Eivind Bering

Stretching, breaking, and dissolution of polymeric nanofibres by computer experiments

Thesis for the degree of Philosophiae Doctor

Trondheim, June 2021

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Natural Sciences
Department of Physics

© Eivind Bering

ISBN 978-82-326-5216-7 (printed ver.)
ISBN 978-82-326-5764-3 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (electronic ver.)

Doctoral theses at NTNU, 2021:176

PRINTED IN
NORWAY
NO - 1598

NORDIC SWAN ECOLABEL
Printed matter
2041 0731

Printed by Skipnes Kommunikasjon AS

Abstract

Bundles of polymeric materials are ubiquitous and play essential roles in biological systems, and often display remarkable mechanical properties. With the never-ending experimental advances in control and manipulation of molecular properties on the nanometric level follows an increasing demand for a theoretical description that is valid at this scale. This regime of nano-scale bundles of small numbers of molecules has not been investigated much theoretically; here chain–chain interactions, surface effects, entropy, nonlinearities, and thermal fluctuations all play important roles.

In this thesis, I present a broad exploration by molecular-dynamics simulations of single chains and bundles under external loading. Stretching and rearrangements of chains are investigated, as well as their breaking and dissolution.

The first article studies the response to external load in nanofibres composed of a typical polymer, polyethylene-oxide (PEO). We cover the full range from unloaded fibres up to and including their breaking, focusing on all features that arise from chain-chain interactions and collective behaviour of the chains. Density functional computations provide a benchmark to gauge and validate the empirical force field approach, and offer an intriguing view of the bundle chemical evolution after breaking. The size-dependence for bundles under moderate loading is systematically explored through structural cross-sections, cohesive energies and elastic properties below and above the glass transition temperature. A remarkably high Young's modulus on the order of 100 GPa was estimated with DF and MD, explained by the semi-crystalline state of the fibres giving mechanical properties comparable to those of carbon nanotubes and of graphene. The bundle breaks whenever the potential energy is raised above its metastability range, but also below that limit due to creep activated by thermal fluctuations. A Kramer's-type approximation for the rate of chain breaking is proposed and compared to simulation data.

Single-molecular systems of the same polymer, PEO, have been systematically studied to understand the applicability of thermodynamics in the small scale limit. Isometric and isotensional single-molecule stretching experiments and their theoretical interpretations have shown the lack of a thermodynamic limit at those scales and the non-equivalence between their corresponding statistical ensembles. This

disparity between thermodynamic results obtained in both experimental protocols can also be observed in the entropy production, as previous theoretical results have shown. In the work presented in the second article, this theoretical framework is applied to find the entropy-production associated with stretching of two different statistical ensembles for two different system sizes. In the smallest system, they are different up to a factor of two, and for the bigger system the difference is smaller, as predicted. In this way, we provide numerical evidence that a thermodynamic description is still meaningful for the case of single molecule stretching.

With the stretching energies for these systems, the third article in this thesis documents for the first time that the Helmholtz and Gibbs energy from single-molecule stretching can be related by a Legendre–Fenchel transform. This opens up a possibility to apply this transform to other systems which are small in Hill's sense, incrementally extending the applicability of the powerful framework of thermodynamics.

The final article is more applied, were we study a long standing problem within the cellulose industry. Cellulose is a highly abundant bio-degradeable and renewable material with a vast range of applications from new material development to biofuels, as a food additive, for biomedical applications and in fabrics and more. Particularly for packaging and fabric applications, it is desirable to be able to dissolve cellulose into smaller units, especially single chains. A considerable amount of effort has gone into research on how this can be achieved in an effective and environmentally friendly manner. Our simulation study seeks to explain one of the most promising methods for dissolving cellulose in laboratory experiments, which makes use of a mixture solvent consisting of water with Na^+, OH^- and urea, which is energetically agitated at temperatures around 258 K.

The simulations explore the influence of the agitation in both water and the mixture solvent. While agitation appears to be essential to obtain dissolution in the simulations, the bundle with no agitation in the mixture solution also swells up with significant amounts of urea entering the bundle, as well as more water than in the bundles in the pure water. Under agitation the bundles dissolve in both solvents, but the dissolution is incomplete and the remaining structures are qualitatively different. Our findings highlight the importance of urea in the solvent, as well as the hydrophobic interactions, and are consistent with experimental results.

List of Articles

Articles included in this thesis

I. **E. Bering** and A. S. de Wijn,

Stretching and breaking of PEO nanofibres. A classical force field and ab initio simulation study,

Soft Matter **16**:11 (2020). DOI: 10.1039/D0SM00089B

Contribution: MD simulations — data analysis — paper drafting, review and editing

II. **E. Bering**, S. Kjelstrup, D. Bedeaux, J. M. Rubi, and A. S. de Wijn,

Entropy Production beyond the Thermodynamic Limit from Single-Molecule Stretching Simulations,

The Journal of Physical Chemistry B **124**:40 (2020). DOI: 10.1021/acs.jpcb.0c05963

Contribution: simulations — data analysis — paper drafting, review and editing

III. **E. Bering**, D. Bedeaux, S. Kjelstrup, A. S. de Wijn, I. Latella, and J. M. Rubi,

A Legendre–Fenchel Transform for Molecular Stretching Energies,

Nanomaterials **10**:12 (2020). DOI: 10.3390/nano10122355

Contribution: simulations — data analysis — paper drafting, review and editing

IV. **E. Bering**, J. Ø. Torstensen, A. Lervik, and A. S. de Wijn

A computational study of cellulose dissolution under agitation in water and a water/NaOH/urea mixture,

Draft

Contribution: simulations — data analysis — paper drafting

Additional scientific contributions

1 J. T. Kjellstadli, **E. Bering**, M. Hendrick, S. Pradhan, and A. Hansen,

Can Local Stress Enhancement Induce Stability in Fracture Processes? Part I: Apparent Stability,

Frontiers in Physics **7**:105 (2019). DOI: 10.3389/fphy.2019.00105.

2 J. T. Kjellstadli, **E. Bering**, M. Hendrick, S. Pradhan, and A. Hansen,

Can Local Stress Enhancement Induce Stability in Fracture Processes? Part II: The Shielding Effect,

Frontiers in Physics **7**:156 (2019). DOI: 10.3389/fphy.2019.00156.

Contribution: These articles are partly built on findings presented in my master's thesis [1], where I worked closely with the other authors. After this I mainly contributed with discussions and minor reviewing.

Preface

This thesis is submitted to the Norwegian University of Science and Technology (NTNU) as a partial fulfillment of the requirements for the degree of Philosophiae Doctor. The work was preformed at the Department of Physics with Professor Alex Hansen as a supervisor, with regular supervision from my co-supervisor Associate Professor Astrid S. de Wijn at the Department of Mechanical and Industrial Engineering, Faculty of Engineering, since the first one and a half year of this doctoral work. Professor Pietro Ballone, currently associated with University College Dublin, Ireland, was the initial principal investigator of this research project. He was a co-supervisor and contributed with intensive supervision for the first one and a half year of this doctoral work, and has been an important resource also after. Signe Kjelstrup and Dick Bedeaux at the Department of Chemistry have played important roles as mentors.

This doctoral work has resulted in three published articles and one article draft, all of which are presented at the end of this thesis. In addition to research, this doctoral program has included coursework amounting to 35.5 ECTS, corresponding to a workload of about one semester, as well as teaching of undergraduate physics laboratory exercises for six semesters, corresponding to a workload of two semesters. A research stay of one month at Weizmann Institute of Science in Israel was scheduled from May 2020, which unfortunately had to be canceled due to the COVID-19 pandemic.

The PhD was funded by the Research Council of Norway under the NANO2021 program as part of project no. 250158 *"Yield kinetics and creep in nanofibrous materials for biomedical applications"* and also through its Centres of Excellence *PoreLab*, with project no. 262644.

Eivind Bering
Trondheim, May 2021

Acknowledgments

"No man is an island entire of itself"

John Donne

First, I would like to extend my profound gratitude to Astrid de Wijn, whom "adopted" me one and a half year in to my PhD, and has been my day-to-day supervisor ever since. Thank you for believing in me, and for giving me the space to grow.

I started my PhD under the supervision of Pietro Ballone, who taught me a great deal about simulations and the work ethics of a dedicated scientist, and for this I am thankful, and I wish him all the best. I can hardly imagine how I would have be able to complete my first article without his support.

I would also like to express my sincere gratitude my main supervisor Alex Hansen for many years of good collaboration, and not least for making PoreLab possible and for convincing me to join the group. Being an integral part of a larger research group has made a great positive impact on me and my work. While I will only mention some of you, I am highly thankful to all of my colleagues and friends here at PoreLab. Thanks to Signe Kjelstrup for your inclusiveness and your inspirational enthusiasm, and to Dick Bedeaux for your many clever remarks. Thanks to Marie-Laure Oliver for always going the extra mile to create the best possible working conditions. Thanks to Anders Lervik and Jonathan Torstensen for many nice discussions on cellulose and more.

Thanks to my colleagues and friends at the Department of Physics, I have not forgotten you, even though I rarely have been present at the Department since I got an office at PoreLab. Thanks for the academic wine club and the D&D, it was good fun! I am also thankful to the other students and post doc's that have been a part of Astrids group during my PhD, I am happy to see you every Friday, with or without cake.

I was lucky enough to attend the MolSim course by CECAM at the University of Amsterdam in the beginning of my PhD, this was a true booster for motivation that I would recommend to anyone that is entering the field of molecular simulation. My only complain would be that it should have lasted for years rather than weeks.

As for computational resources, I would like to thank the NTNU HPC Group for maintaining the Idun cluster, and to UNINETT Sigma2—the National Infrastructure for High Performance Computing and Data Storage in Norway. I would also like to thank Terje Røsten other maintainers of the Linux cluster at the Department of Physics.

Thanks to the people in the slack line community Lineforeningen for providing me with great opportunities for distraction from my office when the weather was nice. Fortunately for this work perhaps, this does not happen too often in Trondheim. And thanks to NTNUI Calisthenics for accepting me in to the group on short notice during my last year of the PhD, it has really improved my quality of life, which I believe in turn makes everything better, including this thesis. Also the guys that I train with at Finalebanen deserves a special thanks, exercising outdoors in variable weather conditions would not have been the same without you!

Thanks to Ailo, Anders and Jonathan for helpful feedback on the thesis and for moral support towards the end of this work. To my other friends, and to my family: Thank you for your continued support and understanding. Finally, thanks to Ronja for all the moments we have shared so far. I am sure there are more to come!

Contents

Abstract i

List of Articles iii

Preface v

Acknowledgments vii

1 Introduction 1
- 1.1 Outline 2

2 Polymers 3
- 2.1 Models 4
  - 2.1.1 Ideal chains 4
  - 2.1.2 Real chains 10
  - 2.1.3 The Fiber Bundle Model 10
- 2.2 Creep 11
- 2.3 Polyethylene Oxide (PEO) 11
- 2.4 Cellulose 14
- 2.5 Applications for the articles 15
  - 2.5.1 Article I 15
  - 2.5.2 Article II-III 16
  - 2.5.3 Article IV 17

3 Molecular Dynamics 19
- 3.1 Length and time scale 19
- 3.2 The initial configuration 20
- 3.3 The potential energy surface 21
  - 3.3.1 A minimalistic reactive force field 22
- 3.4 Verlet algorithms 22
- 3.5 Long-range interactions 23
- 3.6 Thermostats and barostats 24
  - 3.6.1 Nosé-Hoover 25

3.6.2 Langevin thermostat 26
3.6.3 Berendsen 26
3.7 Ergodicity 27
3.8 Challenges 27
3.9 Applications for the articles 28
3.9.1 Article I 28
3.9.2 Articles II-III 28
3.9.3 Article IV 28

4 Thermodynamics of stretching **29**
4.1 Hill's thermodynamics 30
4.2 Non-equilibrium thermodynamics 30
4.3 The entropy production 31
4.3.1 The isometric ensemble 31
4.3.2 The isotensional ensemble 32
4.4 Thermodynamics of non-additive systems 32
4.5 Applications for the articles 33
4.5.1 Article II 33
4.5.2 Article III 33

5 Conclusions **35**

References **39**

Articles **47**
Article I: Stretching and breaking of PEO nanofibres. A classical force field and *ab initio* simulation study 49
Article II: Entropy Production beyond the Thermodynamic Limit from Single-Molecule Stretching Simulations 69
Article III: A Legendre–Fenchel Transform for Molecular Stretching Energies 81
Article IV: A computational study of cellulose dissolution under agitation in water and a water/NaOH/urea mixture 95

1 Introduction

"Have no fear of perfection; you'll never reach it"

Marie Curie

While very small systems and macroscopic systems often can be well described by established theoretical frameworks, there is a lack of understanding at the intermediate levels. The aim of this work has been to explore this intermediate scale: systems that are large enough to be computationally expensive, yet too far from the thermodynamic limit to allow for standard approaches from statistical physics and thermodynamics. This regime of nano-scale bundles of small numbers of molecules has not been investigated much theoretically; here chain–chain interactions, surface effects, entropy, nonlinearities, and thermal fluctuations all play important roles.

The recurring topic of this thesis is the stretching of relatively small polymeric molecules. We know that mechanical properties of polymers play an important role in their function, in both the man-made and natural context. In living organisms, fibres are usually assembled in bundles, inter-linked to form gels, or incorporated into bio-minerals, giving origin to a variety of tissues such as muscles, cartilage and bones. Present day's technology strives to match the properties of these remarkable materials, in many cases attempting to imitate their hierarchical structures, starting from the molecular building blocks to the texture of macroscopic tissues.

This thesis builds a foundation for a deeper understanding for the aforementioned effects by computer simulations, and hopefully these findings can provide useful insight for future studies. This is a largely unexplored area, and in many ways this thesis is a first attempt at theoretically studying realistic nanofibres.

In the first article, we consider nanofibres of a simple paradigmatic polymer, polyethylene oxide (PEO), made of nanometric bundles, stretched between two rigid, planar clamps. We cover the full range from unloaded fibres up to their breaking point, focusing on all features that arise from chain–chain interactions and collective behaviour of the chains.

In the second article, we look at a single chain of PEO, and calculate the entropy production associated with stretching in a force-controlled and length-controlled ensemble. With this, we provide numerical evidence that a thermodynamic description is still meaningful for the case of single-molecule stretching.

The force-elongation curve of these two ensembles are explored in more detail in the third article, were we show that the stretching energies from the Helmholtz and Gibbs ensemble can be related by a Legendre–Fenchel transform. This is a general result that is applicable for a wide range of systems that exhibit non-additivity, where interaction between the parts of the systems is not negligible in comparison with the total energy.

The last manuscript concerns the dissolution of cellulose from bundles to single chains in a solvent mixture. Here computer simulations with oscillating force- and length-controlled stretching and compression of the bundle is used to mimic agitation of the system. Simulations are preformed both with and without agitation in both water and the solvent mixture. While agitation appears to be an absolute criteria to obtain dissolution of cellulose in experiments, it has not yet been accounted for in simulations.

1.1 Outline

The upcoming chapters gives a general background to relevant topics of this work. Chapter 2 gives a short introduction to selected models to study polymers. Chapter 3 elaborates further how molecular-dynamics simulations in particular can be used as a simulation technique. Chapter 4 gives a brief introduction to the thermodynamics of stretching in small polymeric systems. Then follows conclusions made from this work, before the the scientific articles are presented at the end.

2 Polymers

> *"When you change the way you look at things, the things you look at change."*
>
> Max Planck

Polymers are built up of repetitions of molecular units, *monomers*, that are linked together by covalent bonds [2]. They are ubiquitous in everyday life, ranging from familiar synthetic plastics such as polystyrene to natural biopolymers such as DNA and proteins that are fundamental to biological structure and function. The richness of the properties of these materials has since long diverged into a discipline of its own, and this thesis will barely scratch the surface of the complexities this field has to offer.

The term *polymer* stems from the Greek words *poly* (many) and *meros* (part), and it designates molecules made up of repetitions of some simpler unit. While molecules comprising a small number of repeat units technically are classified as *oligomers*, from the Greek *oligo* (a few), we will not be using this term. This chapter gives a brief introduction to a selection of models describing different aspects of the behaviour of polymeric materials, and two specific polymers will be explored in more detail.

A characteristic feature of polymers is that they exhibit entropic elasticity. Thermodynamic systems have a tendency to maximize their entropy [3], from which the phenomenon of an entropic force emerges [4]. This is why a rubber band is hard to stretch. As the entropic force is a highly general feature, it makes no assumption on the underlying forces. A derivation of this force in an idealized model is presented later.

Polymers are commonly classified in three groups: *thermoplastics*, *elastomers* and *thermosets* [5]. Thermoplastics are linear or branched polymers that turns liquid with heat, elastomers are cross-linked rubbery polymers, and thermosets are normally network polymers in which chain motion is greatly restricted by a high degree of cross-linking, making the material rigid. Thermoplastics are then further separated into those which are crystalline and those which are amorphous. In the amorphous state, the conformation of the polymer chain resemble that of cooked spaghetti, and there is no well defined molecular order. In crystalline polymers, on the other hand, the chains fold together and form ordered regions called *lamellae*. Highly crystalline polymers are then characterized by their melting point T_{m}, after which they turn liquid in a first order transition. Amorphous polymers, on the other hand, is characterized by their glass transition temperatures T_{g}, which presents features of a second-order transition at which they transform abruptly from the hard glassy state to the soft rubbery state. Most thermoplastics are however intermediate.

Evidently, the features of a polymeric material are highly dependent on the degree of polymerization and the morphology, such as branching, cross-linking, and supra molecular structure [6]. We will limit this introduction to linear homopolymers, that is, polymeric chains with two end-points with no branching or cross-linking composed of identical monomers.

2.1 Models

Idealized models are powerful in their analytical accessibility, and are highly useful for the verification and understanding of the behaviour of more complex models and experimental results. We will here briefly introduce some popular models for molecular stretching, and derive the temperature dependence of the entropic force for a single ideal chain.

2.1.1 Ideal chains

At the most idealized level, we can describe a polymeric molecule with the Freely-Jointed Chain (FJC) model: a chain consisting of N links of length b, with no correlation among the directions of the individual links [7]. The molecule will then be described by a random walk, and the mean square end-to-end vector is

$$\langle R_{\mathrm{ee}}^2 \rangle = \sum_{i,j} \langle \boldsymbol{b}_i \cdot \boldsymbol{b}_j \rangle = \sum_{i,j} \langle |\boldsymbol{b}_i|^2 \rangle \delta_{ij} = Nb^2, \tag{2.1}$$

since the direction of a step is independent of the direction of the previous step. The angular brackets indicate an ensemble average, and denote an average over all possible states of the system. Another length that characterizes the polymer is the radius of gyration, which can be defined by

$$\langle R_{\mathrm{g}}^2 \rangle = \frac{1}{N+1} \sum_{i=1}^{N+1} \langle |\boldsymbol{r}_i - \boldsymbol{r}_{cm}|^2 \rangle, \tag{2.2}$$

where $\boldsymbol{r}_i$ is the position vector of node i and $\boldsymbol{r}_{cm}$ is the center of mass. It is a quantity that can describe the size of polymers of any architecture, and can be determined experimentally with e.g. static light scattering or small angle neutron-

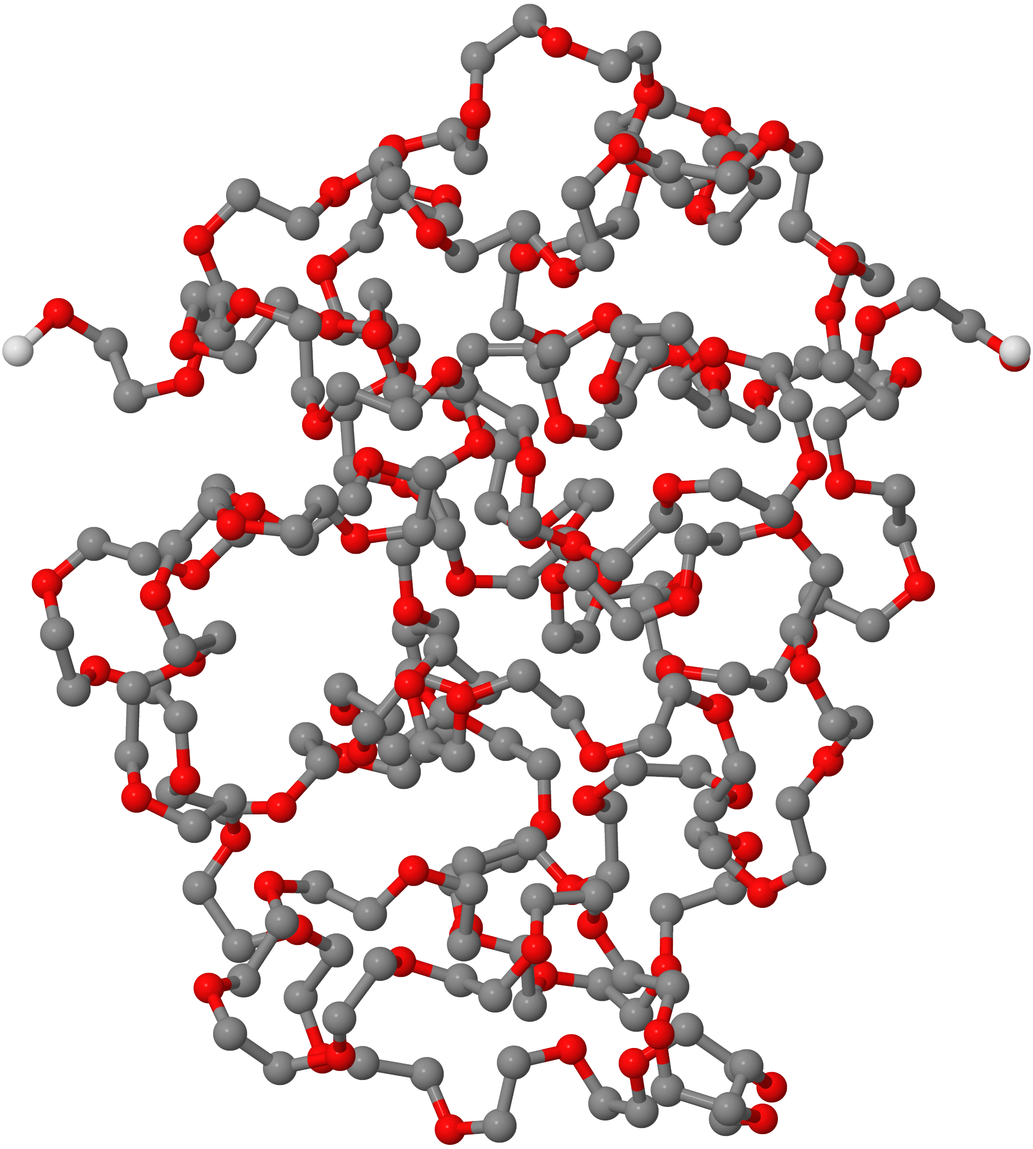


Figure 2.1: A united atom model of a PEO molecule with $n = 108$ monomers. Each monomer is composed of three beads, two methylene groups (gray), and one oxygen atom (red).

and x-ray scattering. For this reason it is a commonly used quantity in polymer physics. In the FJC model one can show that

$$\langle R_{\mathrm{g}}^2 \rangle = \frac{b^2}{6}\left[N - \frac{1}{N}\right], \tag{2.3}$$

and we see that $R_{\mathrm{g}}^2 \simeq R_{\mathrm{ee}}^2/6$ for $N \gg 1$. In its simplicity, the FJC model captures a highly general aspect of polymers: *all chains display ideal chain behaviour for sufficiently long chains, if we only take into account the interactions between neighboring units on the chemical sequence* [8]. That is, as long as we choose b large enough for the segments to be uncorrelated, a sufficently long polymeric chain will be described by a random walk of that step length, neglecting long-range interactions. One can, however, impose a second criterion

$$l_{\mathrm{C}} = Nb \tag{2.4}$$

where we define the contour length l_{C} as the end-to-end distance of the unfolded chain. Combining Eqs. (2.1) and (2.4) gives us a condition for when b is equal to the *Kuhn length* l_{K},

$$b = \frac{\langle R_{\mathrm{ee}}^2 \rangle}{l_C} \equiv l_{\mathrm{K}}, \tag{2.5}$$

and the effective number of segments is then

$$N = \frac{l_{\mathrm{C}}}{l_{\mathrm{K}}} \equiv N_{\mathrm{eff}}, \tag{2.6}$$

such that our polymer is described by a random walk of N_{eff} steps of length l_{K}.

In the framework of Flory [9], the mean-square end-to-end distance of an *unperturbed* chain is characterized by the ratio

$$C_\infty = \frac{\langle R_{\mathrm{ee}}^2 \rangle}{n b_{\mathrm{m}}^2}, \tag{2.7}$$

where n is the number of monomers and

$$b_{\mathrm{m}}^2 = \sum_i a_i^2 \tag{2.8}$$

is the sum of the squares of the length of the backbone bonds a_i in one monomeric unit. The ratio C_∞ is known as the *characteristic ratio* of a polymer chain for a given chemical and structural type. While the ratio $\langle R_{\mathrm{ee}}^2 \rangle / n b_{\mathrm{m}}^2$ is not a constant for low degrees of polymerization, it reaches an asymptotic value C_∞ in the long chain limit. The stiffer the polymer, the higher the characteristic ratio, with typical values in the range of 4 to 12. When the characteristic ratio is known, the chain can be described by a FJC with Kuhn length

$$l_{\mathrm{K}} = \frac{C_\infty n b_{\mathrm{m}}^2}{l_{\mathrm{C}}}. \tag{2.9}$$

The force-elongation curve of the FJC is described by a Langevin function, and is shown in Table 2.1. The FJC model can be extended (FJC+) by replacing the links with elastic springs. In the low force regime, entropic elasticity dominates, and the two models coincide. In the high force regime the segment elasticity E_{S}, with contributions from the potentials for stretching, bending and torsion, will determine the slope of the force-elongation curve. The FJC+ model has been shown to describe single-molecule stretching well in cases with absence of supramolecular assemblies and negligible interchain interactions [10, 11].

Another variation of the FJC model is the Freely Rotating Chain (FRC) model, where we keep the bond angles θ fixed in addition to the bond lengths. This is useful for polymers that are locally very rigid, and becomes random walks only at large length scales. Defining the angle θ as the change in angular direction between consecutive bonds, the mean-square end-to-end distance is

$$\langle R_{\mathrm{ee}}^2 \rangle = \sum_{i,j} \langle \boldsymbol{b}_i \cdot \boldsymbol{b}_j \rangle = N b^2 \frac{1+\cos\theta}{1-\cos\theta}, \tag{2.10}$$

and the reader is referred to e.g. Rubinstein [12] for the derivation. Notice that we again retain $\langle R_{\mathrm{ee}}^2 \rangle^{1/2} \sim N^{1/2}$, which is a main property of ideal chains. The contour length in this model is simply

$$l_{\mathrm{C}} = N b \cos\frac{\theta}{2}. \tag{2.11}$$

The Worm-Like Chain (WLC) model is another important model for polymers, which is also known as the persistent chain model [2]. It is particularly useful for polymers with high stiffness, and that have a uniform flexibility over the whole polymer length. It can be defined from the FRC model in the limit $b \to 0$ and $\theta \to 0$ at constant contour length and *persistence length* l_p [12], related to the Kuhn length by

$$l_{\mathrm{p}} = \frac{l_{\mathrm{K}}}{2}. \tag{2.12}$$

The chain is then described by a curvilinear function with local chain directions given by unit vectors $\boldsymbol{e}(l)$, and the chain flexibility is determined by the orientational correlation function between two points with a distance Δl along the chain

$$K_{\mathrm{or}} = \langle \boldsymbol{e}(l) \boldsymbol{e}(l + \Delta l) \rangle, \tag{2.13}$$

from which we can identify the persistence length by the integral width

$$l_{\mathrm{ps}} = \int_0^\infty K_{\mathrm{or}}(\Delta l) \mathrm{d}(\Delta l), \tag{2.14}$$

which again can be related to the bending modulus E_{b} by

$$l_{\mathrm{p}} = \frac{E_{\mathrm{b}}}{k_{\mathrm{B}} T}. \tag{2.15}$$

Table 2.1: Table summarizing selected statistical-mechanical models for polymer elasticity [13–15]. Here l is the end-to-end length, l_C is the contour length, l_p is the persistence length, K_S is the segment elasticity and ϕ is the specific stiffness of the polymer.

Model	Expression
FJC	$l(f) = l_\mathrm{C}\left(\coth\left(\frac{fl_\mathrm{K}}{k_\mathrm{B}T}\right) - \frac{k_\mathrm{B}T}{fl_\mathrm{K}}\right)$
FJC +	$l(f) = l_\mathrm{C}\left(\coth\left(\frac{fl_\mathrm{K}}{k_\mathrm{B}T}\right) - \frac{k_\mathrm{B}T}{fl_\mathrm{K}}\right) + \frac{nf}{K_\mathrm{S}}$
WLC	$f(l) = \frac{k_\mathrm{B}T}{l_p}\left(\frac{1}{4}\left(1-\frac{l}{l_\mathrm{C}}\right)^{-2} - \frac{1}{4} + \frac{l}{l_\mathrm{C}} - 0.8\left(\frac{l}{l_\mathrm{C}}\right)^{2.15}\right)$
WLC +	$f(l) = \frac{k_\mathrm{B}T}{l_p}\left(\frac{1}{4}\left(1-\frac{l}{l_\mathrm{C}}+\frac{f}{\phi}\right)^{2} + \frac{l}{l_\mathrm{C}} - \frac{1}{4} - \frac{f}{\phi}\right)$

For the force-elongation curve shown in Table 2.1, one can also include the stiffness of the chain as a third fitting parameter to obtain the Extensible Worm Like Chain (WLC+) model, shown the last row in Table 2.1. Note that the last term in the function for the elasticity of the WLC model presented in Table 2.1 was not originally included [7], but reduces the relative error compared to experiments from about 15% to 1% [13].

The entropic force in an idealized chain

From Boltzmann we know that the entropy can be defined as

$$S = -k_\mathrm{B}\sum_i p_i \ln p_i \tag{2.16}$$

where p_i is the probability that a microstate i is occupied, and k_B is Boltzmann's constant. By assuming *equal a priori probability*, that the occupation of any microstate is equally probable, we can write Eq. (2.16) in terms of the possible microstates of the system Ω

$$S = k_\mathrm{B}\ln\Omega, \tag{2.17}$$

since $p_i = 1/\Omega$. This fundamental assumption of statistical thermodynamics holds in general for isolated systems in equilibrium. The number of possible configurations for a random walk of N steps with and end-to-end distance of l can be written as $\Omega = \Omega(l, N)$ [8]. Going to three dimensions, $\Omega(l, N)$ will scale like

$$\Omega(l, N) \simeq N^{-3/2}\exp\left(-\frac{3}{2}\frac{l^2}{Nb^2}\right), \tag{2.18}$$

which gives us

$$S(l, N) = S(0) - \frac{3k_\mathrm{B}}{2Nb^2}l^2 \tag{2.19}$$

for the entropy of the chain in three dimensions at fixed elongation.

We can convert this to Helmholtz energy, which is defined by

$$F(l, N, T) = U(l, N) - TS(l, N). \tag{2.20}$$

In a model where the internal energy U is independent of the chain configuration, the free energy associated with Eq. (2.19) is then

$$F(l, N, T) = F(0) + \frac{3k_\mathrm{B}T}{2Nb^2}l^2. \tag{2.21}$$

We can then readily obtain the entropic contribution to the force associated with the current elongation of the polymer from the derivative of the free energy with respect to the elongation, giving

$$f_\mathrm{S} = \frac{3k_\mathrm{B}T}{Nb^2}l. \tag{2.22}$$

Combined with Eqs. (2.4) to (2.6), we get the entropic force on our FJC of Kuhn segments as

$$f_\mathrm{S} = \frac{3k_\mathrm{B}T}{l_\mathrm{C}l_\mathrm{K}}l. \tag{2.23}$$

The same expression can also be found when one uses the force as a control parameter rather than the end-to-end distance [12]. The partition function of the FJC can be written as a sum over all conformations

$$Z(f, N, T) = \sum_{\text{states}} e^{-U(f,N)/(k_\mathrm{B}T)}, \tag{2.24}$$

where the internal energy is given by

$$U(f, N) = f\sum_{i=1}^{N} b\cos\theta_i, \tag{2.25}$$

where θ_i now denotes the angle of bond i with the end-to-end vector. Equation (2.24) can then be integrated over all possible bond angles of the chain in three dimensions, after which the average end-to-end distance $l(f)$ of a chain with constant N can be found via the derivative of the Gibbs energy

$$l(f) = -\frac{\partial}{\partial f}G(f, N) = k_\mathrm{B}T\frac{\partial}{\partial f}\ln Z(f, N) \tag{2.26}$$

resulting in the well known *Langevin function*

$$l(f) = l_\mathrm{C}\left(\coth\left(\frac{fl_\mathrm{K}}{k_\mathrm{B}T}\right) - \frac{k_\mathrm{B}T}{fl_\mathrm{K}}\right), \tag{2.27}$$

which can be shown to be similar to Eq. (2.23) in the low force limit.

2.1.2 Real chains

While ideal chains only take local correlations into account, the interactions between monomers separated by many bonds may also be important [12]. While the ideal chain is described by a random walk, *real chains* cannot occupy the same position in space more than once. For this reason, the conformations of the real chain are similar to that of a self-avoiding random walk. The behaviour of a real chain in a solvent may be characterized in a parameter for the *excluded volume*, giving the change in volume due to non-local interactions.

When the solvent is good, the conformations of a real chain are determined by the balance of the effective repulsion energy between monomers, and the entropy loss due to swelling. With *Flory theory* one can make rough estimations of the free energy contributions with a basis in ideal chains. Both the energetic and the entropic contribution are overestimated, resulting in a fortunate cancellation of errors. Flory theory leads to a universal scaling law for the end-to-end distance

$$\langle R_{\mathrm{ee}}^2 \rangle^{1/2} \propto N^{\nu}. \tag{2.28}$$

In a good solvent, the excluded volume is greater than zero, and the scaling exponent $\nu \approx 3/5$ is independent of the quality of the solvent.

At some temperature, any polymer-solvent pair will have an attractive contribution that exactly cancels the steric repulsion. The excluded volume is then zero. At this temperature we have $\nu = 1/2$, and the chain has nearly ideal confirmations. This is known as the *theta point.* In a poor solvent, the effective volume is negative, indicating an effective attraction.

2.1.3 The Fiber Bundle Model

The Fiber Bundle Model is an excellent example of a model that can be classified as metaphorical [17, 18]. Rather than attempting to precisely describe reality, it aims to illustrate non-trivial mechanisms which goes much beyond the specifics of the model itself. The strength of this model is that it can reveal mechanisms too complex to study in realistic models, and that can be completely lost in phenomenological models [19].

The model consists of an assembly of bonds, organised into chains and bundles. Bonds deform elastically up to a critical load marking their sudden breaking, followed by the redistribution of the load among the surviving chains. The model is able to merge random bond-breaking processes with collective, self-organised phenomena such as avalanches, and is typically characterized by quenched disorder; disorder that is not evolving with time.

In this model all fibres in the bundle have the same length x and the same elastic constant κ. Each fiber is assigned a individual threshold t_i, which are assumed to be independent random variables from the same probability distribution. The force carried by fiber i is then simply given by

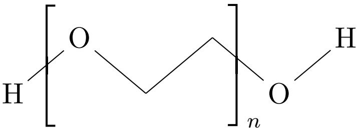


Figure 2.2: The sceletal formula for a PEO molecule composed of n monomers.

$$f_i = \begin{cases} \kappa x & \text{for } x < t_i \\ 0 & \text{for } x \geq t_i. \end{cases} \tag{2.29}$$

There are several different models for redistribution of the load of a failed fiber, the simplest being the *equal-load-sharing model*, where the extra load is distributed equally on all the remaining fibres. This model corresponds to fibres being stretched between rigid clamps. The other extreme is the *local-load-sharing model*, where only the neighboring surviving fibres share the extra load from the broken fibres.

While the fiber bundle model is not a realistic model, or even a phenomenological one, it is still widely acknowledged as one of the most important theoretical frameworks to investigate fracture and breakdown of disordered media, used both by the engineering and physics community, and it exists with numerous extensions [20, 21].

2.2 Creep

While perfectly elastic materials feature stress proportional to the strain, perfectly viscous materials exhibit stress proportional to the *strain rate* [22]. Most polymers are viscoelastic, with characteristics intermediate between perfectly elastic and perfectly viscous behaviour. Time dependent strain is also seen in metals, where it is called anelasticity. The time dependent deformation under static loading is characterized as creep. It is closely related to fatigue, which is the result of cyclic loading. The timescale of deformation is given by the timescale of structure relaxation, and in creep this is typically long, on the order of seconds or even years.

An idealized particle based fiber bundle model has successfully reproduced characteristic stages of creep seen in the experimental investigations of polymeric materials [21]. The smoothened energy landscape is giving rise to much faster system dynamics than convetional more complex models for molecular dynamics, allowing for the study of appropriate time- and length-scales for this phenomena.

2.3 Polyethylene Oxide (PEO)

PEO (also referred to as polyethylene glycol or PEG) is a highly water-soluble thermoplastic polymer with a wide range of applications [23, 24]. With a literature glass transition temperature of 207 K [25], PEO is a liquid or a low-melting

Table 2.2: Parameters Equation 2.33, modelling the force-elongation curve of PEO by a Markovian two-level in equilibrium.

Quantity	Magnitude
n	108
l_{helical}	2.6 Å
l_{planar}	3.3 Å
$G_{\text{planar}} - G_{\text{helical}}$	8 $k_{\text{B}}T$
l_{K}	2.9 Å
K_{S}	62 N/m

solid at room temperature, depending on the degree of polymerization. Very high-purity PEO has been shown to be crystalline, with PEO_{16} having a melting point of 300 K [26]. It has been studied extensively experimentally [27, 28]. Of applications with high relevancy to current events it can be mentioned that PEO is used as an excipient (or "filler") in many pharmaceutical products, among them the Moderna and Pfizer–BioNTech vaccines for SARS-CoV-2 [29, 30]. As a side note, it can also be mentioned that PEO is famous for being a self-pouring liquid. This resembles the flying chain effect that is observed with a chain of beads in a jar, a phenomena that is still not completely understood [31, 32]. While the interpretation of this feature of PEO is not entirely without ambiguity, the analogy to PEO as weakly connected linear chains in a loosely coiled conformation is illustrative for the behaviour observed at ambient conditions with a high degree of polymerization [23].

Another interesting property of PEO is that it undergoes a conformational transition during stretching. For this reason, the force-elongation curve of PEO is typically neither well described by the worm-like chain model (WLC) or the freely jointed chain model (FJC). In the crystalline state PEO assumes a helical conformation, trans-trans-gauche (ttg) [33]. As the molecule is stretched, the equilibrium is shifted towards the elongated planar state, trans-trans-trans (ttt) [27]. This transition could also be induced by solvation effects, and it has been shown that these effects can dominate over backbone stretching energy and the conformational entropy of the polymer [34]. This explains experimental results where PEO can be well described by the WLC or WLC+ model [35].

To obtain the force-elongation curve for PEO in cases where the conformational transition is not induced by solvation effects, the molecule may be modeled as a Markovian two-level system in equilibrium. We will here present such a model by Oesterhelt, Rief and Gaub [27]. As all monomers retain either a planar of a helical

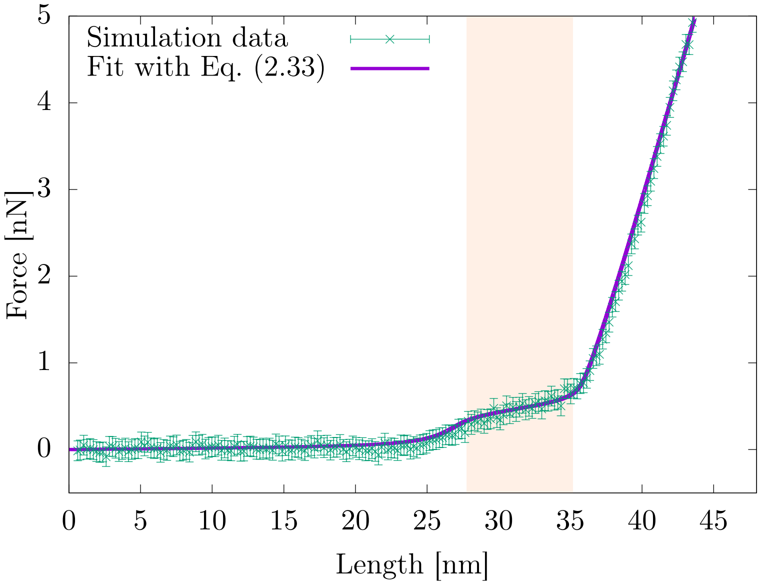


Figure 2.3: Force-elongation curve for a PEO molecule with $n = 108$ monomers, similar to the one shown in Figure 2.1. The orange background marks the range from $l = nl_{\text{helical}}$ to $l = nl_{\text{planar}}$. The curve is averaged over 20 samples at ambient temperature.

confirmation, the contour can be written

$$l_{\text{C}} = n_{\text{planar}} \cdot l_{\text{planar}} + n_{\text{helical}} \cdot l_{\text{helical}} \,. \tag{2.30}$$

The ratio of the populations $n_{\text{helical}} / n_{\text{planar}}$ is Boltzmann distributed,

$$\frac{n_{\text{helical}}}{n_{\text{planar}}} = \mathrm{e}^{+\Delta G / k_{\text{B}} T}, \tag{2.31}$$

where ΔG is the free energy difference between the states. The difference in free energy as a function of the applied force is given by

$$\Delta G(f) = (G_{\text{planar}} - G_{\text{helical}}) - f \cdot (l_{\text{planar}} - l_{\text{helical}}) \,, \tag{2.32}$$

which can be combined with the force-elongation curve for FJC+ in Table 2.1 to obtain[1]

$$\begin{aligned} l(f) =& n \left(\frac{l_{\text{planar}}}{\mathrm{e}^{+\Delta G / k_{\text{B}} T} + 1} + \frac{l_{\text{helical}}}{\mathrm{e}^{-\Delta G / k_{\text{B}} T} + 1} \right) \cdot \\ & \left(\coth \left(\frac{f \cdot l_{\text{K}}}{k_{\text{B}} T} \right) - \frac{k_{\text{B}} T}{f l_{\text{K}}} \right) + \frac{n f}{K_{\text{S}}}. \end{aligned} \tag{2.33}$$

---

[1] In the original article the signs of ΔG is switched, this is a typo.

In Figure 2.3, we show how this model compares to our simulation data of the system shown in Figure 2.1; a chain of PEO with $n = 108$ monomers. In the low force region, we determine the Kuhn length l_K by a linear fit with

$$l_\mathrm{low}(f) = nl_\mathrm{helical} f l_\mathrm{K}/(3k_\mathrm{B}T),$$

and in the high force region we determine the planar length l_planar and the elastic constant K_S by a linear fit with

$$l_\mathrm{high}(f) = nl_\mathrm{planar} + nf/K_\mathrm{S}.$$

All the parameters for Eq. (2.33) is shown in Table 2.2. The ttg segment length of $l_\mathrm{helical} = 2.6$ Å is tuned for best fit with the data, starting from the crystalligraphically measured length of 2.78 Å [33]. The ttt segment length of $l_\mathrm{planar} = 3.3$ Å can be compared to Oesterhelts value of 3.58 Å [27]. Both are off the expected values, but they are still considered to be reasonable close. The limiting cases of Eq. (2.30) is shown with an orange background in Figure 2.3, and illustrate the range from the completely unfolded helical state to the unstretched planar state. The Kuhn length of $l_\mathrm{K} = 2.9$ Å is lower than experimental results of 7.4 Å in a solvent of phosphate buffered saline [35], which may in part be explained by the solvent and the relatively small degree of polymerization.

The force field is similar to that of Article I, except for the addition of an implicit solvent, with `pair_style lj/charmm/coul/charmm/implicit` in LAMMPS, which includes an additional $1/r$ term in the Coulombic formula, serving as a distance-dependent dielectric term accounting for screening-effects. This corresponds to a good solvent, and limits chain-chain interactions. The chain is stretched with constant strain rate, with a velocity of 22 m/s, and the chains are constrained to move between two clamps, described by two moving Lennard-Jones walls 2 Å outside of the endpoints. The temperature is controlled by a Langevin thermostat to $T = 293$ K, and the data is averaged over 20 samples.

2.4 Cellulose

Cellulose is the most abundant polymer in the world [36]. Made out of linked sugars it is an important building block for plant life on earth. It is produced by plants, animals and single-celled organisms, and its popularity is increasing among a broad range of applications from new material development to cosmetics, as a food additive, for biomedical applications and in fabrics. It is composed of glucose units linked by $\beta(1 \rightarrow 4)$ bonds [37]. Each unit is rotated 180° to one another, and two glucose units are referred to as a cellobiose unit. The skeletal formula is show in Figure 2.4 with the Haworth projection.

One particular property of cellulose, is that that it degrades rather than melting at high temperatures. This degradations typically occur above 470 K[38]. In a

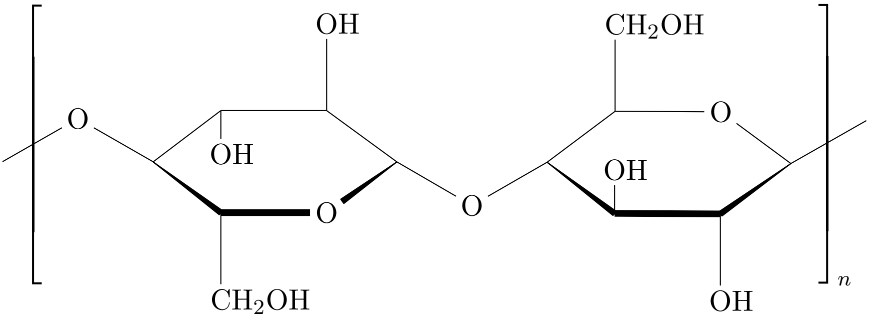


Figure 2.4: The skeletal formula of cellulose, composed of n units of cellobiose.

solvent, cellulose has been shown to undergo a glass transition at temperatures around 500 K [39], when extrapolated to dry conditions.

In nature, cellulose exist as microfibrils structures of sizes down of 36-18 chains [40, 41]. For many applications, particularly for packaging and fabric applications, it would be beneficial to have the cellulose as single chains. There is a great interest in understanding how cellulose might be dissolved. A bundle of cellulose is held together by hydrogen bonds in-plane and van der Waals bonds that act perpendicularly. The perhaps most dominant challenge in obtaining dissolution is that the cellulose bonds are both hydrophilic and hydrophobic [42, 43].

In the laboratory, the cellulose nanofibrils are typically isolated from wood-based fibres by mechanical treatment with high shear forces, with the optional chemical or enzymatic pretreatment step [44]. One common solvent for cellulose is composed of hydroxide and urea dissolved in water, which seems to be most effective at temperatures around 258 K [45]. The exact role of individual solvent constituents have been heavily investigated, but experimental and computational findings are not consistent between studies. One hypothesis, perhaps the most common, is that sodium and urea is thought to penetrate the bundle, where sodium then disrupts the interchain bonding and urea stabilize the dissolved chains to prevent re-agglomeration.

2.5 Applications for the articles

2.5.1 Article I

The first article is a computational study of tensile properties of PEO nanofibres, composed of linear chains of 33 monomers each. The geometry of the system is inspired by the fiber bundle model: the end-points of each chain are attached to stiff parallel planes, such that each chain is elongated by the same amount. The particles at the end-points of each chain are however allowed to move within these planes. The tensile loading was studied both with constant force and constant strain rate. For the constant force simulations, the load on the chains are equally distributed through the clamps, similar to the equal load sharing fiber bundle model. The simulations with constant strain rate focus on the interchain interactions, and do not include any redistribution of forces through the clamps. Going far beyond the complexity of the fiber bundle model, we aim to study polymeric fibres from a material science point of view and in this way probe the limitations of these models.

While the fiber bundle model exhibits quenched disorder, the disorder in these simulations originates in the chaotic potential energy landscape and in fluctuations from a thermal bath, to be discussed more later. The effect of configuration defects is also explored. The covalent bonds are modeled with Morse potentials, allowing for the breaking of bonds. While these computational experiments targets creep, their high energy and limited time scale more closely approach the mechanical breaking process.

Crystallinity plays an important role, as we have a transition from a disordered state to a semi-crystalline state during stretching. In experimental studies of PEO, the samples are highly amorphous, and the tensile properties are dominated by the interchain interactions between the loosely coiled chains. The semi-crystalline fibres in this computational study resemble more closely the structure of polyethylene, to which it display similar mechanical properties.

2.5.2 Article II-III

Continuing with the stretching of PEO nanofibres, the second and third article focus on single-molecule stretching with various degrees of polymerization at theta conditions. Here entropic elasticity is explored in addition do the enthalpic elasticity, and the system is compared to an idealized FJC model. The system is verified by studying the behaviour in the entropic region according to Eq. (2.23), which can be written as

$$f_{\mathrm{S}} = \frac{3k_{\mathrm{B}}T}{N_{\mathrm{eff}}b_{\mathrm{eff}}^{2}}l. \tag{2.34}$$

Note that equation 7 in article II is off by a factor of 6: this expression was obtained by using the radius of gyration rather than the end-to-end distance as the length in the expression for the entropy, shown in 2.19. As the molecule undergoes a transition from helical to planar conformations, the unfolded length relevant for the entropic regime is the length of the molecule before this transition. Accounting for this, one obtains a Kuhn length of 2 Å. In this article $N = 51$ denotes the number of beads in the chain, and the resulting effective number of beads for the FJC model is $N_{\mathrm{eff}} = 12$. I consider this to be reasonable for a PEO-molecule composed of $n = 16$ monomers.

2.5.3 Article IV

The fourth article has a more applied objective, and concerns the dissolution of cellulose. The aim of the work is to understand the specific details concerning the dissolution process of cellulose in a solution of water, sodium hydroxide and urea. As this is a highly complex system, it is not well described by idealized models.

3 Molecular Dynamics

"We adore chaos because we love to produce order"

M.C. Escher

Molecular dynamics (MD) is a incredibly versatile tool for computing the time evolution of systems. The complexity of the systems can be increased far beyond what can be described analytically. Similar to experiments in the laboratory, the outcome is prone to statistical errors. One also has to pay great attention into making sure that the simulation conditions are physical, and that the model is appropriate for what one is trying to achieve. Nonetheless, these brute force computer experiments have been shown to be highly useful tools for a wide range of applications. One attractive feature of molecular dynamics is that it permits direct visualization of the detailed motions of individual atoms in a system, thereby providing a window into the microscopic world. While the instantaneous movements of a singe trajectory might not be representative for the system, they can be useful as a guide toward understanding the mechanisms underlying a given process. Moreover, the average values have been shown to be reliable in most cases. We will here go through some basic principles of molecular dynamics, with an emphasis on properties that have been relevant for this work. For more details see [46].

3.1 Length and time scale

One of the first points to consider when preparing a computer simulation is the relevant length and time scale. With fully atomistic simulations one explicitly solves Newton's equations of motion for every single atom, and one typically work on the nano-scale in length and time. If one is interested in significantly larger scales, it is often a good idea to *coarse-grain*: one particle in our simulation can represent multiple atoms or whole groups of atoms. This reduction of degrees of freedom allows for considerably larger systems or significantly longer simulation times. The interpretation of time scales in coarse grained simulations is however a complicated topic of its own, especially relevant for dynamic processes.

For systems with long-range interactions, the computation of the resulting long-range forces is typically dominating for the computational cost. For a simulation where N particles interact with all other particles, a naive summation over pairs would give a running time of $\mathcal{O}(N^2)$. By making use of Fast Fourier Transforms (FFT), the time complexity is often reduced to $\mathcal{O}(N \log N)$. Some algorithms even obtain linear time algorithms, at the cost of limited accuracy. This will be discussed more in Section 3.5.

A common strategy to reduce the finite size effect of small systems is to make use of periodic boundaries in one or multiple dimensions, such that a particle interacts not only with the other particles in the system, but also replicas of these through the boundaries. As we will discuss more later, having periodic boundaries is even a requirement for using FFT. One should note that the computational cost of FFT-based methods scales with the volume of the simulation box.

When modeling systems such as polymers in a solvent, a common approach to speed up the simulations and increase the available time and length scales is to make use of implicit solvents. While there are different approaches to achieving it, they all aim to modify the non-bonded interaction potential to account for solvents without explicitly including the solvent-particles. This does not however account for the hydrophobic effect, viscosity or hydrogen bonding. If those effects are not relevant, implicit solvents can often be a good choice.

3.2 The initial configuration

In a MD simulation, positions and momenta of all particles at time t are functions of their initial values at $t = 0$:

$$\boldsymbol{r}_i(t) = f[\boldsymbol{r}^N(0), \boldsymbol{p}^N(0); t]. \tag{3.1}$$

However, one should be aware that any perturbation with amplitude ϵ gives rise to exponentially growing deviation

$$|\Delta \boldsymbol{r}_i(t)| \sim \epsilon \exp{(\lambda t)}, \tag{3.2}$$

known as a Lyapunov instability. Moreover, round off errors also causes deviations in the trajectory. While there is no reason to expect an MD simulation to exactly reproduce every detail of a physical system, one has observed that global properties such as total energy and response functions appear to be reliable. It is often safe to assume that

Let us assume for now that the trajectories given by MD are sufficiently realistic, given that the parameters for the interactions between the particles are correct.

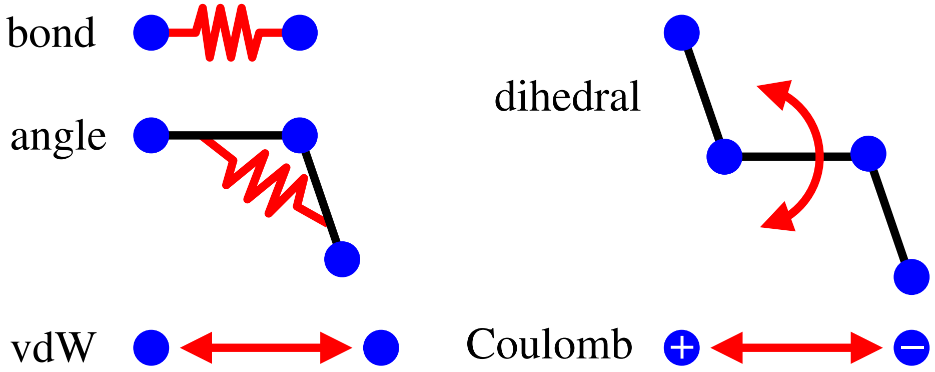


Figure 3.1: Illustration of typical terms in the potential energy surface for molecular simulations. The dihedral illustrates the potential for the torsions, these terms are often used interchangeably.

3.3 The potential energy surface

The interaction potential is an integral part of a MD simulation with classical force fields, and should be discussed in some detail. For a system made of N interacting atoms, we will have a set of coordinates and momenta $\{R_\alpha, \alpha = 1, \ldots, N\}, \{p_\alpha, \alpha = 1, \ldots, N\}$. The potential energy surface will then at all times be a function of the set of positions $U(\{R_\alpha\})$, and the forces on each atom will be given by

$$\boldsymbol{F}_\alpha = -\nabla_{R_\alpha} U(\{R_\alpha\}). \tag{3.3}$$

The potential energy will typically have contributions from bonded and non-bonded interactions,

$$U = U_\mathrm{b} + U_\mathrm{nb}, \tag{3.4}$$

and the the bonded energy is commonly given by the sum of two-, three-, and four-body terms from atoms joined by one, two and three consecutive covalent bonds, as illustrated in Figure 3.1. In the functional form of AMBER [47], the bonded interactions reads

$$U_\mathrm{b} = \sum_{\mathrm{bonds}} k_b(l - l_0)^2 + \sum_{\mathrm{angles}} k_a(\theta - \theta_0)^2 + \sum_{\mathrm{torsions}} \sum_n \frac{1}{2} V_n[1 + \cos(n\omega - \gamma)], \tag{3.5}$$

where k_b, k_a and V_n are suitable force constants. l_0, θ_0 and γ reflect the length, bending and dihedral angles of unstrained bonds. The non-bonded interactions consists of Van der Waals terms based on 6-12 potentials and electrostatic terms for Coulomb's law,

$$U_\mathrm{nb} = \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} f_{ij} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\}, \tag{3.6}$$

where q_i are atomic charges and σ_{ij} and ϵ_{ij} are length and energy constants. The coefficients f_{ij} are weighting coefficients for pairwise energy and force contribution between permanently bonded atoms. In AMBER, this coefficient is 0.0, 0.0 and 0.5 for atoms separated by 1, 2 and 3 bonds respectively, otherwise it is 1. Other families of force fields such as e.g. CHARMM, GROMOS and OPLS all have slightly different functional forms, and different conventions for e.g. the attribution of charges. In AMBER, each atom is attributed partial atomic charges through Restrained ElectroStatic Potential calculations [48].

3.3.1 A minimalistic reactive force field

The harmonic potential is only a reasonable approximation for bonds in proximity to equilibrium. We will here present a minimal reactive force field with a broader range of applicability that even allows for the breaking of covalent bonds.

We can replace the term for the stretching in Eq. (3.5) by a Morse potential of the form

$$U_{\mathrm{str}}(l) = D_{\mathrm{e}} \left[1 - e^{-\alpha(l-l_0)}\right]^2, \tag{3.7}$$

which saturates to a finite value at large separations. The required parameters for the dissociation energy D_{e} are obtained from density functional computations [49], and the parameters for α are found by requiring the Morse potential to have the same curvature as the harmonic bonds [50], i.e. $\alpha = (k_{\mathrm{b}}/2D_e)^{1/2}$.

Another advantage of the Morse potential is that it can account for the anharmonicity found in real molecules. With this potential, bonds shorter than $\ln(2)/\alpha + l_0$ give rise to stable anharmonic oscillations

$$E(v) = h\nu_0(v + 1/2) - \frac{[h\nu_0(v + 1/2)]^2}{4D_{\mathrm{e}}} \tag{3.8}$$

where v is the vibrational quantum level and

$$\nu_0 = \frac{\alpha}{2\pi}\sqrt{2D_{\mathrm{e}}/m_i}, \tag{3.9}$$

where m_i is the mass of particle i. When the initial configuration and the parameters for the potential energy surface is given, the position of the atoms may be updated by a suitable algorithm.

3.4 Verlet algorithms

The simplistic algorithm developed and used for the first MD simulations is still a good choice, upon suitable modifications. Its second order global error δt^2 makes it reasonably long-term stable, and it has a low computational cost in addition to being fairly intuitive. The algorithm can be derived by combining the forward evolution

$$\boldsymbol{r}_i(t+\delta t) = \boldsymbol{r}_i(t) + \delta t \boldsymbol{v}_i(t) + \frac{1}{2}\delta t^2 \frac{\boldsymbol{F}_i(t)}{m_i} \tag{3.10}$$

with the backward evolution

$$\boldsymbol{r}_i(t-\delta t) = \boldsymbol{r}_i(t) - \delta t \boldsymbol{v}_i(t) + \frac{1}{2}\delta t^2 \frac{\boldsymbol{F}_i(t)}{m_i}, \tag{3.11}$$

to obtain the expressions

$$\boldsymbol{r}_i(t+\delta t) = 2\boldsymbol{r}_i(t) - \boldsymbol{r}_i(t-\delta t) + \delta t^2 \frac{\boldsymbol{F}_i(t)}{m_i} \tag{3.12}$$

and

$$\boldsymbol{v}_i(t) = \frac{1}{2}\left[\boldsymbol{r}_i(t+\delta t) - \boldsymbol{r}_i(t-\delta t)\right]. \tag{3.13}$$

A potential downside of this leap-frog scheme is that no consistent estimation of the velocity at time t exists before the positions in the next time step is computed. When forces or other potentials depend on velocities, or when high accuracy with temperature and/or pressure coupling is needed we may use the predictor-corrector integrator velocity Verlet. We then advance time by δt

$$\boldsymbol{r}_i(t+\delta t) \leftarrow \boldsymbol{r}_i(t) + \delta t \boldsymbol{v}_i(t) + \frac{1}{2}\delta t^2 \boldsymbol{a}_i, \tag{3.14}$$

predict velocities at $t + \delta t$

$$\boldsymbol{v}_i(t+\delta t) \leftarrow \boldsymbol{r}_i(t) + \delta t \boldsymbol{a}_i(t), \tag{3.15}$$

compute forces $\boldsymbol{F}^N(t+\delta t)$ at coordinates $\{\boldsymbol{r}^N(t+\delta t)\}$, before we correct the velocities

$$\boldsymbol{v}_i(t+\delta t) \leftarrow \boldsymbol{v}_i(t+\delta t) + \frac{1}{2}\delta t[\boldsymbol{a}_i(t+\delta t) - \boldsymbol{a}_i(t)]. \tag{3.16}$$

The velocity Verlet is the default time integrator in popular software for MD such as LAMMPS [51], while popular alternatives such as GROMACS has leap-frog algorithms as the standard integrator with velocity Verlet as a possible option.

3.5 Long-range interactions

For the non-bonded interaction shown in Eq. (3.6) we saw that every particle in the system interacts with all the other particles, and it is not surprising that this is where MD calculations typically spend most of the computational time. For the dispersion forces, one can introduce a cutoff and a tail correction, however for the Coulombic terms this is not allowed: as the Coulomb potential decays as $1/r$, and the number of interacting particles increase with the the surface area of a sphere $4\pi r^2$, the tail correction to the potential energy clearly diverges. There are however other techniques that allow for significant reductions in the computational cost of the long-range interactions without losing the long-range character of the forces.

On such approach is *Ewald summation.* It is applicable when we have periodic boundary conditions. Each particle with point charge q_i is attributed a Gaussian screening cloud, with a total charge that exactly cancels q_i. To compensate for this, a smooth periodic charge distribution is introduced to account for the original charges and their images. These functions can be represented by a Fourier series which converge rapidly, and artifacts due to self-interaction can be corrected at the end.

The electrostatic potential due to the compensating charge distribution is found by Poisson's equation and computed in reciprocal space. The short-range electrostatic energy due to the screened point charges is then computed in real space, before the self-interaction is computed separately and subtracted from the contribution to the Coulomb energy. Ewald summation has a time complexity of $\mathcal{O}(N^{3/2})$.

The Particle-Particle Particle-Mesh (PPPM) method also splits the Coulomb potential is in two, this time with a switching function. Similar to the Ewald method, the short-range interactions is calculated directly from the particle-particle interactions. For the long range interactions, the charges are first distributed on a mesh. The Poisson's equation is then solved via a FFT technique, before the forces are calculated and assigned back to the particles in the system. As the computational cost of solving the Poisson equation drops significantly when the charges are distributed on a mesh, the PPPM method can obtain a time complexity of $\mathcal{O}(N \log N)$.

The Fast Multipole Method (FMM) also deserves a brief mention, as it promises nothing less than a linear time complexity, $\mathcal{O}(N)$. It has, however, been shown to introduce irregularities in the electrostatic potential and force [52], and for this reason it seems it is not much used in practice as of today.

3.6 Thermostats and barostats

The most common ensemble for MD is perhaps the NVE, or micro-canonical ensemble. However, it is also fairly common to run simulations in the canonical (NVT) ensemble or the isobaric (NPT) ensemble. While the thermodynamic potentials converge to the same value for all ensembles in the thermodynamic limit, the fluctuations in each ensemble will be different. As a consequence, the response functions are different in the different ensembles.

The most popular methods for obtaining canonical dynamics are the "extended phase space" approaches, where the physical phase space is supplemented with additional variables [53]. This method was first introduced by Andersen for constant pressure simulations [54], but the perhaps most common of these methods today is the one developed by Nosé and Hoover.

3.6.1 Nosé-Hoover

The Nosé-Hoover thermostat is a completely deterministic approach to obtain isothermal conditions in MD. An additional artificial time-scaling coordinate s is included that control the fluctuations in the kinetic energy, representing coupling to a heat bath. The variable s plays the role of a time-scaling parameter, and is associated with a quantity Q of units energy×time^2, commonly denoted as an effective "mass".

To avoid fluctuating time-intervals, during simulations, the variable s may be replaced with a thermodynamic "friction" variable ξ, which can be either positive or negative, obtaining equations of motions that can be expressed as [55]

$$\begin{aligned} \dot{\boldsymbol{r}}_i &= \frac{\boldsymbol{p}_i}{m_i} \\ \dot{\boldsymbol{p}}_i &= \boldsymbol{F}_i - \xi \boldsymbol{p}_i \\ \dot{\xi} &= \left(\frac{K}{K_0} - 1 \right) / \tau^2. \end{aligned} \tag{3.17}$$

Here τ is a phenomenological relaxation time in which the system is brought from a kinetic energy K to K_0 corresponding to the sought temperature. For $\tau \to \infty$ we regain the original equation of motion. We stress that this is a *non-Hamiltonian* system, however the equation of motions in Eq. (3.17) can also be derived from a Hamiltonian system using a non-canonical choice of variables [56].

Following along the same lines, the Nosé-Hoover thermostat can also be extended to control the pressure to give a Nosé-Hoover barostat, giving the equations of motion for the positions and momenta

$$\dot{\boldsymbol{r}}_i = \frac{\boldsymbol{p}_i}{m_i} + \frac{p_\epsilon}{W} \boldsymbol{r}_i \tag{3.18}$$

and

$$\boldsymbol{p}_i = \boldsymbol{F}_i - \left(1 + \frac{\mathrm{d}}{\mathrm{d}N} \right) \frac{p_\epsilon}{W} \boldsymbol{p}_i - \frac{p_{\xi_1}}{Q_1} \boldsymbol{p}_i. \tag{3.19}$$

Here N is the number of particles, ξ_1, p_{ξ_1} and Q_1 are the parameters for the thermostat, and ϵ, p_ϵ and W are the parameters for the barostatting. The parameter ϵ is determined by the volume of the system V and the initial volume $V(0)$ by

$$\epsilon = \ln(V/V(0)). \tag{3.20}$$

W in Eq. (3.19) is a mass parameter associated with ϵ, and p_ϵ is the momentum conjugate to ϵ. Note that this is the improved scheme of Martyna *et al.* [57], as the earlier scheme of Hoover only approximate the desired distribution [56, 58].

The Nosé-Hoover algorithm as presented here only generates the correct canonical distribution for systems with only one conserved quantity, or with no external forces and a fixed center of mass. This restriction is avoided by coupling the Nosé-Hoover thermostat to a chain of thermostats [57]. This addition of extra thermostats comes at a low computational cost; only the first thermostat interacts with N particles, the others form a simple one dimensional chain.

One should note that the Nosé-Hoover thermostat is global, and it is non-Galilean invariant; [59] the laws of motion are not the same in all inertial frames.

3.6.2 Langevin thermostat

The Langevin thermostat has a different approach to control the temperature. Here, we mimic the viscous aspect of a solvent to approximate the canonical ensemble. Random collisions between the liquid and the particles account for the heat dissipation by the Langevin equation:

$$\dot{\boldsymbol{r}}_i = \frac{\boldsymbol{p}_i}{m_i} \tag{3.21}$$

$$\dot{\boldsymbol{p}}_i = \boldsymbol{F}_i - \xi \frac{\boldsymbol{p}_i}{m_i} + \beta_i(t) \tag{3.22}$$

where $\boldsymbol{F}_i$ are the forces derived from the potential energy, ξ is the friction coefficient controlling the coupling between the particles and the heat bath, and β_i are random forces, given by a Gaussian probability distribution with correlation function

$$\langle \beta_i(t) \cdot \beta_j(t') \rangle = 6\xi k_B T \delta_{ij} \delta(t - t') \tag{3.23}$$

where T is the target temperature. The force is then uncorrelated with time. For large η we have overdamped Langevin dynamics or Brownian dynamics with no average acceleration. The thermostat can be included in a Verlet-type algorithm following the steps of Grønbech-Jensen and Farago [60].

The Langevin thermostat is also non-Galilean invariant. But in contrast to the Nosé-Hoover thermostat, the Langevin thermostat does not regulate the temperature uniformly. Moreover, it is also profile-unbiased [59], and has been shown to outperform the chained Nosé-Hoover thermostat in non-equilibrium simulations [61].

3.6.3 Berendsen

The Berendsen thermostat represents yet another approach to control the temperature during simulations. Here the system is weakly coupled to a constant temperature bath by adding an extra term in the equations of motion which effects the temperature change

$$\left.\frac{dT}{dt}\right|_{\text{bath}} = \frac{T_0 - T}{\tau_T}, \tag{3.24}$$

where T_0 is the reference temperature, T is the instantaneous temperature and τ_T is a relaxation time. Similarly for the Berendsen barostat, we have

$$\left.\frac{dP}{dt}\right|_{\text{bath}} = \frac{P_0 - P}{\tau_P}, \tag{3.25}$$

where again P_0 is the reference pressure, P is the instantaneous pressure and τ_p is a relaxation time. For every time step the system volume is then rescaled to accommodate Eq. (3.25) [62]. The main drawback of the Berendsen thermostat is perhaps that it does not sample the NVT ensemble [63]. Moreover, due to artifacts associated with velocity rescaling algorithms, the usage of the Berendsen thermostat and barostat is generally not recommended [64].

3.7 Ergodicity

Ergodicity is one of the core assumptions for the application of molecular-dynamics simulations: the notion that we can replace an average over all possible states in an ensemble with an average in time. For a given property $A[\boldsymbol{r}^N, \boldsymbol{p}^N]$, we can write this as

$$\langle A \rangle_{\text{ensemble}} = \lim_{t\to\infty} \frac{1}{t} \int\limits_0^t A[\boldsymbol{r}^N(t'), \boldsymbol{p}^N(t')]\mathrm{d}t'. \tag{3.26}$$

There are, however, a few scenarios where this assumption is not valid. Firstly, the simulations needs to be long enough to have the possibility of visiting a representative fraction of the phase space. Moreover, it will not hold for a system whose potential energy $U(\boldsymbol{r})$ possesses high barriers, as this leads to separatrices in the phase space. This requires the use of advanced sampling techniques such as rare events or umbrella sampling [65, 66].

3.8 Challenges

For MD simulations to be meaningful, it is crucial to have a realistic force field. Moreover, there are some inherent limitations of MD with atomic point charge models. The electron density is a dynamic continuum, and assigning fixed charges to a finite amount of particles or dummy-particles will never be able to fully reproduce the electromagnetic field. For this reason, water has shown to be excruciatingly difficult to model, and the melting temperatures in the most common models are completely off [67]. Polarizability can also be challenging to account for, though there are potentials that take this into account. In general, when quantum effects are dominating one should consider using density-functional approximations. The case of water, however, remains challenging, and most widely used approximations for the exchange-correlation functional does not describe the properties of pure water systems well [68].

3.9 Applications for the articles

Molecular-dynamics simulations are the main tool for the investigations described in all four articles in this thesis. The simulations were all performed in LAMMPS [51], with suitably-chosen force fields.

3.9.1 Article I

The simulations in the first article are preformed with a united atom force field, with potentials accounting for stretching, bending and torsion. The minimalistic reactive force field with Morse potentials from Section 3.3.1 was used to allow for breaking of covalent bonds, with breaking energy from density functional computations from the literature.

When tensile stress or strain is applied to the molecules, work is being done on the end-particles. For this reason, it is preferable to use a thermostat that acts locally, and the Langevin thermostat was the preferred choice.

One of the objectives of this study was to approach system sizes that would be available to experimental applications, and for this reason all boundaries were non-periodic.

In an attempt to sample a representative portion of phase space, the simulations started with a simulated annealing protocol to reach low-energy structures closer to experimental conditions; the samples were heated up to high temperatures before the temperatures was decreased slowly to ambient temperature. This protocol provided useful to overcome barriers in the potential energy landscape.

3.9.2 Articles II-III

The model for the second and third article is similar to the one in the first article, with the notable exception of the disabling of all non-bonded interactions. An annealing protocol similar to the one in the first article was employed together with a combination of relatively long simulations and many samples to sample a representative portion of phase space for different elongations of the molecule. The main focus of these articles is the analysis of the trajectories in thermodynamical framework.

3.9.3 Article IV

The cellulose molecules in the fourth article was modeled with a fully atomistic model, with explicit solvent molecules of sodium hydroxide, urea and water. A force field of the AMBER family was employed with a fully-atomistic model with explicit solvents. Long-range non-bonded interactions plays an important role in this system, and they are calculated with a PPPM solver. The simulations are run with Nosé-Hoover thermostats and barostats to control temperature and pressure.

4 Thermodynamics of stretching

> *"Thermodynamics is a funny subject. The first time you go through it, you don't understand it at all. The second time you go through it, you think you understand it, except for one or two small points. The third time you go through it, you know you don't understand it, but by that time you are so used to it, it doesn't bother you any more."* [69]
>
> ARNOLD SOMMERFELD

The laws of thermodynamics formalize the behaviour of energy conversion in nature. It is a framework of general applicability that opens up for drawing connections that would otherwise be unattainable. Thermodynamics was originally derived and apply to the macroscopic limit, but it can also be extended to deal with small systems [70]. Let us first consider the case of global equilibrium. The energy balance for the internal energy was first introduced by Clausius [3], whom related the internal energy U to the change in entropy S, temperature T, pressure p and change in volume V.

Converted to the case of polymer stretching, the change in internal energy at thermal equilibrium would be

$$\mathrm{d}U = T\mathrm{d}S + f\mathrm{d}l, \tag{4.1}$$

where f is the force pulling on the system and l is the end-to-end distance.

Upon reflection, the idea that systems consisting of $\sim 10^{23}$ particles can be characterized by this handful of parameters is nothing less than astonishing. This fundamental thermodynamic relation holds for both reversible and irreversible processes [71]. Gibbs later extended this to include the chemical potential μ_j of component j in a system with of n components. In the case of polymer stretching, this would be

$$\mathrm{d}U = T\mathrm{d}S + f\mathrm{d}l + \sum_{j=1}^{n} \mu_j \mathrm{d}N_j. \tag{4.2}$$

With this, we can account for change in internal energy due to a change in the number N_j of component j.

4.1 Hill's thermodynamics

While Eq. (4.2) is perfectly valid for macroscopic systems, it does not hold for small systems. The solution proposed by Hill was to introduce a large ensemble of $\mathcal{N}$ replicas of a small system [72]. System properties are then obtained by dividing the ensemble value by $\mathcal{N}$. The change in total internal energy of all replicas in the ensemble is then expressed by

$$\mathrm{d}U_{\mathrm{t}} = T\mathrm{d}S_{\mathrm{t}} + f\mathrm{d}l_{\mathrm{t}} + \sum_{j=1}^{n} \mu_j \mathrm{d}N_{j,\mathrm{t}} j + \epsilon \mathrm{d}\mathcal{N} \tag{4.3}$$

where the subscript t indicates an ensemble property, and ϵ is the partial replica energy, or the *subdivision potential*, as Hill named it. Any system with a non-zero subdivision potential is characterized as small, and should be approached accordingly. With this extension, one can take the effect of shape- and size-variation of small systems into account within the established framework of thermodynamics [73–76].

4.2 Non-equilibrium thermodynamics

Non-equilibrium thermodynamics allows us to describe also the *rate* of energy conversion between different states [77]. This branch of thermodynamics emerging with Onsager describe transport processes in systems that are out of global equilibrium [78, 79]. We still need local equilibrium, but this is a much softer criteria, and it has been shown to be a good assumption even for strong shock waves [80].

In non-equilibrium thermodynamics, the second law is expressed by the local entropy production in the system, σ. It is given by the sum of the product of the conjugate fluxes, J_i, and forces X_i, such that

$$\sigma = \sum_i J_i X_i \geq 0. \tag{4.4}$$

Each flux is a linear combination of all forces

$$J_i = \sum_i L_{ij} X_j, \tag{4.5}$$

where the reciprocal relations

$$L_{ji} = L_{ij} \tag{4.6}$$

apply. When the complete set of extensive variables, A_i, are identified, the conjugate fluxes are

$$J_i = \mathrm{d}A_i/\mathrm{d}t, \tag{4.7}$$

and the forces are

$$X_i = (\partial S/\partial A_i)_{A_{i\neq j}}. \tag{4.8}$$

In combination, Eqs. (4.4) to (4.8) contain all information on the non-equilibrium behaviour of the system [77].

4.3 The entropy production

When a molecule is stretched, one can control either the elongation of the molecule or the external force on the molecule. We will now derive the friction coefficient associated with stretching in both cases [77, 81]. We will make use of the fact that mean value variations of the internal energy in an ensemble of identical, independent replicas follow the same thermodynamic relations as for large systems.

4.3.1 The isometric ensemble

In the case in which the end-to-end distance l and the temperature T are externally controlled, variations in the average energy of the molecule is given by the Gibbs equation

$$\mathrm{d}\bar{U} = T\mathrm{d}S + \bar{f}\mathrm{d}l, \tag{4.9}$$

where $\bar{f}$ is the average force between the endpoints of the molecule. The changes in the average internal energy is given by the first law of thermodynamics as

$$\mathrm{d}\bar{U} = \mathrm{d}Q + \bar{f}_{\mathrm{ext}}\mathrm{d}l, \tag{4.10}$$

where $\bar{f}_{\mathrm{ext}}$ denotes the average external force on the terminals, and $\mathrm{d}Q$ is the heat delivered to the molecule. This change in heat should be equal to minus the heat exchange with the environment, $\mathrm{d}Q = -\mathrm{d}Q_0$. The change in entropy has contributions from the molecule $\mathrm{d}S$, and from the environment $\mathrm{d}S_0$. For a non-equilibrium process, the sum is the total entropy production S_{irr}. From the second law of thermodynamics we get

$$\frac{\mathrm{d}S_{\mathrm{irr}}}{\mathrm{d}t} \equiv \frac{\mathrm{d}S}{\mathrm{d}t} + \frac{\mathrm{d}S_0}{\mathrm{d}t} \geq 0, \tag{4.11}$$

where the equality only holds for completely reversible processes. Since $\mathrm{d}S_0 = \mathrm{d}Q_0/T$, we now obtain

$$\frac{\mathrm{d}S_{\mathrm{irr}}}{\mathrm{d}t} = \frac{1}{T}(\bar{f}_{\mathrm{ext}} - \bar{f})\frac{\mathrm{d}l}{\mathrm{d}t}. \tag{4.12}$$

We can then describe the system with a linear force flux relation via Eqs. (4.4) and (4.5). If we denote the velocity by $v \equiv \mathrm{d}l/\mathrm{d}t$ and the average change in the force by $\Delta\bar{f} \equiv \bar{f}_{\mathrm{ext}} - \bar{f}$, we can formulate a rate law for isometric stretching

$$\Delta\bar{f} = \xi_l(l)v, \tag{4.13}$$

where ξ_l=$\xi_l(l)$ is the friction coefficient associated with stretching specific for the length-controlled case. This is then a rate law for isometric stretching. Once $\xi_l(l)$ is known, the entropy production follows by Eq. (4.12) as $\mathrm{d}S/\mathrm{d}t = v^2\xi_l(l)/T$.

4.3.2 The isotensional ensemble

In the case in which the applied force f_{ext} and the temperature T are externally controlled, variations in the average energy of the molecule now follows the Gibbs equation

$$\mathrm{d}\bar{U} = T\mathrm{d}S + f\mathrm{d}\bar{l}, \tag{4.14}$$

where $\bar{l}$ indicates the average length. The first law now reads

$$\mathrm{d}\bar{U} = \mathrm{d}Q + f_{\text{ext}}\mathrm{d}\bar{l}, \tag{4.15}$$

and the total entropy production is now

$$\frac{\mathrm{d}S_{\text{irr}}}{\mathrm{d}t} = \frac{1}{T}(f_{\text{ext}} - f)\frac{\mathrm{d}\bar{l}}{\mathrm{d}t}. \tag{4.16}$$

With the change in the controlled force given by $\Delta f = f_{\text{ext}} - f$, and the resulting average stretching velocity $\bar{v} = \mathrm{d}\bar{l}/\mathrm{d}t$, the rate law in the force-controlled regime can be written

$$\Delta f = \xi_f(f)\bar{v}. \tag{4.17}$$

Here ξ_f=$\xi_f(f)$ is the friction coefficient at isotensional conditions. In the thermodynamic limit, the two friction coefficients are the same. Away from the limit, they may be different, as the rate laws depend on the set of the environmental control variables in use.

4.4 Thermodynamics of non-additive systems

The principle of *additivity* plays a key role in the thermodynamic framework. If one divides a system in different parts, the system is said to be additive if the interaction between the parts is negligible in comparison with the total energy. Systems consisting of a small number of particles and systems with long-range interactions are intrinsically non-additive [82]; the interaction energy between the different parts of the system cannot be neglected. The subdivision potential ϵ in Eq. (4.3) is then different from zero. The lack of additivity is often associated with concave regions known as curvature anomalies in the thermodynamic potentials.

For non-additive systems, the accessible equilibrium configurations depends on the choice of control parameters used to define its state [83]. This phenomena is called *ensemble inequivalence.*

Partition functions of the different ensembles are known to be related by Legendre transforms. For sufficiently long polymers, this is indeed true, as the usual Legendre transform

$$G\left(T, N, f(l)\right) = F(T, N, l) - f(l)l \tag{4.18}$$

is all we need to transform between the Gibbs energy $G\left(T, N, f(l)\right)$ and the Helmholtz energy $F(T, N, l)$ with respect to l at constant T and N. With the extension of Fenchel, this transformation can be written [84]

$$-G_{\mathrm{LF}}\left(T, N, f(l)\right) = \max_{l'}\left[f(l)l' - F(T, N, l')\right], \tag{4.19}$$

and $-G_{\mathrm{LF}}\left(T, N, f(l)\right)$ would then be the Legendre–Fenchel transform of $F(T, N, l)$. While it reduces to the form shown in Eq. (4.18) for functions that are both differentiable and convex, these are not requirements for its applicability. Thus we are also equipped to transform thermodynamic potentials that exhibit convexity anomalies. When these anomalies do occur, the thermodynamic potentials will not coincide with their convex envelope, leading to ensemble inequivalence [85].

Moreover, when the Helmholtz energy exhibits a curvature anomaly in the canonical ensemble, the states associated with this anomaly is jumped over by a first order phase transition in the isotensional ensemble [83].

4.5 Applications for the articles

Article II and III concern the thermodynamics of stretching of single molecules.

4.5.1 Article II

In the second article, a large amount of replicas of small systems were studied, and the entropy production associated with stretching is computed in the framework presented in Section 4.3.

4.5.2 Article III

The third article presents a general method for transforming the free energies of different ensembles for molecular stretching, namely the Legendre-Fenchel transformation discussed in Section 4.4. With this framework, the Helmholtz stretching energies from article II is transformed to Gibb energies.

5 Conclusions

The aim of this work was to explore the intermediate scale between single-molecule and bulk. A significant part of the time was spent investigating a united atom model of PEO with molecular-dynamics simulations. While the model is greatly simplified, it still displays a richness due to the complex energy landscape.

Fluctuations are important both on the bundle level and on the single chain level, though they play very different roles. On the bundle level, thermal fluctuations can trigger the breaking of bonds in the form of creep. The activation free energy depends on the strain and the temperature; at low strain rates, thermal fluctuations have more time to overcome the barrier, and at high temperatures the fluctuations are more energetic. This was shown to be well described by a Kramer's type approximation for the rate of chain breaking.

Single chains modeled with density functional approximations displayed a broad range of metastability during tensile stretching. With increasing number of chains this range narrowed down due to lateral chain-chain interactions which perturbed the linear arrangement of the chains.

The role of supramolecular structures was investigated systematically in bundles of 1-24 chains of 33 PEO monomers and one case of 100 chains. Stretching the chains between planar clamps produced geometries that were greatly simplified with respect to macroscopic polymers, with the intention of magnifying the role of lateral chain-chain interactions. Our study of the location of the first chain to break revealed that the outer layer of chains are significantly more likely to be the first chain to break than the other chains. A number of defects were studied in more detail; twisting around the elongation axis, non-optimal 2D isomers and 2D twinning of nanocrystals. At and above ambient temperature these defects form and disappear in a dynamical fashion, but the system retains a high degree of structure and a recognizable hexagonal structure, with ordering and tight packing of chains increasing with load. It is evident that the impact of the defects on the ultimate stress of the bundles are non-trivial.

If one were interested in reproducing the experimental tensile properties of PEO one would have to choose a different model, as the amorphous structure of entangled chains completely dominates the tensile properties. The Young's modulus and the tensile strength of amorphous samples are not determined by the stretching of covalent bonds, but rather by weaker chain-chain interactions. For the purpose of this work, the simplistic model has successfully allowed for a more isolated study of the lateral chain-chain interactions in polymers under tensile loading.

On the single chain level the relative size of the fluctuations is of even more significance than for a typical large system. By considering sets of replicas of the small systems according to the method of Hill, extensivity of the thermodynamic potentials was restored. For single molecules composed of 3,7 and 16 monomers of PEO, fluctuations in force and length were shown to give rise to size-dependent ensemble deviations, with the two smallest systems exhibiting convexity anomalies in the thermodynamic potentials. These occurred in a region where the molecule was undergoing torsional unfolding, and originate in local maxima in the potential of mean force accessible only in the length-controlled ensemble.

The difference in fluctuations of the two ensembles of small systems gave rise to a difference in the entropy production. With 7 monomers of PEO, the friction coefficient of isometric stretching is roughly twice the value of that of an ensemble with isotensional stretching. This difference decreased when the size of the system was increased to 16 monomers. This study shows how non-equilibrium properties are affected by the absence of the thermodynamic limit, and the method could be applied systematically to the study of irreversible processes that take place at small scales.

A method was described and numerically verified employing the Legendre–Fenchel transform to manage the curvature anomalies in the small systems of PEO. By this method, the isotensional Gibbs free energy are obtained from simulations in the isometric ensemble in such a way that the states characterized by this free energy are unique. This is of general applicability, and reduces to the usual Legendre transform for larger systems where the free energy is differentiable and convex. Removing these limitations opens up for wider applications of the established framework of thermodynamics.

Moving on to the more applied case of cellulose dissolution, we find that our simulations indicate that additional external forces are required to dissolve bundles of naturally occurring structures of cellulose. This is consistent with experimental observations. In general, the behaviour we observe are in line with current experimental opinion. We find that in the mixture solvent, the urea intercalates into the bundles, increasing their volume significantly. For smaller bundles, we find that sheets of cellulose are stable in the mixture solvent, but in pure water they immediately collapse into bundles. This is believed to be caused by hydrophobic interactions between cellulose chains. When we add periodic external forces to mimic stirring or agitation of the solvent-cellulose mixture, we see that the bundles are being dissolved. We find that the cellulose dissolves more readily in the mixture solvent than in pure water, which is consistent with experimental results. In the water solvent, the dissolution is incomplete. Under boundary conditions that are likely more representative of longer cellulose bundles, they tend to leave sheets of cellulose that are bound together with hydrophobic interactions. This is a further evidence of the stabilizing effect of the urea molecules. Qualitatively, our findings are consistent with the experimental results that the NaOH and urea solvent is able to completely dissolve cellulose under experimental conditions when water only partially dissolves cellulose, except in the case of small degree of polymerization. The different behaviour of water and solvent is likely related to cellulose-cellulose hydrophobic interactions, which appear to be more stable in pure water than in the mixture solvent, where urea is absorbed into the bundles.

References

[1] E. Bering.
"On the Stability of the Local Load Sharing Fiber Bundle Model".
MSc thesis. NTNU, 2016.

[2] G. Strobl. *The physics of polymers: Concepts for understanding their structures and behavior.* 2007. DOI: `10.1007/978-3-540-68411-4`. arXiv: `arXiv:1011.1669v3`.

[3] R. Clausius. "The Mechanical Theory of Heat". In:
London: Taylor and Francis, 1867.

[4] N. Roos. "Entropic forces in Brownian motion".
Am. J. Phys. **82** (2014), pp. 1161–1166. DOI: `10.1119/1.4894381`.

[5] R. J. Young and P. A. Lovell. *Introduction to polymers.* English.
Boca Raton, 2011.

[6] Q. Guo. *Polymer morphology: Principles, characterization, and processing.* 2016. DOI: `10.1002/9781118892756`.

[7] M. Doi and S. F. Edwards. *The theory of polymer dynamics.*
Clarendon Press, 1986, p. 391.

[8] P.-G. de Gennes. *Scaling concepts in polymer physics.*
Ithaca, New York: Cornell University Press, 1979.
DOI: `10.1002/actp.1981.010320517`.

[9] P. J. Flory. *Principles of Polymer Chemistry.*
Ithaca, New York: Cornell University Press, 1953.

[10] M. Rief, M. Gautel, et al. "Reversible unfolding of individual titin immunoglobulin domains by AFM." eng.
Sci. (New York, N.Y.) **276** (1997), pp. 1109–1112.
DOI: `10.1126/science.276.5315.1109`.

[11] S. B. Smith, Y. Cui, and C. Bustamante.
"Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules".
Science **271** (1996), pp. 795–799. DOI: `10.1126/science.271.5250.795`.

[12] M. Rubinstein. *Polymer Physics.* Oxford: Oxford University Press, 2003, p. 456.

[13] R. Petrosyan. "Improved approximations for some polymer extension models". *Rheol. Acta* **56** (2017), pp. 21–26. DOI: `10.1007/s00397-016-0977-9`.

[14] N. I. Abu-Lail and T. A. Camesano. "Polysaccharide properties probed with atomic force microscopy". *J. Microsc.* **212** (2003), pp. 217–238. DOI: `https://doi.org/10.1111/j.1365-2818.2003.01261.x`.

[15] S. Indrehus. "Exploring Biopolymers by the Fibre Bundle Model". MSc thesis. NTNU, 2017.

[16] S. N. Næss, A. Mikkelsen, et al. *Molecular Biophysics.* Trondheim, 2017.

[17] J. Bouchaud. "Econophysics: still fringe after 30 years?" *Europhys. News* **50** (2019), pp. 24–27. DOI: `10.1051/epn/2019103`.

[18] J. T. Kjellstadli. "Burst Distribution by Asymptotic Expansion in the Equal Load Sharing Fiber Bundle Model". *Front. Phys.* **7** (2019), pp. 1–10. DOI: `10.3389/fphy.2019.00201`.

[19] A. Hansen, P. C. Hemmer, and S. Pradhan. *The Fiber Bundle Model: Modeling Failure in Materials.* Weinheim: Wiley-VCH, 2015, pp. 1–236. DOI: `10.1002/9783527671960`.

[20] F. Kun, F. Raischel, et al. "Extensions of Fibre Bundle Models". In: *Model. Crit. Catastrophic Phenom. Geosci. A Stat. Phys. Approach.* Ed. by P. Bhattacharyya and B. K. Chakrabarti. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 57–92. DOI: `10.1007/3-540-35375-5_3`.

[21] G. Linga, P. Ballone, and A. Hansen. "Creep rupture of fiber bundles: A molecular dynamics investigation". *Phys. Rev. E* **92** (2015), p. 022405. DOI: `10.1103/PhysRevE.92.022405`.

[22] M. A. Meyers and K. K. Chawla. *Mechanical Behavior of Materials.* 2nd ed. Cambridge: Cambridge University Press, 2009, pp. 1–27.

[23] J. M. Harris. *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications.* New York: Ed. Plenum Press, 1992.

[24] R. Haag and F. Kratz. "Polymer Therapeutics: Concepts and Applications". *Angew. Chemie Int. Ed.* **45** (2006), pp. 1198–1215. DOI: `https://doi.org/10.1002/anie.200502113`.

[25] B. Ellis and R. Smith. *Polymers: A Property Database, Second Edition.* CRC Press, 2008.

[26] A. C. French, A. L. Thompson, and B. G. Davis. "High-purity discrete PEG-oligomer crystals allow structural insight". *Angew. Chemie - Int. Ed.* **48** (2009), pp. 1248–1252. DOI: `10.1002/anie.200804623`.

[27] F. Oesterhelt, M. Rief, and H. E. Gaub. "Single molecule force spectroscopy by AFM indicates helical structure of poly(ethylene-glycol) in water". *New J. Phys.* **1** (1999), pp. 6.1–6.11.

[28] L. M. Bellan, J. Kameoka, and H. G. Craighead. "Measurement of the Young's moduli of individual polyethylene oxide and glass nanofibres". *Nanotechnology* **16** (2005), pp. 1095–1099. DOI: `10.1088/0957-4484/16/8/017`.

[29] European Medicine Agency (EMA). *COVID-19 Vaccine Moderna.* 2021.

[30] Medicines and Healthcare products Regulatory Agency (MHRA). *Public Assessment Report Authorisation for Temporary Supply COVID-19 mRNA Vaccine BNT162b2 (BNT162b2 RNA) concentrate for solution for injection.* Tech. rep. Department of Health and Social Care (DHSC) Pfizer Limited & BioNTech Manufacturing GmbH, 2020.

[31] E. G. Flekkøy, M. Moura, and K. J. Måløy. "Mechanisms of the Flying Chain Fountain". *Front. Phys.* **6** (2018), p. 84. DOI: `10.3389/fphy.2018.00084`.

[32] E. G. Flekkøy, M. Moura, and K. J. Måløy. "Dynamics of the Fluctuating Flying Chain". *Front. Phys.* **7** (2019), p. 187. DOI: `10.3389/fphy.2019.00187`.

[33] H. Tadokoro. *Structure of Crystalline Polymers.* Vol. 84. 6. New York: John Wiley and Sons: New York, 1979, pp. 604–605. DOI: `https://doi.org/10.1002/bbpc.19800840617`.

[34] S. Liese, M. Gensler, et al. "Hydration Effects Turn a Highly Stretched Polymer from an Entropic into an Energetic Spring". *ACS Nano* **11** (2017), pp. 702–712. DOI: `10.1021/acsnano.6b07071`.

[35] F. Kienberger, V. P. Pastushenko, et al. "Static and Dynamical Properties of Single Poly(Ethylene Glycol) Molecules Investigated by Force Spectroscopy". *Single Mol.* **1** (2000), pp. 123–128. DOI: `10.1002/1438-5171(200006)1:2<123::AID-SIMO123>3.0.CO;2-3`.

[36] A. Dufresne. *Nanocellulose From Nature to High Performance Tailored Materials.* English. [Erscheinungsort nicht ermittelbar], 2012.

[37] A. D. French. "Glucose, not cellobiose, is the repeating unit of cellulose and why that is important". *Cellulose* **24** (2017), pp. 4605–4609. DOI: `10.1007/s10570-017-1450-3`.

[38] F. Shafizadeh and A. G. W. Bradbury. "Thermal degradation of cellulose in air and nitrogen at low temperatures". *J. Appl. Polym. Sci.* **23** (1979), pp. 1431–1442. DOI: `https://doi.org/10.1002/app.1979.070230513`.

[39] L. Szcześniak, A. Rachocki, and J. Tritt-Goc. "Glass transition temperature and thermal decomposition of cellulose powder". *Cellulose* **15** (2008), pp. 445–451. DOI: `10.1007/s10570-007-9192-2`.

[40] D. J. Cosgrove. "Re-constructing our models of cellulose and primary cell wall assembly". *Curr. Opin. Plant Biol.* **22** (2014), pp. 122–131. DOI: `10.1016/j.pbi.2014.11.001`.

[41] K. Heise, E. Kontturi, et al. "Nanocellulose: Recent Fundamental Advances and Emerging Biological and Biomimicking Applications". *Adv. Mater.* **33** (2021). DOI: `10.1002/adma.202004349`.

[42] C. Yamane, T. Aoyagi, et al. "Two different surface properties of regenerated cellulose due to structural anisotropy". *Polym. J.* **38** (2006), pp. 819–826. DOI: `10.1295/polymj.PJ2005187`.

[43] B. Lindman, G. Karlström, and L. Stigsson. "On the mechanism of dissolution of cellulose". *J. Mol. Liq.* **156** (2010), pp. 76–81. DOI: `10.1016/j.molliq.2010.04.016`.

[44] E. B. Heggset, K. Syverud, and K. Øyaas. "Novel pretreatment pathways for dissolution of lignocellulosic biomass based on ionic liquid and low temperature alkaline treatment". *Biomass and Bioenergy* **93** (2016), pp. 194–200. DOI: `10.1016/j.biombioe.2016.07.023`.

[45] J. Cai and L. Zhang. "Rapid dissolution of cellulose in LiOH/urea and NaOH/urea aqueous solutions". *Macromol. Biosci.* **5** (2005), pp. 539–548. DOI: `10.1002/mabi.200400222`.

[46] D. Frenkel and B. Smit. *Understanding Molecular Simulation.* 2002. DOI: `10.1016/B978-012267351-1/50006-7`. arXiv: `arXiv:1409.6842v1`.

[47] L. J. Yang, C. H. Tan, et al. "New-generation amber united-atom force field". *J. Phys. Chem. B* **110** (2006), pp. 13166–13176. DOI: `10.1021/jp060163v`.

[48] C. I. Bayly, P. Cieplak, et al. "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model". *J. Phys. Chem.* **97** (1993), pp. 10269–10280. DOI: `10.1021/j100142a004`.

[49] M. K. Beyer. "The mechanical strength of a covalent bond calculated by density functional theory". *J. Chem. Phys.* **112** (2000), pp. 7307–7312. DOI: `10.1063/1.481330`.

[50] C. Chen, P. Depa, et al. “A comparison of united atom, explicit atom, and coarse-grained simulation models for poly(ethylene oxide)”. *J. Chem. Phys.* **124** (2006), 234901(11). DOI: `10.1063/1.2204035`.

[51] S. Plimpton. “Fast parallel algorithms for short-range molecular dynamics”. *J. Comput. Phys.* **117** (1995), pp. 1–19. DOI: `10.1006/jcph.1995.1039`. arXiv: `nag.2347 [10.1002]`.

[52] D. S. Shamshirgar. “Fast methods for electrostatic calculations in molecular-dynamics simulations”. PhD thesis. 2018.

[53] M. Tuckerman. “Statistical Mechanics: Theory And Molecular Simulation” (2001).

[54] H. C. Andersen. “Molecular dynamics simulations at constant pressure and/or temperature”. *J. Chem. Phys.* **72** (1980), pp. 2384–2393. DOI: `10.1063/1.439486`.

[55] G. Linga. “The Breaking Process of an Idealized Polymeric Bundle under Applied Tensile Stress” (2014), p. 2014.

[56] M. E. Tuckerman and G. J. Martyna. “Understanding Modern Molecular Dynamics: Techniques and Applications”. *J. Phys. Chem. B* **104** (2000), pp. 159–178. DOI: `10.1021/jp992433y`.

[57] G. J. Martyna, D. J. Tobias, and M. L. Klein. “Constant pressure molecular dynamics algorithms”. *J. Chem. Phys.* **101** (1994), pp. 4177–4189. DOI: `10.1063/1.467468`.

[58] W. G. Hoover. “Constant-pressure equations of motion”. *Phys. Rev. A* **34** (1986), pp. 2499–2500. DOI: `10.1103/PhysRevA.34.2499`.

[59] J. Ruiz-Franco, L. Rovigatti, and E. Zaccarelli. “On the effect of the thermostat in non-equilibrium molecular-dynamics simulations”. *Eur. Phys. J. E* **41** (2018). DOI: `10.1140/epje/i2018-11689-4`. arXiv: `1806.06538`.

[60] N. Grønbech-Jensen and O. Farago. “A simple and effective verlet-type algorithm for simulating langevin dynamics”. *Mol. Phys.* **111** (2013), pp. 983–991. DOI: `10.1080/00268976.2012.760055`.

[61] Z. Li, S. Xiong, et al. “Influence of thermostatting on nonequilibrium molecular-dynamics simulations of heat conduction in solids”. *J. Chem. Phys.* **151** (2019). DOI: `10.1063/1.5132543`.

[62] H. J. Berendsen, J. P. Postma, et al. “Molecular dynamics with coupling to an external bath”. *J. Chem. Phys.* **81** (1984), pp. 3684–3690. DOI: `10.1063/1.448118`.

[63] G. Bussi, D. Donadio, and M. Parrinello.
"Canonical sampling through velocity rescaling".
J. Chem. Phys. **126** (2007). DOI: `10.1063/1.2408420`. arXiv: `0803.4060`.

[64] E. Braun, S. M. Moosavi, and B. Smit. "Anomalous Effects of Velocity Rescaling Algorithms: The Flying Ice Cube Effect Revisited".
J. Chem. Theory Comput. **14** (2018), pp. 5262–5272.
DOI: `10.1021/acs.jctc.8b00446`. arXiv: `1805.02295`.

[65] T. S. Van Erp. "Dynamical Rare Event Simulation Techniques for Equilibrium and Nonequilibrium Systems". In:
Kinet. Thermodyn. Multistep Nucleation Self-Assembly Nanoscale Mater.
John Wiley & Sons, Ltd, 2012. Chap. 2, pp. 27–60.
DOI: `https://doi.org/10.1002/9781118309513.ch2`.

[66] G. M. Torrie and J. P. Valleau. "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling".
J. Comput. Phys. **23** (1977), pp. 187–199.
DOI: `https://doi.org/10.1016/0021-9991(77)90121-8`.

[67] C. Vega, E. Sanz, and J. L. Abascal.
"The melting temperature of the most common models of water".
J. Chem. Phys. **122** (2005), pp. 1–9. DOI: `10.1063/1.1862245`.

[68] M. J. Gillan, D. Alfè, and A. Michaelides.
"Perspective: How good is DFT for water?" *J. Chem. Phys.* **144** (2016).
DOI: `10.1063/1.4944633`. arXiv: `1603.01990`.

[69] K. J. Laidler and J. H. Meiser. *Physical chemistry.* English.
Menlo Park, California: Benjamin/Cummings Pub. Co., 1982.

[70] D. Bedeaux, S. Kjelstrup, and S. K. Schnell.
Nanothermodynamics – General Theory. 1st ed.
Trondheim: PoreLab, 2020.

[71] F. Mandl. *Statistical Physics.* New York: John Wiley & Sons, Ltd, 1988.

[72] T. L. Hill. *Thermodynamics of Small Systems.*
New York: W.A. Benjamin Inc, 1963, p. 210.

[73] B. A. Strøm, J. M. Simon, et al. "Size and shape effects on the thermodynamic properties of nanoscale volumes of water".
Phys. Chem. Chem. Phys. **19** (2017), pp. 9016–9027.
DOI: `10.1039/c7cp00874k`.

[74] M. T. Rauter, O. Galteland, et al.
"Two-phase Equilibrium Conditions in Nanopores".
Nanomaterials (2020), pp. 1–18. DOI: `10.3390/nano10040608`.

[75] M. Erdős, O. Galteland, et al. "Gibbs Ensemble Monte Carlo Simulation of Fluids in Confinement: Relation between the Differential and Integral Pressures". *Nanomaterials* **10** (2020). DOI: `10.3390/nano10020293`.

[76] V. Bråten, Ø. Wilhelmsen, and S. K. Schnell. "Chemical Potential Differences in the Macroscopic Limit from Fluctuations in Small Systems". *J. Chem. Inf. Model.* (2021). DOI: `10.1021/acs.jcim.0c01367`.

[77] S. Kjelstrup and D. Bedeaux. *Non-Equilibrium Thermodynamics of Heterogeneous Systems.* Singapore: World Scientific, 2008. DOI: `10.1142/6672`.

[78] L. Onsager. "Reciprocal Relations in Irreversible Processes. I." *Phys. Rev.* **37** (1931), pp. 405–426. DOI: `10.1103/PhysRev.37.405`.

[79] L. Onsager. "Reciprocal Relations in Irreversible Processes. II." *Phys. Rev.* **38** (1931), pp. 2265–2276. DOI: `10.1103/PhysRev.38.2265`.

[80] B. Hafskjold, D. Bedeaux, et al. "Nonequilibrium thermodynamics of surfaces captures the energy conversions in a shock wave". *Chem. Phys. Lett. X* **7** (2020), p. 100054. DOI: `10.1016/j.cpletx.2020.100054`.

[81] J. M. Rubi, D. Bedeaux, and S. Kjelstrup. "Thermodynamics for Single-Molecule Stretching Experiments". *J. Phys. Chem. B* **110** (2006), pp. 12733–12737.

[82] I. Latella, A. Pérez-Madrid, et al. "Thermodynamics of Nonadditive Systems". *Phys. Rev. Lett.* **114** (2015), p. 230601. DOI: `10.1103/PhysRevLett.114.230601`.

[83] A. Campa, L. Casetti, et al. "Concavity, Response Functions and Replica Energy". *Entropy* **20** (2018). DOI: `10.3390/e20120907`.

[84] T. Rockafellar. *Convex Analysis.* Princeton, New Jersey: Princeton University Press, 1972.

[85] M. Süzen, M. Sega, and C. Holm. "Ensemble inequivalence in single-molecule experiments". *Phys. Rev. E* **79** (2009), p. 51118. DOI: `10.1103/PhysRevE.79.051118`.

[86] G. Liu, W. Li, et al. "Molecular dynamics studies on the aggregating behaviors of cellulose molecules in NaOH/urea aqueous solution". *Colloids Surfaces A Physicochem. Eng. Asp.* **594** (2020), p. 124663. DOI: `10.1016/j.colsurfa.2020.124663`.

Articles

Article I

E. Bering and A. S. de Wijn

Stretching and breaking of PEO nanofibres. A classical force field and ab initio simulation study,

Soft Matter **16**:11 2736-2752 (2020).
DOI: 10.1039/D0SM00089B

Article I

Article I

Soft Matter

PAPER

Cite this: *Soft Matter*, 2020, **16**, 2736

Received 14th January 2020, Accepted 10th February 2020

DOI: 10.1039/d0sm00089b

rsc.li/soft-matter-journal

Stretching and breaking of PEO nanofibres. A classical force field and *ab initio* simulation study

Eivind Bering *[a] and Astrid S. de Wijn [b]

The burgeoning development of nanotechnology is allowing us to construct more and more nano-scale systems in the real world that used to only exist in computer simulations. Among them, nanofibres made of only a few aligned polymeric chains in particular might soon have important roles in nanofabrications as well as in nanomedicine. In this work, we present a broad exploration by computer simulations of elastic and inelastic properties of polyethylene-oxide (PEO) nanofibres under load. We cover the full range from unloaded fibres up to their breaking point, focusing on all features that arise from chain–chain interactions and collective behaviour of the chains. We employ both molecular dynamics (MD) simulations and density functional theory (DF). The classical force field is represented by a minimal reactive force field model, allowing for the breaking of covalent bonds. Density functional (DF) computations provide a benchmark to gauge and validate the empirical force field approach, and offer an intriguing view of the bundle chemical evolution after breaking. Force-field based MD is employed for the systematic investigation of bundles of up to 24 chains, and for a single bundle of 100 chains. Low-temperature results for bundles under moderate loading provide a size-dependent sequence of cross-sections, structures, cohesive energies and elastic properties. A remarkably high Young's modulus on the order of 100 GPa was estimated with DF and MD, explained by the semi-crystalline state of the fibres giving mechanical properties comparable to those of carbon nanotubes and of graphene. Breaking is investigated by simulations with constant strain rate or constant stress. The bundle breaks whenever the potential energy is raised above its metastability range, but also below that limit due to creep activated by thermal fluctuations. A Kramer's-type approximation for the rate of chain breaking is proposed and compared to simulation data.

1 Introduction

Polymers underlie a vast variety of industrial applications.[1] They also represent an essential ingredient of life, since proteins, nucleic acids and polysaccharides consist of polymeric molecules.[2] The mechanical properties of polymers play an important role in their function, in both the man-made and natural context.[3] Experimental developments in nanotechnology and biophysics as well as the overwhelming growth of computer power are making it feasible to investigate the mechanical properties of polymer fibres down to the molecular and atomistic scale.

Experimental measurements at the nano scale have been focused very heavily on the single molecule limit, using vibrational spectroscopy, atomic force microscopy and optical tweezers to determine the strain–stress relation in organic polymers and in biopolymers.[4,5] Measurements provide elastic properties and limiting resistance that can be compared with state of the art *ab initio* computations,[6] but also highlight the difference with the same properties measured for macroscopic samples of the same material. In between the single molecule and the bulk are different structures and length scales that play a big role in this difference. This range of multiple chains together evokes the image of rope or yarn, which is often plied, consisting of macroscopic fibres twisted together to increase the strength and integrity of the final material. Similarly, the interaction and collective behaviour of the molecules inside polymer bundles play a crucial role in the mechanical properties of the nanometric bundle. While these intermediate length scales have not been studied nearly as much as the single-molecule case, they are becoming experimentally accessible as well. Advanced fabrication techniques such as electrospinning allow the routine preparation of fibres of sub-μm length and diameter, reaching down to the 10 nm range.[7] This size still corresponds to hundreds of polymer chains, leaving out a relatively wide range that is currently accessible only to self-assembly, that, however, is available only for selected, mainly biological, polymers.

From a theoretical point of view, bundles in this size range are complex and therefore challenging. Important roles are

---

[a] *NTNU, Department of Physics, PoreLab, S.P. Andersens vei 15 B, 7031 Trondheim, Norway. E-mail: eivind.bering@ntnu.no*

[b] *NTNU, Department of Mechanical and Industrial Engineering, Richard Birkelands vei 2B, 7034 Trondheim, Norway. E-mail: astrid.dewijn@ntnu.no*

This journal is © The Royal Society of Chemistry 2020

played by chain–chain interactions, surface effects, entropy, nonlinearities, and thermal fluctuations. At the same time, the systems are so large as to be computationally expensive, yet too far from the thermodynamic limit to allow for standard approaches from statistical physics and thermodynamics. This regime of nano-scale bundles of small numbers of molecules (as opposed to single molecules[8] or bulk[9,10]) has not been investigated much theoretically.

In our computational study, we consider nanofibres of a simple paradigmatic polymer, *i.e.*, polyethylene oxide (PEO), made of nanometric bundles, stretched between two rigid, planar clamps. PEO (also referred to as polyethylene glycol or PEG) has a wide range of applications,[11] and typing "polyethylene glycol" in an internet search engine in October 2019 produces about six million hits. It has been studied extensively experimentally.[5,12] In our simulations, we focus on the elastic and inelastic response of these nanofibres to tensile load, covering the creep regime and up to their limiting resistance and chemistry of breaking.

Basic bonding and elastic properties of very thin samples are analysed by density-functional (DF) simulations. This approach provides quantitative and predictive information on the interplay of different deformation modes such as torsion, bending and stretching along the chain, on the role of electron orbitals and covalent bonds in determining the chain breaking mechanism, deformation under stretching, and the re-bonding of chains after breaking.

We study somewhat larger bundles using a reactive classical force field, stretching them until they break. The stretched bundles exhibit a complex structure in the lateral direction, showing solid-like structures. In addition, we show that there are more complex structural effects in 3D, related to the helicity of the chains. The mechanical properties depend nontrivially on the number of chains in the bundle. Finally, we also investigate the thermal effects on the structure and breaking of bundles. We show, among other things, that imperfections in the structure impact on the mechanical properties. The combination of chemical detail and accuracy provided by DF with the computational efficiency, and rigorous statistical mechanics framework provided by MD simulation offers a comprehensive view of a complex phenomenon such as creep in polymers.[13,14]

2 Method

We perform molecular dynamics (MD) simulations of stretching of a nanofibre of poly-ethylene oxide (PEO) composed of N molecular chains on the form CH_3–[O–CH_2–CH_2]$_n$–O–CH_3 with $n = 33$.

In our study, a tensile load or strain is applied to the PEO fibre through planar clamps. These clamps consist of a geometrical constraint on one of the coordinates, here z, of the terminations. The position of the carbon atoms belonging to the two terminal methyl groups are constrained to lay in the plane defining the clamps. The junction of chain and clamp can move along the plane of the clamp, hence tension combines with inter-chain cohesion to mimic lateral compression of the fibre. The simulations are carried out using the free and open-source LAMMPS Molecular Dynamics Simulator.[15] A sketch of this model is shown in Fig. 1.

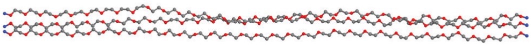


Fig. 1 A snapshot of a bundle with three chains.

2.1 Classical simulations

The polymeric PEO chains are described with a united atom model where each carbon is grouped with its bonded hydrogen atoms to form a united atom. It has been shown that this united-atom representation provides results in reasonable agreement with available experimental data at lower computational cost.[16] The basic force field we use has a functional form that resembles the OPLS model,[17] with some important distinctions. The potential energy surface is dividend into bonded and non-bonded parts,

$$U(\{\mathbf{R}_i,\ i = 1, N\}) = U_{\text{bond}}(\{\mathbf{R}_i,\ i = 1, N\}) + U_{\text{non-bond}}(\{\mathbf{R}_i,\ i = 1, N\}), \tag{1}$$

where $\mathbf{R}_i$ is the position of the i-th atom. The non-bonded part accounts for Coulomb interaction, short range repulsion, and dispersion interactions. The latter two are described by a Lennard-Jones pair potential, so that

$$U_{\text{non-bond}}(\{\mathbf{R}_i,\ i = 1, N\}) = U_{\text{C}}(\{\mathbf{R}_i\}) + U_{\text{LJ}}(\{\mathbf{R}_i\}), \tag{2}$$

where

$$U_{\text{C}}(\{\mathbf{R}_i\}) = k_{\text{e}} {\sum_{i\neq j}}' \frac{q_i q_j}{|\mathbf{R}_i - \mathbf{R}_j|} \tag{3}$$

and

$$U_{\text{LJ}}(\{\mathbf{R}_i\}) = 4 {\sum_{i\neq j}}' \varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{|\mathbf{R}_i - \mathbf{R}_j|}\right)^{12} - \left(\frac{\sigma_{ij}}{|\mathbf{R}_i - \mathbf{R}_j|}\right)^{6}\right], \tag{4}$$

where k_{e} is the Coulomb constant, $\{q_i\}$ are atomic charges, σ_{ij} and ε_{ij} are the length and energy scales of the LJ potential. The prime on each sum indicates that pairs of atoms separated by one, two, or three consecutive bonds are excluded. This is the typical CHARMM convention.[18] The interaction energy between two dissimilar non-bonded atoms is estimated by the Lorentz–Berthelot combination rules, *i.e.* $\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2}$ and $\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}$.[19]

The bonded interaction consist of stretching, bending and torsional contributions

$$U_{\text{bond}}(\{\mathbf{R}_i,\ i = 1, N\}) = U_{\text{str}}(\{\mathbf{R}_i\}) + U_{\text{bend}}(\{\mathbf{R}_i\}) + U_{\text{tors}}(\{\mathbf{R}_i\}). \tag{5}$$

where each term is given by a quadratic function of the deviation of bond lengths and distances from their equilibrium value.

While a conventional quadratic bond stretching potential provides a fair description of the low energy portion of the system potential energy surface, it does not account for the anharmonicity of real bonds under high stress, and in particular it does not allow for bond breaking, which is what we are concerned with in this work. In our minimal reactive force field, we therefore replace the conventional quadratic form of standard bond stretching potential with a Morse potential

$$U_{\text{str}}(\{\mathbf{R}_i,\ i = 1,\ N\}) = D_{ij}[1 - \mathrm{e}^{-\alpha_{ij}(r_{ij}-\bar{r}_{ij})}]^2, \tag{6}$$

which saturates to a finite value at large separations. Here D_{ij} is the dissociation energy, $\bar{r}_{ij}$ is the equilibrium bond distance and α_{ij} gives the width of the potential. If the bond length reaches a certain cut-off distance the bond is permanently removed. For these simulations a cut-off distance of 4 Å was used. For the C–C bond, this effectively reduces the dissociation energy by 2%. As is conventional, we do not take into account the zero point energy of the vibrational levels of around 4 kJ mol^{-1}. The required parameters for the dissociation energy D_{ij} are obtained from density functional computations of the same bond breaking,[20] and the parameters for α_{ij} are found by requiring the Morse potential to have the same curvature in the minimum as the harmonic bond, *i.e.* $\alpha_{ij} = \sqrt{K^{\text{s}}_{ij}/2D_{ij}}$, where K^{s}_{ij} is the force constant in the harmonic potential

$$U_{\text{harmonic}}(\{\mathbf{R}_i\}) = \frac{1}{2}\sum_{\{ij\}} K^{\text{s}}_{ij}\left[r_{ij} - \bar{r}_{ij}\right]^2. \tag{7}$$

The stress on end-bonds joining chains to the clamps is enhanced by inertia effects. To avoid their preferential breaking, these bonds are modeled with the harmonic potential. Our model covers only the C–C and C–O bond breaking events of interest for our study. Bond-order force fields such as ReaxFF[21] could provide a more comprehensive view. However, this simple reactive force-field suffices for our purposes.

The potentials for the bending and torsion read

$$U_{\text{bend}}(\{\mathbf{R}_i\}) = \frac{1}{2}\sum_{\{ijk\}} K^{\text{b}}_{ijk}\left[\theta_{ijk} - \bar{\theta}_{ijk}\right]^2 \tag{8}$$

and

$$U_{\text{tors}}(\{\mathbf{R}_i\}) = \sum_{\{ijkl\}}\sum_{\{c\}} K^{\text{t},c}_{ijkl}\left[\cos\left(\phi_{ijkl}\right)\right]^{c-1}, \tag{9}$$

where i, j, k and l are atoms joined by consecutive covalent bonds, K^{b}_{ijk} and K^{t}_{ijkl} are force constants of bending and torsion energy contributions and $\bar{\theta}_{ijk}$ are equilibrium angles. These are chosen to reproduce molecular properties measured by spectroscopy or computed by *ab initio* methods. Note that the sum over the torsional coefficients includes every possible dihedral. Moreover, bending and torsion terms are not removed whenever a bond breaks. This introduces a slight artifact into the potential energy surface, whose elimination would require recoding the computation of the angular part of energy and forces.

This united atom force field parameterization is taken from van Zon *et al.*,[22] based on a modification of the explicit atom force field of Neyertz *et al.*[23] The set of parameters used in our simulations is reported in Tables 1 and 2. The time step used in the simulations is 1 fs.

The initial configuration of the bundles is generated by placing the end particles randomly within circular cross sections of area of 16 Å^2 per chain, and the rest of the beads in chain are placed at equidistant spacing in the *z*-direction with a random component in the *xy*-plane.

Table 1 Force field parameters for the stretching, bending and for the non-bonded interaction[16] with disassociation energies[20]

Bonds	K^{s}_{ij} [kJ (mol Å^2)$^{-1}$]	D_{ij} [kJ mol^{-1}]	$\bar{r}_{ij}$ [Å]
C–C	2587.4	370.8	1.54
C–O	3094.0	344.5	1.43
Bends	K^{b}_{ijk} [kJ mol^{-1}]	$\bar{\theta}_{ijk}$ [Å]	
O–C–C	727.7	110.0	
C–O–C	1070.1	112.0	
LJ-interaction	σ_{ii} [Å]	ε_{ii} [kJ mol^{-1}]	q_i [q_{e}]
CH$_3$	3.699	1.047	0.174
CH$_2$	3.624	0.831	0.174
O	3.034	0.401	−0.348

Table 2 Force field parameters for the torsion[16]

Torsion [kJ mol^{-1}]	$K^{\text{t},0}_{ijk}$	$K^{\text{t},1}_{ijk}$	$K^{\text{t},2}_{ijk}$	$K^{\text{t},3}_{ijk}$	$K^{\text{t},4}_{ijk}$	$K^{\text{t},5}_{ijk}$	$K^{\text{t},6}_{ijk}$
O–C–C–O	2.211	15.194	17.844	−32.460	−13.871	−1.189	12.322
C–C–C–O	5.183	5.610	6.272	−15.428	−0.678	−4.568	3.567

From the beginning of the simulation the fibers are stretched between the two clamps. This is somewhat artificial, since polymer chains at low strain tend to adopt a (nearly)-Gaussian coil configuration, however it might correspond well to stretching experiments using tweezers or AFM.

From these initial configurations the systems are subjected to simulated annealing to reach low-energy structures closer to equilibrium and what one would expect for experimental samples.[24] During this process the Morse potential for the bond stretching is temporarily replaced by the harmonic potential, and the samples are heated up to about 1000 K before the temperature is gradually decreased to the desired temperature during 1 ns while keeping the force on the ends fixed at 1 nN per chain.

When the molecules are being stretched out, work is done and energy is added to the system through the end-particles. For this reason, it is necessary to use a thermostat that acts locally, and does not regulate the temperature uniformly. The Langevin thermostat is thus the preferred choice rather than the more commonly used Nosé–Hoover. The parameter for the relaxation time in the Langevin thermostat was set to 1 ps.

We stretch the bundles either with constant strain rate, or constant force, corresponding to the experimental conditions of constant strain and constant stress measurements. In the constant strain rate simulations, we increase the separation between the clamps at a constant velocity. In the constant force simulations, the position of the clamps is adjusted to keep the force on the fibre constant at a pre-assigned value. In the simulations with constant strain rate, the force that was applied to a chain before it breaks does not get redistributed over the other chains. This allows us to isolate the interaction between the chains from the interaction mediated by the terminals, which cloud these effects in a system with more realistic boundary conditions.

The constant stress condition, on the other hand, represents a molecular dynamics realisation of the well known fibre bundle model (FBM) with global redistribution of load upon chain breaking.[25] The two modes of operation with constant stress and constant strain rate have been carried out with the LAMMPS options aveforce and move respectively. In one particular sample, torque was added on the atoms on the terminations in the lateral direction with the option addtorque. The molecular dynamics simulations have been carried out with finite boundary conditions in all directions.

For the analysis of trajectories, chains are identified as neighbours if the distance between the endpoints on the terminals are less than 1.5 times the estimated bulk separation of the chains, estimated from simulations with 100 chains. Chains with less than six neighbours are counted as part of the contour. Both criteria were validated by visual inspection of the cross sections for a number of bundles.

2.2 *Ab initio* simulations

Qualitative and quantitative aspects of the breaking of PEO chains at $T = 0$ K have been investigated by density functional (DF) computations. We used the plane wave-pseudopotential formulation of DF theory implemented in CPMD *ab initio* simulation package,[26,27] with the exchange–correlation energy given by the generalised gradient approximation of Perdew, Burke and Ernzerhof (PBE).[28]

The system is enclosed in an orthorhombic simulation cell with periodic boundary conditions. The plane wave basis of reciprocal lattice vectors is included up to a kinetic energy cut-off of 120 Ry, with the sampling of the Brillouin zone limited to the Γ-point. This last approximation is justified by the large size of the simulation box, and by the insulating character of the material. Only valence states are included in the computation, and the core valence electron interaction is represented by *ab initio* norm conserving pseudopotentials of the Troullier–Martins type.[29] Dispersion (van der Waals) interactions are essential to describe lateral chain–chain cohesion. In the present study they are accounted for using the semi-empirical approach by Grimme.[30]

In all samples, chains extended along the entire length L_z of the longest side of the periodic orthorhombic simulation cell to mimic an infinitely long polymeric fibre. The periodicity along x and y at $L_x = L_y = 14.4$ Å has been set to keep the lateral interaction of periodic replicas low. Stretching is imposed by increasing L_z beyond its initial value of $\sim$3.2 Å per monomer, with the $T = 0$ K condition enforced by minimizing the potential energy by means of quenched MD. Additional DF computations have been carried out on crystalline PEO starting from the experimental structure and unit cell of ref. 31.

Most computations have been carried out in the spin-compensated picture. Since this might break down in proximity of the bond breaking, a few test computations have been carried out considering unconstrained spin-orbitals, including the possibility of bare spin polarisation in an open shell electronic structure. These tests did not provide any evidence of local spin-uncompensated domain, hence the results reported in the following section all refer to spin compensated computations.

3 Results

3.1 Chemistry of bundle breaking and *ab initio* simulations

We begin by investigating in detail the chemistry of bundle breaking, which is however only possible for the smallest bundles, as it requires *ab initio* simulations. Hence, density functional simulations of stretching and breaking of chains at $T = 0$ K have been carried out on samples made of one and two PEO chains, each chain consisting of 10 –$(CH_2)_2O$– monomers.

3.1.1 Energy *versus* strain relation. The potential energy per monomer upon stretching a single PEO chain is shown in Fig. 2. A fit of the computational data using the anharmonic functional form:

$$V(\varepsilon) = V_0 + \alpha\varepsilon^2 + \beta\varepsilon^3 + \delta\varepsilon^4 \tag{10}$$

with $\alpha > 0$, where the strain per monomer $\varepsilon = (L_z - L_z^0)/L_z^0$ provides both an estimate of the minimum energy periodicity $L_z^0 = 3.46 \pm 0.05$ Å and an estimate of the potential energy V_0 at periodicity L_z^0. The relatively large uncertainty in L_z^0 is due to a number of reasons. At low strain the potential energy is controlled primarily by weak torsional restoring forces, the energy optimization with respect to the atomic coordinates is slow, the computed potential energy shows small amplitude fluctuations around the minimum, and even the anharmonic fit is not very accurate at negative strain.

To translate V_0 into a binding energy we also computed the ground state energy of the PEO monomer $(CH_2)_2O$ (see Fig. 2). Hence, the computed cohesive energy per monomer $V_c = 103$ kJ mol^{-1} in the single chain refers to the ring-opening polymerisation reaction from the ethylene oxide monomer, which is the simplest epoxide ether. Each atom has the same number and type of bonds in the monomer and in the chain, and the relatively low V_c represents the energy gain in releasing the large strain that is apparent in the monomer ground state geometry.

With length increasing beyond L_z^0, the fit of eqn (10) becomes accurate, and faithfully reproduces the system potential energy up to $\varepsilon = 0.28$. We observe that at low strain ε the dependence of potential energy on ε is remarkably anharmonic, as reflected in the behaviour of $dV(\varepsilon)/d\varepsilon$ reported in Fig. 2, which clearly deviates from Hooke's law in proximity of $\varepsilon = 0$. This is apparently due to the interplay of torsion, bending and stretching energies, whose relative size and role changes progressively with increasing strain.

A linear regime in $dV(\varepsilon)/d\varepsilon$ emerges at intermediate strain $0.04 \leq \varepsilon \leq 0.15$. However, this linear term does not go through the origin as in Hooke's law, but it is given by $dV(\varepsilon)/d\varepsilon = k(L_z^0)^2(\varepsilon - \varepsilon_0)$, with $k = 653$ kJ (mol Å^2)$^{-1}$ and $\varepsilon_0 \sim 0.02$. An effective Young's modulus $\bar{Y}(\varepsilon)$ can be estimated from the second derivative $d^2V(\varepsilon)/d\varepsilon^2$, that represents an effective strain dependent force constant $k(\varepsilon)$. To turn $k(\varepsilon)$ into $\bar{Y}(\varepsilon)$, we need to attribute a nominal cross section S to the single chain. This can be done using experimental data on the chain–chain distance ($d = 4.3$ Å) in crystal PEO.[31] Assuming hexagonal cross sections, one obtains $S = 16$ Å^2. Since the force constant depends on ε, also $\bar{Y}(\varepsilon)$ is a function of strain, growing from $\bar{Y}(0) = 82$ GPa at L_z^0 to $\bar{Y}(\varepsilon = 0.1) = 246$ GPa at $L_z = 3.8$ Å per monomer. Both these

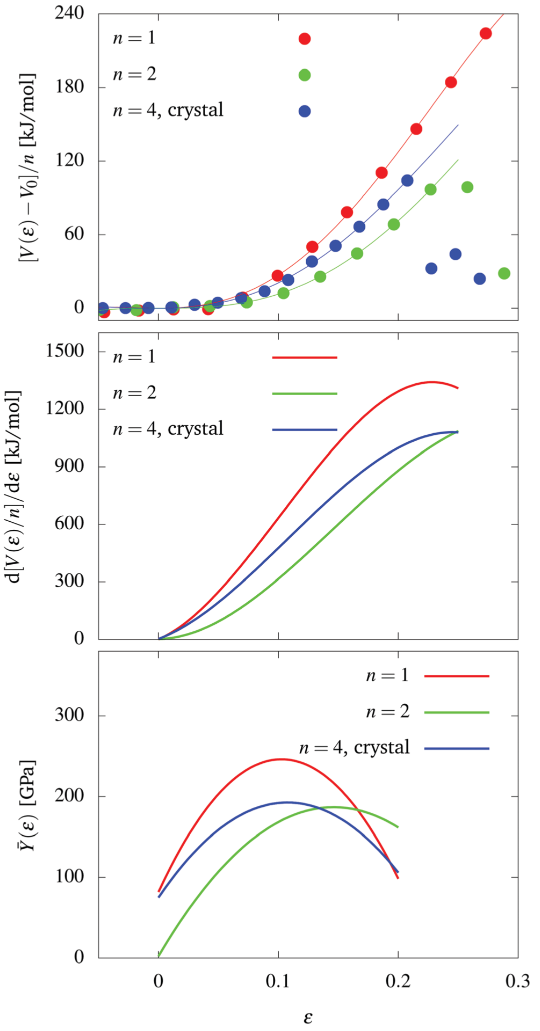


Fig. 2 Top: The potential energy of one and two PEO chains as a function of the strain $\varepsilon = (L_z - L_z^0)/L_z^0$, where L_z is the periodic length corresponding to one PEO monomer per chain and L_z^0 is the deviation from equilibrium at L_z^0 = 3.46 ± 0.05 Å for the single chain; L_z^0 = 3.26 ± 0.05 Å for the double chain; L_z^0 = 3.27 ± 0.05 Å for the crystal sample. The same quantity is reported for the stretching of the PEO crystal cell along the direction of chains, representing the limit of an infinitely extended bundle. Dots: Simulation results (DF); full lines: interpolation by the anharmonic fit in eqn (10). Middle: The stress as a function of the strain, computed by the anharmonic fit in eqn (10). All energies refer to the single monomer. Bottom: The effective Young's modulus, assuming a cross sectional area of S = 16 Å^2 per chain.

values are large, with the high-strain value being comparable to steel. Even the highest value can be understood by considering that it refers to the stage of straining dominated by the stretching of C–C and C–O covalent bonds.

It is worth emphasising that this high estimate of the Young's modulus is not directly related to the curvature of the potential energy around its minimum, but corresponds to a high load elastic regime in which a combination of bending and especially stretching energies provide the restoring force opposing further elongation of the chain.

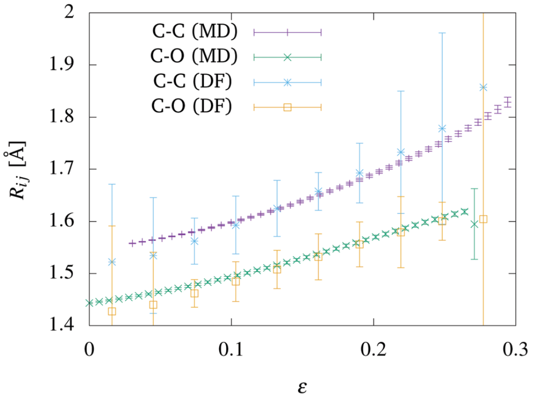


Fig. 3 Bond distances during the stretching of a single PEO chain computed by GGA-DFT and MD.

In addition, the stretching process in aligned chains does not correspond to the elastic deformation of a macroscopic PEO sample, in which the averaging over glassy domain and crystal grains of different orientation give origin to a linear (elastic) regime at low strain, of Young's modulus and elastic constants much reduced with respect to the values computed for oriented, defect-free chains.

3.1.2 Structural changes upon stretching. The non-linear dependence of d$V(\varepsilon)$/dε on load is reflected in the strain dependence of the C–C and C–O bond lengths, reported in Fig. 3, of bending angles, shown in Fig. 4, and of dihedral angles shown in Fig. 5. The plots confirm that, as expected, dihedral angles are the first to manifest sizable strain, followed by bending, and stretching in the last stage. The MD results for

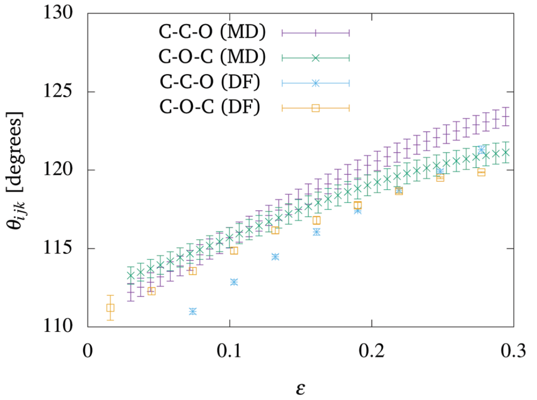


Fig. 4 Bending angles during the stretching of a single PEO chain computed by GGA-DFT and MD.

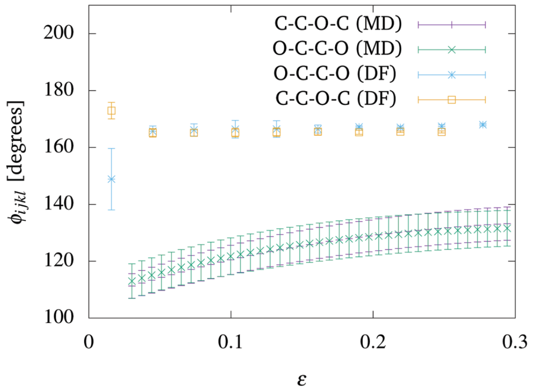


Fig. 5 Torsional angles during the stretching of a single PEO chain computed by GGA-DFT and MD.

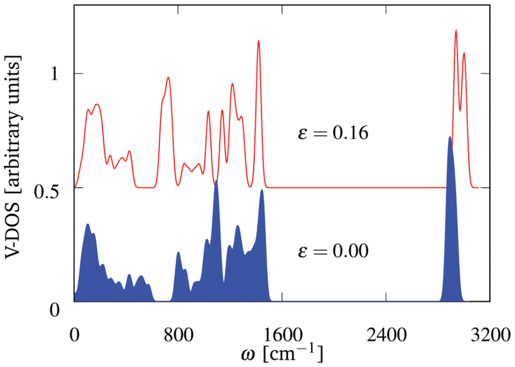


Fig. 6 Comparison of the vibrational density of states (V-DOS) computed by DF at two different values of the longitudinal periodicity: $\varepsilon = 0.00$ (blue line) and $\varepsilon = 0.16$ (red line).

the bond lengths match the DFT results over the entire range. The angles are shifted by about 3 to 4 degrees in the range we are interested in but otherwise also show similar behaviour as well. Only the torsional angles show significant quantitative deviation between MD and DFT calculations. This could be related to a shift in the hybridisation of the C atoms (see below), which is not captured in the classical force field. Regardless, the high-strain behaviour is dominated by bending and stretching, and therefore this difference will likely not have a qualitative effect on the behaviour at high strain, which is what we focus on in this work.

Deviation of the applied force from the linear behaviour at $\varepsilon > 0.19$ corresponds to the onset of high-strain anharmonicity, which dominates the system behaviour in the last stages of stretching. The applied force reaches its maximum at $\varepsilon = 0.22$. Beyond that length, the force decreases with increasing strain, and the system would be unstable under constant stress conditions. If subjected to thermal fluctuations at non-zero temperature, it would break even earlier. The clearest signature of breaking appears in the strain dependence of bond lengths. More in detail, the strain dependence of C–C and C–O is non-monotonic at first, then both bond lengths increase linearly for $0.04 \leq \varepsilon \leq 0.2$, with a relative variation larger for the C–C than for the C–O bond length. Non-linearity in the strain dependence of the bond lengths appears at $\varepsilon \sim 0.23$. Above this length the slope of C–C with strain turns upwards, indicating the incipient failing of C–C bonds. Anharmonicity of the restoring force and non-linearity of bond lengths *versus* strain are fully developed at $\varepsilon \sim 0.26$, an elongation at which the single chain appears broken at several C–C bonds even by visual inspection of simulation snapshots. Remarkably, the trend shown by C–O over the same range of global strain is opposite, since C–O saturates at a fairly high value while C–C increases rapidly.

An interesting view of the same breaking process is obtained by looking at the standard deviation along the chain of individual bond lengths, shown in the error bars of Fig. 3. The minimisation of the chain energy in the anharmonic regime at low strain makes bonds slightly inequivalent from each other despite connecting the same type of atom. Thus, the standard deviation σ of chemically equivalent bond lengths is not negligible at low strain. Then, σ decreases with increasing strain, since restoring forces become stronger, making bond lengths better defined. Eventually, σ shows a rapid and drastic increase above $\varepsilon = 0.23$, providing the most unambiguous sign of chain breaking.

The spread of C–C bond lengths leading to breaking could be reflected in spectroscopic data. To highlight this effect, we computed by DFT the vibrational density of states of the single PEO chain at three different values of the longitudinal periodicity, corresponding to $\varepsilon = 0$, $\varepsilon = 0.16$ and $\varepsilon =$ and $\varepsilon = 0.25$. At the near-equilibrium periodicity of $\varepsilon = 0$, the vibrational density of states shows the typical features of similar organic systems, with a high frequency band at $2800 \leq \omega \leq 3100$ cm^{-1} due to the C–H bond stretching, an intermediate band due to bond bending at $1200 \leq \omega \leq 1600$ cm^{-1}, and a broad background of modes at $0 \leq \omega \leq 1200$ cm^{-1}, of mixed character but consisting primarily of bond-torsion modes (see Fig. 6). With increasing strain, the C–H stretching band moves to somewhat higher frequency and splits into a symmetric (at the highest frequency) and an anti-symmetric (at slightly lower frequency) C–H stretching. Bending modes move to slightly lower frequency, while the range at $\omega \leq 1200$ cm^{-1} shows the major and least predictable changes, that could easily be detected by spectroscopy and could be used to monitor the system evolution up to near breaking conditions. The chain breaking itself, taking place at $\varepsilon = 0.23$, is likely to be a stage too short to be characterised by spectroscopy. Moreover, the changes in the vibrational DOS upon breaking are less easy to predict and to interpret, since new molecular species appear in the system, possibly spin-unpaired radicals, not easily accounted for even at the DF-level.

Notice that with increasing strain bending angles tend to 120° and dihedral angles tend to 180°, values that characterise organic structures of sp^2 bonding. Since bonding angles modulate the size of Hamiltonian matrix elements, their correspondence with the sp^2 geometry suggests that the sp^3 bonding of the unstrained chain turns to sp^2 at high strain. The change of electronic structure with increasing strain is supported by the computation of atomic charges, evaluated by the electrostatic potential (ESP) method.[32] The results show an enhancement of atomic charges with increasing strain, followed by a sudden drop at breaking. These effects are not accounted for by standard reactive classical force fields, such as ReaxFF.[21]

As already stated, the single PEO chain does not collapse upon crossing the length corresponding to the maximum restoring force. In other words, the chain does not snap during simulated breaking, because of the constant strain condition enforced by these DF simulations, and of the absence of lateral perturbations from thermal fluctuations that could undermine the strained geometry in its metastable regime. Moreover, no simple healing process is available to the single chain, which splits in a collection of geometrically equivalent strained epoxy units. In reality, it might be practically impossible to maintain (meta)-stability beyond the linear stages, at $\varepsilon > 0.19$.

3.1.3 The two-chains system. To assess the role of chain–chain interactions the simulation of stretching at $T = 0$ K has been repeated for a sample consisting of two chains. The effect of these interactions turns out to be sizable, and, surprisingly, it has the effect of decreasing the strength of the bundle under tensile load.

Also in this case, strain has been imposed by changing the periodicity of the unit cell containing 2×10 PEO monomers. As before, for any given periodicity, the potential energy $V(\varepsilon)$ has been carefully optimised with respect to the atomic positions by quenched molecular dynamics, and the resulting $V(\varepsilon)$ has been fitted by the same expression of eqn (10). The length of minimum potential energy turns out to be $L_z^0 = 3.26$ Å, *i.e.*, nearly 0.2 Å shorter than for a single chain. This could be due to the effect of the surface tension (surface energy, at $T = 0$ K) per unit length of the bundle, which is stronger for two chains than for one, and tends to reduce the surface area and thus the length of the system.

Comparison of the ground state energy of the single and double PEO chain computed at the DF level allows us to estimate the chain–chain cohesive energy of two chains at 1.2 kJ mol^{-1} per PEO monomer. A similar comparison for the crystal sample gives a cohesive energy of 13.1 kJ mol^{-1} per PEO monomer. We emphasise that this cohesive energy arises from the chain–chain interaction, due to dispersion energy only, and without any reference to the energy gained in forming chains from the epoxide monomer. With increasing strain the evolution of $V(\varepsilon)$ (shown in Fig. 2) qualitatively follows the same behaviour of the one-chain case. The computation of dihedral angles, bending angles and bond stretching shows that also these structural parameters follow the same trends in both cases, consisting at first in the deformation of dihedral angles, then of bending angles and finally of stretching bond distances. A quantitative comparison is not easy, because of the different reference length L_z^0 of the two samples. It is apparent, however, that the two-chains system breaks at a shorter length and at a lower applied force per chain than in the single chain case. In the case of two chains, the apparent Young's modulus computed from $\mathrm{d}^2V(\varepsilon)/\mathrm{d}\varepsilon^2$ depends on strain, starting from about 2 GPa at low strain, and reaching 183 GPa at high strain.

Analysis of configurations shows that the monomers in the two chains never really align in the lateral direction, confirming the sizable role of chain–chain interactions. These interactions are apparently amplified by the helicity of each PEO chain, adding a longitudinal modulation of short-range repulsive interactions in addition to the attractive dispersion forces. The longitudinal corrugation in the interaction energy, in particular, gives rise to multiple local minima in the potential energy, that make it difficult to unambiguously identify the ground state of bundles.

3.1.4 Chemical rebonding after chain breaking. *Ab initio* computations offer a glimpse of the events following the breaking of individual chains in the two-chains system. We find that at first, the separation of one chain into nearly equivalent segments is observed, terminated by planar OCH_2 groups, in which C is in the sp^2 electronic configuration. Then, interaction with the neighbouring chain causes the collapse of the bundle, with a short sequence of chemical changes that carry out the partial healing of the broken two-chains bundle as displayed in Fig. 7. At long times, the healing results in finite chains terminated by OH on one side, and planar OCH_2 on the other side, with the release of a few epoxide ether monomers. The post-breaking evolution given by DF computations might be affected by limitations of the approximate DF approach in tackling open shell species, and, in any case, it is not correctly reproduced by the classical force field model, even at the qualitative level. Perhaps it might be described by some highly sophisticated reactive force field models, that, however, require heavy stages of parametrisation.

3.1.5 Stretching crystalline PEO. To assess the effect of helicity of PEO chains and of lateral chain–chain interactions, simulations have been carried out for the experimental crystal structure of PEO, in which chains are closely packed. In our study, the periodic system is seen as the limiting case of thick bundles. The apparent effect of the condensed environment is to give a complex unit cell, containing four chains and 28 monomers in total, with an average equilibrium length per monomer of 3.27 Å.

Especially at low strain, the $V(L_z)$ curve (see. Fig. 2) for the extended system shows a fine structure, apparently due to the lateral interactions, creating conditions of local minima and sometimes preventing a complete energy optimisation. Nevertheless, the energy increase with increasing periodicity at low strain is more parabolic than in the single and double chain case, since in the condensed phase it is more difficult to decouple torsion, bending and stretching than in the few-chain

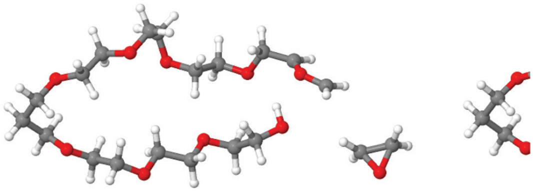


Fig. 7 Partial healing of the broken two-chain bundle.

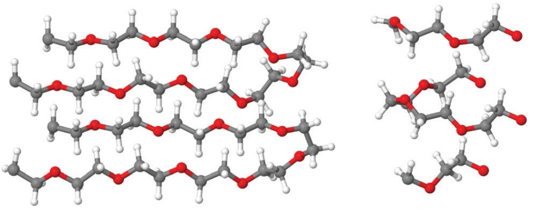


Fig. 8 Final state upon stretching a crystal sample up to the breaking point. The system has been rotated along the long axis to expose chains that otherwise are partially hidden in the trapezoidal cross section. Healing takes place by connecting neighboring chains forming loops returning to a single clamp. One of the chains appears to be broken because of periodic boundary conditions.

samples. As a result, the apparent Young's modulus $\bar{Y}$ computed from the strain-dependent force constant is somewhat more constant than in the one- and two-chain cases, varying from $\bar{Y}$ = 75 GPa at low strain to $\bar{Y}$ = 190 GPa at intermediate strain. The bundle breaks at shorter periodicity per monomer (L_z = 3.7 Å) than the one and two-chain system. Also in this case, the breaking of one chain destabilises neighboring chains, giving rise to chemical reactions. The healing of broken bonds is complete, giving rise to a set of loops breaking the continuous connection between the two sides of the broken bundle, without leaving behind any radical species as shown in Fig. 8.

As a result of the healing process, the final potential energy per monomer is only 3.2 kJ mol^{-1} higher than the ground state energy. The low energy of the final state emphasises the fact that bundles become metastable very soon upon stretching, and only the kinetic barrier encountered by the healing process prevents the breaking at much lower strains and stresses than found in the simulation. The argument is even more compelling for thick bundles, in which healing might occur by reacting with neighboring chains, and for long bundles, since the energy required for a given relative stretching is extensive, and the energy required to break a chain is a constant. There are obviously many things in the chemistry of the chains that are not captured by the classical force field. The most important among them is the fast chemical rebonding of chains following their breaking under strain, effectively removing all dangling bonds from the system. However, from comparing the behaviour of the bond stretching and bending at high strain, we conclude that the classical force field captures the behaviour of the chain under high strain sufficiently well up to the breaking point.

3.2 Stretched bundles with $N > 2$ chains

We continue by investigating the structural properties of stretched bundles of up to 24 chains, now using molecular-dynamics simulations and potential energy minimisation based on the empirical force field described above. At any non-vanishing temperature, entropy effects turn unconstrained chains into coils, whose radius scales like the square root of the number of monomers N. In addition, there are finite-size effects and bead–bead, *i.e.* non-bonded, interactions which cause deviations from the ideal scaling. In our systems, the clamps and the tension applied to the bundle change the picture qualitatively and quantitatively. An example of a stretched bundle at finite temperature is shown in Fig. 1.

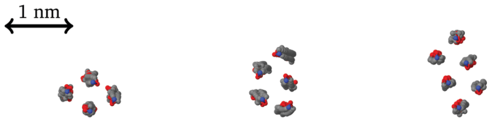
3.2.1 Structure at low and vanishing temperature. To give a first characterisation of structural, cohesive and elastic properties of PEO bundles, and to provide a term of comparison for the results of MD simulations to the DF simulations, we first investigate the bundles in the limit of very low temperature and under moderate tensile load (κ = 1 nN per chain), an order of magnitude below the breaking limit. At low temperature T this moderate tension is sufficient to force the bundle's chains into approximatively linear configurations. Once chains adopt the linear configuration, the bundle structure is determined mainly by the geometry of its cross section, that can be visualised as a 2D cluster, in which the interaction among particles is represented by the integral along z of the lateral chain–chain potential energy. The full 3D system, however, is more complex due to the nonisotropic lateral interactions.

To approach the T = 0 K condition, the energy of N-chain bundles is first minimised with respect to all coordinates (at fixed applied force) by short annealing runs of 1 ns. These cannot guarantee that the simulation will reach the ground state of the system, but the relatively ordered geometry of samples up to at least room temperature suggests that the potential-energy landscape consists of only a few major valleys, and the geometries determined by short annealings are likely to be representative of low temperature structures, providing information also on defects and isomers.

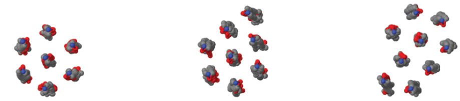
The sequence of clusters obtained in this way is reported in Fig. 9. As already stated, in this figure all bundles are under a 1 nN per chain tensile load. Once again, as expected, the compact hexagonal motif dominates the structure of the 2D clusters. However, non optimal geometries arise because of short annealing and incomplete optimisation. Moreover and more importantly, both unexpected 2D geometries and longitudinal variations in the cross section arise from the spontaneous helicity of PEO chains, which adds a small but complex perturbation to the lateral interactions.

An example of nearly degenerate isomers in the 2D cluster representation, corresponding to different bundle polymorphs, is provided by the structure of the 20-chain bundle (see Fig. 10), found in several different configurations as a result of successive fast optimisation cycles. The cross section may also change along z, when one or more chains are not quite straight, but show a marked winding around some other chain as seen in Fig. 10(b).

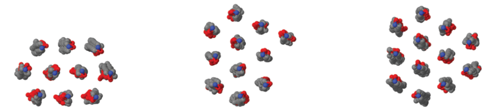
Fig. 11 shows the average force–elongation curves for the three high-energy isomers shown in Fig. 10(a)–(c), and for the three low-energy isomers shown in Fig. 10(d)–(f). The initial slope is the same, but the high-energy isomers are able to sustain a little more force than the low-energy ones. This observation is reminiscent of mechanical hardening in metals. However, the two effects are not obviously related, since hardening in metals is caused by defects blocking the movement of dislocations, while in our system dislocation dynamics does not play an obvious role in the yield. It is thus not clear



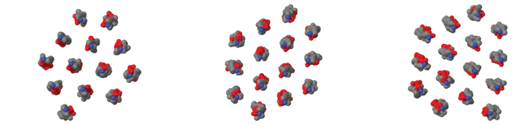
(a) 4 chains (b) 5 chains (c) 6 chains



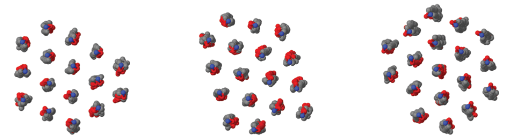
(d) 7 chains (e) 8 chains (f) 9 chains



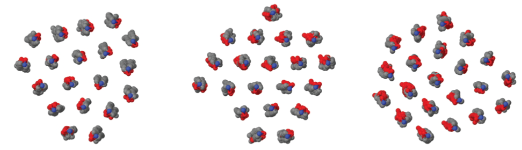
(g) 10 chains (h) 11 chains (i) 12 chains



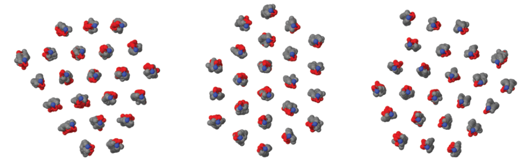
(j) 13 chains (k) 14 chains (l) 15 chains



(m) 16 chains (n) 17 chains (o) 18 chains

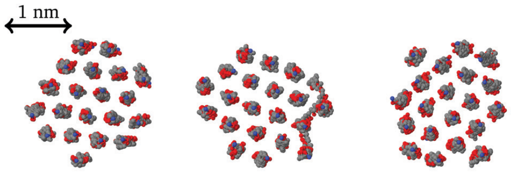


(p) 19 chains (q) 20 chains (r) 21 chains

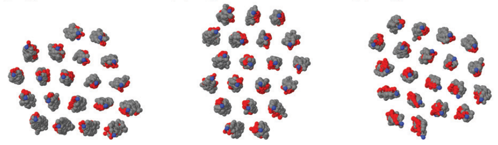


(s) 22 chains (t) 23 chains (u) 24 chains

Fig. 9 Cross sections of bundles of sizes from N = 4 to N = 24 at T = 0 K with a constant force of 1 nN per chain.



(a) $V_c/N = 583$ kJ/mol (b) $V_c/N = 594$ kJ/mol (c) $V_c/N = 597$ kJ/mol



(d) $V_c/N = 641$ kJ/mol (e) $V_c/N = 646$ kJ/mol (f) $V_c/N = 650$ kJ/mol

Fig. 10 Isomers of the 2D cluster representing the structure of the 20-chain bundle at T = 400 K with a constant force of 1 nN per chain obtained by equivalent optimisation cycles started from independent configurations. The cohesive energy per chain associated with each configuration averaged over 2 ns is given in the sub-caption. Figure (a)–(c) represent low cohesive energy configurations, and (d)–(f) represent configurations with high cohesive energy.

what is the origin of this effect, whose magnitude, in any case, is apparently quite small.

The ideal, regular-looking cross-section is not always the one with the lowest energy. In several cases, the atomistic bundles of N chains adopt a cross-section different from the one in the ideal sequence, contradicting the assumptions of cylindrically symmetry and invariance along z. We found that this difference is often not the result of incomplete optimisation of the atomistic bundles, since preparing them according to the ideal cylindrical cross section almost invariably results in a higher potential energy. Structures of this type clearly point to the effect of the PEO chain helicity on the geometry of bundles.

The plot of the cohesive energy per chain in the lowest energy structure for each N displays the characteristic trend found in the size dependence of cohesive energies for atomic and molecular clusters. Cohesive energy increases rapidly at first, and then saturates slowly to the cohesive energy per chain of a hexagonal arrangement of linear and aligned chains. The shortfall of cohesive energy at low N can be attributed to the surface energy, whose fractional weight on the total cohesion is highest for N = 1, and decreases as $\sqrt{N}$ with increasing size.

To leading order, the energy of the bundle can be estimated from the volume and the surface area, *i.e.* the compact surface delimiting the bundle. First, we quantify the surface area, or equivalently, the perimeter of the 2D cluster. To this aim, we attribute an hexagonal cross area to each chain in the bundle, with an interchain distance d = 4.3 Å. Then, the perimeter C_N of the 2D cluster is determined by counting the number of hexagon sides exposed to the vacuum. The size dependence of the cohesive energy of the bundle is represented as:

$$V_c(N) = NV_0 - \sigma C \tag{11}$$

where V_0 and σ represent the bulk energy per chain and the surface energy, respectively, to be determined by fitting the numerical values of the cohesive energy. Since, on average, the contour length scales as $C \propto N^{1/2}$, we recover the expected scaling of the cohesive energy per chain:

$$\frac{V_c(N)}{N} = V_0 - \alpha\frac{\sigma}{N^{1/2}} \tag{12}$$

where the numerical coefficient gives the (average) proportionality of C_N and $N^{1/2}$.

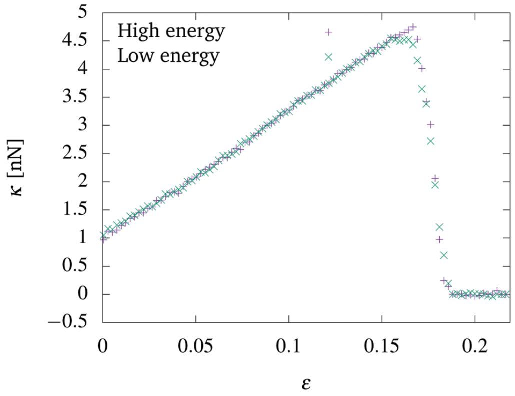


Fig. 11 Force elongation curves averaged over the three high energy configurations showed in Fig. 10(a)–(c) and for the low energy configurations showed in Fig. 10(d)–(f).

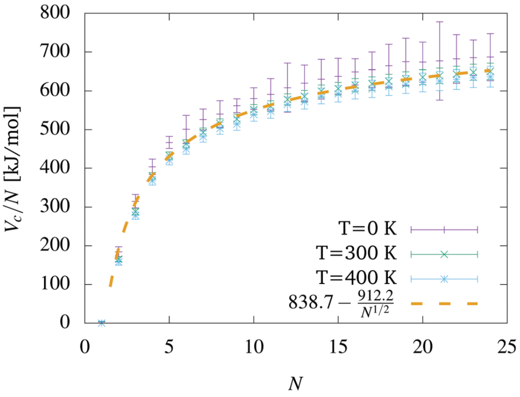


Fig. 12 Computed cohesive energy per chain as a function of size N at T = 0, 300, 400 K, with a fit by the smooth expression in eqn (12), representing the interplay of bulk cohesion and surface energy.

Comparison of the computed $V_c(N)$ with the result of the smooth fit in eqn (12) shows a good match, as seen in Fig. 12. The single sample with N = 100 had V_c/N = 760 kJ mol^{-1} after 1 ns of annealing, which is about 2% higher than the prediction given by the fit for T = 300 K in eqn (12). To investigate the impact of the annealing time, the annealing procedure was repeated with an additional 2 ns of annealing time, which increased the cohesive energy by about 3%. While longer annealing times indeed results in lower energy structures, 1 ns is considered to give a reasonable balance of realism and efficiency.

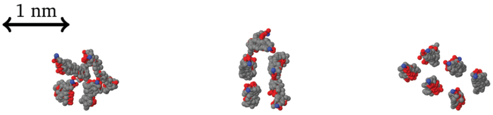
The N and $\sqrt{N}$ scaling of cohesive energy is less apparent in the average bundle length $L(N)$ measured by the separation of the clamps and Young's modulus at non-zero temperature, as we will see in the next Section 3.2.2. In these cases, we see a sequence of peaks and dips, that point to sizes whose packing exposes an optimal or unfavored contour, respectively. The single sample with N = 100 at T = 300 K had a length of L = 114.6 Å after the first annealing run of 1 ns, and a length of L = 114.7 Å after an additional annealing run of 2 ns, which is about 0.2% lower than expected from the smaller systems.

At T = 0 K we determine the relation of length L_z and stress by increasing L_z in regular steps, and minimizing the potential energy at each step with respect to all internal degrees of freedom. Stress κ is computed from the forces on the terminals. The Young's modulus for the breaking of a single chain was calculated from the slope of the beginning of the strain curve to be 144 ± 6 GPa, again assigning the chain a cross sectional area of S = 16 Å^2.

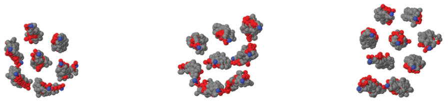
Breaking at T = 0 K takes place in a localised way, focusing strain on a limited stretch of the chain when the system crosses the stability boundary. At T = 0 K, the pictures of breaking at constant stress or constant strain are equivalent. However, these two modes of stretching differ in very important aspects at non-zero T, because of the role of fluctuation and thermal activation.

3.2.2 Nonzero temperature. At finite temperature, we use the simulated annealing protocol described in Section 2 at the desired temperature, to obtain equilibrium conditions at a constant applied force per chain. Typical cross-sections of the bundles obtained in this way are shown in Fig. 13. The figure shows that the hexagonal structure observed at low temperature remains highly ordered even at high temperatures. When the number of chains in a bundle has specific values (such as 3, 7, or 19), a shape of minimal contour forms with complete shells. However, when the number of chains is not one of these "magic" numbers, then there are typically several possible isomers appearing. Nevertheless, the hexagonal lattice structure usually remains. We have observed one special case, for N = 22, where we have found a 7-fold symmetric bundle with full shells. In extended 2D lattices, such defects have to be coupled to a 5-fold symmetry defect, forming the two end points of a dislocation line. In finite systems, one of the two conjugated defects may be annihilated at the surface. Sometimes, but especially if we do not anneal the system and instead run it at constant temperature, we find a more disordered state, where the isomerisation changes along the length of the bundle. We have also observed that isomerisation can also change with time.

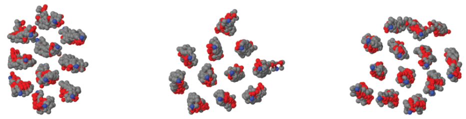
The lengths of the bundles at T = 300 and 400 K with a constant applied force per chain of 1 nN is shown in Fig. 14. For comparison, a single chain at T = 0 K had a significantly longer length of 115.7 ± 0.1 Å. The length depends only weakly on the size of the bundle and levels off quickly for bundles consisting of more than 20 chains. For small bundles, there is a clear signature of the "magic" numbers at 3 and 7, but for larger bundles this effect is smaller and washed out by statistical error and thermal noise. Overall, the effect of the bundle size is significant, but small. The cohesive energies show even less of the structure (see Fig. 12). At this strain, the entropic effects are small, which is evident from the weak temperature dependence of both the average length and the Young's modulus. The general trend is the same as for the *ab initio* computations



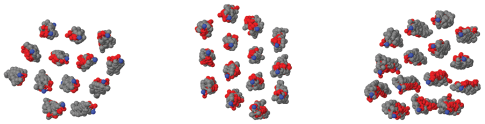
(a) 4 chains (b) 5 chains (c) 6 chains



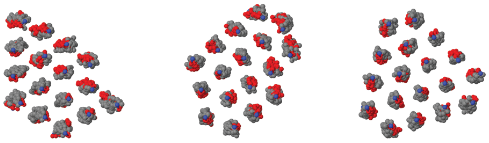
(d) 7 chains (e) 8 chains (f) 9 chains



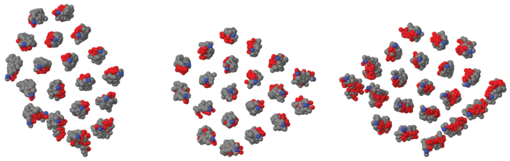
(g) 10 chains (h) 11 chains (i) 12 chains



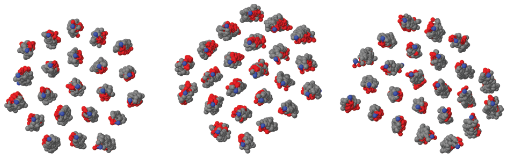
(j) 13 chains (k) 14 chains (l) 15 chains



(m) 16 chains (n) 17 chains (o) 18 chains



(p) 19 chains (q) 20 chains (r) 21 chains



(s) 22 chains (t) 23 chains (u) 24 chains

Fig. 13 Cross sections of bundles of sizes from N = 4 to N = 24 at T = 400 K with a constant force of 1 nN per chain.

presented in Section 3.1 in that the equilibrium length is shorter with increasing bundle size. This can be seen clearly in Fig. 15, where the average length at which the first chain in the bundle breaks is shown as a function of bundle size. For the bundle with 100 chains at T = 300 K the first chain broke at a length of L = 134.8 Å, in good agreement with the picture from Fig. 15.

The force per chain as a function of strain is shown in Fig. 16 for a number of different-sized bundles at various temperatures.

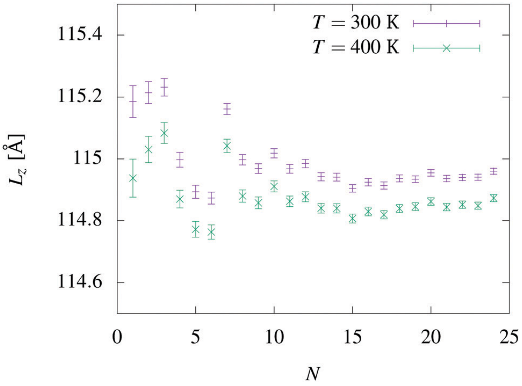


Fig. 14 Average bundle length $L(N)$ for 100 samples as a function of size N with a constant force of 1 nN per chain at T = 300 K and T = 400 K.

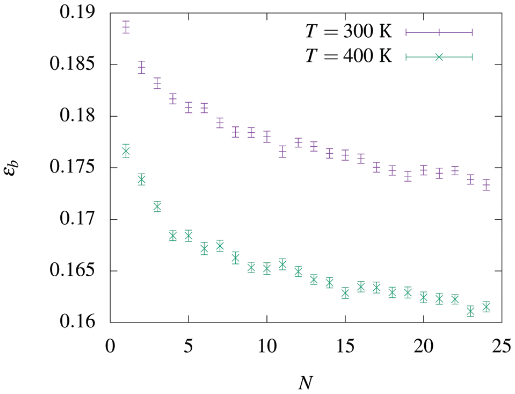


Fig. 15 The average strain at which the first chain in the bundle breaks, indicated as ε_b, for 100 samples at constant strain rate 0.8 m s^{-1} as a function of size N at T = 300 and T = 400 K.

For all sizes, there is an initial linear elastic regime. The force remains almost linear in the strain until the first chain in the bundle breaks. The Young's modulus can be computed from the slope of $\kappa(L)$ at low strain. Here it was estimated by linear regression up to a strain of 0.05, about 30% of the strain at which the first chain breaks. The resulting Young's moduli are presented Fig. 17. The modulus display no significant size dependence for systems of size $N > 9$, also the bundle of 100 chains at T = 300 K had a Young's modulus of Y = 130 GPa. Experiments with Linear PE also indicate that the Young's modulus to a good first approximation is independent of the molecular weight, as the macroscopic deformation is the dominant factor.[33] In agreement with previous findings,[34] one can see that a decrease in temperature will lead to an increase in the modulus.

The estimated modulus is high. The experimental values for the Young's modulus of bulk PEO is typically on the order of 0.1–20 MPa.[35–37] Though the Young's modulus and the tensile

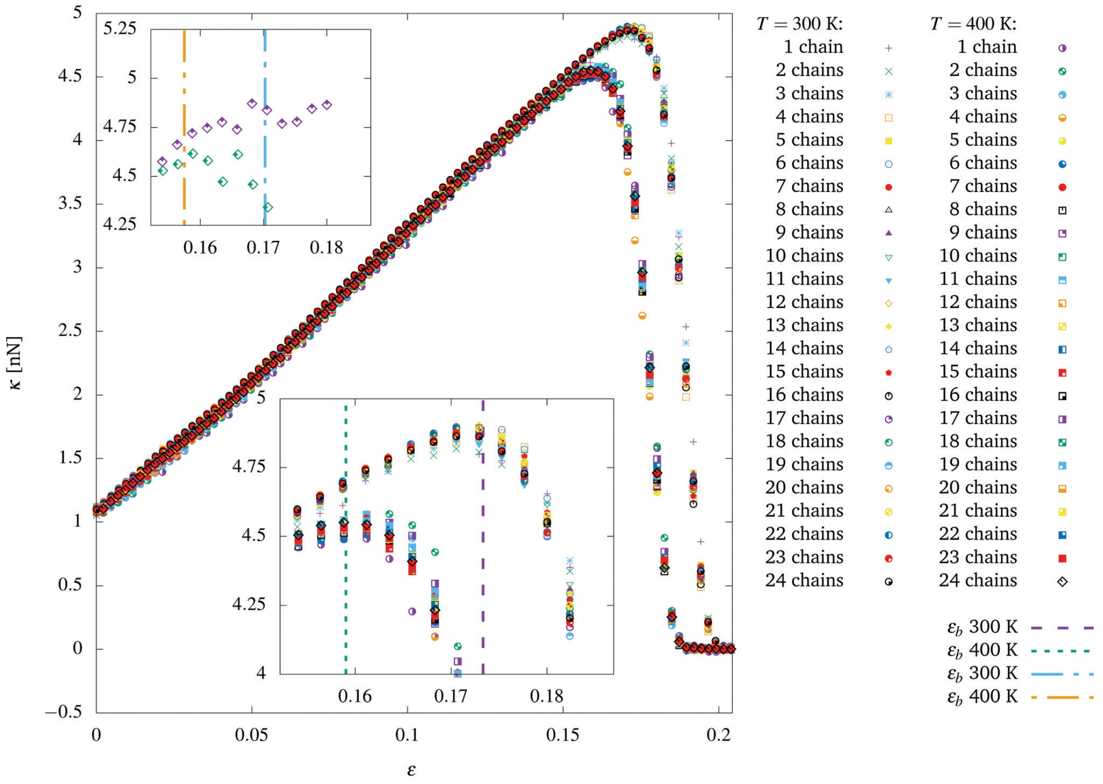


Fig. 16 The force per chain during constant strain rate simulations at 0.8 m s^{-1} for system of bundles of 1–24 chains of at T = 300 K and T = 400 K. The lower inset shows the breaking part zoomed in, with lines indicating when the first chain broke on average in bundles of size N = 24, labeled as ε_b. The upper inlet shows single samples with bundles of size N = 24, and again the lines indicate when the first chain in the bundle breaks.

strength of these amorphous samples are not determined by the stretching of covalent bonds, but rather by weaker interactions.[38]

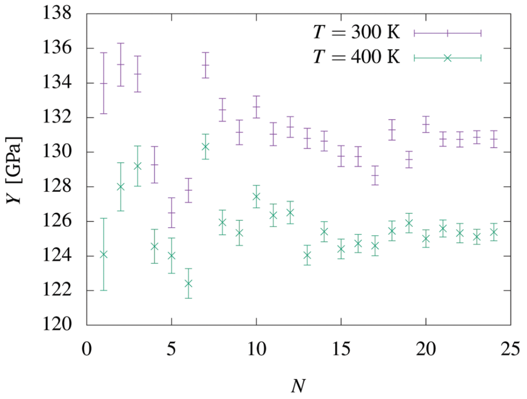


Fig. 17 The Young's modulus of bundles of PEO with a constant strain rate of 0.8 m s^{-1} as a function of size N at T = 300 and T = 400 K. The Young's modulus are estimated from the strain curves, with 100 samples per system.

For the elastic properties of PEO along the chain axis, an experimental value of 10 GPa has been reported,[39] in well agreement with computational results.[9] The values we find from our simulations for PEO are much higher, comparable to the values for PE.[40] Stretching of thermoplastic PE gives a Young's modulus on the order of 1 GPa,[41] while experiments with single crystals gives a modulus in the range 168–278 GPa,[42,43] and even up to 370 GPa for low density films.[39] *Ab initio* computations of single chain PE gives a modulus between 300 and 500 GPa.[10,34,44]

We can understand this by considering that the stretching constants of bulk PEO in experiments is strongly determined by the structure. A major difference between PE and PEO is that the latter is known to have a loosely coiled confirmation.[9] The high modulus and high strength found in fibers relies mainly on high polymer chain orientation and extension, and polymers with helical chain configuration exhibit a much lower theoretical modulus.[45] In our model the chains are extremely extended and ordered, producing systems far from the complexity of macroscopic samples. Our semi-crystalline fibers resemble more closely the structure of PE, and thus display similar mechanical properties.

3.3 Breaking

We now turn to the further stretching and finally breaking of the larger bundles in the classical molecular-dynamics simulations. We discuss here the results of simulations at a constant strain rate, stretching out the bundles, until all the chains have broken. Fig. 18 shows a sequence of snapshots of a bundle of 17 chains being stretched out and breaking. The computational limiting resistance load κ_1 was estimated from the average curves in Fig. 16 to 30.5 ± 0.1 GPa at 300 K 28.4 ± 0.1 GPa at 400 K. For 0 K, κ_1 was estimated to 36.9 ± 0.3 GPa using the same methodology with 10 samples of single chain bundles.

Because our simulations are at constant strain rate, the force that was applied to a chain before it breaks does not get redistributed over the other chains. This allows us to isolate the interaction between the chains from the interaction mediated by the terminals, which cloud these effects in a system with more realistic boundary conditions. If there were no relevant interaction between the chains, then each chain would break independently, and the fraction of broken chains would not depend on the bundle size. We observe only a small systematic size dependence.

There are several important further observations that can be made regarding the dependence of breaking on temperature and bundle size.

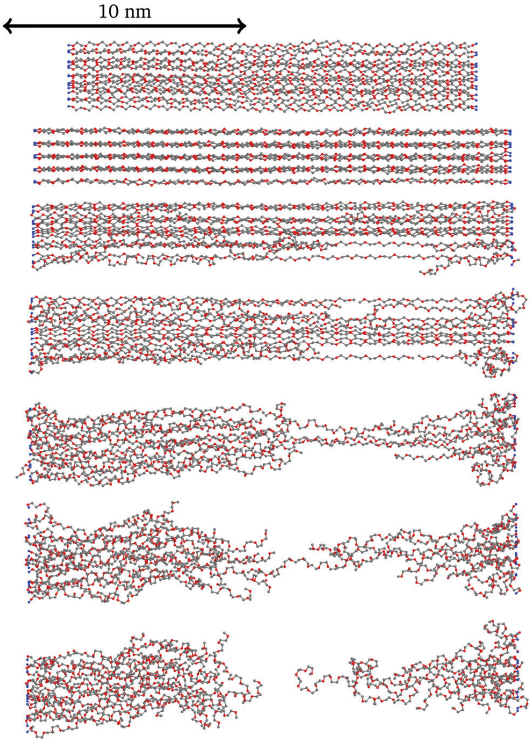


Fig. 18 Snapshots of a bundle of 17 chains under a constant strain rate of 0.8 m s^{-1}, stretching out and breaking at T = 400 K.

3.3.1 Thermally activated breaking. At higher temperature, the chains are slightly shorter and the Young's Modulus is lower, in agreement with previous findings from DF simulations of crystalline polyethylene.[34] Also, chains break earlier and breaking times follow a broader distribution. Ideally, neglecting the weak lateral interaction, at T = 0 K, all chains break at the same length.

The breaking rate of chains from the simulations is shown in Fig. 19. The spread of breaking times and earlier breaking can both be understood from thermal activation. Before the zero-temperature breaking point is reached, the chains can break by thermal fluctuations. At higher temperature, there are more and bigger thermal fluctuations, and hence the chains may break more quickly. The activation free energy depends on the strain. At low strain rates, it takes more time for thermal fluctuations to overcome the higher barrier.

We can make some simple estimates of the thermally activated breaking rate and dependence on strain rate and temperature. For this purpose, we treat the chain as one-dimensional and assume that the chain–chain interaction is negligible. We assume that each breaking bond experiences a mean field response from the rest of the chain equal to the average force in the chain, F_0. The total potential-energy landscape then consists of a linear term and the energy of the bond. For small forces, there is a deep minimum for the intact bond and an even deeper escape with a broken bond, with a high barrier in between. For some critical force F_c this minimum completely disappears at r_c, and even at T = 0 K the bond would break without any thermal activation. When a bond breaks through thermal activation, it can be assumed to be close to this point and thus F_0 is close to F_c.

We expand the potential-energy landscape around the breaking point r_c to the third order, and locate the nearest maximum and minimum at a distance close to the breaking point,

$$\Delta r = \sqrt{\frac{\Delta F}{3C_3}}, \qquad (13)$$

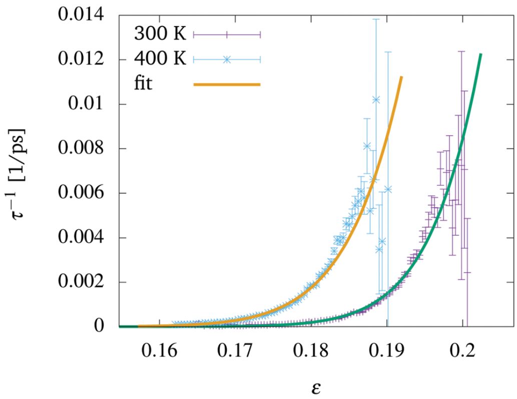


Fig. 19 The rate of chains breaking as a function of strain, for temperatures T = 300 and T = 400 K averaged over bundles of size 16–24 chains with 100 samples each. A fit using the expression for one-dimensional thermally activated breaking, eqn (15), is also included.

where $\Delta F = F_0 - F_c$ and C_3 is the third order expansion coefficient. We can then obtain the escape rate from the Kramers rate given by

$$\tau^{-1} = \frac{\Omega^2}{2\pi\gamma}\exp[-\Delta E/(k_B T)] \tag{14}$$

with Ω being the instantaneous effective oscillation frequency, ΔE being the energy barrier and γ the microscopic friction coefficient (viscous). We find

$$\tau^{-1} = C\Delta r\exp(-6C_3\Delta r^3/(k_B T)), \tag{15}$$

with C a constant. This result contains several parameters that depend on the details of the complex potential landscape and dynamics, which we must obtain indirectly through fitting.

A fit of eqn (15) to the simulation results is included in Fig. 19. The expression fits well, but we note one odd result in the fit parameters. The value found for the parameter r_c, 151.9 ± 0.2 Å, the length at which a break is immediate, is higher than the breaking length found in the 0 K simulations, 147.1 ± 0.2 Å for a single chain. This is probably related to the fact that the real energy landscape is much more complicated than in our simple 1D estimate, and there are several different reaction paths that may be activated at different temperatures. We have also briefly tested the dependence on the strain rate, and found that it behaves as expected: At higher strain rate, the chains break at a higher strain. The breaking rate, however, is not dependent on the strain rate, which implies that we are indeed in the adiabatic regime.

3.3.2 Bundle size and structure. We have investigated the dependence of various observables on the size of the bundle. Neglecting lateral chain–chain interactions, the force *versus* strain curve would be the same for all sizes. Simulation results, however, show that the force per chain for bundles of different sizes do not fall exactly on top of each other, but are quite close. This observation points to the effects of chain–chain interactions, which are small but not completely negligible, as already suggested for the results for the linear regime and for the Young's Modulus.

To deepen our understanding of the bundle effect on chain breaking, we examine the location of the first bond that breaks inside the bundle. As seen in Section 3.1, the C–C bonds are weaker than the C–O bonds, and are therefore more likely to break. At T = 300 K, 94.8% of the first bonds to break in each chain are C–C bonds. This strengthens the connection to PE mentioned in Section 3.2.2. At T = 400 K, 87.4% of the first bonds to break in each chain are C–C bonds. Apart from this, the distribution of breaking bonds is fairly uniform over the length of the chain, with the exception of the end bonds, which we have kept as harmonic, as is described in Section 2. If this is not done, they are about 3 times more likely to break than other bonds.

We can also investigate the relation between the location of the first chain that breaks and the structure of the cross-section of the bundle. We characterise this aspect with the conditional probability of a chain being the first to break given the number of neighbours in the two-dimensional lattice of the cross-section. This is shown in Fig. 20. The outer chains, with few neighbours, are slightly more likely to be the first to break than chains inside the bundle that are surrounded by other chains. As a result, small bundles with relatively many outer chains show signs of yielding earlier than large bundles. Two possible origins of this effect have been considered, *i.e.*, the rugosity of the chain–chain interaction along the longitudinal direction, and the dependence of fluctuations of individual chains on the number of neighbours.

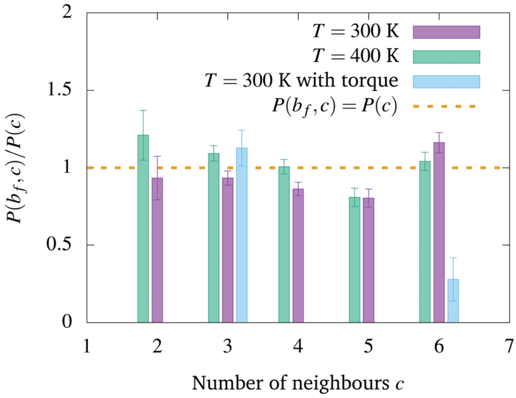


Fig. 20 The ratio of the probability for a chain to break first b_f given that is has c number of neighbouring chains to the probability for a chain to have c neighbouring chains. Two data sets with 100 samples per system with bundles of sizes from N = 7 to N = 24 at T = 300 K and T = 400 K respectively are compared to a data set of 100 samples with 7 chains with an external torque at T = 300 K. The breaking simulations were carried out with a constant strain rate of 0.8 m s^{-1}. The outer chains, with few neighbours, are slightly more likely to be the first to break than chains inside the bundle that are surrounded by other chains. The error bars show the statistical error $\sim 1/\sqrt{n}$.

The effect of the bundle size on the Young's modulus or the yield appears to be relatively small. This is likely due to the fact that covalent bonds are much stronger than the other interactions in the system, such as Coulomb and dispersion forces. In polymers, also steric interactions are important, which could affect the breaking process through the roughness of the longitudinal chain–chain interaction. We suspect that chain–chain effects would be larger in systems with the stronger intermolecular interaction compared to bonding between monomers inside the chains.

3.3.3 Winding and defects. In metals, the increase of the yield strength in small systems is related to the absence of localised and especially extended (dislocation) defects. We therefore also briefly investigate here the effect of defects and disorder on the yielding and creep of the fibre bundles.

Fig. 10 shows results for several bundles with the same number of chains but different order. One can see that the more unordered cross sections are further away from equilibrium. At the same elongation, some of the chains will not be straight, but twisted around in some way, and subjected to higher forces. Thus one would expect these bundles to break at a shorter elongation, however the force-elongation curves in Fig. 11 indicate

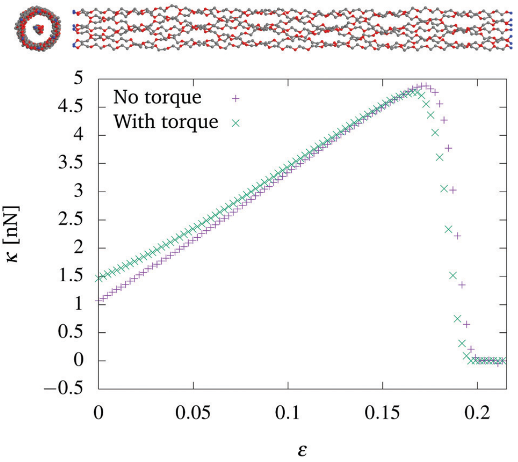


Fig. 21 Comparison of force–elongation curves for bundles with and without added external torque. The samples are stretched at a strain rate of 0.8 m s^{-1} at T = 300 K. Both curves are averaged over 100 samples of bundles with N = 7 chains. Side and front view snapshots are also included at the top.

the opposite. To investigate this further, bundles of 7 chains was deliberately twisted by applying a torque in the longitudinal direction applied to the two constrained planes. The result of these simulations is shown in Fig. 21. With the added torque, the 6 outer chains are twisted around the central chain. The cross section and side view are shown as well. All the outer chains are rotated about two times around the center. The bundle with torque has a slightly different response during the stretching and breaks significantly earlier than the untwisted bundle, as one would expect. The breaking statistics are included in Fig. 20, and indeed show that the outer chains are by far the most likely to break. The inner chain breaks first in only about 4% of cases.

3.3.4 Parameter dependencies. To asses the value and transferability of the reported findings, it is useful to explore the robustness of the chosen parameters. The role of the strain rate was briefly explored by stretching 100 samples of bundles with 7 chains at T = 300 K at a strain rate of 0.4 m s^{-1} rather than 0.8 m s^{-1}. The average strain at which the first chain breaks then decreased by 0.30 ± 0.05% compared to the values presented in Fig. 15, which could be reasonably explained by thermal activation of breaking. The estimated rate of chain breaking was similar for the two strain rates, and the estimated Young's modulus was the same within the accuracy reported here.

The choice of bond disassociation energy was briefly challenged by running the same systems with 7 chains with a 8% reduced bond disassociation energy, again with a strain rate of 0.4 m s^{-1} at T = 300 K. Lowering the disassociation energy increases the role of thermal activation of breaking, and these computations suggests that the rate of chain breaking then is closer to what was reported for T = 400 K in Fig. 19. The average strain at which the first chain in the bundles broke decreased by about 1%, while the Young's modulus again was the same within the accuracy reported here. An in-depth study on the role of the bond disassociation energy on the breaking of nanofibres could be an interesting future project.

4 Summary and conclusions

We have investigated stretching and breaking of nano-scale polymer bundles using computer simulations based on an empirical atomistic force field and on a density functional approximation. Simulated nanometrically thin bundles of PEO chains were subjected to tensile load up to their failing.

The density functional (DF) simulations have been performed for small systems, aiming at providing data to benchmark the empirical force field, and to explore features such as re-bonding after breaking that are not accounted for by the force field model. Over a broad range of strain, the DF computations reveal a complex picture. Stretching is resisted at first by torsional restoring forces, and later by bending and stretching.

Stretching of a single chain consisting of ten PEO monomers at T = 0 K provides an *ab initio* estimate of the energy *versus* strain relation. The results show that in the low-load limit the energy curve is not simply parabolic, because of the interplay of dihedral, bending and stretching terms determined by rather different force constants. At intermediate strain, where bending and stretching dominate the system response, the simulation provides a high estimate of the Young's modulus that approaches the values measured for oriented carbon protuberances and graphene.

Under conditions of constant strain, a single chain has a broad range of metastability even beyond the length of maximum restoring force, due to the absence of lateral interactions that might perturb the linear arrangement of the chain. The picture is already changed by increasing the system size to two parallel chains, which surprisingly break at a shorter length and load than the single chain. The chains are not perfectly in registry along their common longitudinal direction due to the intra-chain interaction and the helical structure of the chains, giving rise to multiple local minima in the potential energy. The proximity of the two PEO chains opens the way to chemical rebonding after breaking, resulting in a sample made of free floating epoxide monomers, and PEO segments terminated by –OH on one side, and by planar $-OCH_2$ on the other side. This partial healing might be affected by the approximate DF approach in tackling open shell species, and by the idealised setting of simulations, neglecting defects and impurities.

At the DF level, the limit of thick bundles has been investigated by stretching a system made of the experimental unit cell periodically repeated in space. In this case the stress–strain relation is more linear down to low load, and the bundle breaks at an even shorter length and load than the two-chains system. There are obviously many effects in the chemistry of the chains that are not captured by the classical force field. However, by comparing the behavior of the bond stretching and bending at high strain, we conclude that the classical force field captures sufficiently well the behavior of the single chain under high strain up to the breaking point.

To investigate larger bundles of up to 24 chains of 33 PEO monomers, and one case of 100 chains, we have used molecular-dynamics simulations. The classical force field in our simulations uses a standard functional form for the system potential energy, slightly modified in its stretching term to allow for breaking of C–C and C–O covalent bonds. The bundles are stretched out using two opposed planar structureless clamps. This geometry is greatly simplified with respect to the structure of macroscopic polymers but it is considered to provide a sufficiently realistic picture of bundles of nanometric and subnanometric diameter under tension. Moreover, it allows us to isolate and study the effect of the interaction between chains.

We have first determined the bundle geometry and structure under moderate tension at zero temperature, at ambient temperature and slightly beyond, reaching up to 400 K. Beyond a moderate load of the order of 1 nN per chain, the PEO chains tend to align in the lateral direction, although they retain in part their spontaneous helicity. When the chains are aligned like this, the cross section of the fibre shows a clear structure of a 2D cluster, that at low temperature is arranged according to a hexagonal pattern.

Bundle properties depend on the number N of chains in a non-monotonic way, and are marked by steps at discrete sizes that correspond to the filling of shells in the 2D hexagonal cluster, representing the analog of magic sizes in the physics of nanoclusters.[46,47] We obtain the strain *versus* load relation of the bundles, from which the Young's modulus, the elastic range and the limiting strength have been determined. Deviations from linear elasticity first occur upwards, corresponding to the stiffening of the bundle due to an-harmonic interactions. At 95% of the limiting load the stress–strain relation bends downwards, and the stress vanishes over a narrow strain range where the bundles break.

Finally, we have simulated the failing of the bundles under load slowly increasing towards the limiting resistance value. This last computational experiment targets creep, but given the large ratio of cohesive to thermal energies, together with the limited time covered by MD, simulations closely approach the mechanical breaking process.

The PEO bundles display a remarkably high tensile strength. The ultimate stress was computed from the MD simulations to be 36.9 ± 0.3 GPa at low temperature, 30.5 ± 0.1 GPa at 300 K and 28.4 ± 0.1 GPa at 400 K, displaying a moderate temperature dependence. The estimated Young's modulus is also high, with DF computations up to 80 GPa at low strain, and about 250 GPa at intermediate strain. The fibres in the MD computations exhibited a modulus of $Y = 144 \pm 6$ GPa at low temperature, around 130 GPa at $T = 300$ K and 125 GPa at $T = 400$ K. As discussed in Section 3.2.2, the high strength and stiffness are due to the semi-crystalline state of the fibres, resulting in samples displaying mechanical properties closer to that of crystalline PE, or even steel. The matching of helical chains that locally is an essential structural feature of extended PEO systems at vanishing or low strain is heavily dominated by the stretching, and the tensile properties are limited primarily by the covalent bonds.

Our study of the relation between the location of the first chain that breaks and the structure of the cross-section of the bundle revealed that outer chains are significantly more likely to be the first to break than chains inside the bundle that are surrounded by other chains. Applying a torque on a set of samples emphasized this effect.

We have investigated the effect of defects in the structure and deviations from the ideal perfect bundles. Defects affect the ultimate strength of the bundles and are present down to the lowest temperatures. We have identified a number of different types of defects, such as chains twisting around the elongation axis, non-optimal 2D isomers and 2D twinning of nanocrystals. At and above ambient temperature all these types of defects form and disappear in a dynamical fashion, rounding the steps at the shell closing sizes, but up to at least 400 K the system retains a high degree of ordering and a recognisable hexagonal structure. Ordering and tight packing of chains increase with increasing load.

The impact of the defects on the ultimate strength of the bundles is nontrivial. In general, non-optimal isomers appear to increase the strength somewhat. This observation is reminiscent of mechanical hardening in metals. However, the two effects are not obviously related, since hardening in metals is caused by defects blocking the movement of dislocations, while in our system dislocation dynamics does not play an obvious role in the yield.

Finally, the thermal activation of chain breaking was investigated, and a Kramer-type expression for the breaking rate was proposed and compared to simulation data. This expression was based on a highly simplified picture, but nevertheless provides a reasonable description of the data. Moreover, it gives us a mean to extrapolate breaking rates and creep to different conditions.

PEO is a relatively simple polymer, and the results obtained in this work demonstrate quite general properties. Thus, we expect similar behaviour to appear in many other more complex materials that form bundles. However, the chain–chain interactions are relative weak in PEO when compared to the bonds inside the chain. The effects of the structure and chain–chain interaction may thus be substantially bigger in bundles consisting of more strongly interacting chains.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work has been supported by the National Infrastructure for Computational Science in Norway (UNINETT Sigma2) with allocation NN9573K. EB acknowledge The Research Council of Norway for NFR project number 250158 and 262644. The authors are grateful to Prof. Pietro Ballone for his support and many interesting and illuminating discussions.

References

1 J. R. Fried, *Polymer science and technology*, Prentice Hall, 2014.

2 J. M. J. M. Berg, J. L. Tymoczko, L. Stryer and L. Stryer, *Biochemistry*, W. H. Freeman, 2002.

3 G. Strobl, *The physics of polymers: Concepts for understanding their structures and behavior*, 2007.

4 C. Bustamante, Z. Bryant and S. B. Smith, *Ten years of tension: single-molecule DNA mechanics*, 2003.

5 F. Oesterhelt, M. Rief and H. E. Gaub, *New J. Phys.*, 1999, **1**, 6.1–6.11.

6 T. Hugel, M. Rief, M. Seitz, H. E. Gaub and R. R. Netz, *Phys. Rev. Lett.*, 2005, **94**, 048301.

7 Kenry and C. T. Lim, *Prog. Polym. Sci.*, 2017, **70**, 1–17.

8 D. S. Boudreaux, *J. Polym. Sci., Part A-2*, 1973, **11**, 1285–1292.

9 H. Matsuura and T. Miyazawa, *Polym. Lett.*, 1969, **7**, 65–66.

10 T. Kurita, Y. Fukuda, M. Takahashi and Y. Sasanuma, *ACS Omega*, 2018, **3**, 4824–4835.

11 J. M. Harris, *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications*, Plenum Press, New York, 1992.

12 L. M. Bellan, J. Kameoka and H. G. Craighead, *Nanotechnology*, 2005, **16**, 1095–1099.

13 G. Linga, P. Ballone and A. Hansen, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2015, **92**, 022405.

14 X. Wu, R. J. Moon and A. Martini, *Cellulose*, 2014, **21**, 2233–2245.

15 S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1–19.

16 C. Chen, P. Depa, V. G. Sakai, J. K. Maranas, J. W. Lynn, I. Peral and J. R. D. Copley, *J. Chem. Phys.*, 2006, **124**, 234901(11).

17 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

18 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 1–31.

19 H. A. Lorentz, *Ann. Phys.*, 1881, **248**, 127–136.

20 M. K. Beyer, *J. Chem. Phys.*, 2000, **112**, 7307–7312.

21 A. C. Van Duin, S. Dasgupta, F. Lorant and W. A. Goddard, *J. Phys. Chem. A*, 2001, **105**, 9396–9409.

22 A. van Zon, B. Mos, P. Verkerk and S. de Leeuw, *Electrochim. Acta*, 2001, **46**, 1717–1721.

23 S. Neyertz, D. Brown and J. O. Thomas, *J. Chem. Phys.*, 1994, **101**, 10064.

24 P. J. Steinbach and B. R. Brooks, *Chem. Phys. Lett.*, 1994, **226**, 447–452.

25 A. Hansen, P. C. Hemmer and S. Pradhan, *The Fiber Bundle Model: Modeling Failure in Materials*, Wiley-VCH, Weinheim, 2015, pp. 1–236.

26 D. Marx and J. Hutter, *Modern Methods and Algorithms of Quantum Chemistry*, John von Neumann Institute for Computing, Jülich, 2000, vol. 3, pp. 329–477.

27 CPMD, Copyright IBM Corp 1990–2019, Copyright MPI für Festkörperforschung Stuttgart 1997–2001, http://www.cpmd.org/.

28 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

29 N. Troullier and J. L. Martins, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1991, **43**, 1993–2006.

30 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 130901.

31 Y. Takahashi and H. Tadokoro, *Macromolecules*, 1973, **581**, 5–8.

32 C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.*, 1990, **11**, 361–373.

33 T. A. C. G. Capaccio and I. M. Ward, *J. Polym. Sci.*, 1976, **14**, 1641–1658.

34 J. C. L. Hageman, R. J. Meier, M. Heinemann and R. A. de Groot, *Macromolecules*, 1997, **30**, 5953–5957.

35 M. A. Al-Nasassrah, F. Podczeck and J. M. Newton, *Eur. J. Pharm. Biopharm.*, 1998, **46**, 31–38.

36 G. Stan, F. W. Delrio, R. I. MacCuspie and R. F. Cook, *J. Phys. Chem. B*, 2012, **116**, 3138–3147.

37 A. Y. Jee, H. Lee, Y. Lee and M. Lee, *Chem. Phys.*, 2013, **422**, 246–250.

38 H. Staudinger, *Die Hochmolekularen Organischen Verbindungen – Kautschuk und Cellulose*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1932, pp. 105–123.

39 I. Sakurada and K. Kaji, *J. Polym. Sci., Part C: Polym. Symp.*, 1970, **31**, 57–76.

40 H. Matsuura and T. Miyazawa, *Bull. Chem. Soc. Jpn.*, 1968, **41**, 1798–1808.

41 Cambridge University Engineering Department, *Materials Data Book*, 2003.

42 P. J. Barham and A. Keller, *J. Mater. Sci.*, 1985, **20**, 2281–2302.

43 B. Du, J. Liu, Q. Zhang and T. He, *Polymer*, 2001, **42**, 5901–5907.

44 S. Suhai, *J. Chem. Phys.*, 1986, **84**, 5071–5076.

45 J. Yao, C. Bastiaansen and T. Peijs, *Fibers*, 2014, **2**, 158–186.

46 W. D. Knight, K. Clemenger, W. A. De Heer, W. A. Saunders, M. Y. Chou and M. L. Cohen, *Phys. Rev. Lett.*, 1984, **52**, 2141–2143.

47 M. Brack, *Rev. Mod. Phys.*, 1993, **65**, 677–732.

This journal is © The Royal Society of Chemistry 2020

Article II

E. Bering, S. Kjelstrup, D. Bedeaux, J. M. Rubi, and A. S. de Wijn

Entropy Production beyond the Thermodynamic Limit from Single-Molecule Stretching Simulations

The Journal of Physical Chemistry B **124**:40 8909–8917 (2020).
DOI: 10.1021/acs.jpcb.0c05963

Article II

This is an open access article published under a Creative Commons Attribution (CC-BY) License, which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

ACS AUTHORCHOICE

THE JOURNAL OF PHYSICAL CHEMISTRY B

pubs.acs.org/JPCB Article

Entropy Production beyond the Thermodynamic Limit from Single-Molecule Stretching Simulations

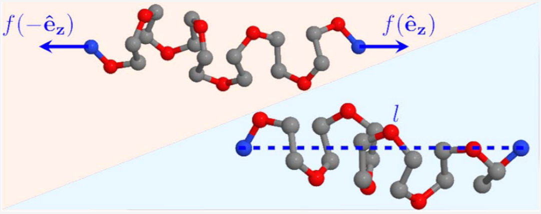
Eivind Bering,* Signe Kjelstrup,* Dick Bedeaux, J. Miguel Rubi, and Astrid S. de Wijn

Cite This: *J. Phys. Chem. B* 2020, 124, 8909−8917

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Single-molecular systems are a test bed to analyze to what extent thermodynamics applies when the size of the system is drastically reduced. Isometric and isotensional single-molecule stretching experiments and their theoretical interpretations have shown the lack of a thermodynamic limit at those scales and the nonequivalence between their corresponding statistical ensembles. This disparity between thermodynamic results obtained in both experimental protocols can also be observed in entropy production, as previous theoretical results have shown. In this work, we present results from molecular dynamics simulations of stretching of a typical polymer, polyethylene-oxide, where this framework is applied to obtain friction coefficients associated with stretching at the two different statistical ensembles for two different system sizes, from which the entropy production follows. In the smallest system, they are different up to a factor of 2, and for the bigger system, the difference is smaller, as predicted. In this way, we provide numerical evidence that a thermodynamic description is still meaningful for the case of single-molecule stretching.


$f(-\hat{\mathbf{e}}_\mathbf{z})$
$f(\hat{\mathbf{e}}_\mathbf{z})$
l

INTRODUCTION

Small systems, unlike those that are in the thermodynamic limit, do not have an extensive internal energy.[1] Because of the small number of particles, they are subjected to large fluctuations. Consequently, it becomes more challenging to obtain relations for average quantities, which are standard in thermodynamics and statistical mechanics of large systems. Gibbs thermodynamics, as we know it from standard texts,[2] ceases to apply for such systems. In view of the numerous and important applications in nanotechnology, for instance, in nanofluidics[3,4] and biology,[5] this situation poses a problem: there is a need to describe energy conversion on the small scale, but a lack of sufficient theoretical understanding. At the most extreme end of the small scale, we are not able to properly describe statistical averages for single molecules. Doubt has thus been raised on the applicability of standard thermodynamic equations to the stretching of single molecules under all conditions.[6]

In general, the energy involved in the stretching of a sufficiently small polymer depends on whether one controls the stretching length or the stretching force. The average force for isometric stretching differs from that for isotensional stretching. In the long polymer limit, they are the same, however, which has been verified experimentally, computationally, and theoretically. A very good discussion of this is given by Süzen et al.[7]

In an earlier paper,[8] some of us extended Hill's theory for thermodynamics of small systems[1] to time-dependent stretching processes, by deriving expressions for the entropy production for isometric and isotensional stretching. This leads to rate laws with friction coefficients that depended on the control variables. The aim of the present work is to calculate such friction coefficients and the corresponding entropy production using computer simulations and to verify that they depend on the control variables. This is the first example of a dynamic coefficient in molecular stretching.

We investigate the molecular stretching numerically using molecular dynamics simulations.[9] As a model, we have chosen to use a united-atom model of poly-ethylene oxide (PEO), cf. Figure 1, well-documented in the literature.[10] This molecular model has all standard modes of movement under tension, translation, rotation, torsion, and, eventually, the breaking of bonds, and lends itself to a testing of the theoretical description.

In our simulations, the stretching process can be controlled by the environment in two different ways. The endpoints of the hydrocarbon chain can be controlled by either an external force, *i.e.*, f_{ext} is a constant, or by fixing the terminal positions of the molecule, *i.e.*, l is a constant. These isometric and

Received: June 30, 2020
Revised: September 16, 2020
Published: September 16, 2020

© 2020 American Chemical Society

https://dx.doi.org/10.1021/acs.jpcb.0c05963
J. Phys. Chem. B 2020, 124, 8909−8917

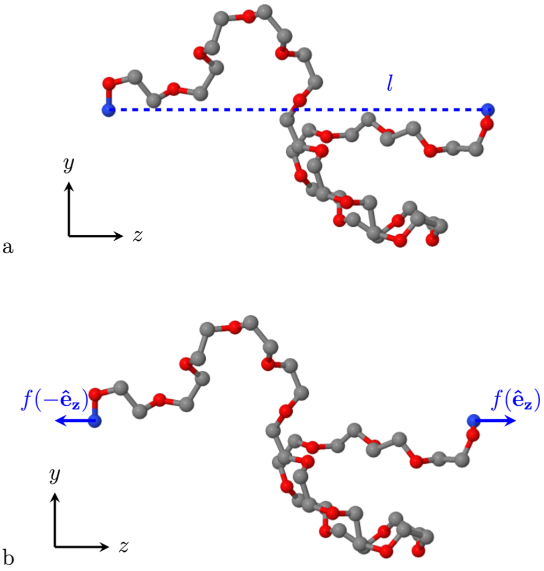

l
y
z
a
$f(-\hat{\mathbf{e}}_\mathbf{z})$
$f(\hat{\mathbf{e}}_\mathbf{z})$
y
z
b

Figure 1. Illustration of the isometric (a) and isotensional (b) simulation mode. Each monomer is composed of three beads, two methylene groups (gray), and one oxygen atom (red).

isotensional ways to operate are illustrated in Figure 1a,b. The figures show molecules that are not fully stretched.

Typically, torsional degrees of freedom are associated with lower energies and forces than bending, which in turn is associated with lower energies and forces than bond stretching. We thus expect the response to the environment to change as each of these different modes of elongating the molecule becomes accessible. From the simulation results, we shall find the appropriate dynamic description and relate the molecular properties to the dissipation.

In the thermodynamic limit, the rate laws of the two modes of operation are the same. Here, we present for the first time detailed numerical evidence that there is a difference in the dynamics in the two cases, as predicted from the method of Hill.[8]

■ THEORY

The thermodynamic basis for our numerical single-molecule stretching experiments was worked out earlier,[8] when we derived the governing equations for isometric and isotensional experiments on single molecules. In the classical thermodynamic limit, the same set of equations applies to both cases. For small systems, however, there are different sets, as each set depends on how the system is controlled by the environment.[1] An introduction to the general idea of Hill and a more extensive explanation on the structure of nano-thermodynamics can be found in a recent book.[11] In the present work, our system is always just one polymer. The length and therefore the number of monomers and the degrees of freedom vary. A bar will be used above a symbol to denote the average property of an ensemble of systems. We recapitulate the results of earlier[8] to provide a basis for the present step, how the equations can be applied to understand simulations and—in a possible next step—experimental results.

Isometric Experiments. In this experiment, we control the temperature T and the length of the molecule, l. The change in the average internal energy of a system is $\bar{U}$, given using the Gibbs equation

$$\mathrm{d}\bar{U} = T\ \mathrm{d}S + \bar{f}\ \mathrm{d}l \qquad (1)$$

where S is the system entropy and $\bar{f}$ is the average internal force working on the terminals, see Figure 1a. The average internal energy can also change by adding heat and work to the system, $\mathrm{d}\bar{U} = \mathrm{d}Q + \bar{f}_{\mathrm{ext}}\mathrm{d}l$. The length change is a result of a change in the average external force on the terminals, $\bar{f}_{\mathrm{ext}}$. By introducing these relations in eq 1, we can identify the entropy change in the surroundings by $\mathrm{d}S = \mathrm{d}Q/T$, while the average entropy production per unit of time for the system (one molecule) becomes

$$\frac{\mathrm{d}S_{\mathrm{irr}}}{\mathrm{d}t} = \frac{1}{T}(\bar{f}_{\mathrm{ext}} - \bar{f})\frac{\mathrm{d}l}{\mathrm{d}t} \qquad (2)$$

We now denote the velocity by $v \equiv \mathrm{d}l/\mathrm{d}t$ and the average change in the force by $\Delta\bar{f} \equiv \bar{f}_{\mathrm{ext}} - \bar{f}$. The rate law for the isometric case becomes

$$\Delta\bar{f} = \xi_{\mathrm{l}}(l)v \qquad (3)$$

Here, $\xi_{\mathrm{l}} = \xi_{\mathrm{l}}(l)$ is the friction coefficient specific for the length-controlled case. This is now of primary interest, one of the two coefficients we want to find.

Once we know the friction coefficient, we can compute the entropy production from eq 2, that is, $\mathrm{d}S/\mathrm{d}t = v^2\xi_{\mathrm{l}}(l)/T$. The entropy production is proportional to the friction coefficient of the length-controlled case. The entropy production is zero when the external force is balanced by the internal force, $\bar{f}_{\mathrm{ext}} = \bar{f}$.

Isotensional Experiments. In isotensional experiments, we control the temperature T and the force of the molecule, f_{ext}. The average internal energy changes as

$$\mathrm{d}\bar{U} = T\ \mathrm{d}S + f\ \mathrm{d}\bar{l} \qquad (4)$$

The length of a single molecule is now fluctuating, and $\bar{l}$ indicates its average. The first law takes the form $\mathrm{d}\bar{U} = \mathrm{d}Q + f_{\mathrm{ext}}\mathrm{d}\bar{l}$. By the same reasoning as above, we obtain the entropy production per molecule

$$\frac{\mathrm{d}S_{\mathrm{irr}}}{\mathrm{d}t} = \frac{1}{T}(f_{\mathrm{ext}} - f)\frac{\mathrm{d}\bar{l}}{\mathrm{d}t} \qquad (5)$$

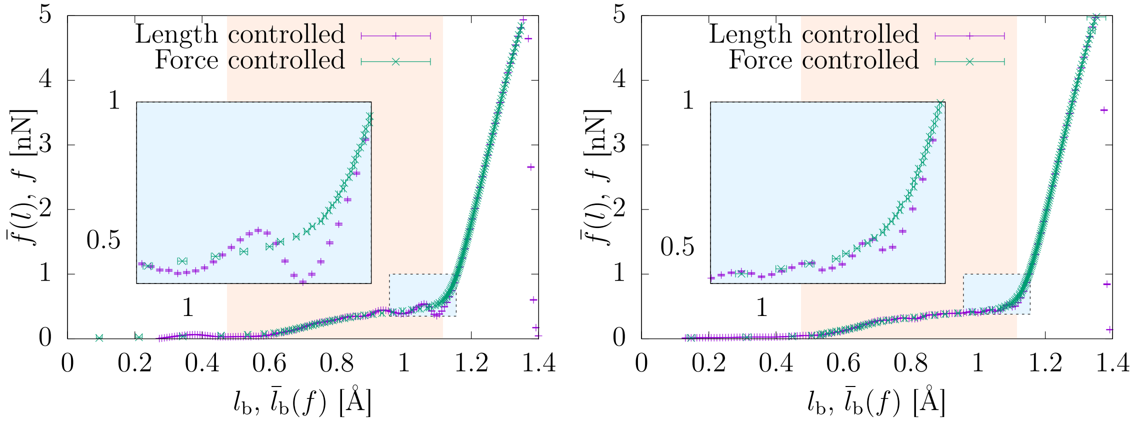
The controlled change in the force is $\Delta f = f_{\mathrm{ext}} - f$, resulting in the average stretching velocity $\bar{v} = \mathrm{d}\bar{l}/\mathrm{d}t$. The rate law in the force-controlled regime becomes

$$\Delta f = \xi_{\mathrm{f}}(f)\bar{v} \qquad (6)$$

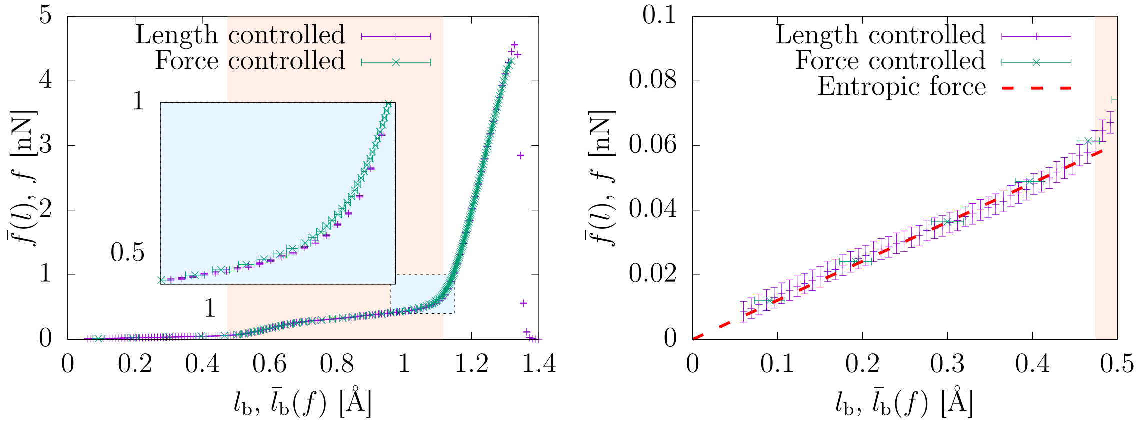
where $\xi_{\mathrm{f}} = \xi_{\mathrm{f}}(f)$ is the friction coefficient under isotensional conditions, the second target of this study. The entropy production then follows as $\mathrm{d}S/\mathrm{d}t = \bar{v}^2\xi_{\mathrm{f}}(f)/T$. The entropy production is now proportional to the friction coefficient of the force-controlled case.

In the thermodynamic limit, the two friction coefficients are the same. Away from the limit, this is not the case, as the rate laws depend on the set of the environmental control variables in use.

We shall find below that the stretching simulations of PEO with the smallest molecule under investigation gives a friction coefficient for the case of Figure 1a which is around twice the value of the coefficient for Figure 1b, confirming the prediction



(a) $N = 12$ (b) $N = 24$



(c) $N = 51$ (d) Entropic region for $N = 51$

Figure 2. Force−elongation curves from the isometric and isotensional simulations for $N = 12$ (a), $N = 24$ (b), and $N = 51$ (c) as a function of the length per bond. The region for the torsional unfolding is marked with an orange background, and the transition region to the monomer-stretching regime is shown more clearly in the insets. In (d), we see that the entropic region for $N = 51$ is well-described by a freely jointed chain with $N_{\text{eff}} = 10$ and $b_{\text{eff}} = 4$ Å.

from the theory that we can expect differences between the two coefficients.

Force in the Entropic Regime. Figure 1 illustrates the molecule for relatively small forces, when it is in the entropic regime. In this regime, the molecule behaves similarly to the thermodynamic limit because it has numerous degrees of freedom for movements.

We assume that the molecule to a good approximation can be modeled as a freely jointed chain in the entropic regime with an effective bead length b_{eff} and an effective number of beads N_{eff}, with an unfolded length $l_{\text{unf}} = N_{\text{eff}} b_{\text{eff}}$.[12] In a system with a solvent, this would correspond to an assumption of theta conditions, that is, the solvent is exactly poor enough to increase the intramolecular forces to perfectly balance out the steric effects. The statistics of the configurations of the system then becomes similar to a random walk, and the radius of gyration, $R_g = l_{\text{unf}}/6N_{\text{eff}}$, gives rise to the entropic force f_S

$$f_S = \frac{18 k_B T l}{N_{\text{eff}} b_{\text{eff}}^2} \tag{7}$$

The length b_{eff} is expected to be close to the length of each monomer.

At larger extensions, the forces will first become dominated by unfurling of the torsional degrees of freedom, then the bending, and finally the stretching of the bonds.[13] In these regimes, the force and dynamics typically display nonlinearities.

Helmholz' and Gibbs' Energies. Away from the entropic regime, we expect to be in the small-system regime. In this regime, there is a nontrivial size dependency of properties which is normally extensive. This is due to the fact that fluctuations in the different ensembles are different and lead to different size effects.

For the isometric experiments, there is a fluctuating force for each length. If we let

Table 1. Force Field Parameters for the Stretching, Bending, and Torsion,[10] with Disassociation Energies[15]

bonds		K^{s}_{ij} [kJ (mol Å^2)$^{-1}$]		D_{ij} [kJ mol^{-1}]		$\bar{r}_{ij}$ [Å]	
C−C		2587.4		370.8		1.54	
C−O		3094.0		344.5		1.43	
bends			K^{b}_{ijk} [kJ mol^{-1}]		$\bar{\theta}_{ijk}$ [Å]		
O−C−C			727.7		110.0		
C−O−C			1070.1		112.0		
torsion [kJ mol^{-1}]	$K^{\mathrm{t},1}_{ijk}$	$K^{\mathrm{t},2}_{ijk}$	$K^{\mathrm{t},3}_{ijk}$	$K^{\mathrm{t},4}_{ijk}$	$K^{\mathrm{t},5}_{ijk}$	$K^{\mathrm{t},6}_{ijk}$	$K^{\mathrm{t},7}_{ijk}$
O−C−C−O	2.211	15.194	17.844	−32.460	−13.871	−1.189	12.322
C−C−O−C	5.183	5.610	6.272	−15.428	−0.678	−4.568	3.567

$$\langle f(t)\rangle_{\mathrm{l}} = \bar{f}\,(l) \tag{8}$$

we can compute the Helmholtz energy from

$$A(l) = \int_{l_0}^{l} \bar{f}\,(l')\mathrm{d}l' \tag{9}$$

That is, the integral along the length axis of the force−elongation curves is shown in Figure 2, giving the area below the curves.

For the isotensional experiments, there is a fluctuating length for each force. If we let

$$\langle l(t)\rangle_{\mathrm{f}} = \bar{l}(f) \tag{10}$$

the Gibbs energy is given by

$$G(f) = \int_{f_0}^{f} \bar{l}\,(f')\mathrm{d}f' \tag{11}$$

That is, the integral along the force axis of the force−elongation curves is shown in Figure 2, giving the area above the curves.

In the thermodynamic limit, A and G are related by a Legendre transformation. With $\Delta l = l - l_0$ and $\Delta\bar{f} = \bar{f}(l) - \bar{f}(l_0)$, we obtain

$$A(l) + G(f = \bar{f}\,(l)) = \Delta\bar{f}\,\Delta l \tag{12}$$

for sufficiently large systems.[8] Small systems in general deviate from this, and the entropy production in the two ensembles is different. However, eq 12 is still valid when the force is linear in the elongation, like it is in the entropic regime.

The nonequivalence between the isometric and isotensional statistical ensembles is the result of the difference between the work done to stretch the molecule, $\bar{f}l$ and $f\bar{l}$, respectively. Considering the nonlinear force−elongation relationship $f = al + bl^2 + \ldots$, with a and b two constant parameters, we can easily show that up to a linear order, both works coincide. The nonlinear term, however, breaks down the equality, thus indicating the failure of the thermodynamic limit.

For the entropy production, it is useful to evaluate the expression $\bar{f}\,\frac{\mathrm{d}l}{\mathrm{d}t} - f\frac{\mathrm{d}\bar{l}}{\mathrm{d}t}$ from eqs 2 and 5, which is greater than or equal to zero in the second order of l for a specific set of lengths and velocities. From this, one would expect the entropy production for the isometric ensemble to be larger than for the intensional ensemble when the force elongation is nonlinear.

■ MODELS AND METHODS

Although the theory presented above is of general applicability, we choose a specific system for our numerical experiments: a chain of poly-ethylene oxide (PEO) of the form CH_3−[O−CH_2−CH_2]$_n$−O−CH_3, modeled with a united atom model where each carbon is grouped with its bonded hydrogen atoms. The PEO monomer consists of one oxygen and two carbons along with their hydrogens. As stated above and illustrated in Figure 1, the endpoints of the chain are controlled by either length (Figure 1a, N, l, T is controlled) or by fixing the endpoints in space (Figure 1b, N, f_{ext}, T is controlled).

The potential energy as a function of the coordinates of the coarse-grained particles has contributions from stretching, bending, and torsion. Using a model that includes these different dynamics allows us to examine the effect of the different modes of stretching and the nonlinearities on the results. The force field is compatible with the LAMMPS[14] simulation package that has been used for all of our computations.

The bond stretching is given using a Morse potential

$$U_{\mathrm{bond}}(\{\mathbf{R}_{\mathbf{i}},\ i = 1, N\}) = D_{ij}[1 - \mathrm{e}^{-\alpha_{ij}(r_{ij} - \bar{r}_{ij})}]^2 \tag{13}$$

which saturates to a finite value at large separations. The parameters used for the dissociation energies D_{ij} were obtained from density functional computations,[15] and the parameters for α_{ij} were found by requiring the Morse potential to have the same curvature as the harmonic bond, that is, $\alpha_{ij} = \sqrt{K^{\mathrm{s}}_{ij}/2D_{ij}}$. The harmonic force field parameterization is taken from van Zon et al.,[16] based on a modification of the explicit atom force field of Neyertz et al.[17] The potentials for the bending and torsion of bonds are

$$U_{\mathrm{bend}}(\{\mathbf{R}_{\mathbf{i}}\}) = \frac{1}{2}\sum_{\{ijk\}} K^{\mathrm{b}}_{ijk}[\theta_{ijk} - \bar{\theta}_{ijk}]^2 \tag{14}$$

and

$$U_{\mathrm{tors}}(\{\mathbf{R}_{\mathbf{i}}\}) = \sum_{\{ijkl\}}\sum_{\{c\}} K^{\mathrm{t},c}_{ijkl}\cos^{c-1}(\phi_{ijkl}) \tag{15}$$

where i, j, k, and l are the atoms joined by consecutive covalent bonds and K^{s}_{ij}, K^{b}_{ijk}, and K^{t}_{ijkl} and $\bar{r}_{ij}$ and $\bar{\theta}_{ijk}$ are force constants and reference values, respectively, of stretching (s), bending (b), and torsion (t) energy contributions, selected to reproduce molecular properties measured by spectroscopy or computed by *ab initio* methods. Note that the sum of the torsional coefficients includes every possible dihedral. Nonbonded interactions were not taken into account, which means that our model polymer is surrounded by an implicit theta solvent. We make this choice because an ideal chain of interacting subunits would deviate from a Gaussian chain even in the thermodynamic limit.[12] The force field parameters we used are presented in Table 1.[10,16,17]

The temperature was controlled with a Langevin thermostat, which mimics the viscous aspect of a solvent. During sampling, the relaxation time was set to 1 ps and the temperature was set

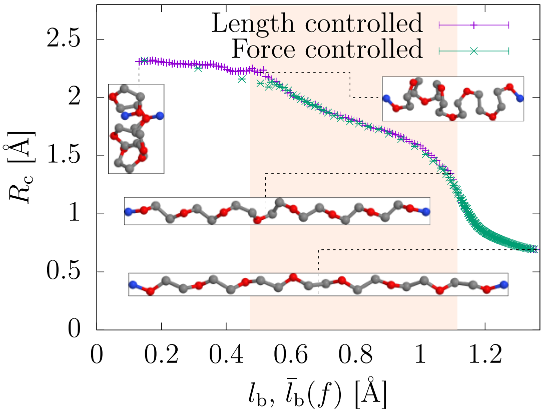


Figure 3. Cylindrical R_c as a function of the length of the molecule. Four snapshots of the molecule are provided to illustrate the different stretching regimes for a molecule of length $N = 24$. The region for the torsional unfolding is marked with an orange background, where the end of the range is found from the inflection point of the shown curve. One can see that in the first two snapshots, the molecule attains a helical ttg order, while in the last two snapshots, the molecule is in a planar all-trans configuration.

to 300 K. The time step used in the simulations was 1 fs. All quantities presented were averaged over 200 samples.

We obtain initial conditions with a low potential energy using a simulated annealing approach. After the initialization setup, all samples are heated up to 2000 K during 0.1 ns before the temperature is slowly decreased during 1 ns.

Case Studies. In the present paper, we present investigations of three different molecule sizes, $N = 12$, $N = 24$, and $N = 51$. Some simulations were also performed with $N = 102$. The forces varied from 0.01 up to 5 nN or up to the failure limit of the molecule. The length-controlled simulations were sampled evenly in the length, while the force-controlled simulations were sampled evenly on a log scale in the force. This was done to distribute the data points more evenly along the force−elongation curve. To ease the comparisons between system sizes, the molecule length will be presented in units of the longitudinal length divided by the number of bonds $l_b \equiv l/(N-1)$ and $\bar{l}_b \equiv \bar{l}/(N-1)$.

■ RESULTS AND DISCUSSION

To obtain an intuitive understanding of the behavior of the molecule during stretching, it is useful to study the cylindrical radius R_c, defined here as the radius of the smallest longitudinal cylinder that can envelop the molecule, shown for $N = 24$ in Figure 3. There is a sequence of collapses, to be elaborated on below. Four snapshots illustrate the molecular conformation in these regimes. At small lengths, we have a regime dominated by the entropic elasticity, here, the radius R_c is 2.3 Å and relatively constant. When the molecule is stretched above $l_b = 0.5$ Å, the torsional degrees of freedom are the first to be confined, and the molecule is unfolded from a helical to a planar configuration. This transition where the C−O−C−C backbone changes from a trans-gauche (ttg) order to an all-trans configuration (ttt) is elaborated in section Torsional Unfolding. This is followed by the unbending and finally the bond stretching. Especially, in regions where several types of dynamics are at play, there is a nonlinear response to stress.

Various Stretching Regimes. In the force−elongation curves shown in Figure 2 for the systems with $N = 12$ (a), $N = 24$ (b), and $N = 51$ (c), we can again identify the different regimes. The entropic regime is shown more clearly for $N = 51$, see Figure 2d, where lengths below 0.05 Å are considered to be close to zero. The data in this region are consistent with a linear curve. The range where torsion plays a role is indicated by an orange background. The nonlinear transition zone to the monomer-stretching regime is also displayed in more detail in the insets.

Entropic Regime. A predominantly linear relation between force and length develops when 0.05 Å < l_b,$\bar{l}_b$ < 0.47 Å. This is the entropic regime, for which results for $N = 51$ are enlarged in Figure 2d. From the slope of this curve, we find the effective length b_{eff} of the neighboring units of the ideal chain that gives the correct force−elongation behavior of the molecule in this regime. Within the accuracy of the data presented in Figure 2d, we see that the elongation behavior in this regime is well-described by an ideal freely jointed chain for forces up to about 0.05 nN. With a persistence length $b_{\text{eff}}/2$ of 2 Å,[18] we effectively have $N_{\text{eff}} = 10$ beads. The persistence (Kuhn) length b_{eff} corresponds to approximately twice the length of the individual monomers, explained by the bending and torsion, which effectively stiffen the chain. The force- and length-controlled cases appear identical in this regime, as the force−elongation curve here is well-described by a linear function. These findings are in line with eq 12.

Torsional Unfolding. As the molecule is stretched further, the degrees of freedom are reduced, and the freely jointed chain model is no longer applicable. The torsional degrees of freedom are the first to be confined, and this occurs in the region 0.47 Å < l_b,$\bar{l}_b$ < 1.1 Å, marked with an orange background in Figures 3 and 2. The beginning of the interval was found by looking at the deviation from linearity in Figure 2d, and the end of the interval was found from the inflection point of Figure 3. PEO strands are known to attain a helical shape in the crystalline state, in which the bonds of the C−O−C−C backbone are folded in a trans-gauche (ttg) order.[19] This can be seen in the first two snapshots in Figure 3 and is also the case for PEO dissolved in water.[20] An increase in the force gives rise to a transition from a helical ttg order to an elongated, planar all-trans configuration (ttt), as seen in the last two snapshots in Figure 3.

From Figure 2a−c, we can see a systematic deviation that varies with molecular size. This is emphasized in the insets. For

N = 12, we observe pronounced oscillations in the force−elongation curve; for N = 24, we observe smaller oscillations; and for N = 51, we observe no oscillation. These oscillations in the length-controlled ensemble are finite size effects that originates from local maxima in the potential of mean force associated with the unfolding of the molecule. Here, the molecule is mechanically unstable, and these modes are not accessible in the force-controlled ensemble.[5] This leads to different fluctuations in the two ensembles.

Monomer-Stretching Regime. As the molecule is extended above $\bar{l}_b$ > 1.1 Å, the individual monomers are elongated. The molecule is unbending, and the potentials for the stretching, bending, and torsion give rise to a molecule-specific segment elasticity,[13] increasingly dominated by the stretching of the covalent bonds.

In this region, a small systematic difference appears in the force−elongation curves between the length-controlled and the force-controlled stretching experiments. This can be seen in the inset of Figure 2a−c. The molecule is straightened out further, illustrated by the cylindrical radius in Figure 3 eventually falling to a value less than half of the shortest bond length. The nonlinear contributions in the Morse potential for the bond stretching become increasingly prominent. From the derivative of the force−elongation curve, shown in Figure 4, we observe a maximum around $\bar{l}_b$ = 1.2 Å. The probability for the bonds to rupture completely is increasing, explaining the force dropping to zero for the last points from the length-controlled simulations, as shown in Figure 2a−c.

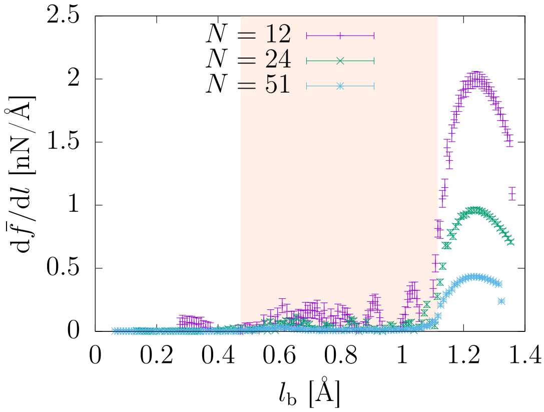


Figure 4. Derivative of the force−elongation curve from the length-controlled simulations, $\mathrm{d}\bar{f}/\mathrm{d}l$ for N = 12, N = 24, and N = 51. The region for the torsional unfolding is marked with an orange background. We see that the maximum values coincide at about l_b = 1.25 Å.

These nonlinearities from the stretching of the Morse potentials give rise to different fluctuations in the two ensembles, and we expect to see an effect of the small system size. The differences between the force−elongation curves shown in Figure 2a−c are the largest in the transition regime to the monomer-stretching regime, emphasized in the insets. The differences are small but they are finite and systematic.

Gibbs and Helmholtz Energies. The free-energy differences, and the deviation from the Legendre transform in eq 12, are computed from the force−elongation curves shown in Figure 2a−c, according to section Helmholz' and Gibbs' Energies, and shown in Figure 5. We divide by the work required to stretch the molecule completely, in order to compare the different system sizes. The largest free-energy difference is observed in the transition from the torsional-unfolding regime to the monomer-stretching regime, see the insets of the force−elongation curves in Figure 2a−c. Both in the case of N = 12 and N = 24, there is a clear correspondence between the deviations in the force−elongation curves in this region and the peak in the free-energy difference, as shown in Figure 5. There is a significant deviation from eq 12, with the smallest system showing the largest deviation, as expected.

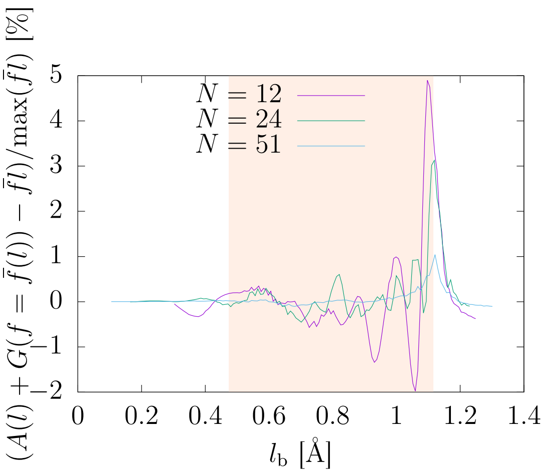


Figure 5. Percentage-wise difference in the Gibbs and Helmholtz energies for N = 12, N = 24, and N = 51 found by integration of the force−elongation curves shown in Figure 2a−c. We see that there is a significant deviation from eq 12, and the relative difference is the largest for the smallest system.

Friction Laws. *Force-Controlled Simulations.* We can now use our simulations to estimate the friction coefficient $\xi_f = \xi_f(f)$ in eq 6. This was done for the systems with N = 24 and N = 51 by perturbing the force and determining the rate of change in the average length. To this end, we first generated 200 independent samples, each equilibrated at 150 different constant forces f_0 for 5 ns. At time t = 0, the force on each of these samples was increased by 140 different force increments in the range 4−28%. The length as a function of time before and after the increase in the force is shown in Figure 6 for three force increments in the system with N = 51, averaged over 200 samples.

From these results, we find that the time scale for the initial linear force response is 0.5 ps for N = 51. As one can see in Figure 6, this does not appear to depend on the magnitude of the force increment. The ratio of the force increment to the increase in the linear response is equal within the accuracy of the data points. A similar investigation of N = 24 results in a time scale of ∼0.2 ps. The time scale for the linear regime is related to the relaxation time of the system, which depends on the length of the molecule. Other time scales in the range 0.1−1 ps was explored and was found to give similar results, although with increased fluctuations, indicating a reasonably good robustness on this parameter. Continuing with the chosen time scales, the linear response $\mathrm{d}\bar{l}/\mathrm{d}t$ was then

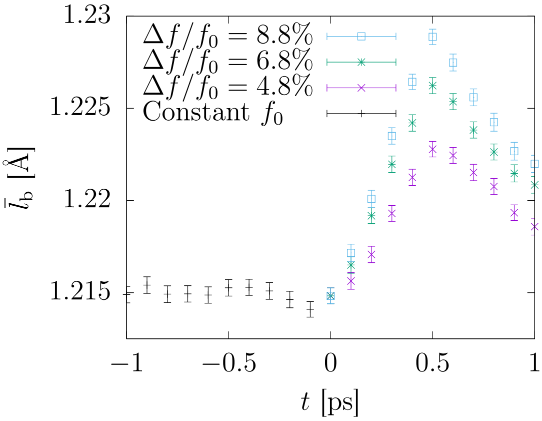
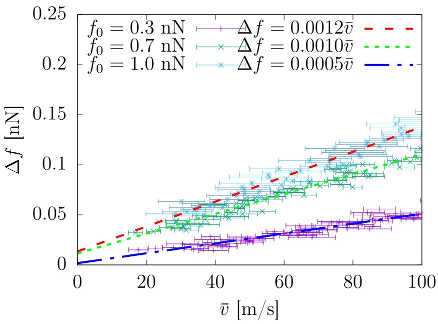
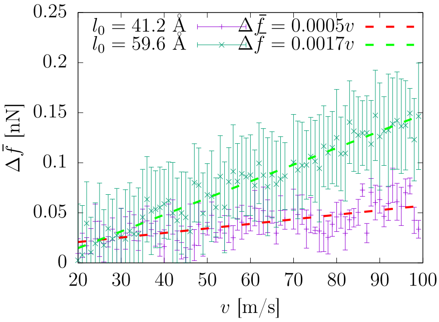


Figure 6. Length as a function of time for chains of length $N = 51$ before and after the force is increased by 4.8, 6.8, and 8.8% from $f_0 = 2.3$ nN at $t = 0$. From this, we conclude that the time scale for the linear response is ~0.5 ps for $N = 51$.

estimated for a range of force increments Δf, as shown in Figure 7 for molecules with $N = 51$ equilibrated at $f_0 = 0.33$, 0.67, and 1.00 nN. The friction coefficient $\xi_\mathrm{f} = \xi_\mathrm{f}(f)$ was found from the slope of the force−velocity curves, cf. eq 6. Unlike what is the case in the thermodynamic limit, the friction coefficient was largely dependent on the value of the force and the length of the polymer.



(a) The estimated change in the velocity $\bar{v}$ from force controlled simulations when the force is increased with Δf.



(b) The estimated change in the force $\Delta \bar{f}$ from length controlled simulations, as the samples are stretched at a constant velocity v.

Figure 7. Relation between the force and the stretching velocity, estimated in the two simulation modes, for molecules of length $N = 51$. Linear trends are observed, from which the friction coefficient is estimated.

Length-Controlled Simulations. To estimate the friction coefficient $\xi_\mathrm{l} = \xi_\mathrm{l}(l)$ in eq 3 for $N = 24$ and $N = 51$, we stretch the molecule in a range of velocities and estimate the increase in the force $\Delta \bar{f}$ associated with each stretching velocity for each sample. A total of 200 independent samples were first equilibrated at 150 different constant lengths l_0 for 5 ns, and at time $t = 0$, the samples were stretched at 80 different constant velocities $v = \mathrm{d}l/\mathrm{d}t$ in the range 20−100 m/s for 1 ps. The force response from the molecule $\Delta \bar{f}$ for each stretching velocity was then averaged over the same time scale as used for estimating the linear response in the force-controlled simulations. The resulting force−velocity curves for molecules $N = 51$ with initial lengths of $l_\mathrm{b} = 0.824$ Å and $l_\mathrm{b} = 1.192$ Å can be seen in Figure 7. Again, we found the friction coefficient $\xi_\mathrm{l} = \xi_\mathrm{l}(l)$ using eq 3 from the slope of these force−velocity curves. The variation in the coefficient with the length of the molecule or the force applied was similar to the results from the isotensional experiments, but the coefficients for force-controlled systems were systematically smaller than those for the length-controlled systems. As the fluctuations increased significantly for shorter lengths, only lengths per bonds larger than 0.4 Å are shown. Both curves showed a maximum near the relative length 1.2 Å per bond, where the Morse potential for bond stretching is strongly nonlinear.

The difference in the friction coefficient can be expected from a dynamical investigation of the system, by considering the time scales and following the approach of Just et al.[21] to obtain the general form of the effective slow dynamics. The length of the molecule acts as the slow variable, and the probability distributions of the fast variables of the internal degrees of freedom of the molecule are different for fixed force and fixed length. This also leads to two different damping constants.

Entropy Production.

The force-controlled friction coefficient $\xi_\mathrm{f} = \xi_\mathrm{f}(f) = \xi_\mathrm{f}(f(\bar{l}))$ found in the section Force-Controlled Simulations and the length-controlled friction coefficient $\xi_\mathrm{l} = \xi_\mathrm{l}(l)$ found in the section Length-Controlled Simulations are presented as a function of the length in Figure 8 for molecules $N = 24$ and $N = 51$. The difference between ξ_f and ξ_l is smaller for the largest molecule, as expected from eq 12.

The entropy production is found by multiplying this coefficient with the constant velocity squared over the temperature. The energy dissipation producing heat in the surroundings is the entropy production times the (constant) temperature. Apart from this trivial rescaling factor, the basic properties are considered to be temperature-independent under the assumption of theta conditions.

For very short lengths, the entropy production by definition should go to zero. Although the uncertainty in this region is rather high, we emphasize that zero is within the margin of error. In the region of torsional unfolding, the ensemble difference is the largest for the smaller system with $N = 24$ compared to the bigger system with $N = 51$. This is as expected from the discussion of the different stretching regimes. The entropy production reaches a maximum around 1.2 Å per bond for both system sizes, well into the monomer-stretching regime. Again, the ensemble difference is significantly larger for the smallest system. This can be explained by the nonlinearity of the Morse potential for the bond stretching, giving rise to

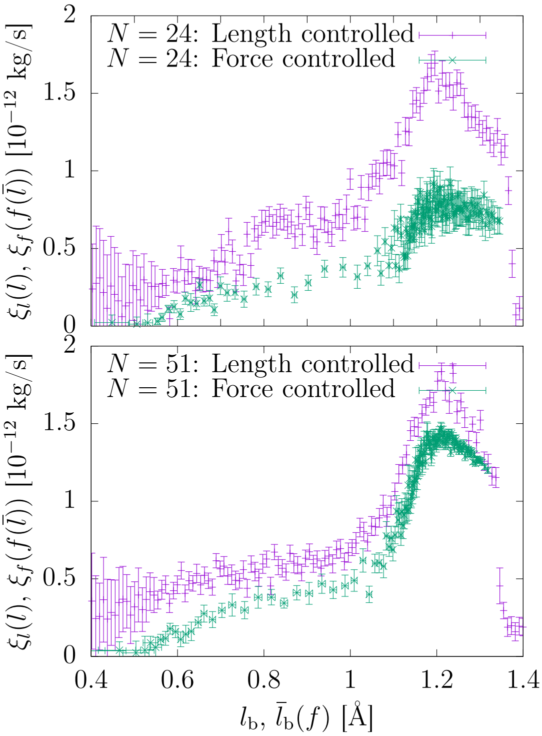


Figure 8. Estimated friction coefficient from the length-controlled and force-controlled simulations for $N = 24$ and $N = 51$. We see that the ensemble deviation is most noticeable in the monomer-stretching regime and that the ensemble deviation is more pronounced for the smallest system.

different fluctuations in the two ensembles. Comparing the derivative of the force−elongation curves presented in Figure 4, we see that the maxima appear to coincide. Moreover, any coupling to low-frequency tangential phonons can also very quickly dissipate energy in this regime.

We have seen above that the magnitude of the friction coefficient differs between the two stretching modes, with the length-controlled process having a higher friction coefficient than the force-controlled process. It follows that the first process dissipates more energy regardless of the length of the molecule, as expected. Note that the force-controlled simulations significantly display larger size dependence than what is seen in the length-controlled simulations.

CONCLUSIONS AND PERSPECTIVES

In small-scale systems, away from the thermodynamic limit, standard thermodynamics is no longer valid. In this case, thermodynamic potentials become nonextensive and statistical ensembles are not equivalent. Even if the system is very small, extensivity can be restored, if one considers the set of replicas of the original system as a large-scale system. Such a procedure, proposed by Hill,[1] makes it possible to apply the method of thermodynamics on very small scales. This method, initially proposed when the system is in equilibrium, was extended[8] to nonequilibrium situations for the case of the stretching of a polymer.

In this article, we have shown that dissipation generated at small scales is sensitive to the lack of equivalence between statistical ensembles at small scales. Based on earlier work,[8] we have carried out simulations well beyond the thermodynamic limit. We have simulated the stretching of a single PEO molecule of length $N = 12$, 24, and 51 under force-controlled and length-controlled ensembles and have extracted friction coefficients for the largest two systems.

We have confirmed systematic finite size effects in the two ensembles of general nature. In the static case, the finite size effects are most pronounced in the region of torsional unfolding and originate in local maxima in the potential of mean force that are accessible only in the length-controlled ensemble. This is visible for $N = 24$ and even more for $N = 12$. In the dynamic case, the finite size effect originates in the two ensembles having different fluctuations. This is predicted by theory and confirmed for the first time for the dynamical coefficient. For short polymers with $N = 24$, the friction coefficient of isometric stretching is roughly twice the value of that of an ensemble with isotensional stretching. The difference between the friction coefficients decreases when the length of the polymer is increased to $N = 51$.

Our study shows how nonequilibrium properties are affected by the absence of the thermodynamic limit. The method presented could be applied systematically to the study of irreversible processes that take place at small scales.

AUTHOR INFORMATION

Corresponding Authors

Eivind Bering − *Department of Physics, PoreLab, Norwegian University of Science and Technology, 7491 Trondheim, Norway;* orcid.org/0000-0003-0263-2469;
Email: eivind.bering@ntnu.no
Signe Kjelstrup − *Department of Chemistry, PoreLab, Norwegian University of Science and Technology, 7491 Trondheim, Norway*; Email: signe.kjelstrup@ntnu.no

Authors

Dick Bedeaux − *Department of Chemistry, PoreLab, Norwegian University of Science and Technology, 7491 Trondheim, Norway*
J. Miguel Rubi − *Department of Condensed Matter Physics, Universitat de Barcelona, 08028 Barcelona, Spain; Department of Physics, PoreLab, Norwegian University of Science and Technology, 7491 Trondheim, Norway*
Astrid S. de Wijn − *Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jpcb.0c05963

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are grateful to the Research Council of Norway for its Center of Excellence funding scheme, project no. 262644, PoreLab, and for the funding to EB through project no. 250158.

REFERENCES

(1) Hill, T. L. *Thermodynamics of Small Systems*; W.A. Benjamin Inc: New York, 1963; p 210.

(2) Callen, H. B. *Thermodynamics and an Introduction to Thermostatistics*; John Wiley and Sons: New York, 1985; p 512.

(3) Rauter, M. T.; Galteland, O.; Erdös, M.; Moultos, O. A.; Vlugt, T. J. H.; Schnell, S. K.; Bedeaux, D.; Kjelstrup, S. Two-phase Equilibrium Conditions in Nanopores. *Nanomaterials* **2020**, *10*, 608.

(4) Braun, E.; Chen, J. J.; Schnell, S. K.; Lin, L.-C.; Reimer, J. A.; Smit, B. Nanoporous Materials Can Tune the Critical Point of a Pure Substance. *Angew. Chem., Int. Ed.* **2015**, *54*, 14349−14352.

(5) Franco, I.; Schatz, G. C.; Ratner, M. A. Single-molecule pulling and the folding of donor-acceptor oligorotaxanes: Phenomenology and interpretation. *J. Chem. Phys.* **2009**, *131*, 124902.

(6) Keller, D.; Swigon, D.; Bustamante, C. Relating Single-Molecule Measurements to Thermodynamics. *Biophys. J.* **2003**, *84*, 733−738.

(7) Süzen, M.; Sega, M.; Holm, C. Ensemble inequivalence in single-molecule experiments. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2009**, *79*, 051118.

(8) Rubi, J. M.; Bedeaux, D.; Kjelstrup, S. Thermodynamics for Single-Molecule Stretching Experiments. *J. Phys. Chem. B* **2006**, *110*, 12733−12737.

(9) Bering, E.; de Wijn, A. S. Stretching and breaking of PEO nanofibres. A classical force field and ab initio simulation study. *Soft Matter* **2020**, *16*, 2736−2752.

(10) Chen, C.; Depa, P.; Sakai, V. G.; Maranas, J. K.; Lynn, J. W.; Peral, I.; Copley, J. R. D. A comparison of united atom, explicit atom, and coarse-grained simulation models for poly(ethylene oxide). *J. Chem. Phys.* **2006**, *124*, 234901.

(11) Bedeaux, D.; Kjelstrup, S.; Schnell, S. K. *Nanothermodynamics—General Theory*, 1st ed.; PoreLab: Trondheim, 2020.

(12) de Gennes, P.-G. *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca N.Y., 1979.

(13) Oesterhelt, F.; Rief, M.; Gaub, H. E. Single molecule force spectroscopy by AFM indicates helical structure of poly(ethylene-glycol) in water. *New J. Phys.* **1999**, *1*, 6.

(14) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.

(15) Beyer, M. K. The mechanical strength of a covalent bond calculated by density functional theory. *J. Chem. Phys.* **2000**, *112*, 7307−7312.

(16) van Zon, A.; Mos, B.; Verkerk, P.; de Leeuw, S. W. On the dynamics of PEO-NaI polymer electrolytes. *Electrochim. Acta* **2001**, *46*, 1717−1721.

(17) Neyertz, S.; Brown, D.; Thomas, J. O. Molecular dynamics simulation of crystalline poly(ethylene oxide). *J. Chem. Phys.* **1994**, *101*, 10064.

(18) Mark, J. E.; Flory, P. J. The Configuration of the Polyoxyethylene Chain. *J. Am. Chem. Soc.* **1965**, *87*, 1415−1423.

(19) Tadokoro, H. *Structure of Crystalline Polymers*; Krieger Pub. Co., 1990; p 465.

(20) Begum, R.; Hiroatsu Matsuura, a. Conformational properties of short poly(oxyethylene) chains in water studied by IR spectroscopy. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 3839−3848.

(21) Just, W.; Gelfert, K.; Baba, N.; Riegert, A.; Kantz, H. Elimination of Fast Chaotic Degrees of Freedom: On the Accuracy of the Born Approximation. *J. Stat. Phys.* **2003**, *112*, 277−292.

Article III

E. Bering, D. Bedeaux, S. Kjelstrup, A. S. de Wijn, I. Latella, and J. M. Rubi

A Legendre–Fenchel Transform for Molecular Stretching Energies

Nanomaterials **10**:12 2355-2365 (2020).
DOI: 10.3390/nano10122355

Article III

nanomaterials

Article

A Legendre–Fenchel Transform for Molecular Stretching Energies

Eivind Bering [1,*], Dick Bedeaux [2,*], Signe Kjelstrup [2], Astrid S. de Wijn [3], Ivan Latella [4] and J. Miguel Rubi [4]

[1] PoreLab, Department of Physics, Norwegian University of Science and Technology, NO–7491 Trondheim, Norway
[2] PoreLab, Department of Chemistry, Norwegian University of Science and Technology, NO–7491 Trondheim, Norway; signe.kjelstrup@ntnu.no
[3] Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, NO–7491 Trondheim, Norway; astrid.dewijn@ntnu.no
[4] Department of Condensed Matter Physics, Universitat de Barcelona, Av.Diagonal 647, 08028 Barcelona, Spain; ilatella@ub.edu (I.L.); mrubi@ub.edu (J.M.R.)
* Correspondence: eivind.bering@ntnu.no (E.B.); dick.bedeaux@ntnu.no (D.B.)

Received: 10 November 2020; Accepted: 25 November 2020; Published: 27 November 2020

Abstract: Single-molecular polymers can be used to analyze to what extent thermodynamics applies when the size of the system is drastically reduced. We have recently verified using molecular-dynamics simulations that isometric and isotensional stretching of a small polymer result in Helmholtz and Gibbs stretching energies, which are not related to a Legendre transform, as they are for sufficiently long polymers. This disparity has also been observed experimentally. Using molecular dynamics simulations of polyethylene-oxide, we document for the first time that the Helmholtz and Gibbs stretching energies can be related by a Legendre–Fenchel transform. This opens up a possibility to apply this transform to other systems which are small in Hill's sense.

Keywords: nanothermodynamics; polymers; molecular simulation; single-molecule stretching

1. Introduction

As we reduce system dimensions from the micro- to the nano-scale, surface properties become increasingly important, and the normal thermodynamic equations (thermodynamic limit properties) cease to apply. Hill [1] proposed a way to restore the structure of ordinary Gibbs' thermodynamics to deal with small systems. His idea was to introduce an ensemble of small systems, for which ordinary thermodynamics again can be applied. For an in-depth discussion, see also [2]. In Hill's description, Legendre transforms and Maxwell relations exist, but only at the level of the ensemble of small systems. A single small system, however, does not obey the normal Legendre transforms. A characteristic of small systems is that extensive properties cease to be extensive due to finite size effects, and the thermodynamic potentials depend on the type of environmental control variables, or the ensemble to which they belong. In other words, in general, statistical ensembles are not equivalent for small systems. This striking property is typically observed also in systems with size comparable with the range of the interactions [3–5]. Ensemble inequivalence in long-range interacting systems is related to the occurrence of curvature anomalies in thermodynamic potentials, which in this case arise because the interaction energy is not additive. It has been shown that Hill's approach for small systems can be implemented for long-range interacting systems as well, and that it naturally takes into account the non-additivity induced by the interactions [6,7]. Such a parallelism between small systems and systems with long-range interactions [8,9] indicates that the methods used to describe long-range interacting systems also may find a wider application in the characterization of small systems, and vice versa.

For small systems, the relative size of the fluctuations will be of more significance than for a typical large system. For sufficiently small polymers with a non-linear force-response, one would expect the difference in fluctuations to give rise to size-dependent ensemble deviations. The energy involved in stretching then depends on whether one controls the stretching length or the stretching force. The average force for isometric stretching of a small molecule differs from that of isotensional stretching. In the long polymer limit, they are the same, however, and this has been verified experimentally, computationally, and theoretically. A detailed discussion of this is given by Süzen et al. [10].

We have also studied this problem [11], and verified that the forces were not the same, as predicted from theory. This resulted in a Helmholtz energy for isometric stretching and a Gibbs energy for isotensional stretching for small molecules that were not related by a Legendre transform, which is also known from experiments by Keller et al. [12]. In addition, ensemble inequivalence has been recently highlighted in pulling experiments by Monge et al. [13].

A question therefore arises: is it then at all possible to transform the small system description from one set of variables to another set, like we normally do when we use Legendre transforms? To be more specific: is it possible to transform the Helmholtz energy of a molecule (which describes isometric stretching) into its Gibbs energy (which applies for isotensional stretching)? The aim of this short communication is to show that this is indeed possible.

We shall use our earlier simulation results [11] and verify that the Helmholtz and Gibbs energies for the stretching of a short polymer can be related to each other using the Legendre–Fenchel transform [14], a generalization of the usual Legendre transform, suitable for free energies that exhibit curvature anomalies. This transform has already proven useful in long-range interacting systems displaying ensemble inequivalence [5,15,16], and here it is applied for the first time to a common stretching phenomenon. The Legendre–Fenchel transform reduces to the usual Legendre transform when the Helmholtz energy is differentiable and convex; in the present case, this happens for large polymers. As we precisely show with our numerical simulations, the Helmoltz energy of the considered small polymers in fact present curvature anomalies under certain conditions, making it impossible to use the conventional Legendre transform.

2. Method

This section is split into two parts. The first part introduces the model and the computational details, and the second part presents the theoretical method.

2.1. Simulation Details

We use the same model as some of us have used previously [11,17] to investigate molecular stretching of poly-ethylene oxide (PEO) on the form $CH_3{-}[O{-}CH_2{-}CH_2]_n{-}O{-}CH_3$ in molecular dynamics simulations. It is a united-atom model with each bead representing either a methyl group, a methylene group or an oxygen atom. This model is based on a common model documented in the literature [18–20], and has all the standard contributions to the potential energy from bond stretching, bending, and torsion, and includes also the breaking of bonds. It therefore lends itself well to a testing of the stretching energies. In this particular force-field, the standard harmonic bond stretching potential is replaced by a Morse potential

$$U_{\mathrm{bond}}(\{\mathbf{R_i}, i = 1, N\}) = D_{ij}\left[1 - e^{-\alpha_{ij}(r_{ij} - \bar{r}_{ij})}\right]^2, \tag{1}$$

where the parameters for the dissociation energies D_{ij} are obtained from density functional computations from the literature [21]. The stiffness of the bond is determined by $\alpha_{ij} = \sqrt{K^s_{ij}/2D_{ij}}$. Furthermore, the potentials for the bending and torsion of bonds read

$$U_{\text{bend}}(\{\mathbf{R_i}\}) = \frac{1}{2}\sum_{\{ijk\}} K^b_{ijk}[\theta_{ijk} - \bar{\theta}_{ijk}]^2 \tag{2}$$

and

$$U_{\text{tors}}(\{\mathbf{R_i}\}) = \sum_{\{ijkl\}}\sum_{\{c\}} K^{t,c}_{ijkl} \cos^{c-1}(\phi_{ijkl}), \tag{3}$$

where i, j, k and l are atoms joined by consecutive covalent bonds. K^s_{ij}, K^b_{ijk}, K^t_{ijkl} are force constants for stretching (s), bending (b) and torsion (t). $\bar{r}_{ij}$ and $\bar{\theta}_{ijk}$ are equilibrium values for bond stretching and bending, respectively. All force-field parameters were tabulated previously [11]. Non-bonded interactions were not taken into account in the current work, which means that our model polymer is surrounded by an implicit theta solvent. The force field is compatible with the LAMMPS [22] simulation package, that has been used for all of our computations.

The temperature was set to 300 K during sampling, and was controlled by a Langevin thermostat with a relaxation time of 1 ps and a time step of 1 fs. The initial configurations were exposed to a simulated annealing protocol prior to sampling, in an attempt to capture a representative portion of the phase space [17,23]. The presented data are averaged over 5 ns for 200 samples.

2.2. Energy Transforms

For the theoretical analysis, consider now an arbitrary polymer with N beads. The energy of the polymer is given by

$$H(\mathbf{r}_1, ..., \mathbf{r}_N; \mathbf{p}_1, ..., \mathbf{p}_N) = \sum_{j=1}^{N} \frac{p_j^2}{2m_j} + V(\mathbf{r}_1, ..., \mathbf{r}_N)\,, \tag{4}$$

where $p_j \equiv |\mathbf{p}_j|$, m_j is the mass of bead j, and $V(\mathbf{r}_1, ..., \mathbf{r}_N)$ is the potential interaction. In our previous work [11,17], we gave an explicit expression for the interaction potential with contributions from bond stretching, bending, and torsion. The polymer is controlled either in the isometric ensemble by fixing the end-to-end distance $x \equiv |\mathbf{r}_N - \mathbf{r}_1|$, or in the isotensional ensemble by applying a stretching force $f \equiv |\mathbf{f}_N - \mathbf{f}_1|$. The canonical partition function in the isometric ensemble is

$$Z(T, N, x) = \frac{1}{\hbar^{3(N-1)}N!}\int' d\mathbf{r}_1...d\mathbf{r}_N \int' d\mathbf{p}_1...d\mathbf{p}_N \exp\left(-\beta H\right)\,, \tag{5}$$

where the end-to-end distance x is controlled, by keeping $\mathbf{r}_N - \mathbf{r}_1$ constant in the integral over the spacial coordinates. The prime for the spacial integrals indicates this. The prime for the momenta indicates that we keep the center of mass fixed. Furthermore $\hbar$ is Planck's constant and $\beta \equiv 1/(k_{\text{B}}T)$, where k_{B} is Boltzmann's constant. Because of the symmetry of the system the partition function Z depends only on x and not on the direction of $\mathbf{r}_N - \mathbf{r}_1$. The partition function for the isotensional ensemble is

$$\Delta(T, N, f) = \beta f \int_0^{x_{\max}} dx Z(T, x) \exp(\beta f x)\,, \tag{6}$$

where now the stretching force is constant, and $x_{\max}$ denotes the length of the unfolded polymer. The Helmholtz energy is given by

$$F(T, N, x) = -k_{\text{B}}T \ln Z(T, N, x)\,, \tag{7}$$

and the Gibbs energy by

$$G(T, N, f) = -k_{\text{B}}T \ln \Delta(T, N, f)\,, \tag{8}$$

in which x and f are the relevant conjugated variables as usully considered in thermodynamics and statistical mechanics of polymer systems [24]. It follows from Equation (6) that

$$\exp\left[-\beta G\left(T, N, f\right)\right] = \beta f \int_0^{x_{\max}} dx \exp\left\{-\beta\left[F(T, N, x) - fx\right]\right\}\,. \tag{9}$$

This makes it possible to calculate the Gibbs energy in the isotensional ensemble from the Helmholtz energy in the isometric ensemble. As the above derivation shows, this transformation is also correct for small polymers.

For sufficiently long polymers, the usual Legendre transform

$$G\left(T,N,f(x)\right) = F(T,N,x) - f(x)x \tag{10}$$

is valid. However, we verified in our first paper [11], using molecular dynamics simulations, that for small polymers, the usual Legendre transform is not valid.

Differences in the Helmholtz energy are calculated using

$$F(T,N,x_1) - F(T,N,x_0) = \int_{x_0}^{x_1} \bar{f}(x)dx\ . \tag{11}$$

Gibbs energy differences are calculated using

$$G(T,N,f_1) - G(T,N,f_0) = -\int_{f_0}^{f_1} \bar{x}(f)df\ . \tag{12}$$

By Equation (9), one may also find the Gibbs energy from $F(T,N,x_1)$ in Equation (11). With $x_0 = 0$ and $f_0 = 0$, Equation (9) gives

$$\exp\left\{-\beta G\left(T,N,f(x_1)\right)\right\} = \beta f(x_1) \int_0^{x_{\max}} dx \exp\left\{-\beta\left[F(T,N,x) - f(x_1)x\right]\right\}\ , \tag{13}$$

where $f(x_1)$ is obtained by means of interpolation of the isotensional force-elongation curve. From a saddle point approximation to compute the integral in Equation (13), one obtains

$$-G_{\mathrm{LF}}\left(T,N,f(x_1)\right) = \max_x\left[f(x_1)x - F(T,N,x)\right]\ . \tag{14}$$

The function $F^*(T,N,f) = -G_{\mathrm{LF}}(T,N,f)$ is known as the Legendre–Fenchel transform [5,15,16] of $F(T,N,x)$ with respect to x at constant T and N.

The Legendre–Fenchel transform is a generalization of the Legendre transform, well known in statistical physics [5,16], and reduces to the latter when the transformed function is differentiable and convex. An important property of the Legendre–Fenchel transform is that it always yields convex functions; thus $-G_{\mathrm{LF}}(f)$ is convex in f at constant T and N. Furthermore, if $F^*(f) = -G_{\mathrm{LF}}(f)$ is transformed again, one has

$$F^{**}(x) = \max_f[fx + G_{\mathrm{LF}}(f)]. \tag{15}$$

Because $-G_{\mathrm{LF}}(f)$ is a convex function, at points f for which $-G_{\mathrm{LF}}(f)$ is differentiable the above transform (15) reduces to the usual Legendre transform, leading to

$$F^{**}(x) = f(x)x + G_{\mathrm{LF}}(f(x)), \tag{16}$$

where $f(x)$ is the unique solution to $dG_{\mathrm{LF}}(f)/df = -x$. Since $F^{**}(x)$ is simply the Legendre transform of $-G_{\mathrm{LF}}(f)$, the former is the isotensional Helmholtz free energy. Moreover, due to the properties of the Legendre–Fencel transform, F^{**} is the convex envelope of the isometric free energy F, namely, the largest convex function such that $F^{**} \leq F$. Thus, the isometric and isotensional ensembles are not equivalent if F does not coincide with its convex envelope F^{**}. In mathematical terms, ensemble inequivalence may arise because the Legendre–Fenchel transform is not necessarily self-dual (or involute), that is, $F^{**} \neq F$ when F is non-convex. In contrast, the convex envelope F^{**} of F has the same Legendre–Fenchel transform as F, meaning that $(F^{**})^* = F^*$ [16].

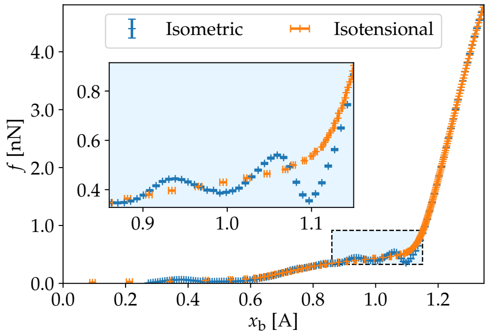
We have outlined a method for obtaining the free energy G_{LF} in the isotensional ensemble from the free energy F in the isometric ensemble. This method applies, in particular, when F is non-convex. By computing the derivative of $G_{\mathrm{LF}} = G_{\mathrm{LF}}(T, N, f(x_1))$ with respect to f, one obtains the force elongation relation

$$\frac{d}{df} G_{\mathrm{LF}} = -x(f) \tag{17}$$

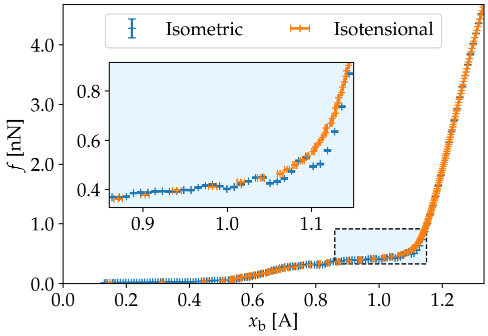
in the isotensional ensemble. The purpose of this paper is now to test these formulas.

3. Simulation Results

In the molecular dynamics simulations, one obtains the average force $\bar{f}(x) = \langle f(t) \rangle_x$ between the end points in the isometric ensemble. In the isotensional ensemble, one obtains the average distance between the endpoints $\bar{x}(f) = \langle x(t) \rangle_f$. The force-elongation curves from the isometric and isotensional ensembles are shown as a function of the length per bond $x_{\mathrm{b}} = x/(N-1)$ for systems of size $N = 12, 24$ and 51 in Figure 1a–c.

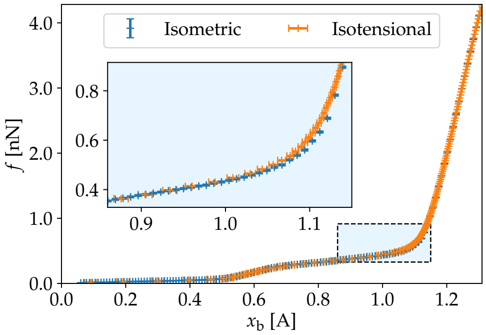


(**a**) $N = 12$



(**b**) $N = 24$

Figure 1. *Cont.*



(**c**) $N = 51$

Figure 1. Force as a function of length per bond from isometric and isotensional simulations for chains of poly-ethylene oxide (PEO) composed of $N = 12$, 24 and 51 united atoms. The ensemble inequivalence is most pronounced for the smallest systems.

It is clear from these figures that the isometric and the isotensional force are different, a fact that is more pronounced for the smaller polymers. In the isometric ensemble, the slope of the curve $\bar{f}(x)$ is not restricted to be a positive quantity, since the Helmoltz free energy $F(T, N, x)$ is not necessarily a convex function with respect to x at fixed T and N.

In other words, the response function $\kappa^{(x)}$ defined through

$$\frac{1}{\kappa^{(x)}} = \left(\frac{\partial \bar{f}}{\partial x}\right)_{T,N} = \left(\frac{\partial^2 F}{\partial x^2}\right)_{T,N} \tag{18}$$

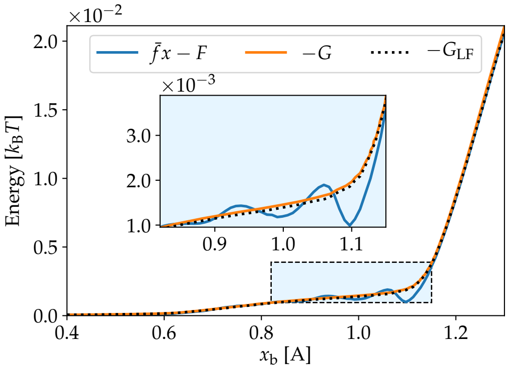
can be negative in the isometric ensemble [25], meaning that the associated system configurations minimize the free energy when the average force between the ends of the polymer decreases for increasing elongation. Under these conditions, interactions between monomers tend to separate them from each other, decreasing internal forces required to keep the polymer in equilibrium. We highlight that negative values of $\kappa^{(x)}$ in this ensemble may be realized because x is always kept fixed at a definite value. Furthermore, in the isotensional ensemble, the end-to-end distance fluctuates at constant applied force. In that case, the slope of $\bar{x}(f)$ cannot be negative, namely,

$$\kappa^{(f)} = \left(\frac{\partial \bar{x}}{\partial f}\right)_{T,N} = -\left(\frac{\partial^2 G}{\partial f^2}\right)_{T,N} \geq 0, \tag{19}$$

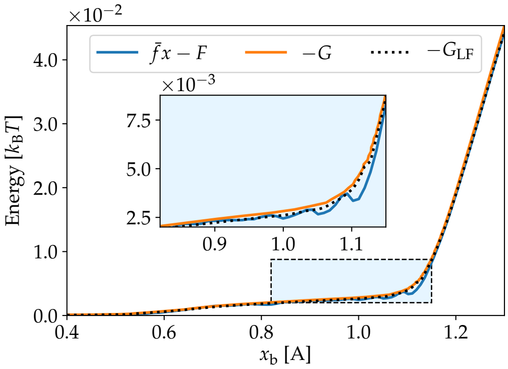
because internal forces under these conditions do not equilibrate with the external force applied on the polymer.

The points of negative slope in the isometric ensemble can be explained by the torsional unfolding of the molecule. These mechanically unstable modes are not accessible in the isotensional ensemble. As a consequence, we see that the ensemble deviation is most pronounced around $x_b = 1.1$, which marks the end of the region for torsional unfolding. This was previously discussed in great detail [11]. Prior to this region, around $x_b < 0.5$, the molecule is twisted helically, and the relation between force and elongation is predominantly linear due to entropic effects. In the last regime, with $x_b > 1.1$, the molecule is planar, and the force-elongation curve is dominated by the stretching of the individual monomers.

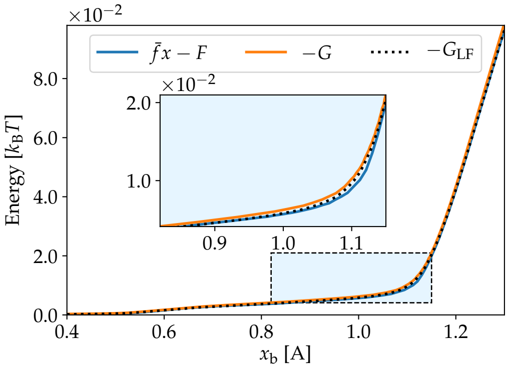
Differences in the Helmoltz energy are found from the isometric ensembles by Equation (11), and the Gibbs energy differences are found from the isotensional ensemble by Equation (12). The Legendre–Fenchel transform of the Helmoltz energy is then found by Equation (14). We present these curves for systems of size $N = 12, 24$ and 51 in Figure 2a–c. The Gibbs energy is shown with an orange line, and is compared to the Legendre transform of the Helmholtz energy in blue. It is clear that the Legendre–Fenchel transform of the Helmholtz energy, shown with a black dotted line, gives an approximation to the Gibbs energy that is far superior that of the Legendre transform. We would also like to stress that the Legendre–Fenchel transform is exact in the limit $N \to \infty$, since the saddle-point approximation is exact in this limit.



(**a**) $N = 12$



(**b**) $N = 24$

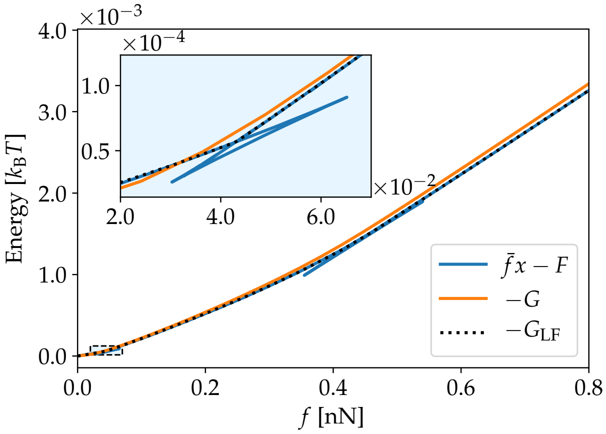


(**c**) $N = 51$

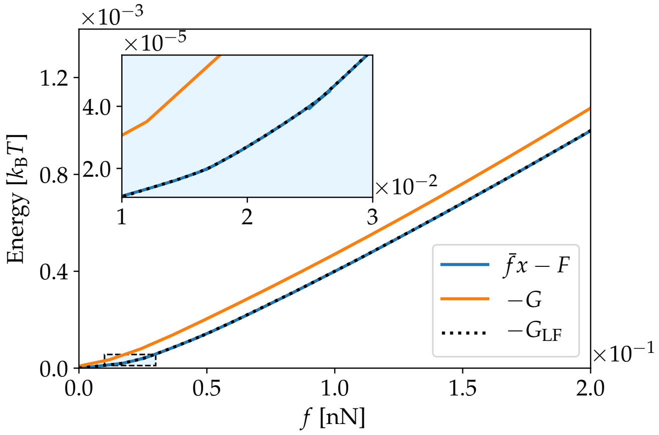
Figure 2. Energy as a function of length per bond for chains of PEO composed of $N = 12$, 24 and 51 united atoms. While the Legendre transform of the Helmholtz energy F is different from minus the Gibbs energy G, we see that the Legendre–Fenchel transform G_{LF} is an excellent approximation in all three cases.

We can see from Figure 2a–c that even for finite N, the free energy $G_{\mathrm{LF}}\,(T, N, f(x_1))$ is in excellent approximation equal to $G\,(T, N, f(x_1))$. This shows that the exact transformation, which follows from the relation between the partition function, given in Equation (13), as well as the approximate Legendre–Fenchel transform, Equation (14), can be used to obtain the Gibbs energy from the Helmholtz energy for the stretching of small polymers. The curve for $\bar{f}(x_1)x_1 - F(T, N, x_1)$ is the result of the isometric simulations and differs from the Gibbs energy curves. This shows clearly that the Legendre transform, given in Equation (10), is not valid for small polymers.

In Figure 3a,b we present the energies from Figure 2a,b as a function of force rather than elongation for systems of size $N = 12$ and 24. As the curves for the Helmholtz energy for these systems are not convex, the corresponding Legendre transformed curves as a function of force is not one-to-one. This is emphasized in the inserts.



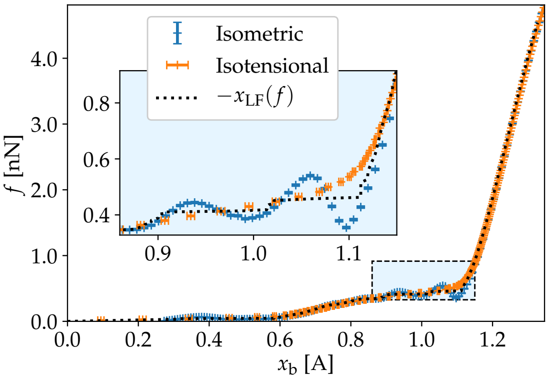
(**a**) $N = 12$



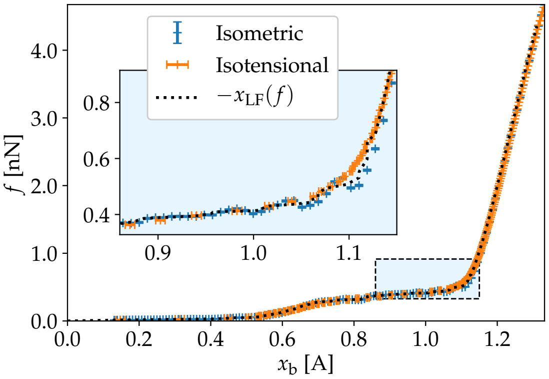
(**b**) $N = 24$

Figure 3. Energy as a function of force for chains of PEO of composed of $N = 12$ and 24 united atoms. The smallest system displays multiple singularities, one of which is emphasized in the insert. Although less pronounced, singularities can be seen also in the system with $N = 24$.

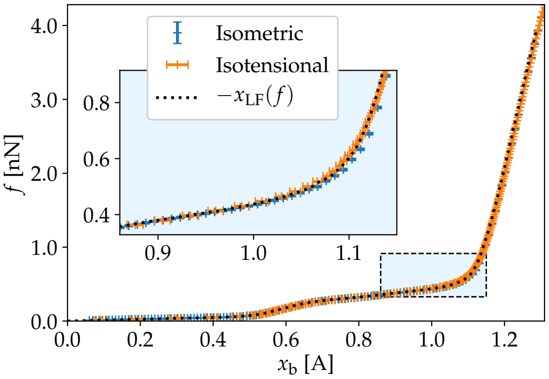
The force elongation relation for the Legendre–Fenchel transformed energy can be obtained by the derivative of G_{LF} with respect to f, cf. Equation (17). This monotonically increasing curve is shown with a black dotted line in Figure 4a–c, with the original force-elongation curves for comparison. It is clear that the Legendre–Fenchel transform is non-involutive for $N = 12$, and that it is involutive for $N = 51$, where it reduces to the Legendre transform [14].



(**a**) $N = 12$



(**b**) $N = 24$



(**c**) $N = 51$

Figure 4. The force-elongation curve x_{LF} computed from the Legendre–Fenchel transform cf. Equation (17) is compared to the force-elongation curves from Figure 1. We recognize the singular points in G_{LF} as jumps in $x_{\mathrm{LF}}(f)$, particularly visible in the smallest system with $N = 12$.

4. Discussion and Conclusions

We have analyzed the stretching of small polymers in which the thermodynamic limit cannot be invoked. We have shown that small size contributions to the isometric Helmholtz free energy induce curvature anomalies in this thermodynamic potential, which disappear as the number of beads in the polymer is increased. We described a method employing the Legendre–Fenchel transform to manage these curvature anomalies and obtain the isotensional Gibbs free energy from simulations in the isometric ensemble, in such a way that the states characterized by this free energy are unique.

The Legendre–Fenchel transform in Equation (14) reduces to the usual Legendre transform (Equation (10)) when the free energy $F(x)$ is differentiable and convex in x at constant T and N. Legendre–Fenchel transforms rather than Legendre transforms must be used in particular because $F(x)$ is non-convex [5,16]. As noted previously, the Legendre–Fenchel transform always yields convex functions and therefore, $-G_{\mathrm{LF}}(f)$ is convex in f at constant T and N. The fact that $-G_{\mathrm{LF}}(f)$ is convex ensures that the slope of the curve $\bar{x}(f)$ is non-negative, as required in equilibrium states under fluctuations of the end-to-end distance. Remarkably, this is the case when the free energy $F(x)$ presents a non-convex anomaly in the isometric ensemble. This implies a negative slope in the curve $\bar{f}(x)$. The Legendre–Fenchel transform maps the states associated with the anomaly into a point f at which $-G_{\mathrm{LF}}(f)$ is non-differentiable. This behavior is exemplified in Figure 3a,b for $N = 12$ and $N = 24$, respectively; in particular in the inserts. Such singularities are not observed for $N = 51$, as $F(x)$ in this case is convex.

We have seen above that the Legendre–Fenchel transform enables us to transform the stretching energy from the isometric to the isotensional ensemble also for small polymers. This removes the limitations set by the Legendre transforms, applicable only in the thermodynamic limit, and opens up a possibility for wider applications. The scheme documented here for molecular stretching energies reduces to the usual Legendre transforms in the thermodynamic limit.

Author Contributions: E.B. contributed software, validation and formal analysis, with A.S.d.W. as supervisor. Conceptualization was by D.B., I.L. and J.M.R. D.B., S.K. and E.B. wrote the original draft. All authors contributed to the writing—review. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Research Council of Norway grant no. 262644 (PoreLab Center of Excellence) and 250158 (yield kinetics).

Acknowledgments: The IDUN cluster at NTNU HPC Group is acknowledged for computational resources. Some computations were also performed on resources provided by UNINETT Sigma2—the National Infrastructure for High Performance Computing and Data Storage in Norway.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Hill, T.L. *Thermodynamics of Small Systems*; W.A. Benjamin Inc.: New York, NY, USA, 1963; p. 210.
2. Bedeaux, D.; Kjelstrup, S.; Schnell, S.K. *Nanothermodynamics—General Theory*, 1st ed.; PoreLab: Trondheim, Norway, 2020.
3. Campa, A.; Dauxois, T.; Ruffo, S. Statistical mechanics and dynamics of solvable models with long-range interactions. *Phys. Rep.* **2009**, *480*, 57–159. [CrossRef]
4. Bouchet, F.; Gupta, S.; Mukamel, D. Thermodynamics and dynamics of systems with long-range interactions. *Phys. A Stat. Mech. Its Appl.* **2010**, *389*, 4389–4405. [CrossRef]
5. Campa, A.; Dauxois, T.; Fanelli, D.; Ruffo, S. *Physics of Long-Range Interacting Systems*; Oxford University Press: Oxford, UK, 2014.
6. Latella, I.; Pérez-Madrid, A. Local thermodynamics and the generalized Gibbs-Duhem equation in systems with long-range interactions. *Phys. Rev. E* **2013**, *88*, 42135. [CrossRef]
7. Latella, I.; Pérez-Madrid, A.; Campa, A.; Casetti, L.; Ruffo, S. Thermodynamics of Nonadditive Systems. *Phys. Rev. Lett.* **2015**, *114*, 230601. [CrossRef]
8. Gross, D.H.E. *Microcanonical Thermodynamics*; World Scientific: Singapore, 2001.

9. Chomaz, P.; Gulminelli, F. Phase Transitions in Finite Systems. In *Dynamics and Thermodynamics of Systems with Long Range Interactions*; Dauxois, T., Ruffo, S., Arimondo, E., Wilkens, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2002; Volume 602, pp. 68–129.
10. Süzen, M.; Sega, M.; Holm, C. Ensemble inequivalence in single-molecule experiments. *Phys. Rev. E* **2009**, *79*, 051118. [CrossRef]
11. Bering, E.; Kjelstrup, S.; Bedeaux, D.; Miguel Rubi, J.; de Wijn, A.S. Entropy Production beyond the Thermodynamic Limit from Single-Molecule Stretching Simulations. *J. Phys. Chem. B* **2020**, *124*, 8909–8917. [CrossRef] [PubMed]
12. Keller, D.; Swigon, D.; Bustamante, C. Relating Single-Molecule Measurements to Thermodynamics. *Biophys. J.* **2003**, *84*, 733–738. [CrossRef]
13. Monge, A.M.; Manosas, M.; Ritort, F. Experimental test of ensemble inequivalence and the fluctuation theorem in the force ensemble in DNA pulling experiments. *Phys. Rev. E* **2018**, *98*, 032146. [CrossRef]
14. Rockafellar, T. *Convex Analysis*; Princeton University Press: Princeton, NJ, USA, 1972.
15. Ellis, R.; Haven, K.; Turkington, B. Large Deviation Principles and Complete Equivalence and Nonequivalence Results for Pure and Mixed Ensembles. *J. Stat. Phys.* **2000**, *101*, 999–1064. [CrossRef]
16. Touchette, H. The large deviation approach to statistical mechanics. *Phys. Rep.* **2009**, *478*, 1–69. [CrossRef]
17. Bering, E.; de Wijn, A.S. Stretching and breaking of PEO nanofibres. A classical force field and ab initio simulation study. *Soft Matter* **2020**, *16*, 2736–2752. [CrossRef] [PubMed]
18. Chen, C.; Depa, P.; Sakai, V.G.; Maranas, J.K.; Lynn, J.W.; Peral, I.; Copley, J.R.D. A comparison of united atom, explicit atom, and coarse-grained simulation models for poly(ethylene oxide). *J. Chem. Phys.* **2006**, *124*, 234901(11). [CrossRef] [PubMed]
19. van Zon, A.; Mos, B.; Verkerk, P.; de Leeuw, S. On the dynamics of PEO-NaI polymer electrolytes. *Electrochim. Acta* **2001**, *46*, 1717–1721. [CrossRef]
20. Neyertz, S.; Brown, D.; Thomas, J.O. Molecular dynamics simulation of crystalline poly(ethylene oxide). *J. Chem. Phys.* **1994**, *101*, 10064. [CrossRef]
21. Beyer, M.K. The mechanical strength of a covalent bond calculated by density functional theory. *J. Chem. Phys.* **2000**, *112*, 7307–7312. [CrossRef]
22. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19. [CrossRef]
23. Steinbach, P.J.; Brooks, B.R. Protein simulation below the glass-transition temperature. Dependence on cooling protocol. *Chem. Phys. Lett.* **1994**, *226*, 447–452. [CrossRef]
24. Hill, T.L. *An Introduction to Statistical Thermodynamics*; Dover Publications Inc.: New York, NY, USA, 1986.
25. Campa, A.; Casetti, L.; Latella, I.; Pérez-Madrid, A.; Ruffo, S. Concavity, Response Functions and Replica Energy. *Entropy* **2018**, *20*, 907. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Article IV

E. Bering, J. Ø. Torstensen, A. Lervik, and A. S. de Wijn

A computational study of cellulose dissolution under agitation in water and a water/NaOH/urea mixture

Draft

Article IV

This paper is awaiting publication and is not included in NTNU Open