Computational Modeling of Biological Membrane and Interface Dynamics

Erik Lindahl

Stockholm 2001
Doctoral Dissertation
Royal Institute of Technology
Department of Physics
Computational Modeling of Biological Membrane and Interface Dynamics

Erik Lindahl

Stockholm 2001

Doctoral Dissertation
Royal Institute of Technology
Department of Physics
Akademisk avhandling som med tillstånd av Kungliga Tekniska Högskolan i Stockholm framlägges till offentlig granskning för avläggande av teknologie doktorsexamen fredagen den 18 maj 2001 kl 10.00 i Kollegiesalen, Administrationsbyggnaden, Kungliga Tekniska Högskolan, Valhallavägen 79, Stockholm.

Faculty opponent: H. Larry Scott, Illinois Institute of Technology, Chicago.

Cover illustration: A snapshot of phospholipid molecules aggregating into a membrane, using data from paper 5. There is a metastable water pore present across the bilayer.

ISBN 91-7283-088-3
TRITA-FYS-8055
ISSN 0280-316X
ISRN KTH/FYS/TEO/R--01/5--SE

© Erik Lindahl, 2001

Printed by Universitetsservice US AB, Stockholm 2001
Abstract

The exponentially increasing power of hardware and rapid advances in numerical algorithms during the last few decades have made it possible to use computer simulations to study structure and dynamics in biomolecular systems.

This thesis presents my research on biological interface dynamics, in particular lipid membranes. The studies were made possible by the development of high performance software for parallel molecular dynamics, resulting in some of the largest and longest biomolecular simulations performed to date. Calculated properties are put in relation to experimental observations, and further used to develop better theoretical models for the underlying collective dynamics. The main scientific achievements are:

- The lateral and normal solvent motions close to macromolecular surfaces are separated into local diffusive events on picosecond scale, and slow dynamics governed by a reduced mobility due to the potential of mean force from the surface.
- For the first time, undulatory & peristaltic deformations are observed in atomic detail simulations of bilayers, comprising systems of 120,000 atoms and linear scales of 20 nm. This makes it possible to calculate mesoscopic membrane properties like bending modulus and area compressibility in very good agreement with experiments.
- The theoretical framework of curvature dynamics is modified and extended. Simulated modes are consistent with macroscopic wave equations, providing a coherent model of membrane motion from 0.25 mm and 1 ms down to lipid sizes and 100 ps.
- The calculation of a local virial is implemented in simulations, and used to resolve pressure profiles across a bilayer. The resulting surface tension is decomposed into energetic and entropic contributions from various interactions. This provides a very detailed description of the microscopic forces keeping the bilayer together.
- Lipid chain NMR relaxation rates and rotational diffusion coefficients are calculated from a bilayer simulation extended to 0.1 µs, in agreement with spectroscopic results. It is further possible to determine a translational diffusion coefficient for long time scales that is consistent with values obtained by fluorescence and NMR experiments.
- The first-ever atomic level simulations of spontaneous aggregation of phospholipids into bilayers are reported, starting from random solutions of lipids. The aggregation is found to occur in several key steps, with the reduction and disappearance of biologically relevant transmembrane water pores as the rate limiting process.

Keywords: Molecular dynamics, solvent diffusion, phospholipid bilayers, mesoscopic dynamics, undulations, peristaltic motions, local virial, lateral pressure profile, NMR relaxation rate, reorientation dynamics, translational diffusion, spontaneous aggregation.
Preface

This thesis presents some of the main results of my research performed at Theoretical Physics, Department of Physics at the Royal Institute of Technology during the period 1996–2001. The contents is divided in two parts, where the first constitutes a background to the field of research and puts the results in a larger context. Chapter 1 is a short introduction to the structure and motion of biological macromolecules, while chapter 2 explains the concept of doing chemistry *in silico*, i.e., with computers. Chapters 3 and 4 cover the physical and chemical processes I have studied with these theoretical tools, and chapter 5 is a somewhat more detailed account for the definition and implementation of local pressure that I have introduced in the molecular dynamics simulation software. Part of this final chapter is necessarily of a more technical nature since it is meant to complement the relatively brief description in paper 3. This introduction is followed by a few commentary remarks on the papers.

The second part of this work contains the scientific publications forming the basis of the thesis, and in which the results of the research are presented in more detail. Have fun!

List of Papers

1. *Solvent diffusion outside macromolecular surfaces*,
   Erik Lindahl and Olle Edholm,

2. *Mesoscopic undulations and thickness fluctuations in lipid bilayers from molecular dynamics simulations*,
   Erik Lindahl and Olle Edholm,

3. *Spatial and energetic–entropic decomposition of surface tension in lipid bilayers from molecular dynamics simulations*,
   Erik Lindahl and Olle Edholm,

4. *Molecular dynamics simulation of NMR relaxation rates and slow dynamics in lipid bilayers*,
   Erik Lindahl and Olle Edholm,

5. *Simulation of the spontaneous aggregation of phospholipids into bilayers*,
   Siewert J. Marrink, Erik Lindahl, Olle Edholm and Alan E. Mark,
   Submitted to J. Am. Chem. Soc.
6. **GROMACS 3.0: A package for molecular simulation and trajectory analysis**, 
   Erik Lindahl, Berk Hess and David van der Spoel, 
   Submitted to J. Mol. Mod.

A large part of my research has involved developing and optimizing biophysical software for parallel computers, resulting in some additional works not included in this thesis:

   David van der Spoel, Aldert van Buuren, Emile Apol, Pieter Meulenhoff, Peter Tieleman, Alfons Sijbers, Berk Hess, Anton Feenstra, Erik Lindahl, Rudi van Drunen and Herman Berendsen, 
   Internet: http://www.gromacs.org

8. **Identification of related proteins on family, superfamily and fold level**, 
   Erik Lindahl and Arne Elofsson, 

9. **Isolated hypervariable regions of streptococcal M protein bind human C4BP with very high specificity: implications for antigenic variation**, 
   Eva Morfeldt, Karin Berggård, Torbjörn Drakenberg, Eskil Johnsson, 
   Erik Lindahl, Sara Linse, Jenny Persson and Gunnar Lindahl, 
   Submitted to EMBO J.
This Author's Contributions to the Papers

Science of international competitive quality is rarely the work of a single man or woman, and this definitely applies to the research in all of the publications quoted above. The most important scientific ideas have all emerged in vivid discussions among the authors, but it is nevertheless possible to roughly define the various contributions. The first idea and computer simulations behind paper 1 came from Olle Edholm, while I performed the statistical data analysis. The theoretical derivations and text were developed by both of us in cooperation. For the research presented in papers 2 through 5, we have used the molecular dynamics simulation package GROMACS, of which I am a coauthor (this package is presented in more detail in paper 6). My main contributions to the software include the optimizing inner loop generator, the assembly code for x86 PC hardware, uniprocessor and parallel versions of the particle mesh Ewald summation, Nosé-Hoover thermostat, Parinello-Rahman barostat, and the automatic package configuration scripts. The mesoscopic membrane studies of paper 2 were initiated by me, and I also performed the parallel simulations and trajectory analysis. The theoretical framework was developed by me, in close cooperation with Olle Edholm and the text was written in a joint effort. For paper 3, I came up with the local pressure approach. I also developed the implementation of local virial for the various potential terms, implemented it in the code and performed the simulations and analysis. The application to surface tension decomposition was due to Olle, while the theory work and manuscript writing was shared between us. In paper 4, I carried out the simulations and the relaxation analysis, while dihedral analysis was developed by my coauthor. Both of us took part in the derivation of theoretical concepts and manuscript writing. The idea behind paper 5 came from Siewert-Jan Marrink who also was responsible for pushing the project to a result; a huge amount of trial simulations were performed in Stockholm and Groningen for the initial part of the project, with the eventual successful aggregation accomplished in Groningen. All authors contributed to the manuscript writing. The final paper 6 is an account for the molecular dynamics simulation package developed in Groningen, Uppsala and Stockholm, with some of my contributions to the code mentioned above. Both I, David van der Spoel and Berk Hess contributed to benchmarking and the final manuscript text. Great software work guys, we have really earned the credit!
Acknowledgments

First of all I’d like to extend my gratitude to my supervisor Dr. Olle Edholm for all the encouragement and support during these years. Apart from being a nice guy and positive coworker, he has always put off time for our discussions and I have learned invaluable amounts of physics from him. I also want to thank professor Clas Blomberg for inviting me to the theoretical biophysics group. The Swedish National Allocations Committee for High Performance Computing and the Center for Parallel Computers have granted me insane amounts of supercomputing resources, and the Royal Institute of Technology made this research possible by providing my financial support during these years through a Fellowship of Excellency.

I have really enjoyed the collaborations with David van der Spoel in Uppsala, Siewert-Jan Marrink and Berk Hess in Groningen, and Arne Elofsson at Stockholm University. They have all shared a lot of their knowledge, and I think we have performed some great research together.

Sharing room with Lars Sandberg and Per Aronsson has simply been terrific, and they have become close friends over these years. I had really fun redesigning the undergraduate course in mathematical methods of physics together with Lars, and it is always rewarding to discuss music with both of them. I just hope I haven’t bored you too much with all those Maria Callas recordings… Fredrik Viklund at the department of Biochemistry has become a kind of virtual room-mate over the years, doing heroic efforts to keep my thoughts away from dull science stuff; we really ought to do some bird watching together soon! A lot of other people have also contributed to make my commitments in the PhD section of the student union and the university board very rewarding experiences. This space is too limited to mention all of them, but nevertheless: Thank you!

It has always been entertaining to discuss unix matters with Ausrius Juozapavicius, Jurij Smakov and Richard Armiento. All the other staff and PhD students at the department have also provided constant companionship and inspiration: Anders Vestergren, Helena Magnusson, Tommy Olsson, Christian Ekstrand, Mathias Ekman, Kristin Persson, Nils Sandberg, and Blanka Magyari-Köpe just to mention a few. Agneta Christiansson and Björn Pettersson have been doing their best to keep us (relatively) sane in this special environment, although I’m sorry to disappoint Agneta by still not getting a real job for a while.

Finally, I’d like to thank my parents for constantly encouraging my odysseys in science, and express my very deepest love to Camilla; thanks for putting up with me during these last few weeks of thesis-writing and grant proposals…

Stockholm, March 2001

[Signature]
CONTENTS

Abstract ............................................... iii
Preface .................................................. iv

I Background 3

1 Introduction 5
  1.1 Life on the Molecular Level ....................... 5
  1.2 From Structure to Dynamics ......................... 8
  1.3 What This Thesis is About ......................... 10

2 Molecular Dynamics 11
  2.1 Doing Chemistry in Silico ......................... 11
     2.1.1 Force Fields .................................. 13
     2.1.2 Integration Algorithms ....................... 16
  2.2 Defining Ensemble Properties ...................... 19
  2.3 Limitations ....................................... 20
  2.4 Constraint Dynamics .............................. 22
  2.5 Dummy Particles ................................. 24

3 Water & Interface Systems 25
  3.1 Solvent Properties ................................ 25
  3.2 Interface Structure & Binding Potentials .......... 27
  3.3 Diffusion, Mobility & Autocorrelations .......... 32

4 Membranes 37
  4.1 Composition & Structure .......................... 37
     4.1.1 Bilayer Experiments & Simulations .......... 38
     4.1.2 Molecular Area & Surface Tension .......... 41
  4.2 Dynamical Bilayer Properties .................... 43
     4.2.1 Lipid Chain Relaxation ...................... 44
     4.2.2 Rotational & Translational Diffusion ....... 47
     4.2.3 Curvature Deformations .................... 49
In fond memory of Marianne Belfrage
Part I

Background
Introduction

Sometimes I think the surest sign that intelligent life exists elsewhere in the universe is that none of it has tried to contact us.

Calvin and Hobbes (Bill Watterson)

1.1 Life on the Molecular Level

Biology as we know it is characterized by an incredible diversity of species, ranging from huge mammals to microscopic bacteria that can survive far above the boiling point of water, not to mention plants. Even more fascinating, all these animals and plants essentially have their basic hierarchical structures and functions in common—they are all built from cells, the smallest independent units classically considered to be “living”, either in the form of free structures (e.g., bacteria) or as multicellular organisms. The size of a human egg cell is in the order of 0.1 mm, making it barely visible to the naked eye. A bacterium is almost 100 times smaller while nerve cells can extend a meter or more [1].

We have learned a lot about science in general and cellular biology in particular during the last century; through the development of better optical and electronic microscopes we now know that inside these structures there are even smaller elements like ribosomes, mitochondria, and the nucleus with chromosomes containing our genetic material. This has further lead to the birth of a molecular biology concerning the individual chemical molecules mediating all the processes of life, on nanometer scales [2]. As of this writing, it was only a couple of weeks ago (February 15/16, 2001) that the International Human Genome Sequencing Consortium and Celera Genomics published their draft mappings of our entire genetic code [3, 4], pushing biology into the information age and marking the beginning of a new post-genomic era. Looking back, science has come a very long way, but we are still far from the eventual goal of fully understanding the properties of life processes from the underlying chemistry and physics of the participating molecules...
Chapter 1. Introduction

Figure 1.1. Left: A short fragment of the DNA double helix carrying our genetic code from which proteins are synthesized. The width is only a few nanometers, but the full molecule contains billions of basepairs and would have a length of about a meter if extended. This was the first three-dimensional structure of a biopolymer when it was determined in 1953 by Watson & Crick using X-ray crystallography data from Rosalind Franklin’s lab. Right: Myoglobin is a relatively small protein molecule (about 1600 atoms, 3 nm in diameter), responsible for storing the oxygen in muscles. This and the similar hemoglobin which carries oxygen in the blood were the first protein structures determined, by Kendrew and Perutz respectively (using X-ray crystallography). Watson/Crick (together with Wilkins) were awarded the Nobel prize in Physiology and Kendrew/Perutz the prize in Chemistry, both in 1962, for these important discoveries. (Images rendered by the author using Raster3D [5].)

An essential difference between the chemistry of nonliving systems and that of the living is the much greater complexity of biological molecules. In addition to smaller organic and inorganic compounds, living cells contain complex macromolecular assemblies consisting of hundreds to millions of atoms. In light of the recent media coverage, the most famous of these is certainly the deoxyribonucleic acid, DNA, in which the genetic code is stored as a sequence of bases (illustrated in Fig. 1.1). The unique design of each living organism is described by the information in this genetic blueprint, which is used to synthesize proteins in the ribosomes (a part of the cell). Proteins are the workhorse molecules of cells and responsible for an abundance of tasks, for example the photosynthesis in plants and the digestive enzymes which degrade food to simple compounds. Further examples are muscle proteins that produce mechanical work from chemical reactions and transport proteins, which move molecules like oxygen to their site of utilization.

The most frequently occurring molecules in this thesis are however neither proteins or sugar phosphates like DNA, but the at first sight somewhat less impressive lipids, which are the main component of all cellular membranes (see Fig. 1.2). These relatively small molecules have a dipolar or sometimes charged “headgroup” soluble in water (hydrophilic)
1.1. Life on the Molecular Level

Figure 1.2. The chemical structure of 1,2-dipalmitoyl-sn-glycero-3-phosphocholine, also known as dipalmitoylphosphatidylcholine, or simply “DPPC”. The tails are drawn as zig-zag lines with each vertex representing a CH$_2$ group. It is a relatively typical lipid present in biological systems, and often chosen for simulations since pure DPPC systems are well studied experimentally.

and one or two extended hydrocarbon tails which dislike water (hydrophobic). At larger concentrations, this makes them aggregate and turn the water-repelling parts towards each other, while the soluble headgroups face the aqueous environment. Depending on the relative sizes of headgroup and tails, they can either form spherical micelles or create planar bilayer structures. The discovery that cellular membranes consist of lipids was due to Overton, who at the turn of the 20th century noted correlations between membrane penetration rates of small molecules and their partition coefficients between oil and water [6]. Gorter and Grendel later pioneered the idea that membranes are arranged as bimolecular leaflets, or lipid bilayers [7]. These bilayers play a central role for both structure and function of all cells. Although biological membranes are slightly more complex with other molecules like cholesterol solved in the bilayer and a supporting cytoskeleton, their characteristics are mainly determined by the properties of the lipids.

The outermost plasma membrane enveloping the cell from its environment is perhaps the most important structure (as illustrated in Fig. 1.3), but all membranes define different compartments in the cell. They determine the nature of all communication through the interface and work as screening devices, allowing and sometimes even assisting the penetration of some molecules but not others. In addition, they provide a matrix supporting membrane proteins which transport e.g. ions in and out of the cell [8]. This barrier function is crucial for the biological activity of the cell; the ordinary antibiotics used to fight bacteria stick to proteins and lipid composition found in bacterial membranes but not human and destroy this barrier by increasing the permeability of the membrane, killing the bacterium.
Figure 1.3. The construction of a cellular membrane. A typical lipid has a headgroup of partially charged atoms soluble in water, and one or several hydrocarbon chains. The length of the molecule is roughly 2 nm. When these aggregate they can form stable bilayers with a thickness of about 5 nm by orienting the chains to face each other. Larger patches of biological membranes usually also contain other molecules like whole integrated proteins. These structures are responsible for all the walls and interfaces in living cells.

1.2 From Structure to Dynamics

The first experimental structures of proteins and DNA were all derived from crystallized forms of the biomolecules. X-ray crystallography is still the highest resolution method available, but it is limited to provide an average structure of the molecule in an environment very different from that of a living cell. It has further relatively long been known from various experimental studies [9, 10] that biomolecules exhibit substantial fluctuations that are essential to their biological activity but not captured in the average structures. Actually, this is one of the main reasons considerably less is known about the detailed properties of biological membranes compared to e.g. DNA and proteins; they do not have any well defined equilibrium structure that is easy to determine. With the advent of new experimental methods like laser spectroscopy and NMR that also work with samples in solution, increasing attention has however been focused on fast collective dynamics and interactions of biological macromolecules (see, e.g., Ref. [11]).

In 1972, Singer and Nicholson demonstrated the dynamics of cellular membranes by a beautiful experiment where they fused cells from human and mice to a single structure. Initially, the proteins from the two species were separated on their respective sides of the fused cell, but after about 40 minutes they were evenly distributed over the whole membrane [12]. This lead them to propose a fluid-mosaic model where the proteins are relatively free to diffuse around in the bilayer, which is better described as a two-dimensional liquid than a fixed structure. The membranes can also be distorted out of the plane in bending motions to deform the cell, and even repair themselves when torn. This flexibility is responsible for the ability of red blood cells to change shape when binding oxygen, and it provides the very
basis for cell division and fusion. It is also useful for enveloped viruses that enter animal cells by membrane fusion and exit by budding. Lipids membranes are thus quite complicated structures, with dynamics stretching over several orders of magnitude as shown in Fig. 1.4.

Most biomolecular structures have also been found to be strongly influenced by their surrounding environment, usually water which constitutes about 70% of most cells. The water molecules close to surfaces like DNA, proteins or membranes are very tightly bound to the structures, and sometimes better characterized as being part of them rather than the bulk solvent. Most biological processes will not work without these solvent molecules. In fact, replacing less than every tenth water molecule in a solution with urea will destroy the structure of most proteins [13]. This importance comes from the very special properties of water as a highly polar environment screening molecular interactions and its rigid network of hydrogen bonds even in the liquid phase.

Our knowledge of biomolecular dynamics is still quite limited, although it is a very active field of research. There are experimental methods that can resolve individual atoms, but they are usually not capable of directly quantifying complex collective motions. Indirect methods like spectroscopy can sometimes be applied in combination with theoretical models for the motions, but this approach is far from generally applicable.
1.3 What This Thesis is About

The above introduction might not have enlightened you a lot about what I have really been doing during these years. Where do physics, computers, and simulations enter in this equation?

Both chemistry and most other experimental sciences usually rely on a top-down approach. That is, measurements are gradually refined to be able to observe smaller structures and faster processes until technical limits are reached. If we had access to very, very powerful computers one could try to reverse this algorithm and do bottom-up modeling instead. This is the central idea of the methods I have used for the research summarized in this thesis; starting from simple (almost trivial) pairwise interactions of atoms, computers can be used to simulate what happens in complex biological molecules on longer timescales. This way it is actually possible to see the atomic motions on a level usually not accessible to experiments. The knowledge gained can then be used to return to the drawing-table and formulate better models for the phenomena observed, to be able to understand and perhaps even manipulate the systems, e.g. transport drug molecules through cellular membranes. Further, when the simulations reach time and length scale where it is also possible to perform experiments (and agree with these), the chemistry and physics of the molecules can be traced all the way from individual atoms up “real-world” macroscopic systems.

Of course, that is the nice, ideal picture—in practice it is hours, days and months of programming, debugging, problems with numerical algorithms and waiting for those long runs to finish. . . But at least you never have to do any wet laboratory work!
Molecular Dynamics

We’ve all heard that a million monkeys banging on a million typewriters will eventually reproduce the entire works of Shakespeare. Now, thanks to the Internet, we know this is not true.

Robert Wilensky, University of California, 1996

2.1 Doing Chemistry in Silico

Physics provides a lot of information and knowledge about the detailed structure of matter at the atomic level. Essentially all interactions and reactions between atoms can be understood from the motions of their electrons, and in principle these could be calculated using quantum mechanics. The equations describing this are actually quite simple and they can even be solved exactly, but only for extremely simple configurations. Any realistic system requires approximate numerical methods using computers.

Ultimately, the fundamentalist minded physicist would thus like to be able to predict the complete structure and dynamics of any system, preferably including the universe, from quantum mechanics by numerically solving the time-dependent Schrödinger equation for the constituent electrons and atomic nuclei. For macromolecules like proteins with thousands to millions of atoms this is unfortunately not an alternative. Despite the extremely fast development of computers in the last few decades the necessary power certainly will not see the light of the day this century, and probably not in future ones either.

In practice this does not matter much, though. Relatively few processes, in particular on the biomolecular level, actually require any quantum mechanical approach—the world is classical! Of course we cannot describe the detailed processes of bond breaking or formation without it, and some of the fast vibrations in bonds and angles might also border to the quantum mechanical region, but apart from this most motions are more than adequately described by sticking to classical mechanics. In fact, it will often prove a much
better alternative regardless of the computational cost, since the models are parameterized to reproduce real experimental observations rather than simplified theories.

With this motivation, the *ab initio* approach can be replaced with a semi-empirical parameterization of classical forces in the system. Since this is many orders of magnitude faster than solving the Schrödinger equation we will be able to simulate far larger systems for far longer times. It also makes it much easier to carry out simulations under realistic biochemical conditions, since many quantum mechanical calculations correspond to simulations in vacuum at a temperature of 0 K, which is slightly different from the real world...

These simplifications make it possible to study chemistry using computers instead of experiments. Since the machines are based on silicon microchips the description *in silico* has appeared in the last few years, in analogy with the *in vivo / in vitro* designations used for biochemical results in living organisms and test tubes, respectively. It is also used for other techniques mixing biology and computers, e.g. bioinformatics.

The concept of molecular dynamics was originally developed by Alder & Wainwright in the early 1950’s as a technique to simulate a system of colliding hard core particles [14]. It was later extended to continuous potentials and uniform time steps by Rahman [15]. The underlying idea is actually rather simple, but it requires a lot of computing power even by today’s standards. Our aim is to reproduce the time development of a system of $N$ interacting particles (e.g. atoms) with masses $m_i$ by directly solving Newton’s equations of motion,

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i, \quad i = 1..N,$$

where $\mathbf{r}_i(t)$ is the position of particle $i$. The momentary force $\mathbf{F}_i$ on each atom should be calculated from the interactions in the system. It is defined as the derivative (i.e., slope) of a potential energy function $V$ that is determined by the positions of all the atoms,

$$\mathbf{F}_i = -\nabla_{\mathbf{r}_i} V(\mathbf{r}_1, \ldots, \mathbf{r}_N).$$

The calculation of this potential function is a central part of the algorithm, which we will approach in the next section. The equations imply that the force will try to move the particles to reach a state of low potential energy, in the same way as the force of gravity acts on a free apple to move it to a state of lower height (i.e., potential energy). The resulting acceleration tells us how the speed is changing, and from the speed variation it is possible to determine approximate positions of the atoms a very short time later. This process is called *integrating the equations of motion*, and repeating the calculation for a huge number of small steps results in a *trajectory* with the development of positions, velocities and forces on all atoms during the simulation. If the potential energy function is a good approximation of the real interactions between the particles, this can provide an extremely detailed description of both the dynamics and equilibrium properties in the system under study. The first-ever such computer simulation of a biomolecule was performed as recently as 1977, but the field is advancing at an incredible pace (see further table 2.1). The available scales in time and space are still limited compared to many biochemical processes, but the performance is virtually exploding due to better parallelization and faster software, not to
2.1. Doing Chemistry in Silico

<table>
<thead>
<tr>
<th>System</th>
<th>Year</th>
<th>Authors</th>
<th>Atoms</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard spheres</td>
<td>1953</td>
<td>Alder &amp; Wainwright [14]</td>
<td>108</td>
<td>-</td>
</tr>
<tr>
<td>Water</td>
<td>1971</td>
<td>Rahman &amp; Stillinger [18]</td>
<td>648</td>
<td>2.2 ps</td>
</tr>
<tr>
<td>Protein in vacuum</td>
<td>1977</td>
<td>McCammon et al. [19]</td>
<td>580</td>
<td>8 ps</td>
</tr>
<tr>
<td>Bacteriorhodopsin</td>
<td>1995</td>
<td>Edholm et al. [21]</td>
<td>18,384</td>
<td>300 ps</td>
</tr>
<tr>
<td>Porine</td>
<td>1998</td>
<td>Tieleman &amp; Berendsen [22]</td>
<td>65,898</td>
<td>1 ns</td>
</tr>
<tr>
<td>Peptide folding</td>
<td>1998</td>
<td>Duan &amp; Kollman [23]</td>
<td>9500</td>
<td>1 µs*</td>
</tr>
<tr>
<td>Lipid bilayers</td>
<td>1999</td>
<td>Lindahl &amp; Edholm [24]</td>
<td>121,856</td>
<td>10 ns</td>
</tr>
<tr>
<td>Micelle equilibrium</td>
<td>2001</td>
<td>Lindahl &amp; Marrink†</td>
<td>549,552</td>
<td>30 ns</td>
</tr>
</tbody>
</table>

Table 2.1. Some examples of the rapid development of molecular dynamics simulations. Although the method is still limited to relatively small systems and time scales, it is advancing at a faster-than-exponential pace. The computational complexity defined as system size multiplied by the simulation time has increased more than 130 times every decade, not counting the more detailed and costly algorithms used in later works. The most recent simulations break this pattern, showing even faster performance improvements. The millisecond is within reach! *The Duan/Kollman simulation used very short cut-offs to be able to reach a microsecond. †Work in progress. (Table adapted from Ref. [25].)

mention extended techniques like dissipative particle dynamics [16, 17]. This author is convinced we will be able to simulate biomolecules for a millisecond before the decade is over, which would make it possible to study processes like DNA transcription and the folding of proteins. An important part of these efforts is the development of higher performance simulation software, by introducing better optimizations and algorithms that enable longer time steps. Paper 6 in this thesis presents the latest release of GROMACS, a versatile package for molecular dynamics that is almost an order of magnitude faster than other existing software. Although still in its infancy, computer simulation is rapidly developing into a powerful tool for many disciplines, both in chemistry and physics.

2.1.1 Force Fields

Once we have determined the equations of motion, the remaining important task is to define and calculate the potential energy \( V(\{r\}) \). There are many possible choices for the implementation of this function, and it is always an approximate compromise between a description as detailed as possible and one which can be evaluated fast on a computer. Its final appearance, together with the actual interaction parameters used in the simulation, is called the force field. Several different force fields have been developed by various research groups, but most of them are closely related as illustrated by Fig. 2.1.

A potential that provides extremely high accuracy in all situations would be slow to calculate and not very useful, so when applying it to a molecular level several approximations are made. First, it is subdivided into contributions from bonded interactions of connected atom (e.g., bonds, angles, and bond rotations) and nonbonded interactions between pairs
of atoms only located close to each other. The potential due to variations in bond length between atoms $i$ and $j$ is usually modeled as a simple harmonic spring,

$$ V_{\text{bond}}(r_{ij}) = \frac{k_{ij}^b}{2} (r_{ij} - r_{ij}^0)^2, $$

where $k_{ij}^b$ is a force constant that describes the stiffness of the actual type of bond and $r_{ij}^0$ is the equilibrium length of the bond. In a similar way, angle stretching for the atoms $i, j, k$ (where $i$ is bonded to $j$ and $j$ bonded to $k$) can be described by

$$ V_{\text{angle}}(\vartheta_{ijk}) = \frac{k_{ijk}^\vartheta}{2} \left( \vartheta_{ijk} - \vartheta_{ijk}^0 \right)^2, $$

with $\vartheta_{ijk}^0$ being the equilibrium angle. The force constant $k_{ijk}^\vartheta$ determines how hard it is to distort the angle. In larger molecules there will also be variations in potential due to the rotation around the middle bond in a sequence of four atoms as shown in Fig. 2.2. This is an effect from the geometrical configuration of the bonding electron orbitals. Although the true interaction is quite complicated, it is usually modeled with a simple periodic dihedral potential,

$$ V_{\text{dihedral}}(\varphi_{ijkl}) = k_{ijk}^{\varphi} \left[ 1 + \cos \left( n\varphi - \varphi_0 \right) \right]. $$

There are only two reasonable alternatives for the zero position of the angle, but of course scientists cannot agree on using one of them. We apply the biochemical convention, where the angle $\varphi$ is zero in the $cis$ state (when atoms 1 and 4 are on the same side of the bond).
2.1. Doing Chemistry in Silico

The geometrical definition of the dihedral angle according to the biochemical convention. The image shows a gauche state. Right: The Ryckaert-Bellemans potential (solid) and an ordinary single-cosine dihedral (dashed). The cis state is located at $\phi = 0^\circ$, trans at $\phi = 180^\circ$ and the two gauche conformations at $\phi = 60^\circ$ and $\phi = 300^\circ$. The cosine potential will be complemented with 1,4 nonbonded interactions.

The multiplicity $n$ determines the number of minima during a full rotation and $\phi_0$ their positions. For aliphatic hydrocarbons (e.g. the noncharged tails in lipid molecules) with four or more consecutive carbons we sometimes need a better potential description which explicitly takes into account that the two gauche states at $\pm 60^\circ$ are slightly less favorable than the trans one at $180^\circ$, and also that transitions between gauche and trans are considerably easier than crossing the cis state due to the steric repulsion between atom 1 and 4. Ryckaert and Bellemans [26] have constructed such a potential which can be written as a sum of six cosine terms (see Fig. 2.2 for the appearance). When using this modified dihedral instead of the ordinary one, the corresponding nonbonded interaction responsible for the repulsion between atom 1 and 4 should be removed. In paper 4, where we study the NMR relaxation rates in lipid chains, it turns out that the dynamics is extremely sensitive to the exact height of this dihedral barrier. This is a very sensitive test of the accuracy of the potential, and suggests that other alternatives like the Kuwajima [27] potential are probable even better descriptions than Ryckaert-Bellemans for extended hydrocarbon chains.

All these interactions concern atoms closely bound to each other. This makes them very local in space, and most important their number will only increase linearly with system size, meaning they are not very costly to calculate. In contrast, there are a lot more of nonbonded interactions between atoms separated by more than three bonds or even located in different molecules. The calculation of these forces is the single most time-consuming part of any molecular dynamics simulation, accounting for roughly 90–95% of the total processor usage, even if we assume all forces to be between pairs of particles and neglect contributions beyond some cut-off distance. These nonbonded interactions are usually described with a repulsive term important on short distances (overlap of wave functions), attractive dispersion forces that are always present, and electrostatic interactions between
charged particles. The dispersive and repulsive components are often combined in the form of Lennard-Jones interactions,

\[ V_{\text{Lennard-Jones}}(r_{ij}) = C_{ij}^{(12)} \left( \frac{1}{r_{ij}^{12}} - \frac{1}{r_{ij}^6} \right), \]

(2.6)

where \( C_{ij}^{(6)} \) and \( C_{ij}^{(12)} \) are parameters that depend on the type of atoms involved, determining the amount of attraction and repulsion respectively. The resulting potential is weakly attractive at large distances but strongly repulsive when the atoms start to overlap. The electronic repulsion is actually better described by an exponential term (Buckingham interactions), but the exponential function is very expensive to calculate on a computer and thus usually replaced by the slightly simpler but much cheaper \( r^{-12} \) expression. For charged pairs of atoms the electrostatic interaction is given by

\[ V_{\text{Coulomb}}(r_{ij}) = \frac{q_i q_j}{4\pi \varepsilon_0 \varepsilon_r r_{ij}}, \]

(2.7)

where \( q_i \) and \( q_j \) are the charges. The permittivity of free space is designated \( \varepsilon_0 \), and \( \varepsilon_r \) is the relative permittivity. The latter is usually set to its vacuum value of unity, which might seem unrealistic when simulating systems like water with an experimental value up to 80 times higher. This macroscopic screening is however largely an effect from rotation and displacement of whole dipoles and molecules. In molecular dynamics, all these motions are reproduced explicitly and the remaining microscopic permittivity is thus considerably smaller, approaching the vacuum value. It is still possible to argue that the value should be higher than unity, but in many force fields this is instead corrected implicitly by parameterizing the charges to yield correct equilibrium properties. As an alternative, a permittivity \( \varepsilon_r = 2 \) is sometimes used to account for the electronic polarizability always present. There are many more elaborate models for the relative permittivity, but eventually one has to accept that it is a macroscopic feature that is not easy to represent on atomic level. The ultimate solution would be to extend the force field description by using explicit bond polarization charges, and there have recently been several successful attempts to do this, starting with pure water [28, 29]. It is presently not routinely used in any major force fields though, either due to difficulties in parameterization and problems with oscillating polarization or because the models are simply too expensive computationally.

### 2.1.2 Integration Algorithms

Once the forces for the configuration of atoms at the current time \( t \) have been calculated the next step is to generate a new configuration at time \( t + \Delta t \) according to the dynamics in Eq. 2.1. For a simple model motion where one desires a solution as exact as possible, a high accuracy algorithm to integrate Eq. 2.1 with small steps would be necessary. It would not matter how often the forces had to be calculated. The situation for the very big systems usually studied with molecular dynamics is however very different. In this case it is useless trying to determine a very detailed solution for individual atoms since the dynamics is
chaotic and small numerical errors will grow exponentially and affect the trajectories. This might at first seem terrible since it strikes against the whole concept of simulations, but it only reflects real systems—equilibrium properties are not sensitive to details of individual trajectories. It is thus futile to try to reproduce motions exactly. Instead, one should make sure that any reasonably long part extracted from a trajectory would be a fair description of a particle with the same initial conditions, called a shadow trajectory [30].

There are several numerical algorithms available with good performance for long time steps that only require a single force evaluation per step. One of the most frequently used (and best for molecular dynamics) was developed by Verlet [31] and has since turned into an entire class of integrators. It is based on the addition and subtraction of the Taylor expansions for the time dependence of the coordinates $r_i$ at the times $t - \Delta t$ and $t + \Delta t$, which together with Eq. 2.1 leads to

$$r_i(t + \Delta t) \approx -r_i(t - \Delta t) + 2r_i(t) + \frac{\Delta t^2}{m_i} F_i$$

(2.10)

$$v_i(t) \approx \frac{1}{2\Delta t} [r_i(t + \Delta t) - r_i(t - \Delta t)].$$

(2.11)

The truncation error introduced with this scheme is of the order $(\Delta t)^4$ for the coordinates and $(\Delta t)^3$ for the velocities. However, a more practical problem with this approach is that the velocities are obtained as the difference of two terms of the same magnitude, making it very sensitive to numerical precision and round-off errors. A slightly modified, but theoretically equivalent, algorithm is the Leap-Frog scheme [32],

$$r_i(t + \Delta t) \approx r_i(t) + \Delta t v_i(t + \Delta t/2)$$

(2.12)

$$v_i(t + \Delta t/2) \approx v_i(t - \Delta t/2) + \frac{\Delta t}{m_i} F_i.$$  

(2.13)

This is still a second order approximation to the equations of motion, but it avoids the difference between large terms when calculating the velocities. The name comes from the generation of positions and velocities at whole and half time steps, respectively. They are leaping like frogs over each others back, see Fig. 2.3. The only major drawback is that the velocities are offset from the positions by half a step, but in the molecular dynamics software we have developed this is circumvented by averaging the velocities at plus and minus half a step to obtain the same value as the original Verlet algorithm would have produced without round-off errors.
Figure 2.3. The Leap-Frog integration scheme. Coordinates $r_i$ (black squares) and velocities $v_i$ (gray squares) of the atoms are calculated at alternating whole and half time steps. Compared to advanced predictor-corrector schemes, this simple algorithm has the advantage of being very stable for long time steps. This means we can advance further in time for each calculation of the forces and thus improve performance significantly since the force calculation accounts for 90-95% of the simulation time. The only major drawback is that the velocities are offset by half a step, but the exact values at integer steps are easily calculated by averaging velocities from plus and minus half a step.

There are several more advanced integrators available, for example the Gear algorithms and other predictor-corrector schemes [15, 33, 34]. In general, these are not very suitable for molecular dynamics. Their main feature is the superior accuracy for very small time steps, but since the expensive force calculation has to be repeated for each step there is no point in using small steps considering the shadow trajectory discussion above. When we are trying to use time steps as long as possible to extend the simulation, these high-order advanced integrator schemes will exhibit much larger energy drifts than simple integrators, which would restrict the possible time step length.

In contrast, the Leap-Frog scheme is very appealing. It has relatively large short time fluctuations but very low long time energy drifts—it can even integrate an harmonic oscillator with only six points per period [35]. Many of the more advanced numerical algorithms also require more than one evaluation of the forces which essentially renders them useless for molecular dynamics, since that is the single most time-consuming part of any simulation. (Consult reference [36] for a good comparison of molecular dynamics integrators.) Reasonable time steps when simulating biomolecules are in the order of 1 fs (femtosecond, $10^{-15}$ s.), during which a typical atom moves about 0.004 nm. A simulation covering just 1 ns will then require 1,000,000 steps of force evaluation and integration, with each step in turn involving maybe 1,000,000 pairwise interactions for a reasonably sized system. Still, with a well optimized program this is quite feasible and takes less than a day on a fast computer—and we will shortly discuss methods to extend the time step.
2.2 Defining Ensemble Properties

Most of the features observed in experiments can not be attributed to the individual atoms, but they are collective properties of a very large number of atoms, for instance temperature and pressure. It is important to be able to control these in a simulation. If we start with a protein and water in a box it should proceed at about 300 k and normal pressure, not 1500 k or thermonuclear pressures, which would cause the biomolecule to fall apart.

The temperature $T$ of a system is directly related to the average kinetic energy of the atoms and the number of degrees of freedom $N_f$. We define a momentary microscopic temperature $T_m$ as

$$T_m(t) = \frac{1}{N_f k_B} \sum_{i=1}^{N} m_i v_i(t) \cdot v_i(t), \quad (2.14)$$

where $k_B$ is Boltzmann’s constant. The effective system temperature is an equilibrium average $T = \langle T \rangle$ of this entity over a suitably long time interval. The initial temperature of the system can thus be set by assigning random velocities with a gaussian distribution of the correct width to the atoms. Ideally, that temperature would be maintained through the simulation, but due to e.g. round-off errors in the integration algorithm and the use of force truncation the temperature needs to be controlled when performing longer simulations. This can be accomplished by rescaling the atom velocities every time step with a calculated factor $\lambda$,

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau_T} \left( \frac{T_0}{T} - 1 \right)}, \quad (2.15)$$

where $T_0$ is the reference temperature and $\tau_T$ a temperature coupling time constant. This weak coupling scheme [37] will produce an exponential relaxation to the reference temperature. There are also more elaborate temperature-coupling techniques like the Nosé-Hoover approach that better maintain the statistical mechanics ensemble of the system [38, 39].

It is not trivial to calculate pressure in condensed systems; the ideal gas law only holds for much lower densities where atoms do not interact, which is clearly not the case in e.g. water. Simulation pressure must instead be derived from the basic definition of force per unit area on the walls of the simulation cell.$^1$ This is not directly accessible in the simulation, but can be calculated with a trick using the virial theorem (See, e.g., Refs. [40, 41]),

$$2K = \Xi, \quad (2.16)$$

where $K$ is the full average translational kinetic energy tensor,

$$K = \frac{1}{2} \sum_{i=1}^{N} m_i v_i \otimes v_i, \quad (2.17)$$

$^1$Actually, the most appropriate definition of pressure is the derivative of the free energy with respect to the volume, but this is equivalent.
and $\Xi$ the virial of the system,

$$\Xi = -\sum_{i=1}^{N} r_i \otimes F_i.$$  \hspace{1cm} (2.18)

This virial has an internal component $\Xi_{int}$ from forces inside the system, and an external contribution $\Xi_{ext}$ caused by the pressure $p$ on the cell surface $S$,

$$\Xi_{int} = -\sum_{i=1}^{N} r_i \otimes F_i^{int}$$  \hspace{1cm} (2.19)

$$\Xi_{ext} = \int_p r \otimes dS,$$  \hspace{1cm} (2.20)

where $p_N$ is the component of the pressure normal to the surface element $dS$. For isotropic pressure this can be simplified by extracting the scalar value $p$ from the integral,

$$\Xi_{ext} = p \int_S r \cdot dS = p \int_V \nabla \cdot r dV = 3 p V,$$  \hspace{1cm} (2.21)

where a third each can be attributed to the $x$, $y$ and $z$ dimensions. In the anisotropic case, a similar relation holds for each component of the full tensor. This provides a formula to calculate the pressure,

$$p = 2 K - \frac{\Xi_{int}}{V},$$  \hspace{1cm} (2.22)

using the momentary velocities, positions and forces in the system. The scalar pressure $p$ equals a third of the trace of this tensor, i.e. the average of $x$, $y$ and $z$ dimensions.

The resulting pressure can be controlled through a coupling method very similar to the temperature, by scaling the coordinates of all particles and the box size every time step with a factor $\mu$ (compare with Eq. 2.15),

$$\mu = \sqrt[3]{1 + \frac{\Delta t}{\tau_p} \beta (P - P_0))}.$$  \hspace{1cm} (2.23)

For any reasonable value of the system compressibility $\beta$, this will produce exponential pressure relaxation, similar to the weak temperature coupling. An extended ensemble alternative is provided by the Parinello-Rahman algorithm \[42, 43\].

### 2.3 Limitations

Although potentially a powerful technique, it is important to realize that molecular dynamics as any other method has limitations which must be considered. First, the molecular
2.3. Limitations

interactions are entirely classical. While usually appropriate, there are exceptions. According to quantum mechanics single atoms may tunnel through potential barriers and we cannot hope to describe chemical reactions in which bonds form or break. Further, hydrogen atoms are very light and their motions border to quantum mechanics. Apart from doing quantum mechanical calculations, which is out of the question, one could apply correction terms to the energy and specific heat of the oscillations in the bonds. Alternatively, these degrees of freedom could be removed and the bonds treated as rigid, approximating the ground state of a quantum mechanical oscillator. This has the additional advantage that the integration time step can be increased significantly (say, from 1 fs to 2 fs). This type of constraint dynamics is described in sections 2.4 and 2.5 below.

Even if we have extended the size of simulated systems considerably in this work they are still fairly small, at least compared to real macroscopic structures. If one tried to simulate an isolated system, this would mean many of the atoms experienced a large and unnatural boundary surface to a vacuum environment. To avoid this, periodic boundary conditions are used. The atoms are placed in a rectangular (or more general, triclinic) simulation cell, and when atoms or interaction vectors cross the boundary they simply reappear on the opposite side, as if the system was a periodic crystal. The effects of this approximation will diminish as the box size increases, but they are evident in many simulations. Paper 2 in this thesis includes a discussion regarding some of the effects on membrane dynamics caused by the limited size of the bilayer patches simulated.

Finally, the accuracy of a simulation is entirely dependent on the accuracy of the underlying force field, which contains several approximations and uncertain parameters. It might seem a good idea to consistently start from quantum mechanics and derive detailed values for individual interaction parameters. Unfortunately this neglects another important approximation—all forces are assumed to act between pairs of atoms. Of course one could include three and four particle interactions, but that would be ridiculously expensive. A smarter alternative that works surprisingly well is to use semi-empirical effective parameters that have been corrected to partially account for the higher order interactions while still being pair-additive. We still neglect all dynamic effects like polarizations in bonds, though. To speed up the calculation of forces, the nonbonded interactions are further usually truncated beyond a distance of 1–2 nm. This is a fair approximation for Lennard-Jones interactions, but not always for electrostatics if there are free charges in the system. We have implemented algorithms like particle mesh Ewald summation that can perform a full summation of the forces between the infinite periodic images of all atoms, but due to the periodic boundary conditions these will over-emphasize the artificial periodicity of the system and sometimes introduce unnatural ordering.

Nevertheless, as long as these approximations are kept in mind and the results carefully checked, molecular dynamics is a very reliable method to study the motions present in biological macromolecules, and the effects of most approximations diminish the larger and longer scale phenomena we are studying.

\(^2\)Actually, this is not entirely true. As you can read in paper 6, there is currently development going on with GROMACS to be able to mix quantum mechanics and molecular mechanics to treat small parts of the system at the quantum level, but that is for the future.
interactions are entirely classical. While usually appropriate, there are exceptions. According to quantum mechanics, single atoms may tunnel through potential barriers and we cannot hope to describe chemical reactions in which bonds form or break\textsuperscript{2}. Further, hydrogen atoms are very light and their motions border to quantum mechanics. Apart from doing quantum mechanical calculations, which is out of the question, one could apply correction terms to the energy and specific heat of the oscillations in the bonds \cite{44}. Alternatively, these degrees of freedom could be removed and the bonds treated as rigid, approximating the ground state of a quantum mechanical oscillator. This has the additional advantage that the integration time step can be increased significantly significantly (say, from 1 fs to 2 fs). This type of constraint dynamics is described in sections 2.4 and 2.5 below.

Even if we have extended the size of simulated systems considerably in this work they are still fairly small, at least compared to real macroscopic structures. If one tried to simulate an isolated system, this would mean many of the atoms experienced a large and unnatural boundary surface to a vacuum environment. To avoid this, periodic boundary conditions are used. The atoms are placed in a rectangular (or more general, triclinic) simulation cell, and when atoms or interaction vectors cross the boundary they simply reappear on the opposite side, as if the system was a periodic crystal. The effects of this approximation will diminish as the box size increases, but they are evident in many simulations. Paper 2 in this thesis includes a discussion regarding some of the effects on membrane dynamics caused by the limited size of the bilayer patches simulated.

Finally, the accuracy of a simulation is entirely dependent on the accuracy of the underlying force field, which contains several approximations and uncertain parameters. It might seem a good idea to consistently start from quantum mechanics and derive detailed values for individual interaction parameters. Unfortunately this neglects another important approximation—all forces are assumed to act between pairs of atoms. Of course one could include three and four particle interactions, but that would be ridiculously expensive. A smarter alternative that works surprisingly well is to use semi-empirical effective parameters that have been corrected to partially account for the higher order interactions while still being pair-additive. We still neglect all dynamic effects like polarizations in bonds, though. To speed up the calculation of forces, the nonbonded interactions are further usually truncated beyond a distance of 1–2 nm. This is a fair approximation for Lennard-Jones interactions, but not always for electrostatics if there are free charges in the system. We have implemented algorithms like particle mesh Ewald summation that can perform a full summation of the forces between the infinite periodic images of all atoms \cite{45, 46}, but due to the periodic boundary conditions these will over-emphasize the artificial periodicity of the system and sometimes introduce unnatural ordering.

Nevertheless, as long as these approximations are kept in mind and the results carefully checked, molecular dynamics is a very reliable method to study the motions present in biological macromolecules, and the effects of most approximations diminish the larger and longer scale phenomena we are studying.

\textsuperscript{2}Actually, this is not entirely true. As you can read in paper 6, there is currently development going on with GROMACS to be able to mix quantum mechanics and molecular mechanics to treat small parts of the system at the quantum level, but that is for the future.
2.4. *Constraint Dynamics*

Figure 2.4. With LINCS, the bond length constraints are applied in two steps. In the initial correction to the updated but unconstrained coordinates, the projection of the new bond on the old one is set to the length of the latter. The second step then corrects for the lengthening due to rotation of the bond.

We now project out the forces that act along the bonds in the equation of motion. This can be performed as a correction to the updated coordinates $r_{n+1}^{\text{unc}}$, producing the intermediate $r_{n+1}^{*}$ according to the derivation by Hess [47],

$$ r_{n+1}^{*} = (I - T_n B_n) r_{n+1}^{\text{unc}} + T_n d $$

$$ = r_{n+1}^{\text{unc}} - M^{-1} B_n \left( B_n M^{-1} B_n^T \right)^{-1} \left( B_n r_{n+1}^{\text{unc}} - d \right). \quad (2.27) $$

The matrix $T$ is short for the expression $M^{-1} B^T (B M^{-1} B^T)^{-1}$. In principle this matrix inversion could be very costly and cause a lot of communication between processors, but the off-diagonal elements in $B M^{-1} B^T$ will only be non-zero when the two corresponding bonds are connected to the same atom, and in this case the matrix can be inverted with a fast power expansion using only a few terms instead of a full calculation [47]. It depends on the matrix being sparse, i.e., most elements outside the diagonal should be zero. This is always valid for pure bond constraints which produce a relatively low connectivity, but the algorithm does not work with coupled bond and angle constraints.³

After this initial correction, the projection of the new bond on the old one has the correct length, but if it has rotated the total length will be too large. In the second step, the rotational lengthening is corrected for by setting the projection of the new bond on the old to

$$ p_i = \sqrt{2d_i^2 - l_i^2} \quad (2.28) $$

³In practice this is not a significant problem, since the angle constraints are usually applied to hydrogens, and for these it is anyway a much better solution to use the dummy particles described in section 2.5.
where \( l_i \) is the current distance between the atoms. We obtain the final corrected coordinates as
\[
r_{n+1} = (I - T_n B_n) r^*_n + T_n p.
\]
(2.29)

The last step is an iterative process which in principle should be repeated until convergence is obtained, but a single iteration is usually sufficient for the accuracy required in molecular dynamics simulations.

### 2.5 Dummy Particles

The idea of replacing fast bond motions with constraints can actually be extended to remove the dynamics of entire particles, while keeping them as a part of the structure in the molecule. For hydrogens, this can be accomplished by first removing all these atoms as independent particles. Instead, their positions are calculated at the beginning of each step from the ideal geometry of bond lengths and angles, using the positions of the heavier atoms to which they are connected. The dummy position is then used just as any other atom for the force calculation, but before integrating the equations of motion the dummy forces are distributed back onto the constructing atoms [37]. In this way it is not necessary to integrate the motion of the light hydrogens explicitly, meaning the time step can be extend to about 4 fs, a factor two more effective than the normal setup, without losing any significant dynamics.

The time step is still not limited by the LINCS constraint algorithm or the dynamics of heavy atoms, but the water motions. Since the center of mass almost coincides with the oxygen atom, the light hydrogens will rotate fast around the oxygen, and if they move too far in a single step the analytical settle [48] constraint algorithm will not be able to maintain the water geometry. Some of this rotation can be canceled if the water center-of-mass is changed by moving some mass from the oxygen to the hydrogens, for instance setting their masses to 4u instead of 1u. This will make it possible to increase the steps to 6 fs, or perhaps even 7 fs. Admittedly, this redistribution of mass is completely artificial and deviates from physical reality. However, it turns out that the water dynamics is not that severely affected by it; the coefficient of diffusion for spc water becomes \( 3.3 \times 10^{-9} \text{m}^2\text{s}^{-1} \) compared to \( 4.1 \times 10^{-9} \text{m}^2\text{s}^{-1} \) with normal masses, while the viscosity changes from \( 4.3 \times 10^{-4} \text{kg m}^{-1}\text{s}^{-1} \) to \( 4.9 \times 10^{-4} \text{kg m}^{-1}\text{s}^{-1} \) [49]. These differences might seem large to the casual observer, but spc is a very simple water model and compared to the experimental results for diffusion (\( 2.3 \times 10^{-9} \text{m}^2\text{s}^{-1} \)) and viscosity (\( 8.0 \times 10^{-4} \text{kg m}^{-1}\text{s}^{-1} \)) the values obtained with redistributed masses are actually closer than the normal ones. The modified water motion does not seem to slow down the dynamics of macromolecular solutes [49], but even if it would, the relatively small changes in water diffusion and viscosity indicate a significant improvement compared to normal time steps. We successfully used this technique for time step extension in paper 5, to be able to perform the hundreds of nanoseconds of simulations on bilayer aggregation systems.
I began to study science because I had a love of certainty and a wish to discover the unvarnished truth about the universe. I now realize that this is quite like becoming an archbishop to meet girls.

Jan Isley

3.1 Solvent Properties

Water is of crucial importance to life. It is even so important that most biomolecules would not work at all or even maintain their three-dimensional structure without the surrounding solvent. How does this come? A single water molecule in its liquid state is basically tetrahedral, a property which might not be entirely evident from the usual way of drawing the oxygen and two hydrogen atoms. This geometry is explained from the electronic configuration of the oxygen which is referred to as $sp^3$ hybridized, meaning there are four equivalent valence orbitals, just as in methane (CH$_4$) or an ammonium ion (NH$_4^+$).

The oxygen atom is located at the center of the tetrahedron, the hydrogen atoms in two of the apices and two pairs of nonbonding electrons form the other two. Since oxygen is a more electronegative atom than hydrogen, the electrons in the bonds will be displaced slightly towards the oxygen. The two bonds in the molecule are therefore polarized and have permanent dipole moments directed from the oxygen (negative) to the hydrogens (positive). Another two dipoles are formed with their positive ends at the oxygen, extending towards each of the nonbonding pairs of electrons. In the presence of other charged or polar molecules all these dipoles will become considerably stronger due to the external field; the vacuum dipole moment of isolated H$_2$O molecules has been measured to 1.86 debye [50], and ab initio computer simulations predict it to increase up to about 3 debye in condensed phases like water and ice [51]. Water is thus not only polar, but also highly polarizable,

An orbital is a chemist’s way of speaking about the spatial distribution of a electrons orbiting the nucleus.
which explains its effective screening and very high dielectricity constant $\varepsilon_r \approx 80$ relative to vacuum. This means it is quite hard to perform accurate simulations of water molecules; it is first with recent models incorporating the polarizability of the bonds (see, e.g., the work of Van Maaren and Van der Spoel [29]) that it is possible to reproduce most of the dynamical water properties simultaneously.

The interplay between water molecules is dominated by electrostatic interactions of these dipoles. Their most favorable configuration is achieved when the dipole in a O–H bond is aligned with the dipole moment of the nonbonding electrons on a different water molecule. Such a dipole-dipole interaction brings the hydrogen and oxygen atoms of the two molecules so close to each other that a weak bond is created. These hydrogen bonds form a very complex network among the water molecules as illustrated in figure 3.1. Even in liquid water there are about 1.7 hydrogen bonds on average per molecule, not far from the full value 2 found in a perfect ice crystal. The thermal motions of the atoms in liquid water at room temperature are large enough that a typical pair of these bonded molecules will separate and form new bonds in clusters with other neighbors [52] in a time of roughly 4 ps. There are also smaller displacements of the molecules, corresponding to deformations or transient breaks in hydrogen bonds, which occur on a shorter time scale. As a result of this underlying molecular mobility, bulk water behaves as a moderately viscous fluid except at times below about 0.1 ps, where the rigidity of the hydrogen bond network becomes apparent [2].
3.2 Interface Structure & Binding Potentials

When a molecule is introduced in an aqueous environment, the solvent will form an envelope around it, very similar to the interface it makes with the air to maintain the number of hydrogen bonds with its neighbors. Interestingly, this essentially conserves the total energy of the system, irrespectively of whether the inserted molecule is soluble in water or not. This envelope or cage-like structure (sometimes referred to as a clathrate, see e.g. Ref. [53]) represents a much more ordered state of the solvent than without the solute, corresponding to a lower entropy $S$. This entropy enters in the free energy of the state,

$$F = U - TS.$$  \hspace{1cm} (3.1)

High free energies correspond to unfavorable states of the system, so if this was the only part of the process it would not occur spontaneously—almost no molecules would be soluble in water. However, many compounds interact strongly with the water and form new hydrogen bonds. This will decrease the energy $U$ and alter the balance. Another important implication is that the envelope does not have to be as ordered when the solvent interacts with the solute. This explains why macromolecules like proteins with many charged residues and dipoles are highly soluble in water, while small uncharged hydrocarbons are not. Since the entropic term of the free energy will decrease when the temperature $T$ increases, it also suggests why the solubility of most compounds increases at higher temperature [54].

The simulations of the spontaneous aggregation of lipids reported in paper 5 is a very nice example of these entropic effects. The initial completely random configuration of lipids in the solvent is extremely unfavorable since almost all the chains are in contact with water, resulting in a large interface area. This is the main reason why the system separates into lipid and solvent regions in about 100 ps. Paradoxically, once the transmembrane water pore has formed it will actually be stabilized by the same phenomena. The structure where the polar lipid headgroups line the pore is still not as favorable as the equilibrium bilayer, but transition to the final bilayer requires the pore to break down and temporarily expose some of the hydrocarbon chains to water. The defect will thus form a metastable state where the system can remain for tens of nanoseconds. These hydrophobic effects that occur between water and lipids also explain why the eventual membranes formed are so stable despite their flexibility. Actually, normal washing powder works through the same phenomena; the main ingredients are surfactant molecules that work as an interface and envelope fat molecules to make them soluble in water.

Experimental methods, in particular field gradient NMR techniques, have made large progress in this area in the last few years, but it is still difficult or sometimes impossible to explicitly resolve the fast solvent motions experimentally, making computer simulations an interesting alternative. The conformations and interactions of the water with a biomolecular surface can be analyzed through several different techniques in these simulations. One of the most fundamental properties is the distribution of the molecules, i.e. the local density of solvent atoms as a function of the distance to the surface. This radial distribution...
Figure 3.2. Left: Radial distribution functions for the oxygen atom in water outside myoglobin for negatively charged protein atoms (dashed), uncharged (solid) and positive ones (dot-dashed). Right: The corresponding radial distribution functions for hydrogen atoms. (Lindahl & Edholm, unpublished results.)

function (RDF) of particles of type A (e.g., water oxygen) relative to those of type B (for instance in a protein) is defined as

$$g_{AB}(r) = \frac{\langle \rho_A(r) \rangle}{\langle \rho_A \rangle_{local}} = \frac{1}{\langle \rho_A \rangle_{local}} \frac{1}{N_B} \sum_{i \in A} \sum_{j \in B} \frac{\delta(r_{ij} - r)}{V(r)},$$

(3.2)

where $\rho_A(r)$ is the density of type A particles at distance $r$ from particle B and $V(r)$ the local volume accessible to them. This is normalized with the average density $\langle \rho_A \rangle_{local}$ in the whole volume. For large values of $r$ the expression will tend towards unity, but the largest distance at which it can be correctly evaluated is limited by the size of the system simulated and the periodic boundary conditions.

We have performed analyses of this type for the distribution of entire water molecules close to a myoglobin protein in paper 1. Slightly more detailed versions of the resulting RDFs are shown in Fig. 3.2, where the local densities of oxygen and hydrogen atoms have been calculated separately outside surface regions of negative, positive and neutral charge. There is a clear layer structure close to the surface with at least two, sometimes three, more or less pronounced peaks. The hydration envelop is most evident outside neutral parts of the surface, where the first peak is located relatively far out, at 0.35 nm from the protein both for hydrogens and oxygens; the waters are not interacting much with the surface. The negatively charged regions attract the hydrogens, resulting in a peak already at 0.2 nm. This will displace the whole water molecule inwards, shifting also the first oxygen peak to 0.28 nm, closer than for the uncharged surface. The behavior close to atoms

---

2See section 2.3 in the chapter on molecular dynamics. Here, $r$ is the distance to the closest particle of type B, and when this corresponds to a surface in the next periodic image there will be an artificial decrease in the RDF.
carrying a positive charge is essentially the same, but with exchanged roles of hydrogen and oxygen. From these densities of atoms it is possible to calculate the corresponding average interactions with the surface by using the Boltzmann distribution,

\[
\rho(r) \propto \exp \left\{ -\frac{U(r)}{k_BT} \right\},
\]

where \( k_B \) is Boltzmann’s constant and \( U(r) \) a potential of mean force (pmf). This equation basically states that regions (actually statistical mechanical states) of low energy will be more populated than those of high energies, with exponential weighting. Solving the Boltzmann distribution expression for the potential gives

\[
U(r) = -k_BT \ln \rho(r) + C,
\]

where \( C \) is an arbitrary constant usually set to yield zero potential far from the surface. The pmf can alternatively be calculated by extracting the normal component of the force on the particles as a function of the distance to the surface and integrate it. These two methods are theoretically equivalent, but due to statistical errors they might yield slightly different results.

The pmf experienced in surface regions of different charge is displayed in Fig. 3.3. From the results in paper 1 it is clear that a whole first layer of hydration water is present within 0.3–0.4 nm of the surface, and the potential of mean force shows that this water is bound to the surface with an average potential of 2–3 kJ/mol. Since the surface has groups with both stronger and weaker bonding this is not uniform; experimental residence times of water molecules close to a protein [55] show a variation with about a factor 20. The corresponding fluctuations in energy are then \( k_BT \ln 20 \approx 7.5 \) kJ/mole. This would imply binding energies between protein and hydration water which range from almost zero to about 10 kJ/mole. Although the size of the average binding energy can be surmounted by thermal motions \( (k_BT \approx 2.5 \) kJ/mole) many waters will thus be considerably harder bound to the surface and restricted in their motions.

The interface structure can be examined in more detail by extracting the orientations of the water molecules in the first layer. Figure 3.4 shows the angular distributions of the vector between the two hydrogens in a water and of the total water dipole vector which is directed from the oxygen to the average of the two hydrogens. In regions where the surface carries a negative charge density the water molecule will turn one of its hydrogens towards the surface as illustrated in Fig. 3.5, which means the overall dipole will make an angle of about \( \pm 125^\circ \) to the normal while the angle of the vector between hydrogens is roughly \( \pm 35^\circ \). (Since the hydrogen–hydrogen vector can be defined in two ways, the latter peaks will be mirrored at \( \pm 145^\circ \).) This conformation explains the additional sharp peak at 0.35 nm in the hydrogen distribution outside a negative surface; it comes from the second hydrogen atom directed away from the surface as shown in the left part of Fig. 3.5. The uncharged regions of the surface do not interact strongly with the water and the corresponding angular distributions are not very distinct. Finally, in the case of surface atoms with positive charge, the water molecule will turn the oxygen atom inwards while the hydrogen-hydrogen vector is parallel to the surface.
Figure 3.3. Top: Potential of mean force experienced outside negatively charged regions of the surface by water oxygens (solid) and hydrogens (dashed). The latter will be more attracted, but the whole water molecule is displaced inwards. Middle: Potential of mean force at uncharged surface parts. The minima for water oxygens and hydrogens coincide since none of them are interacting strongly with the surface. Bottom: When the surface carries a positive charge the water oxygen atoms will experience an attractive potential, shifting also the hydrogen distribution inwards.
3.2. Interface Structure & Binding Potentials

Figure 3.4. Top: Angular distribution of the water dipole (solid) and hydrogen-hydrogen vector (dashed) outside a negatively charged surface region. The water molecule is turning one of its hydrogens towards the surface. Middle: Distributions of the two vectors when the closest atoms are uncharged. Both curves are more spread out and relatively featureless. Bottom: Outside positively charged parts of the surface the water molecules will rearrange so the oxygen is pointing in the direction of the charge. (Lindahl & Edholm, unpublished results.)
3.3 Diffusion, Mobility & Autocorrelations

Even after the system has relaxed to the equilibrium structure described in the preceding section there will be random Brownian motions of the particles both in bulk liquids and in presence of the ordering surface potential.

This translational motion of molecules or atoms is termed diffusion, an example of a transport process, which in general is defined in terms of the response of a system to a perturbation \([41, 56]\), either external or a spontaneous internal fluctuation. Diffusion is best described by the macroscopic (empirical) Fick’s law, which states that the flux \(j\) is proportional to the decrease in concentration \(u(r, t)\) of the same kind of particles,

\[
j = -D \nabla u(r, t),
\]

where the proportionality constant \(D\) is called the coefficient of diffusion. If this is combined with an equation expressing the conservation of particles or density we obtain the diffusion equation (see, e.g., Ref. [57]),

\[
\frac{\partial u(r, t)}{\partial t} - D \nabla^2 u(r, t) = 0.
\]

This is a very common equation in physics, and it is also used to describe many other processes with some conserved quantity, for instance temperature variations caused by the flow of heat.

The general solution to this type of equation can be found by first calculating the answer in the special case when we start with a pointlike particle or density located at origo in free space,

\[
u(r, 0) = \delta(r).
\]
By using Fourier and Laplace transform methods for the space and time variables this initial condition can be shown to give the solution

\[ u(r, t) = \frac{1}{(4\pi D t)^{d/2}} \exp \left\{ -\frac{|r|^2}{4Dt} \right\}, \tag{3.8} \]

where \( d \) is the dimensionality of the system. Free space corresponds to \( d = 3 \), but close to a surface the dynamics can be separated into a motion normal to the surface with \( d = 1 \) and the parallel displacements with \( d = 2 \).

Even if the process is random, the particle will diffuse away from its initial position \( r = 0 \) in the limit of long times. This can be shown by calculating the mean square displacement (msd) of the solution, i.e. the average square distance a particle has moved,

\[ \langle |r(t)|^2 \rangle = \int |r|^2 u(r, t) d^3 r, \tag{3.9} \]

under the additional condition that the total density is maintained,

\[ \int u(r, t) d^3 r = 1. \tag{3.10} \]

After a relatively simple calculation this yields

\[ \frac{\partial \langle |r(t)|^2 \rangle}{\partial t} = 2dD. \tag{3.11} \]

It follows from the theory of Green's functions that the answer is general in free space and not restricted to the simple initial condition above. The limiting behavior of the msd is thus given by

\[ \lim_{t \to \infty} \langle |r(t) - r(0)|^2 \rangle = 2dDt, \tag{3.12} \]

which is known as an Einstein relation\(^3\) for the transport process. This is an important equation, since it makes it possible to calculate the macroscopic coefficient of diffusion from the microscopic displacements of coordinates. Figure 3.6 shows a plot of this msd in bulk water as a function of time.

The assumptions made when deriving this relation should however be kept in mind; the free space initial condition corresponds to vacuum or a continuum bulk liquid, definitely not a surface region [58]. Paper 1 contains a slightly more complex derivation where the presence of a mechanical surface is taken into account, but neither this can reproduce the effects from binding potentials of mean force like those displayed in Fig. 3.3.

To circumvent this, one can try to evaluate the coefficient of diffusion in a slightly different way. The coordinate displacement can be written as the integral of the velocity,

\[ r(t) - r(0) = \int_0^t v(t') dt', \tag{3.13} \]

\(^3\)For once actually an equation named after the one who derived it.
which is used to write the mean square displacement as

$$\langle |r(t) - r(0)|^2 \rangle = \int_0^t \int_0^t \langle \mathbf{v}(t') \cdot \mathbf{v}(t'') \rangle dt' dt''. \quad (3.14)$$

Through a rather technical but straightforward manipulation of the integration variables this can be simplified and rewritten on the form

$$\langle |r(t) - r(0)|^2 \rangle = 2t \int_0^t \left(1 - \frac{s}{t}\right) \langle \mathbf{v}(s) \cdot \mathbf{v}(0) \rangle ds. \quad (3.15)$$

This means the coefficient of diffusion alternatively can be calculated as

$$D = \lim_{t \to \infty} \frac{1}{3} \int_0^t \left(1 - \frac{s}{t}\right) \langle \mathbf{v}(s) \cdot \mathbf{v}(0) \rangle ds = \frac{1}{3} \int_0^\infty \langle \mathbf{v}(s) \cdot \mathbf{v}(0) \rangle ds, \quad (3.16)$$

where the last term is an integral of the velocity autocorrelation function (vac). This describes how long a particle remembers its velocity, i.e., how correlated the velocity is with its own value at an earlier time.

Equation 3.16 is called a *Kubo* or *Green-Kubo* relation, and similar equations exist for quantities like viscosity and conductivity. It is actually only a single example from a wide class of relations between transport phenomena and time-correlation functions.

In the limit of infinite times it is completely equivalent to Eq. 3.12, but not necessarily for shorter times. While the diffusion coefficient calculated from msd is dominated by the asymptotic behavior when $t \to \infty$, the main contribution to the vac integral comes from the short-time motions, as evident from Fig. 3.6. This can make a crucial difference in systems where the coefficient of diffusion is space-dependent, since the particles will not stay in a region of constant $D$ long enough for the asymptotic dependence in the Einstein relation to apply.

In paper 1, this difference between Eqs. 3.12 and 3.16 is utilized to study the motion of solvent atoms close to biological macromolecules. The dynamics in this interface region leads to some conceptual problems compared to a pure substance, since the two equations only are strictly valid in bulk matter. Close to a macromolecule the diffusion will be reduced by the presence of the surface, but more important the mean potential will render the bulk expressions invalid. The ultimate solution would be to solve the general equations for diffusion in an external potential (like the Smoluchowski or Fokker-Planck equations [61, 62, 63]), but this is a very complicated task when the potential is not exactly known but must be determined simultaneously with the diffusion coefficient. A much simpler idea, which we have adopted, is to compare the coefficients of diffusion obtained on the different scales in time and space probed by the msd and vac methods, respectively.

In both cases, the diffusion seems to be significantly reduced close to the protein surface. The calculations using the velocity autocorrelation function show that the diffusion normal to the surface decreases about 50% when going from bulk solvent to the layer closest to the macromolecule. The parallel diffusion is also reduced, but only by 30%. In contrast,
3.3. Diffusion, Mobility & Autocorrelations

**Figure 3.6.** Left: Mean square displacement of the oxygen atom in bulk liquid water at 298 K. The higher slope at times below 1 ps corresponds to the free flight dynamics (small motions where the molecule does not feel the presence of its neighbors). Right: Velocity autocorrelation function of the same water; note the different timescale. Although the initial correlation declines very fast there is a long time tail which usually is fitted analytically when performing the integral to obtain the diffusion coefficient. Both these approaches yield a diffusion coefficient $D \approx 3.5 \times 10^{-9}$ m$^2$/ps for the spc water model [59] in bulk liquid, slightly higher than the experimental value $D = 2.3 \times 10^{-9}$ m$^2$/ps [60]. This discrepancy is a well known fact, explained by the very simple geometry of the model molecule, lack of polarizability and possibly also the limited size of the systems usually simulated.

The merits of the Green-Kubo approach to determine coefficients of diffusion is often underestimated; it is usually claimed that the long time tails make it impossible to calculate an accurate numerical answer and that the MSD should be a better alternative. As we have shown, this is not quite the whole truth; it is easy to determine the initial part of the long time tail from a double logarithmic plot of the autocorrelation, and a suitable correction can then be applied analytically. While it is not necessarily a better method than the mean square displacement, it converges faster and makes it possible to probe fast local diffusion dynamics, for instance the way it is applied to determine lipid diffusion coefficients on short time scales in paper 4.

The mean square displacements indicate that the long time mobility is reduced by 60–65%, both for normal and parallel motions. This shows it is possible to separate the dynamics of the waters into a faster proper Brownian diffusion on scales of a few picoseconds and a slow mobility over hundreds of picoseconds and longer times. The first process essentially occurs inside the external potential minimum, but since it is relatively fast and involves smaller displacements it will be less affected by the potential barriers than the MSD, where the confinement in the potential restricts the mobility. If these findings are combined with our estimate of the binding potential influence from the previous section, we see that the global mobility of solvent water decreases a factor 5–10 close to the surface of a biomolecule, while the reduction in their fast Brownian dynamics is less than a factor of two.
Membranes

The most exciting phrase to hear in science, the one that heralds new discoveries, is not “Eureka!” (I found it) but “That’s funny…”

Isaac Asimov

4.1 Composition & Structure

All biological membranes, regardless of source, contain lipids as well as proteins. The ratio varies enormously; the myelin membrane in nerve cells contains only about 18% proteins, while our inner mitochondrial membrane reaches 76% [64]. The diversity in lipid composition of typical membranes is equally large. The most commonly found molecules are glycerophospholipids, e.g. phosphatidylcholine, phosphatidylethanolamine, phosphatidyglycerol, and the dimeric cardiolipin. Prefixes like “dipalmitoyl” indicate the length and type of the extended hydrocarbon chains attached. Sterol lipids like cholesterol constitute a large fraction of the plasma membrane of mammalian cells, and many membranes also contain carbohydrates bound to proteins or lipids. The reason for this heterogeneity is still not fully understood; some components seem to alter the packing properties of the membrane [65], but many lipids also participate in important biochemical pathways [66]. A functional biomembrane is thus a relatively complex structure. The integral membrane proteins are further very hard to crystallize, so it is only in the last few years we have begun to understand structure and function of components like ion channels [67, 68].

Despite this complexity, the fundamental physiochemical properties of membranes are all determined by the lipid molecules, so in order to better understand membranes it has proven very useful to start by studying pure lipid bilayer systems. It is these molecules that provide the basic barrier functionality, and paradoxically they are also responsible for the extreme flexibility of the membrane. This type of “soft” material properties are much
less studied than static ones like protein structure, since they are rather difficult to resolve experimentally. They are however very important to the biological activity; membranes are responsible for the way a cell maintains or changes its shape, and the way it can adhere to neighbors in multicellular organisms. A striking example is the familiar biconcave shape of erythrocytes, which simply is the mathematical surface that achieves the largest area to exchange oxygen with the surroundings, while simultaneously minimizing the free energy cost of bending the plasma membrane. The lipid bilayers provide many fascinating examples of how biologically important collective properties are explained by basic molecular interactions.

### 4.1.1 Bilayer Experiments & Simulations

Experimentally, it has been possible to obtain high resolution X-ray structures of individual lipids from crystalline bilayers at low temperatures and limited hydration (see, e.g., Refs. [69, 70]), but this is very different from their structure at biological conditions. In these crystals, the headgroups are packed and the lipid chains essentially completely parallel in extended *all-trans* conformations. As the temperature increases the system goes through a transition to a gel phase with more random headgroup orientations and some defects introduced in the chain configurations, but it is still a relatively ordered phase [71]. Before reaching biological conditions, there is a second phase transition to a liquid crystalline state, which is much more disordered with truly mobile lipids and large-scale motions of the surface. These fluctuations make it impossible to determine an atomic structure of biologically relevant bilayers, not because of poor resolution but simply because it does not exist!

However, it is still possible to determine statistical distributions of the densities for different atom types in the bilayer. This can be accomplished with neutron diffraction methods, since the scattering changes dramatically upon deuteration. By specifically replacing hydrogens with deuterium and compare the results it is possible to calculate the difference form factors due to the deuteration and fit gaussian distributions of the corresponding atom densities to the data [72]. The main problem with this approach is that it is difficult to use with realistic water contents; the work of Büldt and coworkers only reached 13–14 waters per DPPC lipid [72], considerably lower than the fully hydrated state. We should also keep in mind that the real atomic distributions are not gaussian, although it provides a fair approximation of the density peak locations that can be compared with simulation results like Fig. 2 in paper 3. Wiener and White have improved the experimental resolution significantly by using combinations of neutron and X-ray data [73], but their hydration level was limited to only 5.4 waters per lipid.

The lack of long-range order in liquid crystals compared to a real crystal severely limits the resolution of X-ray diffraction. It is usually only possible to resolve the first two diffraction peaks, resulting in very rough electron density profiles across the bilayer. There are two important effects that explain this. The first and relatively obvious one is the intrinsic disorder and significant local motions of the system, which will broaden the measured electron density [74]. The second effect comes from the X-ray measurements being performed on multilamellar stacks of bilayers to achieve a periodicity similar to that of a crystal, but for
bilayers it will rather be a smectic liquid crystal. Such a system exhibits large-scale undulations of the bilayers (like those we are able to observe in paper 2) that average the long-range order which would be present in a crystal. This implies that crystal scattering theory is not valid; the normal sharp peaks will be replaced by diffuse scattering distributions. Caillé has developed a theory that predicts power law tails for these peaks, making it possible to perform liquid crystallography on smectic phases [75]. Nagle and coworkers have applied this to correct X-ray data for the long range undulations [76], which resulted in the first structure determination of a fully hydrated liquid crystalline bilayer [77]. The liquid crystallography approach makes it possible to extract at least four X-ray diffraction peaks, which provides much better experimental accuracy for properties like electron density, bilayer thickness and area per lipid [78].

On the molecular level, Electron Spin Resonance (esr) measurements have made it possible to estimate the local ordering and orientation of the lipid chains by attaching spin probes to selected carbon atoms [79, 80, 81]. Since these probes typically consist of 5–10 heavy atoms and a number of hydrogens, there has however been objections that the results are seriously influenced by the probe itself. For this reason, most spectroscopic investigations of bilayers instead rely on Nuclear Magnetic Resonance (nmr) techniques [83, 84]. Due to the restricted motion in bilayers, the magnetic and electric second rank interaction tensors will be incompletely averaged and cause a quadrupolar splitting in deuterium nmr measurements [85]. This “defect” can be used to directly determine the average order in the lipid chains, if hydrogen atoms are selectively replaced with deuterium. The frequency separation between the two resonance lines of the quadrupolar interaction can be calculated to [86]

\[
\Delta \nu = \frac{3}{8} \chi (3 \cos^2 \vartheta - 1),
\]

(4.1)

where \( \chi = \frac{e^2 qQ}{h} \) is the quadrupolar coupling constant (167 kHz for the C–D bond in deuterium nmr), the brackets denote an average and \( \vartheta \) is the angle between the C–D bond and the external field. From this we can define the order parameter of the carbon position as

\[
S_{CD} = \frac{1}{2} (3 \cos^2 \vartheta - 1).
\]

(4.2)

This provides very useful information about the flexibility gradient in the bilayer interior [87, 88], and since \( S_{CD} \) is very sensitive to the lateral packing density it is also possible to relate it to the area per lipid [89]. Nmr can also be used to quantify the dynamics of the corresponding motions, which we will discuss in section 4.2.1.

There are several other experimental techniques like fluorescence, Raman, and IR spectroscopy that successfully have been applied to membrane systems, but they still do not produce a complete picture of the structure and motion of the bilayers. In contrast, computer simulations provide a very appealing alternative for this very type of collective dynamics, at least for relatively small temporal and spatial scales. The first computational membrane

---

*Our initial simulation studies on spin-marked lipids unfortunately seem to confirm this. Local chain dynamics slows down almost an order of magnitude due to the attached probe [82].*
Figure 4.1. A snapshot cross-section of a simulated bilayer consisting of 236 dppc lipids and roughly 6000 water molecules at 323 K, using data from paper 2 in this thesis. The waters are drawn as gray lines. Most of the lipids are shown with black lines, but a few molecules are displayed as space filling atom models to show the large fluctuations in the structure. The lipid chains are quite disordered and the bilayer behaves as a twodimensional liquid.

studies were performed on simplified model systems, with very limited detail in the headgroup and tails, and at first not even water surrounding the bilayer. A simulation of a monolayer was performed by Kox et al. in 1980 [90], while van der Ploeg and Berendsen [91] in 1982 published an 80 ps simulation of a bilayer system with 16 decanoate molecules on each side. The first more realistic simulations were performed a few years later by Bert Egberts and coworkers, who used explicit water molecules and atomic detail in their bilayer models [92, 93]. During the last decade, several different lipid systems have been investigated with full atomic detail, for instance dlpe [94], popc [95], dppc in liquid crystalline [96] and gel phases [97, 98], but also mixtures of phospholipids with e.g. cholesterol [99, 100, 101]. The attainable length and size of simulations have increased significantly, which has made it possible to reproduce true mesoscopic membrane properties (as reported e.g. in this thesis), and to introduce small membrane proteins in simulated bilayers [21, 102]. In contrast to the classical textbook picture with aligned headgroups and parallel hydrocarbon chains, these computer simulations highlight the much more disordered liquid crystalline structure of membranes, as shown by the snapshot structure in Fig. 4.1.

It is not trivial to perform accurate simulations in general, and membrane systems in particular are very sensitive to the choice of force field. The only way to assess results and ascertain that we are at least close to the truth is to compare with available experimental
4.1. Composition & Structure

![Graph showing electron density vs distance from bilayer center](image)

Figure 4.2. The electron density extracted from a simulation of a 64 dppc lipid bilayer agrees very well with the density calculated from the experimental form factors determined by Nagle et al. using liquid crystallography [77] (the smooth curve). The simulation data reproduces the experimental peak-to-peak spacing of 3.6 nm accurately. In a larger system, the long scale undulations would reduce the resolution significantly by averaging the data.

observations; the electron densities from X-ray diffraction and NMR order parameters provide rather stringent tests for dppc bilayers. Since the present simulations reproduce these properties accurately (as illustrated in Figs. 4.2 and 4.3), we have fair reason to believe that also the results not available from experiments should be reasonably correct.

4.1.2 Molecular Area & Surface Tension

The forces keeping a bilayer together are quite complex and a delicate balance between attraction and repulsion in different parts of the system. At extremely small concentrations even surfactants like lipids are separately soluble in water, but already at moderately small volume fractions they will aggregate and form some kind of interface, for example a bilayer as shown in Fig. 4.1. Unless suspended, bilayers in solution normally exist in the form of vesicles, spherical structures that can have radii as small as 10 nm. Depending on the type of lipid and temperature, the attractive and repulsive forces in the different parts of the system might not balance, and then the molecules would instead form micelles or some kind of hexagonal phase [104]. As the number of lipids is increased, more and more will go to the surface and the area per molecule, $A$, in the interface will decrease. This packing cannot continue indefinitely, though, and in practice there will exist a state of lowest free energy at some certain area per molecule, $A_0$. The system will keep adding lipids to the interface, decreasing the area per lipid, until this minimum value is approached. Beyond this, the molecules will maintain their packing density and the extra lipids will instead be accommodated by creating more interface, e.g. by increasing the vesicle size or bending the surface. The equilibrium state is referred to as a saturated membrane. Even when this
structure has equilibrated it will not be closely packed. The lipids exhibit large fluctuations and motions, and in the center of the bilayer their conformations remind more of a pure liquid hydrocarbon than of biological molecules. This random structure is confirmed both by computer simulations and the experimental order parameters in Fig. 4.3 which almost vanish at the end of the chain.

If we study the variation of the free energy with the area per molecule, this equilibrium area density of a bilayer is defined from [105]

$$\gamma = \left( \frac{\partial F}{\partial A} \right) = 0,$$

when the normal pressure and temperature are kept constant. If we force the system to a higher or lower area per molecule than the equilibrium value $A_0$ it is said to be subject to a nonzero surface tension $\gamma$. Since the first derivative per definition disappears at the minimum from Eq. 4.3, the free energy penalty for changing the area by compression or expansion is approximately given by

$$\Delta F \approx \frac{1}{2} F''(A_0) (A - A_0)^2.$$
If we introduce the experimental *area compressibility modulus* defined as

\[ K_A = A \frac{\partial \gamma}{\partial A}, \]  

the free energy variation with area can be written as

\[ \Delta F = \frac{1}{2} K_A \frac{A}{A_0} (A - A_0)^2. \]  

The area in the denominator should in principle be \( A \) instead of \( A_0 \), but in the region where the approximation with a harmonic variation of the free energy holds this difference is not significant. Most lipid bilayers are characterized by a very low area compressibility modulus, meaning they are easy to stretch. In papers 2 and 3 we study systems where we obtain values in the order of 250 mN/m. This corresponds to a compressibility \( 2 \times 10^{-3} \) atm\(^{-1} \), which is a factor 40 larger than the bilayers’ volume compressibility. In practice this mean we can neglect overall volume changes (although there will still be local fluctuations in volume and bilayer thickness) and define the average membrane thickness as

\[ d = V_L / A, \]

where \( V_L \) is the constant volume per lipid.

The optimal value of the equilibrium area \( A = A_0 \) for a membrane is determined from a balance between the molecular interactions in the system, but the details of these forces are difficult to study experimentally since they depend on the local distribution of lateral pressure inside the bilayer and in the water interface. We have approached this problem by implementing a calculation of a space-dependent pressure tensor in the molecular dynamics simulations according to the derivation in chapter 5. Using this (see paper 3) we find that the net surface tension is a sum of several large terms of opposite signs. The head-group region is contracting the bilayer, partly due to interactions between the lipid dipoles (20 mN/m), but the main contribution is a large entropic term originating from the hydration of the headgroups (80 mN/m). These will work as an effective barrier between the solvent waters and the hydrophobic interior of the bilayer. There are also forces in the chain region contracting the bilayer. Since the *trans* configuration of the chain bonds is slightly more favorable than the *gauche* states the chains will achieve a lower energy \(^2\) when they are extended and occupy a smaller area. This will also be favorable to the Lennard-Jones attractions, resulting in a surface tension of about 150 mN/m. However, the packing is extremely unfavorable from an entropic point of view, and these huge terms (-240 mN/m) more than cancel the energy, making the total pressure in the chain region expansive, despite the relatively low density there.

### 4.2 Dynamical Bilayer Properties

Few, if any, biological molecules are static structures, and the rate and extent of their motions are usually very important to their biological function. This is particularly true for

\(^2\)See section 2.1.1.
Chapter 4. Membranes

Lipid membranes. The Singer & Nicholson fluid mosaic model [12], and more recently computer simulations, have helped us understand that membranes are essentially microscopic seas of lipids with integral proteins floating around. One of the most fascinating properties of membranes is that this dynamics is not limited to the individual lipid molecules—the most important properties are instead collective mesoscopic effects of the system, like their bending and stretching elasticity. Even in the absence of external influence, a typical lipid bilayer will exhibit significant internal motions and large-scale three-dimensional thermal fluctuations of the membrane surface.

Many of these collective mesoscopic phenomena are crucial to the function of membranes; Paper 5 in this thesis for instance provides a direct illustration of the collective nature of bilayer stability. The initial separation of water and lipids is extremely fast, but it leaves water defects in the bilayer. Since it is very unfavorable to remove a single lipid or water from this defect, the pore can be metastable for $10-100$ ns, but once the transition state is reached it disappears in less than a nanosecond! It is quite probable that similar transient water pores also occur spontaneously as the result of fluctuations in equilibrium membranes, which would offer e.g. small ions a favorable permeation route independent of channel proteins [106]. Another striking example of the importance of bilayer dynamics is the possible role of plasma membrane fluidity and bending deformations in the outer hair cells in our ears. These exhibit electromotility, i.e., they can convert electrical energy to mechanical motion (as observed by Brownell and coworkers [107]). This is used for active filtering of sound, and enables the extremely sharp frequency discrimination of our hearing. Recently, Raphael et al. proposed a mechanism for this [108], where the electrical depolarization induces mesoscopic bending deformations of the membrane. When connected to the underpinning cytoskeleton, this could result in the observed length change of the outer hair cells, as illustrated in Fig. 4.4. This depolarization also seems to reduce the fluidity of the membrane by up to $51\%$ [109]; we still lack a complete understanding of such phenomena, but it illustrates the significance of bilayer dynamics.

4.2.1 Lipid Chain Relaxation

The fluidity of the hydrocarbon interior is probably the single most prominent feature of a liquid crystalline bilayer. As we discussed above (and in paper 3), it is the complex interactions of these chains that are responsible for the expansive pressure balancing the headgroup attraction in a typical membrane. The motions cover a very broad range, from picosecond vibrations to slow diffusion on the microsecond scale. NMR spectroscopy provides a way to study these processes experimentally and validate the dynamics of computer simulations, which in turn provide additional atomic detail and reveal the interactions responsible for relaxation—it is potentially a very powerful combination! In practice, available computing power has previously limited this type of comparison to the internal dynamics of single lipids combined with heavy modeling [110], but the $100$ ns simulation presented in paper 4 of this thesis makes a direct comparison of computer simulation and NMR lipid bilayer relaxation rates possible.

Without going too deep into details (Ref. [111] is a gold mine in this context), an NMR experiment subjects the sample to a fixed magnetic field $B_0$ that separates the energy levels
4.2. Dynamical Bilayer Properties

Figure 4.4. A hypothetical model for electromotile length changes and membrane curvature in the outer hair cells. Left: The plasma membrane is connected to “pillars” of about 30 nm length with unknown structure. These pillars bind to filaments of actin (thick beads) that run circumferentially around the cell and are cross-linked with spectrin molecules that run longitudinally. Right: When the cell is depolarized, the altered electrical potential across the bilayer will align the headgroup dipoles and induce a curvature, leading to the observed length change of the cell [108]. Figure adopted from Ref. [109] and reproduced with kind permission from William E. Brownell.

of the nuclei (their “spin”). The strength of the field is usually quoted in terms of Larmor frequency, i.e., the frequency of the electromagnetic quanta necessary to induce transitions between these levels, e.g., 600 MHz. By irradiating the sample with an electromagnetic field of this frequency it is then possible to alter the distribution of the spin states among the nuclei. When the second field is turned off, the system will slowly relax due to random molecular interactions and motions on scales of the Larmor frequency. By recording the changes in the magnetization of the system during the process we can determine the rate of this spin-lattice relaxation and gain valuable knowledge about the molecular dynamics of motions whose frequencies correspond to the Larmor frequency (time scales of nanoseconds).

The spin-lattice relaxation rate $R_1$ (reciprocal of the relaxation time) is directly related to the spectral density of the motions in the system. With deuterium NMR it is for instance possible to show [111, 112] that the relaxation rate can be calculated as

$$R_1 = \frac{3}{40\chi^2} [j(\omega_0) + 4j(2\omega_0)].$$

(4.7)

The experiments thus provide pointwise values of the spectral density of the relaxation process. Brown and coworkers have made extensive investigation of these relaxation rates at different frequencies in bilayer vesicles [114, 113, 115, 116], and showed that the chain relaxation is a very complex process with a broad distribution of motions [117]. In theory,

---

1Actually, this assumes the experiment is performed on vesicles so that all orientations of the external field $B_0$ are equally probable. A more detailed account for this and the related expression for $^{13}$C NMR can be found in paper 4 or the work of Brown [113].
one could assemble the whole power spectrum by repeating the experiments on a sufficient number of different Larmor frequencies. The reason why this is interesting is that the Wiener-Kinchin theorem (see, e.g., Ref. [62]) tells us the power spectrum is a simple Fourier transform of the full autocorrelation function of the corresponding motion, which provides a much more direct view of the chain dynamics. Unfortunately, the very limited range of accessible NMR frequencies currently prevents this. The idea of paper 4 is essentially to invert this procedure; by extending computer simulations significantly we are able to extract the full autocorrelation functions (compare with the discussion in section 3.3), and use a Fourier transform to obtain the power spectrum \( j(\omega) \).

It is thus possible to derive the NMR relaxation rates in the chains directly from simulation data and compare the computer simulation and NMR molecular dynamics in the nanosecond range, which poses a much harder test than simulated equilibrium bilayer properties. Our results in paper 4 indicate that the major part of the relaxation is due to trans/gauche isomerizations on time scales of \( 10^{-10} \) s in the hydrocarbon chains, superposed with a much slower overall reorientation of the lipid molecule as illustrated in Fig. 4.5. For an isolated chain, a dihedral potential like those illustrated in Fig. 2.2 would give rise to fast exponential relaxations, without any frequency dependence of the relaxation rate. The neighboring chains in the bilayer however restrict these motions and makes them significantly slower; most transitions will revert back to the initial state in a couple of picoseconds. There are several models that try to describe this dynamics in terms of coupled “kink” defects that can diffuse along the hydrocarbon chains [118, 119, 120], but...
it has not been possible to predict any form for the relaxation. In contrast, the present
computer simulations make it possible to completely separate the internal and overall
motions of the lipids, showing that the chain isomerization is actually non-exponential on all
scales, from picoseconds to tens of nanoseconds. Their autocorrelations instead decay as
power laws with exponents close to $-\frac{1}{2}$, explaining the $1/\sqrt{\omega}$ frequency dependence of
NMR relaxation rates observed by Brown.

### 4.2.2 Rotational & Translational Diffusion

On time scales above a few nanoseconds, the most significant contributions to the bilayer
relaxation come from transport phenomena like rotation of lipid molecules and transla-
tional diffusion over the surface if the NMR experiment is performed on spherical vesicles.
Additional possible effects include the rotation of whole vesicles [121] and long-range col-
lective fluctuations [122], although these are not accurately reproduced in computer simu-
lations due to the limited system size.

This ability of membrane components to reorient and move around laterally in the
system is a central theme of the fluid mosaic model. In paper 4, we directly extract the
rotational motions of the lipid molecule illustrated in Fig. 4.5. We find that the spinning
rotation around the long axis is purely diffusive and has a relaxation time of about $3$ ns.
The long axis also undergoes diffusion, but it is considerably more complicated since it is
also subject to an ordering potential from the rest of the bilayer. As shown by Szabo [112],
there will be a residual order parameter that enters the definition of the diffusive relaxation
(the interested reader can find the equations in paper 4). Using this, the simulations predict
a diffusion of the long axis that is almost an order of magnitude slower than the spinning
motion, which is in good agreement with the experimental NMR observations reported by
Mayer et al. for the similar DMPC lipid [123]. Essman et al. have earlier obtained values
in the same order of magnitude from a $10$ ns simulation at higher temperature and surface
area [124], but the values quoted by these authors are smaller since they do not include
the restricting potential and approximate lipids as free cylindrical Brownian rotors. Their
spinning diffusion is further mixed with a small contribution from the long axis motion,
which makes it slower than our result that comes from completely decomposed dynamics.

Several different experimental techniques have been used to investigate the lateral dif-
fusion of lipids, but the resulting values are not always consistent; methods like neutron
scattering that probe fast short-scale motions result in a diffusion coefficient in the order
of $10^{-6}$ cm$^2$/s$^{-1}$ [125], while spectroscopic methods such as NMR or fluorescence recovery
after photobleaching that probe long-range dynamics on micrometer scales give values of
$0.9\times1.3\times10^{-7}$ cm$^2$/s$^{-1}$ for DPPC at $323$ K, almost an order of magnitude smaller [126, 127].
It has been suggested that this discrepancy is explained by two different types of motion
being present in the bilayer [128]. The translational diffusion of a lipid molecule should
occur when a free volume of critical size appears in the immediate vicinity of the particle.
If the lipid moves, its original location will be free. A new particle could then move into
this space, but it is also quite possible that the first lipid moves back to its old position after
a short time. It is only the first alternative that produces an effective displacement on long
Figure 4.6. Lateral diffusion on different time scales illustrated with stroboscopic snapshots of the center of mass positions in the upper layer of a 64 DPPC lipid system. Top: Translations during the first 5 ns of simulation, with 10 ps separation between successive points. Bottom: Translations during 100 ns of simulation with 100 ps between points. All but two lipids are plotted in gray while the last two are black to better emphasize the typical diffusion length scale.
time scales, but since neutron scattering probes dynamics in the nanosecond range it might rather measure the “rattling” of lipids moving back and forth.

The translational coefficient of diffusion can be calculated from simulations by using methods very similar to those we discussed for water diffusion in section 3.3, but the bilayer is restricted to two dimensions and much longer trajectories are required since the lipids move slower. For this reason, the large-scale diffusion coefficient has previously not been reproduced in simulations. The length of the trajectory in paper 4 however makes it possible to probe this kind of very slow dynamics in the system. Figure 4.6 illustrates the center-of-mass motions in the 64 DPPC system for two different time scales. While there are significant displacements already after 5 ns, it is not until the 100 ns scale that we really simulate the liquid state of the system. The mean square displacement indeed decreases continuously, at least over the first 10 ns of the mean square displacement vs. time, but as reported in paper 4 the asymptotic value using data from 100 ns is in very good agreement with the large-scale experimental results. To test whether the short-scale faster motions are diffusive or not we extract the Green’s functions of Eq. 3.8 over different time intervals, which somewhat surprisingly results in perfectly gaussian distributions for all times above 10 ps. The model with a slow diffusion and faster rattling motions is thus not entirely accurate; the lipids exhibit proper diffusion on almost all scales of time!

Although we do calculate a faster diffusion for short intervals, it is only on the shortest scales we see an order of magnitude difference. On the nanosecond scale probed by neutron scattering we only observe an increase by a factor 1.5–2 compared to the asymptotic large-scale value. This is however in agreement with some more recent experimental investigations that predict lower translational diffusion coefficients, around $2 \times 10^{-7}$ cm$^2$s$^{-1}$, from neutron scattering.[129]

### 4.2.3 Curvature Deformations

The motions discussed in the preceding sections are all important parts of the bilayer dynamics, but they are all relatively localized effects; even the lateral diffusion is restricted to the bilayer plane. This final section will instead focus on some of the most general and complex collective phenomena; global three-dimensional deformation modes of the membrane surface, with length scales that can approach the millimeter range. The investigations and theory development in paper 2 in this thesis concern exactly this type of motions, and provides a beautiful example of complex collective dynamics emerging from individual simple atomic interactions. These deformations are denoted bending or curvature modes and the free energy changes associated with them are known as curvature free energies. We will shortly see that they can be divided into three main types of motions; individual molecules or a small number of lipids moving relative to the monolayer surface, relative motions of the monolayers, and finally the really global motions of the entire membrane.

The analysis of these motions requires a mathematical description of the bilayer, where we model the positions and conformations of the lipid molecules with an idealized geometrical surface. For processes where relative motions of the two layers are important this must be done separately in each layer, while the larger motions only depend on the average conformation of the whole membrane which then can be approximated as an in-
Chapter 4. Membranes

Figure 4.7. On scales below the monolayer correlation length $\lambda_0$, the dynamics is dominated by single and collective temperature excited protrusions, displacing lipids out of an average monolayer surface. They give rise to a restoring force similar to a microscopic surface tension, due to the increased lipid exposure to solvent. This is confirmed by the protrusion surface tension observed in our simulations; it corresponds reasonably well to the experimental free energy per unit area required to expose hydrocarbons to water.

The motions of individual lipids are caused by random thermal fluctuations, but their collective interactions exhibit clear patterns. On the smallest scales for which we can define a surface (ranging from roughly twice the interlipid distance up to a correlation length $\lambda_0$ of the same order as the bilayer thickness), the main feature is seemingly random normal displacements of molecules out of the average surface where its neighbors are located, as depicted in Fig. 4.7. These protrusions will expose more of the lipid to the solvent, and generate a restoring force proportional to the excess local interface area \[ \gamma_{prot} \]. We can estimate the total local surface per unit area in the $(x, y)$ plane by performing a Taylor expansion of the surface element from the conformation,

$$
\frac{dS}{dx}\approx 1 + \frac{1}{2} \left[ (u_x)^2 + (u_y)^2 \right] dx dy.
$$

The last part of this expression is the excess surface area; since the constant of proportionality will have the same units as a surface tension we designate it $\gamma_{prot}$, but it should not be confused with a real macroscopic surface tension in the bilayer. This leads to a free energy variation per unit projected area which varies with the surface form as

$$
\Delta F_{prot} = \frac{1}{2} \gamma_{prot} |\nabla u(x, y)|^2.
$$

The spectral distribution of the protrusions calculated in paper 2 corresponds to a microscopic tension $\gamma_{prot}$ of about 50 mN/m for the whole system, or roughly 25 mN/m in each layer. It is interesting to compare this to the 25 cal mol$^{-1}$Å$^{-2}$ required to expose a hydrocarbon surface to water in experiments [131, 132], which corresponds to a surface tension of
17 mN/m. This agrees quite well with the rough approximation from the simulations (it is a very small effect compared to the simultaneous mesoscopic motions). Long-wavelength protrusions should thus be easier to excite than shorter ones since the area exposed to water will be smaller. This is however only valid below the correlation length \( \lambda_0 \), where these displacements of individual lipids relative to the average surface is the dominant effect.

Above the correlation length \( \lambda_0 \), we have to account for the motions of the entire surface. On these scales, the protrusive motions will only be experienced as a microscopic roughening of the average surface; the dominant effects are instead peristaltic and undulatory motions. Peristaltic deformations \( u_p \) are antisymmetric displacements of the two layers (Fig. 4.8), while the undulations \( u_u \) involve symmetrical motions (Fig. 4.9). If the two surfaces are given by \( u_1(x, y) \) and \( u_2(x, y) \) we define

\[
\begin{align*}
  u_u(x, y) &= \frac{u_1(x, y) + u_2(x, y)}{2}, \\
  u_p(x, y) &= \frac{u_1(x, y) - u_2(x, y)}{2}.
\end{align*}
\]

(4.10) (4.11)

At zero surface tension, these modes are limited by the cost of bending the membrane, and we should thus be able to calculate the free energy required as a function of how curved the membrane is. In the Monge representation the average curvature is defined as \([104, 133]\)

\[
H = \frac{(1 + u_x^2) u_{yy} + (1 + u_y^2) u_{xx} - 2u_x u_y u_{xy}}{2\sqrt{(1 + u_x^2 + u_y^2)^3}} \approx \frac{1}{2} (u_{xx} + u_{yy}),
\]

(4.12)

where we in the second step have assumed that the deformations are relatively small, i.e., \( |\nabla u(x, y)| \ll 1 \). The free energy is a yet unknown function of this curvature, but to second order we can always express it as

\[
F_{\text{curv}}(H) = F_0 + \Delta F \approx F_0 + F_1 H + F_2 H^2.
\]

(4.13)
Figure 4.9. Symmetric out-of-plane perturbations of the bilayer which do not change its overall volume are defined as undulatory excitations. They are limited by the free energy cost of bending the bilayer. Since this decreases fast with increasing wavelength they will dominate the dynamics on large scales in the absence of external surface tensions.

For undulations of a symmetric system like a pure bilayer we can discard the $F_1$ term since the energy cannot depend on which way we bend the membrane. Further, all peristaltic modes of finite wavelength will produce equal regions of positive and negative curvature, and the single infinite volume-changing mode is not curved at all. This means we can assume the curvature excess free energy to be proportional to the square of the average curvature for the two layers,

$$\Delta F_{\text{curv}} = 2k_c H^2 = \frac{1}{2} k_c \left[ \nabla^2 u(x,y) \right]^2,$$

where we introduce the bending modulus constant $k_c$, which can be related to the pressure distribution and derivatives inside the membrane [134]. This applies both to undulatory and peristaltic modes, but depending on the details of the chain interactions their bending moduli may not be identical. For the simulations in paper 2 we calculate the undulatory bending modulus to $k_c = 4 \times 10^{-20}$ J (in very good agreement with the experimental values for similar lipids [135]) while the peristaltic is $k_d = 2 \times 10^{-20}$ J. This indicates it is easier to deform the monolayers in an antisymmetric way, at least on the limited scales of length accessible in simulations. The most probable explanation to this is that the symmetric undulatory deformations involve slipping of the two layers relative to each other [136].

The peristaltic modes will additionally experience a force restoring the membrane thickness to its average value. Since we essentially do not know anything about it, we use the simplest approximation possible and assume it to be harmonic, producing a free energy variation

$$\Delta F_{\text{thick}} = \frac{1}{2} k_e \left[ u_p(x,y) - \frac{d}{2} \right]^2,$$

where $d$ is the average thickness of the membrane. The thickness fluctuations in our simulations correspond to $k_e = 4 \times 10^{-21}$ J/nm$^4$.

---

Note that this is not necessarily the case for a biological membrane; the attached cytoskeleton breaks the assumed symmetry.
4.2. Dynamical Bilayer Properties

The total free energy variation for peristaltic motions is obtained as the sum of Eqs. 4.14 and 4.15, and the amplitude of the modes will be determined from an equipartition between the different degrees of freedom. When a macroscopic external surface tension is applied to the bilayer, it too will of course enter in the equipartitions, as discussed in detail in paper 2.

In that work, we find that the peristaltic modes dominate the global dynamics of systems smaller than about 10 nm. Above this scale, the thickness fluctuations are approximately constant with the limiting mesoscopic value determined by Eq. 4.15. The macroscopic motions of the membrane are thus predominantly undulatory oscillations, or damped deformations for short wavelengths. We are able to calculate these relaxation times for the two longest modes present in our systems and the results agree perfectly with mesoscopic wave equations [137]. Since these are developed from experimental results on millimeter scale it provides us with a coherent model for membrane dynamics covering almost 6 orders of magnitude in both space and time.
Defining Pressure & Local Pressure

Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906, by his own hand. Paul Ehrenfest, carrying on the work, died similarly in 1933. Now it is our turn to study statistical mechanics.

David L. Goodstein, States of Matter

5.1 What is Pressure?

The physical phenomena we refer to as pressure is actually a molecular effect, due to the average momentum transfer which results when a (real or imaginary) surface is bombarded by adjacent molecules in motion. The normal definition of pressure, as first introduced in section 2.2, is derived from the total net forces acting on the surface of a macroscopic part of a system instead of individual particles. It is thus a thermodynamic property, defined through statistical mechanics averages. This is valid in the thermodynamic limit, which essentially means macroscopic systems involving in the order of $10^{23}$ particles. This is not quite in the range of atomic level computer simulations, but a couple of thousand atoms is usually enough for the average to work, although it will make the pressure calculated in a simulation very noisy—it is quite common to have a pressure of 1 bar and fluctuations of a couple of hundred bars.

In physics, the pressure is commonly described as a $3 \times 3$ matrix (tensor) instead of a single scalar value

$$
P = \begin{pmatrix}
P_{xx} & P_{xy} & P_{xz} \\
P_{yx} & P_{yy} & P_{yz} \\
P_{zx} & P_{zy} & P_{zz}
\end{pmatrix}
$$
Each of these elements can be defined as the force per unit area that acts on the surface of an infinitesimal cubic volume that has edges parallel to the $x$, $y$, and $z$ axes. The first index refers to the normal axis of the plane on which the force acts, while the second denotes the direction of the force itself. The diagonal values are thus the “normal” pressures along the $x$, $y$, and $z$ axes, while the off-diagonal elements correspond to forces that try to deform the shape (in contrast to only changing the size) of the system. Using the initial analogy, this corresponds to a gradient of the particle velocity in the direction perpendicular to the imaginary surface. For an isotropic system at equilibrium without viscous forces these shear components are zero, but when they are significant the whole matrix is often called a stress tensor $\sigma$ instead, identical to the pressure but with the opposite sign. This nomenclature is especially common in solid mechanics or material sciences, and for practical reasons that naming is adopted in the following sections (for a more detailed discussion, consult e.g. Refs. [138, 139]).

As an interesting example, pure bilayers in the liquid crystalline phase do not have any shear rigidity since the lipids are free to move in the plane (the off-diagonal elements are thus zero), while real biological membranes often have a more solid underpinning, cytoskeleton, attached that significantly increases the rigidity of the cell wall.

In section 2.2 the pressure was derived as

$$p = 2 \frac{K - Z_{int}}{V}, \quad (5.2)$$

which can be used to introduce the stress tensor

$$\sigma = \sigma_K + \sigma_C, \quad (5.3)$$

with $\sigma_K$ being the contribution from velocities (kinetic stress tensor) and $\sigma_C$ the part from interparticle forces (configurational stress tensor),

$$\sigma_K = -\frac{1}{V} \sum_i m_i v_i \otimes v_i \quad (5.4)$$
$$\sigma_C = \frac{2}{V} \sum_i r_i \otimes F_i \quad (5.5)$$

### 5.2 Local Pressure

In some cases, the macroscopic ensemble description of the stress tensor might not be sufficient to understand the details of the microscopic interactions, e.g. in systems like lipid bilayers where some parts of the system want to expand while other are contracting it. In this section we will show that it is possible to deduce a more general definition, which is also valid on these microscopic scales, making it possible to resolve the local variations of the pressure, or equivalently, surface tension inside a simulation-sized system. This is not trivial since the forces used to define pressure usually have a longer range than the length scale on which the local pressure should be resolved.
The extension is nevertheless relatively easy to perform for the kinetic contribution to the stress tensor; since the kinetic energy is a plain sum over particle velocities it can simply be restricted to the particles located in the small area under study. The configurational/interaction part of the stress tensor is more complicated and has to be reconstructed from scratch, using the conservation of linear and angular momentum in physics.

In analogy to the earlier macroscopic definition of pressure we start by calculating the force from interactions with the surroundings and relate this to the internal stress tensor of a very small volume. Starting with a microscopic system composed of point particles, the first question is how the total external force acting on this part should be defined. Since nothing is known about the detailed interactions, the most reasonable choice should be the rate of change of linear momentum \( \dot{p} \) in a volume \( V \) (which is fixed in space) \([140, 141]\), since we by definition have \( \dot{p} = F \). The advantage with this definition is that the linear momentum itself is a sum of the particle momenta, defined as \( p_i = m_i v_i = m_i \dot{r}_i \). These are all \textit{located} inside the volume, removing the problem with the range of force interactions.

The total linear momentum in the volume can be written as a sum over the particles, which is then converted to an integral

\[
P_V(t) = \sum_{r_i \in V} p_i(t) = \int_V J(R, t) d^3R, \tag{5.6}
\]

where the function \( J(R, t) \) is the local density of momentum, constructed using delta functions located on the particle positions

\[
J(R, t) = \sum_i p_i(t) \delta[R - r_i(t)]. \tag{5.7}
\]

The rate of change of \( P_V \), i.e. the external force, is now obtained from the time derivative of the components of \( J(R, t) \),

\[
\dot{J}(R, t) = - \sum_i p_i [\dot{r}_i \cdot \nabla R \delta(R - r_i)] + \sum_i \dot{p}_i \delta(R - r_i), \tag{5.8}
\]

where the first term has been rewritten using the feature

\[
\frac{d}{dt} \delta[R - r_i(t)] = \dot{r}_i \cdot \nabla R \delta(R - r_i) = -\dot{r}_i \cdot \nabla R \delta(R - r_i). \tag{5.9}
\]

The derivative in Eq. 5.8 has the units of a force density in the volume. When inserted in the integral, the first term on the right side contributes the change of momentum due to particles crossing the boundary of the volume. The form this term is written on corresponds to a gradient of a second-rank tensor \( \sigma_K \),

\[
\sigma_K(R, t) = - \sum_i \frac{p_i \otimes p_i}{m_i} \delta(R - r_i) = - \sum_i m_i v_i \otimes v_i \delta(R - r_i), \tag{5.10}
\]
where \( \otimes \) is an outer product of the two vectors. Since this tensor bears a clear resemblance both in concept and look to the kinetic energy density used when defining macroscopic pressure it can be interpreted as the kinetic part of the local stress tensor. Shortly, this will also be proved mathematically. The rest of the right hand side in Eq. 5.8 can be written as

\[
\sum_i \dot{p}_i \delta(R - r_i) = -\sum_i \left[ \nabla_{r_i} \Phi(r_1, \ldots, r_N) + \nabla_{r_i} \varphi^{\text{ext}}(r_i) \right] \delta(R - r_i),
\]

(5.11)

where \( \Phi \) is the potential function for the particle interactions and \( \varphi^{\text{ext}} \) represents the influence from an externally applied field on the subvolume. The total force on the volume \( V \) from the external potential can be calculated separately using the second term in this expression, resulting in

\[
F^{\text{ext}}_V(R, t) = -\int_V n(R, t) \nabla \varphi^{\text{ext}}(R) d^3R,
\]

(5.12)

where \( n(R, t) \) is the particle density,

\[
n(R, t) = \sum_i \delta[\mathbf{R} - \mathbf{r}_i(t)].
\]

(5.13)

It is possible to rewrite also the remaining first part on the right side of Eq. 5.11 as the gradient of a tensor by using the translational invariance of the interparticle potential \( \Phi \),

\[
\sum_i \nabla_{r_i} \Phi(r_1, \ldots, r_N) = 0.
\]

(5.14)

This means the first term in Eq. 5.11 can be written as

\[
\sum_i \left[ \nabla_{r_i} \Phi(r_1, \ldots, r_N) \right] \delta(R - r_i)
\]

\[
= \sum_i \left[ \nabla_{r_i} \Phi(r_1, \ldots, r_N) \right] [\delta(R - r_i) - \delta(R - R_0)]
\]

\[
= -\nabla_R \left\{ \sum_i \left[ \nabla_{r_i} \Phi(r_1, \ldots, r_N) \right] \otimes \int_{C_{0i}} \delta(R - l) dl \right\}
\]

\[
= -\nabla_R \sigma_C(R, t),
\]

(5.15)

for an arbitrary point in space \( R_0 \), where the line integral is taken along any contour \( C_{0i} \) from \( R_0 \) to \( R_i \). The last step introduces the gradient of a configurational stress tensor \( \sigma_C \), defined as

\[
\sigma_C = \sum_i \left[ \nabla_{r_i} \Phi(r_1, \ldots, r_N) \right] \otimes \int_{C_{0i}} \delta(R - l) dl.
\]

(5.16)
5.2. Local Pressure

Using this, it is now possible to write Eq. 5.8 on the very simple form

\[ \dot{\mathbf{J}}(\mathbf{R}, t) = \nabla \sigma(\mathbf{R}, t) - n(\mathbf{R}, t) \nabla \varphi^{\text{ext}}(\mathbf{R}). \]  

(5.17)

The left hand side force density \( \dot{\mathbf{J}}(\mathbf{R}, t) \) can equivalently be interpreted as the gradient of the force per area, or pressure, apart from the sign. Since the second term on the right side accounts for the forces from external fields the interactions with the rest of the system are represented by the first right hand term. This is already written as the gradient of a tensor, and we have thus deduced a full local stress tensor \( \sigma = \sigma_K + \sigma_C \), where first part is the kinetic contribution and the second the configurational component.

This definition of microscopic stress is not unique, but it will depend on the exact choice of contours \( C_0 \). This is a consequence of the fact that the conservation of momentum defines the gradient of the stress tensor and not the tensor itself. This will unfortunately make the distribution of local stress ambiguous in regions with large variations in density and on scales below about one nanometer, where molecular fluctuations in density are apparent. Irving & Kirkwood [142] showed that this lack of uniqueness appears as a difficulty to specify exactly where the force is acting and simply suggest using straight lines as contours between the interacting particles. This might seem the most obvious and only correct choice, but it has been pointed out both by Harasima [143] and by Ono & Kondo [140] that in inhomogeneous systems, a completely planar liquid-vapor interface can be equally well described with slightly different stress tensors. Applied to general interfaces between two regions, this means the total surface tension as the integrated difference between normal and lateral components of the pressure tensor,

\[ \gamma = \int_z [P_N - P_L] \, dz \]  

(5.18)

is well defined, but the position of the corresponding surface of tension

\[ z_0 = \int_z z \ [P_N - P_L] \, dz \]  

(5.19)

has an inherent uncertainty—it could in principle be placed anywhere in the interface region by an (un)suitable choice of interaction contours. It is however straightforward to prove that the total stress of the system is independent of the integral contours [141], and in the absence of external fields both the force and torque (using the rotational invariance of the potential) acting on the volume \( V \) can be resolved on the bounding surface, yielding expressions compatible with derivations from the statistical mechanics pressure definition presented in section 2.2. According to ordinary macroscopic theory of elasticity [144] these results follow from an assumed short-range nature of the interparticle forces, but the results above show this is an unnecessary assumption since one only needs to demand translational and rotational invariance of the potential energy function, which hold regardless of the range of forces.

\(^1\)Well, everything is relative...
In contrast to the macroscopic stress tensor, this microscopic entity need not even be symmetric, as long as Eq. 5.15 is fulfilled. It is however always possible to use the non-uniqueness of the definition to choose a symmetric representation of the microscopic stress tensor [145].

5.3 Application to Central Pair-Additive Potentials

If the inter-particle potentials can be expressed in the form

$$\Phi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \sum_{i<j} \varphi(\mathbf{r}_{ij}), \quad (5.20)$$

with $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$, Eq. 5.15 can be written as

$$\sum_i \left[ \nabla_{\mathbf{r}_i} \Phi(\mathbf{r}_1, \ldots, \mathbf{r}_N) \right] \delta(\mathbf{R} - \mathbf{r}_i)$$

$$= \sum_{i<j} \nabla_{\mathbf{r}_i} \varphi(\mathbf{r}_{ij}) \left[ \delta(\mathbf{R} - \mathbf{r}_j) - \delta(\mathbf{R} - \mathbf{R}_0) \right]$$

$$= -\nabla_{\mathbf{R}} \left[\sum_{i<j} \frac{\varphi(\mathbf{r}_{ij})}{\mathbf{r}_{ij}} \otimes \int_{C_{ij}} \delta(\mathbf{R} - \mathbf{I}) d\mathbf{l}\right]. \quad (5.21)$$

This applies to all common interactions in molecular dynamics except long-range lattice summation algorithms like particle mesh Ewald [146]. Actually, all common lattice sums are still based on simple pairwise interaction functions, but due to the infinite sums it is not possible to attribute the force to a certain particle and location in space, making the local configurational stress tensor ambiguous. Until this is solved (and it might not ever be), it is not possible to determine local stress from lattice sums, but they must be replaced by long cut-offs in the simulations.

The local configurational stress tensor from these interactions is thus obtained as

$$\sigma_C = \sum_{i<j} \mathbf{r}_{ij} \frac{\varphi'(\mathbf{r}_{ij})}{\mathbf{r}_{ij}} \otimes \int_{C_{ij}} \delta(\mathbf{R} - \mathbf{I}) d\mathbf{l}. \quad (5.22)$$

For central pair forces it is further straightforward to define a symmetric stress tensor. By choosing the linear contours

$$\mathbf{l}(\mathbf{r}_i, \mathbf{r}_j) = \lambda \mathbf{r}_j + (1 - \lambda) \mathbf{r}_i, \quad 0 \leq \lambda \leq 1, \quad (5.23)$$

between interaction sites, the symmetric Irving/Kirkwood definition of the stress tensor [142] is obtained,

$$\sigma = -\sum_i m_i \mathbf{v}_i \otimes \mathbf{v}_i \delta(\mathbf{R} - \mathbf{r}_i) + \sum_{i<j} \mathbf{r}_{ij} \otimes \mathbf{r}_{ij} \frac{\varphi'(\mathbf{r}_{ij})}{\mathbf{r}_{ij}} \int_0^1 \delta \left[ \mathbf{R} - \lambda \mathbf{r}_j - (1 - \lambda) \mathbf{r}_i \right] d\lambda. \quad (5.24)$$
In practical calculations the delta function will be replaced with a small but finite volume element. Comparing this expression to the macroscopic pressure definition in section 2.2 we see that the local stress corresponds to the kinetic contribution from the particles inside the small volume studied and the projection of each virial term between pairs of particles on the subvolume studied, as illustrated in Fig. 5.1.

In paper 3 this local pressure definition is applied to a molecular dynamics simulation of a lipid bilayer to extract the internal surface tension profiles for different types of interactions. In this case, the subvolumes are chosen as thin slices in the (x, y) plane to determine normal and lateral pressures (and thus surface tension) inside the membrane as a function of the z coordinate. The ambiguity of the pressure definition on the shortest scales will limit the significant resolution to about 0.5 nm, but this is sufficient to separate the surface tension contributions from e.g. water, headgroups and lipid chains. It is particularly interesting to be able to relate the distribution of tension in the headgroup-water interface of a bilayer to the molecular conformations and interactions of the solvent molecules earlier discussed in section 3.2.

\footnote[i.e., the physical significance by which the local pressure can be related to the corresponding macroscopic quantity; the local pressure itself has a higher resolution.]
Bibliography


Comments on the Papers

Computers are useless. They can only give you answers.

Pablo Picasso

Solvent diffusion outside macromolecular surfaces

In this paper, we study the ordering and dynamics of water molecules close to a protein. This environment is quite different from a pure bulk liquid, and we find that the mean square displacement of equation 3.12 is not appropriate for calculating diffusion due to the influence from the surface. For short time scales, the square displacement is not linear, and in the limit of slow processes the solvent molecules will have diffused away from the surface. We deduce an analytical expression for the true mean square displacement outside a planar reflecting surface, but also suggest the alternative use of velocity autocorrelations to better resolve diffusion close to surfaces. Using this, we find that it is possible to separate both normal and lateral motions of solvent into local diffusive events on scales of a few picoseconds, and a slow long time dynamics that is governed by a reduction of mobility due to the mean potential from the macromolecular surface.

Mesoscopic undulations and thickness fluctuations in lipid bilayers from molecular dynamics simulations

This work comprises the largest atomic scale computer simulations of lipid membranes published to date, extending both spatial and temporal scales by almost one order of magnitude. It was made possible by the optimized molecular dynamics software presented in paper 6. These simulations enable us to investigate dynamics of bilayers on a mesoscopic scale previously not accessible to computer simulations, and we are able to determine collective properties like bending modulus in very good agreement with experimental observations. We revise the earlier theories of bilayer motions and further develop a corresponding model for the thickness fluctuations between the two monolayers in the membrane, and relate the predictions to simulation results. For the longest length scales probed in the simulations, this model agrees extremely well with mesoscopic wave equations for undulatory motions (which in turn have been confirmed by macroscopic experimental results). This provides a coherent theoretical framework for membrane dynamics on scales ranging from 0.25 mm
down to interlipid (almost atomic) distances, and times from milliseconds to 100 ps. We finally propose that the previously unobserved peristaltic dynamics might explain the system size dependence of area per lipid observed in bilayer simulations.

**Spatial and energetic–entropic decomposition of surface tension in lipid bilayers by molecular dynamics simulations**

Classically, pressure is a macroscopic property that is only defined as an average over billions of billions of atoms. It is still possible to define the average for a smaller system like those in a typical molecular dynamics simulation, but the fluctuations are much larger. In this work, we have extended this one step further by calculating the *local pressure* in different parts of the simulation system. While the definition is relatively straightforward, it is not trivial to apply in realistic systems with bond constraints, etc. The present article is the first work where it has been implemented for general atomic detail systems. This enables us to resolve local pressure profiles inside a bilayer, and thereby obtain the surface tension as a function of the normal coordinate (i.e., across the membrane). In contrast to macroscopic experiments, the computer simulations make it feasible to separate components from different types of interactions, and to determine energetic and entropic contributions. This provides a very detailed description of the dynamics keeping the bilayer together. We find that the system is contracted by an entropic term mainly due to hydration of the headgroups. The energetic interactions in the chains are also contractive, but they are canceled by a huge entropic contribution that makes the total surface tension in the chain region expansive. We also develop a theoretical model for the energetic part of surface tension from Lennard-Jones interactions in the lipid chains, which makes it possible to relate the chain energy to the experimentally measured phase transition enthalpy of the bilayer.

**Molecular dynamics simulation of NMR relaxation rates and slow dynamics in lipid bilayers**

Although computer simulations have developed rapidly, the accessible time scales have usually been much too short to enable direct comparison of the dynamics with spectroscopic experimental techniques like NMR. In this work, we present the longest molecular dynamics simulations ever performed on membranes, reaching 0.1 microseconds. This makes it possible to observe very slow reorientation motions of the lipid molecules in the bilayer, and to determine the relaxation spectrum of the chain C–H vectors. This is closely related to the experimental NMR relaxation, and we are able to calculate and reproduce not only approximate relaxation rates but also the experimentally observed variation with applied NMR Larmor frequency. This connects the very complex relaxation patterns observed in spectroscopic studies to details in the atomic motions not accessible to experiments. The computer simulation further enables us to decompose the relaxation into fast non-exponential internal dynamics in the lipid chains, and slower overall rotational diffusion of the molecules in the bilayer. We also determine values for the rotational and translational diffusion coefficients of the lipids; the length of this trajectory makes it possible to observe both the fast local diffusion and the slow long-range diffusion previously only measured in experiments.
Simulation of Spontaneous aggregation of phospholipids into bilayers

Previously, all computer simulations of realistic membranes have been started from pre-assembled bilayer and water configurations, constructed to match experimental properties. This paper presents the first series of atomic level simulations where the initial setups are instead completely random solutions of different types of lipids in water. We actually observe the collective phase transitions of the molecules into perfect bilayers on time scales of \(10^{-100}\) ns. Not only does it prove that a bilayer is the most stable state for the simulated lipids (i.e., the force fields and models are appropriate), but it also suggests a detailed aggregation pathway and provides approximate time scales for the different processes involved. We also propose that the reduction and destabilization of microscopic hydrophilic water pores spanning the bilayer is the rate limiting step of the whole aggregation process.

GROMACS 3.0: A package for molecular simulation and trajectory analysis

This article introduces the third major release of the parallel molecular dynamics software that has been developed cooperatively by my colleagues in Groningen, Uppsala and me. Compared to other simulation packages, the main feature is that we achieve extremely high and yet unparalleled performance. It is usually one, and in some cases two orders of magnitude faster than available alternatives. This has been accomplished by carefully tuning the code and avoiding all unnecessary calculations in the innermost loops of the program. We have also implemented assembly loops using the special multimedia hardware instructions of modern PC hardware. Although these are primarily aimed at the gaming and graphics market, we have proven the approach to be surprisingly effective also for this kind of scientific applications. The article also provides an overview of the many utility and analysis programs distributed with the package, and benchmarks to demonstrate the performance on both single processors and parallel computers.
Comments on the Papers