TECHNISCHE
UNIVERSITÄT
WIEN
Vienna | Austria

Dissertation

DNA-Based Dendrimers in the Bulk

and at Interfaces

ausgeführt zum Zwecke der Erlangung

des akademischen Grades eines Doktors

der technischen Wissenschaften

unter der Leitung von

Ao.Univ.Prof. Dipl.-Ing. techn. Gerhard Kahl

am Institut für Theoretische Physik

der Technischen Universität Wien

eingereicht an der Fakultät für Physik

der Technischen Universität Wien

von

Dipl.-Ing. Clemens Jochum

Matrikelnummer 00926046

Schrotzbergstraße 6/24

1020 Wien

Wien, 8. Juni 2020 _____

Abstract

Dendrimers are synthetic macromolecules characterised by a highly branched and regular internal architecture. Recently, dendrimer-like DNAs (DL-DNAs) were synthesised *via* enzymatic ligation of Y-shaped DNA building blocks. These electrically charged dendrimers represent a novel macro-molecular aggregate, which holds high promise in bringing about targeted self-assembly of soft-matter systems in the bulk and at interfaces.

In this work, a joint simulational-experimental study of these novel macro-molecules is presented. Based on a bead-spring model for the DL-DNAs, simulations are performed in order to investigate dilute and bulk systems as well as DL-DNAs at interfaces. The results of structural quantities, *e.g.*, the radial distribution function (RDF), extracted from simulations are compared to the corresponding experimental results obtained *via* light scattering experiments.

Furthermore, we compute in the limit of vanishing density an effective, coarse-grained potential between DL-DNAs, based on Widom's particle-insertion method. With this potential at hand, investigations of large-scale systems of DL-DNAs at high concentrations become feasible.

The study of these electrically charged dendrimer systems represents a relevant field of research in the area of soft matter due to the potential role of such particles for various interdisciplinary applications, ranging from molecular cages for drug delivery to the development of dendrimer-based ultra-thin films in the area of nanotechnology.

Kurzfassung

Dendrimere sind synthetische Makromoleküle, die durch ihre stark verzweigte und reguläre interne Struktur charakterisiert werden. Sogenannte "dendrimer-like DNA" (DL-DNA) Moleküle können mittels Selbst-Assemblierung von DNA-Strängen im Labor synthetisiert werden. Diese elektrisch geladenen Dendrimere sind neuartige makromolekulare Aggregate, die vielfältige Anwendungen für gezielte Selbst-Organisation im Bereich der Weichen Materie verheißen.

Im Rahmen dieser Arbeit wird eine theoretische Untersuchung dieser neuartigen Makromoleküle präsentiert und mit experimentellen Ergebnissen ergänzt. Basierend auf einem "bead-spring" Modell für DL-DNAs, bei dem die Basenpaare als Monomere modelliert werden, werden Simulationen dieser Teilchen durchgeführt um solche Systeme bei verschiedenen Dichten sowie an Grenzflächen zu untersuchen. Anhand der Simulationsdaten werden strukturelle Größen wie zum Beispiel die radiale Verteilungsfunktion und der Strukturfaktor berechnet und mit den entsprechenden Ergebnissen von Lichtstreuexperimenten verglichen.

Weiters werden effektive Potentiale dieser Moleküle mit einem Algorithmus, der auf der "Widom insertion" Methode basiert, berechnet. Mittels dieser effektiven Potentiale können größere Systeme über längere Zeiträume simuliert werden.

Aufgrund der vielversprechenden Anwendungsmöglichkeiten in der Nanotechnologie, zum Beispiel ultra-dünne Filme und Molekülkäfige, können von Untersuchungen dieser elektrisch geladenen Dendrimere maßgeblichen Impulse in Bereichen der Nanotechnologie erwartet werden.

Für Marlis.

Preface

Before you read this thesis, a few organisational remarks are necessary. The content presented in this thesis is organised in the following page layout: while the main information of this thesis is located in the body text of the page, additional information , *e.g.*, sidenotes, references, and captions, are provided in the margin[1] of the page.

[1] This here is the margin.

The acknowledgments are located in the back of thesis.

Notation Conventions

Two important notation conventions should be mentioned: vectors are typeset in bold type, *e.g.*, vector $\boldsymbol{a}$, while matrices are typeset in uppercase roman type, *e.g.*, matrix A.

Citation Conventions

Information taken from literature is cited to the author's best knowledge and ability in close proximity to the passage where this information is used in the text. Usually, the corresponding reference is provided in a note in the margin of the page. The first time a reference is cited it is spelled out fully, whereas subsequent citations of this reference are given in an abbreviated form.

Short, verbatim quotes appear in-line highlighted by quotation marks and an appropriate reference, *e.g.*, "This is an exemplary quote",[2] while longer quotes are typeset in italic type and indicated as an individual block of text, again with an appropriate reference.

[2] See exemplary reference: J. Doe. "Article without a Title". In: *J. Obsc. Sci.* 3 (1415), pp. 92–653.

Publications and Funding

---

[3] See exemplary reference: Doe, "Article without a Title".

Everything starts somewhere, though many physicists disagree. But people have always been dimly aware of the problem with the start of things. They wonder how the snowplough driver gets to work, or how the makers of dictionaries look up the spelling of words.

—Terry Pratchett, *Hogfather*

Contents

Contents

1 Introduction and Background

What do we mean by soft matter? Americans prefer to call it "complex fluids". This is a rather ugly name, which tends to discourage the young students.

—Pierre-Gilles de Gennes[4]

In this introductory chapter a brief overview of soft matter physics (Section 1.1) and structural DNA nanotechnology (Section 1.2) will be given. These areas of research represent an overarching background for the topics presented in this thesis.

The rest of this thesis is structured as follows: the relevant theoretical concepts and numerical methods for this work are introduced in Chapter 2, while the examined DNA-based dendrimers, their synthesis, and an appropriate theoretical model are presented in Chapter 3. Furthermore, the remaining chapters are dedicated to the results: Chapter 4 for dilute systems of dendrimer-like DNAs (DL-DNAs), Chapter 5 for effective potentials of DL-DNAs, Chapter 6 for bulk systems of DL-DNAs, and Chapter 7 for two-dimensional systems of DL-DNAs. Finally, a conclusion and some outlook are presented in Chapter 8. The Appendices A, B, and C are dedicated to the simulation packages used in this work, the simulation details and paramaters used for the individual simulations, and complementary results which were deemed important enough to stay in the thesis but did not find a place in the results chapters, respectively.

1.1 Soft Matter

As the topic of this thesis belongs to the domain of soft matter, this field of research will be briefly summarised in the following.

Everyone working in soft matter physics is (at one point or the other) confronted by the same question: "What is soft matter?". I find this question surprisingly difficult to answer, given the fact that I have spent a significant amount of time studying soft matter systems. An easy way out of this conundrum can be provided by just listing all the systems which belong to soft matter and their applications in everyday life: liquid crystals and LCD technology, colloidal suspensions and paints, foams and beer, emulsions and milk, polymers and rubber, gels and cosmetic products.[5] To be exhaustive, this list would probably never end, but I still would not be able to provide a precise definition for soft matter that encompasses all these systems. Nevertheless, in the following I will try to define this field in different ways, hoping to give the reader a satisfactory notion of what soft matter is (and what it is not).

Broadly speaking, soft matter is a subfield of condensed matter. The name suggests that materials belonging to this field of research are "soft", as in "soft to the touch", in contrast to "hard" materials, such as metals. While this definition is not wrong *per se*,[6] it is difficult to apply it to certain liquids, *e.g.*, milk, and some soft matter systems can actually be quite hard, *e.g.*, glasses and plastic.

The term "soft matter" itself was popularised by Pierre-Gilles de Gennes — seen by many as the founding father of soft matter — in his Nobel prize speech in 1991.[7] According to De Gennes the name "soft matter" is a translation of the term "matière molle", which was coined by Madeleine Veyssié and became popular in his lab as a joke due to its double meaning.[8,9]

In De Gennes' opinion soft matter systems can be defined by a common property: they have large response functions, *i.e.*, they undergo significant transformations in response to outside stimuli.[10] While these stimuli can be of mechanical nature as in the definition discussed before, they can also be

[5]See M. Mitov. *Sensitive Matter: Foams, Gels, Liquid Crystals, and Other Miracles*. Trans. by G. Weiss. Harvard University Press, 2012. URL: https://worldcat.org/isbn/9780674064560.

[6]Due to its simplicity this is still a popular definition among soft matter researchers as an explanation to the non-initiated, *e.g.*, T. C. Lubensky's definition as "materials that will not hurt your hand if you hit them", see T. Lubensky. "Soft condensed matter physics". In: *Solid State Communications* 102 (1997), pp. 187–197. DOI: https://doi.org/10.1016/S0038-1098(96)00718-1, p. 187.

[7] See De Gennes, "Soft Matter (Nobel Lecture)".

[8]See De Gennes' comment in the inaugural issue of Soft Matter: P.-G. De Gennes. "Soft matter: more than words". In: *Soft Matter* 1 (2005), pp. 16–16. DOI: 10.1039/B419223K, p. 16.

[9] Due to the author's lack of French language skills this matter could not be explored further.

[10] See De Gennes, "Soft matter: more than words", p. 16.

changes by an electromagnetic field or by temperature. Thus, in De Gennes' view, the "softness" in soft matter should be interpreted as sensitivity or susceptibility to external forces — of all kinds — and not necessarily as "softness" as in "soft to the touch".

To complicate matters even further, there are soft matter systems which also fall outside of this definition *via* response functions. For example, some areas in soft matter deal with active matter, *i.e.*, a system consisting of agents which consume energy to make them move, such as bacteria.[11] Other areas deal with abstract concepts like tesselation or percolation, which *per se* are purely mathematical, but which can readily be applied to many soft matter systems.[12]

Further important properties defining some (but not all) soft matter systems are listed below:

- The system's characteristic energies, *e.g.*, bond energies, are on the scale of thermal energy $k_{\mathrm{B}}T$.[13]

- Topologies and steric constraints of the system play an important role in structure formation.[14]

- The system exhibits non-Newtonian behaviour and higher viscocities than those found in simple liquids, *e.g.*, H_2O.[15]

While the attempts for suitable definitions discussed so far and also the ones which were not discussed are clear and self-contained they usually face two problems: they do not encompass all systems generally considered to belong to soft matter and these definitions are often incompatible with each other, or, putting it differently:[16]

> . . . *any attempt to concretely define soft matter would likely be met with failure. The boundaries are nebulous, topics diverse, and the practitioners widely distributed across academic departments.*

Because such a statement might not be a satisfactory conclusion for the reader, a final attempt at providing some clarity: The braces that hold soft matter together are the methods and concepts which originate from different scientific disciplines and are used to investigate problems that are

[11] See J. Alvarado et al. "Force percolation of contractile active gels". In: *Soft Matter* 13 (2017), pp. 5624–5644. DOI: 10.1039/C7SM00834A.

[12] See Alvarado et al., "Force percolation of contractile active gels".

[13] See G. Gompper et al. "Komplexe Materialien auf mesoskopischer Skala: Was ist Weiche Materie?" In: *Physik in unserer Zeit* 34 (2003), pp. 12–18. DOI: 10.1002/piuz.200390002.

[14] See V. Vitelli and W. Irvine. "The geometry and topology of soft materials". In: *Soft Matter* 9 (2013), pp. 8086–8087. DOI: 10.1039/C3SM90111D.

[15] See G. Gompper et al. "Eine Welt zwischen Fest und Flüssig: Aktuelle Forschung an Weicher Materie". In: *Physik in unserer Zeit* 34 (2003), pp. 19–25. DOI: 10.1002/piuz.200390003.

[16] See J. L. Silverberg. "A Big Tent for Soft Matter". In: *APS News* 24.5 (May 2015). URL: https://www.aps.org/publications/apsnews/201505/ (visited on 04/20/2020).

[17] See De Gennes, "Soft matter: more than words", p. 16.

[18] *E.g.*, M. Doi. *Soft Matter Physics*. Oxford University Press, 2013. DOI: 10 . 1093 / acprof : oso / 9780199652952 . 001 . 0001 and R. A. L. Jones. *Soft Matter Physics*. Oxford University Press, 2002. URL: https://www.worldcat.org/isbn/ 9780198505891.

[19] The name dendrimer originates from the Greek word δένδρον (dendron), which translates to "tree". Initially, the name "arborol" (from *arbor*, the Latin word for tree) was also used, see G. R. Newkome et al. "Micelles. Part 1. Cascade molecules: a new approach to micelles. A [27]-arborol". In: *The Journal of Organic Chemistry* 50 (1985), pp. 2003–2004. DOI: 10.1021/jo00211a052.

[20] See E. Buhleier et al. ""Cascade"- and "Nonskid-Chain-like" Syntheses of Molecular Cavity Topologies". In: *Synthesis* 2 (1978). 155, pp. 155–158. DOI: 10.1055/s-1978-24702.

[21] See D. Astruc et al. "Dendrimers Designed for Functions: From Physical, Photophysical, and Supramolecular Properties to Applications in Sensing, Catalysis, Molecular Electronics, Photonics, and Nanomedicine". In: *Chem. Rev.* 110.4 (2010), pp. 1857–1959. DOI: 10.1021/cr900327d.

[22] See C. C. Lee et al. "Designing dendrimers for biological applications". In: *Nat. Biotechnol.* 23 (2005), pp. 1517–1526. DOI: 10.1038/ nbt1171.

[23] See D. C. Tully and J. M. J. Fréchet. "Dendrimers at surfaces and interfaces: chemistry and applications". In: *Chem. Commun.* (14 2001), pp. 1229–1239. DOI: 10 . 1039 / B104290B.

[24] See N. C. Seeman. "Structural DNA Nanotechnology". In: *Nano-Biotechnology Protocols*. Ed. by S. J. Rosenthal and D. W. Wright. Humana Press, 2005, pp. 143–166. DOI: 10.1385/1-59259-901-X:143.

[25] See N. C. Seeman. "Nucleic Acid Junctions and Lattices". In: *J. Theor. Biol.* 99.2 (1982), pp. 237–247. DOI: 10.1016/0022-5193(82)90002-9.

[26] See P. Rothemund. "Folding DNA to Create Nanoscale Shapes and Patterns". In: *Nature* 440 (2006), pp. 297–302. DOI: 10.1038/nature04586.

[27] See F. Romano and F. Sciortino. "Switching Bonds in a DNA Gel: An All-DNA Vitrimer". In: *Phys. Rev. Lett.* 114 (2015), p. 078104. DOI: 10. 1103/PhysRevLett.114.078104.

the interdisciplinary product of these otherwise distinct disciplines. To quote De Gennes:[17]

> But these byzantine discussions on names are not very important. What matters is a certain unity of concepts, and a certain creativity.

The reader is referred to relevant literature for further information.[18]

1.2 DNA-Based Materials

Dendrimers[19] are synthetic macromolecules possessing a highly branched and regular internal structure. First synthesised in the late 1970s,[20] their tree-like architecture is achieved by repeatedly attaching some basic building unit around a central core, thus generating the radially branched shells called generations.[21] Due to their potential applications in various interdisciplinary contexts, *e.g.*, drug delivery[22] and ultra-thin films,[23] the study of dendrimer systems is an important field of research in the area of soft matter.

The DNA-based dendrimers examined in this work belong together with all other synthetic all-DNA structures to the field of structural DNA nanotechnology.[24] This subfield of nanotechnology was pioneered by Nadrian C. Seeman in the early 1980s[25] when he proposed the use of DNA as a programmable building material in nanostructures. Since then this interdisciplinary research field has experienced tremendous growth giving rise to a diverse assortment of complex all-DNA nanostructures, *e.g.*, DNA origami.[26]

More recently, DNA-based self-assembly has become popular in the soft matter community. This method of self-assembly allows for fabrication of complex all-DNA constructs with tunable shapes and interactions. Many of these novel macromolecular aggregates are ideal candidates for exploring unconventional bulk phase behaviour of soft matter system such as gels[27]

and liquid crystals.[28]

The DNA-based dendrimers investigated in this work were first synthesised in 2004 by Dan Luo and coworkers.[29] These dendrimer-like DNA (DL-DNA) molecules clearly showcase the strengths of novel DNA-based nanostructures that can be assembled with remarkable control and precision using DNA as a building material. Diverse applications, *e.g.*, nano-barcodes,[30] DNA-based vaccine technologies,[31] and structural probes involving multiplexed molecular sensing processes,[32] can be realised using DNA-based dendritic architectures such as the DL-DNA molecules examined in this work.

[28] See M. Salamonczyk et al. "Smectic Phase in Suspensions of Gapped DNA Duplexes". In: *Nat. Commun.* 7 (2016), p. 13358. DOI: 10.1038/ncomms13358.

[29] See Y. Li et al. "Controlled assembly of dendrimer-like DNA". in: *Nat. Mat.* 3 (2003), pp. 38–42. DOI: 10.1038/nmat1045.

[30] See S. H. Um et al. "Dendrimer-like DNA-based Fluorescence Nano-barcodes". In: *Nat. Protoc.* 1 (2006), pp. 995–1000. DOI: 10.1038/nprot.2006.141.

[31] See Y. H. Roh et al. "Multivalent DNA-based Vectors for DNA Vaccine Delivery". In: *Methods. Mol. Biol.* 1143 (2014), pp. 159–179. DOI: 10.1007/978-1-4939-0410-5_11.

[32] See H. M. Meng et al. "DNA Dendrimer: An Efficient Nanocarrier of Functional Nucleic Acids for Intracellular Molecular Sensing". In: *ACS Nano* 8 (2014), pp. 6171–6181. DOI: 10.1021/nn5015962.

2 Theory and Methods

Physics is mathematical not because we know so much about the physical world, but because we know so little; it is only its mathematical properties that we can discover.

—Bertrand Russell, *An Outline of Philosophy*[33]

In this chapter, we provide an overview over the theoretical foundation on which this work is based. Furthermore, essential methods used in this work to produce results are presented.

A summary of statistical mechanics is given in Section 2.1. Based on statistical mechanics is the concept of Molecuar Dynamics (MD), which is the workhorse for generating results in this work and which is introduced in Section 2.2. The derivation of important structural quantities used in this work is given in Section 2.3, while Section 2.4 introduces the Ornstein-Zernike (OZ) relation and the hyper-netted chain (HNC) method, which is then used extensively in Chapter 6. In Section 2.5, we explore the possibility of using effective potential methods in order to reduce the computational cost in simulations. Finally, Section 2.6 introduces important definitions of quantities that characterise molecular shape and size.

As each section only offers a brief review of the particular topic, the reader is provided in the respective section with references to the relevant literature for more information on the topic.

[33] See B. Russell. *An Outline of Philosophy*. G. Allen & Unwin, Ltd., 1927

2.1 Statistical Mechanics in a Nutshell

In this section, a short summary of classical statistical mechanics will be presented. The fundamental idea behind statistical mechanics is the calculation of macroscopic quantities *via* statistical evaluation of the microscopic dynamics.

Whereas the equations of motion governing the microscopic dynamics are derived in Section 2.1.1, statistical ensembles, which are the workhorses of statistical mechanics, are presented in Sections 2.1.2 and 2.1.3. The reader is referred to literature[34,35] for a more in-depth review of these topics.

[34] See H. Goldstein. *Classical Mechanics*. 3rd ed. Pearson, 2002. URL: https://www.worldcat.org/isbn/9780201657029 for more information on Section 2.1.1.

[35] See M. P. Allen and D. J. Tildesley. *Computer Simulations of Liquids*. Oxford University Press, 2017. DOI: 10.1093/oso/9780198803195.001.0001 and J.-P. Hansen and I. R. McDonald. *Theory of Simple Liquids*. Academic Press, 2013. URL: http://www.worldcat.org/isbn/9780123870322 for more information on Sections 2.1.2 and 2.1.3.

2.1.1 Lagrangian and Hamiltonian Mechanics

The full information about the dynamics of a classical, three-dimensional system consisting of N particles and interacting *via* potential $\mathcal{V}$ is contained in its Lagrangian, $\mathcal{L}$, which in the most general case is defined as[36]

[36] See Chapter 2 in Goldstein, *Classical Mechanics*

$$\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = \mathcal{T}(\dot{\boldsymbol{q}}) - \mathcal{V}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \,, \tag{2.1}$$

$$\text{with} \quad \mathcal{T}(\dot{\boldsymbol{q}}) = \frac{1}{2} \sum_{i=1}^{3N} m \dot{q}_i^2 \,. \tag{2.2}$$

Here, $\mathcal{T}(\dot{\boldsymbol{q}})$ denotes the kinetic energy and we assume for the moment a time-dependent, non-conservative potential $\mathcal{V}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$,[37] which is a function of the generalised coordinates of the system, $\boldsymbol{q}$,[38] and their respective time derivatives, $\dot{\boldsymbol{q}}$. Then, the equation of motion for each generalised coordinate q_i of the system can be written down using the Euler-Lagrange equation expressed *via*

[37] *E.g.*, charged particles moving in a time-dependent, external electric field. Note, however, that in this work the interaction potential will be conservative and time-independent: $\mathcal{V} = \mathcal{V}(\boldsymbol{q})$.

[38] In a three-dimensional system containing N point-like particles there are $3N$ generalised coordinates, each corresponding to one spatial coordinate of a single particle.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) - \left(\frac{\partial \mathcal{L}}{\partial q_i}\right) = 0 \quad \text{with} \quad i = 1, \ldots, N \,. \tag{2.3}$$

In its essence, Equation (2.3) is equivalent to Newton's equations of motion. An alternative formulation can be achieved using the Hamiltonian

formalism with the Hamiltonian $\mathcal{H}$ defined as[39]

$$\mathcal{H}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = \sum_i \dot{q}_i p_i - \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \,, \tag{2.4}$$

where p_i is the generalised momentum conjugate to q_i and is given by

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \,. \tag{2.5}$$

The equations of motion are then defined *via*

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \,, \tag{2.6}$$

which again are equivalent to Newton's equations of motion.

For a conservative, time-independent potential ($\mathcal{V} = \mathcal{V}(\boldsymbol{q})$) Equation (2.4) reduces to

$$\mathcal{H}(\boldsymbol{q}, \boldsymbol{p}) = \mathcal{T}(\boldsymbol{p}) + \mathcal{V}(\boldsymbol{q}) \,, \tag{2.7}$$

which corresponds to the total energy of the system, and the equations of motion of Equation (2.6) in Cartesian coordinates reduce to[40]

$$\boldsymbol{p}_n = \dot{\boldsymbol{r}}_n m_n \quad \text{and} \quad \dot{\boldsymbol{p}}_n = -\nabla_{\boldsymbol{r}_n} \mathcal{V} = \boldsymbol{f}_n \,, \tag{2.8}$$

where $\boldsymbol{r}_n$, $\boldsymbol{p}_n$, and $\boldsymbol{f}_n$ denote the position, momentum, and force of particle n, respectively.

Finally, we can write down the total time derivative with respect to time t for a Hamiltonian $\mathcal{H}$ with a conservative, time-independent potential $\mathcal{V}(\boldsymbol{q})$:

$$\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} = \sum_k \left(\frac{\partial \mathcal{H}}{\partial q_i} \dot{q}_i + \frac{\partial \mathcal{H}}{\partial p_i} \dot{p}_i \right) = 0 \,. \tag{2.9}$$

Equation (2.9) reflects one of the fundamental laws of conservation in physics: the conservation of energy.

[39] See Chapter 8 in Goldstein, *Classical Mechanics*

[40] See Allen and Tildesley, *Computer Simulations of Liquids*, p. 96.

2.1.2 Statistical Ensembles

A macroscopic system of $N \sim 10^{23}$ particles will evolve in time and follow a trajectory $\Gamma(t)$ in $6N$-dimensional phase space, where the Cartesian coordinates, q, and the components of momentum, p, each represent $3N$ dimensions. Starting from an initial condition, the dynamics of the system evolution in time is uniquely defined and is governed by the set of $6N$ coupled differential equations given by Equations (2.6) or (2.8). Obviously, even with modern computing power it has proved intractable to numerically solve these equations of motion for a macroscopic system of $N \sim 10^{23}$ particles, not to mention the impossibility of an analytical solution.

An alternative to explicitly solving the equations of motions is given by statistical ensembles: instead of tracking the phase space trajectory $\Gamma(t)$ of a single system, we consider an ensemble, *i.e.,* a multitude[41] of virtual copies of the same system, all of which are compatible with a set of fixed macroscopic parameters, *e.g.,* the number of particles, N, the system volume, V, and the system temperature, T, in the canonical ensemble. Put differently, we assemble all individual phase space states, *i.e.,* the microstates, which are available to the system and consistent with the set of fixed macroscopic parameters into the statistical ensemble.

[41] Depending on the system this number can be infinitely large.

We can then define the probability density or phase space density, $\rho_{\mathrm{ph}}(\Gamma, t)$, where $\rho_{\mathrm{ph}}(\Gamma, t)\,\mathrm{d}\Gamma\mathrm{d}t$ is the probability that at time t the system's phase state is located at position Γ in phase space. This probability density $\rho_{\mathrm{ph}}(\Gamma, t)$ must satisfy

$$\rho_{\mathrm{ph}}(\Gamma, t) \geq 0 \quad \forall\, \Gamma, t \quad \text{and} \tag{2.10}$$

$$\int \rho_{\mathrm{ph}}(\Gamma, t)\,\mathrm{d}\Gamma = 1 \,, \tag{2.11}$$

where the integration is carried out over the full available phase space.[42] The explicit form of $\rho_{\mathrm{ph}}(\Gamma, t)$ depends on the given macrostate, *i.e.,* the particular ensemble.

[42] In this work we only consider continuous systems. For discrete systems the integral in Equation (2.11) is replaced by the corresponding sum.

In equilibrium, function $\rho(\Gamma)$ is independent of time: if all virtual realisations in such an ensemble continue to evolve on their trajectories in time,

their distribution in phase space does not change, *i.e.*,[43]

$$\frac{\mathrm{d}\rho_{\mathrm{ph}}}{\mathrm{d}t} = 0 \,. \tag{2.12}$$

In order to obtain a meaningful result for some macroscopic observable $\mathcal{A}(\boldsymbol{q}, \boldsymbol{p}) = \mathcal{A}(\Gamma)$, which depends on the system's microstate in phase space, Γ, we can take the ensemble average[44]

$$\langle \mathcal{A} \rangle_{\mathrm{e}} = \int \mathcal{A}(\Gamma) \rho_{\mathrm{ph}}(\Gamma) \, \mathrm{d}\Gamma \,, \tag{2.13}$$

$$\text{with} \quad \int \rho_{\mathrm{ph}}(\Gamma) \, \mathrm{d}\Gamma = 1 \,. \tag{2.14}$$

The integrals in Equations (2.13) and (2.14) are calculated over the $6N$-dimensional phase space.

In the following subsections we will present different types of ensembles and their respective probability distributions $\rho_{\mathrm{ph}}(\Gamma)$.

2.1.3 The Canonical Ensemble

In this work, only the canonical ensemble will be of interest. The fixed, macroscopic parameters in this ensemble are the number of particles, N, the system volume, V, and the system temperature, T; therefore it is also referred to as the NVT ensemble and it can be imagined as a system surrounded by a heat bath of fixed temperature T.

It can be shown that the probability distribution of this ensemble in phase space is given by the Boltzmann distribution:[45]

$$\rho_{\mathrm{ph}}(\boldsymbol{q}, \boldsymbol{p}) \propto e^{-\beta \mathcal{H}(\boldsymbol{q}, \boldsymbol{p})} \,, \tag{2.15}$$

with $\beta = 1/k_{\mathrm{B}}T$. Normalising the distribution with respect to the whole phase space while accounting for the possible indistinguishability of the particles and the phase space volume h^{3N} occupied by each individual

microstate[46] yields

$$\rho_{\mathrm{ph}}(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{N!} \frac{1}{h^{3N}} \frac{e^{-\beta \mathcal{H}(\boldsymbol{q},\boldsymbol{p})}}{Z_{NVT}} \,, \tag{2.16}$$

$$\text{with} \quad Z_{NVT} = \frac{1}{N!} \frac{1}{h^{3N}} \int e^{-\beta \mathcal{H}(\boldsymbol{q},\boldsymbol{p})} \, \mathrm{d}\boldsymbol{q}\,\mathrm{d}\boldsymbol{p} \,, \tag{2.17}$$

where the canonical partition function Z_{NVT} is calculated by integrating over the whole phase space. Then, the ensemble average of observable $\mathcal{A}(\boldsymbol{q}, \boldsymbol{p})$ in the canonical ensemble can be obtained *via* substitution of Equation (2.16) into Equation (2.13):

$$\langle \mathcal{A} \rangle_{\mathrm{e}} = \frac{1}{N!} \frac{1}{h^{3N}} \int \mathcal{A}(\boldsymbol{q}, \boldsymbol{p}) \frac{e^{-\beta \mathcal{H}(\boldsymbol{q},\boldsymbol{p})}}{Z_{NVT}} \, \mathrm{d}\boldsymbol{q}\,\mathrm{d}\boldsymbol{p} \,. \tag{2.18}$$

Using Equation (2.18) we can now establish a relation between statistical mechanics and thermodynamics by calculating an appropriate thermodynamic potential. For the canonical ensemble this thermodynamic potential is the Helmholtz free energy, F, which is defined as

$$F = F(N, V, T) = E - TS \,, \tag{2.19}$$

where E and S are the internal energy and the entropy of the system.

It can be shown that in the thermodynamic limit, *i.e.*, $N \to \infty$, $V \to \infty$, and $N/V = \mathrm{const.}$, E and S can be calculated *via*[47]

$$E = \lim_{\substack{N,V \to \infty \\ N/V = \mathrm{c.}}} \langle E \rangle_{\mathrm{e}} = -\frac{\partial}{\partial \beta} \ln (Z_{NVT}) = -\frac{1}{Z_{NVT}} \frac{\partial Z_{NVT}}{\partial \beta} \,, \tag{2.20}$$

$$\text{and} \quad S = \lim_{\substack{N,V \to \infty \\ N/V = \mathrm{c.}}} \langle S \rangle_{\mathrm{e}} = -\frac{1}{T} \frac{\partial}{\partial \beta} \left(\ln (Z_{NVT}) \right) + k_{\mathrm{B}} \ln (Z_{NVT}) \,. \tag{2.21}$$

Finally, we have arrived at an elegant equation linking the canonical partition function Z_{NVT}, which counts the number of microscopic states in the ensemble, with the macroscopic thermodynamic potential F:

$$F = E - TS = -k_{\mathrm{B}} T \ln (Z_{NVT}) \,. \tag{2.22}$$

Other frequently used types of ensembles which are not used in this work include

- the microcanonical (constant NVE) ensemble,

- the isobaric-isothermal (constant NPT) ensemble,

- and the grand-harmonic-isothermal (constant μVT) ensemble,

with pressure P and chemical potential μ. The reader is referred to the literature[48] for more information on these ensembles.

[48] *E.g.*, K. Huang. *Statistical Mechanics*. 2nd ed. New York: Wiley, 1987. URL: https://www.worldcat.org/isbn/9780471815181, B. Widom. *Statistical Mechanics*. Cambridge: Cambridge University Press, 2002. DOI: 10.1017/CBO9780511815836, and F. Schwabl. *Statistische Mechanik*. Berlin: Springer-Verlag, 2006. DOI: 10.1007/3-540-31097-5.

2.2 Molecular Dynamics in a Nutshell

This section gives a brief overview over MD simulations. While the basic principles are explained in Section (2.2.1), some techniques to implement certain aspects of MD simulations and tricks to improve their efficiency are given in Section (2.2.2). The reader is referred to the literature[49] for more information on MD simulations.

[49]See for example D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. 2nd ed. Cambridge University Press, 2004. DOI: 10.1017/CBO9780511816581 and D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press, 2001. DOI: 10.1016/B978-0-12-267351-1.X5000-7.

2.2.1 Principles of Molecular Dynamics

Ultimately, the goal of performing large-scale simulations of physical systems is to gain insight into the system and calculate quantities of interest, *e.g.*, pressure P or heat-capacity C_V, usually with the intention of comparing the results to experimental measurements. While in simulations we have access to the instantaneous microstate of the system, it is reasonable to assume that the experimental measurement of a "macroscopic" observable $\mathcal{A}$ does not correspond to the instantaneous microstate but rather to a time-averaged value over some macroscopic time interval t_{obs}.

In a system of N particles the microstate of the system follows a trajectory, $\Gamma(t)$, in $6N$-dimensional phase space. Then, the instantaneous value of $\mathcal{A}$ at time t is defined by the particular microstate: $\mathcal{A}(t) = \mathcal{A}\left(\Gamma(t)\right)$. Assuming that the observation time t_{obs} is large compared to the characteristic time scale of the system, *e.g.*, some relaxation time $\tau_{\mathrm{rel}} \ll t_{\mathrm{obs}}$, the experimentally

[50] Errors in measurement are ignored in this idealised thought experiment. See Allen and Tildesley, *Computer Simulations of Liquids*, pp. 46–49.

oberved value, $\mathcal{A}_{\text{obs}}$, can be written as[50]

$$\mathcal{A}_{\text{obs}} = \lim_{t_{\text{obs}} \to \infty} \frac{1}{t_{\text{obs}}} \int_{t_{\text{i}}}^{t_{\text{f}}} \mathcal{A}\left(\Gamma(t)\right) \, \mathrm{d}t = \langle \mathcal{A} \rangle_{\text{t}} \; , \tag{2.23}$$

where $t_{\text{obs}} = t_{\text{f}} - t_{\text{i}}$, with t_{i} and t_{f} as the start and end time of the measurement. Here, the time average is denoted as $\langle \dots \rangle_{\text{t}}$ in order to distinguish it from the ensemble average $\langle \dots \rangle_{\text{e}}$ of Equation (2.13).

By numerically integrating the equations of motion in time in a discrete step-by-step fashion and calculating observable $\mathcal{A}(t)$ at regular intervals Δt, we can approximate the right hand side in Equation (2.23) *via*

$$\langle \mathcal{A} \rangle_{\text{t}} = \frac{1}{N_{\text{s}}} \sum_{k=0}^{N_{\text{s}}} \mathcal{A}\left(\Gamma(k\,\Delta t)\right) \; , \tag{2.24}$$

[51] Using modern high-performance computing (HPC) methods MD simulations with particle numbers of the order $N \sim 10^8$ are already achievable, see C. Mei et al. "Enabling and scaling biomolecular simulations of 100 million atoms on petascale machines wit, a multicore-optimized message-driven runtime". In: *SC '11: Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis.* 2011, pp. 1–11.

with the number of samples $N_{\text{s}} = t_{\text{sim}}/\Delta t$ and simulation time t_{sim}. Whereas the limited computing power at our disposal is by far not able to simulate macroscopic systems ($N \sim 10^{23}$) at comparably macroscopic time scales ($t_{\text{sim}} \to \infty$), large-scale simulations with long (but still finite) simulation times are nevertheless possible[51,52] and yield satisfactory results that are comparable to results obtained *via* experiment or other analytic or *in silico* means.

[52] See Section 2.2.2 for technical details of MD simulations.

So far, an important question has still remained open: are time average $\langle \mathcal{A} \rangle_{\text{t}}$ in Equation (2.23) and ensemble average $\langle \mathcal{A} \rangle_{\text{e}}$ in Equation (2.13) equal? Put differently, will a system evolving in time explore every point Γ in phase space with non-zero probability[53] $\rho_{\text{ph}}(\Gamma)$ in the thermodynamic limit, *i.e.*, $t_{\text{obs}} \to \infty$? While it can not be proven explicitly for every system, it can be assumed that the ergodic hypothesis, which can be written as[54]

[53] According to the system's statistical ensemble.

[54] See Frenkel and Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, pp. 15–17.

$$\langle \mathcal{A} \rangle_{\text{e}} \overset{!}{=} \langle \mathcal{A} \rangle_{\text{t}} \; , \tag{2.25}$$

[55] Exceptions are systems which exhibit ergodiciy breaking. Often, these are systems with spontaneous symmetry breaking.

holds true.[55] Ergodicity implies that a system evolving in time will reach every physically possible state in phase space as $t_{\text{obs}} \to \infty$. A weaker form of this property, quasi-ergodicity, means that under the same conditions the system's trajectory will come arbitrarily close to each physically available

point in phase space. The conclusion for MD simulations is that the simulation trajectory has to sample a representative subset of the phase space in order to produce meaningful results for macroscopic observables.

2.2.2 Tricks of the Trade

Modern MD simulations employ a host of numerical algorithms and methods in order to improve the efficiency of the simulation and achieve larger system sizes and longer simulation times. The following list (incomplete and in no particular order) summarises some of the most important techniques applied in this thesis. While some of these methods are only applicable in MD simulations, others can also be used in MC simulations.

- **Numerical integration:** most modern simulation methods use some sort of finite-difference calculation in order to approximate the derivatives with respect to time appearing in Newton's equations of motion.[56] Two basic algorithms to achieve this goal of numerical integration in time are leapfrog integration and (velocity) Verlet integration.[57]

- **Warm-up period:** in order to ensure that the initial state of the system does not correspond to an undesired or unphysical configuration which could result in an unstable simulation (*e.g.*, due to exceedingly large forces), a warm-up process with artificial restrictions to particle interactions and movement can be performed before starting the actual simulation.

- **Equilibration period**: in most cases, the initial state will not correspond to the equilibrium state of the system. While one could start from a constructed initial configuration which is at equilibrium, it is often simpler to start with a more or less arbitrary configuration and wait until the system has equilibrated, *i.e.*, the system retains no memory of the initial state.[58] In practice, this corresponds to discarding the first few percent ($\approx 10\%$) of the simulation time in the analysis.

- **Thermostat:** simulations of ensembles with a heat bath, *e.g.*, NVT ensembles, must ensure that temperature T is held constant. Such a

[56] See Rapaport, *The Art of Molecular Dynamics Simulation*, pp. 60–62.

[57] See L. Verlet. "Computer "Experiments" on Classical Fluids". In: *Phys. Rev.* 159 (1967), pp. 98–103. DOI: 10.1103/PhysRev.159.98.

[58] See Rapaport, *The Art of Molecular Dynamics Simulation*, p. 18.

[59] See Frenkel and Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, pp. 140–158.

[60] See H. J. C. Berendsen et al. "Molecular dynamics with coupling to an external bath". In: *J. Chem. Phys.* 81 (1984), pp. 3684–3690. DOI: 10.1063/1.448118 and Frenkel and Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, pp. 158–160.

[61] See P. P. Ewald. „Die Berechnung Optischer und Elektrostatischer Gitterpotentiale". In: *Ann. Phys. (Berl.)* 369 (1921), S. 253–287. DOI: 10.1002/andp.19213690304.

[62] See Frenkel and Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, pp. 292–300.

[63] See Rapaport, *The Art of Molecular Dynamics Simulation*, pp. 15–17.

[64] For long ranged interactions calculated *via* Ewald summation method the interaction contribution calculated in real space can be treated as short ranged.

[65] See Verlet, "Computer "Experiments" on Classical Fluids".

thermostat can be implemented *via* different algorithms, such as (list not exhaustive) Andersen, Nosé-Hoover, or Langevin thermostats.[59]

- **Barostat:** similarly, simulations of isothermal-isobaric NPT ensembles must implement some kind of pressure bath in addition to the heat bath. An example of such a barostat is the Berendsen barostat.[60]

- **Ewald summation:** a powerful tool to manage long-range interactions $\propto r^{-1}$, *e.g.*, electrostatic interactions, in periodic systems is the Ewald summation method,[61] where the interaction potential is divided into two parts: a short-range contribution, and a long-range contribution which is free of singularities. Whereas the short-range part is calculated in real space, the long-range part is calculated in reciprocal space *via* Fourier transform.[62]

- **Periodic boundary conditions:** an invaluable technique is the method of periodic boundary conditions (PBC), whereby a finite simulation box is made virtually infinite by connecting opposing boundaries, *i.e.*, the topolgy of a cubic box becomes that of a torus.[63] If the box is chosen large enough to avoid unphysical artifacts, such as self-interaction of particles, PBC represent a cheap way of approximating macroscopic system sizes. Modern simulation methods usually combine PBC with the Ewald summation method.

- **Neighbour lists:** As many interactions in simulations are short-ranged,[64] cutting off the interactions at a cut-off distance $r_{\rm c}$ does not significantly affect the numerical accuracy of the simulation. Even if the majority of the particles is not located within interaction range, the inter-particle distances nevertheless have to be calculated. The run time of these calculations scales with $O(N^2)$ for N interacting particles. A useful book-keeping device to reduce the unnecessary computations is the so-called neighbour list, also known as Verlet list.[65] In this method each particle is associated with a list which contains the neighbouring particles located within interaction range $r_{\rm c}$ and some buffer distance known as "skin". A further improvement are the cell (linked-)lists, whereby the simulation box is divided into a

cuboid cell structure in such a way that each particle can only interact with particles in its own cell and adjacent cells. This way, the computational complexity of the force calculation between the particles can be reduced from $O(N^2)$ to $O(N)$.[66]

[66] See Frenkel and Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, pp. 550–558.

2.3 Structural Quantities

In this section some of the structural quantities containing information on the macroscopic structure of the system are presented.

2.3.1 Radial Distribution Function

The structure of an N-particle system of volume V and at temperature T can be characterised by the set of n-body correlation functions $g^{(n)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n)$ defined *via*[67]

[67] See Hansen and McDonald, *Theory of Simple Liquids*, pp. 32–33.

$$g^{(n)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n) = \frac{V^n N!}{N^n (N-n)!} \frac{1}{Z_N} \int_V \cdots \int_V e^{-\beta U} \, d\boldsymbol{r}_{n+1} \ldots d\boldsymbol{r}_N \,, \quad (2.26)$$

$$\text{with} \quad Z_N = \int_V \cdots \int_V e^{-\beta U} \, d\boldsymbol{r}_1 \ldots d\boldsymbol{r}_N \,. \quad (2.27)$$

Here, $\boldsymbol{r}_i$ denotes the coordinates of particle i, whereas $U = U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ denotes the interaction energy between all particles. Further, Z_N is the partition function of the canonical, *i.e.*, NVT, ensemble and $\beta = (k_B T)^{-1}$, with Boltzmann constant k_B.

The simplest non-trivial[68] function of this set of correlation functions for a homogenous system is the pair correlation function $g^{(2)}(\boldsymbol{r}_i, \boldsymbol{r}_j)$:[69]

[68] While the zeroth order correlation is always constant ($g^{(0)} = 1$), the first order correlation also reduces to $g^{(1)}(\boldsymbol{r}_i) = 1$ for homogenous systems.

[69] Often, the superscript of $g^{(2)}(\boldsymbol{r}_i, \boldsymbol{r}_j)$ is dropped and the pair correlation function is denoted as $g(\boldsymbol{r}_i, \boldsymbol{r}_j)$.

$$g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{N(N-1)}{\rho^2} \frac{1}{Z_N} \int_V \cdots \int_V e^{-\beta U} \, d\boldsymbol{r}_3 \ldots d\boldsymbol{r}_N \,, \quad (2.28)$$

with density $\rho = N/V$. In a system of identical particles the choice of $i = 1$ and $j = 2$ is arbitrary. While the set of particle distribution functions $\{g^{(n)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n)\}$, with $n \in [1, N]$, provides a complete description of the system's structure, knowledge of the lower order correlation functions, *i.e.*, the pair correlation function $g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)$, is sufficient to calculate the

equation of state and other important thermodynamical properties of the system.[70]

In homogenous systems the pair correlation function only depends on the relative vectors $r_{ij} = r_j - r_i$, i.e., $g^{(2)}(r_1, r_2)$ can be written as $g^{(2)}(r_{12})$.[71] Without loss of generality, we can assume particle i to be fixed at the origin. Then, the average number of particles included in volume element $\mathrm{d}^3 r$ at coordinate position r_{ij} is given by $g(r_{ij})\rho \, \mathrm{d}^3 r$.

Simply counting the particle pairs separated by vector $r_{ij} = r$[72] is an alternative way of obtaining the value of $g(r)\rho \, \mathrm{d}^3 r$. For this purpose we introduce the pair density $\rho^{(2)}$ [73] via[74]

$$\rho^{(2)}(r', r) = \left\langle \sum_i \sum_{j \neq i} \delta^{(3)}(r' + r - r_i)\delta^{(3)}(r' - r_j) \right\rangle . \qquad (2.29)$$

Here, we take the ensemble average of particle pairs where particles i and j are located at positions $r' + r$ and r', respectively. Again assuming a homogenous system, we arrive at an expression for $g(r)$ by integrating and averaging r' over the system volume V and normalising the results with respect to the system density $\rho = N/V$:[75]

$$g(r) = \frac{V^2}{N^2} \frac{1}{V} \int_V \rho^{(2)}(r', r) \, \mathrm{d}r = \frac{V}{N^2} \left\langle \sum_i \sum_{j \neq i} \delta^{(3)}(r - r_{ij}) \right\rangle . \qquad (2.30)$$

Furthermore, we can write

$$g(r) = \frac{V}{N^2} \left\langle \sum_i \sum_{j \neq i} \delta(r - r_{ij}) \right\rangle . \qquad (2.31)$$

In Equation (2.31) we additionally assumed isotropy in the system, e.g., a liquid system, so that the result only depends on the relative distance $r = |r|$ (with $r_{ij} = |r_j - r_i|$). Finally, we have arrived at the radial distribution function (RDF), $g(r)$, which is of special importance due to its relation to the (static) structure factor, $S(k)$, via the Fourier transform, see Equation (2.43). This function gives the probability of finding a pair of particles separated by distance r relative to the probability of a uniform uncorrelated distri-

[70] See Hansen and McDonald, *Theory of Simple Liquids*, p. 33.

[71] In homogenous systems shifting the arguments of $g^{(2)}(r_1, r_2)$ by an arbitrary constant vector h does not change the outcome: $g^{(2)}(r_1 + h, r_2 + h) = g^{(2)}(r_1, r_2)$. Choosing h to be $-r_1$ yields $g^{(2)}(0, r_2 - r_1) = g^{(2)}(r_{12})$.

[72] Here, we drop the subscripts of r_{ij} for simplicity's sake.

[73] The n-particle density $\rho^{(n)}$ can be defined via $\rho^{(n)}(r_1, \ldots, r_N) = \rho^n g^{(n)}(r_1, \ldots, r_N)$. See Hansen and McDonald, *Theory of Simple Liquids*, p. 33.

[74] See Allen and Tildesley, *Computer Simulations of Liquids*, p. 69.

[75] See Allen and Tildesley, *Computer Simulations of Liquids*, p. 70.

bution of particles, *i.e.*, the distribution of an ideal gas.[76] Note, however, that the simplified definition of $g(r)$ in Equation (2.31) only holds true for homogeneous and isotropic systems.

In Monte Carlo (MC) and MD simulations, $g(r)$ can be computed by evaluating Equation (2.31). This is achieved by appropriately sampling inter-particle distances from simulation data and compiling them into a histogram with binning Δr and normalising the result with respect to a completely uncorrelated system, *i.e.*, an ideal gas.[77]

In addition to structural insight, the RDF also provides a way to calculate the ensemble average $\langle \mathcal{A} \rangle$ of any pairwise defined function $a(r_{ij})$ with

$$\langle \mathcal{A} \rangle = \left\langle \frac{1}{2} \sum_i \sum_{j \neq i} a(r_{ij}) \right\rangle = \frac{N^2}{2V} \int_0^\infty g(r) a(r_{ij} = r) 4\pi r^2 \, dr \, . \qquad (2.32)$$

For example, the internal energy of a system can be expressed using the energy equation expressed by[78]

$$U_{\mathrm{id}} + U_{\mathrm{ex}} = \frac{3}{2} N k_{\mathrm{B}} T + \frac{2\pi N^2}{V} \int_0^\infty g(r) u(r) r^2 \, dr \, , \qquad (2.33)$$

where $u(r)$ is the pairwise potential energy and U_{id} and U_{ex} denote the ideal and excess parts of the internal energy, respectively. Whereas the ideal part corresponds to the internal energy of the ideal gas, *i.e.*, the kinetic energy, the excess part is determined by the interaction potential $u(r)$.

It can be shown that in the low-density limit ($\rho \to 0$) the radial distribution function $g(r)$ is equal to the Boltzmann factor of the pairwise potential $u(r)$:[79]

$$\lim_{\rho \to 0} g(r) = e^{-\beta u(r)} \approx 1 - \beta u(r) \, , \qquad (2.34)$$

where the approximation on the right hand side is only valid for weak interactions $u(r) \ll k_{\mathrm{B}} T$. Equation (2.34) is not only a useful expression for calculating $\lim_{\rho \to 0} g(r)$, but also provides a method of obtaining an unknown pairwise potential *via* $u(r) = -k_{\mathrm{B}} T \ln[g(r)]$ from trajectory data, *e.g.*, using the umbrella sampling method.[80] One immediate result of Equation (2.34) is $g(r) = 1$ for non-interacting particles, *i.e.* an ideal gas, with

[76] By rescaling $g(r)$ with density ρ it is ensured that at large distances $\lim_{r \to \infty} g(r) = 1$.

[77] The application of this procedure is equivalent to replacing the δ functions $\delta(\ldots)$ in Equation (2.31) by a function which is non-zero in a small separation range of Δr.

[78] See Hansen and McDonald, *Theory of Simple Liquids*, p. 25 and Hansen and McDonald, *Theory of Simple Liquids*, p. 36.

[79] See Hansen and McDonald, *Theory of Simple Liquids*, pp. 38–39.

[80] See Section 2.5.

$u(r) = 0$: in this case the particles are completely uncorrelated.

An alternative way of calculating $g(r)$ (and thus also $S(k)$) is given by integral equation theory, *i.e.*, the Ornstein-Zernike (OZ) equation, using closure relations such as the Percus-Yevick (PY) approximation or the hyper-netted chain (HNC) approximation.[81]

2.3.2 Structure Factor

As mentioned before, an attractive feature of the RDF, $g(r)$, is its direct relation to the experimentally measureable structure factor (SF), $S(k)$,[82] which we will become more familiar with in the following.

[82] Not to be confused with F_{hkl}, a quantity also called structure factor, which is only valid for systems with long-range positional order, *i.e.*, crystals.

We consider a system of N identical point-like scatterers (at positions $\{r_1 \ldots r_N\}$): every scattering event of a particle with an incident wave of wavelength λ is characterised by its scattering vector $k_\mathrm{s} = k_2 - k_1$, where the incident wave vector, k_1, and the scattered wave vector, k_2, have the same magnitude[83] $|k_1| = |k_2| = 2\pi/\lambda$. Using the angle between incident and scattered wave vectors denoted by $\theta = (k_1 \cdot k_2)/|k_1||k_2|$, we can write

[83] This only holds true under the assumption of elastic scattering.

$$k_\mathrm{s} = |k_\mathrm{s}| = |k_2 - k_1| = \frac{4\pi \sin\left(\frac{\theta}{2}\right)}{\lambda} \ . \tag{2.35}$$

[84] See M. Born. "Quantenmechanik der Stoßvorgänge". In: *Z. Phys.* 38 (1926), pp. 803–827. DOI: 10.1007/BF01397184.

Under the assumption of weak scattering, *i.e.*, the Born approximation,[84] the amplitude and phase of the wave which is scattered by all N particles are given by[85]

[85] From now on we will drop the subscript of k_s.

$$\psi(k_\mathrm{s} = k) = \sum_m \mathrm{e}^{-i k \cdot r_n} \ . \tag{2.36}$$

[86] Assuming that refraction and multiple scattering can be neglected, *i.e.*, assuming kinematic diffraction, see H. Chen. "Kinematic Diffraction of X-Rays". In: *Characterization of Materials*. 2012, pp. 1–21. DOI: 10.1002/0471266965.com018.pub2.

[87] Experimentally, $S(k)$ can be interpreted as the inter-particle interference since it describes how waves from different particles interfere.

Note that $\psi(k)$ is a function of scattering vector k and thus of angle θ. Using an appropriate experimental setup, the scattered intensity given by $I(k) = \psi^*(k)\psi(k)$,[86] with the complex conjugate $\psi^*(k)$, can be measured for for each k individually. Normalising the ensemble average of the scattered intensity $I(k)$ by the number of scatterers N yields the (static) structure factor $S(k)$:[87]

$$S(\mathbf{k}) = \frac{1}{N} \left\langle \psi^*(\mathbf{k}) \psi(\mathbf{k}) \right\rangle = \frac{1}{N} \left\langle \sum_m \sum_n e^{-i\mathbf{k}\cdot\mathbf{r}_{mn}} \right\rangle , \qquad (2.37)$$

with relative vectors $\mathbf{r}_{mn} = \mathbf{r}_n - \mathbf{r}_m$. Analogous to the RDF, this expression can be evaluated by using appropriately sampling inter-particle vectors $\mathbf{r}_{mn}$ from simulation data, *e.g.*, from MD simulations. Furthermore, Equation (2.37) can be rewritten in terms of $g(\mathbf{r})$ in Equation (2.30) *via*[88]

[88] See Hansen and McDonald, *Theory of Simple Liquids*, p. 105.

$$S(\mathbf{k}) = \frac{1}{N} \left\langle \sum_m \sum_n e^{-i\mathbf{k}\cdot\mathbf{r}_{mn}} \right\rangle = 1 + \frac{1}{N} \left\langle \sum_m \sum_{m\neq n} e^{-i\mathbf{k}\cdot\mathbf{r}_{mn}} \right\rangle \qquad (2.38)$$

$$= 1 + \frac{1}{N} \left\langle \sum_m \sum_{m\neq n} \int_V e^{-i\mathbf{k}\cdot\mathbf{r}} \delta(\mathbf{r} - \mathbf{r}_{mn}) \, d\mathbf{r} \right\rangle \qquad (2.39)$$

$$= 1 + \rho \int_V e^{-i\mathbf{k}\cdot\mathbf{r}} \underbrace{\frac{V}{N^2} \left\langle \sum_{m=1} \sum_{m\neq n} \delta(\mathbf{r} - \mathbf{r}_{mn}) \right\rangle}_{g(\mathbf{r})} d\mathbf{r} , \qquad (2.40)$$

where $\rho = N/V$. Thus, we have established the aforementioned relation between $g(\mathbf{r})$ and $S(\mathbf{k})$ *via* the Fourier transform.

Because $g(\mathbf{r})$ is not normalised, *i.e.*, $\lim_{|\mathbf{r}|\to\infty} g(\mathbf{r}) = 1$, at the $\mathbf{k}$ origin the outcome of Equation (2.40) scales with V:

$$\lim_{V\to\infty} S(\mathbf{k})\big|_{|\mathbf{k}|=0} = \lim_{V\to\infty} \rho \int_V g(\mathbf{r}) \, d\mathbf{r} = \infty . \qquad (2.41)$$

Experimentally, this contribution corresponds to a scattering angle $\theta = 0$, *i.e.*, the radiation which passes through the sample unscattered. Henceforth, we shall ignore this δ function contribution in order to regularise $S(\mathbf{k})$ *via*

$$S'(\mathbf{k}) = 1 + \rho \int_V e^{-i\mathbf{k}\cdot\mathbf{r}} [g(\mathbf{r}) - 1] \, d\mathbf{r} \quad \text{and} \quad S(\mathbf{k}) \equiv S'(\mathbf{k}) . \qquad (2.42)$$

Analogous to Equation (2.31), the assumption of isotropy in liquid systems results in a structure factor $S(k)$ which is a function of $k = |\mathbf{k}|$ and can be expressed *via*

$$S(k) = \frac{1}{N} \left\langle \sum_{m=1}^N \sum_{n=1}^N \frac{\sin(k r_{mn})}{k r_{mn}} \right\rangle = 1 + \rho \int_V 4\pi r^2 \frac{\sin kr}{kr} [g(r) - 1] \, dr , \quad (2.43)$$

where the integral on the right hand side of the equation corresponds to a Fourier transform in spherical coordinates. Now, we can introduce the total correlation function $h(r) = g(r) - 1$ and its Fourier transform $\hat{h}(k)$ in order to write

$$S(k) = 1 + \rho \hat{h}(k) \,. \tag{2.44}$$

An alternative but equivalent definition of $S(k)$ is given by[89]

[89] See Allen and Tildesley, *Computer Simulations of Liquids*, p. 72

$$S(k) = \frac{1}{N} \left\langle \hat{\rho}(\boldsymbol{k}) \hat{\rho}(-\boldsymbol{k}) \right\rangle \,, \tag{2.45}$$

$$\text{with} \quad \hat{\rho}(\boldsymbol{k}) = \int_V e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \sum_m \delta^{(3)}(\boldsymbol{r} - \boldsymbol{r}_m)\,\mathrm{d}\boldsymbol{r} = \sum_m e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_m} \,, \tag{2.46}$$

where $\hat{\rho}(\boldsymbol{k})$ is the Fourier transform of the number density

$$\rho(\boldsymbol{r}) = \sum_m \delta^{(3)}(\boldsymbol{r} - \boldsymbol{r}_m) \,. \tag{2.47}$$

Finally, we consider the low-k limit of the structure factor $S(k)$, *i.e.*, probing the system with a large wavelength λ, *i.e.*, $\lambda \to \infty$ and $k \to 0$. A valuable relation for this case is the compressibility equation given by[90]

[90] See Hansen and McDonald, *Theory of Simple Liquids*, pp. 40–41

$$\rho k_{\mathrm{B}} T \chi_T = 1 + \rho \int_V [g(r) - 1]\,\mathrm{d}\boldsymbol{r} \,, \tag{2.48}$$

where χ_T denotes the isothermal compressibility. With Equation (2.44) at hand, we can rewrite Equation (2.48):[91]

[91] See Hansen and McDonald, *Theory of Simple Liquids*, p. 78

$$\lim_{k \to 0} S(k) = 1 + \rho \hat{h}(0) = \rho k_{\mathrm{B}} T \chi_T = k_{\mathrm{B}} T \left(\frac{\partial \rho}{\partial p} \right)_T \,. \tag{2.49}$$

Here, p denotes the pressure applied to the system. The combination of Equations (2.31), (2.44) and (2.49) provides an elegant way to associate microscopic properties of the system, *i.e.*, the particle coordinates $\boldsymbol{r}_1 \ldots \boldsymbol{r}_N$, which can be extracted from simulations, and experimentally accessible quantities, *i.e.*, the structure factor $S(k)$ obtained *via* static light or neutron scattering experimtens, with macroscopic properties of the system, *i.e.*, the isothermal compressibility χ_T, which describes the response of the density of the system, ρ, to the applied pressure p.

2.3.3 Form Factor

If the system being probed *via* static scattering experiments does not consist of point-like particles but contains larger molecules with finite extension, *e.g.*, polymers or dendrimers, the structure factor $S(k)$ can not be directly inferred from the total (isotropic) scattering intensity $I(k)$ but has to be modified *via*[92]

$$I_{\text{tot}}(k) \propto S(k)F(k) \quad \Rightarrow \quad S(k) \propto \frac{I_{\text{tot}}(k)}{F(k)} \, , \qquad (2.50)$$

where $F(k)$ corresponds to the form factor (FF) which describes the scattering intensity of an isolated particle and depends on the particle's shape. In analogy to Equation (2.38) the form factor $F(k)$ of dendrimeric molecules — which are assumed to consist of point-like particles — can be calculated using[93]

$$F(k) = \frac{1}{N} \left\langle \sum_m \sum_n e^{-i\mathbf{k} \cdot \mathbf{r}_{mn}} \right\rangle = \frac{1}{N} \left\langle \sum_{m=1}^{N} \sum_{n=1}^{N} \frac{\sin(k r_{mn})}{k r_{mn}} \right\rangle \, , \qquad (2.51)$$

with $\mathbf{r}_{mn} = \mathbf{r}_n - \mathbf{r}_m$, where $\mathbf{r}_m$ and $\mathbf{r}_n$ denote the positions of particles m and n, respectively. Note, that in Equation (2.51) particles m and n represent the individual constitutents of macromolecules, *e.g.*, polymers or dendrimers. Each of these macromolecules corresponds to a point-like scattering particle in Equation (2.38). In other words: while Equation (2.38) describes the *inter*-particle interference in a scattering experiment, Equation (2.51) describes the *intra*-particle interference.

In experiments, the form factor $F(k)$ is accessible by performing light scattering experiments on systems in the dilute, *i.e.*, low-density, regime:

> *The intensity of radiation scattered by an ensemble of widely separated particles is thus identical on a relative scale to the mean intensity scattered by one isolated particle.*[94]

It can be shown[95] that in the low-k limit the form factor ($k \to 0$) can be

[92] See A. Guinier and G. Fournet. *Small angle scattering of X-rays*. Trans. by C. B. Wilson. J. Wiley & Sons, 1955. DOI: 10.1002/pol.1956.120199326, p. 35. Note, that $F(k)$ in Eqation (2.50) corresponds to $F^2(h)$ in this reference.

[93] See Guinier and Fournet, *Small angle scattering of X-rays*, p. 12.

[94] Guinier and Fournet, *Small angle scattering of X-rays*, p. 36

[95] See Guinier and Fournet, *Small angle scattering of X-rays*, p. 25 and G. Beaucage. "Approximations Leading to a Unified Exponential/Power-Law Approach to Small-Angle Scattering". In: *J. Appl. Cryst.* 28 (1995), pp. 717–728. DOI: 10.1107/S0021889895005292, pp. 720–721.

approximated in the Guinier regime ($kR_{\mathrm{g}} \lesssim 1$) *via*

$$F(k) \approx N \mathrm{e}^{\frac{-(kR_{\mathrm{g}})^2}{3}} \, , \tag{2.52}$$

which represents a useful relation between the form factor, $F(k)$, the number of scatterers, N, and the radius of gyration, R_{g}, in the Guinier regime ($kR_{\mathrm{g}} \lesssim 1$). Equation (2.52) is known as Guinier's law and can be used to extract R_{g} from experimental form factor data in the low wave vector limit.

2.3.4 Radial Density Profile

A simple yet insightful quantity which provides detailed information structure of macromolecules, such as dendrimers, is the radial density profile defined by[96]

[96] Not to be confused with the number density $\rho(\mathbf{r})$ in Equation (2.47).

$$\rho(r) = \left\langle \sum_i \delta(r - |\mathbf{r}_i - \mathbf{r}_{\mathrm{com}}|) \right\rangle . \tag{2.53}$$

Here, the summation index i runs over all particles pertaining to a particular entity, *e.g.*, all monomers of a particular dendrimer or all counterions, and particle positions $\mathbf{r}_i$ are set in relation to the center of mass of the macromolecule, $\mathbf{r}_{\mathrm{com}}$.

By calculating $\rho(r)$ for each individual subgeneration of a dendrimer, this quantity provides ample information about the complex internal structure of dendrimers. Furthermore, by calculating $\rho(r)$ for ions, the condensation behaviour of ions within a dendrimer can be examined.

2.4 Ornstein-Zernike Relation

[97] See L. S. Ornstein and F. Zernike. "Accidental deviations of density and opalescence at the critical point of a single substance". In: *Proc. Akad. Sci.* 17 (1914), pp. 793–806. URL: https://www.dwc.knaw.nl/DL/publications/PU00012727.pdf.

Originally introduced by Ornstein and Zernike in 1914,[97] the Ornstein-Zernike (OZ) relation defines the direct correlation function $c(\mathbf{r}, \mathbf{r}')$ between two particles located at positions $\mathbf{r}$ and $\mathbf{r}'$, respectively, in terms of the total

pair correlation function $h(\mathbf{r}, \mathbf{r}') = g(\mathbf{r}, \mathbf{r}') - 1$ *via* [98,99]

$$h(\mathbf{r}, \mathbf{r}') = c(\mathbf{r}, \mathbf{r}') + \int c(\mathbf{r}, \mathbf{r}'')\rho(\mathbf{r}'')h(\mathbf{r}'', \mathbf{r}')\, \mathrm{d}\mathbf{r}'' \, . \qquad (2.54)$$

The integral on the right hand side of Equation (2.54), which is given by $\gamma(\mathbf{r}, \mathbf{r}') = h(\mathbf{r}, \mathbf{r}') - c(\mathbf{r}, \mathbf{r}')$, corresponds to the indirect correlations between the two particles. This indirect correlation is mediated *via* intermediary particles, as can be demonstrated by the recursive solution for Equation (2.54):

$$\begin{aligned} h(\mathbf{r}, \mathbf{r}') = c(\mathbf{r}, \mathbf{r}') &+ \int c(\mathbf{r}, \mathbf{r}'')\rho(\mathbf{r}'')h(\mathbf{r}'', \mathbf{r}')\, \mathrm{d}\mathbf{r}'' \\ &+ \int \int c(\mathbf{r}, \mathbf{r}''')\rho(\mathbf{r}''')c(\mathbf{r}''', \mathbf{r}'')\rho(\mathbf{r}'')c(\mathbf{r}'', \mathbf{r}')\, \mathrm{d}\mathbf{r}'''\, \mathrm{d}\mathbf{r}'' \quad (2.55) \\ &+ \dots \, . \end{aligned}$$

Here, the first term in the sum of Equation (2.55) is the direct correlation between two particles located at positions $\mathbf{r}$ and $\mathbf{r}'$, while the second term corresponds to the indirect correlation mediated *via* one intermediary particle located at $\mathbf{r}''$, and the third term corresponds to the indirect correlation mediated *via* two intermediary particles located at $\mathbf{r}''$ and $\mathbf{r}'''$. In order for Equation (2.55) to be exact, this series has to be expanded *ad infinitum*, ultimately taking into account an infinite number of many-body correlations.

For isotropic and homogenous, *i.e.*, rotationally and translationally invariant, systems Equation (2.54) can be rewritten as[100]

$$h(r) = c(r) + \rho \int c(|\mathbf{r} - \mathbf{r}'|)h(|\mathbf{r}'|)\, \mathrm{d}\mathbf{r}' \, , \qquad (2.56)$$

where the integral on the right hand side corresponds to a convolution and is easier to handle in reciprocal space. Therefore, by applying the Fourier transform to both sides of Equation (2.56), we obtain

$$h(r) = c(r) + \rho \, (h * c) \, (r) \qquad (2.57)$$

$$\text{and} \quad \hat{h}(k) = \hat{c}(k) + \rho\hat{h}(k)\hat{c}(k) \, , \qquad (2.58)$$

with convolution operator $*$ and the Fourier transforms of $h(r)$ and $c(r)$,

[98] See Hansen and McDonald, *Theory of Simple Liquids*, p. 73.

[99] In analogy to Section 2.3.1 we choose the notation $g(\mathbf{r}, \mathbf{r}')$ instead of $g^{(2)}(\mathbf{r}, \mathbf{r}')$ as there is no risk of ambiguity. The same notation is applied to $h(\mathbf{r}, \mathbf{r}')$ and $c(\mathbf{r}, \mathbf{r}')$.

[100] See Hansen and McDonald, *Theory of Simple Liquids*, p. 73.

$\hat{h}(k)$ and $\hat{c}(k)$. Combining Equations (2.44) and (2.58) yields a connection between the direct correlation function $c(r)$ and the structure factor $S(k)$:

$$\hat{h}(k) = \frac{\hat{c}(k)}{1 - \rho\hat{c}(k)} , \tag{2.59}$$

$$\text{and} \quad S(k) = \frac{1}{1 - \rho\hat{c}(k)} . \tag{2.60}$$

In order to solve for both $h(r)$ and $c(r)$ (or their Fourier transforms, $\hat{h}(k)$ and $\hat{c}(k)$) additional information is necessary. This additional information is provided by the so-called closure relation which can be expressed via[101]

$$g(r) = e^{-\beta u(r) + h(r) - c(r) + B(r)} , \tag{2.61}$$

where $u(r)$ denotes the pair potential and the bridge function, $B(r)$, depends on the details of the system, especially on the pair potenital, $u(r)$.

A frequently used closure relation for particles with an impenetrable core is the Percus-Yevick (PY) approximation given by[102]

$$B(r) = \ln[1 + \gamma(r)] - \gamma(r) , \tag{2.62}$$

with the indirect correlation function $\gamma(r)$. Using the PY closure, an analytical solution of the integral equation provided by the OZ relation is possible for a system of hard spheres.[103]

Another common closure relation, which is widely used for soft potentials, is the hyper-netted chain (HNC) approximation which can be derived using diagrammatic methods, i.e., graph theory: omitting the elementary clusters[104] in the exact convolution of Equation (2.57) yields a trivial bridge function:[105]

$$B(r) = 0 \tag{2.63}$$

$$\text{and} \quad g(r) = e^{-\beta u(r) + h(r) - c(r)} = e^{-\beta u(r) + \gamma(r)} . \tag{2.64}$$

While no analytical solutions to the OZ equation are available via the HNC approximation, several algorithms to compute numerical solutions have

[101] See T. Morita and K. Hiroike. "A New Approach to the Theory of Classical Fluids. I: Formulation for a One-Component System". In: *Prog. Theor. Phys.* 23 (1960), pp. 1003–1027. DOI: 10.1143/PTP.23.1003.

[102] See J. K. Percus and G. J. Yevick. "Analysis of Classical Statistical Mechanics by Means of Collective Coordinates". In: *Phys. Rev.* 110 (1 1958), pp. 1–13. DOI: 10.1103/PhysRev.110.1.

[103] See M. S. Wertheim. "Exact Solution of the Percus-Yevick Integral Equation for Hard Spheres". In: *Phys. Rev. Lett.* 10 (8 1963), pp. 321–323. DOI: 10.1103/PhysRevLett.10.321.

[104] See T. Morita. "Theory of Classical Fluids: Hyper-Netted Chain Approximation, I: Formulation for a One-Component System". In: *Progress of Theoretical Physics* 20.6 (1958), pp. 920–938. DOI: 10.1143/PTP.20.920.

[105] The HNC method was developed almost simultaneously by several groups, see (list not exhaustive): J. van Leeuwen et al. "New method for the calculation of the pair correlation function. I". in: *Physica* 25.7 (1959), pp. 792–808. DOI: 10.1016/0031-8914(59)90004-7, Morita, "Theory of Classical Fluids: Hyper-Netted Chain Approximation, I: Formulation for a One-Component System", and L. Verlet. "On the theory of classical fluids". In: *Il Nuovo Cimento (1955-1965)* 18 (1960), pp. 77–101. DOI: 10.1007/BF02726040.

been devised.[106]

A well-established method[107] which is used in this work and which we will call HNC method, starts out with an initial estimate for one of the functions, *e.g.*, the indirect correlation $\gamma(r)$. This estimate is then used in combination with Equation (2.64) to obtain $c(r)$ *via*

$$c(r) = h(r) - \gamma(r) = g(r) - \gamma(r) - 1 = e^{-\beta u(r) + \gamma(r)} - \gamma(r) - 1 \,. \qquad (2.65)$$

Now, we use the obtained direct correlation function $c(r)$ with Equation (2.57) *via*

$$\gamma(r) = h(r) - c(r) = \rho\,(h * c)\,(r) \qquad (2.66)$$

in order to obtain a new estimate of $\gamma(r)$.[108] With this new estimate at hand, a new iteration of the cycle we just discussed can be started. This algorithm, which corresponds to a Picard algorithm in numerical analysis,[109] is applied iteratively until convergence of $c(r)$ and $\gamma(r)$ to stable solutions is achieved. If a good initial estimate is available, the problems occurring due to the divergence of the Picard algorithm at higher densities can be solved using Broyles' mixing procedure.[110]

Starting with only the pair potential $u(r)$, the method described above enables us to directly obtain $g(r)$ and $S(k)$ in a fast way.

2.5 Effective Interactions

Many soft matter systems consist of complex macromolecules containing a large number of constituent particles suspended in a bath of smaller molecules. Therefore, it is necessary to capture a wide range of length- and time-scales in order to fully describe emerging phenomena in these systems. Due to the computational effort necessary to run atomistic simulations of such systems on large length- and long time-scales, only simulations of few macromolecules on short time-scales are possible in such a manner.

Two different coarse-graining approaches to solve this problem are presented in Sections 2.5.1 and 2.5.2, whereas a method based on neural net-

[106] *E.g.*, an algorithm based on the Newton-Raphson method, see M. J. Gillan. "A new method of solving the liquid structure integral equations". In: *Mol. Phys.* 38 (1979), pp. 1781–1794. DOI: 10.1080/00268977900102861.

[107] See R. O. Watts. "Integral equation approximations in the theory of fluids". In: *Statistical Mechanics: Volume 1*. Ed. by K. Singer. Vol. 1. The Royal Society of Chemistry, 1973, pp. 1–70. DOI: 10.1039/9781847556929-00001.

[108] This step can be sped up by employing the fast fourier transform (FFT).

[109] See J. Ortega and W. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*. Ed. by J. Ortega and W. Rheinboldt. Academic Press, 1970, pp. 1–6. DOI: 10.1016/B978-0-12-528550-6.50008-9.

[110] See A. A. Broyles et al. "Comparison of Radial Distribution Functions from Integral Equations and Monte Carlo". In: *J. Chem. Phys.* 37 (1962), pp. 2462–2469. DOI: 10.1063/1.1733028.

work potentials (NNPs) is presented in Appendix A.4.

2.5.1 RDF-based Methods

By integrating out the macromolecules' degrees of freedom deemed negligible for the system's essential meso- and macroscopic behaviour, the complexity of the system can be reduced and longer simulations of larger systems become feasible. The simplest approaches reduce the entire macromolecule down to one effective coordinate without orientation assuming an isotropic effective interaction.

Consider two mesoscopic particles, whose effective coordinates are seperated by distance r_{12}, suspended in an infinitely dilute system.The relation connecting the radial distribution function $g(r_{12})$ to the effective interaction between the particles $\phi_{\text{eff}}(r_{12})$ is given by[111]

[111] See Equation (2.34).

$$g(r_{12}) = e^{-\beta\phi_{\text{eff}}(r_{12})} \tag{2.67}$$

$$\text{or} \quad \phi_{\text{eff}}(r_{12}) = -k_{\text{B}}T\ln[g(r_{12})] \, , \tag{2.68}$$

whith reciprocal thermal energy $\beta = 1/k_{\text{B}}T$. Equation (2.68) represents a straightforward way of measuring $\phi_{\text{eff}}(r_{12})$ as $g(r_{12})$ can be measured in MD or MC simulations. Here, one caveat has to be added: as the distance seperating the two complex molecules decreases, their repulsion increases, reaching energies of several $k_{\text{B}}T$ and higher. This means that such configurations where particles are positioned at small seperations are rare in MD and MC simulations, resulting in a large error for the measured effective interaction ϕ_{eff}. In order to reduce this error one would have to spend a huge amount of computational time in order to sufficiently sample these close distances, where almost all of the computational time would be spent on sampling irrelevant and easily accessible configurations. This hindrance to ergodicity and the resulting imbalance in sampling can be mitigated *via* various methods, *e.g.*, bias potentials[112] and umbrella sampling.[113]

[112] See R. Blaak et al. *Accurate coarse-grained potentials for soft matter systems*. Ed. by G. Sutmann et al. Vol. 28. IAS Series. Jülich: Forschungszentrum Jülich, Zentralbibliothek, 2015, pp. 209–258. URL: https://juser.fz-juelich.de/record/188877, pp. 222–223.

[113] See Blaak et al., *Accurate coarse-grained potentials for soft matter systems*, p. 229.

2.5.2 Widom Insertion Method

An alternative approach to calculating $\phi_{\text{eff}}(r_{12})$ is based on the Widom insertion algorithm and will therefore be called Widom insertion (WI) method in this work.[114] A derivation is presented in the following.

Consider again the two identical mesoscopic particles — each possessing M internal degrees of freedom — located at effective coordinates $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ in an infinitely dilute system. We denote the internal degrees of freedom of the system, *i.e.*, the M coordinates of the subparticles (or monomers) constituting our two molecules, as $\{\boldsymbol{s}_{i\alpha}\}$, with indices $\alpha = 1, 2$ and $i = 1 \ldots M$ corresponding to the particular mesoscopic particle and its subparticles, respectively. The total interaction energy within our system, U_{tot}, can now be written in terms of the two sets of monomer coordinates, $\{\boldsymbol{s}_{i1}\}$ and $\{\boldsymbol{s}_{j2}\}$:

$$U_{\text{tot}}(\{\boldsymbol{s}_{i1}\}, \{\boldsymbol{s}_{j2}\}) = U_{\text{intra}} + U_{\text{inter}} \, , \tag{2.69}$$

$$\text{with} \quad U_{\text{intra}} = U_{11}(\{\boldsymbol{s}_{i1}\}) + U_{22}(\{\boldsymbol{s}_{j2}\}) \tag{2.70}$$

$$\text{and} \quad U_{\text{inter}} = U_{12}(\{\boldsymbol{s}_{i1}\}, \{\boldsymbol{s}_{j2}\}) \, . \tag{2.71}$$

Here U_{intra} ($U_{\alpha\beta}$ with particle indices $\alpha = \beta$) represents the *intra*-molecule energy, whereas U_{inter} ($U_{\alpha\beta}$ with particle indices $\alpha \neq \beta$) represents the *inter*-molecule energy.

It can be shown[115] that the effective interaction between the two macromolecules, $\phi_{\text{eff}}(\boldsymbol{r}_{12})$, is defined *via*

$$e^{-\beta\phi_{\text{eff}}(\boldsymbol{r}_{12})} = \frac{\tilde{Z}_{\text{tot}}(\boldsymbol{r}_{12})}{\tilde{Z}_{\text{intra}}^2} \, , \tag{2.72}$$

where $\tilde{Z}_{\text{tot}}(r_{12})$ denotes the configuration part of the partition function due to the total interaction energy $U_{\text{tot}}(\{\boldsymbol{s}_{i1}\}, \{\boldsymbol{s}_{j2}\})$, while $\tilde{Z}_{\text{intra}}$ denotes the configuration part of the partition function of a single molecule due to the intra-particle interaction energy $U_{\alpha\alpha}(\{\boldsymbol{s}_{i\alpha}\})$ which is identical for both molecules. Vector $\boldsymbol{r}_{12}$ is defined by $\boldsymbol{r}_{12} = \boldsymbol{r}_2 - \boldsymbol{r}_1$, where $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are the effective coordinates of the two mesoscopic particles.[116] Note, that the effective coordinate $\boldsymbol{r}_\alpha$ is defined *via* some function of subparticle

coordinates: $r_\alpha = f(\{s_{i\alpha}\})$, *i.e.*, providing effective coordinate r_α and all monomer coordinates except for one, *e.g.*, $\{s_{i\alpha}\} \setminus \{s_{1\alpha}\}$, still determines the coordinates of the missing monomer ($s_{1\alpha}$ in the given example). Without loss of generality, we can assume $r_\alpha = s_{1\alpha}$ in order to calculate $\tilde{Z}_{tot}(r_{12})$ and $\tilde{Z}_{intra}$ *via*[117]

[117] See Blaak et al., *Accurate coarse-grained potentials for soft matter systems*, p. 221.

$$\tilde{Z}_{tot}(r_{12}) = \int \ldots \int \int \ldots \int ds_{11} \ldots ds_{M1} ds_{12} \ldots ds_{M2} \tag{2.73}$$

$$\times e^{-\beta U_{12}(\{s_{i1}\},\{s_{j2}\})} \delta(s_{11} - r_1) \delta(s_{12} - r_2)$$

$$\times e^{-\beta U_{11}(\{s_{i1}\})} e^{-\beta U_{22}(\{s_{j2}\})} = e^{-\beta F_c(r_{12})}$$

and

$$\tilde{Z}_{intra} = \int \ldots \int e^{-\beta U_{\alpha\alpha}(\{s_{i\alpha}\})} \delta(s_{1\alpha} - r_\alpha) \, ds_{2\alpha} \ldots ds_{M\alpha} \,. \tag{2.74}$$

In both Equations (2.73) and (2.74), the effective coordinates r_α and thus monomer coordinate $s_{1\alpha}$ are fixed using δ-functions. Furthermore, the constrained Helmholtz free energy, $F_c(r_{12})$ for two macromolecules fixed at effective coordinates r_1 and r_2 is defined in analogy with Equation (2.22) *via*

$$F_c(r_{12}) = -k_B T \ln[\tilde{Z}_{tot}(r_{12})] \,. \tag{2.75}$$

As $\tilde{Z}_{intra}$ is identical for both particles, the denominator on the right hand side of Equation (2.72), $\tilde{Z}_{intra}^2$, corresponds to the squared configuration part of the partition function due to the total intra-particle interaction energy in the system $U_{intra} = U_{11}(\{s_{i1}\}) + U_{22}(\{s_{j2}\})$.

Inserting Equations (2.73) and (2.74) into Equation (2.72) yields

$$e^{-\beta \phi_{eff}(r_{12})} = \left\langle e^{-\beta U_{12}(\{s_{i1}\},\{s_{i1}\})} \right\rangle_e \,, \tag{2.76}$$

where $r_\alpha = s_{1\alpha}$ and the ensemble average $\langle \ldots \rangle_e$ denotes an ensemble average with respect to a Hamiltonian whose interaction part does not contain $U_{12}(\{s_{i1}\}, \{s_{j2}\})$, *i.e.*, the Hamiltonian contains only contains the self-interactions of the individual macromolecules, U_{intra}, but not the in-

termolecule interaction, U_{inter}. Finally, by radially averaging over inter-molecule vector $\boldsymbol{r}_{12}$ we obtain

$$\phi_{\text{eff}}(r_{12}) = -k_{\text{B}}T \ln \left[\left\langle e^{-\beta U_{12}(\{s_{i1}\}, \{s_{j2}\})} \right\rangle_{\text{e}} \right] , \qquad (2.77)$$

with $r_{12} = |\boldsymbol{r}_2 - \boldsymbol{r}_1|$.

The WI method for calculating $\phi_{\text{eff}}(r_{12})$ can now be implemented in the following way: first a large number (N_{s}) of equilibrated and uncorrelated single-molecule configurations is generated *via* MD simulations. Then, the Widom insertions are performed by combining the N_{s} samples into $\lfloor N_{\text{s}}/2 \rfloor$ pairs and placing each pair of molecules in the simulation box so that their effective coordinates are located at a distance r_{12} from each other. Using Equation (2.77), the effective potential $\phi_{\text{eff}}(r_{12})$ can then be computed from the ensemble of these inserted pairs.

Thus, we have achieved a simple but powerful method of calculating a radially symmetric effective potential $\phi_{\text{eff}}(r_{12})$ between two macromolecules. Note, however, that the WI method is inefficient when interacting and non-interacting entities exhibit stark differences in their spatial configurations, especially at small inter-particle distances.[118]

[118] See Blaak et al., *Accurate coarse-grained potentials for soft matter systems*, p. 227.

2.6 Shape Descriptors

In this section, we will provide an overview over several key observables used in this work for the classification of shape, size and spatial position of the examined molecular structures.

2.6.1 Center of Mass and Geometric Center

In physics, the center of mass, r_{com}, of a system of N particles corresponds to the mass averaged first moment of the particles' positions:

$$r_{\text{com}} := \frac{1}{M} \sum_{i=1}^{N} r_i m_i \, . \tag{2.78}$$

Here, m_i and r_i denote the mass and the position of particle i, respectively, whereas $M = \sum_{i=1}^{N} m_i$ is the total mass.[119]

[119] The total mass M can also be defined as the zeroth moment of position weighted by mass m_0^*, *i.e.*, $M = m_0^* := \sum_{i=1}^{N} r_i^0 m_i$.

[120]The term "geometric center" is also used synonymously with the term "centroid".

[121] Likewise, N is equal to the unweighted zeroth moment of position: $m_0 := \sum_{i=1}^{N} r_i^0 = N$.

Similarly, the geometric center, r_{gc},[120] of a collection of N particles is determined by to the unweighted first moment of position averaged by N:[121]

$$r_{\text{gc}} := \frac{1}{N} \sum_{i=1}^{N} r_i \, . \tag{2.79}$$

In the case of particles with equal mass, the center of mass is the same as the geometric center. Because in this work only macromolecules consisting of identical particles are examined, the term "center of mass" is used synonymously with geometric center.

2.6.2 Gyration Tensor

For a collection of N particles, *e.g.*, a polymer chain, the gyration tensor T_{gyr} contains the averaged central moments and cross-moments of order two:

$$T_{\text{gyr}} := \frac{1}{N} \sum_{i=1}^{N} \left(r^i - r_{\text{gc}} \right) \otimes \left(r^i - r_{\text{gc}} \right) , \tag{2.80}$$

[122] See L. D. Landau and E. M. Lifshitz. *Mechanics, Third Edition: Volume 1 (Course of Theoretical Physics)*. 3rd. Butterworth-Heinemann, 1976. URL: http://www.worldcat.org/isbn/0750628960.

with $r^i - r_{\text{gc}}$ as the position vector of particle i with respect to the particles' geometrical center.[122] The matrix form in Cartesian coordinates of this second order tensor is given by

$$T_{gyr} = \frac{1}{N} \begin{bmatrix} \sum r_1^i r_1^i & \sum r_1^i r_2^i & \sum r_1^i r_3^i \\ \sum r_2^i r_1^i & \sum r_2^i r_2^i & \sum r_2^i r_3^i \\ \sum r_3^i r_1^i & \sum r_3^i r_2^i & \sum r_3^i r_3^i \end{bmatrix} , \tag{2.81}$$

where the summations are performed over all particles, *i.e.*, $\sum = \sum_{i=1}^{N}$, and r_m^i denotes the mth Cartesian coordinate of position vector $\boldsymbol{r}^i$ — again measured relative to $\boldsymbol{r}_{gc}$.

Solving the eigenproblem of the above matrix and expressing T_{gyr} in the eigenbasis yields the diagonalised gyration tensor:

$$T_{gyr}^* = \begin{bmatrix} \lambda_x^2 & 0 & 0 \\ 0 & \lambda_y^2 & 0 \\ 0 & 0 & \lambda_z^2 \end{bmatrix} , \tag{2.82}$$

with eigenvalues λ_i^2 $(i = x, y, z)$, which are also called the principal moments of the gyration tensor. The axes of the eigenbasis are chosen in such an order that $\lambda_x^2 \leq \lambda_y^2 \leq \lambda_z^2$.

In the case of particles with equal mass m the inertia tensor T_{in} is diagonal in the same frame of reference as T_{gyr}^* and is related to T_{gyr}^* *via* $T_{in} = m \left(\mathrm{Tr}\left(T_{gyr}^*\right) \mathbb{1} - T_{gyr}^* \right)$, where $\mathbb{1}$ is the second order identity tensor of rank three.[123]

[123] See K. Šolc. "Shape of a Random-Flight Chain". In: *J. Chem. Phys.* 55 (1971), pp. 335–344. DOI: 10.1063/1.1675527, pp. 335-336 for a more in-depth comparison of the two tensors.

2.6.3 Radius of Gyration

In statistical physics of polymers, the radius of gyration, R_g, is a measure used to describe the size of a molecular object consisting of N sub-particles, *e.g.*, a dendrimer.[124] The squared radius of gyration, R_g^2, equals to the trace of the gyration tensor, $\mathrm{Tr}\left(T_{gyr}\right)$, *i.e.*, the first invariant of T_{gyr}:

[124] See M. Doi. *Introduction to Polymer Physics.* Trans. by H. See. Oxford University Press, 1996. URL: http://www.worldcat.org/isbn/9780198517894, pp. 8-9.

$$R_g^2 = \frac{1}{N} \sum_{i=1}^{N} \left(\boldsymbol{r}_i - \boldsymbol{r}_{gc} \right)^2 = \mathrm{Tr}\left(T_{gyr}\right) , \tag{2.83}$$

where the same definitions as in Equations (2.80), (2.81), and (2.82) apply. Unfortunately several different definitions of a scalar quantity named "ra-

dius of gyration" exist (*e.g.*, the IUPAC definition). It is noted that in this work only the nonweighted form of Equation (2.83) is used.

2.6.4 Hydrodynamic Radius

The hydrodynamic radius, R_H, of a macromolecule in solution corresponds to the radius of a hard sphere that diffuses at the same rate as the solute.[125] If such a hard sphere undergoes diffusion in a viscous fluid, a frictional force f_d, also called drag force, is exerted on this particle. Stoke's law, which relates f_d to the sphere's radius and drift velocity, r and v, and solvent viscocity, η, is given by:[126]

$$f_d = 6\pi\eta r v \, . \tag{2.84}$$

As a result of the fluctuation-dissipation theorem, the diffusion constant D and the particle mobility $\mu = v/f_d$ are connected *via* the Einstein-Smoluchowsky relation:

$$D = \mu k_B T \, , \tag{2.85}$$

with Boltzmann constant k_B and Temperature T. Combining Equations (2.84) and (2.85) yields the Stokes-Einstein equation, which can then be used to calculate R_H after measuring D in experiments:

$$D = \frac{k_B T}{6\pi\eta R_H} \, . \tag{2.86}$$

In addition, we can relate the diffusion constant, D, to the mean squared displacement (MSD), $\langle \Delta \boldsymbol{r}^2(t) \rangle$, using Fick's law of diffusion *via*[127]

$$\langle \Delta \boldsymbol{r}^2(t) \rangle = 2 N_d D t \, , \tag{2.87}$$

where N_d denotes the number of dimensions.

2.6.5 Measures of Shape Anisotropy

The anisotropy of a molecules shape can be determined using the quantities described in the following.[128] In all definitions the eigenvalues are ordered

[125] Note that while there exists a theoretical definition of the hydrodynamic radius *via* the equation for the diffusion constant by Kirkwood (see Doi, *Introduction to Polymer Physics*, p. 85), here, the term hydrodynamic radius exclusively denotes the Stokes-Einstein radius as measured in experiments. Sometimes this term is used synonymously with "effective hydrated radius in solution", see P. W. Atkins and J. de Paula. *Atkins' Physical chemistry*. 11th. Oxford; New York: Oxford University Press, 2018. URL: http://www.worldcat.org/isbn/9780198769866.

[126] Equation (2.84) only holds true for small Reynolds numbers.

[127] See Hansen and McDonald, *Theory of Simple Liquids*, pp. 270–277.

[128] These shape metrics were first introduced in D. N. Theodorou and U. W. Suter. "Shape of unperturbed linear polymers: polypropylene". In: *Macromolecules* 18 (1985), pp. 1206–1214. DOI: 10.1021/ma00148a028, pp. 1207–1208.

by size *via* $\lambda_x^2 < \lambda_y^2 < \lambda_z^2$.

Asphericity b is given by

$$b = \lambda_z^2 - \frac{1}{2}(\lambda_x^2 + \lambda_y^2) = \frac{3}{2}\lambda_z^2 - \frac{R_g^2}{2} \ . \tag{2.88}$$

The case $b = 0$ only occurs for $\lambda_x^2 = \lambda_y^2 = \lambda_z^2$, *i.e.*, the distribution of particles is symmetric with respect to the $x-$, $y-$, and $z-$axes. This condition is met for spherically symmetric conformations, but also for other configurations such as a uniform distribution of particles on the faces of a cube.

Acylindricity, c, on the other hand, is defined *via*

$$c = \lambda_x^2 - \lambda_y^2 \tag{2.89}$$

and is only zero for $\lambda_x^2 = \lambda_y^2$. In analogy to the asphericity b, this case occurs for particle distributions which are cylindrically symmetric or if they are symmetric with respect to the $x-$ and $y-$axes, *e.g.*, a uniform distribution on the faces of a regular prism.

Finally, the relative shape anisotropy, κ^2, is given by

$$\kappa^2 = \frac{3}{2}\frac{\lambda_x^4 + \lambda_y^4 + \lambda_z^4}{(\lambda_x^2 + \lambda_y^2 + \lambda_z^2)^2} - \frac{1}{2} = \frac{b^2 + \frac{3}{4}c^2}{R_g^4} \ , \tag{2.90}$$

with $\kappa^2 \in [0, 1]$. While a spherically symmetric configuration yields $\kappa^2 = 0$, the case $\kappa^2 = 1$ only occurs if all particles are distributed on a single line.

3 DL-DNA: Experimental Synthesis and Computational Modelling

> *Conclusion: Big helix in several chains, phosphates on outside, phosphate-phosphate inter-helical bonds disrupted by water. Phosphate links available to proteins.*

—Rosalind Franklin[129]

The focus of this chapter lies on giving an overview of the investigated DNA-based dendrimers and the relevant theoretical and experimental aspects. First synthesised in 2003,[130] these novel macromolecular aggregates hold high promise in targeted self-assembly of complex soft matter systems in the bulk and at interfaces. In their work[130] Li *et al.* demonstrated that dendrimer-like DNA (DL-DNA) molecules can be fabricated in a controlled fashion *via* enzymatic ligation of Y-shaped DNA (Y-DNA) building blocks. Among the significant advantages of their synthetic strategy are high yield and purity (*i.e.*, monodispersity), unidirectional and stepwise growth, and non-reversible, efficient synthesis of high-generation DL-DNA. Furthermore, the size, structure and morphology of these all-DNA constructs can be changed by alterations to the building block. By attaching different functional groups the dendrimers' interactions can be enhanced.[131]

The rest of this chapter is organised as follows: In Section 3.1 a comprehensive introduction to DL-DNA molecules is provided, whereas two theoretical models with different levels of coarse-graining are described in Sections 3.2 and 3.3, respectively. These models are used as the basis for our simulations in the following Chapters 4 to 7. Finally, the experimental

[129] Lecture notes by Rosalind Franklin from November 1951, see A. Sayre. *Rosalind Franklin and DNA.* 1st ed. New York: W. W. Norton & Co., 2000. URL: https://worldcat.org/isbn/9780393320442.

[130] See Li et al., "Controlled assembly of dendrimer-like DNA"

[131]*E.g.*, peptide- and saccharide-conjugated dendrimers for drug delivery, see J. Liu et al. "Peptide- and saccharide-conjugated dendrimers for targeted drug delivery: a concise review". In: *Interface Focus* 2 (2012), pp. 307–324. DOI: 10.1098/rsfs.2012.0009.

synthesis of DL-DNA is presented in Section 3.4. Segments of this chapter were previously published.[132]

3.1 Dendrimer-Like DNA

Similarly to other dendrimers, this DNA-based dendrimer consists of a core unit and a fundamental building block which is repetitively attached to the free ends of the core, forming a dendritic structure. In this case, this fundamental unit is a Y-shaped DNA structure called Y-DNA consisting of three single-stranded DNA (ssDNA) elements. The nucleic acid sequence of every ssDNA element can be divided into two equally long segments, each of which has a complementary counterpart on one of the other ssDNA elements. The pairing of these segments is done in such a way that the hybridization of the three ssDNA strands produces the three-armed Y-DNA-structure seen in Figure 3.1. In our investigation, each of these arms is endowed with 13 nucleobase pairs (bp) and a single-stranded sticky end consisting of four single nucleobases (b).

FIGURE 3.1: Sketch of a Y-DNA structure: three ssDNA chains (colored red, green, and blue, respectively) assemble to form a double-stranded star-like configuration with single-stranded sticky ends.

FIGURE 3.2: Sketch for the synthesis of the G2 DL-DNA structure: four Y-DNA units, each consisting of three ssDNA chains (colored red, green, and blue, respectively) assemble *via* enzymatic ligation and form a G2 dendrimer.

While a single Y-DNA unit corresponds to the first and smallest dendrimer generation, G1,[133] iteratively attaching further Y-DNA elements yields DL-DNA molecules of higher generations, *e.g.*, the G2 DL-DNA shown in Figure 3.2. The connection between two Y-DNAs is achieved through enzymatic ligation, where single-stranded sticky ends of the arms of the two Y-DNAs form regular double-stranded DNA (dsDNA) *via* base-pairing. In this work, GN denotes a DL-DNA of generation N and its individual sub-

generations will be indicated by g_i, with $i = 1, 2, ..., N$. The constituent particle numbers, *e.g.*, the number of nucleobase pairs, N_{bp}, for DL-DNA of different generation index numbers N are listed in Table 3.1. Distinctive features of DL-DNA include its semi-flexibility — rigidity along the double-helical DNA arms and flexibility at the central junctions of the Y-DNAs —, its self-assembling nature, and its charge. Furthermore, the basic Y-DNA building blocks can be scaled up to sizes of tens of nm by increasing the the number of base pairs along the Y-DNA arms. Synthesizing a similarly versatile dendrimeric structure proves quite difficult for chemical dendrimers.[134]

GN	N_Y	N_{bp}	N_{mon}
G1	1	39	51
G2	4	168	192
G3	10	426	474
G4	22	942	1038
G5	46	1974	2166
G6	94	3666	4422
G7	190	8166	8934
G8	382	16422	17958

TABLE 3.1: The numbers N_Y, *i.e.*, the number of Y-DNAs, N_{bp}, *i.e.*, the number of base pairs, and N_{mon}, *i.e.*, the number of DNA monomers are listed for different DL-DNA generation numbers GN with $N \in \{1, ..., 8\}$.

[134] See Li et al., "Controlled assembly of dendrimer-like DNA", p. 42

3.2 Coarse-Grained Modelling of DL-DNAs

The growth of the constituent unit numbers, such as the number of nucleobase pairs N_{bp}, with generation index number N showcased in Table 3.1 is typical for dendrimers.[135] Especially the doubling of N_{bp} at higher N leads to a characteristic exponential behaviour $N_{bp} \propto 2^N$, see Figure 3.4 for a plot detailling this growth. When undertaking atomistic MD or MC simulations this aspect is often detrimental due to great computational cost of calculating the large number of occurring interactions. Hence, coarse-grained models are employed (or have to be employed) in order to avoid these computational limits and restrictions.[136]

A bead-spring models, a simple and widely used approach for simulating polymeric systems, *e.g.*, the Rouse model,[137] serves as the basis for our computational analysis. In this model nucleobase pairs of dsDNA and single nucleobases of ssDNA are represented by charged beads in a coarse-grained fashion.[138] These beads are connected *via* spring-like bonded interactions and serve as monomers in a polymer chain, which in turn represents the double helix of dsDNA. Accordingly, each Y-DNA arm is modeled as a chain of 17 charged monomers: a single dsDNA junction monomer followed by 12 dsDNA chain monomers, which are then succeeded by a single-stranded end group of four ssDNA monomers.

[135] "They double their molecular weight at each generation growth step, become progressively denser and more compact, and have a very low polydispersity." (Hawker and Fréchet, "Preparation of polymers with controlled molecular architecture. A new convergent approach to dendritic macromolecules", p. 7638)

[136] See H. I. Ingólfsson et al. "The power of coarse graining in biomolecular simulations". In: *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 4 (2014), pp. 225–248. DOI: 10.1002/wcms.1169

[137] See P. E. Rouse. "A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers". In: *J. Chem. Phys.* 21 (1953), pp. 1272–1280. DOI: 10.1063/1.1699180

[138] This particle-based model for dsDNA has been used in previous theoretical studies of polyelectrolytes such as DNA, *e.g.*, A. Wynveen and C. N. Likos. "Interactions between planar polyelectrolyte brushes: effects of stiffness and salt". In: *Soft Matter* 6 (2010), pp. 163–171. DOI: 10.1039/B919808C, p. 164.

FIGURE 3.3: Schematic representation of our DL-DNA model: a representation of a single Y-DNA according to our particle-based model is shown on the left; the middle panel shows how the dendrimer-like structure of DL-DNA which arises due to the union of Y-DNAs *via* enzymatic ligation; the rightmost panel shows the equivalence of two paired ssDNA ends to regular dsDNA at the connection of two Y-DNAs in our model. Shown are base pairs (**dark blue**) of the dsDNA arms and sequences of single nucleobases (light blue).
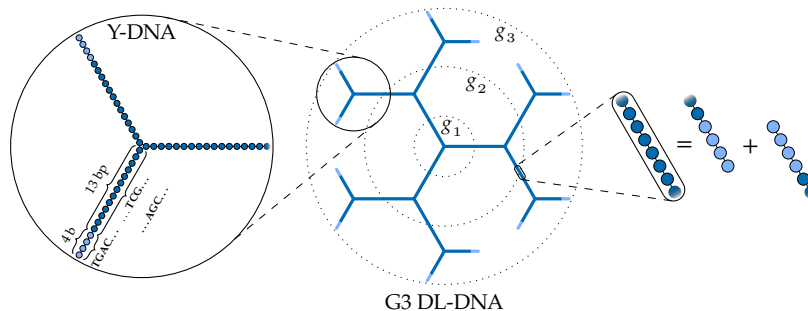

Y-DNA

G3 DL-DNA
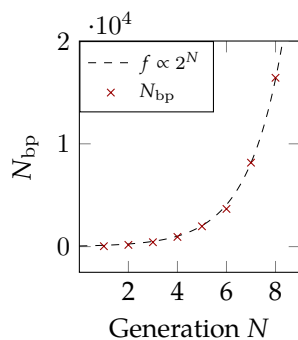
FIGURE 3.4: Exponential growth behaviour of number of bp, N_{bp}, as a function of generation index number N.

[139] See K. Kegler et al. "Forces of Interaction between DNA-Grafted Colloids: An Optical Tweezer Measurement". In: *Phys. Rev. Lett.* 98 (2007), p. 058304. DOI: 10.1103/PhysRevLett.98.058304

While the first thirteen monomers correspond to nucleobase pairs, the last four represent single nucleobases. The connection between two Y-DNAs *via* enzymatic ligation is established by replacing the four + four ssDNA monomers with four dsDNA monomers, see Figure 3.3. The numbers of constituents of DL-DNAs, *e.g.*, the number of Y-DNAs, N_{Y}, with different generation numbers GN are given in Table 3.1. The individual interaction models and their corresponding parameters for electrostatic interaction, steric interaction, and bonded interactions are given in Section 3.2.1. Recent experiments also tested the validity of this computational model in the context of forces between DNA-grafted colloids.[139] Employing coarse-grained models such as the one described in this section are invaluable when employing computational and numerical methods such as MD and MC simulations, as the involved computational cost is then greatly reduced.

3.2.1 Interactions in the Coarse-Grained Model

Electrostatic Interaction

The electrically charged phosphate backbone of dsDNA is emulated by providing each ssDNA- and dsDNA-monomer with a negative charge $q_{\mathrm{M^-}} = -e < 0$, where e is the elementary charge. In order to preserve the global electrostatic charge neutrality of the system, an equal amount of positively charged monovalent counterions with charge $q_{\mathrm{C^+}} = e > 0$ is introduced. Additionally, we also simulate systems with different concentrations of monovalent salt ions, $\mathrm{Na^+}$ and $\mathrm{Cl^-}$ (also denoted as $\mathrm{S^\pm}$), with

Species	m [u]	q [e]
$\mathrm{M^-}$ (ds)	660	-1
$\mathrm{M^-}$ (ss)	330	-1
$\mathrm{C^+}$	20	$+1$
$\mathrm{S^\pm}$	20	± 1

TABLE 3.2: Mass m, given in unified atomic mass units u, and charge q, given in units of the elementary charge e, of the system's particle species as specified in Wynveen and Likos, "Interactions between planar polyelectrolyte brushes: effects of stiffness and salt", p. 164.

the purpose of studying the influence of salt on the conformational characteristics of DL-DNAs. The charges and other physical properties of each particle type are listed in Table 3.2.

Any two charged species α and β then interact *via* the Coulomb interaction

$$V_{\text{Coul}}^{\alpha\beta}(r) = \lambda_{\text{B}}\frac{q_\alpha q_\beta}{r}k_{\text{B}}T \quad \text{with } \lambda_{\text{B}} = \frac{e^2}{4\pi\epsilon k_{\text{B}}T} = 7.0\,\text{Å}, \qquad (3.1)$$

where r denotes the interparticle separation and charges $q_\alpha, q_\beta \in \{e, -e\}$. The calculation of the Bjerrum λ_{B} length yields $7.0\,\text{Å}$ for an aqueous solution with uniform permittivity $\epsilon = 80.1\,\epsilon_0$ at temperature $T = 298\,\text{K}$.
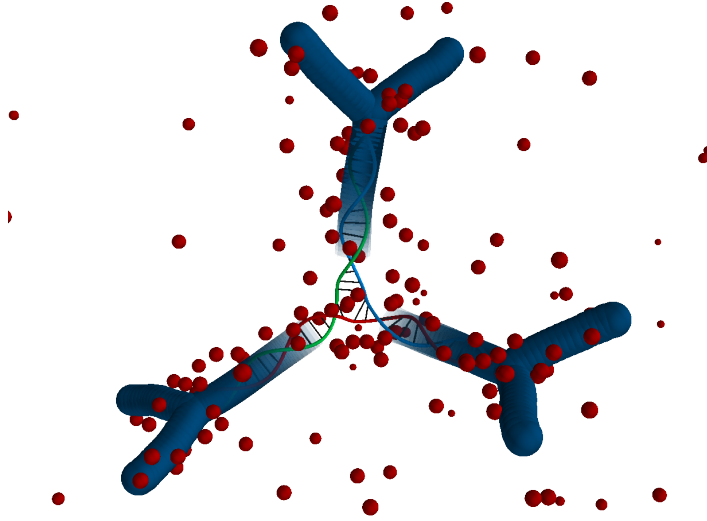

FIGURE 3.5: Representative simulation snapshot of a G2 DL-DNA dendrimer. The helices of dsDNA are represented by blue tubes. Counterions (C^+) are coloured red. Additionally, the actual DNA double helix is indicated in an artistic manner at the central junction of the DL-DNA. In this snapshot only a fraction of the simulation box is shown.

In order to handle the long-range nature of Coulomb interactions in simulations sophisticated methods such as Ewald summation have to be used, see Section 2.2.2. At salt concentrations $c \gtrsim 30\,\text{mM}$ it is computationally unfeasible to simulate large system volumes due to the immense number of salt particles. However, systems with high salt concentrations allow for the application of the Debye-Hückel theory, thus substantially reducing the computational costs as the range of the Coulomb interaction can be truncated by a cut-off range r_{cut} and salt ions do not have to be simulated explicitly. For more information about Debye-Hückel theory the reader is referred to the literature.[140]

[140]See Section 3.1 in C. N. Likos. "Colloidal interactions: From effective potentials to structure". In: *Riv. del Nuovo Cim.* 37 (2014), pp. 125–180. DOI: 10.1393/ncr/i2014-10098-1 and P. Debye and E. Hückel. "Zur Theorie der Elektrolyte. I. Gefrierpunktserniedrigung und verwandte Erscheinungen". In: *Physikalische Zeitschrift* 24 (1923), pp. 185–206.

Steric Interaction

Steric effects are implemented by introducing a truncated and shifted Lennard-Jones potential, which corresponds to the Weeks-Chandler-Andersen potential (WCA),[141] between any two particles separated by distance r:

$$
V_{\mathrm{LJ}}^{\alpha\beta}(r) = \begin{cases} \infty & \text{if } r < r_{\alpha\beta}\,, \\ 4\epsilon_{\mathrm{LJ}}\left[\left(\frac{\sigma_{\mathrm{LJ}}}{r-r_{\alpha\beta}}\right)^{12} - \left(\frac{\sigma_{\mathrm{LJ}}}{r-r_{\alpha\beta}}\right)^{6} + c_{\mathrm{shift}}\right] & \text{if } r_{\alpha\beta} \leq r \leq r_{\mathrm{cut}}\,, \\ 0 & \text{if } r_{\mathrm{cut}} < r\,, \end{cases} \quad (3.2)
$$

Species	Radius r_α $\left[\text{Å}\right]$
M^-	9
C^+	2
$S^\pm$	2

TABLE 3.3: The "van-der-Waals radii", r_α of the different particle species of the system, see Equation 3.2.

α \ β	M^-	C^+	$S^\pm$
M^-	14 Å	7 Å	7 Å
C^+	7 Å	0 Å	0 Å
$S^\pm$	7 Å	0 Å	0 Å

TABLE 3.4: The distance of divergence $r_{\alpha\beta}$ of the steric interaction for different combinations of particle species.

with the following parameters values: characteristic length scale $\sigma_{\mathrm{LJ}} = 4\,\text{Å}$, potential well depth $\epsilon_{\mathrm{LJ}} = 1.0\,\text{kJ}\,\text{mol}^{-1}$, and r-shift $r_{\alpha\beta} = r_\alpha + r_\beta - \sigma_{\mathrm{LJ}}$, with $\alpha, \beta \in \{M^-, C^+, S^+, S^-\}$, referring to monomers, counterions, and salt co- and counterions, respectively. The energy shift $c_{\mathrm{shift}} = \frac{1}{4}$ guarantees the continuity of the potential function at the cutoff radius. This way, the excluded volume interaction between counterion particles reduces to the usual WCA interaction, diverging at zero separation, while the monomer-monomer and monomer-ion steric potentials diverge at center-to-center distances $> 0\,\text{Å}$, accounting for the larger size of the monomers. The radii r_α for different particle types are given in Table 3.3, while the distances of divergence, $r_{\alpha\beta}$, for different combinations of particle species are listed in Table 3.4. Therefore, the steric interaction between DNA-monomers M^- and ions C^+, S^+, S^- acts in the range of

$$
r_{M^-C^+} = 7\,\text{Å} < r \leq \sqrt[6]{2}\sigma_{\mathrm{LJ}} + r_{M^-C^+} = 11.49\,\text{Å}\,. \quad (3.3)
$$

Using the value of radius r^*, where the energy of the steric potential is equal to the energy of thermal fluctuations, *i.e.*, $V_{\mathrm{LJ}}^{M^-C^+}(r^*) = k_{\mathrm{B}}T$, as a measure of the steric range, we find that $r^* = 10.43\,\text{Å}$. This finding matches well the corresponding value of the effective dsDNA helix diameter of approximately $2\,\text{nm}$ from literature.[142]

Bonded interactions

A polymer chain of such DNA-monomers is then realised by joining consecutive monomers with indices i and j along ssDNA- and dsDNA-strands *via* the harmonic bond potential

$$V_{\mathrm{b}}\left(r_{ij}\right) = \frac{k_{\mathrm{b}}}{2}\left(r_{ij} - l_{\mathrm{b}}\right)^2 , \qquad (3.4)$$

with intermonomer distance r_{ij}, equilibrium bond length $l_{\mathrm{b}} = 3.4\,\text{Å}$, and spring constant $k_{\mathrm{b}} = 210\,\mathrm{kJ/(mol \cdot Å^2)}$, corresponding to $\approx 10^3 k_{\mathrm{B}} T \cdot l_{\mathrm{b}}^2$. This choice of parameters results in the typical distance between neighbouring base pairs of B-DNA[143] of $3.4\,\text{Å}$ and yields a dispersion in the rise[144] of about $0.15\,\text{Å}$, which is consistent with the value observed in structural studies of DNA.[145] The basic idea of this harmonic bond is illustrated in Figure 3.6. Comparing the bond length $l_{\mathrm{b}} = 3.4\,\text{Å}$ with the steric monomer-monomer interaction offset $r_{\mathrm{M-M-}} = 14\,\text{Å}$ reveals that neighbouring monomers in a polymer chain of straight configuration are located in the divergent regime of the WCA potential. Therefore, the WCA steric interaction is configured to only act between monomers which do not pertain to the same arm and additionally exclude the steric interaction between the first five monomers located at the Y-DNA junctions and connections.

By additionally introducing a harmonic angle bond potential $V_{\phi}(\phi)$ of form

$$V_{\phi}\left(\phi\right) = \frac{k_{\phi}}{2}\left(\phi - \pi\right)^2 , \qquad (3.5)$$

which acts on angle ϕ between any monomer and its two neighbouring monomers, the rigidity and thus the persistence length of dsDNA are properly replicated. In vivo ($c \approx 150\,\mathrm{mM}$) the persistence length typically assumes a value of $l_{\mathrm{p}} = 500\,\text{Å}$.[146] Here, angle ϕ denotes the angle between intermonomer vectors $\boldsymbol{r}_{ij}$ and $\boldsymbol{r}_{jk}$, connecting three consecutive monomers with indices i, j, and k, see Figure 3.7. If angle ϕ deviates from the equilibrium angle value $\phi = \pi$ the angle potential $V_{\phi}(\phi)$ exerts a torque, forcing monomers i, j, and k back towards a linear configuration.

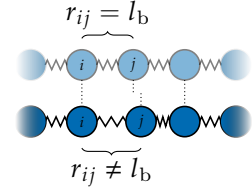In order to reproduce the bending behaviour of ssDNA and dsDNA in the

FIGURE 3.6: An illustration of the harmonic bond as defined by Equation (3.4). The springs in this image represent the forces resulting from a deviation of the intermonomer distance r_{ij} from the equilibrium distance l_{b}.

[143] B-DNA is one of the three possible structures of the DNA double helix, see D. L. Beveridge et al. "The ABCs of molecular dynamics simulations on B-DNA, circa 2012". In: *J. Biosci.* 37 (2012), pp. 379–397. DOI: 10.1007/s12038-012-9222-6.

[144] Rise, *i.e.* the distance between two adjoining nucleobase pairs along the double helix axis, is one of the structural parameters defining the geometry of dsDNA, see R. Dickerson. "Definitions and nomenclature of nucleic acid structure components". In: *Nucleic Acids Res.* 17 (1989), pp. 1797–1803. DOI: 10.1093/nar/17.5.1797, p. 1800.

[145] The parameter value for l_{b} is taken from Wynveen and Likos, "Interactions between planar polyelectrolyte brushes: effects of stiffness and salt", p. 164. For more information on the conformational properties of DNA see A. Wynveen et al. "Helical coherence of DNA in crystals and solution". In: *Nucleic Acids Res.* 36 (), pp. 5540–5551. DOI: 10.1093/nar/gkn514.
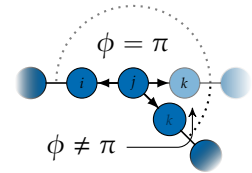

FIGURE 3.7: An illustration of the harmonic angle bond as defined by Equation (3.5). A deviation of angle ϕ from the equilibrium constellation $\phi = \pi$, results in a restoring torque.

[146] See J. F. Marko and E. D. Siggia. "Stretching DNA". in: *Macromolecules* 28 (1995), pp. 8759–8770. DOI: 10.1021/ma00130a008, p. 1860.

polymer chains and at the junctions, the angle bond constant k_ϕ assumes different values:

$$k_\phi = \begin{cases} 750 \, \text{kJ} \, \text{mol}^{-1} & \text{for stiff chains,} \\ 150 \, \text{kJ} \, \text{mol}^{-1} & \text{for sticky ends,} \\ 0 \, \text{kJ} \, \text{mol}^{-1} & \text{for flexible junctions.} \end{cases}$$

Since the persistence length of unpaired ssDNA is lower than for ds-DNA,[147,148] the degree of flexibility of the ssDNA end group is set to a lower value $k_\phi = 150 \, \text{kJ} \cdot \text{mol}^{-1}$. Furthermore, the central junctions of the Y-DNAs are fully flexible, *i.e.*, the bending energy constant is chosen to be zero. In Figure 3.8 a single Y-DNA arm is shown with illustrations of all bond potentials (*i.e.*, harmonic and harmonic angle bond potentials) and the corresponding values of their bond parameters.

[147] See B. Tinland et al. "Persistence Length of Single-Stranded DNA". in: *Macromolecules* 30 (1997), pp. 5763–5765. DOI: 10.1021/ma970381+

[148] See H. Chen et al. "Ionic strength-dependent persistence lengths of single-stranded RNA and DNA". in: *Proc. Natl. Acad. Sci. U.S.A.* 109 (2012), pp. 799–804. DOI: 10.1073/pnas.1119057109

FIGURE 3.8: Representative sketch of a single arm of Y-DNA in the coarse-grained model. Shown are base pairs (**dark blue**) of the dsDNA arm which is terminated by a ssDNA strand (**light blue**).
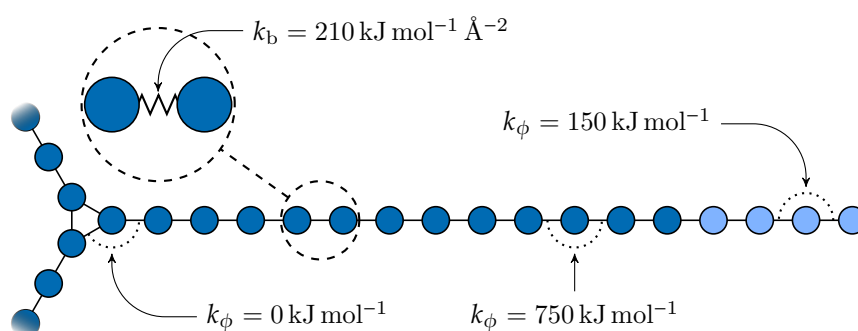
[149] Simulation details and parameters are presented in Appendix B.

[150] LAMMPS simulations were performed by Nataša Adžić from the Faculty of Physics at the University of Vienna, whereas ESPResSo simulations were carried out by the author.

[151] See Appendix A for more information on the three simulation packages used in this work.

The model described above is utilised in our MD simulations,[149] which were performed using the simulation packages ESPResSo and LAMMPS.[150] Depending on the model and simulation package with which the simulations were performed, the results in the following Chapters 4 to 7 will be labelled accordingly, *i.e.*, ESPResSo, LAMMPS, and oxDNA (which will be introduced in Section 3.3).[151] No labelling indicates that the results were obtained using ESPResSo.
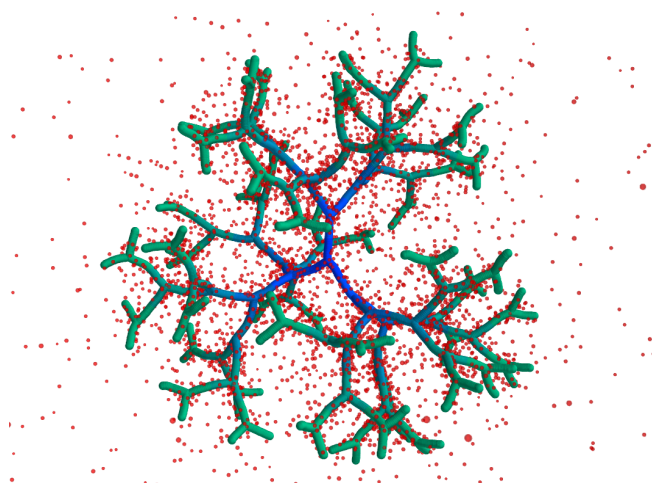
A representative simulation snapshot of a G6 DL-DNA dendrimer obtained *via* ESPResSo is displayed in Figure 3.9.
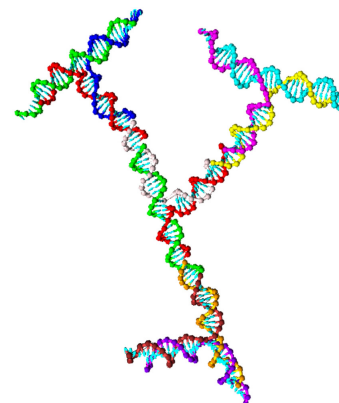

FIGURE 3.10: A simulation snapshot of a G2 DL-DNA with rigid connetions in the oxDNA2 model. Differently coloured strands correspond to the different ssDNA strands listed in Section 3.4. Image courtesy of Nataša Adžić.

3.3 The oxDNA Model

As a complementary method to our coarse-grained approach the oxDNA code was used as a more detailed model, which explicitly represents the double helical structure of DNA.[152,153] In this model DNA is considered as a string of rigid nucleotides, interacting via potentials which depend on the nucleotides' position and orientation.

The reader is referred to Appendix A.3 for a more in-depth review of oxDNA. In the simulations a more recent version of the oxDNA model was used.[154]

3.4 Experimental Synthesis

The experimental part of our FWF-project involved synthesising DL-DNA molecules and measuring their structural properties, *e.g.*, their static form factors and structure factors, using light scattering experiments.[155] This synthesis was performed by employing a one-pot approach and following a two-step assembly process established in previous research. To briefly

[152]oxDNA simulations were performed by Nataša Adžić from the Faculty of Physics at the University of Vienna.

[153]This model was first published in T. E. Ouldridge et al. "Structural, Mechanical, and Thermodynamic Properties of a Coarse-Grained DNA Model". In: *J. Chem. Phys.* 134 (2011), p. 085101. DOI: 10.1063/1.3552946. For a more thorough explanation of the model see T. E. Ouldridge. "Coarse-grained modelling of DNA and DNA self-assembly". PhD thesis. Oxford University, UK, 2011. URL: https://ora.ox.ac.uk/objects/uuid:b2415bb2-7975-4f59-b5e2-8c022b4a3719.

[154] See B. E. K. Snodin et al. "Introducing improved structural properties and salt dependence into a coarse-grained model of DNA". in: *J. Chem. Phys.* 142 (2015), p. 234901. DOI: 10.1063/1.4921957.

[155]Experimental synthesis and analysis were performed by Emmanuel Stiakakis at the Institute of Complex Systems (ICS-3) at Forschungszentrum Jülich.

[156]For the previously published synthtetic procedure see Um et al., "Dendrimer-like DNA-based Fluorescence Nanobarcodes".

summarise the assembly process:[156] first the Y-DNAs are generated *via* self-assembly, then DL-DNAs are synthesised *via* enzyme-assisted assembly. In order to minimise the total number of different strands necessary, the ssDNA sequences were slightly modified in comparison to the previously established procedure by Um *et al.*

Below a list of ssDNA sequences is presented. These sequences are used for synthesizing the Y-DNA and were designed using the program SEQUIN.[157]

[157] See https://www.ncbi.nlm.nih.gov/projects/Sequin/index.html.

The bold letters correspond to the sticky-end sequence and p indicates the position of the phosphate modification. The 5' terms and 3' terms indicate the carbon numbers in the DNA's sugar backbone at that position and determine the directionality of the ssDNA. DNA strands used in this study were purchased from Integrated DNA Technologies, Inc.[158] By measuring the absorbance at 260 nm using a microvolume spectrometer the concentrations of the ssDNA strands could be determined.[159]

[158] See https://www.idtdna.com.

[159] The microvolume spectrometer in use was a NanoDrop™ 2000 by Thermofisher Scientific.

Nucleobase sequences of ssDNA

Strand (a):

- ■ S_{1a} : 5'-p-**TGAC**-TGGATCCGCATGACATTCGCCGTAAG-3'

- ▲ S_{2a}: 5'-p-**GTCA**-TGGATCCGCATGACATTCGCCGTAAG-3'

- ♦ S_{3a}: 5'-p-**ATCG**-TGGATCCGCATGACATTCGCCGTAAG-3'

- ▼ S_{4a}: 5'-p-**GCAA**-TGGATCCGCATGACATTCGCCGTAAG-3'

Strand (b):

- ■ S_{1b}: 5'-p-**TGAC**-CTTACGGCGAATGACCGAATCAGCCT-3'

- ▲ S_{2b}: 5'-p-**CGAT**-CTTACGGCGAATGACCGAATCAGCCT-3'

- ♦ S_{3b}: 5'-p-**TTGC**-CTTACGGCGAATGACCGAATCAGCCT-3'

- ▼ S_{4b}: 5'-p-**GTCA**-CTTACGGCGAATGACCGAATCAGCCT-3'

Strand (c):

- ■ S_{1c}: 5'-p-**TGAC**-AGGCTGATTCGGTTCATGCGGATCCA-3'

- ▲ S_{2c}: 5'-p-**CGAT**-AGGCTGATTCGGTTCATGCGGATCCA-3'

- ◆ S_{3c}: 5'-p-**TTGC**-AGGCTGATTCGGTTCATGCGGATCCA-3'

- ▼ S_{4c}: 5'-p-**GTCA**-AGGCTGATTCGGTTCATGCGGATCCA-3'

By annealing of the three partially complementary ssDNA S_{1a}, S_{1b}, and S_{1c} at equal molar ratio and employing a one-pot approach the core Y-DNA was synthesised. This Y-DNA, which we denote as a DL-DNA of the first generation (G1), has each of its arms terminated by a non-palindromic sticky-end with a length of 4 nucleobases. Analogously, the building blocks Y_i, with $i = 1, 2, 3, 4$, of all higher generation DL-DNA can be synthesised using the corresponding strands S_{ia}, S_{ib}, and S_{ic}. See the construction scheme of Y_i below for a generalised formulation of this process. This self-assembly of Y-DNAs corresponds to the aforementioned first step of the procedure. In the second synthesis step, these Y-DNA building blocks are then connected *via* enzymatic ligation to form the DL-DNA nanostructures. For example: to build a second generation DL-DNA (G2), which is the next generation after G1, we generate the core Y-DNA (Y_1) along with three Y_2 The enzyme-assisted assembly is realised by hybridising the complementary sticky ends of the tri-functional Y-DNA cores using T4 DNA ligase.[160]

Successive generations (G3, ..., G6) can then be created by generating the corresponding Y_i elements *via* self-assembly and synthesising the DL-DNA structure from these elements using the enzyme-assisted assembly. A generalised version of this protocol is presented below. After the synthesis additional steps of purification were performed in order to guarantee a monodisperse sample of DL-DNAs of the desired generation.[161]

[160]The T4 DNA ligase was obtained from the Promega Corporation (https://promega.com).

[161]Purification of G1 and G2 DL-DNAs was achieved using ultra spin columns by the Amicon corporation (https://www.amico.com), while dialysis was used to purify GN DL-DNAs (with $N \geq 3$) and remove salt excess.

Construction scheme

Y-DNAs:

- $Y_1 = S_{1a} + S_{1b} + S_{1c}$

- $Y_2 = S_{2a} + S_{2b} + S_{2c}$

- $Y_3 = S_{3a} + S_{3b} + S_{3c}$

- $Y_4 = S_{4a} + S_{4b} + S_{4c}$

- $Y_5 = Y_2 = S_{2a} + S_{2b} + S_{2c}$

- $Y_6 = Y_3 = S_{3a} + S_{3b} + S_{3c}$

- $Y_i = S_{ma} + S_{mb} + S_{mc}$ with $m = (i \mod 4) + 1$

DL-DNAs:

- 1st generation DL-DNA: $G1 = Y_1$

- 2nd generation DL-DNA: $G2 = G1 + 3 \times Y_2$

- 3rd generation DL-DNA: $G3 = G2 + 6 \times Y_3$

- 4th generation DL-DNA: $G4 = G3 + 12 \times Y_4$

- 5th generation DL-DNA: $G5 = G4 + 24 \times Y_5$

- 6th generation DL-DNA: $G6 = G5 + 48 \times Y_6$

- N^{th} generation DL-DNA: $GN = G(N-1) + 3 \cdot 2^{N-2} \times Y_N$

Finally, the succesful assembly of DL-DNAs was confirmed *via* agarose gel electrophoresis, see Figure 3.11. Figure 3.11 demonstrates the decreasing mobility of the macromolecules with increasing generation number GN. The DNA constructs migrate from the top of the figure towards the bottom in single bands. Bands towards the bottom of 3.11 correspond to entities with a higher mobility and thus a smaller size. The DL-DNA molecules exhibit high flexibility; this explains why the G6 molecule, which, according to Table 3.1, corresponds to roughly $4\,\text{kbp}$, scores higher on mobility than the $3\,\text{kbp}$ marker (third band in lane M from the bottom). The disparity in the precision and sharpness of the bands is a consequence of two phenomena: DL-DNAs of higher generation can explore larger variety of confirmations

with different diffusive behaviours and thus a less well-defined band; dendrimers of lower generations, *e.g.*, G1, exhibit a diffuse band due to the larger distance they travelled through the agarose gel.
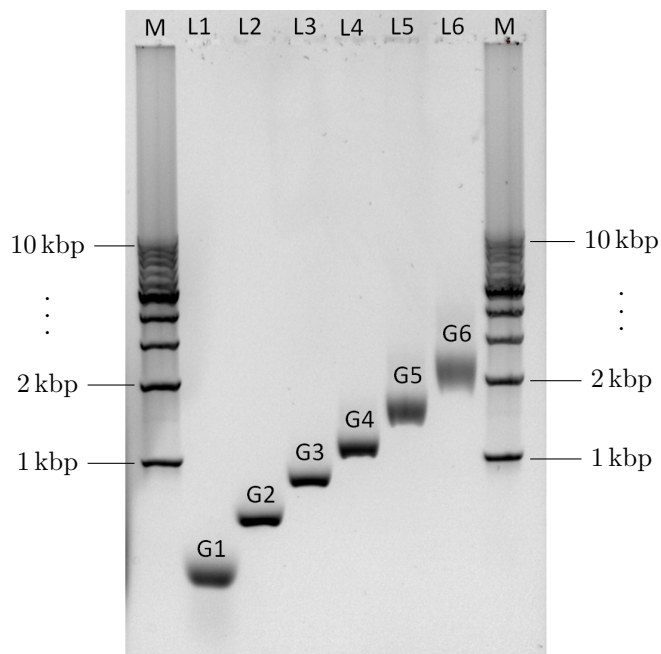

FIGURE 3.11: Gel electrophoresis analysis of DL-DNA structures of generation index GN, with $N = 1, \ldots, 6$. Lanes LN, with $N = 1, \ldots, 6$, show the results for GN, with $N = 1, \ldots, 6$, respectively. The dendrimers electrophoretic mobility is demonstrated by using non-denaturing agarose gel (0.5%). In lane M a DNA-ladder with 10 ds-DNA markers from 1 kbp to 10 kbp with step size 1 kbp (as labelled) is presented. This figure is adapted from Jochum et al., "Structure and stimuli-responsiveness of all-DNA dendrimers: theory and experiment", p. 1608. Image courtesy of Manolis Stiakakis.

4 Dilute Solutions of DNA-Based Dendrimers

In the following chapter a thorough investigation of DL-DNAs in dilute solutions is presented.[162] Due to the large intermolecular distances resulting from the low density we treat the DNA-dendrimers as isolated, non-interacting macromolecules. The goal of this chapter is the comparison of the theoretical and experimental results obtained by investigating such systems. While the theoretical coarse-grained model, which was presented in Section 3.2, is implemented in MD simulations, the experimental synthesis was carried out in accordance with Section 3.4 and light scattering experiments were performed.[163]

The rest of this chapter is organised as follows: In the first Section 4.1 quantities regarding the overall size of DL-DNAs, *i.e.*, hydrodynamic radius, R_{H}, and radius of gyration, R_{g}, are presented, comparing results obtained from simulations and experimental results. The subsequent Section 4.2 contains further information on the static structure of the dendrimers, *e.g.*, the radial distribution function $g_{\mathrm{mm}}(r)$, whereas Sections 4.3 and 4.4 provides a host of conformational details which are solely accessible in simulations, *e.g.*, distance and angle distributions. Finally, Section 4.5 examines the conformational responsiveness of DL-DNAs to salt stimuli. Some of the results presented in this chapter were previously published.[164]

The LAMMPS and oxDNA simulations in this chapter were performed by Nataša Adžić, while ESPResSo simulations were carried out by the author. The reader is referred to Appendix A for more information on the software packages. The relevant simulation parameter values and simulation details

[162]This investigation is the continuation of a preliminary study by Dominic A. Lenz, see D. A. Lenz. "Self-organization of dendrimers and dendrimer-colloid mixtures". PhD thesis. University of Vienna, AT, 2012. URL: http://othes.univie.ac.at/21039/.

[163]Experiments were performed by Emmanuel Stiakakis at the Institute of Complex Systems (ICS-3) at Forschungszentrum Jülich.

[164]See Jochum et al., "Structure and stimuli-responsiveness of all-DNA dendrimers: theory and experiment".

are given in Appendix B.2.

4.1 Comparing Experiment and Simulation

[165]Measuring the friction coefficient (f_d/v in Equation (2.86)) of the molecule in a constant flow field *via* non-equilibrium MD simulations and applying Equation (2.86) yields R_H. See Equation (5) in L. B. Weiss et al. "Computation of the Hydrodynamic Radius of Charged Nanoparticles from Nonequilibrium Molecular Dynamics". In: *J. Phys. Chem. B* 122 (2018), pp. 5940–5950. DOI: 10.1021/ acs.jpcb.8b01153, p. 5493. R_g can be calculated from particle coordinates of an appropriate simulated ensemble using Equation (2.83).

[166]For dynamic properties hydrodynamic interactions have to be taken into account, *e.g.*, *via* dissipative brownian dynamics (DPD), see P. J. Hoogerbrugge and J. M. V. A. Koelman. "Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics". In: *EPL* 19 (1992), pp. 155–160. DOI: 10.1209/0295-5075/19/3/001. This leads to a significant increase in computational cost.

FIGURE 4.1: Hydrodynamic radius R_H (as extracted from experiment) and radius of gyration R_g (as predicted by simulation) as functions of the generation index of DL-DNAs. The experimental values for R_g are provided for $G5$ and $G6$. Experimental data courtesy of Manolis Stiakakis; LAMMPS data courtesy of Nataša Adžić.

[167] The radius of gyration R_g is extracted by employing static light scattering measurements in the Guinier regime ($qR_g < 1$) using a 632 nm laser. However, DL-DNAs at low generations, *e.g.*, G1, do not scatter sufficiently due to their small size and therefore no meaningful value for R_g can be extracted.

[168] See G. R. Strobl. *The Physics of Polymers*. Springer, 1996. DOI: 10. 1007/978-3662032435, p. 290.

In order to characterise the overall size of a single isolated DL-DNA two quantities have proven especially useful: the radius of gyration, R_g, and the hydrodynamic radius, R_H. Both can in principle be determined from simulations;[165] they are also experimentally accessible by means of different scattering techniques, *e.g.*, small-angle neutron scattering (SANS), small-angle X-ray scattering (SAXS), or dynamical light scattering (DLS). However, in practical applications one has to face problems: while R_g is readily accessible in MD simulations, measuring R_H in such simulated systems comes at high computational cost. In simulations the investigation of static properties, *e.g.*, R_g, is usually more feasible than the investigation of dynamic properties, *e.g.*, R_H.[166] Experimentally, the situation is reversed as experimental measurement of R_g is proving quite difficult for dendrimers of lower generation number GN, with $N < 5$.[167] For polymers these two quantities are often similar in magnitude.[168]
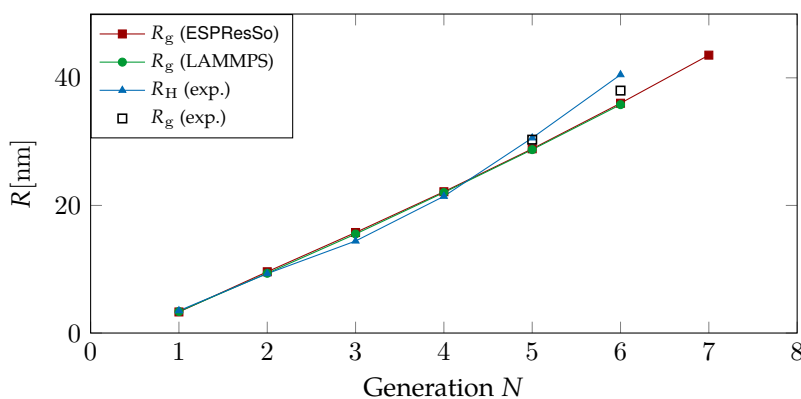
Due to these circumstances, we found it suitable to compare the radius of gyration, R_g, predicted by simulation to the hydrodynamic radius, R_H, extracted from experiment. Though these two radii are different by definition — one measuring spatial extent of the molecule and the other measuring its hydrodynamic drag — they differ in their values only by a small amount

so that a comparison of R_g from simulation with R_H from experiment is an appropriate way to validate the model.

A comparison of the results for R_g and R_H originating from experiment and simulation is presented in Figure 4.1. No salt ions were added in simulations, whereas in experiment an electrostatically negligible amount of salt ($c = 1\,\mathrm{mM}$) was added in order to ensure stability of the macromolecules. The results for radius of gyration, R_g, obtained from simulation were calculated using Equation (2.83). The hydrodynamic radius R_H, on the other hand, was measured in experiments using DLS and determining the diffusion coefficient D of DL-DNAs in a dilute solution. The suitability of the applied model is demonstrated by the excellent agreement between the results of R_g and R_H in the range from G1 to G5. Due to the increasing difference between R_g and R_H for DL-DNAs of higher generation numbers, *i.e.*, for G6 DL-DNAs, additional experimental values of R_g for G5 and G6 DL-DNAs are provided. These supplementary experimental results were assessed *via* static light scattering (SLS). Figure 4.1 shows how experimental results for the radius of gyration and the hydrodynamic radius coincide at generation number G5, *i.e.*, $R_g = R_H$; this observation further justifies our choice of comparing two different quantities that characterise the size of a dendrimers of lower generation numbers.

The aforementioned discrepancies between R_H and R_g can also be understood in terms of the functional forms of $R_H(GN)$ and $R_g(GN)$ as function of generation index GN: while $R_g(GN)$ exhibits a linear behaviour, $R_H(GN)$ clearly assumes a concave shape leading to a diverging difference between the two quantities. This behaviour reflects the non-linear growth of the dendrimer with increasing generation number GN and can be explained by the change in the macromolecules' sphericity: as the monomer density at the periphery of the DL-DNA grows with increasing generation number GN, the molecule assumes a more spherical shape.[169]

[169] As verified by the eigenvalues of the gyration tensor of the DL-DNAs in Table C.3 in Appendix C.4.

An immediate consequence of the different growth dynamics of R_g and R_H, as measured in experiments, is the decrease of their ratio from a value of around 1.0 towards smaller values, *e.g.*, $R_g/R_H = 0.94$ for G6. The

amplification of this trend could be further observed in the regime of even higher generations, *i.e.*, GN, with $N \geq 7$, a region unfortunately outside of our experimental possibilities. Examining relevant literature reveals that the theoretical value of ratio $R_{\mathrm{g}}/R_{\mathrm{H}}$ is 0.778 for a homogeneous hard sphere and 1 for a hollow sphere with an infinitely thin shell.[170] In combination with the polydispersity of the synthesised DL-DNAs in experiment this leads us to assume that the observed difference between the experimental R_{H} and R_{g} obtained from simulation for GN, with $N \geq 5$ is reasonable and in agreement with literature.[171] Further proof could be provided by obtaining experimental results for R_{g}, *e.g.*, by employing SLS, for additional values of generation index GN, with $N < 5$. In general, the experimental results and data obtained from simulation show excellent agreement.

More information on the structural properties of DL-DNAs is provided by the monomer-monomer form factor $F_{\mathrm{mm}}(k)$ (defined in Equation (2.51)). Figure 4.2 shows a comparison of experimental results obtained from SLS and theoretical results originating from simulations. Qualitatively, the results presented in Figure 4.2 are in good agreement.

[170] See G. S. Grest et al. *Advances in Chemical Physics. Polymeric Systems.* Vol. 94. New York: J. Wiley & Sons, 1996, pp. 67–163. DOI: 10.1002/9780470141533 and S. U. Egelhaaf and P. Schurtenberger. "Shape Transformations in the Lecithin-Bile Salt System: From Cylinders to Vesicles". In: *J. Phys. Chem.* 98 (1994), pp. 8560–8573. DOI: 10.1021/j100085a041.

[171] See Grest et al., *Advances in Chemical Physics*.

FIGURE 4.2: Monomer-monomer form factor $F_{\mathrm{mm}}(k)$ for a DL-DNAs of generation G6. Left: experimental data along with the Guinier fit to $F_{\mathrm{mm}}(k)$ as a function of k^2, see Eq. (2.52). Right: comparison of $F_{\mathrm{mm}}(k)$ as a function of k between experimental and simulation data. Experimental data courtesy of Manolis Stiakakis.
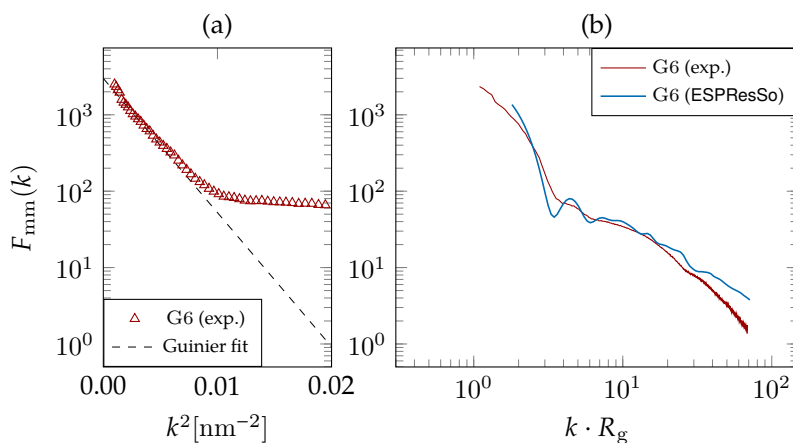
A more detailed analysis of the monomer-monomer form factor $F_{\mathrm{mm}}(k)$ of dendrimers obtained from simulation, which provides a deeper insight into the structural properties of DL-DNA, can be found in Section 4.2.

4.2 Structural Analysis from Simulation

By analyzing the monomer-monomer pair correlation function, also known as the monomer-monomer radial distribution function $g_{\mathrm{mm}}(r)$,[172] additional insight into the conformational features of th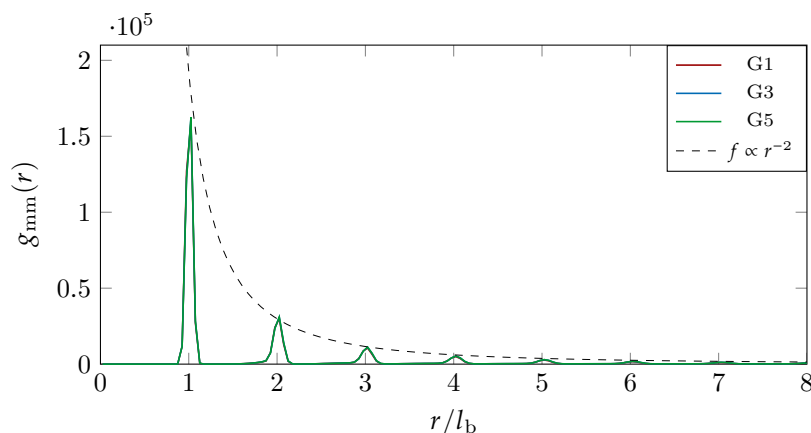e internal structure of the examined dendrimers can be obtained. Figure 4.3 shows the radial distribution function $g_{\mathrm{mm}}(r)$ for DL-DNAs of generations G1, G3, and G5 functions of distance r, given in units of the equilibrium bond length $l_{\mathrm{b}} = 3.4\,\text{Å}$. In this regime of $r \leq 8l_{\mathrm{b}}$ the three curves of $g_{\mathrm{mm}}(r)$ for GN, with $N = 1, 3, 5$, are virtually identical.

[172] See Section 2.3.1.

FIGURE 4.3: Monomer-monomer radial distribution function $g_{mm}(r)$ for DL-DNA of generations G1, G3, and G5 as functions of distance r, given in units of equilibrium bond length l_{b}. A function $f \propto r^{-2}$ fitted to the local maxima is represented by the dashed line.

The well-defined maxima which are located at equidistant positions and seperated by length l_{b} indicate that the bonds between the monomers are rather stiff. The first and largest peak represents the nearest neighbour separation along the Y-DNA arms. The dashed line in Figure 4.3 shows that the height of the local maxima goes with r^{-2}, which is the inverse of the rate at which the volume of the spherical shells of thickness $\mathrm{d}r$ increases with sphere radius r. Thus, one can assume a constant monomer density ρ_{m} along the dendrimer arms.

The radial density density profile, $\rho(r)$, reveals detailed information about the interior scaffolding of DL-DNAs.[173]

[173] See Equation (2.53) for the definition of $\rho(r)$.

FIGURE 4.4: Radial density profiles $\rho(r)\sigma^3$ of monomers and counterions, where σ is the steric interaction distance, as functions of r, given in units of equilibrium bond length $l_{\rm b}$; $\rho(r)$ is shown for different entities of the system (as labelled and see text). Results are shown for DL-DNAs of (a) G2, (b) G4, (c) G6, and (d) G8.
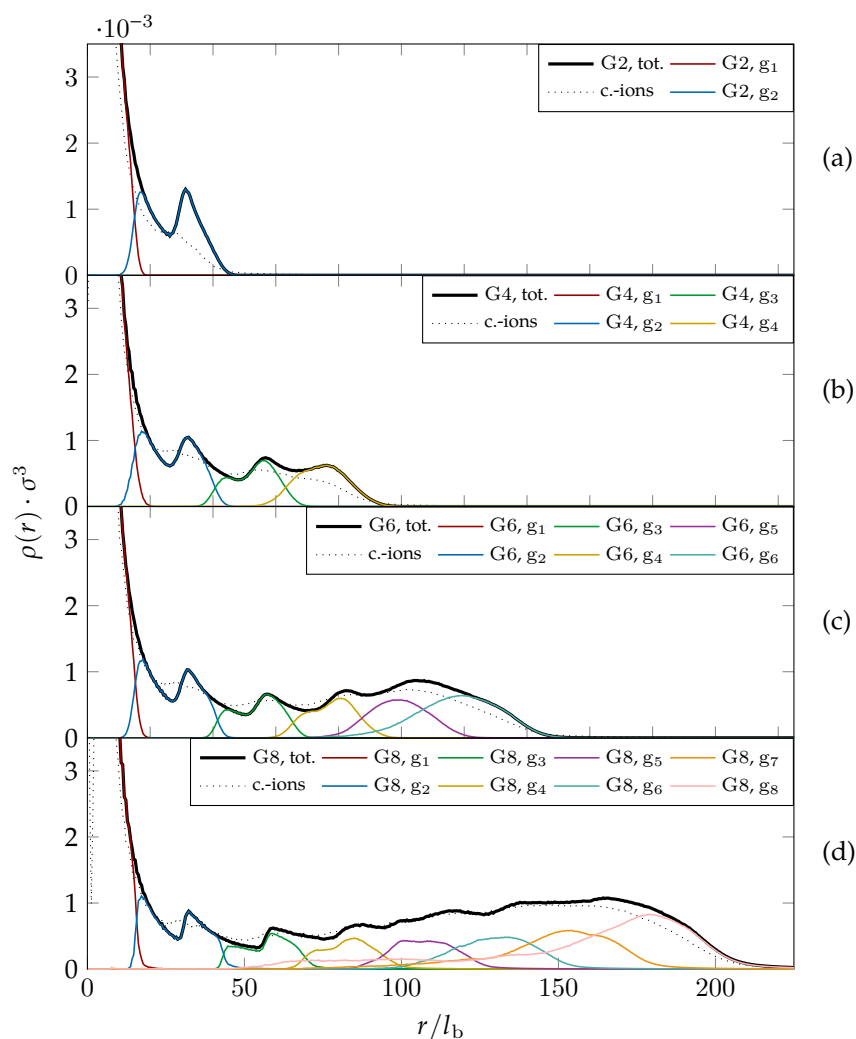
In Figure 4.4 the density profiles, $\rho(r)$ for specific components of DL-DNA molecules are shown, showing dendrimers of sizes G2, G4, G6, and G8. The different entities, that are considered are: (i) all monomers (without distinction; "total"), (ii) the monomers pertaining to a specific subgeneration, g_i, and (iii) the counterions. The monomers are regularly distributed in concentric structures around the core Y-DNA so that only minor overlap between subsequent subgenerations exists. Such a behavior is typically seen in charged macromolecules with rigid bonds:[174] it prevents backfolding of the outer parts of the dendrons, a feature that is in striking contrast to the standard dense-core model of dendrimers with flexible bonds.[175] An exemption to this observation of clearly separated subgenerations is the G8 DL-DNA in Figure 4.4(d). There, it can be seen that parts of the outermost subgenerations g_7 and g_8 curl back towards the interior of the molecule

[174]See R. Blaak et al. "Charge-Induced Conformational Changes of Dendrimers". In: *Macromolecules* 41 (2008), pp. 4452–4458. DOI: 10.1021/ma800283z.

[175]See H. M. Harreis et al. "Can Dendrimers Be Viewed as Compact Colloids? A Simulation Study of the Fluctuations in a Dendrimer of Fourth Generation". In: *J. Chem. Phys.* 118 (2003), pp. 1979–1988. DOI: 10.1063/1.1530577.

due to overcrowding on the outer shell. This rigidity of the Y-DNA arms is additionally reinforced by the Coulomb repulsion between equally charged monomers resulting in a complete suppression of backfolding, even though the junctions of the Y-DNA elements, where the three arms meet, are fully flexible.

The distribution of counterions closely follows the monomer density due to the system's propensity towards local charge neutrality. Furthermore, the spatial structure of the counterions is less pronounced due to an entropic "smearing out" of the profiles. Overall, we obtain, especially for higher generations, molecules with almost constant density in their interior. This observation is in contrast to the usual dense-core, flexible dendrimers[176] whose monomer profiles monotonically drop as one moves from the center of the molecule towards its periphery. Due to their "uniform-density", these DNA constructs lend themselves to analytical description *via* the Poisson-Boltzmann theory,[177] as the constant ion density inside the DL-DNAs simplifies analytical calculations.

The form factors $F_{mm}(k)$[178] for G1 to G5, as extracted from simulation, are shown in Figure 4.5(a) to 4.5(c). In the limit of small wave vectors, *i.e.*, for $k \rightarrow 0$, $F_{mm}(k)$ equals the total number of monomers (scatterers) of the molecule, as is demonstrated by Equation (2.52) and by Figure 4.5(b). Furthermore, oscillations in $F_{mm}(k)$ for $k\sigma \gtrsim 10^{-1}$ can be observed, where the first local minimum becomes more pronounced with increasing generation index GN. This observation signifies that the larger molecules possess a more spherical shape and that the sharpness of the boundary of the molecules at the outermost shell increases.

Another consequence of the model's rigidity is encountered in the large wave-vector behaviour of $F_{mm}(k)$; namely, the form factor satisfies the law $F_{mm}(k) \propto k^{-1}$ in the limit of large wave vectors.[179] This is the typical scaling law derived for scattering on rigid rods, which is in striking contrast to flexible dendrimers, which usually scale with k^{-4} according to Porod's law.[180]

[176] See Harreis et al., "Can Dendrimers Be Viewed as Compact Colloids? A Simulation Study of the Fluctuations in a Dendrimer of Fourth Generation" and M. Ballauff and C. N. Likos. "Dendrimers in Solution: Insight from Theory and Simulation". In: *Angew. Chem. Int. Ed.* 43 (2004), pp. 2998–3020. DOI: 10.1002/anie.200300602.

[177] See J. S. Kłos. "Dendritic Polyelectrolytes Revisited Through the Poisson-Boltzmann-Flory Theory and the Debye-Hückel Approximation". In: *Phys. Chem. Chem. Phys.* 20 (2018), pp. 2693–2703. DOI: 10.1039/C7CP07138H and T. Colla et al. "Equilibrium Properties of Charged Microgels: A Poisson-Boltzmann-Flory Approach". In: *J. Chem. Phys.* 141 (2014), p. 234902. DOI: 10.1063/1.4903746.

[178] The definition of $F_{mm}(k)$ is given in Equation (2.51).

[179] See N. J. Wagner et al. "Structure of Isotropic Solutions of Rigid Macromolecules via Small-Angle Neutron Scattering: Poly(y-benzyl L-glutamate)/Deuterated Dimethylformamide". In: *Macromolecules* 28 (1995), pp. 5075–5081. DOI: 10.1021/ma00118a041.

[180] See I. O. Götze and C. N. Likos. "Conformations of Flexible Dendrimers: A Simulation Study". In: *Macromolecules* 36 (2003), pp. 8189–8197. DOI: 10.1021/ma030137k.

FIGURE 4.5: The form factor $F_{\mathrm{mm}}(k)$ of the DL-DNAs of generations G1 to G6 (as labelled) obtained from simulations, given as functions of the scaled dimensionless wave-vector. Data are shown: (a) on a linear scale with the abscissa given in units of R_{g}; (b) on a double-logarithmic scale with the abscissa given in units of σ; and (c) on a double-logarithmic scale with $\lim_{k\to 0} F_{\mathrm{mm}}(k)$ rescaled to 1 and the abscissa given in units of R_{g}. The dashed lines in panel (c), *i.e.*, $F_{\mathrm{mm}}(k) \propto k^{-1}$, corresponds to the typical scaling law for the scattering from rigid rods Wagner et al., "Structure of Isotropic Solutions of Rigid Macromolecules via Small-Angle Neutron Scattering: Poly(y-benzyl L-glutamate)/Deuterated Dimethylformamide" for large k-values.
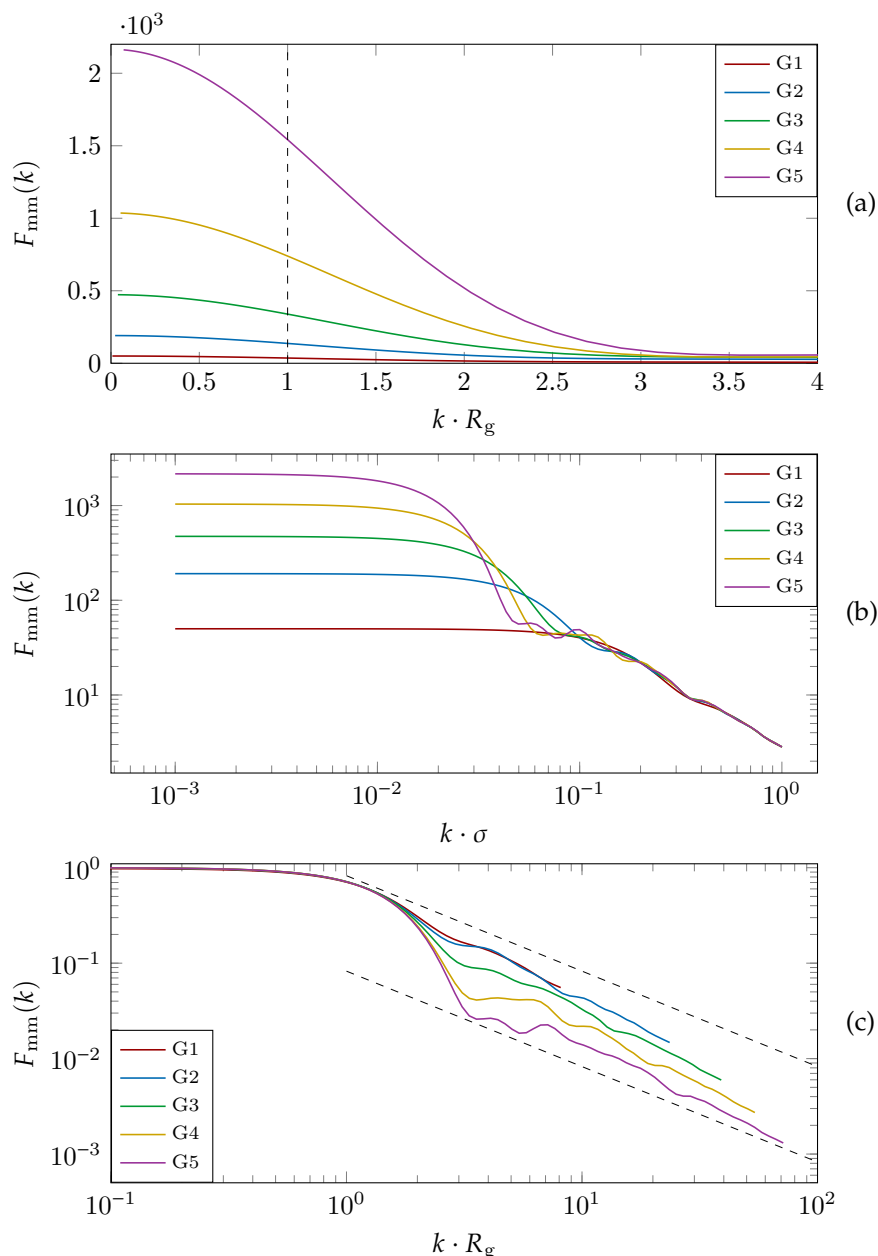
[181] See C. M. Sorensen. "Light Scattering by Fractal Aggregates: A Review". In: *Aerosol Sci. Technol.* 35 (2001), pp. 648–687. DOI: 10.1080/02786820117868, p. 652.

[182] The mass of a fractal object, M_{f}, goes with

$$M_{\mathrm{f}} \propto r_{\mathrm{f}}^{D_{\mathrm{m}}} \, ,$$

where r_{f} is the fractal object's size and D_{m} denotes the objects mass fractal dimension. For non-fractal objects of constant density $D_{\mathrm{m}} = 3$, whereas $D_{\mathrm{m}} < 3$ for fractal objects. This means that the density of fractal objects decreases as their size r_{f} increases. Analogously, the surface area of a fractal object, S_{f}, scales *via*

$$S_{\mathrm{f}} \propto r_{\mathrm{f}}^{D_{\mathrm{s}}} \, ,$$

where D_{s} is the objects surface fractal dimension. While $D_{\mathrm{s}} = 2$ corresponds to a perfectly smooth surface, *e.g.*, a sphere surface, for fractal objects $D_{\mathrm{s}} \neq 2$.

This outcome can be better understood in the context of scattering from fractal aggregates: it has been shown that for an arbitrary system of scatterers the scattering intensity scales with wave vector k *via* [181]

$$F_{\mathrm{mm}}(k) \propto (kR)^{-2D_{\mathrm{m}}+D_{\mathrm{s}}} \quad \text{for} \quad \frac{2\pi}{R} < k < \frac{2\pi}{a} \, , \tag{4.1}$$

where D_{m} and D_{s} are the mass and surface fractal dimensions, respectively.[182] While a denotes the size of a single scatterer, *i.e.*, a monomer in our case, R denotes the size of the system of scatterers, *i.e.*, the size of a

dendrimer R_g. In the case of solid spheres in $d = 3$ Euclidean dimensions, $D_m = d = 3$ and $D_s = d - 1 = 2$, which results in the well-known Porod's law: $F_{mm}(k) \propto k^{-4}$.

However, rigid dendrimers, such as DL-DNAs, can be characterised as fractal aggregates whose mass and surface fractal dimensions are equal: $D = D_m = D_s < d$.[183] Thus, the scaling law in Equation (4.1) becomes

[183] See Sorensen, "Light Scattering by Fractal Aggregates: A Review", p. 652.

$$F_{mm}(k) \propto (kR_g)^{-D} \quad \text{for } kR_g > 1, \tag{4.2}$$

with $D = 1$ (as can be seen in Figure 4.5(c)). Analysing the form factor this way yields a beautiful connection between the structure of examined dendrimers and fractal objects.

4.3 Angle and Distance Analysis

Due to the wealth of data collected in simulations we are able to examine structural details of DL-DNAs, such as internal angles and distances, which are inaccessible in experiments. For the following analysis of fluctuations in the internal structure of the dendrimers two types of angles are introduced: ϕ and θ_i. Here, ϕ is defined as the angle between the bonds of three consecutive monomers within a Y-DNA arm, see Figure 4.6. This definition of ϕ is equivalent to the one given in Equation (3.5). Consequently, the distribution of ϕ is a faithful measure of the rigidity of Y-DNA arms. Angle θ_i, on the other hand, denotes the angle enclosed between the vectors of two Y-DNA arms, whereby these arm vectors are defined as the vectors connecting the first and last monomers of a specific Y-DNA arm, *i.e.*, the arm is assumed to be fully rigid. Note that for each Y-DNA there exist three angles θ_i ($i = 1, 2, 3$); for Y-DNAs of subgenerations gn, with $n \geq 2$, two angles (which we will denote as θ_1 and θ_2) are equivalent, because one arm of the Y-DNA connects to the interior, *i.e.*, the antedecent subgenerations, while two arms extend to the exterior, *i.e.* the subsequent subgenerations. All three angles are indistinguishable, *i.e.*, equivalent, for the innermost subgeneration, g1. See Figure 4.7 for an illustration of angles θ_1, θ_2, and
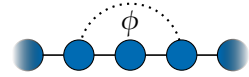

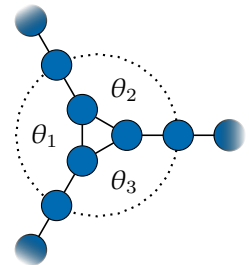**FIGURE 4.6:** Sketch to illustrate the definition of the intermonomer angle ϕ.

FIGURE 4.7: Sketch to illustrate the definition of the interarm angles θ_i, $i = 1, 2, 3$.

θ_3.[184]

FIGURE 4.8: Probability distributions $P(\phi)$ as functions of the intermonomer angle ϕ for DL-DNAs of generation index GN, with $N \in \{1, 3, 5, 7\}$ (as labelled). The angle ϕ is given in units of π. The distributions of (a) the innermost subgeneration g$_1$ and (b) the outermost subgeneration g$_N$ are shown. Each $P(\phi)$ is normalised via $\int_0^\pi P(\phi)\,\mathrm{d}\phi = 1$.
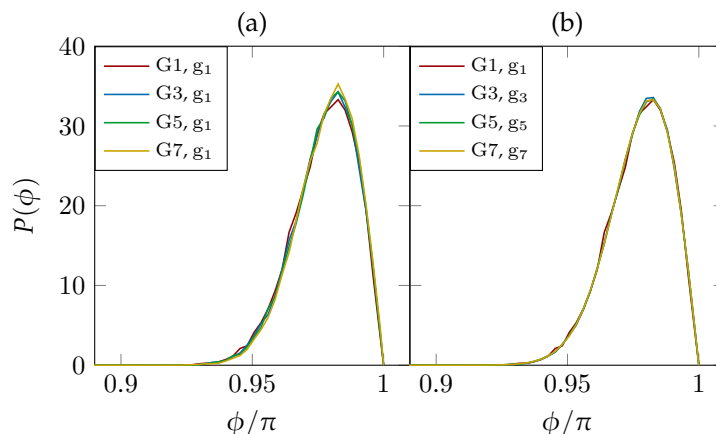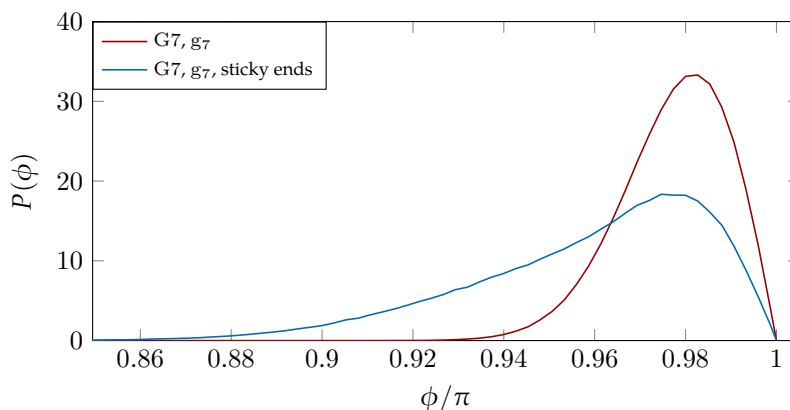

FIGURE 4.9: Probability distributions $P(\phi)$ as function of the intermonomer angle ϕ for a G7 DL-DNA. The angle ϕ is given in units of π. The distributions of the outermost subgeneration g$_7$ and of the sticky ends pertaining to that subgeneration, g$_7$, are shown. Each $P(\phi)$ is normalised via $\int_0^\pi P(\phi)\,\mathrm{d}\phi = 1$.

Measuring these two types of angles, ϕ and θ_i, on the level of the individual subgenerations g$_n$ allows us to gain a deeper understanding of the typical internal conformation of the dendrimers. The examination of a fully rigid Y-DNA would yield $\phi = \pi$ and $\theta_i = 2\pi/3$ for $i = 1, 2, 3$, a configuration illustrated in Figures 4.6 and 4.7.

In Figure 4.8 the probability distribution $P(\phi)$ of the intermonomer angle ϕ is shown for the innermost and outermost subgenerations, i.e., g$_1$ and g$_N$, respectively, of DL-DNAs with different generation numbers GN. All distributions exhibit a fairly similar shape with a pronounced maximum close to the fully rigid Y-DNA arm, i.e., $\phi \approx 0.98\pi$. This feature demonstrates that the interactions of our model, especially the harmonic angle bond of Equation (3.5), tend to keep the monomer chains straight.

The increased flexibility of the sticky ends is clearly recognisable in two distributions $P(\phi)$ for g_7 and for the single-stranded ends of that subgeneration depicted in Figure 4.9. While both functions have their maximum located at the same value of $\phi \approx 0.98\pi$, only distribution $P(\phi)$ for the sticky ends exhibits a pronounced tail towards lower values of ϕ.
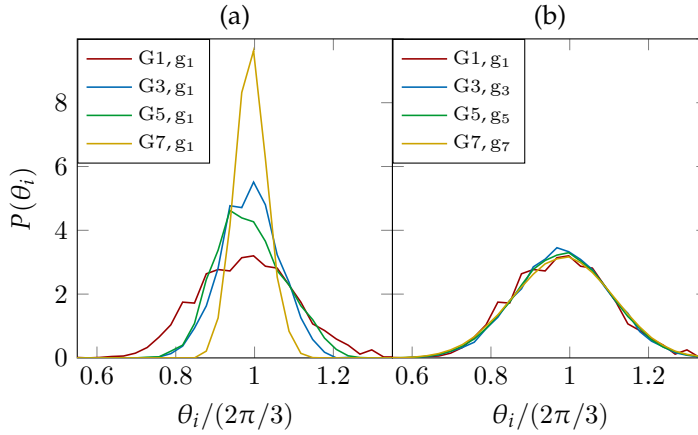

FIGURE 4.10: Probability distributions $P(\theta_i)$ as functions of the inter-arm angles θ_i for DL-DNAs of generation index GN, with $N \in \{1, 3, 5, 7\}$ (as labelled). The angles θ_i are given in units of $2\pi/3$. In this plot $P(\theta_i)$ corresponds to the combined distribution of all three angles θ_i, $i = 1, 2, 3$. The distributions of (a) the innermost subgeneration g_1 and (b) the outermost subgeneration g_N are shown. Each $P(\theta_i)$ is normalised *via* $\int_0^\pi P(\theta_i)\, d\theta_i = 1$.

In Figure 4.10 the corresponding probability distributions $P(\theta_i)$ as functions of the angles θ_i, ($i = 1, 2, 3$) for generation index GN, with $N \in \{1, 3, 5, 7\}$, are displayed. Analogously to Figure 4.8 the results for the innermost subgeneration g_1 and for the outermost subgeneration g_N are shown. Here, we do not distinguish between the three angles θ_i, $i = 1, 2, 3$, since their distributions coincide due to symmetry for the innermost subgeneration, g_1, as shown in Figure 4.11(a). This symmetry is conserved also for the outermost subgeneration, g_N, as can be seen in Figure 4.11(d). Independent of generation index GN the most probable angle θ_i of the innermost subgeneration, g_1, is centered around $\theta_i = 2\pi/3$, confirming the rigidity of the Y-DNA branches, as Figure 4.10(a) reveals. However, the width of these distributions increases with decreasing generation number GN, as the amplitude of the fluctuations in θ_i correlates negatively with the size of the dendrimer branch attached to the corresponding arm. With growing generation number, GN, the number of attached branches grows more rapidly, so that the fluctuations in the angle θ_i become less probable due to the reduced available volume and the restrictions due to the mutual electrostatic repulsions between the different arms. The distributions $P(\theta_i)$

for the outermost subgenerations, on the other hand, are virtually identical for all examined generation numbers GN.

FIGURE 4.11: Probability distributions $P(\theta_i)$ as functions of angles θ_i, $i = 1, 2, 3$ for individual subgenerations g_n of a G7 dendrimer. The angles θ_i are given in units of $2\pi/3$. The corresponding Y-DNA conformations associated with the distributions $P(\theta_i)$ are indicated by sketches of the same color in the upper left corners of the panels. Probability distributions $P(\theta_i)$ are shown for (a) subgeneration g_1, (b) subgenerations g_1 and g_3, (c) subgenerations g_3 and g_5, and (d) subgenerations g_5 and g_7. Each $P(\theta_i)$ is normalised *via* $\int_0^\pi P(\theta_i)\, \mathrm{d}\theta_i = 1$.
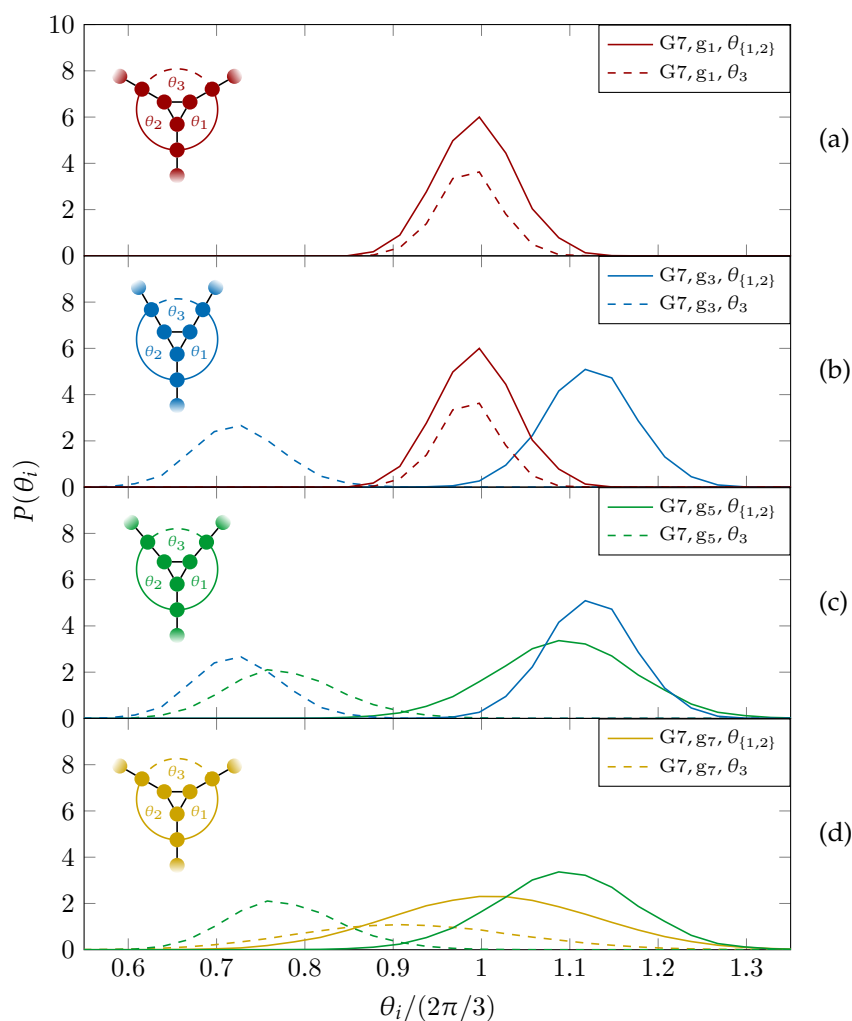
Figure 4.11 provides a more detailed examination of probability distributions $P(\theta_i)$ as functions of the angles θ_i, ($i = 1, 2, 3$), for G7 DL-DNAs. For each subgeneration g_n we collect statistics for two seperate distributions: one distribution containing angle θ_3, the angle between the two arms connecting to the subsequent subgeneration g_{n+1}, and one containing angles θ_1 and θ_2, the angles enclosed on one side by the arm connecting to the preceding subgeneration g_{n-1}. For consistency's sake this definition is also applied to g_1, even though in this case angles θ_i, $i = 1, 2, 3$, are indistinguishable as no preceding subgeneration g_0 exists.

Figure 4.11(a) reveals the most probable angle of the innermost subgener-

ation, g_1, of G7 dendrimers to be $\theta_i \approx 2\pi/3$, $i = 1, 2, 3$. Examining $P(\theta_i)$ for g_3 in Figure 4.11(b) reveals a bimodal distribution. This feature can be explained by the deformation of Y-DNAs from a conformation with $\theta_i \approx 2\pi/3$, with $i = 1, 2, 3$, at g_1 to a configuration with $\theta_i \approx 1.15\pi > 2\pi/3$, with $i = 1, 2$, and $\theta_3 \approx 0.7\pi < 2\pi/3$ at g_3, as indicated by the corresponding Y-DNA sketches in Figures 4.11(a) and 4.11(b). Exemplary reconstructions of these two cases from simulation data can be seen in Figures 4.12 and 4.13. This change is caused by the monomers pertaining to the outer and inner generations that pull and push, respectively, monomers of the intermediate subgenerations outwards *via* steric and electrostatic interactions.

A similar phenomenon, albeit not as pronounced, is found in Figure 4.11(c): in this case the peaks of the distributions $P(\theta_i)$ are located closer to $2\pi/3$, with $P(\theta_3)$ being centered around $\theta_3 = 0.8\pi$ and $P(\theta_i)$, $i = 1, 2$ being centered around $\theta_i = 1.1\pi$. The outermost subgeneration, g_7, does not exhibit such a phenomenon but instead its angles are distributed *via* a unimodal distribution $P(\theta_i)$ with a single peak located at $\theta_i \approx 2\pi/3$, $i = 1, 2, 3$, see Figure 4.11(d). In contrast to the distributions $P(\theta_i)$ for the innermost subgeneration, g_1, shown in Figure 4.11(a), the distributions for g_7 are broader, illustrating the increased orientational freedom of arms not restrained by subsequent subgenerations. For all examined subgenerations g_n, $n \in \{1, 3, 5, 7\}$, the summation of the approximate angle values yields $\sum_{i=1}^{3} \theta_i \approx 2\pi$ indicating a planar conformation of the arms. The results for $\theta_\Sigma = \sum_{i=1}^{3} \theta_i$ shown in Figure 4.14 corroborate this finding.
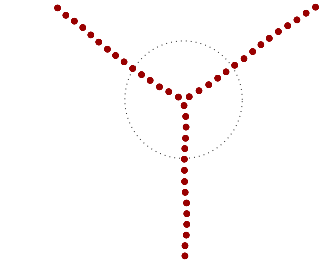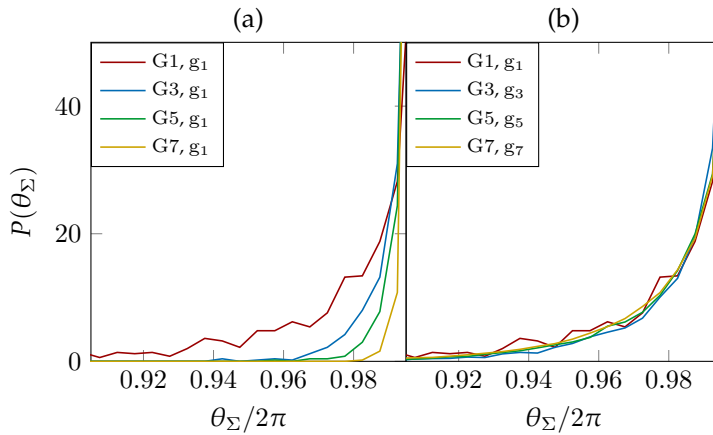

FIGURE 4.12: Reconstruction of an individual g_1-Y-DNA configuration from simulation data of a G7 DL-DNA.
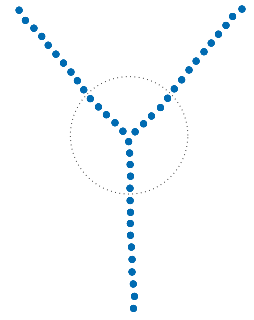

FIGURE 4.13: Reconstruction of an individual g_3-Y-DNA configuration from simulation data of a G7 DL-DNA.

FIGURE 4.14: $P(\theta_\Sigma)$, *i.e.*, the probability distribution of the sum of the three junction angles θ_Σ of the individual subgenerations within a G7 DL-DNA (as labelled). The resulting sum θ_Σ is given in units of 2π. $P(\theta_\Sigma)$ is normalised *via* $\int_0^{2\pi} P(\theta_\Sigma) \, d\theta_\Sigma = 1$.

In Figure 4.14 the distribution $P(\theta_\Sigma)$, where $\theta_\Sigma = \sum_{i=1}^{3} \theta_i$, is shown for the innermost subgeneration, g_1, and the outermost subgeneration, g_N, of DNA-dendrimers with generation index GN ($N \in \{1,3,5,7\}$). The data provide evidence that the Y-DNAs' conformations are almost completely planar, i.e., $\theta_\Sigma \approx 2\pi$, with some differences between the subgenerations. As already indicated in Figure 4.10, the Y-DNAs of the innnermost subgeneration, g_1, whose arms are subject to outward forces caused by the subsequent subgenerations, are more planar than the outermost subgenerations. With increasing generation index, however, the Y-DNAs' deviation from the planar configuration becomes more pronounced, i.e., values $\theta_\Sigma < 2\pi$ become more likely. Underlying to this behaviour are two opposing effects: Coulomb repulsion (Equation (3.1)) and the aforementioned outward forces drive the Y-DNAs towards a planar configuration, but at the same time this planarity reduces the number of configurations available to the Y-DNAs and therefore their entropy.

Finally, we further examine the entropic fluctuations by additionally investigating the distribution of nearest junction-to-junction separations, $P(r_{yy})$. We define separation r_{yy} to be the distance between the respective first monomers (the monomers closest to the Y-DNA's junction) of two Y-DNA arms, belonging to subeqeuent subgenerations g_n and g_{n+1} and which would be connected *via* enzymatic ligation in the experimental synthesis. Since two connected arms which extend from one Y-DNA junction to its neighbouring junction consists of 30 monomers, a fully extended chain would be of length $29l_b$.

FIGURE 4.15: Probability distribution $P(r_{yy})$ of the nearest junction-to-junction separation of successive subgenerations g_n and g_{n+1} within a G6 DL-DNA as function of separation r_{yy}. Distance r_{yy} is given in units of equilibrium bond length l_b. $P(r_{yy})$ is normalised *via* $\int_0^\infty P(r_{yy})\,dr_{yy} = 1$.
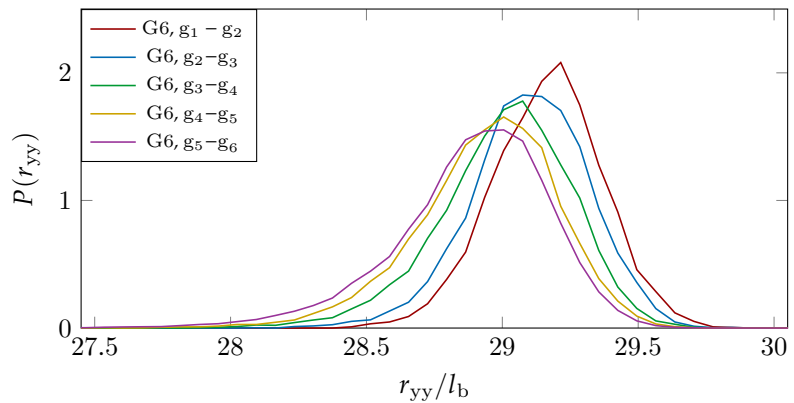
The probability $P(r_{yy})$ of the nearest junction-to-junction separation of successive subgenerations g_i and g_{i+1} within a G6 DL-DNA is shown in Figure 4.15 as a function of distance r_{yy}. Our previous analysis of DL-DNA conformation demonstrated that the individual arms of DL-DNAs are rather straight, *e.g.,* see Figure 4.8. This finding is confirmed by the distributions in Figure 4.15: the distributions' range is mainly limited to $r_{yy} \in [28.5l_b, 29.5l_b]$ and their peaks are located at $r_{yy} \approx 29l_b$, which is consistent with fully extended and almost completely straight arms. While the innermost branches have a more narrow distribution with its peak located slightly above $r_{yy} = 29l_b$, the distribution's peak position decreases monotonically as one moves towards the exterior of the molecule, (*i.e.,* towards the outer subgenerations). The shrinkage of bond lengths belonging to the outer branches is a consequence of osmotic swelling which tends, on one hand, to stretch central (inner) branches, while on the other hand, it allows a slightly higher flexibility of the branches belonging to higher subgenerations. To understand the physics behind this, we need to consider the osmotic pressure from the counterions trapped in the interior of the molecule, which tries to swell the dendrimer by exercising an outward force at a ficticious spherical surface of radius r_{max} (as in Figure 4.18(b)) that surrounds the molecule. This force is transmitted to the interior of the dendrimer but the number of Y-DNA branches along which it is partitioned is halved each time subgeneration index g_i decreases. Accordingly, the innermost generations are pulled more strongly than the outermost ones; thus they are more rigid and straight, an effect observable in the simulation snapshot (see Figure 3.9).

The distributions of the nearest junction-to-junction separation, $P(r_{yy})$, Figure 4.15 can be intepreted as the product of two competing aspects: the variation of the 29 intermonomer distances along the arm from junction to junction, whose distribution, $P(r_{mm})$, is displayed in Figure 4.16 and the corresponding intermonomer angles ϕ (see Figure 4.8).

FIGURE 4.16: Probability distribution $P(r_{\mathrm{mm}})$ of the separation of neighbouring monomers along a Y-DNA arm of subgenerations g_1, g_3, and g_5 within a G5 dendrimer, as labelled. Distance r_{mm} is given in units of equillibrium bond length l_{b}. $P(r_{\mathrm{mm}})$ is normalised *via* $\int_0^\infty P(r_{\mathrm{mm}})\,\mathrm{d}r_{\mathrm{mm}} = 1$.
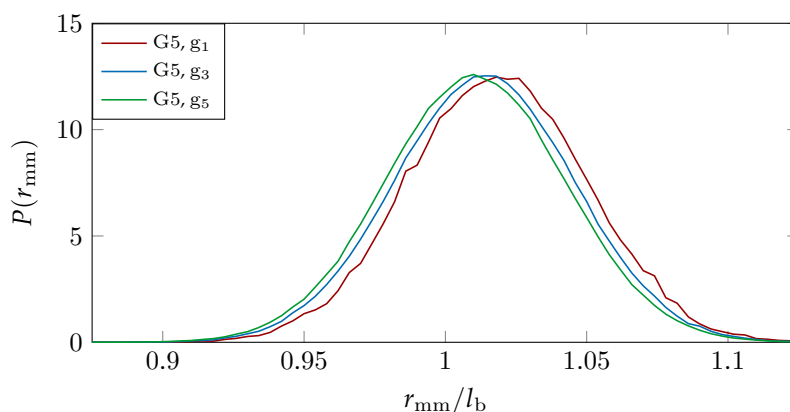
4.4 Counterion Condensation

Following the analysis of counterion distribution in Figure 4.4, we further investigate the condensation of counterions within the DL-DNA structures. In this investigation we imagine each arm of the Y-DNAs to be surrounded by a tube of radius r_{t}; by counting the number of counterions (C^+) captured within these tubes and comparing this number to the total number of counterions we obtain the ratio $Q_{\mathrm{t}}(r_{\mathrm{t}})$. Similarly, a sphere of radius r_{s} is assumed to be centered at the center-of-mass, r_{com}, of the dendrimers. The ratio $Q_{\mathrm{s}}(r_{\mathrm{s}})$ is then defined by the ratio of the number of counterions located within the sphere and the total number of counterions.

FIGURE 4.17: Representative simulation snapshot of a G3 dendrimer (left) and a G6 dendrimer (right). The DNA arms are colored blue. Counterions C^+ within a tube of radius $r_{\mathrm{t}}^* = 7.5\sigma = 30\,\text{Å}$ are colored green, whereas the rest of the counterions is colored red. The ratios $Q_{\mathrm{t}}(r_{\mathrm{t}}^*)$ for the G3 and G6 snapshots amount to 48% and 51%, respectively. These values are consistent with the values of $Q_{\mathrm{t}}(r_{\mathrm{t}}^*)$ shown in Figure 4.18. Only a fraction of the respective simulation boxes is shown in these snapshots.
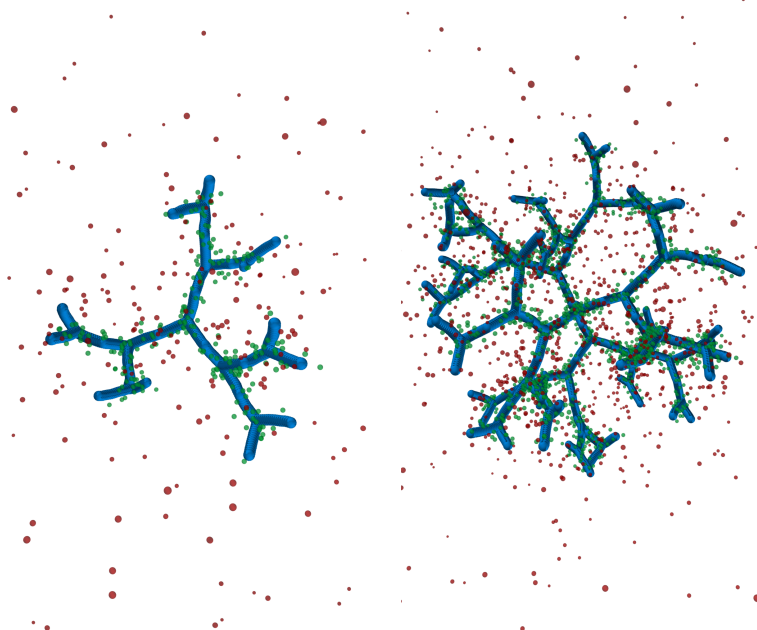
The dependence of these two quantities, $Q_s(r_s)$ and $Q_t(r_t)$, on their respective radii, r_s and r_t, are computed and depicted in Figure 4.18. Even though the considered system is electro-neutral, the values of $Q_t(r_t)$ in Figure 4.18(a) at tube radii $r_t \gtrsim 3\, r_{M^- C^+} = 5.25\, \sigma = 21\,\text{Å}$ vary substantially between dendrimers of different size. For example: for G1 and G5 dendrimers this difference amounts to more than 20%. This observation is a direct consequence of the increase in volume accessible to the counterions for larger dendrimers.
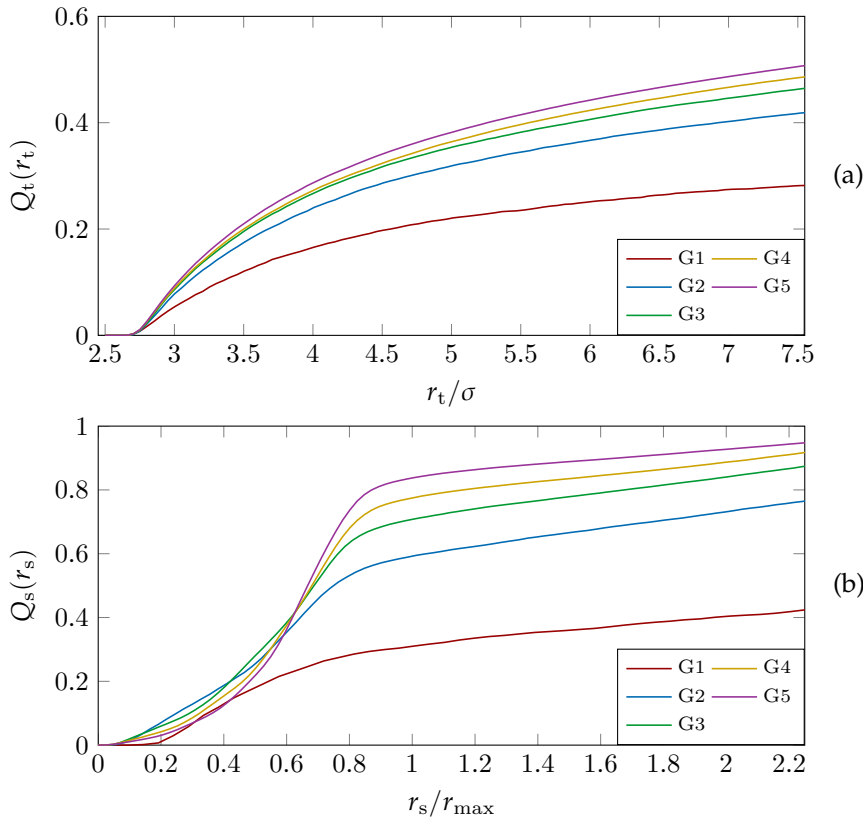
FIGURE 4.18: (a) Percentage of total ions, $Q_t(r_t)$, captured in tubes of radius r_t surrounding each arm of the Y-DNAs as a function of r (in units of the steric interaction length σ). The $Q_t(r)$ are shown for dendrimers G1 to G5, as labelled. (b) Percentage of total ions, $Q_s(r_s)$, captured in spheres of radius r_s centered at the center of mass r_{com} of the DL-DNA as a function of r_s/r_{max}. The maximum extent of the dendrimer, r_{max}, can be extracted from Figure 4.4 as the largest distance r where $\rho(r)$ of the total dendrimer is non-zero. The graphs $Q_t(r_t)$ and $Q_s(r_s)$ are shown for dendrimers G1 to G5, as labelled.

Similarly, the total amount of counterions absorbed by the dendrimers, which corresponds to $Q_s(r_s/r_{max})$, with r_{max} being the maximum extent of the dendrimer, varies depending on generation number GN. While a G1 DL-DNA absorbs around 40% of the total counterions, the amount of enclosed counterions in a G5 DL-DNA approaches 90%. Comparing $Q_s(r_s)$ for G1 and G5 in Figure 4.18(b) reveals a striking difference in the transition of the counterion profile $Q_s(r_s)$ from the interior to the exterior of the dendrimers ($r_s \approx 0.9\, r_{max}$): the transition becomes increasingly sharp

as the generation index GN grows, *i.e.*, the transition is rather smooth for G1, whereas a kink is clearly visible for G5. Consequently, DL-DNAs of higher generations act as osmotic dendrimers which capture the counterions in their interior, in full analogy with the osmotic polyelectrolyte stars.[185] However, DL-DNAs are robust against salinity, maintaining their size essentially and their structure being unaffected by addition of large quantities of monovalent salt, as will be demonstrated in Section 4.5.

[185] See A. Jusufi et al. "Counterion-induced entropic interactions in solutions of strongly stretched, osmotic polyelectrolyte stars". In: *J. Chem. Phys.* 116 (2002), pp. 11011–11027. DOI: 10.1063/1.1480007 and A. Jusufi et al. "Conformations and Interactions of Star-Branched Polyelectrolytes". In: *Phys. Rev. Lett.* 88 (2001), p. 018301. DOI: 10.1103/PhysRevLett.88.018301.

4.5 Responsiveness to Salt Concentration

TABLE 4.1: Comparison of the results for the radius of gyration, R_g, obtained in simulations and for the hydrodynamic radius, R_H, extracted from experiment (as labelled) over six generations of DL-DNA dendrimers. Experiments were performed for a salt concentration of 0.1 mM. Simulations were carried out both for the salt-free regime ($c_0 = 0$ mM) and using a salt concentrations of $c_1 = 1$ mM (as labelled). LAMMPS simulation data courtesy of Nataša Adžić. Experimental data courtesy of Emmanuel Stiakakis.

Generation	$R_{H/c=0.1\,\mathrm{mM}}^{\mathrm{exp}}$	$R_{g/c=0\,\mathrm{mM}}^{\mathrm{ESPResSo}}$	$R_{g/c=0\,\mathrm{mM}}^{\mathrm{LAMMPS}}$	$R_{g/c=1\,\mathrm{mM}}^{\mathrm{LAMMPS}}$
G1	—	3.25	3.37	—
G2	9.31	9.43	9.37	9.6
G3	14.42	15.61	15.52	15.4
G4	22.35	21.68	22.02	21.6
G5	30.82	28.61	28.75	28.2
G6	44.45	35.63	35.8	—

The responsiveness of DL-DNA systems to external stimuli was tested by analysing the effect of finite salt concentrations[186] on the overall size and structure of the dendrimer. Experiments were performed in a wide range of salt concentrations, starting at a low salt regime ($c = 0.1$ mM) and extending up to quite high concentrations of $c \approx 10$ M. Limitations in computational power restricted the salt concentrations in simulations to values $c \lesssim 10$ mM.[187] Results for the hydrodynamic radius, R_H, as obtained from the experiment, are summarised in Figure 4.19. By adding salt essentially no change in the size of the dendrimer is observable up to a concentration of $c = 10$ M.

[186] Monovalent NaCl was used in experiments.

[187] Simulating the system shown in Figure 3.9 with a salt concentration of 500 mM would result in approximately $2 \cdot 10^6$ salt particles.
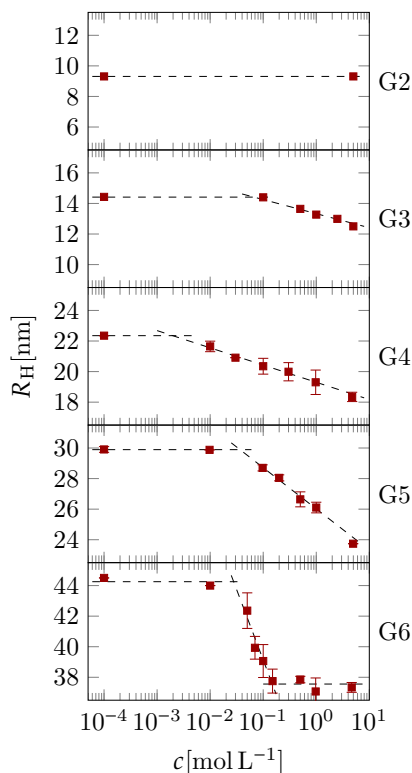
FIGURE 4.19: Plot showing the effect of different NaCl concentrations c on the hydrodynamic radius R_H. Results were obtained by DLS experiments performed by Emmanuel Stiakakis. Regimes of constant size and regimes of monotonic shrinkage are indicated by corresponding fitting lines (dashed lines).

The observation that the dendrimers' shape is almost completely unaffected by salt is confirmed in the simulations (*cf.* Table 4.1). The absence of any shrinking at low salt concentrations is the consequence of the rigidity of the molecule, *i.e.*, the high persistence length. In order to overcome the stiffness of the molecule, one has to proceed to higher salt concentrations, *i.e.*, above $c = 10\,\text{mM}$; under such conditions the screening of the charge of the molecule starts to affect the Coulomb interaction between the monomers, inducing thereby the shrinking of the molecule. This reduction in size is more pronounced for higher dendrimer generations and it can range from approximately 10% to 20% for extremely high salt concentrations (*i.e.*, $c \sim 1\,\text{M}$). Throughout, the decrease of R_H is generation dependent and the critical salt concentration at which molecule starts to shrink differs from generation to generation.

	$c =$ 0 mM	$c =$ 0.1 mM	$c =$ 1 mM	$c =$ 10 mM	$c =$ 30 mM
R_g^{sim} [nm]	21.90	—	21.60	21.00	20.40
R_H^{exp} [nm]	—	22.35	—	21.65	20.91

TABLE 4.2: Comparison between the radius of gyration, R_g, (obtained from simulation) and the hydrodynamic radius, R_H, (extracted from experiment) for different values of salt concentration c for a G4 DL-DNA. LAMMPS simulation data courtesy of Nataša Adžić.

Additionally, we provide a comparison between the experimental measurements on G4 DL-DNAs in Figure 4.19 with results obtained from MD simulations for such dendrimers. The corresponding results are summarised in Table 4.2 and are shown in Figure 4.20. Varying the salt concentration from $c = 1\,\mathrm{mM}$ to $c = 30\,\mathrm{mM}$ results in a shrinkage of the molecule by approximately 7%.

FIGURE 4.20: Hydrodynamic radius R_{H}, as extracted from experiments, and radius of gyration R_{g} obtained from simulation (as labelled), as functions of the NaCl concentration c, given in mM for a G4 DL-DNA. Experimental data courtesy of Manolis Stiakakis; LAMMPS simulation data courtesy of Nataša Adžić.
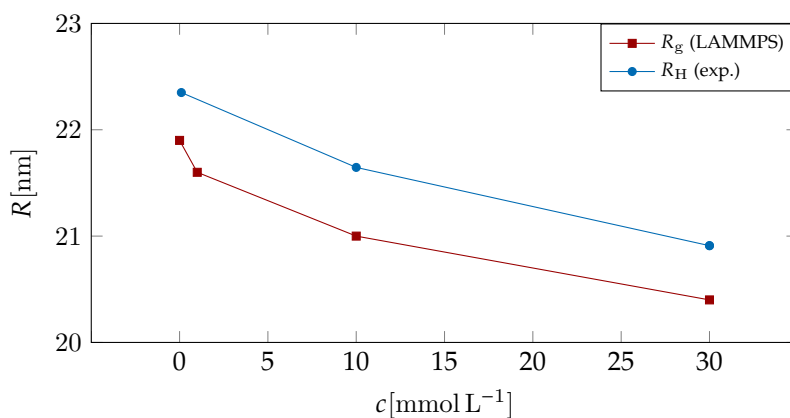
5 Effective Potentials of

DNA-Based Dendrimers

In the following chapter the concept of effective potentials, as introduced in Section 2.5, is applied to DL-DNA molecules of generation index G1, G2, and G3.[188] For the computation of effective potentials we employ the methods of umbrella sampling (Section 5.3) and WI (Section 5.2) in particular. In addition to the DL-DNA design described in Section 3.2 we implement a slightly modified version of the DNA-dendrimer which displays a higher degree of flexibility, see Section 5.1. In these simulations, salt concentrations of $c_1 = 150\,\mathrm{mM}$ and $c_2 = 500\,\mathrm{mM}$ are considered.

Substituting DL-DNA dendrimers with effective, point-like particles interacting *via* such effective potentials allows for bulk simulations of DL-DNA systems in the high density regime,[189] which are difficult or even impossible to access *via* fully atomistic or coarse-grained approaches. Furthermore, with an effective potential of these macromolecules at hand we are able to apply the HNC method[190] in order to obtain structural descriptors, *e.g.*, the pair correlation function, $g(r)$, or the structure factor, $S(q)$, of DL-DNAs in bulk solution. In Chapter 6 the results of simulations using effective potentials and numerical investigations *via* the HNC method of bulk solutions of DL-DNA dendrimers are discussed.

This chapter is structured as follows: The modified version of the DL-DNA structure is introduced in Section 5.1. Effective potentials for the original and the modified DL-DNA dendrimers (with generation numbers G1, G2, and G3) computed *via* the WI method are shown in Section 5.2, whereas some of these results are compared to effective potentials computed *via*

[188]Some of the results presented in this chapter are part of a publication in preparation, see C. Jochum et al. "DNA Stars Confined to an Interface: Planar vs. Tripod Configurations". (to be submitted).

[189] And simulating a sufficiently large simulation box.

[190] See Section 2.4 for an introduction to the HNC method.

71

umbrella sampling in Section 5.3. Finally, in Section 5.4, a so far unfruitful attempt to apply NNPs to bulk solutions of G1 DL-DNA dendrimers is discussed briefly.

The simulations in Section 5.3 were performed by Nataša Adžić using the oxDNA code, while the results of Section 5.2 were obtained by the author using the ESPResSo simulation package. The training of the neural networks in Section 5.4 was executed by Florian Buchner using the n2p2 software package, where the necessary training data was generated by the author using the ESPResSo software. The reader is referred to Appendix A for more information on the software packages. The relevant simulation parameter values are listed in Appendix B.3.

5.1 Modified Design of DL-DNA

One of the main characteristics of the DL-DNA design introduced in Section 3.2 is the stiffness of the Y-DNA arms. The resulting lack of back-folding of dendritic arms in DL-DNA molecules results in open dendrimeric structures containing voids.[191] Introducing a more flexible version of DL-DNAs enables us to investigate the influence of this arm rigidity on the conformations and thus the effective interactions between these dendrimers. In particular, we expect the modified DL-DNAs to exhibit more dense-core characteristics than the original design.

[191] See Section 4.2.

FIGURE 5.1: Connection segment of two Y-DNA arms in the coarse-grained model. While case (a) corresponds to the rigid DL-DNA design in Section 3.2, case (b) corresponds to the design with flexible connections presented in this section. Shown are base pairs (**dark blue**) of the dsDNA arms which are terminated by a ss-DNA strands (light blue).
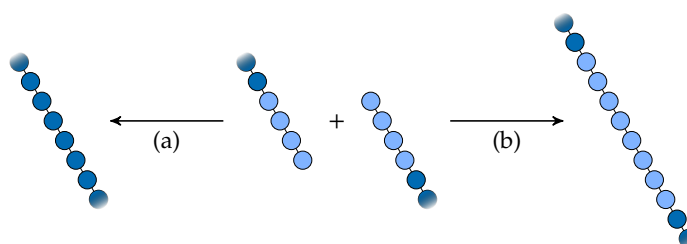
[192] See Rothemund, "Folding DNA to Create Nanoscale Shapes and Patterns"

Due to the versatility of DNA as a building material[192] such a change is relatively easy to realise: instead of terminating the Y-DNA arms with sticky ends of length four, we choose a non-sticky end-sequence of four

nucleobases.[193] Then, instead of base-pairing *via* enzymatic ligation, the ssDNA ends of two different Y-DNAs are connected[194] so that a single strand of eight nucleobases is created, see Figure 5.1. This minor tweak to the original DL-DNA design yields DL-DNAs with enhanced flexibility.[195] See Figure 5.2 for a representative snapshot of a G2 DL-DNA with flexible connections.

As a result we can now distinguish two kinds of connections between two Y-DNA junctions, Y_1 and Y_2, by looking at the composition of the DNA strand linking Y_1 and Y_2 together:

(a) 13 bp (Y_1) - 4 bp (rigid connection) - 13 bp (Y_2) and

(b) 13 bp (Y_1) - 8 single bases (flexible connection) - 13 bp (Y_2).

Henceforth the original DL-DNA design introduced in Section 3.2 (case (a) in Figure 5.1) will carry the suffix "with rigid connections", whereas the modified design presented in this section (case (b) in Figure 5.1) will be identified by the suffix "with flexible connections".

5.2 Widom Insertion Method

In this section the results of applying the Widom insertion (WI) method to DL-DNA molecules of generation numbers G1, G2, and G3 are presented. In these computations the respective center-of-mass coordinates of the dendrimer, r_{com}, were defined as the effective coordinates of the DL-DNAs.[196] For generations G1 and G3 only DL-DNAs with rigid connections are considered, whereas for generation G2 DL-DNAs with rigid and flexible connections are compared.[197] The parameter values used in the WI method and the simulations for sample generations are listed in Appendix B. All investigated systems were studied using NaCl salt concentrations of $c_1 = 150\,\mathrm{mM}$ and $c_2 = 500\,\mathrm{mM}$.

[193] For example: replace the 5'-p-**TGAC** end-sequences of ssDNA strands S_{1a}, S_{1b}, and S_{1c} in Section 3.4 by 5'-p-**TTTT** end-sequences. This way, the formation of base-pairs by unpaired terminal nucleobases of DL-DNA molecules is prevented.

[194] In the synthesis this connection is achieved *via* strain-promoted alkyne-azide cycloaddition (SPAAC), see N. J. Agard et al. "A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems". In: *J. Am. Chem. Soc.* 126 (2004), pp. 15046–15047. DOI: 10.1021/ja044996f

[195] In the coarse-grained model of DL-DNAs with flexible connections the value of the bending constant k_ϕ for sticky ends in Equation (3.5) is set to $k_\phi = 75\,\mathrm{kJ \cdot mol^{-1}}$ in order to match results from oxDNA simulations of DL-DNAs with flexible connnections.
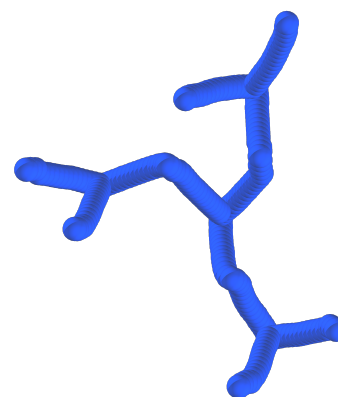

FIGURE 5.2: A simulation snapshot of a G2 DL-DNA with flexible connections in the coarse-grained model. Note the kink in the DL-DNA arms connecting the different Y-DNAs.

[196] See Appendix C.1 for supplementary results using the center-of-mass of the three central junction-monomers of the innermost sub-generation g_1, r_{cy}, as the effective coordinate.

[197] See Section 5.1 for the difference in design between DL-DNAs with rigid and flexible connections.

5.2.1 Effective Potentials (WI) of 1st Generation DL-DNAs

Here, the effective potentials, $\phi_{\mathrm{eff}}(r)$, of G1 DL-DNAs computed *via* the WI method are presented and compared to the corresponding radial density profiles (RDPs): the results for salt concentrations $c_1 = 150\,\mathrm{mM}$ and $c_2 = 500\,\mathrm{mM}$ are displayed in Figures 5.3 and 5.4, respectively.
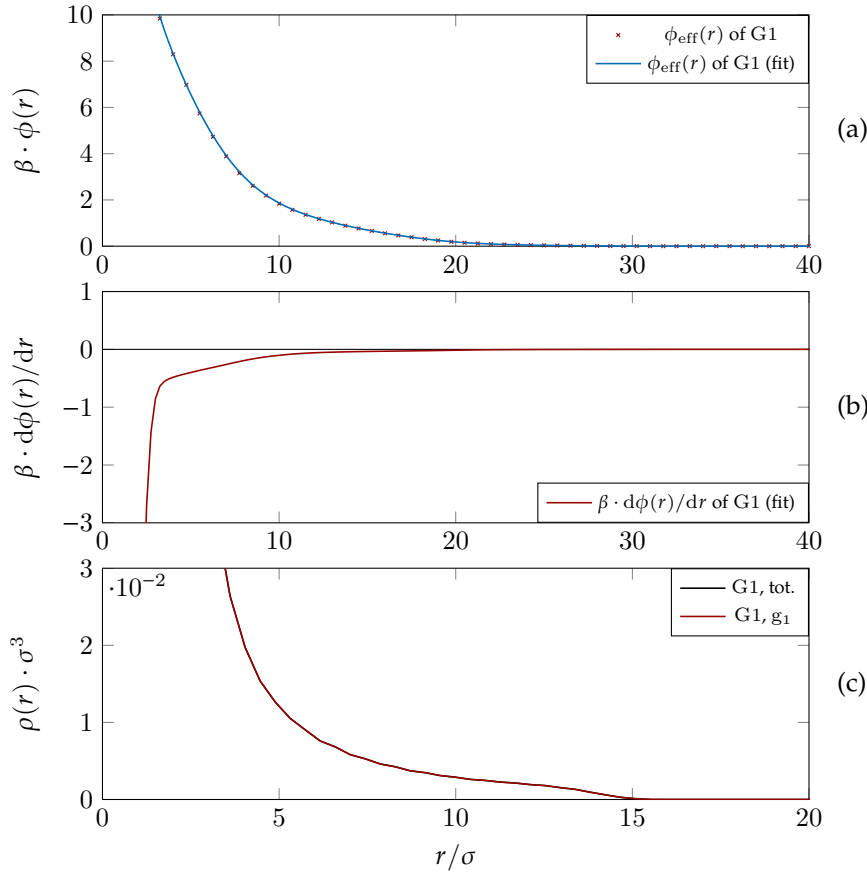
FIGURE 5.3: Effective potential $\phi_{\mathrm{eff}}(r)$, (subplot (a)) in units of $k_{\mathrm{B}}T$ (with $\beta = 1/k_{\mathrm{B}}T$), derivative of the effective potential $\mathrm{d}\phi_{\mathrm{eff}}(r)/\mathrm{d}r$ (subplot (b)) and monomer density profiles $\rho(r)$ (subplot (c)) as functions of r, given in units of the steric interaction length σ. Results are shown for G1 DL-DNA dendrimers with rigid connections and a salt concentration $c_1 = 150\,\mathrm{mM}$. $\rho(r)$ is shown for the whole dendrimer (labelled "tot.") and the individual subgenerations g_i (labelled "g_i") with $i = 1, 2$.
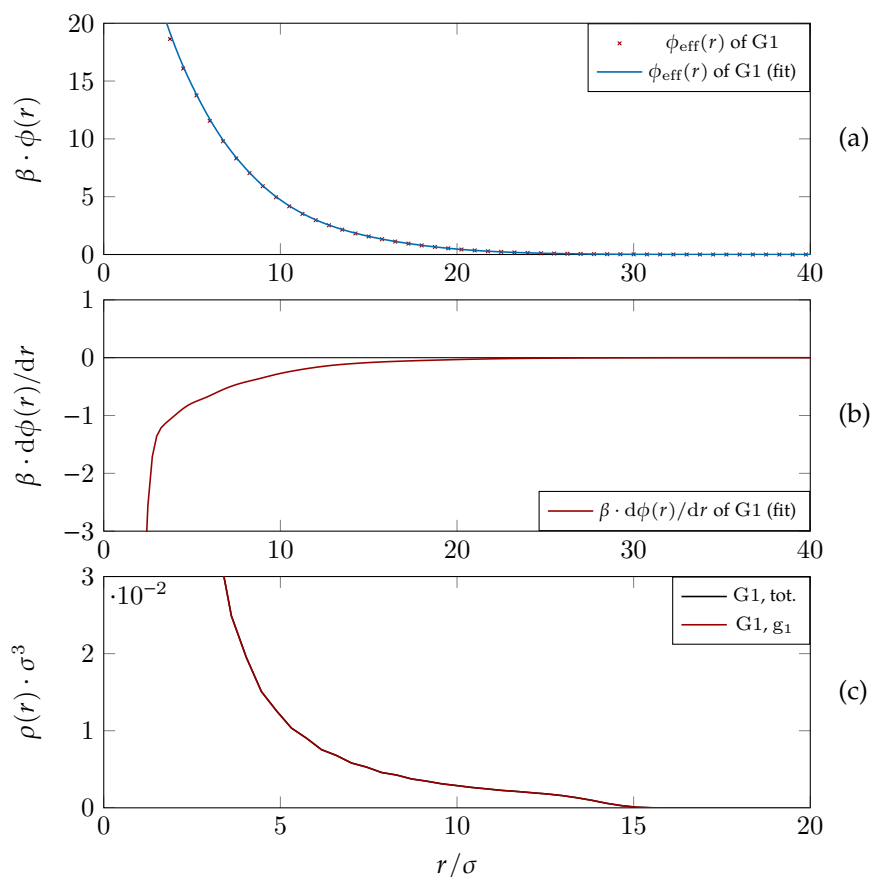
In these computations the center-of-mass coordinates, r_{com}, were defined as the effective coordinates of the G1 dendrimers. Due to the smaller size and the higher symmetry of G1 DL-DNAs the effective coordinate at r_{com} approximately coincides with the central monomers. Thus, the effective potential, $\phi_{\mathrm{eff}}(r)$, diverges at small separations $r \leq 4\sigma$ due to the steric interaction between the DNA monomers close to r_{com}. As the electrostatic interaction does not have a major influence on the conformation of the G1 DL-DNAs (see Figures 5.3(c) and 5.4(c)), we expect the effective potentials $\phi_{\mathrm{eff}}(r)$ of Figures 5.3(a) and 5.4(a) to be of similar shape. The only observ-

able difference between Figures 5.3(a-b) and 5.4(a-b) is the strength of the effective potentials: whereas the effective potential close to the divergence amounts to $\phi_{\text{eff}}(r = 4\sigma) \approx 8.3 k_{\text{B}} T$ for $c_1 = 150\,\text{mM}$, the corresponding value for $c_2 = 500\,\text{mM}$ is $\phi_{\text{eff}}(r = 4\sigma) \approx 18.3 k_{\text{B}} T$. This difference of approximately factor two between results pertaining to different salt concentrations is also observed in G2 and G3 results, which will be presented later on.

FIGURE 5.4: Effective potential $\phi_{\text{eff}}(r)$, (subplot (a)) in units of $k_{\text{B}} T$ (with $\beta = 1/k_{\text{B}} T$), derivative of the effective potential $\mathrm{d}\phi_{\text{eff}}(r)/\mathrm{d}r$ (subplot (b)) and monomer density profiles $\rho(r)$ (subplot (c)) as functions of r, given in units of the steric interaction length σ. Results are shown for G1 DL-DNA dendrimers with rigid connections and a salt concentration $c_2 = 500\,\text{mM}$. $\rho(r)$ is shown for the whole dendrimer (labelled "tot.") and the individual subgenerations g_i (labelled "g_i") with $i = 1, 2$.

The slight downward turn in the monomer density profiles, $\rho(r)$, seen in Figures 5.3(c) and 5.4(c) indicates the higher flexibility of the ssDNA ends. This feature is located at distances $r = 12.5\sigma \approx 14.7 l_{\text{b}}$ to $r = 15\sigma \approx 17.6 l_{\text{b}}$.

Due to its geometry the G1 dendrimers typically assume extremely aspherical conformations.[198] Thus the validity of a radially symmetric effective potential, such as the aforecalculated $\phi_{\text{eff}}(r)$, is limited. Increasing the generation number GN yields larger macromolecules with a higher sphericity.
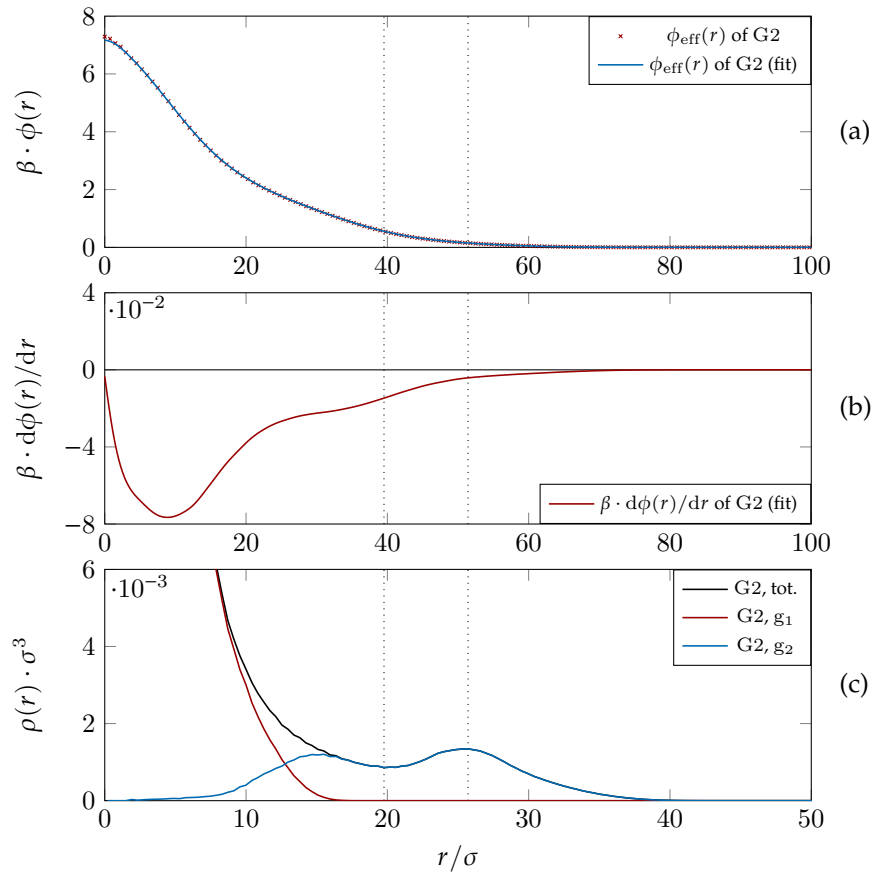
[198] As demonstrated by the eigenvalues of the gyration tensor listed in Table C.3.

The effective interactions of G1 stars in two dimensions are discussed in detail in Chapter 7.

5.2.2 Effective Potentials (WI) of 2nd Generation DL-DNAs

After investigating the effective potentials, $\phi_{\mathrm{eff}}(r)$, of G1 DNA-stars, we now focus on the corresponding effective interactions of G2 dendrimers. Potentials $\phi_{\mathrm{eff}}(r)$ were computed for G2 DL-DNAs with rigid and flexible connections and under salt concentrations $c_1 = 150\,\mathrm{mM}$ and $c_2 = 500\,\mathrm{mM}$. The center-of-mass coordinate of the dendrimer, r_{com}, is used as the effective coordinate of the DL-DNAs. The results for the four different cases are shown in Figures 5.5 to 5.8.

FIGURE 5.5: Effective potential $\phi_{\mathrm{eff}}(r)$, (subplot (a)) in units of $k_{\mathrm{B}}T$ (with $\beta = 1/k_{\mathrm{B}}T$), derivative of the effective potential $\mathrm{d}\phi_{\mathrm{eff}}(r)/\mathrm{d}r$ (subplot (b)) and monomer density profiles $\rho(r)$ (subplot (c)) as functions of r, given in units of the steric interaction length σ. Results are shown for G2 DL-DNA dendrimers with rigid connections and salt concentration $c_1 = 150\,\mathrm{mM}$. $\rho(r)$ is shown for the whole dendrimer (labelled "tot.") and the individual subgenerations g_i (labelled "g_i") with $i = 1, 2$. The dotted vertical lines indicate structural features in $\rho(r)$ and the equivalent inter-dendrimer distances in $\phi_{\mathrm{eff}}(r)$.

The main characteristic feature connecting all resulting potentials, $\phi_{\mathrm{eff}}(r)$, is their ultra-soft nature: all potentials exhibit an upper limit of $\phi_{\mathrm{eff}}(r = 0) \approx 4k_{\mathrm{B}}T$ and $\phi_{\mathrm{eff}}(r = 0) \approx 8k_{\mathrm{B}}T$ for salt concentrations $c_1 = 150\,\mathrm{mM}$ and

$c_2 = 500\,\mathrm{mM}$, respectively.

In analogy to previous results regarding G1 DNA-stars, it can be observed that varying the salt concentration c barely influences the dendrimers structure in the rigid case (see Figures 5.5(c) and 5.6(c)) and in the flexible case (Figures 5.7(c) and 5.8(c)).[199] However, the strength of the effective potentials at $r = 0$ decreases by a factor of two when increasing salt concentration c from $150\,\mathrm{mM}$ to $500\,\mathrm{mM}$, see Figures 5.5(a) and 5.6(a) (rigid connections) and Figures 5.7(a) and 5.8(a) (flexible connections).

[199] This observation corroborates the related results of Section 4.5.
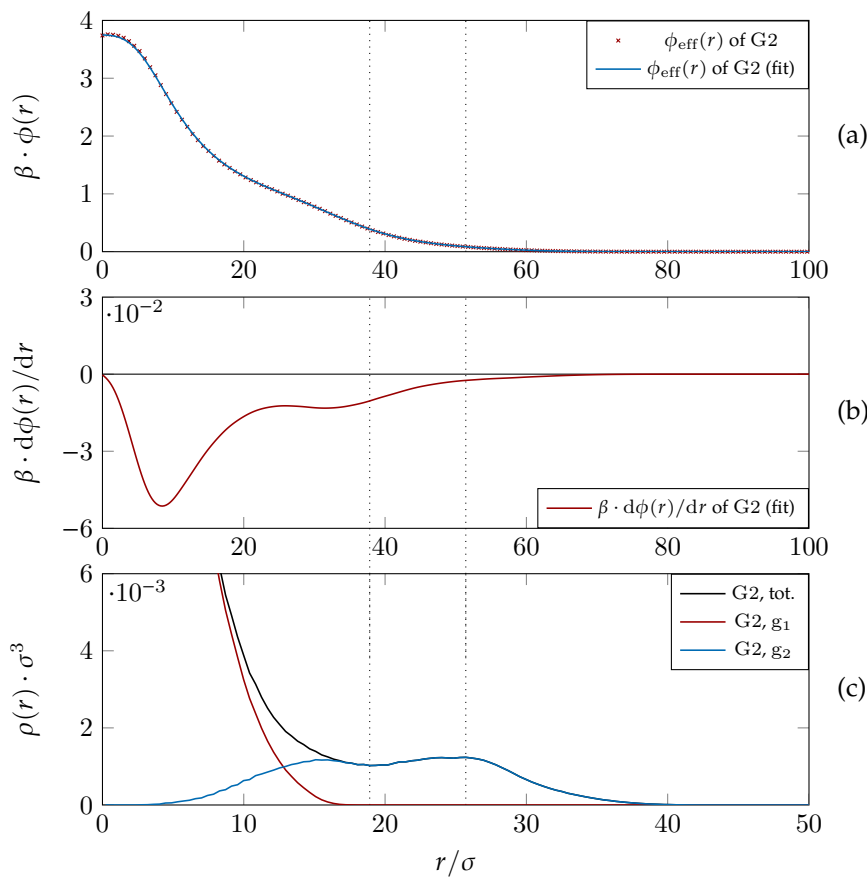
FIGURE 5.6: Effective potential $\phi_{\mathrm{eff}}(r)$, (subplot (a)) in units of $k_{\mathrm{B}}T$ (with $\beta = 1/k_{\mathrm{B}}T$), derivative of the effective potential $\mathrm{d}\phi_{\mathrm{eff}}(r)/\mathrm{d}r$ (subplot (b)) and monomer density profiles $\rho(r)$ (subplot (c)) as functions of r, given in units of the steric interaction length σ. Results are shown for G2 DL-DNA dendrimers with rigid connections and salt concentration $c_2 = 500\,\mathrm{mM}$. $\rho(r)$ is shown for the whole dendrimer (labelled "tot.") and the individual subgenerations g_i (labelled "g_i") with $i = 1, 2$. The dotted vertical lines indicate structural features in $\rho(r)$ and the equivalent inter-dendrimer distances in $\phi_{\mathrm{eff}}(r)$.

Furthermore, a noticeable qualitative difference exists between the two cases of rigid and flexible connections, regardless of salt concentration c: Whereas in the rigid case the effective potentials $\phi_{\mathrm{eff}}(r)$ display two distinctively different length scales in their decay with increasing distance r, e.g., Figure 5.5(a) and 5.5(b), no such phenomenon can be observed in the flexible case, e.g., Figure 5.7(a) and 5.7(b).
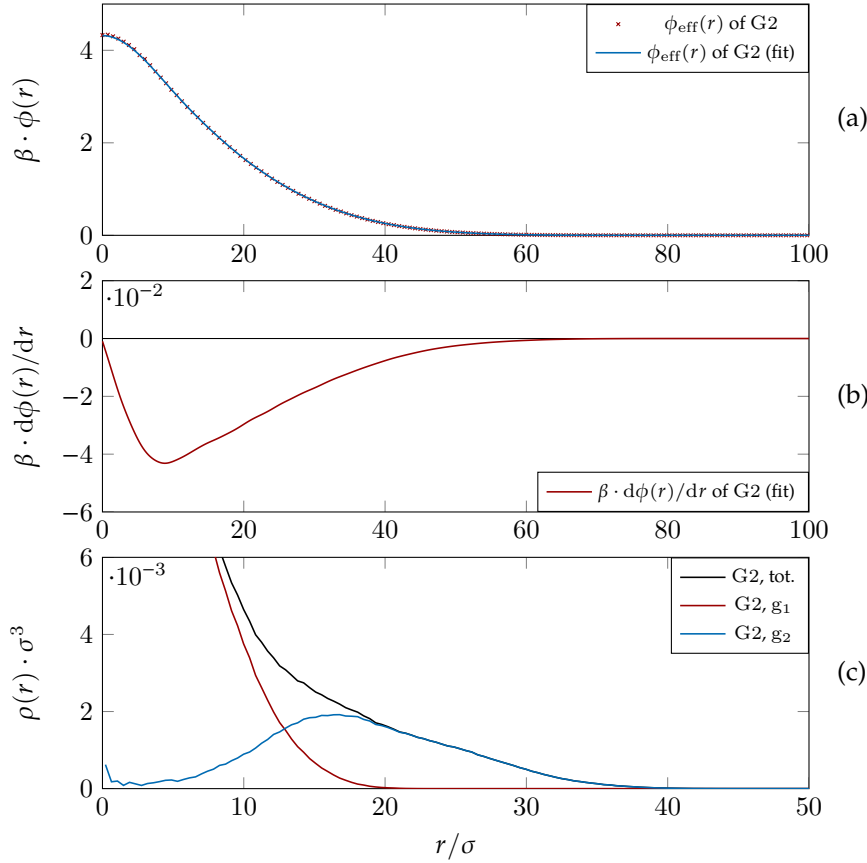
FIGURE 5.7: Effective potential $\phi_{\mathrm{eff}}(r)$, (subplot (a)) in units of $k_{\mathrm{B}}T$ (with $\beta = 1/k_{\mathrm{B}}T$), derivative of the effective potential $\mathrm{d}\phi_{\mathrm{eff}}(r)/\mathrm{d}r$ (subplot (b)) and monomer density profiles $\rho(r)$ (subplot (c)) as functions of r, given in units of the steric interaction length σ. Results are shown for G2 DL-DNA dendrimers with flexible connections and salt concentration $c_1 = 150\,\mathrm{mM}$. $\rho(r)$ is shown for the whole dendrimer (labelled "tot.") and the individual subgenerations g_i (labelled "g_i") with $i = 1, 2$. The dotted vertical lines indicate structural features in $\rho(r)$ and the equivalent inter-dendrimer distances in $\phi_{\mathrm{eff}}(r)$.

The emergence of two length scales in the decay of $\phi_{\mathrm{eff}}(r)$ must be a consequence of the dendrimers' inner architecture: the functions $\rho(r)$ in Figures 5.5(c) and 5.6(c) both exhibit a peak at $r \approx 25\sigma \approx 29 l_{\mathrm{b}}$, a feature which is not observed in the corresponding radial density profiles $\rho(r)$ in Figures 5.7(c) and 5.8(c). In a straight-arm configuration the length of the DNA connection between two junctions of the subsequent subgenerations g_1 and g_2 amounts to $29 l_{\mathrm{b}}$. Thus, the aforementioned peak corresponds to the distance where most Y-junctions of subgenerations g_2 are located with respect to center of mass, r_{com}. The increased flexibility of dendrimers with flexible connections allows for some limited amount of backfolding of the outer subgeneration g_2 which can be inferred from the delocalisation and disappearance of the aforementioned peak of $\rho(r)$ in Figures 5.7(c) and 5.8(c).

FIGURE 5.8: Effective potential $\phi_{\mathrm{eff}}(r)$, (subplot (a)) in units of $k_{\mathrm{B}}T$ (with $\beta = 1/k_{\mathrm{B}}T$), derivative of the effective potential $\mathrm{d}\phi_{\mathrm{eff}}(r)/\mathrm{d}r$ (subplot (b)) and monomer density profiles $\rho(r)$ (subplot (c)) as functions of r, given in units of the steric interaction length σ. Results are shown for G2 DL-DNA dendrimers with flexible connections and salt concentration $c_2 = 500\,\mathrm{mM}$. $\rho(r)$ is shown for the whole dendrimer (labelled "tot.") and the individual subgenerations g_i (labelled "g_i") with $i = 1, 2$. The dotted vertical lines indicate structural features in $\rho(r)$ and the equivalent inter-dendrimer distances in $\phi_{\mathrm{eff}}(r)$.

This structural feature has implications for the effective potentials, $\phi_{\mathrm{eff}}(r)$: as two G2 dendrimers with rigid connections approach each other they start to "feel" each other at an inter-dendrimer distance of $r \approx 60\sigma$. Decreasing the inter-dendrimer distance r even further results in a significant increase of the interaction energy, $\phi_{\mathrm{eff}}(r)$, as the outermost peaks[200] of their respective density profiles $\rho(r)$ start to overlap, see Figures 5.5(c) and 5.6(c). Analogously, the local minimum of $\rho(r)$ at $r \approx 20\sigma$ reduces the increase of the effective potential $\phi_{\mathrm{eff}}(r)$, which can be observed in Figures 5.5(b) and 5.6(b).

The lack of these structural traits in the density profiles $\rho(r)$ of DL-DNAs with flexible connections (see Figures 5.7(c) and 5.8(c)) is mirrored in the corresponding effective potentials $\phi_{\mathrm{eff}}(r)$: the derivatives of the effective potentials of DL-DNAs with flexible connections, $\mathrm{d}\phi_{\mathrm{eff}}(r)/\mathrm{d}r$, shown in Figures 5.7(b) and 5.8(b) display a clear lack of structural features when

[200] These peaks are located at $20\sigma \lessgtr r \lessgtr 30\sigma$.

compared to the corresponding derivatives of DL-DNAs with rigid connections shown in Figures 5.5(b) and 5.6(b).

Such an analysis of $\phi_{\text{eff}}(r)$ and $\rho(r)$ gives us profound insight into the effects of variations in dendrimer flexibility and in salt concentration c on inter-dendrimer interactions: we find that all potentials $\phi_{\text{eff}}(r)$ share their ultra-soft and long-ranged nature; on a quantitative level they differ as a consequence of their sensitivity to dendrimer flexibility and salt concentration c. While not explicitly of Gaussian form, these potentials show similarity to the Gaussian core model (GCM).[201]

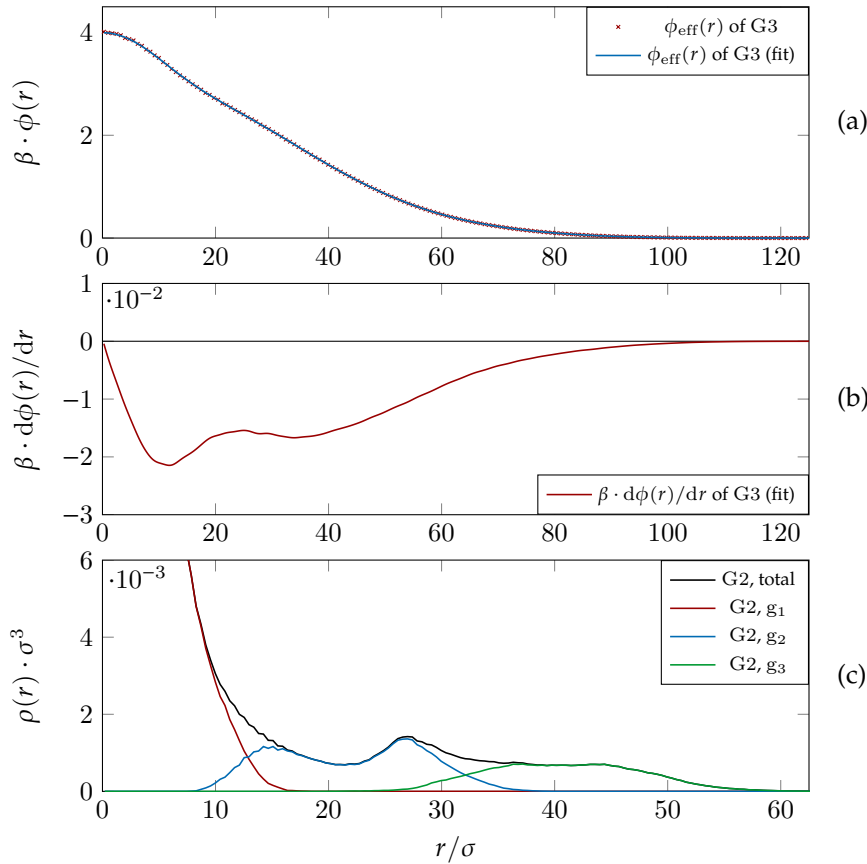### 5.2.3 Effective Potentials (WI) of 3rd Generation DL-DNAs

FIGURE 5.9: Effective potential $\phi_{\text{eff}}(r)$, (subplot (a)) in units of $k_{\text{B}}T$ (with $\beta = 1/k_{\text{B}}T$), derivative of the effective potential $\mathrm{d}\phi_{\text{eff}}(r)/\mathrm{d}r$ (subplot (b)) and monomer density profiles $\rho(r)$ (subplot (c)) as functions of r, given in units of the steric interaction length σ. Results are shown for G3 DL-DNA dendrimers with rigid connections and salt concentration $c_1 = 150\,\text{mM}$. $\rho(r)$ is shown for the whole dendrimer (labelled "tot.") and the individual subgenerations g_i (labelled "g_i") with $i = 1, 2$.
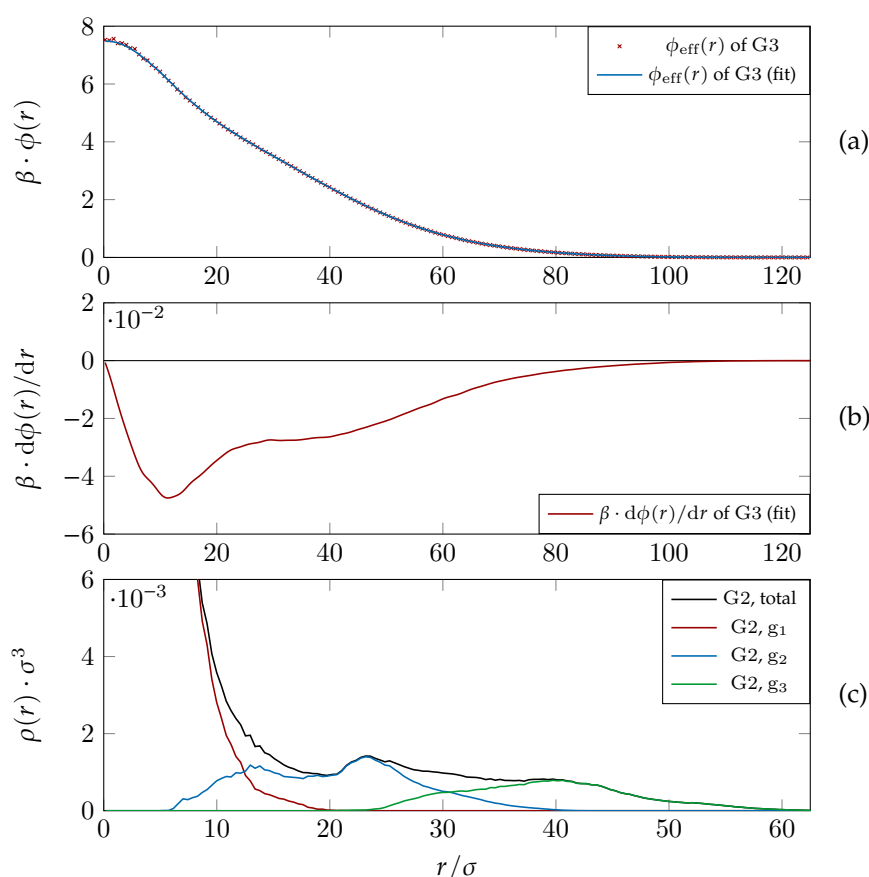
Finally, we apply the WI method to G3 DL-DNAs with rigid connections using salt concentrations of $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$. The results

show similarities to the results of G2 as revealed in Figures 5.9 and 5.10. As previously discovered, the variation of the salt concentration c does not affect the conformation of the dendrimer (see Figures 5.9(c) and 5.10(c)), but has an impact on the strength of the effective potentials ϕ_{eff}. The radial density profile $\rho(r)$ of subgeneration g_2 exhibit peaks as before in the G2 case, while the additional subgeneration g_3 has a more broad, featureless appearence. Due to the additional subgeneration a longer interaction range of ϕ_{eff} can be observed in Figures 5.9(a) and 5.10(a) in comparison to the G2 case. Again, the effective potentials ϕ_{eff} display two length scales in their decay. While not as pronounced as in previous results, this feature is recognisable in Figures 5.9(b) and 5.10(b).

FIGURE 5.10: Effective potential $\phi_{\text{eff}}(r)$, (subplot (a)) in units of $k_{\text{B}}T$ (with $\beta = 1/k_{\text{B}}T$), derivative of the effective potential $\mathrm{d}\phi_{\text{eff}}(r)/\mathrm{d}$ (subplot (b)) and monomer density profiles $\rho(r)$ (subplot (c)) as functions of r, given in units of the steric interaction length σ. Results are shown for G3 DL-DNA dendrimers with rigid connections and salt concentration $c_2 = 500\,\text{mM}$. $\rho(r)$ is shown for the whole dendrimer (labelled "tot.") and the individual subgenerations g_i (labelled "g_i") with $i = 1, 2$.

5.3 Umbrella Sampling

Here, we show additional results computed *via* the umbrella sampling

[202] See Section 2.5.1 for a description of the umbrella sampling method.

method[202] in order to verify the validity of the effective potentials, $\phi_{\text{eff}}(r)$, calculated using the WI method in Section 5.2.

FIGURE 5.11: Effective potential $\phi_{\text{eff}}(r)$, in units of $k_{\text{B}}T$ (with $\beta = 1/k_{\text{B}}T$) of G2 dendrimers computed *via* the WI method and the umbrella sampling method (as labelled). Results are shown for G2 DL-DNAs with rigid connections at a salt concentration $c = 150\,\text{mM}$. Charges q in the system are reduced by a factor of 3. Umbrella sampling data courtesy of Nataša Adžić.
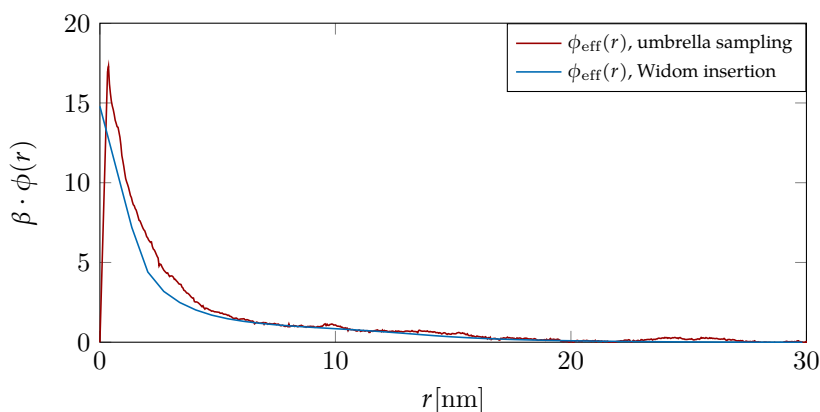
Figure 5.11 shows the effective potentials, $\phi_{\text{eff}}(r)$, for G2 dendrimers with rigid connections at a salt concentration $c = 150\,\text{mM}$ computed *via* the WI

[203] These results were obtained by Nataša Adžić using the oxDNA software.

method and the umbrella sampling method (as labelled).[203] Both methods were implemented *via* the oxDNA code with a charge reduction to 33%.[204]

[204] See Section 6.2 for a discussion of the charge reduction.

The results for $\phi_{\text{eff}}(r)$ obtained by the two different methods show excellent agreement, strengthening the validity of results in Section 5.2. Due to the limited number of samples used for these calculations some noise can be observed in the results computed *via* umbrella sampling, which explains the minor discrepancies between the curves for $\phi_{\text{eff}}(r)$ in Figure 5.11.

5.4 Neural Network Potentials

Until now our approach for determining effective interactions between DL-DNAs consisted of directly computing the two-body effective potential, $\phi_{\text{eff}}(r)$, of the investigated dendrimers *via* the WI method and the umbrella sampling method, discussed in Sections 5.2 and 5.3, respectively.

In this section, we present the results of an attempt to obtain N-body effec-

tive potentials of coarse-grained DL-DNAs computed *via* artificial neural networks (NNs). The technique of using NNs to compute N-body potentials of particles, which is commonly summarised under the name neural network potentials (NNPs), is based on the ability of NNs to reproduce highly complex functions such as free energy landscapes by learning from training data.[206] In this method the goal is to compute the effective potentials using NNs, where we treat each dendrimer as a point-like particle with its center-of-mass, r_{com}, as its effective coordinate. By analysing the training data, which is provided in form of coordinate-trajectories of the centers-of-mass with corresponding forces and energies, the NN tries to correlate the force acting on an individual DL-DNA with the configuration of its neighbouring dendrimers. For efficiency's sake, this environment configuration is additionally encoded in the so-called atomic environment descriptors.[207]

An NNP-approach, such as the one described here, is applicable if the forces acting on the particles are the result of a restricted environment, *i.e.*, no long-range forces are present, and there is a one-to-one mapping of the DL-DNAs' atomic environment descriptors to the resulting forces. By adding a sufficiently high salt concentration c to the system, the limited range of the electrostatic interaction is guaranteed so that the first condition is fulfilled. The second condition, which states that two identical atomic environments of a dendrimer must result in the same force acting on said dendrimer, however is not satisfied in our approach as discussed below.

The results show that our training attempts were ultimately unsuccessful.[208] Additionally, we failed to find significant correlations between the atomic environment descriptors of the coarse-grained, *i.e.*, point-like particles located at the centers-of-mass, dendrimers and their forces in the provided datasets. A possible cause for this phenomenon is that the set of microscopic configurations corresponding to a given coarse-grained configuration actually result in contradicting force values.

For example, two effective particles, representing G1 dendrimers, seperated by less than two Y-DNA arm lengths in the coarse-grained image can corre-

[206] Training data was provided by the author, whereas the actual training of NNs was performed by Florian Buchner using the Neural Network Potential Package (n2p2) written by Andreas Singraber. See F. Buchner. "Coarse-graining of dendrimer-like DNA using high-dimensional neural network potentials". TU Wien, 2019. (unpublished project thesis) for more details.

[207] In the default case, the n2p2 package uses the atom-centered symmetry functions.

[208]See Buchner, "Coarse-graining of dendrimer-like DNA using high-dimensional neural network potentials"

spond to two planar stars in a parallel conformation or they can correspond to two G1 dendrimers with their arms touching.[209] While the first case will only result in a rather small force, depending on salt concentration c, the second case will yield a strong force.

In order to remedy this conundrum one could introduce additional atomic descriptors, *e.g.*, orientation parameters, in addition to the effective coordinate so that the training data provides sufficient information of a one-to-one mapping from coordinate input to forces. Another option would be to manipulate the training data by averaging the force data over all microscopic configurations corresponding to the same coarse-grained configuration.[210] This way the training data set only contains forces uniquely defined by the coarse-grained configuration. This is likely more computationally expensive and will require considerable additional work before the actual training of the neural network.

[209] Figure C.4 in Appendix C.1 shows snapshots obtained when using the WI method. There, the equivalent situation for G2 dendrimers can be seen.

[210] *Cf.* Equation (7) in L. Zhang et al. "DeePCG: Constructing coarse-grained models via deep neural networks". In: *J. Chem. Phys.* 149 (2018), p. 034101. DOI: 10.1063/1.5027645, p. 3

6 Bulk Solutions of DNA-Based Dendrimers

After previously introducing our coarse-grained model of DL-DNAs in Chapter 3 and examining DL-DNAs in dilute solutions in Chapters 4, we now focus on investigating these dendrimers in the bulk.[211] In this chapter the results of investigations of bulk systems of G1, G2, and G3 DL-DNAs with rigid connections at salt concentrations $c_1 = 150\,\mathrm{mM}$ and $c_2 = 500\,\mathrm{mM}$ and at different densities η are presented. On the theoretical side there are three approaches underpinning these investigations:

[211]Some of the results presented in this chapter are part of a publication in preparation, see N. Adžić et al. (in preparation).

(a) MD simulations of DL-DNAs using the coarse-grained model introduced in Section 3.2,

(b) MD simulations where DL-DNAs are substituted with effective, point-like particles interacting *via* the effective potentials calculated in Section 5.2,

(c) and application of the HNC method (see Section 2.4) to the effective potentials of Section 5.2.

Whereas theoretical results were obtained for all three dendrimer generation numbers G1, G2, and G3, experimental results are only available for G2 dendrimers.

The results for bulk systems of G1, G2, and G3 dendrimers are shown in Sections 6.1, 6.2, and 6.3, respectively.

All simulations in this chapter were performed by the author using the ESPResSo simulation package. The reader is referred to Appendix A for more information on the ESPResSo software. The relevant simulation

details and simulation parameter values are given in Appendix B.4.
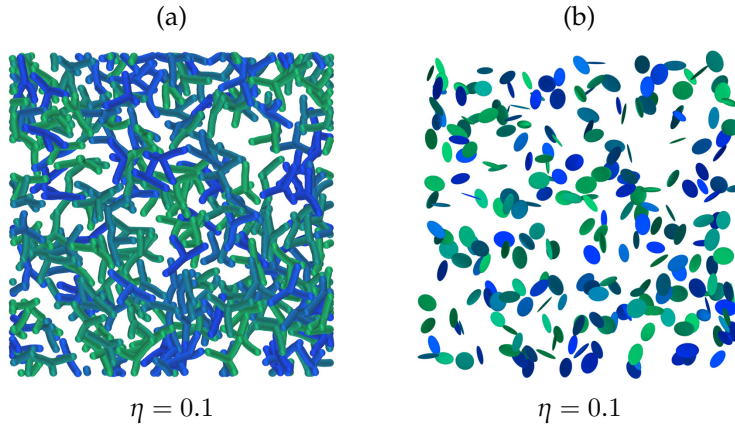
6.1 Bulk Solutions of 1st Generation DL-DNAs

FIGURE 6.1: Simulation snapshots of systems containing G1 DL-DNA dendrimers at density $\eta = 0.001$ and at salt concentration $c_2 = 150\,\text{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).
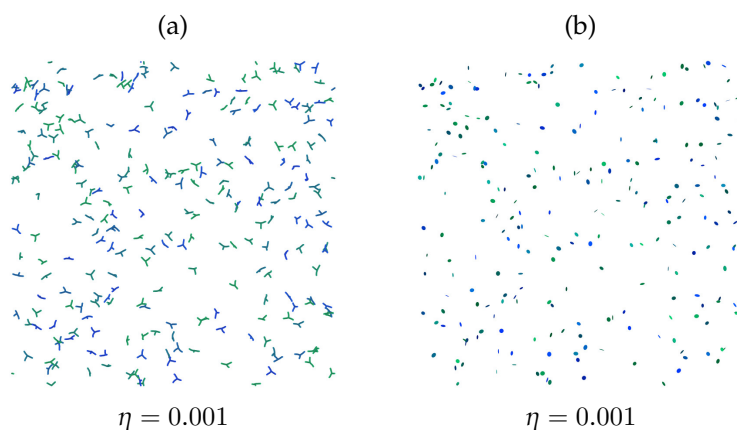
In this section, we present the results of our investigations of systems of G1 DL-DNAs with rigid connections in bulk solutions at three different densities ($\eta = 0.001$, $\eta = 0.1$, and $\eta = 0.5$) and at salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$. In this work, the dimensionless density η is defined by

$$\eta = \rho_\text{d} \frac{4\pi R_\text{g}^3}{3} \quad , \text{ where } \quad \rho_\text{d} = \frac{N_\text{d}}{V_\text{b}} . \tag{6.1}$$

Here, R_g is the radius of gyration, N_d denotes the number of dendrimers in the system, and V_b is the volume of the simulation box. As such, the dimensionless density η resembles a packing fraction in a system of spherical particles where the volume of the system V_b is rescaled by the volume of a sphere with radius R_g. The following results were obtained by employing approaches (a) and (c) discussed in the beginning of this chapter. Simulation snapshots of bulk systems at the examined densities η and at salt concentration $c_1 = 150\,\text{mM}$ are shown in Figures 6.1(a), 6.2(a), and 6.3(a).

The snapshots in Figures 6.1(b), 6.2(b), and 6.3(b) were obtained by replacing the DL-DNA molecules with their gyration ellipsoid, *i.e.*, an ellipsoid where the lengths of its principal axes correspond to the eigenvalues of the gyration tensor, λ_i, with $i = 1, 2, 3$.[212,213]

[212] See Table C.1 in Appendix C.2 for a list of values of λ_i, with $i = 1, 2, 3$, for the investigated systems.

[213] See Figures C.5-C.7 in Appendix C.2 for the corresponding simulation snapshots of systems at salt concentration $c_2 = 500\,\text{mM}$.

(a) (b)

$\eta = 0.1$ $\eta = 0.1$

The pair correlation function $g(r)$ and the structure factor $S(k)$ of DL-DNA systems at densities $\eta \in \{0.5, 2.0, 5.0\}$ and at salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$ were obtained *via* application of the HNC method[214] to the effective potentials $\phi_{\text{eff}}(r)$ shown in Figures 5.3(a) and 5.4(a).

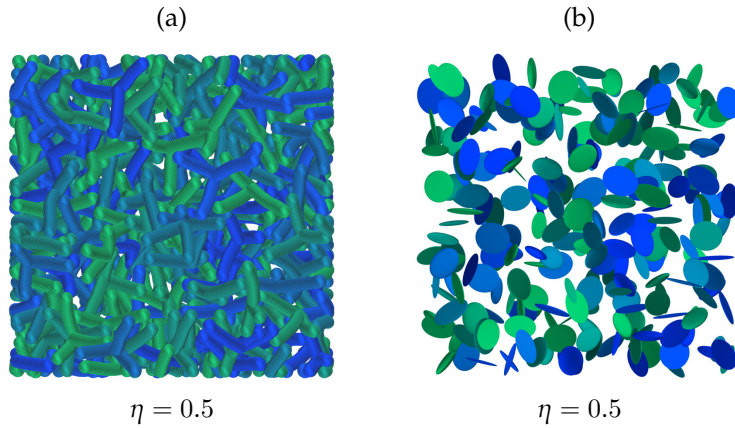[214] See Section 2.4 for an explanation of the HNC method.

(a) (b)

$\eta = 0.5$ $\eta = 0.5$

As shown in Section 2.3.1, the pair correlation function, $g(r)$, in a strongly dilute system, *i.e.*, $\eta \to 0$, can be calculated *via* the Boltzmann factor:[215]

[215] See Equation (2.34).

$$g(r, \eta \to 0) = e^{-\beta \phi_{\text{eff}}(r)} , \qquad (6.2)$$

with $\beta = 1/k_{\text{B}}T$. See Figure 6.4 for the results.

Clearly, the pair correlation functions obtained *via* the HNC method converge towards the results for $g(r)$ obtained using Equation (6.2) for both salt concentrations, c_1 and c_2, corroborating the validity of the HNC re-

sults. In the high density limit, on the other hand, a clear trend can be observed: as the density η increases, the dendrimers get pushed closer together, resulting in an ever more pronounced peak in $g(r)$ at smaller inter-dendrimer distances r. While this observation holds true for systems at both salt concentrations, c_1 and c_2, individually, a difference $g(r)$ in appears when comparing systems at c_1 and c_2 against each other: increasing the salt concentrations c leads to a smaller Debye length λ_D and thus to a stronger screening of the electrostatic interaction. As a consequence the minimum inter-dendrimer distance[216] occurs at $r_{min} \approx 5\sigma$ for $c_2 = 500\,\text{mM}$, while r_{min} approaches 10σ for $c_1 = 150\,\text{mM}$.

[216]Here, we define the minimum inter-dendrimer distance, r_{min}, as the largest value of r where $g(r) = 0$, e.g., $r_{min} \approx 9\sigma$ for c_1 and $r_{min} \approx 5\sigma$ for c_2 in Figure 6.4.

FIGURE 6.4: Radial distribution functions $g(r)$ of G1 DL-DNAs obtained *via* the HNC method as functions of r, given in units of steric interaction length σ. Results are shown at salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$, as labelled. The appropriate effective potentials $\phi_{eff}^{c_1}(r)$ and $\phi_{eff}^{c_2}(r)$ employed in the HNC method are taken from Section 5.2.1. The solid and dashed black lines show the Boltzmann factors for $\phi_{eff}^{c_1}(r)$ and $\phi_{eff}^{c_2}(r)$, which correspond to $g(r)$ in the low-density regime (see text). The investigated densities range from $\eta = 0.5$ to $\eta = 5.0$ (as labelled).
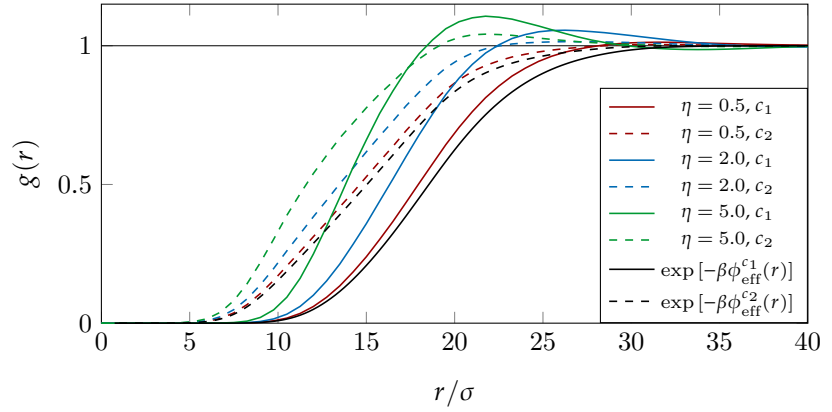
The structure factors $S(k)$ shown in Figure 6.5 reveal similar insights. The isothermal compressibility of a system, χ_T can be obtained *via* the compressibility equation:[217]

[217] See Equation (2.49).

$$\lim_{k \to 0} S(k) = \rho_d\, k_B T\, \chi_T\,,\tag{6.3}$$

with dendrimer density ρ_d. In this low-k limit of $S(k)$ the system is probed over large wavelengths. By comparing the value $S(k = 0)$ in Figure 6.5 for different densities η and for different salt concentrations c, we can see that compressibility χ_T grows with increasing c and with decreasing η.

While the previously discussed results are perfectly consistent with a system of point-like particles interacting *via* the radially symmetric effective interactions $\phi_{eff}(r)$, the structure descriptors listed in Table C.2 in Appendix C.2 do not correspond to spherical conformations.
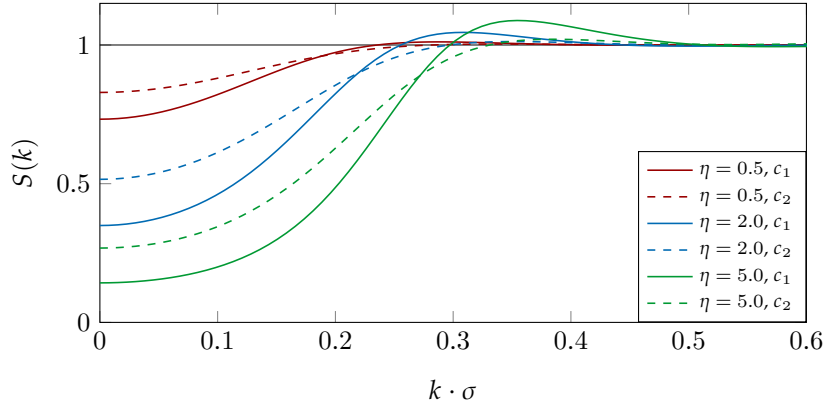
FIGURE 6.5: Structure factors $S(k)$ of G1 DL-DNAs obtained *via* the HNC method as functions of k, given in units of the inverse steric interaction length σ^{-1}; Results are shown for results for salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$, as labelled. The appropriate effective potentials $\phi_{\text{eff}}^{c_1}(r)$ and $\phi_{\text{eff}}^{c_2}(r)$ employed in the HNC method are taken from Section 5.2.1. The investigated densities range from $\eta = 0.5$ to $\eta = 5.0$ (as labelled).

Therefore, we need to gain structural information of the investigated bulk solutions by investigating the MD simulations using our coarse-grained model.[218] The spatial distribution of the center-of-mass positions of the dendrimers was analysed by calculating the respective pair correlation functions $g(r)$ and the structure factors $S(k)$ from MD-simulations at densities $\eta = 0.1$ and $\eta = 0.5$ and at salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$, see Figures 6.6 and 6.7.

[218] *I.e.*, approach (a) of the approaches listed in the introduction to this chapter.
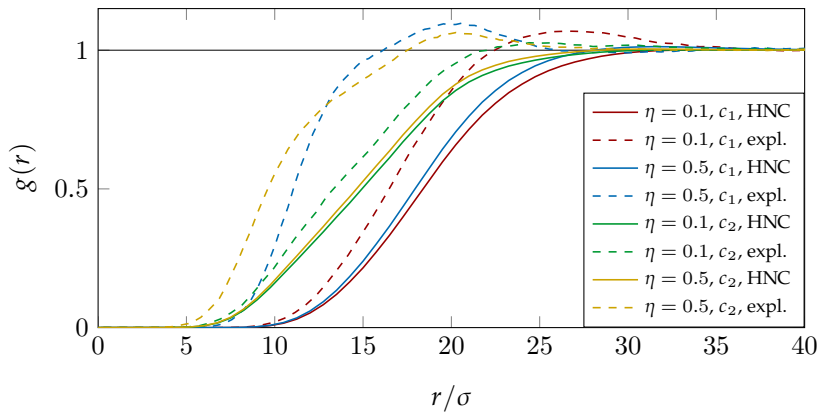
FIGURE 6.6: Radial distribution functions $g(r)$ of G1 DL-DNAs as functions of r, given in units of steric interaction length σ; results were obtained *via* the HNC method (labelled "HNC") and explicit MD simulations (labelled "expl."). The appropriate effective potentials $\phi_{\text{eff}}^{c_1}(r)$ and $\phi_{\text{eff}}^{c_2}(r)$ employed in the HNC method are taken from Section 5.2.1. The systems were investigated at $\eta = 0.1$ and $\eta = 0.5$ (as labelled) and at salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$ (as labelled).

The clear discrepancy between HNC and MD results (see Figures 6.6 and 6.7) is a strong indication that a spherical symmetric effective potential $\phi_{\text{eff}}(r)$ is not adequate to describe G1 DL-DNAs due to their aspherical shape. Whereas the value of the minimum inter-dendrimer distance, r_{min}, obtained *via* HNC and MD results assumes essentially the same value, the stronger increase of the pair correlation function $g(r)$ at $r \geq r_{\text{min}}$ indicates a closer packing of the dendrimers. This observation is corroborated by a considerably smaller value for the compressibility $\chi_T \propto S(k = 0)$ for MD

results as compared to HNC results.

FIGURE 6.7: Structure factors $S(k)$ of G1 DL-DNAs as functions of k, given in units of the inverse steric interaction length σ^{-1}; results were obtained *via* the HNC method (labelled "HNC") and explicit MD simulations (labelled "expl."). The appropriate effective potentials $\phi_{\mathrm{eff}}^{c_1}(r)$ and $\phi_{\mathrm{eff}}^{c_2}(r)$ employed in the HNC method are taken from Section 5.2.1. The systems were investigated at $\eta = 0.1$ and $\eta = 0.5$ (as labelled) and at salt concentrations $c_1 = 150\,\mathrm{mM}$ and $c_2 = 500\,\mathrm{mM}$ (as labelled).
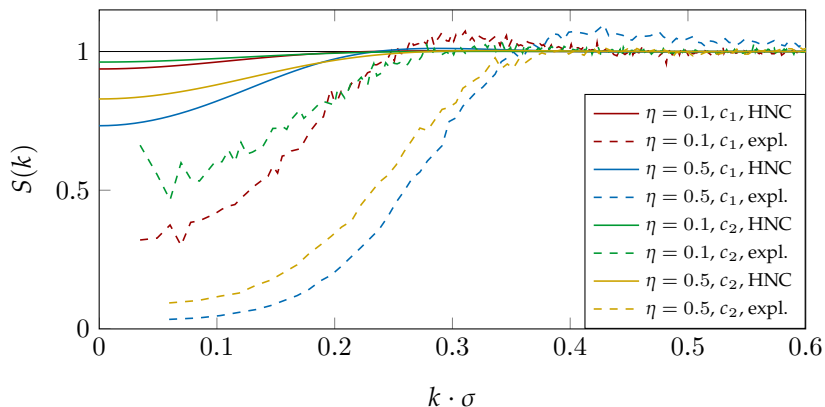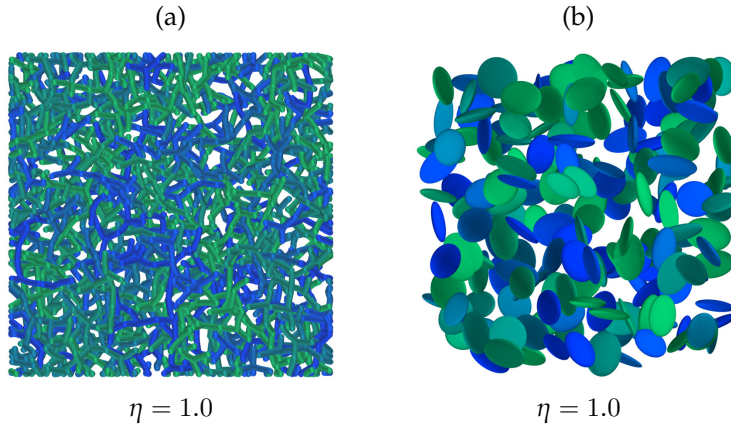

6.2 Bulk Solutions of 2nd Generation DL-DNAs

FIGURE 6.8: Simulation snapshots of systems containing G2 DL-DNA dendrimers at density $\eta = 0.5$ and at salt concentration $c_2 = 150\,\mathrm{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).
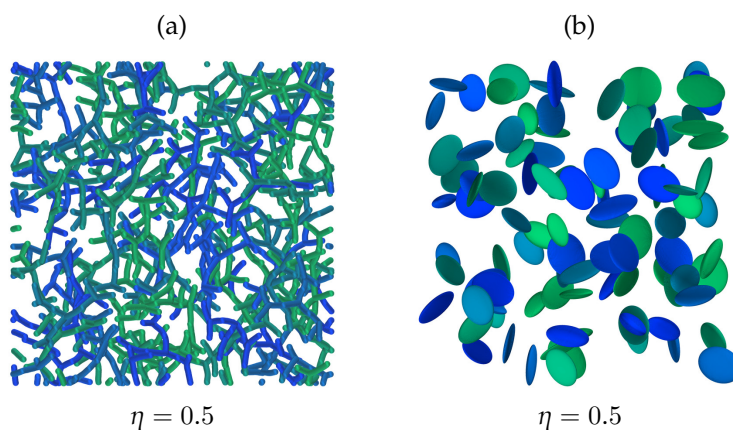
While the conclusion of Section 6.1 was that a spherically symmetric effective potential $\phi_{\mathrm{eff}}(r)$ is unsuitable to describe aspherical dendrimers, such as G1 DL-DNAs, the shape descriptor values listed in Table C.2 in Appendix C.2 indicate an increasing sphericity with increasing generation number GN. Thus, we found it appropriate to reapply the approaches used in Section 6.1, *i.e.*, the HNC method and MD simulations, to bulk solutions of G2 dendrimers.

In this section, we investigate systems of G2 DL-DNAs with rigid connections and with flexible connections in bulk solution at three different densities ($\eta = 0.5$, $\eta = 1.0$, and $\eta = 2.0$) and at salt concentrations $c_1 = 150\,\mathrm{mM}$

and $c_2 = 500\,\text{mM}$.

(a)   (b)

$\eta = 1.0$ $\eta = 1.0$

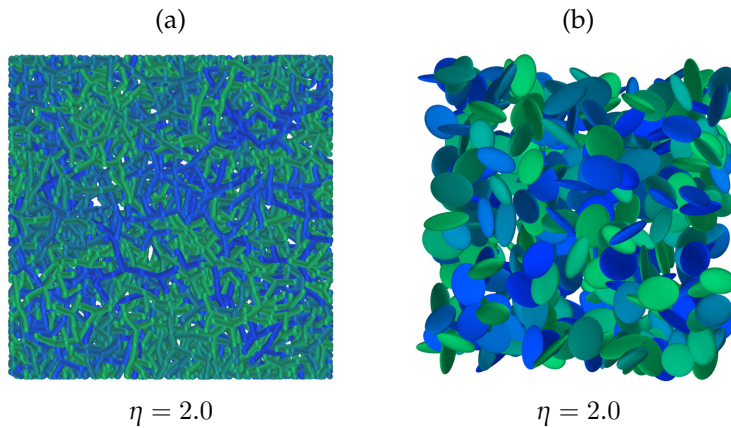**FIGURE 6.9:** Simulation snapshots of systems containing G2 DL-DNA dendrimers at density $\eta = 1.0$ and at salt concentration $c_2 = 150\,\text{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).

These results were obtained by employing approaches (a), (b), and (c) discussed in the beginning of this chapter. Figures 6.8(a)-6.10(a) and 6.8(b)-6.10(b) show snapshots of the examined systems with representations *via* explicit dendrimers and gyration ellipsoids, respectively.[219]

[219] See Figures C.8-C.10 in Appendix C.2 for the corresponding simulation snapshots of systems at salt concentration $c_2 = 500\,\text{mM}$

(a)   (b)

$\eta = 2.0$ $\eta = 2.0$

FIGURE 6.10: Simulation snapshots of systems containing G2 DL-DNA dendrimers at density $\eta = 2.0$ and at salt concentration $c_2 = 150\,\text{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).

By applying the HNC method to the effective potentials $\phi_{\text{eff}}(r)$ shown in Figures 5.5(a), 5.6(a), 5.7(a), and 5.6(a) the pair correlation functions, $g(r)$, and the structure factors, $S(k)$, of G2 DL-DNAs with rigid and with flexible connections at densities $\eta \in \{0.1, 0.5, 1.0, 2.0, 5.0\}$ and at salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$ were calculated, see Figures 6.11 and 6.12.

Looking at the results for $g(r)$ and $S(k)$ for dendrimers with rigid and with flexible connections separately, we observe similar trends as in Figures 6.4 and 6.5: in both cases increasing density η results in a denser packing and a

reduced isothermal compressibility χ_T, whereas increasing the salt concentration c and thus strengthening the screening of the Coulomb interaction leads to a slightly denser packing and a larger isothermal compressibility χ_T.

FIGURE 6.11: Radial distribution functions $g(r)$ of G2 DL-DNAs obtained *via* the HNC method as functions of r, given in units of steric interaction length σ; results are shown for G2 DL-DNAs with rigid connections (labelled "rigid") and with flexible connections (labelled "flex."). The appropriate effective potentials $\phi_{\text{eff}}^{c_1}(r)$ and $\phi_{\text{eff}}^{c_2}(r)$ employed in the HNC method are taken from Section 5.2.2. The investigated densities range from $\eta = 0.1$ to $\eta = 0.5$ (as labelled) with subplots (a) and (b) depicting results for salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$, respectively.
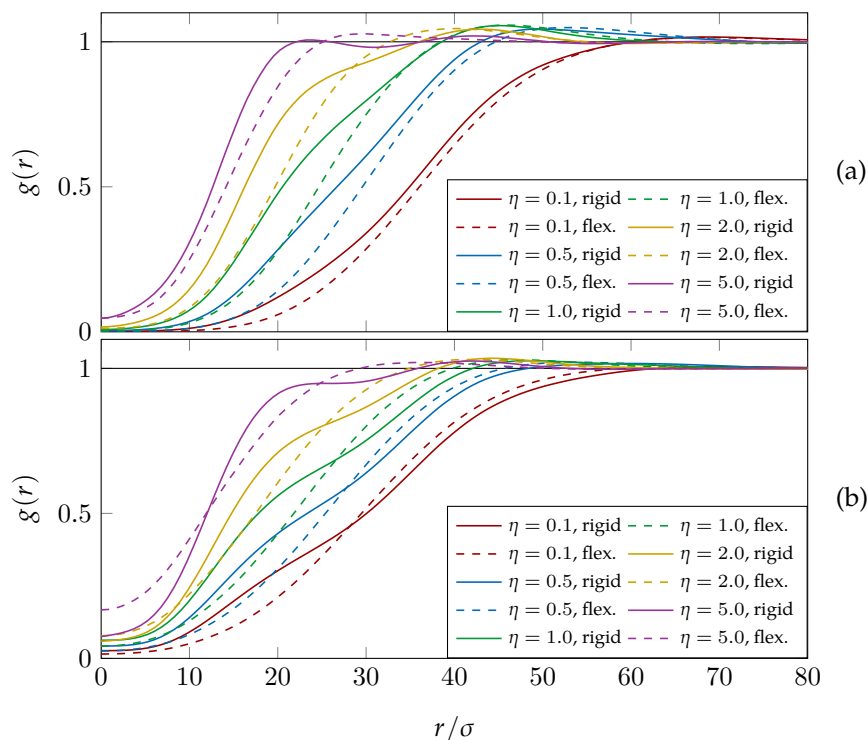
However, when comparing the results for dendrimers with rigid and with flexible connections in Figures 6.11 and 6.12, a striking difference can be observed: while a "shoulder" emerges[220] in $g(r)$ and $S(k)$ with increasing density η for DL-DNAs with rigid connections, no such feature can be found in $g(r)$ and $S(k)$ for DL-DNAs with flexible connections.

This phenomenon is a direct consequence of the two length scales observed in the effective potentials shown in Figures 5.5(a) and 5.6(a), which in turn are a consequence of the structural features of G2 DL-DNAs with rigid connections. The increased flexibility of dendrimers with flexible connections suppresses the emergence of the two length scales shown in the effective potentials in Figures 5.7(a) and 5.8(a) and thus prevents the "shoulder" in $g(r)$ and $S(k)$ from appearing. This observation demonstrates the strong

[220] *E.g.*, at $r \approx 20\sigma$ in Figures 6.11(a) and 6.11(b).

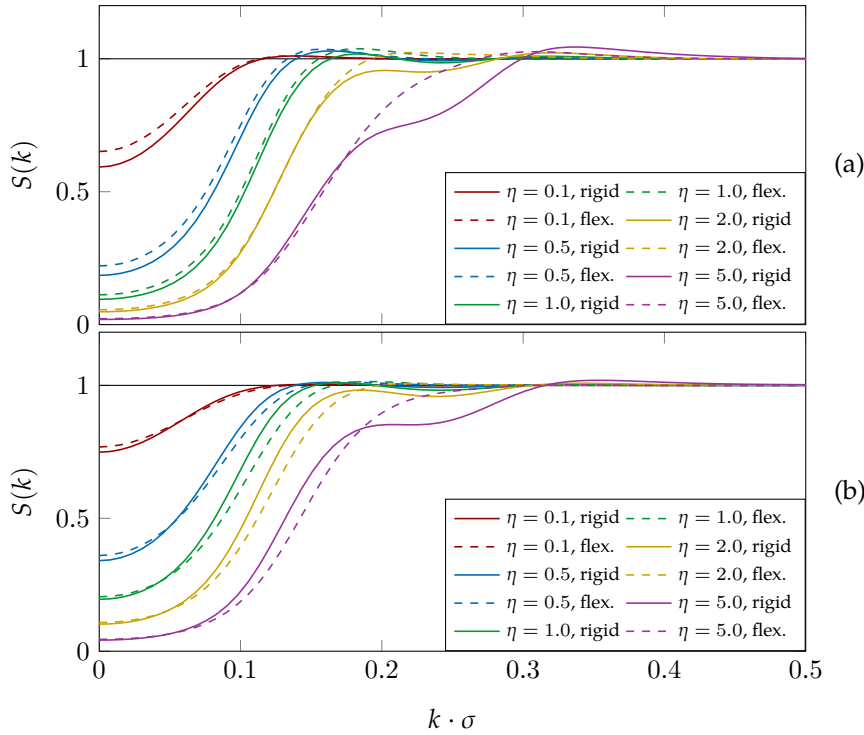impact of a slight change in the structural design of DL-DNA dendrimers.

Analogous to Section 6.2, the results discussed so far in this section are only consistent with a system of point-like particles interacting *via* radially symmetric effective interactions $\phi_{\text{eff}}(r)$. In order to prove the validity of the HNC results shown in Figures 6.11 and 6.12 for a system of proper DL-DNAs, *i.e.*, not effective particles but molecules with internal structure, we compare $g(r)$ and $S(k)$ obtained *via* the HNC method with the corresponding results from MD simulations, see Figures 6.13 and 6.14. There, we show results for $g(r)$ and $S(k)$ for systems of DL-DNA at different densities $\eta \in \{0.5, 1.0, 2.0\}$ and at salt concentration $c_1 = 150\,\text{mM}$ obtained *via* the three different approaches discussed in the introduction to this chapter: results extracted from MD simulations of our coarse-grained model, *i.e.*, approach (a), results extracted from MD simulations of effective particles, *i.e.*, approach (b), and results calculated *via* the HNC method, *i.e.*, approach (c). While the dotted lines shown in Figure 6.13, *i.e.*, the results obtained *via* approach (a), correspond to $g(r)$ of the center-of-mass coordinates of the simulated dendrimers, the corresponding dotted lines presented in Fig-

[221] This indirect approach of calculating $S(k)$ is necessary due to the low number of samples obtained *via* approach (a). See Figure C.13 in Appendix C.2 for a comparison of $S(k)$ obtained by direct calculation from samples and by Fourier transform of $g(r)$.

ure 6.14, *i.e.*, the results for $S(k)$ obtained *via* approach (a), were calculated by Fourier transform of a smoothed version of $g(r)$ obtained *via* approach (a), *i.e.*, the dotted lines shown Figure 6.13.[221]

FIGURE 6.13: Radial distribution functions $g(r)$ of G2 DL-DNAs with rigid connections as functions of r, given in units of steric interaction length σ; results were obtained *via* the HNC method (labelled "HNC"), MD simulations of effective particles (labelled "eff."), and explicit MD simulations (labelled "expl."). The investigated densities range from $\eta = 0.5$ to $\eta = 2.0$ (as labelled) with subplots (a) and (b) depicting results for salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$, respectively.
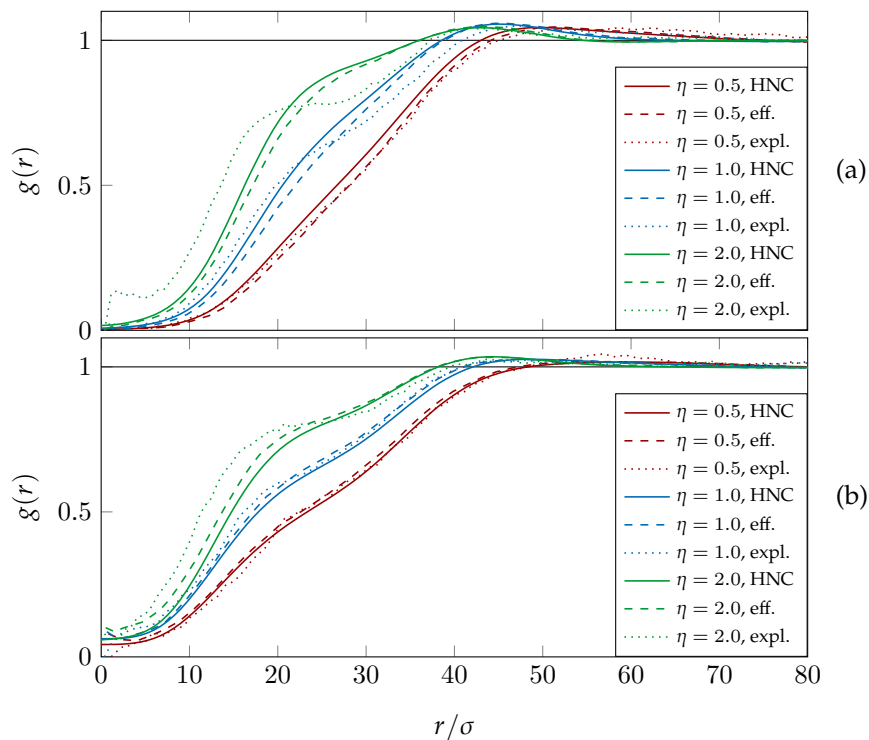
As expected, both $g(r)$ and $S(k)$ computed by the HNC method (approach (a)) and by MD simulations of point-like particles interacting *via* radially symmetric, effective potentials $\phi_{\text{eff}}(r)$ (approach (b)) are consistent. Furthermore, the results obtained *via* MD simulations of our coarse-grained model (approach (c)) shown in Figures 6.13 and 6.14 show excellent agreement with the results obtained *via* approaches (a) and (b), where only the results for $\eta = 2.0$ and $c_1 = 150\,\text{mM}$ presented in Figure 6.13(a) differ somewhat. Therefore, our assumption that G2 DL-DNAs can be represented by point-like, effective particles interacting *via* the spherically symmetric effective potentials ϕ_{eff} proves valid.

As in Chapter 4, we compare our theoretical results with those extracted from experiments, or more exactly, we compare the theoretical structure fac-

tors, $S(k)$, with the corresponding results obtained *via* static light scattering experiments.[222]
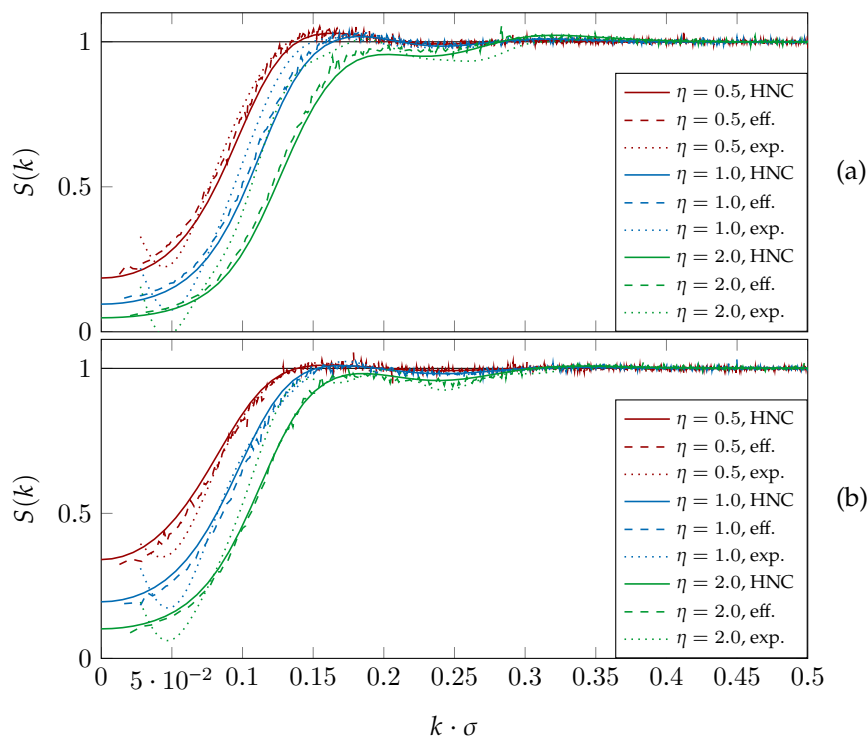
FIGURE 6.14: Structure factors $S(k)$ of G2 DL-DNAs with rigid connections as functions of k, given in units of the inverse steric interaction length σ^{-1}; results were obtained *via* the HNC method (labelled "HNC"), MD simulations of effective particles (labelled "eff.") and explicit MD simulations (labelled "expl."). The investigated densities range from $\eta = 0.5$ to $\eta = 2.0$ (as labelled) with subplots (a) and (b) depicting results for salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$, respectively.

In Figure 6.15 the SLS results at a NaCl salt concentration $c_1 = 150\,\text{mM}$ are shown for densities ranging from $\eta = 0.202$ to $\eta = 5.767$, where R_g in Equation (6.1) is assumed to be $R_\text{g} = 9.4\,\text{nm}$. On a first glance, the most distinctive finding is the emergence of a peak in $S(k)$ with increasing η. Whereas such a development can also be surmised in previously discussed theoretical results for the structure factor $S(k)$,[223] there, this trend is nowhere near as pronounced. Furthermore, the peak in $S(k)$ shown in Figure 6.15 is detected at much higher values of k, indicating that in experiment the DL-DNA dendrimers at high densities η are located at closer inter-dendrimer distances than observed in the theoretical results shown in Figure 6.12.. This conclusion is corroborated by Figure 6.16, where the structure factor, $S(k)$, measured in experiment is compared to $S(k)$ calculated *via* the HNC method.

FIGURE 6.15: Structure factors $S(k)$ of G2 DL-DNAs with rigid connections as functions of k, given in units of the inverse steric interaction length σ^{-1}; results were obtained from experiment *via* SLS at salt concentration $c_1 = 150\,\mathrm{mM}$. Experimental measurements were made at densities ranging from $\eta = 0.202$ to $\eta = 5.767$ (as labelled).
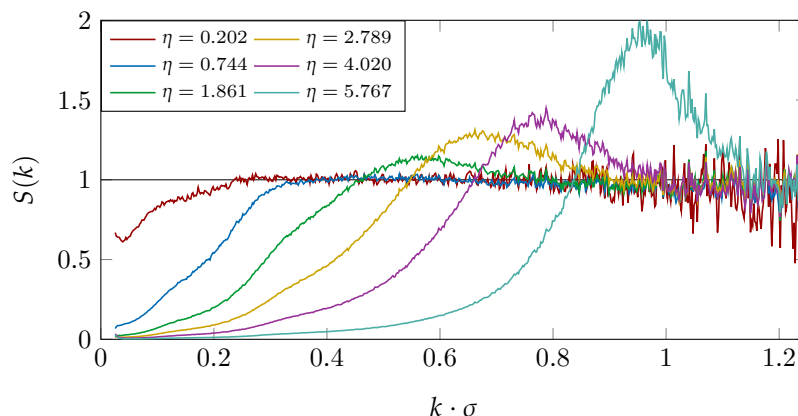
While the results are consistent for low density $\eta = 0.202$ (red lines), the discrepancy between theory and experiment becomes stronger when investigating bulk solutions at larger densities $\eta = 0.744$ and $\eta = 2.200$ (blue and green lines). Still, it is worth noticing that a "shoulder" feature in $S(k)$, as observed in theoretical results shown in Figure 6.12, also seems to appear in experimental measurements for intermediate densities η, *e.g.*, at $\eta \in \{0.744, 1.487, 1.861\}$ in Figure C.12. Due to the good agreement between HNC and explicit MD results in Figure 6.14, we can safely assume that the discrepancy between explicit MD simulations and experiment is of similar magnitude.

Obviously, the coarse-grained model discussed in Section 3.2 does not fully capture the complex behaviour of DL-DNAs at high densities η (while still being valid in the low density regime and for isolated dendrimers, as shown in Chapter 4). Specifically, the experimental results indicate smaller inter-dendrimer distances than observed in theory. One possible reason for this is the reduction of the effective charge carried by the phosphate backbone of DNA due to charge neutralisation by condensation and adsorption of counterions on DNA.[224,225]

Furthermore, the signal in experimental measurements stems from physical scatterers, *i.e.*, particles, which are not necessarily located at the centers-of-mass, r_{com}. On the contrary, r_{com} is often located in "empty" space. To mitigate this effect the center-of-mass of the three central junction-monomers of the innermost subgeneration g_1, r_{cy}, can be used as the effective coordinate of the dendrimers.

[224] This effect is not to be confused with the screening of the electrostatic interaction due to salt ions.

[225] "...the vast majority of the remaining 90% is Manning condensed on the rods ..." (K. Kegler et al. "Polyelectrolyte-Compression Forces between Spherical DNA Brushes". In: *Phys. Rev. Lett.* 100 (2008), p. 118302. DOI: 10.1103/PhysRevLett.100.118302, p. 4). See also A. A. Kornyshev et al. "Structure and interactions of biological helices". In: *Rev. Mod. Phys.* 79 (2007), pp. 943–996. DOI: 10.1103/RevModPhys.79.943, pp. 955-956.
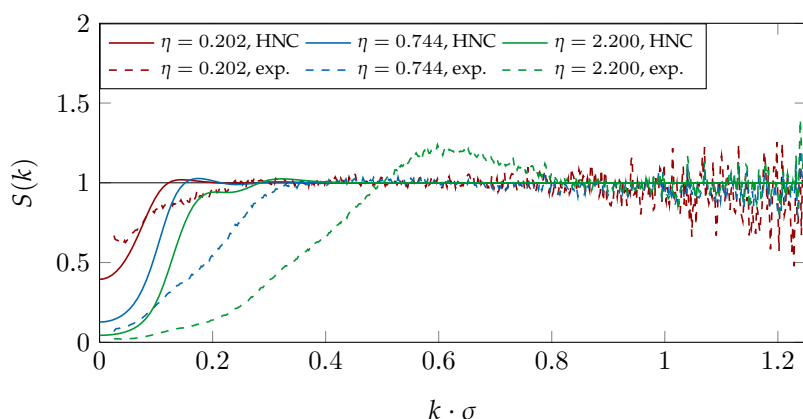
FIGURE 6.16: Structure factors $S(k)$ of G2 DL-DNAs with rigid connections as functions of k, given in units of the inverse steric interaction length σ^{-1}; comparison of results obtained *via* the HNC method (labelled "HNC") with results extracted from experiment (labelled "exp.") at salt concentration $c_1 = 150\,\text{mM}$ and at selected densities η (as labelled).

In order to further investigate these two issues, *i.e.*, reducing the charge of DNA and using r_{cy} as the effective coordinate, an exploratory study was undertaken. In this study, we reduced the charge carried by DNA monomers in our coarse-grained model by different fractions f of the original value, denoted by $q \times f$, with $f \in \{0.25, 1.0\}$, and we used r_{cy} in addition to r_{com} as the effective coordinate of the DL-DNA dendrimers. The effective potentials $\phi_{\text{eff}}(r)$ calculated *via* the WI method using r_{com} as the effective coordinate for different charge reductions $q \times f$, with $f \in \{0.25, 0.75, 1.0\}$, can be seen in Figure C.1 in Appendix C.1, while the corresponding $\phi_{\text{eff}}(r)$ using r_{cy} as the effective coordinate are shown for different charge reductions $q \times f$, with $f \in \{0.25, 1.0\}$ in Figure C.3. A comparison between $\phi_{\text{eff}}(r)$ with r_{com} and r_{cy} as effective coordinate for $q \times 0.25$ is given in Figure C.2.

Applying the HNC method to these potentials yiels the corresponding pair correlation functions $g(r)$ and structure factors $S(k)$, see Figures C.14 to C.18 in Appendix C.2. Whereas implementing these changes results in a development in the desired direction, *i.e.* more pronounced peaks in $S(k)$ (*e.g.*, in Figure C.18), the persisting difference to experimental measurements suggests that these changes are still insufficient.

A promising method for converging towards consistency with experimental results are anisotropic effective interactions:[226] in order to allow small interdendrimer distances in the effective description, the anisotropy, *i.e.* asphericity, of the DL-DNAs must be taken into account. Adopting an anisotropic effective interaction allows for the effective particles' shapes to

deviate from the spherical form. Research in this direction is ongoing.

6.3 Bulk Solutions of 3rd Generation DL-DNAs

FIGURE 6.17: Simulation snapshots of systems containing G3 DL-DNA dendrimers at density $\eta = 0.01$ and at salt concentration $c_2 = 150\,\text{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).

(a)   (b)
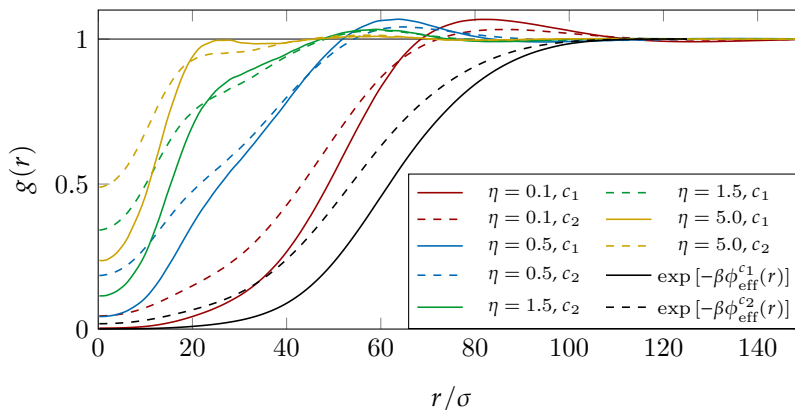
$\eta = 0.01$ $\eta = 0.01$
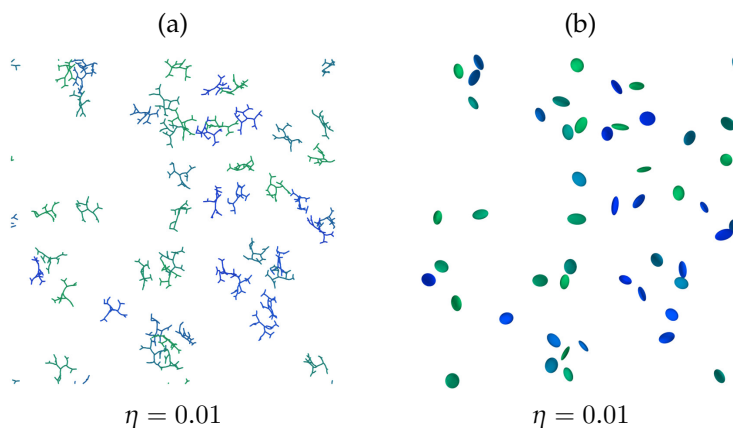
Finally, we present the results obtained by applying the HNC method to the effective potentials $\phi_{\text{eff}}(r)$ of G3 DL-DNAs with rigid connections shown in Figures 5.9(a) and 5.10(a).[227] While performing large-scale MD simulations of such G3 DL-DNAs using our coarse-grained model, *i.e.*, approach (a) discussed in the introduction to this chapter, proved computationally unfeasible at densities $\eta \gtrsim 0.1$, the reader is directed to Figure 6.17 for a visualisation of such systems at a low density of $\eta = 0.01$ and at salt concentration $c_1 = 150\,\text{mM}$.[228] The pair correlation functions, $g(r)$, and the structure factor, $S(k)$, calculated *via* the HNC method for G3 DL-DNAs with rigid connections can be seen in Figures 6.18 and 6.19.

[227] These potentials were calculated for the full DNA charge ($q \times 1.0$) and using the center-of-mass, r_{com}, as the effective coordinate.

[228] These snapshots were generated from exploratory simulations run for the investigations in Section 5.4. See Figure C.11 in Appendix C.2 for the corresponding simulation snapshots of systems at salt concentration $c_2 = 500\,\text{mM}$

FIGURE 6.18: Radial distribution functions $g(r)$ of G3 DL-DNAs obtained *via* the HNC method as functions of r, given in units of steric interaction length σ. Results are shown at salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$, as labelled. The appropriate effective potentials $\phi_{\text{eff}}^{c_1}(r)$ and $\phi_{\text{eff}}^{c_2}(r)$ employed in the HNC method are taken from Section 5.2.3. The solid and dashed black lines show the Boltzmann factors for $\phi_{\text{eff}}^{c_1}(r)$ and $\phi_{\text{eff}}^{c_2}(r)$, which correspond to $g(r)$ in the low-density regime (see text). The investigated densities range from $\eta = 0.1$ to $\eta = 5.0$ (as labelled).
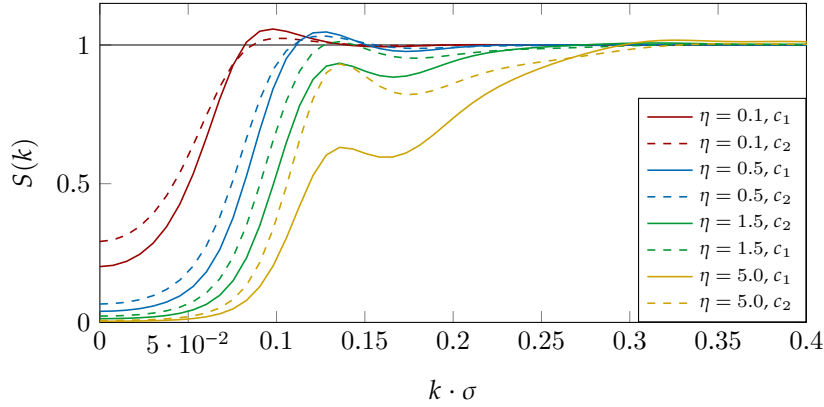

Legend:
$\eta = 0.1, c_1$ — $\eta = 1.5, c_1$
$\eta = 0.1, c_2$ — $\eta = 5.0, c_1$
$\eta = 0.5, c_1$ — $\eta = 5.0, c_2$
$\eta = 0.5, c_2$ — $\exp\left[-\beta\phi_{\text{eff}}^{c_1}(r)\right]$
$\eta = 1.5, c_2$ — $\exp\left[-\beta\phi_{\text{eff}}^{c_2}(r)\right]$

FIGURE 6.19: Structure factors $S(k)$ of G3 DL-DNAs obtained *via* the HNC method as functions of k, given in units of the inverse steric interaction length σ^{-1}; Results are shown for results for salt concentrations $c_1 = 150\,\mathrm{mM}$ and $c_2 = 500\,\mathrm{mM}$, as labelled. The appropriate effective potentials $\phi_{\mathrm{eff}}^{c_1}(r)$ and $\phi_{\mathrm{eff}}^{c_2}(r)$ employed in the HNC method are taken from Section 5.2.3. The investigated densities range from $\eta = 0.1$ to $\eta = 5.0$ (as labelled).

As in Section 6.1 the pair correlation functions, $g(r)$, obtained *via* the HNC method are consistent with $g(r)$ calculated using Equation (6.2). Furthermore, we observe similar trends as seen in Figures 6.11 and 6.12: an increase in density η results in a denser packing and a reduced isothermal compressibility χ_T, whereas an increase in the salt concentration c yields a denser packing and a larger isothermal compressibility $\chi_T \propto S(k = 0)$. However, in comparison to the results for G2 dendrimers with rigid connections in Section 6.2 the "shoulder" feature in $g(r)$ and $S(k)$ is much more pronounced for G3 dendrimers.

7 DNA Stars in Two Dimensions

You are living on a Plane. What you style Flatland is the vast level surface of what I may call a fluid, on, or in, the top of which you and your countrymen move about, without rising above it or falling below it.

—Edwin Abbott Abbott, *Flatland*[229]

Up until now, our investigations focused on bulk solutions of DNA dendrimers at densities ranging from the dilute regime (see Chapter 4) to the dense regime (see Chapter 6). In this chapter, we turn our attention to DL-DNAs at interfaces. For this purpose, we investigate two different types of G1 DNA stars, which have been the subject of previous research.[230,231]

Specifically, we consider planar DNA stars whose spatial displacements are fully restricted to a two-dimensional interface and compare them to DNA stars where only the endpoints of their arms are confined to a two-dimensional plane. Due to its design the latter kind of particles is equipped with the ability to "dip" into the fluid, rendering the system quasi-two-dimensional.[232]

In analogy with the effective interactions of DL-DNAs in the bulk (see Chapter 5), we calculate the effective potentials of DNA-stars in two dimensions in Section 7.2.

While the design of the DNA stars and their effective interactions are presented in Section 7.1 and Section 7.2, respectively, the results obtained from MD simulations of such systems are shown of Sections 7.3 to 7.5.

All simulations in this chapter were performed by the author using the ESPResSo simulation package. The reader is referred to Appendix A for

[229] See E. A. Abbott. *Flatland: A Romance of Many Dimensions*. London: Oxford University Press, 2008. URL: https://worldcat.org/isbn/9780199537501.

[230] See C. Abaurrea Velasco et al. "Effective interactions of DNA-stars". In: *Mol. Phys.* 113 (2015), pp. 2699–2706. DOI: 10.1080/00268976.2015.1048318 and N. Avakyan et al. "Long-Range Ordering of Blunt-Ended DNA Tiles on Supported Lipid Bilayers". In: *J. Am. Chem. Soc.* 139.34 (2017), pp. 12027–12034. DOI: 10.1021/jacs.7b06572.

[231] Some of the results presented in this chapter are part of a publication in preparation, see C. Jochum et al. "DNA Stars Confined to an Interface: Planar vs. Tripod Configurations". (to be submitted).
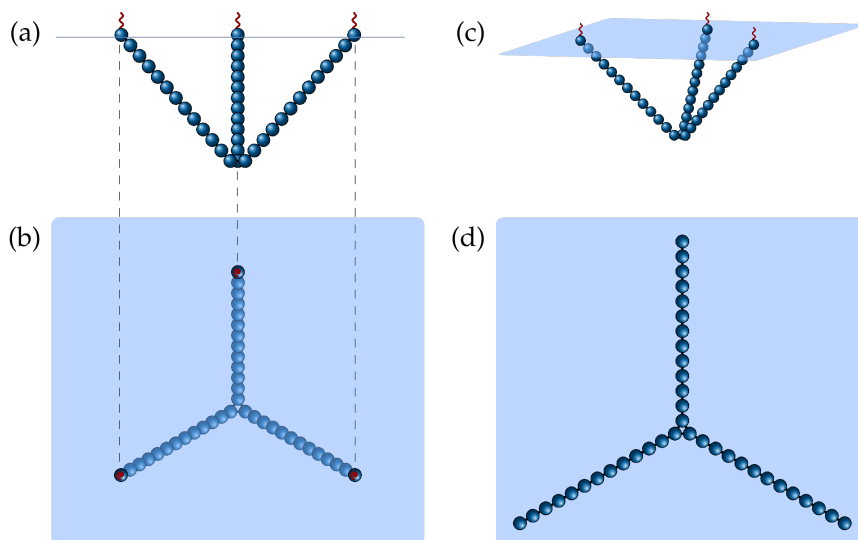
[232] See Section 7.1 for details on this tripod configuration.

more information on the `ESPResSo` software. The relevant simulation details and simulation parameter values are given in Appendix B.5.

7.1 Design and Modelling of DNA Stars

The design of the DNA stars examined in this chapter corresponds to a slightly modified version of G1 DL-DNAs. In contrast to the original DL-DNA architecture presented in Section 3.1, where G1 DL-DNAs possess arms composed of 13 base pairs and a terminal ssDNA sequence of length four, the arms of the modified DNA stars are 13 base pairs long and do not contain any ssDNA elements.

FIGURE 7.1: Illustrations of DNA stars: subplots (a)-(c) show the tripod configuration with the tip immersed into the fluid, whereas the planar configuration can be seen in subplot (d). The blue plane represents the interface, while the blue monomer beads correspond to base pairs (not drawn up to scale). In subplots (a)-(c), the arm endpoints of the DNA star are confined to the interface *via* amphiphilic elements indicated by the red appendices. In subplots (d), the complete DNA star is confined to the interface *via* amphiphilic elements (not shown).


[233]Such configurations can be realised in experiments *via* the appropriate use of amphiphilic elements, see for example Z. Zhao et al. "Amphiphilic DNA Organic Hybrids: Functional Materials in Nanoscience and Potential Application in Biomedicine". In: *International journal of molecular sciences* 19 (2018), p. 2283. DOI: `10.3390/ ijms19082283`.

We assume two different kinds of systems: DNA stars who are completely confined to the interface (labelled "flat" or "planar") and DNA stars where only the end-monomers of the arms are confined to the interface (labelled "tripod").[233] DNA stars in the tripod configuration and the flat configuration are shown in Figures 7.1(a)-(c) and Figures 7.1(d), respectively.

For the computational methods in this chapter, *i.e.*, WI method and MD simulations, we use a theoretical model of DNA stars where the features and interactions in Section 3.2 are adopted.

7.2 Effective Potentials of DNA Stars in Two Dimensions

In this section, the effective potentials $\phi_{\mathrm{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars in two dimensions at salt concentration $c = 100\,\mathrm{mM}$ are examined. In contrast to Chapter 5, we can not assume a radially symmetric r-dependence of ϕ_{eff}, where r denotes the interdendrimer distance. For now, only DNA stars in the flat configuration with rigid, *i.e.*, straight, arms and with interarm angles $\theta_i = 2\pi/3$ ($i = 1, 2, 3$) are considered.[234] Thus, the relevant degrees of freedom of the effective interactions are angle α and the displacements Δx and Δy: while α denotes the angle of rotation between the two stars, Δx and Δy correspond to the distances in x- and y-direction between the centers-of-mass of the stars, see Figure 7.2.
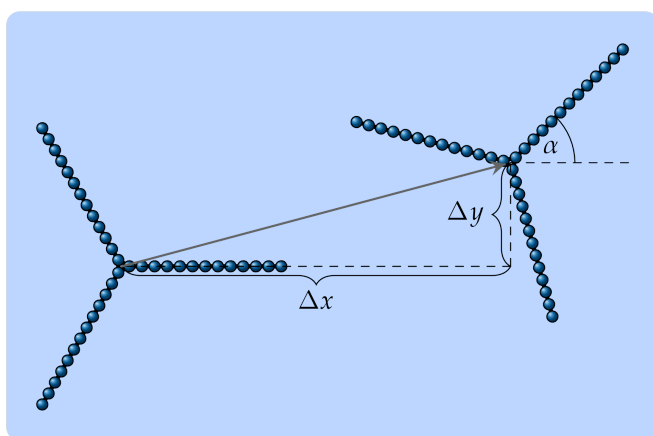
[234]Due to the highly symmetric nature of this DNA star conformation the center-of-mass, r_{com}, of the star conincides with the center-of-mass of the three central junction-monomers, r_{cy}.

FIGURE 7.2: Illustration of the parameters of the effective potential $\phi_{\mathrm{eff}}(\alpha; \Delta x, \Delta y)$ between rigid DNA stars in the flat configuration: rotation angle α and displacements Δx and Δy (as labelled and see text).

Because the stars are assumed to be completely rigid in this investigation, only one single Widom insertion per parameter triple α, Δx, and Δy is necessary in order to determine the interaction energy $\phi_{\mathrm{eff}}(\alpha; \Delta x, \Delta y)$ between the two stars.[235] For comparison, an analytical expression for $\phi_{\mathrm{eff}}(\alpha; \Delta x, \Delta y)$ that can be numerically evaluated *via* fast Fourier-transform taken from literature has been considered.[236]

[235] As every Widom insertion will yield exactly the same energy, statistically averaging over a large number of Widom insertion is rendered pointless. Nevertheless, we will keep the label "Widom insertion method" for consistency's sake.

[236] See Abaurrea Velasco et al., "Effective interactions of DNA-stars". This method will be labelled "reference method", for brevity's sake.

FIGURE 7.3: Effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars as function of center-of-mass shift Δx and Δy and rotation angle α. $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ is given in units of $k_{\text{B}}T$, whereas Δx and Δy are given in Å. Both DNA stars are assumed to be in a planar configuration with rigid arms and interarm angles $\theta_i = 2\pi/3$ ($i = 1, 2, 3$) while being rotated by different values of angle α against each other (as labelled). Two dashed, orthogonal lines in red and green indicate two slices through the energy landscape of the effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$. The energy profiles of these slices, denoted by $\phi_{\text{eff}}(\alpha; \Delta r_1)$ and $\phi_{\text{eff}}(\alpha; \Delta r_2)$, are shown in the plots framed red and green. The data for this figure was obtained *via* the reference method outlined in Abaurrea Velasco et al., "Effective interactions of DNA-stars" (subplots (a) and (c)) and *via* the WI method (subplots (b) and (d)).
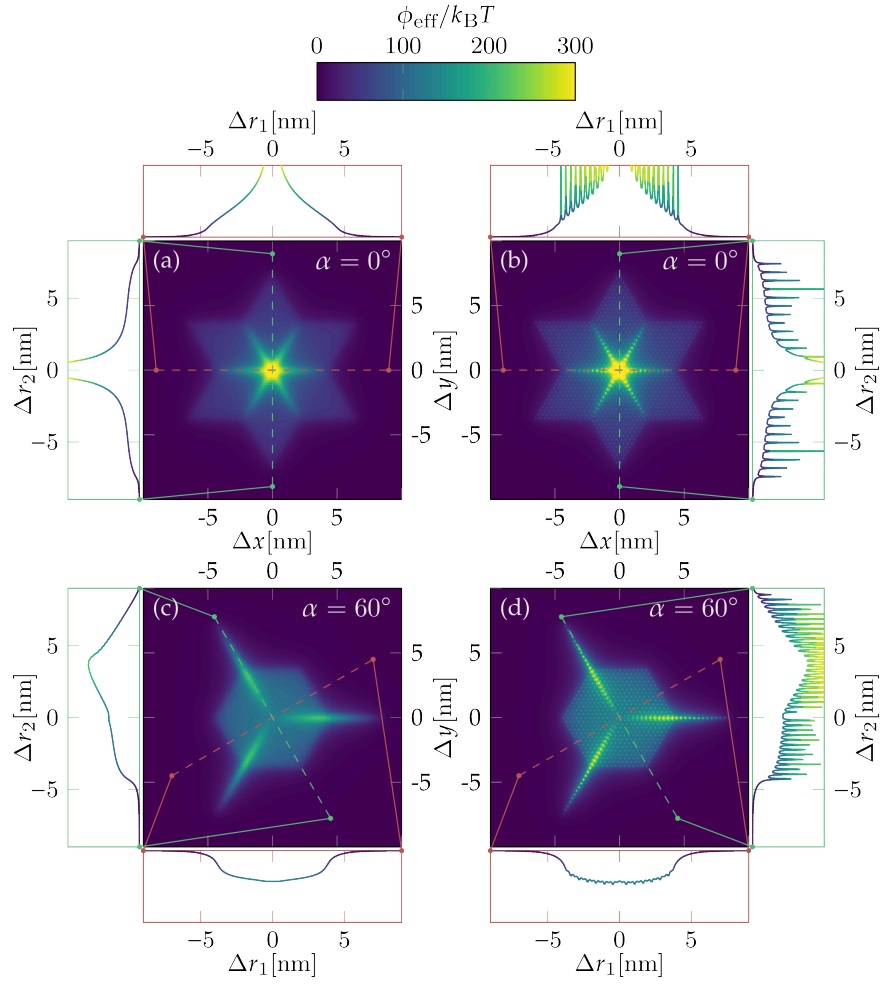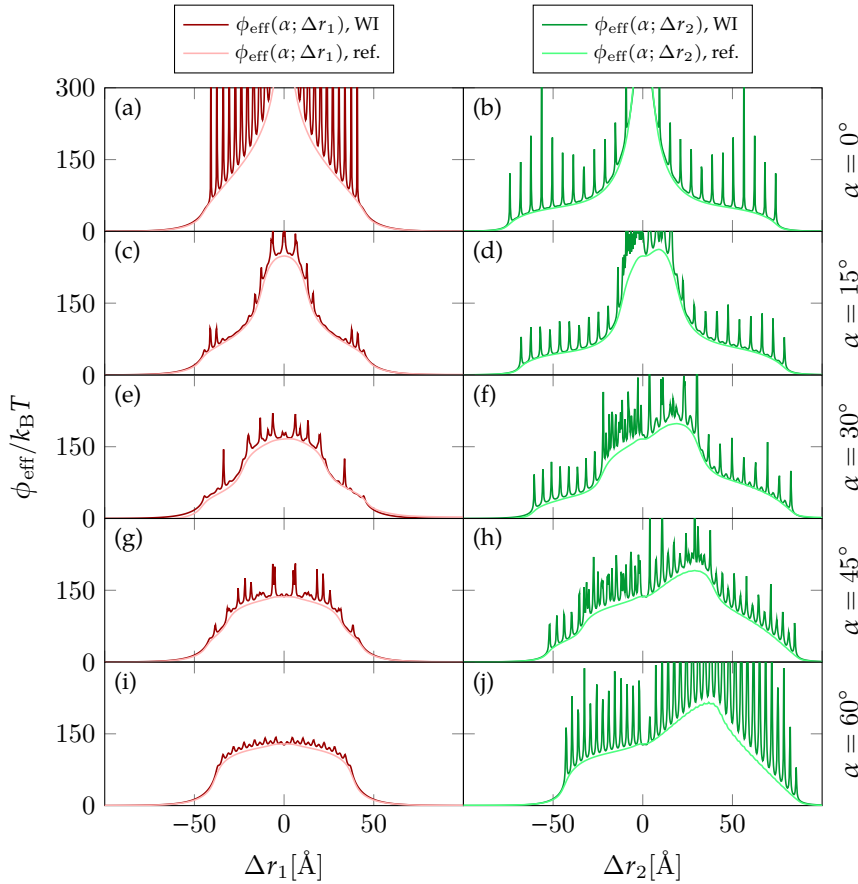

Figure 7.3 shows the effective potentials between two DNA stars, $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$, obtained *via* the reference method and the WI method at angles $\alpha = 0°$ and $\alpha = 60°$, whereas the results for more values of angle α can be seen in Figures C.19-C.23 of Appendix C.3. While the results calculated *via* the two afore-mentioned methods reveal a high level of consistency, qualitatively speaking, closer inspection exposes some markedly quantitative differences: The potentials $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ obtained *via* the reference method display a smooth form, whereas $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ computed *via* the WI method shows granular features due to the periodic divergences in the interaction energy. This discrepancy is the consequence of differences in the underlying models: the model introduced in Section 7.1 assumes point-like charges at the centers of the DNA monomers, while the charges in the model used in the reference method are smeared out with constant charge density along

a cylindrical rod of constant thickness, *i.e.*, the arms of the DNA stars.[237]

A thourough comparison of the energy profiles $\phi_{\mathrm{eff}}(\alpha; \Delta r_i)$, with $i = 1, 2$, of slices through the energy landscape of $\phi_{\mathrm{eff}}(\alpha; \Delta x, \Delta x)$ obtained *via* the reference method and the WI method is discussed later in this section.

Due to the three-fold rotational symmetry of the DNA star's structure we expect the effective potential $\phi_{\mathrm{eff}}(\alpha; \Delta x, \Delta y)$ to exhibit a corresponding rotational symmetry. Indeed, all results display such a three-fold symmetry, see, *e.g.*, Figures 7.3(c) and 7.3(d). For angle $\alpha = 0°$ even a six-fold symmetry can be observed in the potential $\phi_{\mathrm{eff}}(\alpha = 0°; \Delta x, \Delta y)$, see Figures 7.3(a) and 7.3(b), a consequence of the identical orientation of the two stars. Furthermore, the largest peaks in $\phi_{\mathrm{eff}}(\alpha; \Delta x, \Delta y)$ can be observed in the results for $\alpha = 0°$ and $\alpha = 60°$ due to the parallel alignment of one or more arms of the DNA stars and the resulting overlap of the arms.
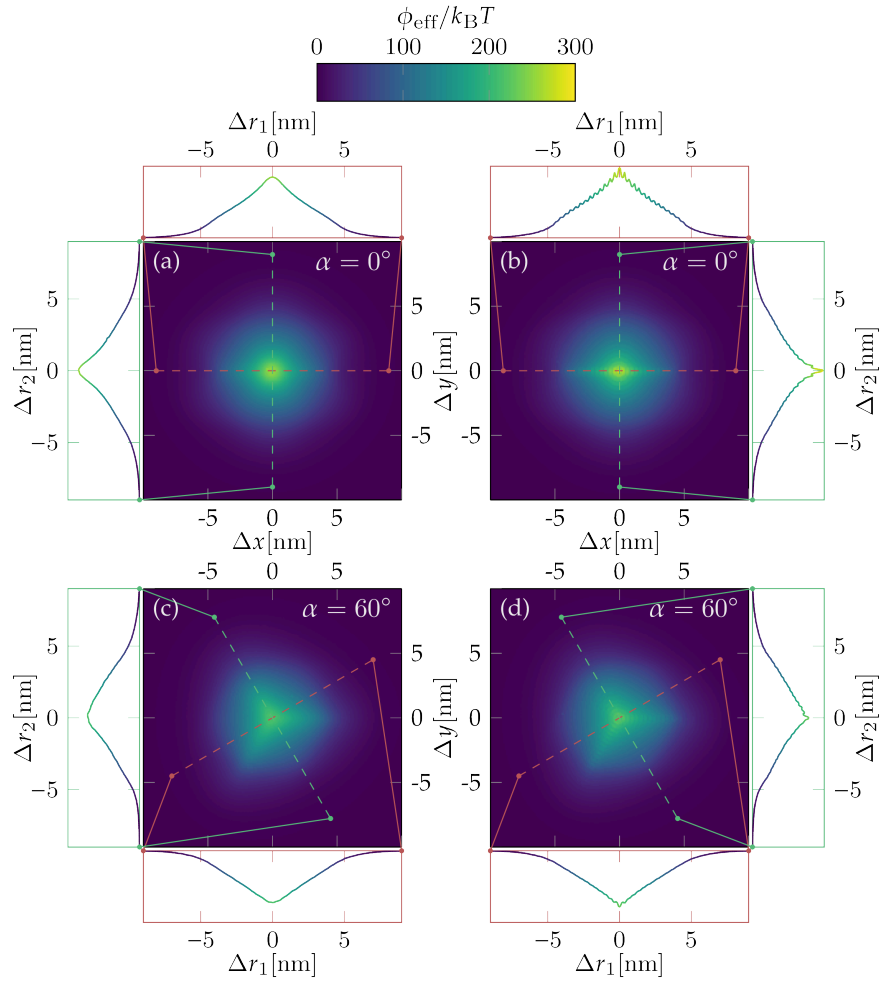
FIGURE 7.4: Energy profiles $\phi_{\mathrm{eff}}(\alpha; \Delta r_1)$ and $\phi_{\mathrm{eff}}(\alpha; \Delta r_2)$ of slices through the energy landscapes of the effective potentials $\phi_{\mathrm{eff}}(\alpha; \Delta x, \Delta y)$ as functions of slice coordinates Δr_1 and Δr_2 and for differen angles α. $\phi_{\mathrm{eff}}(\alpha; \Delta r_i)$ is given in units of $k_{\mathrm{B}}T$, whereas Δr_1 and Δr_2 are given in Å. Results for $\phi_{\mathrm{eff}}(\alpha; \Delta r_i)$ ($i = 1, 2$) are compared between the slices shown in Figure 7.3 and Figures C.19-C.23 (colored red and green). While the orientations of green slices correspond to axes of maximal symmetry in the corresponding $\phi_{\mathrm{eff}}(\alpha; \Delta r_1)$, red slices are positioned orthogonally to green slices, see text. Subplots (a)-(j) show results for rotation angle values $\alpha = 0°$, $15°$, $30°$, $45°$, and $60°$ (as labelled). Data was obtained *via* the Widom insertion method (labelled "WI") and the reference method from Abaurrea Velasco et al., "Effective interactions of DNA-stars" (labelled "ref.").

Finally, we compare the potential energy profiles $\phi_{\text{eff}}(\alpha; \Delta r_i)$ $(i = 1, 2)$ calculated *via* the reference method and *via* the WI method. These energy profiles are obtained for every rotation angle α by cutting a slice through the potential energy landscape $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ along the red and green dashed lines seen in Figures 7.3 and Figures C.19-C.23 in Appendix C.3. In these figures, the orientation of the green dashed line is chosen to correspond to an axis of mirror symmetry of the potential energy landscape $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$, whereby the red dashed lines are oriented orthogonally to the green dashed lines. Figure 7.4 shows the excellent agreement between the results obtained *via* the two afore-mentioned methods: the curves of $\phi_{\text{eff}}(\alpha; \Delta r_i)$ computed *via* the reference method are neatly enveloped by the potentials $\phi_{\text{eff}}(\alpha; \Delta r_i)$ calculated *via* the WI method, with $i = 1, 2$. This envelope clearly shows the periodic, divergent protrusions due to the point-like charges discussed before.

FIGURE 7.5: Effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars as function of center-of-mass shift Δx and Δy and rotation angle α: $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ is given in units of $k_{\text{B}}T$, whereas Δx and Δy are given in Å. Both DNA stars are assumed to be equilibrated while being rotated by different values of angle α against each other (as labelled). Two dashed, orthogonal lines in red and green indicate two slices through the energy landscape of the effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$. The energy profiles of these slices, denoted by $\phi_{\text{eff}}(\alpha; \Delta r_1)$ and $\phi_{\text{eff}}(\alpha; \Delta r_2)$, are shown in the plots framed red and green. The data for this plot was obtained *via* the Widom insertion method with effective coordinate r_{com} in subplots (a) and (c) and effective coordinate r_{cy} in subplots (b) and (d).

Additionally to results presented so far, we started an investigation on the effective potentials $\phi_{\mathrm{eff}}(\alpha; \Delta x, \Delta y)$ of equilibrated stars, where the arms are endowed with flexibility according to our model. There, DNA arms are not completely straight and the interarm angles deviate from the ideal value $\theta_i = 2\pi/3$ ($i = 1, 2, 3$). In order to calculate these effective interactions, we tag one arm per DNA star in order to define the rotation angle α as the angle between the corresponding arm vectors.[238] In contrast to rigid DNA stars, the fluctuations of the conformations of equilibrated DNA stars requires a large number of Widom insertions to average over, see Appendix B.5 for details. The resulting effective potentials are shown in Figure 7.5 and Figures C.24-C.26, with r_{com} and r_{cy} as the effective coordinates of the stars (see caption). While Figure 7.5 suggests that potentials $\phi_{\mathrm{eff}}(\alpha; \Delta x, \Delta y)$ of equilibrated stars exhibit similar features and symmetries as the effective interactions of rigid stars shown in Figure 7.3, the plots for equilibrated stars appear smeared out, in part due to the fluctations of interarm angles $\theta_i = 2\pi/3$ ($i = 1, 2, 3$).[239] Thus, the results for equilibrated stars are inconclusive and seem unsuitable for further use in large scale simulations.

[238] *I.e.*, the vector from the first to the last monomer of the arm.

[239] See Figure 4.10 for the probability distribution $P(\theta_i)$ of the interarm angles θ_i, with $i = 1, \ldots, 3$, as a measure of these fluctuations.

7.3 Dense Systems of Planar DNA Stars

We performed MD simulations of 400 DNA stars with flexible arms in the flat configuration in order to investigate the properties of such two-dimensional systems.[240] Representative simulation snapshots are shown in Figure 7.6.

[240] See Appendix B.5 for details and simulation parameters of the MD simulations.

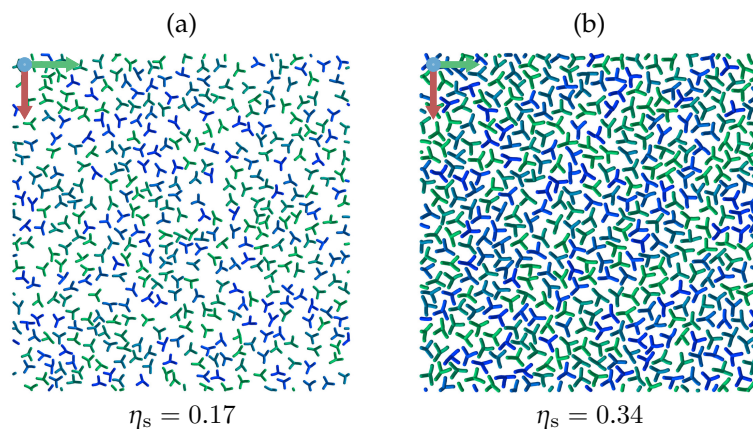(a) $\eta_{\mathrm{s}} = 0.17$ (b) $\eta_{\mathrm{s}} = 0.34$

FIGURE 7.6: Simulation snapshots of two-dimensional systems containing DNA stars in the flat configuration at salt concentration $c = 100\,\mathrm{mM}$ and at densities (a) $\eta_{\mathrm{s}} = 0.17$ and (b) $\eta_{\mathrm{s}} = 0.34$. The DNA stars are shown in eight different colours in order to increase the distinguishability between them. The orientation of the x-, y-, and z-axes of the coordinates are indicated by the red, green, and blue arrows, respectively.

In this section, we present our findings from these investigations. All systems were examined at the same salt concentration $c = 100\,\text{mM}$, while the interface density η_s of Y-DNAs varied between 0.17 and 0.34. This density corresponds to the interface analogue of bulk density η defined in Equation (6.1) and is given *via*

$$\eta_s = 2\pi R_g^2 \rho_d \quad , \text{ with } \quad \rho_d = \frac{N_d}{A_i} \, , \tag{7.1}$$

where R_g is the radius of gyration of the Y-DNA stars, N_d denotes the number of DNA stars in the system, and A_i is the interface area.

FIGURE 7.7: Y-DNA-Y-DNA radial distribution function $g_{yy}(r)$ plotted as function of center-of-mass separation r (given in units of Å). Results are shown for a two-dimensional system of Y-DNAs in the flat configuration and for different values of surface density η_s (as labelled).
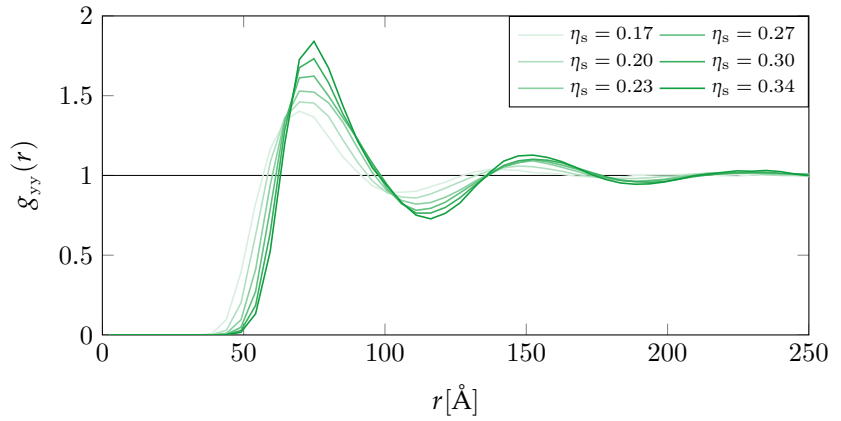
In Figure 7.7 the two-dimensional pair correlation function $g_{yy}(r)$ between the centers-of-mass, r_{com}, of the planar Y-DNA stars is shown for different values of interface density η_s. The smallest distance where $g_{yy}(r)$ is non-zero is located at $r \approx 50\,\text{Å}$, which corresponds to one arm length of the stars. This feature indicates that due to the confinement to two dimensions, the Y-DNAs can not approach each other closer than one arm length. The aforementioned peak, which grows with increasing density η_s, and the subsequent oscillations of the pair correlation function around $g_{yy}(r) = 1$ are typical for a system in the disordered, fluid state.
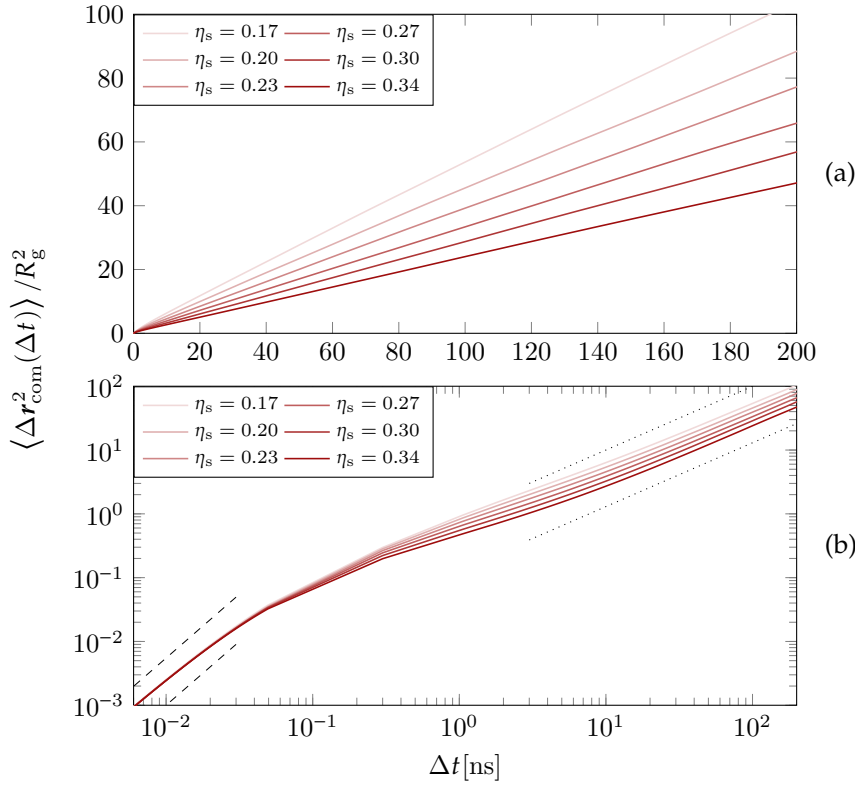
FIGURE 7.8: Mean squared displacement (MSD) as defined in Equation (7.2) of the centers-of-mass, $r_{com}(t)$, of G1 DNA-stars in the flat configuration. Results are given in units of the squared radius of gyration R_g^2 and as as function of time lag Δt, given in ns. The plots are shown in (a) linear scale and (b) log-log scale for different values of surface densitiy η_s, as labelled. Whereas the dashed lines in subplot (b) correspond to MSD $\propto t^2$, the dotted lines in subplot (b) show linear growth (MSD $\propto t$).

To supplement the spatial correlation information expressed *via* $g_{yy}(r)$ we additionally examine the temporal correlation of Y-DNAs in the planar configuration *via* the mean square displacement (MSD) as function of time lag Δt defined as

$$\text{MSD} \equiv \left\langle \frac{1}{t_f - \Delta t - t_i} \int_{t_i}^{t_f - \Delta t} \left(r_{com}(t + \Delta t) - r_{com}(t) \right)^2 \, dt \right\rangle \quad (7.2)$$

$$= \left\langle \Delta r_{com}^2(\Delta t) \right\rangle, \quad (7.3)$$

where $\langle \ldots \rangle$ denotes the ensemble average, *i.e.*, the average over all dendrimers in the MD simulation, while the moving window average in time is calculated *via* the integral on the right hand side in Equation (7.2). Times t_i and t_f correspond to the start and end times of the sampling phase in the simulation. Averaging over the whole ensemble and applying the moving window method ensures a large number of samples for the calculation of MSD and thus statistically meaningful results.

109

These results, which show the MSD of the centers-of-mass of the DNA stars, are shown in Figure 7.8. There, subplot (a) and subplot (b) show the MSD for different values of interface density η_s on a linear scale and on a double logarithmic scale, respectively. Figure 7.8(a) exhibits a linear growth of the MSD with lag time Δt on long time-scales ($\Delta t > 1$ ns). As such the Y-DNAs' MSD demonstrates the behaviour of a random walk according to the fluctuation-dissipation theorem. A clear dependence of the slope of the MSD on density η_s can be seen at higher densities: a smaller displacement is observed. Furthermore, Figure 7.8(b) reveals the ballistic behaviour on short time-scales ($\Delta t < 3.0 \times 10^{-2}$ ns) characterising the DNA stars' trajectory before their first collision after time t.

FIGURE 7.9: Logarithmic derivative, χ, of the mean squared displacements (MSD) as defined in Equation (7.2) of G1 DNA stars in the flat configuration. Results are obtained *via* Equation (7.4) using the MSDs shown in Figure 7.8 and are given as function of lag time Δt, given in ns. See Equation (7.4) for a definition of χ. Results are shown for different values of surface densitiy η_s, as labelled.
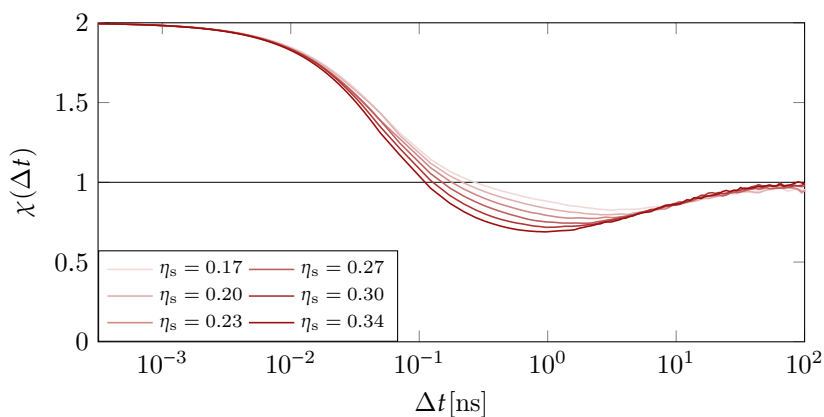
Moreover, the logarithmic derivative of the MSD, χ, is presented in Figure 7.9. This quantity is defined *via*[241]

[241] Transcendental functions, such as the logarithm, are only meaningful if the argument is dimensionless, see C. F. Matta et al. "Can One Take the Logarithm or the Sine of a Dimensioned Quantity or a Unit? Dimensional Analysis Involving Transcendental Functions". In: *J. Chem. Educ.* 88 (2011), pp. 67–70. DOI: 10.1021/ed1000476. To be precise, in Equation (7.4) we have to divide the argument of the logarithm in the denominator by ns and the argument of the logarithm in the numerator by Å². For simplicity's sake, we assume the arguments of the logarithm to be dimensionless without explicitly carrying out the aforementioned divisions.

$$\chi = \frac{\mathrm{d}[\log_{10}\langle \Delta \boldsymbol{r}_{\mathrm{com}}^2(\Delta t)\rangle]}{\mathrm{d}[\log_{10}\Delta t]} \tag{7.4}$$

and provides a deeper insight into the dynamic behaviour of the Y-DNA systems as it determines the dynamic behaviour *via* MSD $\propto t^{\chi}$. In addition to the ballistic regime ($\chi(\Delta t \approx 0\,\mathrm{ns}) = 2$) and the diffusive regime ($\chi(\Delta t > 2\,\mathrm{ns}) = 1$) visible in Figure 7.8, a third region is revealed in Figure 7.9, *i.e.*, the transition between the two aforementioned regimes. Here, the exponent $\chi(0\,\mathrm{ns} < \Delta t < 2\,\mathrm{ns})$ drops below 1 rendering the DNA stars in a sub-diffusive state before converging towards the diffusive state at

$\chi(\Delta t > 2\,\text{ns}) = 1$. While for each value of density η_s the χ-curve starts and ends at the same values, the dynamics of the transition through the sub-diffusive regime differs depending on density η_s: a consequence of higher η_s is a stronger "dip" of $\chi(\Delta t)$ into the sub-diffusive regime.
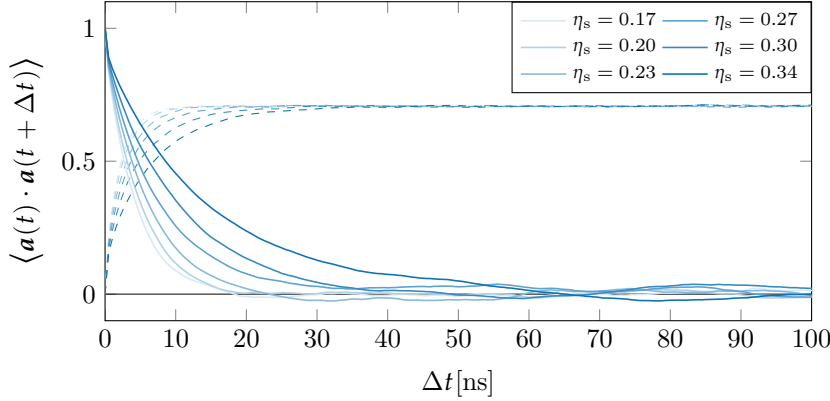

FIGURE 7.10: Auto-correlation function of the Y-DNA arm vector $a(t)$ as function of lag time Δt for the two-dimensional G1 system in the flat configuration. Results are plotted for different values of the surface density η_s (as labelled); lag time Δt is given in ns. While the solid lines show to the mean $\mu(\Delta t) = \langle a(t) \cdot a(t + \Delta t) \rangle$ calculated *via* the moving window method and by averaging over all DNA-stars in the system, the dashed lines represent the corresponding standard deviations, $\sigma(\Delta t)$.

Finally, after considering the translational correlations in the system reflected in the MSD, we now focus on the orientational correlation of the DNA stars in time. More precisely, we tag one arm in each Y-DNA and examine the auto-correlation of the arm vector $a(t)$, *i.e.*, the vector pointing from the center to the final monomer of the arm.

The results, which are shown in Figure 7.10 for different values of surface densitiy η_s, were again calculated using the moving window method and *via* the ensemble average. As expected, the self-correlation of $a(t)$ decays with lag time Δt. Here, density η_s has also an impact on the dynamics of the correlation: increasing the density η_s yields a slower decay of the self-correlation of $a(t)$. For instance, the auto-correlation of $a(t)$ becomes zero at lag times $\Delta t \approx 20\,\text{ns}$ and $\Delta t \approx 70\,\text{ns}$ for interface densities $\eta_s = 0.17$ and $\eta_s = 0.34$, respectively. At these lag times DNA stars at both densities have travelled in the system over a RMS distance of only $4R_g$, see Figure 7.8, indicating a strong decay in the orientational auto-correlation in time as opposed to the translational auto-correlation.

7.4 Dense Systems of Tripod DNA Stars

In addition, we performed MD simulations of 400 DNA stars with flexible arms in the tripod configuration in parallel to the simulations of stars in the flat configuration.[242] Representative simulation snapshots are shown in Figure 7.11.

FIGURE 7.11: Simulation snapshots of two-dimensional systems containing DNA stars in the tripod configuration at salt concentration $c = 100\,\text{mM}$ and at densities $\eta_s = 0.17$ and $\eta_s = 0.54$, as labelled. The DNA stars are shown in eight different colours in order to increase the distinguishability between them. Whereas subplots (c) and (d) show the projections in the xy-plane, the projections in the yz-plane are shown in subplots (a) and (b). The orientation of the x-, y-, and z-axes of the coordinates are indicated by the red, green, and blue arrows, respectively.
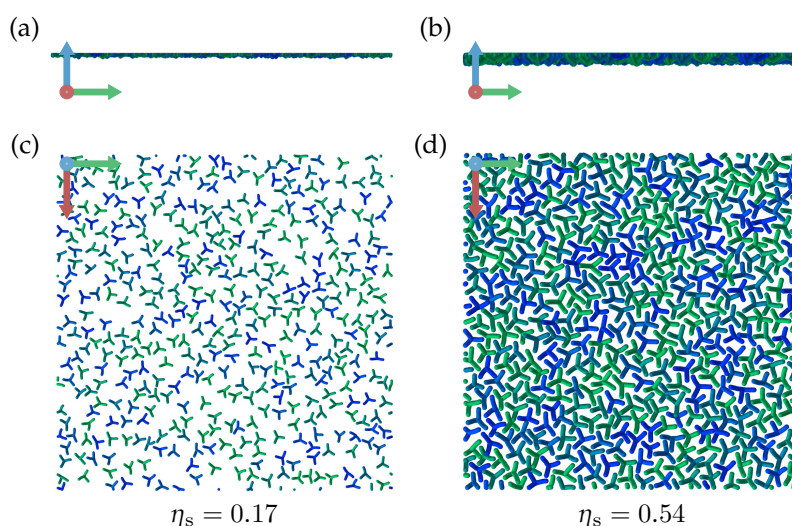
Again, all systems were examined at the salt concentration $c = 100\,\text{mM}$, while the interface density η_s of Y-DNAs varied between 0.17 and 0.58. Due to the ability of these DNA stars to immerse their tip into the fluid, higher interface densities η_s could be reached. By examining the data obtained from these simulations it was possible to carry out an analogous analysis to Section 7.3 whose results are presented in this section.

FIGURE 7.12: Y-DNA Y-DNA radial distribution function $g_{yy}(r)$ plotted as function of center-of-mass separation r (given in units of Å). Results are shown for a two-dimensional system of Y-DNAs in the tripod configuration and for different values of surface density η_s (as labelled).
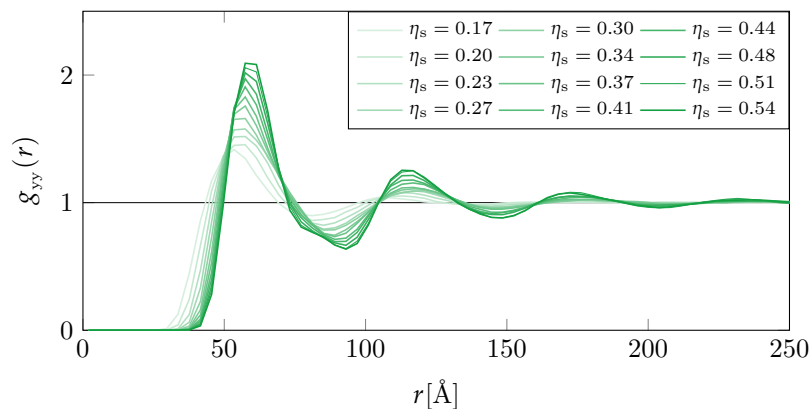
The two-dimensional pair correlation function $g_{yy}(r)$ can be seen in Figure 7.12. Similarly to the Y-DNAs in the flat configuration, a peak which increases with density η_s can be observed, characterising the fluid, disordered state. On the other hand, the smallest r distance where $g_{yy}(r)$ is non-zero is located at $r < 50\,\text{Å}$, which signifies that Y-DNAs in the tripod configuration approach each other more closely than Y-DNAs in the flat configuration.
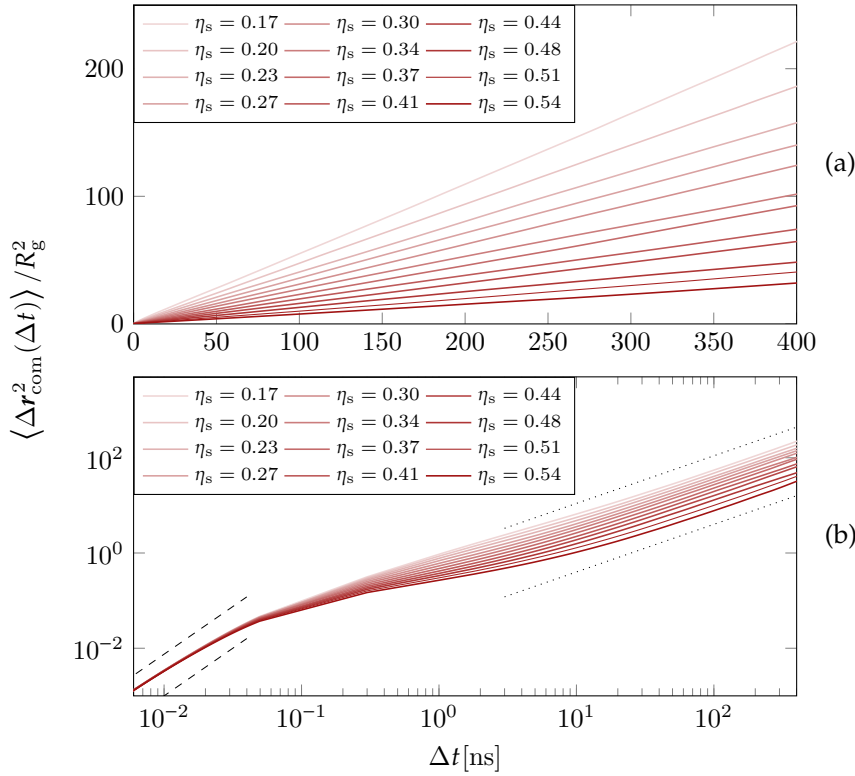

FIGURE 7.13: Mean squared displacement (MSD) as defined in Equation (7.2) of the centers-of-mass, $r_{com}(t)$, of G1 DNA-stars in the tripod configuration. Results are given in units of the squared radius of gyration R_g^2 and as as function of time lag Δt, given in ns. The plots are shown in (a) linear scale and (b) log-log scale for different values of surface densitiy η_s, as labelled. Whereas the dashed lines in subplot (b) correspond to MSD $\propto t^2$, the dotted lines in subplot (b) show linear growth (MSD $\propto t$).

In analogy to Section 7.3, we investigated the translational and orientational correlations in time by calculating the MSD of the centers-of-mass, $r_{com}(t)$, and the auto-correlation of arm vectors $a(t)$ as functions of time lag Δt, see Figures 7.13 and 7.15, respectively. The logarithmic derivative of the MSD, $\chi(\Delta t)$, is shown in Figure 7.14. For each of these results, the same observations as in Section 7.3 hold true. Here, the trend of the η_s dependent behaviour of these results is continued for the higher densities available to Y-DNAs in the tripod configuration. A direct comparison of these quantities obtained for Y-DNAs in the planar configuration and the tripod configuration will be given in the next section.

FIGURE 7.14: Logarithmic derivative, χ, of the mean squared displacements (MSD) as defined in Equation (7.2) of G1 DNA stars in the tripod configuration. Results are obtained *via* Equation (7.4) using the MSDs shown in Figure 7.13 and are given as function of lag time Δt, given in ns. See Equation (7.4) for a definition of χ. Results are shown for different values of surface densitiy η_s, as labelled.
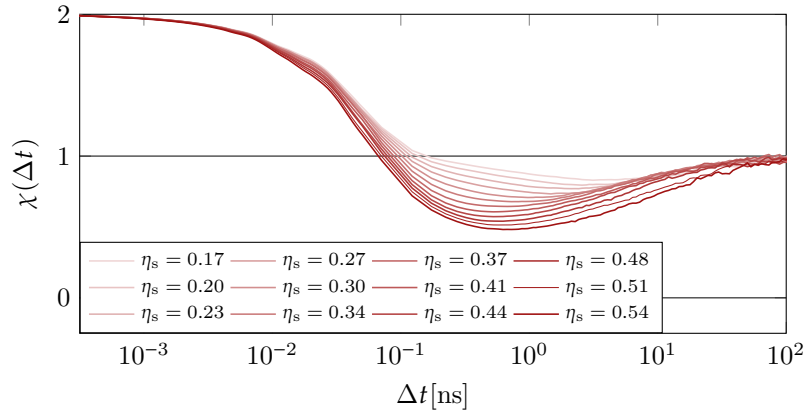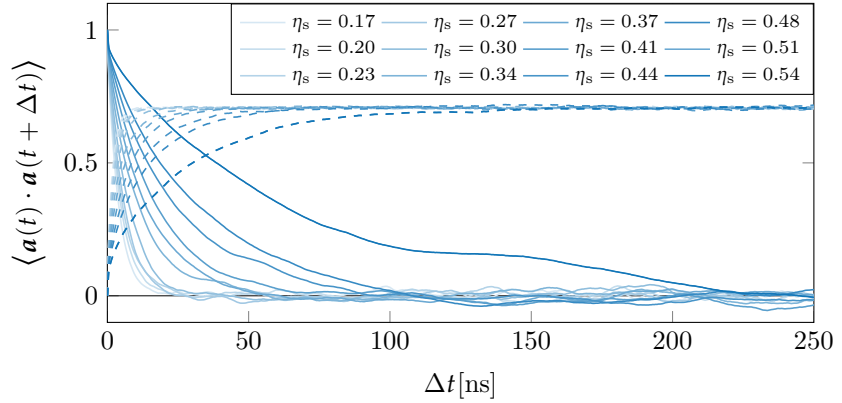

FIGURE 7.15: Auto-correlation function of the Y-DNA arm vector $a(t)$ as function of lag time Δt for the two-dimensional G1 system in the tripod configuration. Results are plotted for different values of the surface density η_s (as labelled); lag time Δt is given in ns. While the solid lines show to the mean $\mu(\Delta t)$ $= \langle a(t) \cdot a(t + \Delta t) \rangle$ calculated *via* the moving window method and by averaging over all DNA-stars in the system, the dashed lines represent the corresponding standard deviations, $\sigma(\Delta t)$.

FIGURE 7.16: Sketch of a DNA star in the tripod configuration showing distance Δz and angle ξ, see text.

In contrast to planar DNA stars, the Y-DNAs in the tripod configuration possess additional degrees of freedom which can be investigated. In Figure 7.17 the probability distributions $P(\xi)$ and $P(\Delta z)$ are shown for different values of density η_s. Whereas Δz corresponds to the immersion depth of the tripod, *i.e.*, the orthonormal-distance of the Y-DNAs junction to the interface, angle ξ denotes the angle between the DNA arms and the interface, see Figure 7.16.
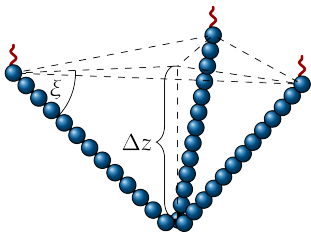
Due to the geometrical properties of tripod Y-DNAs there is a clear correlation between angle ξ and distance Δz as becomes evident from Figures 7.17(a) and 7.17(b). Furthermore, when tracking the trend of these quantities in dependence of the density η_s a striking double peak feature can be observed: for low densities the tripods assume an almost planar form with large peaks in the distributions $P(\xi)$ and $P(\Delta z)$ located at values of angle ξ and depth Δz close to zero. High densities, on the other hand, result in another peak in the distributions $P(\xi)$ and $P(\Delta z)$ at $\xi \approx 0.2\pi$ and $\Delta z \approx 25\,\text{Å}$. This second peak is a consequence of geometric restrictions: The steric interaction between DNA arms does not allow the angle ξ to assume

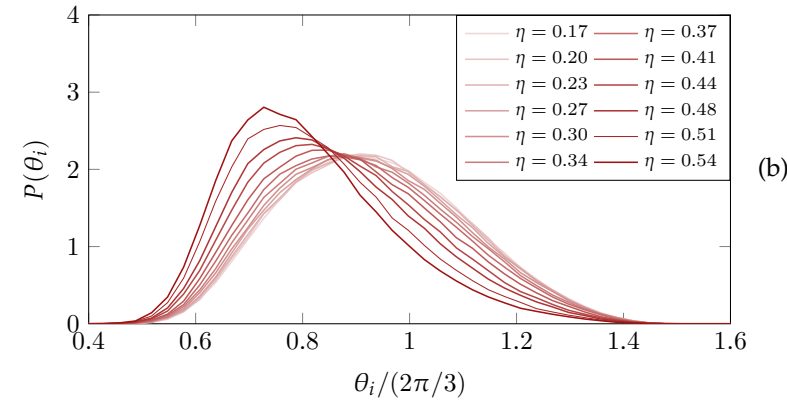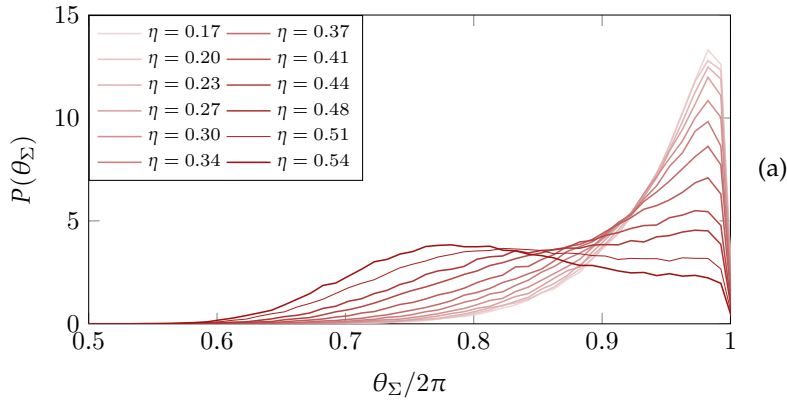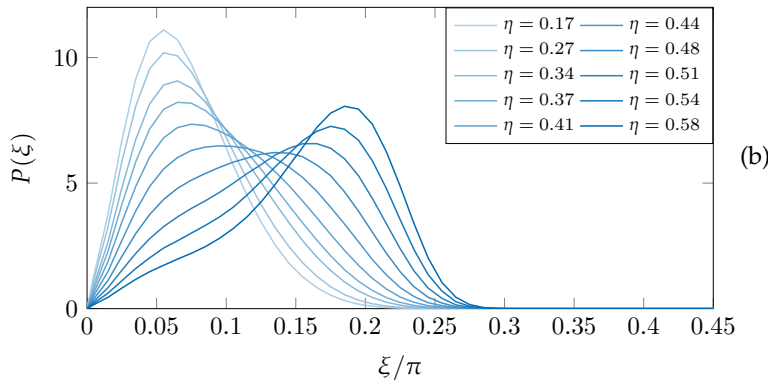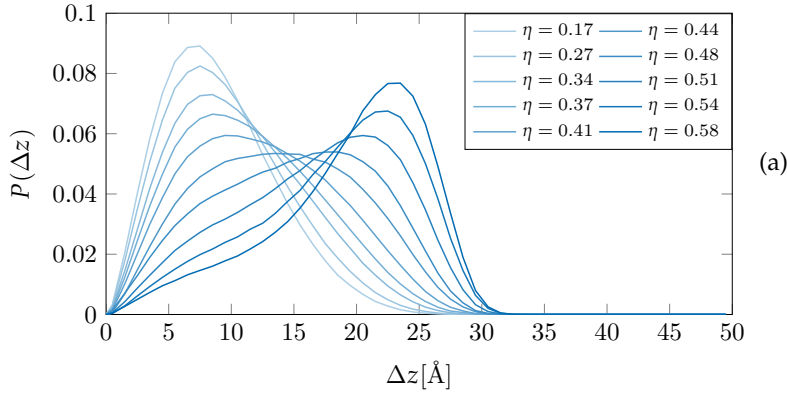values much larger than 0.25π which concomitantly limits the immersion depth Δz.


FIGURE 7.17: (a) Probability distribution $P(\Delta z)$ as function of the immersion depth Δz for the two-dimensional G1-tripod system. Results are plotted for different values of the surface density η_s; Δz is given in Å. The probability distributions are normalised *via* $\int_0^\infty P(\Delta z)\,d(\Delta z) = 1$. (b) Probability distribution $P(\xi)$ as function of the arm-interface angle ξ for the two-dimensional G1-tripod system. Results are plotted for different values of the surface density η_s and ξ is given in units of π. The probability distributions are normalised *via* $\int_0^{\pi/2} P(\xi)\,d\xi = 1$.

FIGURE 7.18: (a) Probability distribution $P(\theta_\Sigma)$ as function of the interarm angle sum θ_Σ for the two-dimensional G1-tripod system. Results are plotted for different values of the surface density η_s, as labelled; θ_Σ is given in units of 2π. The probability distributions are normalised *via* $\int_0^{2\pi} P(\theta)_\Sigma\,d\theta_\Sigma = 1$. (b) Probability distribution $P(\theta_i)$ as function of the interarm angle sum θ_i ($i = 1, 2, 3$) for the two-dimensional G1-tripod system. Probability distributions $P(\theta_i)$ show the cumulative results for all three interarm angles θ_i, $i = 1, 2, 3$. Results are plotted for different values of the surface density η_s, as labelled; θ_i is given in units of $2\pi/3$. The probability distributions are normalised *via* $\int_0^\pi P(\theta_i)\,d\theta_i = 1$.

The deviation of tripod Y-DNAs from the planar configuration can also be analysed *via* the interarm angles θ_i, with $i = 1, 2, 3$, and their sum $\theta_\Sigma = \sum_{i=1}^{3} \theta_i$. The corresponding distributions $P(\theta_\Sigma)$ and $P(\theta_i)$, with $i = 1, 2, 3$, are shown in Figure 7.18. There, a similar trend as in Figure 7.17 manifests with higher values of density η_s yielding a larger deviation from the planar conformation, *i.e.*, $\theta_\Sigma = 2\pi$ and $\theta_i = 2\pi/3$, with $i = 1, 2, 3$.

FIGURE 7.19: Probability distribution $P(\eta_{\mathrm{loc}})$ as function of the local surface density η_{loc} for the two-dimensional G1-tripod system. Results are plotted for different values of the global surface density η_s, as labelled. The probability distributions are normalised *via* $\int_0^\infty P(\eta_{\mathrm{loc}}) \, d\eta_{\mathrm{loc}} = 1$.
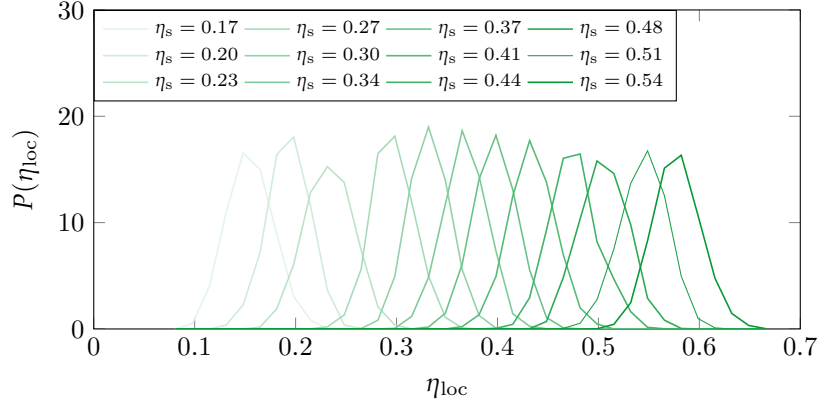
While the surface density η_s denotes the global density for the whole system, we define the local density η_{loc} associated with each Y-DNA as the density within a circular area of radius $r_{\mathrm{loc}} = 200 \, \mathrm{\mathring{A}}$ centered on the center-of-mass of this Y-DNA. In this way, we can account for the fluctuations in density throughout the system. The distributions $P(\eta_{\mathrm{loc}})$ as functions of local density η_{loc} are depicted in Figure 7.19 for different values of global density η_s. While the centers of distributions $P(\eta_{\mathrm{loc}})$ are located at the corresponding global distribution η_s, the spread of η_{loc} is not noticeably different for different values of η_s.

FIGURE 7.20: Probability distribution $P(\Delta z)$ as a function of the immersion depth Δz for the two-dimensional G1-tripod system. Results are plotted for different values of the local surface density η_{loc}, as labelled; Δz is given in $\mathrm{\mathring{A}}$. The probability distributions are normalised *via* $\int_0^\infty P(\Delta z) \, d(\Delta z) = 1$.
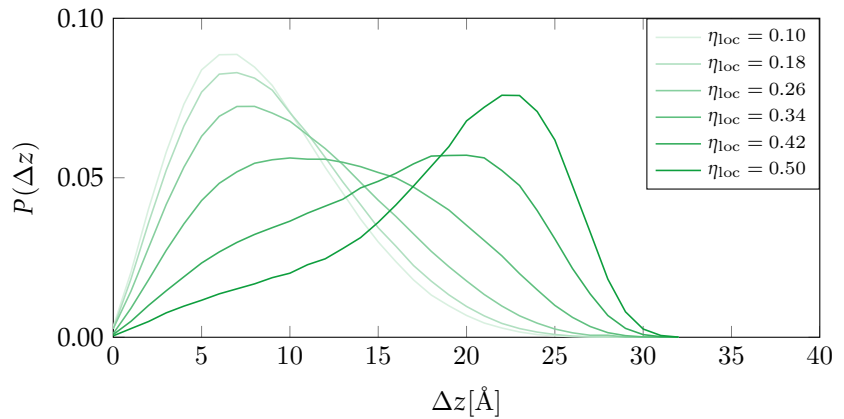
Equipped with this concept of local density η_{loc}, we can now re-analyse the immersion depth Δz with respect to η_{loc}. For this purpose, we calculate Δz and η_{loc} for each Y-DNA individually and repeat this process for systems at different global densities η_{s}. Then, we merge the corresponding results of Δz for each value of η_{loc}. Figure 7.20 shows distributions $P(\Delta z)$ for selected values of local density η_{loc}. Examination of the results shown in Figure 7.20 yields similar conclusions as were previously found for Figure 7.17. Figure 7.21 allows for more insight into the dependence of the immersion depth Δz of a tripod Y-DNA on its surrounding local density η_{loc}: function $\Delta z(\eta_{\mathrm{loc}})$ undergoes a sigmoidal transition from a lower value $\Delta z \approx 2.3 l_{\mathrm{b}}$ (corresponding to the leftmost peak in Figure 7.20) to a higher value of $\Delta z \approx 5.9 l_{\mathrm{b}}$ (corresponding to the rightmost peak in Figure 7.20).
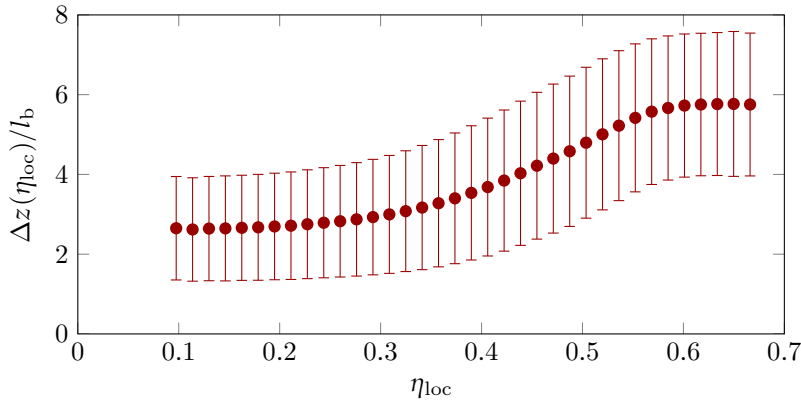

FIGURE 7.21: Immersion depth $\Delta z(\eta_{\mathrm{loc}})$ as a function of the local surface density η_{loc} for the two-dimensional G1-tripod system. $\Delta z(\eta_{\mathrm{loc}})$ is given in Å. The error bars indicate the standard deviations in the distributions of $\Delta z(\eta_{\mathrm{loc}})$.

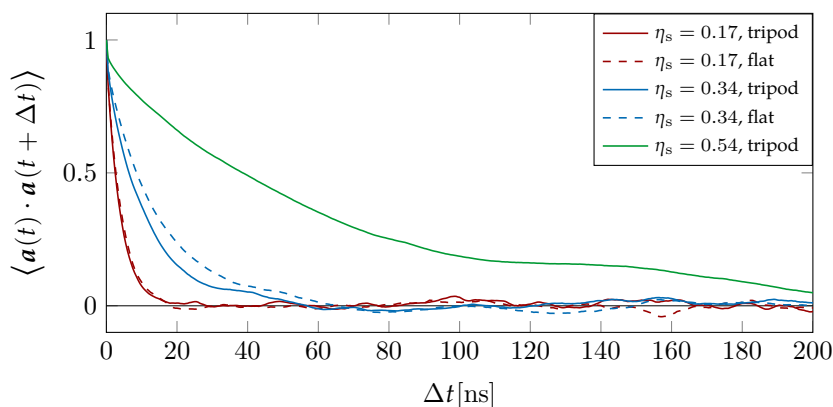## 7.5 Planar DNA Stars vs. Tripod DNA Stars

Finally, we present a comparison for the results presented in the previous sections between planar and tripod Y-DNAs. In Figure 7.22 the two-dimensional pair correlation functions $g_{\mathrm{yy}}(r)$ for both types of systems are shown. For each value of density η_{s} the global maximum is of same height for planar and tripod Y-DNAs while the peak of the latter is slightly shifted to the smaller distances r. This indicates the ability of DNA stars in the tripod configuration to evade each other due to their quasi-two-dimensional nature, a property that the strictly planar Y-DNAs do not have.

FIGURE 7.22: Y-DNA Y-DNA radial distribution function $g_{yy}(r)$ plotted as function of center-of-mass separation r (given in units of Å). Results show the comparison between Y-DNAs in the flat configuration (dashed lines) and Y-DNAs in the tripod configuration (solid lines) for different values of surface density η_s (as labelled).
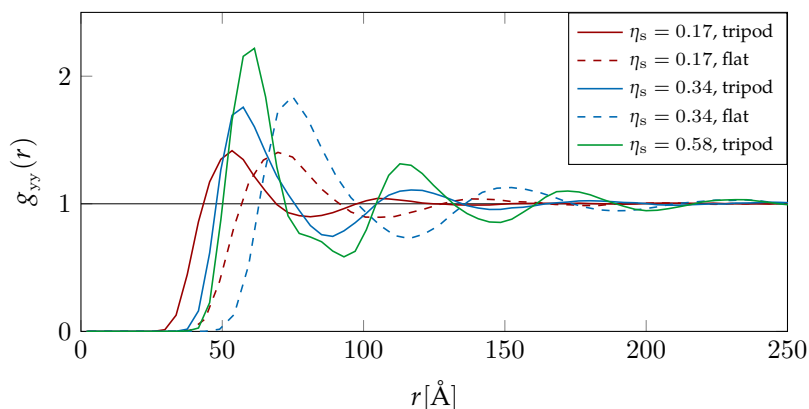
The results of analysing the rotational and the translational correlations in time are presented in the following Figures 7.23, 7.24, and 7.26. In each of these figures the results are compared between planar DNA stars (dashed lines) and DNA stars in the tripod configuration (solids lines). Examining these results shows that the differences between the two kinds of systems are relatively small: whereas the self-correlation of arm vector $a(t)$ shown in Figure 7.23 decays slightly slower for flat Y-DNAs, the MSDs presented in Figure 7.24 indicate that tripod Y-DNAs diffuse somewhat slower.

FIGURE 7.23: Auto-correlation function of the Y-DNA arm vector $a(t)$ as function of lag time Δt given in ns. Results are compared between Y-DNAs in the flat configuration (dashed lines) and Y-DNAs in the tripod configuration (solid lines) for different values of surface density η_s (as labelled). The mean $\mu(\Delta t) = \langle a(t) \cdot a(t + \Delta t) \rangle$ was calculated *via* the moving window method and by averaging over all DNA-stars in the system,

Furthermore, the MSD-dynamics determined by χ (as shown in Figure 7.26) differs only marginally. Here, the transition from the ballistic to the diffusive regime is slightly shallower, *i.e.* less sub-diffusive, for planar Y-DNAs.

118

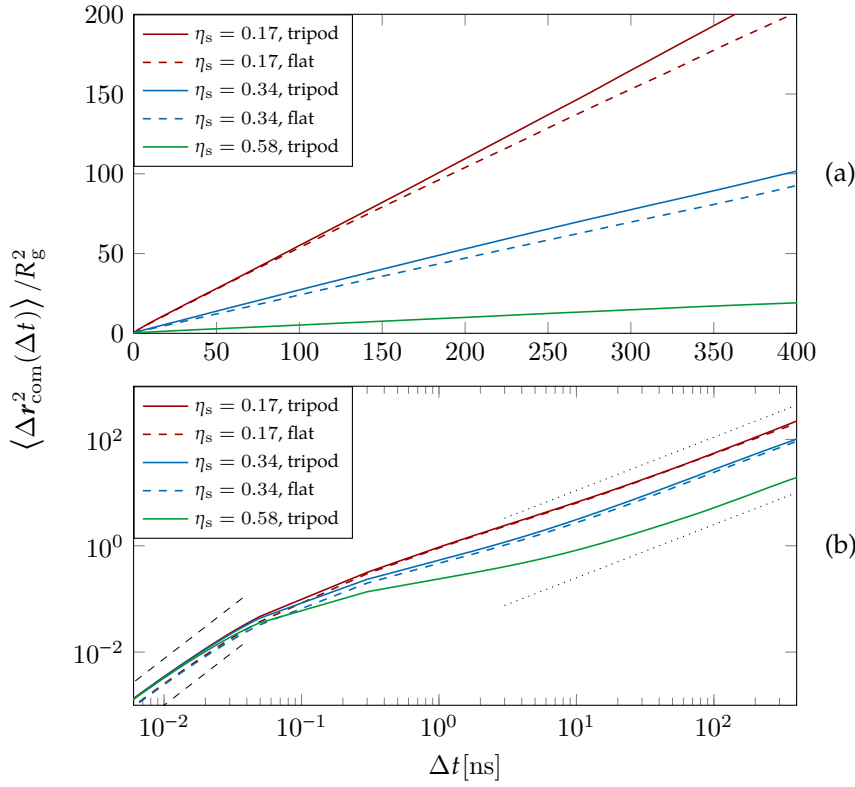**FIGURE 7.24:** Mean squared displacement (MSD) of the centers-of-mass, $r_{com}(t)$, given in units of the squared radius of gyration R_g^2, as function of time lag Δt, given in ns. Results are compared between Y-DNAs in the flat configuration (dashed lines) and Y-DNAs in the tripod configuration (solid lines) for different values of surface density η_s (as labelled). The plots are shown in (a) linear scale and (b) log-log scale. Whereas the dashed lines in subplot (b) correspond to MSD $\propto t^2$, the dotted lines in subplot (b) show linear growth (MSD $\propto t$).
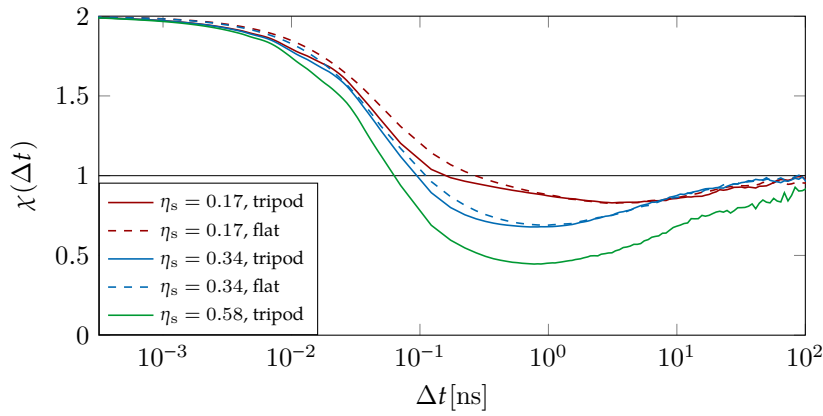

FIGURE 7.25: Logarithmic derivative, χ, of the mean squared displacements (MSD) of the centers-of-mass, $r_{com}(t)$, given in units of the squared radius of gyration R_g^2, as function of lag time Δt, given in ns. See Equation (7.4) for a definition of χ. Results are compared between Y-DNAs in the flat configuration (dashed lines) and Y-DNAs in the tripod configuration (solid lines) for different values of surface density η_s (as labelled).
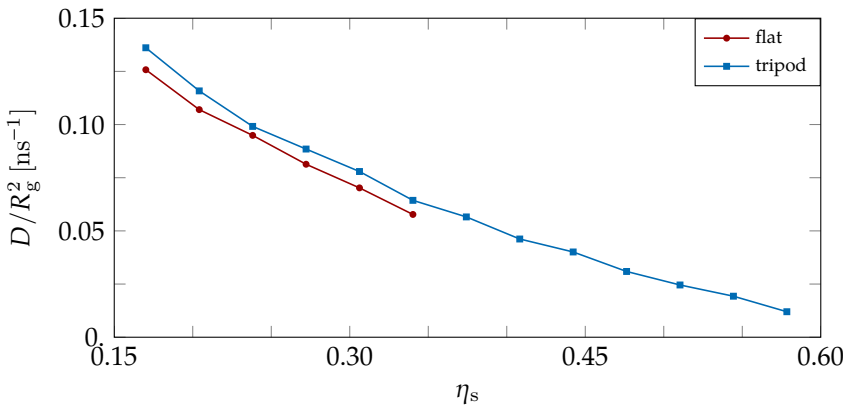

FIGURE 7.26: Diffusion constant of DNA stars, D, rescaled by R_g^2 and given in ns^{-1} as function of interface density η_s. Results are compared between Y-DNAs in the flat configuration (labelled "flat") and Y-DNAs in the tripod configuration (labelled "tripod").

119

Finally, we present a comparison of the diffusion coefficients D as extracted from the MSDs in the diffusive regime presented in Figure 7.24:[243] it can be observed that DNA stars in the tripod configuration are slightly more diffusive due to their aforementioned ability to evade each other.

In general, the results are remarkably similar between the two types of investigated systems in spite of their different range of conformations.

[243] The diffusion constant D can be calculated from the MSD *via* Equation (2.87).

8 Conclusion

. . . and it is probable that there is some secret here which remains to be discovered.

—Charles Sanders Peirce[244]

This final chapter, which concludes the main part of this thesis, serves two purposes:

(a) providing a conclusion of the results presented in this thesis and

(b) giving an outlook on future research which could continue this project.

We presented a coarse-grained bead-spring model of DNA and DL-DNAs in Chapter 3. With this theoretical model at hand we then investigated the structural properties of DL-DNA at the single particle level for different generation numbers GN in Chapter 4. Additionally, we probed the salt-responsiveness of these complex macromolecules and combined our theoretical results with those from experiment, achieving excellent agreement.[245]

In Chapter 5 a comprehensive study of effective potentials calculated at different salt concentrations c using the Widom insertion (WI) method was given for dendrimers of size $G1$ to $G3$. Furthermore, we provided a comparison of these data and results obtained by two complementary methods: umbrella sampling and neural network potentials. While there is good agreement between effective potentials obtained *via* the WI method and umbrella sampling, further research on neural network potentials for complex macromolecules such as DL-DNA is needed. Here, internal degrees of freedom must be considered properly in order to correctly reproduce the pair interactions.[246]

244 C. S. Pierce. *Collected Papers of Charles Sanders Peirce, Volumes V and VI: Pragmatism and Pragmaticism and Scientific Metaphysics*. Ed. by C. Hartshorne and P. Weiss. Belknap Press, 1935. URL: https://worldcat.org/isbn/9780674138025, p. 297.

245 See Jochum et al., "Structure and stimuli-responsiveness of all-DNA dendrimers: theory and experiment".

246 See J. Wang et al. "Machine Learning of Coarse-Grained Molecular Dynamics Force Fields". In: *ACS Cent. Sci.* 5 (2019), pp. 755–767. DOI: 10.1021/acscentsci.8b00913.

We extended our study of DL-DNA systems to bulk systems at high densities in Chapter 6. There, we computed structural quantities, *e.g.*, pair correlation function $g(r)$, of the DL-DNA systems *via* three different methods: MD simulations of our coarse-grained model, MD simulations of effective particles using the effective potentials of Chapter 5, and the HNC method from integral equation theory. While these theoretical results are consistent within themselves, a subsequent comparison to experimental results reveals shortcomings in our theoretical model. These shortcomings could be mitigated using several approaches: adapting the coarse-grained model to accomodate ion condensation effects[247] and replacing the spherically symmetric effective potential with an anisotropic one.[248]

Finally, we presented a purely theoretical study of two-dimensional DNA star systems at interfaces in Chapter 7. A comprehensive comparison between DNA stars in the flat configuration and in the tripod configuration demonstrates how slightly tweaking the design of our DNA stars gives rise to new behaviour. While for this part of the research project experimental results for verification are still lacking, similar systems have already been realised.[249]

Considering the flexibility and adaptability of these DL-DNA molecules, further research exploring the bulk phase behaviour of states of matter such as gels[250] and liquid crystals[251] is especially intriguing. Furthermore, from a fundamental research perspective these ultrasoft DL-DNA particles with open internal architectures and tunable interactions are ideal candidates for the experimental realisation of the recently proposed cluster-crystal structure.[252] In order to explore these possibilities, the molecules could be enhanced *via* PNIPAM polymers so that their interactions can be tuned *via* temperature in addition to salt concentration.[253]

[247]See Kornyshev et al., "Structure and interactions of biological helices", p. 4 and Kegler et al., "Polyelectrolyte-Compression Forces between Spherical DNA Brushes", pp. 955-956.

[248]See Poier et al., "An Anisotropic Effective Model for the Simulation of Semiflexible Ring Polymers" and Heinemann et al., "Angle-resolved effective potentials for disk-shaped molecules".

[249]See Avakyan et al., "Long-Range Ordering of Blunt-Ended DNA Tiles on Supported Lipid Bilayers".

[250] See S. Biffi et al. "Phase behavior and critical activated dynamics of limited-valence DNA nanostars". In: 110 (2013), pp. 15633–15637. DOI: 10.1073/pnas.1304632110.

[251] See M. Siavashpouri et al. "Molecular engineering of chiral colloidal liquid crystals using DNA origami". In: *Nat. Mater.* 16 (2017), pp. 849–856. DOI: 10.1038/nmat4909.

[252]See D. A. Lenz et al. "Monomer-Resolved Simulations of Cluster-Forming Dendrimers". In: *J. Phys. Chem. B* 115 (2011), pp. 7218–7226. DOI: 10.1021/jp109132m and D. A. Lenz et al. "Thermodynamic Stability and Structural Properties of Cluster Crystals Formed by Amphiphilic Dendrimers". In: *J. Chem. Phys.* 144 (2016), p. 204901. DOI: 10.1063/1.4950953.

[253]See G. Paradossi and E. Chiessi. "Solution behaviour of poly(N-isopropylacrylamide) stereoisomers in water: a molecular dynamics simulation study". In: *Phys. Chem. Chem. Phys.* 19 (2017), pp. 11892–11903. DOI: 10.1039/C7CP00808B.

Appendix A

Simulation Software Packages

In order to perform the simulations of Chapters 4 to 7, four different software packages were used: ESPResSo,[254] LAMMPS,[255] oxDNA,[256] and n2p2.[257] A short description of each simulation package is presented in the following. Note that while n2p2 is not a simulation package *per se*, its neural network potentials can be implemented in LAMMPS in order to perform simulations.

A.1 ESPResSo

ESPResSo, short for **E**xtensible **S**imulation **P**ackage for **Res**earch on **S**oft Matter, is a highly versatile software package for performing and analysing MD simulations of many-particle systems using coarse-grained models (*e.g.*, bead-spring models) typically used in soft matter research. First published in 2006,[258,259] using a Tcl-script[260] driven user interface, today the ESPResSo package is used all over the world in research groups with diverse scientific backgrounds, *e.g.*, physics, chemistry, and molecular biology. The recent release of ESPResSo 4.0[261] included a scripting interface *via* the powerful Python language[262] with its great flexibility.

While ESPResSo is mainly developed at the Institute for Computational Physics of the University of Stuttgart, it has contributors from all over the world. ESPResSo is published under the GNU General Public License (GPL3) as free and open-source software.

[254] See http://espressomd.org, not to be confused with Quantum ESPRESSO (https://www.quantum-espresso.org/).

[255] See https://lammps.sandia.gov.

[256] See https://dna.physics.ox.ac.uk.

[257] See https://github.com/CompPhysVienna/n2p2 for the software package and https://compphysvienna.github.io/n2p2/index.html for the documentation.

[258] See H. J. Limbach et al. "ESPResSo - an Extensible Simulation Package for Research on Soft Matter Systems". In: *Comput. Phys. Commun.* 174 (2006), pp. 704–727. DOI: 10.1016/j.cpc.2005.10.005.

[259] See A. Arnold et al. "ESPResSo 3.1 — Molecular Dynamics Software for Coarse-Grained Models". In: vol. 89. Lecture Notes in Computational Science and Engineering. 2013, pp. 1–23. DOI: 10.1007/978-3-642-32979-1_1.

[260] See J. K. Ousterhout. *Tcl and the Tk Toolkit*. Flatbrain Com, 1996. URL: https://dl.acm.org/doi/book/10.5555/524313.

[261] See F. Weik et al. "ESPResSo 4.0 - an extensible software package for simulating soft matter systems". In: *Eur. Phys. J. Special Topics* 227.14 (2019), pp. 1789–1816. DOI: 10.1140/epjst/e2019-800186-9.

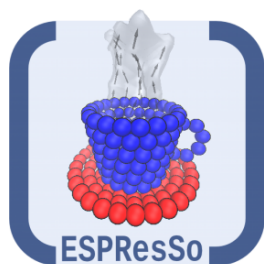[262] See G. van Rossum. *Python tutorial*. Tech. rep. CS-R9526. Amsterdam: Centrum voor Wiskunde en Informatica (CWI), 1995. URL: https://ir.cwi.nl/pub/5007/05007D.pdf.

FIGURE A.1: Logo of the **E**xtensible **S**oftware **P**ackage for **Res**earch on **So**ft Matter (ESPResSo).

[263] See S. Chen and G. D. Doolen. "Lattice Boltzmann Method for Fluid Flows". In: *Annu. Rev. Fluid Mech.* 30 (1998), pp. 329–364. DOI: 10.1146/annurev.fluid.30.1.329.

[264] See Allen and Tildesley, *Computer Simulations of Liquids*, p. 101 and W. C. Swope et al. "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters". In: *J. Chem. Phys.* 76 (1982), pp. 6370–649. DOI: 10.1063/1.442716.

[265] See Allen and Tildesley, *Computer Simulations of Liquids*, pp. 131-132.

[266] See Allen and Tildesley, *Computer Simulations of Liquids*, pp. 35-37.

[267] See Allen and Tildesley, *Computer Simulations of Liquids*, pp. 195-200.

[268] Based on P. P. Ewald's seminal work, see Ewald, "Die Berechnung Optischer und Elektrostatischer Gitterpotentiale".

[269] See M. Deserno and C. Holm. "How to mesh up Ewald sums. II. An accurate error estimate for the particle–particle–particle-mesh algorithm". In: *J. Chem. Phys.* 109 (1998), pp. 7694–7701. DOI: 10.1063/1.477415 and M. Deserno and C. Holm. "How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines". In: *J. Chem. Phys.* 109 (1998), pp. 7678–7693. DOI: 10.1063/1.477414.

[270] See Section 3.2.

[271] See https://vsc.ac.at/.

Typical soft matter systems that can be simulated using the ESPResSo software package include polymers, liquid crystals, colloids, polyelectrolytes, ferrofluids and biological systems, *e.g.*, DNA and lipid membranes. A range of numerical and simulation methods are offered with the package to the users, *e.g.*, a lattice Boltzmann solver in order to solve the Boltzmann equation and account for hydrodynamic interactions.[263]

Simulations performed *via* the ESPResSo software in this work used the velocity Verlet integration scheme,[264] the Langevin thermostat for canonical NVT ensembles,[265] and periodic boundary conditions (PBC).[266] Internal book keeping of particles is managed *via* neighbour lists, more specifically *via* linked cell lists.[267]

For simulations of low particle numbers $N < 10^5$ carrying charges electrostatic interactions are handled using the Ewald summation method,[268] implemented in ESPResSo *via* the so-called particle-particle-particle-mesh algorithm (P^3M).[269]

In the presence of high salt concentrations ($c \geq 100\,\text{mM}$) the charged particle number diverges ($N \gg 10^6$) rendering an explicit treatment of these particles computationally unfeasible. Therefore, in this work, the electrostatic interactions of systems with salt concentrations $c \geq 100\,\text{mM}$ were treated by using an appropriate Debye-Hückel potential.[270] The computations were performed in a highly parallel manner on the Vienna Scientific Cluster (VSC).[271] On this cluster, a spatial-decomposition technique was used in order to partition the simulation volume into smalller cubic subdomains, each of which is assigned to a different processor.

The ESPResSo simulations in this work were performed by the author.

A.2 LAMMPS

Besides ESPResSo this work also presents results obtained *via* the LAMMPS

software package.[272]

According to the documentation:[273]

> *LAMMPS is a classical molecular dynamics code with a focus on materials modeling. It's an acronym for Large-scale Atomic/Molecular Massively Parallel Simulator.*
>
> *LAMMPS has potentials for solid-state materials (metals, semiconductors) and soft matter (biomolecules, polymers) and coarse-grained or mesoscopic systems. It can be used to model atoms or, more generically, as a parallel particle simulator at the atomic, meso, or continuum scale.*
>
> *LAMMPS runs on single processors or in parallel using message-passing techniques and a spatial-decomposition of the simulation domain. Many of its models have versions that provide accelerated performance on CPUs, GPUs, and Intel Xeon Phis. The code is designed to be easy to modify or extend with new functionality.*

While most simulation methods used in the LAMMPS simulations in this work are equivalent to the ESPResSo simulations, one aspect was implemented in a different way: The electrostatic interactions were handled *via* the multilevel summation method (MSM).[274]

LAMMPS is maintained and distributed by researchers at the Sandia National Laboratories and Temple University as an open source code under the terms of the GNU General Public License.

The LAMMPS simulations in this work were performed by Nataša Adžić.

A.3 oxDNA

The oxDNA code was developed by Ouldridge *et al.*[275] and is intended for MD and MC simulations of DNA and RNA on CPUs and GPUs. In our work it was used as a complementary method to our coarse-grained approach, where the double helix structure of DNA is not explicitly represented. In this model DNA is considered as a string of rigid nucleotides, interacting via potentials which depend on the nucleotides' position and orientation.

[272] Introduced in S. Plimpton. "Fast Parallel Algorithms for Short-Range Molecular Dynamics". In: *J. Comp. Phys.* 117 (1995), pp. 1–19. DOI: 10.1006/jcph.1995.1039.
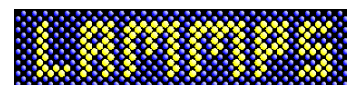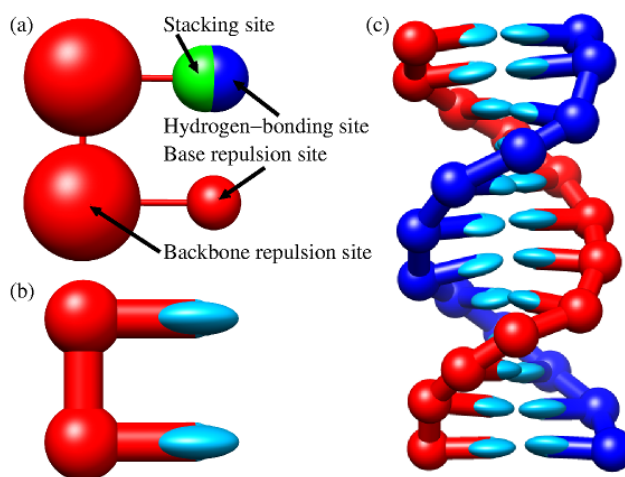
[273] See https://lammps.sandia.gov/.

FIGURE A.2: Logo of the LAMMPS simulation software.

[274] See D. J. Hardy et al. "Multilevel Summation with B-spline Interpolation for Pairwise Interactions in Molecular Dynamics Simulations". In: *J. Chem. Phys.* 144 (2016), p. 114112. DOI: 10.1063/1.4943868 and D. J. Hardy et al. "Multilevel Summation Method for Electrostatic Force Evaluation". In: *J. Chem. Theory Comput.* 11 (2015), pp. 766–779. DOI: 10.1021/ct5009075.

[275] The oxDNA code was first published in 2011, see Ouldridge et al., "Structural, Mechanical, and Thermodynamic Properties of a Coarse-Grained DNA Model". An in-depth explanation of the oxDNA model can also be found in Ouldridge, "Coarse-grained modelling of DNA and DNA self-assembly".

[276]See Snodin et al., "Introducing improved structural properties and salt dependence into a coarse-grained model of DNA".

FIGURE A.4: A simulation snapshot of a G2 DL-DNA with rigid connetions in the oxDNA2 model. Differently coloured strands correspond to the different ssDNA strands listed in Section 3.4. Image courtesy of Nataša Adžić.

[277]See A. Singraber. "Designing and training neural network potentials for molecular dynamics simulations". PhD thesis. University of Vienna, AT, 2018. URL: http://othes.univie.ac.at/55534/.

[278]See J. Behler and M. Parinello. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces". In: *Phys. Rev. Lett.* 98 (2007), p. 146401. DOI: 10.1103/PhysRevLett.98.146401.

These specific interactions include: sugar-phosphate backbone connectivity, steric interaction, hydrogen bonding, nearest-neighbour stacking, and cross-stacking as well as coaxial stacking. The oxDNA model and its interactions are illustrated in Figure A.3. An improved version of the oxDNA software, oxDNA2, which includes new features, *e.g.*, the possibility to simulate systems at salt concentrations above $c = 100\,\text{mM}$, has been released in 2015[276] and was used in this work. An implementation of the oxDNA and oxDNA2 models is included in the LAMMPS repository.

The oxDNA simulations in this work were performed by Nataša Adžić.

A.4 n2p2

The neural network potential package (n2p2) provides the necessary software to implement high-dimensional neural network potentials (HDNNP) for research projects in computational physics and chemistry. Developed by Andreas Singraber during his PhD[277] the theory behind this software package is based on the Behler-Parinello neural network potentials.[278] This package contains the necessary tools to set up and train HDNNPs in order to predict energies and forces from coordinate data of the system the HDNNP was trained on. Furthermore, existing neural network potential parameterizations can be implemented quickly in order to run MD simulations, *e.g.*,

via LAMMPS.[279] Among the systems investigated *via* the n2p2 software so far are H_2O systems[280] and Cu_2S systems.[281] The corresponding training data sets are provided with the official documentation.[282]

The n2p2 calculations in this work were performed by Florian Buchner.

[279] See A. Singraber et al. "Library-Based LAMMPS Implementation of High-Dimensional Neural Network Potentials". In: *J. Chem. Theory Comput.* 15 (2019), pp. 1827–1840. DOI: 10.1021/acs.jctc.8b00770 for the implementation of HDNNPs for MD simulations in LAMMPS.

[280] See T. Morawietz et al. "How van der Waals interactions determine the unique properties of water". In: *Proc. Natl. Acad. Sci. U.S.A.* 113 (2016), pp. 8368–8373. DOI: 10.1073/pnas.1602375113.

[281] See A. Singraber et al. "Parallel Multistream Training of High-Dimensional Neural Network Potentials". In: *J. Chem. Theory Comput.* 15 (2019), pp. 3075–3092. DOI: 10.1021/acs.jctc.8b01092.

[282] See https://compphysvienna.github.io/n2p2/index.html.

Appendix B

Simulation Parameter Values

In this appendix the parameter values for the Molecuar Dynamics (MD) simulations performed *via* the ESPResSo simulation package in Chapters 4, 5, 6, and 7 are presented. All of these simulations were performed by the author.

B.1 Simulation Units in ESPResSo

The fundamental physical quantities in the MD simulations of our project are expressed *via* the units of measurement shown in Table B.1.

Unit	Value
[length]	$1.0\,\text{Å} = 1.0 \times 10^{-10}\,\text{m}$
[mass]	$1.0\,\text{u} = 1.66 \times 10^{-27}\,\text{kg}$
[energy]	$1.0\,\text{kJ mol}^{-1} = 1.66 \times 10^{-21}\,\text{J}$
[time]	$1.0\,\Delta t_{\text{sys}} = 1.0 \times 10^{-13}\,\text{s}$

TABLE B.1: Units of the fundamental physical dimensions used in ESPResSo simulations. These physical dimensions must satisfy Equation (B.1).

While three of these units of measurements may be chosen arbitrarily in ESPResSo, all four must satisfy the relation[283]

$$[\text{time}] = [\text{length}] \sqrt{\frac{[\text{mass}]}{[\text{energy}]}} \ .\tag{B.1}$$

[283] See Weik et al., "ESPResSo 4.0 - an extensible software package for simulating soft matter systems".

The input values of the simulation software pertaining to our model in

Section 3.2 and the simulation output data of the ESPResSo software are communicated *via* the Python interface in units of Table B.1.

B.2 Simulation Parameter Values for Chapter 4

In Chapter 4, a single MD simulation run was performed for each DL-DNA generation with generation numbers from G1 to G8. In these simulations the system contained one single DL-DNA molecule and the counterions. Simulations were run in canonical NVT ensembles with periodic boundary conditions (PBC) at temperature $T = 298\,\mathrm{K}$ ($k_\mathrm{B}T = 2.478\,\mathrm{kJ\,mol^{-1}}$) using a Langevin thermostat with friction coefficient $\gamma = 1\,\mathrm{u}/\Delta t_\mathrm{sys}$, where u and Δt_sys denote the mass and time units in Table B.1. The simulation box size was chosen such that the interaction between the DL-DNA molecules in neighbouring simulation boxes was negligible. While the appropriate amount of counterions was present, no additional salt ions were introduced.

TABLE B.2: Time parameter values of the warmup procedure for MD simulations of dilute solutions of DL-DNAs: number of steps N_steps, time step Δt, and total simulation time t_tot. These parameter values were used for simulations of DL-DNA dendrimers of generation numbers G1 to G8.

GN	N_steps	Δt	t_tot
G1-G8	1.0×10^5	$1.0 \times 10^{-15}\,\mathrm{s}$	$1.0 \times 10^{-10}\,\mathrm{s}$

Approximately the first 10% of the samples were discarded in equilibration, depending on the convergence of quantities of interest, *e.g.*, kinetic and potential energies, towards equilibrium behaviour. In order to avoid unphysical configurations a warmup procedure was performed before the actual simulation, see Section 2.2.2. The accuracy of the P^3M algorithm, with which ESPResSo evaluates the Ewald summations and thus handles electrostatic interactions, was set to 10^{-4}, *i.e.*, the relative force error was less than 10^{-4}.

GN	N_{steps}	Δt	t_{tot}
G1	4.0×10^7	$5.0 \times 10^{-15}\,\text{s}$	$2.0 \times 10^{-7}\,\text{s}$
G2	4.0×10^7	$5.0 \times 10^{-15}\,\text{s}$	$2.0 \times 10^{-7}\,\text{s}$
G3	4.0×10^7	$5.0 \times 10^{-15}\,\text{s}$	$2.0 \times 10^{-7}\,\text{s}$
G4	4.0×10^7	$5.0 \times 10^{-15}\,\text{s}$	$2.0 \times 10^{-7}\,\text{s}$
G5	4.0×10^7	$5.0 \times 10^{-15}\,\text{s}$	$2.0 \times 10^{-7}\,\text{s}$
G6	3.0×10^7	$5.0 \times 10^{-15}\,\text{s}$	$1.5 \times 10^{-7}\,\text{s}$
G7	3.0×10^7	$5.0 \times 10^{-15}\,\text{s}$	$1.5 \times 10^{-7}\,\text{s}$
G8	1.0×10^7	$5.0 \times 10^{-15}\,\text{s}$	$5.0 \times 10^{-8}\,\text{s}$

TABLE B.3: Time parameter values of MD simulations of dilute solutions of DL-DNAs: number of steps N_{steps}, time step Δt, and total simulation time t_{tot}. Parameter values are shown for DL-DNA dendrimers of generation numbers G1 to G8.

Important simulation paramater values are shown in Tables B.2-B.4.

GN	N_{tot}	ρ	V_{b}	l_{b}
G1	102	$5.0 \times 10^{-7}\,\text{Å}^{-3}$	$2.04 \times 10^8\,\text{Å}^3$	$5.89 \times 10^2\,\text{Å}$
G2	384	$5.0 \times 10^{-7}\,\text{Å}^{-3}$	$7.68 \times 10^8\,\text{Å}^3$	$9.16 \times 10^2\,\text{Å}$
G3	948	$5.0 \times 10^{-7}\,\text{Å}^{-3}$	$1.91 \times 10^9\,\text{Å}^3$	$1.24 \times 10^3\,\text{Å}$
G4	2076	$5.0 \times 10^{-7}\,\text{Å}^{-3}$	$4.15 \times 10^9\,\text{Å}^3$	$1.61 \times 10^3\,\text{Å}$
G5	4332	$5.0 \times 10^{-7}\,\text{Å}^{-3}$	$8.67 \times 10^9\,\text{Å}^3$	$2.05 \times 10^3\,\text{Å}$
G6	8844	$5.0 \times 10^{-7}\,\text{Å}^{-3}$	$1.77 \times 10^{10}\,\text{Å}^3$	$2.67 \times 10^3\,\text{Å}$
G7	17868	$5.0 \times 10^{-7}\,\text{Å}^{-3}$	$3.57 \times 10^{10}\,\text{Å}^3$	$3.29 \times 10^3\,\text{Å}$
G8	35916	$5.0 \times 10^{-7}\,\text{Å}^{-3}$	$7.18 \times 10^{10}\,\text{Å}^3$	$4.16 \times 10^3\,\text{Å}$

TABLE B.4: Length and particle number parameter values of MD simulations of dilute solutions of DL-DNAs: total number of particles $N_{\text{tot}} = N_{\text{mon}} + N_{\text{ci}}$ (with DNA monomer number N_{mon} and counterion number N_{ci}), particle density $\rho = N_{\text{tot}}/V_{\text{b}}$, simulation box volume V_{b}, and simulation box length $l_{\text{b}} = \sqrt[3]{V_{\text{b}}}$. Parameter values are shown for simulations of DL-DNA dendrimers of generation numbers G1 to G8.

B.3 Simulation Parameter Values for Chapter 5

In Chapter 5, the samples for the Widom insertion (WI) method were generated *via* MD simulations of DL-DNA dendrimers with generation numbers G1, G2, and G3.

GN	N_{steps}	Δt	t_{tot}
G1-G3	1.0×10^5	$1.0 \times 10^{-15}\,\text{s}$	$1.0 \times 10^{-10}\,\text{s}$

DL-DNA dendrimers with rigid and with flexible connections were simulated at salt concentrations $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$. In order to handle the large amount of salt ions that correspond to such high salt concentrations the electrostatic interactions are handled using the Debye-Hückel potential, where counterions and salt ions are treated implicitly.

c	κ	λ_{D}
$100\,\text{mM}$	$0.106\,\text{Å}^{-1}$	$9.444\,\text{Å}$
$150\,\text{mM}$	$0.126\,\text{Å}^{-1}$	$7.932\,\text{Å}$
$500\,\text{mM}$	$0.230\,\text{Å}^{-1}$	$4.345\,\text{Å}$

These simulations were run in canonical NVT ensembles with PBC at temperature $T = 298\,\text{K}$ ($k_{\text{B}}T = 2.478\,\text{kJ}\,\text{mol}^{-1}$) using a Langevin thermostat with friction coefficient $\gamma = 1\,\text{u}/\Delta t_{\text{sys}}$, where u and Δt_{sys} denote the mass and time units in Table B.1. In order to generate enough samples, the number of simulation runs for G1, G2, and G3 was 4, 20, and 40, respectively. In these simulations the system contained one single DL-DNA molecule. The simulation box size was chosen such that the interaction between the DL-DNA molecules in neighbouring simulation boxes was negligible. This enabled us to use 10^5, $5 \cdot 10^5$, and $5 \cdot 10^5$ Widom insertions for the calculation of the effective potentials $\phi_{\text{eff}}(r)$ for G1, G2, and G3, respectively.

GN	N_{steps}	Δt	t_{tot}
G1	5.0×10^7	$5.0 \times 10^{-15}\,\mathrm{s}$	$2.5 \times 10^{-7}\,\mathrm{s}$
G2	1.0×10^8	$5.0 \times 10^{-15}\,\mathrm{s}$	$5.0 \times 10^{-7}\,\mathrm{s}$
G3	1.0×10^8	$5.0 \times 10^{-15}\,\mathrm{s}$	$5.0 \times 10^{-7}\,\mathrm{s}$

TABLE B.7: Time parameter values of MD simulations of dilute solutions of DL-DNAs used for the computation of effective potentials $\phi_{\mathrm{eff}}(r)$ *via* the WI method: number of steps N_{steps}, time step Δt, and total simulation time t_{tot}. Parameter values are shown for simulations of DL-DNA dendrimers of generation numbers G1 to G3.

Approximately the first 10% of the each sample generation run were discarded in equilibration, depending on the convergence of quantities of interest, *e.g.*, kinetic and potential energies, towards equilibrium behaviour. In order to avoid unphysical configurations a warmup procedure was performed before the actual simulation, see Section 2.2.2. Simulation parameter values were the same, irrespective if the effective coordinate was chosen to be center-of-mass r_{com} or central Y-junction r_{cy}.

GN	N_{mon}	ρ	V_{b}	l_{b}
G1	51	$2.5 \times 10^{-7}\,\text{Å}^{-3}$	$2.04 \times 10^8\,\text{Å}^3$	$5.89 \times 10^2\,\text{Å}$
G2	192	$5.0 \times 10^{-8}\,\text{Å}^{-3}$	$3.84 \times 10^9\,\text{Å}^3$	$1.56 \times 10^3\,\text{Å}$
G3	474	$5.0 \times 10^{-8}\,\text{Å}^{-3}$	$9.48 \times 10^9\,\text{Å}^3$	$2.12 \times 10^3\,\text{Å}$

TABLE B.8: Length and particle number parameter values of MD simulations of dilute solutions of DL-DNAs used for the computation of effective potentials $\phi_{\mathrm{eff}}(r)$ *via* the WI method: number of DNA monomers N_{mon}, particle density $\rho = N_{\mathrm{mon}}/V_{\mathrm{b}}$, simulation box volume V_{b}, and simulation box length $l_{\mathrm{b}} = \sqrt[3]{V_{\mathrm{b}}}$. Parameter values are shown for simulations of DL-DNA dendrimers of generation numbers G1 to G3.

Important simulation paramater values are listed in the following Tables B.6-B.8.

B.4 Simulation Parameter Values for Chapter 6

In Chapter 6, each examined system, *i.e.*, a system of GN DL-DNAs ($N = 1, 2, 3$) at salt concentration c and at density η, was simulated in a single MD simulation run.

GN	N_{steps}	Δt	t_{tot}
G1-G3	1.0×10^5	$1.0 \times 10^{-15}\,\mathrm{s}$	$1.0 \times 10^{-10}\,\mathrm{s}$

TABLE B.9: Time parameter values of the warmup procedure for all MD simulations of bulk solutions of DL-DNAs: number of steps N_{steps}, time step Δt, and total simulation time t_{tot}. These parameter values were used for all simulations bulk solutions of DL-DNA.

Systems were examined at salt concentrations $c_1 = 150\,\mathrm{mM}$ and $c_2 = 150\,\mathrm{mM}$. If not stated otherwise, the presented simulation parameter values were used for systems of both salt concentrations. In order to handle the large amount of salt ions that correspond to such high salt concentrations the electrostatic interactions are handled using the Debye-Hückel potential, where counterions and salt ions are treated implicitly, see Table B.6 for values of the Debye length λ_D.

TABLE B.10: Time parameter values of MD simulations of bulk solutions of G1 DL-DNAs: number of steps N_steps, time step Δt, and total simulation time t_tot. These parameter values were used for simulations of G1 bulk systems at density $\eta \in \{0.001, 0.1, 0.5\}$.

N_steps	Δt	t_tot
1.0×10^8	$5.0 \times 10^{-15}\,\mathrm{s}$	$5.0 \times 10^{-7}\,\mathrm{s}$

TABLE B.11: Length and particle number parameter values of MD simulations of bulk solutions of G1 DL-DNAs: number of DNA monomers N_mon, particle density $\rho = N_\mathrm{mon}/V_\mathrm{b}$, simulation box volume V_b, and simulation box length $l_\mathrm{b} = \sqrt[3]{V_\mathrm{b}}$. Parameter values are shown for simulations of G1 bulk systems at density $\eta \in \{0.001, 0.1, 0.5\}$.

η	N_tot	ρ	V_b	l_b
0.001	12750	$3.37 \times 10^{-7}\,\text{Å}^{-3}$	$3.78 \times 10^{10}\,\text{Å}^3$	$3.36 \times 10^3\,\text{Å}$
0.1	12750	$3.37 \times 10^{-5}\,\text{Å}^{-3}$	$3.78 \times 10^8\,\text{Å}^3$	$7.23 \times 10^2\,\text{Å}$
0.5	12750	$1.69 \times 10^{-4}\,\text{Å}^{-3}$	$7.56 \times 10^7\,\text{Å}^3$	$4.23 \times 10^2\,\text{Å}$

Simulations were run in canonical NVT ensembles with periodic boundary conditions (PBC) at temperature $T = 298\,\mathrm{K}$ ($k_\mathrm{B}T = 2.478\,\mathrm{kJ\,mol^{-1}}$) using a Langevin thermostat with friction coefficient $\gamma = 1\,\mathrm{u}/\Delta t_\mathrm{sys}$, where u and Δt_sys denote the mass and time units in Table B.1. Approximately the first 10% of the samples were discarded in equilibration, depending on the convergence of quantities of interest, *e.g.*, kinetic and potential energies, towards equilibrium behaviour. In order to avoid unphysical configurations a warmup procedure was performed before the actual simulation, see Section 2.2.2. Important simulation paramater values are shown in Tables B.9-B.15.

η	N_{steps}	Δt	t_{tot}
0.5	9.0×10^8	$5.0 \times 10^{-15}\,\text{s}$	$4.50 \times 10^{-6}\,\text{s}$
1.0	4.0×10^8	$5.0 \times 10^{-15}\,\text{s}$	$2.00 \times 10^{-6}\,\text{s}$
2.0	1.5×10^7	$2.5 \times 10^{-15}\,\text{s}$	$3.75 \times 10^{-7}\,\text{s}$

TABLE B.12: Time parameter values of MD simulations of bulk solutions of G2 DL-DNAs: number of steps N_{steps}, time step Δt, and total simulation time t_{tot}. These parameter values were used for simulations of G2 bulk systems at density $\eta \in \{0.5, 1.0, 2.0\}$.

η	N_{tot}	ρ	V_{b}	l_{b}
0.5	19200	$2.58 \times 10^{-6}\,\text{Å}^{-3}$	$7.42 \times 10^8\,\text{Å}^3$	$9.05 \times 10^2\,\text{Å}$
1.0	38400	$5.15 \times 10^{-5}\,\text{Å}^{-3}$	$7.42 \times 10^8\,\text{Å}^3$	$9.05 \times 10^2\,\text{Å}$
2.0	76800	$1.03 \times 10^{-4}\,\text{Å}^{-3}$	$7.42 \times 10^8\,\text{Å}^3$	$9.05 \times 10^2\,\text{Å}$

TABLE B.13: Length and particle number parameter values of MD simulations of bulk solutions of G2 DL-DNAs: number of DNA monomers N_{mon}, particle density $\rho = N_{\text{mon}}/V_{\text{b}}$, simulation box volume V_{b}, and simulation box length $l_{\text{b}} = \sqrt[3]{V_{\text{b}}}$. Parameter values are shown for simulations of G2 bulk systems at density $\eta \in \{0.5, 1.0, 2.0\}$.

η	N_{steps}	Δt	t_{tot}
0.01	1.20×10^8	$2.5 \times 10^{-15}\,\text{s}$	$3.0 \times 10^{-7}\,\text{s}$

TABLE B.14: Time parameter values of MD simulations of bulk solutions of G3 DL-DNAs: number of steps N_{steps}, time step Δt, and total simulation time t_{tot}. These parameter values were used for simulations of G3 bulk systems at density $\eta = 0.01$.

η	N_{tot}	ρ	V_{b}	l_{b}
0.01	23700	$2.89 \times 10^{-7}\,\text{Å}^{-3}$	$8.19 \times 10^{10}\,\text{Å}^3$	$4.34 \times 10^3\,\text{Å}$

TABLE B.15: Length and particle number parameter values of MD simulations of bulk solutions of G3 DL-DNAs: number of DNA monomers N_{mon}, particle density $\rho = N_{\text{mon}}/V_{\text{b}}$, simulation box volume V_{b}, and simulation box length $l_{\text{b}} = \sqrt[3]{V_{\text{b}}}$. Parameter values are shown for simulations of G3 bulk systems at density $\eta \in \{0.5, 1.0, 2.0\}$.

B.5 Simulation Parameter Values for Chapter 7

In Chapter 7, each examined system, *i.e.*, a system of G1 DNA stars at salt concentration $c_3 = 100\,\text{mM}$ and at interface density η_{s}, was simulated in a single MD simulation run.

GN	N_{steps}	Δt	t_{tot}
G1	1.0×10^5	$1.0 \times 10^{-15}\,\text{s}$	$1.0 \times 10^{-10}\,\text{s}$

TABLE B.16: Time parameter values of the warmup procedure for MD simulations of DNA stars in two dimensions: number of steps N_{steps}, time step Δt, and total simulation time t_{tot}. These parameter values were used for simulations of DNA stars in the flat and the tripod configuration at all simulated area densities η.

In order to handle the large amount of salt ions that correspond to such high salt concentrations the electrostatic interactions are handled using the Debye-Hückel potential, where counterions and salt ions are treated implicitly, see Table B.6 for values of the Debye length λ_D.

TABLE B.17: Time parameter values of MD simulations of DNA stars in two dimensions: number of steps N_{steps}, time step Δt, and total simulation time t_{tot}. These parameter values were used for simulations of DNA stars in the flat and the tripod configuration at all simulated area densities η.

GN	N_{steps}	Δt	t_{tot}
G1	2.0×10^7	$5.0 \times 10^{-15}\,\text{s}$	$1.0 \times 10^{-7}\,\text{s}$

Simulations were run in canonical NVT ensembles with periodic boundary conditions (PBC) at temperature $T = 298\,\text{K}$ ($k_B T = 2.478\,\text{kJ}\,\text{mol}^{-1}$) using a Langevin thermostat with friction coefficient $\gamma = 1\,\text{u}/\Delta t_{\text{sys}}$, where u and Δt_{sys} denote the mass and time units in Table B.1. Approximately the first 10% of the samples were discarded in equilibration, depending on the convergence of quantities of interest, *e.g.*, kinetic and potential energies, towards equilibrium behaviour. In order to avoid unphysical configurations a warmup procedure was performed before the actual simulation, see Section 2.2.2. Important simulation paramater values are shown in Tables B.16-B.19.

TABLE B.18: Length and particle number parameter values of MD simulations in two dimensions of DNA stars in the flat configuration: number of DNA monomers N_{mon}, particle density $\rho = N_{\text{mon}}/A_i$, interface area A_i, and simulation box length $l_b = \sqrt[2]{A_i}$. Parameter values are shown for simulations of DNA stars at alt concentration $c_3 = 100\,\text{mM}$ at different area densities η.

η	N_{mon}	ρ	l_b	A_i
0.17	15600	$3.20 \times 10^{-3}\,\text{Å}^{-2}$	$2.19 \times 10^3\,\text{Å}$	$4.80 \times 10^6\,\text{Å}^2$
0.20	15600	$3.90 \times 10^{-3}\,\text{Å}^{-2}$	$2.00 \times 10^3\,\text{Å}$	$4.00 \times 10^6\,\text{Å}^2$
0.24	15600	$4.60 \times 10^{-3}\,\text{Å}^{-2}$	$1.85 \times 10^3\,\text{Å}$	$3.43 \times 10^6\,\text{Å}^2$
0.27	15600	$5.20 \times 10^{-3}\,\text{Å}^{-2}$	$1.73 \times 10^3\,\text{Å}$	$3.00 \times 10^6\,\text{Å}^2$
0.31	15600	$5.80 \times 10^{-3}\,\text{Å}^{-2}$	$1.63 \times 10^3\,\text{Å}$	$2.67 \times 10^6\,\text{Å}^2$
0.34	15600	$6.50 \times 10^{-3}\,\text{Å}^{-2}$	$1.55 \times 10^3\,\text{Å}$	$2.40 \times 10^6\,\text{Å}^2$

η	N_{mon}	ρ	l_{b}	A_{i}
0.17	15600	3.20×10^{-3} Å^{-2}	2.19×10^3 Å	4.80×10^6 Å^2
0.20	15600	3.90×10^{-3} Å^{-2}	2.00×10^3 Å	4.00×10^6 Å^2
0.24	15600	4.60×10^{-3} Å^{-2}	1.85×10^3 Å	3.43×10^6 Å^2
0.27	15600	5.20×10^{-3} Å^{-2}	1.73×10^3 Å	3.00×10^6 Å^2
0.31	15600	5.80×10^{-3} Å^{-2}	1.63×10^3 Å	2.67×10^6 Å^2
0.34	15600	6.50×10^{-3} Å^{-2}	1.55×10^3 Å	2.40×10^6 Å^2
0.37	15600	7.20×10^{-3} Å^{-2}	1.48×10^3 Å	2.18×10^6 Å^2
0.41	15600	7.80×10^{-3} Å^{-2}	1.41×10^3 Å	2.00×10^6 Å^2
0.44	15600	8.40×10^{-3} Å^{-2}	1.36×10^3 Å	1.85×10^6 Å^2
0.48	15600	9.10×10^{-3} Å^{-2}	1.31×10^3 Å	1.71×10^6 Å^2
0.51	15600	9.80×10^{-3} Å^{-2}	1.26×10^3 Å	1.60×10^6 Å^2
0.54	15600	1.04×10^{-2} Å^{-2}	1.22×10^3 Å	1.50×10^6 Å^2
0.58	15600	1.11×10^{-2} Å^{-2}	1.19×10^3 Å	1.41×10^6 Å^2

TABLE B.19: Length and particle number parameter values of MD simulations in two dimensions of DNA stars in the tripod configuration: number of DNA monomers N_{mon}, particle density $\rho = N_{\mathrm{mon}}/A_{\mathrm{i}}$, interface area A_{i}, and simulation box length $l_{\mathrm{b}} = \sqrt[2]{A_{\mathrm{i}}}$. Parameter values are shown for simulations of DNA stars at salt concentration $c_3 = 100$ mM at different area densities η.

Analogous to Chapter 5, the samples for the Widom insertion (WI) method in Chapter 7 were generated *via* MD simulations of planar DNA stars in two dimensions (simulation parameter values not listed). In these simulations the system contained one single DNA star. The simulation box size was chosen such that the interaction between the DL-DNA molecules in neighbouring simulation boxes was negligible. While one single Widom insertions sufficed for rigid DNA stars, 2000 Widom insertions were performed for equilibrated DNA stars with flexible arms.

Appendix C

Complementary Results

This appendix contains complementary results for Chapters 4, 5, and 6.

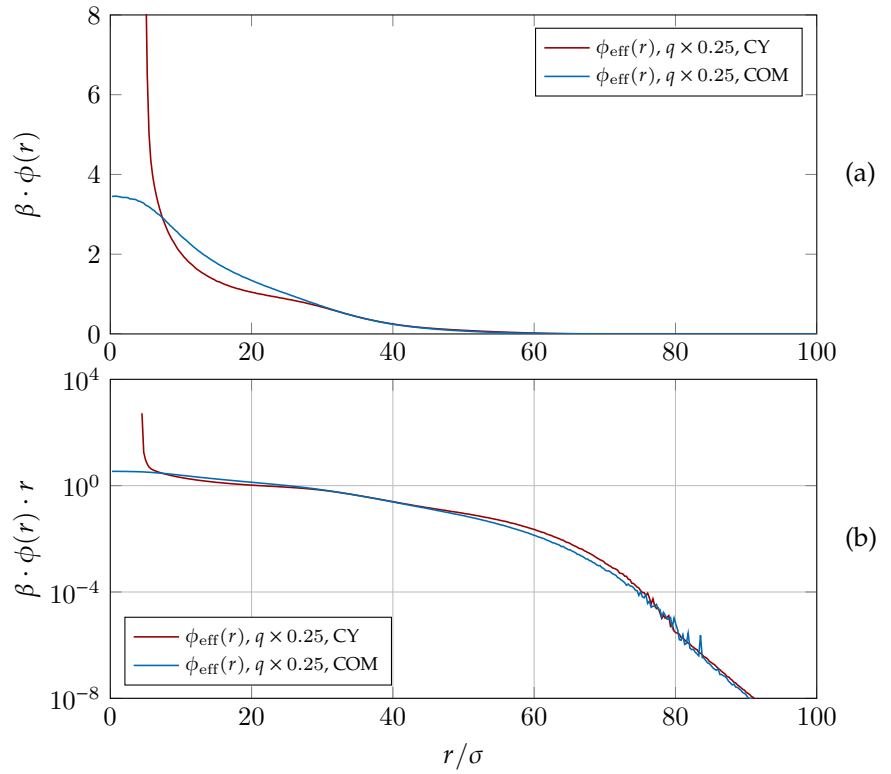## C.1 Complementary Results for Chapter 5
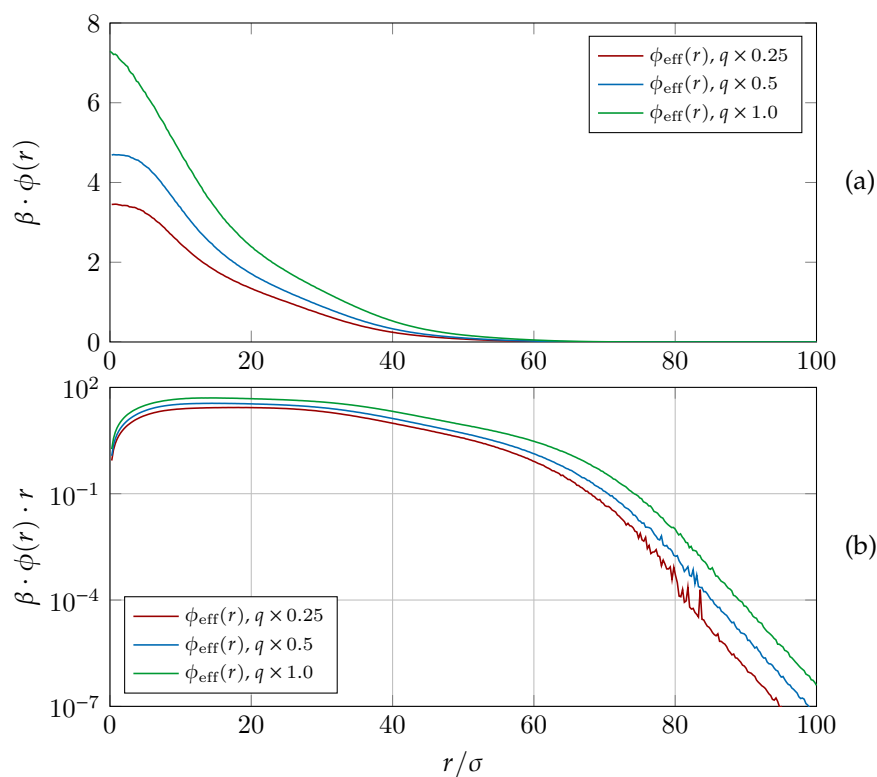
FIGURE C.1: Effective potentials $\phi_{\text{eff}}(r)$ for G2 dendrimers with rigid connections as functions of r, given in units of the steric interaction length σ; Results are shown for salt concentration $c_1 = 150\,\text{mM}$ and at different values of the charge reduction $q \times f$, with $f \in \{0.25, 0.5, 1.0\}$ (as labelled). The effective potentials $\phi_{\text{eff}}(r)$ were calculated with centers-of-mass r_{com} of the DL-DNAs as the effective coordinates.

FIGURE C.2: Effective potentials $\phi_{\text{eff}}(r)$ for G2 dendrimers with rigid connections as functions of r, given in units of the steric interaction length σ; Results are shown for salt concentration $c_1 = 150\,\text{mM}$ and charge reduction $q \times 0.25$. The effective potentials $\phi_{\text{eff}}(r)$ are compared for two different effective coordinates of the dendrimers: centers-of-mass r_{com} (labelled "com") and central Y-junctions r_{cy} (labelled "com").

FIGURE C.3: Effective potentials $\phi_{\text{eff}}(r)$ for G2 dendrimers with rigid connections as functions of r, given in units of the steric interaction length σ; Results are shown for salt concentration $c_1 = 150\,\text{mM}$ and for different values of charge reduction $q \times f$, with $f \in \{0.25, 1.0\}$ (as labelled). The effective potentials $\phi_{\text{eff}}(r)$ were calculated with the positions of the central Y-junctions, r_{cy}, as the effective coordinates.

(a)

(b)

$U_{12} = 7.00 \times 10^{43}\,\mathrm{k}_B T$

$U_{12} = 6.04 \times 10^{-7}\,\mathrm{k}_B T$
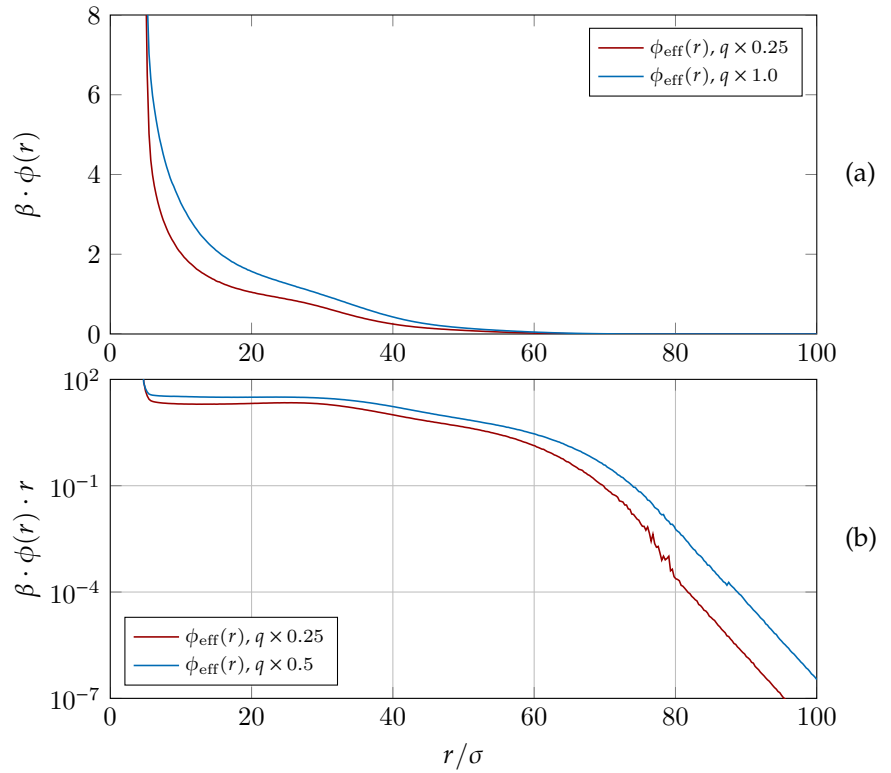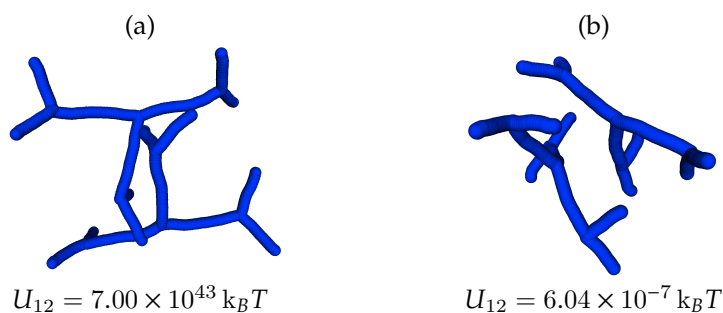
FIGURE C.4: Snapshots of two exemplary Widom insertions for G2 at salt concentration $c_2 = 500\,\mathrm{mM}$ and distance between the centers of mass $r_{12} = 100\,\text{Å}$: (a) shows the case of overlap of the dendrimers, resulting in a very high inter-dendrimer energy $U_{12} = 7.00 \times 10^{43}\,\mathrm{k_B T}$, (b) shows the case of clear separation of the dendrimers, yielding a low inter-dendrimer energy $U_{12} = 6.04 \times 10^{-7}\,\mathrm{k_B T}$.

C.2 Complementary Results for Chapter 6

(a)

(b)
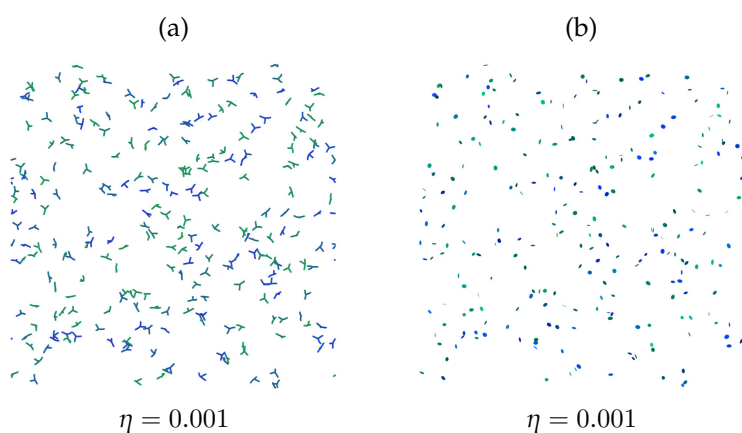
$\eta = 0.001$

$\eta = 0.001$

FIGURE C.5: Simulation snapshots of systems containing G1 DL-DNA dendrimers at density $\eta = 0.001$ and at salt concentration $c_2 = 500\,\mathrm{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).
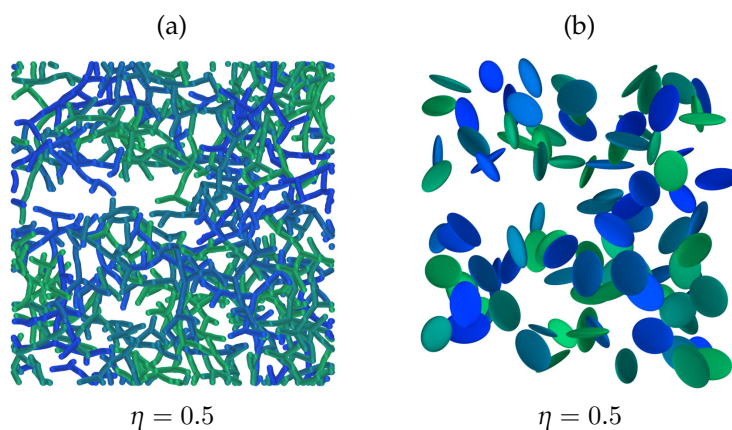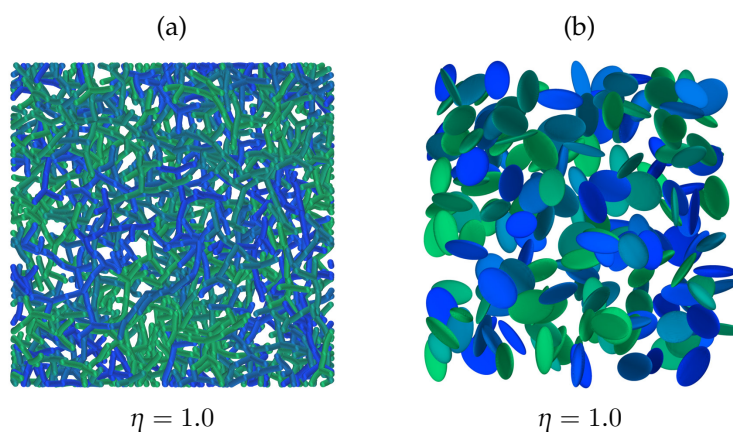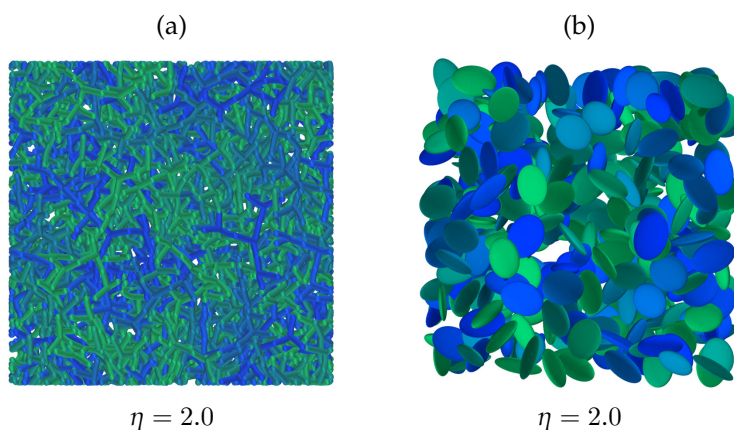
(a)

(b)

$\eta = 0.1$

$\eta = 0.1$
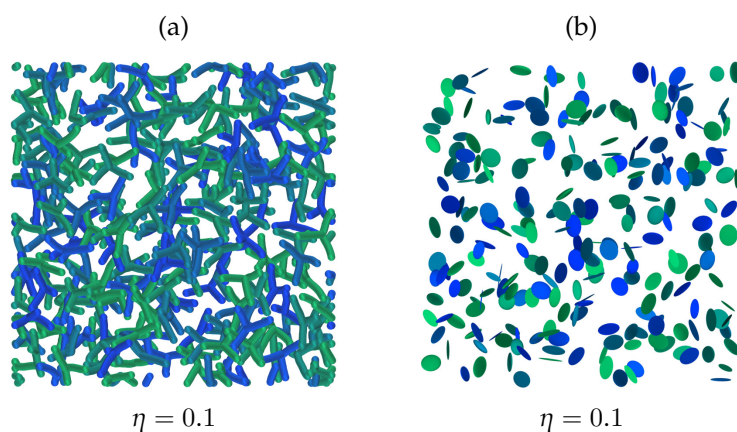
FIGURE C.6: Simulation snapshots of systems containing G1 DL-DNA dendrimers at density $\eta = 0.1$ and at salt concentration $c_2 = 500\,\mathrm{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).
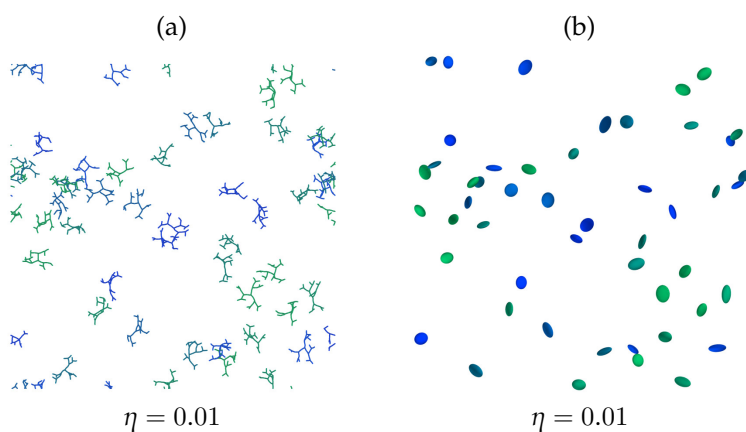
FIGURE C.7: Simulation snapshots of systems containing G1 DL-DNA dendrimers at density $\eta = 0.5$ and at salt concentration $c_2 = 500\,\text{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).


(a)

(b)

$\eta = 0.5$

$\eta = 0.5$
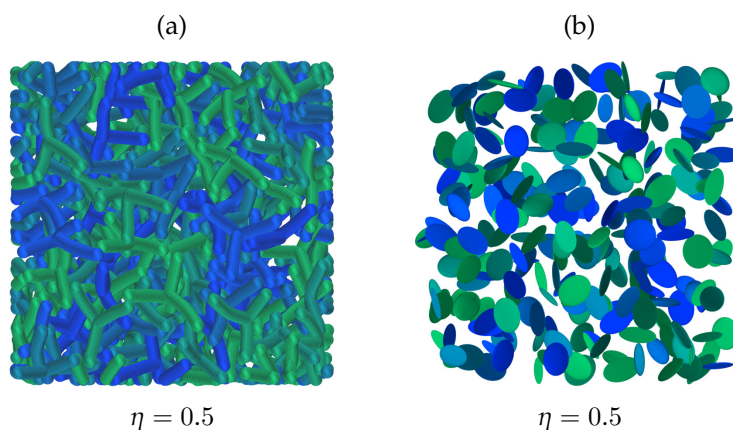
FIGURE C.8: Simulation snapshots of systems containing G2 DL-DNA dendrimers at density $\eta = 0.5$ and at salt concentration $c_2 = 500\,\text{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).

(a)

(b)

$\eta = 0.5$

$\eta = 0.5$

FIGURE C.9: Simulation snapshots of systems containing G2 DL-DNA dendrimers at density $\eta = 1.0$ and at salt concentration $c_2 = 500\,\text{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).

(a)

(b)

$\eta = 1.0$

$\eta = 1.0$

FIGURE C.10: Simulation snapshots of systems containing G2 DL-DNA dendrimers at density $\eta = 2.0$ and at salt concentration $c_2 = 500\,\text{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).

FIGURE C.11: Simulation snapshots of systems containing G3 DL-DNA dendrimers at density $\eta = 0.01$ and at salt concentration $c_2 = 500\,\text{mM}$. Dendrimers are shown in eight different colours in order to increase the distinguishability between them. Whereas subplot (a) explicitly shows the DL-DNA molecules, the dendrimers are represented by their gyration ellipsoids in subplot (b).
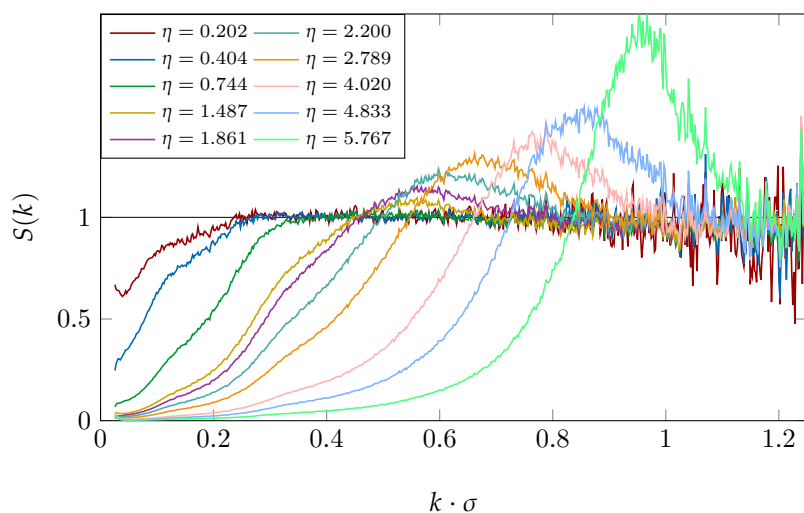
FIGURE C.12: Structure factors $S(k)$ of G2 DL-DNAs with rigid connections as functions of k, given in units of the inverse steric interaction length σ^{-1}; results were obtained from experiment *via* SLS at salt concentration $c_1 = 150\,\text{mM}$. Experimental measurements were made at densities ranging from $\eta = 0.202$ to $\eta = 5.767$ (as labelled). Experimental data courtesy of Manolis Stiakakis.

TABLE C.1: Shape descriptors for DL-DNA dendrimers of generation numbers G1 and G3 in bulk solutions at different densities η and at salt concentrations $c_1 = 150\,\text{mM}$ and $c_1 = 150\,\text{mM}$: values are listed for radius of gyration R_g, eigenvalues of the gyration tensor λ_i^2 (with $i = 1, 2, 3$), asphericity a, acylindricity b, and relative shape anisotropy κ^2, see Section 2.6.5 for definitions. Eigenvalues λ_i^2 are ordered by size, *i.e.*, $\lambda_1^2 < \lambda_2^2 < \lambda_3^2$. The results were obtaind from MD simulations using appropriate time and emsemble averaging and are given in the form $\mu \pm \sigma$, with mean μ and standard deviation σ.

η	$R_g\,[\text{Å}]$	$\frac{\lambda_1^2}{R_g^2}$	$\frac{\lambda_2^2}{R_g^2}$	$\frac{\lambda_3^2}{R_g^2}$	$\frac{\lambda_1^2}{\lambda_3^2}$	$\frac{\lambda_2^2}{\lambda_3^2}$	$\frac{b}{R_g^2}$	$\frac{c}{R_g^2}$	κ^2
				G1, $c = 150$ mM					
0.001	32.61 ± 0.70	0.012 ± 0.015	0.366 ± 0.056	0.622 ± 0.058	0.020 ± 0.027	0.602 ± 0.146	0.433 ± 0.087	0.354 ± 0.059	0.292 ± 0.052
0.1	32.60 ± 0.70	0.012 ± 0.015	0.366 ± 0.056	0.622 ± 0.058	0.020 ± 0.027	0.601 ± 0.146	0.434 ± 0.087	0.354 ± 0.059	0.292 ± 0.052
0.5	32.53 ± 0.74	0.013 ± 0.016	0.362 ± 0.057	0.625 ± 0.059	0.021 ± 0.029	0.593 ± 0.147	0.437 ± 0.088	0.350 ± 0.059	0.293 ± 0.054
				G1, $c = 500$ mM					
0.001	32.60 ± 0.79	0.014 ± 0.018	0.359 ± 0.057	0.627 ± 0.060	0.023 ± 0.032	0.586 ± 0.148	0.441 ± 0.090	0.345 ± 0.060	0.295 ± 0.057
0.1	32.60 ± 0.79	0.014 ± 0.018	0.358 ± 0.057	0.628 ± 0.060	0.023 ± 0.032	0.585 ± 0.148	0.441 ± 0.090	0.344 ± 0.060	0.295 ± 0.057
0.5	32.57 ± 0.81	0.014 ± 0.018	0.357 ± 0.057	0.629 ± 0.060	0.024 ± 0.033	0.582 ± 0.148	0.443 ± 0.090	0.343 ± 0.061	0.295 ± 0.058
				G3, $c = 150$ mM					
0.01	143.24 ± 5.96	0.104 ± 0.041	0.337 ± 0.059	0.558 ± 0.059	0.192 ± 0.087	0.619 ± 0.158	0.338 ± 0.088	0.233 ± 0.084	0.168 ± 0.059
				G3, $c = 500$ mM					
0.01	141.75 ± 6.35	0.106 ± 0.044	0.333 ± 0.059	0.561 ± 0.061	0.195 ± 0.093	0.608 ± 0.158	0.341 ± 0.091	0.226 ± 0.085	0.169 ± 0.062

TABLE C.2: Shape descriptors for DL-DNA dendrimers of generation number G2 in bulk solutions at different densities η and at salt concentrations $c_1 = 150\,\text{mM}$ and $c_1 = 150\,\text{mM}$: values are listed for radius of gyration R_g, eigenvalues of the gyration tensor λ_i^2 (with $i = 1, 2, 3$), asphericity a, acylindricity b, and relative shape anisotropy κ^2. Eigenvalues λ_i^2 are ordered by size, *i.e.*, $\lambda_1^2 < \lambda_2^2 < \lambda_3^2$. The results were obtain from MD simulations using appropriate time and emsemble averaging and are given in the form $\mu \pm \sigma$, with mean μ and standard deviation σ.

| η	R_g[Å]	$\frac{\lambda_1^2}{R_g^2}$	$\frac{\lambda_2^2}{R_g^2}$	$\frac{\lambda_3^2}{R_g^2}$	$\frac{\lambda_1^2}{\lambda_3^2}$	$\frac{\lambda_2^2}{\lambda_3^2}$	$\frac{b}{R_g^2}$	$\frac{c}{R_g^2}$	κ^2
				G2, $c_1 = 150$ mM					
0.5	91.37 ± 2.99	0.032 ± 0.016	0.345 ± 0.065	0.623 ± 0.065	0.053 ± 0.027	0.570 ± 0.162	0.435 ± 0.098	0.312 ± 0.069	0.275 ± 0.063
1.0	91.31 ± 3.02	0.032 ± 0.016	0.344 ± 0.066	0.624 ± 0.066	0.053 ± 0.028	0.568 ± 0.163	0.435 ± 0.098	0.312 ± 0.069	0.276 ± 0.064
2.0	91.27 ± 3.03	0.033 ± 0.016	0.344 ± 0.065	0.623 ± 0.065	0.053 ± 0.028	0.568 ± 0.162	0.435 ± 0.098	0.311 ± 0.069	0.275 ± 0.063
				G2, $c_2 = 500$ mM					
0.5	91.08 ± 3.29	0.034 ± 0.017	0.339 ± 0.066	0.628 ± 0.067	0.055 ± 0.030	0.557 ± 0.164	0.441 ± 0.100	0.305 ± 0.070	0.278 ± 0.067
1.0	91.05 ± 3.31	0.034 ± 0.017	0.339 ± 0.067	0.628 ± 0.067	0.055 ± 0.030	0.557 ± 0.164	0.441 ± 0.100	0.305 ± 0.070	0.278 ± 0.067
2.0	91.07 ± 3.30	0.034 ± 0.017	0.339 ± 0.067	0.628 ± 0.067	0.055 ± 0.030	0.556 ± 0.165	0.442 ± 0.101	0.305 ± 0.071	0.279 ± 0.067

FIGURE C.13: Structure factors $S(k)$ of G2 DL-DNAs with rigid connections as functions of k, given in units of the inverse steric interaction length σ^{-1}; results are obtained from explicit MD simulations (labelled "exp.") and *via* Fourier transform of $g(r)$ computed from explicit MD simulations (labelled "exp. (FT)"). Structure factors $S(k)$ are shown for densities $\eta = 0.5$ (row (a)), $\eta = 1.0$ (row (b)), and $\eta = 2.0$ (row (c)) at salt concentration $c_1 = 150\,\text{mM}$ and $c_2 = 500\,\text{mM}$ (as labelled).
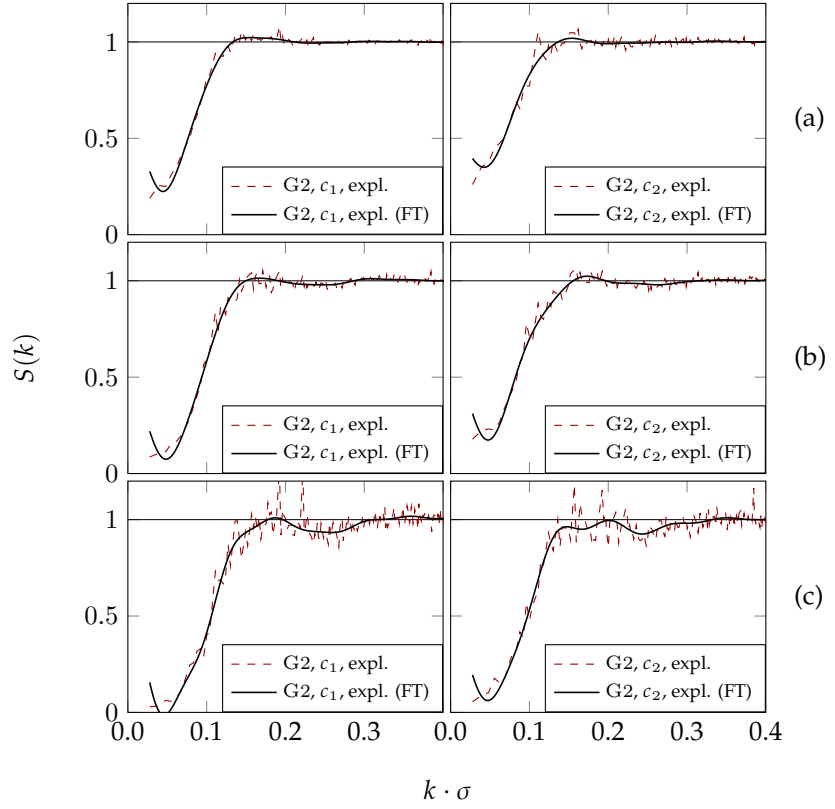
FIGURE C.14: Radial distribution functions $g(r)$ of G2 DL-DNAs obtained *via* the HNC method as functions of r, given in units of the steric interaction length σ; results are shown for G2 DL-DNA with rigid connections at salt concentration $c_1 = 150\,\text{mM}$ and at different values of the charge reduction $q \times f$, with $f \in \{0.25, 1.0\}$ (as labelled). The effective potentials $\phi_{\text{eff}}(r)$ used in the HNC method were calculated with the centers-of-mass, r_{com}, as the effective coordinates. The investigated densities range from $\eta = 0.5$ to $\eta = 5.0$ (as labelled).
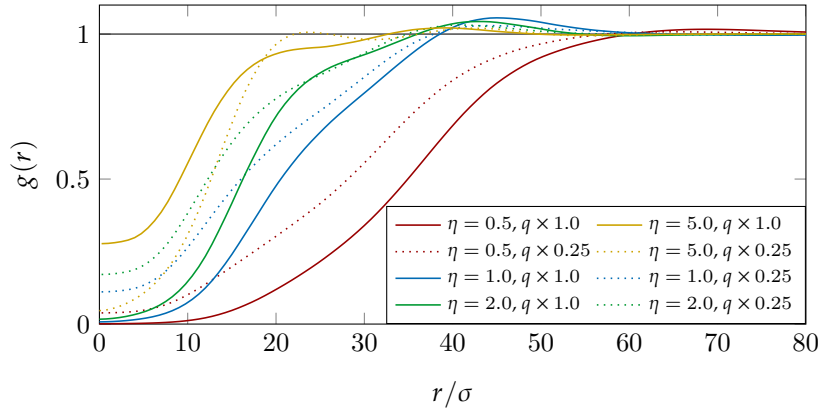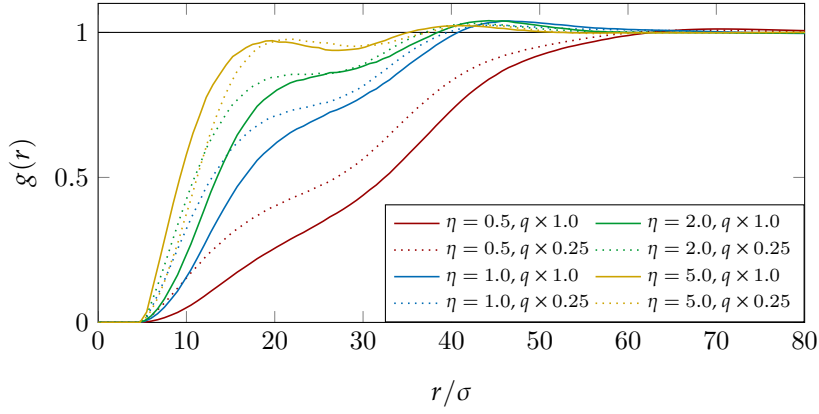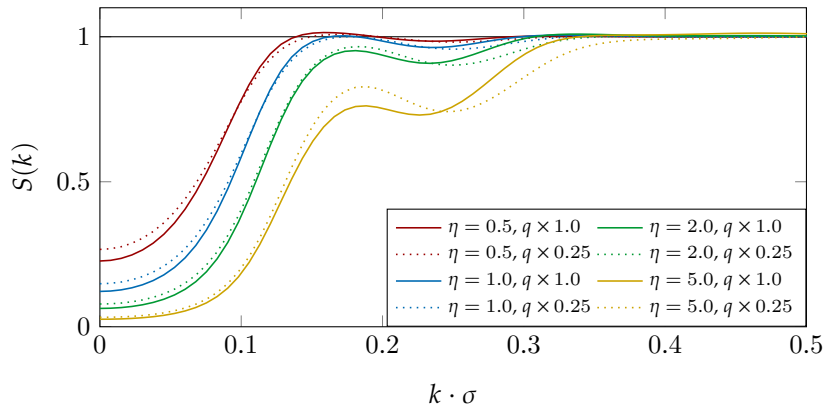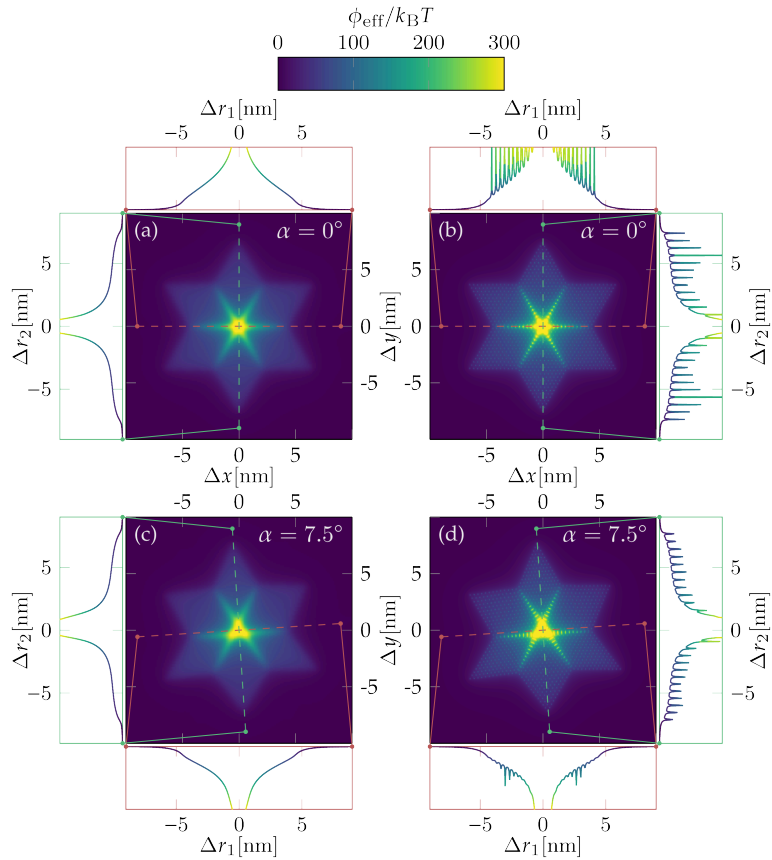
FIGURE C.15: Radial distribution functions $g(r)$ of G2 DL-DNAs obtained *via* the HNC method as functions of r, given in units of the steric interaction length σ; results are shown for G2 DL-DNA with rigid connections at salt concentration $c_1 = 150\,\text{mM}$ and at different values of the charge reduction $q \times f$, with $f \in \{0.25, 1.0\}$ (as labelled). The effective potentials $\phi_{\text{eff}}(r)$ used in the HNC method were calculated with the positions of the central Y-junctions, r_{cy}, as the effective coordinates. The investigated densities range from $\eta = 0.5$ to $\eta = 5.0$ (as labelled).
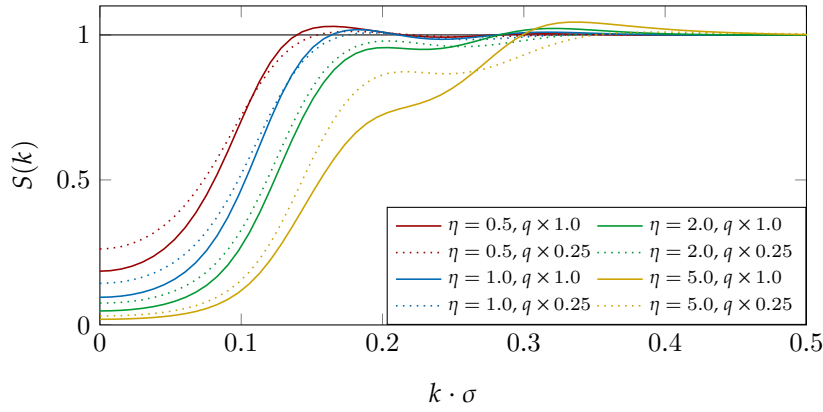
FIGURE C.16: Structure factors $S(k)$ of G2 DL-DNAs obtained *via* the HNC method as functions of k, given in units of the inverse steric interaction length σ^{-1}; results are shown for G2 DL-DNA with rigid connections at salt concentration $c_1 = 150\,\text{mM}$ and at different values of the charge reduction $q \times f$, with $f \in \{0.25, 1.0\}$ (as labelled). The effective potentials $\phi_{\text{eff}}(r)$ used in the HNC method were calculated with the centers-of-mass, r_{com}, as the effective coordinates. The investigated densities range from $\eta = 0.5$ to $\eta = 5.0$ (as labelled).

FIGURE C.17: Structure factors $S(k)$ of G2 DL-DNAs obtained *via* the HNC method as functions of k, given in units of the inverse steric interaction length σ^{-1}; results are shown for G2 DL-DNA with rigid connections at salt concentration $c_1 = 150\,\text{mM}$ and at different values of the charge reduction $q \times f$, with $f \in \{0.25, 1.0\}$ (as labelled). The effective potentials $\phi_{\text{eff}}(r)$ used in the HNC method were calculated with the positions of the central Y-junctions, r_{cy}, as the effective coordinates. The investigated densities range from $\eta = 0.5$ to $\eta = 5.0$ (as labelled).

FIGURE C.18: Structure factors $S(k)$ of G2 DL-DNAs obtained *via* the HNC method as functions of k, given in units of the inverse steric interaction length σ^{-1}; results are shown for dendrimers with rigid connections at salt concentration $c_1 = 150\,\text{mM}$ and at charge reduction $q \times 0.25$. The effective potentials $\phi_{\text{eff}}(r)$ used in the HNC method were calculated with two different effective coordinates of the dendrimers: centers-of-mass r_{com} (labelled "com") and central Y-junctions r_{cy} (labelled "com"). The investigated densities range from $\eta = 0.5$ to $\eta = 5.0$ (as labelled).
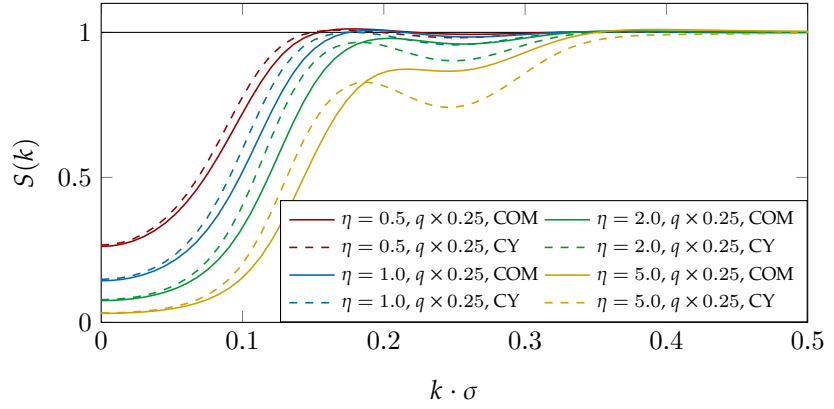
C.3 Complementary Results for Chapter 7

FIGURE C.19: Effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars as function of center-of-mass shift Δx and Δy and rotation angle α. $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ is given in units of $k_B T$, whereas Δx and Δy are given in Å. Both DNA stars are assumed to be in a planar configuration with rigid arms and interarm angles $\theta_i = 2\pi/3$ ($i = 1, 2, 3$) while being rotated by different values of angle α against each other (as labelled). Two dashed, orthogonal lines in red and green indicate two slices through the energy landscape of the effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$. The energy profiles of these slices, denoted by $\phi_{\text{eff}}(\alpha; \Delta r_1)$ and $\phi_{\text{eff}}(\alpha; \Delta r_2)$, are shown in the plots framed red and green. The data for this plot was obtained *via* the reference method outlined in Abaurrea Velasco et al., "Effective interactions of DNA-stars" (subplots (a) and (c)) and *via* the WI method (subplots (b) and (d)).
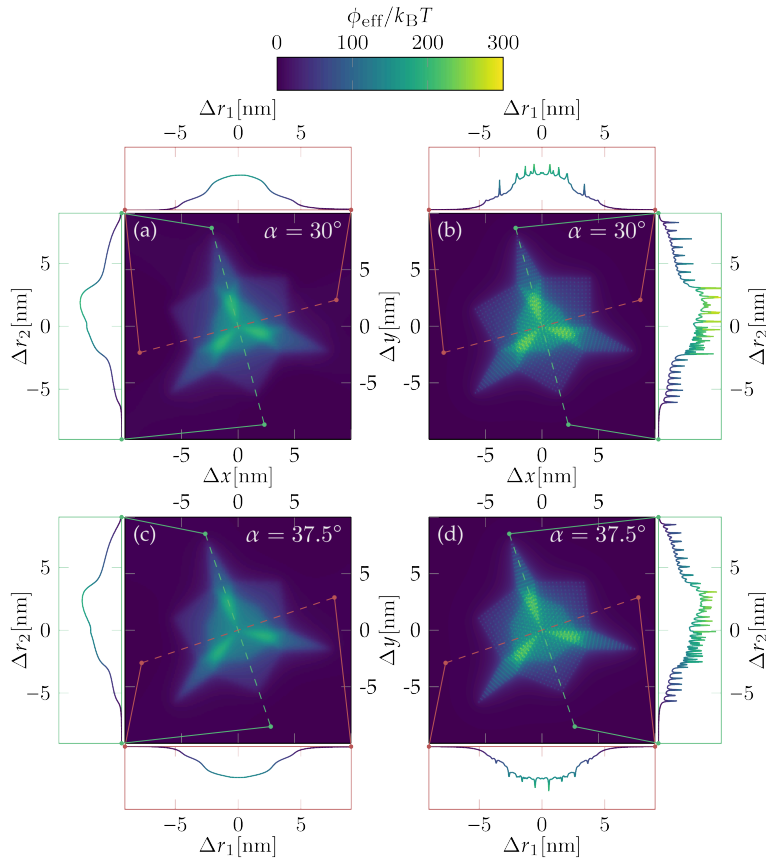
FIGURE C.20: Effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars as function of center-of-mass shift Δx and Δy and rotation angle α. $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ is given in units of $k_{\text{B}}T$, whereas Δx and Δy are given in Å. Both DNA stars are assumed to be in a planar configuration with rigid arms and interarm angles $\theta_i = 2\pi/3$ ($i = 1, 2, 3$) while being rotated by different values of angle α against each other (as labelled). Two dashed, orthogonal lines in red and green indicate two slices through the energy landscape of the effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$. The energy profiles of these slices, denoted by $\phi_{\text{eff}}(\alpha; \Delta r_1)$ and $\phi_{\text{eff}}(\alpha; \Delta r_2)$, are shown in the plots framed red and green. The data for this plot was obtained *via* the reference method outlined in Abaurrea Velasco et al., "Effective interactions of DNA-stars" (subplots (a) and (c)) and *via* the WI method (subplots (b) and (d)).
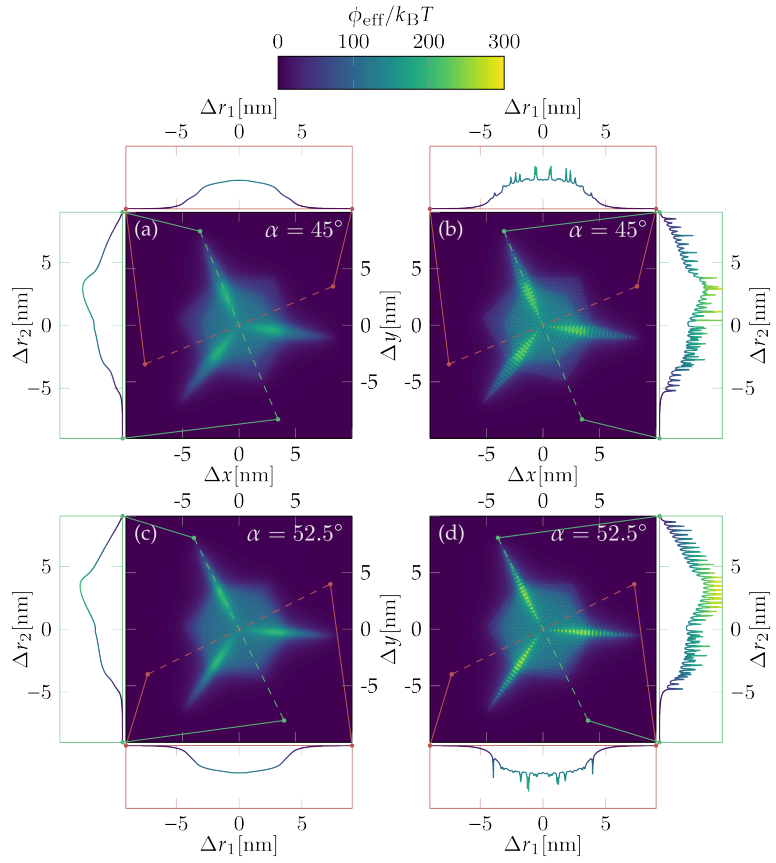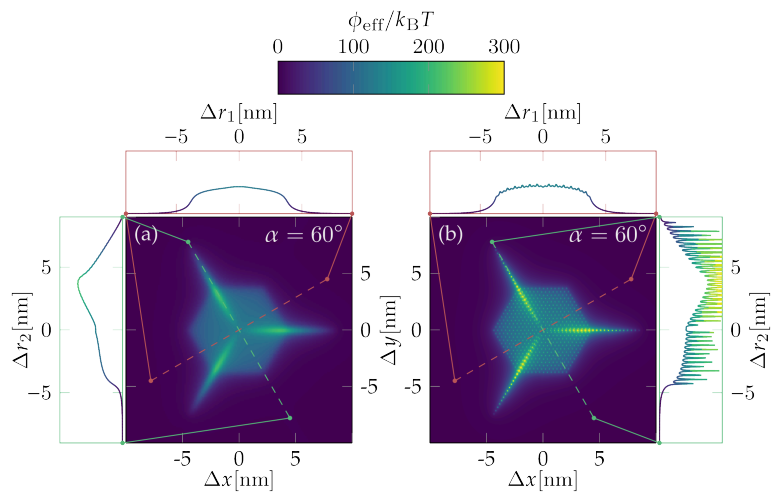
FIGURE C.21: Effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars as function of center-of-mass shift Δx and Δy and rotation angle α. $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ is given in units of $k_{\text{B}}T$, whereas Δx and Δy are given in Å. Both DNA stars are assumed to be in a planar configuration with rigid arms and interarm angles $\theta_i = 2\pi/3$ ($i = 1, 2, 3$) while being rotated by different values of angle α against each other (as labelled). Two dashed, orthogonal lines in red and green indicate two slices through the energy landscape of the effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$. The energy profiles of these slices, denoted by $\phi_{\text{eff}}(\alpha; \Delta r_1)$ and $\phi_{\text{eff}}(\alpha; \Delta r_2)$, are shown in the plots framed red and green. The data for this plot was obtained *via* the reference method outlined in Abaurrea Velasco et al., "Effective interactions of DNA-stars" (subplots (a) and (c)) and *via* the WI method (subplots (b) and (d)).
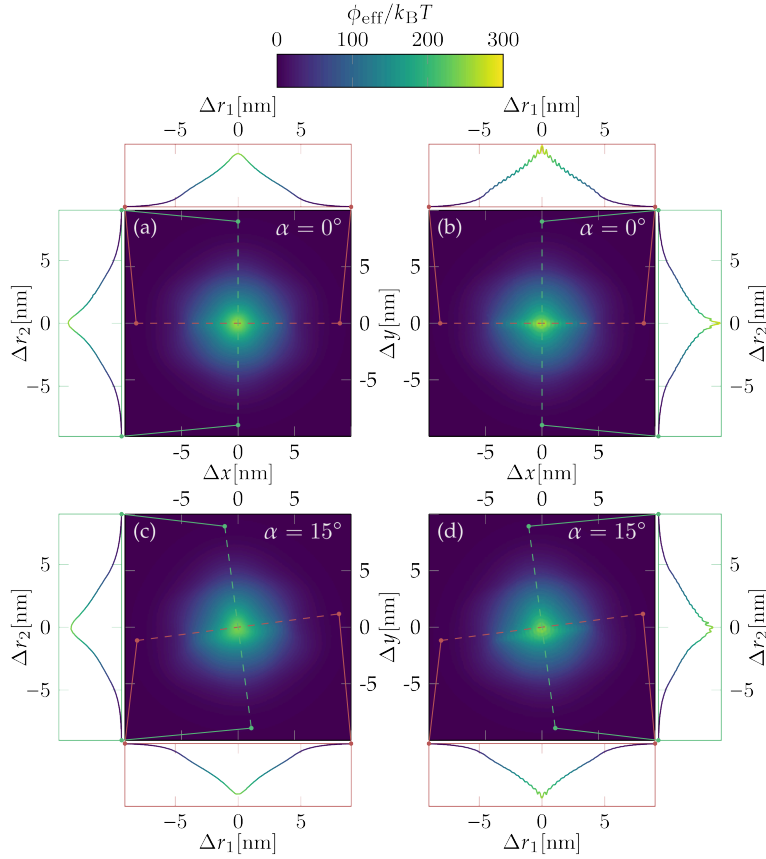
FIGURE C.22: Effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars as function of center-of-mass shift Δx and Δy and rotation angle α. $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ is given in units of $k_{\text{B}}T$, whereas Δx and Δy are given in Å. Both DNA stars are assumed to be in a planar configuration with rigid arms and interarm angles $\theta_i = 2\pi/3$ ($i = 1, 2, 3$) while being rotated by different values of angle α against each other (as labelled). Two dashed, orthogonal lines in red and green indicate two slices through the energy landscape of the effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$. The energy profiles of these slices, denoted by $\phi_{\text{eff}}(\alpha; \Delta r_1)$ and $\phi_{\text{eff}}(\alpha; \Delta r_2)$, are shown in the plots framed red and green. The data for this plot was obtained *via* the reference method outlined in Abaurrea Velasco et al., "Effective interactions of DNA-stars" (subplots (a) and (c)) and *via* the WI method (subplots (b) and (d)).
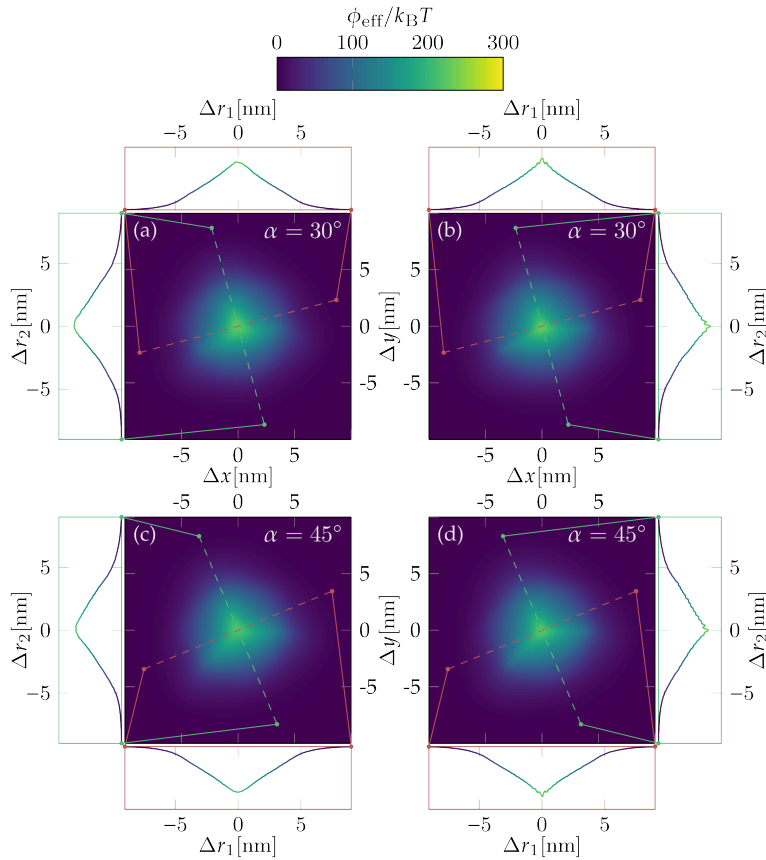

FIGURE C.23: Effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars as function of center-of-mass shift Δx and Δy and rotation angle α. $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ is given in units of $k_{\text{B}}T$, whereas Δx and Δy are given in Å. Both DNA stars are assumed to be in a planar configuration with rigid arms and interarm angles $\theta_i = 2\pi/3$ ($i = 1, 2, 3$) while being rotated by angle $\alpha = 60°$ against each other. Two dashed, orthogonal lines in red and green indicate two slices through the energy landscape of the effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$. The energy profiles of these slices, denoted by $\phi_{\text{eff}}(\alpha; \Delta r_1)$ and $\phi_{\text{eff}}(\alpha; \Delta r_2)$, are shown in the plots framed red and green. The data for this plot was obtained *via* the reference method outlined in Abaurrea Velasco et al., "Effective interactions of DNA-stars" (subplots (a) and (c)) and *via* the WI method (subplots (b) and (d)).

FIGURE C.24: Effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars as function of center-of-mass shift Δx and Δy and rotation angle α. $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ is given in units of $k_B T$, whereas Δx and Δy are given in Å. Both DNA stars are assumed to be equilibrated while being rotated by different values of angle α against each other (as labelled). Two dashed, orthogonal lines in red and green indicate two slices through the energy landscape of the effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$. The energy profiles of these slices, denoted by $\phi_{\text{eff}}(\alpha; \Delta r_1)$ and $\phi_{\text{eff}}(\alpha; \Delta r_2)$, are shown in the plots framed red and green. The data for this plot was obtained *via* the Widom insertion method with effective coordinate r_{com} in subplots (a) and (c) and effective coordinate r_{cy} in subplots (b) and (d).


FIGURE C.25: Effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars as function of center-of-mass shift Δx and Δy and rotation angle α. $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ is given in units of $k_B T$, whereas Δx and Δy are given in Å. Both DNA stars are assumed to be equilibrated while being rotated by different values of angle α against each other (as labelled). Two dashed, orthogonal lines in red and green indicate two slices through the energy landscape of the effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$. The energy profiles of these slices, denoted by $\phi_{\text{eff}}(\alpha; \Delta r_1)$ and $\phi_{\text{eff}}(\alpha; \Delta r_2)$, are shown in the plots framed red and green. The data for this plot was obtained *via* the Widom insertion method with effective coordinate r_{com} in subplots (a) and (c) and effective coordinate r_{cy} in subplots (b) and (d).

151

FIGURE C.26: Effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ between two DNA stars as function of center-of-mass shift Δx and Δy and rotation angle α. $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$ is given in units of $k_{\text{B}}T$, whereas Δx and Δy are given in Å. Both DNA stars are assumed to be equilibrated while being rotated by an angle of $\alpha = 60°$ against each other. Two dashed, orthogonal lines in red and green indicate two slices through the energy landscape of the effective potential $\phi_{\text{eff}}(\alpha; \Delta x, \Delta y)$. The energy profiles of these slices, denoted by $\phi_{\text{eff}}(\alpha; \Delta r_1)$ and $\phi_{\text{eff}}(\alpha; \Delta r_2)$, are shown in the plots framed red and green. The data for this plot was obtained *via* the Widom insertion method with effective coordinate r_{com} in subplots (a) and (c) and effective coordinate r_{cy} in subplots (b) and (d).
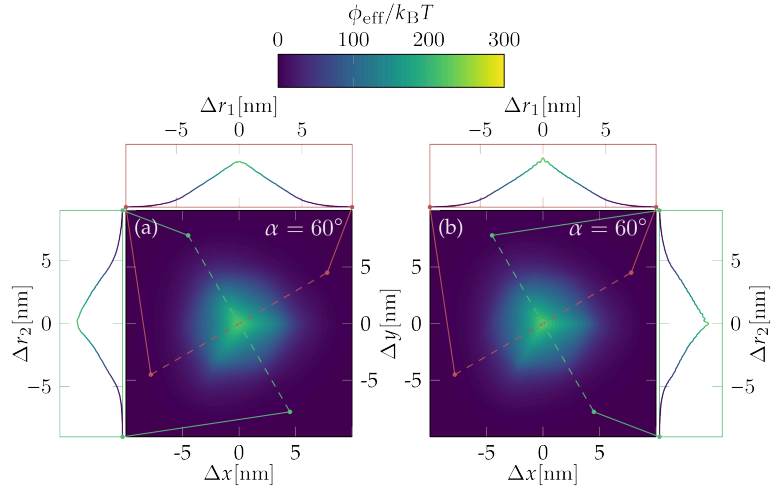
FIGURE C.27: Immersion depth Δz as function of the local number density n_{loc} for the two-dimensional G1-tripod system. Δz is given in Å.
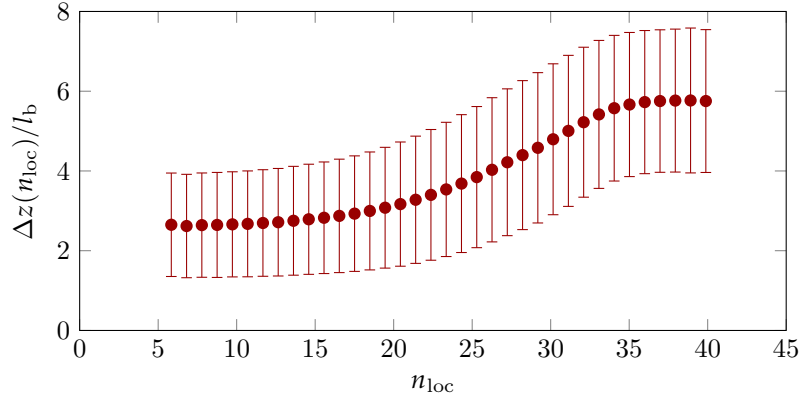
FIGURE C.28: Probability distribution $P(n_{\text{loc}})$ as function of the local number density n_{loc} for the two-dimensional G1-tripod system. Results are plotted for different values of the surface density η_{s}. The probability distributions are normalised *via* $\int_0^\infty P(n_{\text{loc}}) \, \text{d}n_{\text{loc}} = 1$.

C.4 Complementary Results for Chapter 4

TABLE C.3: Shape descriptors for DL-DNA dendrimers (G1 to G8) in the dilute solutions at salt concentrations $c_1 = 150\,\text{mM}$ and $c_1 = 150\,\text{mM}$: values are listed for radius of gyration R_g, eigenvalues of the gyration tensor λ_i^2 (with $i = 1, 2, 3$), asphericity a, acylindricity b, and relative shape anisotropy κ^2, see Section 2.6.5 for definitions. Eigenvalues λ_i^2 are ordered by size, *i.e.*, $\lambda_1^2 < \lambda_2^2 < \lambda_3^2$. The results were obtain from MD simulations using appropriate time averaging and are given in the form $\mu \pm \sigma$, with mean μ and standard deviation σ.

GN	R_g[Å]	$\frac{\lambda_1^2}{R_g^2}$	$\frac{\lambda_2^2}{R_g^2}$	$\frac{\lambda_3^2}{R_g^2}$	$\frac{\lambda_1^2}{\lambda_3^2}$	$\frac{\lambda_2^2}{\lambda_3^2}$	$\frac{b}{R_g^2}$	$\frac{c}{R_g^2}$	κ^2
G1	33.09 ± 0.40	0.006 ± 0.007	0.401 ± 0.048	0.594 ± 0.048	0.010 ± 0.012	0.685 ± 0.133	0.391 ± 0.072	0.395 ± 0.049	0.277 ± 0.034
G2	96.10 ± 1.25	0.024 ± 0.009	0.416 ± 0.037	0.559 ± 0.036	0.044 ± 0.017	0.752 ± 0.111	0.339 ± 0.055	0.392 ± 0.039	0.234 ± 0.022
G3	157.47 ± 2.14	0.146 ± 0.028	0.383 ± 0.028	0.471 ± 0.026	0.312 ± 0.069	0.817 ± 0.087	0.207 ± 0.039	0.237 ± 0.049	0.088 ± 0.025
G4	221.76 ± 2.64	0.206 ± 0.018	0.367 ± 0.017	0.427 ± 0.018	0.484 ± 0.057	0.863 ± 0.064	0.140 ± 0.027	0.162 ± 0.030	0.041 ± 0.011
G5	289.21 ± 3.90	0.254 ± 0.012	0.355 ± 0.012	0.391 ± 0.011	0.650 ± 0.043	0.908 ± 0.046	0.087 ± 0.017	0.101 ± 0.021	0.016 ± 0.005
G6	362.10 ± 7.49	0.268 ± 0.012	0.356 ± 0.009	0.376 ± 0.007	0.715 ± 0.040	0.949 ± 0.028	0.063 ± 0.011	0.088 ± 0.019	0.010 ± 0.004
G7	439.93 ± 12.97	0.261 ± 0.008	0.358 ± 0.006	0.381 ± 0.005	0.685 ± 0.027	0.940 ± 0.020	0.072 ± 0.007	0.097 ± 0.014	0.012 ± 0.003
G8	536.32 ± 35.33	0.267 ± 0.011	0.355 ± 0.007	0.378 ± 0.005	0.706 ± 0.035	0.939 ± 0.016	0.067 ± 0.007	0.088 ± 0.017	0.011 ± 0.003

Bibliography

Abaurrea Velasco, C., C. N. Likos, and G. Kahl. "Effective interactions of DNA-stars". In: *Mol. Phys.* 113 (2015), pp. 2699–2706. DOI: `10.1080/00268976.2015.1048318`.

Abbott, E. A. *Flatland: A Romance of Many Dimensions*. London: Oxford University Press, 2008. URL: `https://worldcat.org/isbn/9780199537501`.

Adžić, N., C. Jochum, E. Stiakakis, G. Kahl, and C. N. Likos. (in preparation).

Agard, N. J., J. A. Prescher, and C. R. Bertozzi. "A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems". In: *J. Am. Chem. Soc.* 126 (2004), pp. 15046–15047. DOI: `10.1021/ja044996f`.

Allen, M. P. and D. J. Tildesley. *Computer Simulations of Liquids*. Oxford University Press, 2017. DOI: `10.1093/oso/9780198803195.001.0001`.

Alvarado, J., M. Sheinman, A. Sharma, F. C. MacKintosh, and G. H. Koenderink. "Force percolation of contractile active gels". In: *Soft Matter* 13 (2017), pp. 5624–5644. DOI: `10.1039/C7SM00834A`.

Arnold, A., O. Lenz, S. Kesselheim, R. Weeber, F. Fahrenberger, D. Roehm, P. Košovan, and C. Holm. "ESPResSo 3.1 — Molecular Dynamics Software for Coarse-Grained Models". In: vol. 89. Lecture Notes in Computational Science and Engineering. 2013, pp. 1–23. DOI: `10.1007/978-3-642-32979-1_1`.

Astruc, D., E. Boisselier, and C. Ornelas. "Dendrimers Designed for Functions: From Physical, Photophysical, and Supramolecular Properties to Applications in Sensing, Catalysis, Molecular Electronics, Photonics, and Nanomedicine". In: *Chem. Rev.* 110.4 (2010), pp. 1857–1959. DOI: `10.1021/cr900327d`.

Atkins, P. W. and J. de Paula. *Atkins' Physical chemistry*. 11th. Oxford; New York: Oxford University Press, 2018. URL: `http://www.worldcat.org/isbn/9780198769866`.

Avakyan, N., J. W. Conway, and H. F. Sleiman. "Long-Range Ordering of Blunt-Ended DNA Tiles on Supported Lipid Bilayers". In: *J. Am. Chem. Soc.* 139.34 (2017), pp. 12027–12034. DOI: `10.1021/jacs.7b06572`.

Ballauff, M. and C. N. Likos. "Dendrimers in Solution: Insight from Theory and Simulation". In: *Angew. Chem. Int. Ed.* 43 (2004), pp. 2998–3020. DOI: `10.1002/anie.200300602`.

Beaucage, G. "Approximations Leading to a Unified Exponential/Power-Law Approach to Small-Angle Scattering". In: *J. Appl. Cryst.* 28 (1995), pp. 717–728. DOI: `10.1107/S0021889895005292`.

Behler, J. and M. Parrinello. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces". In: *Phys. Rev. Lett.* 98 (2007), p. 146401. DOI: `10.1103/PhysRevLett.98.146401`.

Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. "Molecular dynamics with coupling to an external bath". In: *J. Chem. Phys.* 81 (1984), pp. 3684–3690. DOI: `10.1063/1.448118`.

Beveridge, D. L., T. E. Cheatham, and M. Mezei. "The ABCs of molecular dynamics simulations on B-DNA, circa 2012". In: *J. Biosci.* 37 (2012), pp. 379–397. DOI: `10.1007/s12038-012-9222-6`.

Biffi, S., R. Cerbino, F. Bomboi, E. M. Paraboschi, R. Asselta, F. Sciortino, and T. Bellini. "Phase behavior and critical activated dynamics of limited-valence DNA nanostars". In: 110 (2013), pp. 15633–15637. DOI: `10.1073/pnas.1304632110`.

Blaak, R., B. Capone, C. N. Likos, and L. Rovigatti. *Accurate coarse-grained potentials for soft matter systems*. Ed. by G. Sutmann, J. Grotendorst, G. Gompper, and D. Marx. Vol. 28. IAS Series. Jülich: Forschungszentrum Jülich, Zentralbibliothek, 2015, pp. 209–258. URL: `https://juser.fz-juelich.de/record/188877`.

Blaak, R., S. Lehmann, and C. N. Likos. "Charge-Induced Conformational Changes of Dendrimers". In: *Macromolecules* 41 (2008), pp. 4452–4458. DOI: `10.1021/ma800283z`.

Born, M. "Quantenmechanik der Stoßvorgänge". In: *Z. Phys.* 38 (1926), pp. 803–827. DOI: `10.1007/BF01397184`.

Broyles, A. A., S. U. Chung, and H. L. Sahlin. "Comparison of Radial Distribution Functions from Integral Equations and Monte Carlo". In: *J. Chem. Phys.* 37 (1962), pp. 2462–2469. DOI: `10.1063/1.1733028`.

Buchner, F. "Coarse-graining of dendrimer-like DNA using high-dimensional neural network potentials". TU Wien, 2019. (unpublished project thesis).

Buhleier, E., W. Wehner, and F. Vögtle. ""Cascade"- and "Nonskid-Chain-like" Syntheses of Molecular Cavity Topologies". In: *Synthesis* 2 (1978). 155, pp. 155–158. DOI: `10.1055/s-1978-24702`.

Carroll, L. *Alice's Adventures in Wonderland and Through the Looking Glass*. London: Penguin Books, 1998. URL: https://worldcat.org/isbn/9780141439761.

Chen, H., S. P. Meisburger, S. A. Pabit, J. L. Sutton, W. W. Webb, and L. Pollack. "Ionic strength-dependent persistence lengths of single-stranded RNA and DNA". In: *Proc. Natl. Acad. Sci. U.S.A.* 109 (2012), pp. 799–804. DOI: 10.1073/pnas.1119057109.

Chen, H. "Kinematic Diffraction of X-Rays". In: *Characterization of Materials*. 2012, pp. 1–21. DOI: 10.1002/0471266965.com018.pub2.

Chen, S. and G. D. Doolen. "Lattice Boltzmann Method for Fluid Flows". In: *Annu. Rev. Fluid Mech.* 30 (1998), pp. 329–364. DOI: 10.1146/annurev.fluid.30.1.329.

Colla, T., C. N. Likos, and L. Levin. "Equilibrium Properties of Charged Microgels: A Poisson-Boltzmann-Flory Approach". In: *J. Chem. Phys.* 141 (2014), p. 234902. DOI: 10.1063/1.4903746.

De Gennes, P.-G. "Soft Matter (Nobel Lecture)". In: *Angew. Chem. Int. Ed.* 31 (1992), pp. 842–845. DOI: 10.1002/anie.199208421.

De Gennes, P.-G. "Soft matter: more than words". In: *Soft Matter* 1 (2005), pp. 16–16. DOI: 10.1039/B419223K.

Debye, P. and E. Hückel. "Zur Theorie der Elektrolyte. I. Gefrierpunktserniedrigung und verwandte Erscheinungen". In: *Physikalische Zeitschrift* 24 (1923), pp. 185–206.

Deserno, M. and C. Holm. "How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines". In: *J. Chem. Phys.* 109 (1998), pp. 7678–7693. DOI: 10.1063/1.477414.

Deserno, M. and C. Holm. "How to mesh up Ewald sums. II. An accurate error estimate for the particle–particle–particle-mesh algorithm". In: *J. Chem. Phys.* 109 (1998), pp. 7694–7701. DOI: 10.1063/1.477415.

Dickerson, R. "Definitions and nomenclature of nucleic acid structure components". In: *Nucleic Acids Res.* 17 (1989), pp. 1797–1803. DOI: 10.1093/nar/17.5.1797.

Doi, M. *Introduction to Polymer Physics*. Trans. by H. See. Oxford University Press, 1996. URL: http://www.worldcat.org/isbn/9780198517894.

Doi, M. *Soft Matter Physics*. Oxford University Press, 2013. DOI: 10.1093/acprof:oso/9780199652952.001.0001.

Egelhaaf, S. U. and P. Schurtenberger. "Shape Transformations in the Lecithin-Bile Salt System: From Cylinders to Vesicles". In: *J. Phys. Chem.* 98 (1994), pp. 8560–8573. DOI: 10.1021/j100085a041.

Ewald, P. P. „Die Berechnung Optischer und Elektrostatischer Gitterpotentiale". In: *Ann. Phys. (Berl.)* 369 (1921), S. 253–287. DOI: 10.1002/andp.19213690304.

Frenkel, D. and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press, 2001. DOI: 10.1016/B978-0-12-267351-1.X5000-7.

Gillan, M. J. "A new method of solving the liquid structure integral equations". In: *Mol. Phys.* 38 (1979), pp. 1781–1794. DOI: 10.1080/00268977900102861.

Goldstein, H. *Classical Mechanics*. 3rd ed. Pearson, 2002. URL: https://www.worldcat.org/isbn/9780201657029.

Gompper, G., J. K. G. Dhont, and D. Richter. "Eine Welt zwischen Fest und Flüssig: Aktuelle Forschung an Weicher Materie". In: *Physik in unserer Zeit* 34 (2003), pp. 19–25. DOI: 10.1002/piuz.200390003.

Gompper, G., J. K. G. Dhont, and D. Richter. "Komplexe Materialien auf mesoskopischer Skala: Was ist Weiche Materie?" In: *Physik in unserer Zeit* 34 (2003), pp. 12–18. DOI: 10.1002/piuz.200390002.

Götze, I. O. and C. N. Likos. "Conformations of Flexible Dendrimers: A Simulation Study". In: *Macromolecules* 36 (2003), pp. 8189–8197. DOI: 10.1021/ma030137k.

Grest, G. S., L. J. Fetters, J. S. Huang, and D. Richter. *Advances in Chemical Physics. Polymeric Systems*. Vol. 94. New York: J. Wiley & Sons, 1996, pp. 67–163. DOI: 10.1002/9780470141533.

Guinier, A. and G. Fournet. *Small angle scattering of X-rays*. Trans. by C. B. Wilson. J. Wiley & Sons, 1955. DOI: 10.1002/pol.1956.120199326.

Hansen, J.-P. and I. R. McDonald. *Theory of Simple Liquids*. Academic Press, 2013. URL: http://www.worldcat.org/isbn/9780123870322.

Hardy, D. J., M. A. Wolff, J. Xia, K. Schulten, and R. D. Skeel. "Multilevel Summation with B-spline Interpolation for Pairwise Interactions in Molecular Dynamics Simulations". In: *J. Chem. Phys.* 144 (2016), p. 114112. DOI: 10.1063/1.4943868.

Hardy, D. J., Z. Wu, J. C. Phillips, J. E. Stone, R. D. Skeel, and K. Schulten. "Multilevel Summation Method for Electrostatic Force Evaluation". In: *J. Chem. Theory Comput.* 11 (2015), pp. 766–779. DOI: 10.1021/ct5009075.

Harreis, H. M., C. N. Likos, and M. Ballauff. "Can Dendrimers Be Viewed as Compact Colloids? A Simulation Study of the Fluctuations in a Dendrimer of Fourth Generation". In: *J. Chem. Phys.* 118 (2003), pp. 1979–1988. DOI: 10.1063/1.1530577.

Hawker, C. J. and J. M. J. Fréchet. "Preparation of polymers with controlled molecular architecture. A new convergent approach to dendritic macromolecules". In: *J. Am. Chem. Soc.* 112 (1990), pp. 7638–7647. DOI: 10.1021/ja00177a027.

Heinemann, T., K. Palczynski, J. Dzubiella, and S. H. L. Klapp. "Angle-resolved effective potentials for disk-shaped molecules". In: *J. Chem. Phys.* 141 (2014), p. 214110. DOI: 10.1063/1.4902824.

Heisenberg, W. "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik". In: *Z. Phys.* 43 (1927), pp. 172–198. DOI: 10.1007/BF01397280.

Holde, K. E. *Chromatin*. New York: Springer New York, 1989. DOI: 10.1007/978-1-4612-3490-6.

Hoogerbrugge, P. J. and J. M. V. A. Koelman. "Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics". In: *EPL* 19 (1992), pp. 155–160. DOI: 10.1209/0295-5075/19/3/001.

Huang, K. *Statistical Mechanics*. 2nd ed. New York: Wiley, 1987. URL: https://www.worldcat.org/isbn/9780471815181.

Ingólfsson, H. I., C. A. Lopez, J. J. Uusitalo, D. H. de Jong, S. M. Gopal, X. Periole, and S. J. Marrink. "The power of coarse graining in biomolecular simulations". In: *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 4 (2014), pp. 225–248. DOI: 10.1002/wcms.1169.

Jochum, C., N. Adžić, E. Stiakakis, T. L. Derrien, D. Luo, G. Kahl, and C. N. Likos. "Structure and stimuli-responsiveness of all-DNA dendrimers: theory and experiment". In: *Nanoscale* 11 (2019), pp. 1604–1617. DOI: 10.1039/C8NR05814H.

Jochum, C., N. Adžić, E. Stiakakis, G. Kahl, and C. N. Likos. "DNA Stars Confined to an Interface: Planar vs. Tripod Configurations". (to be submitted).

Jones, R. A. L. *Soft Matter Physics*. Oxford University Press, 2002. URL: https://www.worldcat.org/isbn/9780198505891.

Jusufi, A., C. N. Likos, and H. Löwen. "Conformations and Interactions of Star-Branched Polyelectrolytes". In: *Phys. Rev. Lett.* 88 (2001), p. 018301. DOI: 10.1103/PhysRevLett.88.018301.

Jusufi, A., C. N. Likos, and H. Löwen. "Counterion-induced entropic interactions in solutions of strongly stretched, osmotic polyelectrolyte stars". In: *J. Chem. Phys.* 116 (2002), pp. 11011–11027. DOI: 10.1063/1.1480007.

Kegler, K., M. Konieczny, G. Dominguez-Espinosa, C. Gutsche, M. Salomo, F. Kremer, and C. N. Likos. "Polyelectrolyte-Compression Forces between

Spherical DNA Brushes". In: *Phys. Rev. Lett.* 100 (2008), p. 118302. DOI: `10.1103/PhysRevLett.100.118302`.

Kegler, K., M. Salomo, and F. Kremer. "Forces of Interaction between DNA-Grafted Colloids: An Optical Tweezer Measurement". In: *Phys. Rev. Lett.* 98 (2007), p. 058304. DOI: `10.1103/PhysRevLett.98.058304`.

Kłos, J. S. "Dendritic Polyelectrolytes Revisited Through the Poisson-Boltzmann-Flory Theory and the Debye-Hückel Approximation". In: *Phys. Chem. Chem. Phys.* 20 (2018), pp. 2693–2703. DOI: `10.1039/C7CP07138H`.

Kornyshev, A. A., D. J. Lee, S. Leikin, and A. Wynveen. "Structure and interactions of biological helices". In: *Rev. Mod. Phys.* 79 (2007), pp. 943–996. DOI: `10.1103/RevModPhys.79.943`.

Landau, L. D. and E. M. Lifshitz. *Mechanics, Third Edition: Volume 1 (Course of Theoretical Physics)*. 3rd. Butterworth-Heinemann, 1976. URL: `http://www.worldcat.org/isbn/0750628960`.

Lee, C. C., J. A. MacKay, J. M. J. Fréchet, and F. C. Szoka. "Designing dendrimers for biological applications". In: *Nat. Biotechnol.* 23 (2005), pp. 1517–1526. DOI: `10.1038/nbt1171`.

Leeuwen, J. van, J. Groeneveld, and J. de Boer. "New method for the calculation of the pair correlation function. I". In: *Physica* 25.7 (1959), pp. 792–808. DOI: `10.1016/0031-8914(59)90004-7`.

Lenz, D. A. "Self-organization of dendrimers and dendrimer-colloid mixtures". PhD thesis. University of Vienna, AT, 2012. URL: `http://othes.univie.ac.at/21039/`.

Lenz, D. A., B. M. Mladek, C. N. Likos, G. Kahl, and R. Blaak. "Monomer-Resolved Simulations of Cluster-Forming Dendrimers". In: *J. Phys. Chem. B* 115 (2011), pp. 7218–7226. DOI: `10.1021/jp109132m`.

Lenz, D. A., B. M. Mladek, C. N. Likos, G. Kahl, and R. Blaak. "Thermodynamic Stability and Structural Properties of Cluster Crystals Formed by Amphiphilic Dendrimers". In: *J. Chem. Phys.* 144 (2016), p. 204901. DOI: `10.1063/1.4950953`.

Li, Y., Y. D. Tseng, S. Y. Kwon, L. d'Espaux, J. S. Bunch, P. L. McEuen, and D. Luo. "Controlled assembly of dendrimer-like DNA". In: *Nat. Mat.* 3 (2003), pp. 38–42. DOI: `10.1038/nmat1045`.

Likos, C. N. "Colloidal interactions: From effective potentials to structure". In: *Riv. del Nuovo Cim.* 37 (2014), pp. 125–180. DOI: `10.1393/ncr/i2014-10098-1`.

Limbach, H. J., A. Arnold, B. A. Mann, and C. Holm. "ESPResSo - an Extensible Simulation Package for Research on Soft Matter Systems". In:

Comput. Phys. Commun. 174 (2006), pp. 704–727. DOI: `10.1016/j.cpc.2005.10.005`.

Liu, J., W. D. Gray, M. E. Davis, and Y. Luo. "Peptide- and saccharide-conjugated dendrimers for targeted drug delivery: a concise review". In: *Interface Focus* 2 (2012), pp. 307–324. DOI: `10.1098/rsfs.2012.0009`.

Lubensky, T. "Soft condensed matter physics". In: *Solid State Communications* 102 (1997), pp. 187–197. DOI: `https://doi.org/10.1016/S0038-1098(96)00718-1`.

Marko, J. F. and E. D. Siggia. "Stretching DNA". In: *Macromolecules* 28 (1995), pp. 8759–8770. DOI: `10.1021/ma00130a008`.

Matta, C. F., L. Massa, A. V. Gubskaya, and E. Knoll. "Can One Take the Logarithm or the Sine of a Dimensioned Quantity or a Unit? Dimensional Analysis Involving Transcendental Functions". In: *J. Chem. Educ.* 88 (2011), pp. 67–70. DOI: `10.1021/ed1000476`.

Mei, C., Y. Sun, G. Zheng, E. J. Bohm, L. V. Kale, J. C. Phillips, and C. Harrison. "Enabling and scaling biomolecular simulations of 100 million atoms on petascale machines wit, a multicore-optimized message-driven runtime". In: *SC '11: Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*. 2011, pp. 1–11.

Meng, H. M., X. Zhang, Y. Lv, Z. Zhao, N. N. Wang, T. Fu, H. Fan, H. Liang, L. Qiu, G. Zhu, and W. Tan. "DNA Dendrimer: An Efficient Nanocarrier of Functional Nucleic Acids for Intracellular Molecular Sensing". In: *ACS Nano* 8 (2014), pp. 6171–6181. DOI: `10.1021/nn5015962`.

Milton, J. *Paradise Lost*. Penguin Books, 2003. URL: `https://worldcat.org/isbn/9780140424393`.

Mitov, M. *Sensitive Matter: Foams, Gels, Liquid Crystals, and Other Miracles*. Trans. by G. Weiss. Harvard University Press, 2012. URL: `https://worldcat.org/isbn/9780674064560`.

Mladek, B. and D. Frenkel. "Pair interactions between complex mesoscopic particles from Widom's particle-insertion method". In: *Soft Matter* 7 (2011), pp. 1450–1455. DOI: `10.1039/C0SM00815J`.

Morawietz, T., A. Singraber, C. Dellago, and J. Behler. "How van der Waals interactions determine the unique properties of water". In: *Proc. Natl. Acad. Sci. U.S.A.* 113 (2016), pp. 8368–8373. DOI: `10.1073/pnas.1602375113`.

Morita, T. "Theory of Classical Fluids: Hyper-Netted Chain Approximation, I: Formulation for a One-Component System". In: *Progress of Theoretical Physics* 20.6 (1958), pp. 920–938. DOI: `10.1143/PTP.20.920`.

Morita, T. and K. Hiroike. "A New Approach to the Theory of Classical Fluids. I: Formulation for a One-Component System". In: *Prog. Theor. Phys.* 23 (1960), pp. 1003–1027. DOI: 10.1143/PTP.23.1003.

Newkome, G. R., Z. Yao, G. R. Baker, and V. K. Gupta. "Micelles. Part 1. Cascade molecules: a new approach to micelles. A [27]-arborol". In: *The Journal of Organic Chemistry* 50 (1985), pp. 2003–2004. DOI: 10.1021/jo00211a052.

Ornstein, L. S. and F. Zernike. "Accidental deviations of density and opalescence at the critical point of a single substance". In: *Proc. Akad. Sci.* 17 (1914), pp. 793–806. URL: https://www.dwc.knaw.nl/DL/publications/PU00012727.pdf.

Ortega, J. and W. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables.* Ed. by J. Ortega and W. Rheinboldt. Academic Press, 1970, pp. 1–6. DOI: 10.1016/B978-0-12-528550-6.50008-9.

Ouldridge, T. E. "Coarse-grained modelling of DNA and DNA self-assembly". PhD thesis. Oxford University, UK, 2011. URL: https://ora.ox.ac.uk/objects/uuid:b2415bb2-7975-4f59-b5e2-8c022b4a3719.

Ouldridge, T. E., A. A. Louis, and J. P. K. Doye. "Structural, Mechanical, and Thermodynamic Properties of a Coarse-Grained DNA Model". In: *J. Chem. Phys.* 134 (2011), p. 085101. DOI: 10.1063/1.3552946.

Ousterhout, J. K. *Tcl and the Tk Toolkit.* Flatbrain Com, 1996. URL: https://dl.acm.org/doi/book/10.5555/524313.

Paradossi, G. and E. Chiessi. "Solution behaviour of poly(N-isopropylacrylamide) stereoisomers in water: a molecular dynamics simulation study". In: *Phys. Chem. Chem. Phys.* 19 (2017), pp. 11892–11903. DOI: 10.1039/C7CP00808B.

Percus, J. K. and G. J. Yevick. "Analysis of Classical Statistical Mechanics by Means of Collective Coordinates". In: *Phys. Rev.* 110 (1 1958), pp. 1–13. DOI: 10.1103/PhysRev.110.1.

Pierce, C. S. *Collected Papers of Charles Sanders Peirce, Volumes V and VI: Pragmatism and Pragmaticism and Scientific Metaphysics.* Ed. by C. Hartshorne and P. Weiss. Belknap Press, 1935. URL: https://worldcat.org/isbn/9780674138025.

Plimpton, S. "Fast Parallel Algorithms for Short-Range Molecular Dynamics". In: *J. Comp. Phys.* 117 (1995), pp. 1–19. DOI: 10.1006/jcph.1995.1039.

Poier, P., C. N. Likos, A. J. Moreno, and R. Blaak. "An Anisotropic Effective Model for the Simulation of Semiflexible Ring Polymers". In: *Macromolecules* 48 (2015), pp. 4983–4997. DOI: 10.1021/acs.macromol.5b00603.

Pratchett, T. *Hogfather*. London: Corgi, 2013. URL: https://worldcat.org/isbn/9780552167581.

Rapaport, D. C. *The Art of Molecular Dynamics Simulation*. 2nd ed. Cambridge University Press, 2004. DOI: 10.1017/CBO9780511816581.

Roh, Y. H., K. Lee, J. J. Ye, and D. Luo. "Multivalent DNA-based Vectors for DNA Vaccine Delivery". In: *Methods. Mol. Biol.* 1143 (2014), pp. 159–179. DOI: 10.1007/978-1-4939-0410-5_11.

Romano, F. and F. Sciortino. "Switching Bonds in a DNA Gel: An All-DNA Vitrimer". In: *Phys. Rev. Lett.* 114 (2015), p. 078104. DOI: 10.1103/PhysRevLett.114.078104.

Rossum, G. van. *Python tutorial*. Tech. rep. CS-R9526. Amsterdam: Centrum voor Wiskunde en Informatica (CWI), 1995. URL: https://ir.cwi.nl/pub/5007/05007D.pdf.

Rothemund, P. "Folding DNA to Create Nanoscale Shapes and Patterns". In: *Nature* 440 (2006), pp. 297–302. DOI: 10.1038/nature04586.

Rouse, P. E. "A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers". In: *J. Chem. Phys.* 21 (1953), pp. 1272–1280. DOI: 10.1063/1.1699180.

Russell, B. *An Outline of Philosophy*. G. Allen & Unwin, Ltd., 1927.

Salamonczyk, M., J. Zhang, G. Portale, C. Zhu, E. Kentzinger, J. T. Gleeson, A. Jakli, C. de Michele, J. K. G. Dhont, S. Sprunt, and E. Stiakakis. "Smectic Phase in Suspensions of Gapped DNA Duplexes". In: *Nat. Commun.* 7 (2016), p. 13358. DOI: 10.1038/ncomms13358.

Sayre, A. *Rosalind Franklin and DNA*. 1st ed. New York: W. W. Norton & Co., 2000. URL: https://worldcat.org/isbn/9780393320442.

Schwabl, F. *Statistische Mechanik*. Berlin: Springer-Verlag, 2006. DOI: 10.1007/3-540-31097-5.

Seeman, N. C. "Nucleic Acid Junctions and Lattices". In: *J. Theor. Biol.* 99.2 (1982), pp. 237–247. DOI: 10.1016/0022-5193(82)90002-9.

Seeman, N. C. "Structural DNA Nanotechnology". In: *Nano-Biotechnology Protocols*. Ed. by S. J. Rosenthal and D. W. Wright. Humana Press, 2005, pp. 143–166. DOI: 10.1385/1-59259-901-X:143.

Siavashpouri, M., C. H. Wachauf, M. J. Zakhary, F. Praetorius, H. Dietz, and Z. Dogic. "Molecular engineering of chiral colloidal liquid crystals using DNA origami". In: *Nat. Mater.* 16 (2017), pp. 849–856. DOI: 10.1038/nmat4909.

Silverberg, J. L. "A Big Tent for Soft Matter". In: *APS News* 24.5 (May 2015). URL: https://www.aps.org/publications/apsnews/201505/ (visited on 04/20/2020).

Singraber, A. "Designing and training neural network potentials for molecular dynamics simulations". PhD thesis. University of Vienna, AT, 2018. URL: http://othes.univie.ac.at/55534/.

Singraber, A., J. Behler, and C. Dellago. "Library-Based LAMMPS Implementation of High-Dimensional Neural Network Potentials". In: *J. Chem. Theory Comput.* 15 (2019), pp. 1827–1840. DOI: 10.1021/acs.jctc.8b00770.

Singraber, A., T. Morawietz, J. Behler, and C. Dellago. "Parallel Multistream Training of High-Dimensional Neural Network Potentials". In: *J. Chem. Theory Comput.* 15 (2019), pp. 3075–3092. DOI: 10.1021/acs.jctc.8b01092.

Snodin, B. E. K., F. Randisi, M. Mosayebi, P. Šulc, J. S. Schreck, F. Romano, T. E. Ouldridge, R. Tsukanov, E. Nir, A. Louis, and J. P. K. Doye. "Introducing improved structural properties and salt dependence into a coarse-grained model of DNA". In: *J. Chem. Phys.* 142 (2015), p. 234901. DOI: 10.1063/1.4921957.

Šolc, K. "Shape of a Random-Flight Chain". In: *J. Chem. Phys.* 55 (1971), pp. 335–344. DOI: 10.1063/1.1675527.

Sorensen, C. M. "Light Scattering by Fractal Aggregates: A Review". In: *Aerosol Sci. Technol.* 35 (2001), pp. 648–687. DOI: 10.1080/02786820117868.

Stillinger, F. H. "Phase transitions in the Gaussian core system". In: *J. Chem. Phys.* 65.10 (1976), pp. 3968–3974. DOI: 10.1063/1.432891.

Strobl, G. R. *The Physics of Polymers*. Springer, 1996. DOI: 10.1007/978-3662032435.

Swope, W. C., H. C. Andersen, P. H. Berens, and K. R. Wilson. "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters". In: *J. Chem. Phys.* 76 (1982), pp. 6370–649. DOI: 10.1063/1.442716.

Theodorou, D. N. and U. W. Suter. "Shape of unperturbed linear polymers: polypropylene". In: *Macromolecules* 18 (1985), pp. 1206–1214. DOI: 10.1021/ma00148a028.

Tinland, B., A. Pluen, J. Sturm, and G. Weill. "Persistence Length of Single-Stranded DNA". In: *Macromolecules* 30 (1997), pp. 5763–5765. DOI: 10.1021/ma970381+.

Tully, D. C. and J. M. J. Fréchet. "Dendrimers at surfaces and interfaces: chemistry and applications". In: *Chem. Commun.* (14 2001), pp. 1229–1239. DOI: 10.1039/B104290B.

Um, S. H., J. B. Lee, S. Y. Kwon, Y. Li, and D. Luo. "Dendrimer-like DNA-based Fluorescence Nanobarcodes". In: *Nat. Protoc.* 1 (2006), pp. 995–1000. DOI: 10.1038/nprot.2006.141.

Verlet, L. "On the theory of classical fluids". In: *Il Nuovo Cimento (1955-1965)* 18 (1960), pp. 77–101. DOI: 10.1007/BF02726040.

Verlet, L. "Computer "Experiments" on Classical Fluids". In: *Phys. Rev.* 159 (1967), pp. 98–103. DOI: 10.1103/PhysRev.159.98.

Vitelli, V. and W. Irvine. "The geometry and topology of soft materials". In: *Soft Matter* 9 (2013), pp. 8086–8087. DOI: 10.1039/C3SM90111D.

Wagner, N. J., L. M. Walker, and B. Hammouda. "Structure of Isotropic Solutions of Rigid Macromolecules via Small-Angle Neutron Scattering: Poly(y-benzyl L-glutamate)/Deuterated Dimethylformamide". In: *Macromolecules* 28 (1995), pp. 5075–5081. DOI: 10.1021/ma00118a041.

Wang, J., S. Olsson, C. Wehmeyer, A. Pérez, N. E. Charron, G. de Fabritiis, F. Noé, and C. Clementi. "Machine Learning of Coarse-Grained Molecular Dynamics Force Fields". In: *ACS Cent. Sci.* 5 (2019), pp. 755–767. DOI: 10.1021/acscentsci.8b00913.

Watts, R. O. "Integral equation approximations in the theory of fluids". In: *Statistical Mechanics: Volume 1*. Ed. by K. Singer. Vol. 1. The Royal Society of Chemistry, 1973, pp. 1–70. DOI: 10.1039/9781847556929-00001.

Weeks, J. D., D. Chandler, and H. C. Andersen. "Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids". In: *J. Chem. Phys.* 54 (1971), pp. 5237–5247. DOI: 10.1063/1.1674820.

Weik, F., R. Weeber, K. Szuttor, K. Breitsprecher, J. de Graaf, M. Kuron, J. Landsgesell, H. Menke, D. Sean, and C. Holm. "ESPResSo 4.0 - an extensible software package for simulating soft matter systems". In: *Eur. Phys. J. Special Topics* 227.14 (2019), pp. 1789–1816. DOI: 10.1140/epjst/e2019-800186-9.

Weiss, L. B., V. Dahirel, V. Marry, and M. Jardat. "Computation of the Hydrodynamic Radius of Charged Nanoparticles from Nonequilibrium Molecular Dynamics". In: *J. Phys. Chem. B* 122 (2018), pp. 5940–5950. DOI: 10.1021/acs.jpcb.8b01153.

Wertheim, M. S. "Exact Solution of the Percus-Yevick Integral Equation for Hard Spheres". In: *Phys. Rev. Lett.* 10 (8 1963), pp. 321–323. DOI: 10.1103/PhysRevLett.10.321.

Widom, B. *Statistical Mechanics*. Cambridge: Cambridge University Press, 2002. DOI: 10.1017/CBO9780511815836.

Wynveen, A., D. J. Lee, A. A. Kornyshev, and S. Leikin. "Helical coherence of DNA in crystals and solution". In: *Nucleic Acids Res.* 36 (), pp. 5540–5551. DOI: 10.1093/nar/gkn514.

Wynveen, A. and C. N. Likos. "Interactions between planar polyelectrolyte brushes: effects of stiffness and salt". In: *Soft Matter* 6 (2010), pp. 163–171. DOI: 10.1039/B919808C.

Zhang, L., J. Han, H. Wang, R. Car, and W. E. "DeePCG: Constructing coarse-grained models via deep neural networks". In: *J. Chem. Phys.* 149 (2018), p. 034101. DOI: 10.1063/1.5027645.

Zhao, Z., T. Du, F. Liang, and S. Liu. "Amphiphilic DNA Organic Hybrids: Functional Materials in Nanoscience and Potential Application in Biomedicine". In: *International journal of molecular sciences* 19 (2018), p. 2283. DOI: 10.3390/ijms19082283.

Acknowledgements

A grateful mind
By owing owes not, but still pays, at once
Indebted and discharg'd.

—John Milton[284]

No thesis is accomplished by a single person. This is why I want to thank and acknowledge the people, the projects, and the organisations which made my work possible.

My research and my thesis would not be possible without the many individuals who as a collective make up the open source community. These people spend countless hours creating and maintaining the software which many of us use. In the following I will list the most important tools which were used in this work.

While most of my code is written in Python,[285] my large-scale simulations which are based on ESPResSo[286] were performed at the Vienna Scientific Cluster 3 (VSC-3), consuming about 3.5 million CPU core-hours.[287] Many terabytes of simulation data are stored in Hierarchical Data Format 5 (HDF5).[288] Simulation snapshots were created using MayaVi.[289] Like many physics theses this document is typeset using LaTeX[290] — the X$_{\exists}$LaTeX and LuaLaTeX flavours, to be precise. All technical drawings and plots were generated using TikZ[291] and PGFPlots,[292] respectively. Funding by FWF under grant number I 2866-N36 and scholarships for travel and expenses by the TU Wien are acknowledged.

Of course, funding and infrastructure are only some of the necessary ingredients in a successful thesis. Even more important are the people who supported me and who I want to thank.

First, I want to express my sincere gratitude to Gerhard Kahl for giving me the opportunity to work under his supervision and in his group.[293] Not only was I provided with an inspiring and productive working environment, but I also gained many valuable insights into the world of academic research. I am thankful for his patience and for his scientific and moral support I continuously received during the many ups and downs of our projects which culminated in this thesis. While completing a PhD is never easy, his empathy helped overcome the inevitable setbacks I encountered during my

[284] Line 55 of Book IV in J. Milton. *Paradise Lost*. Penguin Books, 2003. URL: https://worldcat.org/isbn/9780140424393.

[285] See https://www.python.org.

[286] See http://espressomd.org.

[287] See https://vsc.ac.at.

[288] See https://www.hdfgroup.org.

[289] See http://docs.enthought.com/mayavi/mayavi.

[290] See https://www.latex-project.org/ for the LaTeX project. For the original TeX system by Donald Knuth, see https://www.tug.org/.

[291] TikZ (acronym for "TikZ ist kein Zeichenprogramm") functions as a user-friendly frontend for the Portable Graphics Format (PGF), see https://www.ctan.org/pkg/pgf.

[292] See https://ctan.org/pkg/pgfplots

[293] *I.e.*, the Soft Matter Theory (SMT) group at the Technische Universität (TU) Wien, see http://smt.tuwien.ac.at/.

work. I wish him all the best and none of the buerocracy in the future.

Furthermore, I would like to thank our collaborators at the University of Vienna and at the Forschungszentrum (FZ) Jülich: I want to thank Christos N. Likos[294] for his experienced guidance in our FWF-project and for the fruitful collaboration between the University of Vienna and TU Wien. His enthusiasm and optimism often provided the much needed encouragement when we hit another dead end in our project. The enrichment of our theoretical results by the experimental work of Manolis Stiakakis[295] definitely made publishing our research easier. Additionally, I am thankful for his answers to my questions about experimental details and for our insightful conversations during my stay at the FZ Jülich. I want to express my special thanks to my colleague Nataša Adžić. Aided by her positive attitude, we persevered together through the many challenges posed by this project. I wish Nataša, Manolis, and Christos success in their future careers.

My gratitude for many hours of pleasant company while working together belongs to all members of the SMT group, whose friendly and kind atmosphere made me feel immediately welcome. For a good part of my studies, this atmosphere was shaped by my PhD cohort consisting of Bene (an inexhaustible source of ideas for side projects), David (our fountain of IT wisdom), and Susanne (whose performances with the TU orchestra are always a pleasure to listen to). Together with my office mate, Gaurav, I made many trips to Knockbox to fight off the post-lunch sleepiness. I also want to thank the rest of the (extended) SMT group: Andreas, Arash, Carina, Dongliang, Emanuela, Florian, and Silvano for creating an environment of scientific appreciation and relaxing lunch breaks.

Of the many other people working at the TU Wien who I have to hank for making my work possible, none are as important as the adminstrative staff at the Institute for Theoretical Physics (ITP),[296] Heike, Ingrid, and Sylvia, who not only took care of the unpleasant buerocracy, but also fostered a homely atmosphere.

I would also like to express my heartfelt appreciation for the company of my friends, with many of whom I have come a long way since starting our physics studies together. Our undertakings, often spurred by our shared interests, never cease to excite me and I hope that we will stay in each other's lives.

To my parents and siblings: I am thankful for the encouragment that I received to pursue my interests. You made me the curious person that I am.

Finally, I am sincerely grateful to Marlis for your support and for the sacrifices you had to make so that I could finish my thesis. I don't believe I could have done it without you.

"Begin at the beginning," the King said, very gravely,
"and go on till you come to the end: then stop."

—Lewis Carroll, *Alice in Wonderland*