MASTER

Torsion profiling of antibody proteins through experiment and molecular simulation

van der Heijden, T.W.G.

Award date:
2015

Link to publication
Torsion profiling of antibody proteins through experiment and molecular simulation

Thijs van der Heijden
0745667

Supervisor:
Cornelis Storm

14 October 2015
Abstract

In this work, we attempted to address how we can use torsion profiling of proteins as an alternative protein identification method. We hypothesised that the distinction between different torsion profiles can be made and that the torsion profile of a protein is determined by the protein structure. Torsion experiments were performed on an immunoglobulin G-protein G complex using magnetic tweezers. Two distinct torsion profiles were found, which were associated with the protein complex and multiply bound beads. For the protein complex a torsion stiffening behaviour was found, at torsion angles of approximately 40 degrees. This is a first experimental indication that the torsion experiments are viable as a method to distinguish between torsion profiles.

After we had tested the feasibility of the technique, we evaluated how the torsion profile of a protein may be determined by its structure; an important step in the development of torsion profiling as a protein identification method, is to be able to understand and/or predict the behaviour of the molecules. Therefore, the torsional properties of immunoglobulin G were assessed specifically, by simulating this molecule subject to external torques, using a coarse-grained simulation method: Fluctuating Finite Element Analysis. Using this method we did not encounter the experimentally found torsion stiffening. For stronger torques the molecule overtwisted: a highly non-physical phenomenon which is the result of the absence of surface-surface repulsion. To prevent the overtwisting behaviour, a surface–surface repulsion interaction was implemented, based on a Lennard-Jones potential. This inhibited the overtwisting, yet the extra deformation did not cause a significant rise in the resistance to torques. Therefore, an alternative constitutive model was explored to describe the material, based on a Gent material. This model showed strong strain stiffening under linear extension, but no divergent torsional resistance was found under axial torque, for the explored torque magnitudes.

Nevertheless, we have shown that, while incorporating both the surface repulsion and Gent material in the model, we were able to identify the most strongly stressed regions within the molecule. We extracted the flexible hinge region of the immunoglobulin G from the full protein and preliminarily evaluated its behaviour subject to torques on a full-atom level, using molecular dynamics simulations. The first results from these simulations are promising: they show a torsion stiffening behaviour, albeit for larger torsion angles and with lower torsional spring constants compared to the experimental results. This may be explained by the fact that we do not consider the full molecule or the presence of a magnetic microparticle. The qualitative similarity between experimental and simulation data is an indication that we can evaluate the torsional behaviour of the protein by analysing its structure. Hence, we developed a scheme to investigate the kinetic behaviour of molecules under external forces on increasingly small time and length scales: we performed torsion experiments on a protein complex using magnetic fields on a fluid cell, where we used the magnetic labels to visualise the location and properties of the proteins. The antibody in the protein complex was then investigated using the coarse-grained simulation method. We used this method to identify the most strongly stressed regions within the antibody molecule, which were subsequently extracted from the full molecule and investigated on a full-atom level, using molecular dynamics simulations.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>Amino Acid</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>CoM</td>
<td>Centre of Mass</td>
</tr>
<tr>
<td>COOH</td>
<td>Carboxylic acid</td>
</tr>
<tr>
<td>cryoEM</td>
<td>cryo-Electron Microscopy</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxynucleic Acid</td>
</tr>
<tr>
<td>EDC</td>
<td>1-Ethyl-3-(3-Dimethylaminopropyl)Carbodiimide</td>
</tr>
<tr>
<td>EDM</td>
<td>Electron Density Map</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-Linked ImmunoSorbent Assay</td>
</tr>
<tr>
<td>ENM</td>
<td>Elastic Network Model</td>
</tr>
<tr>
<td>Fab</td>
<td>Fragment antigen-binding</td>
</tr>
<tr>
<td>Fc</td>
<td>Fragment crystallisable</td>
</tr>
<tr>
<td>FEA</td>
<td>Finite Element Analysis</td>
</tr>
<tr>
<td>FFFEA</td>
<td>Fluctuating Finite Element Analysis</td>
</tr>
<tr>
<td>GMR</td>
<td>Giant MagnetoResistance</td>
</tr>
<tr>
<td>GROMACS</td>
<td>GROningen MAchine for Chemical Simulations</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>LBM</td>
<td>Lattice Boltzmann Method</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular Dynamics</td>
</tr>
<tr>
<td>MES</td>
<td>2-(N-Morpholino)EthaneSulfonic acid</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-Buffered Saline</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PG</td>
<td>Protein G</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>RDF</td>
<td>Radial Distribution Function</td>
</tr>
<tr>
<td>SAXS</td>
<td>Small Angle X-ray Scattering</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris-Buffered Saline</td>
</tr>
<tr>
<td>Symbol</td>
<td>Meaning</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>A</td>
<td>Area</td>
</tr>
<tr>
<td>(A_i)</td>
<td>Surface (i)</td>
</tr>
<tr>
<td>B</td>
<td>Magnetic field</td>
</tr>
<tr>
<td>(B_w)</td>
<td>Weak magnetic field (magnitude)</td>
</tr>
<tr>
<td>C</td>
<td>Covariance matrix</td>
</tr>
<tr>
<td>(C_{ij})</td>
<td>Covariance matrix (element (i,j))</td>
</tr>
<tr>
<td>(c_m)</td>
<td>Proportionality constant (\tau(B))</td>
</tr>
<tr>
<td>(D_O)</td>
<td>Dimension of function space</td>
</tr>
<tr>
<td>d</td>
<td>Separation</td>
</tr>
<tr>
<td>E</td>
<td>Young’s modulus</td>
</tr>
<tr>
<td>(E_p)</td>
<td>Elastic vector</td>
</tr>
<tr>
<td>(e_n)</td>
<td>Eigenvector</td>
</tr>
<tr>
<td>(F)</td>
<td>Deformation gradient tensor</td>
</tr>
<tr>
<td>(F_{ij})</td>
<td>Deformation gradient tensor (element (i,j))</td>
</tr>
<tr>
<td>(F_{LJ})</td>
<td>Lennard-Jones repulsion force</td>
</tr>
<tr>
<td>(f)</td>
<td>Force (magnitude)</td>
</tr>
<tr>
<td>(f_0)</td>
<td>Equilibrium force (magnitude)</td>
</tr>
<tr>
<td>(f_{LJ})</td>
<td>Lennard-Jones pairwise force</td>
</tr>
<tr>
<td>(f_n)</td>
<td>Force per node</td>
</tr>
<tr>
<td>(f_n)</td>
<td>Force per node (magnitude)</td>
</tr>
<tr>
<td>(G)</td>
<td>Shear modulus</td>
</tr>
<tr>
<td>(h)</td>
<td>Box dimension</td>
</tr>
<tr>
<td>(I)</td>
<td>Identity matrix</td>
</tr>
<tr>
<td>(I_1)</td>
<td>Invariant of (F): (I_1 = \text{trace}(FF^T))</td>
</tr>
<tr>
<td>(I_m)</td>
<td>Value of (I_1) for which the strain energy diverges</td>
</tr>
<tr>
<td>(J)</td>
<td>Jacobian determinant of (F): (J = \text{det}(F) = V/V_0)</td>
</tr>
<tr>
<td>(J)</td>
<td>Jacobian determinant of (F): (J = \text{det}(F) = V/V_0)</td>
</tr>
<tr>
<td>(K)</td>
<td>Bulk modulus</td>
</tr>
<tr>
<td>(K_{pq})</td>
<td>Viscous (drag) matrix</td>
</tr>
<tr>
<td>(k_0)</td>
<td>Potential well steepness (spring constant)</td>
</tr>
<tr>
<td>(k_B)</td>
<td>Boltzmann constant ((1.3806488 \cdot 10^{-23} \text{ J K}^{-1}))</td>
</tr>
<tr>
<td>(k_s)</td>
<td>Spring constant</td>
</tr>
<tr>
<td>(k_t)</td>
<td>Torsional spring constant</td>
</tr>
<tr>
<td>(L)</td>
<td>Length</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Symbol | Meaning
--- | ---
$\Gamma(x)$ | Gamma function
$\gamma_{ij}$ | Shear in direction $i$, as function of $j$
$\delta_i$ | Deviation from average position
$\delta_{ij}$ | Kronecker delta (1 for $i = j$, 0 for $i \neq j$)
$\epsilon$ | Lennard-Jones interaction strength
$\epsilon_{ij}$ | Strain tensor (element $i,j$)
$\eta$ | Vector defining triangle
$\eta_i$ | Vector defining triangle (parameter)
$\lambda_i$ | Extension in direction $i$
$\lambda'_j$ | Contribution of node $j$ to Gaussian point $i$
$\mu_m$ | Magnetic moment
$\mu_b$ | Bulk viscosity
$\mu_s$ | Shear viscosity
$\xi$ | Vector defining triangle
$\xi_i$ | Vector defining triangle (parameter)
$\rho$ | Density
$\sigma$ | Stress tensor
$\sigma^e$ | Elastic stress tensor
$\sigma^t$ | Thermal stress tensor
$\sigma^v$ | Viscous stress tensor
$\sigma$ | Lennard-Jones interaction distance
$\sigma_{ij}$ | Stress tensor (element $i,j$)
$\sigma_m$ | Stress magnitude
$\sigma_{th}$ | Threshold stress
$\sigma_{\phi,0}$ | Standard deviation of $\phi$ without external $B$-field
$\tau$ | Torque
$\tau$ | Torque (magnitude)
$\tau_0$ | Equilibrium torque (magnitude)
$\tau_m$ | Magnetic torque
$\tau_n$ | Torque per node
$\tau_{max}$ | Maximum torque (magnitude)
$\phi$ | Torsion angle
$\phi_{down}$ | Maximum torsion angle in the negative rotation direction
$\phi_{eq}$ | Equilibrium torsion angle
$\phi_{max}$ | Maximum torsion angle
$\phi_0$ | Torsion angle at weak $B$-field
$\phi_{up}$ | Maximum torsion angle in the positive rotation direction
$\phi_a$ | Base function
$\omega$ | Frequency
$\partial_i$ | Partial derivative with respect to $i$: $\partial/j\partial i$
List of used terms and symbols
## Contents

1 Introduction  
   1.1 Immunoglobulin G  
   1.2 Simulation methods  
   1.3 Outline  

2 Torsion profiling: experiments and simulations  
   2.1 Experimental set-up  
   2.2 Torsion profiling  
   2.3 Experimental results  
   2.4 Introduction to coarse-grained molecular simulation  
   2.5 Simulating IgG torsion  

3 Surface-surface repulsion interaction  
   3.1 Lennard-Jones repulsion  
   3.2 Implementation  
   3.3 Testing steric repulsion  
   3.4 Steric repulsion in IgG torsion simulations  

4 Deformation, strain energy and stress  
   4.1 Introduction on relevant deformation quantities  
   4.2 Testing the Gent model: Linear extension force  
   4.3 Testing the Gent model: Axial torque  

5 Targeting torsional properties  
   5.1 Targeting strongly strained regions in the IgG molecule  
   5.2 Introduction to molecular dynamics  
   5.3 Preliminary MD simulations on the linker region  
   5.4 Simulation results  

6 Conclusions and outlook  
   6.1 Summary and conclusions  
   6.2 Outlook  
   6.3 Technological relevance  

Bibliography  

Acknowledgements  

A Experimental protocols  

B Derivation FFEA  

C Derivation Gaussian integration points  

D Deformation quantities
**Chapter 1**

**Introduction**

In this chapter, we will introduce the key molecule in this research: immunoglobulin G. This is a so-called antibody molecule, which plays an important role in the animal immune system. We will discuss its structure and key features, and describe the idea of biosensing, in which this molecule can be utilised as well. We will, more specifically, explain the concept of magnetic biosensors, in which, apart from detecting molecules, forces can be actively exerted on these molecules. Additionally, we will describe a few current (biological) simulation methods, which are able to evaluate structures on various time and length scales. Lastly, an outline of this thesis will be given.

### 1.1 Immunoglobulin G

Immunoglobulin G (IgG) is an antibody protein, which is present in the blood and tissue fluids of, among other animals, humans. This section will give a short overview of where IgG resides within the framework of the animal immune system and its molecular structure. Moreover, we will discuss the concept of biosensing and how IgG can be utilised in one type of biosensing device, the magnetic biosensor. The magnetic assay can furthermore be used to actively exert forces on the protein complex in the assay.

#### 1.1.1 The animal immune system

The immune system of animals is a complex system of cells and molecules which all have their own properties and purposes. All of the involved structures arise from stem cells in the bone marrow [1]. A schematic representation of part of the animal immune system is shown in figure 1.1. The hematopoietic stem cells in the bone marrow develop into two more specialised types of stem cells: a common lymphoid progenitor and a common myeloid progenitor. The former will then continue to develop into B- and T-lymphocytes, which are involved in the humoral and cell-mediated immune response, respectively. The humoral immunity attacks intruders by means of excreting molecules into the body.

![Figure 1.1: Strongly simplified representation of part of the animal immune system.](image)

---

[1] Reference to the source of the figure and the text.
Chapter 1. Introduction

Fluids (humor - Greek: χυμός, chymos: juice / sap), whereas the cell-mediated immunity involves activation of phagocytes, cells that attack the foreign structures by ingestion. The most important proteins in the humoral immunity are the antibody molecules, or immunoglobulins (Ig). These are large, roughly Y-shaped molecules produced by plasma B-cells, with the ability to specifically target other molecules. The two identical paratopes (the ends of the top branches of the Y) are able to specifically bind to the epitopes (the part recognised by the immune system) of the target molecules. This bond is of a key-lock type: not covalent, but solely based on non-specific interactions, such as van der Waals and electrostatic interactions.

Five types of Ig molecules are present in mammals (table 1.1 [2]): IgA, IgD, IgE, IgG and IgM, of which IgG is definitely (~75%) the most abundant in the human blood. It occurs at a concentration of around 10 mg/mL in human serum [3, 4], however this concentration does depend on factors such as age, sex and way of life.

Table 1.1: The types of mammalian antibodies and their properties [2].

<table>
<thead>
<tr>
<th>Type</th>
<th># Subtypes</th>
<th>Serum (%)</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>IgA</td>
<td>2</td>
<td>15</td>
<td>High levels in mucosal surfaces (usually as a dimer) and secretions (saliva, breast milk), protection against toxins, viruses and bacteria.</td>
</tr>
<tr>
<td>IgD</td>
<td>1</td>
<td>&lt; 0.5</td>
<td>Likely to perform a unique role in immune response [5], short half-life in serum.</td>
</tr>
<tr>
<td>IgE</td>
<td>1</td>
<td>&lt; 0.01</td>
<td>Lowest serum concentration, shortest half-life; associated with hypersensitivity, allergic reactions and response to parasitic worm infections.</td>
</tr>
<tr>
<td>IgG</td>
<td>4</td>
<td>75</td>
<td>Predominant isotope, longest serum half-life, transported through the placenta, participation in secondary immune response, neutralisation of toxins and viruses.</td>
</tr>
<tr>
<td>IgM</td>
<td>1</td>
<td>10</td>
<td>Usually pentameric, associated with primary immune response.</td>
</tr>
</tbody>
</table>

1.1.2 Structure of Immunoglobulin G

As immunoglobulin G (IgG) is the most prominent antibody in the blood, we are for the purposes of this research mainly interested in this type of antibody. Several representations of the molecule are depicted in figure 1.2. The IgG protein is shown schematically in figure 1.2a. This schematic representation is roughly valid for any type of mammalian antibody (in monomeric form). The structure consists of two identical heavy chains and two identical light chains. Both types of chains have constant and variable regions, of which the latter form the specific antigen-binding site: the paratope. The class of the antibody (A, D, E, G or M) is determined by the structure of the heavy chain. The two major regions of the antibody, the Fab (Fragment antigen-binding) and Fc (Fragment crystallisable) regions, are connected by a rather flexible hinge region, which in the case of IgG contains disulfide bonds that connect the heavy chains.

An atomic structure of an IgG from a house mouse is depicted in figures 1.2b and 1.2c. This model is taken from the Protein Data Bank (PDB: 1IGT [6]). In figure 1.2b, the light chains (A and C, 214 residues) and heavy chains (B and D, 444 residues) can be distinguished by their colour. The IgG has a molecular mass of approximately 148.7 kilodaltons (2.47 · 10⁻²² kilograms). As can also be seen from the differently coloured atomic structure in figure 1.2c, the structure consists of three "bulky" regions (the three branches of the Y-shape) with a thin linker (hinge) region in between. This becomes more apparent in the secondary structure in figure 1.2d.

As the three bulky regions contain several folded structures (such as beta sheets), these are likely to be significantly more rigid than the non-folded amino acid chains in between, which would give rise to a relatively flexible hinge region. This could be interpreted as a very useful feature for the
antibody to fulfil its role in the immune system: a very flexible center would allow for a high variance in motion of the branches. In other words, due to the flexible linker region, the protein would be able to explore a large volume with its paratope, increasing the probability of detecting and binding a foreign structure. When this molecule is being exposed to forces, such as thermal fluctuations, we thus also expect the overall motion of the molecule to concentrate mainly in this flexible hinge region.

1.1.3 IgG in biosensing applications

Due to their specificity, antibodies such as IgG can, aside from their prominent role in the immune system, also be put to use in biosensing devices. Biosensors are self-contained devices which can measure analytes in test samples, by using biochemical or biological mechanisms as a recognition method [7]. A widely used example of a biosensor is the blood glucose sensor, which uses glucose oxidase to detect the concentration of glucose in the blood. As glucose occurs in the blood in a rather high concentration (around 5 mmol/L [8]), the detection of the molecules is fairly straightforward and small errors in the detected values are not problematic. However, various physical conditions are diagnosed by detecting protein levels in body fluids (blood, saliva, urine), which generally occur in much lower concentrations.
Chapter 1. Introduction

One example of such a protein is *cardiac troponin*, present in heart cells [9]. In case of myocardial infarction (a heart-attack), heart muscle cells are damaged or die, which results in the release of troponin into the blood. The rise of troponin concentration in the blood is then used as an indicator for diagnosis. For the sake of the patient, early diagnosis is of vital importance. However, the concentration of troponin in the blood shortly after such an event is in the order of picomolars \(10^{-12} \text{ mol/L}\) [10], which is approximately a billion times smaller than the healthy concentration of glucose in the blood. Conventional detection methods do not work and a transition has to be made to other detection principles.

Apart from the low concentration detection needed for biosensors, a few other requirements have to be taken into account. A few demands could be assigned to such “Point-of-Care” devices, in which the analysis is done at the time of care:

- detect concentrations in the order of (sub-)picomolars;
- measure directly in complex fluids (such as blood plasma, urine or saliva);
- finish analysis in the order of seconds to several minutes;
- detect in small sample volumes.

All of these desired properties are achieved by IgG in the body as a prominent link in the immune system, detecting foreign structures in complex body fluids and at extremely low concentrations. Therefore the possibilities of borrowing its abilities for biosensing purposes are explored; several commercial biosensors nowadays exploit the “specific binding” properties of immunoglobulin in their detection method [11]. The most general method of detection is by formation of a *sandwich assay* in the biosensor, where the target protein is being *sandwiched* between two antibody proteins, of which one is attached to a surface and the other is labelled in some manner (see figure 1.3). Examples of such labels are fluorescent molecules, catalyst molecules (ELISA: Enzyme-Linked ImmunoSorbent Assay) or magnetic particles. These labels can then be detected optically [12, 13] or, in the case of the magnetic particles, by using detection methods based on the magnetic properties of the particles, such as giant magnetoresistance (GMR) [14–16].

One type of biosensing device using magnetic detection is a *magnetic biosensor*, in which the properties of the IgG are combined with the ability to remotely manipulate the motion of the antibody molecules by applying magnetic fields. Its principle and set-up are schematically visualised in figure 1.4. Imagine the protein we are trying to detect has two different epitopes (binding to paratopes of IgG-1 and IgG-2). We assume that having both epitopes is a unique property for a protein. In order to detect this protein, a sensor is built which consists of a fluid cell, of which one surface is coated with a certain type of IgG (IgG-1). Magnetic beads are suspended in the cell, which are coated with the other type of IgG (IgG-2) (figure 1.4A).

The complex fluid (containing many different proteins, among which the target protein), is brought into the biosensor (figure 1.4B). We allow the suspension to mix by diffusion and/or by actively “stirring” the fluid by moving the beads around using magnetic fields. Being able to actively manipulate the...
Figure 1.4: An example of a magnetic biosensor - 

A: One surface of the sensor is coated with IgG-1 and magnetic beads coated with IgG-2 are suspended in the sensor fluid cell; B: the complex fluid is brought into the sensor, containing many different proteins, of which only the target protein has epitopes to bind to IgG-1 as well as IgG-2; C: The suspension is mixed, allowing the proteins to bind to their specific antibodies (if present); D: Using magnetic fields, the beads are pulled towards the coated surface, allowing the formation of a sandwich assay between bead and surface; E: Using opposite magnetic fields, the unbound beads are pulled away from the surface; F: The sensor is flushed, leaving only the bound beads behind. The number of beads eventually bound to the surface is a measure for the concentration of target proteins in the original complex fluid.

motion of the beads speeds up the mixing process considerably. As sensing occurs at micrometre scales, we are dealing with very low Reynolds numbers, so mixing by diffusion alone is very slow. While mixing the target proteins bind with their "epitope 2" to the IgG-2 on the magnetic beads (figure 1.4C). Then, by applying magnetic fields, the beads (and with them, the target proteins) are pulled towards the sensor surface (figure 1.4D). As the beads reach the sensor surface, the target proteins will bind to the IgG-1 on the surface with their other epitope, forming the aforementioned sandwich assay.

After the binding process, opposite magnetic fields are used to pull the excess of beads away from the coated surface, while the beads in the sandwich assays remain bound to the sensor surface (figure 1.4E). The fluid cell is then flushed, removing all excessive molecules and beads and leaving
behind only the bound beads on the surface (figure 1.4F). The number of remaining beads on the sensor surface is a measure for the concentration of target proteins in the original complex fluid. The magnetic properties of the beads thus allow us to actively manipulate the motion of the proteins through the sensor fluid. However, the application of forces to these molecules by magnetic fields could also be used in a different type of experiments, which have a more fundamental interest. Exerting forces on molecules and analysing their response has been the subject of several studies on stretching of DNA [17–19], twisting of DNA [19, 20] and stretching of proteins [21–23]. These experiments are performed to, for example, understand the compactification, transcription, replication and translation of DNA, interactions with proteins, or the structure of proteins [24]. Recently Van Reenen et al. (2013) reported an alternative method to profile proteins; by investigating their resistance to torsion as function of the torsion angle [25].

In this research, we attempt to address how this torsion profiling of proteins can be used as an alternative protein identification method. We hypothesise that a distinction can be made between different torsion profiles and that the torsion profile of a protein is determined by the protein structure.

In order to explore the viability of this profiling method, we will firstly describe in more detail the torsion experiments and discuss the results (see chapter 2). Subsequently, we will make a transition from these experiments to coarse-grained simulations of antibody proteins, to evaluate the behaviour of these molecules with a higher level of detail. Understanding the torsional properties of molecules on a molecular level is an important step towards torsion profiling as a new protein identification method. The remainder of the research is focussed on the adjustments made to the coarse-grained model in order to account for the intramolecular interactions within the antibody molecules. Eventually, in chapter 5, we will show the results of a few preliminary atomistic simulations as well, to analyse the influence of the internal structure of the antibodies. A more detailed outline of this thesis is given in section 1.3.

1.2 Simulation methods

As this thesis mainly describes the simulation of proteins such as immunoglobulin G, we will give a short overview of several existing simulation methods. The current (biological) simulation methods available can roughly be divided into three regimes [26–28]: the relatively small nanoscale, the relatively large macroscale and the intermediate mesoscale. Nanoscale simulations deal with materials on a molecular or atomistic level, with a (sub-)nanometre and (sub-)nanosecond resolution. Processes on these scales include photosynthesis and protein folding (though folding occurs at longer time scales, which can reach $\sim 1$ ms or longer [29, 30]). A simulation method widely used on this level is Molecular Dynamics (MD) [31], which considers the molecules to be classical collections of atoms, which are simulated under influence of thermal forces. Due to the vibrations between atoms, which occur at femtosecond time scales, the typical time step is in the order of femtoseconds. This makes reaching longer (e.g. protein folding) time scales computationally very expensive. An example of coarse-grained nanoscale methods are Elastic Network Models (ENM) [32], which represent the molecules as a collection of nodes, connected by springs. By performing a normal mode analysis on these network models, one finds the bulk motions of the network. This eliminates the problem of simulating on small time scales (as in MD), but it is impossible to account for non-specific interactions such as electrostatics. The macroscale considers processes (almost) visible to the naked eye, such as the function of biological tissues and organs. The length and time scales are found on a millimetre and millisecond or longer scales. The physics on this level is mostly dealt with using continuum mechanics models, in which all atomistic or molecular information is averaged into macroscopic parameters, such as elasticity and viscosity. Thermal influences are accounted for by adjusting the parameters as function of the temperature. One widely used continuum mechanics model is Finite Element Analysis (FEA) [33],
which discretises the continuum material into several elements, on which the continuum equations are evaluated. The region in between these two extremes, the mesoscale, is the regime in which we are, for the purposes of this research, the most interested. The mesoscale is the base of large molecules, such as some proteins (molecular motors, antibodies), with processes on length scales of tens or hundreds of nanometres and time scales of tens or hundreds of nanoseconds. This level also usually deals with low Reynolds number fluids and polymer physics. A simulation method on this scale is **Fluctuating Finite Element Analysis** (FFEA) [34, 35], which we will introduce in chapter 2, or the Lattice Boltzmann Method (LBM) [36], which is used to evaluate fluid flows. Instead of describing fluid particles, this method calculates velocity probability distributions and the collisions of these distributions at nodes on a fixed lattice, representing the particle flow.

A schematic overview of the relevant length and time scales in the three regimes is depicted in figure 1.5.

---

![Figure 1.5: Overview of the currently available (biological) simulation methods.](chart.png)
1.3 Outline

As mentioned in the end of section 1.1, in this research we attempt to address how torsion profiling of proteins can be used as a protein identification method. Therefore, we will give an insight on the performed torsion profiling experiments, as well as the evaluation of the structures using molecular simulation. In more detail:

- In chapter 2, we will describe the experimental assay, the measurement set-up and the results of the torsion experiments on the IgG-PG complex. Subsequently, a transition will be made to a coarse-grained simulation method, *Fluctuating Finite Element Analysis* (FFEA), which we used to perform simulations on the IgG protein. The concept of the method will be described very shortly and we discuss the results of a few simulations on IgG subject to thermal forces. We will continue by describing the simulations on the torsion of IgG, where the thermal fluctuations were turned off, as these do not contribute to the equilibrium torsion angles.

- Chapter 3 will describe the concept and implementation of surface-surface repulsion interactions into the simulation method, in order to prevent non-physical phenomena during the simulation. We will show the results of both a test simulation and a torsion simulation on IgG.

- Chapter 4 will discuss shortly the quantities involved with the deformation of a continuum material: the deformation gradient tensor, the potential energy – or strain energy density – and the Cauchy stress tensor. We will introduce an alternative material model, in order to incorporate torsional stiffening, and show the results of comparative test simulations on the initial and the new material model.

- Chapter 5 will describe a method to target specifically the strongly stressed regions within a molecule, from the coarse-grained simulations. We extracted a strongly stressed region from the molecule and performed a few preliminary atomistic torsion simulations on this region alone. We indicated how to use these simulations to calculate the torsional properties of this region of the molecule and will discuss the similarities and differences between the simulation and experimental data.

- Eventually, in chapter 6, we will summarise the main methods and results of this study and draw relevant conclusions. Furthermore, we will give an outlook for future research and shortly discuss the technological relevance of this work.
To address the question of how to use the torsion profiling as a method to distinguish between different proteins [25], we will firstly describe in more detail the experimental scheme in which the protein complex was investigated. The used assay, the experimental set-up and the results will be described. Subsequently, a transition is made to a coarse-grained simulation method, in order to evaluate the response to torques of IgG, which is often present in such assays, specifically. Understanding and being able to predict the outcome of torsion experiments is an important step towards using the experiments as an identification method. We will therefore shortly introduce the coarse-grained model and discuss the results of simulations of the torsional behaviour of IgG.

2.1 Experimental set-up

We will show the experiments performed in order to probe the torsional properties of a IgG-protein G (IgG-PG) complex, using magnetic tweezers [17, 24]. Protein G (PG) is a protein which binds to the Fc region of IgG [1, 37]. Torsion experiments on IgG-PG had already been performed in [25]. For a detailed description of the used experimental protocols, we refer to Appendix A. The IgG-PG complex was prepared in a fluid cell, as illustrated in figure 2.1.

![Figure 2.1](image-url)  
*Figure 2.1: The fluid cell used in the torsion experiments: IgG was physisorbed to the surface, after which the non-specific binding to the surface was blocked by covering the free surface with BSA. Superparamagnetic microparticles (beads), coated with PG and fluorescent labels, were suspended in the fluid (PBS buffer) and bound to the glass surface through the IgG-PG complex. The non-bound beads eventually sedimented to the bottom surface.*

The sample to be investigated was created as follows: an IgG solution was incubated on a glass cover slip and the free surface was blocked with Bovine Serum Albumin (BSA) solution. We coated superparamagnetic microparticles with a diameter of 2.8 μm with two kinds of protein: protein G and streptavidin:

- Protein G is a cell surface protein from the group G Streptococcal bacteria [1, 37]. As it binds to the Fc region of IgG, more control over the orientation of the assay between microparticle
and surface is achieved; ideally, the particle would be bound as shown in figure 2.1: the PG on the bead would be attached to the Fc region of the IgG, which would, in its turn, be bound to the surface through the Fab region. Naturally, it was also possible for the IgG to orient itself differently, e.g. to bind lying flat on the surface. However, using the aforementioned directional preference, the probability of forming the assay in the desired manner was increased.

- **Streptavidin** is a protein with high affinity to the protein biotin. It was used to bind the fluorescent microspheres, which are biotin-coated, to the beads. These fluorescent labels have a diameter of 0.2 μm and work in the yellow-green part of the light spectrum; their peak excitation and emission wavelengths are at 505 and 515 nm, respectively [38].

The magnetic beads were then incubated on the same surface as the IgG, subject to a low, stationary magnetic field of 2 mT. This in order to align the magnetic moments of the beads parallel to the surface, which would maximise the applied torque in the torsion experiment. During the incubation, the beads were allowed to bind to the surface through the IgG-PG complex. We subsequently closed the fluid cell and flipped the sample, causing the non-bound beads to sediment to the bottom: only the bound beads remained close to the top surface.

The fluid cell illustrated in figure 2.1 was then investigated by using a combination of bright-field and fluorescence microscopy, while immersed in demineralised (demi) water. The fluid cell was placed in the centre of a quadrupole magnet, consisting of four coils controlled by a function generator, and subjected to (rotational) magnetic fields. The microscope was focussed on the top surface of the fluid cell, to only detect the bound particles, and equipped with a CCD camera, which was connected to a computer. This allowed us to visualise and analyse the data digitally. The set-up is depicted in figure 2.2.

![Figure 2.2: The experimental set-up: the fluid cell in figure 2.1 was investigated subject to magnetic fields, by using a combination of bright-field and fluorescence microscopy, immersed in demi water.](image)

With a fluorescence light source, the fluorescent microspheres in the fluid cell were illuminated and they appeared as bright regions on the beads. We needed a fluorescent marker to distinguish the labels from the larger beads, in order to be able to keep track of the rotation of the latter due to the magnetic forces acting on them. As we looked at a two-dimensional projection of such a bead, we tracked the motion of a bright spot (the emitting fluorescent label) on a larger, darker circle with a bright centre (the magnetic bead). A typical image from the microscope is shown in figure 2.3. Upon applying a rotational magnetic field of various field strengths (varying between 4 and 24 mT) in both rotation directions, the bead rotated with the field due to the magnetic torque. We tracked
2.2 Torsion profiling

We will now describe a method to extract the torsional spring constant $k_t$ from the torsion angle of the bead $\phi$, in order to analyse the torsional properties of the protein. Upon applying a rotational magnetic field, the bead started rotating with the field due to the applied torque $\tau_m$ [25]:

$$\tau_m = \mu_m \times B,$$

with $\mu_m$ the magnetic moment of the bead and $B$ the magnetic field. At a certain point, due to the torsional stiffness of the protein complex, the rotating bead could not keep up with the external field anymore; the resistance to torsion had become too strong: equal to the externally applied torque. The magnetic bead remagnetised and “clicked” back to a smaller rotation angle. After some time, the magnetic field caught up, and the bead again rotated with the field.

We studied this clockwork motion, or rather the maximum angle the bead reached (and with it, the maximum applied torque). The ratio between the maximum applied torque $\tau_{\text{max}}$ and the maximum torsion angle $\phi_{\text{max}}$, is defined as the instantaneous torsional spring constant $k_t$ of the protein complex:

$$k_t(\phi_{\text{max}}) = \frac{\tau_{\text{max}}}{\phi_{\text{max}}}, \quad (2.1)$$

since at the maximum torsion angle, the torque exerted by the protein was equal to the maximum externally applied torque.

As we repeated this experiment for different magnetic fields $B$, we could explore the torsional spring constant $k_t$ as function of the torsion angle $\phi$. This function $k_t(\phi)$ may be unique (or distinguishable) between different proteins or protein complexes: this experiment could target the torsional properties of single proteins or protein complexes and eventually be used to identify proteins by their torsion profile.

The maximum magnetic torque $\tau_{\text{max}}$ applied to the bead was dependent on the applied magnetic field: it was, for sufficiently small external fields (the regime which we remain in), linearly proportional to the applied magnetic field $B$ [25]:

$$\tau_{\text{max}}(B) = c_m B, \quad (2.2)$$

with $c_m$ a proportionality constant. It has been shown that even at equal magnetic fields $B$, the magnetic torque on the bead would not be equal for each bead, but showed a standard variation of

Figure 2.3: Typical image of a superparamagnetic microsphere with fluorescent label through the fluorescence microscope.

the rotation of the magnetic bead in order to target the torsional properties of the IgG-PG complex. Occasionally, the bead tended to non-specifically bind to the surface through the fluorescent label. Luckily, this was easily distinguished by the fact that the bright spot (the label) was completely stationary under rotational magnetic fields.
28 percent [25]. Therefore, this proportionality constant $c_m$ was estimated for each individual bead by considering the rotation of the bead without external field ($B = 0$) and at a weak field ($B = 4 \text{ mT}$). Since we assumed to be in the linear spring regime for weak fields (i.e. $k_t$ was equal for both cases), the estimation of $c_m$ was done as follows:

### 2.2.1 Calculating the applied magnetic torque

If no external magnetic field is applied, the torsion angle $\phi$ averages to zero. However, using the variance of the angle (according to the equipartition principle, where every degree of freedom contains $\frac{1}{2}k_BT$ of energy) the following must be valid:

\[
U = \frac{1}{2}k_t\phi^2; \\
\langle U \rangle = \frac{1}{2}k_t\langle \phi^2 \rangle = \frac{1}{2}k_BT; \\
\rightarrow k_t = \frac{k_BT}{\langle \phi^2 \rangle} = \frac{k_BT}{\sigma^2_{\phi,0}}. \tag{2.4}
\]

Expression (2.3) describes the harmonic spring potential connected to the torsional spring constant. The expectation value of the energy contained in this spring equals $\frac{1}{2}k_BT$. By considering the variance of the torsion angle $\langle \phi^2 \rangle$, we derived the torsional spring constant $k_t$ for small angles, as in equation (2.4). $\sigma_{\phi,0}$ represents the standard deviation of the angle $\phi$ without external magnetic field.

We applied a weak magnetic field $B_w$ ($B_w = 4 \text{ mT}$), where we still assumed to be in the linear spring regime: $k_t$ was assumed to be equal between the zero and weak field cases. We combined expressions (2.1) and (2.2), and could derive an expression for the torsional spring constant at this magnetic field $B_w$:

\[
k_t = \frac{c_m B_w}{\phi_w}, \tag{2.5}
\]

where $\phi_w$ is the maximum torsion angle for this weak field $B_w$.

This resulted in two expressions for $k_t$ ((2.4) and (2.5)), from which an expression for the proportionality constant $c_m$ could be extracted:

\[
c_m = \frac{k_BT}{\sigma^2_{\phi,0}} \frac{\phi_w}{B_w},
\]

and from this, an expression for the maximum magnetic torque $\tau_{\text{max}}(B)$:

\[
\tau_{\text{max}}(B) = \frac{k_BT}{\sigma^2_{\phi,0}} \frac{\phi_w}{B_w} B.
\]

### 2.3 Experimental results

In order to calculate the torsion profile of the IgG-PG complex, we tracked the torsion angle $\phi$ of the magnetic particle, subject to rotating magnetic fields with magnitude $B$. The magnitude was varied in steps of 4 mT, between 4 and 24 mT, and the field rotated with a frequency of $\omega = 0.4 \text{ Hz}$. Every 10 seconds, the rotation direction was changed.

A typical measurement of the torsion angle $\phi$ as function of time $t$ is depicted in figure 2.4. As we can see from figure 2.4, the bead indeed rotated with the external field and started clicking to a maximum torsion angle $\phi_{\text{max}}$. In order to accurately determine this angle, the extrema in each rotation cycle were averaged, as depicted in figure 2.5. After the extrema had been calculated, the maximum angles $\phi_{\text{up}}$ and minimum angles $\phi_{\text{down}}$ at each magnetic field were averaged. Subsequently, we divided the difference between the maximum and the minimum by two to arrive at the actual maximum torsion angle $\phi_{\text{max}}$: $\phi_{\text{max}} = \frac{1}{2} \left( \phi_{\text{up}} - \phi_{\text{down}} \right)$. This calculation is done in order to correct for possible offsets in rotation angle (where the bead is not at $\phi = 0$ at the start of the rotation).
2.3. Experimental results

Figure 2.4: The torsion angle $\phi$ as function of time $t$. The vertical bars indicate different magnetic fields $B$.

Figure 2.5: The torsion angle $\phi$ as function of time $t$, with the average and standard deviation of the extrema of the torsion angle indicated.

The torsion angle $\phi_{\text{max}}$ as a function of magnetic field $B$ is shown for five different beads in figure 2.6. From the data, the proportionality constant for the magnetic torque $c_m$ was derived as well, as described in section 2.2, where the temperature was taken at room temperature: $T = 298$ K. The results are shown in table 2.1.
Chapter 2. Torsion profiling: experiments and simulations

Figure 2.6: The maximum torsion angle \( \phi_{\text{max}} \) as function of the magnetic field \( B \), for five different beads.

Table 2.1: Values for the standard deviation of the angle at zero field \( \sigma_{\phi,0} \), the torsion angle at weak field \( \phi_w \) and the derived proportionality constant for the magnetic torque \( c_m \), for the five different beads.

<table>
<thead>
<tr>
<th>#</th>
<th>( \sigma_{\phi,0} (\degree) )</th>
<th>( \phi_w (\degree) )</th>
<th>( c_m \left(10^{-16} \text{ N m T}^{-1}\right) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.823</td>
<td>4.165</td>
<td>3.627</td>
</tr>
<tr>
<td>2</td>
<td>2.139</td>
<td>6.387</td>
<td>0.823</td>
</tr>
<tr>
<td>3</td>
<td>1.826</td>
<td>6.128</td>
<td>1.083</td>
</tr>
<tr>
<td>4</td>
<td>1.657</td>
<td>21.943</td>
<td>4.712</td>
</tr>
<tr>
<td>5</td>
<td>0.761</td>
<td>2.143</td>
<td>2.183</td>
</tr>
</tbody>
</table>

As we found estimated values for the proportionality constant \( c_m \), we calculated the maximum torsion angle \( \phi_{\text{max}} \) as function of the torque \( \tau_m = c_m B \). The result is shown in figure 2.7. From figure 2.7, it becomes apparent that the bond probed in the experiments was likely the same for beads 1 and 5, and for beads 2, 3 and 4. If we now consider a different representation, namely the torsional spring constant \( k_t \) versus the torsion angle \( \phi \), this becomes even more clear; see figure 2.8. In figure 2.8a, we see that indeed bead 1 and 5 showed very similar torsional properties. A strong torsion stiffening occurs at relatively small angles, which indicates firstly that we have probed a different type of bond between the bead and the surface. Secondly, it is likely that this was a multiply bound bead, which accounts for the stiff torsional behaviour. For beads 2, 3 and 4 we see in figure 2.8b that for small angles, the found values for \( k_t \) are of the same order of magnitude as reported in [25]. This is an indication that indeed the IgG-PG complex has been probed. We see a weak torsion stiffening at larger angles (\( \sim 40^\degree \)), but to explore a possible divergence angle, larger torsion angles (with stronger magnetic fields) would need to be investigated.

Using the properties of the magnetic labels, we have found a way to explore the torsion profile of the IgG-PG complex. A value for the torsional spring constant \( k_t \) at small angles has been found and a gradual increase in stiffness is shown as larger torsion angles are reached. From the five beads (five bonds between bead and surface) examined, we can already distinguish two significantly different behaviours: the strongly torsion stiffening bond, likely to be related to a multiply bound bead, and
the weaker, less stiffening bond, probably the IgG-PG complex. This is a first experimental indication that the method can be applied to identify or distinguish the types of bonds between surface and bead: for the purposes of the research a promising result.

Figure 2.7: The maximum torsion angle $\phi_{\text{max}}$ as function of the magnetic torque $\tau_m$, for five different beads.
Figure 2.8: The torsional spring constant $k_t$ as function of the torsion angle $\phi$.

2.4 Introduction to coarse-grained molecular simulation

Now we have tested the viability of the technique to distinguish between torsion profiles, we attempt to understand this torsional behaviour from the protein structure. We therefore make a transition to a coarse-grained simulation method in order to simulate the behaviour of only the immunoglobulin G protein, subject to external torques. As IgG is often present in the protein complexes formed in biosensing devices, this molecule was analysed separately using aforementioned coarse-grained simulations: we attempted to evaluate one of the links in the magnetic assay, in order to be able to use this data in the future analysis of such an assay. If we can understand the torsional behaviour of a molecule by analysing its structure, the evaluation of the torsion profile of each link in the assay can be done more accurately.

The simulations were performed using Fluctuating Finite Element Analysis (FFEA) [34, 35]. FFEA is a coarse-grained, mesoscale simulation method, developed at the University of Leeds, United King-
2.4. Introduction to coarse-grained molecular simulation

and based on Finite Element Analysis (FEA) [33]: a method to numerically solve boundary value problems for differential equations. FEA is used in a wide range of research fields, for example mechanical engineering and fluid dynamics.

In order to simulate a protein using FFEA, it is treated as a continuum material (parametrised by macroscopic parameters, such as density and bulk modulus), as opposed to a collection of atoms: the full-atom model (as described in the Protein Data Bank (PDB), see chapter 1) is converted into a coarse-grained volume mesh, as illustrated in figure 2.9. For more details on the mesh creation, see Appendix B.3. Considering such a continuum material results in a sharp reduction of the number of degrees of freedom, which allows for a faster simulation. The advantage is twofold: one would be able to reach longer time scales and/or to simulate multiple molecules. The consequence of the coarse-graining is, however, that all information on the individual atoms, such as charge and bond strength, is lost. FFEA was used under the assumption that the overall shape of a protein determines its function, rather than the atomistic structure. It has been indicated (see Appendix B.4 and [39]) that the conversion from an atomistic structure to a discrete volume mesh is justified: a volume mesh resembles the dynamics of an atomistic model rather accurately.

2.4.1 FFEA in a nutshell

FFEA is used to numerically solve Cauchy’s continuum equation of motion:

\[
\rho \left( \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} (\nabla \cdot \mathbf{u}) \right) = \nabla \cdot \mathbf{\sigma},
\]  

which is, in essence, the continuum equivalent of Newton’s second equation of motion. It is governed by the velocity field \( \mathbf{u} \) and stresses \( \sigma \) within the material, and macroscopic material parameters, such as the density \( \rho \), bulk and shear viscosities \( \mu_b \) and \( \mu_s \), respectively, and the bulk and shear moduli \( K \) and \( G \). As FFEA is used to evaluate molecules on the mesoscale, on which thermal fluctuations still play an important role in the dynamics of the protein, a thermal stress tensor \( \sigma^t \) is added to the conventional elastic and viscous stress tensors \( \sigma^e \) and \( \sigma^v \), respectively:

\[
\sigma = \sigma^e + \sigma^v + \sigma^t, 
\]  

where we used the Kelvin-Voigt material model for a viscoelastic material [40]. This extra fluctuating stress term gives rise to the name of the simulation method: Fluctuating Finite Element Analysis. The elastic stress tensor \( \sigma^e \) is dependent on the deformation of the material and results from a compressible Mooney-Rivlin material model [41, 42]. This tensor will be discussed in more detail in chapter 4. The viscous stress tensor \( \sigma^v \) is a tensor which includes all internal viscous (drag) forces within the molecule. After evaluating equation (2.6) in the finite element approximation, we arrive at
the following expression:

\[ M_{pq} \frac{\partial v_q}{\partial t} + K_{pq} v_q = E_p + N_p, \]  

which is another representation of Newton’s second law, in matrix form: \( M_{pq} \) is a mass matrix, describing the distribution of mass throughout the structure. \( K_{pq} \) is a drag matrix, describing the dissipation through viscous effects. The elastic force vector \( E_p \) describes the forces due to elasticity and \( N_p \) is a thermal noise vector. For a derivation of equation (2.8), we refer to Appendix B and [34].

### 2.4.2 Performing FFEA simulations on IgG

We applied the FFEA method to the IgG molecule, where we let the molecule move freely, without any constraints, at a temperature of \( T = 298 \) K. An overview of the chosen simulation parameters is shown in table 2.2. A typical FFEA simulation of the IgG molecule is visualised in figure 2.10. As we can see from figure 2.10, the molecule behaved as would be expected from a protein subject to thermal forces: the fluctuation and rotation of the molecule are due to the added thermal stress \( \sigma^t \) to the total stress tensor \( \sigma \), representing the Brownian motion of the molecule, which would usually also be encountered in e.g. molecular dynamics simulations. If we would analyse the variance of the motion of the molecule, we would be able to align the material parameters of the protein to experimental data: the volume explored by the molecule is a measure for the stiffness of the material.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>( \rho )</td>
<td>1500 kg m(^{-3})</td>
</tr>
<tr>
<td>Bulk modulus</td>
<td>( K )</td>
<td>444.4 MPa</td>
</tr>
<tr>
<td>Shear modulus</td>
<td>( G )</td>
<td>307.7 MPa</td>
</tr>
<tr>
<td>Bulk viscosity</td>
<td>( \mu_b )</td>
<td>1 mPa s</td>
</tr>
<tr>
<td>Shear viscosity</td>
<td>( \mu_s )</td>
<td>1 mPa s</td>
</tr>
<tr>
<td>Time step</td>
<td>( \Delta t )</td>
<td>1 fs</td>
</tr>
<tr>
<td>Temperature</td>
<td>( T )</td>
<td>298 K</td>
</tr>
</tbody>
</table>

Figure 2.10: FFEA simulation of the IgG molecule.
2.5 Simulating IgG torsion

As we are, for the purposes of this research, mainly interested in modelling torsional properties of molecules, the thermal noise in our structure was turned off for all subsequent FFEA simulations. Since the molecules were forced externally, the thermal noise would only slightly affect their dynamics. We were only interested, however, in maximum or equilibrium torsion angles, which are not at all affected by the thermal noise. Moreover, turning off the noise in the structure results in cleaner simulation data, without “unnecessary” noise. This made targeting properties such as torsion angles and steric repulsion kinetics (chapter 3) more straightforward.

We simulated the IgG molecule, without thermal influences, subject to an external torque \( \tau \). To model the experimental situation, where the molecule was adhered to a surface (see section 2.1), the ends of the top two branches of the IgG molecule were pinned to their respective positions, as were the nodes that define the torsion axis. The nodes on the torsion axis were fixed to ensure a clean rotation of the molecule subject to torque. The pinned nodes are shown as red squares in figure 2.11.

![Figure 2.11: The pinned nodes on the IgG molecule in the torsion simulations, indicated by the red squares. The dark red region at the bottom indicates the region on which a torque \( \tau \) is applied.](image)

The molecule was twisted by applying a torque \( \tau \) on the bottom region of the molecule (see figure 2.11). Or rather, a force \( f_n \) was applied to all 299 mesh nodes within this region. This force \( f_n \) was calculated from the torque per node \( \tau_n \) according to the standard definition:

\[
\tau_n = r \times f_n \Rightarrow f_n = \frac{\tau_n \times r}{r^2}.
\]

Torque magnitudes per node between \( \tau_n = 1 \cdot 10^{-22} \text{ Nm} \) and \( \tau_n = 8 \cdot 10^{-21} \text{ Nm} \) were explored, summing up to a total torque \( \tau \) between 29.9 and 2392.0 pN nm. We performed the simulations while building up the torque to a value of \( \tau_0 \) quadratically in time \( t \), until a time \( t_b \):

\[
\tau_n(t) = \begin{cases} 
\left( \frac{t}{t_b} \right)^2 \tau_0 & t < t_b, \\
\tau_0 & t > t_b, 
\end{cases}
\]

in order to avoid sudden strong deformations of the material. The total (maximum) applied torque \( \tau \) is therefore given by \( \tau = \sum \tau_n \). We chose \( t_b = 7.5 \text{ ns} \), at 75 percent of the total simulation time of 10 ns.

A typical simulation of the twisting antibody is depicted in figure 2.12. We see in figure 2.12 that the antibody protein gradually twisted in time, until a certain equilibrium torsion angle \( \phi_{eq} \). The torsion angle \( \phi \) as function of time \( t \) is shown in figure 2.13. In figure 2.13 we see that up to a total torque \( \tau \) of 598 pN nm, the molecule rotated nicely and reached an equilibrium after the build-up time of \( t_b = 7.5 \text{ ns} \). From the equilibrium torsion angle \( \phi_{eq} \) and the applied \( \tau \), a torsional spring constant \( k \),
Chapter 2. Torsion profiling: experiments and simulations

Figure 2.12: The torsion of the IgG molecule subject to an applied torque \( \tau = 299 \ \text{pN nm} \).

Figure 2.13: The torsion angle \( \phi \) as function of the time \( t \).

could be calculated, as given in equation (2.1). The torsional spring constant \( k_t \) as function of the equilibrium angle \( \phi_{eq} \) is depicted in figure 2.14. For the weakest torques (smallest torsion angles), no torsion stiffening occurred, as seen in figure 2.14. The torsional spring constant remained constant with increasing torsion angle, which indicates that we were dealing with a linear torsional spring property of the material here. Quantitatively, \( k_t \) did not resemble the experimental value, yet, this could be adjusted by re-parametrisation of the material. Furthermore, the modelled situation does not entirely match the experimental situation, e.g. the influence of the magnetic bead is disregarded. For higher torques, we see that the molecule did not produce a sufficiently large resistance to the torque and it overtwisted: it remained rotating and, instead of reaching an equilibrium torsion angle, the molecule relaxed its internal stresses by twisting back through its own surface; as shown in figure 2.15. We can also see this relaxation of the internal stresses in the potential energy of the molecule, which is shown in figure 2.16. The tinted area in figure 2.16 is zoomed in to show that at the end of the simulation (at a constant external torque \( \tau \), according to equation (2.9)), the molecule kept (over)twisting at a constant rate, which resulted in a periodic behaviour of the potential energy curve. This overtwisting is a highly non-physical phenomenon, which did not result in reliable data. It is likely due to the fact that the surfaces of the molecule do not “feel” each other, which we will try to correct for in chapter 3.
2.5. Simulating IgG torsion

Figure 2.14: The torsional spring constant $k_t$ as function of the equilibrium torsion angle $\phi_{eq}$.

Figure 2.15: Overtwisting of the molecule under a torque $\tau = 1.495$ nN nm.

Figure 2.16: Potential energy of an overtwisting molecule. The tinted area shows a periodic behaviour of the potential energy, where the molecule keeps (over)twisting at a constant rate.
We have seen how in experiments, the torsional properties of proteins can be probed. It has been shown that the properties of different bond types can be distinguished (two distinctively different behaviours). To model the twisting behaviour of one protein in the assay, IgG, in more detail, a transition to a coarse-grained simulation model was made to analyse IgG under torsion. The results of these simulations did not directly show torsion stiffening: for weak torques, the behaviour was linear, for stronger torques, the molecule overtwisted: a highly non-physical behaviour, which is likely due to the lack of surface-surface interactions of the molecule. Therefore, in chapter 3, we will describe how we incorporated a surface-surface repulsion interaction into the FFEA model, in order to try to prevent the material from overtwisting.
In order to deny the surfaces to penetrate each other (section 2.5 and [39]), some sort of surface-surface (steric) repulsion interaction needed to be included into the model. This repulsion should become effective for surfaces of the protein (which are a distance $r$ apart, see figure 3.1), nearing each other within a certain separation $r_c$, the cut-off separation. Several methods to account for the repulsion between surface faces can be imagined, of which one will be described here: a Lennard-Jones repulsion. This method is similar to a widely used repulsion interaction between two hard spheres in e.g. molecular dynamics simulations.

3.1 Lennard-Jones repulsion

In order to characterise the repulsion interaction between two objects, the standard Lennard-Jones potential [43] was used:

$$U_{LJ}(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right] = \epsilon \left[ \left( \frac{r_c}{r} \right)^{12} - 2 \left( \frac{r_c}{r} \right)^{6} \right].$$

where $\sigma$ is the interaction distance and $\epsilon$ the interaction strength. In contrast to the hard sphere repulsion, the given potential, and with it the interaction strength $\epsilon$, represents an energy per squared area (length$^4$), effectively making it a surface-surface interaction energy per unit area of both surfaces. This potential accounts for both the steric (Pauli) repulsion (power 12) and van der Waals attraction (power 6). As we did not desire to include the attractive part, we cut off this potential at its minimum, at $r_c = 2^{1/6} \sigma$ (see figure 3.2), and only the repulsive part of the potential remained:

$$U_{LJ}(r) = \begin{cases} 
4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right] + \epsilon & r < r_c \\
0 & r > r_c 
\end{cases}$$
Chapter 3. Surface-surface repulsion interaction

The potential was shifted by a factor $\epsilon$ in order to ensure a zero steric repulsion energy at distances greater than the cut-off distance.

The total repulsion force $F_{LJ}$ between two surfaces is in fact an integral over both surfaces, which sums the pairwise forces $f_{ij}$ between all points $p$ and $q$ on both surfaces 1 and 2, respectively, as illustrated in figure 3.3. The pairwise force $f_{ij}$ can be represented by a derivative of the standard Lennard-Jones area potential of (3.1) to the distance $r$:

$$f_{ij}(r) = \mp \frac{dU_{LJ}}{dr} = \pm \frac{4\epsilon}{r} \left[ 12 \left( \frac{\sigma}{r} \right)^{12} - 6 \left( \frac{\sigma}{r} \right)^{6} \right] \hat{r},$$

where $\hat{r}$ is a unit vector in the direction of the separation vector: $\hat{r} = \frac{p-q}{|p-q|}$. The total repulsion force then becomes:

$$F_{LJ} = \int_{A_1} \int_{A_2} f_{ij}(r)dA_1dA_2, \quad (3.2)$$

where $A_1$ and $A_2$ are the surfaces 1 and 2, respectively. As equation (3.2) generally cannot be solved analytically, we used a Gaussian quadrature scheme for triangles, where the integral was
3.2. Implementation

approximated by a weighted sum of the values at \( N \) specified Gaussian integration points; the surface-surface interaction was basically converted into a number of particle-particle interactions [27]:

\[
F_{LJ} = \int_{A_1} \int_{A_2} f_{LJ}(r) dA_1 dA_2 \approx \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \lambda_i \lambda_j f_{LJ}(p_i, q_j).
\] (3.3)

Here \( l \) counts over all three nodes of one surface face (the total surface force was being distributed amongst the three nodes), and \( i \) and \( j \) count over all Gaussian integration points on a surface. \( w_i \) is the weight of Gaussian integration point \( i \), and \( \lambda_i \) is the contribution of node \( l \) to Gaussian integration point \( i \) (its \( l \)th barycentric coordinate). \( f_{LJ}(p_i, q_j) \) is the pairwise force between Gaussian points \( i \) and \( j \) on positions \( p_i \) and \( q_j \), respectively. The weights \( w_i \) were calculated in such a way, that \( \sum_i w_i = 1 \) and weights were equal for points with permutations of the same barycentric coordinates \( (\lambda_1, \lambda_2, \lambda_3) \).

The derivation and construction from the barycentric coordinates of the Gaussian points are discussed in detail in Appendix C. The construction is schematically depicted in figure 3.4.

![Figure 3.4: Surface face with six Gaussian integration points. The coordinates of its nodes are depicted by \( x \). The fourth Gaussian point is taken as an example for the construction: the barycentric coordinates \( \lambda_i \) represent the contribution of every node \( l \) to the construction of Gaussian point \( i \). Gaussian points with the same colour are points with permutations of the same barycentric coordinates, and thus have the same weight \( w_i \).](image)

3.2 Implementation

As we found a method to represent the steric repulsion between two surface faces, namely as a discrete sum of point-point forces, distributed over the nodes of the face (3.3), we needed to implement it into the simulation method. As each surface face was treated to be independent, a linked lookup grid [27] was used in order to compute interactions effectively: the simulation box was divided into smaller sub-boxes and the surface faces were placed in these sub-boxes, according to the positions of their centroids. In the resulting grid, an entry points to the next, in order to search the sub-boxes for other faces effectively. The linked lookup grid is shown schematically in figure 3.5.

For clarity, figure 3.5 shows a two-dimensional \( N_x \times N_y \) grid, with sub-box dimension \( h \). When the surface repulsion of one face in the center sub-box, shown in dark green, is calculated, we have to consider all faces in that same sub-box. Additionally, all adjacent sub-boxes need to be considered as well: 8 in the two-dimensional case, 26 in three dimensions, summing to 9 and 27 relevant sub-boxes in two and three dimensions, respectively. This method enforces a lower boundary on the value of \( h \): \( h \) should be greater than or equal to the cut-off distance \( r_c \). If not, the distance between two faces could span more than one sub-box, which would result in possibly not calculating an otherwise existing repulsion.

To calculate all steric interactions, we needed to stepwise consider all surface faces. For each surface face, we firstly determined in which sub-box it had been placed. Subsequently, we stepped through all relevant sub-boxes, for which the distances between each face in the box and the considered
Chapter 3. Surface-surface repulsion interaction

Figure 3.5: Schematic illustration of a two-dimensional \( N_x \times N_y \) linked lookup grid. The red dots represent centroids of faces placed in the sub-boxes. The interactions of one face in the middle region are being calculated, which requires the search for nearby faces in the complete coloured area.

face were evaluated. If the face was closer than the cut-off distance \( r_c \), all pairs of Gaussian points between the two surfaces needed to be considered. For each pair, we again checked whether the separation was smaller than \( r_c \). If yes, the pairwise force and energy were calculated, after which the force was weighted, according to the weight and position of the Gaussian points, and distributed over the nodes of the two surface faces. The process for each face is schematically visualised in the flowchart in figure 3.6. The insets a, b and c illustrate the final three steps in the process, for multiple Gauss points simultaneously.
3.3 Testing steric repulsion

In order to test the repulsion method, we evaluated a test structure. This had several advantages over a more complicated model, such as the IgG protein: the geometry could be made simple and symmetric, allowing for a more intuitive evaluation of the test results; the behaviour of a simple, symmetric structure could be predicted more accurately. Furthermore, we were able to create the structure with relatively few nodes, which allowed for significantly faster simulations than simulations of "real" molecules. We chose a test geometry consisting of two spheres of radius $R$ and separated...
by a distance $d$. The distance between the two centres of mass is denoted by $r_{CoM}$. See figure 3.7.

As described in chapter 2, the molecule was being parametrised by several macroscopic parameters.

The choices for parameters for the material and steric repulsion are shown in table 3.1. We chose the repulsion strength $\epsilon$ at a value of $10^{13}$ J m$^{-4}$, which corresponds to a repulsion strength of order $\sim 10k_B T$ (at $T = 310$ K) between two surface faces of order (\sim 10 nm)$^2$. A force was applied to all the nodes of one of the spheres, towards the other sphere, and thermal fluctuations were disregarded. We expected the spheres to collide and, as they were made of a viscoelastic solid, compress slightly and then bounce back elastically.

To each of the 700 nodes of one sphere, a force of $f_a = 1$ nN was applied, summing to a total external force of $f = 0.7$ μN. Snapshots of the collision test are shown in figure 3.8. As can be seen from figure 3.8, the spheres indeed collided, compressed and bounced back. As we take the snapshots in the center of mass frame of the complete geometry, both spheres move, as opposed to only the forced sphere. In order to evaluate the physical accuracy of this interaction, both the deformation potential energy $U_{pot}$ and the steric energy $U_{steric}$ of the structure were tracked during 10 collision
3.3. Testing steric repulsion

Figure 3.8: Snapshots of the collision test simulation, with an interval of 5 ps.

Simulations. The result was averaged, and the steric energy curve was smoothed: due to the fact that this energy depends on surfaces only (and not on the internal volume) and the strong \( r \)-dependence in the Lennard-Jones potential, this energy showed a rather fluctuating behaviour. The energies as function of time are shown in figure 3.9. We see from figure 3.9 that the potential energy \( U_{\text{pot}} \) slightly

![grafik](grafik.png)

**Figure 3.9: The deformation potential energy \( U_{\text{pot}} \) and steric energy \( U_{\text{steric}} \) as function of the time \( t \).**

lagged the steric energy \( U_{\text{steric}} \) by a few picoseconds. This indicates that the repulsion started working, preventing the front of the sphere to move any further and forcing the sphere to dent at the front, causing a rise in deformation energy. After a while, the steric energy reached a maximum, after which the spheres reached their maximum compression and the structure started to relax. Both steric and potential energy decreased when the spheres started sliding along each other and they slowly returned to their initial shape. The plateau in steric energy between 22 and 32 picoseconds represents this sliding: the spheres kept pressing against each other during the sliding, resulting in a very slowly decreasing repulsion energy. The collision process becomes more insightful if we consider figure 3.10, where the energies are depicted as function of the centre of mass separation \( r_{\text{CoM}} \). We can see that as the spheres reached each other, the steric energy increased, slowly causing the spheres to deform. The turning point between compression and relaxation was reached at a separation
Subsequently, the steric repulsion interaction was incorporated into the simulations on the torsion of IgG. We chose for the repulsion interaction parameters a repulsion strength of $\epsilon = 10^{13}$ J m$^{-4}$ and a cut-off distance of $r_c = 0.3$ nm, while the material parameters were kept as described in table 3.1. The time step was $\Delta t = 1$ fs. A typical simulation is visualised in figure 3.11.

As we can see from figure 3.11, the steric repulsion interactions did effectively resist the penetration of surfaces: the previous overtwisting behaviour. The molecule did not simply relax by twisting through the surface, but was even forced to deform slightly stronger due to the surface-surface repulsion. Again, the torsion angle of the molecule was tracked, and we plot the angle as function of time (as was done in chapter 2). The result is shown in figure 3.12. In figure 3.12 the solid lines represent the data of simulations with steric repulsion incorporated, the dashed lines are data from the original simulations without steric repulsion (see chapter 2, figure 2.13). We see that all steric simulations nicely follow the non-steric data lines, up to a certain torsion angle $\phi$. At this angle, the steric repulsion started to become effective: the surfaces of the molecule repelled each other, causing a rough “wobble”, where the molecule was basically bouncing back and forth between surfaces, trying
3.4. Steric repulsion in IgG torsion simulations

Figure 3.11: Typical simulation of the IgG molecule under a total torque $\tau = 1.196$ nN nm, with steric interactions incorporated.

Figure 3.12: The torsion angle $\phi$ as function of time $t$ for various total applied torques $\tau$. The solid lines are results from the simulations with steric interactions incorporated, the dashed lines are the results from the original simulations without repulsion (figure 2.13).

... to continue rotating due to the applied torque $\tau$. The deviation from the non-steric data is visible from a torsion angle $\phi$ of around $\phi = 220^\circ$. Apart from the collision between different areas on the surface of the molecule, hardly any additional resistance to the applied torque was seen from the twisting; after a short wobble, the molecule remained twisting (albeit without overtwisting) and followed a trend similar to the data from simulations without the steric interactions.
After including the surface-surface repulsion interactions into the model, the molecule was denied to overtwist. The protein indeed did not show this highly non-physical phenomenon anymore, however, the included steric repulsion did not result in stiffening; it did not reach equilibrium angles for the stronger torques. This is likely due to the steric interactions not resulting in strong extra stresses within the material. Therefore, in chapter 4, we will describe how we evaluated the calculation of the stresses in the material and propose an alternative constitutive model. This new material model would include a deformation limit, on which the stresses in the material diverge: the material would stiffen under strong deformation.
Deformation, strain energy and stress

In this chapter, the elastic stress tensor $\sigma$ will be discussed in more detail. We will briefly discuss the origin of the deformation gradient tensor $F$, the deformation potential energy, or rather the strain energy density $W$, and the derivation of the Cauchy elastic stress tensor. For a more extensive exposition of these quantities, see Appendix D.

Moreover, as the steric repulsion alone does not result in torsion stiffening (see chapter 3), we will propose and test an alternative constitutive model, which shows a divergence in the energy upon strong deformations. In this manner, we attempted to mimic the limited extension possibilities of the underlying atomistic structure, in order to incorporate the stiffening of material under strong deformations.

4.1 Introduction on relevant deformation quantities

All deformation quantities (deformation gradient tensor, strain energy density function and Cauchy stress, to be introduced later) are directly related to the instantaneous deformation of the material. This deformation is defined as the operation on an initial position $X$, which maps this position to a new position $x(X)$:

$$X \mapsto x(X).$$

In first order, this deformation can be described by a deformation gradient tensor $F$:

$$F_{ij} = \frac{\partial x_i}{\partial X_j}.$$  

For clarity, we discuss two simple deformations in two dimensions in figure 4.1, in order to illustrate entries of the deformation gradient tensor $F_{2D}$:

$$F_{2D} = \begin{pmatrix} \lambda & \gamma_x \\ \gamma_y & \lambda_y \end{pmatrix}.$$  

We can define a strain energy density function $W(F)$, which is a function of this deformation gradient tensor $F$, as the material deforms, its elastic potential energy rises according to $W$. For the material models discussed in this thesis, $W$ is a function of two invariants of the deformation gradient, $I_1$ and $J$:

$$I_1 = \text{trace} \left( FF^T \right),$$

$$J = \det(F) = \frac{V}{V_0}.$$  

$I_1$ is in essence the sum of all squared entries of the deformation gradient $F$, which automatically implies that $I_1$ is non-negative. The determinant of $F$, $J$, is also known as the Jacobian determinant and describes the relative local volume change in the material ($V$ and $V_0$ are the instantaneous and initial volume of the material, respectively).

Using these two invariants, initially a strain energy density function $W$ was constructed, according to a compressible Mooney-Rivlin material model [41, 42]:

$$W_{\text{Mooney-Rivlin}} = \frac{G}{2}(I_1 - 3) + \frac{3K - 2G}{12}(J^2 - 1) - \left( \frac{3K - 2G}{6} + G \right) \ln J,$$  

where $K$ and $G$ are the bulk and shear modulus, respectively. This energy density function consists of a harmonic shear spring term (first term on the right hand side) and two terms that resist volumetric
expansion and compression, respectively: the second term penalises the expansion of the material (note that $J$ represents the relative volume increase), as a harmonic volumetric spring. The final term penalises a volume decrease; as soon as $J < 1$, this term will produce an energy increase. Note that all terms concerning volume change are related to the bulk modulus $K$.

As we have seen in chapter 3, only considering steric repulsion interactions in the molecule, in order to incorporate torsion stiffening in the material, was not sufficient. As we consider molecules, which underlying atomic structure would only be able to deform to a certain amount, we would like to include a stronger diverging energy function upon strong deformations. Therefore, the Mooney-Rivlin function from expression (4.3) was adjusted slightly, to correspond with an energy function proposed by Gent [46, 47]:

$$W_{\text{Gent}} = -\frac{G_2}{2}(I_m - 3) \ln \left( 1 - \frac{I_1 - 3}{I_m - 3} \right) + \frac{3K - 2G}{12}(J^2 - 1) - \left( \frac{3K - 2G}{6} + G \right) \ln J. \quad (4.4)$$

As we can see from equation (4.4), only the first term on the right hand side was adjusted. Where initially this term was a harmonic spring term, it has been converted into a diverging term. The point of divergence can be tuned by adjusting the parameter $I_m$, which should be greater than or equal to 3, as in the unperturbed state, $I_1$ equals 3 (as $F$ is the identity matrix). Although it appears to be more complicated, the Gent model reduces to the Mooney-Rivlin material, if we take $I_m \rightarrow \infty$:

$$\text{for } x \approx 0: \ln(1-x) \approx -x \Rightarrow \lim_{I_m \rightarrow \infty} \left[ -\frac{G_2}{2}(I_m - 3) \ln \left( 1 - \frac{I_1 - 3}{I_m - 3} \right) \right] = \frac{G_2}{2}(I_1 - 3).$$

To gain more insight in the actual strain energy density function $W$ as function of extension or shear, we evaluate the energy density, using the following deformation gradient tensor:

$$F = \begin{pmatrix} \lambda & y & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (4.5)$$

We plot the energy density functions for the Mooney-Rivlin model ($W_{\text{Mooney–Rivlin}, I_m \rightarrow \infty}$) and the Gent model ($W_{\text{Gent}, I_m = 4, 8, 12, 16}$) in figure 4.2 as function of $\lambda$ (figure 4.2a) and $y$ (figure 4.2b),
4.1. Introduction on relevant deformation quantities

while keeping $\gamma = 0$ and $\lambda = 1$, respectively. Using these strain energy density functions, the elastic Cauchy stress tensor $\sigma^e$ can be derived from this function $W$, if the deformation gradient tensor $F$ is known:

$$
\sigma^e = \frac{1}{J} \frac{\partial W}{\partial F} F^T \quad \text{or} \quad a_{ij} = \frac{1}{J} \frac{\partial W}{\partial F_{ik}} F_{jk},
$$

which gives for the Mooney-Rivlin material:

$$
\sigma^e_{\text{Mooney-Rivlin}} = \frac{G}{J} F F^T + \left[ \frac{3K - 2G}{6} \left( J - \frac{1}{J} \right) - \frac{G}{J} \right] I,
$$

and for the Gent material:

$$
\sigma^e_{\text{Gent}} = \frac{G}{J} \frac{I_m - 3}{I_m - I_1} F F^T + \left[ \frac{3K - 2G}{6} \left( J - \frac{1}{J} \right) - \frac{G}{J} \right] I. \quad (4.6)
$$
Chapter 4. Deformation, strain energy and stress

Strain energy density functions $W$ for the deformation gradient tensor $F$ in expression (4.5), as function of $\lambda$, while keeping $\gamma = 0$.

(a) Strain energy density functions $W$ for the deformation gradient tensor $F$ in expression (4.5), as function of $\gamma$, while keeping $\lambda = 1$.

(b) Strain energy density functions $W$ for the deformation gradient tensor $F$ in expression (4.5), as function of $\gamma$, while keeping $\lambda = 1$.

Figure 4.2: Strain energy density functions $W$ for the deformation gradient tensor $F$ in expression (4.5).
4.2 Testing the Gent model: Linear extension force

In order to test the effect of this new constitutive material model on the dynamics of a material and to clarify the difference between the Mooney-Rivlin and Gent models, in this section and section 4.3, we will describe simulations performed on a test structure. This structure consisted of a single cylinder, with radius \( R \) and length \( L \), as shown in figure 4.3. Our main interest lies with applied forces; more specifically, with linear extension and torsion. Due to the symmetry of the cylinder, we were able to investigate both force directions using the same test geometry, and still intuitively predict the outcome. We took a cylinder of initial length \( L_0 = 12.5 \) nm and radius \( R = 1.5 \) nm. To maximise the contrast between the Mooney-Rivlin and Gent models, we tested both values of \( I_m/\mu \approx 3.01 \) (Gent) and \( I_m \to \infty \) (Mooney-Rivlin).

In order to test the response of the test structure to external linear forces, total applied forces between \( f = 1.36 \) nN and \( f = 1.36 \) μN were explored (spanning two orders of magnitude). The material was pulled from both ends of the cylinder (in both the +z and −z-direction) and the extension \( L/L_0 \) was tracked as function of time, with \( L \) and \( L_0 \) the instantaneous and initial length of the cylinder, respectively. Similar to the simulations described in section 2.5, we applied the forces slowly, to avoid sudden strong deformations: the force was built up quadratically over time, over 75 percent of the simulation. The instantaneous applied nodal force \( f_n(t) \) as function of the build-up time \( t_b \) and maximum force \( f_0 \) was:

\[
f_n(t) = \begin{cases} 
\left(\frac{t}{t_b}\right)^2 f_0 & \text{for } t < t_b \\
0 & \text{for } t > t_b 
\end{cases}
\]

which defines the total (maximum) force \( f = \sum f_0 \). We first discuss simulations on the original material, described by the Mooney-Rivlin model. The extension \( L/L_0 \) as function of time \( t \) for various applied forces \( f \) is plotted in figure 4.4. We see from figure 4.4 that the cylinder nicely followed the quadratic nature of the applied force and stabilised at a certain length after 0.75 ns, at 75 percent of the simulation. To analyse the resistance to force, we plot this equilibrated extension \( (L/L_0)_{eq} \) as function of the applied force \( f \), see figure 4.5. We see that the Mooney-Rivlin material behaved linearly elastic, as was to be expected from the strain energy density \( W \) from expression 4.3 and figure 4.2a.

As we consider single molecules, which underlying atomic structure simply does not contain sufficient “material” to deform to such extent, we subsequently considered a strongly strain-stiffening Gent material (expressions (4.4) and (4.6)), with \( I_m = 3.01 \). The exact same forces were applied, in the exact same manner, and again the extension factor \( L/L_0 \) was tracked as function of time. The result is shown in figure 4.6. As we can see from figure 4.6, the result differs drastically from the Mooney-Rivlin material: at the start of the simulation, we still see the quadratic behaviour of the applied forces, but the material rapidly stiffened and converged towards an equilibrium extension factor. For stronger forces, instead of showing an increasing equilibrium extension factor \( L/L_0 \), we see that the curves converge to a limiting value of around 1.7. This is more clearly visible in figure 4.7. Due to the strong limitation of the material \( (I_m \) was close to its limit value of 3), the cylinder reached its
Figure 4.4: The extension $L/L_0$ of the Mooney-Rivlin material as function of time $t$ for various applied forces $f$.

Figure 4.5: The equilibrated extension $(L/L_0)_{eq}$ of the Mooney-Rivlin material as function of applied force $f$.

maximum extension at fairly weak forces. From the behaviour of both materials, a comparison between the instantaneous spring constants $k_s$ of both materials is made, defined by:

$$k_s \left((L/L_0)_{eq}\right) = \frac{f}{(L/L_0)_{eq}}.$$

This is in principle another way of representing the earlier shown data, but now by comparing the actual resistance (spring constant) of the material to applied forces. The results are depicted in figure 4.8. We used the linear fit of figure 4.5 to interpolate the spring constant $k_s$ for the Mooney-Rivlin material at small extensions, as the spring constant is the reciprocal of the slope. It is clearly seen that for the Gent material, the spring constant $k_s$ diverged at an extension of around 1.7, while the Mooney-Rivlin spring remained in its linear behaviour. By adjusting the value of $I_m$, the divergence
4.2. Testing the Gent model: Linear extension force

Figure 4.6: The extension $L/L_0$ of the Gent material with $I_m = 3.01$ as function of time $t$ for various applied forces $f$.

Figure 4.7: The equilibrated extension $(L/L_0)_{eq}$ of the Gent material with $I_m = 3.01$ as function of applied force $f$.

extension of any (bio-)material one would like to model using the Gent material model may be set to the desired value.
Chapter 4. Deformation, strain energy and stress

Figure 4.8: Comparison of the instantaneous spring constants $k_s$ as function of the expansion factor $L/L_0$ for the Mooney-Rivlin and Gent materials.

4.3 Testing the Gent model: Axial torque

Comparable to the set-up described in the previous section, the response of the structure to applied torques was tested. The torque vector $\tau$ was parallel to the $+z$-direction and on the symmetry ($z$-)axis of the cylinder. Similar to the linear applied force, the torque was applied slowly, and built up quadratically over a time $t_b = 0.75$ ns:

$$\tau_n(t) = \begin{cases} \left(\frac{t}{t_b}\right)^2 \tau_0 & t < t_b, \\ \tau_0 & t > t_b, \end{cases}$$

until it reached a stable value of $\tau_0$. Torques between $\tau = \sum \tau_0 = 71.2$ pN nm and 7.12 nN nm were explored, again two orders of magnitude.

All nodes at one end of the cylinder were pinned to their respective positions, as were the nodes on the torsion axis. On the other end, all points were constrained to the $z$-plane, allowing for a pure rotation, without contraction of the cylinder in the axial direction. The torsion angle between the new and initial states was tracked as function of time. The results for both the Mooney-Rivlin (figure 4.9a) and Gent (figure 4.9b) material models are shown in figure 4.9. We see from figure 4.9 that the Mooney-Rivlin and Gent material models showed similar behaviour: both followed the quadratic nature of the applied force and stabilised at a certain torsion angle $\phi$. Apart from the fact that the Gent material overall appears slightly stiffer than the Mooney-Rivlin material, we see no extra resistance to torques in the Gent material for the explored torque magnitudes. To support this indication, we plot the equilibrium torsion angle $\phi_{eq}$ as function of the applied torque $\tau$ in figure 4.10. Indeed we see no torsion stiffening behaviour occurring in the Gent material in figure 4.10. It is however, slightly stiffer than the original Mooney-Rivlin material. The torsional spring constant $k_t = \tau/\phi_{eq}$ was calculated (with $\phi_{eq}$ in radians), and $k_{t,Mooney-Rivlin} = 1.90 \cdot 10^{-18}$ Nm and $k_{t,Gent} = 2.79 \cdot 10^{-18}$ Nm were found, respectively.

Apparently, the influence of the parameter $I_m$ of the Gent model was way more prominent in the linear extension of the material than in the axial torque. As we may see from expressions (4.3) and (4.4) and figure 4.2, any deviation from the initial structure (where all $\lambda_i = 1$ and all $\gamma_{ij} = 0$) should result in a rise in strain energy. For the Gent model, this would even mean a divergence in the energy,
Figure 4.9: The torsion angle $\phi$ as function of time $t$, for various applied torques $\tau$.

Figure 4.10: The equilibrium torsion angle $\phi_{eq}$ as function of the applied torque $\tau$ for both the Mooney-Rivlin and Gent material models. In this torque regime, we see a linear behaviour for both material models, albeit the Gent model a slightly stiffer one.
resulting in no extra deformation. However, in the case of the axial torsion, it seems that including this divergence hardly had any influence on the behaviour of the material. As torsion of a continuum material is a complex superposition of both extension/compression and shear, the material apparently behaved in such a way that the actual divergence in the energy was not reached for the torques explored. This could be due to a torsion causing a simultaneous shear and compression, for example, that partly cancel each other.

The current Gent material model thus seems rather unsuccessful in incorporating a divergent torsional resistance in the material. A solution to this might be to implement a strain energy density function that deals with extension/compression \( \lambda \) and shear \( \gamma \) more independently. However, we chose to use the Gent model as an extension of the Mooney–Rivlin model, in order to account for the finite deformation character of (bio-)polymer chains. These chains would deal with extension and shear in a similar way: only strong, "outward" deformations should be energetically expensive, due to their finite length. Inward deformations would not need to be as expensive, since such a molecule would be allowed to compress relatively easily. Drastically adjusting the strain energy density function in order to incorporate torsion stiffening alone, would not be directly physically justified.

However, for strong deformations, the material compresses to such extent, that normally the material would start to reach its internal steric limit: the material was allowed to compress immensely, in order to almost nullify the strong penalty on deformations. Compressing to virtually zero volume should not be allowed. A possible solution could therefore be to devise another material model, which incorporates both a diverging energy for strong linear extensions and shear, and a diverging energy for extreme volumetric extensions or compressions. This exploration of the possibilities of a new material model is left for future research.

We have given an introduction to the calculation of the Cauchy stress tensor within the material. As we wanted to include a divergence in the stress, under strong deformations, we proposed a new constitutive material model, based on a Gent model. This Gent model showed strong strain stiffening in linear extension, but no divergence in torsional resistance for axial torque. This is likely due to the torsion being a superposition of shear and compression, which partly cancel each other’s stresses.

We can, however, use the calculated stresses within the material to target strongly deformed regions within the molecule. With this information, the relevant regions (in terms of deformation) in the protein can be evaluated in more detail. We will show in chapter 5 how we used the stresses within the Gent continuum material to further address the torsional properties of molecules.
We can use the information gained by the calculated stresses in the material to evaluate where the strongly deformed regions are located. The strongly deformation-dependent stress of the Gent model does cause larger differences in stress between the strained regions, increasing the contrast between strongly strained regions. Therefore, simulations of antibodies were performed subject to external torques, made of a Gent material with \( I = 3.1 \). The steric repulsion interactions were turned on as well, to deny the penetration of surfaces and incorporate extra deformation of the material, if necessary.

This chapter will introduce a possible analysis method to target specific regions, if distinguishable, for which the stresses are substantially higher than in the rest of the molecule. These regions can be extracted and taken back to the atomistic level, to investigate with higher resolution and all atomistic information re-incorporated, whether certain properties of the molecule can be explained, predicted and/or accounted for by that particular region of the molecule.

### 5.1 Targeting strongly strained regions in the IgG molecule

In order to target the stresses in the material, the stresses in each element were tracked during the simulation, or rather the square root of the double contraction of the stress tensor \( \sigma_m \):

\[
\sigma_m = \sqrt{\text{trace} (\sigma \sigma^T)}.
\]

This quantity \( \sigma_m \) is the square root of the sum of all quadratic elements of the total stress tensor \( \sigma \), a measure that will generally take into account all amounts of elastic, viscous and thermal stresses. As the thermal fluctuations were turned off and the molecule was torqued slowly (and we are in fact only interested in the equilibrium configuration), both the viscous and thermal stress terms would be negligible.

We subsequently took the stress \( \sigma_m \) of all elements and visualised the simulations under influence of torque in a different manner than before. Each surface face of the material was coloured according to the stress \( \sigma_m \) of the underlying element (red for high stresses, green for low), scaled to the maximum stress at that particular time step. A typical simulation is shown in figure 5.1.

As we can see from figure 5.1, the darker regions (representing higher stresses) almost exclusively concentrated in the thin linker region between the three bulky regions of the molecule, when the protein reached its equilibrium torsion angle. We can even distinguish the onset of the torsion; at \( t = 1 \) ns, the relatively high stresses were also present in the lower region, to which the torsion forces were being applied. The stresses were relatively high here, as the molecule needed to get in motion: inertial effects caused slightly more resistance to motion at the start. As the model was sensitive to strong deformations – the potential energy diverged upon reaching a certain amount of deformation – the strongest deforming elements showed up extremely strongly. However, this lowered the contrast between the (almost) undeformed elements and the averagely deformed elements (due to the normalisation to the maximum stress in the material).

Therefore, a transition was made to another representation, for which elements showed up red as the stresses were higher than a certain threshold value \( \sigma_{th} \). This other representation is shown in figure 5.2, for which we set \( \sigma_{th} = 1 \) MPa. In figure 5.2, we can clearly see that the high stresses in the molecule concentrated strongly within the thin linker region between the three branches. As the stresses in other parts of the molecule appeared to be significantly smaller, we extracted the linker region from the molecule and reverted to the atomistic model for this particular part of the IgG. Subsequently, a few preliminary full-atom torsion simulations were performed on the linker region by using molecular dynamics (MD) simulations.
5.2 Introduction to molecular dynamics

In order to evaluate the kinetics of a molecule, it can be modelled by considering all its atoms and their respective interactions and simulating the fine structure by using molecular dynamics (MD). As mentioned in section 1.2, MD is a nanoscale simulation method, with time and length scales in the order of (sub-)nanoseconds and (sub-)nanometres, respectively. The modelling is done by dynamically minimising the potential energy $U$ of the full simulation box. The forces on each atom were calculated by deriving this potential energy to the three-dimensional position. The MD simulation method basically evaluates Newton’s second equation of motion for many particles simultaneously. In general, the total potential energy $U$ is (schematically) represented by the following equation:

$$U = \sum_{\text{bonds}}^{\text{bonded}} U_{\text{stretching}} + U_{\text{bending}} + U_{\text{torsion}} + \sum_{j<i}^{\text{non-bonded}} U_{\text{electrostatic}} + U_{\text{pairs}},$$

(5.1)

where the distinction is made between the interactions between bonded atoms and the non-bonded interactions: we sum over all bonds and over all atom pairs $(i, j)$, respectively. The bonded interactions are schematically depicted in figure 5.3. In general, the bonded interactions are implemented in such
Figure 5.3: A schematic representation of bonded interactions in an atomistic model: stretching, bending and torsion. The preferred (rest) states are indicated by $l_0$, $\theta_0$ and $\phi_0$, which represent the length, bending angle and torsion angle, respectively.

A few preliminary torsion simulations were performed on the linker region of the IgG molecule, in order to obtain more insight in the torsional properties of this protein. As we have seen in section 5.1, the linker region appeared to be responsible for the major part of the deformation under external torques. We took the atomistic structure for IgG from the PDB (1IGT [6]) and extracted the thin region from the atomistic model. The amino acids (AAs) we used to construct the linker region are numbers 229-243, from both chain B and chain D. Due to the fact that a few AAs appeared to be missing in the model of the crystallised IgG molecule (AA 233 and 234), this summed up to 26 amino acids (374 atoms), which is considerably fewer than the 1316 in the total IgG molecule. Missing amino acids generally relate to flexible regions in the molecule (they could not be measured in the X-ray crystallography experiments), which is convenient for us, as the more rigid structures remain present in the model.

The strong reduction in number of amino acids and size of the molecule allowed us to perform some relatively fast MD simulations of the linker region, constrained at the ends to model the relatively heavy, bulky branches and immersed in explicit water and ions (the water molecules and ions are simulated explicitly). As the molecule size was decreased, a smaller simulation box could be used, which meant fewer water molecules and ions to simulate. The atomistic model is shown in figure 5.4.

In the atomistic model, the amino acid chains of the linker region have not folded in any secondary (folding of chains into (alpha) helices or beta sheets) or tertiary (folding of helices and sheets into larger structures) protein structures. The chains were linear and interconnected by three disulfide bonds between the chains (depicted in yellow).

We used the simulation package GROMACS [48, 49] to simulate the behaviour of the linker region under external torques and we chose the Amber99SB-ildn force field [50] as the interaction parameters and potentials in these simulations.
Chapter 5. Targeting torsional properties

5.3.1 Simulation set-up

The linker region was extracted and simulated on the nanoscale, with a full-atom resolution. The top AA of each chain was fixed in place, while the bottom AA of each chain was being constrained to its respective horizontal plane, allowing for a pure rotation (see figure 5.5). We immersed this part of the molecule into explicit water and added ions, with a concentration of 0.1 M, setting the cumulative charge of all molecules to zero.

Before the simulation was started, first the energy of the initial structure needed to be minimised. This was done in order to avoid large forces upon starting the evaluations, as the atomistic model might contain a few out-of-equilibrium bond states. Then, a short \((N,V,T)\) equilibration simulation...
(keeping Number of particles, Volume and Temperature constant) and subsequently a short \((N, P, T)\) equilibration simulation (Number of particles, Pressure and Temperature) were performed, in order to bring the molecules closer towards their physical equilibrium, at a temperature of 298 K and a pressure of 1 bar. The simulation box was subject to periodic boundary conditions and during all equilibration simulations both ends of the chain were fixed in place. After the equilibration steps, we released the bottom AAs into the horizontal plane and started rotating these AAs with a rate \(\omega\) of 10 degrees per picosecond around the torsion axis. In fact, a rotating radial potential well was used (GROMACS: \(V_r \mathbf{m} \mathbf{m}^{-2}\)), which pulled the AAs to the desired positions, as illustrated in figure 5.6. The AAs were pulled towards the center of the well (as illustrated by \(\mathbf{F}\)), which rotated around the rotation axis, depicted by \(\otimes\). The steepness of the potential well was given by a spring constant \(k_0\). This rotation using a potential well resulted in a similar clockwork motion as was seen in experiments (chapter 2), which we will show in section 5.4.

5.4 Simulation results

Simulations were performed on the IgG linker, varying the well steepness \(k_0\). A higher \(k_0\) would result in higher torques \(\tau\) on the hinge region as the molecule was twisted more strongly. A typical twisting simulation, with \(k_0 = 700 \text{ kJ nm}^{-2}\text{mol}^{-1}\), is illustrated in figure 5.7.

![Figure 5.6: Illustration of the rotating radial potential well (GROMACS: \(V_r \mathbf{m} \mathbf{m}^{-2}\)), taken from [49].](image-url)
In figure 5.7, we see that at first, the molecule twisted steadily in time, due to the rotation of the potential well. However, after approximately 100 ns, a saturation angle was reached, which resulted in the "clicking" of the molecule to smaller torsion angles, similar to the remagnetisation of the superparamagnetic microspheres in experiments (see chapter 2): at a certain torsion angle, the resistance to torque became equal to the applied torque, the molecule started lagging behind the potential well, the torque decreased and the molecule clicked to a smaller angle. After some time, the potential caught up with the molecule and the process recurred.

From the positions of the atoms in time, the torsion angle \( \phi \) (the rotation angle of the molecule) could be calculated. Note that the actual coiling of the chains was tracked: the angle between the two strands was evaluated in the lowest possible (horizontal) \( x, y \)-plane they both penetrated. The results are shown in figure 5.8. First of all, in figure 5.8, we can see the clockwork motion in the angle \( \phi \) as function of time \( t \). The simulations seem to be qualitatively similar to the experimental phenomena. However, the (steady) maximum angle reached for this potential strength \( k_0 \) is rather large: around 800 degrees, compared to the approximately 40 degrees in experiments (section 2.3). We have to note several factors that could possibly explain this large torsion angle:

- We only considered two flexible amino acid chains in the simulation, which were allowed to stretch nor compress along the torsion axis. The resistance to torque would likely be larger when stretching the strands. From an experimental point of view, stretching the molecule while applying the torque is rather relevant: a micrometre-sized bead is attached to a surface through a nanometre-sized antibody molecule. The bead hangs from the surface through this IgG molecule, which would be stretched by the gravity on the bead. Furthermore, upon twisting the bead, where thermal forces are also at play, it would likely bounce off the glass surface, resulting in a net stretching of the antibody molecule. Reversely, this means that we also need to be careful about the interpretation of the experimental results: we should identify to what
5.4. Simulation results

Figure 5.8: Torsion angle $\phi$ as function of time $t$ for $k_0 = 700 \text{ kJ nm}^{-2}\text{mol}^{-1}$.

extent the stretching forces are important in the analysis of the torsion profiling experiments.

- We appeared to be missing two amino acids in each chain. Though these amino acids were likely rather flexible (which caused the undetectability in measurements), they would likely contribute to the resistance to torques: the physical size of the chains would increase, resulting in a repulsion between the chains at smaller torsion angles.

- However the bulky regions apparently did not contribute as much to the internal deformation energy as the hinge region, they could still play a role in the resistance to torque: upon twisting, these parts would possibly obstruct each other. This would not result in extra deformation of the material as such, but this steric repulsion might cause a greater resistance to external torques, especially when the potential remains rotating at the same pace: the molecule would lag behind the potential at smaller torsion angles, resulting in the “clicking” occurring sooner.

Despite the apparent inability to directly quantitatively relate the MD results to experimental data, we nevertheless continued to explore multiple rotative potential steepnesses $k_0$, to gain insight into how to analyse the data in order to extract a torsion profile from molecular dynamics simulations.

The torsion angle $\phi$ as function of time $t$ for multiple $k_0$ is shown in figure 5.9. In figure 5.9, we can see, as expected, that the maximum torsion angle $\phi_{max}$ increased for larger $k_0$. We also noted an overshoot of the torsion angle $\phi$, for large $k_0$: the first (few) click(s) occurred at larger torsion angles than the later ones, which could be due to inertial effects. Furthermore, not all simulations showed a nice, periodic clicking behaviour with steady maximum angles. Both issues could be resolved by rotating for longer simulation times at a lower rate $\omega$, which would firstly allow the molecule to properly relax after each click and, secondly, allow us to resolve a steady maximum angle from an equilibrium clicking behaviour.

The magnitude of the torque $\tau$ on the molecule was dependent on the spring constant $k_0$. It was calculated from the position of and the force on all atoms in the rotation group ($\tau = r \times F$). For $k_0 = 700 \text{ kJ nm}^{-2}\text{mol}^{-1}$, this torque $\tau$ is shown, along with the angle $\phi$ as function of time $t$ in figure 5.10. We see from figure 5.10 that, as expected, the torque on the molecule showed a clicking behaviour in time, similar to that of the torsion angle: the torque reached a maximum value right at the click. As in experiments, both the maximum torsion angle $\phi_{max}$ and the maximum torque $\tau_{max}$ were determined from these clicking motions and the torsional spring constant $k_t$ was calculated from these quantities (according to equation (2.1)). We therefore calculated the average and the standard deviation of each maximum angle and maximum torque for every value of $k_0$, as illustrated in figure 5.11.
Subsequently, these found values and deviations were used to calculate values for \( k_t \) at each click, which were, in their turn, averaged into one value of \( k_t \) for each potential strength \( k_0 \).

The torsional spring constant \( k_t \) as function of the torsion angle \( \phi \) is shown in figure 5.12. In figure 5.12, we do indeed see an increase in the torsional spring constant \( k_t \) as the torsion angle \( \phi \) increases, as we have seen in experiments (chapter 2, figure 2.8 and [25]). This is an indication that the flexible region in the atomistic model did show behaviour similar to the full IgG molecule in experiments. Despite the fact that both the torsion angle \( \phi \) and the spring constant \( k_t \) did not directly resemble the experimental values, which may be due to the aforementioned factors, we do see a comparable qualitative behaviour.

The fact that the general trend of \( k_t \) as function as \( \phi \) was much alike in both experiment and simulation, is promising for the future analysis of experimental assays. Apparently, when evaluating the structure using simulation, we can find results qualitatively similar to experimental data. Upon improvement of the model, for example by simulation of the molecule under influence of both torques.
and linear stretching forces (to represent the presence of the magnetic microsphere), we may be able to gain more insight on the physical interpretation of the torsion experiments and eventually even describe and/or predict the outcome of experiments. Naturally, this would require some extension and more careful execution of the existing simulations. At first, we would need to let the molecule equilibrate properly, before applying the external forces. Secondly, these forces would need to be applied in a gentler (slower) fashion, in order to eliminate inertial effects, and simulations would need to be performed for longer times to find a proper steady-state clicking behaviour. Thirdly, as in experiments, the molecule should be rotated in both directions, to find a more accurate value for the absolute torsion angle at a certain torque. These adjustments are left for future research.
We have shown a method to visualise a strongly stressed region within the IgG molecule during the
simulation. To get more insight in the internal structure of this region, it was extracted from the full
molecule and a few preliminary molecular dynamics simulations were performed on the sub-molecule.
While applying a torque to the linker region, we found a slight torsion stiffening at larger torsion
angles. Although the calculated values do not resemble experimental data, we have found a qualitative
agreement between simulations and experiment. This is an indication that the torsion profile can be
evaluated by analysing the protein structure.

The transition from the experimental set-up, via the coarse-grained simulation method, to an atomistic
model, could become a rather powerful approach to analyse the experimental situation. At first, by
evaluating the coarse-grained model, we can simulate multiple molecules on relatively long time
scales. If desired, we can model the motion of a protein complex subject to thermal or external
forces, and, as shown in section 5.1, target the strongly stressed regions within the complex subject
to deformation. These regions can then be extracted and simulated with higher resolution, using
a more fine-grained (atomistic) approach, to analyse the behaviour of these parts separately. This
approach would lead to an understanding of the most relevant regions, in terms of deformation, on
gradually decreasing time and length scales, or increasing resolution. We specifically address these
parts, evaluate their internal structure and relate the results back to the experimental assay and its
properties.

Alternatively, one could devise a hybrid simulation method, combining elements of both an atomistic
and a coarse-grained simulation method, in which the rigid regions are represented by one or multiple
particles, representing groups of atoms, and the more flexible regions are modelled on a full-atom
level. This would require, however, careful re-parametrisation of the MD force fields, concerning the
bond potentials and non-bonded interactions, especially on the boundary between the coarse-grained
and atomistic models.
In this chapter, we summarise the main results of this work and draw the relevant conclusions. Moreover, an outlook for future research is sketched and lastly, we discuss the technological relevance of this work.

6.1 Summary and conclusions

In this research, we attempted to address how we can use torsion profiling of proteins as an alternative protein identification method. We hypothesised that we can distinguish between different torsion profiles and that the torsion profile of a protein is determined by the protein structure. Torsion experiments were performed on an immunoglobulin G–protein G complex using magnetic tweezers, where the complex formed the connection between a magnetic microparticle and a glass substrate, and we found a strong indication that we can indeed distinguish between different types of bonds between particle and substrate: two substantially different torsion profiles were found. One was associated with the immunoglobulin G–protein G complex, the other with a multiply bound bead. The evaluated protein complex showed a torsion stiffening behaviour, at angles of approximately 40 degrees.

After we tested the viability of the technique, we explored how the torsion profile of a protein may be evaluated from its structure: understanding and being able to predict the torsional properties of the protein complex is an important step for the method to be used as an alternative identification scheme. Therefore, the antibody protein immunoglobulin G was assessed specifically, as this molecule is often present in the experimental magnetic assays: if the torsion profile for this molecule is known, it can be used in the future analysis of the protein complexes in the magnetic assays. Its atomistic structure was introduced and we noted that the collection of its two heavy chains and two light chains forms three bulky regions, connected by a flexible linker. This structure was subsequently evaluated using Fluctuating Finite Element Analysis (FFEA), a coarse-grained mesoscale simulation method, which calculates the behaviour of molecules as if they were made from a continuum material, parametrised by macroscopic parameters, such as a density and a bulk modulus.

Upon applying a torque to the FFEA model of the antibody and analysing the data, however, the molecule showed no torsion stiffening behaviour. Moreover, at strong torques, the protein overtwisted: due to the strong torques, the molecule twisted through its own surface to relax the high internal stresses. We hypothesised that this highly non-physical behaviour was due to the absence of surface–surface repulsion: the surface of the molecule did not feel itself, which allowed surface faces to penetrate each other.

In order to prevent the overtwisting behaviour of the antibody molecule, a surface–surface repulsion interaction was implemented, based on a cut-off Lennard-Jones potential. This did deny the surface faces to pass through each other, however, this did not cause a torsion stiffening: the extra deformation caused by the steric interactions did not result in sufficient resistance to the external torques. Steric interactions alone were not sufficient for the FFEA method to model the expected torsion stiffening in the protein.

This led to the transition to an alternative constitutive material model. We investigated the effect of a Gent material model, which contained a divergence in the deformation energy upon reaching a certain deformation: the material would stiffen when it reached this deformation limit. This material model had been proposed to incorporate the behaviour of the underlying atomistic structure of the protein: it would only be able to deform to a certain maximum amount. In contrast to the initial Mooney-Rivlin material model, the Gent material showed a strong strain stiffening when subjected to linear extension forces. Yet, under applied torques, the Gent material showed no divergence in the torsional resistance, although it behaved slightly stiffer than the Mooney–Rivlin material. This is likely due to the torsion stiffening...
resulting in a superposition of both shear and compression: both deformations partly counteracted each other in terms of increasing the deformation energy, which resulted in the energy divergence not being reached for the torques explored.

Despite the inability of the FFEA method to describe the torsion stiffening of the antibody molecule (that is, with the methods described in this thesis), we have shown that it is useful to evaluate torqued molecules with FFEA, in order to target the most strongly deformed regions of the molecule. The immunoglobulin G structure was evaluated, subject to external torques, with both the steric interactions and the Gent material model incorporated. The stresses within the material were tracked during the twisting. Mapping the stresses onto the material by magnitude allowed for a visualisation of the most important regions of the molecule, in terms of deformation. The linker region of the antibody, which appeared to be stressed most strongly, was then extracted from the full protein and analysed subject to torques on a full-atom level, using molecular dynamics simulations. These preliminary simulations demonstrated that the linker region showed a torsion stiffening behaviour, albeit at larger torsion angles and smaller torsional spring constants than shown in experiments. This is an indication that we can qualitatively analyse the torsional properties of a protein by evaluating its structure. The discrepancy between experiments and simulations may be accounted for by considering that we only examine a small part of the full molecule (whereas other regions might play a role in the stiffness, e.g. the bulky regions that sterically repel each other) and are missing several residues in the atomistic model. Furthermore, we do not take into account the effect of the magnetic microsphere, which would stretch the antibody. Reversely, this also means that we should carefully identify the importance of the magnetic bead in the analysis of the torsion profiling experiments.

Conclusively, we have shown an indication that the experimental method is capable of distinguishing different bonds between magnetic microsphere and glass substrate. The torsional properties of the molecules in the assay can be investigated using magnetic tweezers. Furthermore, we can evaluate the proteins in the examined assay using a coarse-grained simulation method, which in principle allows for fast simulation of multiple proteins under arbitrary forces. If necessary, surface-surface repulsion interactions can be incorporated to prevent non-physical phenomena from occurring, or we can change to an alternative constitutive material model for which stresses diverge upon strong deformations. The stresses in the material can be shown during the simulations of the proteins, from which the most strongly strained regions can be identified. These most important regions (in terms of deformation) may then be extracted from the full molecule and investigated on smaller length and time scales, to target their kinetics separately. We have shown that, by analysing the structure of a molecule, a torsion profile was found, which was qualitatively similar to the experimentally found behaviour.

In this thesis, we specifically targeted the torsional properties of antibody molecules, whereas this method can in principle be applied to arbitrary (large) proteins subject to arbitrary forces. Hence, we developed an scheme to investigate the kinetic behaviour of molecules under external forces on increasingly small time and length scales. We performed torsion experiments on a protein complex using magnetic fields on a fluid cell, where we used the magnetic labels to visualise the location and properties of the proteins: a (quasi-)macroscopic approach. The antibody in the protein complex was then investigated using the mesoscale FFEA method. We used this method to identify the most strongly stressed regions within the antibody molecule, which were subsequently extracted from the full molecule and investigated on a nanoscale level, using molecular dynamics simulations.

6.2 Outlook

In this work we described a scheme to examine large proteins on multiple time and length scales. Still, several aspects of the presented methods would benefit from a more careful or more extensive approach. We will note a few factors which could be a motive for future work.

Firstly, concerning the implementation of repulsive interactions, we evaluated all surface faces of the molecule when calculating the steric interactions. However, most of the molecule was rather static, which means that these parts of the antibody were not likely to be interacting with other nearby
6.3 Technological relevance

Surfaces. Defining regions on the molecule which should be sterically repulsive, rather than the full molecule, would speed up steric interaction calculations considerably. Furthermore, we have seen that the current Gent constitutive model did not result in a divergent torsional resistance, which was due to the superposition of shear and compression in the material while twisting it. In order to incorporate the divergent behaviour using the Gent model, we would likely need to penalise both strong deformations and strong compressions in a divergent manner. The physical interpretation would be that for strong extensions, the underlying structure would reach its contour length, which would result in strong resistance to further deformation. For strong compressions, however, the atomistic structure would reach its steric limit: the atoms would start to repel each other, resulting in strong resistances to compression. Incorporating both divergences would likely result in torsion stiffening in both linear extension and axial torque. Future research could include the exploration of alternative constitutive (Gent) material models, in order to model the torsional resistance accurately, yet still based on a physical foundation.

Thirdly, in order to further investigate the torsional properties of the immunoglobulin G, one could more carefully perform the molecular dynamics simulations on the flexible hinge region. To mimic the experimental approach, one would first need to apply the torques in two directions, in order to measure the absolute torsion angle; this corrects for the possible angle offset. Moreover, the frequency of the rotating potential field should be decreased, in order to enable the atoms to equilibrate to their new state, and the simulation should be extended to longer simulation times. This in order to more accurately define the maximum torsion angle and maximum torque in the steady state. In addition to these minor adjustments to the preliminary simulations, the presence of the (large) superparamagnetic bead should be taken into account: the gravity on the heavy bead would stretch the protein complex in the assay. It bouncing off the glass substrate would result in even stronger stretching forces on the molecule. This could be modelled by simultaneously exerting a stretching force and a torque on the linker region, and investigating the torsional spring constant as function of the torsion angle, for multiple applied stretching forces.

Alternatively, one could devise a hybrid molecular dynamics simulation set-up, in which the atomistic simulation of the linker region is combined with coarse-grained modelling of the bulky branches of the antibody molecule: the branches may be represented as one or multiple particles, representing groups of atoms. Note that this would require careful parametrisation of the bond properties in the coarse-grained model, especially on the boundary between the fine-grained and coarse-grained models.

6.3 Technological relevance

As described in the introduction of this thesis, biosensors are being developed in order to push the detection limit to (sub-)picomolar concentrations. The methods described in this work could be applied to investigate specifically the separate molecules present in the biosensors’ sandwich assays, and evaluate their torsional properties. Understanding and being able to predict the behaviour of molecules subject to torques is an important step towards introducing torsion profiling as an alternative protein identification method. Apart from the fundamental information we would gain from investigating torsion responses of proteins, once the torsion profile of a certain protein complex is known, it can be directly applied in biosensing devices: using rotational magnetic fields, we would be able to measure the torsion profile of the proteins in the assay. This would allow us to distinguish between beads bound to the surface in the desired manner (through the sandwich assay) and beads non-specifically bound to the surface. This information can then be used to determine the number of specifically bound beads on the surface, which is a direct measure for the concentration of target protein in the fluid: assessing the type of bond between bead and surface in a biosensor would decrease uncertainties in the measured concentration considerably.
Chapter 6. Conclusions and outlook


I would like to thank several people, who have been a great support during the research that eventually resulted in this thesis. Firstly, many thanks to Kees Storm, my daily supervisor, who has been extremely helpful in structuring and planning this thesis and the research as a whole. Many thanks to Fabiola Gutierrez, whose patience and expertise have helped me through the experimental part of this research. After several lengthy and fruitful discussions, she provided more insight in the experimental context of torsion profiling. Thanks to Leo van Ijzendoorn, for his involvement in both the experimental and the simulation parts. Credits to Stefan Paquay, the group’s programming and Linux troubleshooter, for his effort to help me optimise the simulation code. Thanks to all TPS group members, in or outside the office, for various fruitful discussions and their pleasant company, in particular at lunch and at the weekly Borrel.
Many thanks to my friends at the University of Leeds, who I visited for three months in April 2014 and again for a few days in September 2015, in particular Sarah Harris and Ben Hanson, for their extreme hospitality and involvement in the project. I would also like to express my gratitude towards Daniel Read and Oliver Harlen, from the School of Mathematics, for their support and fruitful discussions concerning the FFEA method during my stay in Leeds.
And lastly, thanks to the members of the graduation committee, for taking the time and making the effort to evaluate this work and to be present at the defence.
A.1 Preparation of buffers

PBS buffer
Protocol for preparing 200 mL, 150 mM, pH 7.4
1. Dissolve 1 PBS tablet in 200 mL of demi water.

BSA buffer
Protocol for preparing 1 mL, 1 mg/mL
1. Weigh 1 mg of BSA;
2. Add 1 mL of demi water.

MES buffer
Protocol for preparing 80 mL, 15 mM, pH 5
1. Weigh 234.4 mg of MES powder;
2. Add 80 mL of demi water;
3. Add acid (e.g. hydrochloric acid) to the solution with a pipette and measure the pH iteratively, until pH 5 is reached (One might consider to add less water in step 2 and fill to 80 mL after adjusting the pH: controlling the pH is easier in smaller volumes);

EDC buffer
Protocol for preparing 300 μL
1. Weigh 2.7 mg EDC powder;
2. Add 300 μL of cold (refrigerated) demi water.

Wash buffer
Protocol for preparing 40 mL, pH 8, 0.1% Tween 20
1. Dilute the TBS-T 10× solution to 1×: add 36 mL of demi water to 4 mL of TBS-T;
2. Add base (e.g. sodium hydroxide) to the solution with a pipette and measure the pH iteratively, until pH 8 is reached (One might consider to add less water in step 1 and fill to 40 mL after adjusting the pH: controlling the pH is easier in smaller volumes);
3. Slowly and carefully (viscous!) add 40 μL of Tween 20.

Storage buffer
Protocol for preparing 40 mL, pH 7.4, 0.01% Tween 20
1. Dilute the TBS-T 10× solution to 1×: add 36 mL of demi water to 4 mL of TBS-T;
2. Slowly and carefully (viscous!) add 4 μL of Tween 20.
Appendix A. Experimental protocols

A.2 Preparation of protein solutions

**Protein G functionalisation solution**

*Protocol for preparing 150 µL, 6 µM*

1. Add 19.44 µL of protein G in solution (1 mg/mL=46.3 µM) to 130.56 µL MES buffer.

**IgG incubation solution**

*Protocol for preparing 1 mL, 10 nM*

1. Add 1.55 µL of IgG in solution (1 mg/mL=6.45 µM) to 998.45 µL of PBS buffer.

**Streptavidin solution**

*Protocol for preparing 50 µL, 3 µM*

1. Add 0.8 µL of streptavidin in solution (10 mg/mL=189 µM) to 49.2 µL MES buffer.

**Fluorescent label solution**

*Protocol for preparing 100 µL*

1. Add 1 µL of biotinilated fluorescent microspheres (0.2 µm) in solution to 99 µL of demi water.

A.3 Functionalisation of the M-270 COOH superparamagnetic beads with protein G

*Protocol for preparing 200 µL of M-270 solution*

**Materials**

- 10 µL of M-270 COOH superparamagnetic beads in stock solution;
- 25 µL streptavidin solution
- 100 µL protein G functionalisation solution
- 200 µL freshly prepared EDC buffer
- 400 µL storage buffer
- 600 µL wash buffer
- 600 µL MES buffer
- Demi water

**Instructions**

1. Add 10 µL of M-270 COOH beads to 200 µL of demi water, after gently shaking the bottle;
2. Wash 3 times in 200 µL of demi water, while vortexing for 5-10 minutes each time;
3. Wash 3 times in 200 µL of MES buffer, while vortexing for 5, 5 and 30 minutes, respectively (prepare the EDC buffer in the meantime);
A.4 Torsion experiments on IgG - protein G assays

Protocol for measuring 1 sample

Materials
- 2 clean glass cover slips
- 1 fluid cell sticker
- 2 μL fluorescent label solution
- 10 μL functionalised M-270 COOH beads
- 100 μL IgG incubation solution
- 100 μL BSA buffer
- 800 μL PBS buffer

Instructions

Preparation of sample (biochemical lab)
1. Place the fluid cell sticker on a glass cover;
2. Incubate 100 μL of IgG incubation solution on the glass in the fluid cell, for 45-60 minutes (prepare the BSA buffer in the meantime);
3. Remove the supernatant and wash twice with 100 μL of PBS buffer;
4. Remove the supernatant and incubate 100 μL of BSA buffer, for 30 minutes;
5. Wash twice with 100 μL of PBS buffer, leave the supernatant on the second time.

Preparation of beads (biochemical lab)
6. Add 10 μL of functionalised M-270 COOH beads to 100 μL PBS buffer;
7. Shortly sonicate the solution;
8. Sonicate the fluorescent label solution, add 2 μL to the beads and vortex shortly (30 seconds);
9. Wash twice with 100 μL of PBS buffer;
10. Remove the supernatant and add 200 μL of storage buffer.
11. Shortly sonicate the solution.

Measurement (microscopy lab)

12. Turn on setup;
13. Set the voltage and current on both channels to 24 V and 1.5 A, respectively;
14. Set the magnification of the microscope to 63\times (immersion);
15. Demagnetise the setup for 10–20 seconds, using a magnetic field of 20 mT;
16. Remove the supernatant from the sample;
17. Remove the top sticker of the fluid cell;
18. Place the sample under the microscope;
19. Incubate 8.5 \muL of labelled M-270 beads for 3–5 minutes, in a constant magnetic field of 2 mT.
   Close the fluid cell with a cover glass;
20. After incubation, flip the sample upside down;
21. Start measurement.
Appendix B

Derivation FFEA

B.1 Finite Element Analysis

In order to describe the mathematical background of FFEA [34], we consider the Cauchy momentum equation (2.6):

$$\rho \left( \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} (\nabla \cdot \mathbf{u}) \right) = \nabla \cdot \sigma,$$

with \( \mathbf{u} \) the velocity and \( \sigma \) the stress at all points in the material, and \( \rho \) the density. At the macroscale level, the stress tensor \( \sigma \) is composed of an elastic \( (\sigma^e) \) and a viscous \( (\sigma^v) \) term:

$$\sigma = \sigma^e + \sigma^v,$$

where the Kelvin-Voigt material model for a viscoelastic material is used [40]: the total stress \( \sigma \) is given by the sum of the partial stresses. The material has an isotropic linear viscous stress, which is dependent on the velocity \( \mathbf{u} \):

$$\sigma^v = \mu_v \left( \nabla \cdot \mathbf{u}^T + (\nabla \cdot \mathbf{u}^T)^T \right) + \left( \mu_b - \frac{2}{3} \mu_v \right) (\nabla \cdot \mathbf{u}) I,$$

where \( \mu_v \) and \( \mu_b \) are the shear and bulk viscosity, respectively, \( \nabla \cdot \mathbf{u}^T \) represents the Jacobian matrix and \( I \) is the identity matrix. In element-wise notation:

$$\sigma^v_{ij} = \mu_v \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) + \left( \mu_b - \frac{2}{3} \mu_v \right) \frac{\partial u_m}{\partial x_n} \delta_{ij}, \quad (B.1)$$

where we sum over identical indices (Einstein summation convention). \( \delta_{ij} \) is the Kronecker delta \( (\delta_{ij} = 1 \text{ when } i = j \text{ and } \delta_{ij} = 0 \text{ when } i \neq j) \).

The viscous stress tensor is dependent on the velocities \( \mathbf{u} \) within the material. The elastic stress tensor, however, is entirely derived from the instantaneous deformation of the material.

B.1.1 FEA at mesoscale

In general, molecules are evaluated on a nanoscale level. Models are atomistic (or coarse-grained atomistic) and temperature plays an important role in the kinetics of the particles. FEA, however, usually describes macroscale structures, where thermal fluctuations average out into macroscopic parameters, such as density and thermal expansion. We will now model molecules with an adjusted version of the standard FEA, where the molecules are evaluated on a mesoscale level: atomistic information such as charge and bond strength are entirely disregarded - the molecule is considered to be a continuous material - but thermal energy still plays an important role in the kinetics. In order to simulate molecules using FEA, we therefore need to include a third term in the stress tensor, on top of the elastic and viscous stress, to account for the thermal stress (2.7):

$$\sigma = \sigma^e + \sigma^v + \sigma^t$$

B.1.2 Fluctuation-dissipation theorem

The fluctuating thermal stress tensor \( \sigma^t \) is being calculated using the fluctuation-dissipation theorem. It relates the relaxation of a structure to its statistical fluctuations. Loosely speaking, the kinetic energy a molecule gains from thermal fluctuations needs to be dissipated by viscous effects. Alternatively, the Brownian motion of a structure is dissipated by viscous drag forces and internal
Appendix B. Derivation FFEA

viscosity. In this case, we relate the correlation of the thermal noise vector \( N_p \) to the drag matrix \( K_{pq} \) as follows [34]:

\[
\langle N_p N_q \rangle = \frac{k_B T}{\Delta t} (K_{pq} + K_{qp}),
\]

where \( k_B T \) is the thermal energy (Boltzmann's constant \( k_B \) times temperature \( T \)) and \( \Delta t \) the size of one time step. The elements of the resulting thermal stress tensor \( \sigma' \) then become:

\[
\sigma'_{ij} = \left( \frac{2k_B T}{V \Delta t} \right)^{1/2} \left( \mu_i^{1/2} X_{ij} + \left( \mu_b - \frac{2}{3} \mu_s \right)^{1/2} X^0 \delta_{ij} \right),
\]

with \( V \) the element volume. \( X_{ij} \) is a symmetric stochastic tensor, with 6 independent elements, and \( X^0 \) is a stochastic variable, independent of \( X_{ij} \). Their properties are:

\[
\langle X_{ij} \rangle = 0, \\
\langle X_{ij} X_{kl} \rangle = \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}, \\
\langle X^0 \rangle = 0, \\
\langle X^0 X_{ij} \rangle = 0.
\]

B.2 Derivation of FFEA

To derive the matrix equation (2.8), we start at the Cauchy momentum equation [34]:

\[
\rho \left( \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} (\nabla \cdot \mathbf{u}) \right) = \nabla \cdot \mathbf{\sigma}, \quad (B.2)
\]

where \( \rho \) is the density, \( \mathbf{u} \) the velocity and \( \mathbf{\sigma} \) the stress tensor at all points in the material. Note that the divergence of the stress tensor means a vector with the divergences of the rows of the stress tensor:

\[
\nabla \cdot \mathbf{\sigma} = \begin{pmatrix} \nabla \cdot \sigma_{x1} \\ \nabla \cdot \sigma_{y1} \\ \nabla \cdot \sigma_{z1} \end{pmatrix} = \begin{pmatrix} \partial_x \sigma_{xx} + \partial_y \sigma_{xy} + \partial_z \sigma_{xz} \\ \partial_x \sigma_{yx} + \partial_y \sigma_{yy} + \partial_z \sigma_{yz} \\ \partial_x \sigma_{zx} + \partial_y \sigma_{zy} + \partial_z \sigma_{zz} \end{pmatrix},
\]

or more generally, the dot product of a vector with a tensor means the dot product of that vector with the tensor rows:

\[
\begin{pmatrix} a \\ b \\ c \end{pmatrix} \cdot \mathbf{\sigma} = \begin{pmatrix} a \sigma_{xx} + b \sigma_{xy} + c \sigma_{xz} \\ a \sigma_{yx} + b \sigma_{yy} + c \sigma_{yz} \\ a \sigma_{zx} + b \sigma_{zy} + c \sigma_{zz} \end{pmatrix}.
\]

Here, \( \sigma_{ij} \) indicates the element in the stress tensor at row \( i \) and column \( j \). Alternatively, this means the stress in the \( j \)-direction on a surface with its normal vector in the \( i \)-direction. The notation \( \partial_i \) represents a partial derivative with respect to \( i \).

The method allows for a structure to be discretised into finite elements, which are taken to be homogeneous and only the boundaries of these elements have to be taken into account when evaluating the interactions with their environment.

In order to solve the partial differential equation (B.2) using FEA, it first needs to be written into a weak formulation:

\[
\int_V \rho w \left[ \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} (\nabla \cdot \mathbf{u}) - \frac{1}{\rho} \nabla \cdot \mathbf{\sigma} \right] dV = 0 \quad (B.2)
\]

\[
\int_V \rho w \left[ \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} (\nabla \cdot \mathbf{u}) \right] dV = \int_V w (\nabla \cdot \mathbf{\sigma}) dV = 0
\]

\[
\int_V \rho w \left[ \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} (\nabla \cdot \mathbf{u}) \right] dV = \int_S w (\mathbf{\sigma} \cdot \mathbf{n}) dS + \int_V \nabla w \cdot \mathbf{\sigma} dV = 0
\]

\[
\int_V \rho w \left[ \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} (\nabla \cdot \mathbf{u}) \right] dV = \int_S w dS + \int_V \nabla w \cdot \mathbf{\sigma} dV = 0,
\]
where integration by parts is used. In element-wise notation:

\[
\int_V \rho w \left[ \frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} \right] dV - \int_S Wf dS + \int_V \frac{\partial w}{\partial x_j} a_{ij} dV = 0, \tag{B.3}
\]

where we sum over identical indices, according to the Einstein summation convention. \( w \) is a weight function, which satisfies the boundary conditions and \( f \) are the external surface force functions, which we will take to be zero for now: surface forces will be calculated separately. Equation (B.3) contains first order derivatives of both \( u \) and \( w \), so both the elements of \( u \) and \( w \) must be differentiable over the domain.

If we subsequently take the weight functions \( w \) to be composed of base functions \( \phi_a \), spanning the solution space:

\[
u_i = \sum_a \nu_{ia} \phi_a.
\]

This composition is made such a way, that the material derivative of \( u_i \) (i.e. the derivative experienced while travelling along with the material: \( \frac{Du_i}{Dt} = \frac{\partial}{\partial t} + (\nabla \cdot u) \)) becomes:

\[
\frac{Du_i}{Dt} = \sum_a \frac{\partial \nu_{ia}}{\partial t} \phi_a.
\]

If we substitute this in equation (B.3), we find the following equation (where we sum over indices \( j \) and \( a \)):

\[
\int_V \rho w \frac{\partial \nu_{ia}}{\partial t} \phi_a dV + \int_V \left[ \mu_s \frac{\partial w}{\partial x_j} \frac{\partial \nu_{ia} \phi_a}{\partial x_j} + \mu_s \frac{\partial w}{\partial x_i} \frac{\partial \nu_{ja} \phi_a}{\partial x_j} + \left( \mu_b - \frac{2}{3} \mu_s \right) \frac{\partial w}{\partial x_i} \frac{\partial \nu_{ja} \phi_a}{\partial x_j} \right] dV
\]

\[
+ \int_V \frac{\partial w}{\partial x_j} \sigma_{ij} dV + \int_V \frac{\partial w}{\partial x_i} \sigma_{ij} dV = 0
\]

If we subsequently take the weight functions \( w \) to be the same as the basis functions \( \phi_a \), (corresponding to the Galerkin formulation) we find:

\[
\int_V \rho \phi_a \phi_b \phi dV \frac{\partial \nu_{ia}}{\partial t} + \int_V \left[ \mu_s \frac{\partial \phi_b}{\partial x_k} \frac{\partial \phi_a}{\partial x_k} \nu_{ia} + \mu_s \frac{\partial \phi_b}{\partial x_j} \frac{\partial \phi_a}{\partial x_j} \nu_{ja} + \left( \mu_b - \frac{2}{3} \mu_s \right) \frac{\partial \phi_b}{\partial x_i} \frac{\partial \phi_a}{\partial x_j} \nu_{ja} \right] dV
\]

\[
+ \int_V \frac{\partial \phi_b}{\partial x_j} \sigma_{ij} dV + \int_V \frac{\partial \phi_b}{\partial x_i} \sigma_{ij} dV = 0
\]

\[
\int_V \rho \phi_a \phi_b \phi dV \delta_{ij} \frac{\partial \nu_{ja}}{\partial t} + \int_V \left[ \mu_s \frac{\partial \phi_b}{\partial x_k} \frac{\partial \phi_a}{\partial x_k} \delta_{ij} + \mu_s \frac{\partial \phi_b}{\partial x_j} \frac{\partial \phi_a}{\partial x_j} + \left( \mu_b - \frac{2}{3} \mu_s \right) \frac{\partial \phi_b}{\partial x_i} \frac{\partial \phi_a}{\partial x_j} \right] dV \nu_{ja}
\]

\[
+ \int_V \frac{\partial \phi_b}{\partial x_j} \sigma_{ij} dV + \int_V \frac{\partial \phi_b}{\partial x_i} \sigma_{ij} dV = 0,
\]

or if we relabel the indices to \( p \) and \( q \), which are both functions of \((i, \beta)\) and \((j, \alpha)\), respectively:

\[
M_{pa} \frac{\partial q_a}{\partial t} + K_{pq} q_a = E_p + N_p, \tag{B.4}
\]
Appendix B. Derivation FFEA

with

\[ M_{pq} = \delta_{ij} \int_V \rho \phi_i \phi_j dV/m \]
\[ K_{pq} = \int_V \left[ \mu_s \frac{\partial \phi_i}{\partial x_j} \frac{\partial \phi_j}{\partial x_i} + \mu_b \left( \frac{2}{3} \mu_s \right) \frac{\partial \phi_i}{\partial x_k} \frac{\partial \phi_k}{\partial x_i} \right] dV, \]
\[ E_p = -\int_V \frac{\partial \phi_i}{\partial x_j} \sigma_{ij} dV, \]
\[ N_p = -\int_V \frac{\partial \phi_i}{\partial x_j} \sigma_{ij} dV. \]

\( M_{pq} \) represents a mass matrix, \( K_{pq} \) is responsible for internal drag forces, and \( E_p \) and \( N_p \) are an elastic and thermal noise vector, respectively, making equation (B.4) the finite element equivalent of Newton’s second law.

B.3 Creating the mesh

We described a simulation method in which we do not need any atomistic information on the molecule to evaluate the dynamics: all information concerning bond length, bond strength, bond angles, charge, atom radius, etc. is discarded, and only the overall shape of the molecule is taken into account. Subsequently, the molecule is parametrised with macroscopic material parameters, such as the density \( \rho \), bulk and shear modulus \( K \) and \( G \), and bulk and shear viscosity \( \mu_b \) and \( \mu_s \). Due to the loss of the atomistic detail, using FFEA to model the kinetics of proteins has several advantages over atomistic simulations. Firstly, a coarse-grained model such as FFEA is, in general, computationally less expensive than an atomistic model such as molecular dynamics. Moreover, FFEA is not only compatible with the same input as a molecular protein model: a Protein Data Bank (PDB) file, containing the positions and types of the protein atoms, but also with less detailed molecular data such as an electron density map (EDM). Both PDB and EDM data result from experiments, although from different methods: PDB data mostly result from X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy experiments [51]. EDM data represent the distribution of electrons in a molecule, indicating the structure with a lower resolution (without knowledge of all individual atoms), and they result from cryo-electron microscopy (cryoEM) experiments [35, 52]. In fact, in order to create a mesh from a PDB file, we need to first convert it into an artificial EDM. We replace the atoms in the PDB file by spheres with a certain radius (representing the electron clouds around the atoms). Subsequently, the outer surface of this EDM is used to create a surface mesh. As the resulting mesh is generally created with too high a resolution, we coarsen the mesh by deleting edges shorter than a certain threshold length: both nodes on a long edge are basically combined into one, in between the nodes, while the volume is conserved [26]. The mesh creation process is depicted in figure B.1. With the coarse mesh obtained, we now fill it with tetrahedra using NETGEN [53], a mesh generator software, in order to create a volume mesh. This volume mesh is ultimately used as the structure input for FFEA simulations.

We further evaluate the accuracy of the created volume mesh by comparing the mesh to a radial distribution function (RDF) obtained from small angle X-ray scattering (SAXS) experiments [54]. This RDF represents the probability of finding atoms at a distance \( r \), weighed by the electron density of the atoms (figure B.3a). Therefore, we calculate the distance \( r \) between all nodes in the IgG volume mesh (see figure B.2) and calculate the probability of finding a certain distance \( r \), see figure B.3b. The comparison is shown in figure B.3. Note that in both cases, the probability is normalised not to unity, but to the maximum probability. As we can see from figure B.3, both RDFs look very similar. The small peaks in the RDF from the volume mesh nodes (figure B.3b), which are absent in the experimental RDF, are likely due to, apart from the mesh itself, the fact that the volume mesh is static, in contrast to a real molecule. Due to thermal motion of the molecule, the experimental RDF gets averaged over time, while the artificial RDF gives slightly more insight into the internal structure, as the peaks indicate a structural increase in probability to encounter a certain distance \( r \).
B.3. Creating the mesh

Figure B.1: A schematic illustration of the mesh creation process.

Figure B.2: All nodes in the IgG volume mesh.

Figure B.3: Radial distribution functions of an IgG molecule.

(a) Radial distribution function from SAXS experiments [54].

(b) Radial distribution function from the volume mesh nodes.
B.4 Relating FFEA to atomistic simulations

Previous work has attempted to evaluate the similarities and differences between FFEA and higher resolution simulations, such as molecular dynamics [39]. This section serves as a summary of the most important methods and results. Part of a different molecule, the stalk of dynein, was investigated, see figure B.4. Dynein is a molecular motor protein; it moves along the cytoskeleton (the network of fibres which provides rigidity of the cell) of eukaryotic cells (cells with a cell nucleus). Dynein transports cargo along microtubules, towards the cell nucleus, and is responsible for the beating of flagella. With the stalk, it binds to the microtubules, while the motor domain dissipates ATP to undergo a conformational change and so produce motion of the molecule.

For the dynein stalk, both MD and FFEA simulations were performed subject to thermal forces, using a full-atom model and coarse-grained mesh, respectively (see figure B.5). The mesh had been created from the full-atom PDB information, as described in section B.3.

Figure B.4: Dynein; a molecular motor. The stalk and motor domain are indicated [39].

Figure B.5: The full-atom model and coarse mesh used in the MD and FFEA simulations, respectively [39].

B.4.1 Principal Component Analysis

After the simulations, the data was analysed using Principal Component Analysis (PCA). This is an analysis method to extract the principal modes, the most prominent motions, of the molecule from the simulation data. The coordinates $x_i$ of all atoms or nodes in the structure are decomposed into their temporal average $\bar{x}_i$ and the deviation $\delta_i(t)$ from that average:

$$x_i(t) = \bar{x}_i + \delta_i(t).$$
B.4. Relating FFEA to atomistic simulations

Subsequently, the elements of the covariance matrix \( C_{ij} \) of the deviations are calculated:

\[
C_{ij} = \langle \delta_i(t) \delta_j(t) \rangle = \frac{1}{N_t} \sum_{k=0}^{N_t} \delta_i(k\Delta t) \delta_j(k\Delta t),
\]

with \( N_t \) the number of time steps considered and \( \Delta t \) the size of a time step. PCA finds a number of largest eigenvalues \( \lambda_n \) of this covariance matrix \( C \). \( C \) is symmetrical and can be diagonalised by matrices \( Q \) and \( R \), where the diagonal matrix \( Q \) contains the eigenvalues \( \lambda_n \) in descending order and the columns of \( R \) contain the corresponding orthonormal eigenvectors \( e_n \): \( C = RQ^T \). The number of the mode is now given by the index \( n \), and the corresponding motion is described by \( e_n \). \( \lambda_n \) represents the total variance explored by mode \( n \).

We compare the principal modes of the two models by calculating and diagonalising both covariance matrices. A visual comparison of the first mode between the MD and FFEA simulations is given in figure B.6. As we can see from figure B.6, the first mode corresponds virtually perfectly, apart from a small deviation in amplitude. In order to make a more formal comparison between the two methods, the dot product between the eigenvectors of the two methods was calculated. As the dimensions of the two vectors are not equal (the number of atoms is greater than the number of mesh nodes), the atoms in the MD structure were mapped onto the FFEA mesh. The result is shown in figure B.7.

From figure B.7, we can see that the first four modes correspond almost perfectly between the two methods, as the dot product is close to unity. These four modes together capture more than 80 percent

---

**Figure B.6:** Comparison of the first mode of the dynein stalk between the MD (green) and FFEA (blue) simulations [39].

**Figure B.7:** Comparison of the dot products between the first 8 eigenvectors of the MD and FFEA modes [39].
of the total variance in both models. Alternatively: the first four modes contain the most important information concerning the motion of the dynein stalk. Moreover, mode number 6 also corresponds reasonably well, as do 5 and 7, only the mode numbers seem to have swapped from one model to the other. The dot product decreases for higher mode numbers, which is explained by the fact that FFEA considers a coarse-grained representation of the same molecule: higher order, less prominent modes in the MD model vanish into the individual element of the coarse FFEA mesh.

Conclusively, for the dynein stalk, the MD and FFEA methods correspond well; coarse-graining this structure into a volume mesh is justified. This is a notable result: in order to simulate the molecule, we can discard all atomistic information and we just consider the overall shape to predict the molecule’s dynamics.
We will, instead of integrating the pairwise forces over the surface of each surface face, use a discrete sum by imposing the Gaussian quadrature method. Similar to conventional discrete integrations, we will define a set of Gaussian integration points for which the forces are calculated and sum these in a weighted manner, depending on their position. We evaluate the following integral over a triangle:

\[ \int_A f(p) \, dA, \]  

with \( f \) an arbitrary function, \( p \) points on the triangle and \( A \) the surface. Let the triangle with node positions \( x_i \) \((i = 1, 2, 3)\) be defined by vectors \( x_1 \) (to define the position), \( \xi \) and \( \eta \) and their respective path parameters \( \xi \) and \( \eta \), both between 0 and 1, as depicted in figure C.1 [55]. The lengths of the vectors \( \xi \) and \( \eta \) are denoted as \( l_\xi \) and \( l_\eta \) respectively. We can also rewrite the points \( p \) in terms of \( \xi \) and \( \eta \), so

\[ f(p) = f(\xi, \eta). \]

An area element of the triangle \( dA \) is constructed by selecting regions \( l_i \, d\xi \) on both vectors, as depicted in figure C.2. The total area \( A \) of the triangle is given by

\[ A = \frac{1}{2} |\xi \times \eta| = \frac{1}{2} l_\xi l_\eta \sin(\angle(\xi, \eta)), \]

where \( \angle(\xi, \eta) \) represents the angle between \( \xi \) and \( \eta \). The area of the surface element \( dA \) is given by [56]:

\[ dA = l_\xi d\xi l_\eta d\eta \sin(\angle(\xi, \eta)) = 2Ad\xi d\eta. \]

So the surface integral in expression (C.1) can be rewritten as:

\[ \int_A f(p) dA = 2A \int_0^1 \int_0^{1-\eta} f(\xi, \eta) d\xi d\eta. \]

We now choose \( f(\xi, \eta) \) as the general product:

\[ f(\xi, \eta) = \xi^i \eta^j. \]
Appendix C. Derivation Gaussian integration points

Figure C.2: Surface element $dA$ of the triangle, spanned by two regions on both vectors: $l_\xi d\xi$ and $l_\eta d\eta$.

where $i$ and $j$ are nonnegative integers. We evaluate the integral:

$$
\int_0^1 \int_0^{1-\eta} f(\xi, \eta) d\xi d\eta = \int_0^1 \int_0^{1-\eta} \xi^i \eta^j d\xi d\eta = \frac{\Gamma(i+1) \Gamma(j+1)}{\Gamma(i+j+3)},
$$

where $\Gamma(x)$ is the Gamma function. As $i$ and $j$ are nonnegative integers, $\Gamma(i+1) = i!$, so:

$$
\int_0^1 \int_0^{1-\eta} \xi^i \eta^j d\xi d\eta = \frac{i! j!}{(i+j+2)!}.
$$

We will now use this identity to construct the Gaussian integration points. The surface integral is being approximated by a sum over all Gaussian points. In general:

$$
\int_A f(p) dA = 2A \int_0^1 \int_0^{1-\eta} f(\xi, \eta) d\xi d\eta = A \sum_{k=1}^{N_p} w_k f(\xi_k, \eta_k),
$$

with $k$ the index of one of the $N_p$ Gaussian integration points. More simplified:

$$
\int_0^1 \int_0^{1-\eta} f(\xi, \eta) d\xi d\eta = \frac{1}{2} \sum_{k=1}^{N_p} w_k f(\xi_k, \eta_k). \tag{C.2}
$$

In order to construct the Gaussian point positions and their weights, we need the approximation of the previously discussed surface integral $\int_0^1 \int_0^{1-\eta} \xi^i \eta^j d\xi d\eta$ to be exact up to a certain order $O_p (i+j \leq O_p)$. The number of Gaussian points needed is dependent on both the order and the required symmetry of the triangle; we need a three-fold symmetry in order to represent the triangle appropriately. However, the lower boundary on the number of points is defined by the dimension $D_O$ of the function space at a certain order. This is given by:

$$
D_O = \frac{(O_p + 1)(O_p + 2)}{2}.
$$

On multiple Gaussian points, we impose the requirement that points with a permutation of $\xi_i$ and $\eta_i$ share the same weight. So: if $(\xi_1, \eta_1) = (a, b)$ and $(\xi_2, \eta_2) = (b, a)$, we make sure that $w_1 = w_2$. 
C.1 Constructing Gaussian points

C.1.1 First order

For the first order, we find three equations:

\[
\begin{align*}
\int_0^1 \int_0^{1-\eta} 1 \, d\xi \, d\eta &= \frac{1}{2} = \frac{1}{2} \sum_{k=1}^{N_p} w_k, \\
\int_0^1 \int_0^{1-\eta} \xi \, d\xi \, d\eta &= \frac{1}{6} = \frac{1}{2} \sum_{k=1}^{N_p} w_k \xi_k, \\
\int_0^1 \int_0^{1-\eta} \eta \, d\xi \, d\eta &= \frac{1}{6} = \frac{1}{2} \sum_{k=1}^{N_p} w_k \eta_k,
\end{align*}
\]

which we can all uniquely satisfy by defining 1 Gaussian point; we will have three equations for three variables, namely \(w_1\), \(\xi_1\) and \(\eta_1\). The solution is given by \(w_1 = 1\), \(\xi_1 = \frac{1}{3}\) and \(\eta_1 = \frac{1}{3}\).

\[
\begin{array}{c}
\text{Figure C.3: Illustration of the first order Gaussian point.}
\end{array}
\]

C.1.2 Second order

In order to construct the second order Gaussian points, we need to satisfy equation (C.2) for the following set of functions:

\[f(\xi, \eta) = \{1, \xi, \eta, \xi^2, \eta^2, \xi \eta\}.\]
Appendix C. Derivation Gaussian integration points

This leads to 6 equations:

\[
\begin{align*}
\frac{1}{2} &= \frac{1}{2} \sum_{k=1}^{N_p} w_k, \\
\frac{1}{6} &= \frac{1}{2} \sum_{k=1}^{N_p} w_k \xi_k, \\
\frac{1}{6} &= \frac{1}{2} \sum_{k=1}^{N_p} w_k \eta_k, \\
\frac{1}{12} &= \frac{1}{2} \sum_{k=1}^{N_p} w_k \xi_k^2, \\
\frac{1}{12} &= \frac{1}{2} \sum_{k=1}^{N_p} w_k \eta_k^2, \\
\frac{1}{24} &= \frac{1}{2} \sum_{k=1}^{N_p} w_k \xi_k \eta_k.
\end{align*}
\]

We could satisfy these equations by choosing 2 Gaussian points (6 equations, 6 variables), but due to the symmetry of the triangle, we will need 3. If we, on top of those 6 equations, require all variables to be positive and define \( w_1 = w_2 = w_3 = \frac{1}{3} \) (due to symmetry), we find a solution as shown in Table C.1.

Table C.1: Weights and coordinates of the second order Gaussian points.

<table>
<thead>
<tr>
<th>( k )</th>
<th>( w_k )</th>
<th>( \xi_k )</th>
<th>( \eta_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \frac{1}{3} )</td>
<td>( \frac{1}{6} )</td>
<td>( \frac{1}{6} )</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{3} )</td>
<td>( \frac{1}{3} )</td>
<td>( \frac{1}{3} )</td>
</tr>
<tr>
<td>3</td>
<td>( \frac{1}{3} )</td>
<td>( \frac{1}{3} )</td>
<td>( \frac{1}{3} )</td>
</tr>
</tbody>
</table>

Figure C.4: Illustration of the second order Gaussian points.
C.1.3 Fourth order

Fourth order Gaussian points are the next and final example of the construction of Gaussian points. We will have to satisfy equation (C.2) for the following set of functions:

\[ f(\xi, \eta) = \{1, \xi, \eta, \xi^2, \eta^2, \xi \eta, \xi^3, \eta^3, \xi^2 \eta, \eta^2 \xi, \xi \eta^2, \xi^3 \eta, \xi^2 \eta^2, \xi \eta^3, \xi^3 \eta^2\} \]

This will result in 15 equations, according to equation (C.2). We will therefore need 6 Gaussian integration points, two pairs of three, taking into account the symmetry. We will not be able to define the system uniquely by using only these 15 equations, so apart from these equations, we impose (as both pairs will have the same weights and contain two points with a permutation):

\[
\begin{align*}
\xi_1 &= \eta_1, & \xi_4 &= \eta_4, \\
\xi_2 &= \eta_3, & \xi_5 &= \eta_6, \\
\xi_3 &= \eta_2. & \xi_6 &= \eta_5.
\end{align*}
\]

We solve these equations (the original 15 and the additional equations) numerically, and find the values for \( w_k, \xi_k \) and \( \eta_k \) as shown in table C.2.

<table>
<thead>
<tr>
<th>( k )</th>
<th>( w_k )</th>
<th>( \xi_k )</th>
<th>( \eta_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.109951743655322</td>
<td>0.091576213509771</td>
<td>0.091576213509771</td>
</tr>
<tr>
<td>2</td>
<td>0.109951743655322</td>
<td>0.816847572980459</td>
<td>0.091576213509771</td>
</tr>
<tr>
<td>3</td>
<td>0.109951743655322</td>
<td>0.091576213509771</td>
<td>0.816847572980459</td>
</tr>
<tr>
<td>4</td>
<td>0.223381589678011</td>
<td>0.445948490915965</td>
<td>0.445948490915965</td>
</tr>
<tr>
<td>5</td>
<td>0.223381589678011</td>
<td>0.108103018168070</td>
<td>0.445948490915965</td>
</tr>
<tr>
<td>6</td>
<td>0.223381589678011</td>
<td>0.445948490915965</td>
<td>0.108103018168070</td>
</tr>
</tbody>
</table>

We solve these equations (the original 15 and the additional equations) numerically, and find the values for \( w_k, \xi_k \) and \( \eta_k \) as shown in table C.2.

Figure C.5: Illustration of the fourth order Gaussian points.

C.2 Transformation to barycentric coordinates

As all positions within surface faces in the model are represented in barycentric coordinates (coordinates as function of the positions of the three nodes of the triangle), we will now describe the
Appendix C. Derivation Gaussian integration points

transformation from positions in terms of $\xi$ and $\eta$ to positions in terms of $\lambda_i$ ($i = 1, 2, 3$), where $\lambda_i$ is the fraction of the position that vector $x_i$ accounts for. Any position $x$ is thus defined in terms of $x_i$:

$$x = \sum_{i=1}^{3} \lambda_i x_i,$$

as opposed to the current representation:

$$x = x_1 + \xi \xi + \eta \eta.$$

We can write the vectors $\xi$ and $\eta$ in terms of $x_i$:

$$\xi = x_2 - x_1, \quad \eta = x_3 - x_1.$$

So:

$$x = x_1 + \xi \xi + \eta \eta$$
$$= x_1 + \xi (x_2 - x_1) + \eta (x_3 - x_1)$$
$$= (1 - \xi - \eta) x_1 + \xi x_2 + \eta x_3,$$

or

$$\lambda_1 = 1 - \xi - \eta,$$
$$\lambda_2 = \xi,$$
$$\lambda_3 = \eta.$$

The Gaussian point parameters will then become as shown in table C.3.

Table C.3: Weights and coordinates of the Gaussian points in barycentric coordinates.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$w_k$</th>
<th>$\lambda_1^k$</th>
<th>$\lambda_2^k$</th>
<th>$\lambda_3^k$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First order</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{1}{3}$</td>
</tr>
<tr>
<td></td>
<td>Second order</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{2}{3}$</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{1}{3}$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{2}{3}$</td>
<td>$\frac{1}{3}$</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{2}{3}$</td>
</tr>
<tr>
<td></td>
<td>Fourth order</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.109951743655322</td>
<td>0.816847572980459</td>
<td>0.091576213509771</td>
<td>0.091576213509771</td>
</tr>
<tr>
<td>2</td>
<td>0.109951743655322</td>
<td>0.091576213509771</td>
<td>0.816847572980459</td>
<td>0.091576213509771</td>
</tr>
<tr>
<td>3</td>
<td>0.109951743655322</td>
<td>0.091576213509771</td>
<td>0.091576213509771</td>
<td>0.816847572980459</td>
</tr>
<tr>
<td>4</td>
<td>0.223381589678011</td>
<td>0.108103018168070</td>
<td>0.445948490915965</td>
<td>0.445948490915965</td>
</tr>
<tr>
<td>5</td>
<td>0.223381589678011</td>
<td>0.445948490915965</td>
<td>0.108103018168070</td>
<td>0.445948490915965</td>
</tr>
<tr>
<td>6</td>
<td>0.223381589678011</td>
<td>0.445948490915965</td>
<td>0.445948490915965</td>
<td>0.108103018168070</td>
</tr>
</tbody>
</table>
**Appendix D**

**Deformation quantities**

### D.1 Deformation gradient tensor

A deformation \( x(X) \) maps points initially located at \( X \) to a new position in target space in time as (equation (4.1)) \(^{44, 45}\)

\[
X \mapsto x(X).
\]

We may expand the deformation \( x(X) \) around some origin \( O \), to obtain

\[
x_i(X) \approx x_i(O) + \left( \frac{\partial x_i}{\partial X_j} \right)_o X_j + \frac{1}{2} \left( \frac{\partial^2 x_i}{\partial X_j \partial X_k} \right)_o X_j X_k + \ldots
\]

We now simply define the **deformation gradient tensor** \( F \) as the first order deformation approximation:

\[
x(X) = x(O) + F \cdot X.
\]

A nonzero \( x(O) \) results in a translation (no deformation), so without loss of generality we may set it to zero. The tensor \( F \) contains elements

\[
F_{ij} = \frac{\partial x_i}{\partial X_j}.
\]

An illustration of the deformation as described here is depicted in figure D.1.

![Illustration of the deformation](image)

**Figure D.1:** Illustration of the deformation \( x(X) \) mapping \( X \) to a new position in target space, as described in equation (4.1). The deformation includes a translation of the origin \( O \) to a new position \( x(O) \), which does not contribute to the deformation gradient tensor \( F \).

In the tensor in (4.2),

\[
F_{2D} = \begin{pmatrix}
\lambda_x & \gamma_{xy} \\
\gamma_{yx} & \lambda_y
\end{pmatrix},
\]

\( \lambda_i \) represents an extension in direction \( i \): the original dimension in direction \( i \) is multiplied by a factor \( \lambda_i \). An extension stretches (or compresses) the material, so surface area (or volume in three dimensions) need not be conserved. The \( \gamma_{ij} \) represents a shear in the \( i \)-direction, as function of \( j \), where surface area is conserved under pure shear. In general, a shear deformation \((i,j) \rightarrow (i',j')\) (where \( i, j = x, y, z \)) can be represented as:

\[
\begin{cases}
  i' = i + \gamma_{ij} \\
  j' = j
\end{cases}
\]
Appendix D. Deformation quantities

D.2 Strain energy density

To simplify expressions in this section and section D.3, we have introduced two widely conventional invariants related to the tensor $F$ [44, 45]:

\[ I_1 = \text{trace} (FF^T), \]

\[ J = \det(F). \]

$I_1$ is the trace of the left Cauchy-Green deformation tensor $FF^T$, which is in essence the sum of the squared entries of $F$. $I_1$ is non-negative. The determinant of $F$, $J$, is also known as the Jacobian determinant and describes the relative local volume change in the material:

\[ J = \det(F) = \frac{V}{V_0}, \]

where $V$ and $V_0$ are the instantaneous and reference (rest) volume of the material, respectively.

The most simple example of a strain energy density function is that for an incompressible neo-Hookean material. This is essentially the continuum version of an incompressible harmonic spring. The function $W$ looks as follows:

\[ W_{\text{neo-Hookean}} = \frac{G}{2} (I_1 - 3), \]  

with $G$ the shear modulus. The subtraction of three is due to the fact that for no deformation, $I_1 = 3$ (as $F$ would be the identity matrix). Hereby we enforce $W$ to be zero for a material in its equilibrium state. This function is constructed in such a way, that $W$ rises quadratically with any stretching or shear, similar to the harmonic spring.

We will now expand this simple expression into a more complicated, more accurate model for compressible materials. Therefore, we use a strain energy density function $W$ for compressible Mooney-Rivlin materials, as proposed by Ciarlet and Geymonat [41, 42]:

\[ W_{\text{Mooney-Rivlin}} = \frac{G}{2} (I_1 - 3) + 3K - \frac{2G}{12}(J^2 - 1) - \left( \frac{3K - 2G}{6} + G \right) \ln J. \]  

As we can see from equation (D.4), the neo-Hookean model from expression (D.3) is still represented, as the first term on the right hand side. The second term penalises the expansion of the material (note that $J$ represents the relative volume increase), as a harmonic volumetric spring. The final term penalises a volume decrease; as soon as $J < 1$, this term will produce an energy increase. Note that all terms concerning volume change, are related to the bulk modulus $K$.

As we consider molecules, which can only be stretched until a certain amount, we would at some point like to include a stronger diverging energy function upon strong deformations. Therefore, we adjust the Mooney-Rivlin function from expression (D.4) slightly, to correspond with the energy function proposed by Gent [46, 47]:

\[ W_{\text{Gent}} = \frac{G}{2} (I_m - 3) \ln \left( \frac{1 - \frac{I_m - 3}{I_m - 3}}{1 - \frac{I_m - 3}{I_m - 3}} \right) + \frac{3K - 2G}{12}(J^2 - 1) - \left( \frac{3K - 2G}{6} + G \right) \ln J. \]

Here we have replaced the neo-Hookean term from (D.3) with a slightly more complicated expression, and we have introduced a new parameter, $I_m$, which is the value of $I_1$ for which the energy diverges. In other words, the smaller $I_m$, the stiffer the material. Note that this parameter $I_m$ should be greater than 3, as the latter corresponds to the unperturbed state of the material. Although it is more complicated, the first term on the right hand side nevertheless reduces to the simple neo-Hookean model, if we
D.2. Strain energy density

take \( I_m \to \infty \):

\[
\begin{align*}
\text{for } x \approx 0: \quad \ln(1 - x) & \approx -x, \\
\text{so for } I_m \to \infty : \\
W & = -\frac{G}{2} (I_m - 3) \ln \left( 1 - \frac{I_1 - 3}{I_m - 3} \right) \\
& \approx +\frac{G}{2} (I_m - 3) \frac{I_1 - 3}{I_m - 3} \\
& = \frac{G}{2} (I_1 - 3).
\end{align*}
\]

To gain more insight in the actual strain energy density function \( W \) as function of extension or shear, we evaluate the energy density, using the following deformation gradient tensor:

\[
F = \begin{pmatrix} \lambda & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{D.5}
\]

We plot the energy density functions for the \textit{Neo-Hookean model} \( (W_{\text{Neo-Hookean}}) \), the \textit{Mooney-Rivlin model} \( (W_{\text{Mooney-Rivlin}}) \) and the \textit{Gent model} \( (W_{\text{Gent}}) \) in figure D.2 as function of \( \lambda \) (figure D.2a) and \( \gamma \) (figure D.2b), while keeping \( \gamma = 0 \) and \( \lambda = 1 \), respectively. For the Gent model we demonstrate three values for \( I_m \): \( I_m = 4 \), \( I_m = 8 \) and \( I_m = 16 \). As we can see from figure D.2a, the incompressible neo-Hookean material model is inadequate for modelling a molecule in the simulations, as the minimum for the strain energy density lies at \( \lambda = 0 \). This corresponds to an expansion factor of zero, which means a collapse of the material. The Mooney-Rivlin model, however, would be adequate, as the energy diverges for both small and large values of \( \lambda \), with its minimum at \( \lambda = 1 \). The Gent model is essentially an adaptation of the latter, where we enforce a divergence of the energy density \( W \) at a certain deformation. As we can see, for smaller values of \( I_m \), the energy diverges at smaller deformations.

The resistance to shear is equal for both the neo-Hookean and the Mooney-Rivlin model, as seen in figure D.2b: for both models, the energy density is a quadratic function of \( \gamma \). The Gent model shows a similar behaviour, however, for smaller values of \( I_m \) again the energy diverges at smaller deformations.
Appendix D. Deformation quantities

![Graph showing strain energy density functions](image)

(a) Strain energy density functions $W$ for the deformation gradient tensor $F$ in expression (D.5), as function of $\lambda$, while keeping $\gamma = 0$.

(b) Strain energy density functions $W$ for the deformation gradient tensor $F$ in expression (D.5), as function of $\gamma$, while keeping $\lambda = 1$.

Figure D.2: Strain energy density functions $W$ for the deformation gradient tensor $F$ in expression (D.5).
D.3 Calculating the Cauchy stress tensor

In order to include the strain energy into the FFEA simulations, we need to calculate the Cauchy elastic stress tensor \( \sigma^e \) from the strain energy density function \( W \) as one of the applied stresses in the simulations (along with a viscous stress \( \sigma^v \) and a thermal stress \( \sigma^t \), see chapter 2). To derive the stress from the energy density, we first consider a linearly elastic material with a modulus of \( E \) (in one dimension). This is the continuum equivalent of a harmonic spring, see figure D.3. The stress \( \sigma^e \)

\[
\begin{align*}
\sigma_{xx} &= E \epsilon_{xx}, \\
W &= \frac{1}{2} \sigma_{xx} \epsilon_{xx} = \frac{1}{2} E \epsilon_{xx}^2.
\end{align*}
\]

So reversely, the stress is calculated from the strain energy density as

\[
\frac{dW}{d\epsilon_{xx}} = E \epsilon_{xx} = \sigma^e_{xx}.
\]

In other words, by taking the derivative of the strain energy density function \( W \) to the strain \( \epsilon \), we find the Cauchy stress tensor \( \sigma^e \):

\[
\sigma^e = \frac{\partial W}{\partial \epsilon} \quad \text{or} \quad \sigma^e_{ij} = \frac{\partial W}{\partial \epsilon_{ij}}.
\]

In terms of the deformation gradient tensor \( F \), this is equivalent to

\[
\sigma^e = \frac{1}{J} \frac{\partial W}{\partial F} F^T \quad \text{or} \quad \sigma^e_{ij} = \frac{1}{J} \frac{\partial W}{\partial F_{ik}} F_{jk}.
\] (D.6)

D.3.1 Examples of Cauchy stress tensors from strain energy density functions

In order to derive the strain energy density function \( W \) to the deformation gradient tensor \( F \), we take the simplified invariants from equations (D.1) and (D.2) and derive those to \( F \). For convenience, we right-multiply these expressions with \( F^T \) to simplify the results drastically:

\[
\begin{align*}
\frac{\partial I_1}{\partial F} F^T &= \frac{\partial [\text{trace}(FF^T)]}{\partial F} F^T = 2FF^T, \\
\frac{\partial J}{\partial F} F^T &= \frac{\partial [\det(F)]}{\partial F} F^T = \det(F) I = J I,
\end{align*}
\]
Appendix D. Deformation quantities

with I the identity matrix.
We use the found expressions to derive the Cauchy stress tensor $\sigma^e$ from the strain energy density function $W$, according to equation (D.6). The most simple stress tensor is found for the incompressible neo-Hookean model:

$$\sigma_{\text{neo-Hookean}}^e = G \frac{1}{J} F F^T.$$ 

The Mooney-Rivlin model includes this term as well, similar to the expression for $W$, and has an extra contribution on the diagonal:

$$\sigma_{\text{Mooney-Rivlin}}^e = \frac{G}{J} F F^T + \left[ \frac{3K - 2G}{6} \left( J - \frac{1}{J} \right) - \frac{G}{J} \right] I.$$

The final expression, the Cauchy stress for the Gent model, includes an extra fraction with parameter $I_m$ in the neo-Hookean term, from which it is clearly visible that the stresses will diverge when $I_1$ reaches the value of $I_m$:

$$\sigma_{\text{Gent}}^e = \frac{G}{J} \frac{l_m - l_1}{l_m - l_1} F F^T + \left[ \frac{3K - 2G}{6} \left( J - \frac{1}{J} \right) - \frac{G}{J} \right] I.$$