size to the y axis.

After the size of each segment was decided and pressure and temperature set; each segment was constructed using MD43. Here each molecule, or rather the mass center, was placed along a grid filling the box volume, the rotation was random. This artificial placement of the molecules will yield large forces in the system as this configuration is unrealistic. The numerical value of these large forces is not of interest when evaluating the system and therefore the forces are limited until they have diminished to an acceptable value. The calculation of molecular density and number of molecules in each phase as outlined in chapter 6.1.1 allows this technique to be effective. If the system size had been too small for the number of molecules the repulsive forces resulting from close proximity and overlapping molecules would have forced the molecular modeling software to terminate regardless. In Figure 6-3 we see the individual parts after their respective equilibrium runs, they are here depicted in the order they are to be combined. The initial equilibration runs were performed using the velocity scaling technique, with delta T set to 5 K, to maintain the system temperature. After each individual box had been equilibrated the main system was constructed by stacking the individual boxes in the z-direction. The main system was then transferred form the local computer clusters to the Hexagon supercomputer. From this start position, several trials were run from the same initial configuration to adjust and fine-tune the main system. The system was again simulated using the velocity scaling technique due to Nosé-Hoovers inability to handle large fluctuations in the system. After this initial equilibration the thermostat was changed to Nosé-

Hoover. The completed system can be seen in Figure 6-4, here the crystal has already rotated away from its initial orientation.

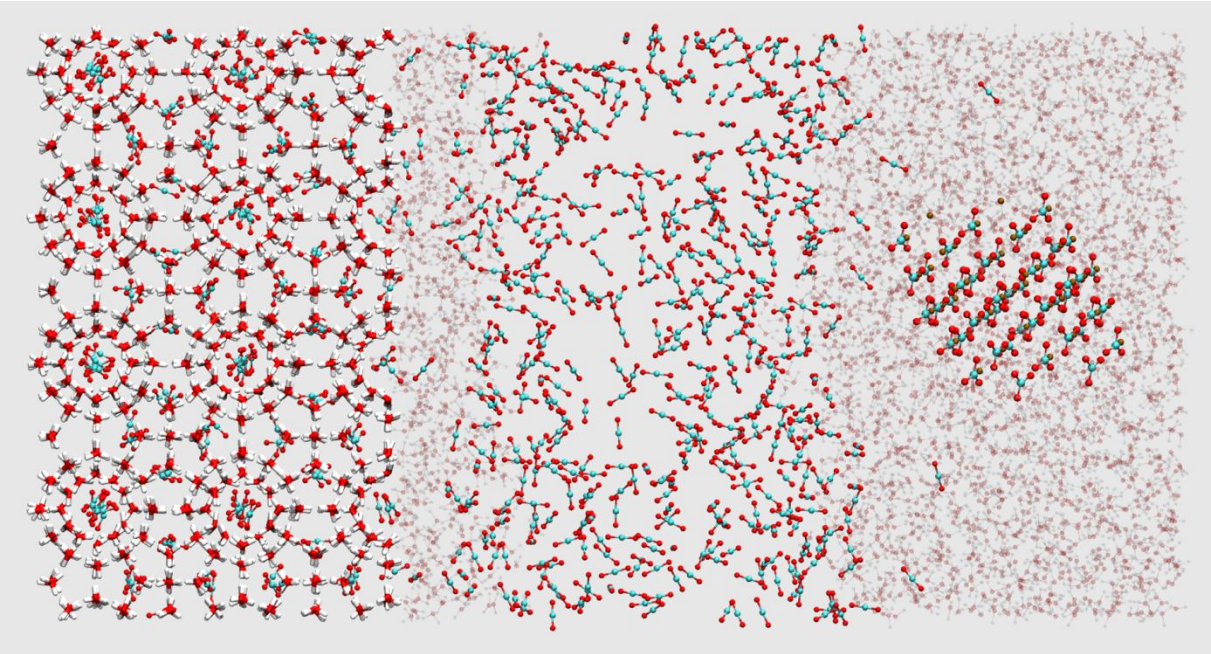

Figure 6-4. The complete system. Here the H_2O layers have been toned down to expose the calcite crystal and the dissolved CO_2

6.1.6 Simulation conditions

The temperature of the simulation was decided in accordance with the prerequisite of the main goal. To decide on a realistic temperature the reservoir Utsira was chosen as a template. The Utsira reservoir is a sandstone formation spanning a large portion of the central North Sea. The Utsira reservoir is the location of the Sleipner project, the world's first commercial application of carbon capture and sequestration technologies for geological storage of CO_2. In this thesis pure CO_2 was used, in reality any CO_2 removed from flue gas will contain impurities, these impurities will affect the density of the CO_2 phase. The injection point at Utsira lies at 1012 meters depth, the temperature profile of the formation is based upon a single downhole measurement of 37°C. This measurement is subject to an uncertainty of up to 10°C due to the fact that the measurement was taken before the fluids in the borehole had reached equilibrium (Nooner et al. 2007). In this thesis the temperature was set to 313.15 K or 40°C and the pressure to 150 bars.

6.1.7 Setup of simulation systems

NVT ensemble was used as the basic for the simulations. Initially simple velocity scaling with 5K as the permitted variation were used for all the systems. This was done since simple veloc-

ity scaling is faster, more robust and does not fail for systems far from equilibrium. After this initial period with simple velocity scaling the Nosé-Hoover thermostat vas switched on with translational and rotational relaxation time set to 100 femtoseconds. The water molecules in hydrate were locked in space but free to rotate around their mass centers. All other were free to move. Short range molecular interactions were modeled using the Lennard-Jones potential with a cut-off distance at 10 Å. Electrostatic interactions were handled by means of Ewald summation, the parameter for real space accuracy (α) is 3,1415 and for reciprocal space 9.81 m/s^2. The time step was set to $0{,}5 \times 10^{-15}$ seconds (0,5 femtosecond). Intermediate averaging were performed for every 5000 steps. Trajectory details were outputted to .DCD files every 0,1 picosecond, meaning a full .DCD file of 500 trajectory steps contains 50 picoseconds of simulation time and md43 calculates 2oo time steps between each trajectory step that is read in VMD. Constraint dynamics in form of the SHAKE algorithm were employed and all the molecules were treated as rigid by setting quaternion treatment to true. Three versions of the system were simulated, System 1 at a temperature of 277 K and the F3C water model. System 2 at 313,15 K with the F3C water model and System 3 at 313,15 K whit the SPC/E water model.

Table 6-I. Number of molecules in each layer.

Layer		Number of molecules	Dimensions [Å]			Density [Kg/m^3]
			X	Y	Z	
Hydrate	H_2O	1839	48.12	48.12	40	
	CO_2	192				
Large Water		2300	48.12	48.12	30	999.87
CO_2		494	48.12	48.12	20	780.3
Thin Water		500	48.12	48.12	6	999.87
Calcite	Ca	45	48.12	48.12		
	CO_3	45				
Total		5415	Z = 96 Å			
Number of atoms		16200				

7 Results and discussion

System 1, 2 and 3 were simulated for a time of 13.8, 14.9 and 14.3 nanoseconds respectively. Each of the systems achieved a computational speed of 10 751 time-steps/hour with 64 processors. A trial run using 8 processors were completed to evaluate the effectiveness of the 64 processors. With 8 processors a computational speed of 1 369 time-steps/hour were achieved, this means that when compared with 64 processors 11 168 time-steps/hour should have been reached if the efficiency of the parallel processing were 100 %. The system therefore scales to 64 processors with 96.3 % efficiency when compared to an "ideal" rate of calculation.

In this chapter RDF has been used to evaluate the average distance to the first layer of adjacent molecules and to study structuring effects above a surface. As mentioned in section 5.10 the RDF calculations performed continuously by md43 has a higher resolution than those calculated by VMD. VMD reads the .DCD files outputted by md43 and these files contain one entry every 200 time step. VMD however allows us to make better selections while md43 always takes the entire system into the calculation. Another issue to keep in mind is that the definition of spherically symmetric atomistic RDFs as used in the analysis throughout this work is only reflecting average values and as such implies uniform distribution. The systems in consideration are multiple phases with distinct interfaces. This makes the analysis more complex since the analysis have to be based on combinations of atomistic correlation functions and even then not with enough information preserved in these sampled g(r). Sampling of molecular correlation functions are very time consuming and would require rewriting of the original software used throughout this work and is outside the scope of this work. A zero order approximation for molecular correlation functions could be reconstructed from the atom to atom correlation function and intra molecular structures using the superimposing principle (Kvamme 1995; Kvamme 2002) and this is recommended in suggestions for further work since the programming involved would be outside the be outside the time limits of this study.

The effect this has on the RDFs shown here is that the height of the peaks cannot be evaluated against RDF data from other systems. The distances on the x axis however are still valid. When comparisons between systems have been done here the numbers of molecules selected are the same for all systems and the peaks can therefore be compared to each other.

7.1 Transfer of water molecules between water layers:

Because the molecules are free to move in the system and are not bound by any constraints keeping them in space the coherency of the phases is important. Especially the thin water layer is of concern as its lack of depth could have warranted its collapse under the effects of the CO_2 phase. In this case we would have seen a large number of water molecules transferred over to the thick water phase. The number of H_2O molecules in each of the water layers has been summarized in Table 7-I. It can be observed that there are limited difference between the initial calculated densities and the densities at the end of the simulation. In system 1 there are at the end of simulation 53 more H_2O molecules in the thick layer compared to initial configuration, in system 2 and 3 there is a difference of 5 and 3 molecules respectively. In all three systems the H_2O molecules has moved from the thin layer over to thick layer. The solubility of H_2O into the CO_2 is limited. But the dynamics of the CO_2/H_2O interface might enable CO_2 to capture H_2O molecules which are loosely attached to the thin layer and transport them through to the thick H_2O layer.

Table 7-I. Number of H_2O molecules i each phase. Transferred molecules describe number of H_2O molecules originally in the opposite phase.

		Original	System1	System 2	System 3
	Thin layer	500	447	495	497
	Thick layer	2300	2353	2305	2303
Transferred					
	Into thick layer	-	107	53	24
	Into thin layer	-	54	48	21

We see that there is no significant change in the number of molecules in each water layer with system 1 having transferred the most with a difference of 53 water molecules at the end of simulation. There is some exchange of molecules between the water phases but this seems to even out with equal numbers escaping from each side suggesting a domino effect where molecules entering one layer of water is followed by another molecule leaving the same layer of water. The effect of the hydrate structure is keeping the whole thin water layer in place and the CO_2 molecules are not able to disturb this water interaction.

7.2 Orientation and motion of calcite crystal

When the system was set up the crystal was immersed into the center of the thick water layer with the $\{10\bar{1}4\}$ surface in the x-y plane. The Ca terminated surface was oriented in the negative y direction. The crystal was centered in the water phase by matching the mass centers and then removing any water within and in close proximity of the calcite. When the simulation was started the calcite quickly rotated away from the initial position. This initial rotation occurred during the initial equilibration of the system and this orientation can clearly be seen as having high energy and is therefore unfavorable for the crystal in this system.

In Figure 7-1 the final orientation for calcite in all three systems are illustrated. Here the $\{10\bar{1}4\}$ surface is facing outwards and the layer of hydrate shown corresponds to area (D) in Figure 7-7. It can be observed that system 3 has the closest resemblance to the initial orientation, with the $\{10\bar{1}4\}$ surface nearly lying in the x-y plane, the Ca corner are here closer to the hydrate surface with the carbonate corner away from it. The Ca terminated surface has however turned in the negative x direction. In systems 1 and 2 the same surface is oriented towards the hydrate surface but the rotation and angle of approach is different, these two systems have the same water model and this might be the reason for the resemblance here.

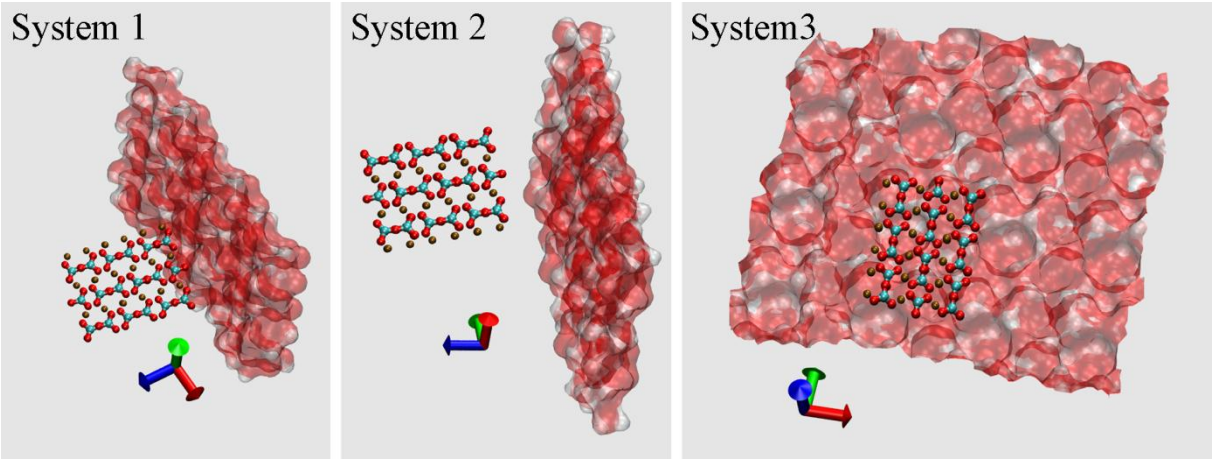

Figure 7-1. Orientation of the systems. Hydrate layer are in the x-y plane (red-green arrows). The $\{10\bar{1}4\}$ surface facing out. Blue arrow in z direction.

The simulation systems are not equilibrium systems in terms of Gibbs phase rule. As such there is no unique phase equilibrium state as asymptotic limit of the simulations. And energy is exchanged with the surroundings continuously through the temperature thermostat so the systems will continuously progress toward lowest possible free energy. There is still a significant degree of motion in the system after the simulation has run for a long time. This is evident in Figure 7-2 where the trajectories for the mass center of the calcite crystal are shown

for 4000 trajectory steps. Here only translational motion and not rotational motion is shown. In Figure 7-2 below, motion is shown in the x-y, y-z and x-z plane. It can clearly be seen that the crystal moves with greatest ease in the x-y plane perpendicular to the interfaces in z directions. In the two other planes system 1 exhibits similar movement patterns while systems 2 and 3 have reduced movement in the x-z plane. This could be because the structure of the hydrate surface in x direction makes a wave shape that can be seen in Figure 6-2 (D) if one outlines the left-hand surface seen there. The same wave shape is not present in the y direction and the calcite might have found a position resulting in more favorable interaction energies. In Figure 7-2, a hydrate cell is shown so that the magnitude of the calcite motion can be seen. The hydrogen bond networks in water are strong and the likely reason that the crystal have low mobility. Another participating factor to the lack of motion in the z direction might be that the distance to the interfaces on each side of the crystal was too short. A simulation setup that investigates, amongst others, this possibility is suggested in further work.

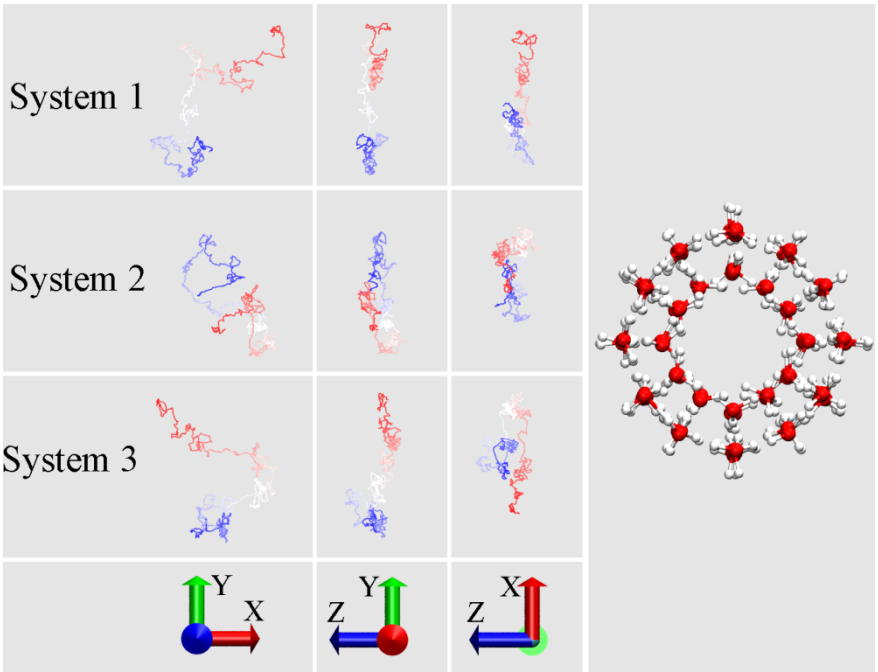

Figure 7-2. Trajectories of center of mass for the calcite crystal in liquid phase. Angle of view is indicated by the system axes. Hydrate cell are shown for size comparison, the trajectories and hydrate cell are 1:1.

7.3 Structuring around calcite crystal.

The charges of the different components are important in efforts to estimate the structuring effects the calcite crystal has on its surroundings. The charges as listed in chapter 5.11 shows us that the crystal faces in this thesis will be surrounded by water molecules and will form hydrogen bonds. It is this structuring that prevents hydrate from forming directly onto the

surface of a material like calcite. The structure of hydrate is never compatible with the water structure imposed by the mineral surface. In the calcite there are three different main types of surfaces as depicted in Figure 6-1: one surface is calcium terminated, meaning that the outer-most layer consists only of Ca atoms, one is predominated by the carbonate group and one is a mix. The large positive charge on calcium means that the CO_2 molecules dissolved in water is expected to orient themselves with one of the oxygen molecules pointing towards the calcium terminated surface. From Figure 7-3 system 2, we see that this is as expected, the curve representing g(r) between calcium and oxygen starts earlier than the curve for calcium and carbon, at 3.45 Å versus 4.35 Å. This is clearest for system 2 and one can see how the first Ca-O peak is followed by a Ca-C peak as one would expect from the linear CO_2 molecule. In system 3 this tendency is not as clear as in system 2 suggesting that the water model used in system 3 needs to be further investigated to see if it can be verified if modeling CO_2 dissolved in water near mineral surfaces.

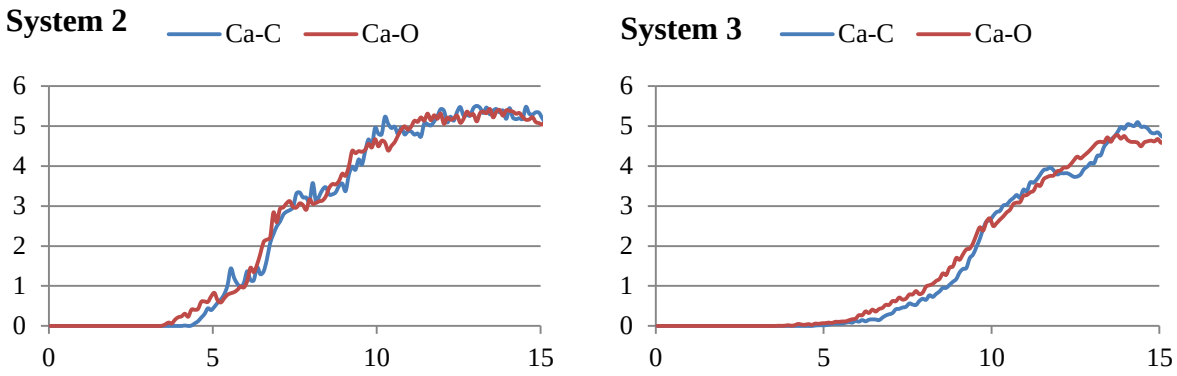

Figure 7-3. g(r) between calcite and CO_2 in systems 2 and 3. From VMD

If we use the Lorentz Berthelot mixing rule in equation (5-31) we find that σ^{AB} for A = Ca and B = O equals 2.2 and 2.05 for B = C. From Figure 7-3 it can be seen that the first-contact value in System 2 is 4.35 for Ca - C and 3.45 for Ca – O (4.75 and 3.65 respectively for system 3). The longer observed distance is likely caused by the adsorbed water layer shown in Figure 7-6 for water, on the right hand side. System 3 has more water molecules packed around the crystal which may cause a shielding effect resulting in the behavior in system 3 seen in Figure 7-3, where the preference for CO_2-oxygen to orient itself toward the crystal is not as pronounced. The difference in first-contact value for system 2 given above tells us that the CO_2 molecule is not orthogonally oriented towards the surface, if that where the case the difference between Ca-O and Ca-C would correspond to the bond length between C-O in CO_2 (1.14 Å).

The number of CO_2 molecules near the calcite crystal during the 500 frames of the last .DCD can be seen in Table 7-II. We see that system 2 has the greatest number of CO_2 molecules closest to the calcite crystal, as would be expected with the increased temperature and improved CO_2 behavior mentioned above for system 2. There were never any CO_2 molecules within 4 Å of the calcite surface as is supported by the radial distribution function in Figure 7-3 and the subsequent discussion.

Table 7-II: Number of CO_2 molecules within [Å] of Calcite during 500 trajectory steps. The average values are the total number of CO_2 molecules within that distance averaged over trajectory steps (500) while the maximum values are the maximum CO_2 molecule within that distance during the trajectory steps.

	System 1		System 2		System 3	
[Å]	average	maximum	average	maximum	average	maximum
5	0.216	3	1.65	9	0.414	3
6	1.698	6	4.362	15	3.252	15
7	4.752	15	9.648	18	7.854	21
8	12.306	27	16.29	24	14.628	33

Due to the charge difference in the calcite crystal the surrounding water layers structure themself differently depending on which calcite side they are closest to. The structuring of a few layers of water can be seen in Figure 7-4. In these pictures the surface of the crystal has been drawn using the surf drawing method in VMD, this method exaggerates the physical size of atoms so that they fill out a surface. This is a simple way of showing how influential the molecules are on the surface. In part (A) of the figure, the calcium terminated surface can be seen to attract the negative oxygen atom in water. This leads to the hydrogen atoms pointing outwards from the crystal. The reverse can be seen on the opposite side (B) as the negative oxygen atoms in the carbonate group attract the hydrogen in water. Both these phenomena can be seen to affect the structure of the superimposing layers of water where the initial effect from the crystal affects the internal structuring in the water further away from the crystal. Above the $\{10\bar{1}4\}$ surface (C, D) the structuring is less evident due to the mixture of calcium atoms and carbonate groups as seen in Figure 7-5, again we see the effects of individual atom types on the edges of the surface this is discussed further down.

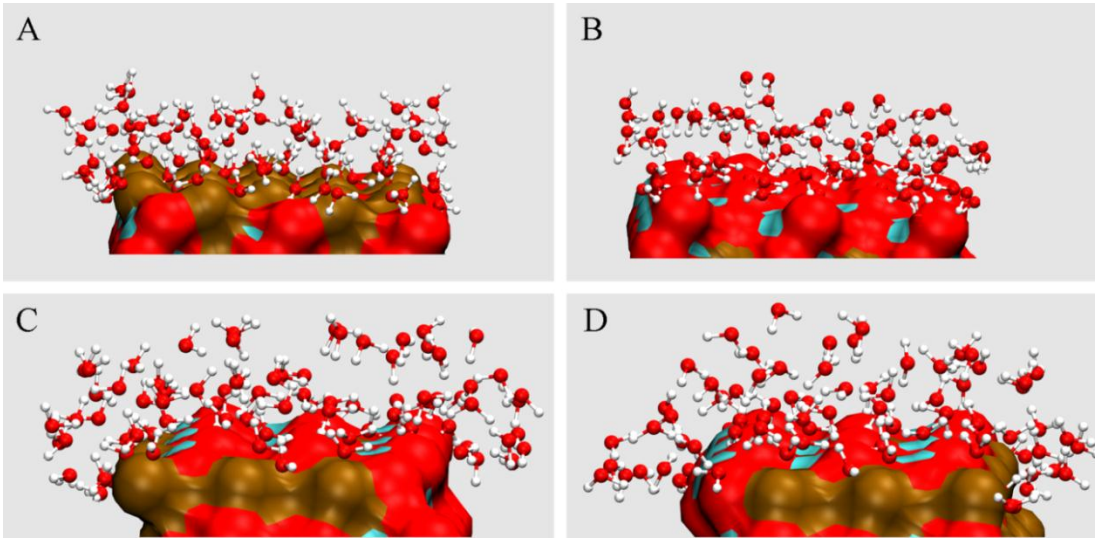

Figure 7-4. H_2O structuring on the different surfaces of the calcite crystal. (A) is of the calcium terminated surface, (B) is of the Carbonate surface, (C) is of one of the $\{10\bar{1}4\}$ surfaces and (D) is of the other $\{10\bar{1}4\}$ surface. In the figures brown is calcium, blue is carbon, red is oxygen and white is hydrogen.

The inner layer of H_2O is quite stable as only 5 of 167 molecules leaves the surface configuration during 3404 frames. Thus the water-calcite interactions appear to be quite favorable for this system. The innermost water layer on the $\{10\bar{1}4\}$ surface is oriented in three manners as seen in Figure 7-5. Near the calcium atoms the water is oriented with the oxygen atom down and both hydrogen atoms straight up. Near the carbonate groups water orients itself in two different ways, some have one hydrogen pointing down toward the oxygen in the carbonate group and the other hydrogen pointed away from the surface. The rest of the water molecules have both hydrogen atoms pointed down towards a carbonate group, these are primarily found on the edges towards the two terminated surfaces. Perdikouri et al (2009) has shown by atomistic force microscopy (AFM) that the growth of a calcite crystal surface contains stepped imperfections (Perdikouri et al. 2009). These imperfections will expose the sublayers, uncovering structures similar to the edges in Figure 7-5.

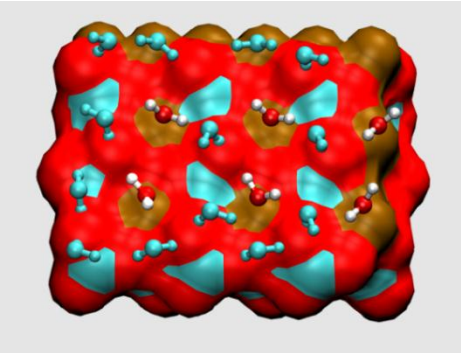

Figure 7-5. H2O molecules adsorbed to the $\{10\bar{1}4\}$ surface. The water molecules have been grouped by color to show the two main orientations. In the crystal brown is calcium, blue is carbon and red is oxygen.

The mixed orientation on the $\{10\bar{1}4\}$ surface compared with that of the previously discussed surfaces lessens the extent of the calcites structuring effects. The inherent disorder in water appears quicker when the water layers are not strictly organized. This means that growth, or existence, of hydrate is more likely to occur at a closer range than above the other crystal surfaces in our crystal. The lessened extent of structuring of this surface is most likely the reason for it being the most dominant in nature. Water is a strong solvent and when it is oriented in the same direction the pull in the surface increases. The layered structure of the $\{10\bar{1}4\}$ cleavage plane protects the surface and the water interaction is there for weaker.

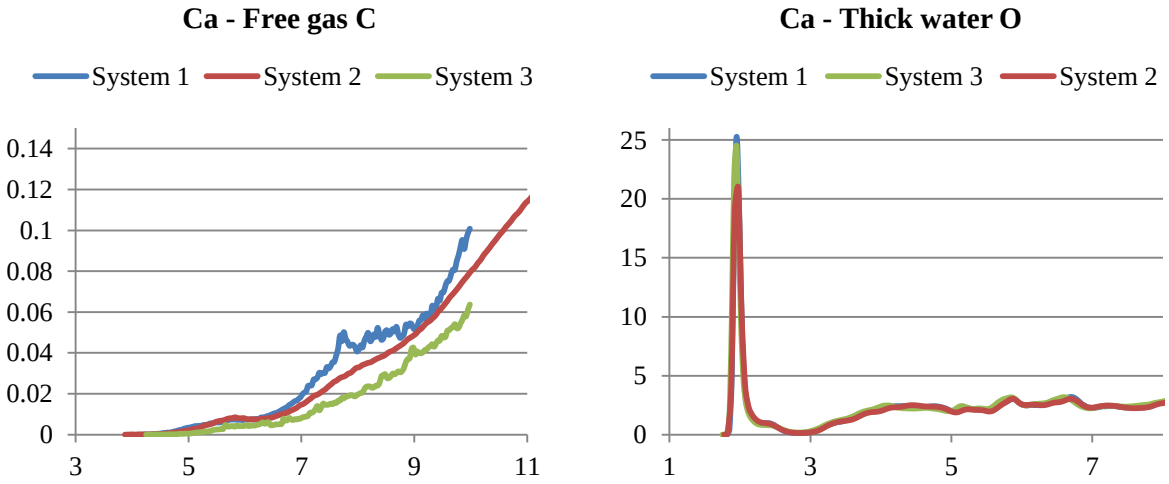

Figure 7-6. g(r) between Ca in calcite and surrounding phase. From md43.

Figure 7-6 illustrates the average CO_2 and water structuring around the calcite crystal. For CO_2 we see that the concentrations close to the crystal are really low. As expected the colder system 1 has more CO_2 closer to the crystal and this system also has a small top around 5 Å. Similar top in the warmer systems is shifted to a larger distance. From g(r) between Ca and O in water we see that the closest water appears at 1.9 Å in all three systems.

7.4 CO_2 adsorbed on hydrate

The shape of the hydrate cages together with the CO_2 trapped in them has a strong structuring effect on the free water. We can clearly see from Figure 7-7 that the cage structure of the hydrate is transferred to the adjacent water. In Figure 7-7, where both sides in the z direction are shown, we see that the smaller cages in (D) have a clear structuring effect on the free water, similar to that of the half cages in (C). There are however, a greater number of CO_2 molecules in (C) as would be expected from their proximity to the gas phase.

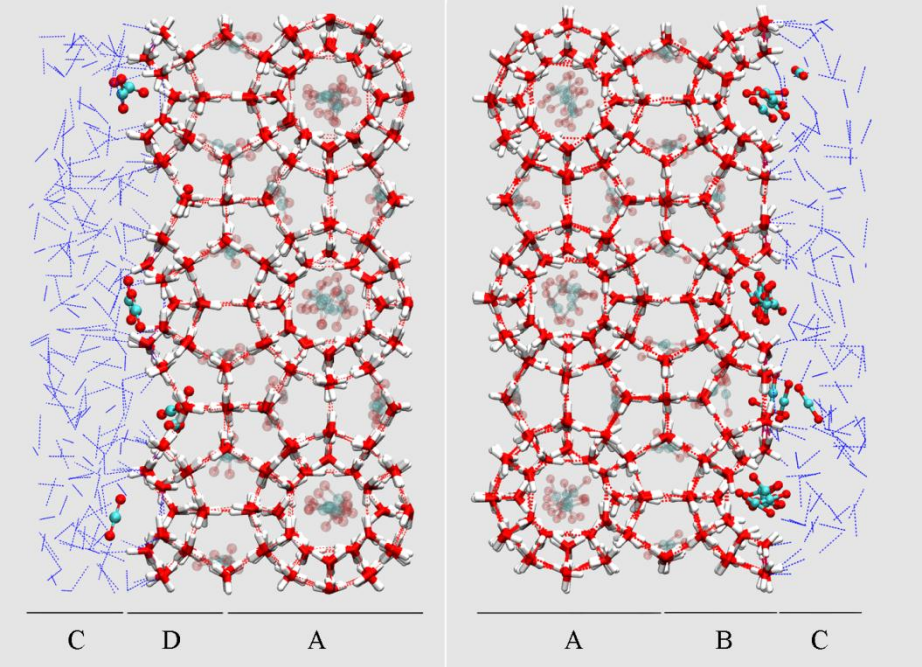

Figure 7-7. The structuring of water on both hydrate surfaces in z direction. Blue lines are hydrogen bonds between water molecules. Clear CO_2molecules are from the free gas phase. Faded CO_2 are guest molecules. (A): Large cages. (B): Half cages. (C): Free water. (D): Smaller incomplete cages.

In Figure 7-7 we see that a number of free CO_2 molecules gather at a position between two half cages in the (C) region, this gathering of free CO_2 was seen in all similar locations for all three systems. Unlike the adsorbed molecules in a similar position in the (D) region these molecules have no equivalently placed CO_2 molecules inside the hydrate. This is most likely because in the (C) region the presence of free water enables the free CO_2 molecules to enter half-cavities that would have been too small for them in a complete hydrate lattice. The presence of such molecules in half-cavities that will not be able to accommodate them in a completely formed hydrate will limit, or slow, the formation of new hydrate. This is because any removing of molecules is associated with a cost, or penalty, in energy. A graph of the number of CO_2 molecules adsorbed in such intermediate positions can be found in Appendix B, Fig-

ure B-2. If the system had contained molecules of proper size to enter these smaller half-cavities, CO_2 molecules might have been exchanged without the penalty being too large, or they may not have entered at all as the smaller half-cages could be occupied by the smaller constituents. The addition of guest molecules suitable for the small half-cavities has been mentioned as suggestion for further work.

The structuring of free water around half cages can, in addition to Figure 7-7, also be seen from the RDFs in Figure 7-8. In Figure 7-8 we look at the interaction between hydrogen and carbon in both complete cages and half cages. The dotted curve (···) shows g(r) between H in hydrate water and carbon in the guest molecule (CO_2). The guest molecules are faded in Figure 7-7 to discern them from free CO_2. We see that there is a clear first peak indicating CO_2 in the main large cavities, this first peak is followed by a void showing that there are no CO_2 molecules until we reach the secondary smaller hydrate cages. These cages are smaller and do not contain as many guest molecules, lowering the magnitude of the peak. It is possible to see the high degree of order that one would expect in a structured grid like hydrate cages. The blue double curve (=) shows g(r) between H in H_2O and C in free CO_2 as seen in Figure 7-7. The selected water in this graph is made up from selection (B) and (C) in Figure 7-7. Due to the CO_2 molecules trapped in the half cage we still see the first peak. However, the curve rises earlier showing that the CO_2 molecules are closer to the H_2O molecules, suggesting that the hydrate cage is compressed. The following dip is on the other hand not as sharp as for the red line, due to the less pronounced structuring in the free water phase. After the second peak any structure is lost due to the movement of free water. The green line is for CO_2 dissolved in bulk water. This line is included to illustrate how CO_2 and H_2O would be structured when there is no hydrate present. We see that this last curve begins to rise at the same time as the curve representing the half cages (blue) and they both start before the curve for hydrate (red).

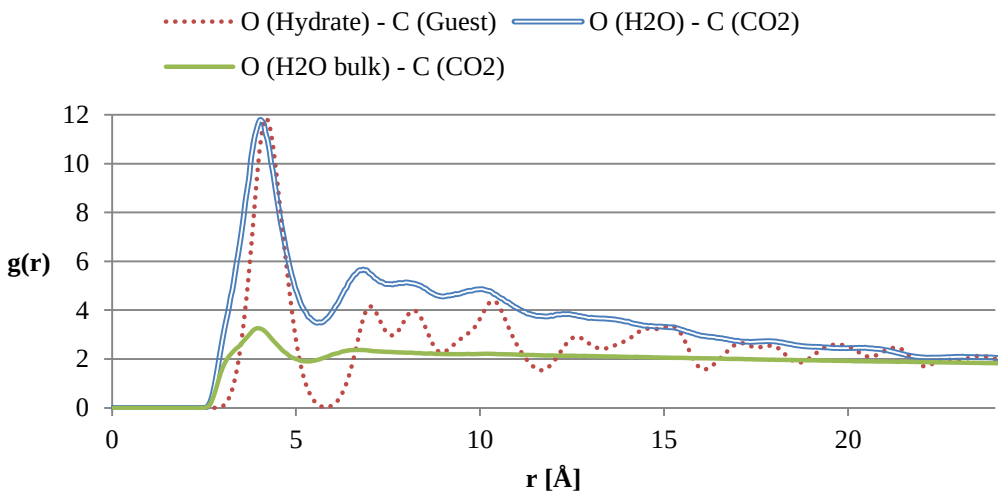

Figure 7-8. g(r) showing the structuring of water on a hydrate surface. From VMD

The structure of CO_2 adsorbed on the surface of hydrate can also be evaluated against the structure of CO_2 molecules trapped within the hydrate. The RDF of this can be seen in Figure 7-9.

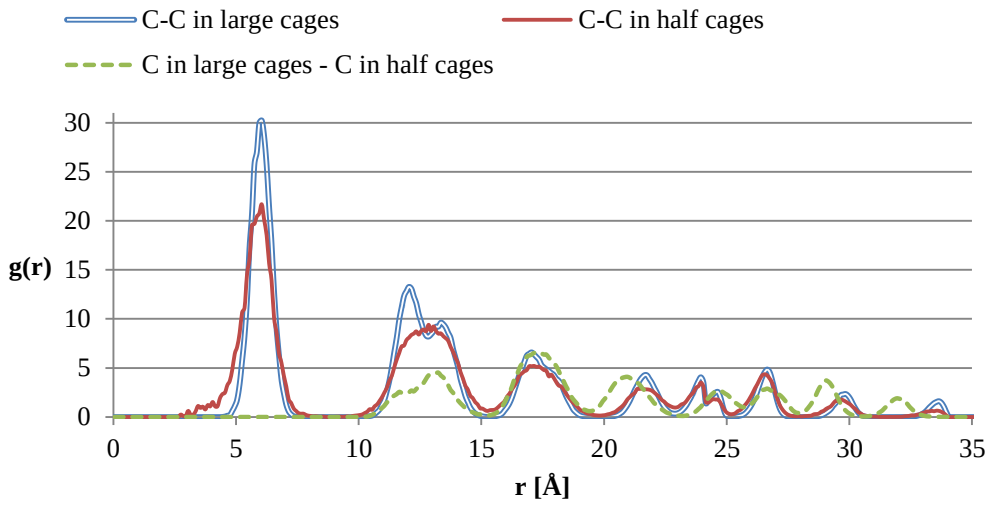

Figure 7-9. g(r) for carbon in CO_2 in and adsorbed on hydrate. From VMD

In Figure 7-9 the blue line is for CO_2 in the large cages (Figure 7-7 A), the red line is for CO_2 in the half cages (Figure 7-7 (C)). The first peak is the CO_2 molecules in the large cell behind (in depth) each other. This is confirmed by the fact that the green curve shoving the structure between CO_2 in large cages and CO_2 in half cages does not have this peak. The start of the red curve shows that the half cage structure is not as tight and perfect as that of the full cage, the CO_2 molecules are in motion and there are some that are not placed in the center of the cage as they would have been in the full cage. The next peak, at 12.5 Å, represents the distance

between the large cages as seen in Figure 7-7 (A). At this peak all the curves matches showing that the adsorbed CO_2 molecules are in an optimal placement for further hydrate growth. The third peak, at 17 Å, corresponds to the distance between two cells diagonally to the upper and lower right. A corresponding curve for oxygen in CO_2 can be found in Appendix A.

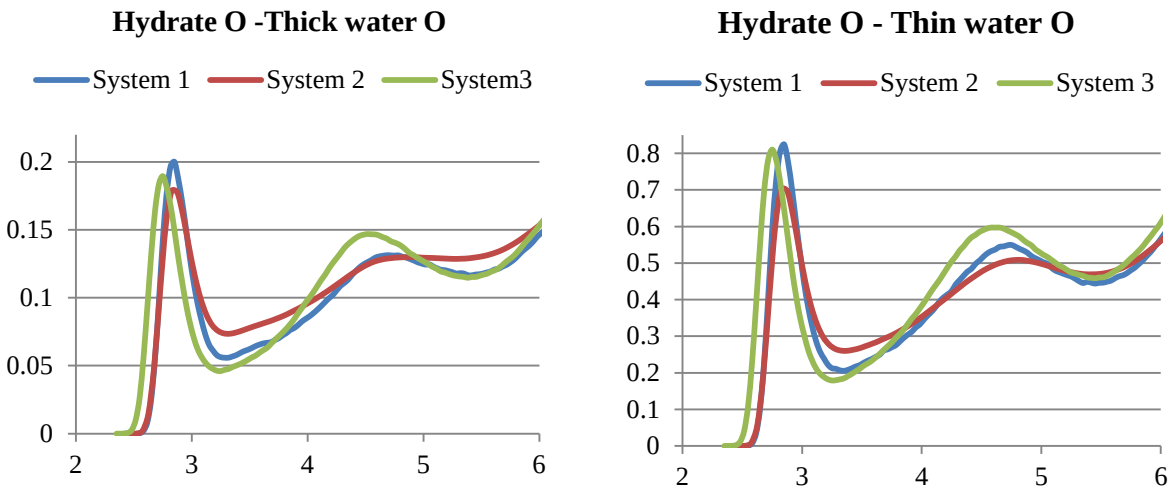

Figure 7-10. g(r) between O in hydrate H_2O and O in the two water layers. From md43.

In Figure 7-10 g(r) between O in hydrate H_2O and O in H_2O in the two water layers can be seen. Since the two layers dominate each side of the hydrate this tells us that the structure of free water on both sides are very similar and that systems 1 and 2 are in agreement that the first water molecules can be found at 2.9 Å and 2.85 Å in the thick and thin layer respectively. System 3 has placed the first peak at 2.75 Å in both water layers.

In Table 7-III the number of CO_2 molecules that have adsorbed onto the hydrate is listed. In order to compile the data basis for these numbers a TCL script was written that opened each of the .DCD files individually, selected the molecules in question, counted and stored them and then closed the .DCD file. This method was chosen because each system generated about 300 .DCD files each, and each .DCD file contains 500 trajectory steps. The .DCD files had to be opened and closed sequentially because the computer could not handle the amount of data in more than maximum 14 .DCD files at the time. For the average value the last 200 .DCD files were used. In Appendix B, graphs illustrating how the number of adsorbed CO_2 molecules changed during the simulation have been included. As expected we see from Table 7-III that the greatest numbers of CO_2 molecules are adsorbed to the front of the hydrate, the thin water layer and the strong interactions between hydrate structure and CO_2 facilitates transport easier than that the thicker water layer with the calcite crystal in it.

Table 7-III. Number of CO_2 molecules adsorbed to hydrate.

	System 1		System 2		System 3	
	Front	Back	Front	Back	Front	Back
Maximum	135	96	111	87	96	78
Average	107	71	81	64	63	36

The uneven motion seen in the start of the red graph in Figure 7-9 combined with the extra width compared to the blue graph can be explained by the variance in motion of the CO_2 molecules in the different types of cages. In Figure 7-12 (A) the trajectory of one CO_2 molecule has been mapped out, showing the motion of the molecule through time. In part (A) we see how the trajectory of the CO_2 molecule in the half cage is much looser than that in the full cage, there are some tendencies for the CO_2 molecule to leave the half cage. How the number of CO_2 molecules adsorbed in these half cages changes with time can be seen in Appendix B, Figure B-2.

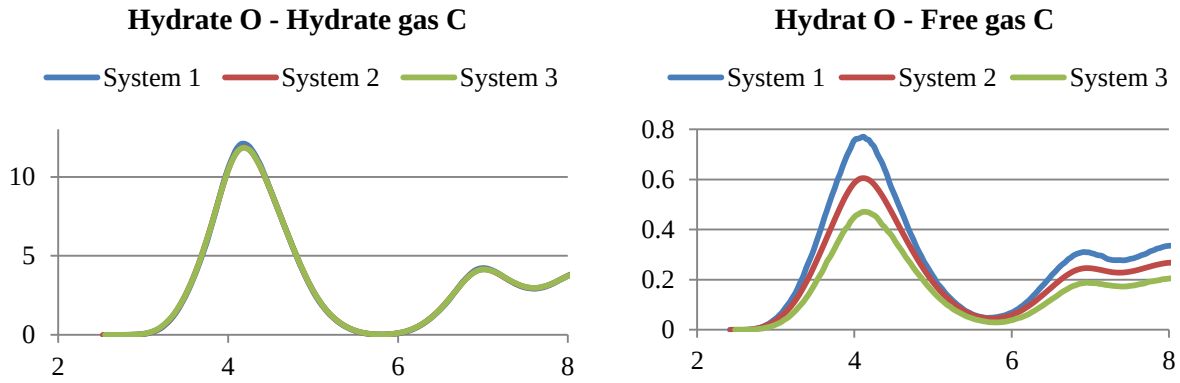

Figure 7-11 g(r) between O in hydrate H_2O and C in CO_2 for the gas in hydrate and free gas. From md43.

In Figure 7-11, g(r) between O in hydrate H_2O and C in CO_2 in both hydrate gas and free gas can be seen. First we see that all systems are in agreement about the first peak in both conditions. The difference between the two cases above is only 0.07 Å at 4.18 and 4.11 in hydrate CO_2 and free CO_2 respectively. The difference in number of adsorbed CO_2 molecules that we saw in Table 7-III are clearly visible in the peaks on the right hand side of Figure 7-11, Hydrate O – Free gas C. The dip at 5.7 Å for all systems are also in excellent agreement with the dip for CO_2 in hydrate.

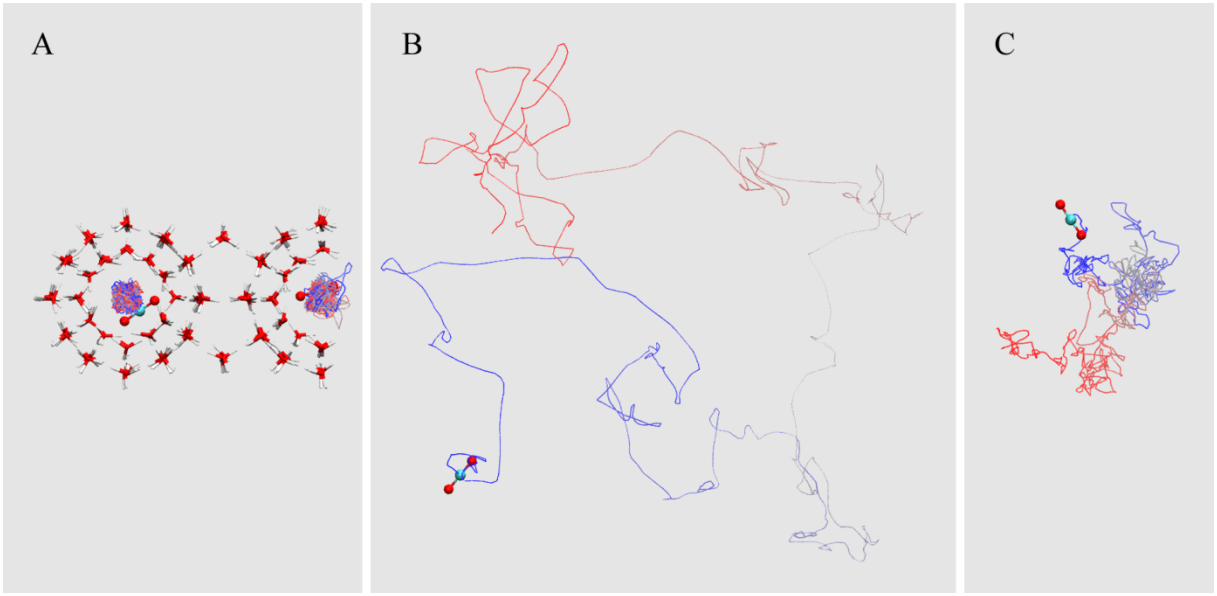

Figure 7-12. Trajectories of CO_2 molecules in different environment. (A): Trajectories for CO_2 in complete and in half cage. (B): Trajectory for CO_2 in free gas phase going between water layers at right and left side. (C): trajectory for CO_2 dissolved in water.

In Figure 7-12 (B) we see how a CO_2 molecule moves in the free gas phase. There is a lot more movement compared with the CO_2 molecule trapped in a half cage. To the left hand side of Figure 7-12 (B) the CO_2 molecule is in and on the interface between gas and thin water layer and we see that as the molecule interacts with water molecules its speed is slowed down and its path becomes more tortious. The same can be seen on the right hand side where the CO_2 molecule interacts with the large water phase, this time the lack of proximity to the hydrate phase and the larger size of the water layer diminishes the mixing and the CO_2 molecule traverses the interface on the gas side. Between the interfaces the path is largely uninhibited and the molecule traverses the volume with ease and speed. In figure (C) the CO_2 molecule is diffusing through the water phase, the path is random and speed of motion is slow.

7.5 Density profile of the system.

The density profile of the system was calculated using a Matlab script that averaged the density of a system over 500 frames from one .DCD file. Figure 7-13 shows the density profile for the whole system in the z direction. We can obviously see clear interfaces between the CO_2 and H_2O phases, as expected in the case of those two fluids (da Rocha et al. 2001). The profile makes it evident that CO_2 density is enhanced at the interface. The calcite crystal can barely be seen on the lower left in Figure 7-13, the low density is partially due to its rotation during the 500 frames. Rotation masks the clear crystal structure one would expect of such an ordered material. The crystal structure of calcite becomes obvious in Figure 7-14 where we

zoomed to facilitate a closer look at the various structural and aggregation phenomena around the crystal.

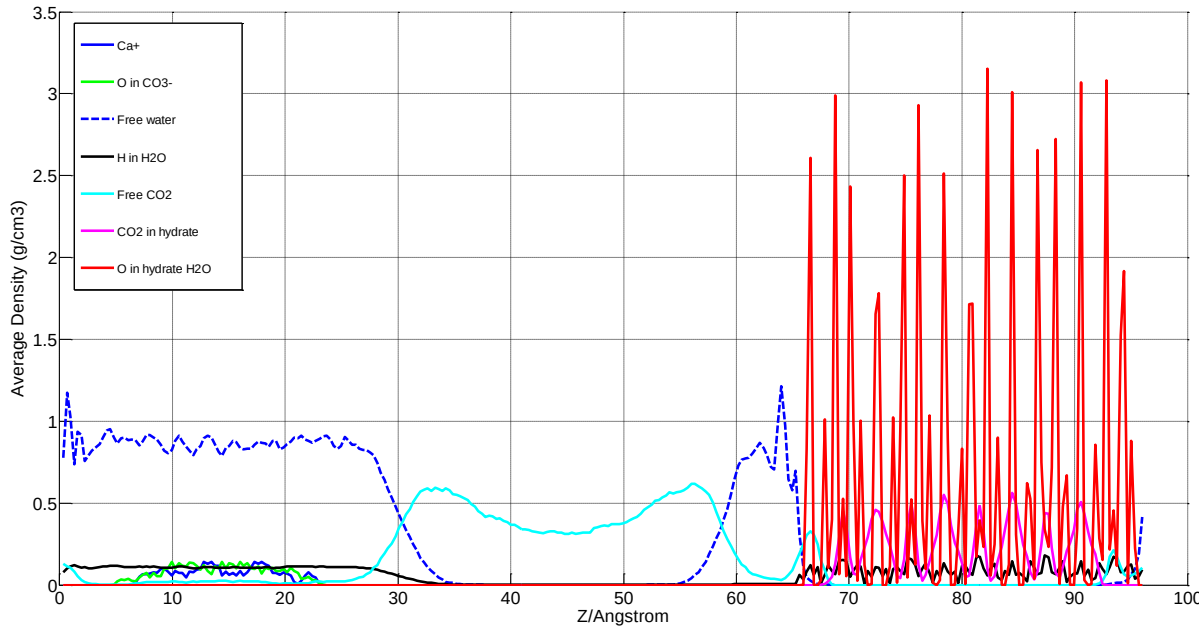

Figure 7-13 Density profile for whole system.

The calcite crystal has rotated away from the original z axis of the system in the course of the simulation; it is therefore difficult to see the clear edges of the crystal. But as Figure 7-14 below shows, the carbonate groups and calcium atoms alternate at the dominant positions, while the locations of carbon in carbonate corresponds well with the oxygen peaks, it can also be seen that the height of the carbon peaks are about one third that of the oxygen in good agreement with the calcite chemical formula. The free CO_2 density profile also varied somewhat in the calcite region; but with the concentrations in the aqueous phases being so low that no clear conclusions could be drawn except for the two pronounced dips directly tied to the hydrate-like free water peaks. In Figure 7-14 one such dip can be seen at 5 Å where the density is near zero, no such dip is apparent at the other side of the crystal and it is therefore believed that the dip is a result of the hydrate structure.

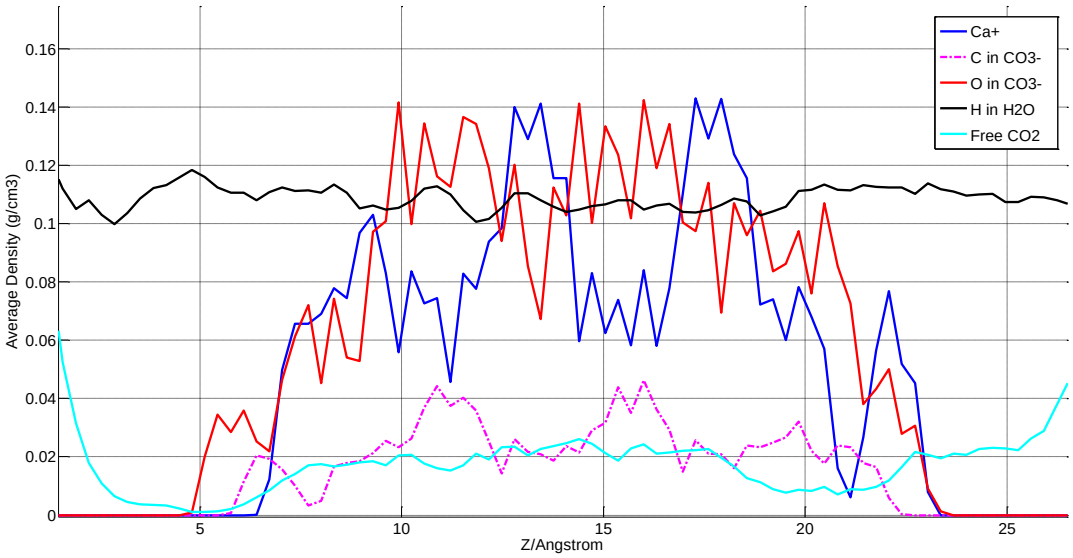

Figure 7-14. Density profile of the calcite crystal (zoomed in).

In Figure 7-15 the density around the hydrate phase can be seen. Note that the Z axis does not correspond to the previous density profiles; in order to display the free water and the free CO_2 curves the system was shifted by 37 Å in z direction. This will not affect the results in any way because of the PBC replicating the system in all directions. It is just a method that allows for better visualization of the actual structuring.

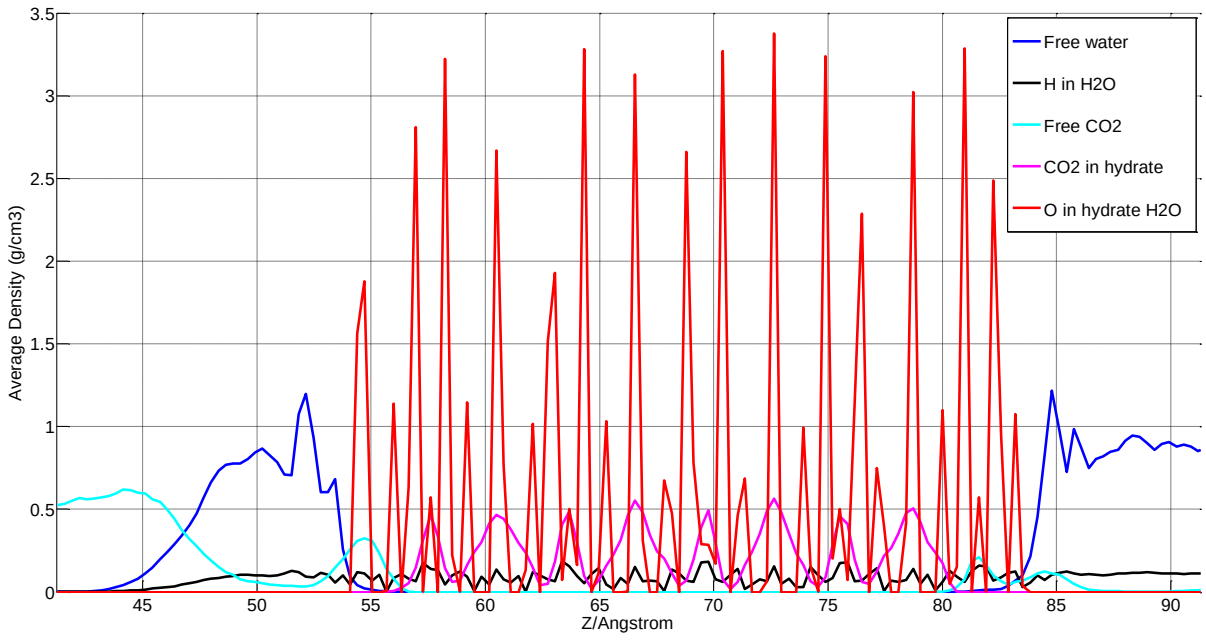

Figure 7-15. Density profile showing the hydrate and neighboring aqueous phases.

In Figure 7-15 the structuring of the free water becomes evident, between 50 and 55 Å the water is emulating the high-low-high peak pattern of hydrate at the same time there is a clear peak in CO_2 at 55 Å. This peak correlates to the free CO_2 molecules adsorbed in the half cag-

es seen in Figure 7-7 (B). The peak is close to the corresponding hydrate CO_2 peaks in both shape and magnitude, and the average number of CO_2 molecules adsorbed in this position was 78, 61 and 52 in to system 1, 2 and 3 respectively. In the hydrate structure the number of CO_2 molecules in the same position is 69. At 80-85 Å the CO_2 curve has two peaks again corresponding to Figure 7-7 (D) where we can see free CO_2 molecules adsorbed to the hydrate in two different locations. Again we see that the peaks agree well with that of hydrate CO_2 with the peak at 82 Å being pointier than that at 55 Å, bearing closer resemblance to the internal peak at 71 Å.

7.6 Intermolecular interactions

Our MD program calculated and reported the interaction energies between different types of molecules. Table 7-IV lists those of them most relevant for the purposes of our investigation averaged over 9,000,000 time steps (4.5 nanosecond). It is important to notice that when md43 calculates these values, all the molecules of a given type are used, independent of their location, so the energies below include, amongst other, bulk, interface and adsorbed energies. For some of the values we see that the uncertainty in the numbers are greater than the value, this is because of the small amount of these in close proximity.

Table 7-IV. Energy of select intermolecular interactions per water molecule, kJ/mol.

	System 1		System 2		System 3	
	Electrostatic	vdW	Electrostatic	vdW	Electrostatic	vdW
Calcite-Thick H_2O	-45.8(5)	0.771(5)	-44.8(1)	0.737(4)	-47.4(0)	1.154(5)
Calcite-Free CO_2	0.0(0)17	-0.0036	-0.04(4)	-0.003(0)	-0.00(8)	-0.0021
H_2O-H_2O in Thick layer	-30.2(2)	4.59(6)	-28.7(4)	4.14(2)	-39.7(1)	8.82(2)
Thick H_2O-Free CO_2	-0.22(5)	-0.63(5)	-0.25(7)	-0.495(6)	-0.05(9)	-0.329(9)
Free CO_2-Free CO_2	-0.9(1)	-3.11(1)	-0.7(1)	-3.13(3)	-0.7(2)	-3.33(2)
CO_2-CO_2 in Hydrate	8.4(6)	-0.636(8)	7.5(4)	-0.637(6)	7.1(8)	-0.643(1)

Just like the pair correlation functions discussed above in Section 7.4, systems 1 and 2 proved to be quite similar. System 2 had slightly less negative and thus less favorable energies than system 1, as one would expect a system at a higher temperature. System 3, at the same temperature as system 2 but utilizing a different water model, seemed to favor water over CO_2 near the calcite crystal compared with system 2. This same trend can be seen between water in the thick layer and water and free CO_2. Again, the "free CO_2" was not only comprised by dis-

solved CO_2 but also of CO_2 in the gas phase and at the interface, as well as adsorbed on hydrate. The energies presented here are therefore more complex than that of a homogenous two-component system. The free CO_2 energy is similar in systems 2 and 3. The comparison of energies in the three systems further supports the observations about water outcompeting CO_2 near the calcite. The rate of CO_2 dissolution into the aqueous phase was never enough to probe for any cumulative properties of CO_2.

8 Conclusions

Molecular dynamics was used to study the impact of calcite in a composite hydrate/water/CO_2 model system. MDynaMix package was chosen as the primary molecular dynamics software platform with VMD as the visual interpreter. The supercomputer facilities at the University of Bergen are very well suited for the simulation of this type, with the opportunity to utilize them proven to be a significant asset. The parallelization algorithms were quite efficient with almost linear upscaling for the chosen number of processors.

The resulting data has been analyzed, and an insight into interactions and adsorption patterns of water and carbon dioxide has been gained. Three systems of the same design but with different settings were studied in order to make the results as flexible as possible. These systems allowed us to study the implications for ordering and potential hydrate formation in the vicinity of calcite and hydrate surfaces.

All the systems have showed only limited variation in the density of the aqueous phase. The transfer was by far the most vigorous in System 1 held at the lowest temperature. This system had ended with 53 additional water molecules in the thick phase at the end of simulation, compared with 5 and 3 for systems 2 and 3, respectively. Systems 1 and 2 displayed a similar orientation of the calcite crystal at the end of simulation, while the calcite in system 3 rotated to align its $\{10\bar{1}4\}$ surface with the interfaces. The translational motion was similar for all three systems. The difference in the final preferred rotation of the calcite crystal suggested that it has not reached a relaxed state in the course of the simulation due to relatively limited duration and the inherently strong network of hydrogen bonds which will slow down a large crystal. Thus allowing the system to evolve further might have resulted in a different orientation. Another contributing factor might be that the distance to the interfaces on each side of the crystal was too short; a simulation setup that investigates, amongst others, this possibility is suggested in further work. Nevertheless, the crystal showed limited motion in the z and x directions suggesting that an energy minimum might lie in this region.

The analysis of pair correlation functions between calcite and the surrounding liquid showed that one layer of water was adsorbed onto the calcite crystal; this layer successfully prevented CO_2 from reaching the calcite surface. This adsorbed layer was found to have a higher density in system 3 where it shielded the interactions between Ca and O in CO_2 to a greater extent than in system 2. This shielding kept the CO_2 from assuming orientation expected from the charge distributions. The increased water layer density in system 3 is supported by the calcu-

lated potential energies between calcite and water being more favorable in system 3 compared to the other two systems.

Visual inspection has shown that ordering imposed on water near the calcite crystal will be incompatible with hydrate growth, and that it is the cleaving plane that determines how far this effect will extend. The orientation of adsorbed water molecules differed depending on the crystal face, with the electrostatic interactions between the partial charges being the decisive factor. The dominant $\{10\bar{1}4\}$ surface has showed to have the most short-range influence on water due to its layered multi-atomic structure. This surface is the most prominently found in nature, making these findings quite important. The shape of the crystal and the $\{10\bar{1}4\}$ surface has allowed us to see how the adsorbed water layer orients itself in the situation corresponding to stepped imperfections of a real calcite crystal surface.

Both visual inspection, analysis of RDFs and density profiles have highlighted the adsorption of CO_2 onto the hydrate surface and strongly hydrate-like ordering of water virtually extending the hydrate structure. CO_2 have adsorbed at different locations on both sides of the hydrate and in numbers approaching those of CO_2 inside complete hydrate cages. This suggests that our model system would facilitate further hydrate growth at these system conditions. We have also observed CO_2 molecules adsorbed on hydrate in positions corresponding to the small hydrate cages. CO_2 molecules located in these locations will be incompatible with further hydrate growth as these cages are too small to encage a CO_2 molecule without destabilizing the hydrate lattice. The appearance of these molecules is most likely a result of the much more relaxed structuring imposed by the free water. If smaller molecules able to fit into small cages were introduced to the system, they might be able to keep the larger CO_2 molecules out of these cages. Inclusion of smaller guest molecules thus has been suggested as a part of further work. Methane is one relevant candidate as this could bring further insight into the CO_2/CH_4 exchange process.

9 Suggestions for further work

Improvements in the model

Investigate the possibility of utilizing a calcite crystal where some or all of the surface atoms are allowed to be free. This might require new potential models to be implemented or a much larger crystal where the self-interactions keep the crystal together. Increased flexibility of other components in the system should also be investigated in order to make the results more accurate. For instance Zhu et al. (2009) has developed a CO_2 model with focus on full flexibility that has been compared favorably with experimental research at conditions similar to this thesis (Zhu et al. 2009) .

This system has contained only pure components, in order to make the results more versatile the inclusion of byproducts of chemical reactions and reservoir fluids should be added. As mentioned in Chapter 3 the addition of CO_2 to a calcite reservoir would result in change of pH and precipitation of Ca^{2+} and HCO_3^- if the pH goes above 5-6. The ions H^+, HCO_3^- and maybe CO_3^{2-} should be included, although CO_3^{2-} is not relevant under normal storage conditions. The use of software that could model these reactions should be considered to find out if they are able to handle the size and complexity of a relevant model system. The addition of salts to the water in order to emulate the brackish water found in reservoirs and many aquifers would also bring the model system closer to reality.

Though the variations in crystal shape and color are large, as mentioned in Chapter 3, this thesis used a perfect calcite crystal. These variations stem from impurities, mostly metal ions included in the crystal lattice structure. The work of Lorens (1981) showed that uptake of cobalt (Co), manganese (Mn) and cadmium (Cd) was rapid and exhausted supply available in the experimental solution. This was suggested to be a result of physical adsorption on active sites and the precipitation with Ca. The addition of such metal ions would alter the surface properties of the calcite crystal. It should be investigated to determine to which extent this will affect the water structure and subsequent hydrate formation potential.

As shown in Section 7.4, the CO_2 molecules adsorbed in small half cages are incompatible with the further growth of hydrate. In order to investigate this effect further, other guest molecules should be added into the mix. These guest molecules should be of a size suitable for small cavities found in hydrate structure I. The addition of such molecules could tell us if

there is competition with the larger CO_2 molecule for these small half cages or if they are simply filled by the small molecules. This could further enhance the water structuring effect shown in this thesis. Methane is the most abundant guest molecule found in nature and thus will be one such guest molecule that could also provide insight into the exchange process of CO_2/CH_4. Replacing the less stable methane hydrate with CO_2 in hydrate reservoirs will result in increased stability of the reservoir. In addition, there will be an environmental gain from sequestrating the CO_2, as well as economic advantages since as methane is a relatively clean energy source.

System set-up

To gain further insight into interactions between calcite/water/CO_2 a system where the hydrate phase has been removed is suggested. This would decrease the number of potential contributions to interactions, and reduce the number of uncertainties. This would also allow the model to contain larger water and CO_2 blocks. And allow for bulk volumes of both phases to be present. The system should be started with the calcite crystal in different locations, one with the crystal immersed in water and one in CO_2 fluid. The collective effects of the surrounding media might change the behavior of the system. Due to the limitations in simulation time sufficient CO_2 may not be able to gather near each other in the water phase in order to show collective properties toward the calcite crystal. If the calcite started out in the gas phase it would not be necessary for the simulations to run for a long time. Due to the difference in water structuring for the surfaces of the calcite crystal the effects of different orientations of the crystal in the two phases should be considered.

As further efforts in investigations of the minimum stability distance between calcite and hydrate at which the hydrate it is proposed to set up a system well within pressure and temperature stability of hydrate but with varying distance between hydrate crystal and calcite. The total system will consist of a hydrate crystal similar to the system used in this study – embedded in liquid water with a calcite crystal. Initial location of the calcite crystal will be varied from close to far and the $\{10\bar{1}4\}$ surface will be suggested. In addition to initial differences in calcite/hydrate distance all systems will of course be set to motion and run in MD. The stability area for the water model used should be evaluated and attempted verified against experimental data.

In order to investigate how a pressure gradient in the gas phase affects the structuring effect on water from both hydrate and calcite another system of two parallel surfaces are suggested. A large stationary calcite surface facing a hydrate surface with water in between, the distance in between should correspond to the combined lengths of structure in water imposed by both surfaces. In order to avoid interactions between interfaces a minimum distance of 5 nanometers is suggested. This system should be allowed to run for a period of time in order to establish the surface layers above each phase. After this initial equilibration run a CO_2 bulk at high pressure should be introduced in one end of the two layers. Some manner of controlling the overall pressure of the system should be used, possible by a less dense area being introduced at the opposite end to the CO_2. This system would allow us to study how the water structures is maintained with a flux of CO_2 passing and pushing by. This system would be physically similar to reservoir conditions where the hydrate and calcite mineral would be facing each other separated by a mobile water layer. The CO_2 bulk would simulate a high pressure injection point.

Software

A rewrite of the software in order to include a zero order approximation for molecular correlation functions would aid in extracting valuable information from the molecular dynamics simulations being run using molecular dynamics. This could be reconstructed from the atom to atom correlation functions and intra molecular structures using the superposition principle (Kvamme, 1995, 2002).

References

Allen, M. P. and D. J. Tildesley (1989). Computer simulation of liquids. Oxford, Clarendon Press.

Andersen, H. C. (1983). "Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations." Journal of Computational Physics **52**(1): 24-34.

Berendsen, H. J. C., J. R. Grigera, et al. (1987). "The Missing Term in Effective Pair Potentials." Journal of Physical Chemistry **91**(24): 6269-6271.

Cramer, C. J. (2004). Essentials of computational chemistry theories and models. Chichester, West Sussex, England Hoboken, NJ, Wiley.

da Rocha, S. R. P., K. P. Johnston, et al. (2001). "Molecular Structure of the Water−Supercritical CO2 Interface." The Journal of Physical Chemistry B **105**(48): 12092-12104.

Demurov, A., R. Radhakrishnan, et al. (2002). "Computations of diffusivities in ice and CO2 clathrate hydrates via molecular dynamics and Monte Carlo simulations." Journal of Chemical Physics **116**(2): 702-709.

Dill, K. A. and S. Bromberg (2003). Molecular driving forces : statistical thermodynamics in chemistry and biology. New York, Garland Science.

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

Espinoza, D. N. and J. C. Santamarina (2010). "Water-CO2-mineral systems: Interfacial tension, contact angle, and diffusion-Implications to CO2 geological storage." Water Resources Research **46**: -.

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

Førrisdahl, O. K. (2002). Computer simulations of natural gas hydrates : equilibrium, melting, inhibition and free energy calculations. [Bergen], Programme of Process Technology Department of Physics University of Bergen.

Graf, D. L. (1961). "Crystallographic tables for the rhombohedral carbonates." The american mineralogist. **46**(November-December).

Harris, J. G. and K. H. Yung (1995). "Carbon Dioxides Liquid-Vapor Coexistence Curve and Critical Properties as Predicted by a Simple Molecular-Model." Journal of Physical Chemistry **99**(31): 12021-12024.

Hoover, W. G. (1985). "Canonical Dynamics - Equilibrium Phase-Space Distributions." Physical Review A **31**(3): 1695-1697.

Humphrey, W., A. Dalke, et al. (1996). "VMD: Visual molecular dynamics." Journal of Molecular Graphics **14**(1): 33-&.

Hwang, S., M. Blanco, et al. (2001). "Atomistic simulations of corrosion inhibitors adsorbed on calcite surfaces I. Force field parameters for calcite." Journal of Physical Chemistry B **105**(44): 10746-10752.

ISAACS. "Interactive structure of amorphous and crystaline systems. Central Michigan University. Authors: Dr. Sébastien Le Roux, Dr. Valeri Petkov ". Retrieved 20.1.2011, from http://isaacs.sourceforge.net/phys/pbc.html.

Jensen, F. (2007). Introduction to computational chemistry. Chichester, Wiley.

Kuznetsova, T. (2001). Molecular modeling for thermodynamic properties of bulk and interfacial systems. [Bergen], University of Bergen Department of Physics.

Kvamme, B. (1995). "Interaction-Site Representation of Polar Mixtures and Electrolyte-Solutions." International Journal of Thermophysics **16**(3): 743-750.

Kvamme, B. (2002). "Thermodynamic properties and dielectric constants in water/methanol mixtures by integral equation theory and molecular dynamics simulations." Physical Chemistry Chemical Physics **4**(6): 942-948.

Kvamme, B. and O. K. Forrisdahl (1993). "Polar Guest-Molecules in Natural-Gas Hydrates - Effects of Polarity and Guest-Guest-Interactions on the Langmuir-Constants." Fluid Phase Equilibria **83**: 427-435.

Kvamme, B., A. Graue, et al. (2007). "Storage of CO2 in natural gas hydrate reservoirs and the effect of hydrate as an extra sealing in cold aquifers." International Journal of Greenhouse Gas Control **1**(2): 236-246.

Kvamme, B., T. Kuznetsova, et al. (2009). "Modelling excess surface energy in dry and wetted calcite systems." Journal of Mathematical Chemistry **46**(3): 756-762.

Levitt, M., M. Hirshberg, et al. (1997). "Calibration and testing of a water model for simulation of the molecular dynamics of proteins and nucleic acids in solution." Journal of Physical Chemistry B **101**(25): 5051-5061.

Lin, S. T., P. K. Maiti, et al. (2010). "Two-Phase Thermodynamic Model for Efficient and Accurate Absolute Entropy of Water from Molecular Dynamics Simulations." Journal of Physical Chemistry B **114**(24): 8191-8198.

Loveday, J. S., R. J. Nelmes, et al. (2001). "Stable methane hydrate above 2 GPa and the source of Titan's atmospheric methane." Nature **410**(6829): 661-663.

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

Mark, P. and L. Nilsson (2001). "Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K." Journal of Physical Chemistry B **105**(43): 24a-24a.

MDynaMix. "(2010). MDynamix hompage,", " Authors: Lyubartsev, A.P. and Laaksonen A.". from http://www.mmk.su.se/~sasha/mdynamix/.

Nooner, S. L., O. Eiken, et al. (2007). "Constraints on the in situ density of CO2 within the Utsira formation from time-lapse seafloor gravity measurements." International Journal of Greenhouse Gas Control **1**(2): 198-214.

Nosè, S. (1984). "A Molecular-Dynamics Method for Simulations in the Canonical Ensemble." Molecular Physics **52**(2): 255-268.

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

Oelkers, E. H., S. R. Gislason, et al. (2008). "Mineral Carbonation of CO2." Elements **4**(5): 333-337.

Palmer, D. "(2010). CrystalMaker Software Limited, Oxford, England . Version 2.2.3. http://www.crystalmaker.com/."

Parker, S. C., J. O. Titiloye, et al. (1993). "Molecular Modeling of Carbonate Minerals - Studies of Growth and Morphology." Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences **344**(1670): 37-48.

Perdikouri, C., C. V. Putnis, et al. (2009). "An Atomic Force Microscopy Study of the Growth of a Calcite Surface as a Function of Calcium/Total Carbonate Concentration Ratio in Solution at Constant Supersaturation." Crystal Growth & Design **9**(10): 4344-4350.

Pickering, P. F., B. Edmonds, et al. "Evaluating New Chemicals And Alternatives For Mitigating Hydrates." http://www.feesa.net/pdf/DiscussionPapers/010918-EvaluatingHydrateInhibitors-RevA.pdf.

Prausnitz, J. M., R. N. Lichtenthaler, et al. (1999). Molecular thermodynamics of fluid-phase equilibria. Upper Saddle River, N.J., Prentice-Hall PTR.

Quinlan, G. D. and S. Tremaine (1992). "On the realiability of gravitational N-body integrations " Monthly Notices of the Royal Astronomical Society **259**(3): 505-518.

Rowley, R. L. and T. Pakkanen (1999). "Determination of a methane intermolecular potential model for use in molecular simulations from ab initio calculations." Journal of Chemical Physics **110**(7): 3368-3377.

Ryckaert, J. P., G. Ciccotti, et al. (1977). "NUMERICAL-INTEGRATION OF CARTESIAN EQUATIONS OF MOTION OF A SYSTEM WITH CONSTRAINTS - MOLECULAR-DYNAMICS OF N-ALKANES." Journal of Computational Physics **23**(3): 327-341.

Seastar2. "Scalable computing: Why and how. White paper by HyperTransport Consortium. A non-profit organisation responsible for managing the open HyperTransport standard. Authors: J. Duato, F. Silla. B. Holden, P. Miranda, J. Underhill, M. Cavalli, S. Yalamanchili, U. Brüning, H. Fröning. ." Retrieved 25.04.2011, from http://www.hypertransport.org/docs/uploads/HNC_WP_33976512.pdf.

Shiraki, R. and S. L. Brantley (1995). "Kinetics of near-Equilibrium Calcite Precipitation at 100-Degrees-C - an Evaluation of Elementary Reaction-Based and Affinity-Based Rate Laws." Geochimica Et Cosmochimica Acta **59**(8): 1457-1471.

Sloan, E. D. (1998). Clathrate hydrates of natural gases. New York, Marcel Dekker.

Susilo, R., S. Alavi, et al. (2008). "Interactions between structure H hydrate formers and water molecules." Journal of Physical Chemistry C **112**(24): 9106-9113.

Udachin, K. A., C. I. Ratcliffe, et al. (2001). "A dense and efficient clathrate hydrate structure with unusual cages." Angewandte Chemie-International Edition **40**(7): 1303-+.

Verlet, L. (1967). "Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules." Physical Review **159**(1): 98.

Xu, T., J. A. Apps, et al. (2002). Reactive geochemical transport simulation to study mineral trapping for CO2 disposal in deep saline arenaceous aquifers.

Zhang, H. and S. J. Singer (2011). "Analysis of the Subcritical Carbon Dioxide−Water Interface." The Journal of Physical Chemistry A: null-null.

Zhu, A. M., X. B. Zhang, et al. (2009). "A Fully Flexible Potential Model for Carbon Dioxide." Chinese Journal of Chemical Engineering **17**(2): 268-272.

Appendix A Supplementary RDFs

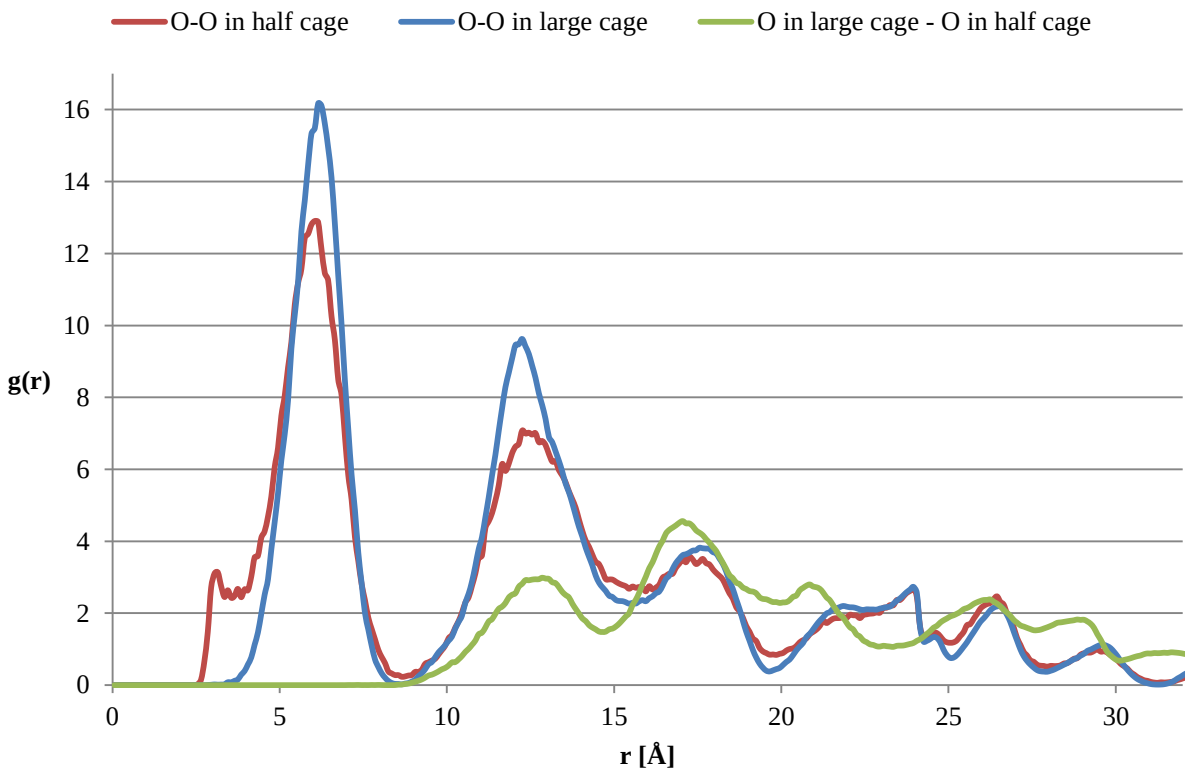

Figure A-1. g(r) for oxygen in CO2 in and adsorbed on hydrate.

Appendix B Number of adsorbed CO_2 molecules

The following graphs show the number of CO2 molecules adsorbed to the hydrate surface at different positions. The graphs have been generated by a computer script that checks all the .DCD files from the simulations. The .DCD files contain the positions of all the molecules. The computer script counts how many atoms of a specified type are present in regions, these regions are specified by the user. Here the number of molecules adsorbed in half cages has been counted approximately every 0.5 picoseconds.

In the figures below the shape of the curves corresponds well to each other, as expected there is less motion in the system with lowest temperature (system 1). While system 2 and 3 follow each other in magnitude, they are not more alike due to the difference in water model.

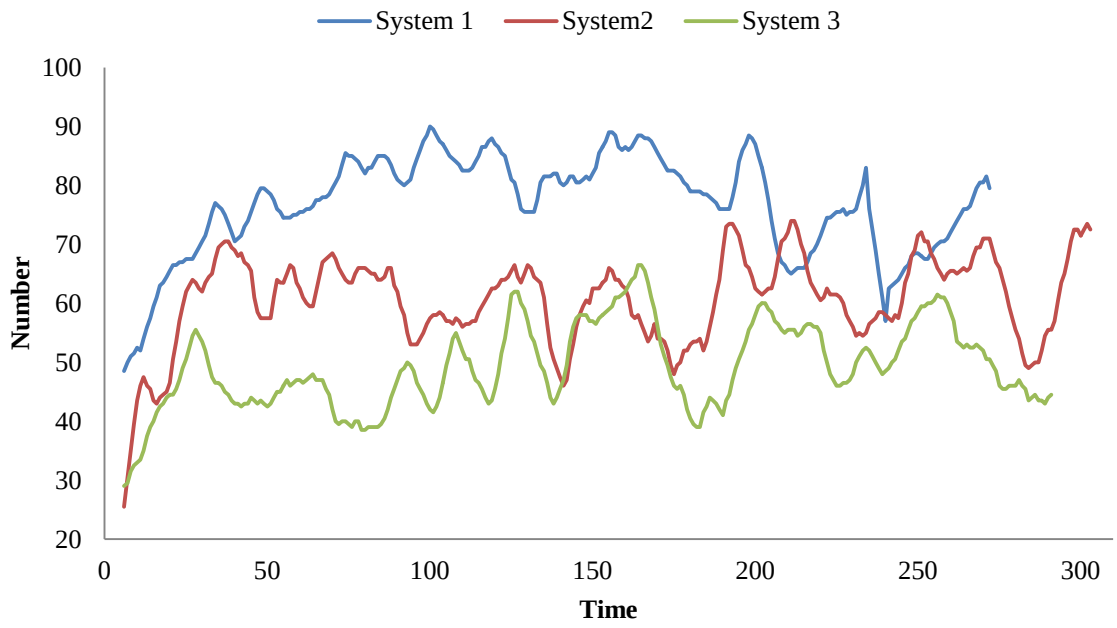

Figure B-1. Number of CO2 molecules adsorbed in half cages on front of the hydrate.

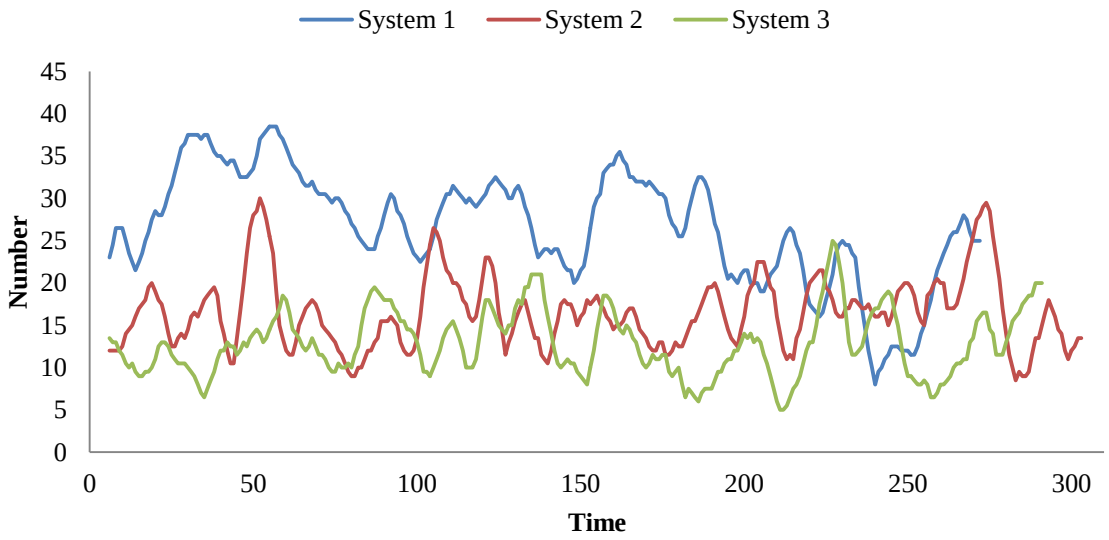

Figure B-2. Number of CO2 molecules adsorbed between half cages on front of hydrate.

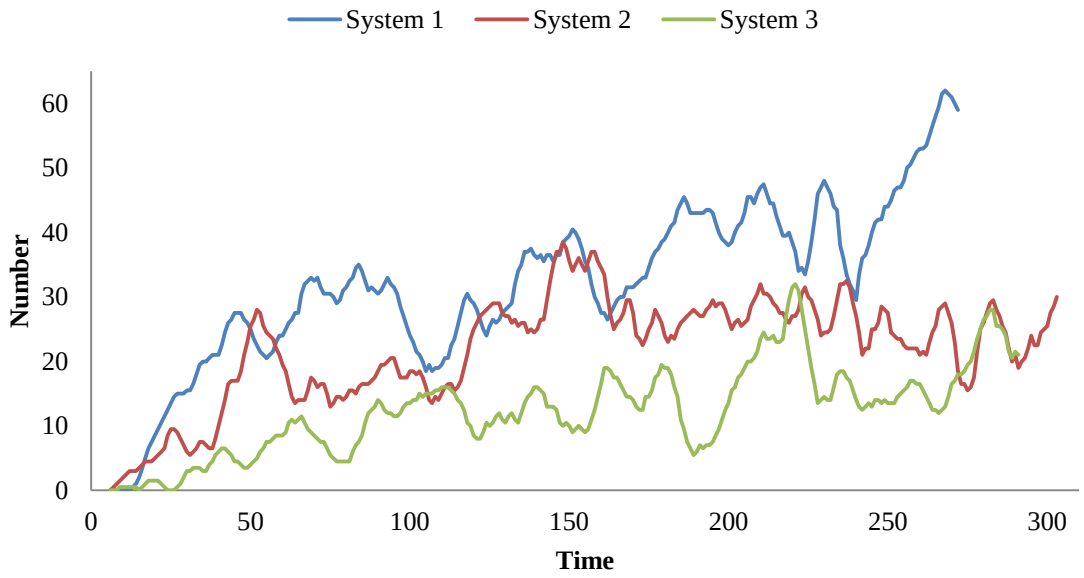

Figure B-3.Number of CO2 molecules adsorbed in half cages on backside of hydrate.

Appendix C Example of MD.input file

The following is the input file used for the simulation of system 3 in md43. Here lines marked with # are comments meant to explain the following choices/settings.

```
#   Sample input file for the MD program, v.>=4.3
#   Simulated system consists of 108 rigid CO2 molecules,
#
#   Lines beginning with "#" are commentaries
#   In this file the commentaries go before the corresponding parameters.
#
#   This file can be directly used as input (see directory sample for
#   short version of the input file)
#
#   Output control parameter
#   Suitable valus 2-10. The less number, the less you see in the output
#   parameters higher 7 used mostly for debug purposes
  5
#
# Full printout file name
  m43_CO2+waterSlice+mixedHydrt
#   Base file name for output files:
#   Other files requested or created by the program have this name with
#   various extensions
  Hyd+CO2313K.H2O_E
#
#   Path to the molecular database
#   This directory contains *.mol files wich describe the molecular
#   structure and the force field
  ../moldb
#
# The program creates and updates periodically a restart file which
# contains configuration of the system and calculated averages.
# The program can be interrupted and then continued from the restart file
# without loosing any information.
#  If "Check only" parameter is true, the program does not run
# simulation. If it is a new run, the program only checks the input.
# If it is continuation of the old run, the program gives calculated results.
#
#  Read from        Dump              Check          Zero counter
# restart file?  restart file?        only?          of cpu time?
     .t.            .t.               .f.              .f.
#
#  The type of statistical ensemble. "Anizotropic NPT" means separate
#  pressure/volume controll in each direction. Use this option only if the
#  system is really anizotropic (a piece of DNA, membrane, liquid crystall,
#  etc). Note, that for "constant pressure"=.t. "constant temperature"
#  must be also .true. (who knows what NVP ensemble is???).
#  Constant temperature?   Constant pressure?    Anizotropic NPT?
```

```
      .t.                .f.                .f.
#
#   Number of molecule types:
  6
#
#   which molecules:
#  The database directory (specified above) should contain files
#  H2O.mol, Na+.mol,  Cl-.mol
#  See README file about format of .mol files
  CalcCryst   H2O_E   CO2_epm2   H2O_E   H2O_E   CO2_epm2
#
#   For each molecule:
#   Number of molecules:
  1          2300       494       500      1839       192
#
#    non-bonded intramolecular interactions:
#   Calculate intramolecular potential (LJ and electrostatic)
#   for non-bound atoms, separated by
#   more than 3 covalent bonds. Normally, it should be .true.
#   unimportant for small molecules)
  .f.          .f.          .f.       .t.   .f.   .f.
#
#    1 - 4 intramolecular interactions:
#   Calculate LJ and electrostatic terms for 1-4 (i.e. separated by
#   3 covalent bonds) intramolecular interactions
  .f.          .f.        .f.       .t.       .f.   .f.
#   Scaling factors for 1-4 LJ interactions
#   These are 0 in AMBER, 1 in CHARMM and 0.25 in GROMOS
#   Not important for small molecules
   1.          1.       1.          1.              1.   1.
#   Scaling factor for 1-4 electrostatic interactions
   1.          1.       1.          1.              1.   1.
#
#    intramolecular potential type
#    Now it should be 0 for all molecules except the water
#    For water 1 is "harmonic" and 2 "anharmonic" flexible SPC water
#    (It tells the program to use a special subroutine)
  0          0          0       0          0     0
#
#   Rules for initial box size / density:
#    - If one of the box sizes is zero, the actual box size (cubic box)
#   is defined by the density.
#    - If the density is also zero, the program run "vaccum
#   simulations". Set Ewald parameters (below) to 0 in the case of
#   vaccum simulations.
#    - If all the three box sizes are not zero, they define initial box size
#   and shape, and so the actual density
#
#   Cell type:
#   0 - rectangular
```

```
#   1 - truncated octahedron: cube of side BOXL centred in 0,0,0 with
#         truncated corners:      |x|+|y|+|z| < 0.75*BOXL
#   2 - hexagonal along Z axis
#
#   temperature(K)     density(g/cm**3)       pressure (atm)
    313.15               0.99982                 1.
#   box size (A)                                    cell type
#         0                  0                 0                          0
       48.12                48.12             96.0                       0
#
#  This is the long time step
#  Time step (s)      Small steps in one long
    0.5d-15              10
#
# Total (long)   Steps for interme-      take averages     dump restart file
#  MD-steps      diate averaging         each .. steps      after .. steps
     1780800              5000                50              5000
#
# Nose-Hoover thermostat parameters:              Meaningful in constant-energy
#                                  simulations (if const.temp.=.f.):
#!!!ADDED: QTR after QT
#  T-termostat param(fs)  P-termostat(fs)   Simple velocity  delta T (K)     Intern.T-termostat
(fs)
#                                  scaling?
    100.0        100.0          5000.            .f.           5.       .f.      50.0
#
#  Rcutoff sets cut off radius for LJ and Real-space electrostatic forces
#  Interaction inside Rcut-fast are recalculated each short time step,
#  interactions between Rcut-fast and Rcutoff each long time step
#
#  Rcutoff(A)      Rcut-fast forces      check neigbours after .. steps
    10.0             5.0                    10
#=========================================
#   Treatment of electrostatic interactions:
#
#   Ewald parameters:
#   alpha and fexp are set from the conditions:
#  erfc(alpha*R)=required precision of the real-space Ewald
#  exp(-fexp)=required precision of reciprocal-space Ewald
#    the rule of thumb:
#    alpha*R               fexp (m/s**2)
     3.14159256           9.81
#   If alpha*R above is negative, the reaction field method is used with
#   -alpha*R as dielectric permitivity, fexp is Debay screening length in Å
#   (setting Debay length to 0 means infinite Debue length, i.e non-conducting
#   solution)
#   If alpha is zero (exactly: between -1 and 0), no special treatment of
#   electrostatic interactions (simple spherical cutoff)
#
#   Be careful by playing with these parameters and understand what are you
```

```
#   doing. Unproper setting can result in funny behavior of the system or too
#   long computation time.
#
#=============================================
#   which molecules move:
#  if .f., the molecules are fixed, but interact with other molecules.
    .t.       .t.       .t.            .t.           .t.    .t.
#   which molecules have translational DoFs:
#  if .f. but LMOVE=.t., the molecules can still rotate
    .t.       .t.       .t.            .t.           .f.    .t.
#   which folecutes have rotational DoFs:
#  if LMOVE=.f. and LMOVETR=.f., the molecules should be fixed, but interact with other
molecules.
    .t.       .t.       .t.            .t.           .t.    .t.
#
#   recalculate list of intramolec. interactions?
#   may be .f., if you have a large molecule with stable conformation
#       Modified cross-inter?
    .t.     .f.
#
#  If true, constrain dynamics with SHAKE algorithm for specified
#  molecular specii will be used. It will keep all the bond lengths constant.
#  No double time step alogorithm in this case
#                                        which molecules considered as rigid (1/0)
#  Constrain dynamics?     tolerance parameter    (does not matter if Constrained dynamics =
.f.)
    .t.                  0.1d-4               1  1  1  1  1  1  1  1  1
# !!!! ADDED
#  Quaternion treatment? (rigid molecules with motion divided into COM-motion and rotation
around COMs)
#  SHOULD BE TRUE ONLY IF CONSTRAINT DYNAMICS IS SPECIDIED AS WELL
#         Exp-6 vdW as well?
 .t.    .t.    .t.   .f.   .f.  .f.          .f.   .f.
# Final positions to pdb-format file?
#        Pressure control for sum of X- and Y- pressure components?
#           Box volume corrections in Z-direction only?
#                       TI over PP?  All molecular species?
#                            All molecular species?
#                                 Which type do the calcs for?
#                                      K_ONE   K_SIX  K_TWELVE  ALAMBDA
     .t.      .t.     .f.   .f.     .f.      2      2       2       4      0.95309d0
#
#  Initial state (from -1 to 4)
#  Values:
#  -1  take initial center-of-mass coordinates of molecules from *.inp file
#  0  take initial atom coordinates from *.inp file
# /*    .inp file should be written in free  x y z format (one atom per line)
#       The order of atoms and molecules:
#       molecular type1                    molecular type 2
#       mol1    mol2    mol3                 mol1    mol2    mol3
```

```
#      at1 at2 at1 at2 at1 at2              at1 at2 at1 at2    ...   */
#  1  start from FCC lattice
#  2  set a cylindrical hole along z-axis. Put molecules with parameter 1
#     specified on the line below into the hole according initial
#     coordinates in .mol file, and distribute others out the hole
#  3  the same for spherical hole
#  4  start from cubic lattice
#                          |           set velocities to 0?
#                          |           (otherwise, Maxwel distribution,
#                          |           or as it is at restart)
    -1                                      .f.
#
#  Parameters for above in cases 2,3 (specify 0 or 1)
    1       0       0     0   0    0
#
#                          If this parameter .t., the atoms described in
#                          file "fixed.atoms" will be
#                          put in a harmonic potential of this radius:
#  Radius of a hole           Keep initial          Permissible     File
#  for large molecules         configuration?        deviation?      name
     10.                         .f.                  3.       fixed.atoms
#
#  If you want to change temparature or density after restart.
#  Change temperature at restart?      Change density?      Scale in z-direction only?
    .f.                               .f.                .t.
#
#  When to start final averaging
#  Final averaging after ..
#  intermediate averaging            XMOL config. file?
     0200                              .f.
#
#  This have a sence if you have to start from a very bad configuration
#  If total force on an atom exceed some level defined by the given
#  parameter, the force will be cut to this level in the internal units
#  Cut large forces?         Parameter
    .t.                     1.d-4
#
#  RDFs have a separate restart file with extension .rdf
#  Calculate RDF?      Dump RDF?     Read restart RDF file?
    .t.                 .t.           .t.
#
#  RDF for all sites?
#  (if .f., they should
#  be specified below)    Cutoff-RDF(A)       Resolution of RDFs
    .f.                    15.                  300
#
#  dump trajectory (0/1/2)?                number of config. in a trajectory
#  0 - no trajectory                       file. Trajectory files have extensions
#  1 - unformatted files                      .001, .002, ...
#  2 - "XYZ" format           interval(s)
```

```
    1                    1.d-13          500
#  dump trajectory of molecules (1/0):
    1    1        1        1    1    1
#
#   TCF calculations
#   ----------------
#  Attention!  Tcf calculations are NOT parallelized and may essentially
#  slow down parallel simulations
#
#  TCFs have a separate restart file with extension .tcf
#  NSTEG is number of points for calculation TCF
#  JUMP - number of MD steps between the points for calculation of tcf.
# calculate tcf?   restart tcf?   dump tsf?    NSTEG      JUMP
    .f.          .f.          .f.          200        5
#
#  which of 12 tcf calculate (0/1/2):
#  12 types:
#  1 - velocity autocorellations
#  2 - angular velocity autocorellations
#  3 - 1 order Legendre polynom for dipole moment
#  4 - 2 order Legendre polynom for dipole moment
#  5 - 1 order Legendre polynom for reorintational tcf defined by vector below
#  6 - 2 order Legendre polynom for reorintational tcf defined by vector below
#  7 - X projection of velocity TCF
#  8 - Y projection of velocity TCF
#  9 - Z projection of velocity TCF
#  10 - X projection of angular velocity TCF
#  11 - Y projection of angular velocity TCF
#  12 - Z projection of angular velocity TCF
#  if "2" specified for tcf 7-9 or 10-12, then tcf projections are
#  calculated in molecular principal coordinate system; otherwize
#  they calculated in laboratorial coordinate system
 1 1 1 1 1 1 2 2 2 0 0 0
#
#   unit vectors for reorientational tcf
#   this vector is defined by 2 selected atoms on each molecule
  1       1    1    1   1   1
  1       2    1    1     1   1
#
#  Other optional parameters
#  if line "add <n>" is omitted, no optional parameters
#  Description of optional parameters is given in file Extra_param
#
add       0
#  Total number of RDF-s
57
#  RDFs are defined by "global" site number
#  First molecule type H2O have 3 sites with number 1,2,3
#  Second molecule Na has one site with number 4
#  and third molecule Cl has one site with number 5
```

```
#   C - C
#  This means RDF between sites 1 and 1, i.e. O atoms on H2O molecule
#   1  1
#   C - O
#  symbol & followed by a number means that these RDF will be averaged
#  (in this case: two hydrogens atoms (atoms 2 and 3 in H2O molecule)
#   are equivalent)
#
#O-O Vann1 1
226 226
#H-H Vann1
&3
227 227
228 228
227 228
#O-H Vann1
&2
226 227
226 228
```

Rest of RDF list has been cut to conserve space.